ИСТОРИЧЕСКИЙ ОЧЕРК.

Еще с середины 40-ых годов , специалисты по радиоэлектроники начали задумываться над возможностью применения специализированных цифровых устройств для решения разнообразных задач ,связанных с обработкой сигналов . Нечего и говорить , что в то время выводы не были благоприятными . С точки зрения стоимости, размеров и надежности предпочтение следовало отдать аналоговой фильтрации и аналоговым методам спектрального анализа . В 50-ых годах теория управления , частично основанная на работе Гуревича ( 1945 г.) , уже утвердилась как самостоятельное научное направление ; были глубоко изучены принципы дискретизации колебаний и возникающие при этом спектральные эффекты , а математический аппарат теории z-преобразования , существовавший еще со времен Лапласа , начал находить применение в радиоэлектроники и смежных дисциплинах . Однако достигнутый уровень развития техники позволял получить практические результаты только в задачах управления медленными процессами и обработке низкочастотных сейсмических сигналов . К середине 60-ых годов были оценены потенциальные возможности интегральных микросхем , что позволило представить полную систему обработки сигналов , для которых наилучшая техническая реализация была бы именно цифровой .

Первый крупный вклад в теорию цифровой обработки сигналов , касающийся анализа и синтеза цифровых фильтров , был сделан Кайзером ( фирма Bell ) ; он показал , как можно рассчитывать цифровые фильтры с нужными характеристиками , используя билинейное преобразование . Примерно тогда же ( 1965 г.) появилась статья Кули и Тьюки о быстром методе вычисления дискретного преобразования Фурье , давшая мощный толчек развитию этого нового технического направления . Позже метод был развит и стал широко известен как быстрое преобразование Фурье ( БПФ ) . Ценность этого метода заключается в сокращении времени вычисления дискретного преобразования Фурье ( на один-два порядка для большинства практических задач ). Опубликование статьи Кули и Тьюки ускорило развитие строгой и достаточно полной теории цифровой фильтрации . Важнейшее значение метода БПФ состояло в том , что он наглядно продемонстрировал , насколько цифровые методы при спектральном анализе могут оказаться экономичнее аналоговых . После создания метода БПФ интенсивность исследований в области цифровой фильтрации резко возросла , и в настоящее время цифровые методы широко используются для спектрального анализа самых разнообразных сигналов , начиная с низкочастотных колебаний в сейсмологии и звуковых колебаний в гидрологии и при анализе речи и кончая видеосигналами в радиолокации .

Первой попыткой исчерпывающего изложения теории цифровой обработки сигналов была книга Гоулда и Рэйдера ( 1969 г.) . Эту книгу применяли в качестве учебного пособия для аспирантов, и как руководство для инженеров ,работающих в промышленности . Естественно , что книга не могла удовлетворить и тех и других . Не нужно доказывать , что хорошее учебное пособие может быть составленно только на основе курса , читавшегося в течении по крайней мере несколько лет , и подходящего набора задач .

**ПРИЧИНЫ ВНЕДРЕНИЯ ЦОС В**

**ЭЛЕКТРОСВЯЗЬ.**

1. Сложность ( нередко невозможность ) решения некоторых задач аналоговым методом .

2. Прогресс в развитии электроники ( создание высокоскоростных многоразрядных АЦП , разработка сигнальных процессоров ) .

3. ЦОС позволяет реализовать универсальные модемы , в которых изменением программы осуществляется переход с одного вида сигнала на другой ( т.е. с одной модуляции на другую ).

4. ЦОС позволяет строить адаптивные радиоприемные устройства, работающие во все усложняющейся электромагнитной обстановке ( т.е. спектр постоянно загружается сигналами ) .

5. Простота , автоматически сменных , алгоритмов ЦОС и высокая точность их реализации .

6. ЦОС позволяет реализовать более сложные алгоритмы радио приема ( разнесенный прием , компенсация и подавление сосредоточенных помех и прием в целом ) .

7. При использование ЦОС значительно меньше влияет разброс параметров и действие дестабилизирующих факторов.

8. Высокая интеграция цифровых микросхем позволяет реализовать очень сложные алгоритмы приема сигналов , сохраняя приемлемый объем и стоимость аппаратуры .

9. Цифровая аппаратура легко поддается миниатюризации. Высокая технологичность и отсутствие регулировки понижает стоимость.

10.Проектирование цифровых устройств легче чем аналоговых и поддается автоматизации ( легко модулируются на ЭВМ ) .

11.ЦОС облегчает работу по созданию спецэфектов на ТВ ( работа режиссеров на теле-студии ) .

12.ЦОС позволяет существенно повысить качество изображения.

**ПРОБЛЕМЫ РАЗВИТИЯ ЦОС .**

1. Для ЦОС необходимо преобразовать аналоговый сигнал в цифровой ( требуется достаточно большой уровень сигнала - порядка 1в ) .

2. Преобразование аналогово сигнала в цифровой приводит к появлению погрешности дискретизации во времени и к погрешности квантования по уровню ( специфические погрешности ) .

3. Процесс обработки сигналов сопровождается погрешностями , вызванными округлениями результатов ( это приводит к ошибкам - шумам ) .

4.Требуется увеличение динамического диапазона и ширины спектра преобразуемых аналоговых сигналов ( т.к. каналы с ограниченной полосой пропускания и сложной помеховой обстановкой ) . Чтобы достигнуть возможности аналоговой техники нужно иметь динамический диапазон АЦП 120-130 дб с df=100 кГц . Таких АЦП пока нет . Реализуемый при df=100 кГц динамический диапазон АЦП 70-80 дб . Для широкополосных сигналов при df=100 Мгц динамический диапазон 6-24 дб .

5. Низкая скорость работы цифровых вычислительных устройств. (Сигнальные процессоры : КМ1813ВЕ11 , ТМS320.10 , ТМS320.20 , ТМS320.30 , ДSР5600 , ТМS320.50 .)

ТЕОРЕМА КОТЕЛЬНИКОВА .

Любой сигнал с ограниченным спектром ( бесконечный во времени ) однозначно определяется своими отсчетами , взятыми через интервал времени dt=1/2F т.е.

, где u(kt)-аналоговая величина;



Эта теорема утверждает , что если сигнал f(t) имеет преобразование Фурье Sf() отличное от нуля при частотах меньших 2Fm . То в отсчетах сигнала f(kt) взятых через интервал t=1/2Fm содержится вся информация о непрерывной функции f(t) . Из теоремы следует , что эти отсчеты содержат информацию о сигнале f(t) в любой момент времени . Однако частота отсчетов должна быть по крайней мере в два раза больше высшей частоты сигнала Fm .

Доказательство.:

Дан сигнал f(t) , его спектр : S()= при <2Fm ,



0 , при >2Fm.

Представим некоторую реализацию сигнала f(t) и его спектр S(f):

Если отсчеты сигнала брать с помощью бесконечно узких импульсов,расположенных в непосредственной близости друг от друга , мы однозначно определим любую функцию . Если интервал между импульсами увеличивать , то где-то мы начнем терять информацию о сигнале . Рассмотрим случай ,когда в качестве отсчетных импульсов используется периодическая последовательность импульсов длительностью  , повторяемых через t=1/2Fm . Временное и спектральное представление этих импульсов:

Спектр отсчетных импульсов можно записать в виде ряда Фурье , т.е. y(t)=A1cost+A2cost+A3cost+............ Процедуру взятия отсчетов удобно рассматривать как умножение функции f(t) на функцию y(t) . Результирующий дискретизованный сигнал можно представить в виде суммы последовательностей импульсов ,амплитуды которых равны значению функции f(t) в момент отсчета , а спектр такого сигнала представляет собой периодически повторяющуюся функцию Sf() с периодом  ,т.е.мы наблюдаем изменение амплитуды импульсов отсчета по закону f(t) и соответственно имеем амплитудную модуляцию каждой гармоники спектра импульсов отсчета сигналa :

Для восстановления првоначального сигнала нам достаточно отфильтровать полученный сигнал ФНЧ с частотой среза расположенной в интервале от Fm до 1/t-Fm . Рассмотрим какова может быть наименьшая частота следования счетных  импульсов, что бы еще имелась возможность отфильтровать полезный сигнал. В случае , если 1/ t=2Fm мы еще имеем возможность отфильтровать полезный сигнал если же 1/t<2Fm ,то произойдет наложение спектральных составляющих и восстановление первоначального сигнала без ошибки станет невозможным. Следовательно , для восстановления сигнала ,полученные отсчетные импульсы необходимо подать на вход ФНЧ с частотой среза равной Fm. Реакция идеального ФНЧ на узкий импульс единичной амплитуду представляет собой функцию вида : (t)=sin2Ft/2Ft

На вход фильтра мы подаем сумму импульсов с амплитудами равными f(kt) Разложение сигнала f(t) в ряд Котельникова указывает на технический способ передачи непрерывной функции (сигнала) f(t)с ограниченным спектром путем передачи отсчетных импульсов ,который сводиться к следующему:

и со сдвигом один относительно другого на t=1/2Fm . Сигнал на выходе фильтра представляет собой сумму откликов ,т.е. Что соответствует ряду Котельникова .



**Восстановление сигналов по его отсчетам .**

1)взятие отсчета f(kt) функции f(t) в моменты kt ;

2)значение полученных отсчетов передаются на приемную сторону с использованием любогометода кодирования и модуляции ;

3)на приемной стороне вырабатываются короткие импульсы ,амплитуды которых пропорциональны принятым значениям отсчетов ;

4)полученные импульсы подаются на идеальный ФНЧ с частотой среза Fм . На выходе фильтра получается функция f '(t) , пропорциональная переданной функции f(t) . Идеальный ФНЧ с полосой пропускания Fм при действии на его вход единичного импульса (t) дает на выходе напряжение ,соответствующее функции : (t)=sin2 Fmt/2Fmt При восстановлении функции f(t) на вход фильтра подают короткие импульсы с амплитудами , соответствующими f(kt) и с интервалами t. На выходе фильтра получается напряжение , соответствующее сумме откликов фильтра на каждый из импульсов . В моменты времени kt функция f(t) восстанавливается совершенно точно , так как в этот момент только одна из отсчетных функций (t-kt) не равна нулю . В остальные моменты времени для точного восстановления необходимо суммировать бесконечное число отсчетных функций .

**Ошибки восстановления сигнала по отсчетам Котельникова.**

Как было отмечено выше , точное восстановление сигнала возможно только при строго ограниченном спектре сигнала и при использовании идеального ФНЧ .НА практике мы имеем дело с сигналами конечными во времени, т.е. бесконечным , теоретически , спектром и для восстановления используем реальные ФНЧ . Рассмотрим ошибки восстановления , вызванные реальностью сигнала (сигнал ограничен во времени , т.е. не ограничен по частоте ). Основная энергия сигнала сосредоточена в диапазоне частот до Fm и только малая доля будет выходить за Fm .

1)На основании т. Котельникова мы не можем восстановить спектральные составляющие , лежащие выше частоты Fm .

2)В спектре восстановленного сигнала появяться дополнительные составляющие , представляющие собой зеркальное отображение " вниз " по частоте спектральных составляющих сигнала относительно оси совпадающей с частотой среза идеального ФНЧ и равной Fm .Поясним

это на рисунке: фнч

S f(f) S1(f) S2(f) S3(f)

0 Fm 3Fm f

Огибающая спектральной плотности сигнала f(t) представляет собой функцию S1(f) . Спектр отсчетных импульсов Sf(f) представляет собой периодически повторяемую функцию S1(f) с периодом 2Fm . Идеальный ФНЧ с частотой среза Fm не пропускает составляющие основного сигнала и пропускает составляющие сектра амплитудно-модулируемой первой гармоники спектра отсчетных импульсов (2Fм) .

3)При восстановлении сигнала конечной длительности следует иметь ввиду что :

а) точность восстановления в средней части сигнала будет наибольшей, а по краям наименьшей;

б) в моменты , соответствующие отсчетам сигнал восстанавливается точно, а в средней части между отсчетными моментами ошибка максимальна

ВЫБОРКИ ИЗ АНАЛОГОВОГО СИГНАЛА.

Схема взятия выборки из аналогового сигнала.

1-Умножитель

2-Схема хранения УВХ

3-Квантователь

4-Преобразователь АЦП

5-Регистр

УВХ-устройство выборки и хранения. Перед умножителем стоит фильтр для уменьшения помех. Квантователь находит ближайший оцифрованный уровень. Устройство хранения дает время квантователю для принятия решения. Устройство хранения-конденсатор,окруженный ключами с большим сопротивлением ( т.е.RC-цепочкой с малой емкостью).Постоянная времени  стремится к единице, это переходный процесс в цепочке (т.е. конденсатор заряжается). За время t изменение сигнала мало,т.к. очень большое входное сопротивление преобразователя.Это и есть хранение. Преобразователь -преобразует вид кода (т.е. переводит его в бинарную систему счисления, за счет пороговых устройств). Регистр-считывает этот код, а за тем последовательно, побитно передает в линию.

ДИСКРЕТНОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ.

Квантование перидического сигнала.



cos, cos2, ... , cosn.

n=3 n=Ґ

Много ли  нужно иметь и от чего это зависит (зависит от того насколько

гладкий сигнал).Если ширина спектра периодического сигнала конечно,

то он описывается конечным числом гармоник .N-кол-во отсчетов на один период.

ДПФ строго описывает периодический сигнал с конечным спектром ( если это не

соблюдается ,то появляется ошибка в представлении сигнала ДПФ ).

N-1

дtе tt, где Т=Nt, t



Ґ

т ttdttфильтрующее свойство функции.

Ґ

Ґ

дt)=е Cnexpj2n Пара преобразований Фурье

Ґ

T

Cn=1тд(t)exp(-j2ntdt

0

t -1

Сn=1/Nt т еt-t)exp(-j2ntdt={сжали ось времени symbol 120 \f "Symbol" \s 10symbol 61 \f "Symbol" \s 10t/symbol 68 \f "Symbol" \s 10tsymbol 125 \f "Symbol" \s 10=

0 

N N-1 N-1 N

N т е exp-j2nn)d=N е  т exp(-j2nN)d

0 =0 =0 0

N-1

=1/N е exp(-j2n

=0

T=Nt

N-1

Cn=1/N е  exp(-j2n Пара дискретного преобразования Фурье

=0

N-1

 е Cn exp(j2n

0

Cn-комплексная гармоника, а N-кол-во отсчетов.

СВОЙСТВА ДИСКРЕТНОГО ПРЕОБРАЗОВАНИЯ ФУРЬЕ.

1. Линейность - если в цепи отклик на сумму воздействий равен сумме откликов.

Спектр суммы сигналов равен сумме спектров сигналов.

N-1

 е Сxn exp(j2n

0 Выборки двух сигналов.

N-1

 е Cyn exp(j2n)

0

Z , Линейность преобразования Фурье

Сzn=Cxn+Cyn ( для интегралов и сумм).

2. Для дискретного сигнала кол-во отсчетов спектра ( Сn) равно кол-ву

отсчетов сигнала.

3.Коэффициент (Со) дает постоянную составляющую.

N-1

Со=1/N е  ѕ это математическое ожидание.

=0

4. Если N-четное ,то тогда

N-1 

Cn/2=1/N е 



5. Если  - вещественные, то Cn ,расположенные симметрично

относительно Cn/2 образуют комплексно сопряженные пары.

N-1 N-1 

C =1/N е  exp(-j2n)/N)=1/N е  exp(j2N)=Cn

N-n =0 =0

Отсчеты выше C повторяют спектр от Co до C .

N/2 N/2

Но мы не нарушаем теорему Котельникова, т.к. Сn комплексное число,

оно требует два числа для своего представления. Следовательно нужно

ровно N отсчетов ,как и по Котельникову ( N=2FT=T/t).

ЦАП и АЦП.

1 3 5

4

2 ЦАП АЦП 2

+5в +15в +5в

6 7 6

1.Стробирующий импульс ( аналоговая величина, соответствующая дис-

кретному слову).

2. N-разрядное дискретное слово (код).

3.Опорное аналоговое напряжение (определяет от какого сигнала ведется

счет т.е. служит для получения единиц измерения в дискретных долях).

4.Аналоговый сигнал.

5.Пуск (внешний сигнал - для конкретного момента времени будет получен

код).

6.Логическое питание.

7.Аналоговое питание.

Отдельные земли обеспечивают подавление импульсных помех ( т.е.возрастает

помехоустойчивость) по питанию.

Входные и выходные сигналы ЦАП и АЦП.

Сигналы ЦАП АЦП

аналоговый на выходе ; напряжение вход ; напряжение ; полярность ;

или ток ; полярность ; ве- величина ; ( есть однополярные

личина ( бывают одно- и двуполярные АЦП ) ;

и двуполярные ЦАП ) (2.5В , 5В , 10В , 10.24В , 20В)

(2.5В,5В,10В,10.24В,20В)

(1мА,1.2мА,1.5мА,2.5мА)

цифровое вход ; послед. или парал. выход ; последовательный

слово ( шина ) включение ; или параллельный ;

логические уровни :

ттл-5В ; эсл- -5В,-2.5В ;

кмоп-3В,15В ; источник

питания : анал.±15,±12В ;

дискр.+5В .

сигналы стробирующий импульс а) входной импульс начала

управления ( при завершении ввода преобразования.

слова , т.е. тактовый ввод) б) вых. “состояние”

( говорит , что на выходе

появился код )

t между сигналами а

и б - это врнмя , затрачи-

ваемое АЦП на преобра-

зование.

опорный эталонное напряжение , эталонное напряжение ;

относительно которого внешнее , внутреннее ;

ведется счет ; можно использавать перемен-

ное

При преобразовании мы можем получать прямой код Uвых. ( 0-10В ), или

двуполярный ( ± 10В ). При использовании ЦАП и АЦП необходимо обра-

тить внимание на используемый код ( т.к. они различны ).

Однополярные : как правило старший разряд обеспечивает 0,5 Uопор. ,

n

следующий разряд 0,25 Uопор. , ... , младший 1/ 2 Uопор. .

Двуполярные : первый разряд дает знак , следующий 0,5 Uопор. ,

n-1

младший 1/2 Uопор. .

-0,51 ё -0,38 ® 000 Декодирование аналогового

-0,38 ё -0,26 ® 001 напряжения в бинарное число

-0,26 ё -0,13 ® 010 U = 0,128 - шаг квантования.

-0,13 ё 0 ® 011 Uразмаха = 1,024 В.

0 ё 0,13 ® 100 ( ошибка не больше 0,5 U ).

0,13 ё 0,26 ® 101

0,26 ё 0,38 ® 110

0,38 ё 0,51 ® 111

ОСНОВНЫЕ ТИПЫ ЦАП.

1.Цепочка взвешенных резисторов.

R R Rвх.оу№ ,Rключей№ (удается реали-

зовать ключи с сопротивлением R» Ом)

2R UвыхеIвх

4R Недостатки:

ОУ Слишком большой разброс сопротивлений

8R Uвых и как следствие трудность в изготовление

их на одной микросхеме .

Влияние Rвх.оу на цепь.

Uопорное

2.Цепочка R-2R .

+Uопорн

2R Uвых Достоинства:

Более технологична ,т.к.всего два

номинала сопротивлений.

-Uопорн 2R R Rн

2R R

2R R

2R

АНАЛОГО-ЦИФРОВЫЕ ПРЕОБРАЗОВАТЕЛИ.

А).АЦП последовательного приближения (скоростные).

i(t) УВХ Ком- схема Umax

пар. управл Uвх

1/2Umax

ЦАП RG

Uвых Umin

цап n

вых 1 2 3 4 5

1.При подачи пускового импульса , после УВХ обнуляется регистр (RG),

затем в старший разряд регистра дается “1”,на выходе ЦАП появляется

напряжение равное 0.5Umax .Если Uвых.увх>Uвых.цап ,то ”1” в старшем

разряде остается ( иначе она стирается ).

2.Опять ставят “1” в следующий разряд регистра .......................................

.......................................................................................................................

Кол-во шагов соответствует кол-ву разрядов АЦП.

Б).АЦП параллельного действия .

Uопорн. В качестве опорного на каждый

компаратор (К) подается сетка

УВХ R напряжений - Uопорн.

i(t) K n

2 n

R Kол-во компараторов = 2 .

K При подачи сигнала на вход АЦП ,

R ДЕКО- все компараторы у которых

K ДЕР Q2 Uопорн.<Uвх.

дадут единици.

R Q1 Наивысший номер компаратора

K находящегося в единичном состоя-

R Q0 нии соответствует выходному коду

K АЦП.

R Недостатки:

K Эти АЦП малоразрядные (4-5)

R (т.к.необходимо много компара-

K торов).

R

K

R

КОМПЛЕКСНОЕ ПРЕДСТАВЛЕНИЕ СИГНАЛОВ.

 t)= (t)+j(t)=E(t)exp{ j(t)} (t) 

 

(t)

пре-

образ.

При определенных правилах связи комплексное число дает аналитический сигнал

( т.е. непрерывный вместе со своими производными ) .

(t)

E(t) Описание сигнала через огибающюю E(t)

и фазу (t) .

(t)

0 (t)

(t)=Re{(t)}=E(t)cos(t)



Нам нужно установить правило выбора сомножителей E(t) и cos((t)) т.к.

если мы узнаем один , то легко найдем другой .

Понятие огибающей очень расплывчато , поясним это на примерах :

( ) -огибающие для одного процесса

разные .

Первый дал понятие огибающей и фазе Гильберт , он дал определение

мнимой составляющей ( т.е. ввел комплексные величины ) .

Ґ

(t)=1/ тi( )/ t- d

Пара преобразований Ґ

Гильберта Ґ

(t)=1/ т ()/ t-. d

Ґ

Преобразование Гильберта - широкополосный фазовращатель , оно

поворачивает все спектральные составляющие на 90° .

ѕѕѕѕѕ

E(t)= Ц (t) + (t) - огибающая понятия применимые

для любого сигнала .

(t)=arctg[ (t)/ (t)] - фаза

(t)=d(t) - частота

dt

(t)=Acosw t ; (t)=Asinw t ( т.е. (t) получается приповороте (t)

на 90° ).

(t)= Acosw t +Asinw t = A

Схема получения АМ ОБП .

 1/2cos(w -)t+1/2cos(w + )t

(t) (t)cosw t

генератор

cosw t

cos(w - )t

+

=/2 =/2

sint sinw t (t)sinw t

1\2cos(w - )t- 1/2cos(w +)t

+ Получили АМ ОБП без использования фильтров .

Мы оперируем комплексными функциями для того

чтобы убрать основную часть энергии несущей .

Огибающие и фаза УПСП (узко-полосного случайного процесса ).

Квадратурные составляющие огибающей .

<<

460 465 470 f,кГц

(t) = t- (t)

 - ( ) (t)

(t)- ( )

t t

Фаза УПСП разбивается на две составляющие флуктуированную (t)

и мат.ожидания t .

(t) =Е(t)cos(t)=E(t)cos(t -(t))=E(t)cos(t)cost+E(t)sin(t)sint

A(t) B(t)

A(t) и B(t) медленно меняющиеся функции . Получаются , как случайные

функции времени .

(t)=A(t)cost + B(t)sint , где A(t) и B(t) - квадратурные составляющие

огибающей .

В этом колебание вектор Е(t) будет колебаться , т.е. показывать флуктуацию.

A(t)

E(t)

(t)

B(t)

Свойства функций :

1. Энергетические спектры G () иG () одинаковые .

2. Законы распределения одинаковые  (x)= (x)=a(x)=б(x).

3. Коррелляционные функции равны B( )=B ( ) .

4. Справедливо свойство ортогональности .

ѕѕѕѕ ѕѕѕ

(t)(t)=0 A(t)B(t)=0

5.Ґ <=A(t) < Ґ ; Ґ <=B(t)<Ґ;E(t)>=0 .

ѕ ѕ

6. Если Гауссовский шум то A(t)=0 и B(t)=0

( Т.е. нулевые мат. ожидания ) .

Если A(t)=F то это значит что в случайном процессе

появилась детерменированная ф-ия .

(t)=A(t)cost + B(t)sint+ Fcost

7. A (t)=B (t) =G - мощность реализации .

ѕ ѕ ѕ

E (t)= A (t)+B (t) =2G - мощность огибающей .

8. Ba()=Bб() ( т.к. скорости изменения одинаковы )

9. B()=Ba()cost

ДИСКРЕТНАЯ СВЕРТКА.

Ґ

f(t)=тy(td - Свертка -интеграл Дюамеля (прохождение

Ґ сигнала через нелинейную инерционную

цепь)

N-1

fm=1/N е m- - Свертка дискретных сигналов.

 m=0,1,2,3,...,N-1.Т.к.число отсчетов описывающее

сигнал Х(t) ,будет описывать и функцию fn.

N-1

еСxn exp(j2N) ;Cxn-амплитуда “n”-ой гармоники спектра.

n

N-1

Ym-е Cyl exp(j2N)

l=0

N-1 N-1 N-1

fm=1/N е  е Cxn exp(j2/N) е Cyl exp(j2lm-

 n=0 l=0

N-1 N-1 N-1

1/N е е CxnCyl exp(j2lm/N) е exp(j2n-l)

n=0 l=0 =0

N-1

При n=l , е exp(j2n-l)=N (Если n№l ,то сумма равна “0”).



Тогда получаем:

N-1

fm= е Cfn exp(j2mn/N) ,где Cfn=CxnCyn

n=0

Если в одном из пространств пары преобразования Фурье мы

производим умножение ,то во втором пространстве будет про-

изводиться свертка .Это требуется для анализа длинной после-

довательности ,где легче перемножить спектры ,а потом взять

обратное преобразование Фурье .

 2 2 2 Y 3

2

1

-1 0 1 2 -1 0 1 2

mY(0-m) еXmY(1-m)

еXmY(2-m) еXmY(3-m)

еXmY(4-m)

fm

12

6

0 1 2 3 4 m

**4.2.2. Дискретизация и квантование изображений**

Сформированное и записанное изображение необходимо преобразовать в форму, пригодную для цифровой обработки. Если изображения записываются фотоэлектронным способом, то это обычно не составляет трудности, так как из сканирующего фотоэлемента поступает электрический ток, пригодный для дискретизации и квантования. Таким образом, данный случай можно рассматривать как распространение соответствующих методов цифровой обработки одномерных сигналов на двумерные сигналы. При этом ошибки квантования можно учесть введением в блок-схему дополнительного .источника шума [11]. Расстояние между отсчетами должно удовлетворять теореме Найквиста для двумерных колебаний [1].

Устройства для дискретизации и квантования изображений основаны на технике микроденситометрии. В подобных системах на пленку проектируется луч света с интенсивностью *I1*. Интенсивность *I2* света, прошедшего сквозь пленку (или отраженного от нее), измеряется фотоумножителем. По *коэффициенту пропускания*

*Т=* (4.16)



с помощью соотношения (4.5) можно вычислить оптическую плотность. После этого световое пятно на пленке можно сместить скачком и таким образом получить отсчеты изображения. Математически этот процесс описывается соотношением

*g1(x, y)* = (4.17)



где g - изображение на пленке; *ha* распределение яркости в сечении луча, освещающего пленку; *g*1 эквивалентное изображение, из которого берутся отсчеты (т.е. в дискретных точках x = jx, y = ky сканирующий фотоприемник измеряет именно g1 *).* Матрица отсчетов *g*1  *(* jx, ky *)* представляет собой *дискретизованное,* или *цифровое,* изображение.



Из равенства (4.17) (справедливого также для случая дискретизации изображений, полученных фотоэлектронными средствами) видно, что в процессе дискретизации записанное изображение подвергается искажениям. За счет правильного выбора распределения *ha* и расстояния между отсчетами изображение можно фильтровать в процессе дискретизации. Фильтрацию, связанную с процессом дискретизации [согласно формуле (4.17)], можно использовать для подавления эффектов наложения, возникающих из-за того, что ширина спектра изображения обычно не ограничена (из-за шума зернистости пленки и других высокочастотных составляющих) [12]. Дискретизация коэффициента пропускания эквивалентна дискретизации яркостного изображения, а дискретизация плотности эквивалентна дискретизации плотностного изображения. Часто можно услышать, что предпочтительнее квантовать плотность, так как логарифмическая зависимость приводит к уменьшению динамического диапазона. Однако подобные упрощенные рассуждения могут приводить к ошибкам [13].

**4.2.3. Восстановление и демонстрация цифровых изображений**

При цифровой обработке одномерных сигналов восстановление аналогового сигнала из последовательности чисел достигается путем низкочастотной фильтрации, что теоретически обосновывается теоремой об интерполяции колебаниями с ограниченным спектром [11]. В идеальном случае для такой интерполяции следует применять функцию вида sin *.* Однако данная функция не имеет двумерного варианта, который можно было бы использовать для восстановления аналоговых изображений, так как импульсный отклик идеального фильтра нижних частот, имеющий вид sin*,* принимает отрицательные значения, а это выдвигает требование получения отрицательного света, невыполнимое при восстановлении изображений.



Аналоговое изображение можно восстановить с помощью устройства, подобного примененному при дискретизации изображения. На чистую пленку проектируется луч света, а интенсивность этого записывающего луча модулируется в соответствии с числовыми значениями изображения. В качестве источника света, а также для непосредственной демонстрации изображений можно также применять электронно-лучевые трубки (ЭЛТ). Световое пятно перемещается по поверхности пленки согласно растровой сетке. Нетрудно увидеть, что процесс восстановления изображения описывается соотношением

*g2(x, y)* = (4.18)



где *hd -* распределение яркости записывающего пятна, g1 - матрица отсчетов функции (4.17), представленная здесь набором взвешенных - импульсов, разнесенных на расстояния (x, y*)* друг от друга, а g2 - восстановленное непрерывное изображение. Распределение яркости записывающего пятна является импульсным откликом интерполирующего фильтра, аналогичного применяемому при восстановлении одномерных аналоговых сигналов. Практически во всех системах восстановления изображений записывающее пятно имеет несложное распределение яркости (например, гауссовское). По этой причине точно восстановить изображение не удается, поскольку простые распределения не дают возможности полностью подавить высокочастотные копии спектра изображения, возникающие при дискретизации. К счастью, обычно это не создает существенных трудностей, и в простых системах получается хорошее изображение.



Из вышеизложенного видно, что при дискретизации и демонстрации изображений возникают искажения спектра. Подобные искажения можно скорректировать в процессе цифровой фильтрации квантованных изображений [12].

Восстановление непрерывных изображений связано с другой проблемой, а именно с проблемой верности изображения. Если число, находящееся в памяти машины, представляет значение оптической плотности изображения в конкретной точке, то абсолютно *верное воспроизведение* получится, если пленка, предназначенная для демонстрации, будет иметь точно такую оптическую плотность, как записано в памяти ЭВМ. (Аналогичные требования можно сформулировать для коэффициента пропускания пленки я для характеристики фотоэлектронной системы ). Подобное устройство

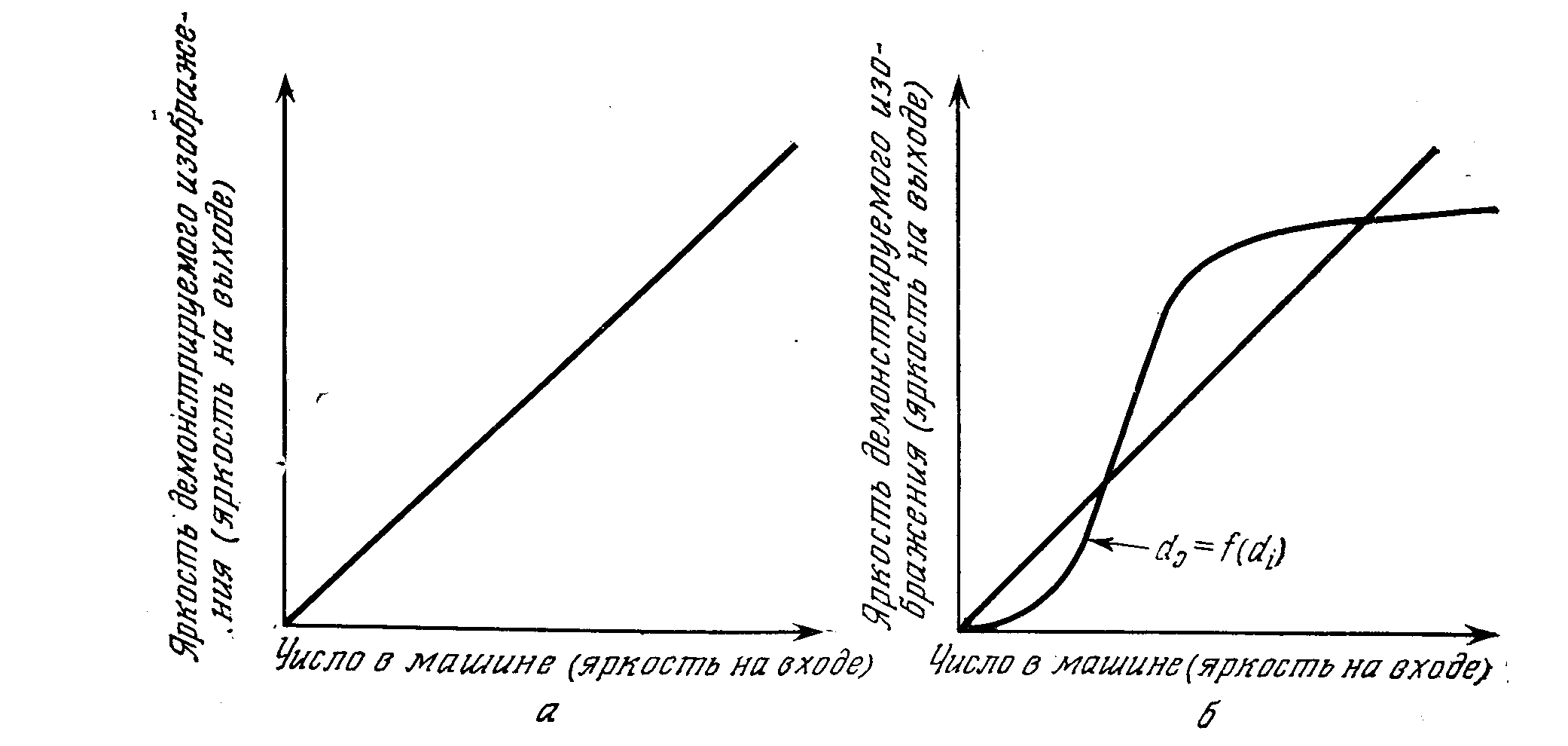


Рис. 4.4.

*а -* сквозные характеристики идеальной системы отображения; б - сквозные характеристики типичной реальной системы отображения.

отображения должно иметь сквозные характеристики» совпадающие с приведенными на рис. 4.4, *а.* Однако такие идеальные характеристики встречаются редко. Характеристики реальных устройств отображения больше напоминают приведенные на рис. 4.4, *б,* где наблюдается значительное отклонение от идеальной прямой с наклоном 45°. Хорошее приближение к идеальной характеристике можно получить путем линеаризации характеристики устройства отображения. Для этого необходимо выполнить следующие действия:

1. Сформировать набор фиксированных значений коэффициента пропускания или оптической плотности, подать их на устройство отображения и измерить фактический его отклик на каждое из значений коэффициента пропускания или оптической плотности.

2. Измерения, полученные на этапе 1, задают *характеристику устройства отображения d0 = f ( di ).* Линеаризованная характеристика описывается соотношением *di = f- -1( d0 ).* Это обратное преобразование можно найти эмпирическим путем и представить в виде таблицы или полинома, вычисленного методом наименьших квадратов.

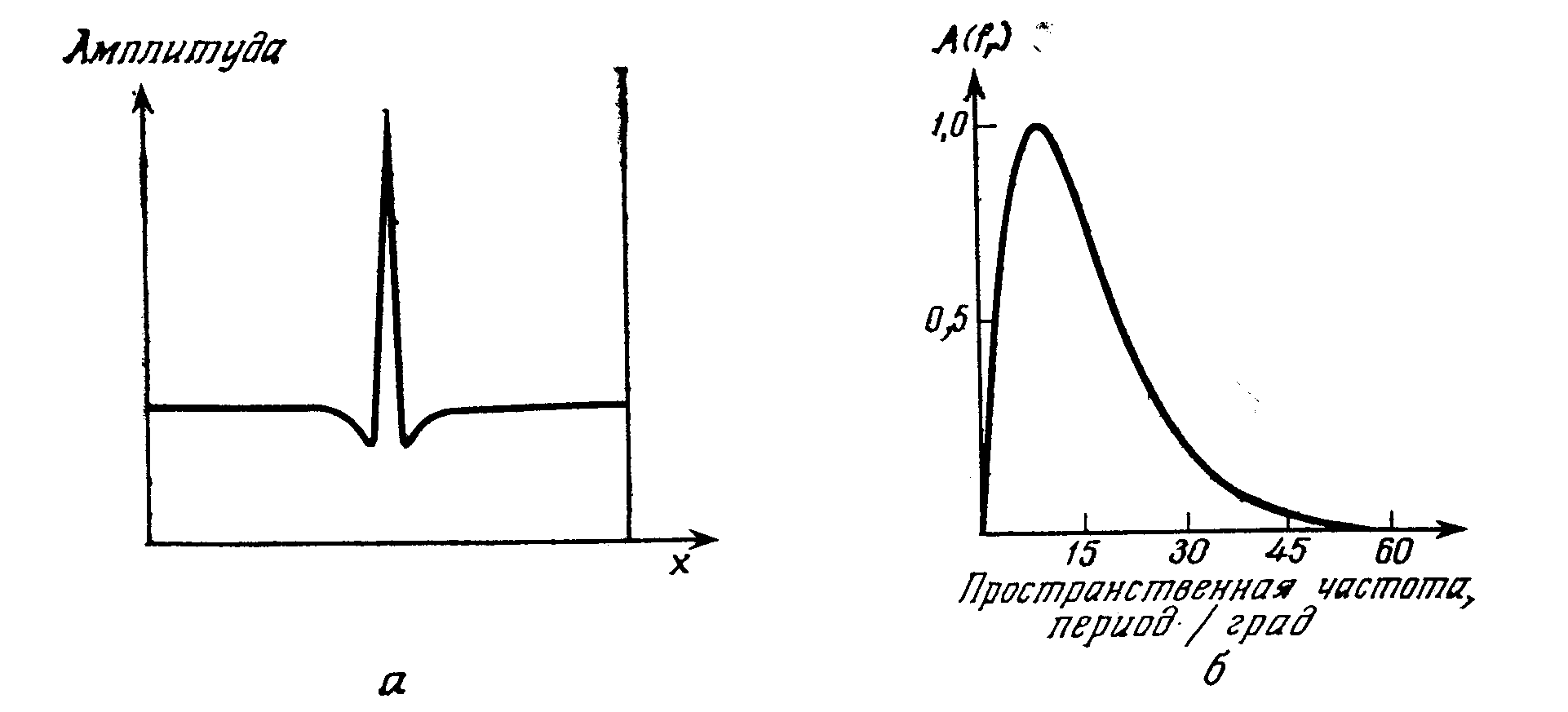
3. Перед демонстрацией изображения числовые данные следует преобразовать согласно функции *f -1.* В результате в них будут введены предыскажения и значения яркостей, записанные в машине, будут воспроизведены на экране без погрешностей.

Метод линеаризации характеристик устройств отображения успешно применялся во многих исследовательских институтах. Точная линеаризация, конечно, невозможна, так как форма нелинейной характеристики изменяется в зависимости от особенностей проявления пленки, чистоты химикатов, старения (или повреждения) люминофора ЭЛТ и т.п. Однако, приложив определенные усилия, можно линеаризировать устройство отображения так, что отклонения от линейности не будут превышать ±5% максимального значения. Следует отметить, что линеаризация характеристики устройства отображения является операцией, применяемой при восстановлении аналоговых изображений; при обработке одномерных сигналов линейными электронными схемами она обычно не используется.

**4.2.4. Свойства системы зрения человека**

Очень часто окончательную оценку изображения делает человек. Если бы человеческое зрение было идеальным и отвечало на световое воздействие с абсолютной точностью и совершенной линейностью, то его можно было бы и не изучать. Однако система зрения человека обладает нелинейной характеристикой, а ее отклик не является абсолютно верным. Важность этих положений .для получения изображений признана довольно давно [13], однако они не в полной мере использовались при обработке изображений.

Одной из характеристик системы зрения человека является способность восприятия яркости света. Эксперименты по определению восприятия людьми минимально различимых градаций яркости света, поступающего от калиброванного источника, показали, что яркость света воспринимается глазом нелинейно. Если начертить график зависимости величины этой минимально различимой градации яркости от эталонной яркости, то при изменении яркости в пределах нескольких порядком этот график имеет логарифмический характер [14]. Такие субъективные экспериментальные результаты согласуются с объективными данными, полученными в экспериментах на животных, в которых было показано, что светочувствительные клетки сетчатки и оптический нерв возбуждаются с частотой, пропорциональной логарифму интенсивности подводимого к ним света [15]. По вполне понятным причинам подобные объективные измерения на людях не проводились. Тем не менее объективные данные для животных и субъективные показания для людей более чем убедительно подтверждают вывод о том, что яркость света



**Рис. 4.5.**

*a* - сечение (осеcимметричной) аппаратной функции глаза человека; б - сечение (осесимметричной) частотной характеристики глаза человека.

воспринимается по логарифмическому закону. Это существенно нелинейный закон.

Другой отличительной характеристикой системы зрения человека является ее пространственночастотный отклик. Импульсный отклик глаза, рассматриваемого в виде двумерной линейной системы (т.е. линейной после начального логарифмического преобразования интенсивности наблюдаемого света), не является - функцией Дирака. Реакция глаза на приходящее световое поле описывается аппаратной функцией, сечение которой показано на рис. 4.5, а [16]. Острый центральный пик и отрицательные боковые лепестки импульсного отклика глаза показывают, что глаз обрабатывает пространственные частоты так же, как фильтр верхних частот. Точная форма частотной характеристики глаза исследовалась с помощью ряда психовизуальных экспериментов; было показано, что глаз подавляет низкие и ослабляет высокие пространственные частоты. В грубом приближении пространственно-частотный отклик глаза имеет полосовой характер. Подобная характеристика ( рис. 4.5,б ), например, была получена в ряде экспериментов, проведенных Манносом и Сакрисоном [17].



Наконец, особенностью зрения человека является способность к насыщению, т.е. к ограничению отклика при очень больших или очень малых интенсивностях наблюдаемого светового потока. Перечисленные свойства системы зрения можно описать моделью, представленной в виде блок - схемы на рис. 4.6. Однако данная модель совершенно не отражает других известных свойств системы зрения. Например, есть сведения, что некоторые стороны процесса восприятия изображения можно объяснить только наличием не одной, как на рис. 4.6, а нескольких линейных систем, включенных параллельно, т.е. в рамках модели с частотными каналами [18]. Другие визуальные явления (такие, как иллюзия одновременного контраста) указывают, что логарифмическое преобразование, введенное в блок-схеме рис. 4.6, является слишком большим упрощением. Но, несмотря на известные недостатки, модель, представленная на рис. 4.6, является полезной, поскольку она

1) объясняет ряд важных явлений, таких, как восприятие яркости света и полосы Маха;

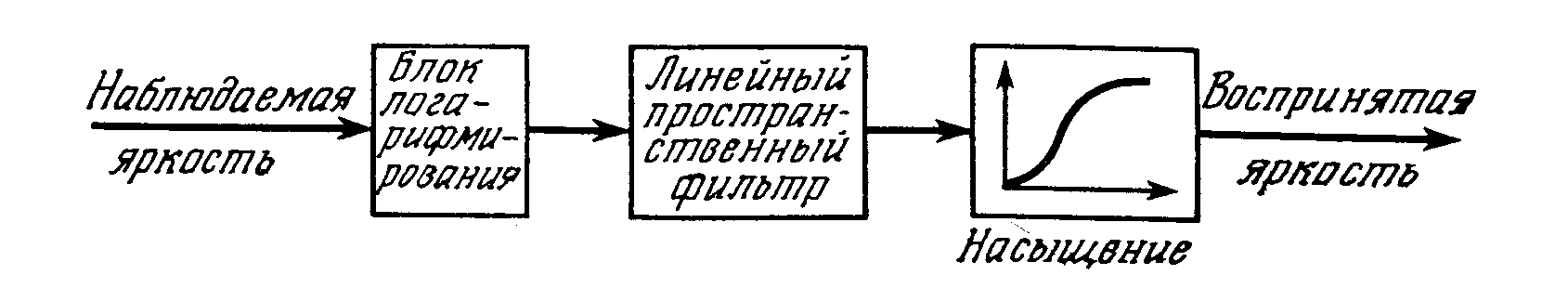


Рис. 4.6. Блок-схема системы зрения человека.

2) указывает, что в системе зрения содержатся некоторые элементы системы обработки информации. В частности система зрения человека, по-видимому, выполняет некоторые операции гомоморфной обработки информации [19].

Полезно связать логарифмическое преобразование изображения, выполняемое глазом, с рассмотренным ранее вопросом о плотностном (и яркостном изображениях. Можно заметить, что поскольку яркость света воздействует на глаз по логарифмическому закону, глаз воспринимает изображение как плотностное, если даже оно представлено (с помощью устройства отображения) в виде яркостного изображения.

Представляется логичным воспользоваться моделями системы зрения человека при анализе возможных применений цифровой обработки изображений. Однако это делать нужно осторожно, так как система зрения человека настолько сложна, что необоснованное применение упрощенных моделей зрения может принести больше вреда, чем пользы. Манное и Сакрисон [17] доказали применимость модели зрения для исследования вопроса о сокращении избыточности изображений. Однако пока еще не определены все области возможного применения моделей зрения.

**4. 3. Применение цифровой обработки для сокращения избыточности изображений**

Сокращение избыточности изображений является первой областью применения цифровой обработки изображений, которая будет здесь рассмотрена. Интенсивное развитие цифровых методов повлияло на все отрасли техники передачи и хранения информации в силу присущих цифровым системам преимуществ в помехозащищенности, возможности исправления ошибок, гибкости при коммутации сообщений, постоянно понижающейся стоимости и увеличивающейся надежности. Одновременно с внедрением цифровой техники расширялось использование изображений в различных областях науки и техники, например в медицине, .экспериментальной физике, бесконтактной дефектоскопии, исследовании природных ресурсов. Такая параллельность развития цифровой техники и расширения области применения изображений привела к естественному результату, а именно к интенсивным исследованиям в области передачи ;и записи изображений цифровыми методами.

Типичное изображение содержит очень много избыточной информации, что заметно даже при беглом взгляде на большинство изображений. Эта избыточность приводит к экономическим потерям. Ширина полосы частот, необходимая для передачи изображения в цифровой форме, зависит от числа отсчетов изображения, разрядности отсчетов, времени, отведенного на передачу, и от мощности передатчика. С расширением полосы увеличиваются необходимая мощность передатчика и расходы. Деньги и энергия не являются проблемой, но электромагнитный спектр предельно загружен. Поэтому сокращение избыточности при передаче изображений является очень важной задачей. Столь же важно оно и для хранения .изображений в цифровом виде. Если бы требовалось хранить только одно изображение, то об этом можно было бы не беспокоиться. Однако во многих существующих и проектируемых системах, таких, как геологоразведочный спутник NASA ERTS (Earth Resources Technology Satellite), получается большое количество изображений, которые целесообразно получать и хранить в цифровой форме. Хотя цифровые запоминающие устройства и становятся все дешевле, число получаемых изображений настолько увеличивается, что сокращение их избыточности является задачей первостепенной важности.

**4.3.1. Некоторые замечания о сокращении избыточности изображений**

Избыточность видеоинформации может быть описана функцией корреляции между отсчетами изображений; она проявляется в высокой степени взаимной статистической прогнозируемости близколежащих отсчетов, взятых из изображения. Конечной целью операции сжатия видеоинформации является устранение этой статистической прогнозируемости (т.е. необходимо в максимально возможной степени уменьшить коррелированность отсчетов). На блок-схеме рис. 4.7 показаны основные операции, выполняемые системой сжатия видеоинформации. Сначала выполняется операция по максимальному уменьшению коррелированности отсчетов изображения. Затем отсчеты должны быть соответствующим образом квантованы. Квантованные отсчеты кодируются в форму, благоприятную для передачи (при этом, конечно, может быть обеспечена возможность обнаружения или исправления ошибок) .

Квантование и кодирование выполняются с учетом общих правил, не зависящих от особенностей схемы декорреляции, выбранной для первого этапа обработки. Поэтому системы сжатия видеоинформации различаются видом схемы, выполняющей операции, относящиеся к первому этапу. В силу этого способам реализации первого блока схемы рис. 4.7 здесь будет уделено больше внимания, чем вопросам

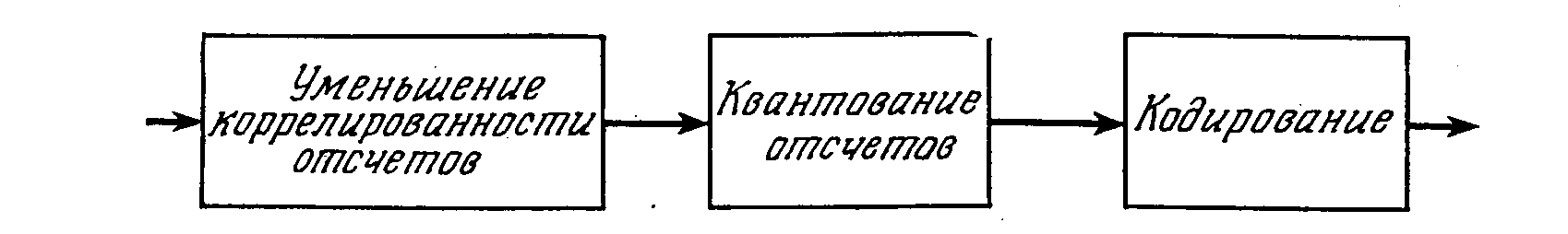


Рис. 4.7. Блок-схема системы сокращения избыточности

видеоинформации.

построения второго и третьего блоков. Такой подход полностью соответствует замыслу данной книги, посвященной техническим применениям цифровой обработки сигналов, т.е. задачам, в основном относящимся к первому блоку.

При разработке принципов реализаций первого блока схемы рис. 4.7 следует учесть ряд соображений. Рассмотрим сначала статистические свойства изображений. Если отсчеты изображения образуют сетку точек размером *NN* и каждый отсчет представлен Р - разрядным двойным числом, то при записи и передаче изображения с помощью обычной импульсно-кодовой модуляции (ИКМ) потребуется *N2P* двоичных разрядов. Однако, как было отмечено выше, типичное изображение имеет большую избыточность. Один из способов, позволяющих измерить эту избыточность и сравнить ее с номинальным числом *N2P* разрядов, заключается в построении гистограммы яркости изображения и вычислении соответствующей энтропии. С помощью Р- разрядных чисел можно описать квантованяе по 2p уровням. Для этого следует проанализировать все *N2* отсчетов и подсчитать, сколько раз встречается каждый уровень квантования. Затем следует построить гистограмму яркости изображения, т.е. для каждого уровня квантования указать число его появлений в изображения. Разделив эти числа на общее число точек *N2* , можно получить аппроксимацию плотности вероятности процесса, порождающего изображение. Если обозначить нормированные частоты через *pi* ( i = 1, 2, ... , 2p ), то энтропия по определению выражается суммой



*h* = \_\_ (4.19)



и равна средней информации (измеряемой числом бит, приходящихся на элемент изображения), содержащейся в каждом элементе изображения. Анализ изображений показал, что типичное значение *h* гораздо меньше числа разрядов *Р,* необходимого для стандартного представления с помощью ИКМ. В работе [20] отмечалось, что энтропия имеет величину порядка 1 бит/точка. Это означает, что разрядность массива, описывающего изображение, можно (хотя бы теоретически) сократить без потерь информации в среднем до 1 бит/точка.

Энтропия служит мерой статистической избыточности, но не дает сведений о ее происхождении. Источником избыточности, как подсказывает наблюдателю его зрение, является высокая степень однородности изображения на малых участках. Эту пространственную избыточность можно определить с помощью ковариационной матрицы изображения. Сначала лексикотрафически преобразуют матрицу из *NN* отсчетов изображения в *N2* - компонентный вектор [т.е. элементы первой строки (или столбца) матрицы *g( j, k )* становятся компонентами вектора с номерами от 1 до *N,* элементы второй строки (столбца) — компонентами с номерами от N+1 до 2 *N* и т.д.]. Затем вычисляют ковариационную матрицу изображения



*[ Cg ] = E { (g - E(g))(g - E(g))T } ,* (4.20)

где *Е -* среднее значение но ансамблю, a g - вектор, построенный из отсчетов изображения. На практике редко оказывается возможным проводить усреднение по ансамблю и ковариационную матрицу получают с помощью оценки пространственной корреляции [21].

Ковариационные структуры, такие, как матрица [ *Cg* ]*,* не имеют взаимно - однозначной связи с исходным изображением. Коул [21] показал, что многие неодинаковые изображения могут оказаться весьма сходными в ковариационном (или спектральном) смысле. Следовательно, имеются основания для замены сложной матричной структуры типа [Cg ] более простой. В частности, рассматривалось (см., например, работу [22] ) применение модели с авторегрессионным марковским процессом *п-го* порядка, где *п* обычно невелико (например, n = 3). Тот факт, что подобные модели оказываются корректными, и применение их оправдано при анализе методов сжатия информации, таких, как дифференциальная импульсно - кодовая модуляция (ДИКМ), указывает на высокую степень взаимосвязи между соседними участками изображения.

При сжатии видеоинформации кроме статистических свойств изображения весьма важно учитывать и особенности получателя изображений. Зрение человека обладает ограниченными возможностями и характеризуется некоторыми известными (отчасти) отличительными особенностями. Использование, конкретных особенностей зрения для сокращения избыточности изображений называется *психофизической обработкой.* Известно, например, что при восприятии яркости света, попадающего в глаз, зрительная система ведет себя как нелинейная система с логарифмической характеристикой. Кроме того, система зрения человека не чувствительна к очень высоким или очень низким пространственным частотам, а в области средних частот ведет себя почти как полосовой фильтр, что обусловлено торможением нервных клеток сетчатки глаза. Нелинейность и частотная зависимость чувствительности зрительной системы позволили создать оптимальные системы сжатия видеоинформации. В этих системах для достижения большей устойчивости к ошибкам, появляющимся при кодировании и передаче, изображение обрабатывается примерно так же, как и в зрительной системе человека. Впервые это предложение было сделано Стокхэмом [23].

Сокращение избыточности информации математически строго обосновывается положениями теории кодирования при заданном критерии точности [24]. Как отмечали Маннос и Сакрисон [17], эффективные теоремы теории кодирования при заданном критерии точности в задачах сжатия видеоинформации применить не удалось. Основной причиной этого явилась сложность выбора критерия допустимой величины ошибок, согласующегося со свойствами системы зрения человека. Манное и Сакрисон смогли показать, что можно пользоваться критерием, связанным с нелинейными и пространственно-частотными свойствами зрения. Их работа имеет очень важное значение для дальнейшего развития методов сокращения избыточности изображений. Введение подходящей предварительной обработки во всех схемах, которые будут рассмотрены ниже, может значительно улучшить качество работы систем сжатия видеоинформации.

**4.3.2. Схемы сокращения избыточности изображений с обработкой в пространственной области**

В одном из возможных вариантов схемы сокращения избыточности видеоинформации в первом блоке (схема рис. 4.7) выполняется операция тождественности, т.е. исходная картинка никак не изменяется, а все сжатие достигается за счет квантования и кодирования. Однако сжатие информации невозможно выполнять без использования критериев, учитывающих особенности наблюдателя и свойства передаваемых данных. Если, например, наблюдателю нужна точность 1/1000 , то необходимое число уровней квантования получается при использовании 10-разрядных двоичных чисел; если же допустима точность 1/8 , то достаточно взять 3-разрядные числа. Следовательно, квантование при сжатии информации играет ограниченную роль. Однако сокращения избыточности можно добиться при кодировании, и одной из основных задач после создания Шенноном теории информации было построение кодов, оптимальных с точки зрения сокращения избыточности информации. Шеннон доказал, что существует код, для которого скорость передачи совпадает со скоростью создания информации источником. Таким образом, для изображений с энтропией порядка 1 бит/точка существуют схемы кодирования, позволяющие построить коды со средней длиной в 1 бит/точка. К сожалению, само по себе существование таких кодов бесполезно, если отсутствуют алгоритмы их построения. Известны алгоритмы построения кодов, приближающихся к оптимальным. Например, кодирование по Хаффмену является эффективной процедурой для согласования кода со статистикой источника информации и позволяет сократить длину сигнала по сравнению со стандартной ИКМ. Однако подобные коды имеют переменное число сим1волов (т.е. при передаче сообщений кодовые слова состоят из различного числа символов); при кодировании и декодировании требуются сложные алгоритмы, связанные с записью, синхронизацией и вспомогательным накоплениям информации. Кроме того, вид подобных кодов очень сильно зависит от вероятности создания символов источником, и любые изменения вероятности могут привести к ухудшению характеристик кода (очень значительному в некоторых случаях). Следовательно, кодирование с квантованием может служить основным средством сжатия видеоинформации лишь в ограниченном числе случаев, так что необходимо искать другие методы.

В качестве метода сжатия видеоинформации в плоскости пространственных координат, выполняемого )в первом блоке схемы рис. 4.7, наиболее широко применяется *дифференциальная импульсно-кодовая модуляция* (ДИКМ). По своей структуре схемы ДИКМ совпадают со схемами кодирования методом линейного предсказания (КЛП), применяемым при сжатии полосы речевых сигналов, и поэтому схемы ДИКМ изображений иногда называют схемами сжатия методом предсказания. Блок-схема ДИКМ приведена на рис. 4.8. В этом методе используется статистическая взаимосвязь яркостей отдельных точек изображения и для каждой точки формируется оценка яркости в виде линейной комбинации яркостей предшествующих точек. Под *предшествующими точками* подразумеваются точки, расположенные перед рассматриваемой точкой при развертке изображения сверху вниз и слева направо (как в телевидении), благодаря чему создается вполне определенный порядок следования точек изображения. Подобная схема, конечно, будет применима и тогда, когда изображение уже «развернуто» методом сканирования. Затем вычисляется и квантуется разность между фактическим значением яркости и ее оценкой. Квантованная разность подвергается кодированию и передается по каналу. На приемном конце символы декодируются, а информация восстанавливается с помощью схемы линейного предсказания *n*-го порядка (конечно, идентичной соответствующей схеме на передатчике), в которой формируются оценки яркости, добавляемые к разностям, полученным по каналу.

Схемы предсказания, изображенные на рис. 4.8, называются *схемами с предсказанием назад ,*  поскольку квантование сигнала

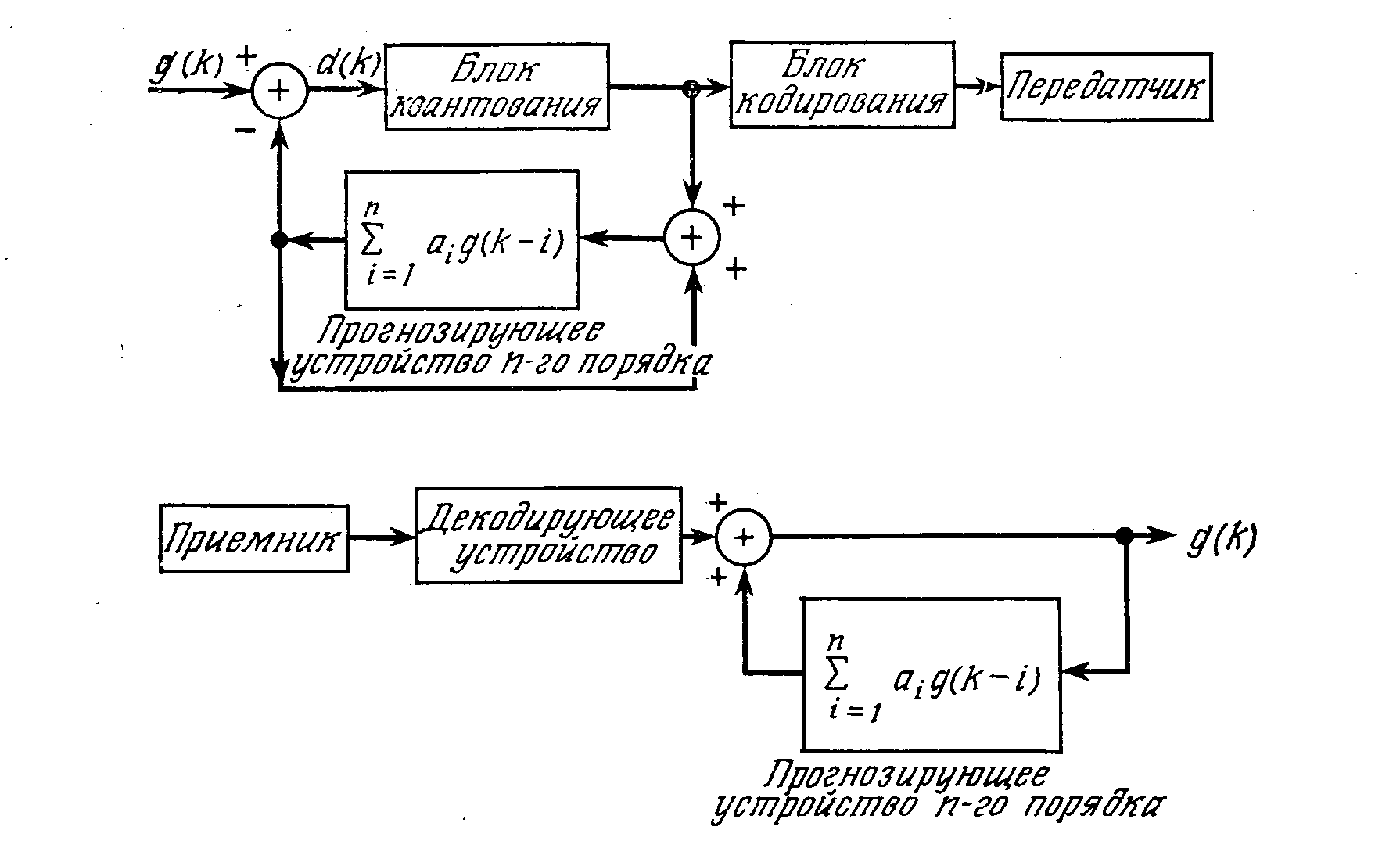


Рис. 4.8. Блок-схема системы сжатия методом ДИКМ с предсказывающим устройством *n-го* порядка.

происходит внутри петли обратной связи, а при восстановлении сигнала предсказанное значение подается по схеме назад. Можно спроектировать схемы ДИКМ, в которых предсказанные значения сигнала подаются вперед, а также создать схемы ДИКМ, где блок квантования расположен вне петли обратной связи. Однако такие системы дают восстановленное изображение с большими ошибками. Схема с предсказанием назад необходима в *приемнике* потому, что символы поступают последовательно. При использовании в *передатчике* аналогичной схемы предсказания назад в случае отсутствия ошибок, связанных с квантованием, можно было бы восстановить изображение с абсолютной точностью. Если схему квантования включить в петлю предсказывающей схемы передатчика, то и в приемнике, и в передатчике предсказание будет осуществляться на основе одинаковых квантованных отсчетов, что позволит уменьшить ошибки восстановления.

Сжатие в схемах ДИКМ достигается за счет вычитания сигналов, поскольку разности имеют значительно меньший динамический диапазон. Предположим, например, что исходное изображение передается методом ИКМ и для представления яркостей его точек нужны числа от 0 до 255. Тогда, если допустимая ошибка равна единице младшего разряда, то необходимо квантование в 8-разрядные числа. Однако значения разностей яркостей соседних точек будут гораздо меньшими; если разности (в том же масштабе) будут изменяться от 0 до 7, то для получения ошибки, равной единице младшего разряда, достаточно квантования в 3-разрядные числа.

Поскольку идея ДИКМ достаточно проста, то, как следует из схем рис. 4.8, характеристики системы сокращения избыточности изображений методом ДИКМ определяются [порядком предсказывающего устройства *п,* значениями коэффициентов прогнозирования *аi,* числом уровней квантования и их расположением.

Порядок предсказывающего устройства зависит от статистических характеристик изображения. Как правило, если последовательность отсчетов может быть промоделирована авторегрессионным марковским процессом *п-го* порядка, то разности, полученные с помощью оптимального предсказывающего устройства *п-го* порядка, будут образовывать последовательность некоррелированных чисел [20]. Изображения, очевидно, не являются марковскими процессами *п-го* порядка, но опыт работы по сжатию изображений показывает, что корреляционные свойства изображений можно описать марковским процессом третьего порядка, а это приводит к предсказывающим устройствам третьего порядка *(п=3)* [22]. Аналогично при моделировании изображений было выяснено, что ДИКМ с предсказывающими устройствами более высоких порядков не дает большего выигрыша в качестве изображения (как по субъективным, так и по объективным данным).

Коэффициенты предсказания *аi* можно определить с помощыо анализа средних квадратических ошибок. Пусть *g(k) -* отсчеты на строке развертки, a *(k) -* предсказанные значения этих отсчетов. Необходимо, чтобы средняя квадратическая ошибка была минимальна, т.е. нужно найти



*min e = E { g(k) - }* (4.21)



по всем k, аi

Это известная задача, и если процесс *g(k)* стационарен, то ее решение имеет вид [25]

*,* (4.22)



где

*r ( j - i ) = E [ g ( k - j ) g (k -i ) ]* (4.23)

обычно называется автокорреляционной функцией процесса *g.* Коэффициенты *ai* получаются решением системы уравнений (4.22).

Оптимальные значения коэффициентов предсказания зависят от взаимосвязей точек изображения, описываемых автокорреляционной функцией. Из определения (4.20) видно, что в случае стационарных данных автокорреляционная функция отличается от вышерассмотренной функции на постоянную величину. При нестационарных данных функция *r* (в уравнении (4.23) зависит от пространственных переменных и оптимальные коэффициенты предсказания должны изменяться в зависимости от пространственных координат. Это характерно для изображений. К счастью, нестационарные статистические характеристики изображений обычно можно достаточно хорошо аппроксимировать стационарными функциями, так что неперестраивающееся линейное устройство предсказания дает вполне хорошие результаты. При сжатии видеоинформации методом ДИКМ ошибки обычно появляются на границах изображаемых предметов, где предположение о стационарности удовлетворяется в наименьшей степени, и на восстановленном изображении воспринимаются визуально как аномально - светлые или темные точки.

Выбор числа уровней квантования и расположения порогов квантования является задачей отчасти количественной и отчасти качественной. Расположение порогов квантования можно найти количественными расчетами. В работе Макса [26] впервые было рассмотрено неравномерное квантование, зависящее от функции распределения квантуемого сигнала и сводящее к минимуму среднее квадратическое значение ошибки, вызванной ограниченностью числа уровней квантования. Алгоритм Макса позволяет найти оптимальное расположение точек перехода для заданного числа уровней квантования. Однако число уровней квантования выбирается исходя из субъективных качественных соображений.

Минимальное число уровней квантования paвно двум (одноразрядные числа) и соответствует такому квантованию изображений, при котором разность яркостей принимает фиксированное (положительное или отрицательное) значение. Этот способ обычно называют *дельта - модуляцией,* схему ДИКМ (рис. 4.8) можно упростить заменой квантователя на ограничитель, а предсказывающего устройства *n-го* порядка на интегратор. При сокращении избыточности изображений методом дельта-модуляции наблюдаются те же недостатки, что и при дельта-модуляции других сигналов, например речевых [27], а именно затягивание фронтов и искажения дробления. Однако если частота дискретизации изображения выбрана намного больше частоты Найквиста, то сжатие методом дельта - модуляции приводит к малым (субъективно замечаемым) ошибкам. Если частота дискретизации приближается к частоте Найквиста, то на изображении в большей степени будут проявляться затягивания фронтов (на контурах изображений) и искажения дробления (на участках с постоянной яркостью). Как и при сжатии речи [27], адаптивная дельта-модуляция позволяет уменьшить эти ошибки. Однако в целом при передаче изображений дельта - модуляция оказалась менее эффективной, что при передаче речи.

Квантование с числом уровней, большим двух, позволяет при сокращении избыточности получить изображения более высокого качества. Система сжатия методом ДИКМ с 8-уровневым (З-разрядным) квантованием при оптимальном размещении порогов дает изображения, качество которых такое же, как в системе с ИКМ, имеющей разрядность от 6 до 8. Исключение составляют ошибки вблизи линий резкого изменения яркости.

Сигнал с выхода устройства квантования, конечно, следует кодировать, поскольку распределение вероятностей «квантованных разностей не является равномерным. При удачном выборе кода (например, кода Шеннона — Фано или Хаффмена) удается дополнительно понизить общую скорость создания информации. Прэтт [28] указывает, что при использовании кода Хаффмена в пределе удается понизить скорость создания информации до 2,5 бит/точка. Это дополнительное понижение скорости требуется сопоставить с увеличением стоимости и сложности запоминающего устройства, синхронизаторов и вспомогательных регистров памяти, необходимых для работы с кодами Хаффмена.

Выше обсуждались вопросы сжатия изображений с помощью ДИКМ при выборе элементов по строке (т.е. для прогноза брались точки, лежащие на текущей строке развертки). В силу двумерного характера изображений возможно (и целесообразно) расширить метод ДИКМ так, чтобы при прогнозе учитывались яркости в точках, лежащих не только на текущей, но и на предшествующих строках развертки. Схемы сжатия методом ДИКМ с таким двумерным предсказанием основаны на тех же принципах, что при одномерном предсказании. Поскольку для изображений характерно наличие двумерных статистических взаимосвязей, можно надеяться, что двумерное предсказание даст лучшие результаты по сжатию изображений, так как декорреляция изображений с помощью операций предсказания и вычитания будет производиться по двум координатам. Действительно, устройства с пространственным предсказанием дают более качественные изображения. Хабиби [22] показал, что с помощью двумерного предсказывающего устройства третьего порядка при 8 - уровневом (3 - разрядном) квантовании получались изображения, которые визуально не удавалось отличить от исходных фотографий, обработанных методом ИКМ с 11- разрядными числами.

Для изображений, состоящих из последовательных кадров, например телевизионных, идеи предсказания и вычитания, связанные с ДИКМ, можно распространить на временную область. В подобных изображениях яркость многих точек от кадра к кадру не изменяется или изменяется медленно. Следовательно, можно построить систему сжатия методом ДИКМ, в которой яркость очередной точки прогнозируется на основе яркостей двумерного набора точек текущего кадра и соответствующих точек предшествующих кадров. На практике порядок временного предсказания не может быть высоким, так как для каждого временного слагаемого необходимо иметь запоминающее устройство, где сохранялся бы весь кадр. Моделирование с предсказывающим устройством третьего порядка, в котором для предсказания использовались точки, расположенные в данном (и предшествующем кадрах слева от рассматриваемой точки и вверх от нее, показало, что можно получить очень хорошие изображения при средней разрядности 1 бит/точка [28].

**4.3.3. Схемы сокращения избыточности изображений с обработкой в области преобразований**

Для пояснения основных операций, выполняемых системой сжатия видеоинформации с обработкой в области преобразований, обратимся к ковариационной матрице, определяемой соотношением (4.20). Матрица [*Cg*] описывает корреляцию отсчетов изображения в плоскости (*х, у),* являющейся координатной плоскостью изображения. Важным методом многомерного статистического анализа служит исследование массива данных не только в их естественных координатах, но и в системах координат с более удобными свойствами. В частности, весьма полезными оказались системы координат, основанные на собственных значениях и собственных векторах ковариационной матрицы

[ Cg ] = [ Ф ] [ ] [ Ф ]T = , (4.24)



где [*Ф*] - матрица, составленная из ортогональных собственных вектор - столбцов *Фi* а [] - диагональная матрица собственных значений.



Преобразование координат, определяемое матрицей собственных векторов [*Ф*], обладает тем свойством, что оно производит преобразование заданного массива чисел в другой с некоррелированными элементами, причем получающиеся компоненты имеют убывающие дисперсии. Пусть собственные значения матрицы [Cg] расставлены в убывающем порядке и пронумерованы так, что

, (4.25)



и пусть собственные векторы, связанные с ними, расставлены в том же порядке. Тогда матрица собственных векторов [*Ф*] обладает тем свойством, что умножение ее на вектор-изображение ***g*** (образованный лексикографической расстановкой) дает вектор

*G = [ Ф ]g ,* (4.26)

имеющий некоррелированные компоненты, причем компоненты вектора *G* оказываются расставленными в порядке убывания их дисперсий [29], что является свойством дискретного варианта разложения Карунена - Лоэва, фактически описанного соотношениями (4.24) - (4.26).

Полезность преобразования Карунена — Лоэва ( КЛ, или ковариационного) для сокращения избыточности изображений очевидна. Массив отсчетов изображения заменяется набором переменных, имеющих различные статистические веса ). Сжатие можно получить, отбрасывая переменные с малым статистическим весом и сохраняя остальные. Если, например, оставить *M<N2* компонент вектора *G* и передать их вместе со специальной информацией о том, какие компоненты сохранены, то можно сузить ширину полосы в *N2/M* раз. В приемнике из принятых *М* чисел образуют *N2* - компонентный вектор путем подстановки нулей вместо *N2-М* непереданных компонент. Из этого нового вектора, обозначенного как *G' ,* с помощью преобразования

*gc = [ Ф ]TG’* (4.27)

восстанавливается исходное изображение. В процессе сжатия возникает средняя квадратическая ошибка

*|| g - gc || =*  (4.28)



особенность КЛ - преобразования состоит в том, что из всех линейных преобразований именно оно обеспечивает минимальную величину этой ошибки.

Из соотношений (4.25) и (4.26) видно, что число операций, необходимых для выполнения КЛ - преобразования, пропорционально *N4,* так как исходный массив содержит *N2* отсчетов. Для типичных значений *N ( N = 256* или 512 ) такое число чрезмерно велико. Еще труднее вычислить собственные значения и собственные векторы ковариационной матрицы [*Cg*] размером *N2N2* Эксперименты показал, что очень многие элементы этой матрицы близки к нулю, т.е. коэффициент корреляции между отсчетами быстро стремится к нулю с увеличением расстояния между соответствующими точками изображений. Расстояние, при котором коэффициент корреляции между яркостями элементов изображения становится настолько малым, что его можно приравнять нулю (например, 5 или 10 % максимального значения), называется *радиусом корреляции отсчетов;* его можно выразить через целое число отсчетов. Зная это расстояние, все изображение можно разбить на блоки, размер которых больше радиуса корреляции, но сравним с ним. Если размер каждого блока равен *Р,* то можно вычислить. ковариационную матрицу всех блоков, имеющую размер *P2P2:*



[Cgp] *=* , (4.29)



где *Q=N/P,* a *gi -* вектор, построенный из отсчетов *i-го* блока. Тогда, если [*Фp*] - матрица собственных векторов, связанных с *P2* собственными значениями, расположенными так же, как в формуле (4.25), то операции по сокращению избыточности для каждого из блоков выполняются по формулам (4.26) и (4.27),. как для полного изображения, но матрица [*Ф*] заменяется на [*Фp*]. Как правило, радиус корреляции большинства изображений имеет такую величину, что Р=16 является разумным компромиссом между размером ковариационной матрицы и скоростью, с которой коэффициент корреляции отсчетов приближается к нулю [30]. Длительность вычислений, выполняемых при сжатии видеоинформации поблочно, пропорциональна *Q2/P4.*

Хотя разложение изображения на блоки и делает сжатие видеоинформации методом КЛ - преобразования реально осуществимым процессом, но эффективность его остается недостаточной. Большой объем вычислений препятствует использованию подобных методов для обработки изображений типа телевизионных.

Создание алгоритмов быстрых преобразований (Фурье, Адамара и т.д.) существенно повлияло на многие области применения цифровой обработки сигналов. Аналогичным образом оно - сказалось и на методах сокращения избыточности изображений. Любое линейное преобразование, подобное разложению Карунена - Лоэва, переводит изображение в новую систему координат. В силу свойств КЛ - преобразования случайные компоненты изображения в новых координатах оказываются некоррелированными. Резонно спросить: будут ли другие преобразования, особенно быстрые типа БПФ, обладать такими же полезными свойствами? К счастью, ответ оказывается положительным. Хотя быстрые преобразования и не приводят к полной некоррелированности компонент, как в случае КЛ - преобразования, но все же они дают очень хорошие результаты. Их достоинства, связанные с быстротой вычислений, полностью компенсируют некоторое понижение эффективности сжатия, характерное для них.

Схемы сжатия на основе быстрых преобразований можно описать примерно так же, как и схемы с КЛ - преобразованием. Дополнительным достоинством быстрых алгоритмов является их разделимость, так что двумерные преобразования можно выполнить с 'помощью одномерных операций. Кроме того, их проще описать математически. Если матрица [W] соответствует оператору ортогонального унитарного одномерного преобразования (как, например, матрицы ядер преобразований Фурье, Адамара и т.д. [31] ), то «поворот» изображения в новую систему координат выполняется по формуле

*[ G ] = [ W ]T [ g ] [ W ] ,* (4.30)

гдe ,[g] - исходная матрица отсчетов изображения размером NN, a [G] - преобразование матрицы [g]. Нетрудно заметить, что формула (4.30) описывает двухэтапное преобразование: сначала по строкам изображения, а затем по столбцам преобразований от строк. Записывая преобразование (4.30) в явном виде через элементы матриц, получим:



G(m,n) =



(4.31)

=



где второе равенство является следствием разделимости ядра преобразования. Свойством разделимости обладает ядро преобразования Фурье, наиболее часто применяемого на практике:

(m,n,j,k) = exp [- ] =



(4.32)

= exp [ - ] exp [ - ] ,



а также ядра менее известных преобразований, таких, как преобразования Адамара и Хаара. Более подробно этот вопрос рассмотрен в работе Эндрюса [31].

Собственные значения *i ,* получаемые методом КЛ - преобразования, соответствуют фактическим величинам дисперсий проекций вектора-изображения на координатные оси пространства, в котором вñå компоненты изображения некоррелированы. В системах координат, получаемых при быстром преобразовании, коэффициенты преобразования (т.е. элементы матрицы [*G*] ) равны проекциям вектора - изображения на оси координат, полученным с помощью матрицы преобразования [W], но не являются дисперсиями. Однако как при КЛ - преобразовании, так и в пространствах быстрых преобразований происходит концентрация энергии. В первом случае наибольшие



дисперсии (и, следовательно, наибольшие энергии) связаны с теми столбцами матрицы [*Ф*] или [*Фр*], которые соответствуют предпочтительным (или «естественным») направлениям наибольшего изменения видеоинформации. Аналогично в пространстве быстрого преобразования наибольшими являются коэффициенты, которые соответствуют предпочтительным (или «естественным») направлениям вектора-изображения. С этой точки зрения сжатие в пространстве преобразований (как для преобразования Карунена - Лоэва, так и для быстрых преобразований) является по существу разложением изображения в ряд по базисным векторам (или базисным изображениям, так как каждый вектор должен описывать двумерную структуру) и таким усечением разложения, при котором ошибка мала, а число отбрасываемых составляющих - большое. Усечение оказывается *воз*можным потому, что небольшое число компонент содержит основную часть энергии изображения.

Для иллюстрации рассмотрим схему сжатия в пространстве-преобразовании, основанную на преобразовании Фурье. Из соотношений (4.31) и (4.32) видно, что (*т, п)-й*  коэффициент преобразования *G(m,n)* является проекцией исходного изображения *g(j,k)* на базисный вектор (или базисное изображение), образованный при помощи (*т, п)-го*  значения ядра Фурье

( m , n) = exp ( ) . (4.33)



Для типичных изображений характерно, что в области пространственных частот элементы с малыми индексами велики по сравнению с элементами с большими индексами. Таким образом, структура изображения обычно имеет низкочастотный характер. Низкочастотные составляющие определяют контуры предметов, а также яркость и контрастность изображения. Высокочастотные - составляющие создают резкие линии и определяют общую четкость изображения, но суммарная энергия их невелика. Так, 95% энергии типичного изображения может приходиться на низкочастотные составляющие, занимающие 5% от общей площади двумерной пространственно - частотной области преобразования Фурье. Сохраняя эти спектральные составляющие и достаточно много высокочастотных компонент, чтобы резкость изображения была приемлема для человеческого глаза, можно добиться существенного уменьшения объема избыточной информации.

После того как установлено, что основной принцип сжатия в пространстве преобразований заключается в избирательном сохранении коэффициентов разложения, задача создания системы сжатия изображений может показаться нетрудной. Сложность построения подобных схем кодирования обусловлена необходимостью сравнения свойств операторов различных преобразований и создания методов выбора коэффициентов преобразования, которые следует оставить. Кроме того, задача усложняется квантованием выбранных коэффициентов и кодированием квантованных чисел. Ниже приведены краткие результаты исследований, посвященных этим вопросам.

Был исследован ряд алгоритмов быстрого преобразования, таких, как преобразования Фурье, Адамара, Xaapa [32], слэнт - преобразование [33], косинусное преобразование [34], преобразование по дискретно - линейному базису [35]. Все алгоритмы сравнивались по эффективности сжатия с преобразованием Карунена - Лоэва (оптимальным). Для выявления оптимального алгоритма необходимо сравнивать все преобразования в одинаковых условиях - при одном и том же входном изображении и одинаковых параметрах схем выбора, квантования и кодирования коэффициентов. Этого не было сделано, но приводимые в литературе данные позволяют сделать следующие выводы.

1. Ни один из алгоритмов быстрого преобразования не обеспечивает оптимальной эффективности сжатия изображения, какая получается (при использовании преобразования Карунена - Лоэва.

2. По таким критериям качества, как средняя квадратическая ошибка, ближайшим к преобразованию Карунена - Лоэва, оказывается слэнт - преобразование, а за ним следуют по порядку преобразования Фурье, Адамара и Хаара, причем сравнение выполнялось для изображений небольшого формата, например 1616 или 3232 отсчета.



3. Разница между наилучшими показателями слэнт - преобразования и наихудшими показателями преобразования Xaapa (как по субъективным, так и по объективным критериям) невелика.

Коэффициенты преобразования, которые необходимо сохранить и передать, можно выбрать двумя способами. При *пороговой* дискретизации устанавливается некоторый уровень (определяемый, как правило, на основе полной средней квадратической ошибки), и коэффициенты, его превышающие, сохраняются для передачи, я все остальные отбрасываются. При *зонной* дискретизации в пространстве преобразований размещается маска (трафарет) и элементы, попавшие в нее, сохраняются, а остальные отбрасываются. Операции, выполняемые в ходе преобразования, обычно упорядочиваются в соответствии с некоторым обобщенным индексом (частотой или порядком базисной функции), и коэффициенты преобразования выстраиваются в ряд в порядке увеличения сложности (т.е. числа колебаний на единицу длины) базисных векторов, причем энергия изображения концентрируется в области низких частот или малых порядков. Следовательно, зонная дискретизация эквивалентна обобщенной низкочастотной фильтрации изображения. Пороговая дискретизация, напротив, позволяет выделить значительные коэффициенты преобразования, расположенные где-либо в пространстве преобразований. В результате оказалось, что пороговая дискретизация при одинаковом числе отброшенных коэффициентов дает более высокое качество восстановленного изображения, чем зонная дискретизация. К сожалению, в схемах с пороговой дискретизацией вместе с каждым отсчетом необходимо передавать и его местоположение в пространстве преобразований. По этой причине объем передаваемой информации может заметно возрасти, если положения отсчетов передаются простыми кодами. Однако коды с переменной длиной дают возможность передать адрес при небольшом увеличении числа разрядов кода [32].

Отсчеты, выбранные из пространства преобразований, необходимо квантовать. К сожалению, обычно они имеют гораздо больший динамический диапазон, чем исходные отсчеты в пространстве преобразований, что подтверждает, например, опыт работы с преобразованием Фурье. Такое явление наводит на мысль об использовании чисел с переменной разрядностью, зависящей от значения коэффициента, но это значительно усложняет процесс обработки. Кроме того, для создания устройства квантования, дающего минимальный шум квантования, необходимо знать плотность вероятности значений отсчетов. Исследования плотности вероятности отсчетов в пространстве преобразований [32, 33] показали, что наилучший компромисс между простотой и точностью обеспечивает квантование, основанное на гауссовской плотности, при фиксированной разрядности отсчетов. В этом случае удается получить высококачественные восстановленные изображения, если число уровней квантования составляет всего 64 (6 разрядов) [32]. По-видимому, это связано с тем, что операторы преобразований линейны и дают взвешенные суммы, а сумма произвольных случайных величин распределена по закону, близкому к гауссовскому.

Способ кодирования коэффициентов разложении зависит от примененного алгоритма выборка коэффициентов. Как уже отмечалось, при пороговой дискретизации необходимо, чтобы код содержал адрес отсчета в пространстве преобразования, причем на каждый отсчет отводится фиксированное число разрядов кодовой комбинации. При зонной дискретизации используется низкочастотный характер изображения, т.е. тот факт, что коэффициенты разложения, соответствующие низким частотам (или малым порядкам базисных векторов), имеют большую величину, чем высокочастотные коэффициенты. Этим можно воспользоваться, уменьшая число разрядов кода, отводимых на отсчет, по мере перехода от низких частот к высоким [33]. Информацию о положении отсчетов передавать не нужно, поскольку форма зоны известна, а порядок выбора и передачи отсчетов .внутри зоны может быть зафиксирован.

В целом сокращение избыточности путем обработки в пространстве преобразований (т.е. преобразование, выбор коэффициентов, их квантование и кодирование) позволяет получить хорошие результаты. На рис. 4.10 сравниваются несколько разных преобразований. Изображения на рис. 4.9 и 4.10 состоят из 256256 точек, представленных 8-разрядными числами. Нетрудно видеть, что сжатие путем обработки в пространстве преобразований дает лучшие результаты, чем сжатие методом ДИКМ.



**4.3.4. Другие аспекты задачи сокращения избыточности**

**видеоинформации**

Поскольку часто изображения состоят из многих последовательно появляющихся кадров (как, например, в телевидении) и изображение от кадра к кадру меняется мало, то в будущем, по видимому, больше внимания будет уделяться межкадровому сжатию (в отличие от внутрикадрового). Как отмечалось ранее, схемы межкадрового сжатия методом ДИКМ уже изучались. Оказалось, что комбинация внутрикадрового и межкадрового кодирования (в тех случаях, где это возможно сделать) может привести к уменьшению объема передаваемой информации в 30—50 раз.

Интерес к цветным изображениям возрастает; уже проводились опыты по сокращению их избыточности методом ДИКМ и методом преобразования [33, 36]. Методы сжатия аналогичны рассмотренным выше, но их реализация усложняется из-за наличия трех цветовых сигналов.

В последнее время созданы гибридные системы сжатия. В них для сжатия по одной координате (обычно по строкам дискретизованного изображения, или по горизонтали) используется схема с преобразованием, а по другой координате (по столбцам, или по вертикали) - схема с ДИКМ. В результате получается более простая система ( b ней не требуются двумерные преобразования), но дающая такую же или лучшую эффективность кодирования, чем системы с ДИКМ или с преобразованием [20].

Отметим, наконец, что все рассмотренные выше схемы не являются адаптивными, т.е. в них не изменяются в зависимости от свойств изображения распределение кодовых разрядов, расположение уровней квантования и т.д. Тешер [37] показал, что за счет адаптивности можно дополнительно повысить эффективность кодирования изображений. При одинаковом качестве восстановленного непрерывного изображения ему удалось сократить объем передаваемой информации почти вдвое по сравнению с другими системами.

**4.4. Повышение резкости изображений**

Задача любой системы, формирующей изображение, состоит в создании резкого, чистого изображения, свободного от искажений. Это не всегда возможно сделать. Во-первых, каждая реальная система формирования изображений обладает определенными ограниченными возможностями; импульсный отклик реальной системы имеет конечную ширину, что приводит к неизбежному снижению разрешающей способности. Если на изображении необходимо выделять важные детали, размер которых близок к ширине импульсного отклика, то необходимо бороться с потерями разрешения. Так, например, с межпланетных космических аппаратов приходят снимки замечательного качества (особенно, если учесть, откуда они получены), но ученые, изучающие планеты, всегда пытаются увидеть на них элементы поверхности планеты, искаженные в силу ограниченного разрешения фотокамер. Во-вторых, изображения могут быть испорчены из-за неудачного стечения обстоятельств. Можно принять все меры предосторожности, чтобы получить высококачественные снимки, но какая-то часть их окажется испорченной либо за счет движения объекта или камеры, либо из-за плохой фокусировки и т.д. Среди плохих снимков всегда находятся столь важные или настолько редкие, что стоит пытаться их исправить. Устранение искажений относится к задачам *повышения резкости* (или *восстановления ) изображений.*

Ниже читатель увидит, что для решения задачи восстановления (или повышения резкости) изображения предложено несколько различных .методов. Для решения же рассмотренной в предыдущем разделе задачи сокращения избыточности изображений было представлено только два существенно различных метода. Как будет показано, восстановление изображений является задачей, не имеющей единственного решения, что и привело к многочисленным попыткам найти лучший способ решения.

**4.4.1. Важные аспекты задачи повышения резкости**

**изображений**

Как показано выше, основное уравнение формирования изображения имеет вид

*g(x, у)=* , (4.34)



где *g -* сформированное изображение, *h -* импульсный отклик или аппаратная функция, а *f* - функция распределения яркости объ-

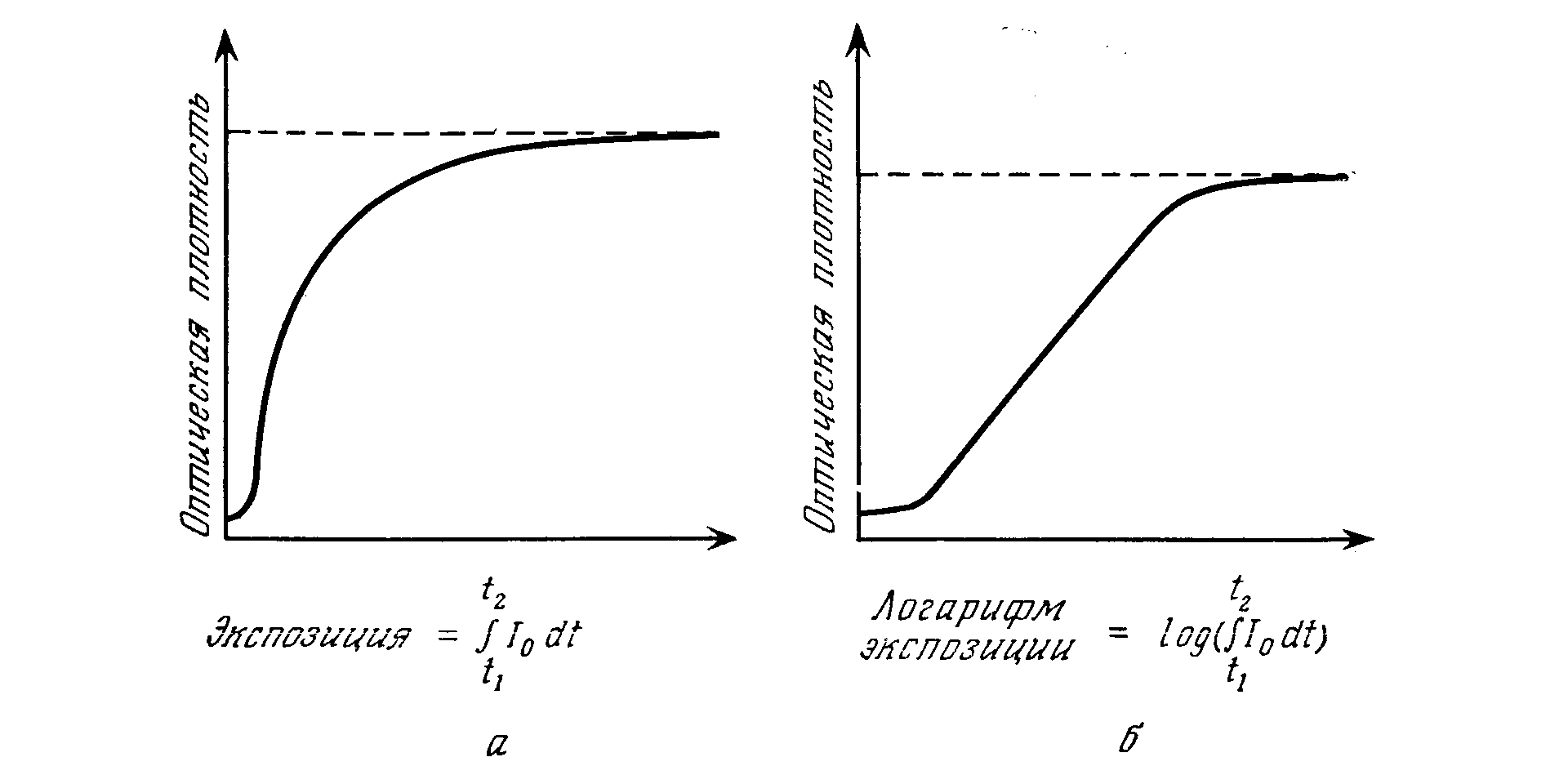


Рис. 4.11.

*а -*  зависимость оптической плотности от экспозиции; б - зависимость оптической плотности от логарифма экспозиции.

екта. Конечно, непосредственно изображение *g* наблюдатель не получает; изображение существует в виде модуляции интенсивности какого-то излучения, исходящего от объекта. Изображение становится известным только после того, как оно будет воспринято и записано некоторым образом (например, на фотопленке, сетчатке глаза, люминесцентном экране). В процессе восприятия и записи в изображение вносятся шумы, поскольку при любом способе записи сигналов присутствуют шумы записывающей системы. Весь процесс получения изображения, т.е. формирование изображения, его восприятие и запись, был показан в виде блок-схемы на рис. 4.3. Таким образом, восстанавливать изображение приходится по записи, содержащей шум.

Задача повышения резкости изображения усложняется особенностями записывающей системы и ее собственным шумом. Как отмечалось в разд. 4.2, наиболее распространенная система для записи изображений, фотопленка, имеет нелинейную характеристику и вносит шум, модулированный сигналом. Типичная характеристика фотопленки приведена на рис. 4.11,а, где показана зависимость оптической плотности, определенной выше, от интенсивности падающего излучения, причем предполагается, что за .время выдержки интенсивность не изменяется. Обычно такую информацию приводят па графиках с логарифмическим масштабом по оси абсцисс (рис. 4.11,б); получающийся в таких координатах график зависимости *D* от log *Е* имеет линейный участок. Вид этой кривой иногда порождает заблуждения относительно резкой нелинейности характеристики фотопленки. Поскольку изображение записывается на пленке зернами серебра, а плотность серебряного слоя пропорциональна оптической плотности, то очевидно, что процессы записи на фотопленку нелинейны по своей сути.

Столь же сложны процессы, создающие шумы пленки. Дисперсия шума, вызванного случайным размещением зерен серебра на пленке, пропорциональна локальной .плотности проявленных серебряных зерен, причем функция пропорциональности усложняется наличием степенного множителя [8]. Таким образом, шум представляет собой флуктуации оптической плотности записи изображения, зависящие от сигнала. В общем случае задача цифровой обработки сигналов с шумом, зависящим от сигнала, еще не решена. Немногочисленные известные результаты [10] указывают, что зависимость шума от сигнала может оказаться не слишком большой проблемой. Поэтому обычно предполагают, что шум создается независимыми от сигнала флуктуациями плотности. Однако логарифмическая зависимость между оптической плотностью пленки, запечатлевшей изображение, и интенсивностью падающего на пленку излучения усложняет задачу. Даже если предположить, что флуктуации оптической плотности не зависят от сигнала, в свете, прошедшем через пленку, будет присутствовать мультипликативный шум, зависящий от сигнала. Аналогичные эффекты наблюдаются в фотоэлектронных системах (типа телевизионных и им подобных), характеристики которых описываются степенным законом.

С учетом вышеизложенного полная модель процессов формирования и записи изображения описывается соотношением

g(x,y) = s [] +



+ n ( x, y ) (4.35)

где *g -* фактически записанное изображение, *s* - характеристика записывающего процесса, а *п —* шум. Предполагается, что *п* не зависит от записанного изображения. Задача повышения резкости (или восстановления) изображения заключается в определении исходного распределения интенсивности излучения объекта *f (x, у)* на основе записи *g(x, у),* содержащей шум.

Формула (4.35) позволяет оценить сложность задачи восстановления изображения. Искажения исходного изображения изменяют интенсивность экспонирующего света, а в нашем распоряжении имеется только нелинейная функция от этих интенсивностей. Если избавиться от нелинейности с помощью преобразования, обратного к *s,* то получим

s-1 { g ( x, y ) = s-1 {s [ h ( x, y ) \*\* f( x, y )] + n ( x, y )} , (4.36)

где знак \*\* означает двумерную свертку. Поскольку обратная нелинейность воздействует на сумму и оператор этой нелинейности недистрибутивен по отношению к оператору сложения, то наличие шумового слагаемого означает, что 1) получить точное обратное преобразование и исходные интенсивности невозможно, 2) при преобразовании *g* с использованием характеристики, обратной к нелинейной характеристике системы записи, будет получена нелинейная комбинация сигнала и шума.

Решения уравнения (4.36) в настоящее время не получены. На практике при восстановлении изображений делают одно из двух предположений. Первое предположение состоит в том, что в соотношении (4.36) операцию *s-1* можно применить к слагаемым ïî отдельности, т.е.

s -1{ g( x, y) } h (x, y) \*\* f ( x, y )+ s -1{ n ( x, y ) } *,* (4.37)



что эквивалентно замене шума на новый шумовой процесс *s -1{п}.* При втором подходе данные обрабатываются в том виде, как они записаны, т.е. предполагается, что нелинейной функцией s можно пренебречь и

g ( x, у ) h ( x, у ) у \*\*  f ( x, у ) + п ( x, у ). (4.38)



Равенства (4.37) и (4.38) были бы справедливы, если бы удовлетворялось любое из двух предположений: 1) функция s имеет *слабую нелинейность* и ее можно аппроксимировать линейной функцией или 2) сигнал *g* изменяется b малых пределах (низкоконтрастное изображение) и операторы *s* и *s-1* близки к линейным. Графики типа изображенных на рис. 4.11 показывают, что нелинейность оператора *s* обычно нельзя считать слабой. При правильном выборе экспозиции диапазон изменения от черного к белому не соответствует предположению о низкой контрастности. Однако на практике изображение удается успешно восстановить в *пространстве яркостей,* связанном с преобразованием (4.37), или в *пространстве плотностей,* связанном с преобразованием (4.38), несмотря на теоретические трудности, возникающие при отбрасывании нелинейностей.

Даже при беглом взгляде на соотношения (4.37) или (4.38) видно, что восстановление изображений относится к задачам *фильтрации* с помощью инверсной свертки (т.е. к обработке записанной информации с помощью оператора *h* , который является обратным по отношению к оператору двумерной свертки с функцией *h).* Эта задача сложна тем, что сводится к решению плохо обусловленных систем линейных уравнений [38]. При повышении резкости изображений задача усложняется за счет того, что для многих изображений аппаратная функция сингулярна, т.е. преобразование Фурье от нее имеет нули, что наблюдается, например, при дрожании съемочной камеры или плохой ее фокусировке. Поэтому при решении задач восстановления изображений значительные усилия затрачиваются на преодоление трудностей, связанных с сингулярностью [2].

Для восстановления изображений цифровыми методами необходимо, чтобы все уравнения были записаны для дискретизованных функций. Поэтому соотношение (4.35) принимает вид

g ( jx, ky ) s {



} + n ( jx, ky ) , (4.39)



где знак (приближения указывает, что дискретные суммы не являются точным представлением исходных интегралов. Аналогичные выражения можно записать для формул (4.37) и (4.38). Интересно отметить, что соотношение (4.39) можно рассматривать как систему уравнений относительно неизвестных значений *f.* Если выполняются предположения, сделанные при выводе соотношения (4.38), то соответствующие дискретные уравнения (где без потери общности можно положить x = y = 1) прекращаются в систему линейных уравнений относительно *f ( p, q ):*



g ( j, k ) , (4.40)



Формула (4.40) подсказывает, что задача восстановления изображений сводится к решению системы линейных уравнений. Это действительно так, и для подтверждения можно представить соотношение (4.40) в виде произведения матрицы на вектор. Поэтому значение цифровых методов обработки сигналов, таких, как линейная фильтрация и БПФ, состоит .в том, что они являются средством для быстрого нахождения точного или приближенного решения очень больших (с *N2* переменными) систем линейных уравнений. Такой .подход очень важен для развития более совершенных методов повышения резкости изображений, но обсуждение его требует применения теории матриц в объеме, чрезмерно большом для данной книги. Подробнее связь между матричными представлениями и восстановлением изображений методом дискретного преобразования Фурье описана в работах [2, 39, 40].

**4.4.2. Основные методы повышения резкости изображений**

Операция дискретной свертки, фигурирующая в формуле (4.40), имеет аналог в пространстве дискретного преобразования Фурье (ДПФ). ДПФ соотношения (4.40) имеет вид

G ( u, v ) = H ( u, v ) F ( u, v ) + N ( u, v ) , (4.41)

где *и, v = 0,* 1, ..., *N—*1, а прописные буквы обозначают ДПФ величин, обозначенных в формуле (4.40) соответствующими строчными буквами. Так, функция

H ( u, v ) = , (4.42)



представляет собой , ДПФ от дискретизованной аппаратной функции; аналогичные соотношения имеются для *G, F* и *N.* Обычно область ненулевых значений аппаратной функции по величине гораздо меньше, чем размеры исправляемого снимка. Следовательно, *N* гораздо больше числа ненулевых отсчетов функции *h ( j, k )* и перед преобразованием эта функция должна быть дополнена соответствующим числом нулей (необходим также надлежащий сдвиг по фазе). Как следует из теории, величину *N* следует выбирать достаточно большой, чтобы устранить нежелательные эффекты заворота изображения, связанные с периодичностью круговой свертки. На практике оказывается, что эти эффекты не очень существенны. Действительно, снимаемая сцена бесконечно велика по сравнению с размерами аппаратной функции и на краях кадра искаженное изображение содержит вклад объектов, расположенных вне кадра, но вносящих в кадр помехи за счет свертки с аппаратной функцией. Краевые эффекты, возникающие в процессе инверсной фильтрации изображения на основе неполной информации (о предметах, находящихся за кадром), более важны, чем завороты. Бэкстер [18] показал, как можно частично исправить эти краевые искажения. При выборе *N* более важно предусмотреть возможность уменьшения заворотов и эффектов усечения, связанных не с самой аппаратной функцией *h( j, k ),* а с функцией, обратной к ней.

Простейшим способом повышения резкости изображения является обработка записи в пространственно-частотной области с помощью обратного фильтра. При этом получается оценка .восстановленного изображения

( u, v ) = Hi ( u, v ) G ( u, v ) = , (4.43)



Хотя это самый простой подход, при его использовании встречаются наибольшие трудности.

1. Для многих видов искажений аппаратная функция такова,. что ее ДПФ имеет нули. Например, для искажений, вызванных движением съемочной камеры в горизонтальном направлении, ДПФ аппаратной функции имеет вид

H ( u, v) = , (4.44)



где *а -* размер размытия, выражаемый числом отсчетов. Если искажения сильные (*а* достаточно велико), так что нули функции (4.44) попадают *в* найквистовский диапазон, то обратный фильтр является сингулярным. Аналогичная трудность возникает в случае 'искажений, вызванных расфокусировкой камеры, когда запись содержит свертку истинного изображения с формой апертуры. Для большинства стандартных форм апертуры (круглой, квадратной и т.д.) соответствующие ДПФ имеют нули; если эти нули попадают в найквистовский диапазон, то обратный фильтр оказывается сингулярным. К сожалению, обычно так и бывает.

2. Столь же неприятно то, что аппаратные функции (даже если они не приводят сингулярности) обычно являются плохо обусловленными, в частности, модуль их преобразования быстро уменьшается вблизи некоторых значений *и* и *v* (обычно в области высоких частот, поскольку искажения связаны с медленными процессами), поэтому обратный фильтр 1/*H* резко увеличивает влияние шумовой составляющей, входящей в формулу (4.41), что ухудшает изображение.

Несмотря на все трудности, обратные фильтры удается приманить для восстановления изображений. На рис. 4.12 приведено цифровое изображение размером 512512 точек, в которое с помощью ЭВМ были внесены искажения, эквивалентные свертке изображения с гауссовской аппаратной функцией. К искаженному снимку был прибавлен шум и отношение сигнал/шум (С/Ш). измеряемое как отношение дисперсий сигнала и шума, стало равным 12000 (33 дБ). Аппаратная функция не содержит нулей, а отношение С/Ш велико, поэтому восстановление методом обратной фильтрации возможно. На рис. 4.13, а приведен результат обработки снимка рис. 4.12 обратным фильтром, реализованным с помощью БПФ. Однако если С/Ш уменьшается за счет роста мощности шума, то восстановление путем обратной фильтрации дает плохие результаты. На рис. 4.13,6 приведено изображение, восстановленное обратной фильтрацией при С/Ш = 200 (23 дБ). Размытое изображение при таком уровне шума выглядит так же, как снимок рис. 4.12, поскольку влияние шума трудно определить визуально, если отношение С/Ш превышает 20 дБ. Шум был прибавлен в области плотностей, связанных с яркостями по логарифмическому закону; изображение искажалось в яростной области, но восстанавливалось в области .плотностей так, как описано в разд. 4.4.4.



Рис. 4.12 и 4.13 иллюстрируют важную мысль: обратный фильтр может работать, но для этого требуется очень большое отношение сигнал/шум и малая степень искажений. К сожалению, нет определенных правил, которыми следует руководствоваться при восстановлении изображений. Но, с другой стороны, обратная фильтрация осуществляется настолько просто, что ее можно проводить, не имея заранее гарантий на успех, и это не приведет к большим затратам, если окажется, что шум или сингулярность воспрепятствуют ее выполнению.

Влиянию помех и сингулярностей менее подвержен другой метод восстановления изображений — винеровская фильтрация. Как следует из названия, такая фильтрация основана на теории оптимальных оценок, предложенной Норбертом Винером. При проектировании фильтра ставится задача найти такую линейную оценку

( j, k ) = L [g ( j, k ) ] ,



(где *L—*линейный оператор), что

E { [ f ( j, k) - ] 2 } ,



имеет минимальную величину. Структуры устройства для получения оценок была найдена многими исследователями. Применительно к обработке изображений первым в явном виде это сделал Хелстром [41] , разработавший пространственный и частотный варианты устройства. При цифровой обработке изображений используется вариант с обработкой b частотной области. Цифровая фильтрация изображений выполняется фильтром с передаточной функцией

Hw( u, v ) = , (4.45)



и

(4.46)



причем звездочка означает комплексное сопряжение, а *Фn*  и *Фf* - энергетические спектры шума и сигнала соответственно. Как и при обратной фильтрации, обработка производится на основе двумерных БПФ и обратное преобразование от (4.46) дает исправленное изображение.

Анализируя соотношение (4.45), можно заметить следующие свойства винеровского фильтра:

1. Если шум очень мал или отсутствует, так что *Фn* 0, то винеровский фильтр переходит в обратный. Таким образом, в пространственно - частотных областях с малым уровнем шума (как правило, это области низких пространственных частот) характеристики винеровского и обратного фильтров совпадают.



2. Если мощность сигнала становится малой, так что Фf  0, то коэффициент передачи винеровского фильтра стремится к нулю. Этим решаются проблемы, связанные с сингулярностью аппаратной функции и с особенностями плохо обусловленных систем уравнений при отсутствии сингулярности.



На рис. 4.14, а и 4.14,6 приведены результаты восстановления изображения, полученные с помощью винеровского фильтра при таких же условиях, как на рис. 4.13, а и *б.* Снимки рис. 4.14, а и рис. 4.13, а похожи, что указывает на эквивалентность винеровского и обратного фильтров при малом уровне шума. Однако по качеству рис. 4.14,6 íaмнîãî превосходит рис. 4.13,6; это свидетельствует о том, что винеровский фильтр лучше подавляет шумы в областях с малым сигналом. Полосы на краях рис. 4.13 и 4.14 вызваны краевыми эффектами свертки, рассмотренными ранее, и заворотами, связанными с аппаратной функцией обратного фильтра..

Внимательное сравнение рис. 4.13,6 и 4.14,6 создает впечатление, что винеровский фильтр восстанавливает изображения с меньшим уровнем шума, чем метод обратной фильтрации, но, возможно, при этом несколько ухудшается резкость и теряются мелкие детали. Все это можно объяснить двумя причинами:

1) критерий минимума средней квадратической ошибки является слишком строгим и его можно ослабить;

2) нелинейные и адаптивные свойства человеческого зрения могут «не согласовываться» с критерием минимальной средней квадратической ошибки.

В разд.4.3.3 будет рассмотрен другой способ построения восстанавливающего фильтра, в котором предъявляются меньшие требования к априорной информации, чем при винеровской фильтрации.

**4.4.3. Повышение резкости изображения «вслепую» и методом уравнивания энергетических спектров**

Кэннон [42] предложил метод, позволяющий восстанавливать изображения лучше, чем с помощью винеровского или обратного фильтров; это так называемый метод уравнивания энергетических спектров или метод *гомоморфной фильтрации.* (Сначала такой фильтр был спроектирован Коулом [21] на основе теории гомоморфных систем, а затем Кэннон нашел эквивалентную форму, пользуясь методом уравнивания энергетических спектров, который и будет описан ниже.) Передаточная функция фильтра определяется из простого условия, гораздо более слабого, чем критерий минимума средней квадратической ошибки в винеровском фильтре. При проектировании фильтра отыскивается такая линейная оценка

( x, y ) = L [ g ( x, у )] ,



*(*где *L -* линейный оператор), чтобы энергетический спектр оценки равнялся энергетическому спектру исходного изображения. Таким образом, условие имеет вид :

( u, v ) = ( u, v ) . (4.47)



Поскольку *g ( х, у )* определяется из соотношений (4.37) или (4.38), то энергетический спектр оценки равен :



( u, v ) = | L( u, v )|2 [ | H ( u, v )|2 Фf( u, v )+Фn( u, v )]*.* (4.48)



Его можно приравнять к правой части соотношения (4.47) и в явной форме найти |*L(u,v)*| *-* модуль пространственно - частотной характеристики линейного фильтры, уравнивающего энергетические спектры (уэс). В результате характеристика фильтра будет иметь вид

Hуэс( u, v ) = | L ( u, v ) | = [ ] , (4.49)



а из соотношения

( u, v ) = Нуэс ( u, v ) G ( u, v ) (4.50)



определяются значения спектра восстановленного изображения. Для получения самого изображения эти значения нужно подвергнуть обратному преобразованию. Как ,и прежде, вычисления следует производить с применением БПФ.

Фильтр с уравниванием энергетических спектров обладает следующими свойствами:

1. При малых шумах, когда Фn 0, характеристика фильтра сводится к модулю характеристики обратного фильтра.



2. При малых сигналах, когда Фf 0, коэффициент передачи фильтра уменьшается до нуля.



3. В промежутке между этими экстремальными условиями коэффициент передачи фильтра больше, чем *Нw ,* но меньше, чем *Hi .* Это обусловлено отсутствием множителя *Н\** в числителе выражения для *Нw* также свойствами операции извлечения квадратного корня. Можно показать, что характеристика фильтра с уравниванием энергетических спектров равна среднему геометрическому от характеристик винеровского и обратного фильтров [21].

Поскольку коэффициент передачи фильтры с уравниванием энергетических спектров больше, чем у винеровского фильтра, но не приводит к увеличению шумов, как в обратном фильтре, то восстановленное изображение содержит большее число мелких деталей, связанных с высокочастотными составляющими, которые винеровский фильтр обычно подавляет. При этом уровень шума увеличивается, но человек обычно согласен мириться с увеличением шума, если получает дополнительную полезную информацию о мелкой структуре изображения.

На рис. 4.15,а и *б* показаны изображения, исправленные фильтром с уравниванием энергетических спектров при условиях, соответствующих рис. 4.13, *а* и *б.* Снимки на рис. 4.13, *a*, 4.14, *a* и 4.15, а похожи; это указывает на то, что при слабых шумах этот фильтр имеет те же свойства, что и обратный. Результат восстановления изображения при сильном шуме представлен на рис. 4.15,6. Сравнение его с рис. 4.13,6, рис. 4.14,6 и с исходным изображением на рис. 4.12,а показывает преимущество данного метода повышения резкости изображений.

Соотношение (4.49) задает только модуль передаточной функции, но можно построить фильтры, в которых задается также определенная фазовая характеристика. Для многих видов искажений (например, для расфокусировки или смазывания изображений при движении) фазовая характеристика имеет значения 0 или *± .* Фильтр, описываемый соотношением (4.49), можно дополнить устройством, вносящим необходимый сдвиг фазы; в результате получается восстанавливающий фильтр, корректирующий в изображении как амплитудные, так и фазовые искажения. Кэннон [42] показал, что подобный метод повышения резкости лучше чисто амплитудной коррекции (как и следует из общих принципов).



Метод уравнивания энергетических спектров проясняет весьма важные особенности задачи восстановления «вслепую», т.е. путем оценивания параметров, необходимых для восстановления искаженного изображения по этому же искаженному изображению. При винеровской фильтрации необходимо заранее знать величины *Фn ,* *Фf* и *Н.* Аналогичное требование можно сформулировать и для случая фильтрации методом уравнивания энергетических спектров. Однако внимательный анализ спектров, фигурирующих в соотношении (4.49), позволяет заметить, что оценки всех необходимых величин можно найти на основе искаженного изображения.

Рассмотрим изображение *g ( x, y ),* формируемое по законам, описываемым формулами (4.37) и (4.38) или им подобными. Разобьем изображение на части размером *ММ,* где *М* велико по сравнению с размерами аппаратной функции, но мало по сравнению с размером исходного изображения *NN.* Типичным значением является *М=64.* Отдельные части изображения могут перекрываться. Если пренебречь краевыми эффектами, то каждую часть изображения можно описать сверткой аппаратной функции с соответствующей областью полного первоначального распределения яркости объекта. Тогда соотношение



*gi (x, у) h(x - x1, у—y1) fi ( x1, y1) dx1 dy1 + ni( x, y )* , (4.51)



аппроксимирует закон формирования изображения в каждой из частей *gi .* После дискретизации *gi* путем взятия *ММ* отсчетов можно найти энергетический спектр каждой части изображения, равный



*Фig( u , v )|H( u , v)|2 Фif (u , v) + Фin ( u , v) ,* (4.52)



где верхний индекс *i* обозначает номер части изображения. Если предположить, что изображение и шум можно аппроксимировать стационарными случайными процессами, то Фif и Фin будут выборками из двух функций, описывающих два энергетических спектра. Поэтому суммирование по *i* приведет к сглаживанию случайных отклонений в энергетических спектрах каждой из частей изображения. Таким образом,



(4.53)



где *Q=N/M,* а и —оценки энергетических спектров сигнала и шума соответственно.



Значение равенства (4.53) определяется следующими «соображениями:

1. Квадрат модуля преобразования Фурье от размытой аппаратной функции приводит к появлению характерных особенностей в усредненною энергетическом спектре частей изображения. Для таких видов искажений, как расфокусировка и смазывание при сдвиге, в спектре остаются характерные признаки, позволяющие установить тип искажения и определить все его существенные параметры, как, например, размер «размазанного» изображения точки и период обращений фазы [42].

2. Знаменатель 'передаточной функции фильтра с уравниванием энергетических спектров полностью совпадает с правой частью равенства (4.53), и если уже нам известна оценка Фf, то можно выполнять операцию восстановления изображения. Оценку Фf можно получить несколькими способами. Bo - первых, можно взять неискаженные изображения, подобные восстанавливаемому, и определить Фf но ним, так как было показано [21], что большинство изображений имеет очень похожие энергетические спектры. Или же можно на основе исправляемого изображения получить оценку *Фn* (обычно это удается сделать на однородных участках изображения), а также определить значения *Н* по его характерным признакам и решить уравнение (4.53) относительно .



На pис. 4.16,а приведен снимок, сделанный камерой, двигавшейся в процессе съемки. Камера двигалась в горизонтальном направлении, и хорошо заметно, что из-за этого мелкие надписи стали совершенно неразличимыми. На рис. 4.16,6 показан тот же снимок после восстановления вышеописанным методом; смазывание за счет сдвига оставило в энергетическом спектре характерные признаки, которые были автоматически распознаны и применены при создании фильтра, восстанавливающего изображение методом уравнивания энергетических спектров. После восстановления текст стал вполне разборчивым. Дальнейшие сведения об этом методе восстановления изображений можно найти в работе [42].

**4.4.4. Замечания о восстановлении изображений в яркостной и плотностной областях**

Все рассмотренные образцы изображений были искажены в пространстве яркостей либо моделированием характеристической кривой пленки на ЭВМ, либо при проведении стационарной съемки в реальных условиях, как снимок рис. 4.16. Однако восстановление изображений проводилось в пространстве плотностей, связанных с яркостями логарифмической зависимостью, как было показано выше. Во-первых, это, очевидно, связано с предположением о линейности, выраженным в виде равенства (4.38). Во-вторых, из практических соображений, относящихся к качеству изображений, восстановление изображений удобнее проводить в пространстве плотностей, а не в пространстве яркостей, где необходимо учитывать соотношение (4.37). Диапазон изменения яркости составляет обычно 2—3 порядка, и в тех местах изображения, где яркость изменяется резко, могут просматриваться боковые лепестки характеристики восстанавливающего фильтра. Динамический диапазон .изменения плотности гораздо меньше одного порядка, и подобный эффект здесь не столь опасен. В обширных работах Кэннона [42] и Коула [21] показано, что изображения, восстановленные в пространстве плотностей, обладают благоприятными для зрения свойствами. Таким образом, предположение о линейности, связанное с равенством (4.38), из практических соображений оказывается более предпочтительным.

**4.4.5. Нелинейные методы восстановления изображений**

С позиций цифровой обработки сигналов все рассмотренные методы сводятся к линейной фильтрации сигналов с применением быстрых преобразований Фурье. Из этого, конечно, не следует, что построение эффективной системы обработки сигналов является тривиальной или несложной задачей. При оптимизации методов фильтрации сигналов и соответствующих машинных программ может потребоваться много труда и изобретательности. Тем не менее, основополагающие принципы фильтрации относятся к области линейной обработки сигналов, и их легко найти в работах по классической цифровой обработке сигналов.

Практические исследования, однако, показывают, что линейная обработка имеет недостатки. Во-первых, реальные изображения обладают рядом свойств, которые не учитываются при линейной обработке. Например, яркости точек изображения всегда положительны, а в схеме с линейной обработкой могут появляться отрицательные числа, связанные с боковыми лепестками характеристики восстанавливающего фильтра. Во - вторых, линейная обработка является лишь приближением к оптимальной обработке, так как средства для записи изображений, такие, как кинопленка, в принципе нелинейны. Поэтому представляют интерес методы повышения резкости изображений, в которых учитывается такая нелинейность.

При нелинейном восстановлении изображений (как почти во всех операциях, связанных с нелинейностями) основная трудность заключается в объеме вычислений. В нелинейных системах эффективность вычислений не такая высокая, как при линейной обработке методом БПФ. В силу этого из всех предлагавшихся методов нелинейного восстановления изображений лишь немногие когда-либо применялись для обработки крупных изображений, так как при большом количестве отсчетов число вычислительных операций чрезмерно возрастает. Решения подобных проблем, по-видимому, b большей степени связаны с математическим анализом, чем с цифровой обработкой сигналов, и поэтому данный раздел будет довольно коротким.

Один из практически реализуемых нелинейных методов относится к восстановлению изображений в пространстве плотностей с учетом предположений, связанных с равенством (4.38). Если яркости искаженного изображения перевести в плотности путем логарифмирования, а затем скорректировать изображение с помощью БПФ и результат пропотенцировать, то получится система с нелинейными характеристиками, но реализованная на основе БПФ. К тому же яркости конечного изображения здесь всегда положительны. Теоретическим основанием подобного метода являются теория гомоморфной обработки сигналов, а также мультипликативная модель процесса формирования изображения [19]. Логарифмическая пространственная фильтрация, по-видимому, согласуется с моделью системы человеческого зрения, представленной в первом разделе главы.

Метод Фридена [43] также гарантирует отсутствие отрицательных значений яркости в восстановленном изображении, которое определяется путем решения системы нелинейных уравнений

*g(j, k) = h( j , k* ) \* \* *ехр* [ - 1+ *h( j*, *k) \* \* ( j*, *k ) +* ] *+*



*+ exp* [ -1 + *( j, k)* ] ,



*P =*  , (4.54)



а исправленное изображение описывается равенством

(4.55)



где символ \*\* обозначает двумерную дискретную свертку, а *Р —* полная энергия исходного изображения. Таким образом, решение оказывается положительным и ограниченным по величине. Однако решить эту систему непросто. Вычисление сверток, фигурирующих в соотношениях (4.54) и (4.55), методом БПФ не очень помогает, поскольку основная трудность заключается в решении системы уравнений относительно неизвестных и (множителей Лагранжа в задачах оптимизации). Данный метод был опробован на очень малых изображениях (размером 50х50 отсчетов) в случае разделимых аппаратных функций, причем для решения нелинейных уравнений применялась итерационная процедура Ньютона - Рафсона. При увеличении размеров снимка решение нелинейных уравнений связано с большими трудностями. Был также предложен, но не реализован практически метод прямой оптимизации [2].



Метод нелинейного восстановления изображений, предложенный Фриденом, исходит из предположений о положительности и ограниченности отсчетов изображения. Нелинейные методы могут также основываться на анализе нелинейности записывающих средств. В работах [40, 44] описан байесовский подход к восстановлению изображений, записанных с помощью нелинейных устройств. Результат обработки получается в виде решения нелинейного матричного уравнения. 0пределение этого решения при большом числе переменных, описывающих квантованное изображение, связано c выполнением множества вычислений, а роль цифровой обработки сигналов сводится к выполнению операций свертки [44]. Такой метод применялся для коррекции изображений размерами до 512х512 отсчетов.

Задача (восстановления изображения в общем случае, т.е. с учетом нелинейности записи и условия, что отсчеты яркости должны быть ограниченными и положительными, сводится к задаче нелинейного программирования [2]. Однако возможности современной техники не позволяют решить в общем случае задачу нелинейного программирования при том числе переменных, которое характерно для изображений, представленных в цифровом виде. Были разработаны и опробованы на «маленьких изображениях (размером, например, 32х32 отсчета) специальные алгоритмы, основанные на симплексном методе и относящихся к нему понятиях математического программирования. К большим снимкам эти алгоритмы еще не применялись. Вообще повышение резкости нелинейными методами является той областью цифровой обработки изображений, где далеко не все сделано.

**4.4.6. Повышение резкости изображений с характеристиками,**

**изменяющимися в пространстве**

В предыдущих разделах была показана роль, которую играют при восстановлении изображений операции свертки и двумерная цифровая фильтрация. Во всех методах требовалось, чтобы законы формирования изображения обладали свойством пространственной инвариантности. Если же процесс формирования изображения не является пространственно - инвариантным, то формула (4.34) принимает вид

(4.56)



и при обращении ее в дискретную форму в ней не появится дискретной свертки. К сожалению, в ряде интересных случаев формирование изображений подчиняется пространственно - нестационарньпм уравнениям, как, например, при неравномерном движении камеры или при наличии оптических аберраций.

Рассмотрим, например, одномерные искажения по горизонтальной оси, когда искажения минимальны в левой части снимка и линейно увеличиваются до максимума в правой части. Если изображение дискретизовать по равномерной сетке, то отсчет, взятый в правой части, будет содержать вклады от большего числа соседних точек, чем отсчет в левой части снимка. Один из «способов устранения этого недостатка состоит в увеличении интервала

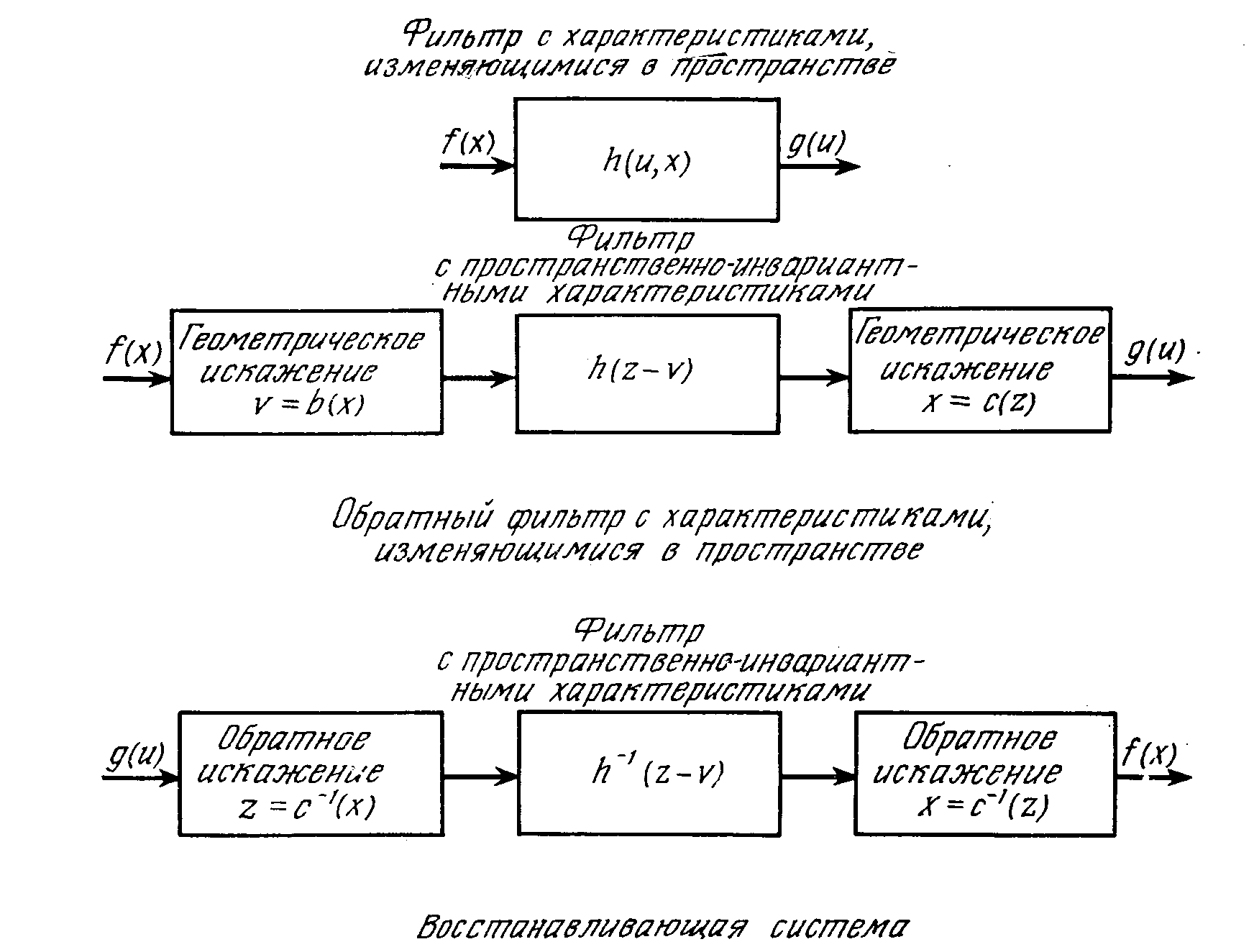


Рис. 4.17. Структурная схема процесса восстановления изображения с пространственно-зависимыми искажениями.

между отсчетами при движении слева направо так, чтобы каждый отсчет содержал взвешенную сумму яркостей одинакового числа соседних точек. Фактически это является преобразованием координат, делающим искажения пространственно-инвариантными. После коррекции исходное изображение восстанавливается с помощью преобразования координат, обратного к первому.

Савчук [46] показал, что этот общий метод можно применять для нескольких видов искажений, вызванных оптическими аберрациями или неравномерным сдвигом во время съемки. Обобщенная схема процесса коррекции показана на рис. 4.17. Пространственно-нестационарные искажения представляются в виде двух геометрических искажений. Сначала снимок переводится в координаты, где искажения оказываются пространственно-инвариантными, а затем осуществляется переход от пространственно-инвариантных координат к координатам, зависящим от положения деталей изображения. Система для восстановления изображения основывается на преобразованиях, обратных к этим двум, и фактическое восстановление выполняется путем линейной обработки в пространстве, где искажения являются пространственно-инвариантными. При этом для быстрой обработки больших изображений можно применять свертку и БПФ, а сами операции преобразования координат требуют малого или приемлемого числа вычислений. Подобная обработка успешно применялась для исправления аберраций типа комы [47] и искажений, вызванных пространственно-неравномерным сдвигом [46].

Атмосферная турбулентность обусловливает смазывания изображений, изменяющиеся с течением времени, поскольку температурные градиенты в воздушных слоях влияют на случайные фазовые задержки b волновых фронтах потоков света, несущих изображение. В силу усреднения за время экспозиции эти флуктуации создают искажения изображений, аппаратная функция которых приближается к гауссовской и стационарна. В ряде случаев делались попытки повышения резкости таких изображений [48]. Но, как правило, искажения оказывались достаточно сильными, и восстановление изображений было малоэффективньм. Недавно Нокс предложил новый метод коррекции атмосферных искажений изображения [49].

В любой момент времени изображение, проходящее сквозь турбулентную атмосферу, формируется под воздействием случайной аппаратной функции. При наблюдении за одним и тем же объектом можно получить набор снимков, описываемых соотношениями типа

(4.57)



в которых аппаратная функция искажений рассматривается как пространственно-инвариантная, но изменяющаяся во времени от момента к моменту (т.е. по переменной *i* ). Усредняя снимки (4.57) по времени, получаем

(4.58)



причем сложение *hi* со случайными фазами создает столь широкую усредненную аппаратную функцию, что большая часть высокочастотных составляющих *f(x у)* теряется.

Если же перед усреднением изображения подвергнуть преобразованию Фурье и возвести спектры в квадрат, то получается другой результат. В этом случае

(4.59)



где звездочка означает комплексное сопряжение. Возведение в квадрат «защищает» высокочастотную информацию, уничтожающуюся при усреднении фаз в равенстве (4.58). Если в наблюдаемом секторе находится точечный источник, то можно найти среднюю квадратическую аппаратурную функцию <*HiHi\**> и выполнить восстановление изображения. Однако в формуле (4.59) теряется фазовая информация, и восстановленное изображение равно обратному преобразованию от квадратного корня из энергетического спектра исходного изображения *f(x, у).*

Другой, но весьма близкий способ обработки состоит в вычислении статистической автокорреляционной функции спектра изображения:



(4.60)



Можно заметить, что при *u*1*=*v1*=*0 равенства (4.60) и (4.59) полностью совпадают. Если, как и прежде, имеется точечный источник, то

(4.61)



поскольку с помощью точечного источника можно найти комплексную автокорреляционную функцию усредненного по времени спектра аппаратной функции. Рассмотрим результат деления обеих частей равенства (4.61) на их модули:



(4.62)



где Ф — фазовая характеристика спектра F - изображения, рассматриваемая на двумерной плоскости фурье - преобразования. В правой части записано двумерное разностное уравнение относительно фазовой характеристики, а слева фигурируют только измеренные величины. Интегрирование этого разностного уравнения по всей плоскости преобразования Фурье даст фазовую характеристику, а в сочетании с модулем *F,* полученным из равенства (4.61), —спектр восстановленного изображения.

Моделирование метода Нокса — Томпсона показало, что он может значительно улучшить разрушение при наблюдениях сквозь турбулентную атмосферу. На рис. 4.18 этот процесс иллюстрируется с помощью изображения астероида, полученного моделированием на ЭВМ. На рис. 4.18, а приведен оригинал, на рис. 4.18,6 показаны четыре отдельных снимка с атмосферными искажениями, а на рис. 4.18, б—восстановленное изображение, причем фазовая информация была получена в соответствии с формулой (4.62) путем обработки 100 снимков, подобных приведенным на рис. 4.18,6.

**4.5. Воспроизведение изображений по проекциям**

Открытие проникающей радиации (рентгеновских лучей, нейтронных пучков и т.д.) дало возможность получать изображения объектов, ранее недоступных или доступных только с помощью грубых и зачастую нежелательных средств; примером такого объекта могут служить внутренние участки мозга. Подобные изображения имеют очень важное значение для развития медицины. Усиление контроля за качеством элементов больших конструкций обусловило важность таких изображений для неразрушающих методов контроля. Однако изображения, полученные путем просвечивания с помощью проникающего излучения, имеют недостаток: они являются двумерными теневыми проекциями трехмерных объектов. Важные особенности пространственного расположения внутренних частей объекта в процессе проектирования в лучшем случае искажаются, а в худшем — теряются вообще. При хирургических операциях (например, при опухолях мозга) незнание внутренней структуры может оказаться, очевидно, роковым обстоятельством.

Задаче воспроизведения внутренней структуры объекта по его проекциям в последнее время уделяется большое внимание, и методы цифровой обработки сигналов оказали заметное влияние на ее решение. Предложен ряд методов воспроизведения изображения по проекциям; список литературы по этому вопросу дан в статье Мерсеро и Оппенгейма [50]. Ниже будет рассмотрена задача воспроизведения изображения по проекциям и дано ее решение методом, характерным для цифровой обработки сигналов, а именно, с помощью преобразования Фурье.

**4.5.1. Образование проекций**

Изображения, получаемые с помощью проникающей радиации, образуются за счет ослабления луча в исследуемом веществе. Чем плотнее вещество, тем слабее интенсивность луча, прошедшего через вещество. Таким образом, изображение, наблюдаемое в проходящих лучах, определяется интегральным влиянием некоторой характеристики вещества объекта на интенсивность луча. Пусть *f(x1, x2,* *x3*) описывает распределение вещества в пространственных координатах *(х1, х2, x3).* Допустим, что просвечивающий луч направлен вдоль оси *х1,* как на рис. 4.19. Тогда распределение интенсивности проникающего излучения в плоскости (*x2, x3)* пропорционально функции *g,* определяемой соотношением

(4.63)



Важное свойство *проекций,* задаваемых формулами типа (4.63), можно заметить при анализе преобразования Фурье от функции *g* (*x2,x3*):

(4.64)



Трехмерное преобразование Фурье от исходного распределения имеет вид

(4.65)



Сравнивая *G* и *F,* видим, что

(4.66)



Таким образом, преобразование от проекции равно преобразованию Фурье от распределения вещества в объекте при *w1=0.* Такую функцию называют сечением преобразования Фурье,

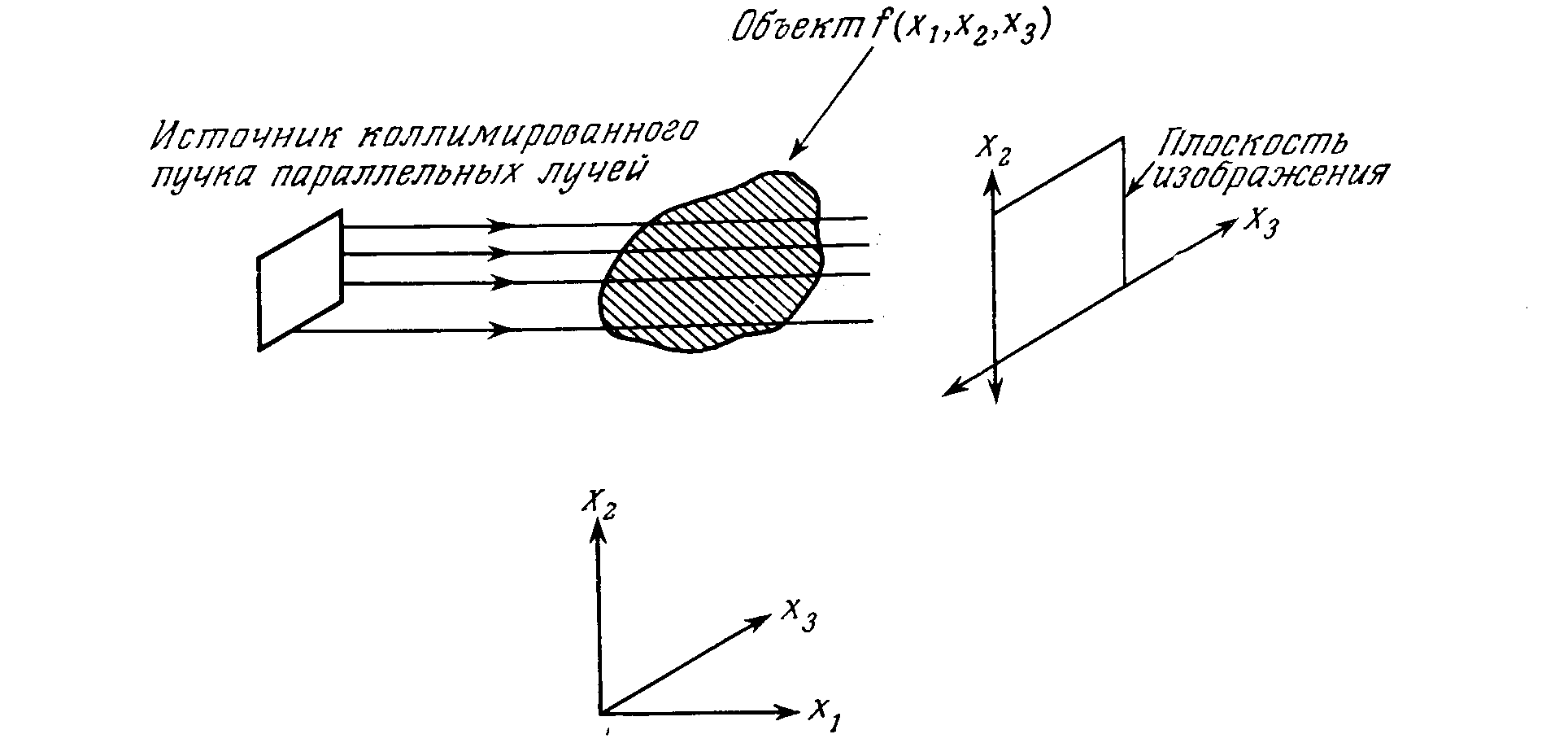


Рис. 4.19. Геометрические соотношения при воспроизведении изображения по проекциям.

12

поскольку она образуется сечением исходного трехмерного преобразования вдоль двумерной плоскости.

Очевидно, что аналогичными свойствами обладают также проекции меньших размерностей. Предположим, что проникающее излучение направлено в виде *плоского* луча (т.е. луч имеет бесконечно малый размер вдоль оси *x3* и перпендикулярен ей, причем вдоль оси *x2* его интенсивность одинакова, а ширина больше размеров объекта). Луч с координатой *x3* будет проектировать все элементы внутреннего строения объекта, лежащего в сечении, совпадающем с плоскостью луча. Если распределение вещества в сечении с координатой *x3* описывается функцией *f3( x1 , x2 )*, то одномерная проекция двумерного сечения имеет вид

(4.67)



и соотношение

(4.68)



как и прежде, описывает связь между преобразованиями проекции и оригинала.

Предположим теперь, что плоский луч проникающего излучения остается перпендикулярным оси *x3*, а источник излучения вращается вокруг некоторого центра, находящегося в объекте, так что угол между направлением луча и осью *x1* равен не нулю, а некоторой величине (см. рис. 4.20). Очевидно, можно сделать преобразование координат так, чтобы ось *и1* была параллельна направлению проектирующего луча. Это преобразование имеет вид



Равенства (4.67) и (4.68) остаются справедливыми и в новой системе координат (*u1*, *u2*), что позволяет сформулировать следующую *теорему о проекции и сечении,* одномерное преобразование Фурье от проекции под углом равно преобразованию Фурье исходного двумерного распределения вещества в двумерной плоскости спектральных переменных вдоль линии, направленной под углом , т.е. является сечением спектра под углом .



С помощью этой теоремы на основе проекций трехмерного тела можно воспроизвести изображение внутренней структуры тела. Если источник плоского луча расположить в точке с координатой *х3,* как показано на рис. 4.20, и изменять угол в интервале 0<<, то, как показано в следующем разделе, из набора одномерных проекций можно воспроизвести структуру тела в сечении с координатой *x3.* Затем координата *х3* изменяется и тем же

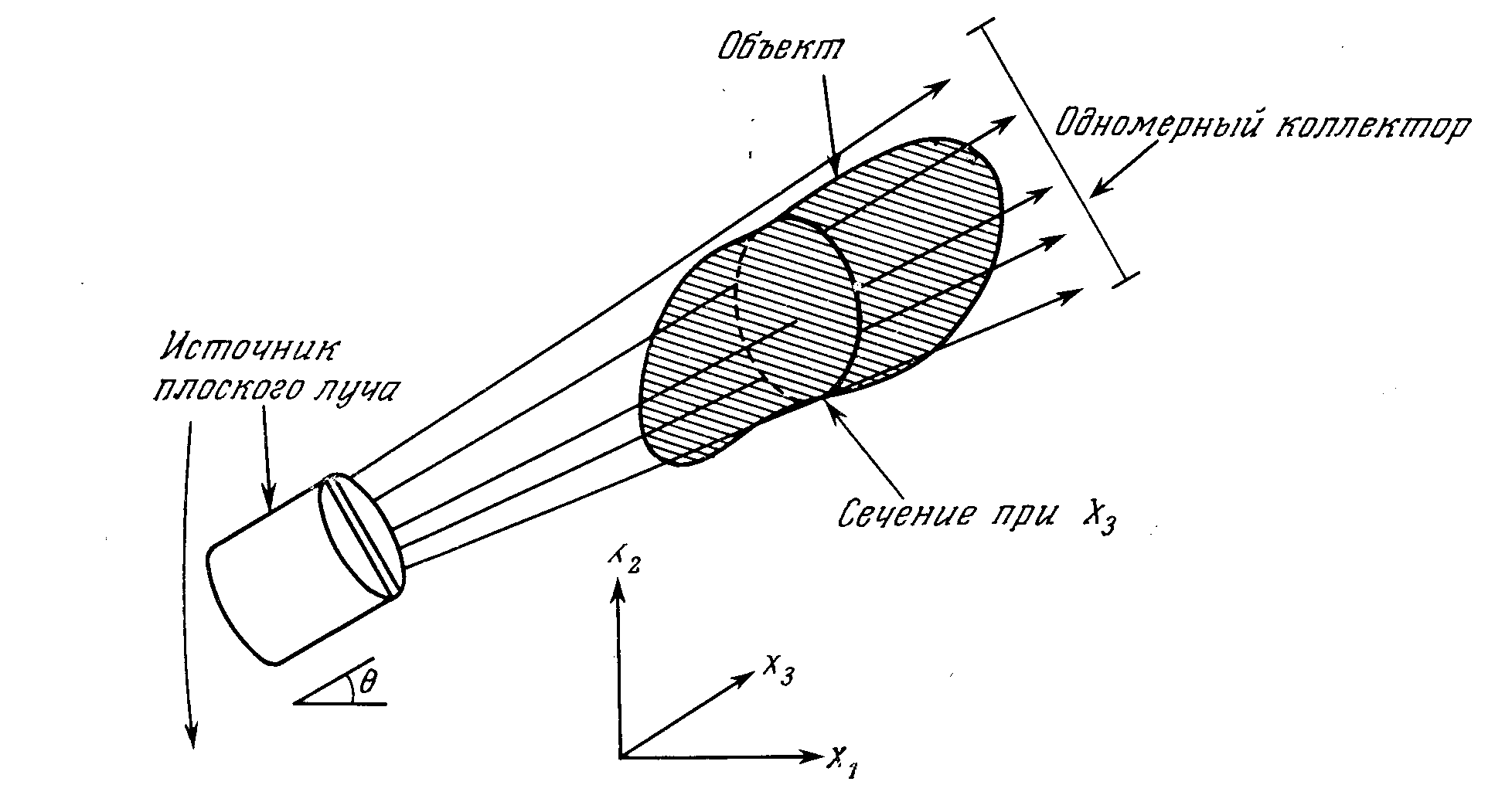


Рис. 4.20. Геометрические соотношения при вращении источника вокруг точки, зафиксированной в объекте, и схема получения двумерных проекций.

способом получается новое сечение. Процесс продолжается до тех пор, пока не будет получен полный набор сечений, показывающих внутреннюю структуру объекта в трех измерениях.

Теоретически, конечно, можно воспроизвести структуру трехмерного объекта по его двумерным проекциям без использования вспомогательных одномерных проекций и связанных с ними сечений, как описано в предыдущем абзаце. На практике, однако, выгоднее пользоваться методом плоских сечений, хотя и связанным с повторением одинаковых операций, но зато не требующих такого огромного числа вычислений, как прямое воспроизведение трехмерной структуры. Поэтому задачу воспроизведения будем в основном рассматривать как задачу выделения двумерной информации из одномерных проекций. Ниже описываются методы цифровой обработки сигналов, позволяющие воспроизвести изображение.

**4.5.2. Методы воспроизведения изображений**

Основная операция при воспроизведении изображения по проекциям следует из теоремы о проекции и сечении. Преобразование Фурье от проекции дает значения преобразования Фурье от исходного распределения вещества в некотором его сечении. Если получить достаточно плотный набор таких сечений, то все пространство спектров будет заполнено сечениями и с помощью обратного преобразования Фурье можно завершить процесс воспроизведения изображения.

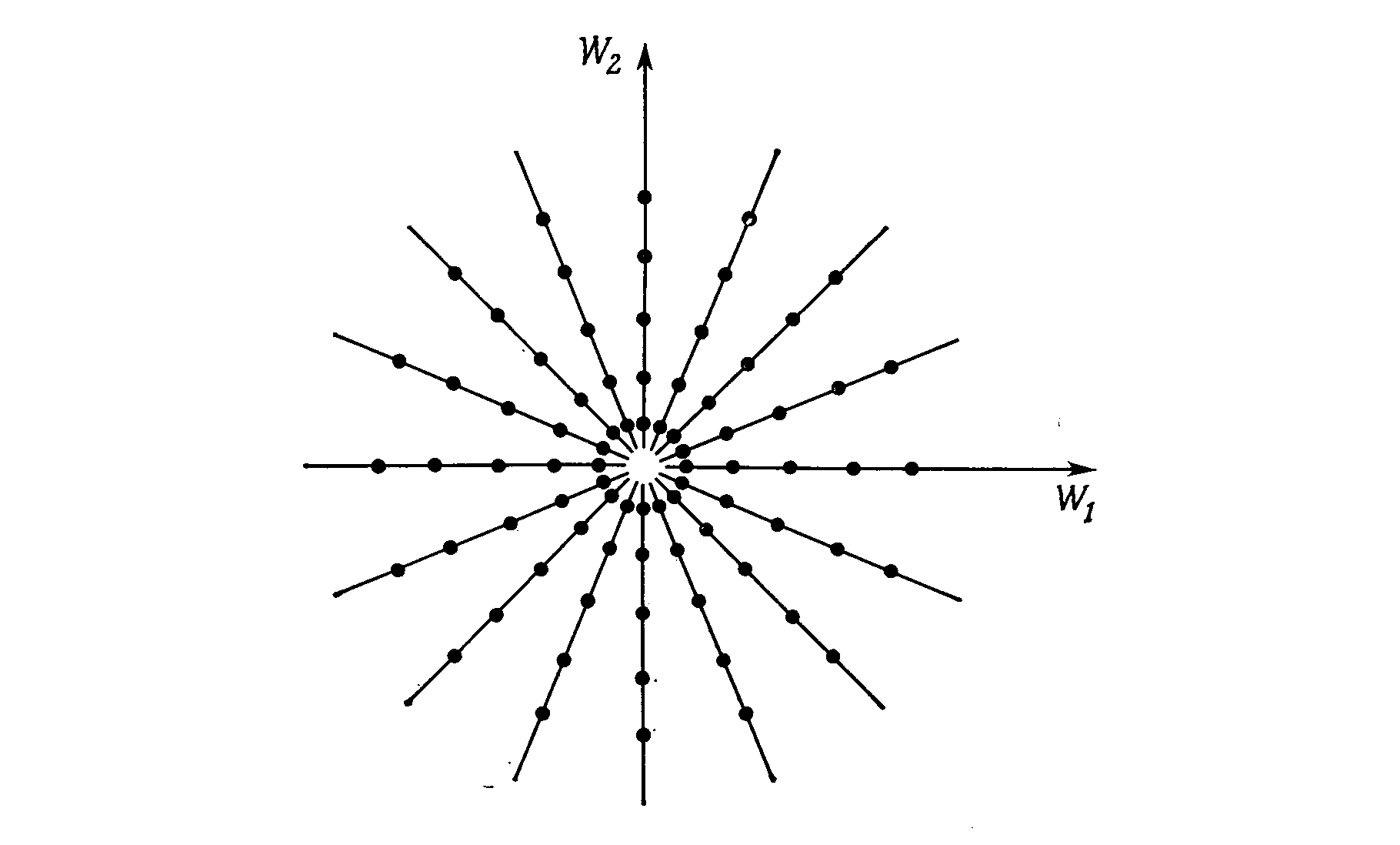


Рис. 4.21. Отсчеты в пространстве преобразований Фурье, полученные из набора проекций.

Каждая точка представляет значение *F(w1,w2),* полученное вычислением ДПФ проекции *g(u)* под углом .



Эта операция достаточно просто описывается, но для фактического ее выполнения необходимы весьма сложные и трудоемкие исследования. При воспроизведении изображения на основе сечений спектра важными являются следующие положения:

1. При вращении источника плоского луча для каждого из углов получается своя проекция. Число проекций, необходимых для достаточно полного заполнения пространства преобразований Фурье, очевидно, зависит от формы преобразования Фурье воспроизводимого распределения. Однако чем меньше шаг по углу, тем плотнее заполняется пространство спектров (см. рис. 4.21). На практике число проекций ограничивается двумя факторами: 1) увеличением объема вычислений, связанным с обработкой дополнительных проекций; 2) увеличением дозы облучения при получении дополнительных проекций (что очень важно в медицинских приложениях метода).



2. Сечения спектрального пространства образуют равномерную сетку в той системе координат, где на одной из осей откладывается угол поворота . Однако размещение отсчетов спектра в каждом из сечений не согласуется с обычными двумерными преобразованиями. Двумерное ДПФ вычисляется на прямоугольной сетке отсчетов, образующих строки и столбцы, а точки, в которых известны ДПФ сечений, образуют полярный растр, показанный на рис. 4.21. Поэтому необходимо с помощью интерполяции перейти к прямоугольному растру, что несложно с точки зрения теории, но нетривиально на практике, поскольку каждый отсчет, находящийся на концентрической полярной сетке, перед интерполяцией необходимо перевести в соответствующую прямоугольную систему координат. Такое преобразование двумерных координат и интерполяцию можно несколько упростить, пользуясь методом так называемого копцентрично - прямоугольного растра, но зато при этом интервалы дискретизации будут не одинаковыми [50].



3. Как видно из рис. 4.21, двумерное пространство преобразований Фурье очень плотно заполняется в области низких частот, а высокие частоты представлены гораздо реже. Поэтому для воспроизведения крупноструктурных элементов, описываемых низкочастотным спектром, требуется небольшое число проекций, например 20—30. Для достаточно четкого представления высокочастотных составляющих необходимо большое число проекций (100 и более). Если число проекций ограничено, то можно воспользоваться имеющимися проекциями и более плотно заполнить пространство спектров путем интерполяции. Это несколько улучшает результаты воспроизведения, но возможности данного метода ограничены, поскольку результаты интерполяции могут не соответствовать реальным значениям спектров. Качество и детали изображения определяются именно теми спектральными составляющими, которые нельзя предсказать путем интерполяции.

Другой метод воспроизведения изображений, в котором применяется цифровая обработка сигналов, называется методом свертки или «размазывания». Для пояснения метода рассмотрим две взаимно перпендикулярные проекции. Если каждую из проекций «размазать» (т.е. образовать двумерную функцию, не изменяющуюся по одной оси и совпадающую с исходной одномерной проекцией по второй оси) и значения обеих функций сложить как ортогональные векторы, то получится очень грубая аппроксимация исходного объекта. Читатель может проверить это, взяв для примера квадрат и построив проекции в направлениях, перпендикулярных его сторонам. В общем представляется возможным получить изображение «размазывая» проекции (это можно сделать путем свертки с подходящей функцией) и образуя сумму после соответствующего взвешивания отдельных слагаемых. Такой метод можно реализовать с помощью цифровой обработки, а также оптическими средствами. Однако данный метод крайне чувствителен к помехам, поскольку операции свертки и взвешивания влияют на обрабатываемые сигналы примерно так же, как дифференцирование.

Как образец результатов, получаемых при воспроизведении изображений цифровыми методами, на рис. 4.22 показано сечение человеческого мозга, полученное с помощью серийного рентгеновского аппарата для воспроизведения изображений внутреннего строения головы.

Этот краткий обзор не исчерпывает всех вопросов, связанных с воспроизведением изображений на основе проекций. С другими методами можно познакомиться по работам, указанным в списке литературы, составленном Мерсеро и Оппенгеймом [50].

**4.6. Повышение качества изображений**

Цель процесса повышения качества изображения состоит в том, чтобы снимок «выглядел лучше». Неудивительно поэтому, что если цель намечена так туманно, то и методы, применяемые для повышения качества изображений, оказываются весьма разнообразными. Субъективные суждения о том, что изображение «выглядит лучше», связаны также с критериями, зависящими от предназначения изображения (изображение должно «выглядеть лучше» применительно к определенной задаче). Если использование изображения связано с точным анализом или количественными измерениями, то радикальные операции, приводящие к значительному изменению пространственных или яркостных соотношений в изображении, могут оказаться неприемлемыми. С другой стороны если изображение применяется только для субъективных целей, то допустимы операции, существенно изменяющие пространственные или яркостные соотношения или же и те и другие, но в целом улучшающие субъективное восприятие изображения. Следовательно, для повышения качества изображения можно применять широкий круг методов; пригодность каждого из них зависит от целей повышения качества данного изображения.

1. **Повышение качества изображения путем пространственно- частотной фильтрации**

Для улучшения изображений часто используется пространственно-частотная фильтрация. Если даже снимок не имеет явных дефектов, можно применить фильтр с небольшим подъемом характеристики на верхних частотах, и снимок будет выглядеть более резким. Столь же полезна режекция постоянной составляющей, когда подавляются или ослабляются некоторые (или все) составляющие, расположенные вблизи нулевой частоты. В результате снижается насыщенность больших черных и белых пятен, а изменение масштабов яркости улучшает различимость мелких деталей. На рис. 4.23, а и *б* приведен пример повышения качества рентгенограммы тепловыделяющего элемента ядерного реактора с помощью пространственно-частотной фильтрации. На улучшенном снимке стало заметно гораздо больше деталей, а также видны ядерное топливо и оболочка элемента.

Особенно интересный метод повышения качества основан на мультипликативной модели формирования изображения в сочетании с гомоморфной фильтрацией [19]. Согласно законам поверхностного отражения, изображение образуется из двух компонент:

(4.69)



где *i’* - распределение освещающего пучка, а *r -* коэффициент отражения освещаемого объекта. Как правило, освещающая компонента образуется из низкочастотных пространственных составляющих, для которых коэффициент отражения приближается к зеркальному и богат деталями. Если прологарифмировать выражение (4.69):

(4.70)



то связь между коэффициентом отражения, освещением и изображением будет выражаться не произведением, а суммой. При фильтрации логарифма изображения режекторным фильтром, настроенным на нулевую частоту, освещающая компонента будет подавлена, а связанный с этим подъем высоких частот улучшит различимость мелкомасштабных элементов. При потенцировании сигнал возвращается в пространство яркостей и образуется изображение, не содержащее отрицательных яркостей. Заметим также, что логарифмирование обусловливает фильтрацию в пространстве плотностей пленки; это является еще одним доводом, дополняющим соображения о предпочтительности обработки в пространстве плотностей, высказанные ранее в разделе о восстановлении изображений.

На рис. 4.24, *а, б* приведен пример повышения качества изображения методом гомоморфной обработки. Заметим, что изображение стало более резким и на нем лучше различаются предметы, расположенные в тени под крышей.

**4.6.2. Повышение качества изображений с помощью**

**точечных операций**

Метод повышения качества изображений, основанный на пространственно-частотной фильтрации, можно противопоставить другим методам, в которых воздействие не распространяется на некоторую область (как для свертки), а все операции являются точечными изображения преобразуются в точки нового изображения независимо друг от друга. Точечные операции можно сгруппировать следующим образом.

**Преобразования контрастности.** Улучшение изображения происходит за счет изменения его контрастности, что достигается нелинейным преобразованием яркостей. Если, например, корректируемое изображение содержит участки, недодержанные при съемке, то можно воспользоваться преобразованием, «растягивающим» область малых яркостей и переводящим ее в интервал яркостей, более удобных для зрения. Наглядными примерами, в которых требуется подобное преобразование, служат операции коррекции неправильно экспонированных пленок, а так же линеаризации характеристик устройств демонстрации изображения, рассмотренные в первом разделе.

**Улучшение на основе статистических данных.** Выбор закона преобразования контрастности можно частично автоматизировать, воспользовавшись для подбора его параметров статистическими характеристиками изображения (например, средним значением или дисперсией яркости). Предельным случаем является метод выравнивания гистограмм. В теории информации показано, что равномерная гистограмма соответствует сообщению с максимальной информацией. Поэтому, если гистограмма квантованного изображения (дающая число отсчетов, попадающих на каждый из уровней квантования) после преобразования контраста становится равномерной (т.е. все уровни квантования проявляются с равной вероятностью), то изображение должно содержать максимальное количество информации. Данный метод обычно дает наилучшие результаты при квантовании яркостных изображений, гистограммы которых, как правило, отличаются наибольшей неравномерностью [13]. В результате можно довольно просто получить значительное улучшение изображения [51].

**Оконтуривание (препарирование) изображений.** При использовании всех рассмотренных методов решается задача повышения качества



Рис. 4.25. Блок-схема устройства отображения, обеспечивающего поточечное улучшение изображении с непосредственным участием оператора.

изображения без существенного его изменения. Если же цель обработки состоит в том, чтобы облегчить восприятие определенной информации, то очень часто применяют методы оконтурирования, когда возможно заметное искажение яркостных и (или) пространственных соотношений. Наиболее распространенным

является метод псевдоцвета, в котором различным яркостям произвольно сопоставляются разные цвета. Демонстрируемое изображение будет содержать отчетливые контуры, проходящие по границам цветных полос. В результате может либо произойти четкое выделение важных деталей, либо получится обманчивое смешение пятен, скрывающее изображение, либо может образоваться забавная цветная картинка, ничего не выделяющая и ничего не скрывающая. В другом методе производится оконтуривание границами черного и белого цвета путем уменьшения числа уровней квантования (обычно до 10 и менее). Отбрасывание от одного до трех старших разрядов также создает контуры, причем картина, получаемая при выделении деталей таким образом, может оказаться очень живописной.

Весьма интересными применительно к точечным операциям повышения качества изображений являются последние образцы цифровых устройств отображения информации (рис. 4.25), позволяющие оперативно корректировать изображение. С помощью быстодействующих постоянных запоминающих устройств (ПЗУ) удается изменять яркости точек изображения при передаче их из ЗУ на электронно-лучевую трубку. Исходное же изображение, записанное на диски, при этом остается неизменным. Таким образом, загрузив в ПЗУ различные функции, описывающие закон изменения яркостей, можно очень быстро переходить от одного способа преобразования яркости к другому. Нужно видеть такое устройство, чтобы полностью оценить его гибкость в улучшении контрастности, коррекции ошибок экспонирования, подстановке псевдоцвета и т.д. Подобные цифровые устройства отображения превращают точечные операции в эффективное и удобное средство улучшения изображений, обеспечивающее возможность взаимодействия человека с машиной.

Автор хотел бы выразить признательность д-ру Томасу Г. Стокхему мл., который внимательно прочитал рукопись и сделал замечания. Автор также глубоко благодарен всем, кто оказал любезность, предоставив приведенные выше снимки: д-ру Вильяму К. Пратту из Университета штата Южная Каролина (рис. 4.9 и 4.10 в разделе о сокращении избыточности изображений), д-ру Т. М. Кэннону из Лос-Аламосской научной лаборатории (рис. 4.16, а, б и 4.22) д-ру Т.Г. Стокхему (мл.) и д-ру Б. Бекстеру из Университета шт. Юта (рис. 4.23, а, *б),* а также д-ру Е. Баррету из фирмы ЕСЛ (рис. 4.18, а - *в).*