

Министерство образования и науки РФ
Национальный исследовательский ядерный университет «МИФИ»

Ю. Г. Древец

Системы реального времени: технические и программные средства

Учебное пособие

Рекомендовано УМО «Ядерные физика и технологии»
для студентов, обучающихся по направлениям
«Ядерная физика и технологии» и
«Информатика и вычислительная техника»,
в качестве учебного пособия

Москва 2010

УДК 681.518 (075)

ББК 32.96-04я7

Д73

Древс Ю.Г. **Системы реального времени: технические и программные средства:** Учебное пособие. М.: МИФИ, 2010. – 320 с.

Описываются принципы действия всех составляющих технического комплекса систем «жесткого» реального времени, даются общие сведения о деятельности лица, принимающего решения, о структуре операционных систем и программировании для систем реального времени. Рассматривается пример структурной схемы программного обеспечения. Даются краткие сведения об условиях эксплуатации систем управления физическими объектами.

Предназначено для студентов инженерных и физических специальностей, изучающих системы управления физическими объектами по направлениям «Ядерная физика и технологии» и «Информатика и вычислительная техника».

Рецензент д-р техн. наук, проф. *В. А. Острейковский*

ISBN 978-5-7262-1310-1

© *Национальный исследовательский ядерный университет «МИФИ», 2010*

Оглавление

1. ОСНОВНЫЕ ПОНЯТИЯ И КЛАССИФИКАЦИЯ СИСТЕМ УПРАВЛЕНИЯ	6
1.1. Основные понятия	6
1.2. Управление как процесс принятия и реализации решения	6
1.3. Классификация систем управления	8
1.4. Структура систем реального времени	11
Контрольные вопросы и упражнения	13
2. ДАТЧИКИ	14
2.1. Структура измерительного канала	14
2.2. Классификация преобразователей	16
2.3. Метрологические характеристики датчиков	19
2.4. Первичные измерительные преобразователи (чувствительные элементы)	27
2.5. Датчики с бинарным и цифровым выходным сигналом	28
2.6. Потенциометрические датчики	31
2.7. Тензометрические датчики	34
2.8. Электромагнитные датчики	38
2.9. Емкостные датчики	47
2.10. Пьезоэлектрические датчики	49
2.11. Терморезисторы	52
2.12. Термоэлектрические датчики	56
2.13. Фотоэлектрические датчики	58
2.14. Струнные датчики	64
2.15. Датчики потока	65
2.16. Антенны	71
Контрольные вопросы и упражнения	73
3. АНАЛОГОВЫЕ СХЕМЫ ПРЕДВАРИТЕЛЬНОЙ ОБРАБОТКИ СИГНАЛОВ	75
3.1. Операционные усилители	75
3.2. Унифицирующие измерительные преобразователи	82
3.3. Фильтры	85
3.4. Мультиплексоры аналоговых сигналов	90
3.5. Схемы выборки и хранения	92
3.6. Детекторы	95
3.7. Схемы гальванической развязки	98
Контрольные вопросы и упражнения	99

4. АНАЛОГО-ЦИФРОВЫЕ ПРЕОБРАЗОВАТЕЛИ	101
4.1. Преобразователи напряжения в частоту	101
4.2. Преобразователи временного интервала, фазы и частоты в код	105
4.3. Преобразователи напряжения в код, построенные по принципу последовательного счета	109
4.4. Преобразователи напряжения в код, построенные по принципу поразрядного кодирования	118
4.5. Преобразователи напряжения в код, построенные по принципу параллельного сравнения	119
4.6. Основные направления совершенствования схем и принципов преобразования	121
4.7. Основные характеристики преобразователей	123
Контрольные вопросы и упражнения	126
5. СРЕДСТВА ПЕРЕРАБОТКИ ИНФОРМАЦИИ	128
5.1. Типы средств переработки информации	128
5.2. Микроконтроллеры	131
5.3. Микропроцессоры	149
5.4. Классификация вычислительных систем	154
5.5. Основные характеристики средств переработки информации ..	174
Контрольные вопросы и упражнения	180
6. СРЕДСТВА ОТОБРАЖЕНИЯ И УПРАВЛЕНИЯ	182
6.1. Средства отображения информации	182
6.2. Средства управления системой	192
6.3. Пульты управления	195
6.4. Деятельность оператора в системах управления	199
Контрольные вопросы и упражнения	205
7. ИСПОЛНИТЕЛЬНЫЕ СРЕДСТВА	207
7.1. Виды исполнительных средств	207
7.2. Основные типы силовых исполнительных органов	208
Контрольные вопросы и упражнения	215
8. СРЕДСТВА СВЯЗИ	216
8.1. Эталонная модель взаимодействия открытых систем	216
8.2. Физические среды передачи информации	221
8.3. Основные способы передачи импульсной последовательности	224
8.4. Архитектура связи между устройствами	228
8.5. Полевые шины	241
8.6. Локальные сети	247

8.7. Канальный уровень передачи	252
Контрольные вопросы и упражнения	253
9. ИСТОЧНИКИ ЭЛЕКТРОСНАБЖЕНИЯ И ХАРАКТЕРИСТИКА УСЛОВИЙ ЭКСПЛУАТАЦИИ	255
9.1. Общая организация системы электроснабжения	255
9.2. Вторичные источники электроснабжения	256
9.2. Характеристика условий эксплуатации	259
Контрольные вопросы и упражнения	262
10. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ СИСТЕМ РЕАЛЬНОГО ВРЕМЕНИ	263
10.1. Операционные системы	263
10.2. Основные свойства задач	269
10.3. Диспетчеризация задач и приоритеты	271
10.4. Принципы организации прерываний	279
10.5. Организация счета времени	281
10.6. Синхронизация задач	286
10.7. Управление оперативной памятью	289
10.8. Примеры операционных систем реального времени	290
10.9. Типовая структура алгоритмов управляющих ЭВМ	293
10.10. Программирование алгоритмов реального времени	302
10.11. Типовая схема распределения оперативной памяти	312
Контрольные вопросы и упражнения	317
ЗАКЛЮЧЕНИЕ	318
ИСПОЛЬЗОВАННАЯ ЛИТЕРАТУРА	319

1. ОСНОВНЫЕ ПОНЯТИЯ И КЛАССИФИКАЦИЯ СИСТЕМ УПРАВЛЕНИЯ

1.1. Основные понятия

Под *системой* принято понимать совокупность связанных между собой элементов. Система характеризуется *структурой* $S = \langle E, C, \lambda \rangle$, где E – множество элементов, C – множество связей между элементами, λ – соответствие между элементами.

При этом *элементом* называется часть системы, не подлежащая дальнейшему делению на части при данном уровне рассмотрения. Элементы могут представлять собой понятия, и тогда мы будем иметь дело с понятийной системой; если элементами системы являются физические объекты, то налицо техническая система. Заметим, что в определении отмечается некоторая совокупность, т.е. некоторая целостность. Нас интересуют не отдельные элементы, а все вместе, взятые в некотором единстве. Принципиальным является необходимость учета взаимодействия элементов, так как именно оно придает совокупности элементов новое качество.

Часть элементов системы, выделенная на основе какой-либо общности элементов, называется *подсистемой*.

Системы, создаваемые человеком, обладают *целью функционирования*. Именно наличие цели служит тем объединяющим фактором, на основе которого объединяются элементы. Такие системы называют *организованными* системами.

1.2. Управление как процесс принятия и реализации решений

В повседневной деятельности мы часто сталкиваемся с терминами «управлять», «регулировать» и понимаем под этими терминами задачу воздействия на параметры объекта для достижения определенных целей. Естественно называть *управлением* целенаправленное изменение параметров объекта управления.

Совокупность наиболее важных параметров, от которых зависит функционирование объекта управления, описывает его состоя-

ние. Знание состояния объекта управления – важнейший фактор, определяющий принятие решения по управлению. Другой важный фактор – возмущения, которым подвергается объект управления со стороны других объектов, состоянием которых мы не можем управлять. Множество таких объектов объединяется понятием «внешняя среда». На основе оценок состояния объекта управления и внешней среды принимается решение по управлению. Конечно, при этом учитывается цель функционирования и ресурсы, которыми мы располагаем. После того, как решение принято, оно реализуется с потреблением ресурсов. Так мы приходим к структурной схеме процесса управления, показанной на рис. 1.1.

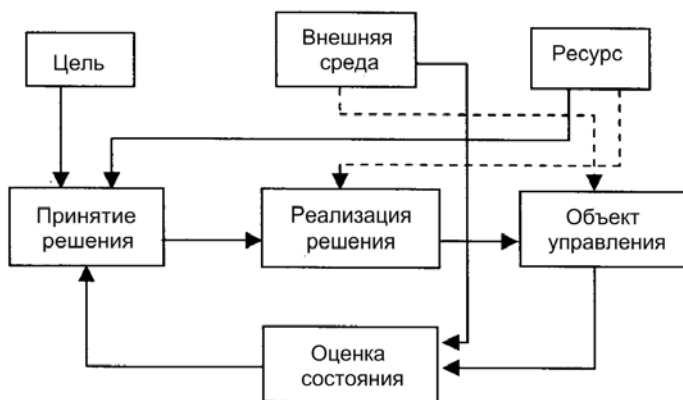


Рис. 1.1

Управление осуществляется путем выдачи на объект управляющих воздействий

$$X(t) = (x_1, x_2, \dots, x_n; t).$$

Кроме них, объект управления подвергается воздействию возмущений со стороны внешней среды; часть из них

$$Z(t) = (z_1, z_2, \dots, z_m; t)$$

измеряется, другая часть

$$F(t) = (f_1, f_2, \dots, f_s; t)$$

неизвестна (часто неизвестно также значение s).

Контролируемые параметры, по которым ведется управление,

$$Y(t) = (y_1, y_2, \dots, y_r; t),$$

называются *управляемыми*. Управляемые параметры дают обобщенную информацию о состоянии объекта управления

$$S(t) = (s_1, s_2, \dots, s_k; t).$$

Смена состояний объекта управления определяет процесс его функционирования:

$$S(t) = \varphi[X(t), Z(t), F(t), Y(t), S(t-1)].$$

Задача управления состоит в том, чтобы найти такую функцию $X(t)$, которая обеспечила бы перевод объекта в заданное конечное состояние S^* , соответствующее цели его функционирования. Этот перевод должен закончиться за конечное число шагов и требовать минимальных затрат ресурсов R .

1.3. Классификация систем управления

Тип системы зависит от того, как принимается решение по управлению. Если в процессе принятия решения участвует человек, то такая система называется *автоматизированной* (АСУ); если человек исключен из процесса принятия решения, то это – *автоматическая* система управления (САУ).

По физической природе объекта управления различают системы управления технологическими процессами, где объект управления представляет собой физический объект, и системы управления экономико-организационного типа, в которых объектом управления является хозяйственная деятельность. В последнее время создаются интегрированные системы управления, в которых присутствуют объекты обоих типов.

Другим признаком, по которому классифицируются системы, является время принятия решения. В системах управления оно зависит от динамики объекта, которая может быть различной: от миллисекунд до часов и дней. Условная граница лежит на уровне 100 мс: если время принятия решения не превышает этой величины, то говорят, что система функционирует в *реальном масштабе времени*; если превышает, то говорят о *произвольном масштабе*. Системы управления технологическими процессами всегда функционируют в реальном масштабе времени. Время в таких системах является одним из факторов, определяющих эффективность системы.

Хорошим примером системы реального времени является промышленный робот, который должен брать что-то с ленты конвейера. Объекты на конвейере движутся, и робот имеет некоторый интервал времени для того, чтобы схватить объект. Если робот опоздает, то объекта уже не будет на месте и поэтому работа будет неверной. Если робот поспешит, то объекта там еще не будет.

Принято различать системы «жесткого» и «мягкого» реального времени.

Системой «жесткого» реального времени называют систему, в которой неспособность обеспечить реакцию на какие-либо события в заданное время является отказом и ведет к невозможности решения поставленной задачи. В качестве условной временной границы допустимого времени реакции обычно принимают 100 мкс. В жесткой системе:

- никакое опоздание неприемлемо ни при каких обстоятельствах;
- результат, выданный с опозданием, бесполезен;
- нарушение крайнего срока времени отклика рассматривается как катастрофический отказ;
- цена превышения заданного времени отклика бесконечно велика.

Точного определения для «мягкого» реального времени не существует, но принято считать, что она иногда может не успевать делать все, что надо, в установленные сроки. В мягкой системе:

- возрастает цена за опоздание результата;
- критическим фактором является низкая производительность, а не опоздания.

В терминах вероятностей эти определения могут быть записаны следующим образом:

$$P\{|t - t_0| > \Delta t\} > 0$$

– для систем «мягкого» реального времени,

$$P\{|t - t_0| > \Delta t\} = 0$$

– для систем «жесткого» реального времени (здесь t – фактический момент выдачи управляющего воздействия, t_0 – заданный момент выдачи управляющего воздействия, Δt – допустимая погрешность).

В зависимости от функций, выполняемых системой управления, различают виды (формы) автоматизации: управление, регулирование, контроль, защиту и блокировку.

Управление представляет собой принятие решений и выдачу на объект управления совокупности воздействий, выполняемых на основании информации о ходе технологического процесса, с целью поддержания требуемого технологического режима или улучшения его параметров. Обычно предполагается, что управление подразумевает поиск и реализацию оптимального (по определенному критерию) решения.

Регулирование представляет собой принятие решений и выдачу на объект управления совокупности воздействий, выполняемых на основании информации о ходе технологического процесса, с целью поддержания параметров производственных процессов постоянными или изменяющимися по заданному закону.

Контроль – это наблюдение за параметрами объекта управления с целью определения его состояния. Параметры, подлежащие контролю, в зависимости от их физической природы различны (температура, давление, расход топлива, число оборотов, сила тока и т.п.). Контроль может быть местным и дистанционным.

Местный контроль дает возможность наблюдать за состоянием параметра непосредственно в контролируемой точке. При дистанционном контроле за состоянием параметров можно следить на расстоянии от контролируемой точки.

Защита – это комплекс средств и мероприятий по предохранению агрегатов и установок при нарушениях технологических режимов.

Блокировка – это комплекс средств и мероприятий, предохраняющих участок установки или агрегат от неправильных операций вследствие невнимательности оператора или ошибочных команд. Различают две группы блокировки: запретно-разрешающую и аварийную.

Запретно-разрешающие блокировки предотвращают неправильные включения и отключения механизмов, а также нарушение установленной технологическими требованиями очередности пуска и остановки различных механизмов.

Аварийные блокировки предназначены для автоматического последовательного отключения (включения) механизмов в соответствии с режимом работы агрегата, подвергшегося аварийному отключению.

Блокировочные устройства применяют также в случаях, когда требуется исключить одновременную подачу команд противоположного знака на один и тот же регулирующий орган из разных мест или от автоматического и дистанционного управления, когда оба вида управления действуют параллельно.

1.4. Структура систем реального времени

Основными составляющими автоматизированных систем реального времени (АСУ ВР) являются комплекс технических средств, программное и информационное обеспечение.

Для принятия решения по управлению нужно знать о состоянии объекта управления и возмущениях со стороны внешней среды. Эта информация вводится в систему с помощью *измерительных средств*, которые являются источниками информации для *средств переработки информации*. Эти средства реализуют алгоритм принятия решения и формируют управляющие воздействия, которые передаются на *исполнительные средства*, непосредственно связанные с изменяемыми параметрами объекта управления.

Для того чтобы человек участвовал в процессе принятия решения, ему также нужно дать возможность следить за состоянием объекта и внешней среды. Для этого он снабжается *устройствами отображения информации*, которые предназначены для преобразования информации в удобную для восприятия форму (чаще всего – в визуальную). Человек может вмешиваться в управление, воздействуя на исполнительные средства и на средства переработки информации с помощью *управляющих средств*.

Поскольку части системы могут находиться на значительных расстояниях друг от друга, часть связей представляет собой *средства передачи данных* – специальные технические устройства, предназначенные для этих целей. Еще одна составляющая комплекса технических средств – *система энергоснабжения*; ее назна-

чение – снабжение всех технических средств электроэнергией заданного напряжения и мощности.

Все эти составляющие комплекса технических средств показаны на рис. 1.2.

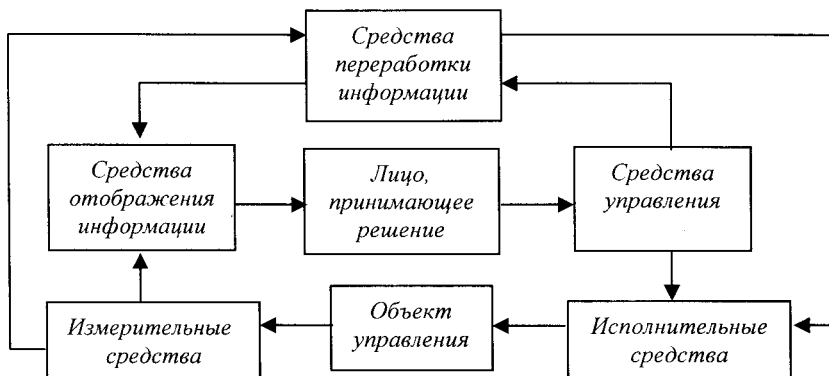


Рис. 1.2

В качестве средств переработки информации обычно применяется узел на основе вычислительной машины. Тип и структура этого узла может быть различной – от простейшего контроллера до кластера, однако в любом случае его применение связано со следующими особенностями:

- он должен «следить» за множеством параллельно протекающих процессов;
- он должен обрабатывать запросы, поступающие в произвольные моменты времени;
- допустимое время принятия решения обычно соизмеримо с временем реализации алгоритма выработки такого решения;
- он должен удовлетворять повышенным требованиям по надежности и достоверности информации;
- состав задач, которые решает этот узел, заранее известен и программное обеспечение для их решения отлажено.

Эти особенности учитываются при построении *программного обеспечения*. Оно включает в себя операционную систему, программы решения функциональных задач и программы контроля и обеспечения устойчивости вычислительного процесса.

Главное назначение операционной системы – обеспечение параллелизма и обработка заявок, поступающих в случайные моменты времени. Программы решения функциональных задач выполняют основное целевое назначение системы – управление объектом. Третья группа программ контролирует работу системы.

Информационное обеспечение содержит данные, необходимые для управления системой.

Контрольные вопросы и упражнения

1. Сформулируйте две постановки задачи выбора оптимального варианта построения системы.
2. Чем управление отличается от регулирования?
3. Приведите примеры систем «жесткого» и «мягкого» реального времени.
4. Укажите на рис. 1.2 связи, где не могут использоваться средства передачи данных.

2. ДАТЧИКИ

2.1. Структура измерительного канала

Формирование сигнала и преобразование его в данные – основная функция измерительного канала. Рассмотрим его типовую структуру.

Информация о некотором параметре исследуемого процесса, поступающая на вход измерительного канала, прежде всего, должна быть нанесена на носитель и изменить его состояние – должен сформироваться сигнал. Для удобства последующего преобразования лучше всего сформировать электрический сигнал. В качестве носителя информации в таком сигнале используются заряд, ток, напряжение или импеданс. Формирование электрического сигнала – функция *первичного измерительного преобразователя* (ПИП).

При измерениях некоторых неэлектрических величин не всегда удается преобразовать их непосредственно в электрическую величину. В этих случаях осуществляют двойное преобразование исходной (первичной) измеряемой величины: сначала с помощью чувствительного элемента – в промежуточную неэлектрическую величину, а затем с помощью измерительного преобразователя – в выходную электрическую. В качестве чувствительного элемента могут использоваться сильфоны, мембраны и т.д. Конструктивную совокупность чувствительного элемента и первичного измерительного преобразователя, размещаемых непосредственно у объекта измерений и преобразующих измеряемые параметры в электрические величины, удобные для дальнейшего преобразования и передачи по каналам связи, называют *датчиком*.

Характерная особенность АСУ РВ – многообразие типов сигналов на входах измерительных каналов. Для удобства построения систем все они приводятся к единому нормированному виду с помощью *унифицирующего (нормирующего) измерительного преобразователя* (УИП).

Унифицированный сигнал – это сигнал определенной физической природы, изменяющийся в определенных фиксированных

пределах независимо от вида измеряемой величины, метода и диапазона ее измерения.

Сигнал с выхода УИП передается на *аналоговый фильтр низких частот* (АФ), применение которого в измерительном канале базируется на том факте, что амплитудно-частотная характеристика идеального низкочастотного фильтра имеет ярко выраженный нелинейный характер: $k(\omega) = 1$ при $\omega < \omega_0$ и $k(\omega) = 0$ при $\omega > \omega_0$. За счет этого частоты, большие частоты ω_0 , через такой фильтр не проходят, так что спектральная плотность мощности процесса, поступающего на последующие каскады преобразования $S(\omega) \approx 0$ при $\omega > \omega_0$. Это используется для решения двух задач.

1. В процессе передачи и преобразования сигнал подвергается воздействию шумов. Они порождаются как внутренними причинами (например, тепловые флуктуации электронных взаимодействий), так и внешними (например, воздействие посторонних электромагнитных полей на линию связи). Поэтому следует учитывать, что во всех этапах преобразования участвуют одновременно два процесса – сигнал как функция времени, подчиняющаяся определенным статистическим закономерностям, и помеха (шум), которую обычно представляют в виде случайного процесса типа «белого шума». Взаимодействие этих двух процессов может иметь сложный характер. Простейший случай: когда два случайных процесса – сигнал и помеха – образуют аддитивную смесь (суммарный случайный процесс). Если они различаются по спектру (в частотном представлении) и спектр помехи лежит выше (на оси частот) спектра сигнала, имеет смысл устранить или уменьшить ее до окончательного преобразования в данные. С этой целью в состав тракта вводят аналоговые фильтры низкой частоты.

2. При последующей дискретизации аналогового сигнала в спектральной области возникает бесконечное число копий его спектральной плотности мощности, отстоящих друг от друга на расстояние $\Delta\omega = 2\pi/T$ (T – интервал дискретизации). При больших значениях T (что оправдано с точки зрения инженерной практики) $\Delta\omega$ мало, и высокие частоты одной копии совмещаются с низкими частотами другой, что приводит к искажению сигналов. Ограничение спектра частотой ω_0 исключает такое наложение. Несмотря на

возможное при этом искажение сигнала, этот прием используют для того, чтобы уменьшить частоту дискретизации и избежать при этом эффекта наложения спектров.

После фильтрации унифицированные по информативному параметру непрерывные величины поступают на вход *преобразователя «непрерывная величина – код»* (ПНК), функциональное назначение которого ясно из названия. Сначала сигналы дискретизируются, превращаясь в ступенчатые сигналы тока или напряжения, а затем квантуются, превращаясь в данные.

Часто с целью более рационального использования технических средств для ввода информации применяют многоканальные системы, стремясь, по возможности, использовать общие для всех каналов ресурсы. Обычно в этом качестве выступает ПНК. Для разделения этого ресурса во времени на его входе устанавливается коммутатор – *мультиплексор аналоговых каналов* (МАК).

Кроме перечисленных элементов, в состав измерительного тракта входят и другие, вспомогательные элементы: буферные регистры, схемы памяти, источники управляющих сигналов, система энергоснабжения и т.п.

2.2. Классификация преобразователей

Важнейшая функция системы сбора информации заключается в восприятии сведений о состоянии объекта или внешней среды и их обработке для ввода в ЭВМ. Технические средства и человек могут воспринимать только ту информацию, которая «материализована», т.е. превращена в доступное для измерений изменение параметров какой-либо физической среды или объекта. В качестве такой среды можно рассматривать воздушную среду, электрическое поле, магнитное поле; объектом может служить столбик ртути, кристалл пьезоэлемента и т.д. Параметрами могут служить цвет, геометрические размеры, величина напряжения и другие характеристики объектов. Эти величины принято называть *носителями информации*. Носитель с представленной на нем информацией называют *сигналом*. Таким образом, сигнал есть физическая величина, отображающая информацию.

О количестве и разнообразии входных физических величин можно судить по тому факту, что система единиц СИ включает в себя 6 единиц пространства и времени, 14 механических, 40 электрических и магнитных, 11 тепловых, 15 световых, 14 акустических, 11 единиц молекулярной физики и физической химии.

В настоящее время существует примерно следующее распределение доли измерений различных физических величин в промышленности: температура – более 50 %, расход вещества и механические величины (перемещение, сила, давление и др.) – по 15 %, количество, время и состав вещества – по 5 %, электрические и магнитные величины – менее 5 %.

В 1960 г. было принято решение о создании Государственной системы промышленных приборов и средств автоматизации (ГСП), которой были унифицированы основные конструкции датчиков. В ГСП все контролируемые величины разбиты на пять следующих групп: теплоэнергетические, электроэнергетические, механические величины, химический состав и физические свойства.

Теплоэнергетические величины: температура, давление, перепад давлений, уровень и расход.

Электроэнергетические величины: постоянные и переменные ток и напряжение, мощность (активная и реактивная), коэффициент мощности, частота и сопротивление изоляции.

Механические величины: линейные и угловые перемещения, угловая скорость, деформация, усилие, вращающие моменты, число изделий, твердость материалов, вибрация, шум и масса.

Химический состав: концентрация, состав, химические свойства.

Физические свойства: влажность, электропроводность, плотность, вязкость, освещенность и др.

Устройства, в которых первично преобразуется измеряемая физическая величина, называют *первичными измерительными преобразователями* (ПИП). Измерительные преобразователи (ИП) работают с естественным и унифицированным выходными сигналами.

Естественный выходной сигнал формируется первичными ИП естественным путем и может представлять собой угол поворота, перемещение, усилие, сопротивление, напряжение, емкость и т.д.

Унифицированный сигнал – это сигнал определенной физической природы, изменяющийся в определенных фиксированных пределах независимо от вида измеряемой величины, метода и диапазона ее измерения. Для получения унифицированных аналоговых сигналов применяют ИП, называемые *нормирующими*.

Измерительные преобразователи можно классифицировать:

- по виду измеряемой физической величины – различают ИП линейных и угловых перемещений, давления, температуры, концентрации вещества и т.д.;
- по виду используемой энергии – электрические, механические, пневматические и гидравлические;
- по соотношению между входной и выходной величинами:
 - преобразование неэлектрической величины в неэлектрическую (рычаги, редукторы, мембраны, пружины и т.д.);
 - преобразование неэлектрической величины в электрическую (потенциометры, термопары, емкостные и индуктивные ИП и др.);
 - преобразование одной электрической величины в другую (датчики тока, напряжения, фазочувствительные схемы и усилители);
- по виду выходного сигнала – аналоговые (потенциальные, токовые, частотные, фазовые), дискретные (амплитудно-, время- и числоимпульсные и др.), релейные, с естественным или унифицированным выходным сигналом;
- по характеру преобразования входной величины в выходную – параметрические, генераторные, компенсационные, частотные и фазовые.

Параметрические преобразователи – это преобразователи, в которых изменение входной неэлектрической величины преобразуется в изменение какого-либо электрического параметра выходной цепи (активного сопротивления, индуктивности, емкости). В них для получения сигнала требуется внешний источник энергии.

Генераторные преобразователи – это преобразователи, в которых входная величина преобразуется в ЭДС на выходе (датчики термоЭДС, пьезоэлектрические, фотоэлектрические, тахометрические и др.). В них формирование сигнала осуществляется за счет энергии самого сигнала.

Компенсационные преобразователи – это преобразователи, в которых входная величина (часто после предварительного преобразования) компенсируется другой величиной, имеющей ту же физическую природу. Для непрерывной компенсации осуществляется отрицательная обратная связь.

Частотные и *фазовые* преобразователи – это преобразователи, в которых различные физические величины на входе (перемещение, скорость, расход) изменяют частоту переменного тока, частоту следования импульсов или фазу.

ИП можно также классифицировать по конструктивному исполнению, по величине погрешности и по другим признакам.

Наибольшее распространение в системах автоматизации получили параметрические преобразователи неэлектрических величин в электрические. В общем случае такой преобразователь состоит из первичного измерительного преобразователя ПП (чувствительного элемента), преобразующего контролируемую величину x в величину x_1 , удобную для измерения, и измерительного преобразователя П, в котором величина x_1 преобразуется в электрический сигнал u за счет подводимой извне энергии.

Устройство, которое, подвергаясь воздействию измеряемой физической величины, выдает эквивалентный сигнал, являющийся функцией измеряемой величины, называют **датчиком**. Другими словами, датчик преобразует один вид носителя информации в другой, обычно – в электрическую величину (напряжение, ЭДС или ток) или в параметр электрической цепи (импеданс, частота, фаза).

2.3. Метрологические характеристики датчиков

Датчик является основным источником электрического сигнала, тогда как остальная часть цепи должна обеспечить его обработку и использование. Следовательно, от качества датчика в первую очередь зависят как соответствие между истинным значением измеряемого параметра и полученным при измерении значением, так и пределы вносимых в полученную величину погрешностей.

Основными характеристиками датчиков являются статические, динамические и эксплуатационные. Статические характеристики относятся к установившемуся режиму, динамические определяют

его поведение при изменениях входной величины, эксплуатационные определяют характеристики эксплуатации (объем, габаритные размеры, потребляемая мощность, достоверность информации, надежность работы, потребляемая мощность, срок службы, условия эксплуатации, стоимость).

2.3.1. Градуировка датчиков

Градуировка датчика охватывает совокупность операций, позволяющих установить соответствие между значениями измеряемой величины и электрическими величинами на выходе с учетом всех дополнительных факторов, которые могут изменить выходной сигнал датчика. Основная характеристика $y = \varphi(x)$ – зависимость выходной величины от входной в установившемся режиме – представляется либо в графической, либо в алгебраической форме (*градуировочная* или *характеристическая* кривые датчика соответственно). К дополнительным факторам, влияющим на выходной сигнал, относятся:

а) связанные с измеряемой величиной физические величины, к которым чувствителен датчик (знак и скорость ее изменения, физические свойства ее материального носителя);

б) независимые от измеряемой величины физические факторы, воздействию которых подвержен работающий датчик;

в) параметры окружающей среды (температура, влажность) или параметры, связанные с питанием (амплитуда и частота напряжения).

По характеристической кривой датчика можно определить такие важнейшие статические характеристики, как чувствительность и линейность.

В общем случае *чувствительность* S определяется для диапазона около некоторого постоянного значения измеряемой величины x как отношение приращения сигнала Δs на выходе к изменению Δx измеряемой величины, которое вызвало это приращение.

Чувствительность может зависеть от размера датчика или особенностей его устройства, от амплитуды и частоты напряжения питания, от температуры окружающей среды. Важным фактором,

влияющим на чувствительность датчика, является частота измененной измеряемой величины. В зависимости от частоты различают два режима работы датчиков, с которыми связывают соответствующие параметры чувствительности и способы градуировки.

Статический режим. Измеряемая величина постоянна или меняется медленно. При градуировке устанавливается ряд последовательно возрастающих постоянных значений измеряемой величины m_i и определяются соответствующие им после достижения установившихся значений электрические сигналы s_i . Отношение величины s_i на выходе к соответствующему значению m_i измеряемой величины называют *статическим коэффициентом преобразования*.

Динамический режим. Измеряемая величина изменяется быстро, в установившемся режиме выходной сигнал s имеет ту же периодичность, что и измеряемая величина.

Пусть измеряемая величина описывается выражением

$$m(t) = m_0 + m_1 \cos \omega t,$$

где m_0 – постоянная составляющая, на которую накладывается синусоидальная переменная с амплитудой m_1 и частотой $f = \omega/2\pi$. Выходной сигнал датчика при этом приобретает форму

$$s(t) = s_0 + s_1 \cos(\omega t + \varphi),$$

где s_0 – постоянная величина, соответствующая m_0 (обе эти величины определяют рабочую точку Q_0 на статической градуировочной кривой); s_1 – амплитуда переменной составляющей выходного сигнала, возникающая под действием переменной составляющей измеряемой величины; φ – сдвиг фазы между изменениями на выходе и на входе. Чувствительность определяется в этом случае выражением

$$S(\omega, Q_0) = s_1(\omega)/m_1(\omega),$$

где величина $S(\omega)$ является *частотной характеристикой* датчика. Ее обычно определяют, рассматривая в совокупности датчик и элементы измерительной цепи, непосредственно с ним связанные. В динамическом режиме при изменении измеряемой величины с частотой ω чувствительность меняется по закону

$$S(\omega) = \frac{S(0)}{\sqrt{1 + (\omega / \omega_c)^2}},$$

где ω_c – частота среза, или граничная частота.

Таким образом, линейность в динамическом режиме зависит от чувствительности статического режима $S(0)$ и параметров частотной характеристики, которые не зависят от значений измеряемой величины в диапазоне, где чувствительность $S(0)$ постоянна.

Чувствительность характеризуется величиной *порога чувствительности* – максимальным изменением измеряемой величины, которое еще не вызывает обнаруживаемого изменения выходного сигнала датчика.

Линейность оценивается величиной разности между фактической и линейной зависимостью выходной величины от входной в процентах от максимального значения измеряемой величины в данном диапазоне измерений.

Говорят, что система линейна в определенном диапазоне измеряемых величин, если ее чувствительность не зависит от значения измеряемой величины, т.е. остается постоянной. В случае линейности характеристической кривой датчика значительно упрощается последующая обработка результатов измерений. Преобразование нелинейной характеристики датчика в линейную называют *линеаризацией*. Этот процесс направлен на то, чтобы сделать сигнал пропорциональным изменениям измеряемой величины.

Линейность в статическом режиме определяется наличием и диапазоном прямолинейного участка статической характеристики; для этого вводят *диапазон измерения* – разность между максимальным и минимальным значением измеряемой величины, где сохраняется паспортная характеристика линейности.

В динамическом отношении датчики представляют собой инерционные или колебательные звенья. Большинство датчиков по характеру динамических свойств относят к безынерционным и апериодическим звеньям первого и более высоких порядков.

Поведение датчика при быстрых изменениях входной величины определяется его динамической характеристикой – кривой переходного процесса при скачкообразном изменении входной величины. По ней можно определить *быстродействие* датчика – параметр, позволяющий оценить, как выходная величина следует во времени за изменениями измеряемой величины.

Параметр, используемый для количественного описания быстрого действия, – это *время установления*, т.е. интервал времени, который должен пройти после ступенчатого изменения измеряемой величины, чтобы сигнал на выходе датчика достиг уровня, отличающегося на определенную величину ε (%) от установившегося значения. Обычно время установления характеризуют *постоянной времени* – промежутком времени, за который выходная величина достигает уровня 0,63 от установившегося значения при ступенчатом изменении входного сигнала.

Для ориентировочных оценок можно принимать, что постоянная времени системы $\tau = 1/2\pi f_c$, а время нарастания или убывания сигнала $t = 2,2\tau = 0,35/f_c$.

Еще одной динамической характеристикой служит полоса пропускания – диапазон частот, для которого чувствительность $S > S_{\max}/\sqrt{2}$.

Различают простую и комплексную градуировку датчиков.

Простая градуировка проводится, когда измеряемая величина определяется единственным физическим параметром, а датчик не чувствителен к влияющим на измерения величинам или не подвержен их воздействию. В этих условиях градуировка состоит в установлении связи точно известных измеряемых величин с соответствующими электрическими величинами на выходе. Градуировка осуществляется одним из следующих способов.

Прямая или абсолютная градуировка. Различные значения измеряемых величин задаются эталонами или образцовыми средствами, имеющими существенно более высокую (на два порядка) точность по сравнению с точностью измерений градуируемого датчика.

Косвенная или сравнительная градуировка. Используется образцовый датчик, градуировочная кривая которого известна, а ее стабильность высока. Образцовый и градуируемый датчики в одинаковых условиях подвергаются, по возможности одновременно, действию одинаковых измеряемых величин, значения которых определяются образцовым датчиком.

Комплексная градуировка требуется, когда воздействие на датчик измеряемой величины без учета дополнительных условий

не позволяет с достаточной точностью провести градуировку. В этом случае требуется такая процедура, при которой датчик подвергается воздействию упорядоченной последовательности номинальных значений измеряемой величины. Эта процедура состоит в следующем: а) датчик устанавливается в нуль; б) определяют значения выходного сигнала сначала при возрастании измеряемой величины, а затем – при ее убывании, изменяя, во-первых, частоту входного воздействия (при постоянной амплитуде) и, во-вторых, его амплитуду (при фиксированной частоте, выбранной внутри полосы пропускания).

Сходную процедуру применяют при наличии влияющих величин. Если, например, реакция датчика зависит от температуры, то проводят серию градуировок (каждую — при постоянной температуре); значения температур выбирают так, чтобы они были распределены во всем диапазоне возможных применений датчика.

2.3.2. Погрешности измерений

Воздействие на датчик определяется истинным значением измеряемой величины, а экспериментатор воспринимает общую реакцию всей измерительной цепи. Разность между измеренным и истинным значениями величины называется *погрешностью измерений*.

Погрешности из-за несовершенства средств измерений составляют группу *инструментальных* погрешностей. Обычно они подразделяются на основные и дополнительные.

Основная погрешность – это максимальная разность между измеренным значением выходного сигнала и его истинным значением, определяемым по идеальной статической характеристике для данной входной величины при нормальных эксплуатационных условиях, т.е. при номинальных значениях всех величин, влияющих на результат измерения (температуры, влажности, давления, напряжения питания и т.п.). За номинальные внешние условия обычно принимаются следующие: температура 20 ± 5 °С, атмосферное давление 760 ± 20 мм рт.ст., относительная влажность 60 ± 20 %.

Дополнительная погрешность – это погрешность, вызываемая условиями внешней среды и внутренними процессами во внутренних деталях датчика; она возникает при отличии значений влияющих величин от номинальных.

По характеру изменения погрешности при повторных измерениях различают систематические и случайные погрешности.

Систематические погрешности остаются постоянными или закономерно изменяются при повторных измерениях одной и той же величины. Они вносят постоянное расхождение между истинным и измеренным значениями величины. Частные случаи возникновения систематических погрешностей датчиков сводятся к следующим:

- *погрешности значения опорной величины*, такие, как смещение нуля усилителя, ошибка при определении опорной температуры термодпары, неточная величина напряжения питания. Они могут быть уменьшены путем тщательной проверки и настройки используемой аппаратуры;

- *погрешности, связанные с определением характеристик датчика* (например, связанные с градуировочной кривой) выявляются при выборочном измерении линейности или коэффициента преобразования, когда он измеряется для одного датчика из партии, а результаты распространяются на всю партию. Еще один источник этих погрешностей – старение или ухудшение параметров датчика под воздействием вредных условий окружающей среды или измеряемого процесса;

- *погрешности, связанные со способом измерения*. Присутствие датчика может изменить измеряемую величину, что приведет к искажению результата. Другая возможная причина подобных погрешностей – завершение измерений до момента достижения установившегося режима. Сюда же можно отнести такую причину, как неправильное введение поправок в результаты измерений (когда не учитываются нелинейность характеристик, самонагрев терморезистора, теплопроводность корпуса датчика и др.).

Систематические погрешности характеризуются величинами разности между результатом измерения a и истинным значением измеряемой физической величины a_0 , выраженной в виде *абсолютной* $\Delta a = a - a_0$ или *относительной* $\Delta a/a = (a - a_0)/a$ ошибок.

Случайные погрешности в момент измерения неизвестны, они изменяются случайным образом при повторных измерениях одной и той же величины. Частные случаи подобных погрешностей сводятся к следующим:

- *погрешности, связанные с собственными параметрами измерительной аппаратуры.* К ним можно отнести флуктуации порога чувствительности датчика и погрешности гистерезиса, которую определяют как половину максимальной разности выходных сигналов, соответствующих одной и той же измеряемой величине, когда она получена в процессе возрастания и убывания этой величины. Наличие подобной погрешности делает выходной сигнал в определенной мере зависимым от предшествующих измерений. К этой же группе можно отнести погрешности из-за воздействия внутренних шумов, таких, например, как шумы, возникающие в результате теплового возбуждения носителей заряда в резисторах или активных элементах;

- *погрешности из-за воздействий внешней среды* (наводки от электромагнитных полей, флуктуации напряжения питающих устройств и т.п.).

Величина случайных погрешностей может быть уменьшена с помощью соответствующих устройств (термостабилизация, антивибрационные основания, стабилизация напряжения питания, заземление экранов и установок, фильтры и т.д.) или экспериментальных методов (статистическая обработка, дифференциальные измерения, применение двухпроводных линий).

При анализе распространения случайных ошибок в последовательности n измерений пользуются средними значениями

$$\bar{a} = \sum_{i=1}^n a_i / n \text{ и дисперсиями } \sigma_a^2 = \sum_{i=1}^n (a_i - \bar{a})^2 / n - 1 \text{ отдельных ре-}$$

зультатов a_i .

Датчики, применяемые в системах электроавтоматики, должны соответствовать определенным требованиям: иметь необходимый диапазон изменения входных и выходных сигналов и линейность статических характеристик; достаточную чувствительность; малые инерционность и погрешность; непрерывную зависимость выход-

ной величины от входной; достаточную мощность выходного сигнала; наименьшее влияние датчика на измеряемый параметр; малые габариты и массу, работать в заданных условиях окружающей среды; обладать высокой надежностью работы.

2.4. Первичные измерительные преобразователи (чувствительные элементы)

Большинство неэлектрических величин, которые измеряются и регулируются в системах автоматики, удобно преобразовать в механическое перемещение, а затем механическое перемещение преобразовать в электрический сигнал.

Давление определяется по упругой деформации сильфона, мембраны или трубчатой пружины, соответственно для малого, среднего и большого давления.

Сильфон представляет собой тонкостенную металлическую оболочку с поперечной гофрированной боковой поверхностью. Сильфон расширяется или сжимается подобно пружине вдоль оси под действием разности давления внутри и снаружи.

Мембрана – это закрепленная по контуру обычно круглая пластина, которая изгибается под действием давления газа или жидкости.

Трубчатая пружина изготавливается из металла (латунь, сталь и др.), обычно в поперечном сечении имеет овальную форму. Под действием избыточного давления трубчатая пружина стремится разогнуться.

Температура определяется по изгибу биметаллической пружины, давлению жидкости или газа, заполняющего баллон термометра. *Биметаллическая пружина* представляет собой двухслойную пластину, изготовленную из металлов с разными коэффициентами линейного расширения. При нагревании пластина изгибается в сторону того слоя, материал которого имеет меньший коэффициент.

Уровень и расход жидкости определяются по перемещению *поплавка*. Так, при движении жидкости снизу вверх в трубке переменного сечения поплавок поднимается, чтобы обеспечить большее поперечное сечение для прохода жидкости. Чем больше рас-

ход, тем выше должен подняться поплавок. Плотность жидкости определяется по глубине погружения поплавка.

Усилие определяется по величине упругой деформации пластины или витой пружины.

2.5. Датчики с бинарным и цифровым выходным сигналом

Датчики с бинарным выходным сигналом – это датчики параметрического типа, в которых дискретно изменяется электрическое сопротивление при изменении той или иной механической величины: при достижении измеряемой величиной определенного значения замыкаются или размыкаются электрические контакты, включенные в те или иные цепи, сигнализирующие, что перемещение больше или меньше определенного значения. Поэтому эти датчики называются *контактными* и применяются в основном в системах автоматического контроля и сортировки размеров.

Для уменьшения погрешностей используют рычажную систему на плоских пружинах, не имеющих зазоров между отверстием и осью.

Материалы для контактов выбираются в зависимости от контактного давления и условий работы датчика. У высокочувствительных маломощных контактных датчиков давление на контактах изменяется от 0,001 до 0,02 Н. Контакты таких датчиков выполняются из драгоценных металлов (платины, золота и их сплавов). При контактных давлениях 0,05...1 Н применяют серебряные контакты. У мощных контактных датчиков контактные усилия составляют несколько ньютонов, и в качестве материала контактов используют вольфрам, молибден и их сплавы, обладающие высокой твердостью и износостойкостью.

Контактные датчики применяются для определения положения при механических перемещениях, для подсчета элементов в дискретных потоках, для контроля достижения предельного уровня или давления или крайних положений подвижных частей.

В качестве *датчиков положения* используются выключатели. Они состоят из электрических контактов, которые механически

размыкаются или замыкаются, когда какая-либо переменная (положение, уровень) достигает определенного значения.

Простейшим выключателем является механический нормально разомкнутый однополюсный выключатель, показанный на рис.2.1: когда выключатель разомкнут, с резистора снимается напряжение +5 В, воспринимаемое схемой как одно из логических состояний. Если контакт замкнут, выходной сигнал равен потенциалу «земли», что воспринимается как другое логическое состояние.

Замыкание механического выключателя обычно вызывает проблемы, поскольку контакты вибрируют («дребезжат») несколько (порядка пяти) миллисекунд, прежде чем замкнуться. Применение цепи, обеспечивающей небольшое запаздывание выходного сигнала, является одним из способов преодоления эффекта дребезга.

Размыкание выключателя также связано с искровым разрядом, возникающим при разрыве цепи с индуктивностью. На рис. 2.2 показаны две схемы, используемые для борьбы с этим явлением. Первая схема (рис. 2.2, а) используется в цепях постоянного тока, вторая (рис. 2.2, б) – в цепях переменного тока.

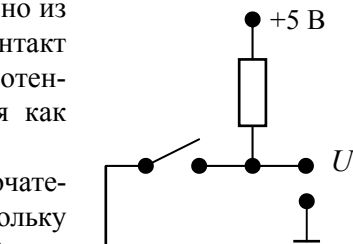


Рис. 2.1

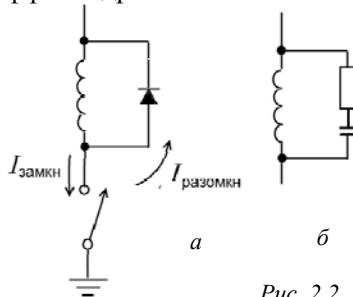


Рис. 2.2

Однополюсный двухпозиционный выключатель может быть типа «разрыв перед замыканием» и «замыкание перед разрывом» (рис. 2.3.).



Рис. 2.3

При переключении в первом случае (рис. 2.3, а) оба контакта разомкнуты на короткое время, во втором (рис. 2., б) через оба контакта ток кратковременно протекает. Борьба с эффектом дребезга контактов здесь ведется с помощью триггерных схем.

Существуют другие методы определения положения с помощью контактных датчиков, некоторые из которых приведены ниже.

Ртутные выключатели состоят из небольших герметически запаянных стеклянных трубок с контактными выводами. Трубка содержит достаточное количество ртути, чтобы замкнуть контакты. Выключатель размыкает и замыкает контакты при изменении положения трубки.

Магнитоуправляемое герметичное язычковое реле – геркон – состоит из двух плоских пружин, запаянных в небольшую стеклянную трубку. Свободные концы пружин находятся друг над другом с очень небольшим зазором между ними. Когда к трубке приближается магнит, пружины намагничиваются в разных направлениях, притягиваются друг к другу и замыкаются.

Достоинства контактных датчиков: простота и дешевизна конструкции, простота регулировки чувствительности, высокая точность, возможность работы в цепях постоянного и переменного тока. К недостаткам относят трудность обеспечения высокой надежности (из-за наличия искрения) и возможность ложных срабатываний при наличии вибраций и ударной нагрузки.

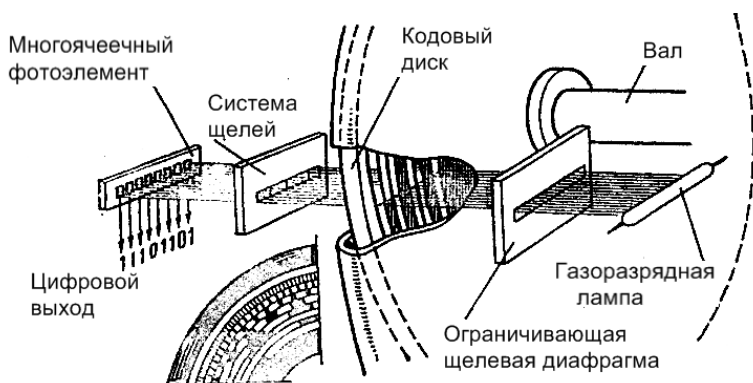


Рис. 2.4

Существует целый ряд датчиков, непосредственно преобразующих входной сигнал в цифровой код. Примером может служить датчик угла поворота, изображенный на рис. 2.4. Кодовый диск разделен на дорожки, которые подсвечиваются газоразрядной лампой. Дорожки имеют прозрачные и непрозрачные сектора. Общий световой поток модулируется этими секторами, в результате чего на многоточечный фотозаэлемент попадают отдельные потоки, несущие цифровой сигнал (рис. 2.5).

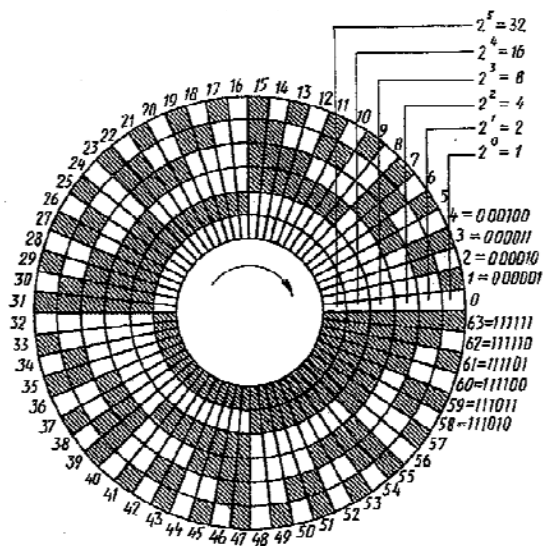


Рис. 2.5

2.6. Потенциометрические датчики

Резисторный потенциометр со скользящим контактом – один из наиболее простых и эффективных преобразователей перемещения. Он представляет собой резистор постоянного сопротивления, выполненный из высокоомной проволоки или пленочного проводящего пластика, по которому перемещается скользящий электрический контакт. Этот контакт механически соединяется с деталью, перемещения которой датчик должен измерять; от нее скользящий контакт электрически изолируется. Величина сопротивления R ме-

жду скользящим контактом и одним из концов резистора является функцией, с одной стороны, положения скользящего контакта и, с другой стороны, конструкции резистора.

В зависимости от формы резистора и, следовательно, от траектории перемещения скользящего контакта, различают потенциометры линейного и углового перемещения. Если резистор имеет однородную конструкцию, то потенциометр линеен. Значение сопротивления потенциометра чаще всего заключено в интервале от 20 Ом до 200 кОм. Для обмотки используют провод из константана, манганина, нихрома, а также сплавы из платины, серебра и т.п. Проводящий пластик имеет более высокую температурную зависимость по сравнению с проволокой.

Функция преобразования потенциметрического датчика имеет вид:

$$U_{\text{вых}} = (U/l)X = kX,$$

где $U_{\text{вых}}$ – выходное напряжение; U – питающее напряжение; l – длина намотки обмотки; X – перемещение движка реостата; k – коэффициент преобразования.

Рабочий диапазон потенциометра – это область изменения x , в которой $R(x)$ является линейной функцией перемещения. Нелинейность составляет 0,1...1 % для проволочных сопротивлений и около 0,05 % для проводящего пластика. Разрешение определяется максимальным перемещением, необходимым для перехода скользящего контакта из определенного положения в ближайшее соседнее; оно зависит от формы и размеров проволоки и скользящего контакта и может меняться по мере их износа. Для проволочных сопротивлений оно составляет 1,5...0,5 мм. Разрешение датчика можно улучшить по сравнению с разрешением потенциометра, если посредством рычага или зубчатой передачи увеличить масштаб перемещения скользящего контакта по сравнению с перемещением точки касания датчика.

Ширина контактной полосы, по которой перемещается движок, в 2...3 раза превышает диаметр провода. Ее получают полировкой вдоль витков тонкой наждачной бумагой.

Подвижный контакт (щетка, движок) выполняют в виде двух–трех параллельных проволочек диаметром 0,1...0,2 мм, несколько

отличающихся по длине, или из специально профилированной пластины с разрезами. Максимально допустимая скорость движения скользящего контакта определяет максимальную частоту перемещений, при которой еще допустимы измерения с помощью потенциометра; эта частота обратно пропорциональна амплитуде перемещений и имеет порядок кГц. Срока службы потенциометра зависит от его типа и лежит в пределах от 10^6 до 10^8 перемещений.

Для получения выходного сигнала, изменяющегося по определенному закону, применяют функциональные потенциометрические датчики. В этих датчиках зависимость сопротивления обмотки от перемещения движка является нелинейной. Требуемая нелинейность обеспечивается различными способами: изменением профиля каркаса; изменением материала или размера провода; изменением шага намотки или длины витка.

На рис. 2.6 показан потенциометрический датчик угловых перемещений: 1 – каркас, 2 – обмотка, 3 – щетка, 4 – дополнительная щетка для снятия сигнала, 5 – токосъемное кольцо.

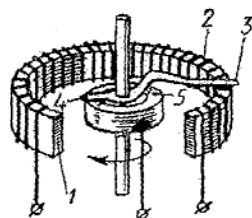


Рис. 2.6

К достоинствам проволочных реостатных преобразователей можно отнести: простоту конструкции; малые размеры и массу; возможность получения необходимых функциональных зависимостей относительно простыми средствами; получение высокоточных линейных статических характеристик; стабильность характеристик.

Потенциометрические схемы имеют и недостатки.

Первый недостаток – нелинейность характеристики – можно ослабить, шунтируя верхнее плечо потенциометра резистором $R = R_n$. На рис. 2.7 показана схема такого включения, а на рис. 2.8 – соответствующие зависимости ($\alpha = X/l$, $\beta = R_n/R_n$).

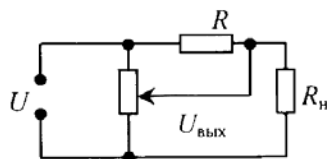


Рис. 2.7

Второй недостаток – ступенчатая зависимость (с перемежающимися малыми и большими скачками) напряжения от перемещения движка из-за периодического замыкания двух соседних витков

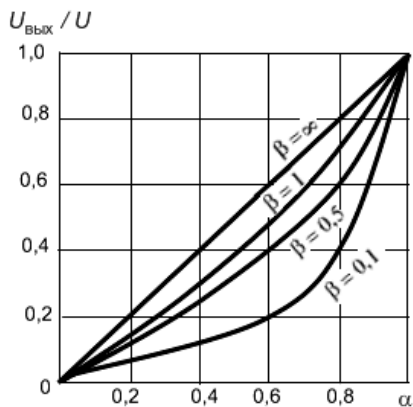


Рис. 2.8

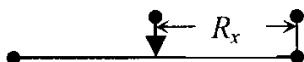


Рис. 2.9

при определенных положениях движка. Его можно уменьшить путем применения более тонкого провода (практический предел – 50 витков/мм, что соответствует разрешению 20 мкм) или применением реохорда (рис. 2.9).

Третий недостаток – наличие скользящего контакта – может стать причиной отказа из-за окисления контактной дорожки, перетирания витков или отгибания движка;

Также к недостаткам можно отнести:

- сравнительно небольшой коэффициент преобразования и

высокий порог чувствительности;

- наличие шума; подверженность электроэрозии под действием импульсных разрядов;

- ограниченную скорость линейного перемещения или вращения (до 100...200 об/мин) токосъемника вследствие его вибраций при переходе с витка на виток и повышение при этом уровня динамического шума;

- низкую износостойчивость.

2.7. Тензометрические датчики

Эти преобразователи используют *тензорезистивный эффект* – изменение активного сопротивления проводниковых и полупроводниковых материалов под действием механической деформации. Тензодатчик состоит из тонкой проволоки, фольги или полоски полупроводника, прикрепленной к подложке или нанесенной на нее (например, напылением) (рис. 2.10).

Датчик жестко закрепляется на деформируемой поверхности, причем жесткость подложки должна быть большей по сравнению с жесткостью чувствительного элемента, но меньшей по сравнению с жесткостью поверхности измеряемого объекта. Это обеспечивает равенство деформаций датчика и образца в области закрепления датчика.

В простейшем случае датчик состоит из сетки, образованной нитевидными проводниками с удельным сопротивлением ρ , площадью поперечного сечения s и длиной nl ,

где l – длина одного нитевидного элемента, n – их количество; для металлических датчиков $n = 10 \dots 20$, для полупроводниковых $n = 1$. Сетка наносится на изолирующую подложку – бумагу или пластик, которая наклеивается на исследуемый образец. Датчик испытывает такую же деформацию $\Delta l/l$, как и образец, в направлении, параллельном нитям.

Для изготовления металлических резисторных датчиков обычно используют сплавы на основе никеля; диаметр проволоки составляет около 20 мкм, толщина подложки – около 0,03 мм для пластика и 0,1 мм для бумаги.

Сопротивление преобразователя длины L с удельным сопротивлением ρ равно

$$R = \rho \int_0^L \frac{dl}{dS(l)},$$

где $S(l)$ – сечение. При растяжении изменяется L , S и ρ (из-за деформации кристаллической решетки) и при $S(l) = S$ сопротивление датчика изменяется на величину

$$\Delta R = \Delta \rho \frac{l}{S} + \Delta l \frac{\rho}{S} - \Delta S \frac{\rho l}{S^2}.$$

Введем коэффициент Пуассона $\sigma = -(\Delta D/D)/(\Delta L/L)$ – отношение относительного изменения диаметра к относительному изменению длины и получим относительное изменение R :

$$\Delta R/R = [(1+2\sigma)\Delta L/L + \Delta \rho/\rho].$$

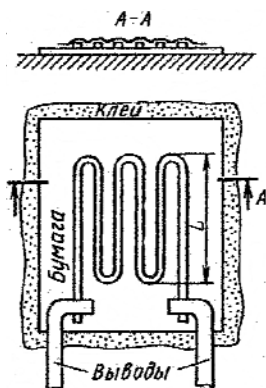


Рис. 2.10

Коэффициент тензочувствительности

$$G = (\Delta R/R)/(\Delta L/L) = 1 + 2\sigma + (\Delta\rho/\rho)/(\Delta L/L)$$

для большинства металлов ($\sigma \approx 0,3$) лежит в пределах 1,6...6. Значительно большие значения у полупроводниковых материалов (из-за одновременного изменения ρ), например, для германия $G \approx 100$, однако это достоинство «компенсируется» большой зависимостью ρ от температуры и большей нелинейностью.

Неправильная установка тензодатчика на объекте может привести к значительным погрешностям из-за ползучести материалов, влияния влаги, плохой электрической изоляции от контролируемого участка и ряда других факторов.

Важное значение имеет и ориентация тензодатчика. Поскольку между измерительными нитями имеются участки проводников, перпендикулярные общему направлению, возникает чувствительность к поперечным деформациям. Для проволоочных датчиков соответствующий коэффициент тензочувствительности $G_t \approx 0,02G$. Чтобы уменьшить этот эффект, поперечные части проводника делаются шире продольных.

Частотные ограничения зависят от способа закрепления и размеров датчика: длина резисторных полосок должна быть гораздо меньше длины волны механических колебаний с тем, чтобы измеряемая под датчиком деформация была практически однородна. Учитывая это правило, максимальная частота вибрации при использовании датчика длиной l определяется соотношением $f_{\max} = V/10 l$, где V – скорость звука в материале образца (например, для стали $V = 5810$ м/с).

Многokратное повторение противоположных по знаку деформаций приводит к постепенному возрастанию собственного сопротивления датчика. *Предел усталости* датчика понимают как число N циклов деформации заданной амплитуды, которые приводят к изменению сопротивления, эквивалентному деформации 10^{-4} . Он зависит от материала датчика и амплитуды деформаций и лежит в диапазоне от 10^4 до 10^8 циклов.

Для измерения сигнала тензодатчиков применяют мостовые схемы включения, причем следует иметь в виду, что тензодатчики – элементы с низким сопротивлением (от 120 до 350 Ом); это озна-

чает, что падение напряжения в проводах, соединяющих источник питания с мостом, может стать причиной погрешностей измерения. Полный выходной сигнал типичного тензометрического датчика мостового типа равен 3...4 мВ на 1 В напряжения питания мостовой схемы.

Для повышения чувствительности применяют дифференциальную схему включения: датчики с сопротивлениями R_1 и R_3 устанавливаются в области растяжения от действия силы F , а датчики с сопротивлениями R_2 и R_4 – в области сжатия (рис. 2.11, а). Условие баланса моста ($R_1R_3 = R_2R_4$) нарушается весьма сильно (рис. 2.11, б).

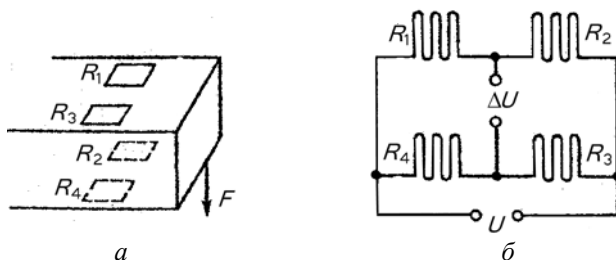


Рис. 2.11

Погрешности тензорезисторов могут быть вызваны изменениями температуры. Для уменьшения этого влияния применяют компенсационную схему включения (рис. 2.12): рабочий датчик с сопротивлением R_p накладывается в направлении действия усилия, а компенсирующий R_k – в перпендикулярном направлении. При деформации детали изменяется R_p , при изменении температуры изменяются оба сопротивления.

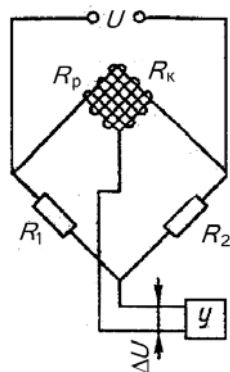


Рис. 2.12

К достоинствам тензорезисторов можно отнести незначительную массу, малые размеры, простоту конструкции, возможность измерения статических и динамических процессов; к недостаткам – относительно невысокую чувствительность, возможность только разового использования, необходимость использования мостовой измерительной схемы и компенсации температурных воздействий.

2.8. Электромагнитные датчики

2.8.1. Принцип действия

Как известно, обмотка, имеющая w витков, при протекании по ней тока I , создает магнитный поток в сердечнике

$$\Phi = Iw/R_M,$$

где R_M – магнитное сопротивление сердечника:

$$R_M = l_{cp}/\mu S,$$

где l_{cp} – средняя длина магнитной линии, μ – магнитная проницаемость сердечника, S – площадь поперечного сечения сердечника.

Электромагнитные датчики предназначены для преобразования перемещения в электрический сигнал за счет изменения параметров электромагнитной цепи. Эти изменения могут заключаться в увеличении или уменьшении магнитного сопротивления датчика или потокосцепления обмотки. В результате таких перемещений изменяется индуктивность обмотки L или ее взаимдукция M с обмоткой возбуждения. Поэтому электромагнитные датчики часто называют *индуктивными*.

Электромагнитные датчики обычно рассматриваются как параметрические, поскольку изменяются величины L и M , но электромагнитные датчики с изменяющейся взаимдукцией можно отнести и к генераторному типу, поскольку в результате изменяется и ЭДС обмотки. Так как эта ЭДС появляется за счет изменения коэффициента взаимдукции с обмоткой возбуждения, то такие электромагнитные датчики называют *трансформаторными*.

К генераторным относятся и *индукционные* датчики, в обмотках которых генерируется ЭДС в зависимости от скорости перемещения. Изменение индуктивности и взаимдукции может происходить и под влиянием механических напряжений в сердечнике электромагнитного датчика. Такие напряжения приводят к изменению магнитной проницаемости ферромагнитного материала сердечника. Электромагнитные датчики, основанные на таком физическом явлении, называются *магнитоупругими* датчиками.

Индуктивность L катушки с сердечником зависит от количества ее витков w и магнитного сопротивления магнитной системы R_M :

$L = w^2/R_m$. В свою очередь, величина R_m зависит от длины магнитопровода и магнитной проницаемости μ . Так, магнитопровод в форме П-образного сердечника с якорем при наличии воздушного зазора δ между сердечником и якорем имеет магнитное сопротивление

$$R_m = \frac{l}{\mu_c S} + \frac{2\delta}{\mu_b S},$$

где l – общая длина сердечника, S – площадь сечения магнитопровода, μ_c и μ_b – магнитная проницаемость сердечника и воздуха соответственно.

Поскольку $\mu_c \gg \mu_b$, то для разомкнутого магнитопровода $L \approx \frac{w^2}{2\delta} \mu_b S$, а для замкнутого магнитопровода $L \approx \frac{w^2}{l} \mu_c S$. В перемещении, которое предполагается измерить, участвует один из элементов магнитного контура. Перемещение якоря (изменение величины зазора δ) изменяет индуктивность и, следовательно, может быть измерено. Для измерения индуктивностей и их изменений используются различные балансные мостовые схемы. На результат оказывают влияние такие факторы как наличие распределенной емкости, потери энергии в проводе обмотки и в сердечнике (из-за токов Фуко и магнитные потери).

К недостаткам индуктивных датчиков относят:

- постоянно действующую на якорь электромагнитную силу, стремящуюся притянуть его к ядру; она пропорциональна квадрату напряжения питания и обратно пропорциональна квадрату частоты питания;

- независимость фазы выходного сигнала от направления перемещения якоря.

Общие замечания по применению индуктивных датчиков:

1) поскольку ферромагнитные материалы – самые малошумящие, на их основе можно изготовить очень чувствительные датчики (до 10^{-5} мм);

2) по самой своей природе индуктивные датчики, с одной стороны, чувствительны к внешним электромагнитным полям, а с другой – способны сами их индуцировать. Поэтому индуктивные дат-

чики необходимо помещать внутри кожуха, служащего магнитным экраном;

3) индуктивные датчики имеют малое выходное сопротивление и высокий КПД;

4) габариты индуктивных датчиков обратно пропорциональны частоте возбуждения.

2.8.2. Дифференциальные индуктивные датчики

Дифференциальные индуктивные датчики представляют собой совокупность двух одинарных нереверсивных датчиков с общим якорем [1]. Они обладают реверсивной статической характеристикой и компенсируют электромагнитную силу притяжения якоря. Оба сердечника в такой схеме (рис. 2.13) идентичны по своим конструктивным и магнитным параметрам. Расположенные на них обмотки w_1 и w_2 имеют также одинаковые параметры и включены встречно. Сопротивление R_n включается между средней точкой трансформатора и средней точкой обмоток преобразователя.

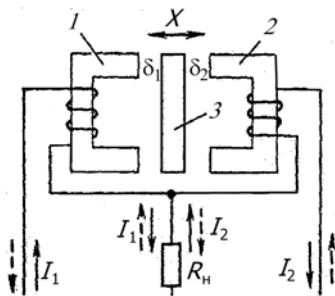


Рис. 2.13

В такой схеме ток $I_{\text{вых}}$, протекающий по сопротивлению нагрузки, равен разности токов правой и левой половин схемы. В векторной форме это равенство примет следующий вид:

$$\dot{I}_{\text{вых}} = \dot{I}_1 - \dot{I}_2,$$

а выходное напряжение

$$\dot{U}_{\text{вых}} = R_n (\dot{I}_1 - \dot{I}_2).$$

При отсутствии входного сигнала зазоры δ_1 и δ_2 между якорем и ярмом одинаковы: $\delta_1 = \delta_2 = \delta_0$. Равны и индуктивности L_1 и L_2 обеих половин датчика, определяемые размерами зазоров. Следовательно, выходное напряжение преобразователя $U_{\text{вых}} = 0$. При перемещении якоря на расстояние X зазоры δ_1 и δ_2 становятся неравными: $\delta_1 = \delta_0 + X$; $\delta_2 = \delta_0 - X$. Изменение зазоров приводит к изменению

индуктивностей: увеличение δ_1 ведет к уменьшению L_1 , а уменьшение δ_1 – к увеличению L_2 . Изменение индуктивностей плеч датчика приводит к дисбалансу токов I_1 и I_2 , в результате чего через сопротивление нагрузки потечет ток $I_H = I_{\text{вых}}$, и появится выходное напряжение.

Если изменяется направление перемещения якоря, то фаза выходного напряжения сдвигается на 180° относительно напряжения питания, являющегося опорным.

Примерные технические параметры датчиков с дифференциальным включением индуктивностей: диапазон измерения – от 0,1 до 200 мм; чувствительность – от 10 до 100 мВ/мм; отклонение от линейности – от 0,5 до 2 % диапазона измерения.

Для получения реверсивной статической характеристики используют и мостовую схему включения индуктивных датчиков (рис. 2.14).

Индуктивные датчики с переменным воздушным зазором применяются для измерения перемещений в диапазоне 0,1...1,0 мм. При больших перемещениях статическая характеристика датчика становится нелинейной. Для измерения больших перемещений (до 50 мм) применяют индуктивные датчики в виде катушки с подвижным внутренним сердечником. Если сердечник полностью введен внутрь катушки, на которую намотана обмотка, то ее индуктивное сопротивление максимально, а ток в обмотке имеет минимальное значение. При выводе сердечника из катушки индуктивное сопротивление уменьшается, а ток соответственно увеличивается. Индуктивные датчики в виде катушки с перемещающимся внутри нее сердечником получили название *плунжерных* датчиков.

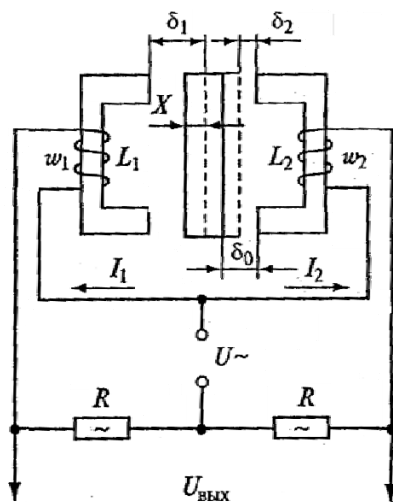


Рис. 2.14

2.8.3. Трансформаторные датчики

Схема трансформаторного датчика приведена на рис. 2.15. Обмотки цепи питания w_1 и w_2 одинаковы и включены таким образом, что, когда по ним протекает рабочий ток, создаваемые ими магнитные потоки Φ_1 и Φ_2 направлены встречно в центральном стержне, на котором расположена выходная обмотка w_0 [2].

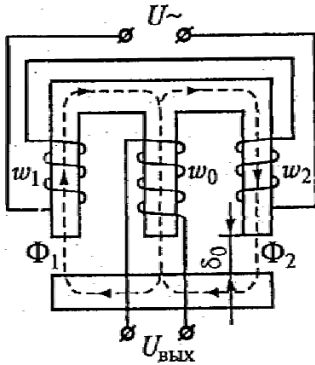


Рис. 2.15

В этом случае выходное напряжение

$$U_{\text{вых}} = 4,44f w_0 (\Phi_2 - \Phi_1),$$

где f – частота питающего напряжения.

Если считать ток в цепи питания постоянным, то магнитные потоки пропорциональны индуктивностям обмоток w_1 и w_2 :

$$\Phi_1 = IL_1 / w_1, \quad \Phi_2 = IL_2 / w_2.$$

Так как в рассматриваемой конструкции преобразователя зазор остается постоянным, а якорь перемещается вдоль зазора, то значение индуктивности зависит от сечения магнитного сердечника, определяемого площадью перекрытия S стержней, которая изменяется при перемещении якоря [2].

В нейтральном положении при $X = 0$ площади перекрытия стержней равны: $S_1 = S_2 = S_0$, что приводит к равенству потоков $\Phi_1 = \Phi_2$ и, следовательно, $U_{\text{вых}} = 0$.

При перемещении якоря влево на расстояние X площадь перекрытия правого крайнего стержня изменится на ΔS , и для потоков в крайних стержнях можно записать:

$$\begin{aligned} \Phi_1 &= I\mu_0 w_1 S_0 / \delta_0, \\ \Phi_2 &= I\mu_0 w_2 (S_0 - \Delta S) / \delta_0. \end{aligned}$$

Выходное напряжение датчика можно описать линейной зависимостью

$$U_{\text{вых}} = K\Delta S = K_1 X,$$

так как площади перекрытия стержней пропорциональны перемещению якоря.

Модификацией рассмотренной схемы является дифференциально-трансформаторный датчик плунжерного типа, показанный на рис. 2.16, а. Это электромеханическое устройство, состоящее из первичной и двух совершенно одинаковых вторичных обмоток, расположенных симметрично относительно первичной на цилиндрическом каркасе [1]. Первичную обмотку питает синусоидальное напряжение. Свободно движущийся внутри обмоток магнитный сердечник в форме стержня обеспечивает связь этих обмоток через магнитный поток, а его перемещение x связано с измеряемым объемом.

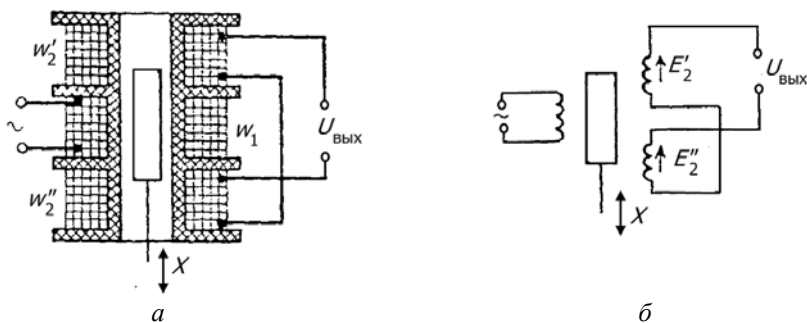


Рис. 2.16

Схема включения обмоток показана на рис. 2.16, б. При возбуждении первичной обмотки с помощью внешнего источника напряжения U_1 переменного тока (обычно с частотой от 50 до 25000 Гц) в двух вторичных обмотках наводятся ЭДС, определяемые взаимной индуктивностью контуров. Вторичные обмотки включены последовательно навстречу друг другу, так что индуцируемые в них напряжения для выходной цепи являются разнополярными. Поэтому результирующий выходной сигнал U_0 равен нулю, когда сердечник находится в центральной позиции. При перемещении X максимальный сигнал формируется на той выходной обмотке, вблизи которой находится сердечник, и где, следовательно, выше взаимная индуктивность. Выходной сигнал дифференциально-трансформаторного датчика плунжерного типа – гармоническое колебание, амплитуда которого модулируется низкочастотным перемещением сердечника.

Порядок величин параметров дифференциального трансформатора следующий: линейные перемещения – от 1 до 500 мм; чувствительность – от 1 до 500 мВ/мм; отклонение от линейности – от 0,05 до 1% диапазона измерения; порог чувствительности – от 0,002 до 0,05% диапазона чувствительности, частоты возбуждения – от 50 Гц до 30 кГц.

Похожий принцип используется в *резольверах*. Резольвер измеряет угловые перемещения и работает по принципу измерения взаимоиנדукции между двумя обмотками. Ротор резольвера соединен с вращающимся объектом, и на его обмотку подается переменное напряжение U . Статор имеет две обмотки, повернутые на 90° друг относительно друга, с которых снимается выходное напряжение соответственно

$$U_1 = KU \sin \varphi,$$
$$U_2 = kU \cos \varphi,$$

где φ – угловое положение ротора.

Выход резольвера есть тригонометрическая функция угла. Эта нелинейность не всегда является недостатком: например, при управлении вращающимися моментами робота требуются именно такие функции угла.

Резольвер дает хорошее разрешение и высокую точность, имеет высокий уровень сигнала и низкий выходной импеданс. Проблемы могут возникать только из-за щеток ротора (износ, дополнительные шумы и механические нагрузки).

К достоинствам рассмотренных индуктивных трансформаторных датчиков следует отнести достаточно высокую выходную мощность; высокие чувствительность и разрешающую способность; сравнительную простоту конструкции; высокую надежность; малые массу и размеры при расчете на напряжение повышенной частоты; невысокую стоимость.

Недостатками рассмотренных измерительных преобразователей являются трудность регулировки и компенсации начального напряжения на их выходе; необходимость экранирования для уменьшения уровня помех; возможность работы только на переменном токе; ограниченность диапазона линейной статической характеристики.

2.8.4. Индукционные датчики

Индукционные датчики предназначены для преобразования скорости линейных и угловых перемещений в ЭДС. Это датчики генераторного типа [1].

Принцип их действия основан на законе электромагнитной индукции. Выходным сигналом индукционных датчиков является ЭДС, которая пропорциональна скорости изменения магнитного потока, пронизывающего витки катушки. Это изменение происходит за счет перемещения катушки в постоянном магнитном поле или за счет вращения ферромагнитного индуктора относительно неподвижной катушки.

Основным отличием индукционных датчиков от индуктивных является то, что в них используется постоянное магнитное поле, а не переменное. Постоянное магнитное поле создается двумя способами: постоянными магнитами или катушкой, обтекаемой постоянным током.

На рис. 2.17, *а* показана схема датчика с обмоткой w_2 , размещенной в воздушном зазоре, в котором постоянный магнитный поток Φ создается катушкой w_1 , включенной на постоянное напряжение U_+ . При перемещении катушки в магнитном поле в ней индуцируется ЭДС, пропорциональная скорости перемещения:

$$E = k \Phi \frac{dx}{dt},$$
 где k – коэффициент пропорциональности, зависящий

от числа витков w_2 и конструктивных параметров датчика.

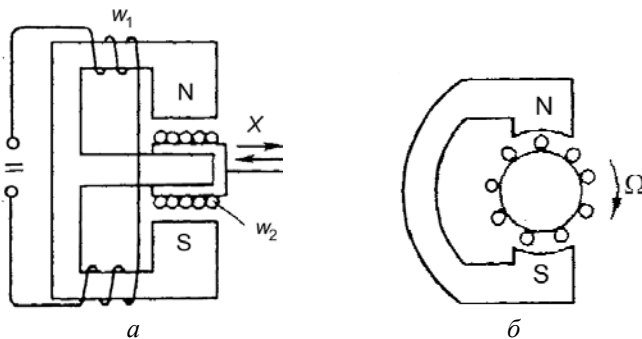


Рис. 2.17

На рис. 2.17, б показан датчик, в котором постоянный магнитный поток создается с помощью постоянного магнита с полюсными наконечниками. ЭДС, индуцируемая во вращающейся катушке, пропорциональна скорости вращения Ω : $E = k\Phi\Omega$.

В обоих этих датчиках катушки подвижны, поэтому для отвода от них выходного сигнала (ЭДС) необходимы гибкие токоподводы или контактные кольца.

2.8.5. Магнитоупругие датчики

Принцип действия магнитоупругих датчиков основан на *магнитоупругом эффекте* – физическом явлении, проявляющемся в виде изменения магнитной проницаемости ферромагнитного материала в зависимости от механических напряжений в нем. Магнитоупругие датчики используются для измерения силовых параметров: усилий, давлений, крутящих и изгибающих моментов, механических напряжений и т.п.

Конструктивно магнитоупругие датчики представляют магнитопровод с одной или несколькими обмотками. Магнитное сопротивление сердечника $R_m = l/(s\mu)$, где l и s – длина и площадь сечения сердечника. Если к сердечнику приложено механическое усилие F , то магнитная проницаемость μ изменится. Следовательно, изменятся и магнитное сопротивление сердечника, и индуктивность обмотки на сердечнике.

Зависимость магнитной проницаемости от механических напряжений имеет нелинейный характер. Однако при определенных значениях напряженности магнитного поля H в сердечнике можно получить близкую к линейной зависимость изменения магнитной проницаемости $\Delta\mu$ сердечника от относительной деформации $\Delta l/l$ или нормального механического напряжения σ в зоне линейных деформаций. Наиболее заметен магнитоупругий эффект в пермалловых (железокобальтовых и железоникелевых) сплавах.

К достоинствам магнитоупругих датчиков следует отнести высокую чувствительность и возможность измерения больших усилий (до нескольких тысяч тонн). В то же время магнитоупругие датчики имеют и серьезные недостатки: 1) наличие температурной

погрешности; 2) наличие погрешности, вызванной влиянием гистерезиса (как магнитного, так и механического, связанного с остаточной деформацией); 3) наличие погрешности, вызванной колебаниями напряжения питания.

2.9. Емкостные датчики

Принцип действия емкостных измерительных преобразователей основан на изменении емкости конденсатора под воздействием входной преобразуемой величины.

Емкостной датчик представляет собой плоский или цилиндрический конденсатор, одна из обкладок которого испытывает подвергаемое контролю перемещение, вызывая изменение емкости. Емкость C связана с толщиной диэлектрика d и его площадью S зависимостью $C = \epsilon \epsilon_0 S/d$, где ϵ – относительная диэлектрическая проницаемость; ϵ_0 – диэлектрическая проницаемость вакуума. Из приведенного соотношения видно, что на емкость конденсатора можно влиять изменением площади перекрытия пластин S (рис. 2.18, а), расстояния между ними d (рис. 2.18, б), диэлектрической проницаемости ϵ вещества, находящегося в зазоре между обкладками конденсатора (рис. 2.18, в). Выбор того или иного измеряемого параметра зависит от характера измеряемой величины.

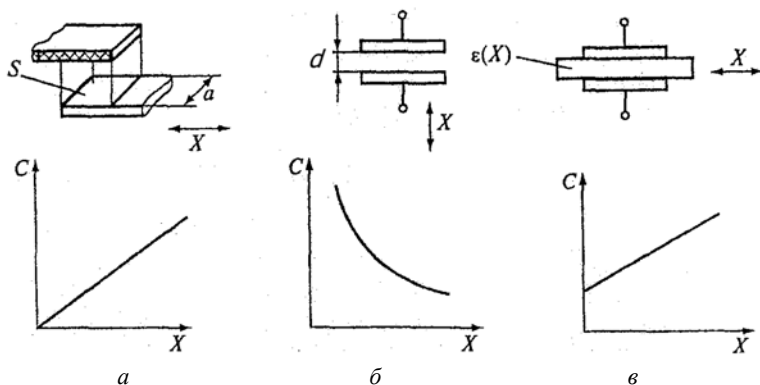


Рис. 2.18

Емкостные преобразователи используют для измерения угловых и линейных перемещений, линейных размеров, уровня, усилий, влажности, концентрации и др. Конструктивно они могут быть выполнены с плоскопараллельными, цилиндрическими, штыревыми электродами, с диэлектриком между пластинами и без него.

Для повышения чувствительности и линейности характеристик используют дифференциальные преобразователи, у которых изменение состояния контролируемой величины приводит к изменению емкости одновременно двух чувствительных элементов, включаемых в разные плечи мостовой измерительной схемы [1]. Более высокую чувствительность позволяет получить резонансная схема (рис. 2.19, а). Высокочастотный генератор ГВЧ имеет частоту напряжения f_{Γ} и питает индуктивно связанный с ним контур, состоящий из индуктивности L_{κ} , подстроечного конденсатора C_0 и емкостного датчика $C_{\text{д}}$. Напряжение U_{κ} , снимаемое с контура, усиливается усилителем $У$ и измеряется прибором Π , шкала которого может быть проградуирована в единицах измеряемой величины.

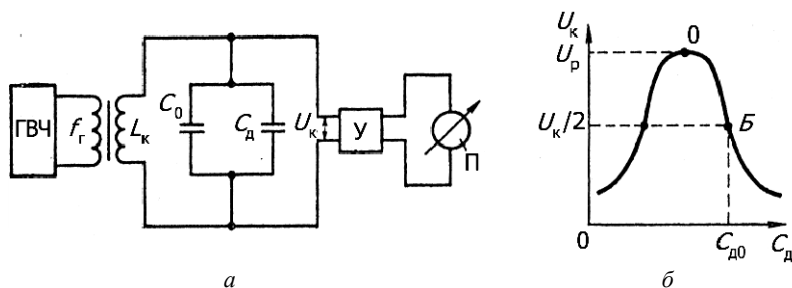


Рис. 2.19

При помощи подстроечного конденсатора контур настраивается на частоту f_0 , близкую (но не равную) к частоте генератора. Настройка производится при средней емкости датчика в диапазоне возможных изменений измеряемой величины (рис. 2.19, б)

$$C_{\text{д}0} = (C_{\text{макс}} + C_{\text{мин}})/2,$$

в результате чего выходное напряжение примерно в два раза меньше напряжения при резонансе (точка А). Малейшее перемещение пластины конденсатора приводит к изменению $C_{\text{д}}$ и, следовательно, к резкому изменению выходного напряжения.

Резонансная частота контура определяется из условия резонанса (равенства емкостного и индуктивного сопротивлений):

$$2\pi fL = 1/2\pi fC, \quad f_p = \frac{1}{2\pi\sqrt{L_k(C_0 + C_d)}}.$$

Емкостные датчики просты, прочны и надежны по конструкции. Диэлектриком обычно служит воздух, так что параметры конденсатора зависят только от геометрических характеристик и не зависят от свойств используемых материалов, если они правильно подобраны. Необходимо защищать датчик от тех факторов окружающей среды, которые могут ухудшить изоляцию между обкладками: от пыли, коррозии, влажности, ионизирующих излучений.

Датчики с изменяющимся зазором обладают большой чувствительностью, поэтому их используют для контроля очень малых перемещений (от 0 до 1 мм). Датчики с изменяющейся площадью электродов применяют для измерения перемещений, больших 1 мм. Емкостные датчики с изменяющейся диэлектрической проницаемостью применяют для контроля уровня, состава и концентрации жидких, а также толщины и влажности твердых диэлектриков и т.д.

К достоинствам емкостных измерительных преобразователей можно отнести простоту конструкции, малые размеры и массу, высокую чувствительность, большую разрешающую способность при малом уровне входного сигнала, отсутствие подвижных токоъемных контактов, высокое быстродействие, возможность получения необходимого закона преобразования за счет выбора соответствующих конструктивных параметров, отсутствие влияния выходной цепи на измерительную.

Недостатки емкостных измерительных преобразователей состоят в относительно низком уровне выходной мощности сигналов, нестабильности характеристик при изменении параметров окружающей среды, влиянии паразитных емкостей.

2.10. Пьезоэлектрические датчики

Эти преобразователи используют явление *пьезоэффекта*, когда приложенная сила генерирует электрический заряд. Пьезоэффект обусловлен тем, что при деформации несимметричной кристалли-

ческой решетки происходит перераспределение зарядов, при котором положительные и отрицательные заряды смещаются к противоположным поверхностям кристалла. Плотность зарядов пропорциональна создаваемому механическому напряжению.

К классу пьезоэлектриков относятся как природные кристаллы (кварц, сегнетова соль), так и искусственные (титанат бария, твердые растворы цирконата-титаната свинца).

Свойства различных пьезоэлектрических материалов определяются относительно трехмерной системы координат. Вектор электрической поляризации (т.е. величина и знак пьезоэффекта) зависит от формы и ориентации кристалла преобразователя по отношению к полярной оси кристалла и расположения электродов, а также от знака деформации.

Связь между деформацией твердого тела и силой устанавливается (в определенных пределах упругой деформации) законом Гука, который для продольного растяжения (сжатия) имеет вид

$$\Delta l = Fl/ES, \quad (2.1)$$

где l – первоначальная длина, Δl – изменение длины при нагрузке F , E – модуль Юнга, S – площадь сечения. Перемещение преобразуется в силу упругим чувствительным элементом.

Соотношение (2.1) показывает, что пьезоэлектрический преобразователь удобно использовать для измерения силы и давления. Конструктивно пьезодатчик представляет упругий элемент, воздействующий на пластину пьезоэлектрика, помещенную между двумя металлическими обкладками. Поскольку пьезоэлектрический преобразователь по существу является диэлектриком с высоким сопротивлением утечки, в качестве его модели удобно рассматривать плоский конденсатор. Полный заряд, индуцированный в нем, прямо пропорционален приложенной силе F : $q = dF$, где d – пьезоэлектрический коэффициент, измеряемый в Кл/Н. Типичная амплитудно-частотная характеристика пьезоэлектрического преобразователя показана на рис. 2.20, где по оси абсцисс отложена частота внешнего воздействия ($\omega_{рез}$ – резонансная частота), а по оси ординат – отношение выходного напряжения $U_{вых}$ к деформирующей силе F .

Для измерения сигнала пьезоэлектрического преобразователя используют два метода: усиление напряжения и усиление заряда. При **усилении напряжения** усилитель должен обладать высоким входным импедансом (приблизительно 100 МОм, 20 пФ), так как наличие дополнительной емкости соединительного кабеля (обычно 200 пФ/м) приводит к снижению фактической чувствительности: напряжение U на конденсаторе прямо пропорционально заряду q и обратно пропорционально емкости C .

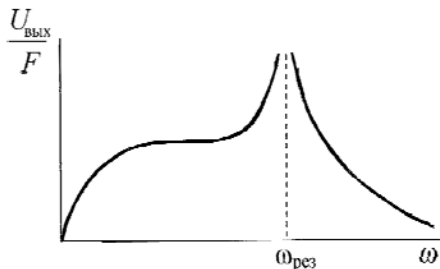


Рис. 2.20

Более распространенный метод — **усиление заряда**, когда весь генерируемый заряд накапливается на конденсаторе обратной связи усилителя.

Верхняя граница допустимого частотного диапазона определяется в основном механическими параметрами преобразователя. Пьезоэлектрические преобразователи могут быть выполнены с частотой собственных колебаний до 100 кГц, что позволяет измерить механические величины, изменяющиеся с частотой до 7...10 кГц.

Пьезоэлектрические датчики применяются для измерения давления, силы, ускорения. На рис. 2.21 показано устройство пьезоэлектрического датчика давления. Оно действует на мембрану 1. Кварцевые пластины 2 зажаты между металлическими прокладками 3. Средняя прокладка соединена с выводом 4, проходящим через экранированную втулку 5 из изоляционного материала. Крышка 6 соединяется с корпусом и через шарик 7 передает давление пластинам.

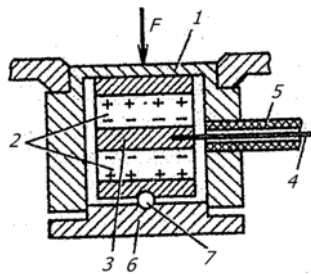


Рис. 2.21

На рис. 2.22 показан пьезокристаллический датчик ускорения. Пьезоэлемент 1 из титаната бария расположен в корпусе прибора 2 между инерционной массой 3 и подпятником 4. Масса имеет относительно большие размеры и изготовлена из вольфрама. Пакет из

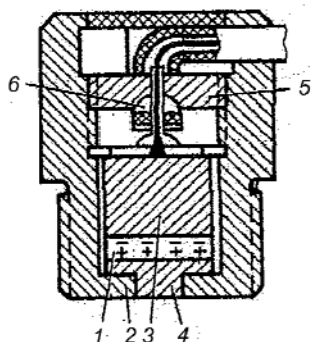


Рис. 2.22

инерционной массы, пьезоэлемента и подпятника прижат к основанию корпуса гайкой 5 через сферическую пятю 6, изоляционную прокладку, пружинную шайбу и контактную пластину. Вывод сигнала выполнен с помощью специального антивибрационного кабеля.

При построении пьезоэлектрических датчиков могут использоваться резонансные свойства пьезоэлемента: его резонансная частота

$$f_p = \frac{n}{2l} \sqrt{\frac{c}{\rho}},$$

где c – коэффициент упругости; l – размер; n – количество полу-волн на размере l ; ρ – плотность, изменяется под действием измеряемой величины.

Перестройка частоты пропорциональна величине создаваемых напряжений σ : $\Delta f/f_p = k\sigma$, $k = \text{const}$.

К достоинствам пьезоэлектрических преобразователей необходимо отнести простоту конструкции, малые размеры и стоимость, высокую надежность, возможность измерения быстропротекающих процессов, к недостаткам – невысокую чувствительность, непригодность к измерению статических величин, относительно невысокий уровень выходного сигнала.

2.11. Терморезисторы

Терморезисторы – это термопреобразователи сопротивления. Принцип их действия основан на зависимости электрического сопротивления проводников от температуры. С ростом температуры сопротивление металлических терморезисторов увеличивается.

Для измерения температуры в диапазоне от -200 до $+850$ °C чаще всего применяются датчики температуры из никеля или платины. Для платины

$$R_t = R_0(1 + At + Bt^2) \text{ при } 0 < t < +650 \text{ °C,}$$

$R_t = R_0 [1 + At + Bt^2 + Ct^3(t - 100)]$ при $-200 < t < 0$ °С,
 где R_t – сопротивление терморезистора при t ; R_0 – сопротивление при 0 °С (обычно порядка 100 Ом); $A = 3,968477 \times 10^{-3} (1/^\circ\text{C})$; $B = -5,8477 \times 10^{-7} (1/^\circ\text{C})^2$; $C = -4,227 \times 10^{-12} (1/^\circ\text{C})^3$.

Платиновые терморезисторы выпускаются серийно в виде неизолированной проволоки диаметром 0,05...0,07 мм с сопротивлением от 1,7 до 333 Ом.

При малых температурах (до 280 °С) применяют более дешевые медные терморезисторы, у которых

$$R_t = R_0(1 + \alpha t) \text{ при } -50 < t < 180 \text{ }^\circ\text{C},$$

$$\alpha = 4,26 \times 10^{-3} (1/^\circ\text{C})^{-1}.$$

Медные терморезисторы выполняются из проволоки диаметром не менее 0,1 мм, покрытой для изоляции эмалью с сопротивлением от 42 до 177 Ом.

Простейшая разновидность измерительной схемы для таких датчиков – измерительный мост (рис. 2.23), в котором терморезистор R_T включен в диагональ моста, а $R_1 = R_2$. Если установить R_3 так, что ток $I_0 = 0$, то окажется, что $R_3 = R_T$ (поскольку $R_1 = R_2$). В технических приложениях поступают иначе: обеспечивают постоянство R_3 , используют высокоомный вольтметр, чтобы между точками 1 и 2 не протекал ток, и калибруют напряжение между точками 1 и 2

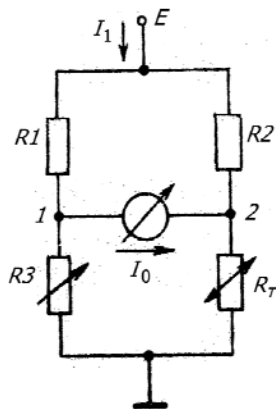


Рис. 2.23

$$U_{12} = E \left[\frac{R_1}{R_1 + R_2} - \frac{R_3}{R_3 + R_T} \right]$$

в единицах температуры. При желании добиться высокой чувствительности измерений приходится существенно увеличивать ток через терморезистор. Допустимые токи лежат в пределах: для медно-марганцевых – до 10 мА, для кобальтово-марганцевых – до 5 мА.

Обычно для подключения терморезисторов используется трехпроводная схема (рис. 2.24): питание датчика осуществляется через

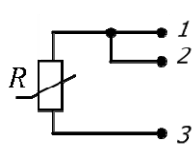


Рис. 2.24

выходы 1 и 3, а сигнал снимается с выходов 2 и 3. Это позволяет уменьшить погрешность, возникающую при изменении сопротивления проводов (например, при изменении их температуры). По сравнению с металлическими терморезисторами более высокой чувствительностью обладают полупроводниковые терморезисторы (*термисторы*).

При весьма малых размерах они имеют высокое сопротивление (до 1 МОм), изменяющееся по закону

$$R_T = A \exp(B/T),$$

где T – абсолютная температура; A, B – константы. С ростом температуры сопротивление термисторов уменьшается.

Особый интерес для массового применения представляют кремниевые термисторы, так как они значительно дешевле других датчиков и имеют большой температурный коэффициент.

Характеристика кремниевого датчика достаточно точно описывается функцией

$$R_T = R_0(1 + a\Delta T + b\Delta T^2),$$

где R_T – сопротивление при температуре T (°C); R_0 – сопротивление при эталонной температуре T_0 , обычно $T_0 = 25$ °C; $\Delta T = T - T_0$, a и b – температурные коэффициенты с типичными значениями $0,78 \times 10^{-2} \text{ } ^\circ\text{C}^{-1}$ и $1,84 \times 10^{-5} \text{ } ^\circ\text{C}^{-2}$ соответственно.

Верхний диапазон измеряемых температур для медно-марганцевых термисторов – 120 °C, а для кобальтово-марганцевых – 180 °C.

К недостаткам термисторов относят нелинейность функции преобразования и плохую воспроизводимость характеристик.

Постоянная времени большинства серийно выпускаемых термисторов лежит в диапазоне от десятых долей до единиц секунд.

При использовании терморезисторов в качестве датчиков различают два основных режима. В первом режиме температура терморезистора практически определяется только температурой окружающей среды. Ток, протекающий через терморезистор, очень мал и практически не нагревает его. Во втором режиме терморезистор нагревается проходящим по нему током, а температура терморезистора определяется изменяющимися условиями теплоотдачи, на-

пример интенсивностью обдува, плотностью окружающей среды и т.п. Поскольку при повышении температуры у термисторов сопротивление уменьшается, ток увеличивается, что приводит к еще большему выделению теплоты, что позволяет получить в схемах с термисторами своеобразные характеристики релейного типа.

Примером применения термистора с собственным нагревом является автоматическое измерение скорости газового потока с помощью термоанемометра [1]. Датчик этого прибора (рис. 2.25) состоит из термистора, представляющего собой тонкую пластинку 1, припаянную к двум марганциновым стержням 2, закрепленным в изоляционной втулке 3.

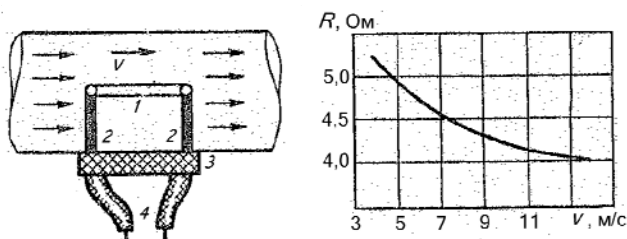


Рис. 2.25

С помощью выводов 4 термистор включается в измерительную схему. Через термистор пропускается ток, вызывающий его нагрев. Но температура (а следовательно, и сопротивление) термистора будет определяться скоростью газового потока, в который помещен датчик. Чем больше будет эта скорость, тем интенсивнее будет отводиться теплота от термистора.

В качестве датчиков температуры могут использоваться термометры на p - n -переходах. Такие датчики характеризуются линейной зависимостью выходного сигнала от температуры в определенном интервале температур. Так, падение напряжения на диоде в прямом включении $\Delta U = 4,6 kT(\ln M - \ln I)/q$, где T – температура, k – постоянная Больцмана, q – заряд электрона, I – прямой ток, M – константа. Поскольку величина напряжения ΔU зависит от тока, то для такого датчика требуется высокостабильный источник тока. Точность диодных термометров доходит до $\pm 0,002$ К в диапазоне от 14 до 300 К. Инерционность составляет порядка 10 с.

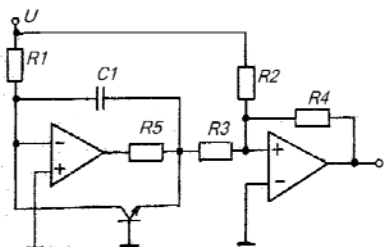


Рис. 2.26

Транзистор – тоже хороший датчик температуры (информативным параметром является напряжение база–эмиттер при фиксированном токе коллектора). Схема включения такого датчика показана на рис. 2.26. Резистор $R1$ задает коллекторный ток транзисторного датчика ($p-n$ -перехода), резистор $R4$ служит для калибровки (регулировка напряжения на выходе при $T = 0\text{ }^\circ\text{C}$). Цепочка $R2-C1$ предотвращает самовозбуждение усилителя.

2.12. Термоэлектрические датчики

В основе работы термопары лежит эффект Зеебека: если два химически однородных проводника (или полупроводника) из разных материалов A и B соединены концами в замкнутый контур и места соединений находятся при разных температурах T_1 и T_2 , то в контуре возникает ток, причина которого – разные энергии выхода носителей. В качестве материалов, используемых для создания таких преобразователей, применяют медь, железо, платину, медно-никелевые и платинородиевые сплавы.

Соединение двух материалов называют *спаем*, сами проводники – *термоэлектродами*, а преобразователь – *термопарой*. Спай, находящийся в среде с измеряемой температурой, называют *горячим* (рабочим), другой спай, относительно которого измеряется температура – *холодным* (свободным).

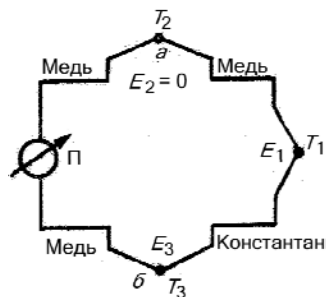


Рис. 2.27

Термопара является источником термоЭДС, то есть датчиком генераторного типа. Чтобы измерить термоЭДС, контур нужно разомкнуть и включить в разрыв электрический прибор. Однако если сделать это, не принимая дополнительных специальных мер, показания прибора не будут соответствовать действительности. На рис. 2.27 приведена

схема такого включения (для примера термопара составлена из медных и константановых термоэлектродов). Если у прибора П медные выводы и места его подключения находятся при температурах T_2 и T_3 , то на эквивалентной схеме появляются два новых спая и дополнительная термопара при температуре T_3 . Для определения T_1 нужно знать неизвестную температуру T_3 , поскольку

$$T_1 = f(E_1 - E_3) = f(\Delta E),$$

где

$$\Delta E = F(T_1 - T_3).$$

Если холодный спай поместить в ледяную ванну ($T_3 \approx 0^\circ\text{C}$), то $\Delta E = F(T_1)$. На практике используют схемные методы компенсации холодного спая. На рис. 2.28 показан вариант схемы компенсации холодного спая.

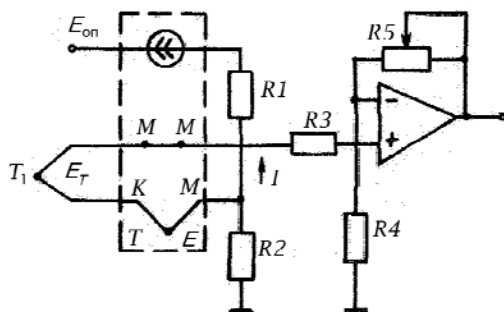


Рис. 2.28

В состав схемы входит датчик температуры холодного спая, источник опорного напряжения $E_{\text{он}}$ и операционный усилитель. При увеличении температуры T возрастает ток I . Схема построена так, что напряжение, создаваемое током I на сопротивлении R_1 , противоположно по полярности термоЭДС E , и поэтому

$$E_0 = E_T - E + (IR_1 + E_{\text{он}})/(1 + R_1/R_2) - E_{\text{он}} \approx E_T.$$

Температура связана с термоЭДС нелинейно. Так, функция преобразования медь-константановых термопар в диапазоне температур от -200 до $+300^\circ\text{C}$ с погрешностью $+2$ мкВ описывается эмпирической формулой $E = At + Bt^2 + Ct^3$, где A , B , C – постоянные, определяемые путем измерения термоЭДС при трех известных температурах, t – температура рабочего спая при $T = 0^\circ\text{C}$.

Чувствительность термопары в зависимости от ее типа лежит в пределах от 1,5 до 70 мкВ/°С, диапазон измеряемых величин: от – 50 до +1500 °С. В диапазоне обычных температур (до 100 °С) ЭДС термопары составляет единицы милливольт, поэтому прибегают к последовательному их включению, набирая «батарею» термопар.

Как правило, внутреннее сопротивление термопары лежит в диапазоне от десятых долей до единиц ом. Постоянная времени термопары зависит от конструкции и качества теплового контакта рабочего спая со средой. Для промышленных датчиков она может составлять десятки секунд; у специальных малоинерционных – от единиц до сотен миллисекунд. Проблема усиления сигналов термопар связана с низким выходным напряжением, большими синфазными помехами промышленной частоты и радиочастотными наводками. Усилитель должен хорошо подавлять синфазные помехи и иметь стабильное дифференциальное усиление; его выходное сопротивление должно быть достаточно высоким (порядка 10 кОм или более).

К достоинствам термопар необходимо отнести возможность измерений в большом диапазоне температур, простоту устройства, надежность в эксплуатации; к недостаткам – невысокую чувствительность, большую инерционность и необходимость поддержания постоянной температуры холодного спая

2.13. Фотоэлектрические датчики

Фотоэлектрические датчики позволяют преобразовывать в электрические сигналы информацию, доставляемую видимым, инфракрасным и ультрафиолетовым светом. При этом измеряемая величина изменяет один из физических параметров излучения, воздействуя на него непосредственно или косвенно; этот процесс может рассматриваться как первичное преобразование измеряемой величины. Физический параметр излучения, на который оказано воздействие, является вторичной измеряемой величиной; к ней и чувствителен оптический датчик.

Для эффективного использования фотоэлектрического датчика требуется, чтобы он соответствовал характеристикам принимаемо-

го излучения – спектру, потоку, частоте модуляции. Свойства излучения определяются, в первую очередь, источником излучения.

К основным источникам излучений фотоэлектрических датчиков можно отнести:

- *лампы накаливания с вольфрамовой нитью*. Вольфрамовая нить помещается в стеклянный или кварцевый баллон. Нить нагревается током; сила тока определяет температуру нити. Достоинства таких ламп накаливания – значительный по величине поток и непрерывный спектр излучения. Недостатки – большая тепловая инерция, ограниченный срок службы, хрупкость конструкции;

- *электролюминисцентные диоды (светодиоды)*. В диоде этого типа энергия, освобождаемая при рекомбинации электрона и дырки в *p-n*-переходе, проявляется в форме фотона. Достоинства светодиодов – малая постоянная времени, определенный и ограниченный спектр излучения, высокая надежность и прочность. Недостатки – слабый по величине поток и его зависимость от температуры.

- *лазеры*. Лазер – источник с очень высокой монохроматичностью и яркостью, обладает очень узкой направленностью и высокой когерентностью.

Выходной величиной фотоэлектрического датчика обычно является оптический ток, поэтому ток и его изменения в функции различных параметров чаще всего служат для характеристики свойств датчиков.

Темновой ток. Это – ток, постоянно протекающий на выходе фоточувствительного устройства, помещенного в темноту и находящегося при определенных условиях питания. Его причины – освобождение носителей заряда за счет теплового возбуждения и фотоэффект при чувствительности датчика к инфракрасному излучению окружающей среды. Следует так выбирать датчик, чтобы его темновой ток был существенно меньше минимального фототока, обусловленного сигналом, и помещать датчик в охлаждаемый корпус.

Чувствительность. Чувствительность датчика определяется как отношение изменения выходного тока ΔI к вызвавшему его изменению измеряемого светового потока ΔF . Для линейного датчика статическая чувствительность $S = I/F$.

Различают спектральную и интегральную чувствительность. *Спектральная* чувствительность – это чувствительность датчика, когда воспринимаемый им поток монохроматичен с длиной волны λ : $S(\lambda) = \Delta I / \Delta F(\lambda)$. *Интегральная* чувствительность – это чувствительность датчика, принимающего немонохроматический оптический сигнал. Она зависит как от спектральной чувствительности датчика к различным составляющим воздействующего излучения, так и от спектрального распределения самого излучения. Принимаемый интегральный поток

$$F(\lambda) = \int_{\lambda_1}^{\lambda_2} \frac{\Delta F(\lambda)}{\Delta \lambda} d\lambda ;$$

интегральный ток описывается формулой

$$I = \int_{\lambda_1}^{\lambda_2} S(\lambda) \frac{\Delta F(\lambda)}{\Delta \lambda} d\lambda ,$$

откуда легко получить теоретическое выражение для интегральной чувствительности. Очевидно, что интегральная чувствительность всегда меньше максимальной спектральной чувствительности.

Постоянная времени. Этот параметр характеризует быстроту, с которой сопротивление фоторезистора принимает новое значение при скачкообразном изменении воздействующего на него потока.

Обнаружительная способность. Она характеризует возможности датчика обнаруживать слабые сигналы при наличии собственных шумов. Источники шума всегда присутствуют в любой оптической системе. Самым распространенным является *тепловой* шум, связанный с тепловым возбуждением носителей заряда. *Генерационно-рекомбинационный* шум обусловлен флуктуациями скорости генерации и рекомбинации свободных носителей зарядов в полупроводниковых материалах, что приводит к случайным изменениям средних концентраций свободных носителей и, следовательно, к флуктуациям напряжения. *Дробовой* шум обусловлен флуктуациями тока, связанными с дискретностью носителей заряда и случайным характером электронной эмиссии.

В приемниках светового потока фотоэлектрических датчиков используется явление фотоэффекта – изменение свойств материала

при изменении его освещенности. Различают вентильный, внешний и внутренний фотоэффекты.

При **вентильном фотоэффекте** между слоями освещенного проводника и неосвещенного полупроводника, разделенными тонким изоляционным слоем, возникает ЭДС, которая зависит от освещенности.

Внешний фотоэффект состоит в том, что под влиянием потока излучения электроны вылетают из катода электронной лампы и ток эмиссии зависит от освещенности катода. Приемниками излучения на основе внешнего фотоэффекта являются электровакуумные или газонаполненные фотоэлементы, фотоэлектронные умножители и передающие электронно-лучевые трубки. Когда световой поток падает на катод, покрытый активным слоем, электроны получают энергию, позволяющую им вылететь из катода. Под действием источника питания с ЭДС E между катодом и анодом создается электрическое поле, которое и заставляет электроны перемещаться от катода к аноду. В электрической цепи создается электрический ток; когда действие света прекращается, ток в фотоэлементе и внешней электрической цепи исчезает.

В фотоэмиссионных устройствах фототок и освещенность связаны линейной зависимостью в широком диапазоне освещенности. Для достижения высокой линейности необходимо стабилизировать напряжение источника питания. От его стабильности зависит также постоянная времени прибора, которая может быть меньше 1 нс. Порядок величины спектральной чувствительности в максимуме – от 10 до 100 мА/Вт.

Внутренний фотоэффект проявляется в том, что активное сопротивление полупроводникового материала зависит от его освещенности. К приемникам излучения на основе внутреннего фотоэффекта относятся фоторезисторы, фотодиоды и фототриоды.

Фоторезисторы выполняются из полупроводниковых материалов, таких как CdS, CdSe, PbS, Ge, Si и других. Сопротивление фоторезистора R быстро падает с увеличением освещенности E (рис. 2.29). Сопротивление освещенного фоторезистора зависит от температуры, однако температурная чувствительность с ростом освещенности ослабевает. Максимальная рассеиваемая мощность

составляет около милливатта на 1 мм^2 чувствительной поверхности. Максимальная спектральная чувствительность в зависимости от материала фоточувствительного элемента составляет от 10^{-1} до 10^2 А/Вт при величине приложенного напряжения 10 В и площади приемной поверхности, равной 1 см^2 . Максимум спектральной чувствительности CdS приблизительно приходится на свет с длиной волны $500... 550 \text{ нм}$, что соответствует середине зоны чувствительности человеческого зрения.

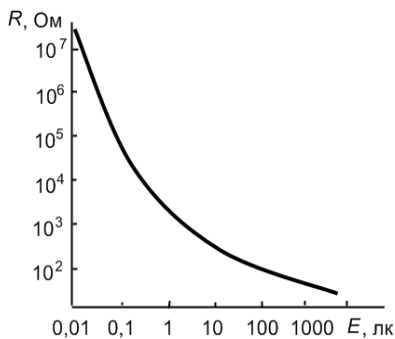


Рис. 2.29

Постоянная времени фоторезистора зависит от природы материала фоторезистора, способов его изготовления и амплитуды перепада потока ΔF , воздействующего на фоторезистор. Она составляет $0,1 \text{ мкс}$ для InSb и InAs, от 1 до 5 мкс для PbSe, от 40 до 1000 мкс для PbS и от 25 до 150 мс для CdS и CdSe. При воздействии на фоторезистор периодически изменяюще-

гося потока его чувствительность уменьшается. Этим изменением можно пренебречь в диапазоне до $10^3...10^4 \text{ Гц}$ (в зависимости от материала). Порядок величины удельной обнаружительной способности для монохроматического излучения в зависимости от материала составляет от 10^8 до $10^{11} \text{ (см} \cdot \text{Гц}^{1/2})/\text{Вт}$. Измерение сопротивления фоторезистора или обнаружение изменения его величины производится с помощью различного типа схем формирования сигналов, соединенных с резистивными датчиками: потенциометрической схемы, питаемой постоянным током, мостовой схемы, операционного усилителя и RC-генератора.

Фотодиоды работают в двух режимах: фотогальваническом и фотодиодном. В *фотогальваническом* режиме не требуется источник питания, поскольку при освещенности p - n -перехода появляется ЭДС, под действием которой возникает ток во внешней цепи. В этом режиме фотодиод непосредственно преобразует энергию света в электрическую энергию. При освещенности $8 \cdot 10^3 \text{ лк}$ фотоЭДС составляет около $0,1 \text{ В}$. Отличительные особенности фотогальва-

нического режима – линейная или логарифмическая зависимость выходного напряжения от нагрузки, минимальный фоновый шум, сравнительно большое время запаздывания (порядка сотен наносекунд) и уменьшенная полоса пропускания, а также существенная температурная зависимость характеристики при логарифмической реакции.

Пример реализации фотогальванического режима показан на рис. 2.30. Эта схема измеряет напряжение U_0 : выходное напряжение $U = (1 + R_2/R_1)U_0$. К этой же группе детекторов относятся лавинные фотодиоды, имеющие высокую чувствительность и высокое быстродействие (менее 0,5 мкс).

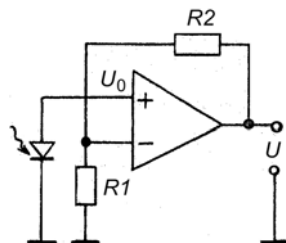


Рис. 2.30

В *фотодиодном* режиме к фотодиоду прикладывается напряжение обратной полярности; при освещении обратный ток резко увеличивается, фотодиод начинает проводить ток в обратном направлении. Спектральная чувствительность лежит в пределах от 0,1 до 1 мкА/мкВт. Временные характеристики зависят от параметров эквивалентных схем фотодиода и нагрузки и имеют порядок единиц наносекунд. Эти же параметры определяют частотную характеристику фотодиодного детектора.

Обнаружительная способность при нормальных условиях применения фотодиода лежит между 10^{10} и 10^{13} (см·Гц^{1/2})/Вт (в зависимости от его типа). Режим работы фотодиода выбирают в зависимости от требуемых характеристик и задачи проводимых исследований. Отличительные особенности фотодиодного режима – линейность, малое время запаздывания, широкая полоса пропускания.

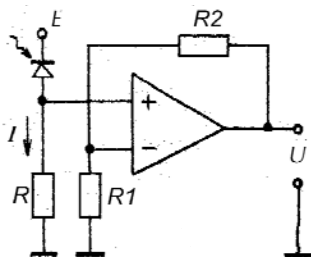


Рис. 2.31

Примером схемы, работающей в фотодиодном режиме, может служить схема, представленная на рис. 2.31. Выходное напряжение $U = R(1 + R_2/R_1)I$. Увеличение сопротивления R в этой схеме позволяет уменьшить шум, но ценой ухудшения быстродействия.

Большой входной импеданс усилителя позволяет избежать уменьшения эффективной нагрузки фотодиода.

Фототранзистор можно рассматривать как фотодиод, соединенный с транзистором; первый задает базовый ток, второй обеспечивает его усиление в β раз. Темновой ток фототранзистора при температуре 25 °С имеет порядок $10^{-8} \dots 10^{-9}$ А. Он зависит от напряжения коллектор-эмиттер и от температуры. Порядок величины спектральной чувствительности в максимуме составляет от 1 до 100 А/Вт. Численные значения временных характеристик зависят от типа фототранзистора, его рабочей точки и нагрузки и меняются в пределах от нескольких микросекунд до нескольких десятков микросекунд. Частотный диапазон – от десятков до сотен кГц.

Фототранзистор может использоваться либо как коммутатор, либо в линейном режиме. В коммутационном режиме его преимущество перед фотодиодом заключается в том, что можно непосредственно управлять относительно большими токами. В линейном режиме фотодиод имеет лучшую линейность, однако в ряде случаев полезным является усиительное свойство фототранзистора.

2.14. Струнные датчики

Для измерения неэлектрических величин применяется частотный метод, при котором измеряемая величина преобразуется в переменное напряжение, частота которого зависит от этой величины. Достоинство частотного метода состоит в том, что в процессе передачи информации по этому методу не возникает дополнительной погрешности.

Наибольшее распространение для преобразования неэлектрических величин в частоту получили струнные датчики. Принцип действия струнного датчика основан на зависимости собственной частоты колебаний f натянутой струны длиной l и массой m от силы натяжения F :

$$f = 0,5 \sqrt{F / ml} .$$

Струнный датчик (рис. 2.32) состоит из струны 1, возбудителя 2 и приемника 3. Одним концом струна жестко закреплена, а другим

концом соединена с первичным преобразователем, например мембраной 4. При изменении давления изменяется сила натяжения струны. С помощью возбuditеля, которым может быть электромагнит, струна выводится из состояния равновесия и начинает колебаться с частотой, определяемой давлением (поскольку длина струны зависит от него). Приемник преобразует перемещения струны в электрический сигнал. В качестве приемника обычно применяют электромагнитный датчик, так как он может попеременно выполнять функции то возбuditеля, то приемника. Когда на его обмотку подается напряжение, то он создает электромагнитную силу, и возбуждает ее колебания. Когда струна уже колеблется, то с этой обмотки снимается переменное напряжение, частота которого равна частоте колебаний струны.

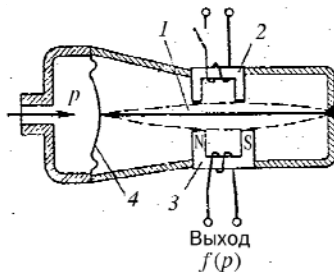


Рис. 2.32

Наибольшее распространение в струнных датчиках получили стальные струны из круглой проволоки диаметром 0,1...0,3 мм. При длине в 40...60 мм в таких струнах возбуждаются колебания с частотой 700...2000 Гц.

Струнные датчики обладают высокой чувствительностью и виброустойчивостью, однако в них трудно обеспечить линейность: нелинейность составляет обычно 2...3 % от диапазона изменения частоты.

2.15. Датчики потока

Под измерением потока подразумевается либо измерение средней или локальной скорости потока, либо измерение массового или объемного расхода жидкости (газа). Устройства, измеряющие скорость потока, называют *анемометрами*; устройства, измеряющие расход – *расходомерами*.

Скорость обычно измеряют косвенно, определяя воздействие, которое она оказывает: а) на физические характеристики чувствительного элемента; б) на определенные физические явления, объектом которых является чувствительный элемент.

Если чувствительный элемент является элементом датчика, помещенного в жидкость, то показателем ее скорости может служить: а) температура и, следовательно, сопротивление нагретой нити, питаемой постоянным током; б) скорость вращения помещенной в поток крыльчатки.

2.15.1. Тепловые анемометры

В тепловых измерителях скорости используется нагреваемый элемент, устанавливаемый на пути потока и обтекаемый этим потоком. Тепло передается от этого элемента к текучей среде с интенсивностью P , определяемой разностью температур ΔT элемента и текучей среды, удельной теплоемкостью c и скоростью v последней, а также профилем потока. На принципе передачи тепла от нагреваемого элемента в поток основаны два метода измерения потока. В *конвекционном* методе измеряется количество тепла, рассеиваемого нагреваемым элементом, тогда как в методе *стационарной тепловой инжекции* определяется изменение температуры текучей среды, связанное с инъекцией тепла в поток.

Конвекционный измеритель потока обеспечивает измерение локальной скорости жидкости или газа путем измерения количества тепла, которое рассеивает нагреваемый элемент, обтекаемый потоком. В качестве такого элемента можно использовать любой резистивный нагреватель. Он нагревается выше температуры текучей среды электрическим током. Мощность теплопередачи в этом случае приближенно задается законом Кинга:

$$P = A (T_n - T_n)(C_0 + C_1 v^{1/2}), \quad (2.2)$$

где P – мощность теплопередачи (Вт); A – эффективная площадь нагреваемого элемента (м^2); T_n – температура нагревательного элемента ($^{\circ}\text{C}$); T_n – температура потока жидкости или газа ($^{\circ}\text{C}$); v – скорость жидкости или газа (м/с); C_0, C_1 – эмпирические коэффициенты, определяемые структурой нагреваемого элемента и удельной теплоемкостью текучей среды.

Если элемент и текучая среда находятся в тепловом равновесии, то количество тепла, ежесекундно передаваемого в поток, равно джоулевой мощности, выделяемой в элементе:

$$P = I^2 R, \quad (2.3)$$

где P – выделяемая в элементе мощность (Вт); I – электрический ток через элемент (А); R – электрическое сопротивление элемента (Ом).

Из соотношений (2.2) и (2.3) следует:

$$I^2 = [A (T_n - T_0)(C_0 + C_1 v^{1/2})]/R.$$

Обычно поддерживается постоянной величина сопротивления R , при этом, очевидно, постоянна и температура резистивного элемента. Зная температуру текучей среды, можно найти из последнего уравнения скорость как функцию тока.

На рис. 2.33 показана простая схема, которая поддерживает температуру (и сопротивление) проволочного элемента на постоянном уровне.

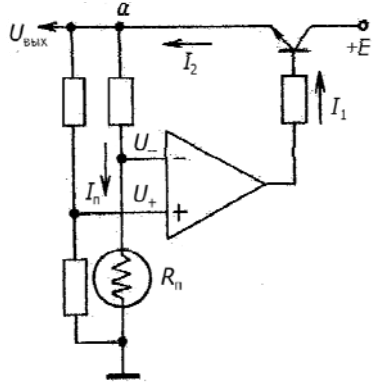


Рис. 2.33

Предположим, например, что в исходном состоянии резистивный мост сбалансирован, так что $U_- = U_+$ при данной скорости потока. Пусть теперь скорость потока возрастает, вызывая охлаждение накаливаемой проволоки. При этом уменьшается ее сопротивление R_n , что приводит к уменьшению напряжения U_- . Напряжение на выходе ОУ, а с ним и ток I_1 возрастают, в результате чего возрастает ток I_2 . Увеличение тока I_2 приводит к увеличению тока I_n через накаливаемую проволоку зонда; проволока нагревается до тех пор, пока не восстанавливается баланс моста. Но если полное сопротивление моста между точкой U_0 и землей остается постоянным, то, значит, напряжение U_0 пропорционально току I_n и его измерение позволяет рассчитать скорость потока.

2.15.2. Инжекционные расходомеры

Средний массовый расход текучей среды можно определить путем инъекции в поток известного количества тепла и измерения температуры этой среды за нагревателем.

Средний массовый расход рассчитывается по формуле

$$F = q/c(T_2 - T_1),$$

где F – массовый расход (кг/с); q – скорость стационарной инъекции тепла (Вт); c – удельная теплоемкость среды (Дж/кг·град); T_1 – температура среды перед нагревателем (°С); T_2 – температура среды за нагревателем (°С). Значения температур можно измерить с помощью термисторов или термопар.

Метод стационарной тепловой инъекции весьма прост, но его точность на практике оказывается низкой по следующим причинам: 1) могут происходить паразитные утечки тепла, например, через стенки теплопровода; 2) датчик температуры T_1 должен находиться достаточно далеко от нагревателя, там, где устанавливается однородное распределение температуры, но это еще больше усложняет проблему паразитных утечек тепла.

2.15.3. Механические расходомеры и анемометры

Механические измерители потока работают на принципе физического воздействия жидкости или газа на вращающийся элемент (например, на крыльчатку турбины или пропеллер) или на элемент, тормозящий поток.

Турбины. Турбинные измерители потока измеряют средний объемный расход жидкости или газа. При турбулентном движении среды скорость n (об/с) вращения крыльчатки (колеса) турбины, занимающей полное сечение теплопровода, зависит от среднего объемного расхода Q (м³/с); диаметра теплопровода D (м) и кинематической вязкости текучей среды η (м²/с):

$$Q/nD^3 = f(nD^2/\eta).$$

Эта зависимость показана на рис. 2.34. Как видно из графика, учет конкретного вида функции f важен только при низких скоростях вращения турбины. Обычно она работает в линейной области, где

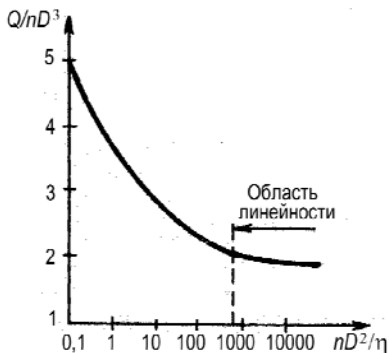


Рис. 2.34

величина Q/nD^3 постоянна, так что скорость вращения турбины прямо пропорциональна среднему объемному расходу жидкости или газа.

Пропеллеры. Принцип их работы ничем не отличается от принципа турбинного измерителя, однако, поскольку пропеллер не охватывает все сечение потока, скорость его вращения пропорциональна скорости потока. Скорость вращения измеряется с помощью оптических или магнитных датчиков. При соответствующей калибровке вращающиеся измерители обеспечивают достаточно точное измерение (0,1...0,5 %). Однако эти устройства не обладают необходимой механической прочностью и быстро выходят из строя при измерении мощных потоков и наличии в потоке посторонних твердых частиц.

Измерители потока с тормозящим элементом. Действие преобразователей данного типа основано на использовании кинетической энергии движущейся струи. Если в поток опустить тормозящий его элемент, закрепленный на пружинящем рычаге, то сила, действующая на этот элемент и равная

$$F_T = C_T S \rho v^2 / 2,$$

где C_T – коэффициент торможения; S – площадь поперечного сечения элемента (m^2); ρ – плотность текучей среды (kg/m^3); v – скорость потока (m/s), отклонит его, что приведет к соответствующему изгибу рычага. Если на нем закрепить тензодатчики, то эту силу и, следовательно, скорость можно измерить. Симметричные тормозящие элементы этого типа являются двунаправленными измерителями; они хорошо работают вблизи точки изменения направления потока. Измерители с тормозящим элементом – прочные устройства. Они часто используются в тех случаях, когда текучая среда содержит взвешенные твердые частицы. С их помощью можно измерять скорость потока как жидкости, так и газа.

2.15.4. Гидродинамические (аэродинамические) анемометры и расходомеры

Трубки Пито. Этот прибор относится к группе преобразователей скоростного напора и обеспечивает измерение локальных скоростей жидкости или газа в потоке. В трубке (рис. 2.35) имеется

два типа отверстий, открытых для текущей среды. На статических входах, выполненных в виде отверстий, оси которых перпендикулярны направлению потока, действует статическое давление текущей среды p_c . Вход на конце трубки собирает заторможенную массу жидкости или газа, находящуюся под полным давлением p_n . Если трубка Пито расположена параллельно потоку, то разность этих двух давлений p (Па) находится из уравнения Бернулли:

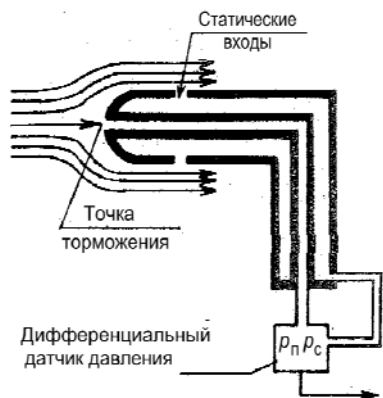


Рис. 2.35

где ρ – плотность текущей среды (кг/м^3); v – скорость потока (м/с).

$$p = p_n - p_c = \rho v^2 / 2,$$

где ρ – плотность текущей среды (кг/м^3); v – скорость потока (м/с).

Трубка Пито чаще всего применяется для измерения скорости воздуха в вытяжных трубах и на самолетах. Это очень точное и прочное устройство, требующее минимального технического обслуживания. Главный недостаток трубки Пито – низкая чувствительность при малых скоростях потока и нелинейность связи между разностью давлений и скоростью.

Преобразователи давления. Устройства этого типа определяют объемный расход ($\text{м}^3/\text{с}$) жидкости или газа. В потокопроводе размещается препятствие с известными характеристиками, которое изменяет поперечное сечение потока. Для любого типа сужающего устройства скорость потока перед ним уменьшается, давление же, обусловленное скоростным напором, растет. За сужающим устройством – обратная картина. На основании уравнения Бернулли расход Q и перепад давлений p связаны соотношением:

$$Q = k (p/\rho)^{1/2},$$

где k – коэффициент, зависящий от формы сужения и физических свойств среды.

На практике применяют две модификации расходомеров с сужающими устройствами: 1) расходомеры переменного давления, где используется зависимость $Q = f(p)$ при $k = \text{const}$; статическая характеристика в этом случае имеет явно выраженный нелинейный характер; 2) расходомеры постоянного давления, когда поддержи-

вается неизменным перепад давления $p = \text{const}$; статическая характеристика линейна и имеет вид: $Q = f(k)$.

Наиболее простыми и дешевыми СУ являются диафрагмы, однако они больше всего подвержены влиянию твердых включений в потоке, что приводит к изменению коэффициента k и, как следствие, к трудно учитываемому росту погрешности. Кроме того, при использовании для измерений диафрагм сталкиваются с существенными (от 40 до 80 %) потерями давления, обусловленными турбулентностью. Сопло меньше подвержено влиянию внешних факторов типа твердых включений на точность при одинаковой с диафрагмой потере давления.

Ротаметр. Это – типичный расходомер постоянного давления. Он представляет собой трубку, расширяющуюся в направлении снизу вверх, в которую помещен поплавок конической формы, максимальный диаметр которого равен внутреннему диаметру трубки в ее нижней части. Жидкость или газ движутся по трубке снизу вверх, увлекая за собой поплавок. По мере его подъема кольцевой зазор между его краями и внутренней поверхностью трубки увеличивается до тех пор, пока подъемная сила, создаваемая разностью давлений снизу и сверху поплавка, не станет равна его весу.

Положение поплавка линейно зависит от расхода. Для минимизации трения с потоком поплавков имеет спиральную нарезку, что способствует его вращению в потоке для сохранения центрального положения относительно корпуса. Обычно погрешность ротаметров составляет 2...3 %. На точность также влияют изменения плотности и вязкости среды, которые являются причинами изменения подъемной силы и силы трения. Ротаметры используются для измерения малых расходов жидкостей и газов (10^{-7} ... 10^{-6} м³/с). Если верхнюю часть поплавка использовать как плунжер дифференциального трансформатора, то можно в качестве выходного сигнала получить напряжение.

2.16. Антенны

Известно, что если на пути распространения радиоволн встречается какое-либо тело с электрическими параметрами, отличными от параметров среды, в которой распространяются радиоволны, то

это тело, облученное электромагнитной энергией, становится источником вторичного электромагнитного поля, что упрощенно можно рассматривать как явление отражения электромагнитных волн от тел. Аппаратуру, обеспечивающую обнаружение и определение координат объектов, называют *радиолокационной станцией* (РЛС).

Передающая антенна РЛС излучает электромагнитные волны в окружающее пространство. Если на пути распространения электромагнитной волны находится какой-либо объект, то часть энергии отражается обратно, в сторону РЛС. С помощью чувствительного приемника можно обнаружить эту волну, а следовательно, и объект. По времени $t_d = 2D/c$ запаздывания отраженного сигнала относительно излученного определяем наклонную дальность цели $D = ct_d/2$ (здесь $c = 3 \cdot 10^8$ м/с – скорость света). Обзор пространства производится путем кругового механического вращения антенны.

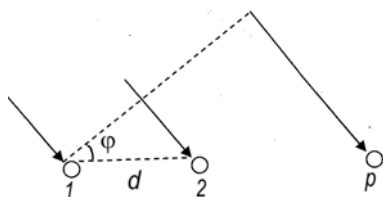


Рис. 2.36

Пусть теперь радиолокационная антенна имеет вид прямолинейной решетки из p вибраторов, отстоящих один от другого на расстоянии d (рис. 2.36). Значительная удаленность цели от РЛС позволяет считать, что лучи, идущие от цели к

вибраторам, направлены параллельно под углом φ к решетке, а амплитуды ЭДС, наводимых в отдельных вибраторах, равны между собой. В этих условиях ЭДС соседних вибраторов отличаются только сдвигом по фазе ψ , вызванным разностью хода волн $d \sin \varphi$:

$$\psi = 2\pi d \sin \varphi / \lambda c.$$

Изменяя величину сдвига по фазе, можно «настроить» антенну на другие углы φ , «поворачивая» ее луч в пространстве при постоянстве расположения вибраторов.

Аналогичные принципы используются и при построении гидроакустических станций – средств, основанных на использовании явления распространения акустических волн в океанах, морях и других водоемах. В качестве отдельных преобразователей в гидроакустических антеннах применяют пьезокерамические элементы.

Возможность применения радиолокационных и гидроакустических станций для решения конкретных задач определяется путем анализа ее основных показателей и параметров:

- зоной обзора – областью пространства, в которой происходит обнаружение целей;

- максимальной дистанцией обнаружения – наибольшим расстоянием между станцией и целью, при котором соблюдаются заданные вероятности правильного обнаружения и ложной тревоги;

- радиус мертвой зоны – минимальное расстояние от станции, на котором можно обнаружить цель и определить ее координаты с заданной точностью;

- разрешающей способностью – возможностью раздельно наблюдать близко расположенные цели;

- точностью измерения координат, характеризуемой случайными ошибками измерения;

- помехоустойчивостью станции – способностью станции выполнять свои функции в условиях преднамеренно создаваемых помех;

- надежностью – способностью станции работать в течение требуемого времени с сохранением всех заданных показателей.

Контрольные вопросы и упражнения

1. Постройте зависимость напряжения от перемещения движка потенциометра, учитывающую замыкание двух соседних витков обмотки.

2. Какую форму имеет сигнал на выходе линейного дифференциального трансформатора?

3. В чем заключается основное отличие индукционных датчиков от индуктивных?

4. Объясните принцип действия электромагнитных датчиков.

5. В чем заключается тензорезистивный эффект?

6. Как компенсировать температуру «холодного» спая термопары?

7. Для компенсации холодного спая термопары требуется еще один датчик температуры. Объясните, какой же смысл применять термопару.

8. Разработайте схему датчика уровня на основе терморезистора с критичной характеристикой.

9. По какому закону изменяется сопротивление термисторов?

10. Как вы представляете себе конструкцию ротаметра? Разработайте блок-схему измерительного блока с ротаметром.
11. Объясните принцип измерения потока тепловым анемометром.
12. Как изменяется собственная частота струнного датчика в зависимости от длины струны?
13. Как осуществляется поворот луча в фазированной антенной решетке?
14. Какие из рассмотренных в этом разделе датчиков относятся к параметрическим? к генераторным?
15. Предложите вариант блок-схемы системы для выполнения простой градуировки датчика.
16. Какие математические методы следует использовать при построении характеристической кривой датчика? при определении значения физической величины по выходному сигналу датчика?

3. АНАЛОГОВЫЕ СХЕМЫ ПРЕДВАРИТЕЛЬНОЙ ОБРАБОТКИ СИГНАЛОВ

Перед вводом различных сигналов от датчиков в ЭВМ часто требуется их предварительная обработка, которая упрощает и облегчает дальнейшие операции над этими сигналами. Характер предварительной обработки зависит от вида датчика и от параметров его выходного сигнала. В состав схем предварительной обработки сигналов могут входить следующие функциональные устройства:

- 1) усилители, предназначенные для усиления маломощных сигналов датчиков;
- 2) преобразователи сопротивления, преобразующие высокое выходное сопротивление датчика в низкое;
- 3) преобразователи тока в напряжение;
- 4) преобразователи действующих значений – для преобразования выходного сигнала датчика в виде переменного тока в сигнал постоянного тока с тем же действующим значением;
- 5) схемы логарифмического сжатия, назначение которых – сжатие динамического диапазона выходных сигналов датчика с помощью логарифмического усилителя;
- 6) фильтры, функция которых – исключение эффекта наложения спектров и подавление шумовых составляющих выходного сигнала датчика;
- 7) схемы линеаризации, предназначенные для коррекции выходного сигнала датчика, имеющего нелинейную характеристику;
- 8) регуляторы усиления, выполняющие функцию переключения коэффициента усиления для обеспечения максимальной чувствительности измерительного тракта.

Рассмотрим основные принципы некоторых схемотехнических решений.

3.1. Операционные усилители

Операционный усилитель (ОУ) – это дифференциальный усилитель постоянного тока с очень большим коэффициентом усиления

ния (обычно $10^5 \dots 10^6$). На рис. 3.1 показана упрощенная эквивалентная схема ОУ. В представленной модели ОУ имеет входное сопротивление $R_{\text{вх}}$, включенное между двумя его входами. Выходная цепь состоит из управляемого источника (генератора) напряжения и включенного последовательно с ним выходного сопротивления $R_{\text{вых}}$. Из-за разности напряжений между входами возникает ток через сопротивление $R_{\text{вх}}$. Дифференциальное входное напряжение $u_2 - u_1$ умножается на коэффициент усиления K ; в результате получается величина выходного напряжения.

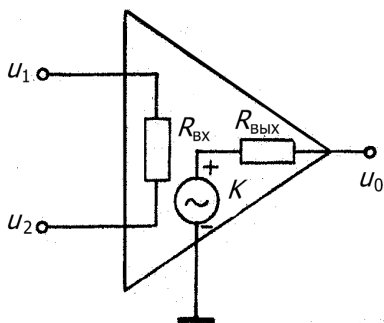


Рис. 3.1

Поскольку ОУ обладает чрезвычайно большим коэффициентом усиления по напряжению, даже малое изменение дифференциального напряжения (порядка долей милливольт) вызывает изменение выходного напряжения в пределах его полного диапазона. Очевидно, что $u_2 - u_1 = u_0/K$. Если u_0 имеет конечное значение, а коэффициент усиления K велик, то $u_2 - u_1 \approx 0$ и

$u_1 \approx u_2$. Входное сопротивление $R_{\text{вх}}$ очень велико, поэтому током через $R_{\text{вх}}$ можно пренебречь. Отсюда следуют две особенности процесса функционирования ОУ:

- 1) при работе ОУ в линейной области на двух его входах действуют одинаковые напряжения;
- 2) входные токи для обоих входов ОУ равны нулю.

Кроме того, для упрощения расчетов и анализа схем предполагается, что идеальный ОУ имеет следующие характеристики:

- коэффициент усиления, при разомкнутой обратной связи равен бесконечности;
- входное сопротивление, равное бесконечности;
- выходное сопротивление, равное нулю;
- ширину полосы пропускания, равную бесконечности.

Рассмотрим, в чем отличие характеристик реальных ОУ от идеальных.

Входной ток. Небольшой ток, называемый *входным током смещения*, протекает через входные цепи ОУ и равен половине суммы входных токов, измеренных при соединении входов между собой (это базовые токи входных транзисторов). Роль входного тока смещения состоит в том, что он создает падение напряжения на резисторах цепей обратной связи и смещения, а также на сопротивлении источника сигнала. От величины этих сопротивлений зависит влияние тока смещения на параметры схемы по постоянному току и отклонения выходного напряжения.

Промышленность выпускает ОУ с входными токами смещения порядка наноампер и ниже для схем с входами на биполярных транзисторах и порядка нескольких пикоампер для схем со входами на полевых транзисторах.

Разность двух входных токов называют *входным током сдвига*. В отличие от входного тока смещения ток сдвига обусловлен отклонениями в технологическом процессе изготовления ОУ. Обычно ток сдвига составляет примерно одну десятую часть тока смещения.

Входной импеданс. Эта величина определяется входным сопротивлением для дифференциального сигнала (импеданс со стороны одного из входов при заземлении другого) и составляет примерно 10^{12} Ом для ОУ со входом на полевом транзисторе и 2 МОм – со входом на биполярном транзисторе.

Входной диапазон сигнала. Для того чтобы ОУ работал правильно, синфазное напряжение на его входах должно находиться в пределах определенного диапазона значений, который обычно не превышает полного диапазона напряжения питания. Это же относится к дифференциальным входным сигналам, однако для некоторых ОУ допустимое напряжение между входами ограничено диапазоном 0,5 В.

Выходное сопротивление – это собственное выходное сопротивление ОУ без обратной связи. Оно лежит в пределах от нескольких десятков до нескольких тысяч ом. Обратная связь делает выходное сопротивление пренебрежимо малым, поэтому большое значение имеет максимально допустимый выходной ток, равный, как правило, примерно 20 мА.

Входное напряжение сдвига. Технологические погрешности приводят к тому, что входные каскады ОУ имеют разбалансировку. Разность входных напряжений, необходимая для того, чтобы выходное напряжение стало равно нулю, называют *входным напряжением сдвига*. Обычно в ОУ предусматривается возможность уменьшения входного напряжения сдвига до нуля (установка нуля).

Для точных систем большое значение имеет дрейф входного напряжения сдвига под влиянием температуры и времени. Типичные значения: для напряжения сдвига – 1 мВ, для температурного дрейфа – 5 мкВ/°С.

Скорость нарастания выходного напряжения. Наличие емкостей и небольшие внутренние токи ограничивают скорость изменения выходного напряжения даже при условии большого разбаланса входов. Предельную скорость изменения выходного напряжения называют *скоростью нарастания*. У маломощных ОУ она не превышает 1 В/мкс, быстродействующие ОУ имеют скорость нарастания порядка 100 В/мкс, а сверхбыстродействующие – до 6000 В/мкс. Для синусоидального сигнала, частота которого равна ω (Гц), а амплитуда – A (В), минимальная скорость нарастания должна составлять $A\omega$ (В/с).

Шумы усилителя. Для построения прецизионной аппаратуры большое значение имеет уровень собственных шумов усилителя. Источники шумов – это тепловой шум резисторов; шумовое напряжение, которое возникает из-за флуктуации сопротивления при протекании через него «внешнего» тока; дробовой шум, возникающий как следствие квантованности зарядов.

В большинстве схем ОУ применяется в конфигурации с замкнутой обратной связью. Она снижает его коэффициент усиления, но обеспечивает реализацию ряда полезных свойств и характеристик, которые используются в устройствах предварительной обработки сигналов.

Инвертирующий усилитель. На рис. 3.2 представлена схема такого усилителя. Проведем ее анализ. Потенциал точки B равен потенциалу земли, следовательно, потенциал точки A также равен потенциалу земли. Это означает, что: а) падение напряжения на

резисторе R_2 равно $u_{\text{ВЫХ}}$; б) падение напряжения на резисторе R_1 равно $u_{\text{ВХ}}$. Поскольку входной ток ОУ равен нулю (см. выше), справедливо равенство:

$$u_{\text{ВЫХ}}/R_2 = u_{\text{ВХ}}/R_1,$$

откуда следует, что коэффициент усиления по напряжению

$$K = u_{\text{ВЫХ}}/u_{\text{ВХ}} = -(R_2/R_1).$$

Обычно коэффициент усиления по напряжению для постоянного тока лежит в пределах от 10000 до 100000.

Поскольку инвертирующий вход ОУ практически заземлен, входным сопротивлением является R_1 . При его увеличении уменьшается коэффициент усиления, который можно повысить, увеличивая R_2 , однако, поскольку возможности увеличения сопротивлений ограничены, входное сопротивление инвертирующего усилителя обычно невелико.

Неинвертирующий усилитель. Рассмотрим схему на рис. 3.3. Напряжение u_A снимается с делителя напряжения:

$$u_A = u_{\text{ВЫХ}}R_1/(R_1+R_2).$$

Поскольку $u_A = u_{\text{ВХ}}$, то

$$K = u_{\text{ВЫХ}}/u_{\text{ВХ}} = 1 + R_2/R_1.$$

Входной импеданс этого усилителя превышает 10^8 Ом, выходной равен долям ома.

Дифференциальный усилитель. На рис.3.4 показана принципиальная схема простого дифференциального усилителя на одном ОУ. Проведем ее анализ. Поскольку входные токи для обоих входов ОУ равны нулю, ток, определяемый напряжением u_2 , протекает через резисторы R_1 и R_2 на землю. Делитель напряжения на этих резисторах определяет величину напряжения на неинвертирующем входе ОУ:

$$u_3 = u_2R_2/(R_1 + R_2).$$

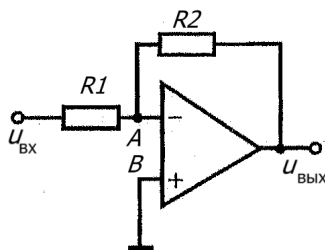


Рис. 3.2

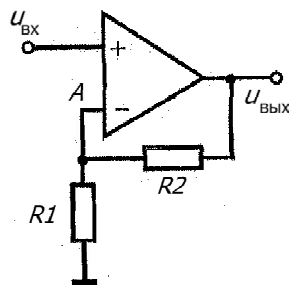


Рис. 3.3

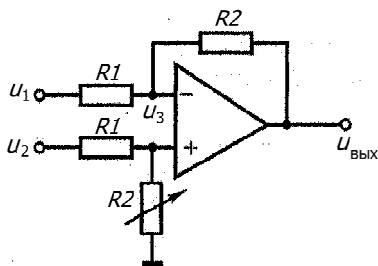


Рис. 3.4

Напряжение на инвертирующем входе ОУ должно быть равно напряжению на его неинвертирующем входе. Следовательно, верхняя половина схемы действует как инвертирующий усилитель. Рассчитывая ток в ней так, как это было сделано для инвертирующего усилителя, получаем:

$$i = (u_1 - u_3)/R_1 = (u_3 - u_{\text{ВЫХ}})/R_2.$$

Подставляя соотношение для u_3 , получаем окончательно:

$$u_{\text{ВЫХ}} = (u_2 - u_1)R_2/R_1.$$

Полученное соотношение определяет коэффициент усиления дифференциального усилителя. Если на обоих входах такого усилителя действует одинаковое напряжение, то теоретически его выходное напряжение равно нулю; если напряжения различны, то коэффициент усиления дифференциального (разностного) напряжения равен R_2/R_1 .

Реальный дифференциальный усилитель не обеспечивает полного подавления синфазного напряжения. Количественной мерой способности дифференциального усилителя подавлять синфазное напряжение служит *коэффициент ослабления синфазного сигнала* (КОСС), определяемый как отношение коэффициента усиления дифференциального напряжения к коэффициенту усиления синфазного напряжения. Величина КОСС может находиться в пределах от 100 до 10000 и обычно измеряется в децибелах.

Дифференциальный усилитель на одном ОУ имеет низкое входное сопротивление; его основное назначение – подавление помех, общих для обоих его входов.

В особую группу выделяют **инструментальные** усилители, представляющие собой структуру из трех операционных усилителей. Характерным примером такого усилителя может служить схема INA118 фирмы BURR–BROWN. В основе этой схемы – дифференциальный усилитель, выполненный на одном ОУ, но для увели-

чения входного сопротивления на двух его входах включены ОУ с защитой от перегрузки при помощи полевых транзисторов. Они обеспечивают малое последовательное сопротивление при нормальных рабочих условиях, а в случае приложения избыточного входного напряжения (более +4 В) ограничивают входной ток на уровне 1,5...5 мА. Коэффициент усиления такого усилителя регулируется внешним сопротивлением в диапазоне от 1 до 10000.

Если усилитель используется для усиления сигнала с большим динамическим диапазоном или для усиления нескольких сигналов, поступающих от разных датчиков и имеющих существенно разные амплитуды, то целесообразно использовать инструментальные усилители с программируемым коэффициентом усиления. Он устанавливается путем изменения параметров обратной связи входных ОУ по сигналам процессора. Например, в инструментальных усилителях AD620 и AD625 фирмы Analog Device программируемый коэффициент усиления лежит в диапазоне 1...10000.

Повторитель напряжения. Если в инвертирующем усилителе (см. рис. 3.2) положить $R_1 = \infty$ и $R_2 = 0$, то входное напряжение будет непосредственно передаваться на выход схемы, т.е. такой усилитель будет повторять входное напряжение. Повторитель напряжения иногда называют *буферным* усилителем или *преобразователем сопротивления*, так как он имеет большой входной и малый выходной импеданс и, следовательно, обладает изолирующими свойствами.

Компаратор. Это специализированный ОУ с дифференциальным входом и парафазным цифровым выходом. Входной каскад компаратора построен аналогично схемам ОУ и работает в линейном режиме. На выходе компаратора формируются сигналы высокого логического уровня, если разность входных сигналов меньше напряжения срабатывания компаратора, или низкого логического уровня, если разность входных сигналов превышает напряжение срабатывания компаратора (рис. 3.5).

Основные параметры компаратора: чувствительность (точность, с которой компаратор может различать два входных сигнала), бы-

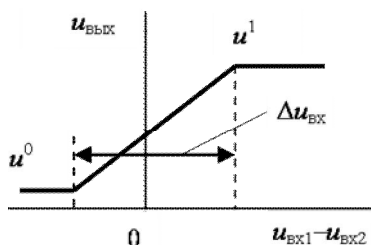


Рис. 3.5

стродействие (скорость отклика, определяемая задержкой срабатывания и временем нарастания сигнала), нагрузочная способность (способность компаратора управлять определенным числом входов цифровых микросхем).

Чувствительность компаратора измеряется единицами милливольт; быстродействие в зависимости от типа схемы лежит в диапазоне от единиц наносекунд до единиц микросекунд; нагрузочная способность зависит от выходного каскада схемы компаратора: для ее повышения включается эмиттерный повторитель.

3.2. Унифицирующие измерительные преобразователи

Естественные сигналы, поступающие от датчиков, различаются по динамическим диапазонам и по мощности, могут быть сигналами напряжения и сигналами тока. В соответствии с ГОСТ 26.011-80 сигналы *напряжения* – это сигналы, поступающие от источника с выходным сопротивлением, значение которого много меньше входного сопротивления приемника сигнала; сигнал *тока* – это сигнал, поступающий от источника с выходным сопротивлением, значение которого много больше входного сопротивления приемника сигнала.

Наиболее существенное соображение, которое нужно принимать во внимание, – сигнал должен быть малочувствительным к электрическим возмущениям. При передаче сигнала напряжением нужно учитывать, что входной импеданс подключенного устройства – не бесконечность, и поэтому по линии связи будет протекать ток. Следовательно, возникнет падение напряжения. Если амплитуда сигнала будет изменяться, то потечет ток перезарядки распределенных емкостей и возникнут помехи. Один из способов передачи сигналов напряжением является организация трехпроводной системы.

Главная причина популярности напряжения как носителя сигнала – простота и доступность устройств для обработки сигнала. Однако напряжение не слишком пригодно для передачи данных на большие расстояния.

Для передачи сигналов на значительные расстояния лучше воспользоваться током, так как он постоянен по всей длине кабеля. На конце линии токовый сигнал преобразуется в напряжение с помощью высокоточного шунтирующего резистора. На передающем конце выходное напряжение датчика преобразуется операционным усилителем в ток. Токовые сигналы, как правило, используются на низких частотах до 10 Гц. При переменном токе становится заметным влияние емкостного эффекта: часть тока уходит либо в обратный провод, либо в заземленный экран.

Измерительная система, использующая для передачи сигнала ток и датчик, гальванически изолированный от выходного сигнала, имеет следующие преимущества:

- достаточно хорошо работает на длинных линиях;
- допускает простую процедуру проверки, если допустимый ток ограничен по минимуму, например, 4 мА; если ток меньше, это значит, что линия отключена или разомкнута;
- обеспечивает хорошую защиту от помех;
- обеспечивает передачу сигнала по двум проводам.

Для согласования источников информации с другими устройствами измерительно-вычислительного комплекса их сигналы должны быть унифицированы, т.е. должны быть преобразованы так, чтобы соответствовать определенным требованиям.

Эти требования установлены ГОСТ 26.011-80 и ГОСТ 26.013-81. В качестве носителя информации в них приняты электрические сигналы постоянного и переменного тока и напряжения, причем оговорены следующие диапазоны их изменения:

- для сигналов тока $-5...+5$ мА (на сопротивлении нагрузки $R_{нагр} = 2,5$ кОм); $0...+20$ мА ($R_{нагр} = 1,0$ кОм); $+4...+20$ мА ($R_{нагр} = 1,0$ кОм); $-20...+20$ мА ($R_{нагр} = 1,0$ кОм); $-20...+20$ мА ($R_{нагр} = 1,0$ кОм); $-100...+100$ мА ($R_{нагр} = 250$ Ом);

- для сигналов напряжения : $0 \dots 1$ В; $-1 \dots +1$ В; $0 \dots 5$ В ($R_{\text{нагр}} = 1,0$ кОм); $1 \dots 5$ В ($R_{\text{нагр}} = 1,0$ кОм); $0 \dots 10$ В ($R_{\text{нагр}} = 2,0$ кОм) и т.д.

Для преобразования естественных выходных сигналов датчиков в унифицированные применяются унифицирующие измерительные преобразователи (УИП). Различают индивидуальные, групповые и многоканальные УИП. Индивидуальные следует использовать для унификации сигналов при сравнительно небольшом числе измеряемых параметров и при ограниченном времени измерения. Они позволяют преобразовывать один унифицированный сигнал в другой, обеспечивать гальваническую развязку входных цепей, размножать входной сигнал по нескольким выходам.

Групповые УИП обслуживают группу датчиков, выходные сигналы которых однородны. Групповой УИП должен применяться вместе с коммутатором и его параметры должны настраиваться в соответствии с параметрами выходного сигнала подключенного датчика.

Многоканальные УИП представляют собой конструктивно объединенные индивидуальные УИП.

Унифицирующие измерительные преобразователи могут выполнять как линейные, так и нелинейные преобразования сигналов. К линейным относятся операции масштабирования и установки нуля, к нелинейным – линеаризация.

Масштабирование и установка нуля. Если выходной сигнал датчика возрастает от y_1 до y_2 , а унифицированный сигнал на выходе УИП должен лежать в пределах от нуля до z ($z > y_2 - y_1$), то для совмещения начала динамических диапазонов УИП и датчика к сигналу датчика должен быть добавлен сигнал $-y_1$, а затем суммарный сигнал должен быть усилен в $K = z/(y_2 - y_1)$ раз. Величина y_1 может изменяться, например, при изменении температуры. В этом случае ее снимают с первичного преобразователя.

Линеаризация. Характеристическая кривая датчика $y = F(x)$, как правило, нелинейна. Для упрощения алгоритмов обработки можно использовать схемные методы линеаризации, реализовав с помощью УИП операцию $x = F^{-1}(y)$. Пусть, например, датчик реа-

лизует функцию $y = ax^2$. Если УИП выполняет операцию $z = ky$, то для определения искомой величины x следует вычислить $x = \sqrt{z/ak}$. Если же коэффициент преобразования УИП выбран так, что $z = \sqrt{y/a} = x$, то никаких вычислений делать не нужно.

Линеаризация чаще всего обеспечивается кусочно-линейной аппроксимацией функции преобразования и построением ОУ с управляемой обратной связью.

3.3. Фильтры

Важным элементом измерительного тракта является схема фильтрации. Ее назначение – изменение частотных характеристик входного сигнала, что необходимо для уменьшения влияния шумовой составляющей на результат его обработки и для исключения эффекта наложения спектров при его дискретизации.

Если сигнал и помеха имеют разный частотный состав, то для уменьшения влияния помехи достаточно пропустить смесь сигнала и помехи через фильтр, который «вырежет» из нее информативную полосу частот.

Эффект наложения спектров заключается в следующем. Как известно, спектр $S_T(\omega)$ дискретизованного сигнала представляет собой последовательность спектров $S(\omega)$ исходного сигнала $s(t)$, сдвинутых относительно друг друга на $\omega_0 = 2\pi/T$ и убывающих по закону

$$\sin\left(\frac{n\pi\tau_0}{T}\right) / \frac{n\pi\tau_0}{T} = \operatorname{sinc}\left(\frac{n\pi\tau_0}{T}\right),$$

где T – шаг дискретизации, τ_0 – длительность импульсов синхронизации. Если шаг дискретизации выбран так, что $\omega_m < \pi/T$, т.е. $T < \pi/\omega_m < f_m/2$, то отдельные спектры не перекрываются (рис. 3.6, а, б).

При пропускании такого спектра через четырехполосник с П-образной амплитудно-частотной характеристикой (идеальный фильтр нижних частот) с полосой пропускания $(-\pi/T, +\pi/T)$ (рис. 3.6, в) из спектра дискретизованного сигнала будет выделяться средняя часть, которая с точностью до постоянного мно-

жителя совпадает со спектром исходного непрерывного сигнала (рис. 3.6, г).

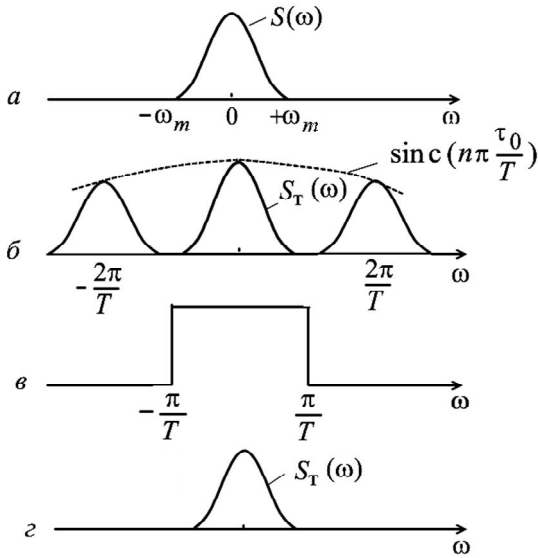


Рис. 3.6

Однако если исходное непрерывное колебание таково, что его спектр с ростом частоты не обращается строго в нуль, то при любом выборе интервала дискретизации соседние составляющие спектра дискретизованного колебания будут частично перекрываться (рис. 3.7, а).

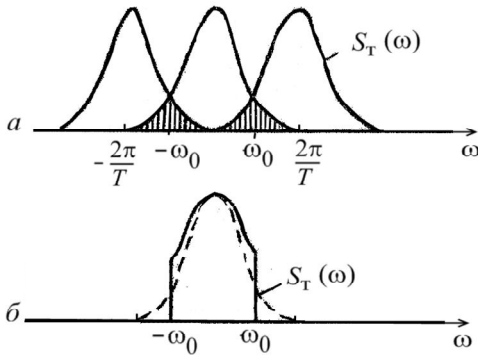


Рис. 3.7

Для предотвращения этого явления можно предварительно пропустить непрерывный сигнал через низкочастотный фильтр с полосой пропускания $(-\pi/T, +\pi/T)$. При этом будут потеряны частоты, лежащие вне этого диапазона, но эффект наложения частот будет исключен.

Фильтры нижних частот могут быть построены как на пассивных, так и на активных элементах.

Пассивные фильтры. Рассмотрим схему, изображенную на рис. 3.8, а. При подаче на вход напряжения

$$u = u_0 \exp j(\omega t + \varphi)$$

мгновенное значение комплексного тока в цепи

$$i = C(du/dt) = j\omega C u_0 \exp j(\omega t + \varphi) = -u_0 / jX_c,$$

где $X_c = 1/\omega C$ – емкостное сопротивление.

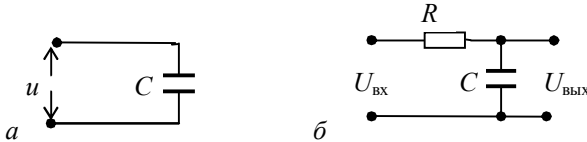


Рис. 3.8

Используем этот результат при анализе схемы на рис. 3.8,б. В ней реактивное сопротивление $Z = R - j/\omega C$, поэтому комплексный ток можно представить в виде:

$$\dot{I} = \frac{\dot{U}_{\text{вх}}}{R - j/\omega C} = \frac{\dot{U}_{\text{вх}}(R + j/\omega C)}{R^2 + 1/\omega^2 C^2}.$$

Тогда комплексное выходное напряжение

$$\dot{U}_{\text{вых}} = \dot{I} X_c = \frac{\dot{U}_{\text{вх}}(R + j/\omega C)}{R^2 + 1/\omega^2 C^2} (-j/\omega C).$$

Модуль выходного напряжения

$$U_{\text{вых}} = (\dot{U}_{\text{вых}} \dot{U}_{\text{вых}}^*)^{1/2} = \dot{U}_{\text{вх}} (1 + R^2 \omega^2 C^2)^{-1/2}.$$

Зависимость модуля коэффициента передачи этой цепи от частоты показана на рис. 3.9. Ее вид свидетельствует о том, что такая цепь пропускает низкие частоты и подавляет высокие, поэтому она является *фильтром низких частот*.



Рис. 3.9

Если поменять местами конденсатор и резистор в схеме фильтра низких частот, то получится простейший пассивный *фильтр высоких частот*. Анализ его характеристик можно выполнить аналогичным образом.

Активные фильтры. Часто возникает необходимость в фильтрах с более плоским участком характеристики в полосе пропускания и с более крутыми склонами. Простое каскадное соединение не даст нужного результата, так как входное сопротивление каждого звена будет служить нагрузкой для предыдущего звена. Применение преобразователей сопротивления увеличивает наклон, но «излом» амплитудно-частотной характеристики остается плавным. Требуемые характеристики можно получить, используя активные фильтры – ОУ с частотно-зависимой обратной связью.

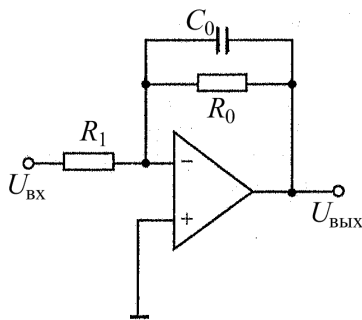


Рис. 3.10

Рассмотрим несколько измененную схему инвертирующего усилителя с реактивным сопротивлением Z во входной цепи и импедансом Z_0 в цепи обратной связи (рис. 3.10). Комплексный коэффициент усиления такой схемы $K = U_{\text{ВЫХ}} / U_{\text{ВХ}} = -Z_0 / Z$. Когда импеданс обратной связи образован параллельно включенными резистором R_0 и конденсатором C_0 , величина $Z_0 = R_0 / (1 + j\omega R_0 C_0)$.

Заменяя $R_0 C_0 = 1/\omega_0 = 1/2\pi f_0$ и полагая, что импеданс $Z = R$, получаем, что комплексный коэффициент усиления

$$K = -(R_0/R)[1 + j(\omega/\omega_0)]^{-1} = |K| \exp(j\varphi), \quad (3.1)$$

где

$$|K| = (R_0/R)[1 + (\omega/\omega_0)^2]^{-1/2}, \quad (3.2)$$

$$\varphi = \pi - \operatorname{arctg}(\omega/\omega_0) = \pi + \Delta\varphi.$$

В логарифмическом масштабе нормированное соотношение (3.2) запишется в виде¹

$$|K| \text{ (дБ)} = -20 \lg [1 + (\omega/\omega_0)^2]^{-1/2}.$$

Эта зависимость – логарифмическая амплитудно-частотная характеристика – изображена на рис. 3.11 линией 1:

$$|K| \text{ (дБ)} \approx 0 \text{ при } f/f_0 \leq 1,$$

$$|K| \text{ (дБ)} \approx -20 \lg(f/f_0) \text{ при } f/f_0 \geq 1.$$

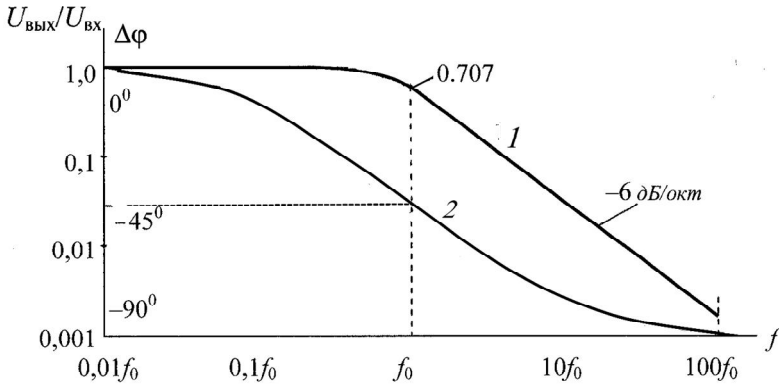


Рис. 3.11

Излом этой характеристики соответствует $f = f_0$, а наклон крутой части составляет $20 \lg 2 = 6$ дБ/окт, или $20 \lg 10 = 20$ дБ/дек. Заметим, что точное значение нормированной характеристики в точке $f = f_0$ равно $1/\sqrt{2}$ (т.е. -3 дБ).

Кривая 2 на рис. 3.11 представляет частотную зависимость фазового сдвига: $\Delta\varphi = 0$ при $f \leq 0,1f_0$; $\Delta\varphi = -45^\circ [1 + \lg(f/f_0)]$ при $0,1f_0 \leq f \leq 10f_0$; $\Delta\varphi = -90^\circ$ при $f \geq 10f_0$. В точке излома $f = f_0$ отклонение $\Delta\varphi = -45^\circ$, абсолютная величина наклона составляет $45^\circ \lg 2 = 13,55$ град/окт или $45^\circ \lg 10 = 45$ град/дек.

¹ При построении частотных характеристик значение числа N в децибелах (дБ) равно $20 \lg 10N$, октава и декада соответствуют двукратному и десятикратному изменению частоты.

Соотношение (3.2) определяет амплитудно-частотную характеристику фильтра нижних частот, а именно – фильтра Баттерворта. Этот тип фильтра имеет максимально плоскую характеристику в полосе пропускания. Другие распространенные типы активных фильтров нижних частот – фильтры Чебышева и Бесселя.

Заметим, что при замене $-(R_0/R) = a_0$, $1 = b_0$, $1/R_0C_0 = b_1$, $p = j\omega$ комплексный коэффициент усиления (3.1) записывается в виде:

$$K = a_0/(b_0 + pb_1),$$

что представляет собой частный случай более общего соотношения

$$K_u = \frac{a_0 + a_1p + \dots + a_m p^m}{b_0 + b_1p + \dots + b_n p^n},$$

описывающего характеристику рекурсивного фильтра.

Основное преимущество активных фильтров – очень низкий выходной импеданс, что позволяет каскадировать подобные схемы и добиваться требуемых частотных характеристик.

3.4. Мультиплексоры аналоговых сигналов

Мультиплексор аналоговых сигналов (МАС) – это схема, обеспечивающая временное разделение каналов на входе обрабатывающего устройства. Она связывает определенный, выбранный в данный момент времени, источник сигналов с устройством его последующего преобразования. Применение МАС позволяет использовать общие преобразователи для всех входных сигналов, сокращая тем самым затраты на измерительную подсистему.

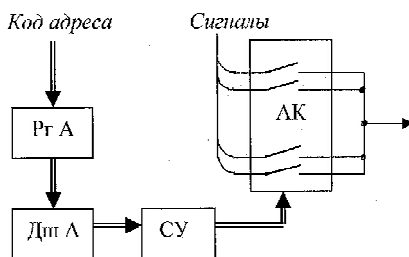


Рис. 3.12

Современные МАС состоят из следующих основных элементов (рис. 3.12): аналоговых ключей АК, адресного регистра РгА, дешифратора адреса ДшА, схемы управления аналоговыми ключами СУ. Принцип действия МАС прост: код адреса выбранного канала, поступивший в

РГА, преобразуется ДША в позиционный код; СУ формирует соответствующие сигналы, управляющие АК, которые подключают к выходу МАС заданный входной канал и тем самым осуществляют коммутацию источников сигналов.

Код адреса аналогового канала может вводиться в РГА в параллельной или в последовательной форме. Вторым вариантом реализован в интегральной схеме (ИС) МАС типа К1104КН1; его недостаток – низкое быстродействие.

По способу адресации каналов МАС можно разделить на МАС с адресным, циклическим и комбинированным опросом. При *адресном* опросе адрес каждого канала определяется потребителем информации. При *циклическом* опросе управляющее устройство запускает коммутатор каналов, работающий в автоматическом режиме с возможностью прерывания цикла. При *комбинированном* опросе управляющее устройство задает зону адресов, в пределах которой осуществляется циклический опрос. Для этого используются два регистра: один – для хранения длины зоны опроса, другой – для циклического режима.

Выпускаемые в настоящее время мультиплексоры имеют ограниченное число аналоговых ключей. Для увеличения числа коммутируемых каналов нужно использовать многоступенчатый принцип коммутации (обычно двухступенчатый). Все ключи первой ступени адресуются одним адресным регистром, вторая ступень коммутируется специальным стробирующим (управляющим) импульсом.

Функциональная полнота схемы МАС определяется числом внешних элементов, необходимых для организации его работы. В качестве примера функционально полной схемы можно привести ИС К1104КН1.

Основными факторами, определяющими динамические свойства мультиплексора, являются время включения, время установления и время выключения, причем наилучший вариант – когда время включения превышает время выключения: это гарантирует отключение ранее выбранного канала, прежде чем произойдет включение следующего. Время установления выходного сигнала мульт-

типлектора зависит не только от его параметров, но и от параметров источников входных сигналов.

Наиболее распространенные ИС МАС имеют следующие характеристики: ключи выполняются по КМОП-технологии, число коммутируемых каналов – от 8 до 16, диапазон коммутируемых напряжений ± 15 В, диапазон коммутируемых токов 5...50 мА, сопротивление открытого ключа от 50 до 500 Ом, время включения от 0,3 до 2,5 мкс.

Кроме твердотельных элементов, в технике коммутации аналоговых сигналов могут применяться электромагнитные коммутаторы с магнитоуправляемым контактом (герконы). Их основные достоинства – высокая надежность и устойчивость к воздействию перенапряжений (в частности, из-за воздействия помех, которые в промышленных установках могут достигать 100 В в диапазоне частот до 2000 Гц); отсутствие гальванической связи между управляющими и коммутируемыми цепями; низкое сопротивление замкнутых контактов и очень высокое – разомкнутых; стабильная работа в широком диапазоне температур. Лучшие образцы герконовых ключей требуют для переключения мощность, меньшую 100 мВт, при скорости переключения порядка 1 мс и полосе частот до 1 кГц.

Современные аналоговые ключи и мультиплексоры характеризуются хорошим согласованием сопротивления каналов, возможностью работать с одним источником питания, а также сохранением работоспособности при существенном превышении сигнала над уровнем источника питания. Применение таких ключей и мультиплексоров обеспечивает защиту других микросхем от перенапряжений и такой широко встречающейся причины выхода из строя аналоговых микросхем как наличие сигнала на входе при отсутствии питания.

3.5. Схемы выборки и хранения

Схемы выборки и хранения аналоговых сигналов служат для фиксации мгновенных значений напряжения в заданные моменты времени. В измерительных трактах эти схемы необходимы для

согласования динамических характеристик непрерывных входных сигналов с динамическими характеристиками аналого-цифровых преобразователей. Благодаря им удается снизить динамическую погрешность, выполнить точную временную привязку значения кода к фактическому моменту получения отсчета и, соответственно, упростить дальнейшую обработку сигнала. Схемы выборки и хранения нужны также для многоканальных систем, где они обеспечивают хранение отсчета по одному каналу, в то время как мультиплексор переключается на другой канал.

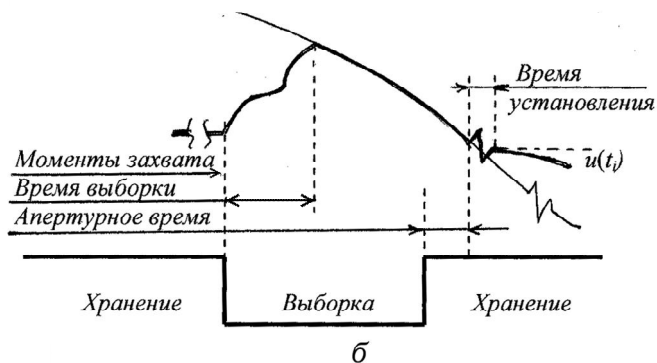
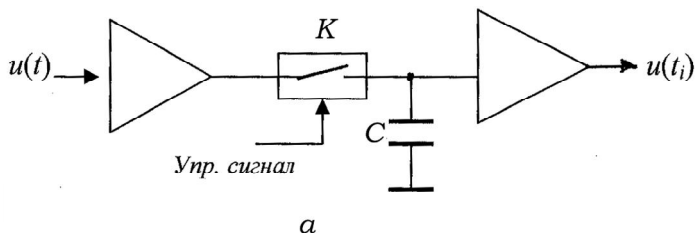


Рис. 3.13

На рис. 3.13 показаны упрощенная принципиальная схема (а) и форма сигналов для типичной схемы выборки и хранения (б). Аналоговый ключ K замыкается при поступлении команды ВЫБОРКА. Конденсатор C начинает заряжаться (или разряжаться) до уровня входного сигнала. После определенной задержки, называемой *временем выборки*, напряжение на конденсаторе достигает границ узкой области вблизи уровня входного сигнала, ширина которой

определяется точностью работы схемы, и остается внутри этой области. Команда ХРАНЕНИЕ вызывает размыкание ключа K , но фактически ключ размыкается после короткого временного интервала, называемого *апертурным временем задержки*. Типичное значение этого времени – несколько десятков наносекунд. После размыкания ключа должно пройти некоторое «*время установления*», прежде чем выходной сигнал стабилизируется, т.е. придет к установившемуся значению и будет готов для преобразования. Однако и после стабилизации уровень этого сигнала не остается постоянным, а медленно спадает со временем из-за утечки заряда. Кроме того, наличие паразитной емкости может привести к появлению на выходе схемы выбросов напряжения, представляющих собой ослабленный отклик схемы на любые большие изменения входного сигнала.

Характеристики схемы выборки и хранения определяется запоминающим конденсатором и режимом его работы. При выборе конденсатора с малой емкостью уменьшается время заряда, но увеличивается спад напряжения в режиме хранения. С другой стороны, конденсатор с большой емкостью более точно хранит напряжение, но при этом долго заряжается. Обычно его номинал выбирается так, чтобы получить минимальное время выборки при условии, что спад напряжения за цикл преобразования не должен превышать единицы младшего разряда аналого-цифрового преобразователя. Типичное значение времени выборки – 5 мкс при точности 0,1 % или от 10 до 20 мкс при точности 0,01 %; скорость спада – от 0,1 до 5 В/с.

Диапазон параметров схем выборки–хранения очень широк. Так, микросхема SHC298 (фирма BURR–BROWN) имеет время установления 19 мкс, апертурную задержку порядка 200 нс, скорость разряда 0,2 мкВ/мкс и скорость нарастания 10 В/мкс. Это самая дешевая схема.

Сверхбыстродействующая схема SHC605 той же фирмы имеет следующие параметры: время установления 60 нс, апертурную задержку 1,7 нс, скорость разряда 8000 мкВ/мкс и скорость нарастания 140 В/мкс.

Реальные схемы выборки и хранения состоят из интегратора с высокоомной нагрузкой и малыми токами утечки и ключевых схем и могут быть построены на операционных усилителях. Так, например, микросхема КР1100СК2 содержит два операционных усилителя с входным сопротивлением более 10 МОм и схему управления, обеспечивающую токовое управление ключами. Для функционального завершения к микросхеме КР1100СК2 необходимо подключить высококачественный (с диэлектриком из полипропилена, полистирола или тефлона) конденсатор C , емкость которого выбирается исходя из приведенных выше соображений и обычно лежит в пределах от 20 до 1000 пФ.

Часто схемы выборки-хранения конструктивно объединяют с преобразователем «напряжение-код».

3.6. Детекторы

Существуют датчики, выходной сигнал которых представляет собой модулированное колебание определенной частоты. Извлечение информации из такого сигнала основано на детектировании – процессе выделения модулирующего сигнала на фоне «несущей».

3.6.1. Амплитудная модуляция и синхронное детектирование

Амплитудная модуляция используется в различных типах параметрических датчиков. В самом общем случае частотный сигнал, несущий в себе информацию, можно представить в виде

$$a(t) = A(t)\cos [\omega_0 t + \theta(t)] = A(t)\cos\psi(t),$$

в котором амплитуда A или фаза θ изменяются по закону передаваемого сообщения. Если изменяется амплитуда – имеет место амплитудная модуляция, если фаза – угловая (частотная или фазовая). В амплитудно-модулированном сигнале фаза $\theta(t) = \theta_0$ и, следовательно, $a(t) = A(t)\cos (\omega_0 t + \theta_0)$.

Пусть модулирующая функция является гармоническим колебанием:

$$s(t) = S_0 \cos(\Omega t + \gamma).$$

Тогда огибающую модулированного колебания можно представить в виде

$$A(t) = A_0 + k_{\text{ам}} s(t) = A_0 + \Delta A_m \cos(\Omega t + \gamma),$$

где Ω – частота модуляции; γ – начальная фаза огибающей; $k_{\text{ам}}$ – коэффициент пропорциональности; $\Delta A_m = k_{\text{ам}} S_0$ – амплитуда изменения огибающей относительно среднего уровня A_0 .

Отношение $M = \Delta A_m / A_0$ называется *коэффициентом модуляции*. Таким образом, мгновенное значение модулированного колебания

$$a(t) = A_0 [1 + M \cos(\Omega t + \gamma)] \cos(\omega_0 t + \theta_0). \quad (3.3)$$

Перепишем выражение (3.3) в форме

$$a(t) = A_0 [\cos(\omega_0 t + \theta_0) + M \cos(\Omega t + \gamma) \cos(\omega_0 t + \theta_0)].$$

Второе слагаемое

$$\begin{aligned} & M [\cos(\Omega t + \gamma) \cos(\omega_0 t + \theta_0)] = \\ & = \frac{1}{2} \{M [\cos((\omega_0 + \Omega)t + (\theta_0 + \gamma))] + M [\cos((\omega_0 - \Omega)t + (\theta_0 - \gamma))]\}, \end{aligned}$$

после чего развернутое выражение колебания принимает вид

$$\begin{aligned} a(t) = A_0 \cos(\omega_0 t + \theta_0) + (A_0 M / 2) \{ & [\cos((\omega_0 + \Omega)t + (\theta_0 + \gamma))] + \\ & + [\cos((\omega_0 - \Omega)t + (\theta_0 - \gamma))] \}. \end{aligned}$$

Первое слагаемое в правой части представляет собой исходное немодулированное колебание с частотой ω_0 . Второе и третье слагаемые соответствуют новым гармоническим колебаниям, которые являются следствием модуляции амплитуды. Частоты этих колебаний $\omega_0 + \Omega$ и $\omega_0 - \Omega$ называются соответственно *верхней* и *нижней боковыми частотами модуляции*.

Детектирование заключается в выделении сигнала, который в неявной форме содержится в модулированном колебании. Детектирование – это процесс, обратный модуляции, поэтому для обозначения этого процесса часто используют термин «демодуляция». Соответственно основным видам модуляции различают амплитудное, частотное и фазовое детектирование.

Амплитудное детектирование может быть выполнено с помощью однополупериодного выпрямителя (последовательное соеди-

нение диода с параллельно соединенными активным сопротивлением R и емкостью C). Ток через диод возможен только в течение отрезков периода, когда положительная полуволна входного напряжения превышает выходное напряжение. В промежутках между импульсами тока через диод происходит разряд конденсатора через резистор и напряжение на выходе убывает; во время импульса конденсатор подзаряжается и выходное напряжение растет. При достаточно большой (по сравнению с периодом высокой частоты $T = 2\pi/\omega_0$) постоянной времени RC зубцы практически отсутствуют, напряжение на выходе воспроизводит огибающую амплитуд входного напряжения. Еще одно условие, необходимое для нормальной работы детектора: постоянная времени цепи нагрузки должна быть мала по сравнению с периодом модуляции [3]. Детектор, удовлетворяющий этим условиям, называют *линейным детектором*.

Если на вход детектора вместе с модулированным сигналом подать еще один сигнал – гармонический вспомогательный сигнал с частотой, равной частоте модуляции, то выходное напряжение будет совпадать по форме с законом модуляции, т.е. с изменением информативного параметра. Такой принцип называется *синхронным детектированием*. Оно обладает рядом преимуществ по отношению к линейному детектированию (большая точность и скорость), однако реализация этого принципа связана со значительными трудностями, поскольку нужно обеспечить синхронность двух сигналов [3].

3.6.2. Фазовый детектор

Фазовый детектор – устройство, которое осуществляет сравнение двух входных частот и формирует выходной сигнал, пропорциональный их фазовой разности. Самый простой фазовый детектор – схема ИСКЛЮЧАЮЩЕЕ ИЛИ с RC -цепочкой на выходе. Его выходное напряжение линейно зависит от фазовой разности опорного и исследуемого сигналов (импульсных), подаваемых на входы схемы.

Фазовый детектор второго типа, чувствительный к взаимному расположению фронтов опорного и исследуемого сигналов, можно построить на основе операционного усилителя, используя его в качестве инвертора (рис. 3.14).

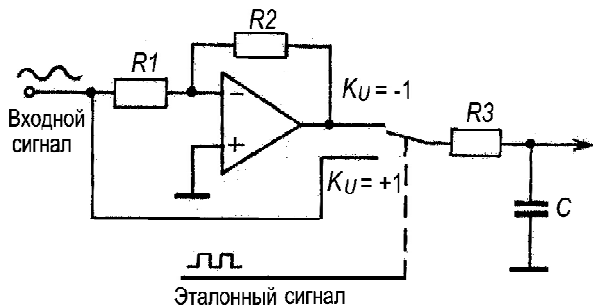


Рис. 3.14

Для того чтобы проанализировать работу такого фазового детектора, допустим, что на его вход подается сигнал $A\cos(\omega_0 t + \theta)$ и опорный сигнал представляет собой прямоугольное колебание, причем изменение полярности опорного сигнала происходит в моменты перехода через нуль функции $\sin\omega_0 t$. Предположим далее, что выходной сигнал усредняется с помощью фильтра низких частот, постоянная времени которого превышает величину одного периода $T = 2\pi/\omega_0$. Тогда можно показать, что в этом случае усредненный выходной сигнал будет пропорционален амплитуде A и синусу относительного сдвига фаз.

3.7. Схемы гальванической развязки

В некоторых случаях при соединении различных элементов измерительного тракта, находящихся на значительном расстоянии, уровни заземления могут различаться на много киловольт. В этих случаях непосредственное соединение будет приводить к потере работоспособности. Для соединения в таких условиях требуются схемы гальванической развязки. Для этого используются три основных метода.

Трансформаторная развязка. Для этого несущий сигнал высокой частоты подвергается модуляции узкополосным сигналом, который требуется изолировать. Усилители с такой развязкой обеспечивают изоляцию до 3,5 кВ и имеют типичную полосу пропускания порядка 2 кГц.

Развязка за счет емкостной связи по высокочастотной несущей, модулированной по частоте сигналом, который требуется изолировать; обеспечивается изоляция до 3,5 кВ и полоса до 70 кГц.

Оптоэлектронная передача сигнала. Связь идет при помощи *оптрона* – излучателя на светодиоде, расположенного в непосредственной близости от фотодетектора. Преобразование входного электрического сигнала происходит при изменении входного тока, проходящего через светоизлучающий диод, которое вызывает его свечение. При этом изменяется освещенность приемника излучения, его сопротивление и, соответственно, ток на выходе оптопары.

Основные достоинства оптопар: отсутствие электрической связи между входом и выходом и обратной связи между фотоприемником и излучателем; широкая полоса частот пропускаемых колебаний; высокая помехозащищенность оптического канала.

Контрольные вопросы и упражнения

1. Почему в схеме неинвертирующего усилителя $U_{\Lambda} = U_{\text{вх}}$?
2. Докажите, что идеальный дифференциальный усилитель полностью устраняет аддитивную синфазную помеху.
3. Для чего служат унифицирующие измерительные преобразователи?
4. Определите экономические условия, при которых целесообразно применять групповые УИПы.
5. Найдите коэффициент преобразования УИП для линеаризации функции $y = ax + bx^2$.
6. Каковы функции аналогового фильтра низких частот?
7. В чем заключается эффект наложения спектров? Что является его причиной?

8. Найдите аналитические соотношения для расчета амплитудно-частотной характеристики пассивного фильтра высоких частот.
9. От каких параметров схемы активного фильтра нижних частот зависит точка излома его амплитудно-частотной характеристики?
10. Перечислите способы адресации в мультиплексорах аналоговых сигналов.
11. Перечислите все погрешности, которые вносит схема выборки-хранения.
12. Что такое «апертурное время» схемы выборки-хранения?
13. Объясните эффект от использования схем выборки-хранения в многоканальных системах.
14. Нарисуйте схему однополупериодного выпрямителя, вид импульсов тока через диод и напряжение на выходе.
15. Используя графики вольт-амперной характеристики диода и входного амплитудно-модулированного сигнала, объясните принцип амплитудного детектирования.
16. Получите соотношения, описывающие выходное напряжение фазового детектора в случае, когда входной сигнал имеет ту же частоту, что и опорный.
17. Получите соотношения, описывающие выходное напряжение фазового детектора в случае, когда частота входного сигнала близка (но не равна) частоте опорного сигнала (указание: это означает, что $\theta = t\Delta\omega$).
18. Для чего нужны схемы гальванической развязки?

4. АНАЛОГО-ЦИФРОВЫЕ ПРЕОБРАЗОВАТЕЛИ

Если измерительную информацию предполагается обрабатывать в системе с помощью цифровой вычислительной машины или с помощью цифрового процессора, то необходимо преобразовать исходный непрерывный (аналоговый) сигнал в дискретный и квантованный (цифровой) сигнал. Для этих целей используются аналого-цифровые преобразователи (АЦП). Выходной сигнал обычно представляется в форме двоичного позиционного кода. Носителями информации во входном сигнале могут быть напряжение, ток, временной интервал, фаза и частота переменного тока. Эти же величины могут служить промежуточными величинами, если по каким-либо причинам прямое преобразование входной величины в код нецелесообразно.

Аналого-цифровые преобразователи классифицируют по виду информативного параметра на его входе и выходе.

4.1. Преобразователи напряжения в частоту

Эти преобразователи относятся к аналого-цифровым преобразователям, поскольку в них изменяющийся во времени аналоговый сигнал – напряжение преобразуется в последовательность импульсов, частота которых прямо пропорциональна текущему его значению. Число этих импульсов за фиксированный промежуток времени представляет собой число-импульсный код, который при необходимости может быть преобразован в двоичный или десятичный. В ряде случаев такие преобразователи обладают определенными преимуществами перед АЦП других типов.

Преобразователи напряжения в частоту могут быть прямого и импульсно-компенсационного типа.

Преобразователи прямого преобразования. Преобразователя такого типа работают по следующему принципу. Входное напряжение $u_{вх}(t)$ поступает на вход интегратора. Его выходное напряжение с помощью схемы сравнения сравнивается с некоторым за-

данным напряжением u_0 . Сигнал равенства этих напряжений возвращает интегратор в исходное состояние, что обеспечивает развертку процесса во времени.

Общее условие, из которого можно найти выходную частоту, имеет следующий вид:

$$\int_0^T u_{\text{ВХ}}(t) dt = u_0, \quad (4.1)$$

где T – период следования выходных импульсов. Поскольку входное напряжение поступает с выхода схемы выборки–хранения, оно может считаться постоянным в течение времени преобразования.

Временные диаграммы выходного напряжения интегратора для разных постоянных входных напряжений $u'_{\text{ВХ}}$ и $u''_{\text{ВХ}}$ показаны на рис. 4.1. Выходная частота, т.е. частота следования сигналов совпадения на выходе схемы сравнения, пропорциональна среднему значению входного сигнала.

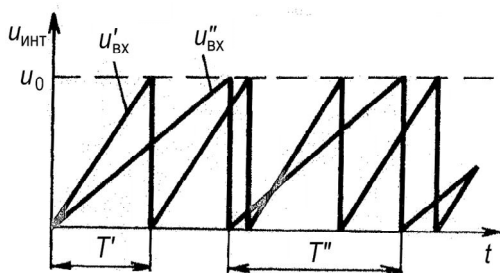


Рис. 4.1

На практике мгновенный возврат интегратора в исходное состояние неосуществим. Поэтому применяют изменение направления интегрирования, формируя обратный ход «пилы» за определенное время. Если параметры интегрирования в прямом и обратном направлении одинаковы, то выходная частота будет в два раза меньше по сравнению с идеальным случаем мгновенного сброса.

Изменение направления интегрирования производится одним из следующих способов [4]:

- 1) подачей параллельно входному напряжению интегратора выходного сигнала инвертирующего масштабного усилителя;
- 2) коммутацией полярности входного напряжения интегратора;

3) изменением знака разности напряжений между инвертирующим и неинвертирующим входами ОУ интегратора.

При подаче на вход переменного входного сигнала пилообразные выходные напряжения интегратора становятся нелинейными и текущая частота изменяется от периода к периоду. Рассмотрим для примера прямое преобразование линейно нарастающего сигнала

$$u_{\text{вх}}(t) = u_{\text{вх}} + (du_{\text{вх}}/dt)t.$$

Интеграл напряжения представляет собой сумму двух функций – линейной и квадратичной. Временные диаграммы показаны на рис. 4.2. При анализе следует иметь в виду, что на основании общего условия (4.1) площади трапеций равны друг другу, т.е. $s_1 = s_2 = s_3 = \dots$, что позволяет определять текущие частоты последующих колебаний по частоте предыдущих.

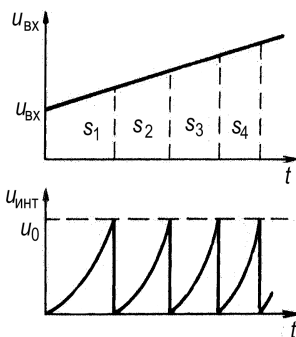


Рис. 4.2

Импульсно-компенсационные преобразователи. Если входное напряжение изменяется во времени, а компенсирующее напряжение $u_{\text{к}}(t)$ имеет форму прямоугольных импульсов длительности $t_{\text{к}}$, то компенсация за период следования импульсов T осуществляется средним значением импульсного напряжения:

$$\frac{1}{T} \int_0^T u_{\text{вх}}(t) dt + \frac{1}{T} \int_0^{t_{\text{к}}} u_{\text{к}}(t) dt = 0. \quad (4.2)$$

Очевидно, что амплитуда компенсирующих импульсов должна быть больше среднего значения входного напряжения, а их полярность – противоположной.

Преобразователь (рис. 4.3, а) включает в себя суммирующий интегратор, нуль-индикатор НИ, сигнализирующий об окончании компенсации, оговоренной условием (4.2), и генератор стабильных одиночных импульсов ГСОИ, запускаемый выходным сигналом НИ. Временная диаграмма работы преобразователя показана на рис. 4.3, б. В результате возрастания выходного напряжения интегратора под действием входного сигнала срабатывает НИ и запус-

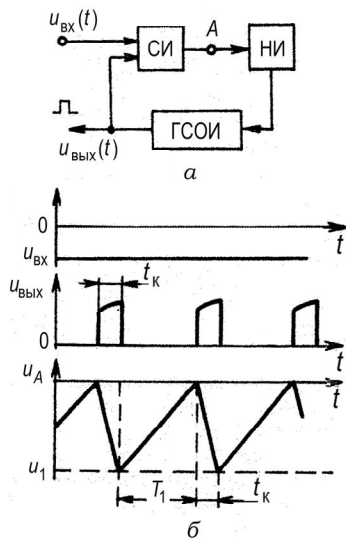


Рис. 4.3

основных функциональных блоков преобразователей напряжения в частоту.

Суммирующий интегратор. Интегратор такого типа интегрирует алгебраическую сумму нескольких подаваемых на разные входы напряжений. Схема суммирующего интегратора для двух напряжений показана на рис. 4.4. Работа этой схемы описывается дифференциальным уравнением

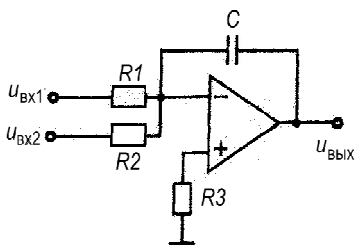


Рис. 4.4

В результате решения этого уравнения относительно $u_{\text{вых}}$ получим:

$$u_{\text{вых}} = -\frac{1}{C} \left(\frac{1}{R_1} \int_0^t u_{\text{вх1}}(t) dt + \frac{1}{R_2} \int_0^t u_{\text{вх2}}(t) dt \right).$$

Возврат интеграторов в исходное положение может производиться быстрым разрядом конденсатора или подключением через

ГСОИ, который генерирует импульс длительности t_k . В течение этого импульса напряжение снижается до значения u_1 . После окончания компенсирующего импульса напряжение на выходе интегратора опять возрастает. Когда оно достигнет нулевого значения, опять сработают НИ и ГСОИ, и начнется новый цикл работы преобразователя.

Максимальная рабочая частота преобразователя «напряжение–частота» достигает 1 МГц, нелинейность не превышает 0,01 %.

Рассмотрим подробнее принципы построения основных функциональных

резистор к одному из входов ОУ такого напряжения, при котором изменяется направление интегрирования.

Источники опорного напряжения. Простейшая схема такого источника может быть выполнена на прецизионном кремниевом стабилитроне. Однако на стабильность напряжения в такой схеме влияют изменения сопротивления нагрузки и тока через стабилитрон. Для улучшения характеристик стабилитрон подключают к неинвертирующему входу ОУ и питают его от стабилизатора тока.

Коммутаторы полярности напряжения. Для коммутации полярности напряжения применяются схемы с транзисторным ключом в цепи его неинвертирующего входа или интегральные ключи на КМОП-структурах. На рис.4.5 показана схема переключения полярности. Она содержит двухканальный интегральный ключ, инвертирующий усилитель на ОУ с единичным усилением и инвертор.

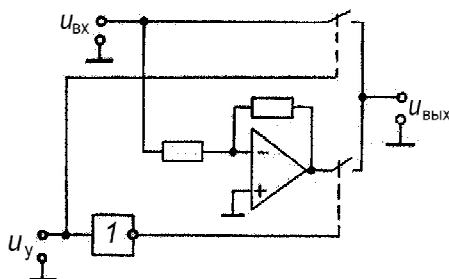


Рис. 4.5

Когда управляющее напряжение u_y представляет собой логическую единицу, входной сигнал $u_{вх}$ проходит на выход через замкнутый верхний ключ схемы, при этом нижний ключ разомкнут. Когда на управляющий вход поступает логический нуль, верхний ключ разомкнут, нижний – замкнут и выходной сигнал имеет полярность, противоположную полярности входного сигнала.

4.2. Преобразователи временного интервала, фазы и частоты в код

Эти преобразователи применяются как в качестве самостоятельных схем, так и в составе других типов АЦП, использующих двойное преобразование. В ряде случаев такой вариант построения оказывается предпочтительным.

Преобразователи временного интервала в код. В этих преобразователях используется метод последовательного счета. Времен-

ной интервал заполняется импульсами тактового генератора стабильной частоты, и его величина характеризуется количеством импульсов, которые укладываются в этом интервале. Схема такого преобразователя приведена на рис. 4.6. Началу временного интервала соответствует старт-импульс. Он устанавливает триггер T в 1, благодаря чему импульсы от генератора ГИ с частотой $f = 1/t$ начинают поступать на счетчик через схему И. Концу временного интервала соответствует стоп-импульс, который переводит триггер в 0, после чего поступление импульсов на вход счетчика прекращается. Следовательно, код счетчика представляет собой длину временного интервала $T = (n-1)t$, где n – число импульсов, поступивших на счетчик.

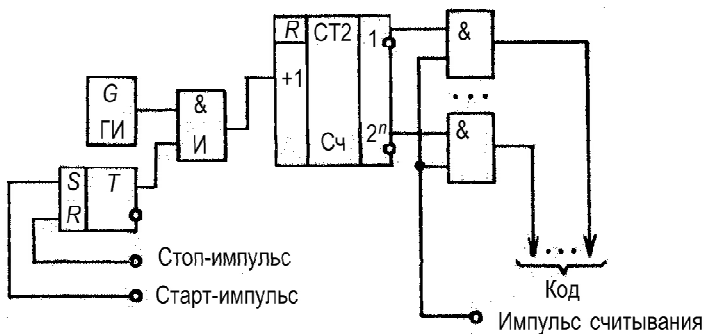


Рис. 4.6

Поскольку старт- и стоп-импульсы несинхронны по отношению к тактовым импульсам, могут возникнуть погрешности: Δt_1 – задержка первого тактового импульса относительно стартового и Δt_2 – задержка стоп-импульса относительно последнего тактового. Поэтому истинное значение временного интервала

$$T^* = (n-1)t + \Delta t_1 + \Delta t_2.$$

Если первую составляющую погрешности можно устранить, применяя специальные схемотехнические приемы, то вторая составляющая принципиально неустранима; ее можно только сократить.

Преобразование фазы в код. Обычно в качестве преобразуемой величины используется разность фаз опорного сигнала $f_0(t)$ и

входного сигнала $f_{\text{вх}}(t)$, информативным параметром которого является фаза. Как правило, имеют дело с синусоидальными сигналами одинаковой частоты.

Наиболее распространенный метод – переход от сдвига фаз к временному интервалу с последующим преобразованием его в код. Для этого фиксируются моменты перехода через нулевой уровень опорного и входного сигналов; соответствующие импульсы служат старт- и стоп-импульсами (см. выше).

Периоду T опорного и входного сигналов (он задан) соответствует сдвиг фаз, равный 2π ; сдвигу фаз φ – длительности временного интервала $T_{\text{и}}$ (он измеряется). Составляя пропорцию, имеем: $T_{\text{и}} = \varphi T / 2\pi$. Переходя далее к кодам и полагая, что длительности T соответствует число импульсов N , а длительности $T_{\text{и}}$ – n , имеем: $\varphi = 2\pi n / N$.

На точности преобразования сказывается наличие высокочастотных помех во входном сигнале. Их влияние уменьшается путем выбора для фиксации сдвига фаз нулевого уровня: здесь крутизна синусоиды максимальна и аддитивная высокочастотная помеха меньше искажает результат сравнения.

Точность можно также повысить, применяя статистическую обработку результатов нескольких измерений (если только изменениями измеряемой фазы за это время можно пренебречь). В простейшем случае это обеспечивается вычислением сдвига фаз во время прохода обоих сигналов через нулевой уровень дважды: в прямом и в обратном направлениях.

Модификация рассмотренного метода: синусоидальные сигналы заменяются прямоугольными импульсами соответствующей полярности. С помощью элемента И легко выделить последовательность разнополярных импульсов, длительность которых $T_{\text{и}}$, пропорциональная измеряемому сдвигу фаз, может быть измерена рассмотренным выше преобразователем.

Количество импульсов ГИ, поступивших с частотой $f = 1/t_{\text{ги}}$ за один период преобразования, равно $T_{\text{и}}/t_{\text{ги}}$. Если выбрать интервал усреднения $T_{\text{ус}} = kT$, а частоту f так, чтобы $T = k_1 t_{\text{ги}}$ (k и k_1 – посто-

янные, выбранные из условия неизменности сдвига фаз или по требуемой точности, если анализу подлежит средняя частота), то общее количество импульсов, поступивших за время T_{yc} ,

$$N = (T_{yc}/T)(T_w/t_{ги}) = kT_w/t_{ги} = k\varphi T/2\pi t_{ги} = \varphi k k_1/2\pi.$$

Видно, что результат измерений инвариантен относительно изменения частоты ГИ и частоты синусоидальных сигналов, т.е. этот метод позволяет исключить два источника погрешностей.

Преобразователи фазы в код часто являются составной частью измерителей перемещений, в которых перемещение сначала преобразуется в фазовый сдвиг, а затем – в код.

Преобразователи частоты в код. Для преобразования частоты в код можно в качестве измеряемой величины использовать: 1) период измеряемой частоты; 2) количество периодов за фиксированный промежуток времени; 3) приращение фазы за фиксированный временной интервал.

Первый случай соответствует задаче преобразования временно-го интервала в код (см. выше), однако следует учитывать, что вычисление частоты по измеренному периоду требует наличия функциональных (гиперболических) преобразователей или процессора.

Реализация второго способа связана с подсчетом периодов (если входной сигнал гармонический, то предварительно его рекомендуется преобразовать в последовательность импульсов) и заданием фиксированного временного интервала. Его длительность должна выбираться с учетом требуемой точности измерения и может изменяться при изменении измеряемой частоты. Для измерения высоких частот используется принцип гетеродинирования.

Третий способ основан на том, что частоту можно определить как отношение приращения фазы к отрезку времени, за который произошло это приращение. Таким образом, проведя два преобразования фазы в код на известном и фиксированном временном интервале, можно определить значение частоты. Обязательным условием, обеспечивающим точность измерения, является строго фиксированная и неизменная форма входного сигнала (предпочтительно – синусоида).

4.3. Преобразователи напряжения в код, построенные по принципу последовательного счета

Способы преобразования напряжения в код. Процесс преобразования напряжения в код можно реализовать разными способами. На практике наибольшее распространение получили три метода:

- *метод последовательного счета*, при котором входная аналоговая величина уравнивается последовательно накапливаемой суммой одинаковых минимальных эталонов, называемых *квантами*. Момент равенства входной величины и суммы эталонов определяется с помощью одного сравнивающего устройства. Результат преобразования характеризуется числом квантов, накопленных при формировании суммы;

- *метод поразрядного кодирования*, при котором входная величина последовательно сравнивается с суммой эталонов, накопленных за предыдущие циклы сравнения. «Вес» каждого эталона равен i квантам, где i последовательно принимает значения $n-1, n-2, \dots, 2, 1, 0$ (n – число разрядов АЦП). Если результат сравнения на i -м шаге положителен (входная величина больше накопленной суммы), то соответствующий эталон прибавляется к сумме, и в i -м разряде выходного регистра устанавливается 1. В противном случае эталон не добавляется, разряд устанавливается в 0 и следует переход к следующему, более «легкому» эталону. Сравнение входной величины и суммы эталонов выполняется последовательно одним сравнивающим устройством, результат преобразования характеризуется совокупностью единиц и нулей в выходном регистре;

- *метод параллельного сравнения*, при котором входная величина одновременно сравнивается с полной совокупностью эталонов, равных $1, 2, 4, 8, \dots, n-1$ квантам. Результат преобразования характеризуется количеством компараторов, отметивших равенство или превышение входной величины по отношению к данному эталону, и представляется последовательностью единиц и нулей. Для получения выходного позиционного кода требуется соответствующее преобразование этого промежуточного результата.

Как видно, каждый метод требует разных аппаратурных затрат (числа эталонов и компараторов) и обеспечивает разное время преобразования. В зависимости от системных требований выбирается тот или иной метод.

Преобразователь последовательного счета (с обратной связью). Для реализации принципа последовательного счета, который был кратко описан выше, требуется один эталон, равный кванту. Он многократно используется для уравнивания входной аналоговой величины. Именно в таком виде этот принцип реализован в преобразователе «напряжение – код» (ПНК), структурная схема которого показана на рис. 4.7.

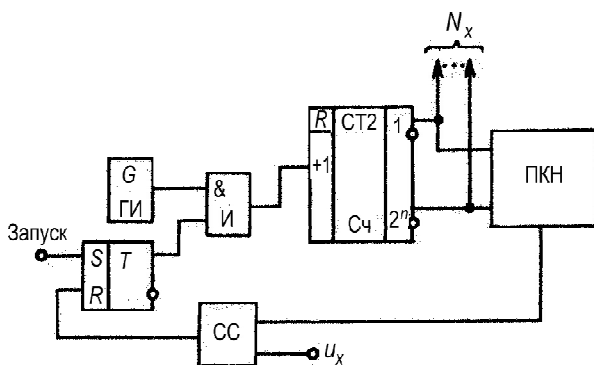


Рис. 4.7

Импульс запуска через элемент И поступает на счетчик. Обратная связь реализована с помощью преобразователя код–напряжение (ПНК): выходы счетчика соединены с входами ПНК и при поступлении на счетчик очередного импульса вместе с увеличением на единицу его содержимого возрастает на одну «ступеньку» напряжение на выходе ПНК. Величина приращения напряжения за один такт равна кванту. Компаратор (схема сравнения СС) закроет элемент И, когда выходное напряжение ПНК достигнет уровня входного сигнала u_x . Состояние счетчика в этот момент является цифровым выходным сигналом ЦАП.

Основной недостаток этого простого способа – зависимость времени преобразования от уровня входного сигнала, причем это

время может быть довольно велико (2^n периодов тактовых импульсов для n -разрядного преобразователя при входном сигнале, близком по уровню к величине полного диапазона преобразования).

В модифицированном варианте АЦП последовательного счета – так называемом «следающем» АЦП – используется реверсивный счетчик. Схема сравнения выдает разные сигналы в зависимости от соотношения между входным сигналом и выходным напряжением ПКН (с точностью до величины кванта). В зависимости от знака рассогласования между $u_{\text{вх}}$ и $u_{\text{ПКН}}$ импульс, поступающий на вход счетчика, либо суммируется с его содержимым, либо вычитается из него.

Для реализации метода последовательного приближения необходим цифро-аналоговый преобразователь типа «код – напряжение». Обычно преобразователи этого типа выполняют по параллельной схеме: преобразование во всех разрядах происходит одновременно с учетом «веса» разряда. Если на вход такого ПКН подается код

$$N = a_n 2^{n-1} + a_{n-1} 2^{n-2} + \dots + a_2 2 + a_1$$

и максимальному значению кода $N_{\text{max}} = 2^{n-1} \approx 2^n$ соответствует выходное напряжение u_{max} , то коду N будет соответствовать напряжение

$$u_{\text{ПКН}} = Nu_{\text{max}}/N_{\text{max}} \approx u_{\text{max}}(a_n 2^{-1} + a_{n-1} 2^{-2} + \dots + a_2 2^{-(n-1)} + a_1 2^{-n}).$$

Видно, что выходной сигнал состоит из n слагаемых, каждое из которых имеет свой коэффициент деления, определяемый номером данного разряда.

Последнее соотношение реализуется двумя способами: 1) матрицей с полным набором весовых сопротивлений и 2) матрицей с двумя весовыми сопротивлениями.

В первом случае весовые сопротивления выбираются равными $R_x 2^{n-i}$. На рис. 4.8, а приведена схема такого ПКН. Входной код поступает на разрядные триггеры Тг1, Тг2, ..., Тг $_n$, управляющие разрядами ПКН через двухпозиционные ключи. Если в i -м разряде кода имеется 1, то ключ K_i подсоединит резистор $R_x 2^{n-i}$ к входу суммирующего ОУ. В противном случае этот вход будет заземлен.

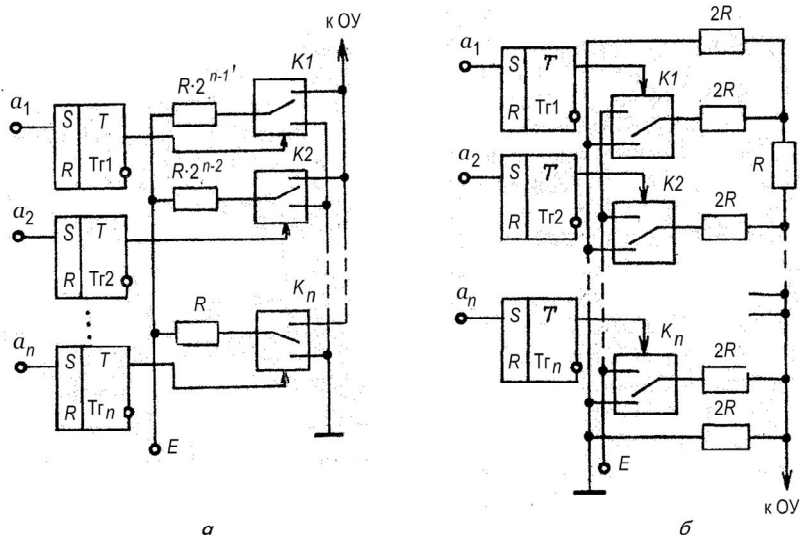


Рис. 4.8

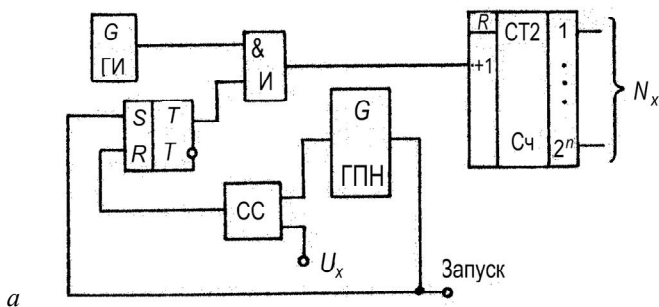
Недосток преобразователя с полным набором весовых сопротивлений – необходимость точного подбора сопротивлений в широком диапазоне номиналов. Особенно трудно сделать это при реализации ПКН посредством интегральной технологии. Этот недостаток отсутствует во втором типе преобразователя – с двумя номиналами весовых сопротивлений (R и $2R$), структурная схема которого показана на рис. 4.8, б.

В этом типе ПКН коэффициент деления реализуется путем соответствующего расположения ветви, поскольку коэффициент передачи напряжения от одной ветви к следующей равен $1/2$. Каждая ветвь составляет для источника питания нагрузку $3R$, а выходное сопротивление ПКН постоянно и равно $2R/3$ независимо от значения кода на входе ПКН. Наличие в i -м разряде единицы создаст на входе ОУ составляющую напряжения, равную по величине $2^{-(n-i)}E/3$.

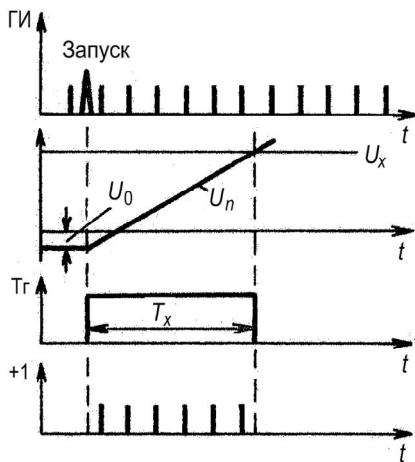
При использовании ПКН необходимо учитывать, что при срабатывании токовых переключателей различия во времени коммутации, сдвиги уровня выходного аналогового сигнала и паразитные

емкости вызывают появление выбросов выходного напряжения. Выброс максимальной амплитуды до 50% значения полной шкалы происходит, когда ключи всех разрядов срабатывают одновременно (например, при переходе от кода 011...1 на код 100...0). Выброс может длиться несколько сотен наносекунд и в быстродействующих схемах нужно предусматривать в составе таких ЦАП специальные сглаживающие устройства.

Преобразователь с промежуточным преобразованием напряжения во временной интервал. Схема преобразователя такого типа изображена на рис. 4.9, а; временная диаграмма его работы показана на рис. 4.9, б.



а



б

Рис. 4.9

Такой ПНК состоит из преобразователя временного интервала в код и двух дополнительных узлов: генератора пилообразного напряжения ГПН и схемы сравнения СС, выполняющих преобразование напряжения во временной интервал.

Импульс запуска по ступает на ГПН, который начинает вырабатывать пилообразное напряжение $U_{\text{п}}$; одновременно импульс запуска с помощью триггера Тг открывает элемент И и импульсы ГИ поступают на счетчик Сч. В момент равенства $U_{\text{п}}$ и U_x срабатывает схема СС, формируя признак конца временного интервала. Соответствующий импульс ставит триггер Тг в нулевое состояние, и импульсы ГИ перестают поступать на счетчик (работа преобразователя временного интервала в код была рассмотрена выше).

Для рассмотренного типа преобразователя справедливо соотношение:

$$U_{\text{п}} = U_{\text{п}0} + at, \quad (4.3)$$

где $U_{\text{п}0}$ – начальный уровень пилообразного напряжения; a – скорость изменения напряжения. Так как при $U_x = U_{\text{п}}$ имеем $t = T_x$, то из (4.3) можно найти значение искомого временного интервала:

$$T_x = (U_x - U_{\text{п}0})/a. \quad (4.4)$$

Если не учитывать погрешности преобразования временного интервала в код, то из (4.4) следует, что выходной код

$$N = T_x/t_{\text{ги}} = (U_x - U_{\text{п}0})/at_{\text{ги}},$$

где $t_{\text{ги}}$ – период следования импульсов ГИ.

Интегрирующие преобразователи. В преобразователе такого типа использовано накопление заряда конденсатором при протекании среднего тока, поступающего на аналоговый вход. Этот заряд сравнивается с зарядом, обеспечиваемым эталонным током, после чего анализируются соответствующие временные соотношения. Простота интегрирующих преобразователей обеспечивает им наилучшее соотношение между высокой точностью и малой стоимостью по сравнению с любыми аналого-цифровыми устройствами.

Наиболее совершенным является метод двойного, или компенсационного интегрирования. Традиционная структура интегрирующих АЦП содержит два основных блока: аналоговый, выпол-

няющий преобразование входного напряжения в длительность временного интервала, и цифровой, обеспечивающий измерение временного интервала и формирования сигналов управления аналоговым блоком. Структурная схема такого преобразователя приведена на рис. 4.10; временные диаграммы – на рис. 4.11.

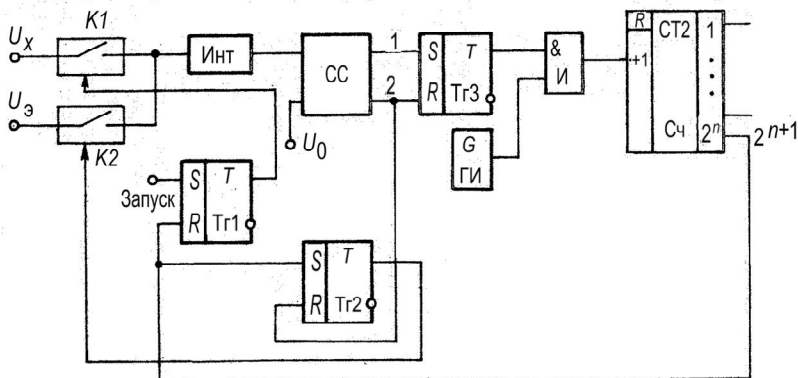


Рис. 4.10

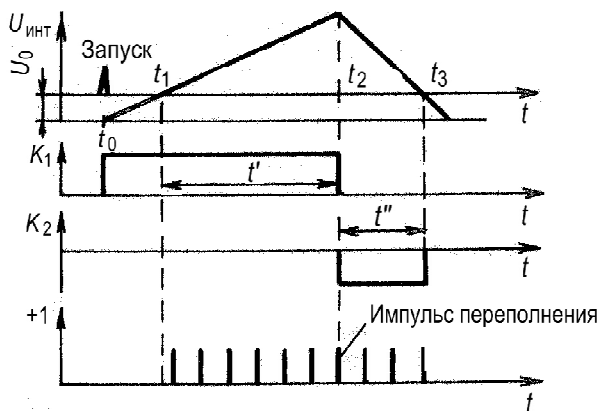


Рис. 4.11

Импульс запуска поступает в момент времени t_0 и, воздействуя на триггер $Tr1$, с его помощью открывает ключ $K1$, после чего преобразуемое напряжение U_x подается на вход интегратора Инт.

Выходное напряжение интегратора изменяется по следующему закону:

$$U_{\text{инт}} = \frac{1}{RC} \int_0^T U_x dt,$$

где RC – эквивалентная постоянная времени интегратора. Выходное напряжение интегратора подается на один из входов схемы сравнения СС, на второй вход которой подается постоянное напряжение U_0 . До начала интегрирования $U_{\text{инт}} < U_0$, а в процессе интегрирования $U_{\text{инт}}$ растет. В момент времени t_1 напряжение $U_{\text{инт}}$ становится равным U_0 и схема сравнения срабатывает. Выходной импульс схемы СС (выход 1) перебрасывает триггер Тг3, в результате чего открывается элемент И и импульсы генератора ГИ начинают поступать на счетчик Сч. Интегрирование напряжения U_x производится в течение фиксированного промежутка времени $\Delta t = t_2 - t_1$. Конец этого временного интервала фиксируется счетчиком, который в момент времени t_2 выдает импульс переполнения, поступающий на триггеры Тг1 и Тг2.

К моменту времени t_2 напряжение на выходе интегратора достигает величины

$$U_2 = U_0 + \frac{1}{RC} \int_{t_1}^{t_2} U_x dt. \quad (4.5)$$

Вследствие появления импульса переполнения счетчика Сч и переброса триггеров Тг1 и Тг2 ключ $K1$ закрывается, а ключ $K2$ открывается; при этом вход интегратора переключается с напряжения U_x на напряжение U_3 , имеющее полярность, обратную полярности напряжения U_x . После момента t_2 импульсы генератора ГИ продолжают поступать на счетчик Сч, а напряжение на выходе интегратора начинает уменьшаться по следующему закону:

$$U_{\text{инт}} = U_2 - \frac{1}{RC} \int_{t_2}^t U_x dt. \quad (4.6)$$

Когда напряжение на выходе интегратора уменьшится до значения U_0 , схема сравнения перейдет в исходное положение и по

выходу 2 перебросит триггеры Тг2 и Тг3 в нулевое состояние. При этом сигнал с выхода триггера Тг2 разомкнет ключ $K2$ и напряжение U_3 отключится от входа интегратора, а сигнал с выхода триггера Тг3 запретит подачу импульсов ГИ на счетчик. В счетчике будет зафиксирован код $N = \Delta T / t_{\text{ти}}$, где ΔT – временной интервал между моментами t_2 и t_3 .

Если (4.6) представить для момента времени $t = t_3$, то получим:

$$U_3 = U_0 = U_2 - \frac{1}{RC} \int_{t_2}^{t_3} U_x dt$$

или, подставив значение U_2 из (4.5),

$$U_3 = U_0 = U_0 + \frac{1}{RC} \int_{t_1}^{t_2} U_x dt - \frac{1}{RC} \int_{t_2}^{t_3} U_3 dt. \quad (4.7)$$

Если считать, что напряжения U_x и U_3 постоянны во время интегрирования, то из (4.7) получим:

$$U_x(t_2 - t_1) = U_3(t_3 - t_2),$$

а так как $t_2 - t_1 = 2^n t_{\text{ти}}$, где n – число разрядов в счетчике, то

$$N = (U_x / U_3) 2^n.$$

Можно отметить следующие положительные свойства этого типа ПНК:

а) получающийся в результате преобразования код прямо пропорционален входному напряжению;

б) медленные вариации частоты генератора и медленные изменения эквивалентной постоянной времени не сказываются на точности преобразования;

в) влияние всех видов помех во входном сигнале уменьшается; если время интегрирования выбрано равным периоду частоты питающей сети, то аддитивная сетевая помеха полностью подавляется, поскольку интеграл от синусоидальной функции в пределах периода равен нулю.

Основной недостаток преобразователей подобного типа состоит в том, что на результатах преобразования сказываются неидеальность интегратора (отклонение закона интегрирования от линейного) и появление дрейфа в его выходном напряжении.

4.4. Преобразователи напряжения в код, построенные по принципу поразрядного кодирования

Алгоритм работы преобразователей этого типа сводится к последовательному выполнению операции деления входного напряжения на эталонное напряжение старшего разряда, затем – полученного остатка на эталонное напряжение следующего разряда и т.д. Для двоичного кода i -й такт преобразования состоит в определении разрядного коэффициента кода a_i , удовлетворяющего соотношению

$$U_{\text{эт}} a_i 2^{-i} = U_{\text{вх}} - \sum_{j=1}^{i-1} U_{\text{эт}} a_j 2^{-j} .$$

Генерируемое дискретное напряжение сходится к уровню входного аналогового сигнала путем ряда последовательных приближений, причем каждый шаг равен половине предыдущего шага (рис. 4.12).

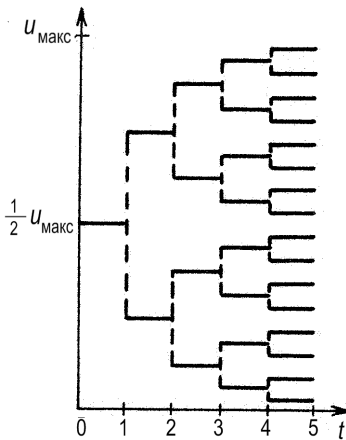


Рис. 4.12

Алгоритм последовательного приближения обычно реализуется следующим схемным способом.

Процессом аналого-цифрового преобразования управляет регистр последовательной аппроксимации РПА (он состоит из сдвигающего регистра Рг и триггеров памяти Тг₁ – Тг_{*n*}), ЦАП, формирующего выходной код, и логических схем управления).

Преобразование начинается с поступлением импульса запуска. Триггер старшего разряда Тг_{*n*} устанавливается в 1, а все остальные – в 0. Одновременно 1 записывается в сдвигающий регистр.

Если $U_{\text{вх}} > U_n = U_{\text{эт}}/2$ (U_n – входное напряжение ЦАП, соответствующее единице старшего разряда), то на выходе компаратора импульс отсутствует и 1 в Тг_{*n*} сохранится. Если $U_{\text{вх}} < U_n$, то выходной импульс компаратора через соответствующую схему И установит Тг_{*n*} в состояние 0. Одновременно произойдет сдвиг в сдвигающем регистре и в следующий триггер памяти запишется 1. Та-

ким же образом выполняются все остальные циклы. Следует заметить, что регистры последовательной аппроксимации выпускаются в виде отдельных микросхем (типа К155ИР17).

Скорость АЦП последовательного приближения определяется временем срабатывания логических схем, внутреннего ЦАП и компаратора и при фиксированной частоте линейно зависит от числа разрядов. Этот способ обеспечивает недорогое при среднем (порядка 1 МГц) быстродействии техническое решение. Он использован в отечественных микросхемах типа К1108ПВ1 и К1108ПВ2 [5].

4.5. Преобразователи напряжения в код, построенные по принципу параллельного сравнения

Преобразователи параллельного сравнения обладают, в принципе, самым высоким быстродействием. Они построены на основе сетки разрешенных уровней напряжения. Эти напряжения формируются с помощью делителей эталонного напряжения E . Каждое из опорных напряжений поступает на первые входы 2^{n-1} схем сравнения, а на все вторые входы этих схем подано входное напряжение U_x .

В качестве примера преобразователя такого типа рассмотрим микросхему К1107ПВ1. Функциональная схема этой БИС объединяет делитель опорного напряжения, 64 стробируемых компаратора напряжения КН (А0 – А63), синхронизирующие буферные каскады, логические схемы шифрации и управления, выходной буферный регистр Рг (рис. 4.13).

Делитель опорного напряжения представляет собой набор низкоомных резисторов с номиналами $R = 1,04...1,24$ Ом. Резисторы образуют шину деления опорного напряжения с двумя самостоятельными выводами 1 и 2. По выводу 1 возможно проведение коррекции напряжения смещения нуля на входе, а по выводу 2 – абсолютной погрешности преобразования в конечной точке шкалы.

Каждый из 64 КН обладает быстродействием порядка 7 нс (типовая задержка) и небольшим коэффициентом усиления. Перепад логических уровней на выходе составляет 200...400 мВ. Управление

ние работой КН, логических элементов, буферного регистра осуществляется от одних и тех же тактовых импульсов ТИ.

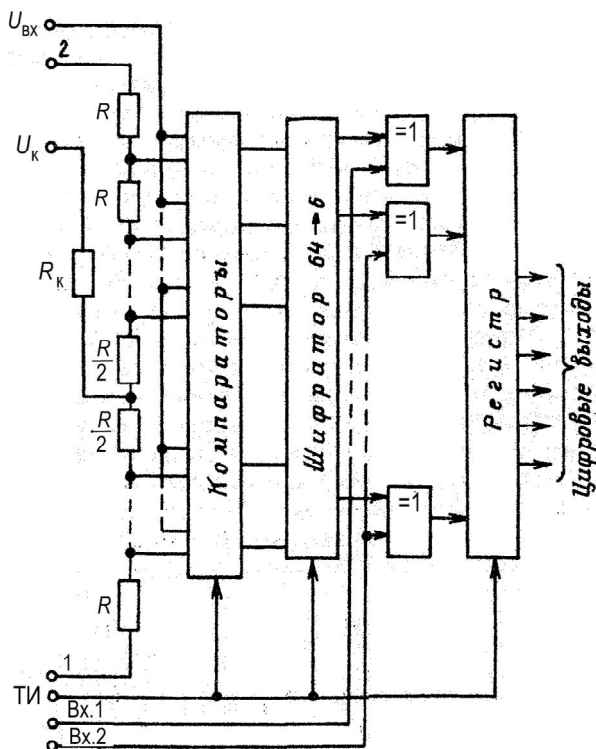


Рис. 4.13

Выборка аналогового сигнала производится через 10...22 нс после появления фронта ТИ на соответствующем входе КН. Шифрация сигналов с выходов 64 КН проводится с момента поступления среза ТИ. Результат шифрации, который является также конечным результатом процесса кодирования, записывается в выходной буферный регистр с появлением фронта следующего ТИ. Задержка, вносимая буферным регистром, не превышает наименьшей длительности периода ТИ, равной 50 нс. Это делает возможным производить выборку аналогового сигнала одновременно с появлением на выходе АЦП результата преобразования. В тот момент, когда на выходе микросхемы появляется результат n -й выборки, на входе

производится $(n+2)$ -я выборка, то есть имеет место конвейеризация обработки сигналов. Частота преобразования МС типа К1107ПВ1 – до 20 МГц.

4.6. Основные направления совершенствования схем и принципов преобразования

Среди современных АЦП все реже можно найти схемы, построенные «традиционным» способом (параллельные, последовательных приближений, интегрирующие). Для обеспечения лучших параметров (быстродействие, разрешающая способность) используются комбинированные структуры.

Практически все АЦП имеют последовательно-параллельную структуру со схемой выборки-хранения на входе и работают по принципу постоянного опроса, когда преобразование сигнала происходит постоянно с максимальной скоростью.

Это обеспечивается, в частности, применением дельта-модуляции. Преобразователь, реализующий этот принцип, состоит из компаратора и квантователя K в основном тракте обработки сигнала и из интегрирующего усилителя в цепи обратной связи (рис. 4.14, а). Алгоритм работы такого преобразователя поясняется рис. 4.14, б.

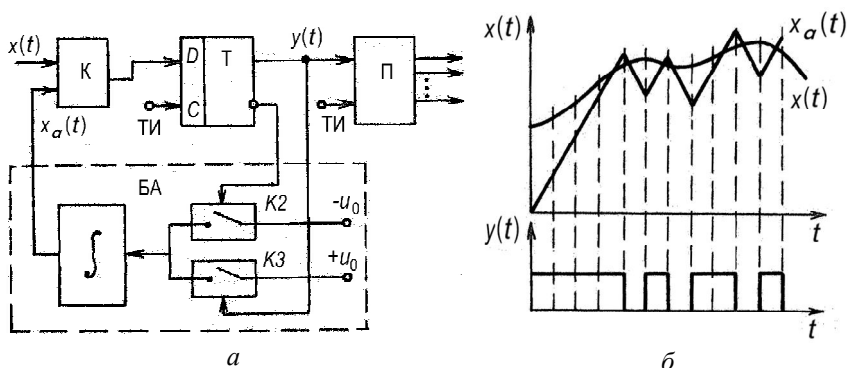


Рис. 4.14

Входной сигнал $x(t)$ сравнивается с выходным сигналом $x_a(t)$ блока аппроксимации, выполненном на интегрирующем усилителе. На выходе компаратора формируется логический сигнал, соответствующий знаку разности между текущим и аппроксимированным значениями сигнала. Он определяет знак величины опорного напряжения u_0 , подаваемого на вход интегратора, и, следовательно, управляет изменением напряжения на его выходе.

Выходной сигнал дельта-модулятора – последовательность значений логических сигналов 1 и 0. Она должна быть преобразована в параллельный двоичный код. Это можно сделать с помощью реверсивного счетчика, который выполняет операцию сложения при поступлении 1 и вычитания при поступлении 0.

Данному преобразователю свойственна динамическая погрешность, возникающая в случае, когда скорость изменения входного сигнала превышает скорость изменения напряжения на выходе интегратора. Определяющими параметрами при этом служат частота тактовых импульсов и величина u_0 .

Развитием этого принципа являются алгоритмы адаптивной дельта-модуляции.

Вариант параллельно-последовательной структуры преобразователя можно представить следующим образом. Входное напряжение подается одновременно на вход АЦП1 и на один из входов усилителя. АЦП1 выполняет грубое кодирование входного сигнала, формируя старшие разряды выходного кода. Выходы АЦП1 поступают на входы ЦАП, который формирует соответствующее напряжение. Это напряжение подается на инвертирующий вход усилителя. Усиленная разность входного сигнала и этого напряжения, представляющая собой погрешность квантования АЦП1, после фиксации кода и окончания переходных процессов в аналоговых цепях преобразуется с помощью АЦП2, выходной код которого объединяется с выходным кодом АЦП1.

Большинство АЦП имеет встроенный источник опорного напряжения и работает от одного источника питания (обычно 5 В).

Современные ЦАП характеризуются снижением энергопотребления и тенденцией к снижению стоимости. В связи с этим расширяется сфера применения ЦАП: в частности, они используются

вместо переменных и подстроечных резисторов. Для такого рода приложений достаточно иметь точность 8...10 разрядов и среднее быстродействие. Так, например, ЦАП типа AD8802/04 фирмы Analog Device, область применения которого – электронная подстройка (замена потенциометров), имеет 8 разрядов при 12 каналах установки с временем 0,6 мкс. Он потребляет мощность 55 мВт, работает от источника 3 В и имеет последовательный интерфейс для связи с системой.

4.7. Основные характеристики преобразователей

В зависимости от источника возникновения принято различать методические и инструментальные погрешности.

Методические погрешности АЦП являются следствием квантования дискретной величины сигнала по уровню и связанного с этим округления результата из-за конечного числа разрешенных уровней квантования. Исследованию этих погрешностей посвящено большое количество работ, поскольку в общем случае эти погрешности имеют сложный характер и многообразные формы проявления.

В простейшем случае, когда шкала уровней равномерна и шаг квантования постоянный, максимальная погрешность квантования зависит от способа отождествления сигнала с уровнем квантования. Для случая отождествления с ближайшим меньшим уровнем максимальная погрешность $\Delta x_k = \max |x(t_i) - y_k| = q$; для случая отождествления с ближайшим уровнем $\Delta x_k = 0,5q$. Рассматривая сигнал $x(t_i) = x$ как реализацию случайной величины X с плотностью распределения $f(x)$, найдем для этого случая математическое ожидание погрешности:

$$M[\Delta x_k] = \int_{x_k - q/2}^{x_k + q/2} (x - x_k) f(x) dx.$$

Полагая, что величина q мала по сравнению с диапазоном изменения сигнала, примем, что $f(x) = f(x_k)$. При этом условии $M[\Delta x_k] = 0$. Соотношение для оценки дисперсии погрешности

$$D[\Delta x_k] = \int_{x_k - q/2}^{x_k + q/2} (x - x_k)^2 f(x) dx$$

при аналогичном условии получает следующий вид:

$$D[\Delta x_k] = q^3 f(x_k) / 12.$$

Обозначим вероятность попадания величины X в k -й интервал через p_k . При малом q можно принять $p_k \approx f(x_k)q$. Поэтому

$$D[\Delta x_k] = q^2 p_k / 12.$$

Суммируя по всему диапазону ($k = 1, 2, \dots, N$), получаем, что

$$D[\Delta x] = q^2 / 12.$$

Инструментальные погрешности возникают вследствие отклонения параметров и характеристик элементов АЦП от идеальных и (или) заданных. К их числу относятся: смещение нуля и изменение крутизны передаточной характеристики (вход–выход) АЦП, а также погрешности, вызванные нелинейностью. Эти погрешности относятся к наиболее трудно устранимым, поскольку их нельзя устранить регулировкой.

Различают интегральную и дифференциальную нелинейность.

Интегральная нелинейность – это максимальное отклонение разности двух входных аналоговых сигналов, вызывающих последовательную смену выходных кодов АЦП, от значения, соответствующего единице младшего разряда кода при нулевых значениях погрешностей смещения нуля и изменения крутизны. Она выражается в долях младшего разряда или в процентах от предела шкалы. Интегральная нелинейность (или просто нелинейность) служит оценкой кривизны передаточной характеристики АЦП в целом.

Если каждый переход от одного кода к смежному происходит при одинаковых приращениях входного сигнала, равных единице младшего разряда, то интегральная нелинейность равна нулю.

Дифференциальная нелинейность – это отклонение передаточной характеристики АЦП от идеальной в точках, каждая из которых равноудалена от двух соседних уровней квантования. Дифференциальная нелинейность измеряется в долях или процентах от младшего разряда. Она характеризует девиацию шага квантования и монотонность преобразования. Если дифференциальная нели-

нейность превышает единицу младшего разряда, то возможен пропуск кода.

На рис. 4.15 абсцисса – входное напряжение, а ордината – двоичный код на выходе преобразователя. Пунктирная прямая – это передаточная характеристика идеального преобразователя. Сплошной линией показано смещение нуля преобразователя. В диапазоне входных напряжений, соответствующих выходным кодам 00000...00100, преобразователь работает идеально. В диапазоне кодов 00100...01001 определяется «монотонная нелинейность» – изменение крутизны передаточной характеристики. В диапазоне 01001...10011 нелинейность имеет немонотонный характер, который описывается дифференциальной и интегральной нелинейностями. Случай пропуска кода показан для кода 10010.

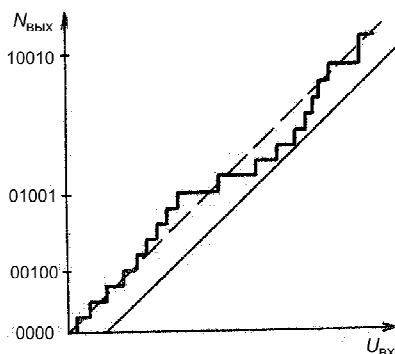


Рис. 4.15

Дифференциальная нелинейность характерна для преобразователей последовательной аппроксимации, что объясняется использованием в ЦАП взвешенных источников тока. Для интегрирующих АЦП преобладающей является интегральная нелинейность, обусловленная главным образом нелинейностью интегратора.

Изменение температуры окружающей среды является причиной изменения величины смещения нуля, крутизны передаточной характеристики и нелинейности.

Обычно справочные данные по АЦП (в том числе – по нелинейности) приводятся для статического режима, т.е. они измеряются при неизменном входном сигнале. Однако в динамическом режиме передаточная характеристика существенно изменяется.

Динамические свойства АЦП определяются переходными процессами в аналоговых электронных схемах, временем развертки значений меры и формирования цифрового эквивалента. Суммарное время составляет *время преобразования*. Полное время цикла преобразования включает в себя также время на установку преоб-

разователя в исходное состояние и время для синхронизации с началом такта.

Одно из проявлений переходных процессов в аналоговых цепях – задержка истинного момента запуска АЦП относительно номинального (*апертурное время* t_a). Поскольку апертурное время – величина случайная, то оно вносит неустранимую погрешность: за время t_a входной сигнал может измениться на величину $t_a x'(t)$, где $x'(t)$ – скорость изменения входного сигнала в момент t . Полная динамическая погрешность определяется полным временем цикла преобразования. Динамическая погрешность связывает характеристики преобразователя и свойства измеряемого процесса, поскольку максимальная производная x'_{\max} связана с частотой среза f_c спектра входного сигнала, имеющего максимальное значение x_{\max} , соотношением $x'_{\max} \leq 2\pi f_c |x_{\max}|$.

Динамическая погрешность проявляется только в том случае, если она превышает разрешающую способность преобразователя.

Разрешающая способность определяется как минимальная величина изменения входного сигнала, вызывающая изменение выходного кода на единицу младшего разряда. Она может задаваться в процентах от полного диапазона, в единицах измерения входной величины или, как это чаще всего делается, числом разрядов преобразователя.

Контрольные вопросы и упражнения

1. Постройте временные диаграммы, поясняющие работу преобразователя последовательного счета и появление задержек Δt_1 и Δt_2 .
2. Разработайте блок-схему преобразователя фазы в код.
3. Разработайте блок-схему АЦП последовательного счета.
4. Запишите в формализованном виде алгоритм метода поразрядного кодирования.
5. Разработайте блок-схему АЦП следящего преобразования.
6. Предложите принцип адаптивного преобразования и блок-схему преобразователя (указание: нужно задаться величиной его разрешающей способности).

7. Предложите схему простого интегрирующего преобразователя (с однократным интегрированием). Укажите основные источники погрешностей этого преобразователя.

8. Напишите формулу для оценки погрешности интегрирующего преобразователя, возникающей из-за нелинейности интегратора.

9. Выведите зависимость между частотой среза входного сигнала и его максимальной производной.

10. Нарисуйте, как зависит от величины входного сигнала погрешность квантования:

а) для двух способов отождествления сигнала с уровнем квантования при идеальной передаточной характеристике;

б) при смещении нуля передаточной характеристики;

в) при изменении крутизны передаточной характеристики;

г) при наличии интегральной нелинейности;

д) при наличии дифференциальной нелинейности.

11. Нарисуйте временную диаграмму работы преобразователя и покажите на ней временные характеристики АЦП и период дискретизации входного сигнала. Что ограничивает сверху и снизу частоту дискретизации?

12. Проведите анализ совместной работы схемы выборки-хранения (СВХ) и АЦП. Какие погрешности характеризуют их совместную работу? При каких условиях можно исключить СВХ на входе АЦП?

13. От каких параметров АЦП зависит его стоимость? Предложите варианты этой зависимости.

5. СРЕДСТВА ПЕРЕРАБОТКИ ИНФОРМАЦИИ

5.1. Типы средств переработки информации

Функции по переработке информации возлагаются на вычислительный комплекс, который в простейшем случае представляет собой одиночную ЭВМ, а в сложных АСУ РВ – многомашинную или многопроцессорную вычислительную систему. Можно выделить следующие виды средств переработки информации в АСУ реального времени.

1. *Контроллер на базе персонального компьютера.* Основные сферы использования – небольшие специализированные системы автоматизации в медицине, научных лабораториях, средствах коммуникации и для АСУТП небольших объектов с числом входов/выходов порядка нескольких десятков. Выполняемые функции – обработка измерительной информации.

2. *Локальный контроллер.* Используются два типа таких контроллеров: 1) встраиваемый в оборудование и являющийся его неотъемлемой частью (станки с числовым программным управлением, автомашины, аналитические приборы); 2) автономные, реализующие функции контроля и управления небольшим изолированным объектом. Контроллеры первого типа выпускаются без кожухов, второго – в защитных корпусах для разных условий окружающей среды. Локальные контроллеры рассчитаны на десятки входов/выходов, их вычислительная мощность невелика. Обычно они имеют интерфейсы для связи с другими средствами автоматизации.

3. *Сетевой комплекс контроллеров.* Это наиболее распространенный тип. Его минимальный состав – набор контроллеров, дисплейные пульта операторов, промышленная сеть для передачи коротких сообщений. Функции этих контроллеров – контроль и управление с числом входов/выходов порядка тысяч.

4. *Распределенные системы управления малого масштаба.* Основное отличие от предыдущего типа – большее разнообразие мо-

дификаций, блоков ввода/вывода, панелей оператора; большая мощность центральных процессоров контроллеров, позволяющая им обрабатывать более 10000 входных/выходных сигналов; выделение отдельных конструктивов удаленных блоков ввода/вывода, рассчитанных на работу в различных условиях окружающей среды. Эти контроллеры могут поддерживать несколько уровней сетей – полевые сети для связи с датчиками, сети первого уровня для связи контроллеров между собой и сети второго уровня для связи с оператором всего участка производства. В дополнение к обычным функциям управления эти контроллеры могут реализовывать сложные алгоритмы управления (например, задачи статической и динамической оптимизации).

5. Полномасштабные распределенные системы управления.

Этот тип контроллеров практически не имеет ограничений ни по выполняемым функциям, ни по объему автоматизируемого объекта. Дополнительно к рассмотренным ранее возможностям эти контроллеры имеют следующие свойства:

- наличие промышленных сетей, позволяющих подсоединять к одной шине сотни узлов (контроллеров и пультов);
- существование весьма мощных модификаций контроллеров;
- использование информационных сетей для связи пультов операторов друг с другом, для связи с серверами, с корпоративной сетью предприятия;
- взаимодействие пультов управления в клиент-серверном режиме;
- наличие в составе программного комплекса пакетов, реализующих функции эффективного управления отдельными агрегатами (многосвязное регулирование, оптимизация и т.п.), диспетчерского управления участками производства (поддержка принятия управленческих решений), технического учета и планирования производства в целом.

Одиночные ЭВМ могут представлять собой микроконтроллеры, скалярные, суперскалярные и векторные процессоры с системами памяти, устройствами управления и средствами ввода/вывода.

Микроконтроллер – это законченная микропроцессорная система обработки информации, реализованная в виде одной БИС и объединяющая в пределах одного полупроводникового кристалла основные блоки управляющей системы: центральный процессор, постоянное запоминающее устройство, оперативное запоминающее устройство, периферийные устройства для ввода и вывода информации.

Индустриальная рабочая станция — это системный блок, дисплей и клавиатура в одном конструктиве.

Скалярный процессор – это программно-управляемое устройство для обработки цифровой информации и управления процессом обработки, содержащее минимально необходимый набор устройств для выполнения этих операций: операционное (арифметико-логическое) устройство с устройством управления, внутреннюю сверхоперативную память и внутреннюю шину для обмена информацией.

Суперскалярный процессор имеет несколько операционных устройств, что позволяет ему вести параллельную обработку информации.

Векторный процессор имеет структуру и систему команд, обеспечивающие возможность выполнения векторных операций.

Рост потребностей пользователей и развитие идей параллелизма привело к созданию многомашинных и мультипроцессорных вычислительных систем.

Многомашинная система состоит из нескольких машин, каждая из которых имеет свою внутреннюю память и работает под управлением своей операционной системы, и средства обмена информацией между машинами.

Мультипроцессорные системы имеют следующие отличительные особенности: наличие единой операционной системы и общих периферийных устройств; возможность перераспределения программ между процессорами; возможность прерывания работы одних процессоров в зависимости от работы других.

5.2. Микроконтроллеры

Микроконтроллер – это законченная микропроцессорная система обработки информации, реализованная в виде БИС и объединяющая в пределах одного полупроводникового кристалла центральный процессор, постоянное запоминающее устройство (ПЗУ), основную память – оперативное запоминающее устройство (ОЗУ) и устройства ввода/вывода информации.

Однокристальному микроконтроллеру (ОМК) присущи следующие особенности:

- система команд ориентирована на выполнение задач управления и регулирования;
- алгоритмы, которые реализуются на ОМК, могут иметь много разветвлений в зависимости от внешних сигналов;
- данные, с которыми оперируют ОМК, имеют небольшую разрядность;
- схемная реализация систем управления на базе ОМК является несложной и имеет невысокую стоимость;
- универсальность и возможность расширения функций управления значительно ниже, чем в системах с однокристальными микропроцессорами.

Широкое разнообразие моделей микроконтроллеров и возможность разрабатывать и производить их в короткие сроки обеспечивается модульным принципом построения ОМК: все ОМК одного семейства содержат в себе одинаковый базовый функциональный блок и различные изменяемые функциональные блоки. Базовый блок принято называть *процессорным ядром* ОМК.

В состав процессорного ядра входят:

- центральный процессор;
- внутренние шины адреса, данных и управления;
- схемы синхронизации, формирующие многофазные импульсные последовательности для тактирования центрального процессора и шин;
- схема управления режимами работы ОМК.

Ядро современных реализуют как на основе CISC-архитектуры, так и на основе RISC-архитектуры.

ОМК с CISC-архитектурой имеют однобайтовый, двухбайтовый и трехбайтовый (реже четырехбайтовый) формат команд. Выборка команды из памяти осуществляется побайтно в течение нескольких машинных циклов. Время выполнения каждой команды с учетом времени выборки в большинстве случаев составляет от 1 до 10 циклов.

ОМК с RISC-архитектурой имеет фиксированный формат команды (12, 14 или 16 бит). Выборка из памяти и исполнение подавляющего большинства команд осуществляется за один машинный цикл.

Модульная структура ОМК показана на рис. 5.1. Изменяемый функциональный блок включает модули основной и постоянной (ОЗУ и ПЗУ) памяти, генератор схемы синхронизации (Г) и модули внешних устройств. К внешним устройствам обычно относятся: порты ввода/вывода, таймеры, аналого-цифровые преобразователи (АЦП), контроллеры последовательного ввода/вывода (К).

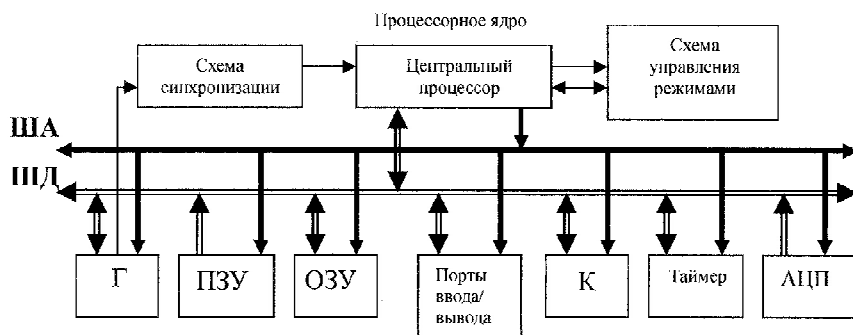


Рис. 5.1

В последних разработках в состав изменяемых модулей включают модули контроля за напряжением питания и ходом выполнения программы и модули внутрисхемной отладки и программирования.

Каждый модуль имеет выводы для подключения его к шинам адреса (ША) и данных (ШД) процессорного ядра.

5.2.1. Функционирование узлов микроконтроллера CISC-типа

Узлы ОМК и организацию их взаимодействия рассмотрим на примере микроконтроллера семейства HC08 (рис. 5.2).

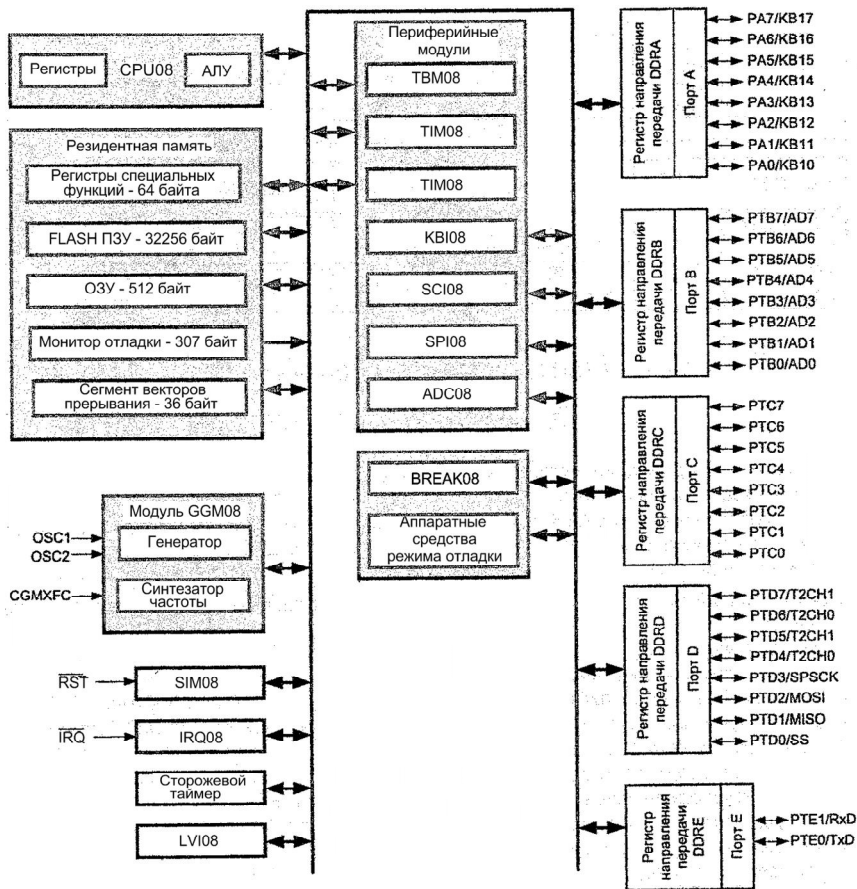


Рис. 5.2

Центральный процессор выполняет операции над 8-разрядными операндами, поддерживает 11 способов адресации, имеет команды умножения и деления. Он содержит АЛУ и пять регистров:

1) *аккумулятор* – 8-разрядный регистр для хранения операндов и результатов операций;

2) *индексный 16-разрядный регистр* участвует в формировании адреса операнда, а также может использоваться для промежуточного хранения результатов и как источник одного из операндов в операциях умножения и деления;

3) *программный счетчик* – 16-разрядный регистр, содержащий адрес текущей команды или адрес операнда, используемого в текущей команде;

4) *указатель стека* – 16-разрядный регистр, содержащий адрес вершины стека;

5) *регистр слова состояния процессора* – 8-разрядный регистр, содержит шесть флагов условий: переноса, нулевого результата, отрицательного результата, переполнения, дополнительного переноса и бит глобальной маски прерывания.

Система команд процессора CPU08 включает 90 команд, которые делятся на шесть групп [6]:

- команды загрузки и пересылки (всего 13 команд) осуществляют перемещение данных между ячейками памяти и регистрами центрального процессора;

- арифметические команды (14 команд) выполняют операции сложения, вычитания, инкремента и декремента, однобайтового и двухбайтового сравнения, умножения и деления;

- логические команды и команды сдвигов (15 команд) – логическое И, ИЛИ, Исключающее ИЛИ, различные операции сдвига, операция взятия обратного и дополнительного кода;

- команды битового процессора (всего 6 команд) устанавливают в определенные состояния отдельные биты в ячейках ОЗУ или в регистрах;

- команды управления ходом вычислительного процесса (40 команд), в состав которых входят: команды условных и безусловных переходов, вызова подпрограмм и обслуживания прерываний;

- команды перехода к режимам пониженного энергопотребления (2 команды) переводят МК в режим ожидания или останова.

Режимы работы

1. *Режим исполнения прикладной программы.* Работа при номинальной частоте генератора; активизированы все необходимые периферийные модули.

2. *Режим ожидания.* Для экономии энергии понижается частота таймера и может быть отключена часть периферийных устройств.

3. *Режим останова.* Вводится с помощью команды STOP. Вся периферия неактивна.

4. *Режим отладки.* Обеспечивает программирование и отладку МК с помощью последовательного интерфейса с ведущим персональным компьютером.

Резидентная память. Она состоит из энергонезависимого запоминающего устройства (FLASH ПЗУ), основного запоминающего устройства (ОЗУ), регистров специальных функций, монитора отладки и сегмента векторов прерываний.

Энергонезависимая память в составе ОМК имеет не только режимы хранения и чтения информации, которая была записана до начала эксплуатации на этапе программирования, но и режимы стирания и программирования под управлением прикладной программы в процессе управления объектом и без перевода ОМК в режим программирования. Функциональная программа обычно хранится в памяти типа FLASH, константы – в перепрограммируемом постоянном запоминающем устройстве с электрическим стиранием (EEPROM ПЗУ), которое имеет режим побайтного программирования. Такая организация позволяет в ходе выполнения программы изменять содержимое EEPROM ПЗУ, что делает его идеальным средством для хранения перепрограммируемых констант (например, различных настроек, которые должны сохраняться при отключении питания). Для выполнения перепрограммирования в процессе работы энергонезависимая память снабжается дополнительными схемами управления. Время перепрограммирования – менее 1 с.

В состав ОЗУ ОМК входит оперативное ЗУ статического типа. Важно обратить внимание, что ОЗУ *статическое*: современные ОМК допускают снижение частоты тактирования до сколь угодно

малых значений с целью снижения энергопотребления, и содержащее статического ЗУ, в отличие от динамического, при этом сохраняется.

ОМК семейства HC08 имеют объединенное адресное пространство памяти программ, данных и регистров специальных функций периферийных модулей. Внутренняя шина позволяет адресовать 64К памяти. В верхних адресах памяти располагаются регистры специальных функций (64 байта): регистры параллельных портов ввода/вывода, регистры данных и регистры управления периферийных модулей. Далее следуют ОЗУ емкостью 512 байт и память программ в виде FLASH-памяти емкостью 32 Кбайта.

Модуль FLASH-памяти имеет в своем составе повышающий преобразователь напряжения, который позволяет выполнять стирание и программирование FLASH ПЗУ под управлением программы, размещенной в ОЗУ, без подключения внешнего источника напряжения программирования. В нижних адресах памяти располагаются монитор отладки (307 байт) и сегмент векторов прерываний (36 байт), где хранятся адреса начальных адресов подпрограмм прерывания.

Модуль памяти монитора отладки позволяет организовать режим внутрисистемного программирования с помощью дополнительных программных и аппаратных средств – симуляторов. В частности, на отдельной ЭВМ может быть реализована программная модель ОМК; в этом случае монитор отладки осуществляет связь программы, записанной в резидентной памяти ОМК, с этой моделью. Процесс программирования и отладки функциональной программы становится экономичнее.

В сегменте векторов прерывания по два байта памяти отведено каждому источнику прерываний. Пользователь на этапе программирования МК заносит в эти ячейки адреса начала программ прерывания. При переходе к выполнению определенной программы ее адрес автоматически загружается в счетчик команд центрального процессора.

Таймер ТВМ08 представляет собой 16-разрядный счетчик количества циклов со схемой управления. Направление счета – толь-

ко прямое; регистры доступны для чтения и записи. При переполнении формируется сигнал прерывания.

Процессоры событий. Модуль процессора событий (TIM08) состоит из 16-разрядного таймера-счетчика и некоторого количества связанных с ним полностью идентичных каналов захвата/сравнения. Каждый канал захвата/сравнения связан с одним из выводов ОМК.

В составе ОМК HC80 два модуля процессора событий и каждый из модулей имеет по два канала захвата/сравнения.

Действие канала входного захвата состоит в подсчете внешних событий. Схема «детектора событий» следит за уровнем напряжения на входе. При изменении уровня логического сигнала вырабатывается строб записи, и текущее состояние счетчика-таймера записывается в 16-разрядный регистр данных канала захвата.

Действие канала выходного сравнения аналогично режиму задержек. Многоразрядный цифровой компаратор непрерывно сравнивает изменяющийся во времени код счетчика таймера с кодом, который записан в 16-разрядном регистре канала сравнения. В момент равенства кодов на одном из выводов ОМК устанавливается заданный уровень логического сигнала. Этот режим можно использовать для получения широтно-модулированного импульсного сигнала: при наступлении события выходного сравнения на выходе устанавливается низкий логический уровень, а при переполнении таймера-счетчика уровень сигнала на выходе инвертируется.

В схеме ОМК предусмотрена возможность пуска и останова таймера-счетчика под управлением программы и, кроме того, возможность считывания кода. При переполнении таймера-счетчика устанавливается флаг переполнения и генерируется запрос на прерывание (если он разрешен).

Аналого-цифровые преобразователи. Характерная особенность HC08 – интегрированный на кристалле ОМК модуль многоканального аналого-цифрового преобразователя ADC08 (АЦП). Многоканальный аналоговый коммутатор на его входе служит для подключения одного из источников аналоговых сигналов. Выбор

источника сигнала для измерения осуществляется посредством записи номера канала коммутатора в соответствующие разряды регистра управления АЦП.

Собственно АЦП выполняется по схеме последовательного приближения или поразрядного уравнивания. Разрешающая способность в разных моделях ОМК составляет от 8 до 12 бит, время преобразования – от 5 до 32 мкс, число каналов – от 8 до 12.

Модуль системной интеграции SIM08 управляет всеми подсистемами микроконтроллера. Он формирует множество сигналов, которые обеспечивают совместное функционирование центрального процессора, блоков памяти и периферийных модулей. В частности, он содержит *контроллер прерываний*, который производит сбор и ранжирование внутренних и внешних запросов, формирует вектор текущего прерывания в соответствии с таблицей приоритетов, управляет процедурой перехода к выполнению подпрограммы перехода и выходом из нее.

Прерывания изменяют текущую последовательность выполнения команд для обслуживания внешнего или внутреннего события, вызвавшего прерывание. Прерывание не останавливает выполнение текущей команды; она завершается, а затем реализуется механизм перехода к подпрограмме прерывания:

- в стеке сохраняется содержимое: регистра признаков, аккумулятора, индексного регистра, счетчика команд;
- бит глобальной маски прерывания в регистре признаков устанавливается в 1, запрещая обслуживание любых других запросов на прерывание до тех пор, пока не будет завершена текущая подпрограмма прерывания;
- счетчик команд загружается адресом начала подпрограммы прерывания (вектором прерывания).

Обработка прерываний осуществляется модулем внешних прерываний. Вектору прерывания отведена специальная область – сегмент вектора прерываний. Каждому источнику запросов на прерывание в этой области отведено по два байта памяти с фиксированными адресами. Пользователь на этапе программирования МК заносит в эти ячейки адрес начала подпрограммы прерывания и

при переходе к ней этот адрес автоматически загружается в счетчик команд центрального процессора.

В ОМК HC80 входят 16 источников событий, требующих прерывания: по одному – по запросам таймера, АЦП, модуля сканирования клавиатуры, модуля формирования тактовой частоты и вектора прерывания по входу «Запрос прерывания»; пять – от схем последовательных интерфейсов; шесть – по событиям процессоров событий. Один источник – вектор программного прерывания – немаскируемый.

Подсистема прерываний ОМК семейства HC80 имеет жесткое распределение приоритетов, которое нельзя скорректировать программными настройками.

Модуль управления питанием CGM08 состоит из генератора и синтезатора частоты. Он обеспечивает тактирование от генератора кварцевого резонатора или от синтезатора частоты, который позволяет снизить частоту резонатора.

Порты ввода/вывода. В процессоре HC08 имеется 33 двуправленные линии ввода/вывода: порт А – для сканирования клавиатуры, порт В – аналоговые входы встроенного АЦП; порты С, D, Е – для связи с выходами периферийных модулей.

Контроллеры последовательного ввода/вывода. Задачи, которые решаются с помощью контроллеров последовательного ввода/вывода, могут быть условно разделены на три группы:

1) связь встраиваемой микропроцессорной системы с системой управления верхнего уровня;

2) связь с внешними по отношению к ОМК периферийными устройствами, а также с датчиками физических величин с последовательным выходом;

3) интерфейс связи с локальной сетью в мультимикропроцессорных системах.

В составе ОМК семейства HC08 имеются два контроллера последовательного интерфейса – синхронного и асинхронного. Модуль синхронного интерфейса (SPI08) предназначен для высокоскоростного обмена между ОМК и периферийными микросхемами, такими как АЦП и ЦАП, FLASH-память большой емкости и др.

Основные элементы контроллера SPI08 – 8-разрядный сдвиговый регистр и два буферных регистра данных, программно доступных по одному и тому же адресу (один регистр для передатчика, один – для приемника). Поскольку SPI08 поддерживает два режима работы – ведущего и ведомого, он имеет четыре линии: передачи данных от ведущего к ведомому, передачи данных от ведомого к ведущему, сигнала синхронизации, выбора ведомого. Направление передачи каждой линии определяется выбором режима контроллера.

Модуль асинхронного интерфейса (SCI08) организует обмен с устройствами управления верхнего уровня. Аппаратные средства SCI08 делятся на две подсистемы: приемника и передатчика. Основу каждой подсистемы составляют сдвиговый регистр и буферный регистр данных.

Модуль сброса при некорректных операциях и адресах IRQ08 срабатывает, когда появляется запрещенная команда (она не может интерпретироваться существующей системой команд) или несуществующий (для данной серии микроконтроллеров) адрес.

Модуль сторожевого таймера генерирует сигнал RESET через определенное время, если его не сбросить внутри этого временного интервала. Для осуществления этого сброса программист должен регулярно вставлять в программу команду сброса этого таймера.

Модуль защиты от пониженного напряжения источника питания LVI08 постоянно контролирует напряжение источника и вызывает сброс, когда оно падает ниже порогового значения.

Модуль связи с клавиатурой KBI08 обеспечивает обмен с клавиатурой.

Аппаратные средства режима отладки формирует необходимые уровни сигналов для перевода микроконтроллера в отладочный режим работы (IRQ, RST, частота тактирования и т.д.). Модуль BREAK08 реализует остановки по контрольным точкам, в которых можно проверить содержимое регистров. Для этого в регистр модуля заносится абсолютный адрес контрольной точки и запускается программа, которая сравнивает его с кодом на внутренней магистрали адреса. При равенстве кодов генерируется прерывание, микроконтроллер переходит на программу монитора отладки, ко-

торая считывает его внутренние программно доступные регистры в память и отображает их на экране дисплея.

5.2.2. Функционирование узлов микроконтроллера RISC-типа

Типичными представителями RISC-процессоров являются 8-разрядные микроконтроллеры общего назначения фирмы Atmel Corp. Их архитектуру рассмотрим на примере БИС AT90S8535 (рис. 5.3).

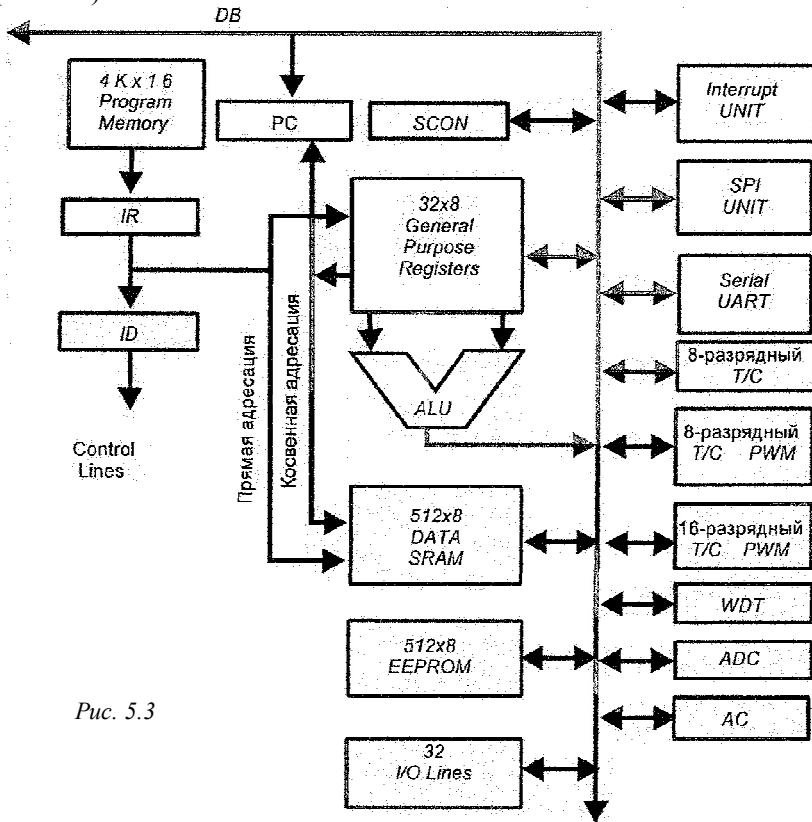


Рис. 5.3

Основой архитектуры являются отдельные шины и области памяти для команд и данных (гарвардская архитектура). Разряд-

ность шины памяти программ расширена до 16 бит. В микроконтроллере используется технология конвейеризации, вследствие чего цикл «выборка-выполнение» команды существенно сокращен. Одноуровневый конвейер в общем потоке позволяет выполнять короткие команды за один машинный цикл – за один период тактовой частоты.

Структурная схема контроллера содержит (см. рис. 5.3):

- арифметико-логическое устройство (ALU), выполняющее 133 различных команды; разрядность обрабатываемого слова – 8 бит;

- регистровый файл (General Purpose Register), состоящий из регистров общего назначения, занимает младшие 32 байта в общем адресном пространстве; шесть из 32 регистров могут использоваться как три 16-разрядных указателя адреса при косвенной адресации данных;

- память команд (Program Memory), которая содержит 4К 12-разрядных ячеек FLASH-памяти с числом циклов перезаписи не меньше 1000 и с способностью самопрограммирования, то есть микроконтроллер способен самостоятельно, без внешнего программатора изменять содержимое ячеек памяти программ. Это дает возможность записать во внешнюю энергонезависимую память несколько рабочих версий программы, а потом по мере необходимости или по реакции на внешние или внутренние логические условия перегружать рабочие программы в тот же микроконтроллер без извлечения его из печатной платы;

- память данных, состоящая из внутренней оперативной статической памяти SRAM и энергонезависимой памяти с электрическим стиранием EEPROM. Этот тип памяти используется для хранения промежуточных данных, констант, таблиц перекодировок, калибровочных коэффициентов и т.п. Число циклов перезаписи – не менее 100000;

- периферия, традиционно состоящая из 8-разрядных портов последовательного ввода/вывода (Serial UART) с блоком управления (SCON), таймеров/счетчиков (T/C) и контроллера прерываний (PC). Число независимых линий портов ввода/вывода – до 53. Каж-

дый разряд порта может быть запрограммирован на ввод или на вывод информации. Мощные выходные драйверы обеспечивают токовую нагрузочную возможность 20 мА по линии порта при максимальном значении 40 мА. Таймеры/счетчики могут работать как таймеры от внутреннего источника опорной частоты и как счетчики внешних событий с внешним тактированием.

Периферийными устройствами являются только блок последовательного периферийного интерфейса (SPI), сторожевой таймер (WDT) и аналоговый компаратор (AC). Блок SPI предназначен для последовательного ввода и вывода данных и используется для программирования микроконтроллера после его установки на печатную плату. Сторожевой таймер предназначен для перезапуска программы при появлении сбоя в ходе ее выполнения: программа, которая работает без сбоя, периодически сбрасывает сторожевой таймер, не допуская его переполнения. Аналоговый компаратор сравнивает по величине напряжения сигналы, поступающие на два его входа; результат используется во внутренних цепях микроконтроллера;

- аналого-цифровой преобразователь (ADC) построен по классической схеме поразрядного уравнивания с устройством выборки/хранения. Каждый из аналоговых входов может быть соединен со входом устройства выборки/хранения через аналоговый мультиплексор. Разрядность АЦП составляет 10 бит. АЦП может работать в двух режимах – однократное преобразование по любому выбранному каналу и последовательный циклический опрос всех каналов.

В режиме широтно-импульсной модуляции (ШИМ-генератор) на выходе блока формируется последовательность импульсов с периодом, равным периоду работы счетчика временной базы, и с длительностью, пропорциональной коду в регистре данных (рис. 5.4). Назначение счетчика команд (R), контроллера прерываний (PC) и дешифратора команд (D) общеизвестно.

Многие МК имеют специализацию, т.е. специальные встроенные функциональные особенности архитектуры.

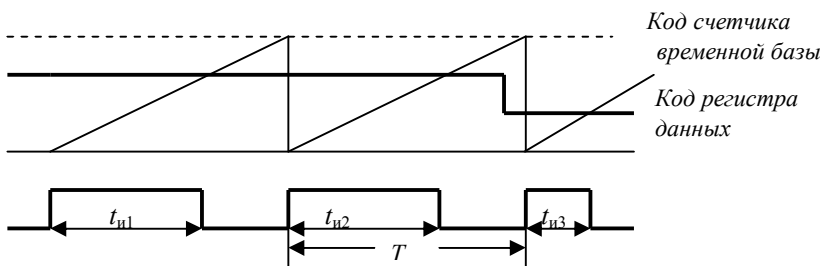


Рис. 5.4

5.2.3. Специализированные микроконтроллеры

Однокристалльные микроконтроллеры с интегрированными устройствами высокоскоростного ввода/вывода ориентированы на решение задач управления объектами и процессами в реальном масштабе времени. Структурная схема содержит:

- модуль центрального процессора (ЦП);
- контроллер памяти;
- контроллер прерываний;
- сервер периферийных транзакций;
- ПЗУ;
- блок аналого-цифрового преобразователя (АЦП) с устройством выборки/хранения и мультиплексором;
 - встроенный трехканальный ШИМ-генератор;
 - последовательный порт;
 - порты ввода/вывода;
 - сторожевой таймер;
 - модуль высокоскоростного ввода/вывода.

Модуль ЦП состоит из арифметико-логического устройства (АЛУ), регистра признаков, устройства микропрограммного управления и регистрового файла, который представляет собой массив регистров, т.е. ОЗУ статического типа. Центральный процессор связан с контроллером памяти, контроллером прерываний и встроенными периферийными устройствами внутренней 16-разрядной шиной. Кроме того, существует дополнительная 8-разрядная шина

для непосредственной передачи байтов команды из контроллера памяти в регистр команд, которая находится в устройстве микропрограммного управления.

Контроллер памяти осуществляет операции выборки/записи данных во внешнюю или внутреннюю память данных или команд. В состав контроллера памяти входит 4-байтная очередь команд.

Контроллер прерываний предназначен для эффективного управления в реальном времени. Обработки запросов прерываний осуществляется с учетом установленного уровня приоритета. В качестве запросов прерываний могут использоваться:

- внешние сигналы;
- запросы от интегрированных периферийных устройств;
- команды вызова подпрограмм обработки прерываний.

Сервер периферийных транзакций выполняет функции периферийного процессора ввода/вывода и освобождает ЦП от выполнения некоторых типичных операций ввода/вывода. Сервер периферийных транзакций работает по прерываниями, но обслуживает их не программно, а аппаратно.

Постоянное запоминающее устройство масочного типа или однократно программируемое емкостью от $8K \times 8$ до $32K \times 8$ байт.

В однократно программируемом ПЗУ (*OTPROM – One-Time Programmable ROM*) программированию подлежат только те разряды, которые после программирования должны содержать «0». Технология программирования состоит в многократном приложении импульсов повышенного напряжения к элементарным ячейкам, подлежащим программированию.

Аналого-цифровой преобразователь предназначен для ввода аналоговых сигналов и состоит из аналогового мультиплексора, устройства выборки/хранения и АЦП, выполненного по способу поразрядного уравнивания. Разрядность преобразователя – 8 двоичных разрядов. Время выборки – 1 мкс, время преобразования 10 – 20 мкс.

Модуль высокоскоростного ввода/вывода предназначен для формирования периодических во времени сигналов (например,

сигнала ШИМ для прямого цифрового управления инверторами напряжения приводов переменного или постоянного тока) и непосредственного измерения временных интервалов между внешними событиями. Информация о событии и времени генерации события в виде команды записывается в ассоциативную память. Каждая команда задает:

- тип события (сброс или установка потенциала на внешнем контакте, запуск АЦП, запуск программных таймеров);
- способ реагирования ЦП на событие – информация о том, будет ли генерироваться прерывание для обслуживания события;
- номер таймера (0 или 1), используемый для контроля текущего реального времени;
- время генерации события.

В каждом машинном такте (100 нс при тактовой частоте 20 МГц) происходит сравнение времени генерации события с реальным временем, которое отсчитывает таймер. Если они совпадают, то генерируется заданное командой событие.

Встроенный трехканальный ШИМ-генератор имеет ограничения на частоту сигналов (39,1 и 19,5 кГц) и не имеет ограничений на скважность импульсов: она регулируется от 1/255 до 1 с дискретностью 1/255. Он эффективно используется для прямого цифрового управления электронными коммутаторами, а также для формирования аналоговых задающих сигналов в цифроаналоговых системах управления.

Последовательный порт представляет собой универсальный синхронно-асинхронный приемопередатчик, который может настраиваться на один синхронный режим и три асинхронных.

Однокристалльный микроконтроллер с интегрированными процессорами событий отличаются наличием мощного процессора событий, модифицированным блоком встроенных таймеров/счетчиков, большей емкостью прямо адресованной внешней памяти (до 16 Мбайт), расширенными интерфейсными возможностями.

Модификация блока двух встроенных таймеров/счетчиков состоит в следующем:

- оба таймера являются реверсивными и могут работать как с внутренним, так и с внешним тактированием;
- есть возможность каскадного включения, когда второй таймер тактируется сигналами переполнения от первого таймера.

Массив процессора событий реагирует на такие типы внешних событий: переход сигнала на внешнем выводе из логической единицы в логический нуль и наоборот; переключение уровня сигнала на внешнем выводе. Модуль процессора событий содержит 16-разрядный счетчик временной базы и некоторое количество универсальных каналов захвата/сравнения.

Универсальные каналы захвата/сравнения в процессоре событий могут работать в одном из трех режимов: 1) режим входного захвата; 2) режим выходного сравнения; 3) режим широтно-импульсной модуляции (ШИМ).

Принцип действия входного захвата такой (рис. 5.5): схема детектора событий «наблюдает» за уровнем напряжения $U_{вх}$ на одном из входов МК. При изменении уровня логического сигнала вырабатывается строб записи, и текущее состояние счетчика таймера записывается в 16-разрядный регистр данных. Разность кодов даст длительность входного сигнала.

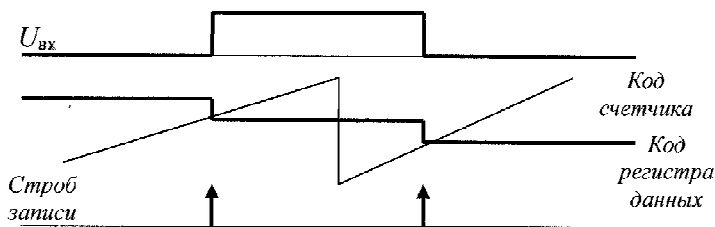


Рис. 5.5

Принцип действия канала выходного сравнения таков (рис. 5.6): многоразрядный цифровой компаратор непрерывно сравнивает изменяющийся во времени код счетчика таймера с кодом, который записан в 16-разрядном регистре канала сравнения. В момент

равенства кодов на одном из выводов МК устанавливается заданный уровень логического сигнала.

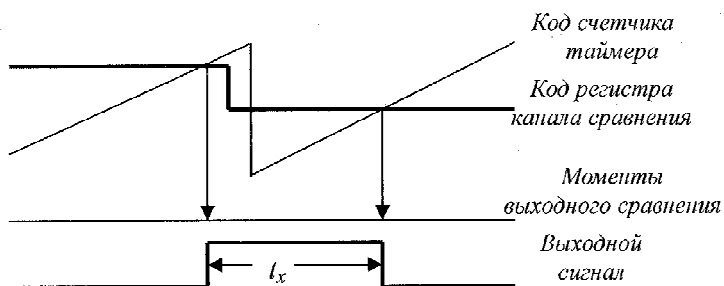


Рис. 5.6

В режиме ШИМ на выводе ОМК формируется последовательность импульсов с периодом, равным периоду работы счетчика временной базы. Длительность импульса прямо пропорциональна коду в регистре данных канала.

Однокристалльные микроконтроллеры с интегрированными средствами управления двигателями предназначены для прямого цифрового управления преобразовательными устройствами электроприводов. Кроме рассмотренных выше блоков ОМК, схема содержит дополнительные линии портов и сложный генератор периодических сигналов.

Линии семи портов имеют такое функциональное назначение: входы АЦП; четыре линии последовательного ввода/вывода; две линии последовательного ввода/вывода, два канала захвата/сравнения, четыре канала сравнения процессора событий; совмещенная шина адреса/данных; передача сигналов управления; функция генератора периодических сигналов.

Генератор периодических сигналов содержит: блок цифрового опорного генератора, который задает пилообразную или треугольную форму опорного цифрового сигнала модулятора; блок трехканального цифрового компаратора – драйвера формирования сигналов фаз, в котором на базе заданного значения скважности формируются три независимых ШИМ-сигнала; блок управления выходами генератора периодических сигналов.

Однокристалльный микроконтроллер с интегрированными средствами цифровой обработки имеет традиционную архитектуру, но дополнительно содержит цифровой сигнальный процессор для реализации функции цифрового регулятора. Это позволило:

- повысить производительность ЦП за счет уменьшения времени выборки из памяти программ и данных;
- использовать конвейерный режим с одновременным выполнением четырех команд для уменьшения времени машинного цикла;
- применить аппаратные средства умножения и деления ;
- дополнить систему команд специальными командами (умножение с накоплением, повторение, автоматизация обновления данных в таблицах последовательных выборок), оптимизированными для решения задач построения цифровых фильтров и регуляторов.

5.3. Микропроцессоры

5.3.1. Структура микропроцессора i8086

Микропроцессор i8086 – это 16-разрядное устройство с быстродействием до 2,5 млн. оп/с и объемом адресуемой памяти 1 Мбайт. Он ориентирован на параллельное выполнение циклов исполнения команд и выборки последующих команд и условно может быть разделен на две части, работающие асинхронно: устройство обработки операндов и устройство обработки команд.

Устройство обработки операндов предназначено для выполнения операций по обработке данных. В функции устройства обработки команд входит: формирование 20-разрядного физического адреса памяти, выборка команд и операндов из памяти, организация очередности команд, запоминание результатов выполнения команд в памяти.

В состав микропроцессора входят следующие блоки и устройства (рис. 5.7):

- 1) арифметико-логическое устройство (АЛУ);
- 2) устройство управления, содержащее

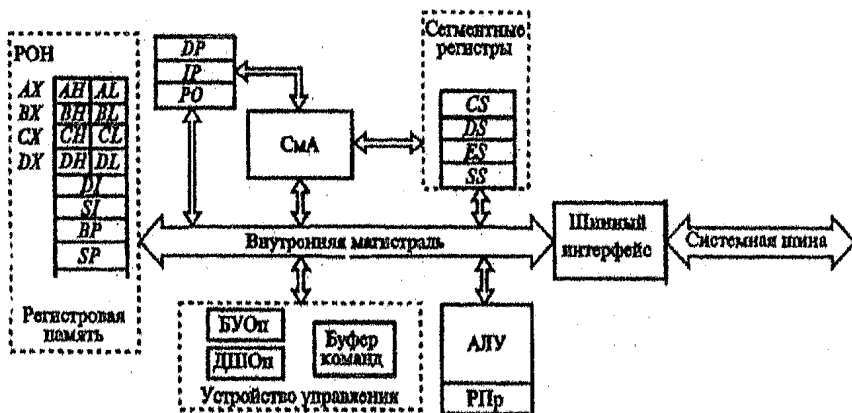


Рис. 5.7

- буфер команд – регистровую память объемом 6 байт, предназначенную для хранения очереди команд (команды, выполняемой в данный момент) и заполняющуюся очередными командами, считываемыми из основной памяти по мере освобождения;

- дешифратор операций (ДшОп);

- блок управления операциями (БУОп), который формирует управляющие сигналы для всех блоков микропроцессора в соответствии с расшифрованным кодом операции;

3) внутренняя регистровая память, состоящая из восьми регистров общего назначения (РОН), четыре из которых допускают раздельное использование младших и старших разрядов, что позволяет работать как с байтами, так и с 16-разрядными словами;

4) регистр адреса команды (указатель команд *IP*), в который помещается смещение адреса выполняемой команды относительно базового адреса;

5) регистр адреса данных (указатель данных *DP*), определяющий смещение адреса данных относительно базового адреса;

6) регистр признаков (РПр), содержащий признаки результата выполненной команды (получение нулевого результата, получение отрицательного результата, наличие переноса из третьего разряда в четвертый или займа из четвертого разряда в третий младшего

байта 16-разрядного числа (используется командами десятичной арифметики), наличие переноса из старшего бита числа в знаковый разряд при сложении или вычитании, наличие переполнения, четность числа единиц в коде результата) и дополнительную управляющую информацию;

7) блок из четырех сегментных регистров, содержащих базовые адреса оперативной памяти: кодового сегмента *CS*, где содержится программа; сегмента данных *DS*; сегмента стека *SS* и дополнительного сегмента данных *ES*;

8) сумматор адреса (*СМА*);

9) регистр обмена (*РО*);

10) шинный интерфейс – схемы, обеспечивающие связь внутренней магистрали микропроцессора с системной шиной.

Система команд микропроцессора содержит 135 команд, которые по функциональному назначению делятся на команды передачи данных, арифметические, поразрядной обработки данных, обработки строк данных, передачи управления, управления процессором.

Команды передачи обеспечивают пересылку операнда-источника на место операнда-приемника без содержательного преобразования. Арифметические команды (сложение, вычитание, умножение, деление, преобразование форматов) выполняют операции над 8- и 16-разрядными целыми двоичными числами с фиксированной и плавающей точками и двоично-десятичными числами. Команды поразрядной обработки предназначены для выполнения логических операций булевой алгебры, операции линейных и циклических сдвигов вправо и влево. Команды обработки строк позволяют обрабатывать как числа, так и тексты (длина обрабатываемых строк может достигать 64 Кбайт). Команды передачи управления позволяют осуществлять безусловные переходы, вызовы и возвраты, условные переходы, управление циклами и обслуживание прерываний. Команды управления процессором изменяют состояния определенных разрядов регистров признаков и используются для синхронизации работы с внешними событиями и процессорами.

Система команд – двухадресная, поэтому в ее составе нет команд, в которых оба операнда извлекаются из основной памяти. Допустимы только следующие сочетания: RR , RS , RI , SI (R – регистр регистровой памяти, S – оперативная память, I – операнд, заданный в адресном поле команды). При обращении к регистровой памяти указывается номер регистра; при обращении к основной памяти исполнительный (физический) 20-разрядный адрес формируется путем сложения *эффективного* адреса с увеличенным в 16 раз содержимым определенного сегментного регистра.

Эффективный адрес может

– быть частью команды;

– находиться в одном из регистров BX , DI , SI ;

– быть равен сумме содержимого одного из:

регистров BX , BP , DI , SI и 8- или 16-разрядного смещения;

базовых регистров (BX , BP) и одного из индексных регистров (DI , SI);

базовых регистров (BX , BP), одного из индексных регистров (DI , SI) и 8- или 16-разрядного смещения.

5.3.2. Микропроцессор i80386

Этот микропроцессор стал родоначальником семейства 32-разрядных схем и принятые в нем архитектурные решения до настоящего времени остаются характерными для последующих микропроцессоров фирмы Intel.

Микропроцессор i80386 оптимизирован для многозадачных ОС и для совместной работы пользователей, использующих разные ОС. Эти возможности поддерживаются архитектурой со встроенными устройствами управления памятью и защиты.

Микропроцессор содержит шесть блоков, реализующих выполнение команд, сегментацию, страничную организацию памяти, сопряжение с шинами, декодирование и упреждающую выборку команд (рис. 5.8). Все эти блоки и устройства работают в режиме конвейера: во время выполнения одной команды декодируется вторая и из памяти выбирается третья. Дополнительным средством

повышения производительности служит специальный блок быстрого умножения/деления.

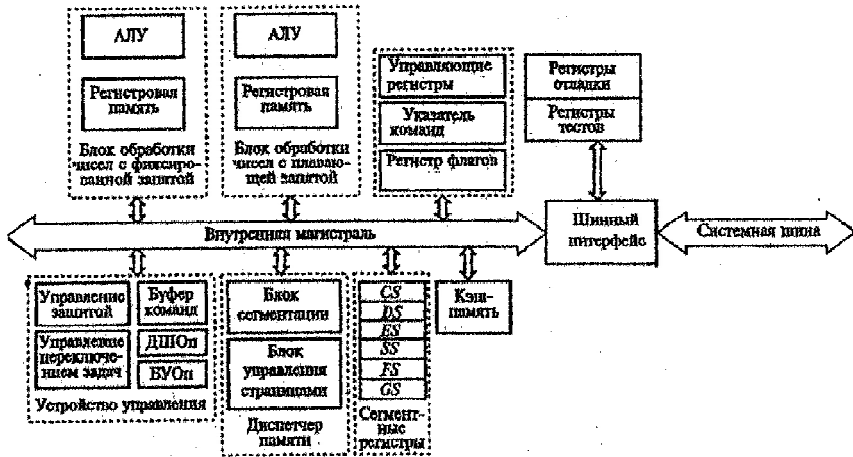


Рис. 5.8

Блоки сегментации и управления страницами позволяют управлять логическим адресным пространством, обеспечивая перемещаемость программ и данных и эффективное распределение памяти между задачами пользователей. Для этого память организована в виде одного или нескольких блоков – сегментов, каждый из которых, в свою очередь, разделяется на одну или несколько страниц.

Сегмент – это часть адресного пространства переменного размера, которая выделяется программе. Каждый сегмент имеет свой дескриптор (описатель), содержащий базовый адрес сегмента, размер сегмента, тип информации (стек, программа, данные) и поле прав доступа. Это поле определяет, как можно использовать информацию данного сегмента. Например, сегменты, содержащие команды, всегда защищаются по записи и могут быть также защищены по чтению; сегменты данных могут быть защищены по записи, но всегда открыты для чтения. Другие биты этого поля показывают, где размещается данный сегмент – в основной или во внешней памяти, и также разрешен или нет доступ к нему текущей программой.

В микропроцессоре имеется 32 регистра: регистры общего назначения, сегментные, указатель команд и флаги, управления, системных адресов, отладки, тестов. Эти регистры являются расширением регистров микропроцессора i8086.

В восьми 32-разрядных РОН хранятся адреса и данные, в шести сегментных регистрах – значения селекторов, указывающих адресуемые в данный момент сегменты памяти (один – текущий сегмент программы, один – текущий стековый сегмент, четыре – сегменты данных).

Указатель команд содержит смещение адреса следующей команды относительно базового адреса сегмента, где хранятся коды команд программы. Регистр флагов имеет то же назначение, что и в микропроцессоре i8086.

Внутренняя кэш-память включается между регистрами процессора и оперативной памятью; она хранит наиболее часто используемую информацию (команды и данные). В различных моделях микропроцессоров емкость внутренней кэш-памяти находится в диапазоне от 8 Кбайт до 128 Кбайт.

5.4. Классификация вычислительных систем

Первые вычислительные системы включали большие и дорогие ЭВМ, в связи с чем их количество в такой системе не превышало трех – четырех. Появление персональных компьютеров привело к построению распределенных систем управления, в которые могут быть вовлечены десятки и сотни компьютеров – персональных рабочих станций.

Эффективность использования вычислительных систем во многом определяется возможностями организации параллельных вычислений на уровне алгоритмов и программ.

Когда параллелизм реализуется внутри процессора (с помощью, например, конвейерного АЛУ или нескольких АЛУ), то такая вычислительная структура определяется как *векторный* или *суперскалярный* процессор.

Если параллелизм обеспечивается путем объединения нескольких ЭВМ или процессоров, то такая вычислительная структура определяется как *вычислительная система*.

Вычислительная система, в которой отдельные ЭВМ или процессоры объединяются вместе посредством специального коммуникационного оборудования (каналов связи), представляет собой *вычислительную сеть*.

Отличительной особенностью вычислительной системы по отношению к ЭВМ является наличие в них нескольких вычислителей (процессоров или ЭВМ), реализующих параллельную обработку. Создание ВС преследует следующие основные цели: повышение производительности системы за счет ускорения процессов обработки данных, повышение надежности и достоверности вычислений, предоставление пользователям дополнительных сервисных услуг и т.д.

В системах управления реального времени используются специализированные вычислительные системы, ориентированные на решение узкого класса задач. Специализация ВС может обеспечиваться различными средствами:

- во-первых, сама структура системы (количество параллельно работающих элементов, связи между ними и т.д.) может быть ориентирована на определенные виды обработки информации: матричные вычисления, решение алгебраических, дифференциальных и интегральных уравнений и т.п. Практика разработки таких ВС показала, что, чем уже класс задач, для решения которых предназначается специализированная ВС, тем большую производительность можно обеспечить;
- во-вторых, специализация ВС может закладываться включением в их состав специального оборудования и специального программного обеспечения.

Специализированные ВС по *виду составляющих элементов* принято подразделять на многомашинные и многопроцессорные. *Многомашинная* ВС (ММС) содержит несколько ЭВМ, каждая из которых имеет свою оперативную память и работает под управле-

нием своей операционной системы. Обмен между ЭВМ идет с помощью специальных дополнительных программных и аппаратурных средств. Многомашинные ВС относятся к классу *систем с распределенной памятью*.

Многопроцессорная (мультипроцессорная) система (МПС) в качестве общего ресурса имеет общую оперативную память (ООП). Все процессоры работают с единым адресным пространством. Параллельная работа процессоров и использование ООП обеспечивается единой операционной системой. Все процессоры «разделяют» общую память, поэтому многопроцессорные ВС относятся к классу *систем с разделяемой памятью*.

По *типам ЭВМ или процессоров*, из которых состоит ВС, различают однородные и неоднородные системы. *Однородные* системы предполагают комплексирование однотипных ЭВМ (процессоров), *неоднородные* – разнотипных. В однородных системах значительно упрощается разработка и обслуживание технических и программных средств. В них обеспечивается возможность стандартизации и унификации соединений и процедур взаимодействия элементов системы. Упрощается обслуживание систем, облегчается модернизация и их развитие. Вместе с тем существуют и неоднородные ВС, в которых комплекслируемые элементы очень сильно отличаются по своим техническим и функциональным характеристикам и могут представлять собой специализированные процессоры.

По *характеру пространственного распределения* вычислительных модулей ВС делятся на системы *сосредоточенного (локального)* и *распределенного* типов. Обычно такое деление касается только ММС.

В локальных и распределенных ММС сильно различается оперативность взаимодействия в зависимости от удаленности ЭВМ. Время передачи информации между соседними ЭВМ, соединенными простым кабелем, может быть много меньше времени передачи данных по каналам связи. Как правило, все выпускаемые в мире ЭВМ имеют средства прямого взаимодействия и средства подключения к сетям ЭВМ.

По *методам управления элементами ВС* различают *централизованные, децентрализованные и со смешанным управлением*. Помимо параллельных вычислений, производимых элементами системы, необходимо выделять ресурсы на обеспечение управления этими вычислениями. В *централизованных ВС* за это отвечает главная, или диспетчерская, ЭВМ (процессор). Ее задачей является распределение нагрузки между элементами, выделение ресурсов, контроль состояния ресурсов, координация взаимодействия. Структура ВС может быть одноуровневой или иерархической. Централизованный орган управления в системе может быть жестко фиксирован, или эти функции могут передаваться другой ЭВМ (процессору), что способствует повышению надежности системы. Централизованные системы имеют более простые операционные системы (ОС). В *децентрализованных системах* функции управления распределены между ее элементами. Каждая ЭВМ (процессор) системы сохраняет известную автономию, а необходимое взаимодействие между элементами устанавливается по специальным наборам сигналов. С развитием ВС и, в частности, сетей ЭВМ интерес к децентрализованным системам постоянно растет.

В системах со *смешанным управлением* совмещаются процедуры централизованного и децентрализованного управления. Перераспределение функций осуществляется в ходе вычислительного процесса, исходя из сложившейся ситуации.

По *принципу закрепления вычислительных функций* за отдельными ЭВМ (процессорами) различают системы с *жестким и плавающим закреплением функций*. В зависимости от типа ВС следует решать задачи статического или динамического размещения программных модулей и массивов данных, обеспечивая необходимую гибкость системы и надежность ее функционирования.

5.4.1. Многомашинные вычислительные системы

Исторически многомашинные вычислительные системы появились первыми. Уже при использовании ЭВМ первых поколений

возникали задачи повышения производительности, надежности и достоверности вычислений.

Многомашинный комплекс состоит из нескольких ЭВМ, каждая из которых имеет свою внутреннюю память и работает под управлением своей операционной системы, и средства обмена информацией между машинами. В отличие от них мультипроцессорные ВС имеют общую память и общие периферийные устройства; они работают под управлением единой операционной системы.

Машины работают параллельно; каждая из них обрабатывают собственные потоки заданий, а в системах управления решает собственное подмножество функциональных задач. Несколько ЭВМ только тогда представляют собой вычислительную систему, когда они могут обмениваться информацией. Для этого требуется их совместимость.

Аппаратурная совместимость означает, что:

- подключаемые друг к другу элементы (узлы, блоки) должны иметь единые стандартные, унифицированные средства соединения: кабели, число проводов в них, единое назначение проводов, разъемы, заглушки, адаптеры, платы и т.д.;

- параметры электрических сигналов, которыми обмениваются элементы (узлы, устройства), тоже должны соответствовать друг другу: амплитуды импульсов, полярность, длительность и т.д.;

- должна быть решена проблема синхронизации вычислений;

- должна быть согласована форма представления чисел и их разрядность.

Программная совместимость предполагает, что:

- существует программа-диспетчер, которая оперативно распределяет ресурсы ВС между задачами, обеспечивает загрузку программ и управляет вводом-выводом;

- программы, которыми могут обмениваться ЭВМ, входящие в состав ВС, должны одинаково интерпретироваться ими (должна быть согласована адресация, разрядности кодов команд коды операций и др.). Эта проблема решается достаточно просто для однотипных ЭВМ.

Программная совместимость может обеспечиваться на уровнях: 1) внутреннего языка; 2) языка ассемблера; 3) языка высокого уровня.

На первом уровне наиболее распространены такие типы организации совместимости, как «общая наименьшая машина», иерархического типа, частичная и с эмуляцией недостающих команд.

Если на ЭВМ1 можно выполнять множество команд А, а на ЭВМ2 – множество команд В и пересечение этих множеств С не пусто, то оно соответствует возможностям «общей наименьшей машины».

Если одно множество полностью входит в другое, то имеет место иерархическая совместимость. Например, если взаимодействующие ЭВМ относятся к одному и тому же семейству, но стыкуются разные модели, то в таких моделях совместимость обеспечивается «снизу вверх», т.е. ранее созданные программы могут выполняться на более поздних моделях, но не наоборот.

Частичная совместимость возникает тогда, когда только часть программ, составленных для одной ЭВМ, может выполняться на другой. Эмуляция отсутствующих команд служит для увеличения степени частичной совместимости: если выполнение программы на малой ЭВМ прерывается из-за отсутствия в ней команд, существующих в большой ЭВМ, то эти команды могут эмулироваться программными средствами малой ЭВМ.

Совместимость на уровне машинных команд имеет следующие недостатки:

- разработка и отладка программ для обеспечения совместимости на этом уровне требует больших затрат времени и сил, особенно, для управляющих ЭВМ;
- эта совместимость приводит к неполному использованию специального оборудования и специальных команд, которые имеются не во всех ЭВМ;
- для надежного обмена программами требуется централизация их контроля и исправления ошибок.

Наиболее высокая степень программной совместимости – на третьем уровне – уровне алгоритмических языков, когда обмен программами, по существу, заменяется обменом алгоритмами. Совместимость на втором уровне имеет промежуточный характер.

Информационная совместимость предполагает, что передаваемые информационные массивы одинаково интерпретируются стыкуемыми модулями ВС. Должны быть стандартизированы алфавиты, разрядность, форматы, структура и разметка файлов, томов и т.д.

Эффективность ММС во многом зависит от организации обмена информацией и командами между ними. Характеристики обмена определяются параметрами технических средств, обеспечивающих взаимодействие устройств, и уровня управления процессами, на котором это взаимодействие осуществляется. В общем случае возможны следующие уровни комплексирования:

- 1) прямого управления (процессор – процессор);
- 2) общей оперативной памяти;
- 3) комплекслируемых каналов ввода-вывода;
- 4) устройств управления внешними устройствами;
- 5) общих внешних устройств.

Чаще всего комплексирование ЭВМ в ММС осуществляется на первом, втором и третьем уровнях.

Уровень прямого управления служит для передачи коротких приказов-сообщений. Последовательность взаимодействия процессоров сводится к следующему. Процессор ЭВМ-инициатора обмена передает в блок прямого управления сообщение и подает команду «Прямая запись». У процессора другой ЭВМ эта команда вызывает прерывание, относящееся к классу внешних. В ответ он вырабатывает команду «Прямое чтение» и записывает передаваемый байт в свою память. Затем принятая информация расшифровывается и по ней принимается решение. После завершения передачи прерывания снимаются, и оба процессора продолжают вычисления по собственным программам. Очевидно, что уровень прямого управления не может использоваться для передачи больших массивов данных,

однако оперативное взаимодействие отдельными сигналами широко используется в управлении вычислениями.

Уровень комплексиремых каналов ввода-вывода предназначен для передачи больших объемов информации между блоками оперативной памяти сопрягаемых ЭВМ. Обмен данными между ЭВМ осуществляется с помощью адаптера «канал-канал» и команд «Чтение» и «Запись». Адаптер – это устройство, согласующее скорости работы сопрягаемых каналов. Обычно сопрягаются селекторные каналы машин как наиболее быстродействующие, но можно сопрягать мультиплексные каналы, а также селекторный и мультиплексный. Скорость обмена данными определяется скоростью самого медленного канала. Скорость передачи данных по этому уровню составляет несколько Мбайт/с.

Уровень управления внешними устройствами предполагает использование встроеного в устройство управления ВУ (УВУ) двухканального коммутатора и команд «Зарезервировать» и «Освободить». Коммутатор позволяет подключать УВУ одной машины к селекторным каналам различных ЭВМ. По команде «Зарезервировать» канал-инициатор обмена получает доступ через УВУ к любым накопителям на дисках или на магнитных лентах. Обмен канала с накопителями продолжается до полного завершения работ и получения команды «Освободить». Лишь после этого УВУ может подключиться к конкурирующему каналу. Такая дисциплина обслуживания требований позволяет избежать конфликтных ситуаций.

Пять уровней комплексирования получили название *логических*, так как они объединяют на каждом уровне разнотипную аппаратуру, имеющую сходные методы управления. Каждое из устройств может иметь логическое имя, используемое в прикладных программах. Этим достигается независимость программ пользователей от конкретной физической конфигурации системы. Связь логической структуры программы и конкретной физической структуры ВС обеспечивается операционной системой по указаниям-директивам пользователя, при генерации ОС и по указаниям дис-

петчера-оператора вычислительного центра. Различные уровни комплексирования позволяют создавать самые различные структуры ВС.

Если в составе ВС имеется дорогое уникальное внешнее устройство, то возможно совместное его использование всеми машинами вычислительного комплекса – это еще один уровень комплексирования.

Многомашинные ВС применяются в системах управления сложными объектами. В таких системах с помощью ММС решают две задачи: обеспечение требуемой производительности и надежности. Применение ММС для обеспечения высоких требований по надежности функционирования и по достоверности информации обобщается по следующей схеме.

Машины решают одну и ту же задачу, одновременно выполняя одинаковые операции над одними и теми же данными и периодически сверяя результаты. Если они совпадают, работа ММС продолжается. В случае несовпадения с помощью тестов или по результатам контроля определяется неисправная ЭВМ, ее выходы блокируются и работу системы обеспечивает исправная машина. После ремонта ММС вновь функционирует в нормальном режиме. Для этого в отремонтированную ЭВМ должна быть загружена обновленная информация и вновь обеспечен синхронный режим совместной работы.

Структурная организация ММС сильно зависит от конкретного применения, но можно указать основные концепции их построения:

- многоуровневая иерархическая организация структуры, соответствующая иерархическому характеру комплекса решаемых функциональных задач;
- динамическое перераспределение задач по ресурсам системы;
- модульность структуры и конструкции элементов ВС, обеспечиваемая унифицированной системой модулей и протоколов обмена;
- наличие развитых средств контроля и диагностики.

5.4.2. Кластеры

Кластером называют группу взаимосвязанных совместно работающих ЭВМ, представляющих для пользователя единый ресурс, единый центр по обработке информации. В качестве вычислительного узла обычно используются двух- или четырехпроцессорные ЭВМ, хотя могут устанавливаться и однопроцессорные ЭВМ, и локальные сети. Каждый узел работает под управлением своей операционной системы. Состав и мощность узлов могут меняться даже в рамках одного кластера, что дает возможность создавать неоднородные системы. Возможно включение в конфигурацию кластера специализированного компьютера, например, файл-сервера. Поскольку все ЭВМ кластера могут работать самостоятельно и в качестве узлов кластера могут выступать вычислительные системы, то кластер фактически представляет собой вычислительную систему более высокого уровня иерархии.

Концепция кластера появилась как следствие двух факторов: 1) стремления повысить производительность вычислительного ресурса, представляемого пользователю и 2) потребностью в высоконадежных средствах обработки информации.

Достоинства кластерных структур определяются следующим образом.

1. Можно создавать кластеры любых размеров, превышающие по суммарной производительности любые самые мощные ЭВМ. Кластер может состоять из десятков и сотен ЭВМ, каждая из которых может быть многопроцессорной вычислительной системой. Это свойство наращивать мощность (*масштабируемость* кластера) обеспечивается аппаратурной, программной и информационной совместимостью его элементов и автоматизацией процесса распределения нагрузки.

2. Пользователь может наращивать технические средства постепенно, и при этом вычислительная мощность кластера будет возрастать линейно, пропорционально приращению характеристик технических средств. Для других типов ВС подобная зависимость

имеет нелинейный характер вследствие роста числа конфликтов и усложнения системы обмена.

Возможны два решения проблемы масштабируемости. *Вертикальное* масштабирование осуществляется добавлением ЭВМ, модулей памяти, процессоров и средств обмена в каждый узел кластера. *Горизонтальное* масштабирование состоит в увеличении числа узлов и в добавлении внешней памяти.

3. Кластер обладает повышенной надежностью и готовностью. Это свойство обеспечивается способностью каждого узла кластера работать автономно, но в любой момент переключаться на выполнение работ другого узла в случае его отказа.

4. Поскольку в качестве «строительных блоков» кластера используются широко распространенные ЭВМ, его стоимость, как правило, меньше стоимости одиночной уникальной ЭВМ с характеристиками, аналогичными обобщенным характеристикам кластера.

Простейший способ обеспечить свойства масштабируемости, надежности и готовности – построить кластер, состоящий из двух узлов, соединенных высокоскоростной линией связи. Эта связь может представлять собой локальную вычислительную сеть, которая используется либо совместно с другими ЭВМ, либо специально для нужд кластера. Дальнейшее развитие этой схемы заключается в дополнительной связи между ЭВМ через дисковую подсистему. Использование избыточной внешней памяти типа RAID (*Redundant Array of Inexpensive Disk* – избыточный массив недорогих дисков) типично для кластеров, оно обеспечивает их высокую надежность.

Для RAID-массивов было разработано шесть разных конфигураций, отвечающих различным задачам. Они названы уровнями RAID, хотя никакой иерархии не представляют.

RAID 0 – это базовая конфигурация дискового массива, предназначенная для повышения производительности системы. Один большой файл разбивается на несколько частей, которые записываются на разные диски. Это называется *расслоением данных*. При обращении к такому файлу для чтения диски могут передавать дан-

ные параллельно, так что общее время его пересылки сокращается по сравнению со временем пересылки при хранении на одном диске.

Архитектура RAID 1 позволяет повысить надежность хранения данных путем записи их идентичных копий на двух дисках. Такие диски называют *зеркальными*. Если один из них выходит из строя, операции чтения и записи продолжают с зеркальным диском.

Уровни RAID 2, RAID 3 и RAID 4 предназначены для повышения надежности системы с помощью различных схем контроля четности, не требующих полного дублирования дисков. Вся информация, предназначенная для контроля четности, хранится на диске. В RAID 5 она распределяется между всеми дисками.

Архитектура кластеров используется при построении серверов для распределенных вычислительных систем.

В распределенных ВС появляются распределенные базы данных. Распределенная база данных (РБД) – это база, отдельные части которой размещены (возможно, с дублированием) на нескольких ЭВМ, объединенных сетью. Она должна удовлетворять следующим требованиям:

- быстрая обработка типовых и незапланированных запросов;
- обеспечение безопасности данных;
- обеспечение логической и физической независимости данных;
- обеспечение «прозрачности»:

- распределенной структуры данных (независимость конечных пользователей и программ от варианта размещения информации на ЭВМ сети);

- совместного доступа к данным (поддержка целостности РБД при одновременной модификации одних и тех же данных несколькими пользователями);

- распределенной обработки (независимость конечных пользователей и прикладных программ от типа сети и применяемой сетевой операционной системы).

Администрирование и доступ пользователей к РБД осуществляются с помощью системы управления РБД, которая, исходя из

упомянутых требований, обеспечивает выполнение следующих функций:

- автоматическое определение ЭВМ, на которой хранятся требуемые в запросе данные;
- декомпозицию распределенных запросов на частные подзапросы к БД отдельных ЭВМ;
- планирование обработки запросов;
- передачу частных подзапросов и их исполнение на удаленных ЭВМ;
- прием результатов выполнения частных подзапросов, полученных в результате декомпозиции запросов на поиск и чтение данных из РБД, и композицию общего результата;
- поддержание в согласованном состоянии копий дублированных данных на различных ЭВМ сети;
- управление параллельным доступом к РБД многих пользователей;
- обеспечение целостности РБД.

Реализация перечисленных выше функций в распределенной ВС поручается **серверам**. Вообще под термином «сервер» принято понимать объект, предоставляющий сервис другим объектам по их запросам; под *сервисом* понимается процесс обслуживания объектов, а наиболее распространенными видами сервиса являются такие как: хранение и поиск данных; передача сообщений и блоков данных; электронная почта; организация и управление диалогом пользователей; предоставление соединений; проведение сеансов взаимодействия прикладных процессов.

Существует множество типов серверов:

- сервер баз данных, выполняющий обработку запросов, направляемых базе данных;
- файловый сервер, обеспечивающий функционирование распределенных ресурсов, включая файлы и программное обеспечение;
- сервер доступа, обеспечивающий возможность клиенту одной локальной сети связываться с другой сетью;

- коммуникационный сервер, управляющий терминалами и осуществляющий маршрутизацию сообщений.

Распределенные ВС можно строить по принципам, получившим названия «файл – сервер» и «клиент – сервер».

В ВС, построенной по принципу «**файл–сервер**», файловый сервер не принимает участия в реализации прикладной программы пользователя. Он просто хранит файлы, хотя в ряде случаев служит концентратором для совместного использования периферийных устройств, организуя очереди к периферийным устройствам. Так как файловый сервер является независимым компьютером сети, то его обычно специализируют, устанавливая, например, большой объем дисковой памяти. Функционирующая на рабочей станции прикладная программа считывает и записывает файлы, обмениваясь ими с сетевым файловым сервером, но все операции выполняет самостоятельно.

Во многих случаях файлы передаются по сети между рабочей станцией и сервером целиком, хотя пользователю нужна только их часть. Это приводит к неэффективной загрузке сети. Другой недостаток систем типа «файл – сервер»: в них невозможен одновременный доступ к одному набору данных многими пользователями.

Присущие локальной сети проблемы породили идею кооперативных вычислений по принципу «**клиент–сервер**». *Клиент* в такой системе – это процесс, который с помощью определенных операций вызывает сервисную функцию. Им может быть программа или пользователь. Сервисная функция в архитектуре «клиент–сервер» описывается комплексом прикладных программ, в соответствии с которым выполняются прикладные процессы. Функции сервера реализуются в виде самостоятельного прикладного процесса либо части общего прикладного процесса. После выполнения каждого задания сервер посылает результаты клиенту, пославшему это задание.

Функции сети связи известны. В этой системе нагрузка на нее значительно ниже, чем в системе с файловым сервером, поскольку обмен идет не целыми массивами данных, а отдельными фрагментами, например, строками таблицы.

Достоинства системы организации вычислений типа «клиент–сервер» определяются, во-первых, тем, что клиентская и серверная часть могут работать на разных по стоимости и по производительности машинах, что приводит к снижению затрат на систему. Во-вторых, такие системы обладают хорошей адаптацией и гибкостью: замена сервера не приводит к нарушению клиентских приложений, а расширение числа пользователей (в определенных пределах) не влияет на работу сервера. В-третьих, каждый функциональный элемент, в том числе и клиентское приложение, легко может быть специализировано для выполнения тех или иных специфических функций.

По общему признанию, вычисления по схеме «клиент–сервер» более дешевые по сравнению с централизованными системами с большими ЭВМ.

В то же время нельзя забывать о том, что отказы сервера делают всю систему неработоспособной, а применение этой схемы для системы, обрабатывающей сотни и тысячи одинаковых приложений (типа заказов билетов), сводят на нет все ее достоинства.

5.4.3. Многопроцессорные вычислительные системы

Многопроцессорная система (МПС) – это совокупность процессорных элементов, которые взаимодействуют в процессе решения одной вычислительной задачи под управлением общей операционной системы. Известно много различных структур МПС и несколько принципов их классификации. Наибольшее распространение получила классификация М.Флинна (M.Flynn), построенная на концепции потока.

Если анализировать ВС с точки зрения сочетания потоков данных и потоков команд (понимая под *потоком команд* последовательность команд одной программы, а под *потоком данных* – последовательность данных, обрабатываемых одной программой), то возможны следующие варианты сочетаний: 1) одиночный поток команд – одиночный поток данных (ОКОД, или SISD – *Single*

Instuction Stream/ Single Data Stream); 2) множественный поток команд – одиночный поток данных (МКОД, или MISD – *Multiple Instuction Stream/Single Data Stream*); 3) одиночный поток команд – множественный поток данных (ОКМД, или SIMD – *Single Instuction Stream/ Multiple Data Stream*); 4) множественный поток команд – множественный поток данных (МКМД, или MIMD – *Multiple Instuction Stream / Multiple Data Stream*).

Архитектура ОКОД (рис. 5.9, а) не представляет интереса для анализа как вычислительная система: это последовательная однопроцессорная скалярная или суперскалярная структура, возможно, с конвейерной обработкой. На одном процессоре выполняется один поток команд; операции выполняются над данными, которые хранятся в одной области памяти.

К **архитектуре МКОД** (рис. 5.9, б) относят многопроцессорные конвейерные структуры.

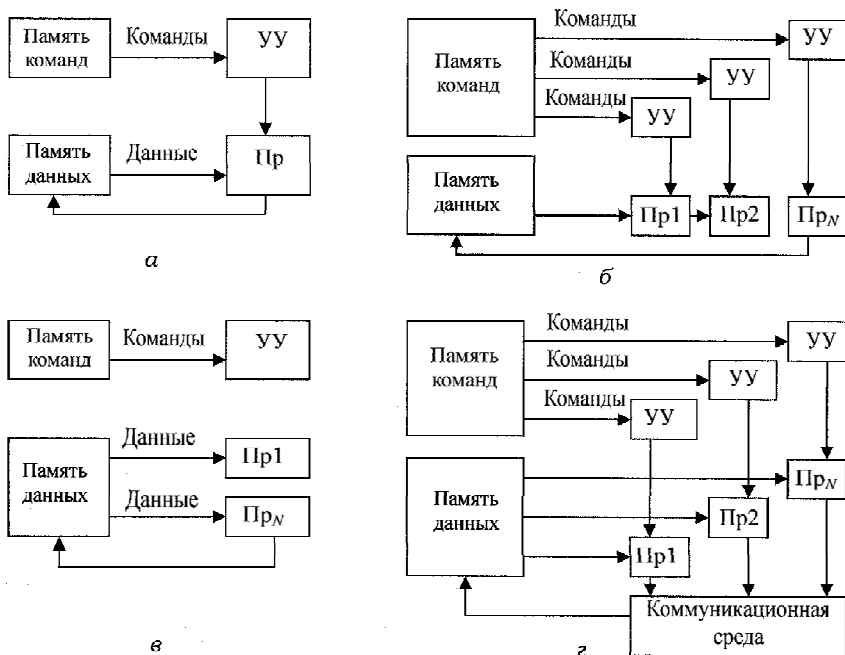


Рис. 5.9

Последовательность данных Д передается набору процессоров Пр1, Пр2,...,Прn, каждый из которых выполняет свою последовательность команд. Примером может служить устройство для сложения двух векторов, компоненты которых – числа с плавающей точкой – последовательно поступают на вход Пр1. Процессоры Пр2, ..., Прn выполняют операции сравнения порядков, выравнивания, сложения мантисс, нормализации в режиме конвейера (параллельно) над всеми компонентами вектора. Следует заметить, что единого мнения по классу MISD нет: есть утверждения, что подобные структуры на практике еще не были реализованы; возможно, этот класс пуст.

Архитектура ОКМД – многопроцессорная матричная структура (рис. 5.9, в). Массив процессоров обрабатывает поток команд, поступающих от центрального управляющего устройства. Каждая команда выполняется всеми процессорами одновременно над разными потоками данных.

Примером применения матричной ВС может служить выполнение операций типа вычисления определенного интеграла:

$$\int_a^b y(x)dx \approx h[(y_0 / 2) + y_1 + y_2 + \dots + y_{n-2} + (y_{n-1} / 2)],$$

где $y_i = y(x_i)$, $x_0 = a$, $x_n = b$, $h = (b-a)/n$. Каждый процессор выполняет одну и ту же операцию $hy(x_i)$.

Главное отличие матричной ВС от векторного процессора заключается в том, что во втором случае высокая производительность обеспечивается с помощью интенсивной конвейеризации, а в первом – за счет высокой степени параллелизма.

Вычислительная система с **архитектурой типа МКМД** (рис. 5.9, г) выполняет одновременно и независимо друг от друга несколько программных ветвей с различными наборами данных, обеспечивая в определенные моменты времени обмен данными. В схеме МКМД процессоры должны быть универсальными, поскольку они должны выполнять все команды, необходимые для соответствующего преобразования данных.

В классе МКМД различаются слабо- и сильносвязанные системы.

Слабосвязанные МКМД-системы – структуры с распределенной памятью. Десять-пятнадцать лет тому назад этот класс ВС состоял исключительно из многомашинных вычислительных комплексов и сетей ЭВМ. Вычислительные системы с такой структурой в качестве узлов используют законченные ЭВМ, включающие микропроцессор, память и подсистему ввода-вывода. Они объединяются коммуникационной средой, которая обеспечивает взаимодействие процессоров посредством простых операций ввода/вывода (рис. 5.10, а). Каждый процессор Pr_1, \dots, Pr_N имеет свой модуль памяти и устройство ввода-вывода. Обмен идет через коммуникационную среду.

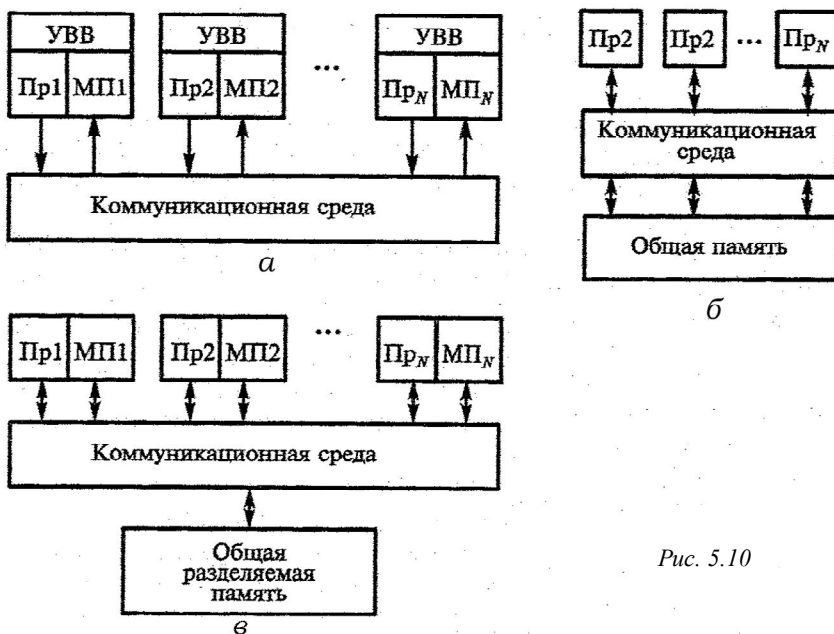


Рис. 5.10

Сейчас в качестве элементов МКМД-систем используются микропроцессоры. Пример современной слабосвязанной МКМД-системы – Cray T3D/T3E. В нем память физически распределена,

так как каждый процессор содержит свою локальную память. В то же время память разделяется всеми процессорами, поскольку каждый из них через свой сетевой интерфейс может обращаться к памяти любого другого процессора, не прерывая его работы. Коммуникации осуществляются с помощью средств операционной системы. Передача данных из одного локального адресного пространства к другому произойдет только в том случае, если посылка сообщения со стороны процесса-отправителя будет востребована процессом-получателем сообщения. В последнее время успехи микроэлектроники сделали возможным построение слабосвязанных систем, состоящих из сотен и тысяч процессорных элементов с возможностью наращивания их количества. Подобные структуры называют **системами с массовым параллелизмом** (MPP-системы – *Mass-Parallel Processing*). Увеличение количества процессорных элементов в ВС не может не привести к росту сложности коммуникационной системы, а также к новым идеям в области управления вычислительным процессом: от управления потоком команд можно перейти к управлению от потока данных, когда операторы программы активизируются по мере готовности данных.

У вычислительных систем с распределенной памятью соотношение цена/производительность ниже, чем у ВС других классов. Для пользователя еще одно достоинство ВС этого класса заключается в возможности точно подобрать конфигурацию в зависимости от бюджета и своих требований к характеристикам вычислительных средств.

Сильносвязанные МКМД-системы – структуры с разделяемой общей памятью. Процессоры совместно используют общую память и каждый из них имеет равные права доступа к программам и данным, которые там хранятся; процессоры могут обмениваться информацией через единое адресное пространство памяти. Такие ВС называются *мультипроцессорными системами с общей памятью*. Существует два варианта организации общей памяти:

1) все процессоры связываются с модулями общей памяти через коммуникационную среду; время обращения к памяти у всех про-

цессоров одинаково и не зависит от того, к какому модулю памяти процессор обращается (см. рис. 5.10, б); такие МПС называют *системами с однородным доступом*;

2) каждый процессор имеет доступ к общей разделяемой памяти, но, кроме того, он может обращаться и к собственной локальной памяти (см. рис. 5.10, в); системы этого типа называют *системами с неоднородным доступом*. Доступ к локальной памяти идет намного быстрее, чем к общей, что повышает производительность ВС.

В классе сильносвязанных МКМД-систем выделяют два типа архитектуры: асимметричную и симметричную.

В асимметричной архитектуре разные процессоры могут отличаться как своими характеристиками (производительностью, надежностью, системой команд и т.д.), так и функциональной ролью, которая поручается им в системе. Например, одни процессоры могут предназначаться для работы в качестве вычислительных модулей, другие – для управления вводом/выводом.

Симметричная архитектура предполагает однородность всех процессоров и единообразие их включения в общую схему МВС.

В качестве проблем, которые характерны для параллельных ВС с общей памятью указываются:

- высокая стоимость;
- неоднородность доступа к памяти: в то время как один процессор обращается к локальной памяти, другой может обращаться к удаленной – время доступа различно и это должно учитываться при организации общего вычислительного процесса;
- необходимость обеспечения согласованности содержимого кэш-памяти.

Сравнение многопроцессорных и многомашинных вычислительных систем приводит к следующим выводам.

Многопроцессорные ВС обеспечивают более быстрый обмен между процессорами, более высокую надежность и живучесть по сравнению с ММС.

Однако МПС имеют и существенные недостатки. Они, в первую очередь, связаны с использованием ресурсов общей оператив-

ной памяти. При большом количестве комплекслируемых процессоров возможно возникновение конфликтных ситуаций, когда несколько процессоров одновременно обращаются с операциями типа «чтение» и «запись» к одним и тем же областям памяти. Помимо процессоров к общей памяти, подключаются все каналы (процессоры ввода/вывода), средства измерения времени и т.д. Поэтому вторым серьезным недостатком МПС является проблема коммутации абонентов и доступа их к общей памяти.

Создание коммутаторов оперативной памяти представляет сложную техническую задачу: для разрешения конфликтных ситуаций необходимы схемы приоритетного обслуживания; они должны быть дополнены буферами для организации очередей запросов. До настоящего времени в номенклатуре технических средств вычислительной техники отсутствуют высокоэффективные коммутаторы общей памяти.

От того, насколько удачно решаются эти проблемы, и зависит эффективность применения МПС. Это решение обеспечивается аппаратно-программными средствами. Процедуры взаимодействия процессоров очень сильно усложняют структуру операционной системы МПС. Накопленный опыт построения подобных систем показал, что они эффективны при небольшом числе комплекслируемых процессоров (от 2 до 10). В отечественных системах «Эльбрус» обеспечивалась возможность работы до 10 процессоров, до 32 модулей оперативной памяти, до четырех процессоров ввода-вывода и до восьми коммутаторов оперативной памяти.

Еще одна сложная проблема – эффективное распараллеливание вычислительного процесса или программы.

5.5. Основные характеристики средств переработки информации

Возможности и эффективность применения средств переработки информации определяются следующими характеристиками.

Быстродействие – количество элементарных операций, выполняемых в единицу времени. В качестве таких операций рассматривают короткие (требующие минимального времени) операции типа сложения. Единица измерения – миллион операций в секунду в формате с фиксированной точкой (*Million Instactions Per Second – MIPS*) или с плавающей точкой (*Millions Floating Point Operation per Second – MFLOPS*). Быстродействие можно оценить при обработке бесконечной последовательности не связанных и не конфликтующих между собой при доступе к памяти команд. Эта характеристика имеет ограниченное распространение, поскольку не дает полного представления о возможностях ЭВМ, так как:

- современные ЭВМ могут выполнять сотни различных операций, длительности которых существенно различаются;

- во многих применениях для пользователя ЭВМ более важно время решения задачи, которое часто больше зависит от алгоритма и структурных особенностей ЭВМ, чем от ее быстродействия.

Производительность – объем вычислительной работы, выполняемой ЭВМ в единицу времени. Оценка и сопоставление производительности ЭВМ – сложная задача, не имеющая до настоящего времени удовлетворительного решения. Причина в том, что на производительность влияет слишком много факторов: характеристики элементов, особенности программы, логические возможности системы команд, особенности операционной системы, состав и характеристики оперативной и внешней памяти, способ обмена данными по шине и др. Поэтому более или менее достоверные оценки производительности можно получить только для тех ЭВМ, которые решают ограниченный круг задач.

Известны следующие типы характеристик производительности:

- пиковая*, или *техническая* – производительность процессора без учета времени обращения к основной памяти за операндами. Это теоретический максимум быстродействия при идеальных условиях;

- номинальная* – производительность процессора с оперативной памятью;

системная – производительность базовых технических и программных средств, входящих в комплект поставки ЭВМ;

эксплуатационная – производительность на реальной рабочей нагрузке, формируемой используемыми пакетами прикладных программ.

Методы оценки производительности разделяются на три группы:

расчетные, основанные на информации, получаемой теоретическим или расчетным путем;

экспериментальные, основанные на информации, получаемой с использованием аппаратно-программных измерительных средств;

имитационные, реализуемые с помощью моделирования.

Основные единицы оценки производительности:

абсолютная, определяемая количеством элементарных работ, выполняемых в единицу времени;

относительная, определяемая для оцениваемой ЭВМ относительно базовой в абсолютных единицах в виде индекса производительности.

Пиковая производительность определяется как среднее число команд типа передачи из регистра в регистр, выполняемых в единицу времени без учета их статистического веса в выбранном классе задач.

Номинальная, так же как и системная производительность определяется средним числом команд с учетом их статистического веса. В настоящее время производительность ЭВМ, предназначенных для решения широкого класса задач, оценивается с помощью специально сконструированных тестовых программ. Они формируются из смесей команд, в которых частота появления каждой операции определяется путем статистического анализа программ решения задач определенного (и достаточно широкого) класса, например, научно-технических задач или задач информационного поиска. На основании такого анализа операциям присваиваются определенные весовые коэффициенты.

При использовании смесей команд производительность определяется по формуле

$$P = \sum_{i=1}^k a_i / \sum_{i=1}^k a_i t_i,$$

где a_i и t_i – соответственно весовой коэффициент и длительность выполнения i -ой команды; k – число различных команд в смеси.

В качестве упрощенной оценки производительности можно использовать время выполнения программы

$$T = NS/R,$$

где N – количество команд, которые будут реально выполнены в ходе реализации программы (в общем случае оно не совпадает с количеством команд в тексте программы); S – число шагов, необходимых для выполнения одной усредненной команды; R – частота, на которой работает процессор (предполагается, что длительность одного шага совпадает с длительностью одного такта процессора).

Тестовые программы для оценки производительности, существующие в настоящее время, можно разделить на три группы:

- 1) наборы тестов фирм-изготовителей для оценки качества собственных изделий;
- 2) стандартные универсальные тесты для ЭВМ, предназначенных для сложных вычислений;
- 3) специализированные тесты для конкретных областей применения.

Результаты оценивания производительности ЭВМ по разным тестам несопоставимы.

Емкость запоминающих устройств – количество структурных единиц информации, которое можно одновременно хранить в памяти. Обычно эти данные указываются отдельно для оперативной и внешней памяти. Наименьшая структурная единица – бит (один разряд, или одна двоичная цифра).

Как правило, емкость указывают в байтах (8 бит) или в более крупных единицах:

- килобайтах (1 Кбайт = 2^{10} байт = 1024 байта),
- мегабайтах (1 Мбайт = 2^{10} Кбайт),
- гигабайтах (1 Гбайт = 2^{10} Мбайт).

Емкость памяти в значительной степени определяет класс задач, которые может решать ЭВМ; большое значение для эффективности работы имеет распределение общего объема памяти между внешней и оперативной: поскольку все программы и все данные невозможно хранить одновременно в оперативной памяти, для нормального хода вычислительного процесса приходится вести обмен между ней и внешней памятью и чем меньше объем ОП, тем чаще приходится выполнять такой обмен.

Надежность – способность ЭВМ при определенных условиях выполнять требуемые функции в течение заданного времени (стандарт ISO-2382/14-78). Надежность ухудшается из-за появления неисправностей в аппаратуре. Моменты их возникновения случайны и их поток принято характеризовать средним временем между появлением неисправностей, а в качестве неисправностей рассматривают такие, которые для своего устранения требуют ремонта оборудования. Их называют *отказами*. Величину λ , обратную среднему времени между отказами, называют *интенсивностью отказов*; это основная характеристика надежности. Ее смысл – среднее количество отказов в единицу времени, размерность – $[ч^{-1}]$.

Если предположить, что отказы элементов, составляющих устройство ЭВМ, взаимонезависимы и отказ любого элемента приводит к отказу всего устройства, то интенсивность его отказов будет равна

$$\Lambda = \sum_{i=1}^m \lambda_i c_i ,$$

где λ_i – интенсивность отказов элементов i -го типа, c_i – количество элементов i -го типа, входящих в состав устройства.

На основе интенсивности отказов вычисляются другие характеристики надежности, в частности – вероятность исправной работы в течение заданного времени t . Если допустить, что интенсивность отказов не изменяется во времени (что имеет место на определенных этапах эксплуатации техники), то эта вероятность может быть описана экспоненциальным законом: $p(t) = \exp(-\lambda t)$. Соответст-

венно вероятность возникновения отказов в течение заданного времени $q(t) = 1 - \exp(-\lambda t)$.

Совершенствование технологии производства, соблюдение и облегчение условий работы аппаратуры приводят к уменьшению интенсивности отказов элементов и, следовательно, к увеличению надежности устройств ЭВМ.

Следует иметь в виду, что проблема надежности ЭВМ не исчерпывается задачей оценки ее характеристик. Во-первых, выход из строя некоторых деталей может не приводить к отказу устройства (например, сгоревшая лампа индикатора работы дисководов не исключает возможности его использования). Во-вторых, отказ устройства в течение некоторого времени может не приводить к искажению результатов его работы (если, например, к отказавшему элементу не было обращений). В-третьих, своевременное обнаружение факта и места возникновения неисправности при определенных условиях дает возможность изменить структуру устройства или ЭВМ, исключив тем самым влияние отказавшего блока на результат решения.

Для описания этих особенностей и возможностей при системном проектировании используются дополнительные характеристики, такие, как например, достоверность информации и живучесть системы.

Разрядность – длина машинного слова. Она определяется количеством бит информации, которые можно одновременно обрабатывать, и может быть различной для АЛУ, памяти, внешних устройств и при передачах между устройствами. Обычно под разрядностью ЭВМ понимают разрядность АЛУ.

Разрядность определяет *точность* вычислений – возможность различать почти равные значения (стандарт ISO-2382/2-76). Поскольку требования к точности могут варьироваться в широких пределах в зависимости от области использования ЭВМ, то выбор требуемой разрядности представляет сложную системную задачу. В некоторых случаях применяют специальные программные методы управления разрядностью.

Стоимость – затраты на разработку, изготовление, установку и эксплуатацию ЭВМ. Обычно пользователь имеет дело не со стоимостью, а с ценой, которую устанавливает производитель или торгующая и обслуживающая организация. Выбор ЭВМ представляет собой многоальтернативную задачу, где определяющим критерием является отношение «качество/цена», причем при оценке качества следует учитывать все характеристики ЭВМ, важные для предполагаемой области ее применения, а при оценке затрат – все этапы эксплуатации, принимая во внимание затраты на обслуживание и возможные ремонты.

Для производителя оценка затрат имеет еще более сложный характер. Во-первых, они зависят от структуры ЭВМ и варианта распределения функций между аппаратурой и программами и, следовательно, влияют на принимаемые технические решения. Во-вторых, приходится учитывать и оценивать целесообразность изготовления отдельных комплектующих элементов и узлов своими силами или их приобретения у другого изготовителя. В-третьих, от стоимости зависит конкурентоспособность продукции. Можно привести еще много подобных аргументов; все они приводят к однозначному выводу: стоимость является определяющей характеристикой как при выборе ЭВМ для конкретного применения, так и при выборе технических решений в процессе создания ЭВМ.

К сожалению, в настоящее время не существует достоверных оценок функциональных зависимостей стоимости от других характеристик и параметров ЭВМ, поэтому задача выбора рациональных технических решений по критерию стоимости дает скорее качественные, нежели количественные результаты.

Контрольные вопросы и упражнения

1. Что такое «микроконтроллер»? Из чего состоит его ядро? функциональный блок?
2. Как работают процессоры событий?
3. Как действуют аппаратные средства режима отладки
4. Каким образом осуществляется специализация микроконтроллеров?

5. Как действует режим широтно-импульсной модуляции в микроконтроллере с интегрированными процессорами событий?
6. Что такое «сегмент»? Какова структура дескриптора сегмента?
7. Расскажите о классификации вычислительных систем.
8. Как реализуется аппаратная совместимость, программная совместимость?
9. Чем кластер отличается от многомашинной системы?
10. В чем заключаются достоинства кластерных структур?
11. Какие функции обеспечивает система управления распределенной базой данных в кластерных структурах?
12. Чем принцип «файл–сервер» отличается от принципа «клиент–сервер»?
13. В чем заключается основной недостаток сильносвязанных МКМД-систем?
14. Какие виды производительности ЭВМ вам известны?

6. СРЕДСТВА ОТОБРАЖЕНИЯ И УПРАВЛЕНИЯ

6.1. Средства отображения информации

6.1.1. Классификация средств отображения

С помощью средств отображения информации (СОИ) человек-оператор получает осведомительную информацию о состоянии объекта управления. Необходимая для рассмотрения общих инженерно-психологических требований классификация СОИ приводится на рис. 6.1. По *функции* выдаваемой информации СОИ делятся на целевые (командные) и ситуационные. Первые отображают цель управления и дают сведения о необходимых действиях (например, командные табло «Стойте», «Идите»); ситуационные индикаторы обрисовывают ситуацию как таковую (датчик температуры).



Рис. 6.1

По *способу использования* показаний СОИ разделяются на следующие группы:

– для контрольного (проверочного) чтения. С помощью таких индикаторов оператор решает задачу типа «да–нет»: работает машина или нет, в норме или нет какие-либо параметры и т.п.;

– для качественного чтения. Такие индикаторы дают информацию о направлении изменения управляемого параметра: возрастает он или падает; если отклоняется, то в какую сторону, и т.д.;

– для количественного чтения. Такие индикаторы передают информацию в виде численных значений. К этой группе относится абсолютное большинство используемых приборов и индикаторов.

По *форме сигнала* различают абстрактные и изобразительные СОИ. В первом случае сигналы передаются в виде абстрактных символов (цифры, буквы, геометрические фигуры), отображающих в закодированном виде состояние объекта. Выбор символов определяется из соображений обеспечения оптимальной эффективности кодирования. Во втором случае передача сигналов осуществляется в форме изображений. В этом случае качество передачи определяется полнотой изображения – степенью схематизации, детализации и количеством воспроизводимых свойств.

По *степени детализации* информации СОИ могут быть интегральными и детальными. На интегральных индикаторах информация выдается оператору в обобщенном виде, чтобы сократить или вообще исключить время на ее выделение, поиск и синтезирование. При использовании интегральных методов следует помнить, что у оператора может возникнуть потребность в детализации обобщенной информации. Для этого должна быть предусмотрена возможность вызова информации на средства отображения информации по запросу оператора.

6.1.2. Виды средств отображения информации и их основные характеристики

Элементы систем отображения. Для отображения состояния отдельных объектов используются разнообразные элементы индикации: лампы накаливания, стрелочные измерительные приборы, оптико-механические проекционные приборы, газоразрядные и электролюминесцентные индикаторы, светодиоды и т.п.

Так, для отображения бинарных состояний системы, а также для опознания объекта, для предупреждения, предостережения удобно применять лампы накаливания или газоразрядные индика-

торы: эти элементы имеют достаточную яркость, четкость и надежны в работе.

Для отображения сообщения, словесной инструкции, последовательности операций наиболее подходят сигнальные оптические табло с различными трафаретами (транспаранты) или электролюминесцентные индикаторы.

Для отображения количественных показателей необходимы устройства типа счетчика, и лучше использовать цифровые индикаторы (вакуумные, газоразрядные, электролюминесцентные).

Для анализа работы оборудования рекомендуется применять устройства на основе электронно-лучевых трубок (ЭЛТ) и жидких кристаллов (*Liquid Crystal Display* – LCD-индикаторы), так как с их помощью можно показать связи между многими параметрами.

Стрелочные приборы преобразуют электрический сигнал в пропорциональный ему механический поворот стрелки. При этом

$$M_d = M_B = M_{дм} + M_{си},$$

где M_d – возбуждающий момент, действующий на вращающуюся часть системы и стрелку прибора, M_B – возвращающий момент, $M_{дм}$ – угловой демпфирующий момент, $M_{си}$ – момент сил инерции.

В установившемся режиме, когда стрелка покоится, $M_d = K\theta$ (K – постоянная пружины, θ – отклонение стрелки), т.е. действующий момент (мера входной величины) зависит от свойств пружины, примененной в приборе.

В люминесцентных дисплеях способность испускать видимый свет реализуется при возвращении электронов на исходный энергетический уровень. Возбуждение электронов может быть осуществлено разными способами: с помощью электрического разряда в газе, посредством ион-электронной рекомбинации или путем бомбардировки электронами высокой энергии.

Дисплеи, в которых реализуется разряд в газе, состоят из стеклянной трубки, наполненной газом, и нескольких электродов, имеющих форму символов, которые расположены близко один за другим. Эти электроды служат катодами трубки, в которой имеется также один анод. Напряжение зажигания – порядка 120 В, рабочее напряжение 80 В, рабочий ток 1 мА.

В светоизлучающих диодах (светодиодах) возбуждение электронов в атомах осуществляется посредством рекомбинации. Диоды собираются в матрицы, обычно размером 5×7 , напряжение питания 1,6...3 В, рабочий ток 150 мА.

Электрооптические дисплеи, в которых используется люминесценция за счет бомбардировки быстрыми электронами, – это ЭЛТ. Электронный луч формируется в ЭЛТ катодом, покрытым материалом с малой работой выхода. Когда катод нагревается нитью накала, самые быстрые электроны вылетают с его поверхности. Их подхватывает электрическое поле, и они с ускорением летят вдоль центральной оси трубки в сторону анода. На управляющей сетке, расположенной между катодом и анодом, поддерживается отрицательное напряжение; с ее помощью регулируется количество электронов, достигших анода (ток луча). Из-за наличия сил отталкивания между электродами сечение луча оказывается размытым; для того, чтобы собрать его в точку, применяется фокусирующая система. Отклонение электронного луча можно выполнять с помощью электрического и магнитного полей. В быстродействующих системах предпочтение отдается электрическому полю. Экран ЭЛТ покрыт слоем люминофора, атомы которого легко возбуждаются при бомбардировке электронами луча. По прошествии времени релаксации они испускают свет и в течение некоторого времени после окончания возбуждения продолжают светиться.

Базовый принцип LCD-технологии, основанной на жидких кристаллах, очень прост: используется панель, содержащая множество точек (пикселей) разного цвета, причем ее задняя стенка освещается смонтированными на ней лампами дневного света (с холодным катодом). Количество ламп – от 1 до 4; питание – от источника переменного тока напряжением 355...755 В частотой 45...60 кГц. Следует заметить, что для начального зажигания лампы необходимо напряжение в 2...3 раза больше рабочего.

Чтобы отдельный пиксел (его размер составляет 0,2...0,33 мм) начал светиться, необходимо открыть как бы маленькую «заслонку», препятствующую прохождению света через него. Жидкие кристаллы могут изменять свою молекулярную структуру, что позволяет регулировать проходящий через них световой поток. Для точ-

ного регулирования количества проходящего света и создания нужных световых оттенков обычно применяется многослойная конструкция, состоящая из двух поляризационных фильтров, цветных фильтров и двух ориентирующих слоев. Когда на такой ориентирующий слой подается напряжение, возникает электрическое поле, меняющее ориентацию жидких кристаллов, причем плоскость поляризации светового луча поворачивается на 90° при прохождении одного слоя. Для вывода цветного изображения используется три поляризационных фильтра – для красного, зеленого и синего цветов.

Лучших результатов с точки зрения стабильности, качества разрешения, яркости и контрастности изображения обеспечивают LCD-экраны с активной матрицей. В активной матрице каждый элемент имеет запоминающий транзистор, который хранит цифровую информацию о состоянии пикселя. При этом изображение сохраняется до тех пор, пока не поступит другой управляющий сигнал. Кроме того, в активной матрице для каждой ячейки экрана используются отдельные усилительные элементы, которые компенсируют влияние емкости ячеек и позволяют значительно уменьшить время изменения их прозрачности (до 50 мс). Запоминающие транзисторы, изготавливаемые из прозрачных материалов, позволяют световому потоку проходить сквозь них.

Надежность LCD-экрана (наработка на отказ) определяется как ресурсом лампы подсвета (от 10 до 50 тысяч часов), так и надежностью самой LCD-панели (30 тысяч часов).

Системы отображения информации. С помощью рассмотренных элементов создаются разнообразные системы отображения информации. Они могут быть выполнены в виде табло, мнемосхем, приборных панелей и щитов.

На *табло* данные отображаются в виде таблиц, информационное поле которых строится с учетом назначения табло, а также с учетом вида и объема отображаемой информации.

Мнемосхема представляет собой условное графическое изображение производственного процесса в виде комплекса символов, изображающих его отдельные элементы с их взаимными связями.

Основное достоинство мнемосхемы – высокая наглядность отображения информации.

Мнемосхемы строятся на базе плазменных панелей постоянного и переменного тока. Панель представляет собой два стекла, между которыми размещена керамическая решетка, образующая прямоугольную матрицу ячеек (рис. 6.2). В керамическую решетку вмонтированы вертикальные и горизонтальные электроды. По периметру панели расположен герметизирующий шов. Пространство между стеклами заполнено смесью инертных газов. При приложении к двум перпендикулярным электродам управляющего воздействия, в ячейке возникает плазменный разряд, излучающий ультрафиолет.

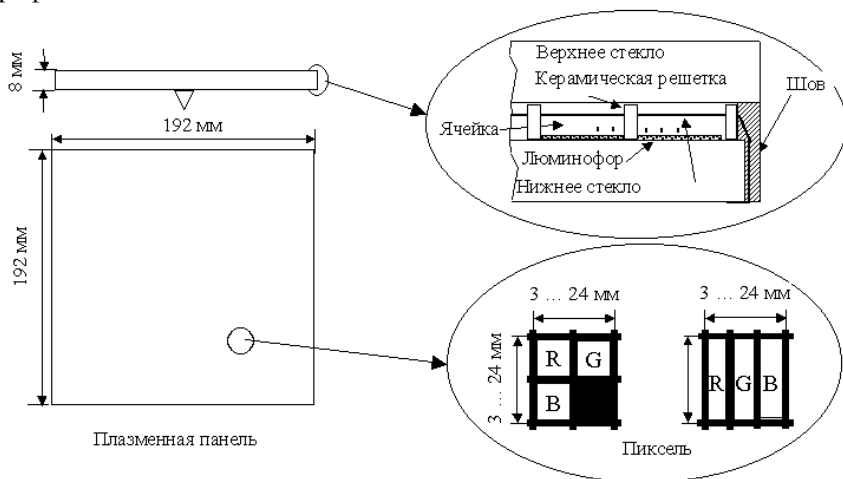


Рис. 6.2

Этот ультрафиолет возбуждает люминофор с красным, синим или зеленым цветом свечения. Люминофор обычно наносится на дно и боковые стенки ячейки, а иногда – тонким полупрозрачным слоем и на потолок ячейки. Верхняя система электродов делается прозрачной или достаточно узкой для того, чтобы не препятствовать выходу света. Три или четыре ячейки с разными цветами свечения образуют квадратный пиксель.

Конструкция панелей позволяет стыковать их друг с другом с минимальными нарушениями слитности изображения, создавая

при этом модульные экраны. Для того чтобы изображение на модульном экране смотрелось слитно, панели для модульных экранов должны обеспечивать стыковку друг с другом с минимальной потерей шага пикселей.

Двойной размер шва складывается из следующих компонентов: толщина барьера, толщина двух швов герметизации, технологический зазор между панелями. Сегодня размер шва составляет около 3 мм. В перспективе двойной размер шва может быть уменьшен до 1...1,5 мм. Используя оптическую маскировку швов, можно достичь практически их полного визуального исчезновения. Но все же наличие швов, проявляющихся на изображении в виде регулярной сетки, ограничивает минимальный размер пикселя на уровне 3...4 мм. При размере пикселя 3 мм экран (640×480 пикселей) будет иметь 95-дюймовую диагональ.

Основные *достоинства* плазменных экранов:

1) высокое качество изображения, достигаемое благодаря высокой яркости и контрастности, способности отображать большое число градаций яркости (до 256 по каждому цвету); отсутствие спада яркости, цветности и геометрических искажений по краям экрана;

2) модульность, позволяющая наращивать экран до любого размера с шагом около 20 см; относительно небольшой вес, позволяющий легко подвешивать и наклонять экран;

3) малая толщина экрана (около 10 см), позволяющая минимизировать пространство, необходимое для его обслуживания; обслуживание в процессе эксплуатации, как правило, не требуется, отсутствует необходимость юстировки;

4) отсутствие мерцания изображения и широкий угол обзора как по горизонтали, так и по вертикали (160° градусов и более).

Основные *недостатки* плазменных экранов:

1) относительно низкая разрешающая способность (шаг пикселя 3 мм);

2) высокое энергопотребление при низком КПД и значительном выделении тепла; «капризность» управляющих плат и блоков зажигания, которые часто перегорают;

3) деградация (эрозия) электродов и выгорание люминофора, особенно при отображении стационарной картинке; подвержен-

ность люминофора ионизирующим излучением; необходимость полной замены панелей после выработки ресурса (порядка десяти тысяч часов для панелей постоянного тока и нескольких десятков тысяч часов для панелей переменного тока).

Другой принцип построения мнемосхем и больших экранов из них – проекционный: он строится из проекционных кубов, каждый из которых представляет собой законченную проекционную систему. Разделение экрана на отдельные кубы вызвано стремлением уменьшить глубину экрана, которая для проекционных систем приблизительно равна размеру диагонали. В проекционных системах свет от лампы проходит через модулятор, формирующий изображение, и проецируется на экран. Известны несколько типов модуляторов.

Наиболее распространенным модулятором для проекторов является жидкокристаллическая матрица. Также появился целый ряд *микроэлектромеханических матриц*, основанных на электрически управляемом изменении угла наклона или формы микрозеркал. Такие системы имеют очень высокое быстродействие (время переключения микрозеркала около 20 нс). Размер каждого зеркала составляет около 16×16 микрон, зазор 1 микрон.

Разновидностью микроэлектромеханических матриц являются матрицы тонкопленочных микрозеркал. В них отклонение зеркал производится не электростатическим, а пьезоэлектрическим способом. Характерный размер такого микрозеркала – 97×97 микрон, зазор 3 микрона.

Для получения цветного изображения используется временное или пространственное совмещение цветов. В первом случае модулятор поочередно засвечивается источниками разных цветов (чаще всего через вращающийся диск-фильтр с секторами разных цветов: красного, зеленого и синего), во втором – используются три модулятора.

Основные *достоинства* проекционных экранов:

- 1) высокая разрешающая способность, яркость и контрастность, широкий угол обзора;
- 2) минимальный зазор между кубами (не более 1 мм) с возможностью сделать его практически незаметным за счет оптических эффектов;

3) независимость от характера изображения, что позволяет круглосуточно проецировать стационарную картинку.

Основные *недостатки* проекционных экранов:

1) относительно низкий ресурс ламп (6...10 тыс. ч); замена ламп обходится дешевле замены панелей в плазменных экранах, однако может потребовать дополнительной юстировки;

2) трудноразрешимые оптические погрешности, проявляющиеся в виде астигматизма (размытость изображения по краям), абберации (появления на изображении цветных колец); низкий коэффициент пропускания, требующий использования просветленной оптики; неравномерность яркости (спад яркости по краям); механические (вследствие недостаточной юстировки) и тепловые (вследствие теплового расширения модулятора) погрешности;

3) большая глубина экрана и необходимость организации сзади пространства для обслуживания;

4) высокая чувствительность к механическим воздействиям;

5) современные оптические системы весьма сложны и дороги.

Большинство СОО коллективного пользования имеют модульную наращиваемую конструкцию. Поэтому достижима любая конфигурация экрана, размеры которого уточняются на основании выбранного типа и модификации оборудования.

Приборная панель представляет собой совокупность отдельных приборов и индикаторов, каждый из которых несет информацию об одном или нескольких параметрах объекта. Приборные панели целесообразно применять в том случае, когда число контролируемых оператором параметров невелико и они слабо взаимосвязаны.

Щит управления представляет собой комплекс оборудования, предназначенного для мониторинга работы оборудования определенного функционального назначения (например, работа реактора, управление основными технологическими режимами, вентиляционными системами, радиометрическим контролем и т.п.).

В абстрактных СОО, представляющих оператору информацию в виде сигналов-символов используются три основные формы зрительной индикации: стрелочная, знаковая, графическая. Рассмотрим основные требования, предъявляемые к каждому из этих видов индикации.

Стрелочная индикация. Этот вид индикации является еще довольно распространенным. Большое влияние на считывание показаний оказывают отдельные элементы стрелочного прибора.

Установлено, что шкалы приборов, несущих наиболее важную информацию, должны иметь диаметр 120...130 мм, менее важную – 70...80 мм, а остальные – 50 мм. Клиновидная стрелка имеет преимущество перед другими; оператор быстрее и точнее ориентируется в показаниях прибора, если кончик стрелки находится на расстоянии от деления, но не более 1,6 мм.

Оптимальная ширина штриха для малых приборов – 0,8...1,0 мм, для больших – 1,2...1,5 мм. Оцифрованные штрихи должны быть в 2...4 раза толще и в 2...2,5 раза длиннее остальных. Внешняя освещенность от 20 до 700 лк не влияет на чтение показаний приборов и не маскирует различия в их читаемости.

Знаковая индикация. В условиях неограниченной экспозиции оперативный порог различения контура знака находится в пределах 9...15'. При ограничении времени экспозиции размер контура знака должен быть 60'.

Видимость цифр на электролюминесцентных индикаторах зависит от яркости их свечения и внешней освещенности. По средним данным изменение освещенности в пределах от 100 до 1000 лк почти не влияет на скорость опознания. При освещенности от 1 до 10 лк время опознания существенно увеличивается.

Качество изображения существенно зависит от разрешения экрана, которое выражается количеством пикселей по горизонтали и вертикали. Она зависит от: режима работы видеоконтроллера, частоты строк, частоты кадров и объема доступной видеопамати.

При использовании ЭЛТ-мониторов нужно обращать внимание на частоту мелькания. При яркости знака 120 нит критическая частота мелькания (нижний предел частоты, при котором изображение воспринимается непрерывным) составляет 35 Гц.

Графическая индикация применяется для грубой оценки параметров, поскольку имеет ограниченную точность.

Выбирая ту или иную категорию кода, необходимо обеспечить различие зрительных сигналов по своим физическим параметрам. Для этого в допустимых пределах должны быть величины яркости, контраста, угловых размеров изображения.

Яркость экрана измеряется в килоджоулях на квадратный метр. Максимальная яркость LCD-дисплеев определяется характеристиками ламп дневного света и составляет до 259 кд/м².

Контрастность вычисляется как соотношение самого яркого и самого темного участка на дисплее. Чтобы экран был черным, жидкие кристаллы должны полностью блокировать прохождение света. Сделать это невозможно, поэтому хорошим соотношением считается 120:1.

Угловые размеры изображения могут характеризоваться:

- *размером экрана по диагонали*; обычно он выражается в дюймах. Геометрический размер диагонали в LCD-мониторах совпадает с диагональю видимого изображения, в то время как в обычных мониторах он на 1 дюйм больше.

- *углом обзора* (в градусах по вертикали и горизонтали) считается тот, где величина контрастности падает до соотношения 10:1. Для ЭЛТ-мониторов он составляет более 150°. Для LCD-дисплеев это критический параметр, поскольку свет от задней стенки дисплейной панели проходит через поляризационные фильтры, жидкие кристаллы и ориентирующие слои. Из монитора он выходит большей частью вертикально поляризованным. Обычно он не превышает 110°. Для устранения этого недостатка производители LCD-дисплеев применяют улучшенные технологии, что расширяет угол до 170°.

6.2. Средства управления системой

С помощью средств управления оператор осуществляет исполнительные (управляющие) воздействия: ввод командной (цифровой и логической) информации, установку требуемых режимов работы аппаратуры, регулировку различных параметров, вызов информации для контроля и т.п. Для решения этих задач используются различные типы органов управления, которые могут быть классифицированы по ряду признаков.

По *характеру движений*, выполняемых человеком, различают:

1) органы управления, требующие движений включения, выключения или переключения. Движения в этом случае простые,

хотя двигательный акт складывается из довольно значительного количества микродвижений пальцев;

2) органы управления, требующие повторяющихся движений: вращательных, нажимных, ударных (например, работа на клавиатуре терминала). Значительную роль здесь играет темп движения;

3) органы управления, требующие точных дозированных движений, например, для настройки параметров. Движения при этом дозируются по силовым, пространственным и временным координатам.

По *характеру перемещения* органы управления могут быть линейными, вращающимися или смешанными.

По *назначению и характеру использования* оператором органы управления разделяются на следующие группы:

- оперативные (основные), используемые постоянно для программного управления, установки режимов работы, регулирования параметров системы, ввода управляющей и командной информации;

- используемые периодически (обычно это вспомогательные органы управления для включения и выключения аппаратуры, периодического контроля ее работоспособности и выполнения других операций, не требующих высокой скорости управляющих воздействий);

- используемые эпизодически, связанные с настройкой, калибровкой основной аппаратуры и регулировкой работы вспомогательного оборудования, регламентными работами, подключением к индикаторам датчиков измеряемых параметров и выполнением других эпизодических операций.

По *конструктивному исполнению* органы управления разделяются на подгруппы: кнопки, тумблеры, рукоятки, селекторные переключатели, педали и т.п.

Все эти средства управления разделяются на два класса: для реализации непрерывного и дискретного принципа управления. Первый широко используется для непрерывного изменения параметров процесса; второй – для ввода различных команд. Выбор того или иного принципа зависит, в первую очередь, от вида изменяемого параметра. Кроме того, он определяется допустимым временем управления и удобством работы оператора.

При непрерывном управлении каждый орган управления связан с изменяемым параметром. Обычно оператор может наблюдать за этим изменением в зависимости от изменения положения органа управления. Время ввода зависит от величины управляющего воздействия.

При дискретном управлении тип управляющего сигнала и его величина кодируются определенными кодами, которые могут вводиться с кнопочной панели или при помощи селекторного переключателя. Время ввода зависит от разрядности и в пределах заданной разрядности не зависит от величины управляющего воздействия.

Все органы управления должны обладать достаточным сопротивлением: упругим, фрикционным, вязкостным, инерционным. Упругое сопротивление противодействует отведению от нуля, возрастает с увеличением отвода и возвращает орган управления в исходное положение. Фрикционное сопротивление обеспечивает удержание органа управления в заданном положении и возрастает при увеличении скорости движения. Вязкостное сопротивление препятствует резким движениям и возрастает с увеличением скорости. Наконец, инерционное сопротивление препятствует быстрым изменениям скорости. Сопротивление обеспечивает точность, скорость и плавность движения и формирует «чувство управления», а также уменьшает возможность их случайного включения.

При установке органов управления необходимо учитывать привычные для человека стереотипы движений. Например, положениям «Пуск», «Включено», «Увеличение», «Подъем», «Открытие» или движениям «Вперед», «Вправо», «Вверх» должны соответствовать перемещения рычагов вверх, от себя, вправо, повороты маховичков или ручек по часовой стрелке, а для кнопок – нажатие верхних, передних или правых кнопок.

Для облегчения управления, уменьшения ошибок и времени поиска органа управления можно использовать различные методы их кодирования: с помощью пояснительных надписей, формы и цвета.

Для связи средств управления с другими средствами системы используются два принципа.

Циклический принцип характеризуется тем, что состояния всех управляющих средств вводятся в систему с определенным постоянным циклом (обычно 0,5...2 с), после чего процессор сравнивает их с предыдущими состояниями. При обнаружении несоответствия включаются соответствующие программы управления. *Апериодический принцип* подразумевает ввод только тогда, когда состояния изменились. Эта работа выполняется дополнительным оборудованием средств управления. Апериодический принцип требует несколько больших затрат оборудования, но меньше влияет на производительность процессора.

При построении средств управления очень важно предусмотреть средств обеспечения заданной достоверности информации. Она зависит от двух условий: 1) от правильности набора оператором вводимой величины; 2) от правильности работы аппаратуры, воспринимающей и передающей набранные данные в систему.

Первое условие определяется квалификацией оператора и психофизиологическими условиями его работы. Очень важно заранее определить последствия неверных действий и предусмотреть исключение при этом аварийных ситуаций. В ряде случаев можно предусмотреть дополнительные действия и логические операции, повышающие внимательность оператора (ввод рассчитанного значения контрольной суммы, повторный ввод и т.п.), однако нужно учитывать, что эти меры увеличивают время реакции оператора.

Второе условие обеспечивается использованием качественных комплектующих и оперативным контролем работы аппаратуры.

При проектировании средств отображения и управления очень важно определить функции человека в системе. Затем выбирается распределение органов управления по типам, после чего устанавливается их размещение на щите управления.

6.3. Пульты управления

Пульт управления – рабочее место оператора. При организации рабочего места должны быть соблюдены следующие условия:

- достаточное рабочее пространство для оператора, позволяющее осуществлять все необходимые движения и перемещения при эксплуатации и техническом обслуживании оборудования;

- достаточные информационные связи между оператором и оборудованием, а также между операторами;
- оптимальное размещение рабочих мест и достаточные проходы для операторов;
- оптимальное размещение оборудования на рабочем месте;
- необходимое освещение и допустимый уровень акустического шума, вибраций и других факторов производственной среды;
- надежная индикация отказов электропитания и аппаратуры пульта.

К основным инженерно-психологическим характеристикам пультов управления относятся его форма и геометрические размеры. На практике применяются следующие формы пультов:

- фронтальная, когда можно все органы управления разместить в пределах зон максимальной и допустимой досягаемости, а индикаторы – в пределах зоны центрального и периферического зрения;
- трапецевидная, когда это невозможно сделать; в этом случае органы управления и индикаторы частично размещаются на боковых панелях, развернутых относительно фронтальной под углом $90 \dots 120^\circ$;
- многогранная, когда средств отображения и индикаторов много; боковые панели располагаются так, чтобы они были перпендикулярны линии взора оператора.

Геометрические размеры пульта управления устанавливаются исходя из антропометрических характеристик того контингента операторов, которым предстоит работать за данным пультом. При этом нужно учитывать, что размер панели, т.е. расстояние между наиболее удаленными приборами по обе стороны от центрального прибора, составляет $0,7d$ (d – дистанция наблюдения). Дистанция наблюдения при вертикальных досках не должна быть менее 330 мм и более 750 мм. Полное поле зрения оператора охватывает пространство на 60° ниже уровня глаз и на 30° в любую сторону от средней плоскости тела.

В качестве объектов восприятия для оператора чаще всего используются различные знаки, штрихи, цифры. Оптимальные угловые размеры должны составлять: для отметок шкал – от 10 до $25'$,

для цифр или букв – 30', для контурных знаков – 17', для сложных знаков – 35'.

Разница между яркостью фона и яркостью объекта является обязательным условием видимости объекта. Она характеризуется яркостным контрастом, который выражается следующим отношением разности между яркостью объекта B_0 и фона B_ϕ :

$$K = (B_0 - B_\phi)/B_\phi \text{ или } K = (B_\phi - B_0)/B_0.$$

Как показали исследования, оптимальная величина контраста лежит в пределах 0,85...0,90. Оптимальная величина яркости экрана принята равной 35 нт с допустимыми отклонениями в пределах от 25 до 50 нт.

При построении информационной модели важно соблюдать последовательность организации внимания: расположение элементов модели должно соответствовать наиболее вероятной последовательности изменений состояния управляемых объектов.

Основные требования к средствам управления касаются, в основном, их размеров и перемещающих усилий. Так, для кнопок (нажатие одним пальцем) диаметр должен лежать в пределах от 12 до 18 мм, допустимое сопротивление – 0,15...0,3 кг, ход – 3...6 мм; для поворотных ручек и селекторных переключателей (регулирование двумя или тремя пальцами) при высоте от 12 до 25 мм диаметр должен составлять 6...25 мм, сопротивление – 0,065...0,1 кг. Клавишные переключатели должны иметь ширину не менее 18...20 мм, длину не менее 20...30 мм, величину рабочего хода не менее 5...10 мм, усилие нажатия не должно превышать 2 кгс.

При размещении органов управления следует учитывать зоны досягаемости рук человека. Оперативные, т.е. наиболее важные и часто используемые органы управления необходимо размещать в пределах оптимальной зоны, которая обеспечивает быстрые, точные и наименее утомительные движения. Вспомогательные органы управления могут размещаться в пределах допустимой или даже максимальной зоны досягаемости. Эпизодически используемые органы управления могут размещаться и за пределами зон досягаемости или быть скрыты под лицевой или боковой панелью пульта управления, но при условии обеспечения легкого доступа к ним.

При размещении органов управления рядом с теми средствами отображения, к которым они относятся, необходимо, чтобы рука оператора при манипулировании органом управления не закрывала индикационную часть соответствующего средства отображения. Направление перемещения органа управления должно правильно сочетаться с изменением показаний соответствующего индикатора. Для повышения точности и скорости действий оператора большое значение имеет правильное расположение индикаторов и органов управления в зоне деятельности оператора. Расположение этих элементов может проводиться с использованием следующих принципов: функционального соответствия, объединения, совмещения стимула и реакции, последовательности действий, важности и частоты использования.

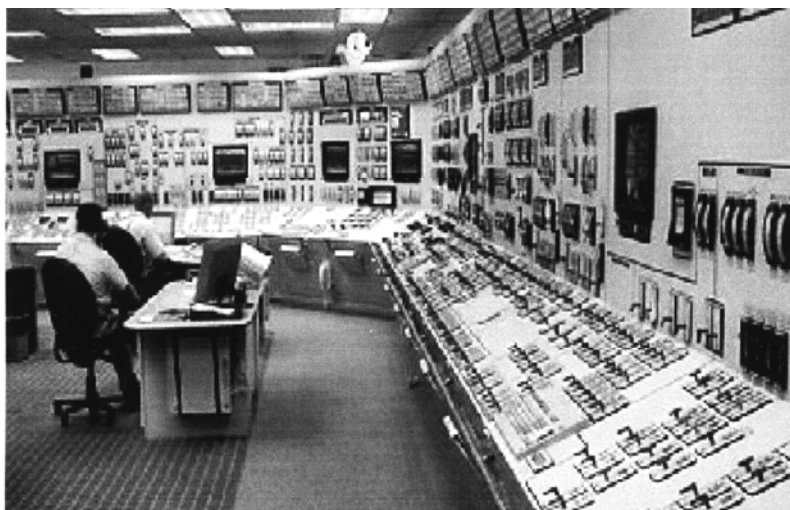


Рис. 6.3

Принцип *функционального соответствия* требует, чтобы каждой подсистеме, входящей в состав системы управления, соответствовала отдельная часть пульта управления. Принцип *объединения* требует объединения в одну группу однотипных элементов, принимающих одно и то же состояние на некотором отрезке времени. Принцип *совмещения стимула и реакции* требует пространственно-

го соотношения элементов управления и индикации. Согласно принципу *последовательности действий* элементы пульта управления следует размещать в последовательности, соответствующей алгоритму управления системой. Последний принцип предусматривает размещение наиболее *часто используемых* и *важных* индикаторов и органов управления в удобном для оператора месте.

Пример пульта управления представлен на рис. 6.3.

6.4. Деятельность оператора в системах управления

Автоматизированные системы управления различаются по своему целевому назначению. Соответственно, различается и деятельность оператора в них. Операторы связаны с различными переменными системы и внешней среды, необходимыми для осуществления процесса управления. Характер этих переменных (скорость изменения, периодичность, величина, взаимосвязь), организация рабочего процесса и поста управления в целом (имеется в виду преимущественное использование тех или иных индикационных устройств и органов управления), условия работы оператора и, наконец, требования к самому оператору также могут существенно различаться.

С этой точки зрения автоматизированные системы управления можно разделить на следующие группы [7]:

- 1) управление движущимися объектами;
- 2) управление диспетчерского типа;
- 3) управление средствами сбора информации;
- 4) управление энергетическими установками;
- 5) управление технологическими процессами циклического типа.

В системах первого типа целью, которой подчинено решение оперативных задач, является обеспечение достижения объектом управления заданного пункта в течение заданного времени. Главные задачи – это достижение и удержание заданных или выбор и удержание оптимальных значений переменных движения, а также обеспечение безопасности движения. Основными управляющими переменными являются курс, скорость, высота (глубина), крен,

дифферент и т.д. Кроме того, оператор должен получать информацию о параметрах внешней среды, характеризующих метеорологические условия, наличие объектов в зоне движения, их характеристиках и т.п.

Органы управления в рассматриваемых системах должны обеспечивать плавные, легко дозируемые по скорости и амплитуде движения. В основном для этой цели используются органы управления вращательно-рычажного типа.

Работа операторов, управляющих движением объекта, имеет свои характерные особенности. Они обусловлены высокой скоростью изменения переменных движения, внезапным развитием предельных состояний или критических ситуаций, высокой вероятностью развития неожиданных изменений внешних условий в зоне движущегося объекта и т.п. Операторы подвергаются воздействию ускорения, перепадов давления, температуры, вибрации, качки, шума и т.п.

Рабочий процесс операторов систем управления движущимся объектом, наряду с высокоактивными периодами, может включать очень монотонные, однообразные малоактивные периоды.

В системах второго типа основное назначение – обслуживание заявок. Оператор должен получать информацию, позволяющую ему выносить решения, во-первых, о значимости той или иной заявки (курс, высота, удаленность, тип и состояние самолета, запрашивающего разрешение на посадку); во-вторых, о состоянии каналов обслуживания. В связи с этим наиболее характерными средствами отображения информации являются экраны, планшеты и мнемосхемы, которые позволяют получать наглядную картину ситуации. В качестве основных органов управления на пульте диспетчера используются переключатели.

Работа операторов-диспетчеров в течение всей смены обычно бывает очень интенсивной; при этом наблюдается неравномерность поступления заявок на обслуживание, исключающая возможность ритмичности в работе. К оператору-диспетчеру предъявляются высокие требования в отношении объема внимания и оперативной памяти, тактического и стратегического мышления и, наконец, способности быстрого принятия решений.

Системы третьего типа предназначены для получения информации о наличии объектов в определенных зонах воздушной или водной среды. Главные оперативные задачи – это обнаружение объекта, его опознание и наблюдение за изменениями местоположения объектов в пределах границ контролируемого пространства.

Основными переменными системы, оптимальное регулирование которых позволяет добиваться отображения сигнала-носителя на средствах индикации, являются такие, как положение (направленность) и чувствительность приемных устройств.

Характерными средствами отображения информации в системах наблюдения за обстановкой являются индикаторные устройства экранного типа, обеспечивающие круговой или секторный обзор пространства. Органы управления представлены настроечными вращательными рукоятками, маховичками и т.п.

Условия работы в системах наблюдения требуют постоянного внимания. Как правило, наблюдение за экраном ведется при пониженной внешней освещенности, что способствует повышению чувствительности зрения и повышает контраст «сигнал – фон». Однако при пониженной освещенности ухудшаются условия выполнения других действий. Из-за необходимости местной подсветки элементов оборудования возникает неравномерность яркости поля зрения, вызывающая дополнительную нагрузку на механизмы адаптации зрительной системы человека. В течение смены могут быть резкие перепады интенсивной и малоактивной, монотонной деятельности. Оператор системы наблюдения также должен обладать способностью к длительной концентрации внимания и большим объемом оперативной памяти.

Системы четвертого типа обеспечивают оптимальный или заданный режим источника энергии в течение заданного времени. Оперативные задачи в системе можно представить как контроль и регулирование энергетического оборудования, контроль и регулирование нагрузки, контроль энергоресурсов, защита энергетической установки.

Основными переменными контроля и регулирования для оператора являются переменные энергии (напряжение, сила тока, давление, температура и т.п.) и переменные оборудования (техническое

состояние и рабочее положение отдельных элементов, скорость вращения, изоляция, сопротивление, вибрация и т.п.).

Основные средства отображения – это мнемосхемы, а также стрелочные цифровые и контрольно-измерительные приборы; органами управления являются переключатели и регуляторы.

Условия работы характеризуются достаточно большим однообразием в силу высокой степени автоматизации. Основная функция управления – контроль, поэтому очень важна проблема сохранения готовности к действию. В критических ситуациях от оператора требуется быстрое принятие решения, основывающегося на правильном техническом диагнозе.

Когда посты операторов располагаются в непосредственной близости от агрегатов, операторы подвергаются воздействиям шума, вибраций, высокой температуры, ускорений, качки и т.п.

Оператор автоматизированной системы управления энергетической установкой должен обладать глубокими знаниями устройства оборудования, взаимодействия отдельных устройств. Он должен хранить в памяти большое число количественных характеристик различных переменных и параметров. Оператор должен уметь реализовывать эти данные при постановке технического диагноза, уметь осуществлять ручное управление.

Пятый тип систем управления предназначаются обычно для массового выпуска каких-либо изделий или вещества. В течение одного цикла технологического процесса начальная форма продукта может доводиться либо до конечной формы, либо только до промежуточной формы, пригодной для последующей обработки.

В качестве оперативных задач следует назвать регулирование работы блоков автоматического оборудования, выполнение отдельных операций с использованием механизированных блоков оборудования, обеспечение заданного ритма циклов технологического процесса, контроль качества продукции.

Основными переменными контроля и регулирования являются переменные физического состояния продукта (температура, размер, вес, цвет и т.п.), а также переменные состояния оборудования. В качестве средств отображения информации наиболее часто используются световые и звуковые сигнализаторы, а также шкальные ин-

дикационные устройства. Типичными органами управления, используемыми на постах управления систем данного типа, являются кнопки, педали, маховички и рычаги.

Условия работы оператора характеризуются монотонностью, выраженной тем сильнее, чем чаще осуществляются циклы технологического процесса.

Успешность деятельности лица, принимающего решения по управлению, в первую очередь зависит от того, насколько информационная модель системы по форме и композиции соответствует его задачам по управлению и его психофизиологическим возможностям по приему и переработке информации.

Прежде всего, это соответствие скорости выдаваемой моделью информации пропускной способности оператора. Эта характеристика существенно зависит от организации его работы. Ее можно характеризовать *полным временем оператора*. Под этим термином понимается время, необходимое оператору для выполнения одной логически законченной операции, которая может состоять из нескольких частных операций. Условно полное время оператора можно представить в виде следующего выражения:

$$T_{\text{оп}} = T_{\text{зп}} + T_{\text{п}}N + T'_{\text{п}} + T_{\text{оу}} + T_{\text{м}} + T_{\text{к}},$$

где $T_{\text{зп}}$ – суммарное время зрительного поиска; $T_{\text{п}}$ – время приема информации и принятия решения в тех случаях, когда не требуется осуществление исполнительных действий; N – количество индикаторов, проконтролированных оператором до того, как он принял информацию, требующую исполнительных действий; $T'_{\text{п}}$ – время приема информации и принятия решения в том случае, когда требуется выполнение определенных исполнительных действий; $T_{\text{оу}}$ – время обнаружения органа управления; $T_{\text{м}}$ – время моторного акта, непосредственно связанного с движением органа управления; $T_{\text{к}}$ – время контроля результата действия.

Обычно пропускная способность не превышает 30 дв. ед./с; реально она лежит в пределах 1...5 дв. ед./с.

Организация потоков информации должна исключать как перегрузку, так и недогрузку оператора.

Для уменьшения перегрузки необходимо:

- предоставлять информацию оператору с необходимым упреждением к началу исполнения;
- сократить поток информации до необходимого минимума;
- предусмотреть возможность фильтрации информации;
- разработать рациональную схему деятельности оператора;
- позволить оператору использовать для принятия решения максимальное время (в пределах общего времени, отведенного на выполнение задачи);
- сохранять на индикаторе информацию по желанию оператора на необходимое время.

Недогрузка оператора вызывает ослабление внимания, что приводит к потере ритма и к ошибкам. Для уменьшения недогрузок необходимо:

- сократить до минимума время от запроса до воспроизведения информации, а также время формирования изображения;
- обеспечить достаточную интенсивность потока информации (при интенсивности потока 1...10 сигналов в час уже может наблюдаться заметное ослабление внимания);
- принять меры к повышению «заметности» вновь появляющейся информации;
- ограничить площадь размещения информации;
- обеспечить оператору возможность контроля за правильностью своих действий.

Большое значение имеет выбор системы кодирования информации. Применительно к деятельности оператора *кодированием* называется способ представления информации с помощью условных символов. Оптимальным является такой способ кодирования, который обеспечивает максимальную скорость и надежность приёма информации человеком.

Задача оптимального кодирования заключается в правильном выборе категории кода, длины алфавита сигналов, уровня кодирования. Категория кода определяется способом кодирования: видом символов, яркостью, цветом, частотой мелькания, размерами. В задаче опознания наибольшую эффективность обеспечивает цвет, в задачах поиска – цвет и вид символов. Наиболее эффективным в

условиях большой нагрузки на оперативную память являются цифровойой и буквенный коды.

Для привлечения внимания оператора целесообразно использовать кодирование различной частотой мелькания. Привычные ассоциации сложились у человека по отношению к различным цветам: красный цвет обычно ассоциируется с опасностью, желтый – с необходимостью получения какого-либо предупреждения, зеленый – со спокойной обстановкой.

Человек-оператор допускает большое число ошибок при управлении сложными системами. По его вине происходит около 40 % общего количества отказов при испытаниях ракет, 64 % отказов в военно-морском флоте и до 70 % отказов в авиации. Надежность оператора обеспечивается правильной организацией режима труда и отдыха, а также подготовкой операторов, обеспечивающей умение регулировать работоспособность в зависимости от требований, диктуемых различными условиями работы [8].

Важнейшими показателями организации режима труда операторов являются: коэффициент загрузки, период занятости, длина очереди [9].

Коэффициент загрузки представляет величину

$$\eta = 1 - (\tau_p / T_{\text{деж}}),$$

где τ_p – общее время, в течение которого оператор не занят обработкой поступающей информации; $T_{\text{деж}}$ – общая продолжительность дежурства. Рекомендуемое значение $\eta \leq 0,75$.

Под периодом занятости понимается время непрерывной работы; рекомендуется, чтобы оно не превышало 15 мин.

Вероятность работы в условиях очереди не должна превышать 0,4; при этом длина очереди не должна превышать трех сигналов.

Контрольные вопросы и упражнения

1. Расскажите о классификации средств отображения. Приведите примеры всех типов устройств отображения.
2. Нарисуйте структуру электронно-лучевой трубки и объясните назначение отдельных элементов.
3. Нарисуйте структуру LCD-панели и объясните принцип ее действия.

4. Перечислите основные требования к информационной модели системы управления.
5. Что нужно делать для уменьшения перегрузки?
6. Что нужно делать для уменьшения недогрузки?
7. Расскажите о классификации средств управления системой.
8. Каково обязательное требование к средствам управления?
9. Какие принципы используются для связи средств управления с другими средствами системы? В чем они заключаются?
10. Каким образом компонуются средства управления со средствами отображения на пультах управления?
11. Предложите способы уменьшения полного времени оператора.

7. ИСПОЛНИТЕЛЬНЫЕ СРЕДСТВА

7.1. Виды исполнительных средств

Исполнительные средства предназначены для создания управляющего воздействия на регулирующие органы объекта управления с целью изменения состояния объекта в соответствии с алгоритмом управления. Состояние объекта определяется состоянием или положением регулирующих органов.

В зависимости от управляющего воздействия на выходе исполнительного элемента они делятся на *силовые* и *параметрические*. Изменение пространственного положения регулирующего органа возможно только если исполнительное устройство создает управляющее воздействие в виде силы или момента. Такие исполнительные органы получили название *силовых*. К ним относятся электромагниты, муфты, двигатели. Исполнительные элементы, изменяющие состояние регулирующего органа (сопротивление, напряжение, ток), называются *параметрическими*. К ним относятся электромагниты, контакторы, тиристорные и транзисторные реле.

Основные требования к силовым исполнительным устройствам заключаются в следующем:

- их максимальная сила или момент должны быть больше, чем перестановочное усилие на регулирующем органе;
- они должны иметь максимальный КПД;
- их статические характеристики должны быть близки к линейным;
- они должны иметь минимальный порог чувствительности, большую мощность управления, высокую надежность и долговечность, небольшие размеры и массу.

Выходные порты компьютера имеют очень низкую мощность и не могут управлять каким-либо физическим устройством; сигнал необходимо усиливать. На уровнях мощности до нескольких сотен ватт можно использовать операционные усилители мощности. При больших мощностях чаще используются программируемые источники энергоснабжения. Для этих целей служит *широтнопulseвая модуляция (ШИМ)*. Принцип ШИМ заключается в том, что выходное напряжение переключается между двумя постоянными значениями с высокой частотой, обычно в диапазоне не-

скольких килогерц. Средний уровень напряжения поддерживается изменением ширины импульса.

7.2. Основные типы силовых исполнительных органов

Электромагниты – это преобразователи электрического сигнала в механическое движение. Они делятся на удерживающие и приводные. *Удерживающие* электромагниты служат для фиксации положения ферромагнитных тел (например, электромагнитные плиты для фиксации деталей на металлообрабатывающих станках). Они не совершают никакой работы, от них требуется лишь определенная сила, на которую они рассчитываются. *Приводные* электромагниты служат для перемещения регулирующих органов (например, клапанов, заслонок), а также используются в контакторах, электромагнитных муфтах и др. Эти электромагниты совершают определенную работу и поэтому рассчитываются на определенную силу и перемещение.

Электромагнитные муфты служат для передачи вращающего момента с ведущего входного вала (чаще всего приводного двигателя) на соосный ведомый выходной вал (вал приводимого регулирующего органа). Для быстрого включения, выключения и реверса приводимых механизмов, а также для регулирования их скорости и ограничения передаваемого момента применяют различные управляемые электромеханические сцепные муфты. Рассмотрим некоторые примеры таких муфт.

Фрикционные муфты служат для жесткого сцепления валов.

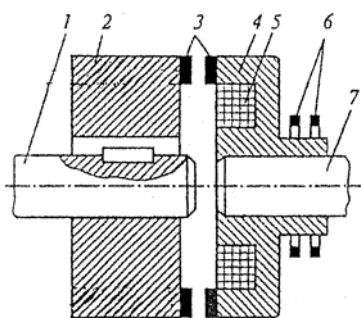


Рис. 7.1

На рис. 7.1 показана однодисковая неререверсивная фрикционная муфта, состоящая из двух цилиндрических полумуфт [2]. На ведущем валу 7 жестко посажена полумуфта 4, являющаяся сердечником и ярмом электромагнита. Его якорем служит полумуфта 2, которая соединена с ведомым валом 1 скользящей посадкой и, следовательно, может перемещаться в осевом направлении. С

помощью колец 6 и щеток на обмотку 5 подается управляющее напряжение. Возникающее тяговое усилие вызывает притяжение якоря (полумуфты 2) к сердечнику (полумуфте 1) и плотное сцепление фрикционных дисков 3 из материалов с высоким коэффициентом трения, за счет чего обеспечивается сцепление валов. Основное достоинство однодисковой фрикционной муфты – простота, однако с ростом передаваемого вращающего момента значительно увеличиваются ее размеры.

Ферропорошковые муфты служат для гибкого сцепления валов. В них полумуфта на ведомом валу посажена жестко, и неизменный воздушный зазор заполнен магнитодиэлектриком, обладающим свойством тиксотропии – способностью становится студенистым, все более загустевая вплоть до затвердения по мере усиления магнитного поля, а при снятии его возвращаться в исходное состояние.

На рис. 7.2 показана схема цилиндрической ферропорошковой муфты с двумя цилиндрическими поверхностями 9 и 10. Кольцевое пространство между ними заполнено порошковой смесью 8. Детали 6 выполнены из ферромагнитных материалов. На внутренней ведущей полумуфте (3 – ведущий полувал) расположена обмотка возбуждения 7, выведенная на контактные кольца 2, к которым прижимаются щетки 4. Крышки 5 и 11, изготовленные из немагнитного материала, позволяют направить большую часть магнитного потока 1 через порошковый слой, уменьшив поток рассеяния, и снизить массу ведомой полумуфты.

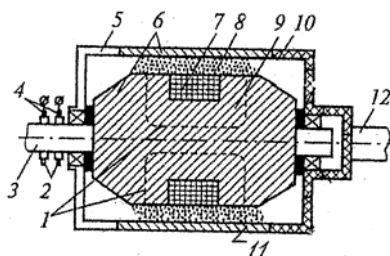


Рис. 7.2

Исполнительные электродвигатели находят широкое применение в системах управления. Теоретические основы электрических двигателей вкратце сводятся к следующему.

У машины постоянного тока нормального исполнения цилиндрический ротор с расположенной на нем обмоткой, называемый *якорем*, вращается в неподвижном магнитном поле, создаваемом обмоткой возбуждения, расположенной на статоре, на которую подается постоянное напряжение (рис. 7.3). Протекание тока в обмот-

ке якоря обусловлено напряжением U , направление которого совпадает с направлением тока. Это напряжение подается на якорь через щетки и коллектор – совокупность медных изолированных друг от друга пластин, образующих цилиндр, по которому скользят щетки. Наличие коллектора является отличительной особенностью машин постоянного тока. Его пластины образуют полукольца, к которым присоединяются начала и концы витков обмотки.

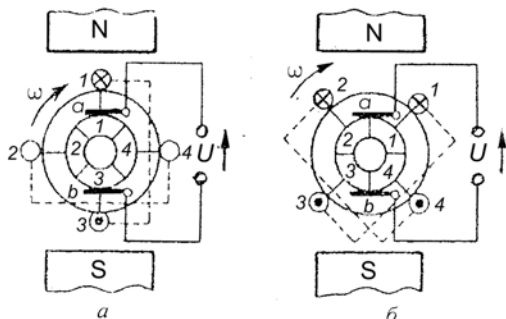


Рис. 7.3

На проводники с током действует электромагнитная сила, приводящая в движение якорь. В положении *a* развиваемый двигателем момент будет максимальным, так как стороны *1–3* витка пересекают магнитные линии перпендикулярно. При дальнейшем вращении момент начнет убывать. После поворота якоря на 45° щетки перейдут с пластин *1–3* коллектора на пластины *2–4* и подключатся к следующему витку (рис. 7.3, *б*). Момент будет сначала возрастать до максимума, а затем начнет уменьшаться, пока щетки снова не перейдут на пластины *3–1* и момент снова начнет возрастать и т.д. Рост и уменьшение момента происходят за время, соответствующее повороту якоря на угол $\pi/2$, на протяжении которого каждая щетка касается одной и той же коллекторной пластины.

Управлять двигателем постоянного тока можно четырьмя способами: 1) подачей переменного напряжения на якорь при постоянстве напряжения на обмотке возбуждения; 2) включением дополнительного сопротивления последовательно с якорем; 3) подачей переменного напряжения на обмотку возбуждения при постоянстве напряжения на обмотке якоря; 4) изменением сопротивления в цепи обмотки возбуждения.

С точки зрения использования в системах управления лучшим является первый метод.

У машин переменного тока используется принцип вращающегося магнитного поля. Рассмотрим две катушки, оси которых сдвинуты в пространстве на $\pi/2$. Будем питать катушки синусоидальными токами, причем ток в горизонтальной катушке, создающей индукцию B_y , будет отставать по фазе на $\pi/2$ от тока в вертикальной катушке, создающей индукцию B_x . Амплитудные значения магнитной индукции, создаваемой катушками в центре, будут равны:

$$B_x = B_{xm} \sin \omega t, \\ B_y = B_{ym} \sin(\omega t - \pi/2) = -B_{ym} \cos \omega t.$$

Угол, который вектор магнитной индукции B будет составлять с осью x , определится выражением:

$$\alpha = \arctg(B_y/B_x) = -\arctg(\sigma \operatorname{ctg} \omega t),$$

где $\sigma = B_{ym}/B_{xm}$.

Направление вектора магнитной индукции будет непрерывно изменяться с течением времени, т.е. поле будет вращающимся. При одинаковых амплитудах токов в катушках имеем:

$$B_{xm} = B_{ym} = B_m; \quad \sigma = 1; \\ B_x^2 = B_y^2 = B_m^2.$$

Получается круговое вращающееся поле, вектор магнитной индукции которого, оставаясь неизменным по величине, вращается с постоянной скоростью.

Машина переменного тока состоит из статора и ротора. Статор представляет собой полый цилиндр, составленный из листов электротехнической стали. В пазах на внутренней поверхности статора, закладывается статорная обмотка. Она выполняется так, что при включении ее в сеть переменного тока внутри цилиндра образуется магнитное поле, вращающееся вокруг оси статора с постоянной скоростью.

В *асинхронной* машине ротор имеет вид цилиндра, обычно набранного из круглых листов стали. У поверхности ротора вдоль его образующих расположены проводники, составляющие обмотку ротора, никак не связанную с внешней электрической цепью. Токи в обмотке ротора возникают в результате того, что ротор при вращении отстает от вращающегося поля, как бы «скользит» относительно поля. Величина этих токов определяется скоростью вращения магнитного поля относительно ротора.

Для оценки этой скорости вводится понятие *скольжения* асинхронной машины:

$$s = (n_0 - n) / n_0,$$

где n_0 – скорость вращения магнитного потока или синхронная скорость; n – скорость вращения ротора.

Поскольку условием возникновения токов в роторе является неравенство скоростей, ротор не может вращаться со скоростью, равной синхронной, чем и объясняется название *асинхронная* машина.

Движущий момент возникает как результат взаимодействия вращающегося магнитного потока с индуктируемым им в роторе токами. Скорость вращения магнитного поля статора (синхронная скорость) равна

$$n_0 = 60f/p,$$

где f – частота переменного тока; p – число пар полюсов.

Для управления асинхронной машиной используются: 1) амплитудный метод, когда меняется амплитуда на всех фазах двигателя; 2) частотный метод, когда меняется частота питающих напряжений; 3) амплитудно-частотный метод, когда одновременно меняются амплитуда и фаза напряжений.

В *синхронной* машине ротор представляет собой электромагнит, обмотка которого питается постоянным током. Два конца этой обмотки выводятся на контактные кольца, укрепленные на валу машины и вращающиеся вместе с ротором. Постоянный ток подводится к ротору извне через неподвижные щетки, скользящие по контактным кольцам. Ротор вращается синхронно с вращающимся магнитным полем статора, чем и объясняется название этих машин.

Регулирование скорости возможно только частотным методом.

К положительным качествам двигателей *постоянного тока* относятся:

- возможность получения теоретически любых, сколь угодно малых и больших частот вращения;
- возможность простого, плавного, экономичного и в широком диапазоне регулирования частоты вращения;
- устойчивость работы практически при любых частотах вращения;
- линейность механических и регулировочных характеристик;
- значительный пусковой момент;

- сравнительно небольшая электромеханическая постоянная времени;

- малые габаритные размеры и масса.

Основным недостатком наиболее широко распространенных коллекторных исполнительных двигателей постоянного тока, ограничивающим области их применения, является наличие скользящих контактов – коллектора и щеток.

Достоинства двигателей *переменного тока* заключаются в том, что они экономичны, имеют простую и надежную конструкцию, простое энергопитание; у асинхронных двигателей отсутствуют вращающиеся коммутирующие элементы. Основной особенностью синхронных двигателей является постоянство частоты вращения при неизменной частоте питающей сети.

Шаговые двигатели – это электромеханические устройства, которые преобразуют электрические импульсы напряжения управления в дискретные угловые и линейные перемещения ротора с возможной его фиксацией в нужных положениях. Они относятся к синхронным двигателям; поворот их ротора на определенный угол вызывается подачей сдвинутых на определенный угол питающих напряжений на обмотки или группы обмоток. Их рабочий режим – стопорный. Обычный угловой шаг – 15^0 при частоте вращения от 110 до 600 шагов в секунду; некоторые современные шаговые двигатели имеют угловой шаг 3^0 при скорости вращения до 1000 шагов в секунду.

К достоинствам шаговых двигателей относят:

- высокую точность без обратной связи, т.е. без датчика угла поворота;

- отсутствие механических вращающихся коммутирующих элементов.

Их недостатки:

- малый вращающий момент;

- ограниченная скорость;

- высокий уровень вибраций из-за скачкообразного движения;

- большие ошибки при потере импульсов в системах с разомкнутым контуром управления.

Устройство, содержащее двигатель, редуктор и элементы управления двигателем называется *приводом*. Структурная схема привода представлена на рис. 7.4. Двигатель Д через редуктор Р

перемещает регулирующий орган РО. Сигнал U_c , поступающий на вход привода, имеет обычно недостаточную мощность для управления двигателем, поэтому он предварительно усиливается усилителем У. Концевые выключатели КВ служат для ограничения перемещения РО. Оператор может устанавливать РО с помощью устройств дистанционного управления ДУ, контролируя его положение прибором П, а при неисправности ДУ – штурвалом ручного управления РУ. Датчики обратной связи по положению ОСП, выполняемые в виде потенциометров, индуктивных датчиков или линейных индукционных потенциометров, и датчики обратной связи по скорости ОСС, выполняемые в виде тахогенераторов постоянного или переменного тока, служат для ввода дополнительных сигналов, необходимых для получения требуемых характеристик от привода.

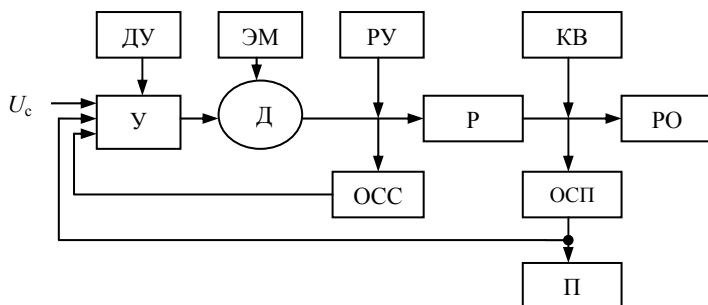


Рис. 7.4

Номинальный момент M на выходном валу и время T полного оборота выходного вала, т.е. быстродействие, являются основными характеристиками привода. Мощность P на валу двигателя, необходимая для обеспечения заданных времени T и момента M , определяется по формуле

$$P = 61,5M/T\eta,$$

где η – КПД редуктора.

Инерционность привода, определяемая временем от начала движения РО до установления полной скорости, зависит от соотношения между пусковым моментом двигателя и моментом инерции привода. Пусковой момент обычно в 2...2,5 раза больше номинального. Важной характеристикой привода является время запаз-

дывания – время от момента подачи сигнала до начала вращения выходного вала.

При проектировании исполнительных элементов для систем управления в первую очередь подлежит выбору вращающий момент, достаточный для создания нужного перестановочного усилия регулирующего органа. Затем устанавливаются требования к габаритам и массе устройств и к их надежности. Выбор делается с учетом стоимости и затрат на эксплуатацию.

Контрольные вопросы и упражнения

1. В чем различие между силовыми и параметрическими исполнительными органами?
2. В чем различие между удерживающими и приводными электромагнитами?
3. Для чего служат электромагнитные муфты?
4. В чем различие между фрикционными и ферропорошковыми муфтами?
5. Расскажите о принципе действия электродвигателя постоянного тока.
6. Как можно управлять двигателем постоянного тока? Объясните принципы управления.
7. Объясните принцип работы электродвигателя переменного тока.
8. В чем различие между асинхронной и синхронной машинами?
9. Для чего служат концевые выключатели в приводах?
10. Для чего служат датчики обратной связи по положению и по скорости?

8. СРЕДСТВА СВЯЗИ

8.1. Эталонная модель взаимодействия открытых систем

Основой для разработки проектов взаимодействия технических средств систем управления реального времени является многоуровневая распределенная система управления (PCY).

К нижнему уровню PCY относятся механизмы и устройства, которые непосредственно «соприкасаются» с процессом – датчики и исполнительные средства. Эта так называемая «полевая» аппаратура образует **уровень локального управления**. На более высоком уровне – **уровне управления процессом** – находятся контроллеры, регуляторы и другие «интеллектуальные» устройства, способные решать простейшие задачи сбора и локальной обработки информации, а также вывода управляющих сигналов на объект.

На среднем уровне PCY – **уровне координации механизмов** – обрабатывается управляющая программа системы, т.е. координируется работа отдельных механизмов. Для этого здесь установлены ЭВМ, которые могут обмениваться между собой информацией. Кроме того, этот уровень обменивается информацией в виде опорных и текущих значений с вышестоящими и нижележащими уровнями.

На следующем уровне PCY – **уровне координации процессов** – координируется деятельность нескольких участков для достижения равномерного потока материалов или энергии.

Верхний уровень PCY – **уровень стратегического управления** – это уровень принятия общих решений, которые влияют на работу всего предприятия.

Построенные по такому принципу системы могут быть интегрированы в единые интегрированные системы управления.

Очевидно, что на каждом уровне происходит сбор информации и обмен ею с выше- и нижележащими уровнями. Объем этой информации на каждом уровне различен, однако процесс обмена в общем виде можно представить следующим образом.

Все процессы связи включают в себя передатчик и приемник. Передатчик передает сообщение, состоящее из последовательности

символов, приемнику по каналу связи, который является общим для передатчика и приемника. Для передачи сообщения применяется какой-либо код, в соответствии с которым передатчик изменяет некоторые физические свойства канала, а приемник восстанавливает сообщение по изменениям, которые он обнаружил в канале. На канал обычно влияет шум, искажающий сообщение и затрудняющий распознавание приемником изменений в канале и правильную интерпретацию сообщения.

Основной проблемой передачи данных является доставка за заданное время сообщения от пункта А в пункт Б при минимизации влияния шума. При передаче сообщений оба участника сетевого обмена должны принять множество соглашений. Например, они должны согласовать уровни и форму электрических сигналов, способ определения длины сообщений, договориться о методах контроля достоверности и т.п.

Средства сетевого взаимодействия могут быть представлены в виде иерархически организованного множества модулей. Процедура взаимодействия двух узлов может быть описана в виде набора правил взаимодействия каждой пары соответствующих уровней обеих участвующих сторон. Формализованные правила, определяющие последовательность и формат сообщений, которыми обмениваются сетевые компоненты, лежащие на одном уровне, но в разных узлах, называются *протоколом*.

Модули, реализующие протоколы соседних уровней и находящиеся в одном узле, также взаимодействуют друг с другом в соответствии с четко определенными правилами и с помощью стандартизованных форматов сообщений. Эти правила называют *интерфейсом*.

Современные технологии располагают разнообразными средствами для передачи больших объемов данных в виде иерархически организованного набора протоколов – *стека коммуникационных протоколов*.

Для того чтобы преодолеть трудности, возникающие из-за большого количества несовместимых стандартов, Международная организация по стандартизации (ISO) разработала эталонную модель взаимодействия открытых систем (ВОС), которая представля-

ет не только еще один стандарт, но и базу для разработки новых стандартов. Модель ВОС не связана с конкретными реализациями и описывает процесс коммуникации только в абстрактных понятиях. Практической целью модели ВОС является обеспечение совместимости и взаимозаменяемости.

Совместимость означает, что обмен данными не потребует непропорциональных расходов на их преобразования.

Взаимозаменяемость означает, что устройства, выпускаемые различными производителями для выполнения одной и той же функции, могут замещать друг друга без каких бы то ни было проблем при условии, что их работа основывается на одних и тех же принципах и правилах.

В модели ВОС определены семь функциональных уровней. Объекты, расположенные на одном уровне в разных узлах коммуникационной сети, называются *одноранговыми*. Эти объекты взаимодействуют между собой на основе протоколов, которые определяют форматы сообщений и правила их передачи.

1. **Физический уровень** представляет собой физическую среду передачи. Все детали, касающиеся среды передачи, уровня сигналов и частот, рассматриваются на этом уровне.

2. **Канальный уровень** обеспечивает функции, связанные с формированием и передачей кадров от одного узла к другому, обнаружением и исправлением ошибок, возникающих на физическом уровне. Если несколько устройств используют общую среду передачи, то на этом уровне также осуществляется управление доступом к среде.

3. **Сетевой уровень** устанавливает маршрут и контролирует прохождение сообщений от источника к узлу назначения.

4. **Транспортный уровень** решает задачу независимости верхних уровней от физической структуры сети и от маршрута доставки сообщений. Этот уровень несет ответственность за проверку правильности передачи данных от источника к приемнику и доставку данных к прикладным программам.

5. **Сеансовый уровень** отвечает за установку, поддержку синхронизации и управление соединением между объектами уровня представления данных.

6. Уровень представления данных обеспечивает кодирование и преобразование неструктурированного потока бит в формат, понятный приложению-получателю, т.е. восстановление исходного формата данных (сообщение, текст, рисунок и т.п.).

7. Прикладной уровень организует решение прикладных задач – передачу файлов, операции с распределенными базами данных и удаленное управление.

Объекты обмениваются данными в соответствии с протоколами, определенными для соответствующих уровней, начиная с последнего, седьмого уровня. Перемещаясь вниз по уровням, сообщение «обрастает» служебной информацией. Наконец, оно достигает нижнего, физического уровня, который и передает его по линиям связи машине-адресату. Когда сообщение по сети поступает по сети на машину-адресат, оно принимается ее физическим уровнем и последовательно перемещается с уровня на уровень. Каждый уровень анализирует и обрабатывает заголовок своего уровня, выполняя соответствующие данному уровню функции, а затем удаляет этот заголовок и передает сообщение вышележащему уровню.

Физический уровень – единственный, имеющий материальное воплощение; остальные уровни представляют собой наборы правил или описание вызовов функций, реализованные программными средствами.

В последнее время в вычислительных сетях промышленной автоматизации находит широкое применение протокол TCP/IP (*Transmission Control Protocol/Internet Protocol*). Уровни этого протокола определены следующим образом.

Уровни 1 и 2 явным образом не специфицированы; предполагается, что существует физическое соединение и соответствующее управление каналом.

Уровню 3 соответствует межсетевой протокол IP, который обеспечивает передачу дейтаграмм по сети от одной станции к другой. На этом уровне получатель не должен подтверждать доставку дейтаграммы.

Уровень 4 включает в себя средства, расширяющие базовые службы IP. Он осуществляет транспорт дейтаграмм между прило-

жениями. Каждое приложение должно самостоятельно осуществлять проверку и исправление ошибок.

Уровень 5 представляет собой совокупность служб, предназначенных для пользователя сети.

Если рассматривать уровни РСУ с позиций организации ВОС или протокола ТСР/IP, то можно сделать вывод о том, что на двух нижних уровнях требуется решать задачи ввода/вывода информации в простейшие контроллеры и процессоры, и это соответствует первым двум уровням модели ВОС. На более высоких уровнях РСУ приходится обмениваться информацией между машинами одного уровня, и здесь требуются локальные сети, в которых взаимодействие узлов организовано так, как предписывается более высокими уровнями модели ВОС.

Основными элементами сети являются стандартные компьютеры, не имеющие ни общих блоков памяти, ни общих периферийных устройств. Каждый компьютер работает под управлением собственной операционной системы, а какая-то общая операционная система, распределяющая работу между компьютерами сети, отсутствует. Взаимодействие между компьютерами сети происходит путем передачи сообщений через сетевые адаптеры и каналы связи. С помощью этих сообщений один компьютер запрашивает доступ к локальным ресурсам другого компьютера.

В АСУТП используются три основные структуры сетевого обмена: архитектура «клиент–сервер», архитектура «ведомый–ведущий» («Master/Slave») и обмен с помощью широковещательных пакетов. Рассмотрим их особенности с точки зрения удовлетворения требованиям АСУТП реального времени (РВ). В нашем случае примем, что система работает в режиме РВ, если она реагирует в течение предсказуемого интервала времени на хаотический поток внешних событий.

В системе «клиент–сервер» клиент посылает запрос, который попадает в буфер и находится там некоторое время, прежде чем поступить в процессор для обработки. При этом сервер приостанавливает свою работу, чтобы ответить на данный запрос. Время ответа на запросы увеличивается по экспоненте с ростом числа клиентов, обращающихся к серверу. Могут образоваться очереди

заявок со стороны клиентов на сервере – источнике информации. Но в этой архитектуре получаем гарантированную доставку сообщений, хотя и медленную. О режиме РВ не может быть и речи, так как время нахождения запроса в очереди заранее неизвестно.

Архитектура «Master/Slave» применяется в основном на нижнем (контроллерном) уровне промышленной автоматизации. Центральная ведущая машина – Master (ПК или ПЛК) периодически опрашивает все необходимые ведомые станции – Slave. Получив запрос, станции выполняют какие-либо действия в зависимости от полученных данных и посылают ответ на запрос. Master получает ответ, обрабатывает его и переходит к следующему узлу. Здесь нет очередей, так как Master сам инициирует обмен, а не ждет запроса от клиентов. Время обмена информацией линейно возрастает с увеличением числа станций Slave. Так как известно, сколько времени будет происходить обмен данными станции Master со станцией Slave, то можно сказать, что здесь реализуется режим РВ.

Третья архитектура – обмен с помощью широковещательных пакетов – применяется для периодического обмена какими-либо данными (входные–выходные сигналы, расчетные параметры и т.д.), в частности, для рассылки данных от управляющей станции к остальным станциям сети. Определяющий параметр – время получения информации – не зависит от числа потребителей. Недостатки данной архитектуры: 1) отсутствие гарантии, что пакет достигнет приемника, так как здесь идет обмен пакетами без подтверждения; 2) наличие очередей со стороны приемника информации, так как при передаче данных с большой скоростью или от нескольких источников буфер на прием быстро заполнится, и остальные пакеты просто не будут приниматься.

8.2. Физические среды передачи информации

Наиболее распространенная среда передачи цифровой информации – электрический кабель. Простота, низкая стоимость и отработавшая техническая база делают кабель наиболее удобным носителем для передачи информации на ограниченные расстояния. Оптический кабель экономически эффективен при передаче больших

объемов информации на значительные расстояния в условиях помех. Если кабельное оборудование использовать невозможно из-за его удаленности или подвижности, то в качестве среды передачи можно использовать радиоволны.

Основным параметром, характеризующим канал связи, является его *пропускная способность* – количество информации, которое можно передать за единицу времени. Она тесно связана с *полосой пропускания*, которая определяется как диапазон частот, которые канал способен передавать с затуханием менее чем 3 дБ (что соответствует 50 % уровня мощности). Обычная телефонная линия имеет полосу пропускания 3,1 кГц, типичная полоса пропускания телевизионного канала – 5,5 МГц.

Важным фактором, отрицательно влияющим на связь, являются помехи (шум). Связь между максимальной пропускной способностью канала $R_{\text{макс}}$ и отношением средних мощностей исходного сигнала и шума S/N выражается формулой Шеннона:

$$R_{\text{макс}} = W \log_2(1+S/N).$$

Наиболее распространенные типы электрических проводников, используемых для связи, – это витая пара и коаксиальный кабель. Витая пара более чувствительна к электромагнитному шуму, полоса ее пропускания ограничена несколькими мегагерцами. Однако благодаря своей простоте и низкой стоимости она часто применяется в качестве среды передачи. Экранированная витая пара состоит из двух пар скрученных проводов, экранированных проводящей оплеткой, которая заземляется. Она хорошо защищает от внешних помех, но удорожает кабель и усложняет его прокладку, так как требует выполнения качественного заземления.

Коаксиальный кабель обеспечивает большую устойчивость к помехам и имеет пропускную способность до 500 МГц. Он обычно применяется для передачи на большие расстояния нескольких сигналов.

В последнее время в качестве среды передачи все чаще используется оптический кабель. Волоконно-оптические кабели состоят из центрального проводника (сердцевины) – стеклянного волокна, окруженного другим слоем стекла – оболочкой, обладающей меньшим показателем преломления, чем сердцевина. Распространяясь

по сердцевине, лучи света не выходят за ее пределы, отражаясь от покрывающего слоя оболочки, как от зеркала. Такое полное внутреннее отражение близко к идеальному, благодаря чему свет может распространяться на значительные расстояния с минимальными потерями.

В зависимости от профиля распределения показателя преломления от центра к периферии и от величины диаметра сердечника различают:

- многомодовое волокно со ступенчатым изменением показателя преломления. Оно имеет резкую границу между сердцевинной и оболочкой и большое затухание (порядка 2,5 дБ/км). Показателя преломления всего два: в сердцевине и в оболочке. Свет источника состоит из волн различной длины, которые при распространении по волокну проходят различный путь. В результате передаваемый импульс по мере прохождения по волокну «размывается» и ослабляется. Это приводит к фактическому слиянию следующих друг за другом импульсов, что ограничивает частоту передачи при увеличении длины линии. Так, например, для типового волокна этот предел равен 15 МГц·км, что означает: видеосигнал с частотой 5 МГц может быть передан не более чем на 3 км;

- многомодовое волокно с плавным изменением показателя преломления. В нем обеспечиваются различные показатели преломления в сердцевине кабеля: от максимального в середине до минимального на периферии. Это приводит к двум следствиям: 1) свет проходит по слегка изгибающемуся пути и 2) задержки распространения различных мод минимальны. Это объясняется тем, что высокие моды, входящие в волокно под большим углом и проходящие больший путь, распространяются с большей скоростью по мере их удаления от центра. Низкие моды, распространяющиеся вблизи сердцевины и проходящие меньший путь, распространяются с меньшей скоростью. Увеличение скорости компенсирует больший путь. Затухание в многомодовых линиях с градиентным изменением показателя преломления порядка 0,8 дБ/км и система наиболее эффективна при передаче на большие расстояния;

- одномодовое волокно имеет толщину сердечника порядка 8 мкм (т.е. она сопоставима с длиной волны) и поэтому допускает

передачу только на одной длине волны; оно имеет затухание порядка 0,4 дБ/км.

В многомодовых наиболее употребительны размеры 62,5/125 мкм и 50/125 мкм (внутренний диаметр/внешний диаметр).

В качестве источников излучения применяются светодиоды (выходная мощность до 0,1 мВт, ширина полосы пропускания 30–80 нм, максимальная скорость передачи – 8 Мбит/с) и полупроводниковые лазеры (выходная мощность до 10 мВт, ширина полосы пропускания 5 нм; скорость передачи до 10 Гбит/с). Для практического применения установлены три рабочих диапазона оптоволоконной связи при длине волны 850–900, 1300 и 1550 нм.

Волоконно-оптические кабели обладают отличными электромагнитными и механическими характеристиками, имеют малое затухание сигнала, однако их серьезный недостаток – сложность соединения волокон с разъемами и между собой при необходимости наращивать длину кабеля.

Еще одним физическим носителем для передачи данных являются радиоволны. Это реальная альтернатива в тех случаях, когда нельзя установить проводное соединение. Радиоканалы работают на частотах от 100 кГц до 100 МГц.

8.3. Основные способы передачи битовой последовательности

Существуют два основных способа передачи битовой последовательности по физическому каналу:

- посылка бит в линию в непосредственном или закодированном виде при сохранении цифрового характера данных (цифровое кодирование);
- модуляция несущей по амплитуде/частоте/фазе и передача модулированного сигнала.

Цифровое кодирование является наиболее простым. При цифровом кодировании применяют потенциальные и импульсные коды. В потенциальных кодах для представления единиц и нулей используется только значение потенциала сигнала; импульсные коды позволяют представить двоичные данные либо импульсами опре-

деленной полярности, либо частью импульса – перепадом потенциала определенного направления.

При использовании для передачи дискретной информации прямоугольных импульсов необходимо выбрать такой способ кодирования, который

- имел бы наименьшую ширину спектра кодированного сигнала при одной и той же скорости передачи;
- обеспечивал синхронизацию между передатчиком и приемником;
- обладал способностью распознавать ошибки;
- обладал низкой стоимостью реализации.

При непосредственном кодировании, например, уровень напряжения 0 В представляет логический «0», а +10 В – логическую «1». Широко используется полярное кодирование – сигналы, соответствующие «0» и «1» имеют противоположные знаки по отношению к общей базе (рис.8.1,а). Такой **метод без возвращения к нулю** (*Non-Return to Zero – NRZ*) прост в реализации, но не обладает достаточной помехоустойчивостью и свойством самосинхронизации. Кроме того, недостатком этого кода является наличие низкочастотной составляющей.

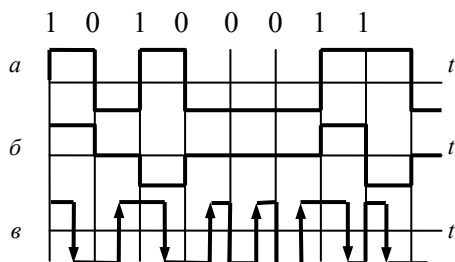


Рис. 8.1

Метод биполярного кодирования с альтернативной инверсией (рис. 8.1, б) для кодирования «0» использует нулевой уровень, а для кодирования «1» используется либо «+», либо «-», причем потенциал каждой следующей единицы противоположен предыдущему. В последовательности «1» недостатки метода NRZ устраняются, при передаче «0» – нет.

При **прямом манчестерском кодировании** бит «0» представляется переходом от уровня низкого напряжения к уровню высокому; бит «1» – в обратном направлении (рис. 8.1, в). Этот метод обладает свойством самосинхронизации: в каждом такте на середине периода имеется переход от одного уровня к другому; эти импульсы можно использовать для синхронизации.

Поскольку физическая среда передачи данных на большие расстояния представляет собой длинную линию, состоящую из последовательно соединенных сопротивлений и индуктивностей с параллельным соединением конденсаторов, передача по ней прямоугольных импульсов приводит к их искажениям. Лучше превратить эти импульсные сигналы в модулированный синусоидальный сигнал.

Аналоговая модуляция является таким способом физического кодирования, при котором информация кодируется изменением амплитуды, частоты или фазы несущей частоты. При амплитудной модуляции спектр сигнала состоит из синусоиды несущей частоты f_c и двух боковых гармоник: (f_c+f_m) и (f_c-f_m) , где f_m – частота измерения информационного параметра сигнала. Этот способ редко используется в чистом виде, но часто применяется в сочетании с другим видом модуляции – фазовой модуляцией.

При частотной модуляции каждый уровень напряжения модулирующего напряжения соответствует определенной частоте. Этот способ модуляции не требует сложных схем и обычно применяется в низкоскоростных сетях на скоростях 300 или 1200 бит/с. При фазовой модуляции значениям данных 0 и 1 соответствуют сигналы одинаковой частоты, но с различной фазой.

Изменение параметра частоты можно связать с последовательностью бит. Так, при фазовой модуляции можно дать непосредственное соответствие: «0» – фазовый сдвиг 0° , «1» – 180° ; можно сделать иначе: 00, 01, 10, 11 закодировать соответственно фазовыми сдвигами 0° , 90° , 180° и 270° .

Число изменений параметра несущей частоты в секунду называется *скоростью передачи* и измеряется в бодах. Бод совпадает с битом только тогда, когда это изменение соответствует одному биту. Как видно из приведенного примера, это не всегда так.

Когда несущая частота модулируется импульсным сигналом, модуляция называется *манипуляцией*.

Импульсные методы передачи можно использовать для непосредственной связи между устройствами на расстоянии до трех километров.

Наиболее простым способом разделения физического канала между несколькими устройствами является *мультиплексирование* – разделение каналов во времени или по частоте. При *временном мультиплексировании* каждое устройство имеет доступ к каналу только в закрепленные за ним временные интервалы. При *частотном мультиплексировании* полоса пропускания канала делится на диапазоны частот, каждый из которых закрепляется за одним виртуальным каналом. Осуществляется модуляция несущей частоты, которая находится в середине выделенного диапазона. Временное мультиплексирование имеет два преимущества перед частотным мультиплексированием:

- 1) вся обработка проводится на цифровом уровне, и нет необходимости в установке и обслуживании высокочастотного оборудования для модуляции и демодуляции сигналов несущей частоты;
- 2) допускается управление распределением временных интервалов между отдельными устройствами; если устройство не передает данные, его свободные интервалы можно передать другому устройству.

Если передача данных по каналу осуществляется всегда в одном направлении, то режим передачи называется *симплексным*; если в каждый момент времени каналом пользуется одно устройство, то передача называется *полудуплексной*; при *полном дуплексном* режиме передача происходит одновременно в обоих направлениях. Такой режим может осуществляться только с помощью частотного мультиплексирования, при котором каждая сторона имеет свой выделенный частотный диапазон.

Для выявления ошибок при передаче данных были разработаны различные методы контроля. Основной подход – добавить к исходному сообщению дополнительную информацию, позволяющую приемнику установить наличие ошибок при передаче. Информация, добавляемая к сообщению для контроля ошибок, состоит из

одного или нескольких байт, вычисленных по содержимому исходного сообщения с помощью определенного алгоритма. Она называется *контрольной суммой*. Самый распространенный способ определения контрольной суммы – разбить данные на блоки известной длины (от нескольких байт до нескольких тысяч байт) и разделить каждый блок на двоичное число длиной один или два байта. Контрольной суммой является остаток от деления. Двоичное число, используемое при делении, называется *порождающим полиномом*. В нем указываются только степени, отличные от нуля. Например, стандартный порождающий полином в соответствии с рекомендацией МККТТ V.41 имеет вид

$$x^{16} + x^{12} + x^5 + 1,$$

что соответствует двоичной последовательности 1000100000010001. Если данные не повреждены, то операция деления данных вместе с контрольной суммой на порождающий полином должна дать в остатке нуль. Другой результат указывает на повреждение данных в блоке и необходимость его повторной передачи.

8.4. Архитектура связи между устройствами

Связь между отдельными устройствами в составе системы управления может быть организована либо по шинной, либо по канальной архитектуре.

8.4.1. Шинная архитектура доступа к периферийным устройствам

Шина – это совокупность линий – проводников, имеющих определенное функциональное назначение. Требования к шинам определяются их функциональным назначением. В соответствии с ним все шины можно разделить на следующие категории: 1) процессорная шина, связывающая процессор с основной памятью; 2) шина ввода/вывода, специально предназначенная для связи процессора или ОП с ПУ; 3) системная шина, которой могут подключаться все устройства.

Использование системной шины – это простейшая схема подключения ПУ к ядру ЭВМ (процессору и основной памяти). Системная шина состоит из трех наборов линий: для передачи адресов, данных и управляющих сигналов. Кроме линий, в архитектуру шины входит контроллер шины. Его назначение – управление обменом данными и служебными сигналами. Для связи с шиной каждое устройство имеет собственный блок управления (контроллер). Шина, таким образом, представляет собой *общий интерфейс*, а контроллеры служат для его согласования со спецификой функционирования отдельных устройств.

Все устройства передают и (или) принимают адресную, числовую и управляющую информацию посредством одинакового набора сигналов. Набор сигналов и правил, на основании которых объединяются устройства, соединенные шиной, называют *протоколом шины*.

Обмен данными между двумя устройствами, подключенными к общей шине, производится по принципу ведущий – ведомый (задатчик – исполнитель). Ведущее устройство активно, оно инициирует обмен; ведомое – пассивно, оно ждет запросов. Чаще всего в качестве ведущего устройства выступает процессор или контроллер ПДП. Могут быть и другие варианты, однако устройства памяти и вывода всегда пассивны и работают только как исполнители.

Операцию передачи данных по шине называют *транзакцией*. Если в операции передачи участвуют устройства ввода/вывода, то различают транзакции ввода (аналогично операциям чтения) и транзакции вывода (аналогично операциям записи). Промежутком времени, в течение которого выполняется одна транзакция, называется *циклом* шины.

Для связи с шиной устройство имеет блок управления – контроллер: задатчик – *передатчик* шины, а исполнитель – *приемник* шины. Для устройств, которые могут быть и задающими, и исполнительными, требуется *приемопередатчик* шины.

При разработке шины важнейшими вопросами являются: выбор ширины шины, ее синхронизация и арбитраж.

Под *шириной* шины принято понимать количество линий в ней, т.е. количество одновременно передаваемых бит данных. При вы-

делении в шине отдельных адресных линий их количество определяет возможности процессора по обращению к памяти – размер ее адресуемого пространства. Ширина шины данных, т.е. число линий, по которым передаются разряды слова данных, влияет на *пропускную способность* шины – количество бит данных, передаваемых по шине в единицу времени (обычно в секунду).

Для широких шин требуется больше проводов и более широкие разъемы; следовательно, она более дорогая. Для того чтобы не увеличивать ширину шины и не удорожать тем самым ее стоимость, разработчики применяют мультиплексную шину. В ней нет пространственного разделения линий на линии данных и адресные линии. Это разделение временное: сначала передается адрес, а затем по тем же линиям – данные. Объединение этих линий сокращает ширину и стоимость шины, но снижает ее пропускную способность. В зависимости от требований к этим характеристикам делается соответствующий выбор.

Поскольку шина соединяет для совместной работы два устройства, работающих независимо друг от друга, требуется решать задачу синхронизации обмена между ними. Можно реализовать синхронный или асинхронный режим обмена. Шина, предназначенная для работы в синхронном режиме (*синхронная* шина) содержит линию, по которой идут тактирующие сигналы, формируемые кварцевым генератором. Все транзакции жестко «привязаны» к этим сигналам. Шина, работающая в асинхронном режиме (*асинхронная* шина) не имеет такой синхронизирующей линии, поэтому начало и конец транзакции не имеют жесткой «привязки» к определенным моментам времени.

В конфликтной ситуации, когда одновременно два или более активных устройств пытаются принять управление шиной на себя, в действие вступает **арбитр**. *Арбитраж* – принятие решения о том, какому из них передать функции временного «хозяина» шины – реализуется схемами анализа приоритетов. Устройство с высоким уровнем приоритета всегда имеет превосходство по отношению к устройству с низким уровнем. Процессору присваивается, как правило, самый низкий уровень приоритета, потому что его работу можно без ущерба прервать, какие бы операции он не выполнял.

Другие же устройства могут иметь движущиеся механические части или быть связаны с процессом, протекающим в реальном времени, что в любом случае требует немедленной реакции на запрос.

Особенность унифицированной шины – способ адресации внешнего устройства: используется одинаковый формат команд как для ввода/вывода, так и для команд обращения к памяти. Как правило, для простого периферийного устройства используется всего два адреса: один для регистра данных и второй для регистра управляющего слова. Для более сложной аппаратуры типа ЗУ на магнитном диске число адресов возрастает до шести.

К достоинствам структуры с общей шиной относят простоту организации многопроцессорных вычислительных систем и относительную легкость модернизации и наращивания состава периферийных устройств. Главный недостаток – невозможность обеспечить параллелизм в работе процессора и внешних устройств.

Стандартизация сигналов и протоколов обмена позволяет подключать без каких бы то ни было переделок одни и те же устройства к разным процессорам. Стандартизации подлежат шина ввода/вывода и системная шина; процессорная шина не может быть стандартизована: ее структура тесно связана с архитектурой ядра ЭВМ, с электрическими схемами его элементов, с их рабочей частотой.

В одной ЭВМ может использоваться несколько различных шин. Это дает возможность лучше учитывать особенности работы устройств разных типов. Например, очевидно, что организация обмена с клавиатурой и с монитором предъявляет различные требования к скорости, к объему передаваемой информации, к управляющим сигналам и т.д.

Рассмотрим наиболее распространенные стандарты шин, применяемых в ЭВМ, рабочих станциях и сетевых серверах.

ISA (Industry Standard Architecture). Этот промышленный стандарт системной шины был разработан более двух десятилетий назад, тем не менее, в основной массе простых ЭВМ и в особенности в промышленных контроллерах до сих пор находит применение из-за своей простоты.

В этом стандарте все устройства: процессор, память и устройства ввода/вывода объединены одной магистралью. Для синхронизации процессора, ОП и сравнительно медленно действующих устройств, таких как внешние запоминающие устройства или устройства ввода/вывода, имеется оборудование, которое обеспечивает добавление определенного количества циклов ожидания. Так, при обращении к основной памяти может быть добавлен 1 цикл шины, при обращении к портам ввода/вывода – 4 цикла, при обращении в цикле прямого доступа к памяти (ПДП) – 1 цикл и т.п.

В качестве управляющих сигналов для схемы добавления тактов ожидания, на основе которых принимается решение о количестве добавляемых циклов, выступают команды чтения/записи, сигналы выбора основной памяти и другие. Анализируя эти сигналы, схема в случае необходимости переводит процессор в состояние неготовности и удерживает его в этом состоянии соответствующее число циклов.

Разрядность шины ISA 8/16– 8 или 16 бит для данных и 16 или 24 – для адреса; тактовая частота 4,77 или 8,33 МГц; пропускная способность составляет 8,33 и 16,7 Мбайт/с. Шина поддерживает синхронный протокол обмена.

Позднее шина ISA была расширена до 32 разрядов данных; разрядность адреса также составила 32 разряда. Такая шина получила название EISA (Extended Industry Standard Architecture – расширенная архитектура промышленного стандарта). Максимальная пропускная способность этой шины при пересылке одиночных слов и тактовой частоте 33 МГц составляет 33,3 Мбайт/с, при групповой пересылке слов ее пропускная способность возрастает до 132 Мбайт/с.

PCI (Peripheral Component Interconnect). Это шина ввода/вывода. Шина PCI поддерживает все функции, типичные для процессорной шины, но в стандартизированном формате, и поэтому является процессорно-независимой. Разрядность адреса и данных 32 или 64. При длине слова 64 бита, тактовой частоте 66 МГц и групповой пересылке шина обеспечивает пропускную способность 528 Мбайт/с.

Характерная особенность этой шины в том, что ее можно связывать с другими шинами посредством «мостов». Подключенные к этой шине ПУ представляются процессору непосредственно соединенными с его собственной шиной. Специальный контроллер заботится о разделении управляющих сигналов локальной шины процессора и PCI-шины. Периферийным устройствам назначаются адреса из адресного пространства памяти процессора.

В стандарте PCI предусмотрены специальные средства для конфигурирования ЭВМ. При подключении к ней нового устройства операционной системе нужно знать его тип и сведения о таких характеристиках устройства, как скорость передачи данных, использование контрольных разрядов и т.д. Эти данные должны храниться в специальной конфигурационной памяти устройства. Программное обеспечение, выполняющее инициализацию шины, считывает содержимое этой памяти и назначает адрес каждому устройству.

Чтобы не увеличивать число линий, адресные и информационные шины линии объединены и работают в режиме мультиплексирования.

Шина PCI управляется централизованным арбитром.

Все обмены по шине буферизованы. PCI легко совместима с большинством известных шин: для этого разработаны и реализованы в виде микросхем «мосты» – PCI/ISA, PCI/EISA и др.

SCSI (Small Computers System Interconnect). Шина SCSI также является шиной ввода/вывода. Она стандартизована (стандарт ANSI X3-0131) в различных модификациях. «Узкая» шина имеет разрядность данных 8 бит и работает на частотах 5 или 10 МГц; «широкая» шина – 16-разрядная. Адаптер SCSI содержит управляющий процессор, способный управлять ПУ разных типов (дисками, принтерами, магнитными лентами, оптическими дисководами с перезаписью и др.). Максимальная скорость передачи по конкретной шине зависит от длины соединительного кабеля и количества подключенных устройств. При максимальной длине соединительного кабеля (25 м) и семи подключенных устройствах «узкая» шина обеспечивает пропускную способность 1,5 Мбайт/с при асинхронной передаче и 5 Мбайт/с – при синхронной. При длине

кабеля не более 1,6 м пропускная способность шины «широкого» стандарта SCSI-3 составляет 320 Мбайт/с.

В отличие от устройств, подключаемых к шине процессора, устройства, подключаемые к шине SCSI, не являются частью его адресного пространства. Шина SCSI соединяется с шиной процессора через SCSI-контроллер. Для пересылки пакетов данных от основной памяти и обратно применяется технология прямого доступа к памяти.

Для шины SCSI применяется схема распределенного арбитража.

Типы внешних устройств, подключаемых к процессору, многообразны. Различны скорости, объемы и временные характеристики обмена. В широком диапазоне варьируются расстояния между устройствами. В середине 1990-х годов была сделана попытка разработать стандарт обмена, обеспечивающий гибкость и универсальность применения при низкой стоимости. Так появилась шина USB (Universal Serial Bus).

Это универсальная шина с последовательной передачей данных, однако ее архитектура не похожа на стандартную шинную, а приближается к канальной. Для того чтобы к шине USB можно было подключать большое количество устройств, она имеет древовидную структуру.

В узлах дерева располагаются устройства, называемые **концентраторами**, или **хабами** (от англ. *hub* – концентратор), действующими как промежуточные управляющие блоки между ядром ЭВМ и периферийными устройствами. Корневой хаб соединяет все дерево с ядром. Каждый хаб имеет несколько выходов, к которым можно подключать ПУ и другие хабы. В нормальном режиме хаб копирует полученное из ядра ЭВМ входное сообщение в свои выходные регистры и оно передается всем устройствам, но отвечает на него только адресуемое. В этом отношении шина USB функционирует подобно общей шине. Однако в отличие от этой шины сообщение от ПУ пересылается только вверх, в направлении корневого узла, и другие устройства его не получают.

Шина USB представляет собой ряд каналов от центрального хаба к ПУ. Каждое устройство может иметь несколько адресуемых

регистров для хранения служебной информации (например, вектора состояния). Для обращения к каждому такому адресуемому элементу используется отдельный подканал в общем канале, ведущем к этому устройству. Максимальное число подканалов – 16. В каждом канале или подканале данные перемещаются от центрального концентратора к устройству и обратно. Между двумя устройствами ввода/вывода обмена информацией не происходит. В основе функционирования шины USB лежит принцип опроса устройств: они могут передавать сообщения только в ответ на запрос процессора.

Каждому устройству на шине USB присваивается адрес. Он никак не соотносится с адресами, используемыми на шине процессора.

Передача данных по шине USB идет кадрами, каждый из которых состоит из пакетов. Центральный концентратор шины каждую миллисекунду передает кадр, синхронизирующий все устройства во времени. В этом в кадре всего один пакет SOF (Start Of Frame – начало кадра) и он посылается всем устройствам. Только этот кадр передается от концентратора к устройству. Следующие пакеты кадра могут передаваться в любом направлении.

Шина USB поддерживает три режима функционирования: низкоскоростной (1,5 Мбит/с), полноскоростной (12 Мбит/с) и высокоскоростной (480 Мбит/с). У разработчика вычислительной системы есть выбор, как организовать структуру обмена. Можно к одному хабу подключить несколько низкоскоростных устройств, а к другому – высокоскоростные. Стандарт допускает также присоединение к одному хабу низкоскоростных и высокоскоростных устройств с разделением времени их обслуживания подобно тому, как это делается в мультиплексных каналах.

8.4.2. Архитектура доступа с процессорами ввода/вывода

Шинная система связи характеризуется непредсказуемыми задержками обмена из-за наличия общего ресурса – шины. Для того, чтобы свести их к минимуму, система ввода/вывода может строиться путем централизации аппаратуры управления на основе специализированных процессоров ввода/вывода, часто называемых

каналами ввода/вывода. Эти устройства способны реализовывать достаточно сложные процедуры ввода/вывода, что позволяет полностью разгрузить процессор от управления операциями ввода/вывода.

Аппаратура канала может в определенной степени учитывать специфику подключаемого с его помощью внешнего устройства, поэтому этот принцип связи относится к архитектуре с **множественными интерфейсами**.

Для управления вводом-выводом используется набор команд ввода/вывода, который обеспечивает управление работой любого периферийного устройства, стандартным образом подключенного через канал к ядру ЭВМ. Все команды ввода/вывода имеют следующие поля: код операции ввода/вывода, адрес канала и адрес ПУ в канале.

Операции передачи данных между ОП и ПУ начинаются по команде НАЧАТЬ ВВОД/ВЫВОД, в которой указывается адрес канала А и адрес устройства В, участвующего в данной операции. Коды операций, которые будет выполнять канал совместно с ПУ, и адреса областей памяти, которые будут использованы для хранения вводимой (выводимой) информации, задаются управляющими словами канала (УСК). На момент начала операции они размещаются в ОП в соседних ячейках памяти, образуя цепочку УСК, и выбираются каналом по мере необходимости. Последовательность УСК составляет *программу канала*.

Формат УСК такой: код команды канала, адрес данных, признаки, счетчик. Код команды определяет операцию, которая должна быть выполнена в канале. Адрес задает первую ячейку памяти, с которой начинается область, выделенная для выполнения команды. Признаки дают дополнительную информацию, необходимую для работы канала: наличие цепочки данных и цепочки команд, неправильная длина блока данных и др. Счетчик фиксирует число передаваемых символов, т.е. размер памяти, выделенной в данном УСК для заданной операции.

Основная функция канала – связать периферийное устройство с основной памятью. Поэтому организация работы канала зависит от соотношения скоростей работы периферийного устройства и ОП.

Известны два режима работы каналов: монопольный и мультиплексный, или режим разделения времени.

В **монопольном** режиме после установления связи с каналом ПУ захватывает канал на все время, пока полностью не завершится инициированная процессором программа обмена. Другие ПУ доступа к каналу в это время не имеют. Если во время работы канала с ПУ от процессора поступает команда обращения к другому ПУ, то выполнение этой команды задерживается до момента окончания операции в канале. В режиме **разделения времени** программные и аппаратные средства канала по очереди обслуживают несколько ПУ. Во время работы канала с одним ПУ другие могут выполнять работу, не требующую использования средств канала (например, могут работать механические элементы устройств ввода/вывода). Интервалы времени, относящиеся к различным ПУ, чередуются в порядке поступления из внешних устройств запросов на обслуживание.

В соответствии с этими режимами работы различают каналы мультиплексные и селекторные.

Мультиплексный канал ввода/вывода предназначен для работы с медленными устройствами в режиме отдельных сеансов связи с каждым из них. За один сеанс передается набор управляющих сигналов, символ или группа символов. Если одновременно запрашивают обслуживания несколько ПУ, то канал выбирает одно из них в соответствии с принятой системой приоритетов. Остальные ПУ, готовые к обмену, ждут своей очереди. Способность ожидания обслуживания без потери информации – обязательное условие подключения устройства мультиплексному каналу.

В физическом отношении мультиплексный канал представляет собой цифровое вычислительное устройство, специализированное для выполнения команд ввода/вывода, поступающих от процессора, и операций по обслуживанию устройств ввода/вывода в процессе передачи данных.

Из рассмотрения структуры мультиплексного канала следует, что его организация идентична организации процессора с микропрограммным управлением и с этих позиций мультиплексный канал можно рассматривать как специализированную ЭВМ, подклю-

чаемую к высокопроизводительному процессору для выполнения действий по обмену информацией.

Операции в *селекторном канале* выполняются исключительно в монопольном режиме, при котором он в течение достаточно большого промежутка времени используется для обслуживания только одного ПУ. Селекторный канал можно рассматривать как канал, содержащий только один подканал.

Процедуры работы селекторного канала во многом похожи на процедуры мультиплексного канала, однако, принимая во внимание высокую скорость обмена, которую должен обеспечить селекторный канал, в его аппаратуру вводят некоторые особенности.

Во-первых, для хранения и модификации текущих параметров операции в селекторном канале используются регистры, а не память, как в мультиплексном канале. Во-вторых, в селекторных каналах применяют предварительную выборку управляющего слова. В-третьих, введением буферных регистров обеспечивается возможность совмещения во времени обмена с ПУ и обмена с ОП.

При проектировании АСУ важно определить структуру обмена, т.е. решить, какую – шинную или канальную архитектуру – следует выбрать, после чего выбрать наиболее подходящий стандарт обмена.

8.4.3. Сбор данных от датчиков

В системах, где периферия представлена автоматическими датчиками и исполнительными устройствами, не всегда целесообразно использовать сложные сетевые структуры. Существует три более простых способа сбора данных от датчиков и передачи их к более высоким уровням иерархии.

Телеметрия. Это способ передачи данных, когда они передаются непрерывно в заранее определенном формате. После завершения одного цикла передачи начинается новый. Каждый параметр определяется его положением в потоке данных.

Опрос. Управляющий компьютер циклически опрашивает текущее состояние датчиков и периодически обновляет данные в своей внутренней базе данных.

Прерывание. Этот метод заключается в передаче только тех переменных, которые изменились по сравнению с предыдущим циклом. Прерывания инициируются датчиками, когда они должны передавать информацию.

На локальном уровне управления используются как линии связи «точка-точка», так и полевые шины. Рассмотрим сначала линии связи, соединяющие между собой два устройства.

Интерфейс с токовой петлей (Current Loop) относится к классу универсальных двухточечных радиальных интерфейсов удаленного последовательного доступа к контроллерам. Он широко применяется в промышленном оборудовании, так как позволяет осуществить связь по физическим линиям на дальние расстояния (до трех км) без использования аппаратуры передачи данных (модемов).

Интерфейс CL представляет собой двухпроводную линию, образующую токовую петлю с дискретно подключаемым источником тока и приемником (рис. 8.2).

Последовательные данные от источника к приемнику передаются побитно и побайтно асинхронным способом. Международный стандарт IEC 381 рекомендует для передачи сигналов диапазон токов 4...20 мА. Ток, превышающий 17 мА, представляет логическую единицу, а ток, меньший, чем 2 мА – логический нуль. Одно из взаимодействующих устройств должно быть активным и служить источником тока, а другое – пассивным (приемником).

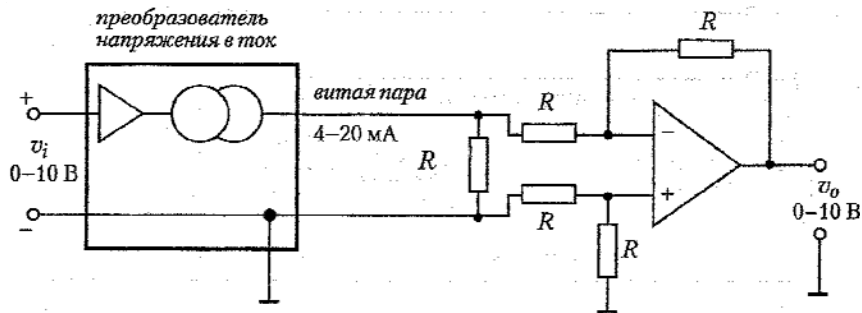


Рис. 8.2

Интерфейс CL имеет, как правило, протяженную линию передачи, которая подвержена влиянию внешних помех и перенапряжений. Поэтому схемы передатчика и приемника должны быть гальванически развязаны за счет использования оптронов и изолированных источников питания. Максимальная скорость передачи по токовой петле – 966 бит/с при длине линии связи до 300 м. Снижая скорость передачи, можно почти пропорционально увеличить длину линии: на скорости 1200 бит/с длина линии увеличивается до 2000 м. Токовая петля обычно используется для сопряжения одного приемника и одного передатчика, но в принципе она может охватывать и несколько последовательно соединенных пассивных приемников.

Токовая петля позволяет передавать данные по двухпроводной линии в одном направлении (симплексная связь): от передатчика к приемнику. Для дуплексной связи (одновременной передачи в двух противоположных направлениях) используется четырехпроводная линия. Интерфейс содержит цепь 1 «Передаваемые данные» и цепь 2 «Принимаемые данные». Этот интерфейс гарантирует передачу сигналов со скоростью 966 бит/с на расстояние до 500 м (на больших расстояниях пропорционально снижается скорость), но не регламентирует типы применяемых кабелей и разъемов.

Другой тип наиболее широко используемого интерфейса – интерфейс стандарта Ассоциации электронной промышленности США (EIA) **EIA-232D** (ранее *RS – Recommended Standard*). Этот тип интерфейса применим как для синхронной, так и для асинхронной связи между двумя устройствами в симплексной, полудуплексной и дуплексной режимах. Стандарт регламентирует состав, назначение и обозначение линий интерфейса, их нумерацию, электрические характеристики и уровни сигналов, скорости передачи и тип разъемов.

В зависимости от условий конкретного применения используется конкретное число линий интерфейса. Так, для асинхронного обмена по физическим линиям нужно три цепи: данные передатчика TD, данные приемника RD и сигнальная земля GND.

Скорость передачи данных по интерфейсу RS-232D составляет от 50 до 19200 бит/с, а максимальная длина линий связи при максимальной скорости не превышает 16 м.

Позднее были разработаны новые стандарты, позволившие улучшить согласование линий, увеличить расстояние и скорость передачи данных, реализовать более сложную структуру соединения приборов. Стандарт **EIA-422C** ориентирован на использование дифференциальной сбалансированной линии передачи с импедансом 50 Ом, что повышает помехоустойчивость интерфейса и скорость передачи (10 Мбит/с при длине кабеля до 13 м и 100 Кбит/с – при длине 1300 м). Кроме того, этот стандарт допускает нагрузку до 10 приемников. Более поздний стандарт **RS-485A**, являющийся усовершенствованием EIA-422C, ориентирован при тех же скоростных характеристиках на нагрузку до 32 источников и 32 приемников данных.

8.5. Полевые шины

Рассмотренные выше протоколы решают задачу сбора информации от небольшого количества датчиков. При наличии множества устройств ввода/вывода подобная организация была бы очень неэффективной. Требуется такая организация ввода/вывода, которая позволяла бы при наличии минимума соединений обеспечивать прием или выдачу информации множеству однотипных источников. Такая задача решается с помощью локальных шин. Локальные шины, объединяющие «полевую» аппаратуру, называют *«полевыми шинами»*.

Для уровня датчиков необходимо выполнение следующих требований: шина должна передавать данные в соответствии с жестким временным регламентом; объем данных должен быть минимальным, чтобы обеспечить работоспособность в критические по нагрузкам моменты. Шина на уровне датчиков должна обеспечивать интерфейс не только для простых, «неинтеллектуальных» пассивных устройств, но и для сложных устройств и контроллеров.

С середины 1980-х годов предпринимаются попытки выработать единый стандарт шины, обеспечивающей взаимодействие контроллеров, устройств связи с объектом, датчиков и исполнительных механизмов, подключенных к одной шине. Шина такого назначения получила название «полевой шины» (fieldbus). Подобный

стандарт должен устанавливать требования к открытому цифровому протоколу обмена, который обеспечивал бы возможность совместной работы средств автоматизации различных производителей.

Такая шина является цифровой, двунаправленной, многоточечной последовательной коммуникационной шиной, используемой для связи таких field-устройств, как контроллеры, датчики, силовые приводы и т.п. Каждое field-устройство обладает самостоятельным вычислительным ресурсом; оно способно самостоятельно выполнять ряд функций по самодиагностике, контролю и обслуживанию двунаправленной связи. Доступ к нему возможен не только со стороны инженерной станции, но и со стороны аналогичных ему устройств. Поэтому технология fieldbus – это нечто большее, чем просто замена интерфейса с токовой петлей.

Fieldbus должна отвечать ряду требований:

- жесткая детерминированность (предсказуемость) поведения;
- обеспечение функций реального времени;
- работа на длинных линиях с использованием недорогих физических средств (например, витая пара);
- повышенная надежность физического и канального уровней передачи данных для работы в промышленной среде (например, при больших электромагнитных помехах);

Общее количественное снижение коммуникационного оборудования делает систему не только проще в эксплуатации, но и надежнее за счет уменьшения потенциальных отказов аппаратуры.

Переход на fieldbus-технологии обещает улучшение качества, снижение затрат, повышение эффективности конечной системы. Серьезный ценовой выигрыш получается за счет проводников и монтажных работ: аналоговая технология связи требует, чтобы каждое устройство имело собственный набор проводов и собственную точку соединения. Fieldbus устраняет эту необходимость, так как использует всего одну витую пару проводников для объединения всех активных (контроллеры) и пассивных (датчики) устройств.

В настоящее время существует несколько десятков стандартов промышленных шин Fieldbus и коммуникационных протоколов

(Bitbus, Profibus, Interbus-S, CAN, Modbus, Lon Works, Fieldbus Foundation и т.д.) и от правильности выбора шины в каждом конкретном случае зависит, насколько успешно будет решена поставленная задача.

Шина **Bitbus** разработана для построения распределенных систем, в которых должны обеспечиваться высокая скорость передачи, детерминированность и надежность доставки данных. Структура шины всегда включает один ведущий (MASTER) узел и до 249 ведомых (Slave) узлов. Физический уровень основан на RS 485, но возможна работа и по оптическому кабелю. Классическая реализация протокола выполнена на микроконтроллере i8044 (Intel), поддерживающем функции удаленного доступа и имеющего небольшое ядро ОС РВ.

К одной шине можно подключить до 28 устройств, а несколько шин объединить с помощью повторителей. Возможные скорости передачи – 62,5 Кбит/с; 375 Кбит/с и 2,4 Мбит/с. При наименьшей скорости допустимое расстояние между повторителями составляет 1200 м.

Шина **Bitbus** иерархически структурирована: одно из присоединенных устройств является ведущим, остальные – ведомыми. Ведущее устройство всегда управляет процессом передачи сообщений с помощью механизма опроса: ведущий посылает запрос ведомому, который должен ответить; ведомое устройство не может передать запрос по своей инициативе, а должно ждать запроса ведущего. Ведомый отвечает ведущему в течение определенного интервала времени. Если ведомый не отвечает, ведущий повторяет запрос несколько раз. Если ответ не получен, ведущий переводит ведомого в разряд недоступных и виртуально выводит его из работы. Протокол **Bitbus** ориентирован на одного ведущего и поэтому не предусматривает способ передачи его прав.

Структура пакетов данных следующая:

1 поле	8 бит	Значение 01111110	Начальный флаг
2 поле	8 бит		Адрес
3 поле	8 бит		Управляющее поле
4 поле	Произвольная длина		Данные
5 поле	16 бит		Контрольная сумма
6 поле	8 бит	Значение 01111110	Заключительный флаг

Ограничивающие начальный и заключительный флаги кодируются уникальной последовательностью.

В сети Vitbus пакеты обрабатываются и пересылаются как межзадачные сообщения программ ведущего и ведомого устройств. Прикладной программист, однако, не работает непосредственно на уровне пакетов, поскольку функции взаимодействия с шиной реализованы в виде набора стандартных процедур программного обеспечения. Эти процедуры включают чтение и запись в память ведомого устройства из памяти ведущего, загрузку задач, запуск и остановка задач и т.д.

Шина Vitbus наиболее эффективна, когда необходимо придерживаться определенной последовательности операций и обращаться к простым ведомым устройствам. Она менее эффективно поддерживает асинхронные операции с более или менее независимыми, т.е. обладающими интеллектом, устройствами

Физическая среда шины **PROFIBUS** – это экранированная витая пара в соответствии со спецификацией интерфейса RS-485 с максимальной длиной 1200 м (до 4800 м с использованием повторителей). Скорости передачи данных – 9,6; 19,2; 187 и 500 Кбит/с. Тип передачи – полудуплексный (т.е. требуется переключение с приема на передачу), асинхронный (посылается запрос о готовности к передаче).

В новых конструкциях предусмотрено применение оптоволоконного кабеля и, соответственно, более высокие скорости.

Шина PROFIBUS может работать как с одним постоянным ведущим устройством, так и с несколькими ведущими. К шине можно подключить до 127 активных и пассивных станций. Сообщение может иметь максимальную длину 256 байт. Шина PROFIBUS позволяет осуществлять настройку различных параметров, которые делятся на параметры шины, одинаковые для всех станций, и параметры устройств, отличающиеся от станции к станции. Например, параметр шины «Время обращения маркера» характеризует максимальное время полного обращения по шине; он должен выбираться таким образом, чтобы активные станции не простаивали слишком долго в ожидании доступа к среде передачи.

В PROFIBUS доступ осуществляется в соответствии с архитектурой MASTER/SLAVE («ведущий/ведомый») с переменным ведущим (гибридный метод).

Взаимодействие между ведущими узлами (активными станциями – MASTER) осуществляется на основе передачи права доступа к среде передачи. Это право передается телеграммой – маркером, который циркулирует между всеми MASTER-узлами по логическому кольцу от узла с меньшим адресом к узлу с большим адресом в рамках определенного времени оборота маркера. Когда активная станция получает маркер, ей на определенное время дается право выполнять на шине функции ведущего устройства (MASTER). При этом MASTER периодически опрашивает все необходимые станции Slave. Получив запрос, станции выполняют какие-либо действия в зависимости от полученных данных и посылают ответ. Получив ответ, MASTER его обрабатывает и переходит к следующей станции.

Обмен сообщениями осуществляется в цикле, который состоит из телеграммы-вызова (передача или/и запрос) активного абонента и соответствующей телеграммы-подтверждения или телеграммы-ответа активного или пассивного абонента. Данные могут передаваться как в телеграмме-вызове (передача), так и в телеграмме-ответе (ответ). Телеграмма-подтверждение не содержит данных пользователя.

В PROFIBUS телеграммы могут быть различного вида: с фиксированной длиной информационного поля, с переменной длиной, без поля данных и др. Для примера рассмотрим формат телеграммы-вызова фиксированной длины с полем данных:

– биты синхронизации, минимально 33 бита состояния покоя (холостой ход);

- стартовый байт (стартовый разделитель);
- байт адреса пункта назначения;
- байт адреса источника;
- контрольный байт (контроль кадра);
- поля данных; фиксированная длина 8 байтов;
- проверочный байт (порядок контроля кадра);
- конечный байт (разделитель конца).

Формат телеграммы-ответа отличается отсутствием битов синхронизации. В телеграмме с переменной длиной информационного поля может содержаться до 246 байтов данных.

Важная задача для любого протокола – сохранение целостности передаваемых данных. Для этого используются специальные начальные и конечные разделители, жесткая синхронизация и бит контроля паритета (четности) для каждого байта.

PROFIBUS может работать как в режиме «точка-точка», так и в режиме широкого вещания. В последнем случае сообщение передается всем ведомым узлам, причем ведущий узел не получает никакого подтверждения.

С помощью шины PROFIBUS можно соединять в гибридную систему несколько независимых устройств и интеллектуальных датчиков.

Interbus использует процедуру доступа к шине по схеме «ведущий/ведомый» (Master/Slave). При этом Master обеспечивает одновременный интерфейс к высокоуровневой управляющей системе и выполняет функции управления шиной. Топология Interbus – это физическое и логическое кольцо, у которого физический уровень построен на основе стандарта RS-485. Это дифференциальный интерфейс, использующий витую пару для информационных передач. Для реализации кольца кабель использует две витые пары (для дуплексного режима) плюс дополнительный провод для передачи сигнала «логическая земля». Такая структура позволяет организовать сеть, работающую на скорости 500 Кбит/с на расстоянии 400 м между двумя соседними узлами сети. Включенная в каждое сетевое устройство функция повторителя сигнала позволяет расширить систему до 13 км. Общее число устройств сети ограничено – 512 узлов.

Преимущество протокола Interbus – небольшие накладные расходы на информационную телеграмму (небольшой заголовок). Поэтому при невысокой скорости передачи относительный объем передаваемых данных за единицу времени довольно высокий. Временные соотношения в Interbus хорошо детерминированы: общее время зависит от объема данных и хорошо просчитывается.

8.6. Локальные сети

8.6.1. Сетевые топологии

Развитие техники управления и появление сложных распределенных систем привело к дальнейшему развитию средств коммуникации. Для объединения нескольких устройств таким образом, чтобы каждый из них мог общаться с любым другим, применяют локальные сети (ЛВС). Они обычно строятся на базе кабельных соединений – витая пара, коаксиальный или оптоволоконный кабель. Современная распределенная система управления строится, как правило, на основе открытых промышленных стандартов. Открытые технологии для потребителя означают возможность компоновки информационной системы из модулей различных производителей, что позволяет выбрать наилучший вариант решения собственной задачи.

Наиболее важные сетевые топологии: шина – станции разделяют общий физический тракт передачи; звезда – центральный узел (концентратор) соединен непосредственно с каждым узлом по двухточечному принципу; кольцо – каждая станция соединена с двумя другими, а вместе они образуют кольцо; древовидная (иерархическая) архитектура – отдельные станции (концентраторы) соединены каскадно с другими концентраторами или оконечными станциями; ячеистая архитектура – между каждой парой узлов установлено соединение «точка-точка»; смешанная архитектура – одновременное использование нескольких технологий.

8.6.2. Метод доступа Ethernet

Этот метод поддерживает шинную топологию. На физическом уровне Ethernet работает как с узкополосным, так и с широкополосным кабелем. Чаще используется первый: данные передаются в канал методом манчестерского кодирования; модуляция несущей частоты не требуется. Допускается использование оптоволоконного кабеля и витой пары. Техническая скорость передачи варьируется от 10 Мбит/с (для коаксиального кабеля) до 1 Гбит/с (для оптоволокна).

Кадр начинается с 8-байтовой преамбулы 101010..., затем следует адрес получателя и адрес источника, каждый из которых имеет длину 6 байт. Адрес получателя может относиться к одному устройству (индивидуальный адрес), к определенному подмножеству устройств (групповой адрес) или ко всем абонентам (широковещательный адрес). Далее следует 2-байтовое поле, интерпретация которого зависит от реализации (длина поля данных или тип протокола верхнего уровня). Пользовательские данные начинаются после этой предварительной информации. Длина поля данных – от 46 до 1500 байт, к сообщению короче 46 байт добавляются нули. Кадр заканчивается 4-байтовой контрольной суммой.

Метод, использованный в Ethernet для координации доступа к среде, называется множественным доступом с прослушиванием несущей и обнаружением коллизий. Этот метод позволяет работать без выделенного устройства, управляющего доступом, и состоит в следующем. Станция, которая собирается начать передачу, должна проверить, что канал свободен – прослушать несущую. Если несущей нет, то можно начать передачу. Несколько станций могут начать передачу одновременно (множественный доступ); тогда возникнет коллизия – наложение сигналов от нескольких передатчиков. Поскольку все станции «слышат» все передачи, включая свою собственную, то передатчик сразу обнаруживает коллизию, так как данные, которые он «слышит» в канале, отличаются от исходных; затем он прекращает передачу текущего кадра и начинает посылать в канал специальную комбинацию бит, чтобы все станции обнаружили коллизию. Для начала новой передачи станция должна снова прослушивать несущую, при этом момент прослушивания определяется случайным образом для каждой передачи в рамках некоторого ограниченного периода. Если кадр не удастся передать в течение определенного времени, то передача прекращается, и информация об этом передается протоколу верхнего уровня.

8.6.3. Маркерная шина

В этой сети все абоненты присоединены к сети с помощью шины. Физической средой является коаксиальный кабель или витая

пара, поддерживающая скорость передачи данных в диапазоне 1...10 Мбит/с. В сети с маркерной шиной доступ к сети осуществляется строго детерминированным образом, т.е. только одна станция в данный момент может начать передачу сообщения. Право на передачу дается маркером – специальной битовой комбинацией, которая последовательно циркулирует между устройствами. Устройство, получившее маркер, имеет право на передачу в течение определенного интервала и затем должно переслать маркер следующему. Если у устройства нет данных для передачи, оно просто передает маркер дальше. Логически маркер последовательно перемещается по кольцу, хотя фактической топологией является шина.

Преимущество маркерной шины заключается в том, что все параметры процесса связи имеют детерминированный характер и время ожидания в наилучшем варианте можно точно определить.

8.6.4. Маркерное кольцо

Отличие этой сети от сети с маркерной шиной заключается в том, что здесь кольцо является не только логическим, но и физическим. Когда станция, имеющая сообщение для передачи, получает маркер, она удаляет его из кольца и в течение максимально разрешенного времени передает свои пакеты данных. Физическая среда передачи – экранированная витая пара со скоростью передачи 4 или 16 Мбит/с.

Каждая станция проверяет поле получателя в поступающих сообщениях. Если сообщение не предназначено для данной станции, оно передается дальше. Когда станция в кольце не включена или не готова к работе по другой причине, ее выходные и входные соединители к шине закорачиваются друг с другом через реле, для того чтобы обеспечить беспрепятственную циркуляцию сообщений по шине.

8.6.5. Локальная контроллерная сеть CAN

CAN (Controller Area Network) является последовательной сетью, которая эффективно поддерживает распределенное управле-

ние в реальном масштабе времени с высоким уровнем защиты данных. Физической средой для сети CAN служит двухпроводная шина с терминаторами на концах. Роль шины выполняет кабель, содержащий одну экранированную витую пару или две (вторая витая пара может использоваться для питания интерфейса CAN-контроллера); плоский кабель; оптоволокно; радио- и ИК-каналы. Основным ограничением протяженности шины является предельно допустимая задержка (при передаче на скорости 1 Мбит/с – 30 м, 10 кбит/с – 5000 м).

В сети CAN не применяется адресация абонентов: сообщение принимается всеми узлами. Каждое сообщение имеет свой идентификатор, стоящий в начале кадра и определяющий его содержание. Сообщение, передаваемое одним узлом, рассылается по сети всем станциям (широковещательная передача), а каждый узел, принявший сообщение, на основании полученного идентификатора решает, будет ли оно обработано или проигнорировано. Возможные коллизии (при одновременном запросе шины несколькими узлами), разрешаются по приоритетности сообщений. Она определяется поразрядным сравнением битов идентификаторов. Все проигравшие узлы автоматически станут приемниками и не сделают попытку передачи, пока шина не освободится.

Каждое CAN-сообщение содержит 0...8 байт данных. Большие блоки данных передаются благодаря принципу сегментации. Передача сообщений идет отдельными кадрами. Каждый кадр состоит из полей: арбитража, контроля, данных и протокола.

Поле арбитража включает в себя идентификатор, однозначно определяющий содержание и приоритет сообщения. Оно состоит из 11-ти или 29-ти разрядного идентификатора и бита RTR (Remote Transmission Request – запрос удаленной передачи). Идентификатор определяет приоритет данного сообщения. Сообщение, передаваемое узлом, рассылается по сети всем станциям, а каждый узел, принявший сообщение, на основании полученного идентификатора решает, будет ли оно обработано или проигнорировано. В сети каждый идентификатор уникален; не существует двух разных сообщений, которые бы имели одинаковые идентификаторы. Стандартным форматом сообщений предусмотрен 11-битный идентификатор.

Поле данных содержит передаваемые данные (не более 8 байт, количество байт указывается).

Протокольные поля следующие:

- стартовое поле содержит один бит;
- управляющее поле «Контроль» содержит 6 битов, из которых 4 бита показывают количество байтов данных, которые будут передаваться в поле данных; 2 бита – резерв для следующих редакций протокола;
- поле контроля (15 бит) обеспечивает контроль данных по четности;
- поле подтверждения сигнализирует передающему узлу, что передача прошла нормально и повтор не требуется;
- поле конца кадра содержит 7 битов.

Возможные коллизии, связанные с одновременным запросом шины несколькими узлами, разрешаются по приоритетности сообщений, определяемой поразрядным сравнением битов идентификаторов. Право на работу с шиной получает тот узел, который передает сообщение с наивысшим приоритетом. Все проигравшие узлы автоматически становятся приемниками и не делают попытку передачи, пока шина не освободится. Механизм арбитража реализован на аппаратном уровне.

На аппаратном уровне реализован контроль сообщений (используется пять механизмов проверки передаваемых сообщений). CAN-контроллер регистрирует ошибки и оценивает их статистическими методами для принятия соответствующих мер – вплоть до отключения от сети ошибающегося узла.

CAN-сеть широко используется в промышленности, в энергетике и т.п. ввиду высокой коммуникационной надежности, небольшой стоимости подключения каждого узла к сети, доступности CAN-контроллеров для нескольких производителей.

Таким образом, можно констатировать, что сформировался рынок промышленных сетей: существует много сетевых технологий и готовых изделий. Ведущие компании объединены в различные ассоциации по поддержке того или иного вида сети. Постоянно ведутся работы по стандартизации сетевых протоколов. При выборе коммуникационной технологии следует руководствоваться количе-

ственными параметрами: объемом передаваемых данных, максимальной длиной шины, допустимым числом узлов на шине, помехозащищенностью, стоимостью, а также эффективностью решения задачи, простотой конфигурации и т.д.

8.7. Канальный уровень передачи

На канальном уровне решаются задачи формирования кадров, защиты от ошибок и управления доступом к среде передачи. Для этих целей разработаны специальные протоколы.

Протокол высокоуровневого управления каналом передачи данных (High-Level Data Link Control – HDLC) разработан ISO и используется в качестве основы для других протоколов.

Управляющие поля в кадре HDLC имеют фиксированное расположение и длину; переменную длину имеет только поле данных. Ограничивающие начальный и заключительный флаги кодируются уникальной последовательностью 01111110. Для обеспечения уникальности такой последовательности передатчик автоматически вставляет 0 после пяти идущих подряд 1 во все поля, кроме флагов. Приемник удаляет любой 0, следующий непосредственно за пятью идущими подряд 1.

Адресное поле имеет смысл только в том случае, если несколько станций могут принимать одно и то же сообщение – например, в локальных сетях, где все станции разделяют общий физический тракт передачи.

Управляющее поле определяет тип кадра: информационный, супервизорный или нумерованный. Информационный кадр содержит пользовательские данные; его управляющее поле идентифицирует порядковый номер текущего и следующего кадров. Супервизорный кадр используется для координации действий приемника и передатчика и выполняет управляющие функции – подтверждение кадров (квитирование), запрос на повторную передачу информационного кадра, если нарушена последовательность кадров. Нумерованные кадры используются для инициализации и разъединения звена и для других целей управления.

Поле данных может иметь любую длину. Поле контрольной суммы вычисляется с помощью многочлена $x^{16}+x^{12}+x^5+1$.

Структура кадра HDLC соответствует семиуровневой структуре ВОС: начальный и заключительный флаги необходимы на физическом и канальном уровнях; адресное и управляющее поля и поле контрольной суммы относятся к 2-му и 3-му уровням; передаваемые данные – к 7-му.

Протокол телеметрии IEC-870 – широко распространенный стандарт для мониторинга и управления распределенными процессами. В протоколе телеметрии IEC-870 единицей передачи информации является телеграмма, состоящая из байтовых последовательностей по 8 бит. Каждый байт передается в виде 11-битовой последовательности (1 стартовый бит, 8 бит данных, 1 бит четности, 1 бит стоповый). Кадры передаются последовательно в каждой телеграмме в следующем порядке:

- заголовок (1 или 4 байта);
- пользовательские данные (переменная длина, до 253 байт);
- контрольная сумма (1 байт);
- символ конца сообщения (1 байт).

Заголовок указывает, какой тип телеграммы следует за ним – фиксированной или переменной длины. Телеграмма фиксированной длины имеет длину 6 байт. В телеграмме переменной длины заголовок содержит начальный байт, за которым следует длина телеграммы. В четырех байтах заголовка в качестве меры безопасности повторяются начальный символ и длина сообщения. В пользовательских данных два байта отведено для указания адреса, затем следуют данные.

Контрольные вопросы и упражнения

1. Расскажите о модели взаимодействия открытых систем.
2. Какие существуют три основные структуры сетевого обмена?
3. Нарисуйте амплитудную, фазовую и частотную манипуляцию.
4. Почему бод не всегда совпадает с битом?
5. Нарисуйте временное и частотное мультиплексирование.
6. Что такое «полевая шина»?

7. Расскажите о работе шины Bitbus.
8. Расскажите о работе шины Profibus.
9. Как работает метод множественного доступа с прослушиванием несущей?
10. Какой принцип доступа осуществляется в маркерной шине?
11. Как работает сеть с маркерным кольцом?
12. Расскажите о работе сети CAN.
13. Как работает шина USB?

9. ИСТОЧНИКИ ЭЛЕКТРОСНАБЖЕНИЯ И ХАРАКТЕРИСТИКА УСЛОВИЙ ЭКСПЛУАТАЦИИ

9.1. Общая организация системы электроснабжения

Современные системы управления требуют хорошо организованных средств электроснабжения, или, как обычно говорят, средств питания. Обычно применяется следующая схема. Через сети и подстанции, составляющие средства внешнего питания, электроэнергия передается по кабелю или по воздушной линии на вход трансформаторной подстанции объекта, где расположена система управления. Напряжение в этой сети обычно составляет 6,3; 10,5; 21; 38,5 кВ. На объекте организуется система внутреннего энергоснабжения, для чего на подстанции напряжение понижается до 380/220 В с частотой 50 Гц или 220 В с частотой 427 Гц. Выход этой системы рассматривается как *первичный* источник питания.

Для питания системы управления средства энергоснабжения должны удовлетворять определенным требованиям. Система питания должна обеспечивать: выработку требуемых напряжений заданной мощности с заданной стабильностью (обычно 2...3%); непрерывность энергоснабжения при авариях; быстрое восстановление номинальных уровней при импульсных изменениях нагрузки; нечувствительность к помехам в системах питания и к внешним помехам.

Для обеспечения этих требований в системе питания должны быть распределительные устройства, вторичные источники питания, токопроводящие шины, схема защиты и устройство управления.

Распределительное устройство принимает питание от первичного источника и распределяет его между отдельными приборами системы управления, осуществляя в случае необходимости резервирование первичного источника.

Вторичные источники питания вырабатывают постоянное питание заданного напряжения, стабильности и мощности.

Схемы защиты формируют сигналы для автоматического отключения питания от нагрузки, если: отсутствует один из номина-

лов напряжения, установленных для питания схем; ток, потребляемый устройством, превысил допустимый уровень (обычно на 20...30%); превышена температура вторичного источника; повышено или понижено напряжение питания.

Устройство управления предназначено для реализации определенной последовательности включения и отключения номиналов напряжения, а также для обеспечения различных режимов работы устройства питания (профилактика системы питания, штатный режим системы питания, профилактика системы управления).

9.2. Вторичные источники электроснабжения

Любая электронная схема требует для своей работы одного или нескольких стабильных источников питания постоянного тока; в стабильных напряжениях нуждаются также и электромеханические устройства, входящие в состав устройств ввода/вывода, накопителей и др. Они должны обеспечивать стабильность токов и напряжений, достаточную мощность и удовлетворять другим эксплуатационным требованиям.

Вторичные источники питания постоянного тока обычно состоят из трансформатора Т, выпрямителя В, фильтра Ф и стабилизатора Ст (рис. 9.1, а).

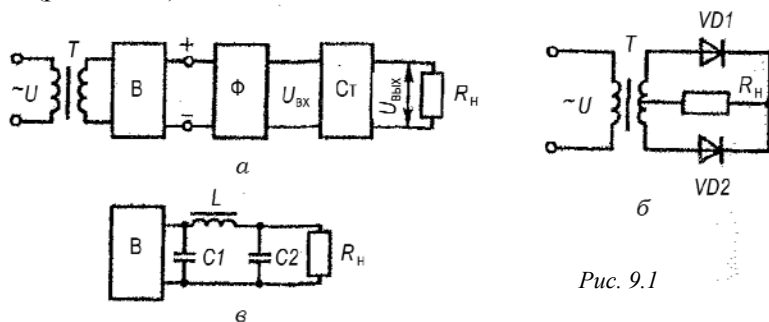


Рис. 9.1

Для преобразования переменного тока в постоянный используются, как правило, двухполупериодные выпрямители (рис. 9.1, б); фильтр сглаживает пульсации на выходе выпрямителя, к которому подключается нагрузка, представленная в виде нагрузочного со-

противления R_n . Наиболее эффективными являются LC -фильтры (рис. 9.1, в). Для стабилизации выпрямленного напряжения используют параметрические и компенсационные стабилизаторы.

В **параметрических** стабилизаторах стабилизирующее действие основано на наличии нелинейного элемента в схеме. Обычно это полупроводниковый стабилитрон, схема включения которого (а) и его нелинейная рабочая характеристика (б) показаны на рис. 9.2. Видно, что изменение потребляемого тока в достаточно большом диапазоне ΔI_n приводит лишь к малому изменению выходного напряжения ΔU . Диод VD2, включенный последовательно со стабилитроном VD1, служит для компенсации температурной неустойчивости схемы; балластное сопротивление $R_{бал}$ задает ток через стабилитрон.

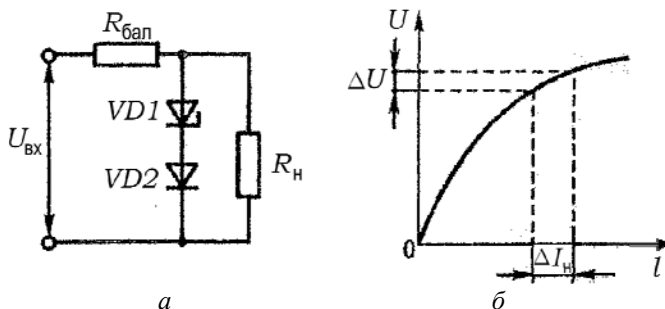


Рис. 9.2

Стабилитроны выпускаются на целый ряд значений напряжения – от двух до 200 В с допустимой мощностью рассеяния от долей ватта до 50 Вт и допуском на напряжение стабилизации от 1 до 20 %.

Компенсационные стабилизаторы – это системы регулирования по отклонению. Они состоят из источника эталонной величины, схемы сравнения, усилительных и регулирующих устройств. Среди компенсационных стабилизаторов широко применяются (особенно в системах электроавтоматики) *линейные* транзисторные стабилизаторы, в которых регулирующий транзистор работает в линейном режиме и представляет собой переменное сопротивление между источником питания и нагрузкой, и *импульсные*, в которых

регулирующий транзистор работает в ключевом режиме с определенной скважностью, определяемой устройствами обратной связи и параметрами схемы стабилизатора.

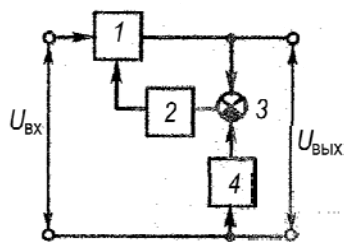


Рис. 9.3

Структурная схема линейного стабилизатора приведена на рис. 9.3. Регулирующий элемент 1 включается последовательно с нагрузкой на выходе. Выходное напряжение сравнивается схемой сравнения 3 с напряжением эталонного источника 4. Отклонение напряжения от заданного значения с выхода элемента 3 воздействует через

усилитель 2 на регулирующий элемент 1. Выходное напряжение всегда ниже нестабилизированного входного напряжения и на управляющем элементе рассеивается некоторая мощность.

Принцип работы импульсного стабилизатора иллюстрирует рис. 9.4. Транзистор $VT1$, работающий в режиме насыщенного ключа, периодически прикладывает к катушке индуктивности L полное нестабилизированное напряжение. При каждом импульсе в катушке возникает ток и в ее магнитном поле запасается энергия, которая поступает на конденсатор C сглаживающего фильтра и поддерживает напряжение и ток в нагрузке между импульсами заряда. Как и в линейных стабилизаторах, выход по цепи обратной связи сравнивается с эталонным напряжением $U_{эт}$, однако управление выходом осуществляется путем изменения длительности импульсов генератора или частоты переключения.

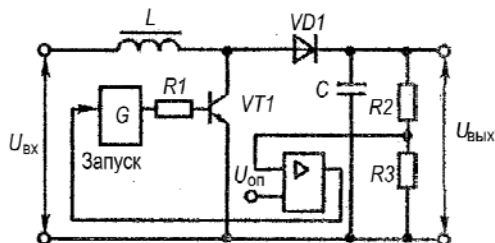


Рис. 9.4

Импульсные стабилизаторы получили широкое распространение, в основном, из-за малой мощности рассеяния: управляющий

элемент в них либо выключен, либо обладает низким сопротивлением (во включенном состоянии – сопротивление транзистора в режиме насыщения).

Еще одно достоинство импульсного стабилизатора заключается в том, что его можно подключить непосредственно к выпрямленному и отфильтрованному напряжению переменного тока, исключив понижающий сетевой трансформатор. Генератор Γ может питаться от нестабилизированного постоянного входного напряжения $U_{вх}$; схемы управления обратной связью (усилитель ошибки, эталонный источник) – от стабилизированного выходного напряжения или от вспомогательного источника. Поскольку отсутствие трансформатора означает, что вход постоянного тока не изолирован от сети, в схеме должны быть предусмотрены соответствующие развязки: их можно реализовать с помощью трансформатора. Наличие трансформаторов в изолирующих цепях вовсе не сводят на нет упомянутое преимущество импульсных стабилизаторов. Дело в том, что размеры трансформаторов определяются размерами сердечника, а они резко снижаются на высоких частотах. В результате импульсные источники с питанием от сети намного меньше и легче, чем аналогичные линейные источники. Кроме того, они работают в более благоприятном тепловом режиме.

9.3. Характеристика условий эксплуатации

В зависимости от особенностей применения системы подразделяются на стационарные, наземные, автомобильные, судовые, авиационные, космические и т.д. Условия эксплуатации будут различными для каждого из перечисленных видов. Например, системы стационарного типа не будут испытывать механических нагрузок; могут существенно отличаться при этом и климатические условия.

В зависимости от временного режима использования различают системы разового действия, дежурные системы и системы непрерывного использования. Системы разового действия используются по целевому назначению только один раз; режим работы дежурных систем включает в себя период ожидания и период использования по назначению (например, система слепой посадки самолетов);

системы непрерывного действия используются по назначению непрерывно в течение всего заданного срока эксплуатации.

Системы реального времени в зависимости от режима эксплуатации подразделяются обслуживаемые, когда в процессе эксплуатации возможно проведение профилактических и ремонтных работ, и необслуживаемые.

Системы автоматизированного управления эксплуатируются в условиях воздействия на них различных факторов, из которых можно выделить две группы: объективные, определяемые внешней средой и условиями эксплуатации, и субъективные, определяемые обслуживанием системы.

Объективные факторы. Они подразделяются на внешние и внутренние.

Среди *внешних* факторов воздействия прежде всего следует выделить климатические факторы.

Температура окружающего воздуха $+20$ °С принимается нормальной. Пределы ее изменения от -87 до $+57,5$ °С. Существенно изменяется она и увеличением высоты: на высоте 20 км она составляет около -50 °С.

Влажность воздуха на уровне земли колеблется от $0,1$ г/м³ в полярных районах до 30 г/м³ в тропиках. Обычно влажность воздуха выражают в относительных единицах, при этом нормальная относительная влажность равна 65%; во влажных тропиках относительная влажность достигает 98%.

Нормальное атмосферное *давление* равно 760 мм рт. ст. ($1,01 \cdot 10^5$ Па). Вблизи поверхности Земли зафиксированы его изменения от 684 мм рт. ст. ($0,91 \cdot 10^5$ Па) до 807,7 мм рт. ст. ($1,08 \cdot 10^5$ Па).

Ветровые нагрузки создаются движением воздушной среды и изменяются в широких пределах; у поверхности Земли скорость ветра изменяется от 0 до 200 км/ч.

Вода, выпадающая в виде атмосферных *осадков*, содержит неорганические и органические частицы. В приморских зонах особенно характерны примеси хлористого натрия, а в тропических – повышенное содержание азотной кислоты. Снег содержит больше азотистых соединений, чем дождь.

Опасны для работоспособности систем *пыль* и *песок*. Попадая в подвижные части, они вызывают повреждения. Кроме того, пыль способствует увеличению электростатических зарядов, что приводит к росту помех, а в отдельных случаях к взрывам.

К наиболее характерным факторам воздействия *биологической среды* на конструкции систем относятся грибковые образования (плесень), особенно интенсивно развивающиеся при повышенной влажности неподвижного воздуха (более 85%) и температуре от 20 до 30 °С. Некоторые виды насекомых, например, термиты, пожирают органические материалы, особенно изоляционные; в этом отношении опасны и грызуны.

Значительное влияние на работу систем оказывает *радиоактивное излучение*. Это влияние особенно сильно проявляется в воздействиях на материалы кристаллической структуры, воздух, изоляцию, стекло и электролиты. Радиация ионизирует воздух, уменьшает проводимость между точками монтажа и может нарушить нормальную работу систем.

Механические воздействия – ускорения, вибрации и удары могут действовать как отдельно, так и в совокупности. Амплитуда вибраций может составлять до нескольких десятков миллиметров на частотах от 5 до 150 Гц, ускорения – до 30g.

К *внутренним* факторам относятся старение, износ, перегрев и паразитные связи.

Субъективные факторы – это факторы, связанные с особенностями обслуживания: условия хранения и эксплуатации, квалификация обслуживающего персонала, уровень профилактики и ремонта и условия работы операторов.

Технической основой специализированных ЭВМ являются промышленные компьютеры – изделия с повышенной устойчивостью к воздействиям внешней среды. Для борьбы с пылью создают избыточное давление внутри системного блока, а всасывающие вентиляторы снабжают сменными пылеулавливающими фильтрами. В моделях для пультового монтажа применяется брызгозащита по лицевой панели. С вибрациями и ударами борются, применяя прочные шасси и корпуса, закрепляя платы расширения дополнитель-

ными амортизирующими скобами и размещая дисковые накопители на специальной виброударостойкой подвеске.

В клавиатурах предусматривается пылевлагозащита; ответственные кнопки типа «сброс» и «выключение питания» прячут за запираемыми на ключ дверцами.

Контрольные вопросы и упражнения

1. Нарисуйте структурную схему системы питания. Повторите назначение отдельных блоков.
2. Повторите требования к системе питания.
3. Почему опасно повышение и понижение напряжения питания относительно номинальных значений?
4. Почему габариты и вес импульсных стабилизаторов оказываются меньше, чем у обычных?
5. Как подразделяются системы реального времени в зависимости от режима эксплуатации?
6. Почему старение, износ, перегрев и паразитные связи относят к внутренним факторам?
7. Как изменяются условия эксплуатации для наземных, судовых, авиационных и космических систем?
8. Как изменяются требования, предъявляемые к системам разового действия, дежурным системам и системам непрерывного использования?

10. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ СИСТЕМ РЕАЛЬНОГО ВРЕМЕНИ

10.1. Операционные системы

Операционные системы реального времени (ОС РВ) характеризуются тем, что значение времени принятия решения используется как одна из переменных, от которых зависит результат. То есть при опоздании результатов они либо могут быть бесполезными, либо ущерб в результате опоздания может быть бесконечно велик. Основная задача операционных систем общего назначения – эффективное разделение ресурсов ЭВМ между несколькими одновременно выполняющимися программами. Операционные системы реального времени разрабатываются в расчете на наличие внешних источников данных. Основная задача ОС РВ – своевременно обработать запрос; все остальные аспекты функционирования ЭВМ отходят на второй план.

Напомним, что существуют системы «мягкого» и «жесткого» реального времени. Поэтому иногда говорят, что та или иная ОС «мягкая» или «жесткая». Нет «мягких» и «жестких» ОС. Одна и та же ОС РВ может лишь служить основой для построения таких систем.

По своей внутренней структуре ОС можно условно разделить на монолитные ОС, ОС на основе микроядра (модульные) и объектно-ориентированные ОС.

В **монолитных ОС** (рис. 10.1) все обрабатываемые программы сосредоточены вместе в едином процессе, которое исполняется в едином адресном пространстве.

Монолитные архитектуры можно представить в виде прикладного уровня, состоящего из работающих прикладных процессов, и системного уровня, состоящего из монолитного ядра. В нем можно выделить

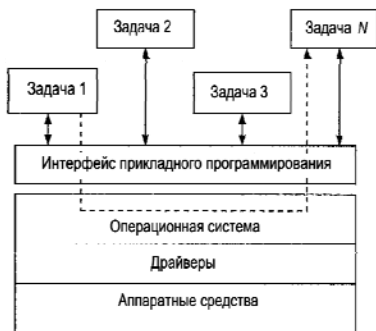


Рис. 10.1

следующие части: интерфейс между приложениями и ядром, собственно ядро и интерфейс между ядром и оборудованием (драйверы устройств). Интерфейс в таких системах играет двойную роль:

- управление взаимодействием функциональных задач с операционной системой;
- обеспечение непрерывности выполнения программ операционной системы.

Основное преимущество монолитной архитектуры – относительная быстрота работы по сравнению с другими архитектурами за счет написания значительных частей системы на языке ассемблера.

Недостатки монолитной архитектуры:

- системные вызовы, требующие переключения уровней привилегий (от пользовательской задачи к ядру), должны быть реализованы как прерывания или специальный вид исключения, что сильно увеличивает время их работы;
- ядро не может быть прервано пользовательской задачей; это может приводить к тому, что высокоприоритетная задача может не получить управления из-за работы низкоприоритетной;
- сложность переноса системы на новые архитектуры процессора из-за значительных ассемблерных вставок;
- недостаточная гибкость и сложность развития: изменение части ядра требует его полной перекомпиляции.

Поскольку в такой системе невозможно выполнение системных вызовов в то время, когда реализуется какой-либо поток ядра, режим реального времени невозможен.

Модульная (микроядерная) архитектура (рис. 10.2) является альтернативой классическому способу построения ОС. Она была разработана с целью убрать интерфейс между приложениями и ядром и облегчить модернизацию системы и перенос ее на новые процессоры. В микроядерных ОС в привилегированном режиме остается работать только очень небольшая часть ОС, называемая *микроядром*. При такой структуре микроядро выполняет две основные функции:

- управление взаимодействием частей системы (например, менеджеров процессов и файлов);

- обеспечение непрерывности выполнения ядра ОС (то есть отсутствие переключения задач во время выполнения микроядра).

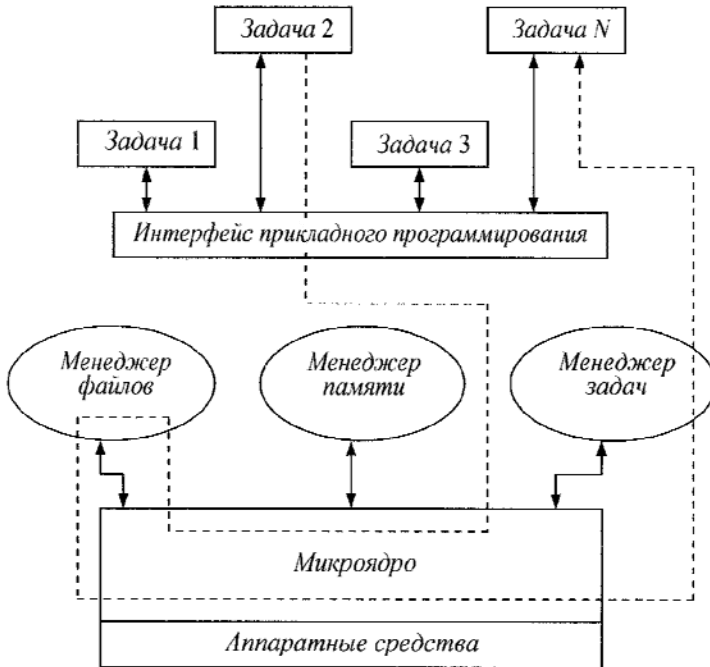


Рис. 10.2

Микроядерные ОС удовлетворяют большинству требований, предъявляемых к современным ОС, обладая *переносимостью* (легкостью переноса с одной аппаратной платформы на другую), *расширяемостью* (добавление и изменение функций), *надежностью* (защита серверов ОС друг от друга) и создавая хорошие предпосылки для поддержания распределенных приложений. За эти достоинства приходится платить снижением производительности, что является основным недостатком микроядерной архитектуры.

В микроядерной архитектуре ядро отвечает только за базовые примитивы ОС (создание и уничтожение потоков, диспетчеризация и синхронизация потоков, поддержка механизма прерываний, часов и таймеров). Это действительно микроядро, так как его объем, например, в системе QNX, всего 12 Кбайт. Все остальные компо-

ненты системы: драйверы, файловые системы, стеки протоколов, пользовательские приложения – выполняются вне пределов ядра как отдельные процессы, каждый в своем защищенном адресном пространстве. Программы, реализующие эти сервисные функции, называются *администраторами ресурсов*. Все эти компоненты используют для общения друг с другом единый, четко детерминированный механизм – обмен сообщениями. Он образует между компонентами системы виртуальную «программную шину», позволяющую подключать к ней или, наоборот, отключать любой компонент на лету. Связь между микроядром и администраторами ресурсов осуществляется с помощью *администратора процессов*.

Недостатки модульной структуры фактически те же, что и у монолитной. Проблемы перешли с уровня интерфейса на уровень микроядра. Системный интерфейс по-прежнему не допускает переключения задач во время работы микроядра, только сократилось время пребывания в этом состоянии, проблемы с переносимостью ядра уменьшились (в связи с сокращением его размера), но остались.

В объектно-ориентированной ОС (рис. 10.3) все системные процессы распределены по отдельным администраторам, которые включаются менеджером прерываний и менеджером задач.

Объектная архитектура строится на основе объектов-микроядер. В ней интерфейс между приложениями и ядром отсутствует вообще. Взаимодействие между компонентами системы (микроядрами) и пользовательскими процессами осуществляется посредством обычного вызова функций, поскольку и система, и приложения написаны на одном языке. Это обеспечивает максимальную скорость системных вызовов. Фактическое равноправие всех компонент системы обеспечивает возможность переключения задач в любое время. Объектно-ориентированный подход обеспечивает модульность, безопасность, легкость модернизации и повторного использования кода.

В отличие от предыдущих систем, не все компоненты самой операционной системы должны быть загружены в оперативную память. Если микроядро уже загружено для другого приложения, то оно повторно не загружается, а используется код и данные уже

имеющегося микроядра. Поскольку разные приложения разделяют одинаковые микроядра, то они должны работать в одном адресном пространстве. Следовательно, система не может использовать виртуальную память и поэтому работает быстрее, так как исключаются задержки на трансляцию виртуального адреса в физический.

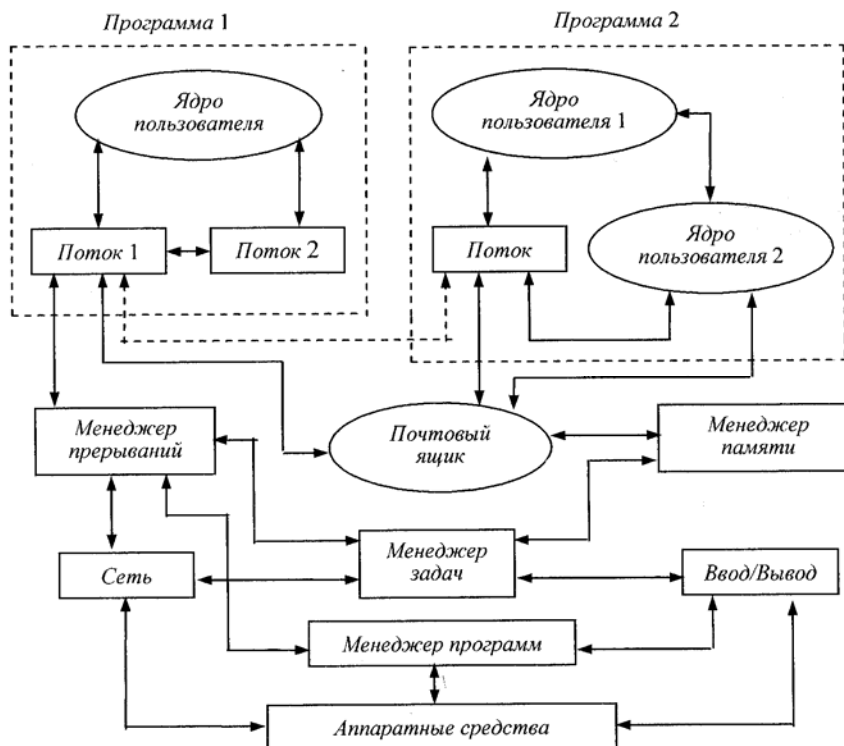


Рис. 10.3

При описании работы программного обеспечения важное значение имеет понятие задачи. *Задача* – это набор операций, предназначенный для выполнения логически законченной функции системы. Задача конкурирует с другими задачами за получение контроля над ресурсами вычислительной системы. Различают две разновидности задач: процессы и потоки.

Процесс – это отдельный загружаемый программный модуль (файл), который во время исполнения имеет в памяти свои незави-

симые области для кода и данных. Процесс – это выполняемая программа. Процесс включает код и данные программы, а также различную дополнительную информацию – переменные системного окружения и т.п.

Поток – это отдельный загружаемый программный модуль (файл), который может использоваться общими участками кода и данных в рамках единого программного модуля. Поток – это фрагмент процесса, содержащий непрерывную последовательность команд, которые могут выполняться параллельно с другими потоками того же и других процессов.

Процесс является, по сути дела, контейнером потоков и содержит как минимум один поток.

К преимуществам потоков можно отнести:

- экономию ресурсов как внешней, так и внутренней памяти, так как множество потоков может размещаться внутри одного EXE-модуля;

- использование потоками общей области памяти, что позволяет эффективно организовать межзадачный обмен сообщениями (достаточно передать указатель на сообщение). Процессы не имеют общей области памяти, поэтому ОС должна либо целиком скопировать сообщение из области памяти одной задачи в область памяти другой (что для больших сообщений весьма накладно), либо предусмотреть специальные механизмы, которые позволили бы одной задаче получить доступ к сообщению из области памяти другой задачи;

- меньше время переключения между задачами-потоками, чем между задачами-процессами, так как, как правило, контекст потоков меньше, чем контекст процессов;

- значительное упрощение использования программ-отладчиков, так как все потоки, а иногда и само ядро ОС РВ размещаются в одном EXE-модуле.

К недостаткам потоков можно отнести то, что:

- как правило, потоки не могут быть подгружены динамически. Чтобы добавить новый поток, необходимо провести соответствующие изменения в исходных текстах и перекомпилировать приложение. Процессы, в отличие от потоков, подгружаемы, что по-

зволяет динамически изменять функции системы в процессе ее работы. Кроме того, так как процессам соответствуют отдельные программные модули, они могут быть разработаны различными компаниями, чем достигается дополнительная гибкость и возможность использования ранее наработанного программного обеспечения;

- потоки имеют доступ к областям данных друг друга, это может привести к ситуации, когда некорректно работающий поток способен испортить данные другого потока. В отличие от этого процессы защищены от взаимного влияния, а попытка записи в «не свою» память приводит, как правило, к возникновению специального прерывания по обработке «исключительных» ситуаций.

10.2. Основные свойства задач

Как правило, вся важная с точки зрения ОС информация о задаче (имя и номер задачи, верхняя и нижняя границы стека, ссылка на очередь сообщений, статус задачи, приоритет и т.п.) хранится в унифицированной структуре данных – управляющем блоке (Task Control Block – TCB).

Приоритет – целое число, присваиваемое задаче и характеризующее ее важность по сравнению с другими задачами, выполняемыми в системе. Приоритет используется, в основном, планировщиком задач для определения того, какая из готовых к работе задач должна получить управление. Различают системы с динамической и статической приоритетностью. В первом случае приоритет задач меняется в процессе исполнения, во втором – жестко задается на этапе разработки или во время начального конфигурирования системы.

Контекст задачи – набор данных, содержащий всю необходимую информацию для возобновления выполнения задачи с того места, где она была ранее прервана. Часто контекст хранится в TCB и включает в себя такие данные как содержимое счетчика команд, указатель стека, содержимое регистров процессора и т.п. Планировщик задач в случае необходимости сохраняет контекст текущей активной задачи и восстанавливает контекст задачи, назначенной к

исполнению. Такое переключение контекстов и является, по существу, основным механизмом ОС РВ при переходе от выполнения одной задачи к выполнению другой.

Состояние (статус) задачи – режим использования задачи: *активная* задача – задача, выполняемая системой в данный момент времени; *готовая* задача – задача, готовая к выполнению и ожидающая у планировщика своей очереди; *блокированная* задача – задача, выполнение которой приостановлено до наступления определенных событий (освобождение ресурса, поступление ожидаемого сообщения, завершение интервала ожидания и т.п.).

Пустая задача – задача, запускаемая самой ОС, когда в системе нет готовых для выполнения задач. Пустая задача запускается с самым низким приоритетом и, как правило, представляет собой бесконечный цикл «ничего не делать». Наличие пустой задачи представляет ОС удобный механизм отработки ситуаций, когда нет ни одной готовой к выполнению задач.

Как правило, многозадачные ОС позволяют запускать несколько копий одной и той же задачи. При этом для каждой такой копии создается свой ТСВ и выделяется своя область памяти. В целях экономии памяти может быть предусмотрено совместное использование одного и того же исполняемого кода для всех запущенных копий. В этом случае программа должна обеспечивать повторную входимость (реентерабельность). Кроме того, программа не должна использовать временные файлы с фиксированными именами и должна корректно осуществлять доступ к глобальным ресурсам.

Повторная входимость (реентерабельность) позволяет без негативных последствий временно прервать выполнение какой-либо функции или подпрограммы, а потом вызвать эту функцию или подпрограмму снова. Частным случаем реентерабельности является рекурсия, когда тело подпрограммы содержит вызов самой себя. Классическим примером нереентерабельности является DOS, а типичной причиной нереентерабельности служит использование глобальных переменных.

Предположим, что у нас есть функция, реализующая запись на диск, и пусть она использует глобальную переменную WRITE_SECTOR, которая устанавливается в соответствии с пара-

метром, передаваемым этой функции при вызове. Предположим, что задача *A* вызывает эту функцию с параметром 3, т.е. хочет записать данные в сектор 3. Допустим, что когда переменная WRITE_SECTOR уже равна 3, но сама запись еще не произведена, выполнение задачи *A* прерывается и начинает выполняться задача *B*, которая вызывает ту же функцию, но с аргументом 10. После того, как запись в сектор 10 будет произведена, управление рано или поздно вернется к задаче *A*, которая продолжит работу с прерванного места. Однако теперь переменная WRITE_SECTOR равна 10, и все данные задачи *A* будут записаны в сектор 10. Из приведенного примера видно, что ошибки, связанные с нереентерабельностью, трудно обнаружить, а последствия они могут вызвать самые катастрофические.

10.3. Диспетчеризация задач и приоритеты

Наиболее важную роль в организации мультипрограммного режима работы управляющей ЭВМ играет центральная программа-диспетчер ОС. Эта программа определяет последовательность удовлетворения заявок на включение других основных программ ОС и программ решения функциональных задач и включает в счёт эти программы.

Дисциплина обслуживания заявок определяет порядок выполнения тех или иных программ. Хотя при отсутствии потерь заявок выбор той или иной дисциплины не влияет на уровень суммарной загрузки ЭВМ, от этой дисциплины зависит величина задержки информации в ЭВМ. Правильный выбор соответствующей дисциплины обслуживания заявок позволяет перераспределять время ожидания в очереди между заявками различных типов, то есть уменьшать величину задержки наиболее важной информации за счёт увеличения времени ожидания заявок на решение менее важных задач. Более того, показано, что в ряде случаев использование определённой дисциплины обслуживания сокращает суммарное время ожидания всех заявок в ЭВМ. Таким образом, центральный диспетчер (ЦД) путём рациональной организации вычислительного процесса повышает эффективность использования производительности управляющей ЭВМ.

Наиболее распространёнными дисциплинами обслуживания заявок в алгоритмах являются:

- кольцевая дисциплина обслуживания заявок;
- обслуживание заявок в порядке их поступления;
- приоритетное обслуживание без прерывания вычислений (относительные приоритеты);
- приоритетное обслуживание с прерыванием вычислений (абсолютные приоритеты).

Перечисленные дисциплины могут также иметь некоторые разновидности, отличающиеся режимом обслуживания заявок каждого типа:

- одиночное обслуживание (обслуживается только одна заявка данного типа);
- обслуживание с ограничением по времени или по числу последовательно обслуженных заявок данного типа.

Кроме того, приоритет программ может быть как фиксированным, так и динамическим. Динамический приоритет изменяется в процессе работы в зависимости от времени ожидания или длительности обслуживания заявок. Часто применяется также принцип группового приоритета, когда один и тот же приоритет обслуживания назначается одновременно нескольким типам заявок, порядок обслуживания которых определяется внутренней подпрограммой-диспетчером данной группы подпрограмм.

В большинстве случаев диспетчер реализует сочетание нескольких упомянутых дисциплин обслуживания. Это позволяет рационально использовать достоинства каждой из дисциплин.

Так как обращения к ЦД производятся после завершения каждой основной программы или группы подпрограмм, то есть весьма часто, то одним из основных требований к диспетчеру является требование экономичности по затратам производительности ЭВМ. Во многих случаях также требуется, чтобы состав, размещение и очерёдность работы основных программ вычислительной машины не были жёстко фиксированными, то есть диспетчер должен иметь элементы настройки на конкретный состав используемых основных подпрограмм. Рассмотрим некоторые варианты программы-диспетчера ОС ЭВМ, удовлетворяющие этим требованиям.

Программа-диспетчер с кольцевым обслуживанием заявок.

Эта программа является простейшей и поэтому наиболее распространённой. Она применяется при обработке входной информации ЭВМ, поступающей в буферные зоны от однотипных источников, при организации обмена с аппаратурой передачи данных, а также при включении в счёт равноценных основных подпрограмм.

Для реализации кольцевой дисциплины обслуживания заявок в ОП ЭВМ выделяются две зоны, одна из которых (зона заявок) используется для хранения кодов заявок Z_j на включение основных программ, а вторая (зона адресов) – для хранения начальных адресов A_j^H этих программ (рис. 10.4).

Каждой основной подпрограмме в зоне заявок и в зоне адресов выделяется определённая ячейка памяти, соответствующая номеру данной программы в системе диспетчеризации.

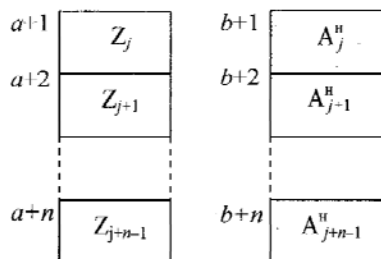


Рис. 10.4

Таким образом, максимальное количество включаемых данным способом программ определяется выбранными размерами зон заявок и адресов.

При необходимости вызова для работы одной из основных программ в процессе функционирования ЭВМ, осуществляется запись в закреплённую за этой программой ячейку зоны заявок определённого кода Z_j , называемого *безусловной* заявкой. Программа-диспетчер, последовательно анализируя содержимое ячеек зоны заявок, находит код безусловной заявки, стирает его и включает в счёт соответствующую основную программу, используя для этой цели её начальный адрес A_j^H , хранящийся в зоне адресов. В отдельных случаях необходимо, чтобы включение определённой программы происходило лишь в тех случаях, когда имеет место совпадение требований на её работу от нескольких других подпрограмм. В этом случае последние записывают в закреплённую за вызываемой программой ячейку зоны заявок условные заявки, ко-

ды которых подбираются таким образом, что совпадение необходимого количества условных заявок образует безусловную заявку. Сущность циклического обслуживания заявок заключается в том, что диспетчер, включившись в счёт после выполнения некоторой j -й основной подпрограммы, дальнейший анализ зоны заявок начинается с последующей $(j + 1)$ -й ячейки этой зоны, если $j + 1 \leq n$, или же с начала зоны, если $j + 1 > n$, где n – общее количество ячеек в зоне заявок.

Каждая задача, представляющая собой отдельную подпрограмму, выполняется циклически. При этом надо придерживаться следующих правил:

- подпрограммы не должны содержать циклов ожидания;
- подпрограммы должны выполнять свою работу как можно быстрее, чтобы дать возможность работать следующей подпрограмме;
- при необходимости подпрограмма должна сохранять свое окружение и текущие результаты, чтобы в следующем цикле возобновить работу.

Можно отметить следующие преимущества циклического алгоритма:

- простота использования и прозрачность для понимания;
- если исключить из рассмотрения прерывания, система полностью детерминирована. Задачи всегда вызываются в одной и той же последовательности, что позволяет достаточно просто произвести анализ «наихудшего» случая и вычислить максимальную задержку;
- минимальные размеры кода и данных;
- отсутствуют ошибки, обусловленные «гонками».

К недостаткам циклического алгоритма можно отнести отсутствие приоритетности и очередей. К тому же задачи вызываются независимо от того, должны ли они в данный момент что-либо делать или нет, а на прикладного программиста ложится максимальная ответственность за работоспособность системы.

Иногда (например, при обработке информации, поступающей в ЭВМ от аппаратуры передачи данных) одна и та же основная программа может обслужить не одну, а несколько заявок, количество которых учитывается в некоторой зоне ОП ЭВМ. В этих случаях

может устанавливаться некоторое ограничение на количество заявок m_j , обслуживаемых каждой программой при одном включении. Это ограничение может учитываться либо подпрограммами ОС, выполняющими внутреннюю диспетчеризацию процесса обработки данных определённого типа, либо собственно программой, осуществляющей эту обработку. Описанные выше ограничения имеют целью лимитировать длительность работы основных программ в определённых пределах с тем, чтобы не допускать длительных задержек в обслуживании заявок на решение других задач. При этом предполагается, что средняя продолжительность обслуживания одной заявки заранее известна.

В некоторых случаях ограничение может накладываться непосредственно на длительность работы каждой включаемой в счёт программы τ_j . Этот способ, предложенный Ф. Корбатом, следует применять в тех случаях, когда длительность работы основных программ заранее неизвестна или же может колебаться в значительных пределах. При реализации данной разновидности дисциплины диспетчеризации выполнение программы, превысившей заданный лимит времени τ_j , прерывается сигналом счётчика относительного времени. Одновременно с этим в зону заявок записывается признак прерывания P_j и заявка на продолжение выполнения прерванной программы. В соответствии с кольцевой дисциплиной обслуживания заявок эта заявка будет обслужена диспетчером в последнюю очередь после анализа всех других ячеек зоны заявок.

Программа-диспетчер с обслуживанием заявок в порядке поступления. Эта программа требует для своей реализации учёта последовательности поступления заявок на включение основных программ. Она может быть реализована двумя способами.

При первом способе заявки на включение основных программ записываются в порядке поступления в специальную буферную зону в виде номера вызываемой программы. Диспетчер последовательно анализирует содержимое буферной зоны заявок и включает в счёт вызванные подпрограммы, выбирая необходимый начальный адрес из зоны адресов по номеру подпрограммы i , указанному в зоне заявок. Перед включением подпрограммы этот номер стирается программой-диспетчером. При применении данного способа

диспетчеризации затрудняется использование принципа условных заявок.

$a+1$	B_j	t_j
$a+2$	B_{j+1}	t_{j+1}
$a+3$	B_{j+2}	t_{j+2}
$a+4$	B_{j+3}	t_{j+3}
$a+5$	t^{\max}	
$a+6$	B_{j+5}	t_{j+5}
$a+7$	t^{\max}	
...		
$a+n-1$	t^{\max}	
$a+n$	B_{j+n-1}	t_{j+n-1}

Рис. 10.5

При втором способе (рис. 10.5) заявки на включение каждой основной программы записываются в одну определённую ячейку зоны заявок, как и при кольцевой системе обслуживания. Однако в данном случае кодом заявки является показание СчРВ ЭВМ t_j в момент поступления данной заявки. Стирание заявки осуществляется путем записи вместо данной заявки кода t^{\max} , превышающего максимальное показание СчРВ.

Реализацию принципа условных заявок можно в данном случае осуществить при помощи признаков блокировки включения той или иной программы B_j , которые размещаются в старших разрядах ячейки зоны заявок

и в совокупности представляют собой позиционный код с числом разрядов, зависящим от количества условий включения данной подпрограммы. При записи условных заявок к текущему показанию СчРВ ЭВМ приформировываются соответствующие признаки блокировки, которые рассматриваются программой-диспетчером как старшие разряды этого счётчика. Поэтому наличие любого признака блокировки предотвращает включение соответствующей основной подпрограммы. Стирание признаков блокировки осуществляется основными или вклинивающимися подпрограммами ЭВМ при выполнении соответствующего условия.

Программа-диспетчер с фиксированными относительными приоритетами. Для размещения информации, обеспечивающей работу данной программы, в оперативной памяти ЭВМ выделяются две зоны. В одной из них хранится таблица установленного приоритета основных программ (рис. 10.6). Каждой основной подпрограмме(или группе подпрограмм равного приоритета) в этой таб-

лице соответствуют две ячейки памяти, в одной из которых указывается порядковый номер и приоритет данной подпрограммы, а в другой ячейке хранится адрес начальной команды этой подпрограммы.

Вторая зона – шкала приоритетов (рис. 10.7) – используется для хранения заявок на работу тех или иных программ и их порядковых номеров, необходимых для обращения в таблицу приоритетов за начальным адресом программы. Каждой основной программе в этой зоне соответствует одна ячейка памяти, расположенная в порядке убывания установленного для программ приоритета. Кроме того, в фиксированной ячейке ОП ЭВМ указывается общее количество приоритетов n , которое используется в данном режиме работы программ ЭВМ. Выбор программы для включения в счёт в данный момент времени осуществляется диспетчером путём последовательного анализа кодов заявок в шкале приоритетов и начинается всегда с заявки самого высокого приоритета (с начала зоны). Заявки равного приоритета анализируются в порядке их поступления, а их количество учитывается в отдельной буферной зоне памяти.

В рассматриваемой приоритетной диспетчерской программе так же, как и в рассмотренных ранее бесприоритетных диспетчерах, в общем случае может быть наложено ограничение на длительность выполняемых основных программ различного типа или же на количество заявок равного приоритета, обслуживаемых при одном включении основной программы.

Описанная выше программа-диспетчер имеет сравнительно небольшой объём (порядка 40 одноадресных команд) и позволяет легко изменять количество используемых основных программ, их приоритет и размещение в памяти ЭВМ. Для этого достаточно лишь соответствующим образом скорректировать информацию в таблице приоритетов.

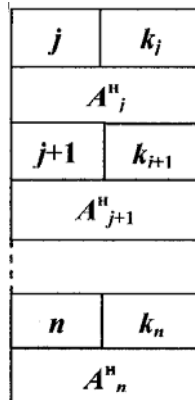


Рис. 10.6

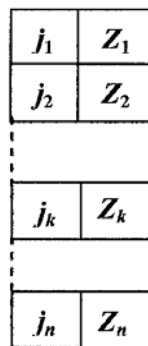


Рис. 10.7

Программа-диспетчер с фиксированными абсолютными приоритетами. В системах, работающих в режиме реального времени, временные ограничения могут быть очень жёсткими. Поэтому нужно обеспечить прерывание работы процессора в момент появления срочного запроса (даже при выполнении цикла процессора) и выполнить переход к соответствующей программе реакции, которая должна выполняться быстро и, по возможности, не прерываться. Иерархия приоритетов различных программ устанавливается в зависимости от их степени важности и от уверенности в том, что в них не допущены ошибки программирования.

Обычно в системах реального времени имеются два класса планирования: реального и разделенного времени. Класс планирования, как правило, дается не отдельным нитям, а целиком процессам. Процессы *реального* времени не прерываются по сигналам таймера и могут быть вытеснены только активизацией более приоритетной нити реального времени. Нити процессов *разделенного* времени вытесняются и друг другом по сигналам таймера, и процессами реального времени по мере их активизации.

Легко заметить, что планирование по методу приоритетной многозадачности может привести к определенным проблемам: к ситуации, когда низкоприоритетные задачи не смогут получить управление. Для решения этой проблемы применяется прием, получивший название «равнодоступность».

Принцип *равнодоступности* означает, что реализуется адаптивная приоритетность: приоритет задачи, которая выполняется слишком долго, постепенно уменьшается, позволяя менее приоритетным задачам получить свою долю процессорного времени.

Равнодоступность применяется главным образом в многопользовательских системах и редко – в системах реального времени.

В общем случае алгоритмы планирования должны соответствовать критериям оптимальности функционирования системы. Однако, если для систем «жесткого» реального времени такой критерий очевиден (всегда все делать вовремя), то для систем «мягкого» реального времени это может быть, например, минимальное максимальное запаздывание или средневзвешенная своевременность выполнения операций. В зависимости от критериев оптимальности

могут применяться алгоритмы планирования, отличные от рассмотренных. Например, ЦД может анализировать моменты выдачи критичных во времени управляющих воздействий и запускать на выполнение ту задачу, которая отвечает за ближайшие из них.

Не стоит особенно увлекаться приоритетами. Можно присвоить высокий приоритет «критической» задаче и низкий – всем остальным. Если «критических» задач несколько, то стоит рассмотреть многопроцессорную систему. Вообще говоря, следует иметь в виду, что сложные динамические системы трудно анализировать и отлаживать, поэтому лучше заплатить за более мощный процессор, чем иметь проблемы из-за непредвиденного поведения системы. В связи с этим большинство существующих систем реального времени представляют собой статические системы с фиксированными приоритетами.

10.4. Принципы организации прерываний

На время выполнения текущей программы внутри ЭВМ и в управляемом ей процессе могут возникать события, требующие немедленной её реакции: она должна прервать обработку текущей программы и перейти к выполнению другой программы, специально предназначенной для этого события. Этот процесс прерывания программы имеет случайный характер, потому что моменты возникновения событий, требующих прерывания, заранее неизвестны. Каждое событие, требующее прерывания, сопровождается сигналом, оповещающим ЭВМ, – запросами прерывания. Программа, затребованная запросом прерываний, называется *прерывающей* программой. Причины прерываний, возникающих в ЭВМ:

- сбой в аппаратуре;
- переполнение разрядной сетки;
- деление на нуль;
- выход за пределы области памяти;
- требование операции ввода-вывода.

Фактически возникновение прерывания означает, что параллельно развивается несколько процессов. Для реализации прерывания в ЭВМ есть специальные аппаратурные и программные средств-

ва, совокупность которых называется *системой прерывания* программ или *контроллером прерывания*.

Основные функции системы прерывания:

- запоминание состояния прерываемой программы;
- осуществление перехода к прерывающей программе;
- восстановление состояния прерванной программы;
- возврат к прерванной программе.

При наличии нескольких источников запросов прерывания должен быть установлен определённый порядок (дисциплина) в обслуживании запросов, то есть, между запросами должны быть установлены приоритетные соотношения, определяющие, какой из нескольких поступивших запросов пойдёт на обработку в первую очередь, и устанавливающие, имеет ли право данный запрос прерывать ту или иную программу.

Основные характеристики системы прерывания:

- общее число входов в систему прерывания (число источников запросов прерывания);
- глубина прерывания – максимальное число программ, которые могут прерывать друг друга;
- время реакции на запрос прерывания t_p ;
- время, необходимое для запоминания состояния прерываемой программы;
- время, необходимое для восстановления состояния прерванной программы.

Когда число причин прерывания превышает сотни, их разделяют на отдельные классы или уровни. Совокупность запросов, инициирующих одну и ту же прерывающую программу, образует класс или уровень прерывания. Объединение запросов в классы прерывания позволяет уменьшить объём аппаратуры, но связано с замедлением работы системы.

Информация, необходимая для перехода к прерывающей программе (в частности, ее начальный адрес) называется *вектором прерывания*. Векторы прерывания обычно хранятся в специальных ячейках памяти.

Центральное место в процедуре перехода к прерывающей программе занимают передача из регистра процессора в память (часто

– в стек) на сохранение текущего вектора состояния прерываемой программы и загрузка в этот регистр вектора прерывания прерывающей программы.

В простейшем случае вектор прерывания состоит только из начального адреса прерывающей программы. Для организации прерываний необходимо вводить анализ приоритетов. *Простейший вариант*: номер приоритета равен номеру линии запросов к входу системы прерывания. Здесь приоритет жёстко фиксирован.

Прерывание с опросом флажков – определение местоположения крайней слева единицы в регистре запросов путём опроса. Расположенный на данном месте разряд соответствует самому приоритетному запросу.

При *циклическом опросе источников прерывания* при помощи дешифратора и элементов «И» в каждом такте проверяется наличие запроса прерывания, номер которого совпадает с кодом счётчика.

Программно-управляемый приоритет позволяет изменить приоритеты в процессе работы системы. Есть два способа: порог прерывания и маска прерывания.

Первый способ – изменение программным способом уровня приоритета программы, которая в данный момент выполняется процессором. Это, в свою очередь, определяет минимальный уровень приоритета запросов, которым разрешается прерывать программу, идущую на процессоре.

Второй способ используется чаще. Маска прерывания – это двоичный код, разряды которого поставлены в соответствие запросам или классам прерывания. Изменяя специальный код маски, можно включать и выключать разные приоритеты. Замаскированное прерывание либо игнорируется, либо запоминается.

10.5. Организация счёта времени

Управление ходом вычислительного процесса в режиме реального времени подразумевает выдачу сигналов включения и отключения определённых подпрограмм в заданные моменты. Это достигается путём использования в процессе вычислений счёта времени.

Счёт времени, или счётчики времени (СчВр) можно организовать различными способами, которые требуют различных затрат

ресурсов, и при больших программных комплексах выбор того либо иного способа может влиять на производительность и надёжность ЭВМ или информационно-вычислительного комплекса (объединение рабочих станций и ЭВМ). Выделяют три способа организации СчВр: программный, аппаратный и смешанный. Основой счёта времени при любом из способов являются метки времени – сигналы генератора стабильной частоты Г. Они поступают на входы счётчиков непосредственно или через схему пересчёта.

Программный способ не требует специального оборудования. В этом случае сигналы прерывания генератора Г поступают в схему прерывания ЭВМ, причём данный канал имеет наивысший приоритет. По сигналу прерывания приостанавливается выполнение текущей программы, включаются подпрограммы СчВр, контроля состояния и передачи управления. Первые две программы выполняются для каждого программного счётчика при поступлении каждой метки времени, третья – передачи управления – в случае необходимости. Счётчиками времени при данной организации являются определённые ячейки ОП.

Период включения подпрограммы счёта времени $T_{с.в}$ определяется частотой следования меток времени и зависит от требуемой точности управления. Подпрограмма счёта времени прибавляет или вычитает 1 из слов памяти, в которых организуются счётчики. Подпрограмма контроля сравнивает, не достигли ли значения счётчиков заданного значения. В случае достижения выполняется подпрограмма передачи управления. Она реализует функции подготовки заявок на включение программ, формирование признаков их включения и пр.

Любые изменения временных диаграмм решения задач при рассматриваемом способе организацию управления требуют дополнительного программирования.

Аппаратный способ организации требует увеличения объёмов аппаратуры по сравнению с программным. Устройство, реализующее аппаратный способ формирования временной диаграммы, состоит из схемы счёта времени и формирования сигналов о состояниях счётчиков, схемы формирования сигналов управления ходом

вычислительного процесса и схемы управления. Возможны различные варианты построения этих схем.

Код, подаваемый i -й регистр в момент установки счётчика, должен представлять значение времени, являющееся суммой заданного интервала включения i -й программы T_i и текущего значения времени, описываемое текущим состоянием счётчика меток времени. То есть, в данной схеме на этапе подготовки информации для записи в счётчики, перед заданием каждого периода включения программы, ЭВМ необходимо вычислять код, который подаётся на регистр.

Выбор того или иного варианта построения схемы счёта времени и формирования сигналов о предельных состояниях счётчиков определяется наличием соответствующих БИС и микропроцессоров.

Структурная схема формирования сигналов управления на основе регистров состоит из регистров адреса и схем «И», разрешающих запись в регистр и считывание из регистра. В исходный момент времени i -й регистр устанавливается в нулевое состояние подачей сигнала «Уст. 0» по соответствующей шине. Затем, одновременно с установкой заданного кода времени в i -й счётчик схемы счёта времени, в данный регистр по разрешающему сигналу шины «Уст. адр.» записывается адрес перехода на включаемую программу. Сигнал нулевого состояния i -го счётчика времени (когда заданный интервал времени истек), поступающий из схемы счёта времени, считывает из регистра адрес, который по шине «Выд. адр.» передаётся в устройство управления.

Аппаратный способ требует значительно меньших затрат времени в цикле работы ЭВМ. ЭВМ должна записать в ОП коды, устанавливаемые в регистрах и на счётчиках схемы счёта времени, а также коды адресов переходов, заносимых в регистры адресов схемы формирования сигналов управления. Данная подготовка информации производится однократно на интервале 100 циклов и более. (При втором варианте схемы построения счёта времени необходимо вычислять код, подаваемый на регистр кода времени перед заданием каждого периода включения.) Другая составляющая работы ЭВМ со счётчиками – выполнение прерывания при включе-

нии подпрограмм. Оно производится только по завершению интервала включения каждой из подпрограмм. При программном же способе обработка прерываний и показаний счётчиков в ОП производится по поступлению каждой метки времени, то есть, для программного способа организации счёта времени требуется больше машинного времени.

Программно-аппаратный способ организации СчВр сочетает достоинства аппаратного и программного способов: значительно сокращает время на выполнение программ счёта и управления и незначительно увеличивает затраты аппаратуры. Можно предложить много вариантов реализации этого способа. Каждый вариант должен учитывать специфику конкретной используемой ЭВМ. Рассмотрим пример реализации программно-аппаратного способа формирования сигналов управления (рис. 10.9).

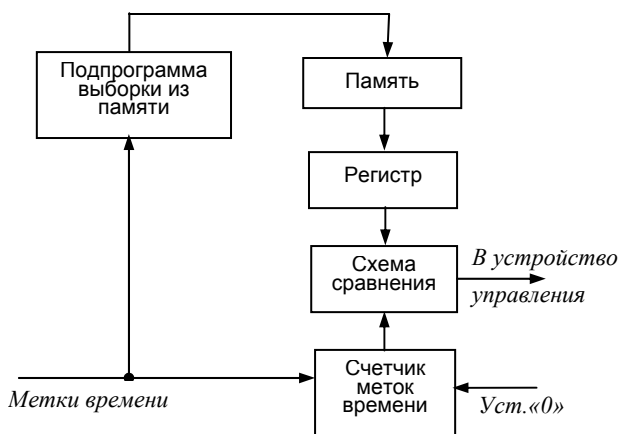


Рис. 10.9

Схема такого устройства работает следующим образом. С поступлением на счётный вход счётчика очередной метки времени он устанавливается в новое состояние. Этот же сигнал включает подпрограмму, которая выбирает из памяти последовательно коды заданных моментов времени, соответствующих временной диаграмме (временам включения программ). Эти коды последовательно поступают на регистр и сравниваются с содержимым счётчика.

Признак совпадения кодов направляется в устройство управления, где используется для формирования команды перехода. Здесь ЭВМ освобождается от необходимости после поступления каждой метки времени программным способом сравнивать состояния счётчика с временами включения программ и программно увеличивать или уменьшать на единицу слова памяти, отведённые под СчВр. Это делает аппаратура. На программное обеспечение ЭВМ ложатся только функции включения подпрограммы выдачи массива после поступления очередной метки времени и последовательной передаче кодов моментов времени на регистр.

Выбор того или иного способа организации СчВр может производиться на основе анализа затрат на единицу времени работы ЭВМ, затрат на создание и эксплуатацию аппаратуры. При определённых условиях вопрос выбора рационального принципа построения временной диаграммы может решаться исходя из реальных возможностей ЭВМ. Например, если ЭВМ не имеет достаточной производительности, то единственно возможным способом является аппаратный, поскольку замена типа процессора или увеличение их количества в составе средств вычислительной техники приводит к большему затратам по сравнению с затратами на аппаратуру счётчиков.

Итак, счётчики времени в ЭВМ считают метки времени, которые формируются Г, входящим в состав ЭВМ. Эта система – «внутренние часы» ЭВМ. В других объектах системы управления тоже могут быть свои «внутренние часы». Во многих случаях для обеспечения управления необходимо обеспечить взаимную синхронизацию «внутренних часов» разных компонентов АСУ. Для этого все часы «привязываются» к единому астрономическому времени. Это значит, что периодически в определённые моменты астрономического времени все «внутренние часы» устанавливаются в соответствующее состояние, после чего работают автономно до следующего момента корректировки.

Корректирующие сигналы единого времени посылаются в эфир специальными радиотехническими средствами – станциями сигналов единого времени (СЕВ). Каждая станция имеет расписание излучаемых частот и времени работы. Настраивая приёмник системы

управления в определённый интервал времени на определённую частоту, устанавливают связь со станцией СЕВ, по которой затем производится привязка [10]. При коррекции СчВр учитывается задержки СЕВ при распространении радиоволн по эфиру и в приёмном тракте.

10.6. Синхронизация задач

Хотя каждая задача в системе выполняет отдельную функцию, часто возникает необходимость во временной согласованности (синхронизации) действий, выполняемых различными задачами. Такая синхронизация необходима в следующих случаях:

1) функции, выполняемые различными задачами, связаны друг с другом. Например, если одна задача готовит исходные данные для другой, то последняя не выполняется до тех пор, пока не получит от первой задачи соответствующего сообщения;

2) необходимо упорядочить доступ нескольких задач к разделяемому ресурсу;

3) необходима синхронизация задачи с внешними событиями;

4) необходима синхронизация задачи во времени. Диапазон вариантов очень широк: от привязки момента выдачи какого-либо воздействия к точному астрономическому времени до простой задержки выполнения задачи на определённый интервал времени.

Рассмотрим эти варианты более подробно.

Связанные задачи. В дальнейшем будем называть *сообщениями* любой механизм явной передачи информации от одной задачи к другой. Объем сообщения может меняться от 1 бита до всей свободной емкости памяти системы.

Связь осуществляется передачей сообщений между задачами. Обмен может быть синхронным и асинхронным. В первом случае доставка сообщений производится после того, как она в плановом порядке получит управление. Обмен сразу инициирует работу планировщика и в действие вступает система приоритетов. Иногда бывает полезным ввести непосредственное управление приоритетами сообщений. Так, например, если после отправки нескольких сообщений на печать система решила отменить ее, сообщение об отме-

не должно встать в начало очереди и должно быть обработано как наиболее приоритетное.

Во втором – циркуляция сообщений оказывает непосредственное влияние на планирование задач. Так, задача, пославшая сообщение, блокируется, если для продолжения работы ей необходимо дожидаться ответа.

Обмен может идти через очередь сообщений или путем передачи сообщений.

Очередь сообщений представляет собой буфер данных; очереди удобны для буферизации данных при обмене между задачами или процедурами обслуживания прерываний. Иногда сообщения передаются через отведенный для этого буфер определенного размера («почтовый ящик»). При этом, как правило, новое сообщение затирает старое, даже если последнее не было обработано.

Наиболее часто используется принцип, когда каждая задача имеет свою очередь сообщений, в конец которой ставится всякое вновь полученное сообщение. Стандартный принцип обработки очереди сообщений по принципу FIFO не всегда оптимально соответствует поставленной задаче. В некоторых ОС РВ предусматривается возможность, когда сообщение от высокоприоритетной задачи обрабатывается в первую очередь (в этом случае говорят, что сообщение наследует приоритет пославшей его задачи).

Сообщение может содержать как сами данные, предназначенные для передачи, так и указатель на такие данные. В этом случае обмен может производиться с помощью разделяемых областей памяти, разделяемых файлов и т.п.

Механизм **передачи сообщений** обеспечивает возможность обмена данными между двумя задачами путем прямого копирования данных из одной задачи в другую. Сообщение может содержать как сами данные, так и указатели на данные.

Общие ресурсы. *Ресурс* – это физическое устройство или область памяти, которые могут одновременно использоваться только одной задачей. При работе с общими ресурсами очень важно обеспечить правильную последовательность передачи общего ресурса от задачи к задаче. Для решения этой проблемы есть несколько методов.

1. Использование специального сервера ресурса, ответственного за упорядочивание доступа к нему. В этом случае запрос на изменение значения глобальных данных посылается серверу в виде сообщения, а он устанавливает порядок доступа.

2. Запрет прерывания на время доступа к ресурсу.

3. Использование для упорядочивания доступа к глобальным данным байта блокировки, или семафора.

Семафор – это целочисленная переменная типа счетчика. Поток, перед тем как использовать ресурс, проверяет его состояние и затем вычитает из содержимого счетчика 1. Если значение семафора было равно нулю, поток проходит сквозь семафор успешно и захватывает ресурс. Если значение семафора было меньше нуля, поток останавливается и ставится в очередь.

Закончив работу, нить увеличивает значение семафора на единицу, открывая его. Первая из стоящих в очереди нитей активизируется и вычитает из значения семафора единицу, снова закрывая его. Если же очередь была пуста, семафор остается открытым. Тогда первая нить, подошедшая к семафору, успешно пройдет через него.

В борьбе за общие ресурсы могут происходить неприятности.

1. «Зацикливание» (или «зависание»). Например, задача *A* захватила ресурс клавиатуры и ждет, когда освободится ресурс дисплея; задача *B*, успев захватить ресурс дисплея, ждет, когда освободится клавиатура. Метод решения – освобождение всего, что было захвачено задачей и повторная попытка или использование сервера ресурса.

2. Инверсия приоритетов. Есть три задачи: высокоприоритетная *A*, среднеприоритетная *B* и низкоприоритетная *C*. Пусть задачи *A* и *B* в начальный момент времени блокированы в ожидании некоторого внешнего события. Тогда задача *C* получает семафор и не успевает его отдать, так как прерывается задачей *A*. Но поскольку семафор занят, задача *A* будет также блокирована и управление получит задача *B*. Она будет занимать процессорное время сколько угодно долго.

Синхронизация с внешними событиями. Это выполняется при помощи аппаратурных прерываний. Тенденции здесь следующие.

1. Стремление обеспечить максимально быструю и детерминированную реакцию на внешнее событие.
2. Стремление добиться максимально возможных периодов времени, когда прерывания в системе запрещены.
3. Реализовать минимально возможное количество функций прерывания за минимальное время.

Синхронизация во времени. Для этих целей используются таймеры. Для точной синхронизации с астрономическим временем применяются СЕВ с подстройкой по радиосигналам точного времени.

10.7. Управление оперативной памятью

Наиболее важным ресурсом после процессора является оперативная память. Большинство программ и данных хранится на жестком диске и загружается в оперативную память по мере необходимости. Поскольку в системах реального времени задержек исполнения быть не должно, все необходимые модули должны быть загружены предварительно. Кроме того, может возникнуть необходимость выгрузки содержимого части оперативной памяти на диск.

При выборе страницы на выгрузку могут быть использованы различные критерии, смысл которых сводится к тому, что на диск выталкивается страница, к которой в будущем, начиная с данного момента, дольше всего не будет обращений. Решение обычно принимается на основе эмпирических критериев. Наиболее популярным является число обращений за последний интервал времени.

Операционная система ведет для каждой страницы программный счетчик. Всякий раз, когда происходит обращение к какой-либо странице, процессор устанавливает в единицу признак доступа в относящейся к данной странице записи таблицы страниц. ОС периодически просматривает признаки доступа всех страниц во всех существующих в данный момент записях таблицы страниц. Если какой-либо признак оказывается равным 1 (было обращение), то система сбрасывает его в 0, увеличивая при этом на 1 значение связанного с этой страницей счетчика обращений. Когда возникает необходимость удалить какую-либо страницу из памяти, ОС нахо-

дит страницу, счетчик обращений к которой имеет минимальное значение. Для того чтобы критерий учитывал интенсивность обращений за последний период, ОС с соответствующей периодичностью обнуляет все счетчики.

Интенсивность страничного обмена может быть снижена в результате упреждающей загрузки, в соответствии с которой при возникновении страничного прерывания в память загружается не одна страница, содержащая адрес обращения, а сразу несколько прилегающих к ней страниц. Здесь используется эмпирическое правило: если обращение произошло по некоторому адресу, то велика вероятность того, что следующие обращения произойдут по соседним адресам.

10.8. Примеры операционных систем реального времени

Как уже отмечалось выше, операционные системы реального времени имеют ряд особенностей по сравнению с операционными системами обычных систем.

Ниже приводится краткий обзор некоторых ОС РВ.

Операционная систем **UNIX** представляет собой многозадачную многопользовательскую операционную систему и является в настоящее время одной из наиболее распространенных в мире. Она состоит из небольшого ядра, управляющего системными ресурсами (процессор, память и ввод/вывод), а остальная часть процедур операционной системы работает как пользовательские процессы. Типичная операционная система содержит 10000–20000 строк на языке С и 1000–2000 строк машинно-ориентированных программ на ассемблере, которые разрабатываются отдельно для каждой аппаратной платформы. Ядро представляет собой единую резидентную программу размером от 100 Кбайт до 1 Мбайт в зависимости от платформы и выполняемых функций.

Процессы в UNIX протекают с разделением времени. Для того чтобы дать возможность всем процессам исполняться, применяется динамическое распределение приоритетов. Процессу, готовому для исполнения, сначала присваивается его номинальный приоритет.

Затем, по мере исполнения, значение этого приоритета уменьшается до тех пор, пока он не становится меньше приоритета следующего из ожидающих процессов, который после этого выбирается для исполнения. В результате процессы с более высоким начальным приоритетом получают большую долю процессорного времени, и все процессы периодически выполняются.

Внешние устройства ввода/вывода рассматриваются в UNIX как файлы. Это упрощает программы, так как можно перенаправить ввод/вывод между файлами или внешними устройствами без изменения кода программы. Это важно и с точки зрения машинной независимости программ.

В последних версиях UNIX поддерживаются такие функциональные элементы систем реального времени, как семафоры, разделяемая память, обмен сигналами между процессами, приоритетное управление задачами и прямой доступ внешних устройств к памяти. Недостаток UNIX – это недружественный пользовательский интерфейс.

В России наиболее распространена операционная система **QNX**, отчасти из-за большой распространенности архитектуры x86 фирмы Intel. QNX – это разработка канадской компании QNX Software System Ltd. Впервые она появилась на рынке в 1981 году. Система построена по технологии FLEET:

- Fault-tolerance (отказоустойчивая);
- Load-balancing (регулирующая нагрузку);
- Efficient (эффективная);
- Extensible (расширяемая);
- Transparent (прозрачная).

Система допускает объединение в одну сеть разнородных Intel-совместимых компьютеров, соединенных через Arcnet, Ethernet, Token Ring или через последовательный порт, причем возможно участие одного компьютера в трех сетях одновременно.

Файловая система полностью соответствует стандарту POSIX. QNX обеспечивает работу с различными типами файловых систем: POSIX, Embedded (Flash, ROM, SRAM), CD-ROM (с поддержкой стандарта ISO 9660 и его расширения Rock Ridge), DOS (доступ ко всем носителям информации в формате DOS), NFS (доступ к раз-

личным типам удаленных файловых систем), SMB (прозрачный доступ к Windows 95 или NT-серверам).

Операционная система **OS-9** относится к классу Unix-подобных ОС. Она имеет модульную структуру, поэтому для исполнения программы в целевой системе требуются лишь некоторые части операционной системы. Весь код, генерируемый в OS-9, не зависит от адреса загрузки, он реентерабельный и может быть записан в ПЗУ. Для связи между процессами можно использовать сигналы, события, каналы и модули данных. Когда сигнал направляется от одного процесса к другому, процесс-адресат исполняет процедуру прерывания. Канал представляет собой последовательный поток данных от одного процесса к другому. Наиболее быстрый обмен данными осуществляется через модули – разделяемые области оперативной памяти, к которой могут обращаться все процессы. Среди других функций режима реального времени система также поддерживает аппаратные прерывания.

Основные характеристики OS-9:

- * многозадачная (65535 процессов, 65535 уровней приоритета);
- * многопользовательская (255 пользователей);
- * переносимость приложений: ANSI C/C++, POSIX 1003.1, TCP/IP (NFS/RPC), X Windows X11.R6 (OSF Motif), JAVA;
- * объектно-ориентированный модульный дизайн;
- * полностью вытесняемое детерминированное ядро с минимальным временем реакции на прерывание;
- * развитые сетевые средства: Arcnet, Ethernet, OMNinet, X.25, ISDN T1/E1, ATM, NFM, TCP/IP, IPX, Profibus, CAN, MIL STD 1553;
- * графические оконные интерфейсы GUI.

Операционная система **Windows NT** способна обеспечить гарантированное время отклика 5–10 мс. Тем не менее время реакции системы зависит от используемых драйверов, от того, запрещена подкачка страниц или нет, и от множества других факторов. Более того, механизмы обработки прерываний Windows NT имеют определенные особенности, которые мешают использовать эту ОС в приложениях «жесткого» реального времени. Например, пользовательский процесс, независимо от его приоритета, может быть пре-

рван для обработки аппаратного прерывания, даже если последнее пришло от такого источника, как манипулятор типа «мышь». Как правило, подпрограммы обработки аппаратных прерываний делают только самые необходимые действия, после чего основные процедуры обработки ставятся в очередь для последующего выполнения с использованием механизма *вызова отложенных процедур* (DPC — Deferred Procedure Call). К сожалению, DPC обслуживает очередь по принципу FIFO без учета приоритетов как прерываний, так и процессов, с которыми эти прерывания связаны. Более того, DPC не функционирует до тех пор, пока не будут отработаны все аппаратные прерывания, даже если они приходят от «мыши». В настоящее время ряд продуктов, таких как LA-SPOX, RTX, Falcon, Hyperkernel, приспособлен для реального времени на платформе Windows NT.

RTKernel 4.5 – многозадачное ядро реального времени для работы в среде MS-DOS. При инициализации создаются две задачи: основная и «пустая», которая нужна для нормального функционирования планировщика (ему нужна хотя бы одна задача, находящаяся в состоянии готовности). Планирование происходит согласно правилам:

1. Из всех задач, находящихся в состоянии готовности, в активное состояние переводится задача с наивысшим приоритетом.
2. Если в состоянии готовности пребывают несколько задач, имеющих одинаковый приоритет, в активное состояние переводится задача, не исполнявшаяся в течение наиболее длительного интервала времени.
3. Если несколько задач находятся в ожидании события, порядок их активизации при наступлении события осуществляется в порядке убывания их приоритетов.

Можно упомянуть и такие ОС реального времени как VxWorks, rSOS, LynxOS, VRTX.

10.9. Типовая структура алгоритмов управляющих ЭВМ

Вариант типовой структурной схемы управляющего алгоритма, включающей элементы операционной системы и специального математического обеспечения, представлен на рис. 10.8 и включает:

- программы обмена с внешними абонентами;
- программы управления и диспетчеризации вычислений;
- программы подготовки к обработке и подготовки к выдаче сообщений;
- программы контроля;
- программы решения функциональных задач, составляющих специальное математическое обеспечение.

Кроме того, если в системе применяется несколько ЭВМ, образующих вычислительный комплекс, то в состав операционной системы входят также программы взаимодействия комплексированных ЭВМ.

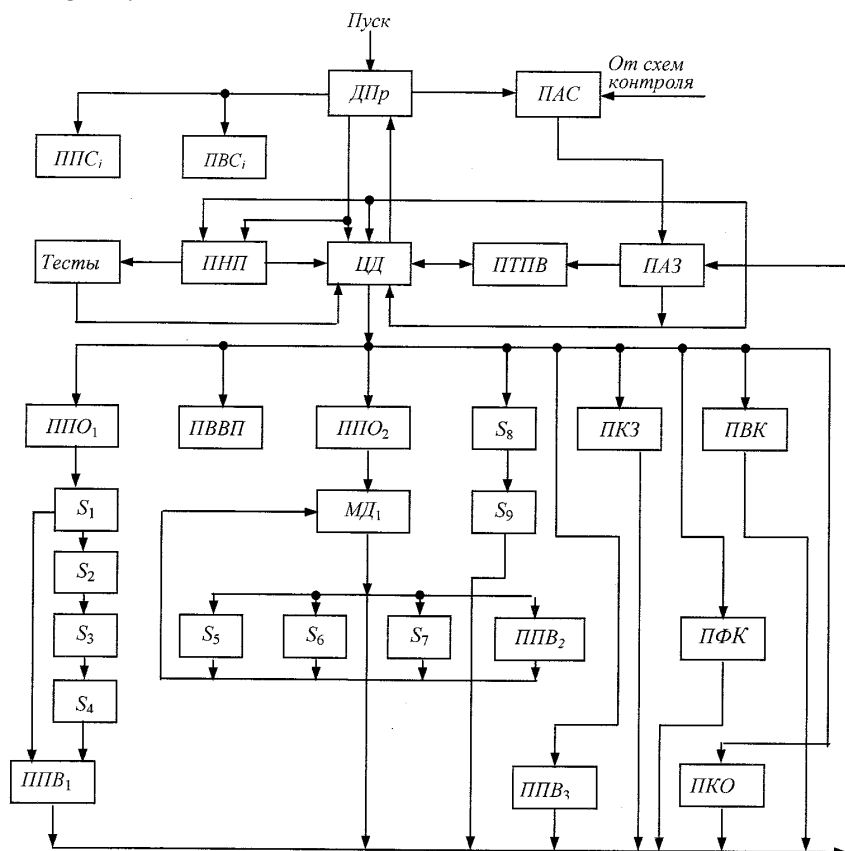


Рис. 10.8

Программы обмена с внешними абонентами можно разделить на две группы:

- программы приема сообщений (ППС_{*i*});
- программы выдачи сообщений (ПВС_{*i*}).

Включение этих программ обычно осуществляется схемно по инициативе внешних устройств, подготовивших сообщение к вводу в ЭВМ или освободившихся для приема сообщений из ЭВМ и для передачи управляемым объектам. Управление вклинивающимися программами обмена производится диспетчером прерываний (ДПр), который управляет также программой анализа сбоев и, возможно, другими программами с абсолютным приоритетом. Если сообщения внешних абонентов не образуют случайного потока и всегда готовы к вводу в ЭВМ (например, показания датчиков), то включение программы приема сообщений может производиться по инициативе ЭВМ от соответствующего диспетчера. В случае выдачи сообщений для включения ПВС_{*i*} необходимо выполнение двух условий: 1) готовности *i*-го канала к передаче сообщения; 2) наличия подготовленного к выдаче сообщения *m*-го типа (под типом сообщения подразумевается номер абонента, которому предназначено сообщение).

Алгоритмы приема сообщений решают следующие задачи:

- преобразование последовательного кода поступающих сигналов в параллельный код;
- первичное декодирование информации с учетом особенностей применяемого помехозащитного кода;
- накапливание поступающей информации в виде отдельных информационных слов и сообщений;
- привязка к поступающему сообщению условного номера внешнего абонента (маркировка сообщения для упрощения последующих процедур ввода этого сообщения в ЭВМ и диспетчеризации его обработки) и времени поступления сообщения в ЭВМ;
- определение подпрограммы приема информации в ЭВМ и буферной зоны памяти для хранения поступившего сообщения (по номеру внешнего абонента, типу сообщения и другим признакам);
- определение адреса свободного места в буферной зоне памяти ЭВМ;

- ввод информации в буферную зону памяти ЭВМ;
- изменение адреса свободного места в буферной зоне;
- сигнализация программе операционной системы ЭВМ о поступлении в машину очередной порции входной информации;
- контроль заполнения буферной зоны памяти входной информацией и сигнализация программе операционной системы о переполнении буферной зоны.

Алгоритмы выдачи сообщений разделяются на алгоритмы программной выдачи основными программами и вклинивающимися программами и алгоритмы внепрограммной выдачи.

Алгоритмы программной выдачи основными программами работают так. При включении в счет программа формирует в рабочей зоне ОЗУ очередное сообщение и при помощи стандартной программы обмена выдает его внешнему абоненту. При этом в справочную зону внешних абонентов записывается признак занятости абонента, который стирается либо специальной периодически включаемой программой наблюдения за состоянием внешних абонентов, либо вклинивающейся программой, включаемой в счет через систему прерывания по сигналу готовности внешнего абонента.

Алгоритмы программной выдачи вклинивающимися программами работают следующим образом. Процесс подготовки сообщения к выдаче заканчивается посылкой в соответствующее устройство обмена сигнала готовности к выдаче информации. Вторым условием для включения программы является сигнал готовности абонента. Операция непосредственного вывода информации осуществляется вклинивающейся программой во время выполнения машиной других алгоритмов системы управления.

Внепрограммная выдача информации отличается тем, что реализация процедуры непосредственного вывода информации осуществляется аппаратурой устройства обмена.

Алгоритмы выдачи сообщений должны осуществлять:

- проверку соответствия шифров подготовленных сообщений и освободившихся каналов передачи данных;
- выбор конкретного сообщения, подлежащего передаче, из совокупности накопленных в буфере, с учетом времени их подготовки и уровня их приоритета;

- выдачу слов сообщения в систему передачи данных;
- регистрацию завершения выдачи сообщения, адреса освобожденного места в буферном накопителе и его объема.

Программы управления и диспетчеризации включают:

- программу начального пуска (ПНП);
- программу тактировки периодических вычислений (ПТПВ);
- центральный диспетчер (ЦД);
- местные диспетчеры (МД).

Основное назначение *программы начального пуска* состоит в формировании, контроле и корректировке исходной необходимой информации в соответствии с заданным режимом функционирования системы или при его изменении. Она включается оператором с пульта управления периодически или при изменении внешних условий для контроля и корректировки исходных данных режима управления.

Основными функциями программы начального пуска являются:

- приведение аппаратуры ЭВМ в исходное состояние для перевода в рабочий режим;
- оперативный контроль исправности аппаратуры ЭВМ и устройств сопряжения при помощи обнаруживающих тест-программ;
- начальный ввод информации в ЗУ ЭВМ в соответствии с заданным режимом работы системы;
- введение астрономического времени в систему и установка начальных состояний всех датчиков времени в системе;
- программное подключение ЭВМ к устройствам сопряжения с внешней аппаратурой и линиями связи.

Программа тактировки периодических вычислений осуществляет контроль счетчиков реального времени и запись заявок на включение периодических программ в соответствии с заданным для них темпом. Она включается по сигналу прерывания от определенного разряда счетчика времени путем записи заявки в шкалу приоритетов центрального диспетчера. Возможно также программное включение путем анализа текущего состояния счетчика времени после завершения каждой программы.

Программа тактировки выполняет следующие функции:

- контролирует ход счетчиков реального времени с различным масштабом и возможность их переполнения;

- подготавливает условные или безусловные заявки на включение периодических программ с учетом времени последнего включения и заданной периодичности;
- учитывает время выдачи заявок на включение периодических программ и корректирует время очередного включения;
- формирует признаки о происшедших включениях периодических программ.

Диспетчеры регламентируют последовательность включения функциональных и управляющих программ в соответствии с заданной дисциплиной, типом программы и ее положением в шкале приоритетов. Наиболее распространенными дисциплинами в алгоритмах диспетчеризации являются:

- обслуживание заявок в порядке их поступления;
- приоритетное обслуживание без прерывания вычислений;
- приоритетное обслуживание с прерыванием вычислений.

Центральный диспетчер управляет включением совокупности программ, решающих крупную функциональную или вспомогательную задачу. *Местные диспетчеры* управляют последовательностью подключения подпрограмм в процессе решения задачи, заданной центральным диспетчером. Центральный диспетчер включается после завершения каждой включаемой им программы, а при отсутствии заявок и сообщений – после завершения анализа всей шкалы приоритетов. Местные диспетчеры включаются центральным диспетчером или функциональными подпрограммами после их завершения.

Заявки на включение программ диспетчерами могут появляться в основной программе в результате:

- приема сообщений от внешних абонентов;
- завершения решения некоторых задач, после чего должны быть вызваны определенные функциональные алгоритмы или подпрограммы;
- поступления данных от программы тактировки периодических вычислений;

Диспетчерские программы выполняют следующие функции:

- анализируют наличие заявок на включение групп подпрограмм в соответствии с заданной дисциплиной обслуживания;

- при обнаружении заявки определенного типа или приоритета проверяют наличие аналогичных заявок;
- корректируют таблицу с заявками при выборе из них очередной для включения подпрограмм;
- передают управление головной подпрограмме или местному диспетчеру для решения заданной функциональной задачи.

Программы подготовки к обработке (ППО_г) выполняют задачи, тесно связанные с основными функциональными задачами, однако их целесообразно выделить в самостоятельную группу, так как они носят характер вспомогательных служебных программ. Эти программы осуществляют:

- необходимое перекодирование и масштабирование квазинепрерывных величин;
- дешифрирование кодовой и признаковой информации;
- контроль и защиту от ложной и сбойной информации;
- перезапись информации из зоны входных сообщений в зону формуляров, подлежащих обработке в данный момент времени.

ППО_г часто выполняет роль местного диспетчера и определяет перечень и последовательность включения функциональных программ, необходимых для обработки данного сообщения.

Программы подготовки к выдаче (ППВ_г) формируют сообщения m -го типа в зоне формуляров выдачи и осуществляют их перезапись в буферную зону выдаваемых сообщений. При этом схемно формируется сигнал или записывается в оперативную память признак наличия подготовленного сообщения. Эти программы определяют дисциплину заполнения буферных зон и в значительной степени дисциплину выдачи, так как формируют адрес очередного сообщения, подлежащего выдаче для каждого абонента.

Программы контроля (ПК) можно разбить на две группы: 1) программы контроля ЦВМ в процессе рабочего функционирования системы в реальном масштабе времени; 2) программы функционального контроля всей системы управления вне рабочего режима. Кроме того, в состав программ контроля входят некоторые части основных функциональных программ, обеспечивающих защиту от сбоев, а также *программа анализа сбоев* (ПАС). Она включается либо схемно при выявлении сбоев методами аппаратного

контроля, либо теми участками программ, на которых обнаружена ошибка.

Подпрограмма анализа сбоев выполняет следующие функции:

- осуществляет регистрацию и накопление данных о всех выявленных искажениях вычислительного процесса и информации;
- селектирует искажения по их повторяемости и источникам;
- вырабатывает решения по ликвидации последствий или уменьшению ущерба от выявленного искажения путем включения соответствующих подпрограмм или переключениями в аппаратуре;
- выдает на индикацию для операторов характеристики выявленных искажений и мероприятий, принятых для ликвидации их последствий;
- ведет накопление наиболее часто встречающихся искажений и условий их появления для обеспечения возможности поиска частичных отказов в аппаратуре и ошибок в программе в режиме профилактики и восстановления системы.

Темп включения ПК при наличии случайных потоков сообщений и значительном изменении объема решаемых задач целесообразно изменять в зависимости от загрузки ЭВМ. Для этого центральный диспетчер включает *программу анализа загрузки* (ПАЗ), которая либо ведет учет холостого времени работы центрального диспетчера, либо подсчитывает суммарное время работы по основным программам и программам обмена в течение некоторого наблюдательного времени. В зависимости от текущей загрузки ЭВМ устанавливается темп включения программ контроля машины, но не ниже некоторого минимально допустимого. Кроме того, ПАЗ корректирует темп решения периодических программ, допускающих снижение темпа при перегрузках, а также управляет переходом на упрощенные методы решения функциональных задач, для которых предусмотрено изменение алгоритма при перегрузках.

Программа контроля ЭВМ в процессе рабочего функционирования может включать программы двух типов:

- программный тест контроля основных устройств ЭВМ без изменения информации, накопленной в оперативной памяти при решении функциональных задач;

– контрольную задачу, имитирующую решение основных функциональных задач по подготовленным и зафиксированным сообщениям с точным сравнением результатов с известным эталоном.

Контрольная задача позволяет проверить ЭВМ в условиях, наиболее близких к реальному функционированию в управляющей системе, в то время как тесты проверяют отдельные устройства ЭВМ при значениях и сочетаниях кодов, существенно отличающихся от тех, которые наиболее вероятны в реальных условиях.

Программы контроля обмена (ПКО) включаются периодически и обеспечивают сравнение с эталонами контрольных сообщений, принятых от внешних абонентов, формирование и подготовку к выдаче контрольных сообщений для внешних абонентов, а также обобщение и индикацию результатов обмена за некоторый интервал времени.

Программа функционального контроля (ПФК) может включаться при приеме сообщений функционального контроля внешних устройств или периодически для формирования сообщений, обеспечивающих функциональный контроль аппаратуры. Кроме того, эта программа может включаться при выявлении сбоев в поступающей информации для выявления источника этих сбоев. Данная программа имеет два режима работы:

– нормальный рабочий режим при решении основных задач управления;

– форсированный режим при решении задач функционального контроля перед включением рабочего режима или при появлении сбоев.

Подпрограмма взаимодействия комплексированных ЭВМ или процессоров (ПВК) обеспечивает межмашинный обмен информацией и распределение функциональных задач по вычислительным машинам с целью обеспечения необходимой пропускной способности управляющей системы или с целью повышения надежности решения функциональных задач. Особо сложными для данной подпрограммы являются задачи выявления ситуаций, когда необходимо изменить распределение задач и режим взаимодействия ЭВМ, а также осуществлять переход между режимами функ-

ционирования различного количества ЭВМ. Для решения перечисленных задач ПВК должны выполнять следующие функции:

- осуществлять начальный ввод исходной информации в память подключаемых ЭВМ и пуск их программ с учетом режима функционирования основной ЭВМ и требуемого режима вычислительной системы;
- проводить регулярный контроль и обмен между ЭВМ данными о состоянии аппаратуры и режиме ее функционирования;
- распределять заявки на включение функциональных алгоритмов между машинами и исходную информацию;
- производить периодический обмен резервируемой информацией о состоянии объектов и ходе процесса управления;
- вырабатывать и реализовывать решения на переключения в вычислительной системе при изменении потоков информации и соответственно при изменении загрузки в данном режиме функционирования;
- оперативно изменять структуру вычислительной системы при отказах ЭВМ или их частей, а также после восстановления аппаратуры.

Программы решения функциональных задач представлены условными блоками S_i , содержание которых, назначение и логические связи полностью определяются типом и задачами управляющей системы. Взаимодействие функциональных блоков по включению в работу может производиться двумя методами: через местный диспетчер или непосредственной передачей управления от блока к блоку.

10.10. Программирование алгоритмов реального времени

Программирование в реальном времени представляет собой раздел мультипрограммирования, связанный с разработкой взаимосвязанных параллельных процессов и управлением временными характеристиками процессов, взаимодействующих с внешней средой. Между программами реального времени и обычными последовательными программами имеются существенные различия:

- логика исполнения программы определяется внешними событиями и может зависеть от времени;
- программа работает не только с данными, но и с сигналами, поступающими от внешней среды в случайные моменты времени;
- программа работает в жестких временных ограничениях;
- результат выполнения программы зависит от общего состояния системы и его нельзя предсказать заранее;
- программа обычно работает в многозадачном режиме поэтому нуждается в средствах синхронизации и обмена данными;
- исполнение программы не заканчивается по окончании входных данных – она всегда ждет поступления новых данных.

Программное обеспечение систем реального времени, в зависимости от их масштаба и функционального назначения изменяется от программ, записанных в ПЗУ, до больших программных комплексов, обеспечивающих как разработку программ, так и их исполнение. Небольшие программы могут не иметь средств разработки и программы для них должны создаваться на более мощных ЭВМ с последующей загрузкой в исполняющую систему. То же относится к микропрограммам, «защитым» в ПЗУ оборудования производителем – они разрабатываются на ЭВМ, отличной от той, на которой исполняются.

10.10.1. Языки программирования

Программирование в реальном времени требует специальных средств, которые не встречаются в обычных языках последовательного программирования. Язык или операционная система для программирования в реальном времени должны предоставлять следующие возможности:

- описание параллельных процессов;
- переключение процессов на основе динамических приоритетов, которые могут изменяться, в том числе, и прикладными процессами;
- синхронизация процессов;
- обмен данными между процессами;
- функции, связанные с часами и таймером, абсолютное и относительное время ожидания;

- прямой доступ к внешним аппаратным портам;
- обработка прерываний;
- обработка исключений.

Не существует наилучшего языка – для каждого приложения и среды необходимо подбирать свои средства и при этом учитывать квалификацию и предпочтения разработчиков.

Первым полным языком программирования в реальном времени является ADA, структура которого похожа на структуру языка Pascal: процессу соответствует задача, которая выполняется независимо от других задач на выделенном виртуальном процессоре, то есть параллельно с другими задачами. Задачи могут быть связаны с отдельными прерываниями и исключениями и работать как их обработчики.

Новым понятием, введенным в ADA, является пакет – модуль со своими собственными описаниями типов данных, переменных и подпрограмм, в котором явно указано, какие из программ и переменных доступны извне. Пакеты могут компилироваться отдельно с последующим объединением в один исполняемый модуль. Это средство поддерживает модульную разработку программ и создание прикладных библиотек. В начале 1990-х годов язык ADA был пополнен новыми функциями для объектно-ориентированного программирования и программирования в реальном времени.

Основным недостатком ADA является его сложность; существующие компиляторы – дорогие продукты, которые требуют мощных процессоров.

Язык программирования C стал популярным для всех приложений, требующих высокой эффективности, в частности для систем реального времени. Это слаботипизированный язык, который позволяет программисту делать почти все вплоть до манипуляции с регистрами и битами. Небольшое количество заранее определенных функций и типов данных делает программы легко переносимыми между разными системами.

Язык C предпочтителен для написания программ с обращениями к функциям операционной системы, так как он обладает отличной совместимостью между логикой определения переменных и синтаксисом обращения к системе.

В языке С++ значительно улучшена абстракция данных с помощью понятия класса, похожего на абстрактный тип данных с четким разделением между данными и операциями. Классы С++ значительно легче использовать на практике, чем аналогичные понятия в других языках, поскольку С++ поддерживает объектно-ориентированное программирование и поэтапное уточнение типов данных.

Главным преимуществом языка С++ является его способность поддерживать разработку легко используемых библиотек программ. Программирование в реальном времени непосредственно в С++ не поддерживается, но может быть реализовано с помощью специально разработанных программных модулей и библиотек классов.

В задачах контроля объектов управления и стабилизации их параметров используются специализированные языки Техно ST, Техно SFC, Техно FBD, Техно IL. Эти языки являются модификациями языков ST (*Structured Text*), SFC (*Sequential Function Chart*) FBD (*Functional Block Diagram*), LD (*Ladder Diagram*), IL (*Instruction List*) стандарта IEC61131-3.

Программа на языке Техно ST или ее компонент разрабатываются в текстовом виде и определяются как последовательность предложений, составленных из операндов, операторов и функций. Аналогичную форму имеют и программы, написанные на языке Техно IL. Язык Техно SFC позволяет создавать программы в виде алгоритма, состоящего из SFC-шагов и SFC-переходов. Для SFC-шагов задаются выполняемые действия, для SFC-переходов – условия переходов между шагами. Для перехода от одного шага к другому условие SFC-перехода должно быть истинным. Вид алгоритма в редакторе языка Техно SFC – это последовательность шагов с указанием условий перехода. Сама программа представляет собой текст, содержащий расшифровку шагов и переходов.

Язык Техно FBD – это язык функциональных блоков. Функциональный блок – это графическое изображение вызова встроенной функции Техно FBD (например, в системе Trace Mode имеется более 150 простых и сложных функций) или функции, определенной пользователем. Создание связей между блоками производится

методом *drag-and-drop*, созданная связь обозначается на диаграмме линией. Аргументы и переменные FBD-программы задаются с помощью табличных редакторов.

10.10.2. Отладка программных комплексов реального времени

Доказательство правильности работы программы является обязательным шагом в ее разработке. Под *отладкой* понимается процесс, позволяющий получить программу, функционирующую с требуемыми характеристиками в заданной области входных данных. Основным методом обнаружения ошибок при отладке программ является их тестирование. Затраты на тестирование для обнаружения ошибок достигают 30...40% общих затрат на разработку программ и в значительной степени определяют качество созданного программного продукта.

Программы как объекты тестирования имеют ряд особенностей, которые отличают процесс их тестирования от традиционного, применяемого для проверки аппаратуры. Во-первых, для сложных программных продуктов практически всегда отсутствует полностью определенный и точный эталон для всех тестовых наборов. Поэтому для тестирования в качестве эталонов используются косвенные данные, которые не полностью отражают функции и характеристики отлаживаемых программ.

Во-вторых, для сложных программ недостижимо исчерпывающее тестирование, гарантирующее их абсолютно полную проверку. Поэтому тестирование обычно проводится в объемах, минимально необходимых для проверки в ограниченных пределах изменения параметров и условий функционирования.

В-третьих, в ряде случаев процесс исполнения программ и получаемые результаты зависят от непредсказуемого изменения входных и промежуточных данных, а также от реального времени. Вследствие этого невозможно создать единственный универсальный метод тестирования и приходится применять ряд значительно различающихся категорий тестов. Целесообразно выделить три стадии тестирования: для обнаружения ошибок; для диагностики и

локализации причин обнаруженных искажений результатов; для контроля выполненных корректировок программ и данных.

Основная цель тестирования для обнаружения ошибок – выявление всех отклонений результатов функционирования реальной программы от заданных эталонных значений. Задача состоит в обнаружении максимального числа ошибок, в качестве которых принимается любое отклонение от эталонов.

Ошибки в алгоритмах и программах условно можно разделить на четыре типа.

Технологические ошибки документации и фиксирования программ в памяти ЭВМ составляют 5...10% от общего числа ошибок, обнаруживаемые при отладке. Многократный перекрестный контроль соответствия данных исходным документам позволяет доводить в отдельных случаях вероятность технологической ошибки в программе до уровня 10^{-7} ... 10^{-8} .

Программные ошибки определяются степенью автоматизации программирования и глубиной формализованного контроля текстов программ. Их количество зависит от квалификации разработчиков, от общего объема комплекса программ и от степени взаимодействия модулей. На начальных этапах разработки и автономной отладки модулей программные ошибки составляют около 1/3 всех ошибок. На этапах комплексной отладки и эксплуатации удельный вес программных ошибок падает и составляет около 15% (ошибки в переменных) и 3% (ошибки в организации циклов) от общего количества ошибок, выявляемых в единицу времени. Каждая программная ошибка влечет за собой необходимость изменения около шести команд.

Алгоритмические ошибки – это ошибки, обусловленные некорректной постановкой функциональных задач, и неполным учетом всех условий решения. К ним также следует отнести ошибки связей модулей и функциональных групп программ. Они проявляются также в неполном учете диапазонов изменения переменных, в неправильной оценке точности; особую часть составляют просчеты в использовании доступных ресурсов вычислительных систем. На начальных этапах проектирования алгоритмические ошибки составляют около 30% общего количества ошибок. Для исправления

каждой алгоритмической ошибки приходится изменять в среднем около 14 команд.

Системные ошибки определяются неполной информацией о реальных процессах, происходящих в источниках и потребителях информации. Кроме того, эти процессы зачастую зависят от используемых алгоритмов и поэтому не могут быть достаточно точно определены заранее. На начальных этапах проектирования не всегда удается точно сформулировать задачи основных групп программ, поэтому технические задания и спецификации на отдельные программы уточняются и конкретизируются в ходе работы, для чего фиксируются отклонения от уточненного задания. Эти отклонения также могут квалифицироваться как системные ошибки.

При автономной и в начале комплексной отладки доля системных ошибок невелика (около 10 %), но она существенно возрастает (до 35...40 %) на завершающих этапах комплексной отладки. В процессе эксплуатации системные ошибки являются преобладающими (около 80 % всех ошибок); при исправлении одной системной ошибки приходится корректировать около 25 команд.

Никакие тесты не могут убедить, что выявлены все ошибки. Как заметил Эдсгер Дейкстра, «Тестирование может доказать только наличие ошибок, но не их отсутствие». При тестировании СРВ основная сложность – это наличие большого количества возможных взаимосвязей между задачами.

После тестирования для обнаружения ошибок применяется тестирование для их диагностики и локализации. Основная задача – точно установить место искажения программы или данных, явившегося причиной отклонения результатов от эталонных.

После локализации и устранения обнаруженных ошибок применяется контрольное тестирование, задача которого состоит в подтверждении правильности корректировки программы. В этом случае успешность тестирования определяется отсутствием проявления ранее обнаруженной, локализованной и устраненной ошибки, а также отсутствием вторичных ошибок, которые могут появиться при корректировке. Вероятность внесения новой ошибки при исправлении старой очень велика – имеющийся опыт разработки программ размером свыше 10000 строк дает вероятность в пределах от 15 до 50 %.

Существует две стратегии наращивания числа компонент, подлежащих проверке. При *восходящем* тестировании сначала проверяются модули нижних иерархических уровней, к которым постепенно подключаются вызывающие их модули. При этом обеспечивается работоспособность вызываемых компонент и функции группы программ проверяются при их естественном исполнении. Все участвующие в тестировании модули нижних иерархических уровней тестируются детально и независимо, что обеспечивает их высокую корректность при подключении к вызывающим модулям.

При *нисходящем* тестировании проверки начинаются с программ управления и организации вычислительного процесса. Первоначально тестируются управляющие программы и программы решения функциональных задач, размещенные на верхних иерархических уровнях. Постепенно к ним подключаются программы более низких уровней. Если некоторые из них не разработаны или не протестированы, то вместо них включаются программные имитаторы – «заглушки». Преимущество такого метода – возможность сохранения и развития тестовых данных по мере подключения программ нижнего уровня и полная проверка межмодульного интерфейса. Однако такое тестирование может требовать больших затрат на обнаружение простейших ошибок в модулях нижних иерархических уровней по мере их подключения.

На практике оба метода используются совместно с учетом сложности тестируемых групп программ. Модули и группы программ многократного использования преимущественно тестируются по восходящему методу. Управляющие и уникальные модули с малым числом вызываемых программ при небольшом числе иерархических уровней целесообразнее тестировать по нисходящему методу.

В многозадачных системах программные модули сначала тестируются раздельно. Должно быть проверено, что каждая строка программы выполняется хотя бы один раз. На этой фазе тестирования обычно полезны отладчики. Они позволяют непосредственно просматривать и изменять регистры процессора и области памяти при исполнении машинного кода. Отладчик вставляет в машинный код программы точки останова, в которых можно проверить со-

стояние регистров и переменных и сравнить их со значениями, требуемыми логикой процесса управления.

Обычно такие пошаговые отладчики не позволяют полностью отладить взаимодействие между несколькими параллельными процессами, но являются полезными и необходимыми средствами при отработке программ на ассемблере.

Только после того как все модули проверены по отдельности и все обнаруженные ошибки исправлены, можно приступать к их параллельному исполнению для отладки взаимодействия. Многочисленные взаимосвязи программных модулей могут привести к ошибкам в работе системы даже при их правильной работе по отдельности.

Основная задача комплексной отладки программ состоит в завершении разработки всего комплекса программ и в доведении его характеристик до значений, заданных требованиями технического задания. Должны быть:

- завершены и проверены сопряжение и взаимодействие по передаче управления и по информации всех компонент, входящих в комплекс;

- проверена возможность получения в процессе рабочего функционирования программного комплекса всех характеристик, заданных требованиями технического задания;

- проверены полнота и состав технической документации, а также точность соответствия ее изделию – комплексу программ.

Наиболее сложным является обеспечение возможности функционирования программ с характеристиками, заданными требованиями технического задания. Система должна удовлетворять всем требованиям не только в диапазоне типичных условий функционирования, но и при предельных, критических сочетаниях значений всех параметров. Каждая функционально законченная группа программ проходит отладку «от простого к сложному». Заранее должны быть определены значения всех частных критериев качества для всех компонент, реализация которых необходима для выполнения всех требований технического задания.

Важнейшим принципом комплексной отладки является последовательное сопряжение программных модулей, начиная с наибо-

лее простых по решаемым задачам и имеющих минимальные связи с другими программами. Для сокращения объема и уменьшения количества отладочных тестов целесообразно сопрягать программы в порядке увеличения объема информации, передаваемой при взаимодействии. В этом случае предшествующая по исполнению программа может являться источником значительной части данных для последующей, что сокращает объем подготовки необходимых тестов и имитаторов.

Комплексная отладка подразделяется на три этапа: 1) статическая комплексная отладка функциональных групп программ вне реального времени; 2) комплексная отладка в реальном времени функциональных групп программ и всего программного комплекса без использования реальных объектов и источников информации; 3) динамическая комплексная отладка программного комплекса в реальном времени и в реальной системе управления.

Статическая комплексная отладка обеспечивает проверку и корректировку условий сопряжения, определяющих взаимодействие отдельных компонент по информации и по управлению. Она состоит в контроле и обеспечении идентичности состава и характеристик информации, подготавливаемой каждой из сопрягаемых программ и данных, которые необходимы для правильного функционирования другой стыкующейся программы.

При статической отладке контролируются:

- структурная схема программного комплекса, определяющая логику взаимодействия и иерархию исполнения программ по задачам управления;
- информационная схема программного комплекса, отражающая реальные связи программ через глобальные переменные и константы;
- список всех возможных маршрутов исполнения групп функциональных программ, входящих в состав программного комплекса.

Комплексная отладка в реальном времени функциональных групп программ и всего программного комплекса без использования реальных объектов и источников информации предполагает использование программных моделей и имитаторов. Если затраты времени на генерацию тестов и обработку результатов малы, то

тесты могут производиться незадолго до их выдачи вне связи с реальным временем. Подготовленная при имитации тестовая информация должна содержать время, на которое она рассчитана и когда она должна быть введена для использования в программном комплексе. Если время, необходимое для генерации тестов и обработки информации, велико, то имитируемая информация заранее накапливается в имитирующей технологической ЭВМ со значениями времени, которым она соответствует (псевдореальное время), а затем переписывается на специальные носители.

При динамической комплексной отладке программного комплекса в реальном времени и в реальной системе управления важно учитывать не только выходные данные программ, но и информацию о реализации процесса их исполнения. Для этого должна быть обеспечена возможность регистрации и селекции любых промежуточных данных.

10.11. Типовая схема распределения оперативной памяти

В управляющих ЭВМ с непосредственным доступом к памяти в рабочем режиме оперативная память не предназначена для хранения программ и используется для переменной информации и информации, характеризующей режимы работы. В соответствии с функциональным назначением она делится на зоны, которые можно сгруппировать следующим образом:

- 1) для информации управляющих и диспетчерских программ;
- 2) для входной информации;
- 3) для выдаваемой информации;
- 4) для хранения результатов обработки информации;
- 5) для хранения контрольной информации;
- 6) для хранения программ.

В **первую** группу входят зоны, в которых хранятся

- исходные данные текущего режима работы (признак включенного режима, характеристики функционирующих элементов системы, перечень решаемых задач);
- таблица приоритетов (адреса начальных команд программ, включаемых центральным диспетчером);

- шкала приоритетов (заявки на включение определенных программ и адреса сообщений, подлежащих обработке);
- таблица периодических программ (темп и время последнего включения периодически исполняемых программ);
- таблица выдачи сообщений (адреса этих сообщений).

Информация, хранящаяся в этих зонах, определяет последовательность удовлетворения заявок на включение программ и включает эти программы. Наиболее распространенные дисциплины обслуживания – это беспriorитетное обслуживание, обслуживание заявок в порядке поступления, приоритетное обслуживание без прерывания (относительные приоритеты) и приоритетное обслуживание с прерыванием (абсолютные приоритеты).

Беспriorитетное обслуживание осуществляется по кольцевому принципу. Для его реализации в памяти выделяются две зоны, одна из которых (зона заявок) хранит коды заявок на включение подпрограмм, а вторая (зона адресов) – начальные адреса этих подпрограмм. При необходимости вызова подпрограммы в закрепленную за ней ячейку записывается код вызова. Программа-диспетчер, последовательно анализируя содержимое ячеек зоны заявок, находит этот код, стирает его и включает в счет соответствующую подпрограмму, используя для этого ее начальный адрес из зоны адресов. Кольцевой принцип обеспечивается тем, что после выполнения j -й подпрограммы дальнейший анализ зоны заявок начинается с $(j+1)$ -й или с первой ячейки зоны.

Обслуживание в порядке поступления требует для своей реализации учета последовательности поступления заявок.

При **первом** способе заявки на включение записываются в порядке поступления в зону заявок в виде номера вызываемой подпрограммы. Диспетчер анализирует содержимое этой зоны и включает подпрограмму, выбирая начальный адрес подпрограммы из зоны адресов по ее номеру. Перед включением подпрограммы этот номер (то есть собственно заявка на включение) стирается программой-диспетчером.

При **втором** способе заявки на включение записываются в соответствующие ячейки зоны заявок, однако заявкой является показание счетчика реального времени в момент поступления заявки.

Программа-диспетчер анализирует содержимое зоны заявок и выбирает минимальное значение времени, после чего включает соответствующую подпрограмму. Стирание осуществляется записью в данную ячейку кода, равного максимальному показанию счетчика реального времени (нулевое значение является одним из возможных значений).

Приоритетное обслуживание также требует для своей реализации двух зон: в одной из них хранится таблица установленного приоритета подпрограмм, в другой – заявки на работу подпрограмм и их порядковые номера.

В первой зоне (таблице приоритетов) каждой подпрограмме отводится две ячейки. В одной из них указывается порядковый номер и приоритет данной подпрограммы, в другой – адрес начальной команды этой подпрограммы.

Во второй зоне (шкале приоритетов) каждой подпрограмме соответствует одна фиксированная ячейка, где хранится ее номер. Эти ячейки расположены в порядке убывания установленного для этой подпрограммы приоритета,

Анализ выполняется по шкале приоритетов программой-диспетчером в порядке убывания приоритетности, начиная с самого высокого приоритета.

Программа тактировки периодических вычислений включается периодически по сигналам от одного из разрядов счетчика реального времени ЭВМ. При каждом включении контролируются показания счетчиков реального времени

$$t_{j(n+1)}^k = t_{jn}^k + T_j^k,$$

где $t_{jn}^k, t_{j(n+1)}^k$ – показания k -го счетчика реального времени в моменты n -й и $(n+1)$ -й тактировки j -й периодической подпрограммы; T_j^k – период тактировки j -й периодической подпрограммы с использованием k -го счетчика реального времени. Эти коды хранятся в специальной зоне оперативной памяти. В определенные моменты времени программа тактировки записывает заявки на включение периодических программ в шкалы приоритетов.

Во **вторую** группу входят зоны:

- внешних сообщений – буферные накопители сообщений от внешних устройств;
- накопления итогов решения функциональных задач, служащих исходными данными для других функциональных задач;
- контрольных сообщений внешних устройств, которые в частном случае могут входить в зону внешних сообщений.

Если значения интенсивности потоков, объемы сообщений и времени передачи примерно одинаковы, то выгодно иметь общую зону; в противном случае целесообразно выделять зоны для определенных типов сообщений.

В третью зону входят зоны:

- выдачи информационных сообщений;
- выдачи контрольных сообщений.

Если устройства характеризуются близкими значениями потока, объема и времени передачи сообщения, то для такого потока целесообразно строить единую зону.

Основной принцип деления **четвертой** группы зон определяет-ся временем хранения и темпом изменения информации в этих зонах:

- рабочая память текущих расчетов характеризуется быстрой сменой информации, причем любой ячейкой может пользоваться любая подпрограмма и перед использованием следующей подпрограммой ячейка считается свободной;
- стандартные ячейки определенных подпрограмм могут содержать только конкретную информацию, подготовленную каждой подпрограммой для собственных последующих включений или для следующих за ней подпрограмм;
- зона долговременного хранения результатов процесса управления является памятью состояния всей системы управления, необходимого для прогнозирования управления и принятия основных решений; эта зона должна быть особенно тщательно защищена от случайных сбоев и отказов ЭВМ;
- зона формуляра для обработки используется для хранения сообщения после того, как программа-диспетчер назначила его на обработку и оно должно быть исключено из зоны входных сообщений;

- зоны формуляров для выдачи используются для формирования сообщений внешним устройствам; после того, как сообщение будет подготовлено для выдачи, оно переписывается в одну из зон выходной информации.

Пятая группа содержит зоны:

- контроля управляющей ЭВМ;
- контроля внешнего обмена;
- функционального контроля системы управления в целом.

Первая зона содержит информацию о текущем техническом состоянии ЭВМ, о результатах контроля и ходе решения контрольных задач, накапливает данные по частоте сбоев отдельных устройств.

Вторая зона хранит данные о подключении и режиме работы каждого устройства, о результатах текущего обмена контрольными сообщениями, о характере и частоте сбоев.

Третья зона обеспечивает данными подпрограммы накопления и анализа информации о состоянии основных устройств управляющей системы, изменения темпа их контроля, анализа отказов и выработки рекомендаций на переключение устройств.

Шестая группа зон содержит зоны:

- программ частого использования;
- буферные (для переписи подпрограмм из внешних накопителей в оперативную память).

Первая зона содержит программы, которые необходимы для управления вычислительным процессом, для обмена с внешними устройствами, для решения часто встречающихся задач и расчета некоторых стандартных функций. Она может содержать и типовые константы.

Вторая зона служит для хранения редко включаемых подпрограмм на время их переписи из внешних накопителей в оперативную память. Ее объем должен обеспечивать хранение достаточно больших массивов подпрограмм, так как частое обращение к внешним накопителям может значительно снизить производительность управляющей ЭВМ.

Контрольные вопросы и упражнения

1. Каковы достоинства и недостатки монолитных операционных систем? микроядерных? объектно-ориентированных?
2. Что такое «процесс»? «поток»?
3. В чем заключаются преимущества и недостатки потоков?
4. Что понимается под «реентерабельностью»?
5. Объясните существо принципов диспетчеризации задач.
6. Как решаются задачи определения приоритетов?
7. Когда необходима синхронизация задач?
8. Из каких соображений выбирается размер страницы?
9. Назовите основные методы решения проблемы передачи ресурсов от задачи к задаче.
10. Приведите примеры операционных систем реального времени.
11. Объясните назначение элементов типовой структурной схемы программного обеспечения систем реального времени.
12. Расскажите о функциях программ контроля хода вычислительного процесса и обеспечения устойчивости вычислений.
13. На какие этапы делится комплексная отладка программ реального времени?
14. Какие зоны входят в первую группу зон оперативной памяти систем реального времени?
15. Из каких зон состоит четвертая группа?

ЗАКЛЮЧЕНИЕ

Технические средства систем реального времени непрерывно развиваются: совершенствуется их практическая реализация, улучшаются характеристики, появляются новые принципы построения. В данном пособии изложены только основные идеи, выдержавшие проверку временем. После их изучения следует просматривать технические журналы и другую техническую документацию, чтобы быть в курсе новейших достижений.

Новые принципы управления появляются не так часто, но это также имеет место. Достаточно привести такие примеры как управление при помощи нечеткой логики или решение ряда задач с использованием генетических алгоритмов. Однако все-таки в этой области расширение идет по линии увеличения количества задач управления, а не принципов их решения. Поэтому программное обеспечение систем реального времени остается достаточно стабильным.

В целом построение систем реального масштаба времени отличается гораздо большим разнообразием применяемых принципов, технических и программных средств, чем построение информационных систем или систем управления экономико-организационными объектами. И это разнообразие с каждым годом все более и более расширяется.

В данном учебном пособии не рассматриваются вопросы проектирования систем управления реального времени. Это самостоятельная и сложная задача, которую разные генеральные конструкторы решают по-разному. Недаром говорят о различных школах проектирования, о существующих традициях. Однако здесь также прослеживается тенденция: все больше параметров теряют ограничения, и на первое место среди критериев эффективности выдвигается задача экономии ресурсов.

Использованная литература

1. Келим Ю.М. Типовые элементы систем автоматического управления: Учебное пособие для сред. проф. образования. М.: ФОРУМ: ИНФРА-М, 2004.
2. Шишмарев В.Ю. Типовые элементы систем автоматического управления: Учебник для сред. проф. образования. М.: Издательский центр «Академия», 2004.
3. Гоноровский, И.С. Радиотехнические цепи и сигналы: Учебник для вузов – 4-е изд., перераб. и доп. М.: Радио и связь, 1986.
4. Воросклевский В.И., Пинягин Н.Я. Преобразователи напряжения в частоту и их применение в технике измерений и управления. М.: Энергоатомиздат, 1994.
5. Датчики. / М.П. Цапенко, И.Ф. Клисторин, А.Ф. Олейников. // Датчики и системы, 1990. № 1. С. 17–18.
6. Микропроцессорные системы: Учебное пособие для вузов / Е.А. Александров, Р.И. Грушвицкий, М.С. Куприянов и др./ Под общ. ред. Д.В. Пузанкова. – СПб.: Политехника, 2002.
7. Крылов, А.А. Человек в автоматизированных системах управления. Л.: Изд. ЛГУ, 1972.
8. Тищенко, Н.М. Введение в проектирование сложных систем автоматики. М.: Энергия, 1976.
9. Основы инженерной психологии. / Б.А. Душков, Б.Ф. Ломов, В.Ф. Рубахин и др./ Под ред. Б.Ф. Ломова. М.: Высш. школа, 1977.
10. Математическое обеспечение управляющих ЦВМ. / В.В. Липаев, К.К. Колин, Л.А. Серебровский. М.: Сов. радио, 1972.

Юрий Георгиевич Дреус

СИСТЕМЫ РЕАЛЬНОГО ВРЕМЕНИ:
ТЕХНИЧЕСКИЕ И ПРОГРАММНЫЕ СРЕДСТВА

Учебное пособие

Редактор *Е.Н. Кочубей*
Макет подготовлен *Е.Н. Кочубей*

Подписано в печать 4.08.2010. Формат 60×84 1/16.
Печ.л. 20,0. Уч.-изд.л. 20,0 Тираж 100 экз.
Изд. № 111-1. Заказ № 241.
Национальный исследовательский ядерный университет «МИФИ».
Типография НИЯУ МИФИ. 115409, Москва, Каширское ш., д.31