

Министерство образования Республики Беларусь

Учреждение образования
«Белорусский государственный университет
информатики и радиоэлектроники»

Кафедра электронных вычислительных машин

КУРС ЛЕКЦИЙ

по дисциплине «Цифровая обработка сигналов и изображений»
для студентов специальности I-40 02 01
«Вычислительные машины, системы и сети»
всех форм обучения

Минск 2007

СОДЕРЖАНИЕ

	Введение. Предмет курса, его задачи, структура.....	4
1.	Сигналы и их особенности.....	6
2.	Ортогональные преобразования в задачах обработки сигналов и изображений. Основные свойства дискретного преобразования Фурье.....	12
3.	Матричное представление корреляции и свертки.....	23
4.	Техника быстрого преобразования Фурье (БПФ) и основные подходы к его реализации.....	30
5.	Алгоритмы формирования ортогональных прямоугольных функций.....	41
6.	Представление сигналов в виде функций Уолша. Алгоритм быстрого преобразования Уолша.....	44
7.	Многомерные преобразования для обработки изображений. Простые вейвлеты. Масштабирующие функции и вейвлет – функции. Прямое и обратное вейвлет-преобразование.....	54
8.	Алгоритмы для вычисления преобразования Хаара.....	73
9.	Стационарные случайные процессы.....	78
10.	Функции спектральной плотности и определения спектров с помощью корреляционных функций.....	85
11.	Определение спектров с помощью фильтрации.....	91
12.	Цифровые методы анализа: оценивание корреляционных функций методом БПФ, оценки спектральной плотности мощности и алгоритмы вычисления.....	94
13.	Получение изображения. Основные понятия предварительной обработки изображений.....	110
14.	Основные преобразования изображения.....	114
15.	Основные взаимосвязи между пикселями изображения. Метрические свойства изображения.....	117
16.	Методы пространственной и частотной области	122

	применительно к предварительно обработке изображения.....	
17.	Сегментация изображения посредством выделения границ областей.....	126
18.	Основные свойства двумерного преобразования Фурье.....	131
19.	Основы фильтрации в частотной области.....	138
20.	Сжатие данных.....	144
21.	Выбор признаков и распознавание образов. Принцип обучения.....	153
22.	Виды разделяющих функций. Классификатор по минимальному расстоянию.....	157
23.	Структура системы распознавания образов с применением ортогональных преобразований. Принцип обучения в двухклассовой системе распознавания.....	162
24.	Задача трех классов образов по критерию минимального расстояния.....	167
25.	Метод отображения по алгоритму наименьших квадратов.....	171
26.	Классификатор для распознавания трёх классов образов по критерию наименьшего среднеквадратичного расстояния.....	175
27.	Классификатор для распознавания k классов образов по критерию наименьшего среднеквадратного расстояния.....	179
28.	Модель нейронной сети. Обучающий алгоритм для перцептрона.....	182
29.	Искусственные нейронные сети.....	193
	Литература.....	221

ВВЕДЕНИЕ. ПРЕДМЕТ КУРСА, ЕГО ЗАДАЧИ, СТРУКТУРА.

Основной целью дисциплины «Цифровая обработка сигналов и изображений» является подготовка студентов по следующим направлениям данной предметной области:

- элементы прикладной теории обработки изображений идентификации и распознавания образов;
- методы, алгоритмы и программные средства обработки изображений и распознавания образов;
- техника решения класса актуальных прикладных задач, включающих основные этапы распознавания образов.

Настоящий курс дает основы технологии, методов и средств обработки изображений, распознавания образов.

Содержание курса ориентировано на теоретическую и прикладную подготовку студентов в рамках создания программных комплексов на базе ЭВМ для решения задач в различных областях народного хозяйства республики Беларусь:

- на всех уровнях организационного управления при создании автоматизированных систем обработки информации;
- в отраслях промышленности при разработке систем идентификации и контроля;
- в медицине при создании автоматизированных рабочих мест для идентификации грудной клетки, анализа крови;
- в экологии при разработке комплексов обработки изображений картографических данных и аэрофотоснимков территории республики;
- в военном строительстве при создании систем специального назначения;

- для правоохранительных органов при разработке систем идентификации сигнатур (подписи), портретов лиц, отпечатков пальцев.

Дисциплина базируется на современных достижениях в области обработки изображений и распознавания образов и имеет прикладную направленность, т.е. ориентирована на решение актуальных народнохозяйственных задач на основе новых информационных технологий.

Перечисленные цели достигаются посредством лекций, лабораторных занятий, консультаций и зачета.

1. СИГНАЛЫ И ИХ ОСОБЕННОСТИ

Проблемы извлечения, передачи и обработки информации являются центральными для многих областей науки и техники, например, связь, автоматическое управление и регулирование, радиолокационные системы, распознавание образов. С физической точки зрения сигнал создается определенным процессом, протекающим во времени.

Важнейшими формами аналитического выражения сигнала является представление записи этого сигнала с помощью колебаний или спектра (временное или частотное представление).

Колебание $s(t)$ описывает сигнал как функцию от времени. Это основная форма аналитического выражения, поскольку t - естественная координата, с которой связаны все реальные явления. В свою очередь, сигнал можно рассматривать как совокупность элементарных функций $\eta_k(t)$, умноженных на коэффициент C_k и составляющих систему функций $\{\eta_k(t)\}$ определенного типа:

$$s(t) = \sum_{k=0}^{\infty} C_k \eta_k(t). \quad (1.1)$$

Система функций $\{\eta_k(t)\}$ носит название базисной системы, а представление сигнала в виде (1.1) называется разложением сигнала по системе базисных функций. Основными задачами теории сигналов являются:

1. анализ сигналов (изучение их свойств);
2. синтез сигналов (нахождение сигнала, обладающего заданными свойствами).

Для отображения исходной функции времени $x(t)$ в спектральной области служит, к примеру, прямое преобразование Фурье:

$$x(j\omega) = \int_0^{\infty} x(t) e^{-j\omega t} dt, \quad (1.2)$$

а для восстановления функции по ее спектру – обратное преобразование Фурье:

$$x(t) = \frac{1}{2\pi} \int_0^{\infty} x(j\omega) e^{j\omega t} d\omega. \quad (1.3)$$

Структуру, с помощью которой можно воспроизвести задачи анализа и синтеза, можно представить (рис. 1.1):

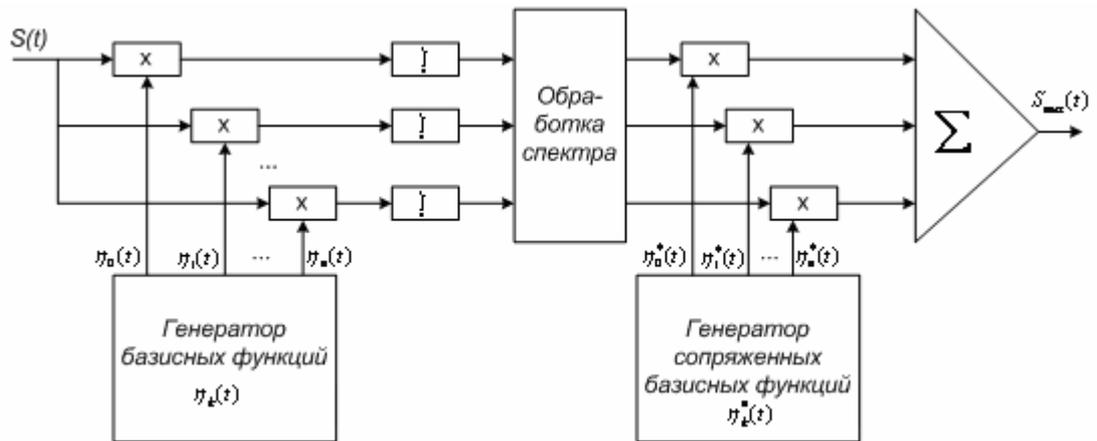


Рис. 1.1. Задачи анализа и синтеза сигналов

В математическом смысле задача спектрального анализа сигнала эквивалентна выполнению прямого преобразования Фурье по системе базисных функций $\eta_k(t)$, а задача синтеза – выполнению обратного преобразования Фурье по системе комплексно сопряженных $\eta_k^*(t)$. В результате спектрального анализа сигнала $s(t)$ получаем спектр $\{C_k\}$, а по спектру синтезируем новый сигнал с заданной точностью (так строятся устройства для фильтрации, сжатия данных). Реальные сигналы всегда ограничены во времени, т.к. мы не можем вести их наблюдение бесконечно долго и, как правило, являются в какой-то мере случайными.

Случайным называется сигнал, значение которого в каждый момент времени является случайной величиной.

Сигнал, известный в любой момент времени, называется *детерминированным*.

1. *Каузальным* называется колебание, имеющее начало во времени, которое можно рассматривать как причинное, т.к. оно может определяться как следствие какой-то причины. При описании каузальных колебаний удобно совместить начало отсчета времени с началом колебаний, которое будет равно 0 при $t < 0$ (рис. 1.2).

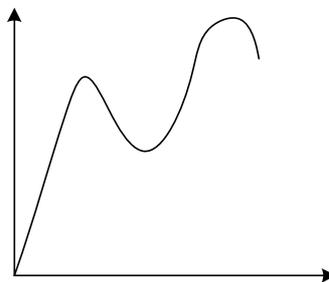


Рис.1.2. Каузальное колебание

2. *Периодическим* называется колебание, которое задается на интервале $-\infty \leq t \leq +\infty$ и любое значение повторяется через интервалы времени, равные T (период):

$$S(t) = S(t + T).$$

3. *Финитным* называется колебание, локализованное во времени, т.е. колебание равно нулю вне некоторого ограниченного интервала времени $t_A \leq t \leq t_B$ (рис. 1.3).

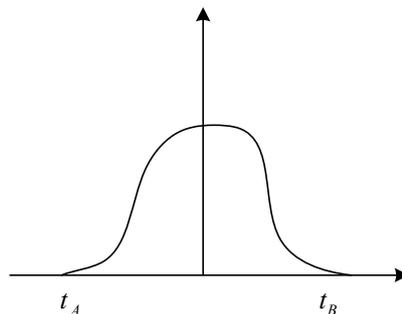


Рис. 1.3. Финитное колебание

4. *Непрерывным* называется колебание, которое рассматривается в каждой точке оси времени, т.е. такое колебание задано на несчетном временном интервале (рис. 1.4).

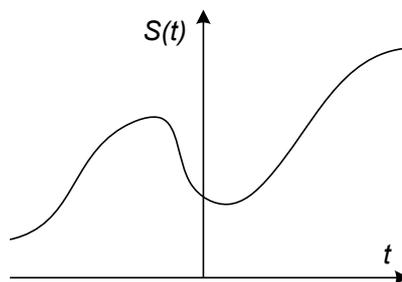


Рис. 1.4. Непрерывное колебание

5. *Дискретное* колебание рассматривается только в фиксированный момент времени, т.е. заданное на счетном множестве временных точек (рис. 1.5).

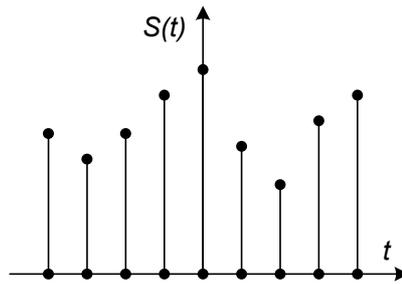


Рис. 1.5. Дискретное колебание

Обычно в динамических системах важным свойством является способность преобразовывать анализируемые гармоники по амплитуде и фазе:

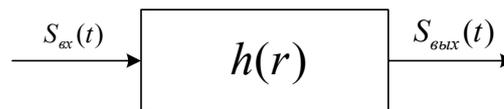


Рис. 1.6. Преобразование гармоник по амплитуде и фазе

Соотношение входного и выходного сигналов может описываться

$$S_{\text{вых}}(t) = \int_0^{\infty} S_{\text{вх}}(r)h(t-r)dr \quad (1.4)$$

Если на входе идеального дискретизатора, осуществляющего взятие отсчетов с заданной скоростью, составляющей N отсчетов в секунду, действует сигнал $x(t)$, то на его выходах формируется дискретизированный во времени сигнал, определенный как:

$$\tilde{x}(t) = \Delta t \sum_{m=0}^{N-1} x(m\Delta t)\delta(t-m\Delta t), \quad (1.5)$$

где $\delta(t)$ - функция Дирака.

$\delta(t)$ - функция, равная 0 всюду, кроме начала координат. В начале координат данная функция принимает бесконечное значение и интеграл от нее по любому интервалу, содержащему начало координат, будет равен 1:

$$\delta(t) = 0 \text{ при } t \neq 0; \delta(0) = \infty$$

$$\int_{-\varepsilon}^{+\varepsilon} \delta(t) dt = 1 \text{ при любом } \varepsilon > 0 \quad (1.6)$$

Функцию, обладающую такими свойствами, можно получить как предел прямоугольного импульса, имеющего единичную площадь, когда длительность этого импульса стремится к 0.

2. ОРТОГОНАЛЬНЫЕ ПРЕОБРАЗОВАНИЯ В ЗАДАЧАХ ОБРАБОТКИ СИГНАЛОВ И ИЗОБРАЖЕНИЙ. ОСНОВНЫЕ СВОЙСТВА ДИСКРЕТНОГО ПРЕОБРАЗОВАНИЯ ФУРЬЕ

2.1. Ортогональность сигналов

Множество непрерывных функций действительного переменного $\{U_n(t)\} = \{U_0(t), U_1(t), \dots\}$ называется ортогональным на интервале $[t_0, t_0 + T]$, если

$$\int_{t_0}^{t_0+T} U_m(t)U_n(t)dt = \begin{cases} c, \forall m = n, \\ 0, \forall m \neq n \end{cases} \quad (2.1)$$

При $c = 1$ множество $\{U_n(t)\}$ называется ортонормированным.

Для вычисления сигнала через коэффициенты разложения используется:

$$x(t) = \sum_{n=0}^{\infty} a_n U_n(t) \quad (2.2)$$

Коэффициент разложения из указанного соотношения можно определить, если умножить обе части (2.2) на $U_m(t)$ и проинтегрировать в интервале $[t_0, t_0 + T]$

$$\int_{t_0}^{t_0+T} x(t)U_m(t)dt = \sum_{n=0}^{\infty} a_n \int_{t_0}^{t_0+T} U_n(t)U_m(t)dt \quad (2.3)$$

В силу условий ортогональности будем иметь

$$a_n = \frac{1}{C} \int_{t_0}^{t_0+T} x(t)U_n(t)dt \quad (2.4)$$

Запишем соотношение (2.8) в матричной форме:

$$\vec{x} = U\vec{a}, \quad (2.5)$$

где \vec{x} - вектор - столбец, U – матрица ортогональных функций.

Умножим обе части на матрицу U^{-1}

$$U^{-1}\vec{x} = U^{-1}U\vec{a} \quad (2.6)$$

так как $U^{-1}U = UU^{-1} = CI_d$, где I_d - единичная матрица,

$$\text{тогда } a = \frac{1}{C}U^{-1}\vec{x} \quad (2.7)$$

2.2. Теорема Парсеваля

Для доказательства теоремы Парсеваля возведем обе части соотношения (2.2) в квадрат:

$$x(t)^2 = \sum_{n=0}^{\infty} U_n(t)^2 a_n^2 + \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} a_p a_q U_p(t) U_q(t) \quad (2.8)$$

Проинтегрируем обе части

$$\frac{1}{T} \int_T x(t)^2 dt = \sum_{n=0}^{\infty} a_n^2 \frac{1}{T} \int_T U_n(t)^2 dt + \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} a_p a_q \int_T U_p(t) U_q(t) dt \quad (2.9)$$

По условию ортогональности

$$\frac{1}{T} \int_T x(t)^2 dt = \frac{C}{T} \sum_{n=0}^{\infty} a_n^2 \quad (2.10)$$

Если $x(t)$ - напряжение или ток, приложенный к концам резистивной нагрузки равной 1 Ом, то левая часть выражения представляет мощность рассеиваемой энергии, а $\frac{C}{T} \sum_{n=0}^{\infty} a_n^2$ является распределением мощности в $x(t)$.

2.3. Ряд Фурье

Для теории формирования и обработки сигнала особое значение имеет возможность разложения заданного в виде функции сигнала по различным ортогональным системам функций.

Впервые в 1807 году французский математик и физик Жан Батист Жозеф Фурье показал, что любую произвольную функцию $x(t)$ можно представить в виде бесконечной суммы синусных и косинусных членов:

$$x(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos n\omega_0 t + \sum_{n=1}^{\infty} b_n \sin n\omega_0 t, \quad (2.11)$$

где ω_0 (рад/с) – основная угловая частота, которая связана с периодом T функции соотношением $T = 2\pi/\omega_0$. Частоты $n\omega_0$ называют гармониками, так как они кратны основной частоте ω_0 . В данном случае речь идет о системе ортогональных функций вида $\{1, \cos n\omega_0 t, \sin n\omega_0 t\}$.

Коэффициенты $\{a_0, a_n, b_n\}$ из формулы (2.11) можно вычислить с учетом ортогональности множества функций $\{\cos n\omega_0 t, \sin n\omega_0 t\}$ на периоде T ;

$$\int_T \cos n\omega_0 t \cos m\omega_0 t dt = \begin{cases} T/2, & m = n, \\ 0, & m \neq n \end{cases} \quad (2.12)$$

$$\int_T \cos n\omega_0 t \sin m\omega_0 t dt = 0, \forall m, n \quad (2.13)$$

$$\int_T \sin n\omega_0 t \sin m\omega_0 t dt = \begin{cases} T/2, m = n, \\ 0, m \neq n \end{cases} \quad (2.14)$$

С учетом этих соотношений:

$$a_0 = \frac{1}{T} \int_T x(t) dt; \quad (2.15)$$

$$a_n = \frac{2}{T} \int_T x(t) \cos n\omega_0 t dt; \quad (2.16)$$

$$b_n = \frac{2}{T} \int_T x(t) \sin n\omega_0 t dt. \quad (2.17)$$

Раздел математики, устанавливающий соотношение между функцией $x(t)$ и коэффициентами a_n и b_n , называется гармоническим анализом, а представление (2.11) – рядом Фурье.

Компоненты ряда Фурье называются гармониками. Любая четная функция может быть разложена в ряд Фурье, состоящий из косинусов, а любая нечетная функция раскладывается в ряд из синусов. Для некоторых функций ряд Фурье может состоять лишь из нечетных гармоник.

В целом, любая полная система ортогональных функций может быть применена для разложения в ряды, которые соответствуют рядам Фурье. Например, часто используется разложение в ряды по функциям Уолша, Хаара, Лагерра, Бесселя и т.д.

2.4. Семейство преобразований Фурье

Преобразование Фурье (Fourier transform) – это разложение функций на синусоиды (далее косинусные функции также будем называть синусоидами, т.к. они отличаются от «настоящих» синусоид только фазой). Анализ Фурье закладывает основы многих методов, применяющихся в цифровой обработке сигналов и изображений (ЦОСиИ). По сути, преобразование Фурье (ПФ) позволяет сопоставить сигналу, заданному во временной области, его эквивалентное представление в частотной области. Обратное, если известна частотная характеристика сигнала, то обратное преобразование Фурье позволяет определить соответствующий сигнал во временной области.

Семейство преобразований Фурье (преобразование Фурье, ряды Фурье, дискретные ряды Фурье и дискретное преобразование Фурье) представлено на рис. 2.1 – 2.4.

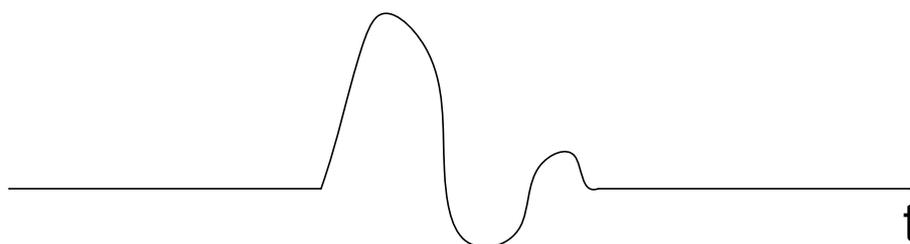


Рис. 2.1. Преобразование Фурье: сигнал непрерывный и аperiodический

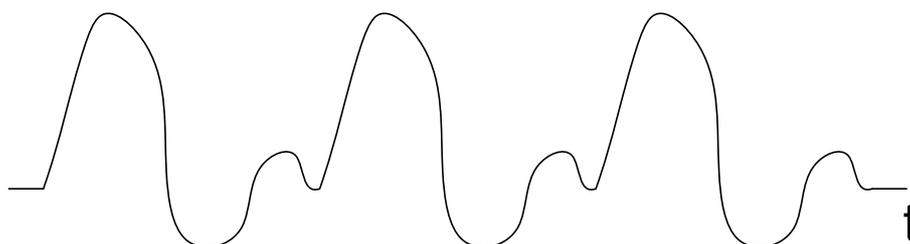


Рис. 2.2. Ряды Фурье: сигнал непрерывный и периодический

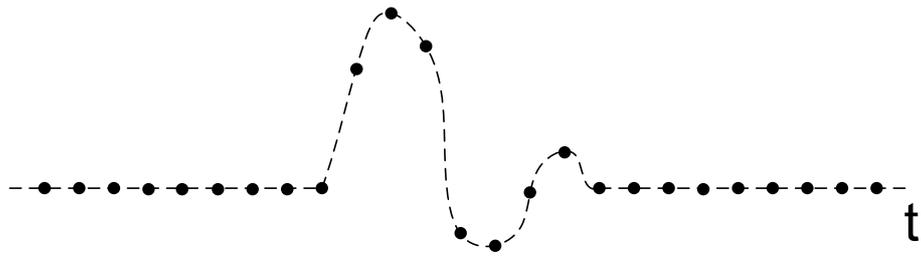


Рис. 2.3. Дискретные ряды Фурье: сигнал дискретный и аperiodический

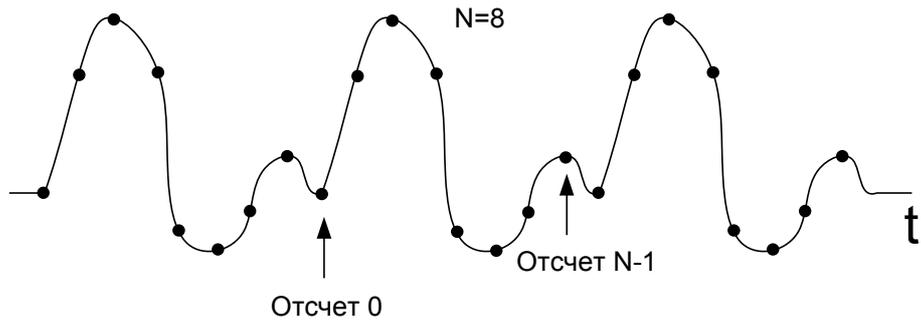


Рис. 2.4. Дискретное преобразование Фурье: (дискретные ряды Фурье) сигнал дискретный и периодический

2.5. Дискретное преобразование Фурье (ДПФ)

Из описанного семейства преобразований к цифровой обработке сигналов и изображений имеет отношение дискретное преобразование Фурье, которое оперирует дискретной по времени выборкой периодического сигнала во временной области. Для того, чтобы быть представленным в виде суммы синусоид, сигнал должен быть периодическим. Но в качестве набора входных данных для ДПФ доступно только конечное число отсчетов (N).

Основная идея ДПФ ни чем не отличается от ПФ (см. рис. 2.5).

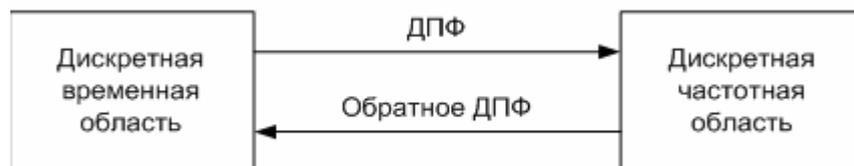


Рис. 2.5. Основная идея ДПФ

Для получения представления $x(t)$ (2.11) рядом Фурье в комплексной форме необходимо использовать соотношения в виде формулы Эйлера:

$$\begin{aligned}\cos n\omega_0 t &= \frac{1}{2}(e^{in\omega_0 t} + e^{-in\omega_0 t}); \\ \sin n\omega_0 t &= \frac{1}{2i}(e^{in\omega_0 t} - e^{-in\omega_0 t}); \quad i = \sqrt{-1}.\end{aligned}\tag{2.18}$$

Тогда

$$\begin{aligned}x(t) &= a_0 + \frac{1}{2} \sum_{n=1}^{\infty} \{a_n(e^{in\omega_0 t} + e^{-in\omega_0 t}) - ib_n(e^{in\omega_0 t} - e^{-in\omega_0 t})\} = \\ &= a_0 + \frac{1}{2} \sum_{n=1}^{\infty} \{(a_n - ib_n)e^{in\omega_0 t} + (a_n + ib_n)e^{-in\omega_0 t}\}\end{aligned}\tag{2.19}$$

Введем коэффициент

$$C_n = \frac{1}{2}(a_n - ib_n).$$

Тогда

$$C_n = \frac{1}{T} \int_T x(t) [\cos n\omega_0 t - i \sin n\omega_0 t] dt \quad \text{или} \quad C_n = \frac{1}{T} \int_T x(t) e^{-in\omega_0 t} dt;$$

$$C_{-n} = C_n^* = \frac{1}{2}(a_n + ib_n).\tag{2.20}$$

Следовательно,

$$x(t) = a_0 + \sum_{n=1}^{\infty} [C_n e^{in\omega_0 t} + C_n^* e^{-in\omega_0 t}] = a_0 + \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} C_n e^{in\omega_0 t} ;$$

$$C_0 = \frac{1}{T} \int_T x(t) dt = a_0 ;$$

$$x(t) = \sum_{n=-\infty}^{\infty} C_n e^{in\omega_0 t} . \quad (2.21)$$

Таким образом, если $\{X(m)\}$ означает последовательность $X(m)$ конечных действительных или комплексных чисел, где $m = \overline{0, N-1}$, то дискретное преобразование Фурье этой последовательности определяется как

$$\boxed{C_x(k) = \frac{1}{N} \sum_{m=0}^{N-1} X(m) W^{km}}, \text{ где } k = \overline{0, N-1}, \quad W = e^{-i2\pi/N}, \quad i = \sqrt{-1}, \quad (2.22)$$

$$\boxed{X(m) = \sum_{k=0}^{N-1} C_x(k) W^{-km}} . \quad (2.23)$$

Выражения (2.22), (2.23) составляют *пару преобразований Фурье*.

Функции W^{km} являются N -периодическими, т.е. $W^{km} = W^{(k+N)m} = W^{k(m+N)}$.

Следовательно, последовательности $\{C_x(k)\}$, $\{X(m)\}$ также являются N -периодическими, т.е.

$$X(\pm m) = X(SN \pm m);$$

$$C_x(\pm k) = C_x(SN \pm k).$$

Рассмотрим *основные свойства дискретного преобразования Фурье*:

а) *теорема линейности*: дискретное преобразование Фурье является линейным, т.е. если $X(m) \leftrightarrow C_x(k)$, $Y(m) \leftrightarrow C_y(k)$ и $Z(m) = aX(m) + bY(m)$, то $C_z(k) = aC_x(k) + bC_y(k)$;

б) *теорема комплексной сопряженности*: если $\{X(m)\} = \{X(0), X(1), \dots, X(N-1)\}$ – такая последовательность действительных чисел, что $N/2$ – целое число и $X(m) \leftrightarrow C_x(k)$, то

$$C_x\left(\frac{N}{2} + l\right) = \overline{C_x\left(\frac{N}{2} - l\right)}, \forall l = \overline{0, N/2}. \quad (2.24)$$

Из (2.22) следует, что $C_x(k) = \frac{1}{N} \sum_{m=0}^{N-1} X(m)W^{km}$, где $W = e^{-i2\pi/N}$.

Тогда, подставляя вместо $k - (N/2 + l)$, будем иметь

$$C_x\left(\frac{N}{2} + l\right) = \frac{1}{N} \sum_{m=0}^{N-1} X(m)W^{(N/2+l)m} = \frac{1}{N} \sum_{m=0}^{N-1} X(m)W^{-(N/2-l)m} W^{Nm} = \overline{C_x\left(\frac{N}{2} - l\right)},$$

т.к. $W^{Nm} = e^{-i2\pi m} \equiv 1$.

в) *теорема сдвига*: если $Z(m) \leftrightarrow C_z(k)$ и $Z(m) = X(m+h)$, $h = \overline{0, N-1}$, то

$$C_z(k) = W^{-kh} C_x(k). \quad (2.25)$$

Доказательство:

$$Z(m) \leftrightarrow C_z(k), \text{ т.е. } C_z(k) = \frac{1}{N} \sum_{m=0}^{N-1} Z(m)W^{km}, \quad k = \overline{0, N-1}.$$

С учетом подстановки $Z(m) = X(m+h)$, будем иметь $C_z(k) = \frac{1}{N} \sum_{m=0}^{N-1} X(m+h)W^{km}$.

Осуществляя замену переменных $m+h=r$, указанное соотношение будет иметь вид $C_z(k) = W^{-kh} \left\{ \frac{1}{N} \sum_{r=h}^{N+h-1} X(r)W^{kr} \right\}$.

Так как

$$\sum_{m=p}^q X(m)W^{km} = \sum_{m=0}^{N-1} X(m)W^{km} ,$$

$$\sum_{m=p}^q C_x(k)W^{-km} = \sum_{m=0}^{N-1} C_x(k)W^{-km} , \text{ когда } p \text{ и } q \text{ удовлетворяют условию}$$

$$|p-q|=N-1, \text{ то } C_z(k)=W^{-kh}C_x(k).$$

Аналогично при $Z(m)=X(m-h)$, $C_z(k)=W^{kh}C_x(k)$.

Можно выделить следующие области применения ДПФ:

- цифровой спектральный анализ
 - анализаторы спектра
 - обработка речи
 - обработка изображений
 - распознавание образов
- проектирование фильтров
 - вычисление импульсной характеристики по частотной
 - вычисление частотной характеристики по импульсной
- быстрое преобразование Фурье (БПФ) – простой алгоритм для эффективного вычисления ДПФ.

ДИСКРЕТНОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ (ДПФ)

- Периодический сигнал может быть разложен на сумму выбранных должным образом косинусоидальных и синусоидальных функций (Жан Батист Жозеф Фурье, 1807).
- ДПФ работает с конечным числом (N) оцифрованных по времени отсчетов $X(m)$. Когда эти группы отсчетов повторяются, они становятся периодическими с точки зрения преобразования.
- Комплексный спектральный выход ДПФ $S(k)$ является результатом

свертки входных отсчетов с базисными функциями синуса и косинуса.

3. МАТРИЧНОЕ ПРЕДСТАВЛЕНИЕ КОРРЕЛЯЦИИ И СВЕРТКИ

Свертка – это математический способ комбинирования двух сигналов для формирования третьего сигнала. Это один из самых важных методов ЦОС. Пользуясь стратегией импульсного разложения, системы описываются сигналом, называемым импульсной характеристикой. Свертка важна, так как она связывает три сигнала: входной сигнал, выходной сигнал и импульсную характеристику.

Корреляция, так же, как свертка, использует два сигнала для получения третьего. Этот третий сигнал называется корреляционным сигналом двух входных сигналов.

3.1. Теорема свертки

Если $\{X(m)\}$ и $\{Y(m)\}$ – последовательности действительных чисел, для которых $X(m) \leftrightarrow C_x(k)$, $Y(m) \leftrightarrow C_y(k)$, и свертка этих последовательностей определяется как

$$Z(m) = \frac{1}{N} \sum_{h=0}^{N-1} X(h)Y(m-h), \quad m = \overline{0, N-1}, \quad (3.1)$$

то

$$C_z(k) = C_x(k)C_y(k).$$

Доказательство

Вычисляя $Z(m)$, получим

$$C_z(k) = \frac{1}{N} \sum_{m=0}^{N-1} Z(m)W^{km}. \quad (3.2)$$

Подставляя в (3.2) соотношение свертки (3.1), получим

$$C_z(k) = \frac{1}{N^2} \sum_{m=0}^{N-1} \sum_{h=0}^{N-1} x(h)Y(m-h)W^{km} = \frac{1}{N} \sum_{h=0}^{N-1} x(h) \frac{1}{N} \sum_{m=0}^{N-1} Y(m-h)W^{km} .$$

Согласно теореме сдвига, имеем

$$\frac{1}{N} \sum_{m=0}^{N-1} Y(m-h)W^{km} = W^{kh} C_y(k) .$$

Таким образом,

$$C_z(k) = C_y(k) \frac{1}{N} \sum_{h=0}^{N-1} x(h)W^{kh} = C_x(k)C_y(k) .$$

Эта теорема утверждает, что свертка временных последовательностей эквивалентна умножению их коэффициентов, полученных после дискретного преобразования Фурье.

3.2. Теорема корреляции

Если $X(m) \leftrightarrow C_x(k)$ и $Y(m) \leftrightarrow C_y(k)$, а их функция корреляции определяется соотношением

$$\hat{Z}(m) = \frac{1}{N} \sum_{h=0}^{N-1} X(h)Y(m+h) , \text{ где } m = \overline{0, N-1} , \quad (3.3)$$

то $C_z(k) = \overline{C_x(k)} C_y(k)$,

где $\overline{C_x(k)}$ - комплексное сопряжение $C_x(k)$

Доказательство

По определению имеем

$$C_z(k) = \frac{1}{N} \sum_{m=0}^{N-1} \hat{Z}(m)W^{km} . \quad (3.4)$$

Подставляя (3.3) в (3.4) и меняя порядок суммирования, получаем

$$C_z^{\wedge}(k) = \frac{1}{N} \sum_{h=0}^{N-1} X(h) \left\{ \frac{1}{N} \sum_{m=0}^{N-1} Y(m+h) W^{km} \right\}$$

Применяя теорему сдвига, будем иметь

$$C_z^{\wedge}(k) = C_y(k) \left\{ \frac{1}{N} \sum_{h=0}^{N-1} X(h) W^{-kh} \right\}.$$

Так как $C_x(k) = \frac{1}{N} \sum_{h=0}^{N-1} X(h) W^{kh}$, то $\frac{1}{N} \sum_{h=0}^{N-1} X(h) W^{-kh} = \overline{C_x(k)}$.

Таким образом, $C_z^{\wedge}(k) = C_y(k) \overline{C_x(k)}$.

Если последовательности $\{X(m)\}$ и $\{Y(m)\}$ идентичны друг другу, то

$$C_z^{\wedge}(k) = |C_x(k)|^2, \quad k = \overline{0, N-1}.$$

Обратное ДПФ последовательности $C_z^{\wedge}(k)$ есть $\hat{Z}(m) = \sum_{k=0}^{N-1} C_z^{\wedge}(k) W^{-km}$.

Тогда

$$\frac{1}{N} \sum_{h=0}^{N-1} X^2(h) = \sum_{k=0}^{N-1} |C_x(k)|^2, \quad \text{т.е. справедлива теорема Парсеваля.}$$

3.3. Матричное представление корреляции и свертки

Если $\{X(m)\}$ и $\{Y(m)\}$ – две N -периодические последовательности действительных чисел, то операции корреляции и свертки определяются соответственно как

$$\hat{Z}(m) = \frac{1}{N} \sum_{h=0}^{N-1} X(h)Y(m+h),$$

$$Z(m) = \frac{1}{N} \sum_{h=0}^{N-1} X(h)Y(m-h).$$

Если $N=4$, то

$$4\hat{Z}(0) = X(0)Y(0) + X(1)Y(1) + X(2)Y(2) + X(3)Y(3);$$

$$4\hat{Z}(1) = X(0)Y(1) + X(1)Y(2) + X(2)Y(3) + X(3)Y(0);$$

$$4\hat{Z}(2) = X(0)Y(2) + X(1)Y(3) + X(2)Y(0) + X(3)Y(1);$$

$$4\hat{Z}(3) = X(0)Y(3) + X(1)Y(0) + X(2)Y(1) + X(3)Y(2);$$

$$\begin{bmatrix} \hat{Z}(0) \\ \hat{Z}(1) \\ \hat{Z}(2) \\ \hat{Z}(3) \end{bmatrix} = \frac{1}{4} \begin{bmatrix} X(0) & X(1) & X(2) & X(3) \\ X(3) & X(0) & X(1) & X(2) \\ X(2) & X(3) & X(0) & X(1) \\ X(1) & X(2) & X(3) & X(0) \end{bmatrix} \begin{bmatrix} Y(0) \\ Y(1) \\ Y(2) \\ Y(3) \end{bmatrix}$$

При $N=4$ для свертки будем иметь

$$4Z(0) = X(0)Y(0) + X(1)Y(-1) + X(2)Y(-2) + X(3)Y(-3);$$

$$4Z(1) = X(0)Y(1) + X(1)Y(0) + X(2)Y(-1) + X(3)Y(-2);$$

$$4Z(2) = X(0)Y(2) + X(1)Y(1) + X(2)Y(0) + X(3)Y(-1);$$

$$4Z(3) = X(0)Y(3) + X(1)Y(2) + X(2)Y(1) + X(3)Y(0).$$

Так как $[Y(k)]_{\text{mod } N} = Y(k - N)$ при $k > 0$,

$[Y(k)]_{\text{mod } N} = Y(k + N)$ при $k < 0$, то

$$\begin{bmatrix} Z(0) \\ Z(1) \\ Z(2) \\ Z(3) \end{bmatrix} = \frac{1}{4} \begin{bmatrix} X(0) & X(1) & X(2) & X(3) \\ X(1) & X(2) & X(3) & X(0) \\ X(2) & X(3) & X(0) & X(1) \\ X(3) & X(0) & X(1) & X(2) \end{bmatrix} \begin{bmatrix} Y(0) \\ Y(3) \\ Y(2) \\ Y(1) \end{bmatrix}.$$

В общем виде корреляцию двух последовательностей можно записать как

$$\begin{bmatrix} \hat{Z}(0) \\ \hat{Z}(1) \\ \hat{Z}(2) \\ \vdots \\ \hat{Z}(N-2) \\ \hat{Z}(N-1) \end{bmatrix} = \frac{1}{N} \begin{bmatrix} X(0) & X(1) & X(2) & \dots & X(N-1) \\ X(N-1) & X(0) & X(1) & \dots & X(N-2) \\ X(N-2) & X(N-1) & X(0) & \dots & X(N-3) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ X(2) & X(3) & X(4) & \dots & X(1) \\ X(1) & X(2) & X(3) & \dots & X(0) \end{bmatrix} \begin{bmatrix} Y(0) \\ Y(1) \\ Y(2) \\ \vdots \\ Y(N-2) \\ Y(N-1) \end{bmatrix}.$$

В свою очередь, соотношение свертки можно записать в общем виде как

$$\begin{bmatrix} Z(0) \\ Z(1) \\ Z(2) \\ \vdots \\ Z(N-2) \\ Z(N-1) \end{bmatrix} = \frac{1}{N} \begin{bmatrix} X(0) & X(1) & X(2) & \dots & X(N-1) \\ X(1) & X(2) & X(3) & \dots & X(0) \\ X(2) & X(3) & X(4) & \dots & X(1) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ X(N-2) & X(N-1) & X(0) & \dots & X(N-3) \\ X(N-1) & X(0) & X(1) & \dots & X(N-2) \end{bmatrix} \begin{bmatrix} Y(0) \\ Y(N-1) \\ Y(N-2) \\ \vdots \\ Y(2) \\ Y(1) \end{bmatrix}.$$

Если последовательности $\{X(m)\}$ и $\{Y(m)\}$ аналогичны друг другу, то

$$Z(m) = \frac{1}{N} \sum_{h=0}^{N-1} X(h)X(m+h), \text{ где } m = \overline{0, N-1}.$$

Это соотношение определяет автокорреляцию последовательности $\{X(m)\}$.

С использованием БПФ схема вычислений корреляции будет иметь вид рис.3.1.

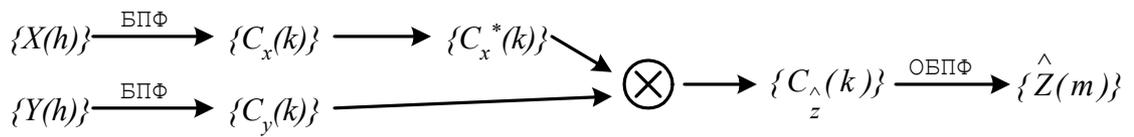


Рис. 3.1. Схема вычисления корреляции

В свою очередь, схему вычисления свертки можно представить как показано на рис.3.2.

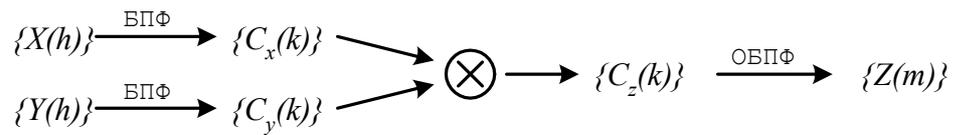


Рис. 3.2. Схема вычисления свертки

Схема вычисления свертки аperiodической последовательности

Рассмотрим случай, когда последовательности, по отношению к которым применяются операции свертки, являются аperiodическими и обозначены:

$$\hat{X}(m), m = 0, 1 \dots M$$

$$\hat{Y}(m), m = 0, 1 \dots P$$

$$M, P < N$$

Свертка последовательностей $\hat{X}(m)$ и $\hat{Y}(m)$ определяется с помощью соотношения:

$$\hat{Z}(m) = \frac{1}{M + P + 1} \sum_{h=0}^{M+P} \hat{X}(h) \hat{Y}(m-h)$$

$$\hat{Y}(l) = 0 \text{ при } l < 0$$

Последовательность $\hat{Z}(m)$ можно получить с помощью БПФ следующим образом:

1. пусть N – наименьшая степень 2, большая, чем $M + P$.
2. образуем последовательности $X(m)$, $Y(m)$ с периодом N , для которых

$$X(m) = \begin{cases} \hat{X}(m), & \text{при } m = 0, 1, \dots, M \\ 0, & \text{при } m = M + 1, M + 2, \dots, N - 1 \end{cases}$$

$$Y(m) = \begin{cases} \hat{Y}(m), & \text{при } m = 0, 1, \dots, P \\ 0, & \text{при } m = P + 1, P + 2, \dots, N - 1 \end{cases}$$

3. применим БПФ для вычисления свертки для последовательностей

$$X(m), Y(m), \text{ где } Z(m) = \frac{1}{N} \sum_{k=0}^{m-1} X(k)Y(m-k)$$

4. получим последовательность $\hat{Z}(m)$: $(M + P + 1) = Z(m)N$, тогда

$$\hat{Z}(m) = \frac{N}{M + P + 1} Z(m)$$

4. ТЕХНИКА БЫСТРОГО ПРЕОБРАЗОВАНИЯ ФУРЬЕ (БПФ) И ОСНОВНЫЕ ПОДХОДЫ К ЕГО РЕАЛИЗАЦИИ

Быстрое преобразование Фурье (FFT) является не более, чем алгоритмом для ускоренного вычисления ДПФ путем сокращения требуемого числа операций умножения и сложения. Данное преобразование было предложено в 1960-ых годах. Алгоритм быстрого преобразования Фурье значительно сокращает количество арифметических операций и объем памяти, необходимой для вычисления ДПФ. ДПФ может быть сильно упрощено, если использовать свойства симметрии и периодичности коэффициентов поворота.

При вычислении N-точечного ДПФ требуется N^2 вычислений с комплексными числами, а при реализации N-точечного БПФ $(N/2)\log_2(N)$ вычислений с комплексными числами. Вычислительная эффективность БПФ по сравнению с ДПФ становится весьма существенной, когда количество точек БПФ увеличивается до нескольких тысяч (табл. 4.1).

Таблица 4.1

Эффективность БПФ

N	Умножений при ДПФ	Умножений при БПФ	Эффективность БПФ
256	65 536	1 024	64 : 1
512	262 144	2 304	114 : 1
1 024	1 048 576	5 120	205 : 1
2 048	4 194 304	11 264	372 : 1
4 096	16 777 216	24 576	683 : 1

Если необходимо рассчитать только несколько точек спектра, ДПФ может быть более эффективным. Вычисление одного выходного отсчета

спектра с использованием ДПФ требует только N умножений с комплексными числами.

Мы будем предполагать далее, что $N=2^n$. При этом общность не теряется, так как N выбирается достаточно большим для того, чтобы удовлетворять теореме дискретизации Котельникова, т.е.

$$N \geq 2BT,$$

где B – полоса частот сигнала $x(t)$; T – его длительность.

Теорема Котельникова-Найквиста-Шеннона: если сигнал таков, что его спектр ограничен частотой F , то после дискретизации сигнала с частотой не менее $2F$ можно восстановить непрерывный сигнал по полученному цифровому сигналу абсолютно точно. Для этого нужно проинтерполировать цифровой сигнал «между отсчетами» специальными функциями.

Рассмотрим случай вещественно-значной последовательности $\{X(m)\}$ при $N=8$. Из свойства комплексной сопряженности ДПФ следует, что

$$C_x(4+l) = C_x^*(4-l); \quad l = \overline{1, N/2-1}.$$

Тогда

$$C_x(k) = \frac{1}{8} \sum_{m=0}^7 X(m) W^{km}; \quad k = \overline{0, N-1};$$

$$W = e^{-i2\pi/8} = e^{-i\pi/4};$$

$$8C_x(k) = \sum_{m=0}^7 X(m) \cos\left(\frac{mk\pi}{4}\right) - i \sum_{m=0}^7 X(m) \sin\left(\frac{mk\pi}{4}\right). \quad (4.1)$$

Используя свойство N -периодичности экспонент, для $N=8$ матрица будет иметь вид

$$F = \begin{bmatrix} W_0 & W_0 \\ W_0 & W_1 & W_2 & W_3 & W_4 & W_5 & W_6 & W_7 \\ W_0 & W_2 & W_4 & W_6 & W_0 & W_2 & W_4 & W_6 \\ W_0 & W_3 & W_6 & W_1 & W_4 & W_7 & W_2 & W_5 \\ W_0 & W_4 & W_0 & W_4 & W_0 & W_4 & W_0 & W_4 \\ W_0 & W_5 & W_2 & W_7 & W_4 & W_1 & W_6 & W_3 \\ W_0 & W_6 & W_4 & W_2 & W_0 & W_6 & W_4 & W_2 \\ W_0 & W_7 & W_6 & W_5 & W_4 & W_3 & W_2 & W_1 \end{bmatrix}.$$

Из свойства симметрии экспоненциальных функций следует, что

$$W_{k+N/2} = -W_k, \text{ где } k = 0, N/2 - 1.$$

То есть $W_4 = -W_0$;

$$W_5 = -W_1;$$

$$W_6 = -W_2;$$

$$W_7 = -W_3.$$

Тогда матрица F будет иметь вид

$$F = \begin{bmatrix} W_0 & W_0 \\ W_0 & W_1 & W_2 & W_3 & -W_0 & -W_1 & -W_2 & -W_3 \\ W_0 & W_2 & -W_0 & -W_2 & W_0 & W_2 & -W_0 & -W_2 \\ W_0 & W_3 & -W_2 & W_1 & -W_0 & -W_3 & W_2 & -W_1 \\ W_0 & -W_0 & W_0 & -W_0 & W_0 & -W_0 & W_0 & -W_0 \\ W_0 & -W_1 & W_2 & -W_3 & -W_0 & W_1 & -W_2 & W_3 \\ W_0 & -W_2 & -W_0 & W_2 & W_0 & -W_2 & -W_0 & W_2 \\ W_0 & -W_3 & -W_2 & -W_1 & -W_0 & W_3 & W_2 & W_1 \end{bmatrix}.$$

Используя двоичную инверсию (перестановку) строк,

$(0,1,2,3,4,5,6,7) \rightarrow (0,4,2,6,1,5,3,7)$ будем иметь

$$\begin{aligned}
F &= \begin{bmatrix} W_0 & W_0 \\ W_0 & -W_0 & W_0 & -W_0 & W_0 & -W_0 & W_0 & -W_0 \\ W_0 & W_2 & -W_0 & -W_2 & W_0 & W_2 & -W_0 & -W_2 \\ W_0 & -W_2 & -W_0 & W_2 & W_0 & -W_2 & -W_0 & W_2 \\ W_0 & W_1 & W_2 & W_3 & -W_0 & -W_1 & -W_2 & -W_3 \\ W_0 & -W_1 & W_2 & -W_3 & -W_0 & W_1 & -W_2 & W_3 \\ W_0 & W_3 & -W_2 & W_1 & -W_0 & -W_3 & W_2 & -W_1 \\ W_0 & -W_3 & -W_2 & -W_1 & -W_0 & W_3 & W_2 & W_1 \end{bmatrix} \cdot \begin{bmatrix} X_0 \\ X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \\ X_6 \\ X_7 \end{bmatrix} = \begin{bmatrix} A_1^l & A_1^l \\ B_1^l & -B_1^l \end{bmatrix} \bar{X} = \\
&= \begin{matrix} A_1^l \left\{ \begin{matrix} X_0 + X_4 \\ X_1 + X_5 \\ X_2 + X_6 \\ X_3 + X_7 \end{matrix} \right\} \\ B_1^l \left\{ \begin{matrix} X_0 - X_4 \\ X_1 - X_5 \\ X_2 - X_6 \\ X_3 - X_7 \end{matrix} \right\} \end{matrix} .
\end{aligned} \tag{4.2}$$

В свою очередь, матрицы A_l^l и B_l^l можно представить в виде, где верхний индекс представляет собой номер шага процедуры БПФ

$$\begin{aligned}
A_1^l &= \begin{vmatrix} A_1^2 & A_1^2 \\ A_2^2 & -A_2^2 \end{vmatrix}; \\
B_1^l &= \begin{vmatrix} B_1^2 & W_2 B_1^2 \\ B_2^2 & -W_2 B_2^2 \end{vmatrix}.
\end{aligned} \tag{4.3}$$

Подставляя выражения для A_l^l и B_l^l в (4.2) получим

$$\begin{aligned}
&A_1^2 \left\{ \begin{matrix} (X_0 + X_4) + (X_2 + X_6) \\ (X_1 + X_5) + (X_3 + X_7) \end{matrix} \right\}; \\
&A_2^2 \left\{ \begin{matrix} (X_0 + X_4) - (X_2 + X_6) \\ (X_1 + X_5) - (X_3 + X_7) \end{matrix} \right\}; \\
&B_1^2 \left\{ \begin{matrix} (X_0 - X_4) + W_2(X_2 - X_6) \\ (X_1 - X_5) + W_2(X_3 - X_7) \end{matrix} \right\};
\end{aligned}$$

$$B_2^2 \left\{ \begin{array}{l} (X_0 - X_4) - W_2(X_2 - X_6) \\ (X_1 - X_5) - W_2(X_3 - X_7) \end{array} \right\}. \quad (4.4)$$

Наконец, на последнем шаге получим

$$\begin{aligned} & [(X_0 + X_4) + (X_2 + X_6)] + [(X_1 + X_5) + (X_3 + X_7)] \\ & [(X_0 + X_4) + (X_2 + X_6)] - [(X_1 + X_5) + (X_3 + X_7)] \\ & [(X_0 + X_4) - (X_2 + X_6)] + W_2[(X_1 + X_5) - (X_3 + X_7)] \\ & [(X_0 + X_4) - (X_2 + X_6)] - W_2[(X_1 + X_5) - (X_3 + X_7)] \\ & [(X_0 - X_4) + W_2(X_2 - X_6)] + W_1[(X_1 - X_5) + W_2(X_3 - X_7)] \\ & [(X_0 - X_4) + W_2(X_2 - X_6)] - W_1[(X_1 - X_5) + W_2(X_3 - X_7)] \\ & [(X_0 - X_4) - W_2(X_2 - X_6)] + W_3[(X_1 - X_5) - W_2(X_3 - X_7)] \\ & [(X_0 - X_4) - W_2(X_2 - X_6)] - W_3[(X_1 - X_5) - W_2(X_3 - X_7)] \end{aligned} \quad (4.5)$$

Описанный алгоритм удобно представить графически (рис. 4.1).

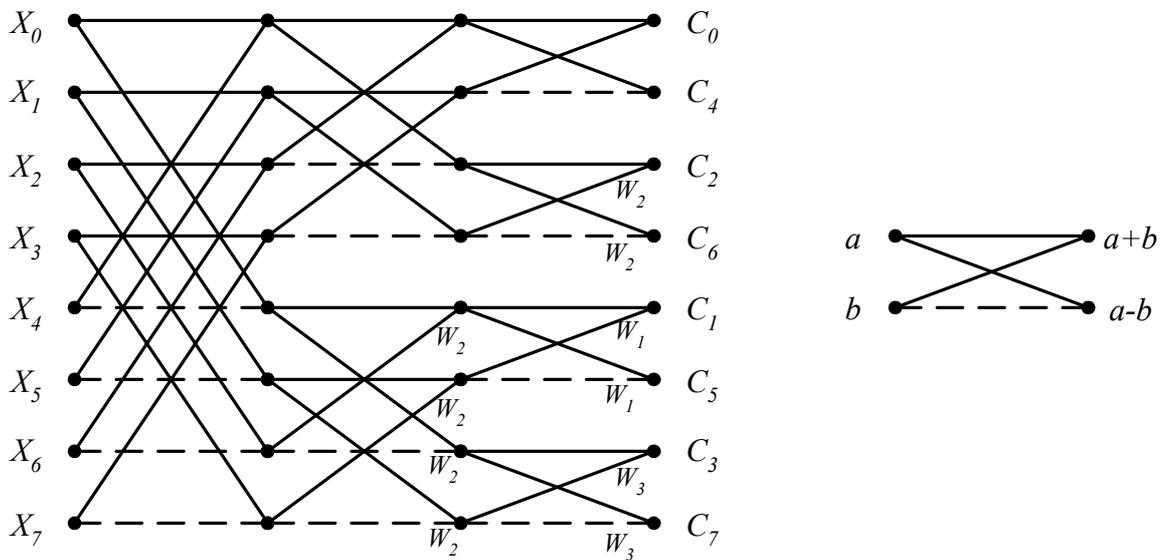


Рис. 4.1. Граф-схема быстрой процедуры вычисления коэффициентов преобразования Фурье

Для определения степеней W на каждом шаге необходимо выразить последовательность $l=0, 1, 2, \dots, N/2-1$ в виде $(n-1)$ – разрядных двоичных последовательностей. В результате для $N=16$, к примеру, получим множество

$$S_1=(000,001,010,011,100,101,110,111).$$

Для получения S_2 необходимо выполнить двоичную инверсию каждой $(n-1)$ -разрядной последовательности множества S_1 , т.е.

$$S_2=(000,100,010,110,001,101,011,111),$$

и записать двоичную последовательность в виде десятичных чисел

$$S_3=(0,4,2,6,1,5,3,7),$$

и таким образом имеем $W_0, W_4, W_2, W_6, W_1, W_5, W_3, W_7$. (табл. 4.2).

Итерация r для БПФ состоит из 2^{r-1} групп, где $r = \overline{1, n}$ ($N=2^n$). Для $N=16$, $r = \overline{1, 4}$.

Таблица 4.2

Значения степени W

Номер итерации	Степени W ($N=16$)
1	$W_0 W_0 W_0$
2	$W_0 W_0 W_0 W_0 W_0 W_0 W_0 W_0 W_4 W_4 W_4 W_4 W_4 W_4 W_4 W_4$
3	$W_0 W_0 W_0 W_0 W_4 W_4 W_4 W_4 W_2 W_2 W_2 W_2 W_6 W_6 W_6 W_6$
4	$W_0 W_0 W_4 W_4 W_2 W_2 W_6 W_6 W_1 W_1 W_5 W_5 W_3 W_3 W_7 W_7$

Процедура, обеспечивающая инверсию, может быть описана в матричном виде следующим образом.

Первый элемент первой строки таблицы равен нулю. Последующие первые элементы каждой из строк определяются как $n_s=N/2^s$, где $s = \overline{1, n}$,

$N=2^n$. Каждая k строка таблицы получается прибавлением элемента n_{k-1} к каждому элементу предыдущих строк. Тогда таблица будет иметь вид

0				
n_1				
n_2	(n_1+n_2)			
n_3	(n_1+n_3)	(n_2+n_3)	$(n_1+n_2+n_3)$	
n_4	(n_1+n_4)	(n_2+n_4)	$(n_1+n_2+n_4)$...
.				
.				
n_k	...			

Требуемая последовательность L_n , соответствующая двоичной инверсии, определяется как $L_n=(0, n_1, n_2, (n_1+n_2), n_3, (n_1+n_3), \dots, n_k, \dots)$. В качестве примера рассмотрим случай для $N=16$. Тогда $n_1=8, n_2=4, n_3=2, n_4=1$, т.е. таблица будет иметь вид

0							
8							
4	12						
2	10	6	14				
1	9	5	13	3	11	7	15

$L_n=(0,8,4,12,2,10,6,14,1,9,5,13,3,11,7,15)$.

Для обработки исходных данных (которые предполагаются комплексными) с помощью алгоритма БПФ требуется $2N$ ячеек оперативной памяти. Поэтому выходной массив может храниться в тех же ячейках памяти, что и исходный массив. Процедура перестановки данных может потребовать дополнительно $2N$ ячеек памяти. Таким образом, для алгоритма БПФ необходимо примерно $4N$ ячеек. В противоположность этому прямой метод

требует приблизительно $2N^2$ ячеек памяти, т.к. необходимо запомнить N^2 значений степеней W .

В общем виде матрицу преобразования Фурье в факторизованной форме можно представить как

$$F_N = \prod_{i=1}^n F_i D_i. \quad (4.6)$$

Для $N=8$ $F_8 = F_1 D_1 \cdot F_2 D_2 \cdot F_3 D_3$, где $F_1 = \begin{bmatrix} I_4 & I_4 \\ I_4 & -I_4 \end{bmatrix}$; I_4 – единичная матрица

размерностью 4×4 ; D_1 – диагональная матрица с элементами W_0 ;

$$F_2 = \begin{bmatrix} I_2 & I_2 & & \\ I_2 & -I_2 & & \\ & & I_2 & I_2 \\ & & I_2 & -I_2 \end{bmatrix}.$$

D_2 -диагональная матрица с элементами W_0, W_2 :

$$D_2 = \begin{bmatrix} W_0 & & & & & & & \\ & W_0 & & & & & & \\ & & W_0 & & & & & \\ & & & W_0 & & & & \\ & & & & W_0 & & & \\ & & & & & W_0 & & \\ & & & & & & W_2 & \\ & & & & & & & W_2 \end{bmatrix} ;$$

$$F_3 = \begin{bmatrix} I_1 & I_1 & & & & & & \\ I_1 & -I_1 & & & & & & \\ & & I_1 & I_1 & & & & \\ & & I_1 & -I_1 & & & & \\ & & & & I_1 & I_1 & & \\ & & & & I_1 & -I_1 & & \\ & & & & & & I_1 & I_1 \\ & & & & & & I_1 & -I_1 \end{bmatrix} ;$$

$$D_3 = \begin{bmatrix} W_0 & & & & & & & \\ & W_0 & & & & & & \\ & & W_0 & & & & & \\ & & & W_2 & & & & \\ & & & & W_0 & & & \\ & & & & & W_1 & & \\ & & & & & & W_0 & \\ & & & & & & & W_3 \end{bmatrix} .$$

Факторизованная форма – это такая форма, когда в каждой строке матрицы, являющейся множителем, присутствует не более 2 значащих элементов, а остальные равны нулю.

Алгоритм быстрого преобразования Фурье можно выразить следующим образом. Общая схема алгоритма состоит в повторяющемся сведении ДПФ вектора длины N к векторам длины N/2 и объединении результатов. Базисом рекурсии служат векторы длины 1, для которых ДПФ – сам вектор.

АЛГОРИТМ БПФ(a, N)

{

1. Если длина вектора равна 1, вернуть a .
2. Разбить вектор a на четную часть $a^{чет} = (a_0, a_2, \dots, a_{N-2})$ и нечетную $a^{нечет} = (a_1, a_3, \dots, a_{N-1})$.
3. Рекурсивно вызвать БПФ на каждой из частей
 $b^{чет} = \text{БПФ}(a^{чет})$
 $b^{нечет} = \text{БПФ}(a^{нечет})$
4. Объединение результатов.

- a. (инициализация) Присвоить ω_N значение главного корня из единицы

$$\omega_N = e^{\frac{2\pi}{N}} = \cos \frac{2\pi}{N} + i \sin \frac{2\pi}{N}$$

- b. (инициализация) Присвоить $\omega = 1$

- c. В цикле вычислить левую и правую часть одновременно:

for($j=0; j < N/2; j++$)

{

$$y_j = b_j^{чет} + \omega b_j^{нечет}$$

$$y_{j+N/2} = b_j^{чет} - \omega b_j^{нечет}$$

$$\omega = \omega * \omega_N$$

}

5. Вернуть вектор y .

}

Ниже изображено дерево рекурсий, рис. 4.2. Каждый уровень, начиная снизу, соответствует проходу алгоритма по всему вектору и объединению сначала одиночных элементов в пары, затем пар в четверки и так далее до конца. Обратите внимание на то, что порядок индексов на верхнем уровне не соответствует нижнему. Это естественно, если учесть, что нечетные индексы после бабочки идут в правую половину вектора, а четные – в левую.

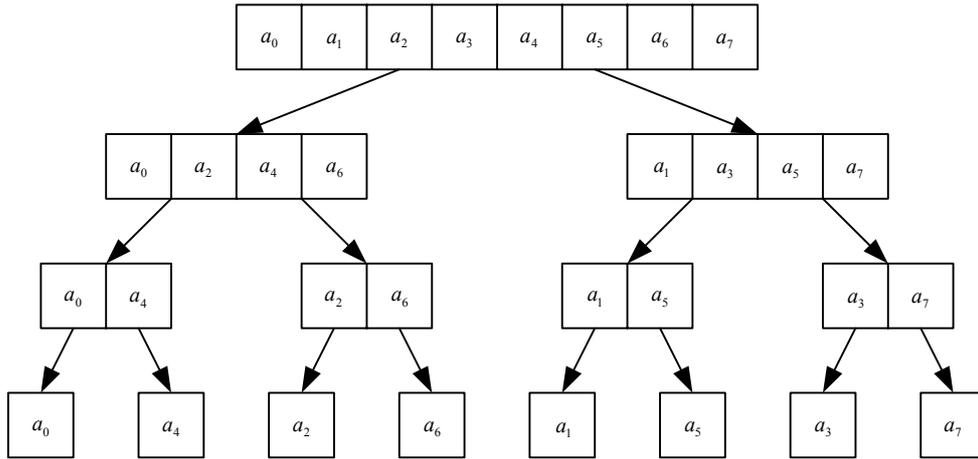


Рис. 4.2. Дерево рекурсий для 8 элементов

Чтобы техника работала на вход алгоритма должен поступать вектор нижнего уровня $(a_0, a_4, a_2, a_6, a_1, a_5, a_3, a_7)$ – тогда объединения произойдут в правильном порядке, как изображено выше, и дадут в результате вектор (a_0, a_1, \dots, a_7) . Таким образом, перед вызовом процедуры следует провести перестановку элементов массива.

5. АЛГОРИТМЫ ФОРМИРОВАНИЯ ОРТОГОНАЛЬНЫХ ПРЯМОУГОЛЬНЫХ ФУНКЦИЙ

Особый класс систем ортогональных функций составляют системы кусочно-постоянных функций, таких как функции Уолша, Адамара и Хаара. Эти системы имеют большое практическое значение, особенно для цифровых систем, поскольку они характеризуются высокоэффективными алгоритмами быстрых преобразований.

В 1923 году американский ученый Уолш получил полную систему ортонормированных функций, которая дополняет систему функций Радемахера. Множество функций Уолша обычно разделяется на три группы, отличающиеся порядком расположения в системе.

Общеприняты следующие упорядочения (способы определения функций Уолша):

- упорядочение по частоте (по Уолшу с помощью функций Радемахера);
- упорядочение по Пэли (диадическое);
- упорядочение по Адамару,

где под частотой функции понимается число пересечений нулевого уровня в единицу времени.

Рассмотрим способ, основанный на взаимосвязи функций Уолша с функциями Радемахера и способ, основанный на матрицах Адамара.

Функции Радемахера получаются из синусоидальных функций с помощью соотношения

$$r_k(t) = \text{sign}[\sin(2^k \pi t)], \quad 0 \leq t < 1, \quad (5.1)$$

где аргумент t – безразмерное время, т.е. время, нормированное к произвольному интервалу T_0 , а целое положительное число k – порядок функции. Символом sign (сигнум-функция) обозначается функция:

$$\operatorname{sign} x = \begin{cases} 1 & \text{при } x > 0, \\ -1 & \text{при } x \leq 0. \end{cases} \quad (5.2)$$

Функции Радемахера принимают одно из двух значений ± 1 и имеют вид меандра.

Функции Радемахера ортонормированны с единичной весовой функцией на интервале $0 \leq t < 1$, т.к. для любых двух функций $r_m(t)$, $r_n(t)$ имеют место соотношения:

$$\int_0^1 r_m(t)r_n(t)dt = \begin{cases} 1 & \text{при } m = n, \\ 0 & \text{при } m \neq n. \end{cases} \quad (5.3)$$

Все функции Радемахера являются нечетными относительно середины интервала определения и не могут быть использованы для аппроксимации сигналов, четных относительно этой точки. Иными словами, система функций Радемахера – неполная. На рис. 5.1 приведены функции Радемахера для $N = 8$.

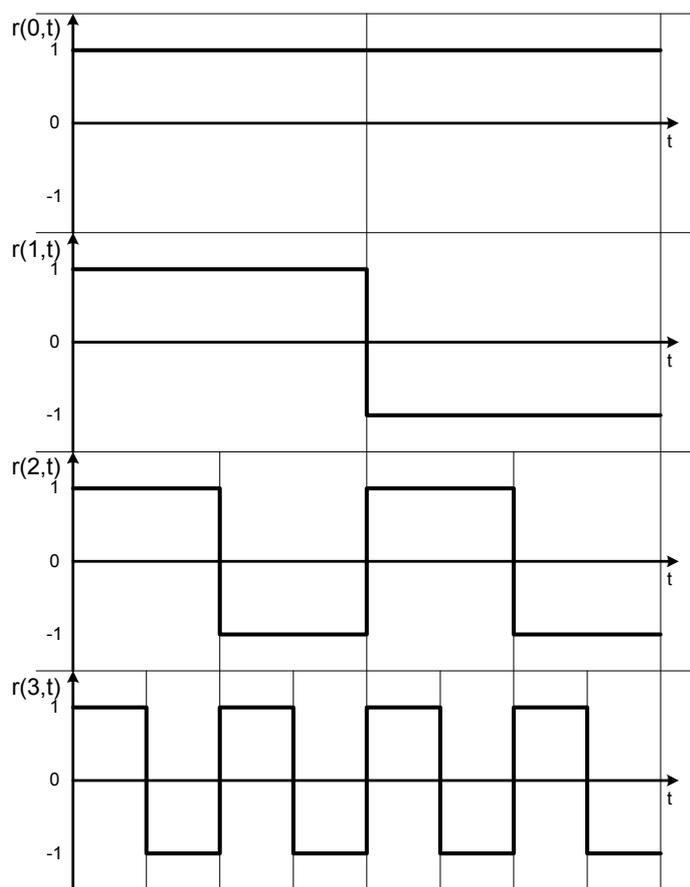


Рис. 5.1 Функции Радемахера для $N=8$

Обозначив для краткости $r(m,t)=r_m(t)$, для $N=8$ будем иметь:

Если $r_1(t) = \quad ++++----$

$r_2(t) = \quad ++--++--$

$r_3(t) = \quad +-+-+--+ ,$

где “+” соответствует +1, а “-” соответствует -1.

6. ПРЕДСТАВЛЕНИЕ СИГНАЛОВ В ВИДЕ ФУНКЦИЙ УОЛША. АЛГОРИТМ БЫСТРОГО ПРЕОБРАЗОВАНИЯ УОЛША

Функции Уолша, образующие полную ортонормированную систему, можно сформировать, образуя произведения соответствующих функций Радемахера. Первые восемь функций Уолша представлены на рис. 6.1.

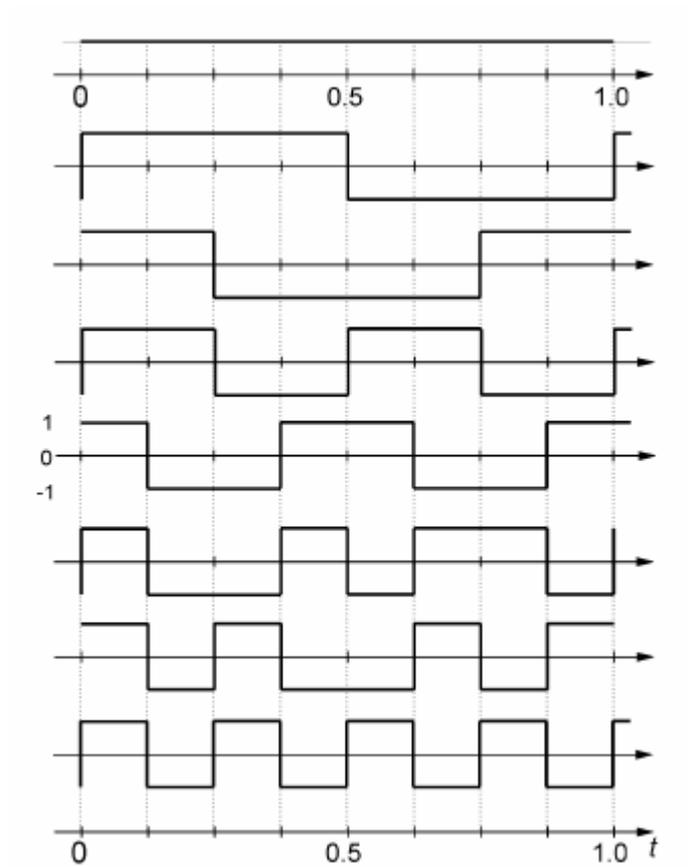


Рис. 6.1. Первые восемь функций Уолша

В свою очередь функции Уолша можно представить следующим образом:

$$\begin{aligned}
 wal(0,t) &= && ++++++++ \\
 wal(1,t) &= r_1 && ++++---- \\
 wal(2,t) &= r_1 r_2 && ++-----++ \\
 wal(3,t) &= r_2 && +- - - + + - - \\
 wal(4,t) &= r_2 r_3 && + - - + + - - +
 \end{aligned}$$

$$\begin{aligned}
wal(5,t) &= r_1 r_2 r_3 = +---+--+-- \\
wal(6,t) &= r_1 r_3 = +-+---+--+ \\
wal(7,t) &= r_3 = +-+---+--+
\end{aligned}$$

Сопоставление этих функций с функциями Радемахера позволяет составить очевидные соотношения, согласно которым каждая функция Уолша $wal(n,t)$ с номером n , входящая в систему из $N = 2^r$ функций, является произведением степеней первых n функций Радемахера.

Принцип нахождения показателей этих степеней определяется двоичным представлением номера функции n . Примем следующие обозначения для разрядов, составляющих двоичное представление числа n : n_1 - первый разряд, n_2 - второй разряд, и так далее до n_r , то есть r -го разряда двоичного представления натурального числа n . При такой нумерации n_1 оказывается старшим разрядом числа n , а n_r - младшим. n_i может принимать одно из двух значений – нуль или единица. Будем считать, что $n_0 = 0$ по определению. Используя символ \oplus для обозначения операции поразрядного сложения по модулю 2, способ построения функций Уолша можно выразить аналитически для любого $N = 2^r$ в виде следующего соотношения:

$$wal(n,t) = \prod_{k=1}^r [r_k(t)]^{n_{r-k+1} \oplus n_{r-k}} \quad (6.1)$$

Поясним применение данной формулы на примере шестой функции Уолша ($n=6$), входящей в систему размером $N = 2^3 = 8$. Произведение состоит из трех множителей вида:

$$\begin{aligned}
&\text{при } k=1 [r_1(t)]^{n_3 \oplus n_2}, \\
&\text{при } k=2 [r_2(t)]^{n_2 \oplus n_1}, \\
&\text{при } k=3 [r_3(t)]^{n_1 \oplus n_0}.
\end{aligned}$$

На основе двоичного представления числа $n=6$ несложно установить, что $n_1 = 1, n_2 = 1, n_3 = 0$.

Таким образом, $n_3 \oplus n_2 = 0 \oplus 1 = 1, n_2 \oplus n_1 = 1 \oplus 1 = 0, n_1 \oplus n_0 = 1 \oplus 0 = 1$ и по формуле:

$$wal(6,t) = r_1(t)r_2^0(t)r_3(t) = r_1(t)r_3(t).$$

Функции Радемахера перемножаются при использовании кода Грея. В некоторых практических приложениях, например в аналого-цифровых преобразованиях, желательно использовать коды, у которых все следующие друг за другом кодовые слова различаются только одной цифрой в некотором разряде. Коды, обладающие таким свойством, называются циклическими.

Очень важным циклическим кодом является код Грея. Двоичное представление числа может быть легко преобразовано в код Грея с помощью полусумматоров.

Пусть $g_{n-1}g_{n-2}\dots g_2g_1g_0$ – кодовое слово в n -разрядном двоичном коде Грея, соответствующее двоичному числу $b_{n-1}b_{n-2}\dots b_2b_1b_0$. Тогда g_i может быть получена как

$$g_i = b_i \oplus b_{i+1}, 0 \leq i \leq n-2;$$

$$g_{n-1} = b_{n-1},$$

где \oplus означает сложение по модулю два, которое определяется как

$$0 \oplus 0 = 0$$

$$1 \oplus 0 = 1$$

$$0 \oplus 1 = 1$$

$$1 \oplus 1 = 0$$

Например, код Грея, соответствующий двоичному числу 101101, может быть образован как на рис. 6.2. Трехразрядный код Грея показан в табл. 6.1.

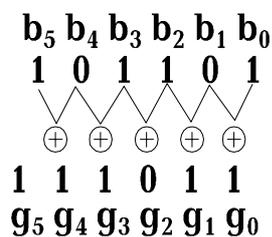


Рис. 6.2. Преобразование двоичного кода в код Грея

Таблица 6.1

Трехразрядный код Грея

Десятичное число	Код Грея			Двоичный код		
	g_2	g_1	g_0	b_2	b_1	b_0
0	0	0	0	0	0	0
1	0	0	1	0	0	1
2	0	1	1	0	1	0
3	0	1	0	0	1	1
4	1	1	0	1	0	0
5	1	1	1	1	0	1
6	1	0	1	1	1	0
7	1	0	0	1	1	1

Преобразование кода Грея в двоичный код начинается с цифры самого левого разряда и движения вправо, принимая $b_i = g_i$, если число единиц, предшествующих g_i , четно и $b_i = \bar{g}_i$ (черта обозначает инвертирование), если число единиц, предшествующих g_i , нечетно. При этом нулевое число единиц считается четным. Пример двоичного числа, соответствующее коду Грея 1001011, показан на рис. 6.3.

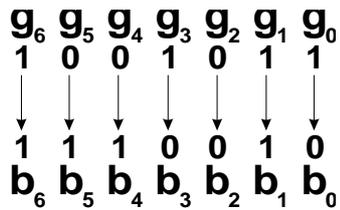


Рис. 6.3. Преобразование кода Грея в двоичный код

Приведем некоторые *свойства функций Уолша*.

1. Функции Уолша *ортонормированны* на интервале $0 \leq t \leq 1$:

$$\int_0^1 wal(k,t)wal(i,t)dt = \begin{cases} 1 & \text{при } k = i, \\ 0 & \text{при } k \neq i. \end{cases}$$

2. Функции Уолша обладают свойством *мультипликативности*, т.е. перемножение двух функций Уолша дает другую функцию Уолша, причем верно соотношение:

$$wal(k,t)wal(i,t) = wal(k \oplus i,t).$$

3. Функции Уолша обладают свойством *симметрии*, проявляющимся в том, что все выводы относительно i справедливы также и относительно t . Например, свойство мультипликативности с учетом свойства симметрии запишется в виде

$$wal(i,t_1)wal(i,t_2) = wal(i,t_1 \oplus t_2).$$

4. Умножение любой функции самой на себя дает функцию нулевого порядка $wal(0,t)$, так как в результате получаются только произведения вида $(+1)(+1)$ и $(-1)(-1)$. Таким образом,

$$wal(i,t)wal(i,t) = wal(0,t) .$$

5. Очевидно также, что умножение $wal(i,t)$ на $wal(0,t)$ не изменяет функцию $wal(i,t)$.

Способ нумерации функций в системе называется упорядочением. Функции Уолша, сформированные в соответствии с выражением (6.1), упорядочены по Уолшу.

В ряде практических задач целесообразно пользоваться иными способами упорядочения. Часто применяются функции Уолша, упорядоченные по Адамару [$had(h,t)$] и по Пэли [$pal(p,t)$].

Независимо от упорядочения функции Уолша, составляющие систему из $N = 2^n$ функций, всегда можно представить в виде произведения степеней первых n функций Радемахера. Принцип же нахождения показателей этих степеней индивидуален для каждого упорядочения.

Остановимся на упорядочении по Адамару. При $N=2^n$ матрица Адамара может быть получена с помощью соотношения

$$H(n) = \begin{bmatrix} H(n-1) & H(n-1) \\ H(n-1) & -H(n-1) \end{bmatrix};$$

$$H(0) = 1.$$

Матрица Адамара также может быть получена из ядра $H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ с

помощью кронекеровского произведения, т.е.

$$H(2) = H_1 \otimes H_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

Далее (табл. 6.2, 6.3) приводятся нумерация функций Уолша в базисе из 16 функций при различных способах упорядочения и нумерация для базиса из 8 функций. Сравнение таблиц показывает, что нумерация одних и тех же функций в упорядочении Адамара меняется в зависимости от размерности базиса.

Таблица 6.2

Нумерация функций Уолша
при различных способах
упорядочения, N=16

wal(n,t)	had(h,t)	pal(p,t)
0	0	0
1	8	1
2	12	3
3	4	2
4	6	6
5	14	7
6	10	5
7	2	4
8	3	12
9	11	13
10	15	15
11	7	14
12	5	10
13	13	11
14	9	9
15	1	8

Таблица 6.3

Нумерация функций Уолша
при различных способах
упорядочения, N=8

wal(n,t)	had(h,t)	pal(p,t)
0	0	0
1	4	1
2	6	3
3	2	2
4	3	6
5	7	7
6	5	5
7	1	4

Быстрое преобразование Уолша можно получить с помощью технологии разбиения матриц. Рассмотрим вывод алгоритма для $N=8$. При $N=8$ в матричном виде преобразование Уолша можно записать как

$$C_x(3) = \frac{1}{8} H(3) X(3).$$

Используя соотношение $H(k) = \begin{bmatrix} H(k-1) & H(k-1) \\ H(k-1) & -H(k-1) \end{bmatrix}$, $H(3)$ можно выразить через $H(2)$, что приводит к

$$\begin{bmatrix} C_x(0) \\ C_x(1) \\ C_x(2) \\ C_x(3) \\ C_x(4) \\ C_x(5) \\ C_x(6) \\ C_x(7) \end{bmatrix} = \frac{1}{8} \begin{bmatrix} H(2) & H(2) \\ H(2) & -H(2) \end{bmatrix} \begin{bmatrix} x(0) \\ x(1) \\ x(2) \\ x(3) \\ x(4) \\ x(5) \\ x(6) \\ x(7) \end{bmatrix}.$$

Из разбиения матрицы следует, что

$$\begin{bmatrix} C_x(0) \\ C_x(1) \\ C_x(2) \\ C_x(3) \end{bmatrix} = \frac{1}{8} H(2) \begin{bmatrix} x_1(0) \\ x_1(1) \\ x_1(2) \\ x_1(3) \end{bmatrix} \quad \text{и} \quad \begin{bmatrix} C_x(4) \\ C_x(5) \\ C_x(6) \\ C_x(7) \end{bmatrix} = \frac{1}{8} H(2) \begin{bmatrix} x_1(4) \\ x_1(5) \\ x_1(6) \\ x_1(7) \end{bmatrix},$$

где $x_1(l) = x(l) + x(4+l); l = \overline{0,3}$;

$x_1(l) = x(l-4) - x(l); l = \overline{4,7}$.

Подставляя вместо $H_2 = \begin{bmatrix} H(1) & H(1) \\ H(1) & -H(1) \end{bmatrix}$, будем иметь

$$\begin{bmatrix} C_x(0) \\ C_x(1) \\ C_x(2) \\ C_x(3) \end{bmatrix} = \frac{1}{8} \begin{bmatrix} H(1) & H(1) \\ H(1) & -H(1) \end{bmatrix} \cdot \begin{bmatrix} x_1(0) \\ x_1(1) \\ x_1(2) \\ x_1(3) \end{bmatrix} \quad \text{и}$$

$$\begin{bmatrix} C_x(4) \\ C_x(5) \\ C_x(6) \\ C_x(7) \end{bmatrix} = \frac{1}{8} \begin{bmatrix} H(1) & H(1) \\ H(1) & -H(1) \end{bmatrix} \cdot \begin{bmatrix} x_1(4) \\ x_1(5) \\ x_1(6) \\ x_1(7) \end{bmatrix}.$$

Тогда получим

$$\begin{bmatrix} C_x(0) \\ C_x(1) \end{bmatrix} = \frac{1}{8} H(1) \begin{bmatrix} x_1(0) + x_1(2) \\ x_1(1) + x_1(3) \end{bmatrix} = \frac{1}{8} H(1) \begin{bmatrix} x_2(0) \\ x_2(1) \end{bmatrix};$$

$$\begin{bmatrix} C_x(2) \\ C_x(3) \end{bmatrix} = \frac{1}{8} H(1) \begin{bmatrix} x_1(0) - x_1(2) \\ x_1(1) - x_1(3) \end{bmatrix} = \frac{1}{8} H(1) \begin{bmatrix} x_2(2) \\ x_2(3) \end{bmatrix};$$

$$\begin{bmatrix} C_x(4) \\ C_x(5) \end{bmatrix} = \frac{1}{8} H(1) \begin{bmatrix} x_1(4) + x_1(6) \\ x_1(5) + x_1(7) \end{bmatrix} = \frac{1}{8} H(1) \begin{bmatrix} x_2(4) \\ x_2(5) \end{bmatrix};$$

$$\begin{bmatrix} C_x(6) \\ C_x(7) \end{bmatrix} = \frac{1}{8} H(1) \begin{bmatrix} x_1(4) - x_1(6) \\ x_1(5) - x_1(7) \end{bmatrix} = \frac{1}{8} H(1) \begin{bmatrix} x_2(6) \\ x_2(7) \end{bmatrix}.$$

Так как $H(1) = \begin{bmatrix} H(0) & H(0) \\ H(0) & -H(0) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, то окончательно получим

$$8C_x(0) = x_2(0) + x_2(1) = x_3(0);$$

$$8C_x(1) = x_2(0) - x_2(1) = x_3(1);$$

$$8C_x(2) = x_2(2) + x_2(3) = x_3(2);$$

$$8C_x(3) = x_2(2) - x_2(3) = x_3(3);$$

$$8C_x(4) = x_2(4) + x_2(5) = x_3(4);$$

$$8C_x(5) = x_2(4) - x_2(5) = x_3(5);$$

$$8C_x(6) = x_2(6) + x_2(7) = x_3(6);$$

$$8C_x(7) = x_2(6) - x_2(7) = x_3(7).$$

Для $N=2^n$:

1. Общее число итераций равно $n=\log_2N$. Индекс r принимает значения $r=1,2,\dots, n$.

2. В r итерации участвует 2^{r-1} групп по $N/2^{r-1}$ элементов. Половина элементов в каждой группе связана с операцией сложения, а другая половина – с операцией вычитания.

3. Общее число арифметических операций, необходимое для вычисления всех коэффициентов преобразования, равняется приблизительно $N\log_2N$.

На рис. 6.4 приведена граф-схема процедуры быстрого алгоритма Уолша.

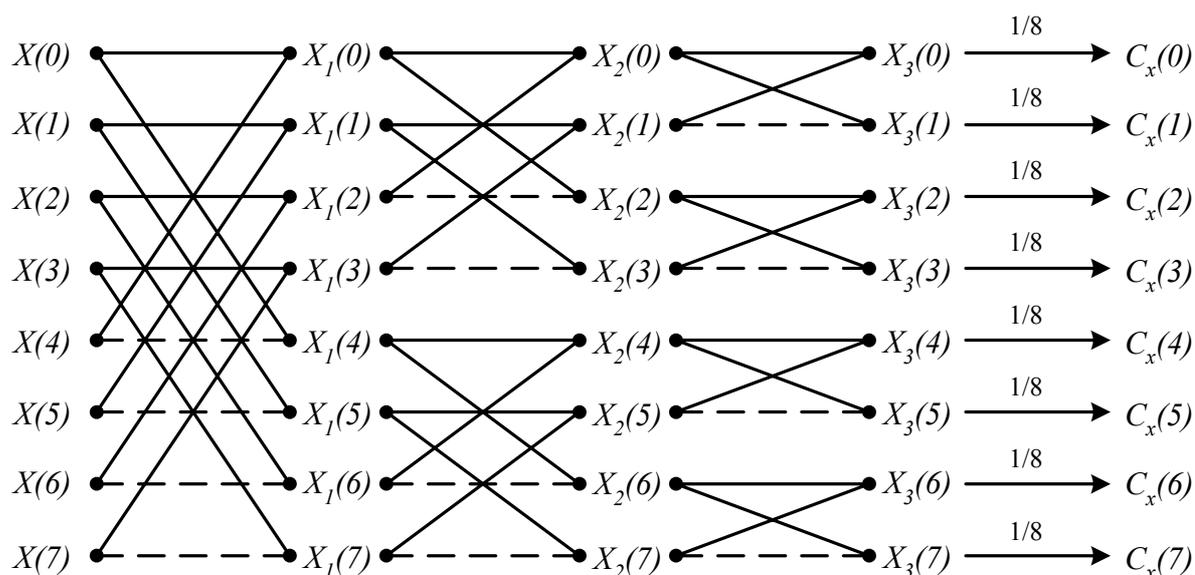


Рис. 6.4. Процедура быстрого алгоритма Уолша

7. МНОГОМЕРНЫЕ ПРЕОБРАЗОВАНИЯ ДЛЯ ОБРАБОТКИ ИЗОБРАЖЕНИЙ. ПРОСТЫЕ ВЕЙВЛЕТЫ. МАСШТАБИРУЮЩИЕ ФУНКЦИИ И ВЕЙВЛЕТ – ФУНКЦИИ. ПРЯМОЕ И ОБРАТНОЕ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЕ

7.1. Обработка стационарных и нестационарных сигналов

Большинство сигналов, встречающихся на практике, представлено во *временной области*, т.е. сигнал есть функция времени. Таким образом, при отображении сигнала на графике одной из координат (независимой) является ось времени, а другой координатой (зависимой) – ось амплитуд. Таким образом, получаем *амплитудно-временное представление* сигнала. Для большинства приложений обработки сигналов это представление не является наилучшим. Во многих случаях наиболее значимая информация скрыта в частотной области сигнала. Частотный спектр есть совокупность частотных (спектральных) компонент и он отображает наличие тех или иных частот в сигнале. Частота измеряется в Герцах (Гц), или в числе периодов в секунду. На рис. 7.1 представлены три косинусоиды: 3, 10 и 50 Гц соответственно.

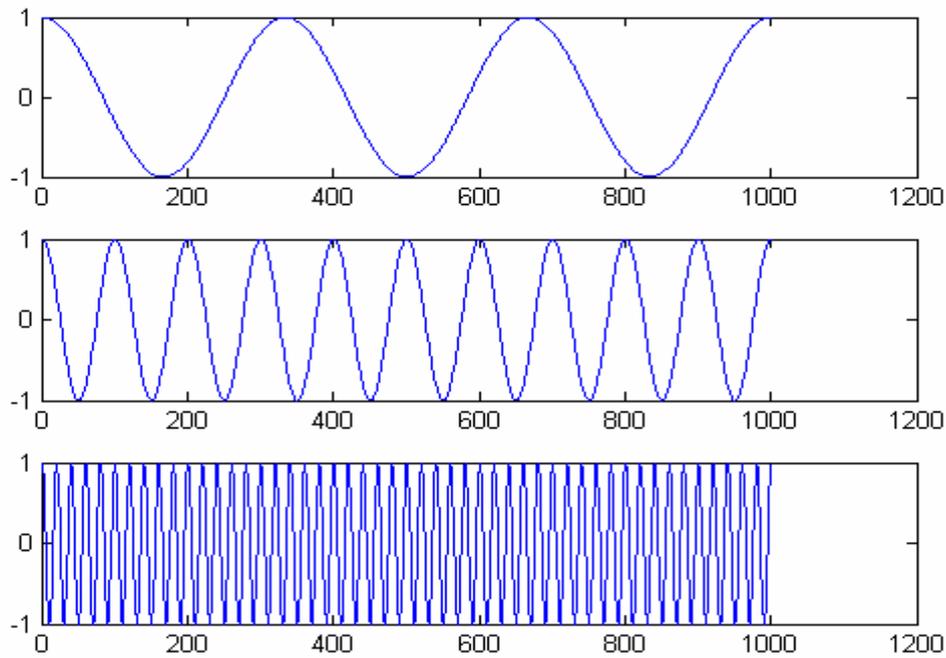


Рис. 7.1 Косинусоиды с частотой 3, 10 и 50 Гц соответственно

Классическим методом частотного анализа сигналов является преобразование Фурье.

Результат преобразования Фурье – амплитудно-частотный спектр, по которому можно определить присутствие некоторой частоты в исследуемом сигнале. Результаты преобразования Фурье трех косинусоид (3, 10 и 50 Гц соответственно) представлены ниже на рис. 7.2.

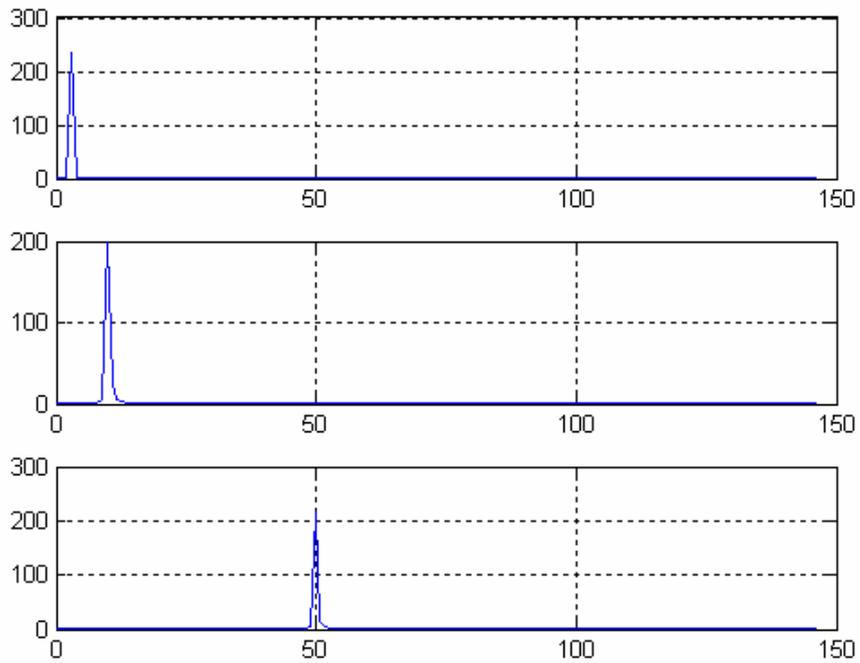


Рис. 7.2. Результаты преобразования Фурье косинусоид с частотой 3, 10 и 50 Гц соответственно

Приведем еще один пример стационарного сигнала

$x(t) = \cos(2\pi 10t) + \cos(2\pi 25t) + \cos(2\pi 50t) + \cos(2\pi 100t)$, его график и результаты преобразования Фурье (рис. 7.3).

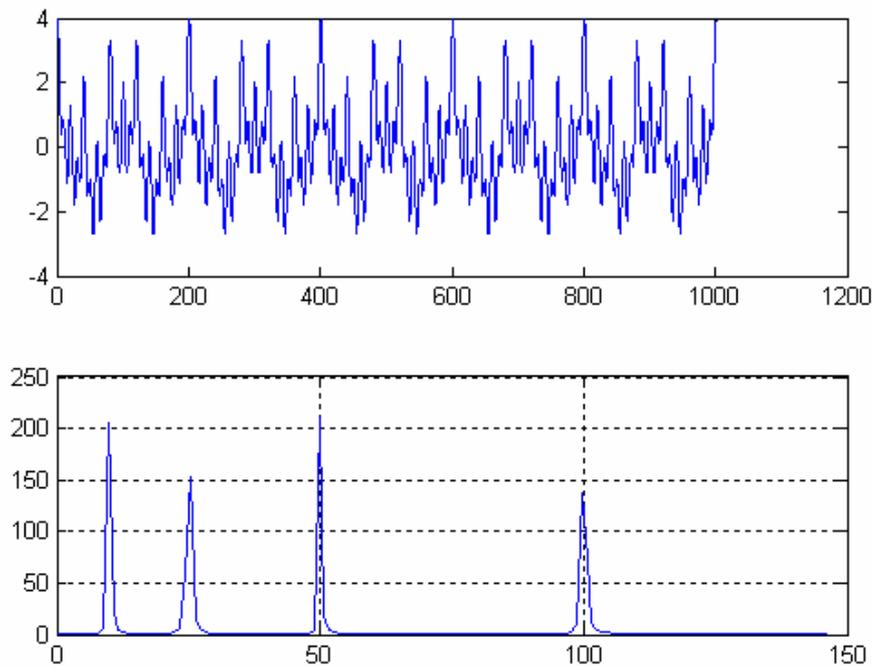


Рис. 7.3. Стационарный сигнал, результаты ПФ данного сигнала

Большинство реальных сигналов имеет сложные частотно-временные характеристики. Как правило, такие сигналы состоят из близких по времени, высокочастотных компонент и долговременных, близких по частоте низкочастотных компонент.

Сигнал на рис. 7.4 нестационарный. Его частота непрерывно изменяется во времени (четыре частотные компоненты соответствуют частотам 100, 50, 25, 10 Гц.)

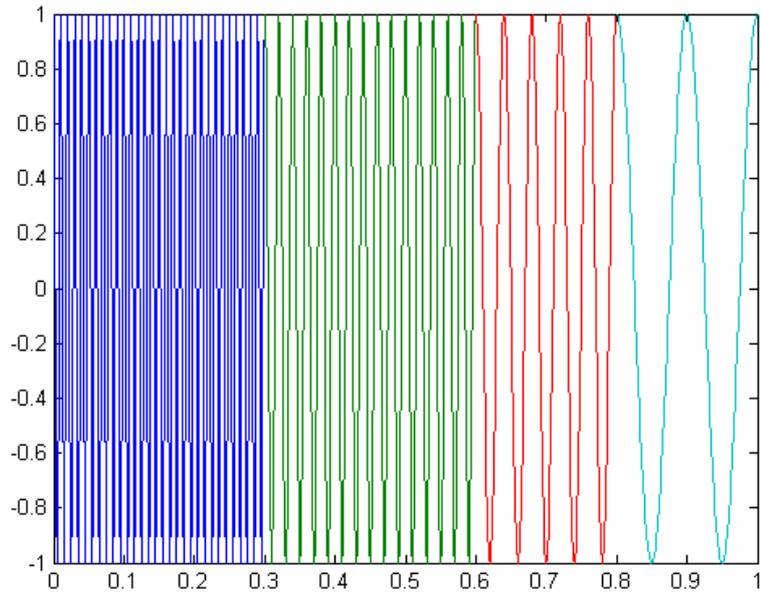


Рис. 7.4. Нестационарный сигнал

Спектр этого сигнала (ПФ) будет иметь вид (рис. 7.5):

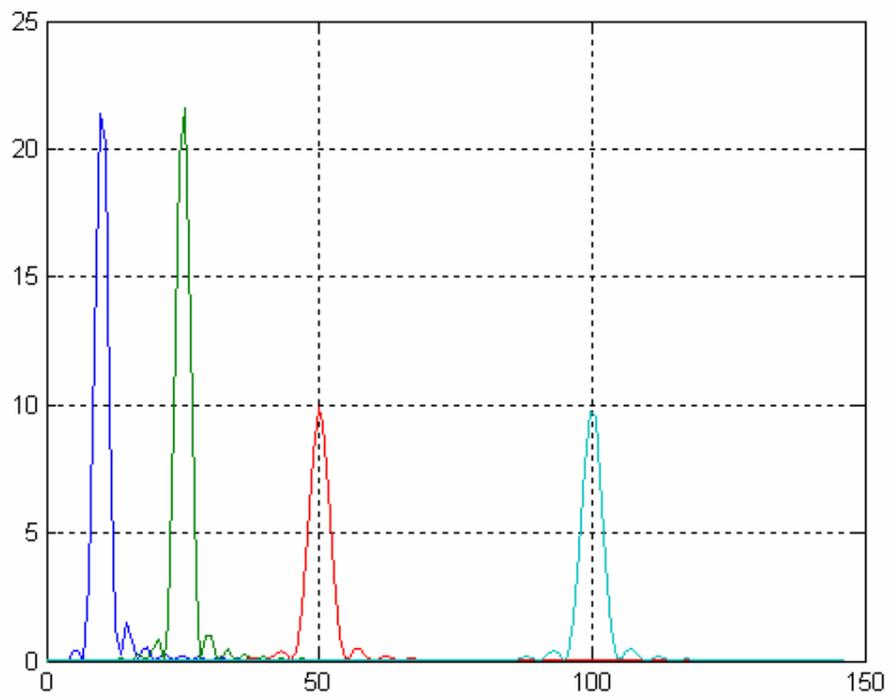


Рис. 7.5. Спектр нестационарного сигнала

В данном случае появляются небольшие «ложные» частоты и неодинаковости амплитудных пиков, поэтому ПФ не подходит для нестационарных сигналов.

Для анализа таких сигналов необходим метод, способный обеспечить хорошее разрешение и по частоте, и по времени. Первое требуется для локализации низкочастотных составляющих, второе – для разрешения компонент высокой частоты. Вейвлет - преобразование благодаря хорошей приспособленности к анализу нестационарных сигналов стало мощной альтернативой преобразованию Фурье.

Преобразование Фурье представляет сигнал, заданный во временной области, в виде разложения по ортогональным базисным функциям (синусам и косинусам), выделяя, таким образом, частотные компоненты. Недостаток преобразования Фурье заключается в том, что частотные компоненты не могут быть локализованы во времени, что накладывает ограничения на применимость данного метода к ряду задач (например, в случае изучения динамики изменения частотных параметров сигнала на временном интервале).

Существует два подхода к анализу нестационарных сигналов такого типа. Первый – оконное преобразование Фурье (short-time Fourier transform). Следуя по этому пути, мы работаем с нестационарным сигналом, как со стационарным, предварительно разбив его на сегменты (окна), статистика которых не меняется со временем. Вторым подходом – вейвлет преобразование. В этом случае нестационарный сигнал анализируется путем разложения по базисным функциям, полученным из некоторого прототипа путем сжатий, растяжений и сдвигов. Функция прототип называется материнским или анализирующим вейвлетом.

7.2. Краткий обзор оконного преобразования Фурье

В случае, когда не возникает вопрос о локализации временного положения частот, метод Фурье дает хорошие результаты. Но при необходимости определить временной интервал присутствия частоты приходится применять другие методы.

Одним из таких методов является обобщенный метод Фурье (оконное преобразование Фурье, ОПФ). Этот метод состоит из следующих этапов:

1. в исследуемой функции создается «окно» – временной интервал, для которого функция $f(x) \neq 0$, и $f(x) = 0$ для остальных значений;
2. для этого «окна» вычисляется преобразование Фурье;
3. «окно» сдвигается, и для него также вычисляется преобразование Фурье.

При ОПФ сигнал делится на отрезки («окна»), в пределах которых его можно считать стационарным. «Пройдя» таким «окном» вдоль всего сигнала, получается некоторая трехмерная функция, зависящая от положения «окна» и частоты (рис. 7.6).

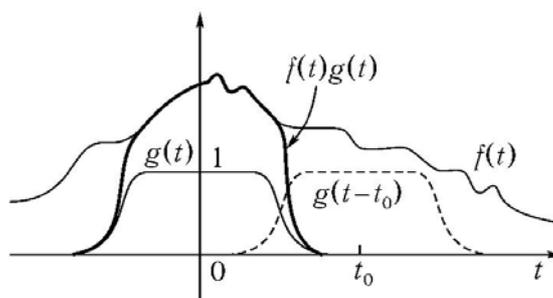


Рис. 7.6. Оконное преобразование Фурье

Функция $f(t)$ перемножается с оконной функцией $g(t)$, и вычисляются коэффициенты произведения $f(t)g(t)$. Затем процедура повторяется для сдвигов окна $g(t-t_0)$, $g(t-2t_0)$, ...

Данный подход позволяет определить факт присутствия в сигнале любой частоты и интервал ее присутствия. Это значительно расширяет возможности метода по сравнению с классическим преобразованием Фурье,

но существуют и определенные недостатки. Согласно следствиям принципа неопределенности Гейзенберга в данном случае нельзя утверждать факт наличия частоты ω_0 в сигнале в момент времени t_0 - можно лишь определить, что спектр частот (ω_1, ω_2) присутствует в интервале (t_1, t_2) . Причем разрешение по частоте (по времени) остается постоянным вне зависимости от области частот (времен), в которых производится исследование. Узкое окно обеспечивает временное разрешение, а широкое – частотное. Проблема состоит в том, что приходится выбирать окно «раз и навсегда», то есть для анализа сигнала, тогда как разные его участки могут потребовать применения разных окон.

Поэтому, если, например, в сигнале существенна только высокочастотная составляющая, то увеличить разрешение можно только, изменив параметры метода. В качестве метода, не обладающего подобного рода недостатками, интерес представляет аппарат вейвлет - анализа.

7.3. Основные положения вейвлет-анализа

Различают дискретный и непрерывный вейвлет-анализ, аппарат которых можно применять как для непрерывных, так и для дискретных сигналов.

Сигнал анализируется путем разложения по базисным функциям, полученным из некоторого прототипа путем сжатий, растяжений и сдвигов. Функция-прототип называется анализирующим (материнским) вейвлетом.

Вейвлет-функция должна удовлетворять двум условиям:

1. Среднее значение (интеграл по всей прямой) равно 0.
2. Функция быстро убывает при $t \rightarrow \infty$.

Обычно функция-вейвлет обозначается буквой ψ . В общем случае *непрерывное вейвлет-преобразование (CWT)* функции $f(t)$ выглядит так:

$$\Psi_x^\omega(\tau, s) = \frac{1}{\sqrt{|s|}} \int x(t) \psi^*\left(\frac{t-\tau}{s}\right) dt, \quad (7.1)$$

где τ - параметр сдвига, s – параметр масштаба, функция $\psi^*\left(\frac{t-\tau}{s}\right)$ – это «материнский вейвлет», то есть функция преобразования.

Слово вейвлет означает маленькая волна. Под маленькой понимается то, что эта функция (окно) имеет конечную ширину. Термин «материнский» означает, что функции с различной шириной носителя, используемые в преобразовании, порождаются одной базовой функцией – материнским вейвлетом. То есть материнский вейвлет является прототипом для всех оконных функций.

Основные вейвлетообразующие функции, или материнские вейвлеты, приведены в табл. 7.1.

Наиболее распространенные вещественные базисы конструируются на основе производных функции Гаусса ($g_0(t) = \exp(-t^2/2)$). Это обусловлено тем обстоятельством, что функция Гаусса имеет наилучшие показатели локализации как во временной, так и в частотной областях.

На рис. 7.7 показаны вейвлеты первых четырех порядков и модули их спектральной плотности. При $n=1$ получаем вейвлет первого порядка, называемый WAVE-вейвлетом с равным нулю нулевым моментом. При $n=2$ получаем МНАТ-вейвлет, называемый «мексиканская шляпа» (mexican hat – похож на сомbrero). У него нулевой и первый моменты равны нулю. Он имеет лучшее разрешение, чем WAVE-вейвлет.

Совместное использование вейвлетов $g_1 - g_4$ для вейвлет-преобразования (ВП) существенно повышает точность вейвлет-анализа.

Основные материнские вейвлеты

Вейвлеты	Аналитическая запись $\psi(t)$	Спектральная плотность $\psi(\omega)$
Вещественные непрерывные базисы		
Гауссовы: – первого порядка, или WAVE-вейвлет, – второго порядка, или МНАТ-вейвлет «мексиканская шляпа» – <u>mexican hat</u>), – n -го порядка,	$-t \exp(-t^2/2)$ $(1-t^2) \exp(-t^2/2)$ $(-1)^n \frac{d^n}{dt^n} [\exp(-t^2/2)]$	$(i\omega) \sqrt{2\pi} \exp(-\omega^2/2)$ $(i\omega)^2 \sqrt{2\pi} \exp(-\omega^2/2)$ $(-1)^n (i\omega)^n \sqrt{2\pi} \exp(-\omega^2/2)$
DOG – difference of gaussians	$e^{-t^2/2} - 0,5e^{-t^2/8}$	$\sqrt{2\pi}(e^{-\omega^2/2} - e^{-2\omega^2})$
LP-Littlewood & Paley	$(\pi t)^{-1} (\sin 2\pi t - \sin \pi t)$	$\begin{cases} (2\pi)^{-1/2}, \pi \leq t \leq 2\pi, \\ 0, \text{ в противном случае} \end{cases}$
Вещественные дискретные		
НААР-вейвлет	$\begin{cases} 1, 0 \leq t \leq 1/2, \\ \geq -1, 1/2 \leq t \leq 1, \\ 0, t < 0, t > 0. \end{cases}$	$i e^{i\omega/2} \frac{\sin^2 \omega/4}{\omega/4}$
ФНАТ-вейвлет, или «французская шляпа» (<u>French hat</u> – похож на цилиндр)	$\begin{cases} 1, t \leq 1/3, \\ \geq -1/2, 1/3 \leq t \leq 1, \\ 0, t > 1. \end{cases}$	$\frac{4 \sin^3 \omega/3}{3 \omega/3}$
Комплексные		
Морле (Morlet)	$e^{i\omega_0 t} e^{-t^2/2}$	$\sigma(\omega) \sqrt{2\pi} e^{-(\omega-\omega_0)^2/2}$

Пауля (Paul) (чем больше n , тем больше нулевых моментов имеет вейвлет)	$\Gamma(n+1) \frac{i^n}{(1-n)^{n+1}}$	$\sigma(\omega) \sqrt{2\pi} (\omega)^n e^{-\omega}$
---	---------------------------------------	---

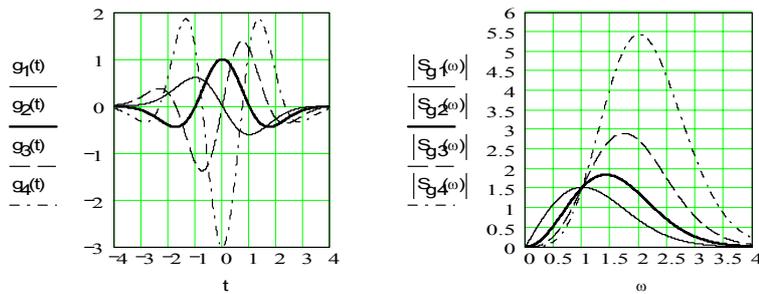


Рис. 7.7. Вейвлеты первых четырех порядков и модули их спектральной плотности

Наиболее простой пример *дискретного* вейвлета – это HAAR-вейвлет. Недостатком его являются несимметричность формы и негладкость – резкие границы в t -области, вследствие чего возникает бесконечное чередование «лепестков» в частотной области, хотя и убывающих как $1/\omega$.

LR -вейвлет, имеющий, наоборот, резкие границы в ω -области, можно считать другим предельным случаем.

Среди комплексных вейвлетов наиболее часто используется базис, основанный на хорошо локализованном и во временной, и в частотной областях вейвлете Морле. Характерный параметр ω_0 позволяет изменять избирательность базиса. Вещественная и мнимая части $\psi(t)$ – это амплитудно-модулированные колебания.

Выбор конкретного материнского вейвлета (будь то непрерывный или дискретный) целиком зависит от характера поставленной задачи и от конкретного анализируемого сигнала. Разные сигналы удастся анализировать тем или иным способом и критерием успеха обычно служит простота

получаемого разложения. При этом решающим фактором оказываются интуиция и практический опыт исследователя.

При ВП мы не имеем частотного параметра, как это было при ДПФ. Вместо него здесь имеется параметр масштаба, который можно определить как величину, обратную частоте. Параметр масштаба в вейвлет-анализе имеет аналогию с масштабом географических карт. Большое значение масштаба соответствуют малому количеству деталей, глобальному представлению сигнала, а низкие значения масштаба позволяют различить детали. Аналогично, в терминах частоты, низкие частоты соответствуют глобальной информации о сигнале (которая содержится на всей его протяженности), а высокие частоты – детальной информации, скрытым особенностям, которые имеют обычно малую протяженность. Масштабирование, как математическая операция, расширяет или сжимает сигнал. Поэтому, при $s > 1$ расширяет сигнал, а при $s < 1$ сжимает его.

В настоящее время большинство вычислений выполняют компьютеры. Это означает, что ни ПФ, ни ОПФ, ни НВП не могут быть практически вычислены путем взятия интегралов. Так как интегралы при вычислении заменяются суммами, то достаточно логичным будет применить дискретизацию частотно-временной области, то есть перейти к **дискретизированному НВП**. Параметр масштаба s дискретизируется на логарифмической сетке, а параметр времени дискретизируется в соответствии с параметром масштаба, что означает, что на разных масштабах имеет место различная частота дискретизации. Чем больше s , тем меньше временных отсчетов мы берем – поэтому для низких частот получаем хорошую частотную и скудную временную информацию, а для высоких частот наоборот. Дискретизированное НВП будет определяться следующим выражением:

$$\Psi_x(j, k) = \int x(t)\psi_{j,k}(t)dt, \quad (7.2)$$

Из-за некоторых недостатков от дискретизированного НВП перешли к **дискретному вейвлет-преобразованию (ДВП)**.

Для начала определим понятие *субполосное кодирование (subband coding)*. Это результат свертки сигнала с несколькими полосовыми фильтрами и децимацией результата. Совокупность набора фильтров и дециматоров называется *банком фильтров*. Каждый получившийся в результате преобразования сигнал несет в себе информацию о спектральной составляющей исходного сигнала при некотором пространственном (временном) масштабе. Для обратного синтеза сигнала выполняется операция интерполяции субполосных сигналов, фильтрация и их сложение. Использование субполосного кодирования вкупе с *кратномасштабным анализом (multiresolution analysis)*, который представляет собой процесс декомпозиции сигнала на различных частотах и различном разрешении одновременно, позволяет получить масштабно-временное представление сигнала. Разрешение сигнала изменяется за счет фильтрации сигнала и является мерой количества детальной информации в нем, масштаб изменяется за счет интерполяции и децимации. Интерполяция соответствует увеличению частоты дискретизации сигнала, а децимация понижению. Эти процедуры осуществляются добавлением новых отсчетов между существующими отсчетами сигнала и удалением некоторых отсчетов из сигнала соответственно.

В ходе ДВП сигнал анализируется в различных частотных полосах с различным разрешением путем декомпозиции на грубую аппроксимацию (полусуммы соседних значений сигнала) и детали (полуразности соседних значений сигнала). Таким образом, определены два множества функций: масштабирующие функции и вейвлеты, соответствующие низкочастотным (НЧ) и высокочастотным (ВЧ) фильтрам.

$$\text{масштабирующие коэффициенты: } a_{j-1,k} = \frac{1}{\sqrt{2}}(a_{j,2k} + a_{j,2k+1}) \quad (7.3)$$

$$\text{вейвлет коэффициенты: } b_{j-1,k} = \frac{1}{\sqrt{2}}(a_{j,2k} - a_{j,2k+1}) \quad (7.4)$$

Эта особенность была замечена Маллатом и изложена им в одной из своих работ.

$$Y_{low}[k] = \sum_n x(n)h(2k - n) \quad (7.5)$$

$$Y_{high}[k] = \sum_n x(n)g(2k - n) \quad (7.6)$$

Эти два выражения определяют два фильтра: низкочастотный и высокочастотный. Добавляя к этой группе фильтров два дециматора, использование которых обусловлено избыточностью данных получаемых с выходов фильтров, получаем банк фильтров анализа (рис. 7.8). Следует отметить, что операция децимации в данном случае удаляет из сигнала нечетные компоненты и оставляет четные. Единичное применение такого банка фильтров к сигналу делит его на две части, в одной из которых содержатся высокие частоты сигнала, а в другой низкие. В матричном виде это выглядит так:

$$\begin{bmatrix} A_n \\ D_n \end{bmatrix} x = \begin{bmatrix} a_{n-1} \\ d_{n-1} \end{bmatrix}, \quad (7.7)$$

где A_n, D_n фильтрующие матрицы размерностью $2^{n-1} \times 2^n$, построенные на основании коэффициентов фильтра в зависимости от разрядности x и претерпевшие децимацию – удаление нечетных строк; x – исходный сигнал; a_{n-1} и d_{n-1} – низкочастотная и высокочастотная составляющие сигнала x .

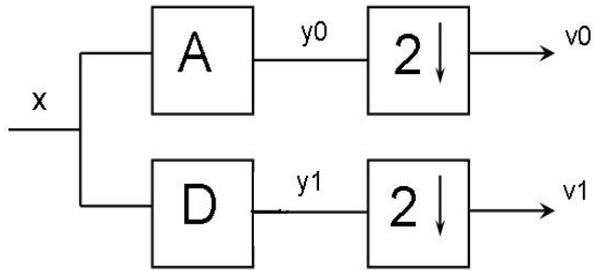


Рис. 7.8. ДВП. Процедура анализа:

A, D – фильтрующие матрицы (низкочастотная и высокочастотная),
 $2\downarrow$ - дециматор, v_0 и v_1 - a_{n-1} и d_{n-1} соответственно

Так как строки комбинированной матрицы (состоящей из A_n, D_n) взаимно ортогональны, так же, как и столбцы, то квадратная матрица инвертируема. Процедура инвертирования в этом случае соответствует процедуре транспонирования матрицы. Полученная матрица $[A^T \ D^T]$ представляет собой банк фильтров синтеза сигнала. Умножение этой матрицы на исходную справа дает нам единичный вектор I . Следовательно, мы имеем ортогональный банк фильтров синтеза – анализа (рис. 7.9).

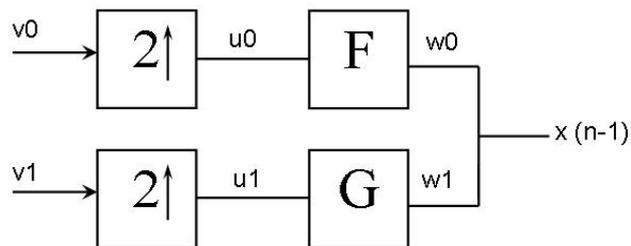


Рис. 7.9. ДВП. Процедура анализа:

v_0 и v_1 - a_{n-1} и d_{n-1} соответственно, $2\uparrow$ - интерполятор,
 F и G – добавляющий и вычитающий фильтры соответственно

Интерполятор представляет собой узел, который вставляет между двумя значениями сигнала третье – ноль. Работа фильтров F и G восстанавливает исходный сигнал x . Таким образом, две эти схемы вместе образуют банк фильтров ДВП.

Аналогично появлению в свое время быстрого ПФ (БПФ), появилось *быстрое вейвлет – преобразование (БВП, FWT)*. БВП вытекает из классического ДВП рекурсивным применением последнего, что возможно благодаря принципу кратномасштабного анализа (рис. 7.10).

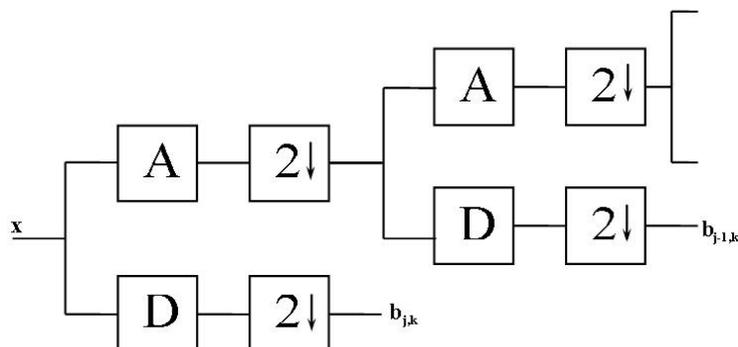


Рис. 7.10. БВП. Логарифмическое дерево банков фильтров

Сигнал x поступает на вход первого банка фильтров в дереве. После преобразования и децимации коэффициенты ВЧ фильтра $b_{j,k}$ сохраняются, так как они являются выходом на своем уровне «логарифмического дерева». Выходы же НЧ фильтра после децимации, когда в выходном сигнале остались только четные компоненты, поступают на вход следующего банка фильтров, где процедура повторяется. Выходом второго уровня являются коэффициенты $b_{j-1,k}$. Эта схема будет повторяться $L = \log_2 N$ раз, где N это число компонент в исходном сигнале. В итоге мы получим $L-1$ векторов с ВЧ фильтров и общее среднее – выход последнего НЧ фильтра, то есть полную декомпозицию сигнала. Применение такого рекурсивного алгоритма и есть БВП. Это преобразование подходит для любого сигнала, длина которого есть степень двойки, а также для любых вейвлетов, относящихся к классу ортогональных. Следует также добавить, что БВП обладает и обратным преобразованием, которое строится с использованием инверсного ДВП, когда в соответствие каждому блоку анализа на каждом уровне ставится блок синтеза. Таким образом, имеет место обратная рекурсия. Впервые такая связь между вейвлетами и фильтрами была открыта Стивеном Маллатом, поэтому

в его честь этот алгоритм называют «алгоритмом Маллата» или «пирамидой Маллата».

Другим подходом к реализации БВП является *лифтинговая схема (lifting scheme)*. Подобная реализация БВП позволяет конструировать очень многие типы вейвлетов, включая ортогональные, биортогональные, вейвлет Хаара и другие, вне зависимости от преобразования Фурье.

Лифтинговая схема, как и алгоритм Маллата, включает в себя два блока: блок анализа и блок синтеза сигнала. Рассмотрим блок анализа (рис. 7.11).,

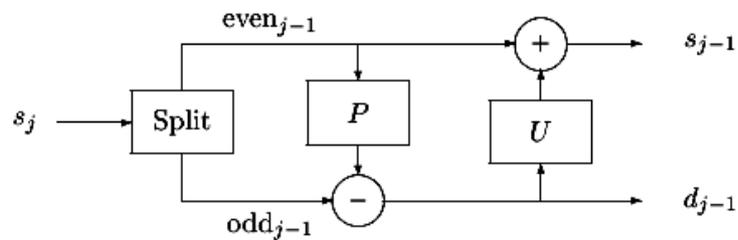


Рис. 7.11. Лифтинговая схема, блок анализа

На рисунке показано, что вектор входных значений s_j поступает на вход *блока разделения (Split)*, который представляет собой устройство разбиения входного сигнала на две подпоследовательности по определенному закону с целью максимально «декоррелировать» сигнал. Одним из необходимых условий при проведении такой операции является возможность обратного восстановления сигнала из двух его подпоследовательностей, что используется на этапе анализа в *блоке объединения (Merge)* (рис. 7.12). Достаточно часто применяется разделение сигнала на две последовательности, состоящие из четных (even) и нечетных (odd) отсчетов исходного сигнала. Такая процедура получила название *Lazy wavelet Transform* – «ленивое» (не требующее больших усилий) вейвлет преобразование, которое задается выражением:

$$(even_{j-1}; odd_{j-1}) := Split (s_j) \quad (7.8)$$

Следующий блок в схеме анализа – это *блок предсказания (Predict)*. На предыдущем шаге мы разделили исходный сигнал на две подпоследовательности четных и нечетных отсчетов. Если структура сигнала имеет локальную корреляцию, то четные и нечетные отсчеты будут сильно коррелированы. То есть, другими словами, имея одну подпоследовательность, должна быть возможность предсказать другую с заданной точностью. Всегда используется подпоследовательность четных отсчетов, чтобы предсказать подпоследовательность нечетных. В данном случае (вейвлет Хаара) нечетный отсчет $s_{j;2l+1}$ будет использовать своего левого соседа $s_{j;2l}$ как предсказание. Тогда уточняющая разность $d_{j-1;l}$ будет являться разностью нечетного отсчета и предсказания:

$$d_{j-1;l} = s_{j;2l+1} - s_{j;2l} \quad (7.9)$$

Таким образом, операция предсказания определяется следующим выражением:

$$d_{j-1} = odd_{j-1} - P(even_{j-1}) \quad (7.10)$$

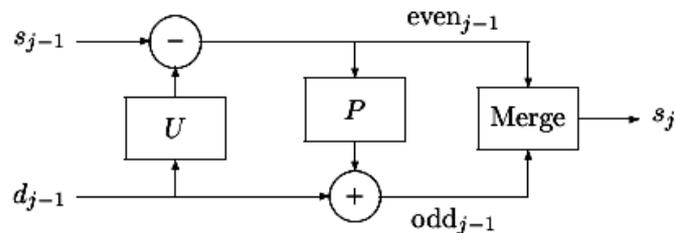


Рис. 7.12. Лифтинговая схема, блок синтеза

И, наконец, последний блок – *блок обновления (Update)*. Коэффициенты, полученные на предыдущем этапе, необходимы чтобы

«поднять» или «подтянуть» («обновить») оставшуюся часть отсчетов сигнала, так как вторая половина будет сохранена и анализироваться не будет. После процедуры четная подпоследовательность будет содержать информацию обо всем сигнале, так как, используя оператор предсказания, множество s будет дополняться информацией о множестве d , поскольку к этому моменту вместо множества d в памяти хранятся коэффициенты d^* преобразования (характеризующие множество d). Обновление четной подпоследовательности определяется следующим выражением:

$$s_{j-1} = even_{j-1} + U(d_{j-1}). \quad (7.11)$$

Получаемые коэффициенты s будут рекурсивно поступать на вход схемы или переходить на более низкий уровень точно такой же схемы с целью произвести полную декомпозицию сигнала. Процедура будет повторяться до тех пор, пока множество s не будет содержать единственный последний коэффициент s_{00} , который теперь будет именоваться – *общее среднее*.

8. АЛГОРИТМЫ ДЛЯ ВЫЧИСЛЕНИЯ ПРЕОБРАЗОВАНИЯ ХААРА

8.1. Ортогональные функции Хаара

Множество функций Хаара $har(n,m,t)$ (рис. 8.1), образующих периодическую, ортонормированную и полную систему функций, было предложено в 1910 году. Рекуррентное соотношение, позволяющее получить $har(n,m,t)$, имеет вид

$$har(0,0,t) = 1, \forall t \in [0,1);$$

$$har(r,m,t) = \begin{cases} 2^{r/2} & , \frac{m-1}{2^r} \leq t < \frac{m-1/2}{2^r}, \\ -2^{r/2} & , \frac{m-1/2}{2^r} \leq t < \frac{m}{2^r}, \\ 0 & , \text{при остальных } t \in [0,1), \end{cases}$$

где $0 \leq r < \log_2 N$, $1 \leq m \leq 2^r$

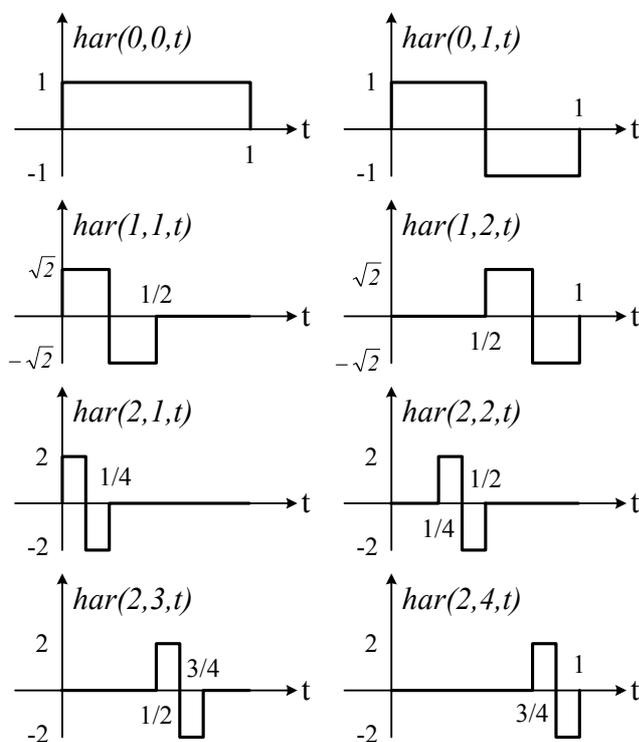


Рис. 8.1. Функции Хаара для $N=8$

Коэффициенты преобразования Хаара $Y(k)$, $k = \overline{0, N-1}$, соответствующие входной последовательности $\{X(m)\}$, получаются в результате преобразования

$$\vec{Y}(n) = \frac{1}{N} H^*(n) \vec{X}(n), \text{ где } H^*(n) \text{ – матрица Хаара размерностью } N \times N.$$

Для $N=8$

$$H^*(3) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} & & & & \\ & & & & \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} \\ 2 & -2 & & & & & & \\ & & 2 & -2 & & & & \\ & & & & 2 & -2 & & \\ & & & & & & 2 & -2 \end{bmatrix}.$$

Рассматривая $H^*(n)$, можно отметить, что $N/2$ коэффициентов преобразования Хаара соответствуют корреляции двух соседних точек в пространстве входных последовательностей, $N/4$ коэффициентов соответствуют связям четырех соседних точек и так далее до N/N коэффициентов, соответствующих всем N координатам пространства входных последовательностей. Это означает, что область преобразования в случае преобразования Хаара обладает свойством как локальной, так и глобальной чувствительности. При дискретном преобразовании Фурье и преобразовании Уолша-Адамара каждый коэффициент преобразования является функцией всех координат пространства входных последовательностей (свойство глобальной чувствительности), а в преобразовании Хаара это относится к первым двум коэффициентам.

Для осуществления преобразования Хаара требуется $2(N-1)$ операций сложения / вычитания и N операций умножения.

8.2. Быстрое преобразование Хаара

Для $N=8$ алгоритм преобразования Хаара типа Кули-Тьюки может быть представлен следующим образом. В исходной матрице преобразования Хаара необходимо переупорядочить столбцы $H(3)$, пользуясь последовательно двоичной инверсией при $N=8, 4, 2$.

Шаг 1. Переставим столбцы $H^*(3)$ в соответствии с двоичной инверсией их номеров при $N=8$, т.е. $\{0, 1, 2, 3, 4, 5, 6, 7\} \rightarrow \{0, 4, 2, 6, 1, 5, 3, 7\}$, что приведет к

$$H^*(3) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ \sqrt{2} & 0 & -\sqrt{2} & 0 & \sqrt{2} & 0 & -\sqrt{2} & 0 \\ 0 & \sqrt{2} & 0 & -\sqrt{2} & 0 & \sqrt{2} & 0 & -\sqrt{2} \\ \boxed{2 & 0 & 0 & 0} & \boxed{-2 & 0 & 0 & 0} \\ \boxed{0 & 0 & 2 & 0} & \boxed{0 & 0 & -2 & 0} \\ \boxed{0 & 2 & 0 & 0} & \boxed{-2 & 0 & 0 & 0} \\ \boxed{0 & 0 & 0 & 2} & \boxed{0 & 0 & 0 & -2} \end{bmatrix}.$$

0 1 2 3 0 1 2 3

Шаг 2. Переставим столбцы (4×4) матриц, заключенных в квадраты в соответствии с двоичной инверсией номеров столбцов при $N=4$, т.е. $\{0, 1, 2, 3\} \rightarrow \{0, 2, 1, 3\}$, что приведет к матрице

$$H^*(3) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ \boxed{\sqrt{2}} & \boxed{0} & \boxed{-\sqrt{2}} & \boxed{0} & \boxed{\sqrt{2}} & \boxed{0} & \boxed{-\sqrt{2}} & \boxed{0} \\ \boxed{0} & \boxed{\sqrt{2}} & \boxed{0} & \boxed{-\sqrt{2}} & \boxed{0} & \boxed{\sqrt{2}} & \boxed{0} & \boxed{-\sqrt{2}} \\ 2 & 0 & 0 & 0 & -2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & -2 \end{bmatrix}.$$

Шаг 3. Переставим столбцы (2×2) матриц, заключенных в квадраты в соответствии с двоичной инверсией номеров столбцов при $N=2$, т.е. $\{0,1\} \rightarrow \{1,0\}$, что приводит к матрице, совпадающей с $H^*(3)$.

Граф-схемы прямого и обратного быстрого преобразования Хаара представлены соответственно на рис 8.2 и 8.3.

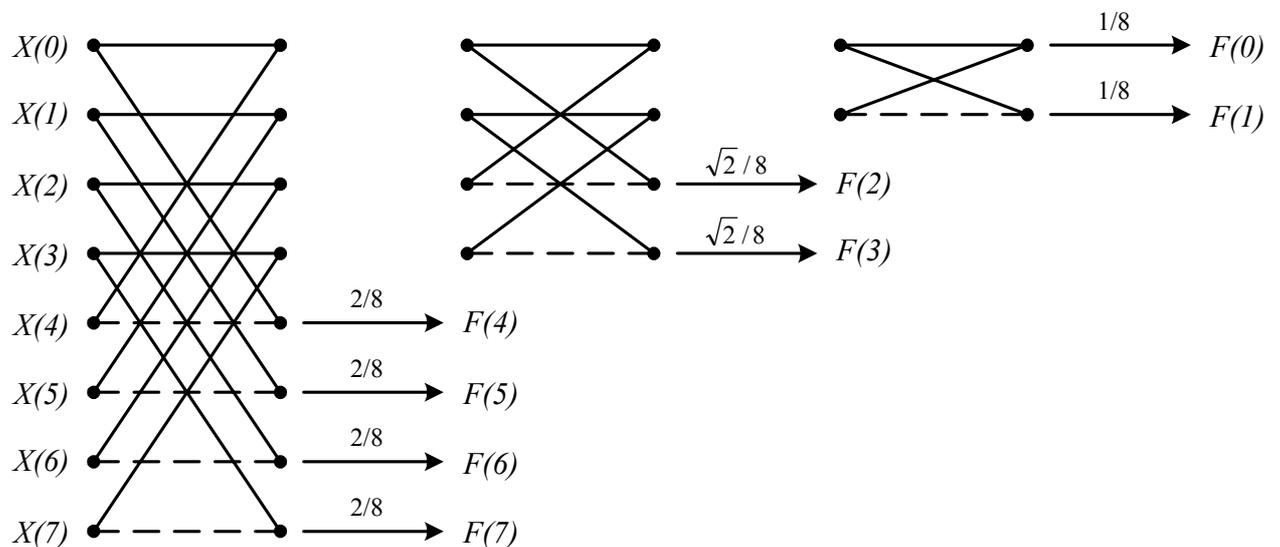


Рис. 8.2. Граф-схема прямого быстрого преобразования Хаара

Тогда спектр можно определить как

$$P_h(0) = F^2(0);$$

$$P_h(1) = F^2(1);$$

$$P_h(2) = \sum_{k=2}^3 F^2(k);$$

$$P_h(3) = \sum_{k=4}^7 F^2(k).$$

В общем виде можно записать

$$P_h(0) = F^2(0);$$

$$P_h(r) = \sum_{k=2^{r-1}}^{2^r-1} F^2(k); \quad r = \overline{1, n}.$$

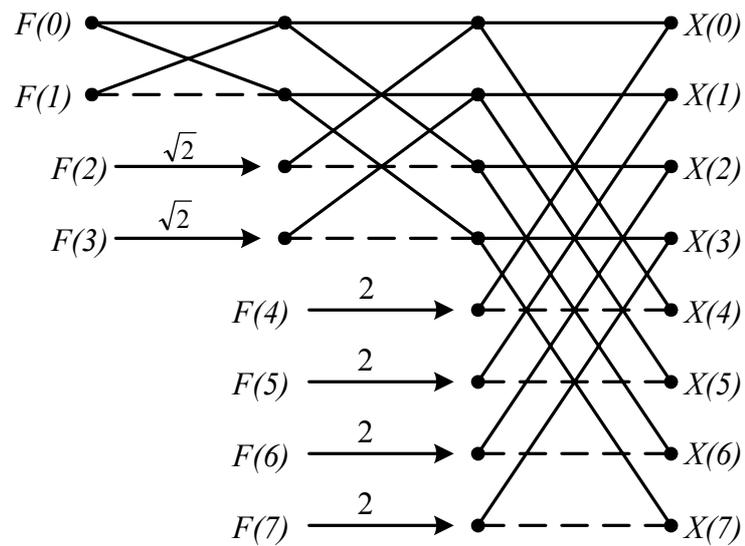


Рис. 8.3. Граф-схема быстрого обратного преобразования Хаара

9. СТАЦИОНАРНЫЕ СЛУЧАЙНЫЕ ПРОЦЕССЫ

Случайный процесс $\{x_k(t)\}$ ($-\infty < t < \infty$), иначе временной ряд или стохастический процесс, обозначается символом $\{\cdot\}$ - это совокупность действительных (или комплексных) функций, которую можно охарактеризовать её вероятностной структурой. Каждая отдельная функция $x_k(t)$, где t - переменная, а k фиксировано, называется выборочной функцией. С практической точки зрения, выборочную функцию (или некоторый отрезок выборочной функции конечной длины) можно считать наблюдаемым результатом отдельного эксперимента. Возможное число экспериментов определяется выборочным пространством индексов k , причём последнее может быть счётным или несчётным (рис. 9.1).

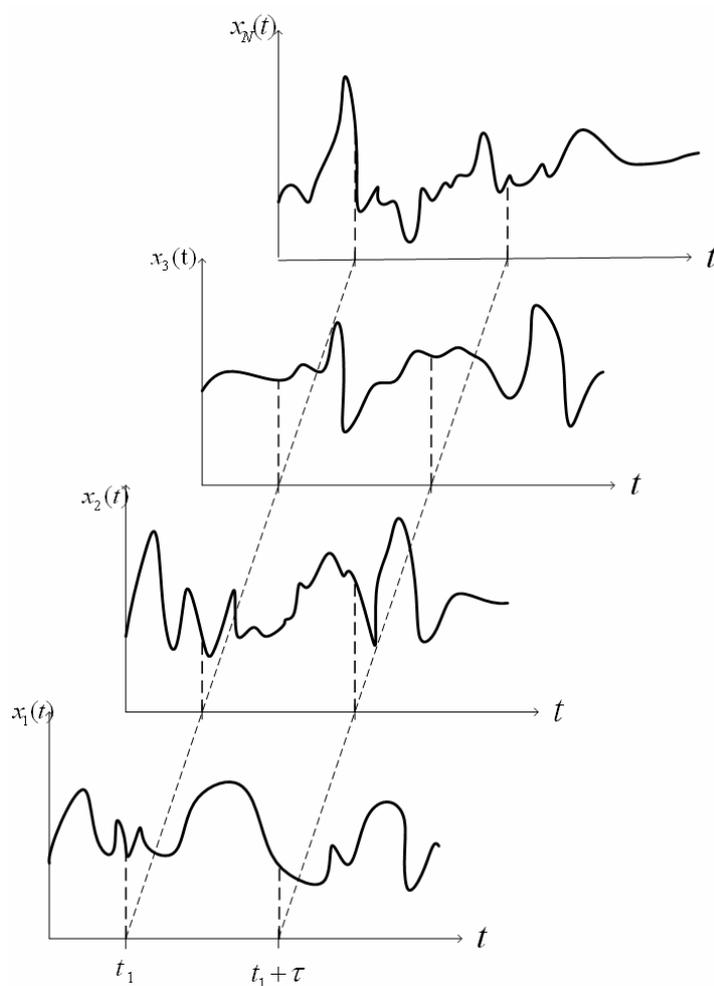


Рис. 9.1. Ансамбль реализаций, задающий случайный процесс

Рассмотрим два произвольных случайных процесса $\{x_k(t)\}$ и $\{y_k(t)\}$. В первую очередь, представляют интерес средние значения по ансамблю в произвольный фиксированный момент времени t , где $x_k(t)$ и $y_k(t)$ - случайные величины от аргумента k . Такие средние значения определяются:

$$\mu_x(t) = E[x_k(t)] \quad \mu_y(t) = E[y_k(t)], \quad (9.1)$$

где E – символ математического ожидания.

Следующие важные статистические характеристики – это корреляционные функции, определённые для произвольных фиксированных моментов времени $t_1 = t$ и $t_2 = t + \tau$.

$$\begin{aligned} C_{xx}(t_1, t_2) &= C_{xx}(t, t + \tau) = E\{[x_k(t) - \mu_x(t)][x_k(t + \tau) - \mu_x(t + \tau)]\} \\ C_{yy}(t_1, t_2) &= C_{yy}(t, t + \tau) = E\{[y_k(t) - \mu_y(t)][y_k(t + \tau) - \mu_y(t + \tau)]\} \\ C_{xy}(t_1, t_2) &= C_{xy}(t, t + \tau) = E\{[x_k(t) - \mu_x(t)][y_k(t + \tau) - \mu_y(t + \tau)]\} \end{aligned} \quad (9.2)$$

В общем случае эти величины различны для различных комбинаций t_1 и t_2 .

При $\tau = 0$:

$$\begin{aligned} C_{xx}(t, t) &= E[(x_k(t) - \mu_x(t))^2] = \sigma_x^2 \\ C_{yy}(t, t) &= E[(y_k(t) - \mu_y(t))^2] = \sigma_y^2 \\ C_{xy}(t, t) &= E\{[x_k(t) - \mu_x(t)][y_k(t) - \mu_y(t)]\} \end{aligned} \quad (9.3)$$

Следовательно, корреляционные функции $C_{xx}(t, t)$ и $C_{yy}(t, t)$ совпадают с обычными дисперсиями процессов $\{x_k(t)\}$ и $\{y_k(t)\}$ в фиксированный момент

времени, в то время как $C_{xy}(t, t)$ представляет собой корреляцию случайных величин $\{x_k(t)\}$ и $\{y_k(t)\}$.

По ансамблю реализаций можно определить и другие статистические характеристики, зависящие от трёх и более моментов времени.

Если средние значения $\mu_x(t)$ и $\mu_y(t)$ вместе с корреляционными функциями $C_{xx}(t, t + \tau)$, $C_{yy}(t, t + \tau)$, и $C_{xy}(t, t + \tau)$ принимают одинаковые значения при всех фиксированных значениях t , то случайные процессы $\{x_k(t)\}$ и $\{y_k(t)\}$ называются слабо стационарными. Если всевозможные распределения, связанные с $\{x_k(t)\}$ и $\{y_k(t)\}$, не зависят от выбора начала отсчёта времени, то такие процессы называются строго стационарными.

Для стационарных случайных процессов $\{x_k(t)\}$ и $\{y_k(t)\}$ средние значения постоянны и не зависят от t . То есть:

$$\mu_x = E[x_k(t)] = \int_{-\infty}^{\infty} xp(x)dx \qquad \mu_y = E[y_k(t)] = \int_{-\infty}^{\infty} yp(y)dy \qquad (9.4)$$

Где $p(x)$, $p(y)$ - плотности вероятности случайных величин $x_k(t)$ и $y_k(t)$ соответственно.

В свою очередь, для произвольных фиксированных величин t и τ :

$$\begin{aligned} R_{xx}(\tau) &= E[x_k(t)x_k(t + \tau)] \\ R_{yy}(\tau) &= E[y_k(t)y_k(t + \tau)] \\ R_{xy}(\tau) &= E[x_k(t)y_k(t + \tau)] \end{aligned} \qquad (9.5)$$

Величины $R_{xx}(\tau)$ и $R_{yy}(\tau)$ называются автокорреляционными или ковариационными функциями процессов $\{x_k(t)\}$ и $\{y_k(t)\}$ соответственно, а $R_{xy}(\tau)$ называется взаимной ковариационной функцией процессов $\{x_k(t)\}$ и $\{y_k(t)\}$.

Для того, чтобы $R_{xx}(\tau)$ была ковариационной функцией слабо стационарного случайного процесса $\{x_k(t)\}$, необходимо и достаточно, чтобы $R_{xx}(\tau) = R_{xx}(-\tau)$ и $R_{xx}(\tau)$ была положительно определённой функцией.

В терминах плотностей вероятности:

$$\begin{aligned}
 R_{xx}(\tau) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 p(x_1, x_2) dx_1 dx_2 \\
 R_{yy}(\tau) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y_1 y_2 p(y_1, y_2) dy_1 dy_2 \\
 R_{xy}(\tau) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 y_2 p(x_1, y_2) dx_1 dy_2
 \end{aligned}
 \tag{9.6}$$

При произвольных μ_x и μ_y ковариационные функции связаны с корреляционными с помощью соотношений:

$$\begin{aligned}
 R_{xx}(\tau) &= C_{xx}(\tau) - \mu_x^2 \\
 R_{yy}(\tau) &= C_{yy}(\tau) - \mu_y^2 \\
 R_{xy}(\tau) &= C_{xy}(\tau) - \mu_x \mu_y
 \end{aligned}
 \tag{9.7}$$

Следовательно, корреляционные функции совпадают с ковариационными, если средние значения равны нулю.

В свою очередь, два стационарных процесса не коррелированы, если $C_{xy}(\tau) = 0$ для всех τ . Следовательно, если $R_{xy}(\tau) = 0$, то два процесса не коррелированы только тогда, когда хотя бы одно из средних $\mu_x(t)$ или $\mu_y(t)$ равно нулю.

Из предположения стационарности следует, что ковариационные функции $R_{xx}(\tau)$ и $R_{yy}(\tau)$ - чётные функции от τ , т.е.

$$R_{xx}(-\tau) = R_{xx}(\tau) \qquad R_{yy}(-\tau) = R_{yy}(\tau)
 \tag{9.8}$$

В то же время взаимная ковариационная функция удовлетворяет соотношению:

$$R_{xy}(-\tau) = R_{yx}(\tau) \quad (9.9)$$

Доказательство:

По определению: $R_{xy}(-\tau) = E[x(t)y(t-\tau)]$, где зависимость от k опущена для простоты обозначений.

Поскольку результат не зависит от сдвига начала отсчёта, то t можно заменить на $t+\tau$ до взятия математического ожидания. Поэтому:

$$R_{xy}(-\tau) = E[x(t+\tau)y(t+\tau-\tau)] = E[x(t+\tau)y(t)] = R_{yx}(\tau)$$

Что и требовалось доказать.

При $x = y$ справедливо:

$$\left. \begin{aligned} R_{xx}(-\tau) &= R_{xx}(\tau) \\ R_{yy}(-\tau) &= R_{yy}(\tau) \end{aligned} \right\} \quad (9.10)$$

Т.е. эти соотношения – частный случай формулы (9.9).

Корреляционные свойства стационарных случайных процессов $\{x_k(t)\}$ и $\{y_k(t)\}$ описываются четырьмя функциями $R_{xx}(\tau)$, $R_{yy}(\tau)$, $R_{xy}(\tau)$ и $R_{yx}(\tau)$, которые целесообразно вычислять только для значений $\tau \geq 0$, поскольку из соотношений (9.9) и (9.10) легко получить их значения для $\tau < 0$.

В случае ансамбля выборочных функций среднее значение (математическое ожидание) и ковариационная функция определяются как:

$$\mu_x(t_1) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N x_k(t_1)$$

$$R_{xx}(t_1, t_1 + \tau) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N x_k(t_1) x_k(t_1 + \tau) \quad (9.11)$$

То есть среднее значение (первый момент) случайного процесса в момент t_1 можно вычислить, выбрав мгновенные значения всех выборочных функций ансамбля в момент времени t_1 , сложив эти значения и разделив на число выборочных функций.

Аналогично ковариационная функция вычисляется путём усреднения по ансамблю произведений мгновенных значений в моменты времени t_1 и $t_1 + \tau$.



Рис. 9.2. Классификация случайных процессов.

Однако в большинстве случаев, характеристики стационарного случайного процесса (ССП) можно вычислить, усредняя по времени в пределах отдельных выборочных функций, входящих в ансамбль.

В этом случае среднее значение $\mu_x(k)$ и ковариационная функция $R_{xx}(\tau, k)$, вычисленные по k -ой реализации будут равны:

$$\mu_x(k) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x_k(t) dt$$

$$R_{xx}(\tau, k) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x_k(t) x_k(t + \tau) dt \quad (9.12)$$

Случайный процесс называется эргодическим, если он стационарен, а математическое ожидание и ковариационная функция, вычисленные при усреднении по ансамблю реализаций и по времени совпадают. Отметим, что свойством эргодичности могут обладать только стационарные процессы. К счастью, на практике ССП обычно оказываются эргодическими и именно по этой причине свойства ССП можно определить по одной наблюдаемой реализации.

К нестационарным процессам относятся все случайные процессы, не удовлетворяющие условию стационарности, то есть такие процессы, основные характеристики которых зависят от времени.

Идеальный белый шум – это стационарный случайный сигнал, обладающий фундаментальным свойством статистической независимости любых двух его отсчётов. Кроме того, он имеет нулевое среднее значение, а автокорреляционная функция такого сигнала равна

$$R_{xx}(\tau) = a\delta(\tau) \quad (9.13)$$

где $\delta(\tau)$ - дельта-функция Дирака, равная нулю всюду, кроме точки $\tau = 0$, и интеграл от неё равен единице по любому интервалу, содержащему начало координат; коэффициент a - скалярная величина, называемая уровнем мощности или интенсивностью белого шума.

10. ФУНКЦИИ СПЕКТРАЛЬНОЙ ПЛОТНОСТИ И ОПРЕДЕЛЕНИЯ СПЕКТРОВ С ПОМОЩЬЮ КОРРЕЛЯЦИОННЫХ ФУНКЦИЙ

Функции спектральной плотности можно определить:

1. с помощью ковариационной функции
2. с помощью фильтрации, возведения в квадрат и усреднения.

Пусть существуют ковариационные и взаимные ковариационные функции $R_{xx}(\tau)$, $R_{yy}(\tau)$ и $R_{xy}(\tau)$, задаваемые формулами.

$$\begin{aligned}R_{xx}(\tau) &= E[x_k(t)x_k(t+\tau)] \\R_{yy}(\tau) &= E[y_k(t)y_k(t+\tau)] \\R_{xy}(\tau) &= E[x_k(t)y_k(t+\tau)]\end{aligned}\tag{10.1}$$

Предположим далее, что конечны интегралы от их абсолютных величин:

$$\int_{-\infty}^{\infty} |R(\tau)| < \infty\tag{10.2}$$

На практике это условие всегда выполняется и для реализаций конечной длины. Тогда:

$$\begin{aligned}S_{xx}(f) &= \int_{-\infty}^{\infty} R_{xx}(\tau)e^{-i2\pi f\tau} d\tau \\S_{yy}(f) &= \int_{-\infty}^{\infty} R_{yy}(\tau)e^{-i2\pi f\tau} d\tau \\S_{xy}(f) &= \int_{-\infty}^{\infty} R_{xy}(\tau)e^{-i2\pi f\tau} d\tau\end{aligned}\tag{10.3}$$

где $S_{xx}(f)$, $S_{yy}(f)$ называются функциями спектральной плотности мощности процессов $\{x_k(t)\}$ и $\{y_k(t)\}$, а $S_{xy}(f)$ называется взаимной спектральной плотностью мощности процессов $\{x_k(t)\}$ и $\{y_k(t)\}$.

Обратные преобразования дают:

$$\begin{aligned}
 R_{xx}(\tau) &= \int_{-\infty}^{\infty} S_{xx}(f) e^{i2\pi f\tau} df \\
 R_{yy}(\tau) &= \int_{-\infty}^{\infty} S_{yy}(f) e^{i2\pi f\tau} df \\
 R_{xy}(\tau) &= \int_{-\infty}^{\infty} S_{xy}(f) e^{i2\pi f\tau} df
 \end{aligned}
 \tag{10.4}$$

Соотношения называют формулами Винера – Хинчина, которые в начале 30-х годов независимо установили связь между ковариационными функциями и спектральными плотностями через преобразование Фурье.

Из свойств симметрии ковариационных функций следует, что

$$\begin{aligned}
 S_{xx}(-f) &= S_{xx}(f) \\
 S_{yy}(-f) &= S_{yy}(f) \\
 S_{xy}(-f) &= S_{yx}(f)
 \end{aligned}
 \tag{10.5}$$

Следовательно, спектральные плотности $S_{xx}(f)$ и $S_{yy}(f)$ - действительные чётные функции от f , а взаимная спектральная плотность $S_{xy}(f)$ - комплексная функция от f .

По определению:

$$S_{xy}(-f) = \int_{-\infty}^{\infty} R_{xy}(\tau) e^{i2\pi f\tau} d\tau$$

$$S_{xy}^*(f) = \int_{-\infty}^{\infty} R_{xy}(\tau) e^{i2\pi f\tau} d\tau \quad (10.6)$$

То есть $S_{xy}(-f) = S_{xy}^*(f)$.

Выполнив в первом интеграле замену переменных, положив $\tau = -u$, $d\tau = -du$. Тогда:

$$S_{xy}(-f) = \int_{-\infty}^{\infty} R_{xy}(-u) e^{-i2\pi f u} du$$

Так как $R_{xy}(-u) = R_{yx}(u)$, то тогда:

$$S_{xy}(-f) = \int_{-\infty}^{\infty} R_{yx}(u) e^{-i2\pi f u} du = S_{yx}(f)$$

Спектральное соотношение можно преобразовать к виду

$$S_{xx}(f) = \int_{-\infty}^{\infty} R_{xx}(\tau) \cos 2\pi f\tau d\tau = 2 \int_0^{\infty} R_{xx}(\tau) \cos 2\pi f\tau d\tau \quad (10.7)$$

Обратные преобразования имеют вид

$$R_{xx}(\tau) = 2 \int_0^{\infty} S_{xx}(f) \cos 2\pi f\tau df$$

$$R_{yy}(\tau) = 2 \int_0^{\infty} S_{yy}(f) \cos 2\pi f\tau df \quad (10.8)$$

Односторонние спектральные плотности $G_{xx}(f)$, $G_{yy}(f)$, где f изменяется только в пределах интервала $(0, \infty)$ определяется формулами

$$\begin{aligned}
G_{xx}(f) &= 2S_{xx}(f) & 0 \leq f < \infty \\
G_{yy}(f) &= 2S_{yy}(f) & 0 \leq f < \infty,
\end{aligned}
\tag{10.9}$$

а вне указанной области значений f имеем

$$G_{xx} = G_{yy} = 0$$

Односторонние спектральные плотности $G_{xx}(f)$, $G_{yy}(f)$ связаны с ковариационными функциями $R_{xx}(\tau)$, $R_{yy}(\tau)$ соотношениями

$$\begin{aligned}
G_{xx}(f) &= 4 \int_0^{\infty} R_{xx}(\tau) \cos 2\pi f \tau d\tau & 0 \leq f < \infty \\
G_{yy}(f) &= 4 \int_0^{\infty} R_{yy}(\tau) \cos 2\pi f \tau d\tau
\end{aligned}
\tag{10.10}$$

Обратные преобразования имеют вид:

$$\begin{aligned}
R_{xx}(\tau) &= \int_0^{\infty} G_{xx}(f) \cos 2\pi f \tau df \\
R_{yy}(\tau) &= \int_0^{\infty} G_{yy}(f) \cos 2\pi f \tau df
\end{aligned}
\tag{10.11}$$

При $\tau = 0$ получим

$$\begin{aligned}
R_{xx}(0) &= E[x^2(t)] = \psi_x^2 = \int_0^{\infty} G_{xx}(f) df \\
R_{yy}(0) &= E[y^2(t)] = \psi_y^2 = \int_0^{\infty} G_{yy}(f) df
\end{aligned}
\tag{10.12}$$

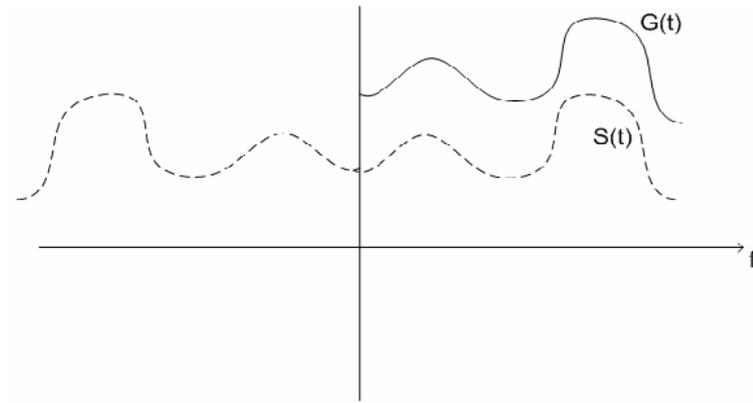


Рис. 10.1

В свою очередь односторонняя взаимная спектральная плотность определяется как:

$$G_{xy}(f) = 2S_{xy}(f) \quad 0 \leq f < \infty$$

$$G_{xy}(f) = 2 \int_0^{\infty} R_{xy}(\tau) e^{-i2\pi f\tau} d\tau = C_{xy}(f) - iQ_{xy}(f) \quad (10.13)$$

где $C_{xy}(f)$ - называется спектральной плотностью (коспектром), а $Q_{xy}(f)$ - квадратурной спектральной плотностью (квадратурный спектр).

Взаимная ковариационная функция определяется как:

$$R_{xy}(\tau) = \int_0^{\infty} [C_{xy}(f) \cos 2\pi f\tau + Q_{xy}(f) \sin 2\pi f\tau] df \quad (10.14)$$

При $\tau = 0$

$$R_{xy}(\tau) = \int_0^{\infty} C_{xy}(f) \cos 2\pi f\tau df . \quad (10.15)$$

Одностороннюю взаимную спектральную плотность можно представить в комплексной полярной форме

$$G_{xy}(f) = |G_{xy}(f)|e^{-i\theta_{xy}(f)} \quad (10.16)$$

где $|G_{xy}(f)| = \sqrt{C_{xy}^2(f) + Q_{xy}^2(f)}$

$$\theta_{xy}(f) = \operatorname{arctg} \frac{Q_{xy}(f)}{C_{xy}(f)}$$

Знаки членов $C_{xy}(f)$ и $Q_{xy}(f)$ могут быть как положительными, так и отрицательными и их сочетание определяет квадрант, в котором расположен фазовый угол.

11. ОПРЕДЕЛЕНИЕ СПЕКТРОВ С ПОМОЩЬЮ ФИЛЬТРАЦИИ

Другой способ вычисления спектральной плотности подразумевает применение аналоговых устройств и включает операции, схематически представленные на рис. 11.1:

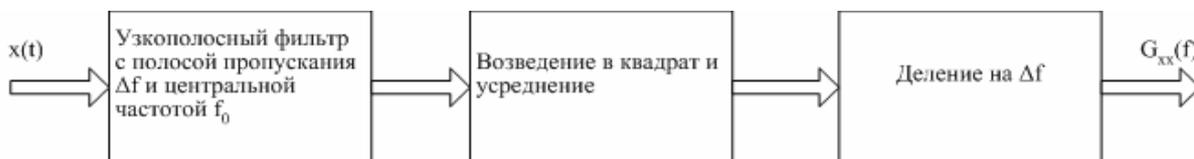


Рис. 11.1. Вычисление спектральной плотности

1. Частотную фильтрацию сигнала $x(t)$ с помощью узкополосного фильтра с полосой пропускания Δf и центральной частотой f_0 . В результате этой операции получается величина $x(f_0, \Delta f, t)$.
2. Возведение в квадрат мгновенного значения отфильтрованного сигнала
3. Усреднение квадрата мгновенного значения по реализации длины T ; в результате этой операции получается оценка среднего квадрата отфильтрованного сигнала
4. Деление на ширину полосы пропускания фильтра Δf , в результате чего получается оценка скорости изменения среднего квадрата в зависимости от центральной частоты f_0 .

Следовательно, оценка СПМ имеет вид:

$$\hat{G}_{xx}(f_0) = \frac{1}{(\Delta f)T} \int_0^T x^2(f_0, \Delta f, t) dt \quad (11.1)$$

Для вычисления взаимной спектральной плотности аналоговыми методами применяется прямое обобщение описанной процедуры на случай двух сигналов $x(t)$, $y(t)$, которая включает следующие операции:

1. раздельная частотная фильтрация обоих сигналов с помощью узкополосных фильтров, имеющие одинаковые полосы пропускания Δf и одну и ту же центральную частоту f_0 : в результате получаются $x(f_0, \Delta f, t)$ и $y(f_0, \Delta f, t)$
2. перемножение мгновенных значений двух отфильтрованных сигналов при отсутствии сдвига фазы между ними с целью получения синфазных членов, необходимых для вычисления коспектра.
3. перемножение мгновенных значений двух отфильтрованных сигналов, причём $y(f_0, \Delta f, t)$ сдвигается по фазе на 90° по отношению к $x(f_0, \Delta f, t)$ с целью получения смещённых по фазе членов, что необходимо для вычисления квадратурного спектра.
4. усреднение каждого из этих произведений по реализации длиной T с целью построения оценок средних произведений оценок синфазных и смещённых по фазе членов.
5. деление каждого из средних произведений на ширину полосы пропускания Δf для получения оценок $\tilde{C}_{xy}(f_0)$ и $\tilde{Q}_{xy}(f_0)$.

Тогда оценка взаимной спектральной плотности имеет вид:

$$\tilde{G}_{xy}(f_0) = \tilde{C}_{xy}(f_0) - i\tilde{Q}_{xy}(f_0), \quad (11.2)$$

где

$$\tilde{C}_{xy}(f_0) = \frac{1}{(\Delta f)T} \int_0^T x(f_0, \Delta f, t)y(f_0, \Delta f, t)dt$$

$$\tilde{Q}_{xy}(f_0) = \frac{1}{(\Delta f)T} \int_0^T x(f_0, \Delta f, t)y^\circ(f_0, \Delta f, t)dt$$

Здесь $y^o(f_0, \Delta f, t)$ - сдвинутый по фазе отфильтрованный сигнал $y(f_0, \Delta f, t)$.

В случае конечного преобразования Фурье

$$G_{xx}(f) = 2 \lim_{T \rightarrow \infty} \frac{1}{T} E[|x_k(f, T)|^2]$$

Величина

$$|x_k(f, T)|^2 = \left[\int_0^T x_k(t) \cos 2\pi f t dt \right]^2 + \left[\int_0^T x_k(t) \sin 2\pi f t dt \right]^2 \quad (11.3)$$

действует как фильтр, выделяющий значение среднего квадрата $x_k(t)$ в узкой полосе частот, окружающей центральную частоту f .

Чтобы доказать это, определим для любой реализации $x_k(t)$ процесса $\{x_k(t)\}$ величину

$$x_k(t, T) = \begin{cases} x_k(t), & 0 \leq t \leq T \\ 0, & \text{в остальных случаях} \end{cases} \quad (11.4)$$

Тогда средний квадрат конкретной реализации $x_k(t)$ можно определить как:

$$\psi_x^2(k) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x_k^2(t) dt = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} x_k^2(t, T) dt \quad (11.5)$$

$$\text{Согласно теореме Парсеваля} \quad \int_{-\infty}^{\infty} f^2(t) dt = \int_{-\infty}^{\infty} |F(f)|^2 df. \quad (11.6)$$

12. ЦИФРОВЫЕ МЕТОДЫ АНАЛИЗА: ОЦЕНИВАНИЕ КОРЕЛЯЦИОННЫХ ФУНКЦИЙ МЕТОДОМ БПФ, ОЦЕНКИ СПЕКТРАЛЬНОЙ ПЛОТНОСТИ МОЩНОСТИ И АЛГОРИТМЫ ВЫЧИСЛЕНИЯ

Для определения оценок спектральной плотности мощности используют корреляционный метод, основанный на использовании соотношения

$$\hat{G}(f) = 4 \int_0^T \tilde{R}(\tau) \cos 2\pi f \tau d\tau \quad (12.1)$$

То есть в начале с использованием (12.1) рассчитывалась оценка ковариационной функции, а затем оценка спектральной плотности мощности вычислялась как преобразование Фурье при соответствующем значении максимального сдвига.

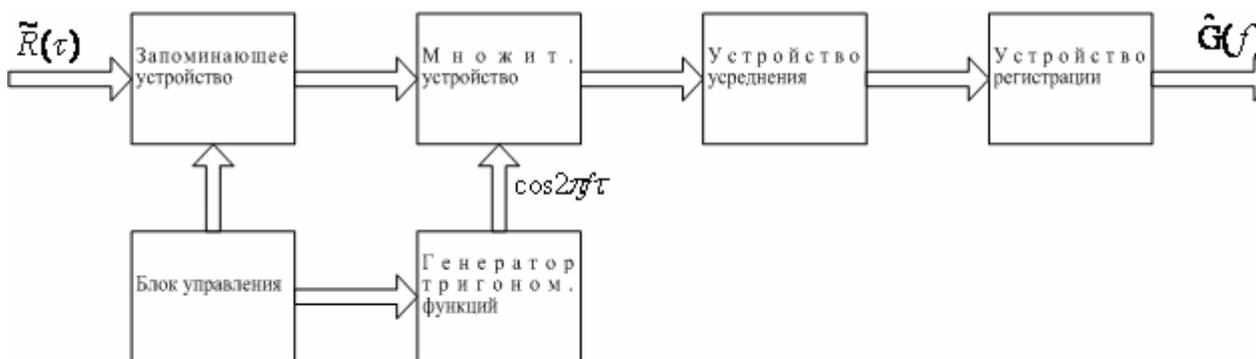


Рис. 12.1.

Исходная оценка функции ковариации вводится в ЗУ, с помощью которого она периодически воспроизводится и анализируется затем на заданных частотах с помощью соотношения (12.1). Синхронизация воспроизведения функции $\tilde{R}(\tau)$ с фазой выходных колебаний генератора

производится с помощью блока управления и получаемые значения СПМ фиксируются в устройстве регистрации.

Разобьём реализацию на n_d смежных отрезков длиной T каждый и обозначим эти отрезки через x_i :

$$(i-1)T \leq t \leq iT \quad i = \overline{1, n_d}$$

Тогда оценка спектральной плотности мощности на произвольной частоте f может быть записана как:

$$\tilde{S}_{xx}(f) = \frac{1}{n_d T} \sum_{i=1}^{n_d} |x_i(f, T)|^2 \quad (12.2)$$

где $X_i(f, T) = \int_0^T x_i(t) e^{-i2\pi ft} dt$

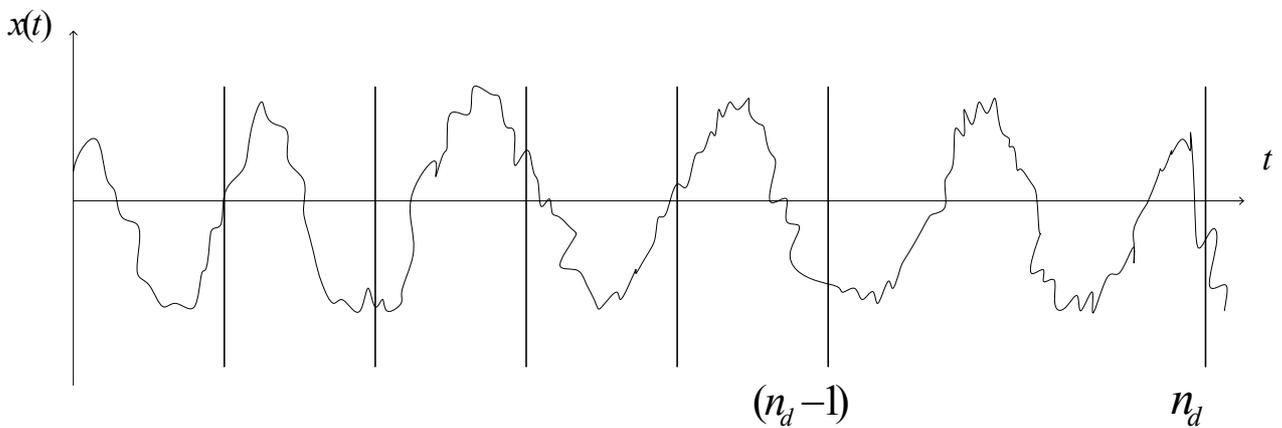


Рис. 12.2

При дискретном временном параметре каждая реализация $x_i(t)$ будет представлена N значениями временного ряда $\{x_{in}\}$, $n = \overline{0, N-1}$. Конечное

преобразование Фурье даст соответствующее соотношению (12.2) значения спектральной плотности мощности на дискретных частотах

$$f_R = \frac{k}{T} = \frac{k}{N\Delta t} \quad k = \overline{0, N-1}$$

При этом коэффициенты Фурье для каждого отрезка определяются как

$$x_i(f_k) = \Delta t \sum_{n=0}^{N-1} x_{in} \exp\left(\frac{-j2\pi kn}{N}\right) \quad (12.3)$$

При этом оценка двусторонней спектральной плотности принимает вид

$$\hat{S}_{xx}(f_k) = \frac{1}{n_d N \Delta t} \sum_{i=1}^{n_d} |x_i(f_k)|^2 \quad k = \overline{0, N-1} \quad (12.4)$$

В свою очередь, односторонняя функция СПМ оценивается как:

$$\hat{G}_{xx}(f_k) = \frac{2}{n_d N \Delta t} \sum_{i=1}^{n_d} |x_i(f_k)|^2 \quad k = \overline{0, N/2} \quad (12.5)$$

Число отсчётов N в соотношении для СПМ при каждом БПФ называют размером блока и оно служит ключевым параметром, определяющим разрешающую способность по частоте

$$\Delta f = \frac{1}{T} = \frac{1}{N\Delta t}$$

Кроме того, при осуществлении БПФ по алгоритму Кули – Тьюки размер выборки удобно выбирать как целую степень числа 2, т.е. $N = 2^p$.

Определённое формулой (12.3) конечное преобразование Фурье функции $x(t)$ можно рассматривать как преобразование Фурье заданной на бесконечном интервале функции $v(t)$, умноженной на прямоугольное «временное окно» $u(t)$, имеющее вид:

$$u(t) = \begin{cases} 1, & 0 \leq t \leq T \\ 0, & \text{в остальных случаях} \end{cases}$$

Иными словами реализацию $x(t)$ можно рассматривать как произведение $x(t) = u(t)v(t)$.

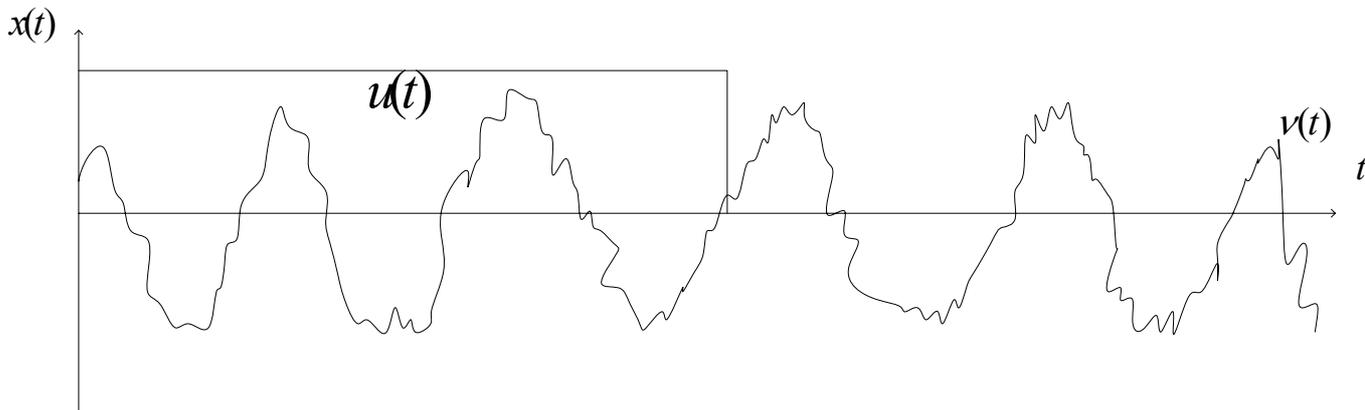


Рис. 12.3

Следовательно, преобразование Фурье функции $x(t)$ есть свёртка преобразования функций $u(t)$ и $v(t)$:

$$x(f) = \int_{-\infty}^{\infty} u(\alpha)v(f - \alpha)d\alpha$$

В свою очередь, преобразование Фурье прямоугольной функции $u(t)$ будет иметь вид:

$$u(t) = T \left(\frac{\sin \pi f T}{\pi f T} \right) e^{-i \pi f T}$$

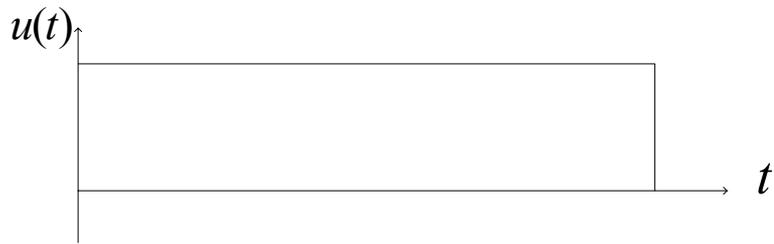


Рис. 12.4

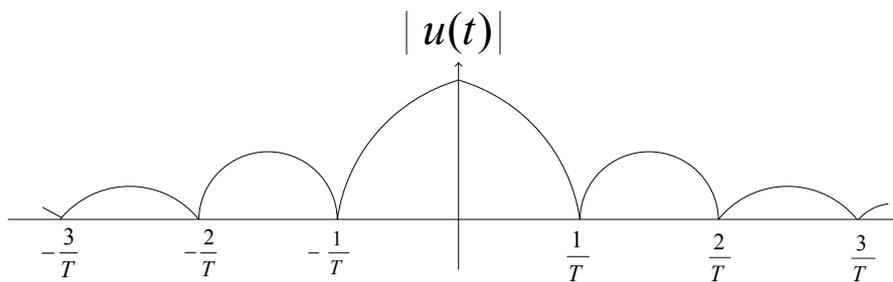


Рис. 12.5

Отметим, что описанная функция представляет собой основное спектральное окно, используемое при анализе. Большие боковые максимумы функции $|u(t)|$ обуславливают большое просачивание энергии с частот, далёких от главного максимума спектрального окна и могут привести к сильным искажениям спектральной оценки, особенно заметным для процессов с узкополосным спектром. Проблема просачивания не возникает при анализе периодических процессов (с периодом T_p).

Для подавления просачивания обычно используют временные окна, сглаживающие исходную реализацию таким образом, чтобы подавить резкие вариации на её начальном и конечном участках.

Существует множество таких окон, но одно из наиболее ранних и широко используемых – косинусоидальное сглаживающее окно, называемое окном Ханна.

$$U_h(t) = \begin{cases} \frac{1}{2} \left(1 - \cos \frac{2\pi t}{T}\right) = 1 - \cos^2\left(\frac{\pi t}{T}\right), & 0 \leq t \leq T \\ 0, & \text{в остальных случаях} \end{cases}$$

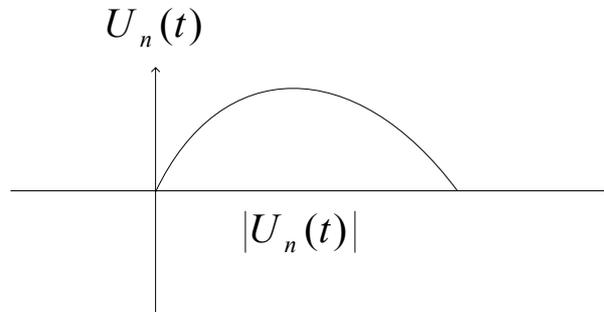


Рис. 12.6

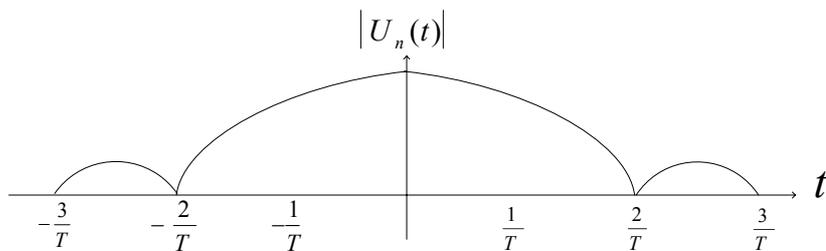


Рис. 12.7

В этом случае $x(t) = u_h(t)v(t)$.

Её преобразование Фурье есть

$$x(f) = \int_0^T x(t) e^{-2\pi f t} dt = \int_{-\infty}^{\infty} u_h(\alpha) v(\alpha) d\alpha$$

Для дискретных значений частоты $f_k = k/T$ при $k = \overline{0, N/2}$ получим:

$$x(f_k) = \frac{1}{2} v(f_k) - \frac{1}{4} v(f_{k-1}) - \frac{1}{4} v(f_{k+1})$$

$$\text{где } v(f_k) = \int_0^T v(t) e^{-i2\pi k t/T} dt$$

Предположим, что в пределах каждой полосы частот шириной $\Delta f = 1/T$ функция $v(t)$ ведёт себя как ограниченный по частоте белый шум. Тогда математическое ожидание произведения $V^*(f)$ и $V(\zeta)$, вычисленное для любых частот f и ζ из набора $k\Delta f = k/T$

$$E[V^*(f)V(\zeta)] = \begin{cases} 0, & f \neq \zeta \\ 1, & f = \zeta \end{cases}$$

С учётом этого свойства

$$E[|x(f_k)|^2] = \left(\frac{1}{2}\right)^2 + \left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^2 = \frac{3}{8}$$

при любых $f_k = k/T$.

Эта величина характеризует потери, вызванные использованием спектрального окна Ханна для оценки спектральной плотности при помощи конечного преобразования Фурье. Поэтому при оценивании спектров необходимо умножение на масштабный множитель $\sqrt{8/3}$. В частности при вычислении спектральной плотности по формуле (12.3) с использованием окна Ханна следует использовать соотношение

$$x_i(f_k) = \Delta t \sqrt{\frac{8}{3}} \sum_{n=0}^{N-1} x_{in} \left(1 - \cos^2 \frac{\pi n}{N}\right) \exp\left[-j \frac{2\pi k n}{N}\right]$$

Для вычисления оценки СПМ необходимо использовать следующие этапы:

1. Имеющаяся реализация x_n разбивается на n_d отрезков, каждый из которых содержит N отсчётов.
2. При необходимости подавить просачивание энергии через боковые максимумы каждый отрезок $\{x_n\} \quad n = \overline{0, N-1}$ сглаживается фильтром, или весовым окном Ханна, или другим подходящим образом выбранным фильтром (к примеру, Баттерворта, Сорта, Хэмминга, Дольфа-Чебышева и т.д).
3. Для каждого отрезка реализации выбирается N значений БПФ $x(f_k)$ $k = \overline{0, N-1}$.
4. Значения $x(f_k)$ умножаются на масштабный коэффициент для компенсации изменений, внесённых сглаживанием (при использовании окна Ханна этот коэффициент равен $\sqrt{8/3}$).
5. По формулам (12.4) и (12.5) вычисляется соответственно оценка двусторонней и односторонней спектральной плотности для n_d отрезков.

Косвенный способ оценивания ковариационной функции основан на использовании соотношения Винера – Хинчина. Оценка ковариационной функции находится как обратное преобразование Фурье спектральной плотности.

В случае дискретной последовательности $\{x_n\} \quad n = \overline{0, N-1}$ циклическая ковариационная функция принимает вид:

$$\hat{R}_{xx}^c(r\Delta t) = \frac{N-r}{N} \hat{R}_{xx}(r\Delta t) + \frac{r}{N} \hat{R}_{xx}[(N-r)\Delta t]$$

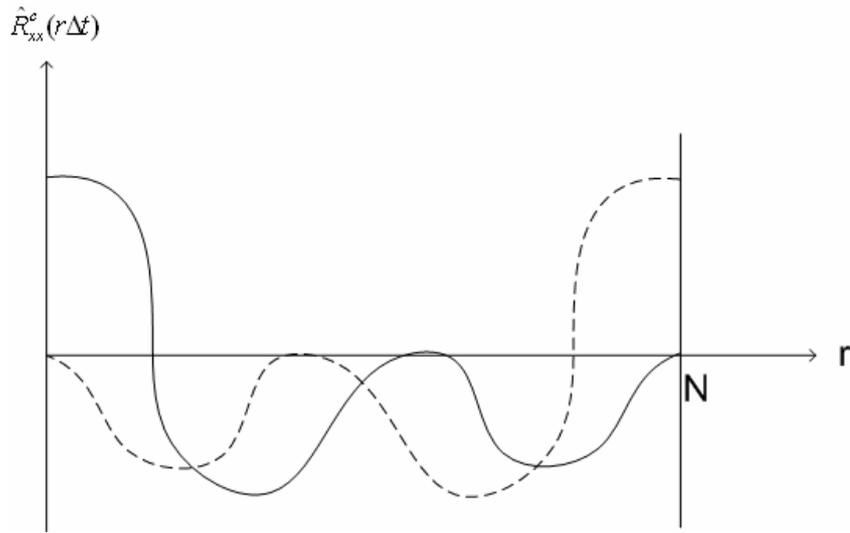
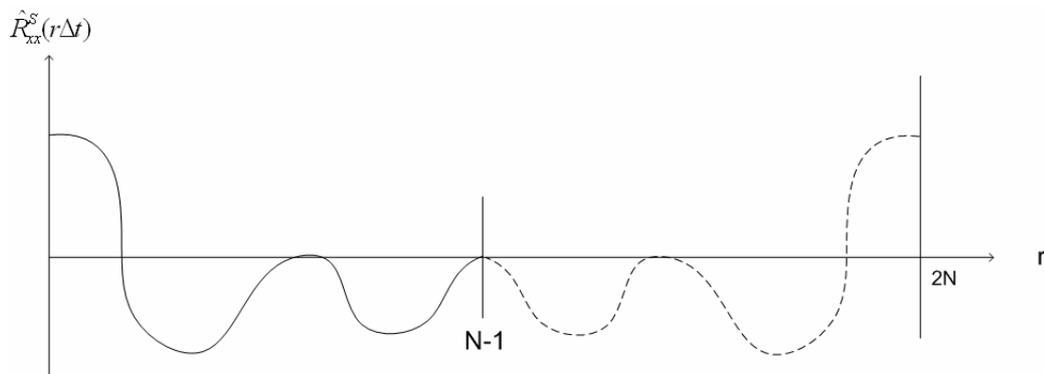


Рис. 12.8

Как видно из рисунка, имеет место эффект наложения. Этого осложнения можно избежать, дополняя исходный временной ряд нулями. В результате этого происходит разделение обеих частей циклической оценки ковариационной функции.

В частности, при добавлении N нулей к исходному временному ряду той же длины будет иметь место полное разделение, т.е.

$$\hat{R}_{xx}^S = \begin{cases} \frac{N-r}{N} \hat{R}_{xx}(r\Delta t), & r = 0, 1, \dots, (N-1) \\ \frac{r-N}{N} \hat{R}_{xx}[(2N-r)\Delta t], & r = N, (N+1), \dots, (2N-1) \end{cases}$$



Заметим, что первая половина этой оценки при $0 \leq r \leq N-1$ представляет собой значения ковариационной функции при положительных значениях сдвига ($0 \leq r \leq m$), а вторая ($N \leq r \leq 2N-1$) при отрицательных значениях сдвига. Но поскольку ковариационная функция есть функция чётная, то вторую половину можно отбросить и окончательно несмещённая оценка ковариационной функции находится в виде

$$\hat{R}_{xx}(r\Delta t) = \frac{N}{N-r} \hat{R}_{xx}^s(r\Delta t)$$

Полученная оценка статистически эквивалентна прямой оценке (непосредственно по реализации случайного процесса). Однако, в зависимости от величины её такой косвенный метод с использованием БПФ может потребовать заметно меньшего объёма вычислений. Действительно, косвенный метод требует нахождения вначале оценки спектральной плотности мощности, для чего осуществляется БПФ по n_d независимым реализациям, каждая из которых содержит N отсчётов и дополнена N нулями, так что БПФ производится для реализации длиной $2N$. Затем $2N$ значений сглаженной спектральной оценки подвергаются обратному преобразованию Фурье, что в общей сложности даёт (n_d+1) быстрых преобразований Фурье, каждое из которых требует $4Np$ операций над действительными числами, где $N = \log_2 N$. Для реализации прямого метода потребуется $m n_d N$ операций. Следовательно, при одном и том же максимальном сдвиге $m=N$ коэффициент ускорения вычислений равен:

$$к.у.в. = \frac{n_d N^2}{(n_d + 1)4Np} \approx \frac{N}{4p}$$

$$\text{При } N=1024 \text{ к.у.в.} = \frac{1024}{40} \approx 26.$$

Оценки взаимных корреляционных функций так же, как и оценки автоковариационных функций могут быть получены двумя основными способами: непосредственным вычислением и с использованием БПФ.

- Непосредственное вычисление.

В этом случае несмещённые оценки взаимных ковариационных функций при сдвигах $r = \overline{0, m}$ ($m < N$) находятся в виде

$$\hat{R}_{xy}(r\Delta t) = \frac{1}{N-r} \sum_{n=0}^{N-r-1} x_n y_{n+r}$$

$$\hat{R}_{yx}(r\Delta t) = \frac{1}{N-r} \sum_{n=0}^{N-r-1} y_n x_{n+r}$$

Выборочную взаимную ковариационную функцию $\hat{R}_{xy}(r\Delta t)$ можно нормировать так? что значения её будут находиться между -1 и +1, для чего необходимо разделить её на величину $\sqrt{R_{xx}(0)}\sqrt{R_{yy}(0)}$. Таким образом, получается нормированная взаимная корреляционная функция:

$$\hat{\rho}_{xy}(r\Delta t) = \frac{\hat{R}_{xy}(r\Delta t)}{\sqrt{R_{xx}(0)}\sqrt{R_{yy}(0)}} \quad r = \overline{0, m} ?$$

которая теоретически должна удовлетворять равенству $-1 \leq \hat{\rho}_{xy}(r\Delta t) \leq 1$.

Аналогично определяется функция $\hat{\rho}_{yx}(r\Delta t)$.

- Для получения взаимной ковариационной функции при помощи БПФ можно рекомендовать следующую последовательность вычислений. Пусть объём каждой выборки из реализаций $x(t)$ и $y(t)$ составляет $N = 2^p$ чисел. Взаимную ковариационную функцию в этом варианте находят по спектральной мощности процесса, для чего дважды применяют БПФ – к функции $x(t)$ и к функции $y(t)$. Оба преобразования

можно получить одновременно, приравняв одну реализацию действительной, а другую мнимой части некоторой комплексной реализации:

$$z(n) = x(n) + jy(n) \quad n = 0, 1, \dots, N-1.$$

Тогда преобразование Фурье функции $z(n)$ будет иметь вид:

$$Z(k) = \sum_{n=0}^{N-1} [x(n) + jy(n)] \exp\left[-j \frac{2\pi kn}{N}\right]$$

и может быть вычислено по алгоритму БПФ. Следует отметить, что

$$\exp\left[j \frac{2\pi m(N-k)}{N}\right] = \exp\left[-j \frac{2\pi mk}{N}\right]$$

Поскольку $\exp(j2\pi m) \equiv 1$ при любых n .

Следовательно, комплексно сопряжённая функция равна:

$$Z^*(N-k) = \sum_{n=0}^{N-1} [x(n) - jy(n)] \exp\left[-j \frac{2\pi nk}{N}\right]$$

Тогда

$$Z(k) + Z^*(N-k) = 2 \sum_{n=0}^{N-1} x(n) \exp\left[-j \frac{2\pi nk}{N}\right] = 2X(k)$$

$$Z(k) - Z^*(N-k) = 2j \sum_{n=0}^{N-1} y(n) \exp\left[-j \frac{2\pi nk}{N}\right] = 2jY(k)$$

Таким образом, преобразования Фурье $X(k)$ и $Y(k)$ двух действительных реализаций $x(n)$ и $y(n)$ будут иметь вид:

$$X(k) = \frac{Z(k) + Z^*(N-k)}{2} \quad Y(k) = \frac{Z(k) - Z^*(N-k)}{2j}$$

Этим же приёмом можно воспользоваться для расчёта преобразования на интервале вдвое большей длины. Для этого вещественная реализация $v(n)$ ($n = 0, 1, \dots, 2N-1$) разделяется на две части: одна состоит из значений $v(n)$ при чётном n , другая – при нечётном. Иными словами:

$$x(n) = v(2n) \quad y(n) = v(2n+1) \quad n = \overline{0, N-1}$$

Рекомендуемая в итоге последовательность операций для получения оценки взаимной ковариационной функции основана на предположении, что длина исходных реализаций $x(n\Delta t)$, $y(n\Delta t)$ есть $n_d N$, $N = 2^p$.

1. Выбирается максимальный сдвиг m и исходные реализации разбиваются на n_d отрезков, каждый из которых состоит из $N > m$ отсчётов.
2. Каждый отрезок $\{x_n\}$, $\{y_n\}$, $n = \overline{1, N}$ дополняется N нулями, в результате чего новый отрезок содержит $2N$ отсчётов.
3. По отрезкам, содержащим по $2N$ точек в соответствии с

$$Z(k) = \sum_{n=0}^{2N-1} [x(n) + jy(n)] \exp\left[-j \frac{2\pi kn}{2N}\right] \quad k = \overline{0, 2N-1}$$

Методом БПФ вычисляются коэффициенты $Z(k)$.

4. На основе соотношений

$$X(k) = \frac{Z(k) + Z^*(N-k)}{2} \quad Y(k) = \frac{Z(k) - Z^*(N-k)}{2j}$$

$$k = \overline{0, 2N-1}$$

Вычисляются $X(k)$ и $Y(k)$.

5. По методу оценивания взаимной спектральной плотности мощности вычисляется оценка $\hat{S}_{xy}(f_k)$ двухсторонней спектральной плотности.

6. На основе соотношения

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} x_k \exp \left[j \frac{2\pi kn}{N} \right]$$

вычисляется обратное БПФ последовательности $\hat{S}_{xy}(f_k)$, дающее оценку $\hat{R}_{xy}^S(r\Delta t)$, $r = \overline{0, 2N-1}$.

7. Несмещённая оценка $\hat{R}_{xy}(r\Delta t)$ взаимной ковариационной функции при положительном сдвиге $r = \overline{0, N-1}$ находится умножением $\hat{R}_{xy}^S(r\Delta t)$ на масштабные коэффициенты $N/(N-r)$.

8. Несмещённая оценка $\hat{R}_{xy}(r\Delta t)$ взаимной ковариационной функции при отрицательном сдвиге $r = \overline{0, N-1}$ находится умножением $\hat{R}_{xy}^S(r\Delta t)$ на масштабные коэффициенты $N/(r-N)$ ($r = \overline{(N+1), 2N}$).

По аналогии с оцениванием автоспектров можно рассматривать следующие этапы оценивания функций взаимной спектральной плотности:

1. Имеющиеся реализации разделяются на n_d отрезков x_n, y_n , каждый из которых содержит N отсчётов.
2. При необходимости подавить просачивание энергии через боковые максимумы каждый отрезок x_n, y_n ($n = \overline{0, N-1}$) сглаживается фильтром Ханна или любым другим подходящим образом выбранным фильтром.

3. Сглаженные значения x_n запоминаются как действительная часть, а сглаженные значения y_n - как мнимая часть последовательности $z(n) = x(n) + jy(n)$ ($n = \overline{0, N-1}$).

4. По формуле

$$Z(k) = \sum_{n=0}^{2N-1} [x(n) + jy(n)] \exp\left[-j \frac{2\pi kn}{2N}\right] \quad k = \overline{0, N-1}$$

для каждого отрезка вычисляются N значений БПФ $Z(k)$.

5. По формулам

$$\begin{cases} X(k) = \frac{Z(k) + Z^*(N-k)}{2} \\ Y(k) = \frac{Z(k) - Z^*(N-k)}{2j} \end{cases} \quad k = \overline{0, N-1}$$

для каждого отрезка вычисляются коэффициенты $X(k)$ и $Y(k)$, $k = \overline{0, N-1}$.

6. Значения $X(k)$ и $Y(k)$ умножаются на масштабный коэффициент для компенсации внесённых сглаживанием изменений (при использовании окна Ханна этот коэффициент равен $\sqrt{8/3}$).

7. Для каждой пары отрезков по значениям $X^*(f_k) = \Delta t X^*(k)$, $Y(f_k) = \Delta t Y(k)$ вычисляется оценка двухсторонней взаимной спектральной плотности

$$\tilde{S}_{xy}(f_k) = \frac{1}{N\Delta t} [X^*(f_k)Y(f_k)] \quad k = \overline{0, N-1}$$

или односторонней взаимной спектральной плотности

$$\hat{G}_{xy}(f_k) = \frac{2}{N\Delta t} [X^*(f_k)Y(f_k)] \quad k = \overline{0, N/2}$$

8. Осреднение оценок по n_d парам отрезков даёт окончательную сглаженную оценку $\hat{S}_{xy}(f_k)$ ($k = \overline{0, N-1}$) или $\hat{G}_{xy}(f_k)$ ($k = \overline{0, N/2}$).

13. ПОЛУЧЕНИЕ ИЗОБРАЖЕНИЯ. ОСНОВНЫЕ ПОНЯТИЯ ПРЕДВАРИТЕЛЬНОЙ ОБРАБОТКИ ИЗОБРАЖЕНИЙ

Визуальная информация преобразуется в электрические сигналы с помощью видеодатчиков. После пространственной дискретизации и квантования по амплитуде эти сигналы дают цифровое изображение.

Основными устройствами, используемыми в системах компьютерного видения, являются телевизионные камеры, состоящие из электронной трубки или твердотельного чувствительного элемента с соответствующими электронными блоками. Твердотельные чувствительные элементы представляются, как правило, приборами с зарядовой связью (ПЗС). Они обладают рядом преимуществ перед электронными трубками (например, имеют меньшие вес и размеры, больший ресурс и малую энергоемкость). Однако, разрешающая способность некоторых трубок пока выше возможности твердотельных камер.

Трубка видикон представляет собой цилиндрическую стеклянную оболочку, содержащую с одного конца электронную пушку, а с другого - экран и мишень. Электронный луч фокусируется и отклоняется с помощью напряжения, прикладываемого к катушкам. Отклоняющий контур обеспечивает сканирование луча по внутренней поверхности мишени для считывания изображения. В процессе сканирования в металлическом слое образуется ток, который подается на выводы трубки. Величина тока пропорциональна числу перемещающихся электронов и следовательно интенсивности светового потока в соответствующем месте. Это изменение тока во время сканирования электронного луча после его обработки в электронном блоке камеры формирует видеосигнал, пропорциональный интенсивности входного изображения. Причем, электронный луч, сканирует по всей поверхности мишени 30 раз в 1 секунду. Полный объем сканирования (называемый кадром) состоит из 525 линий, 480 из которых содержат информацию об изображении.

При рассмотрении устройств на ПЗС удобно подразделить датчики на два типа: датчики линейного сканирования и датчики с плоскостной структурой. Основной частью ПЗС-датчиков является ряд кремниевых чувствительных элементов, называемых фотоячейками. Фотоны от отображаемого объекта проходят через входную прозрачную полукристаллическую кремниевую структуру и образуют пары электрон-дырка. Полученные фотоэлектроны собираются на фотоячейках, при этом величина заряда на каждой фотоячейке пропорциональна соответствующей интенсивности светового потока.

ПЗС-датчики с плоскостной структурой аналогичны датчикам линейного сканирования, с тем отличием, что в них фотоячейки расположены в форме матрицы, а между рядами фотоячеек имеется комбинация переходных регистров. Разрешающая способность датчиков с плоскостной структурой составляет 32×32 элементов на нижней границе и 256×256 элементов для среднего диапазона.

Современные ПЗС-датчики имеют 1024×1024 элементов и больше.

Обозначим через $f(x,y)$ двумерное изображение, получаемое телевизионной камерой или ПЗС-датчиком, дающим изображение:

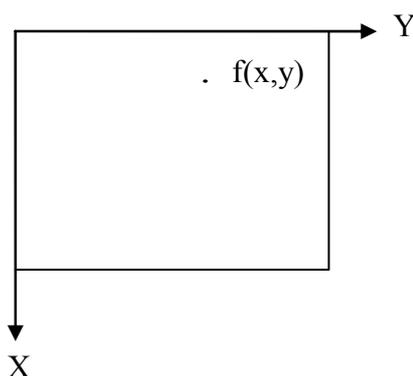


Рис. 13.1

Здесь x и y – пространственные координаты, т. е. координаты плоскости изображения, а величина f в произвольной точке (x,y) пропорциональна яркости (интенсивности) изображения в этой точке.

Для получения удобной с точки зрения вычисления формы, функции изображения $f(x,y)$, ее необходимо дискретизировать как пространственно, так и по амплитуде (интенсивности). Дискретизацию по пространственным координатам (x,y) будем называть дискретизацией изображения, а амплитудную дискретизацию – квантованием по интенсивности или квантованием по уровню серого. Последний термин означает тот факт, что изображение меняется по тону от черного до белого в серых оттенках.

Предположим, что непрерывное изображение дискретизировано равномерно на N строк и M столбцов, причем каждая дискретная величина проквантована по интенсивности. Такая система называется цифровым изображением и может быть представлена в виде:

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0, M-1) \\ f(1,0) & f(1,1) & \dots & f(1, M-1) \\ \dots & \dots & \dots & \dots \\ f(N-1,0) & f(N-1,1) & \dots & f(N-1, M-1) \end{bmatrix}$$

где x и y – дискретные переменные:

$$x = 0, 1, \dots, N-1; y = 0, 1, \dots, M-1.$$

Каждый элемент системы называется элементом изображения, элементом картинки или пикселом.

В соответствии с рис.13.1 $f(0,0)$ является пикселем начала координат изображения, $f(0,1)$ – правый от него пиксел, и так далее.

Обычно на практике величины N , M и число уровней дискретной интенсивности каждого квантованного пиксела задают в виде целых степеней числа 2.

Степень дискретизации и число уровней интенсивности, необходимые для получения требуемого в системе изображения объекта зависят от самого изображения и от вида его использования.

Необходимо отметить, что для получения качественной черно-белой телевизионной картинки требуется 512×512 пикселей со 128 уровнями интенсивности. Как правило, универсальная система компьютерного видения должна иметь, как минимум, разрешающую способность порядка 256×256 пикселей с 64 уровнями интенсивности.

14. ОСНОВНЫЕ ПРЕОБРАЗОВАНИЯ ИЗОБРАЖЕНИЯ

Основными преобразованиями при обработке изображений являются вращение, изменение масштаба и смещение (сдвиг) изображения. Все преобразования записываются в трехмерной декартовой системе координат, в которой координаты точки записываются как (X, Y, Z) . В случае двумерных изображений для обозначения координат пиксела используется сокращенная запись (x, y) .

Смещение (сдвиг) изображения

Предположим, что требуется сместить точку с координатами (X, Y, Z) в новое место, используя перемещения (X_0, Y_0, Z_0) . Смещение выполняется в соответствии с соотношениями:

$$X^* = X + X_0, \quad Y^* = Y + Y_0, \quad Z^* = Z + Z_0,$$

где X^*, Y^*, Z^* - координаты новой точки.

В матричном виде это можно записать как $V^* = TV$, то есть:

$$\begin{bmatrix} X^* \\ Y^* \\ Z^* \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & X_0 \\ 0 & 1 & 0 & Y_0 \\ 0 & 0 & 1 & Z_0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & X_0 \\ 0 & 1 & 0 & Y_0 \\ 0 & 0 & 1 & Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Изменение масштаба изображения

Масштабирование на коэффициенты S_x, S_y, S_z по осям X, Y, Z производится с помощью матрицы преобразования:

$$S = \begin{bmatrix} S_x & 0 & 0 & 0 \\ 0 & S_y & 0 & 0 \\ 0 & 0 & S_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Вращение изображения

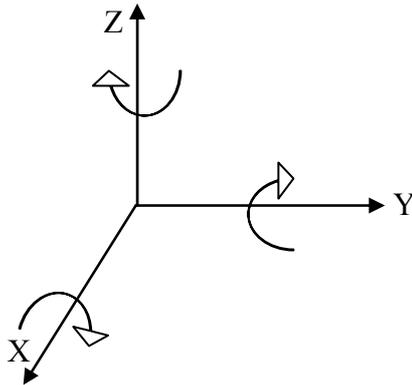


Рис. 13.2

Вращение точки относительно координатной оси Z на угол θ реализуется с помощью преобразования

$$R\theta = \begin{bmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Вращение точки вокруг оси X на угол α выполняется с помощью преобразования

$$R\alpha = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha & 0 \\ 0 & -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Вращение точки вокруг оси Y на угол β реализуется с помощью преобразования

$$R\beta = \begin{bmatrix} \cos \beta & 0 & -\sin \beta & 0 \\ 0 & 1 & 0 & 0 \\ \sin \beta & 0 & \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

15. ОСНОВНЫЕ ВЗАИМОСВЯЗИ МЕЖДУ ПИКСЕЛАМИ ИЗОБРАЖЕНИЯ. МЕТРИЧЕСКИЕ СВОЙСТВА ИЗОБРАЖЕНИЯ

Основные взаимосвязи между пикселями изображения

Пиксел p с координатами (x, y) имеет четыре горизонтальных и вертикальных соседних пиксела с координатами:

$$(x+1, y), (x-1, y), (x, y+1), (x, y-1)$$

Эта группа пикселей, называемая четыре соседа p , обозначается через $N_4(p)$. Четыре диагональных соседних пиксела по отношению к пикселу p имеют координаты

$$(x+1, y+1), (x+1, y-1), (x-1, y+1), (x-1, y-1)$$

и обозначаются через $N_d(p)$. Эти четыре точки вместе с четырьмя указанными выше называются восемь соседей p и обозначаются через $N_8(p)$. Некоторые из точек $N_4(p)$, $N_d(p)$, $N_8(p)$ могут выходить за пределы изображения, если (x, y) находится на границе изображения.

Рассмотрим три типа связей пикселей:

1. Четырехсвязный. Два пиксела p и q с определенными значениями интенсивности являются четырехсвязными, если q относится к группе $N_4(p)$.
2. Восьмисвязный. Два пиксела p и q со значениями интенсивности из V , где V -ряд значений интенсивности, являются восьмисвязными, если q относится к группе $N_8(p)$.
3. m -связный(смешанная связь). Два пиксела p и q со значениями интенсивностей из V являются m -связными, если:
 - а) q относится к группе $N_4(p)$

б) q относится к группе $N_d(p)$ и множество $N_4(p) \cap N_4(q)$ – пустое.

Иными словами, это множество пикселей, являющихся четырьмя соседними, как по отношению к p , так и по отношению к q со значениями интенсивности из V .

Метрические свойства для изображений

Расстояния между двумя векторами X и Y размерностью $n \times 1$ удовлетворяют трем аксиомам:

- 1) $d(X, Y) = 0$, если $X = Y$
- 2) $d(X, Y) = d(Y, X)$
- 3) $d(X, Y) \leq d(X, Z) + d(Z, Y)$, для всех Z

Наиболее важной метрикой является l^p -метрика

$$l^p(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^p \right)^{1/p} \quad \text{для } p \geq 1$$

Если $p=2$, мы получим классическую Евклидову метрику.

Определим для пикселей p, q, z соответственно с координатами (x, y) , (s, t) и (u, v) функцию расстояния или метрику D следующим образом:

- 1) $D(p, q) \geq 0$ [$D(p, q) = 0$, если $p = q$]
- 2) $D(p, q) = D(q, p)$ (симметрия)
- 3) $D(p, z) \leq D(p, q) + D(q, z)$ (неравенство треугольника)

Евклидово расстояние между p и q определяется из соотношения:

$$D_E(p, q) = [(x-s)^2 + (y-t)^2]^{1/2}$$

При измерении этого расстояния пиксели, имеющие расстояние? меньшее или равное некоторой величине r от (x,y) располагаются в окружности радиусом r с центром в (x,y) .

Расстояние D_4 , называемое модульным между p и q , определяется соотношением:

$$D_4(p,q) = |x-s| + |y-t|$$

Метрику D_4 иногда называют метрикой городских кварталов, из-за того, что расстояния измеряются параллельно осям координат.

Покажем, что функция D_4 удовлетворяет к примеру, условию 3.

Запишем соотношение:

$$D_4(p,q) + D_4(q,z) = |x-s| + |y-t| + |s-u| + |t-v|$$

Но, согласно неравенству треугольника для вещественных чисел будем иметь:

$$|x-s| + |s-u| \geq |x-u|,$$

$$|y-t| + |t-v| \geq |y-v|,$$

$$D_4(p,q) + D_4(q,z) \geq |x-u| + |y-v| \geq D_4(p,z), \text{ т.е.}$$

$$D_4(p,z) \leq D_4(p,q) + D_4(q,z), \text{ что и требовалось доказать.}$$

В этом случае пиксели, имеющие расстояние D_4 , меньшее или равное некоторой величине r от (x,y) образуют ромбовидную структуру с центром в (x,y) (центральная точка). Например, пиксели с расстоянием $D_4 < 2$ от (x,y) образуют следующие контуры с равными от центра расстояниями

```

      2
    2 1 2
  2 1 0 1 2
    2 1 2
      2

```

Расстояние D_8 , называемое также шахматным между p и q определяется по формуле:

$$D_8(p,q) = \max \{ |x-s|, |y-t| \}$$

При этом пиксели с расстоянием, меньшим или равным некоторой величине r образуют квадрат с центром в (x,y) . Например, пиксели с расстоянием $D_8 \leq 2$ от центральной точки образуют следующие равноудаленные от центра контуры:

```

    2 2 2 2 2
  2 1 1 1 2
2 1 0 1 2
  2 1 1 1 2
    2 2 2 2 2

```

Геометрические места точек, удаленных на единичное расстояние от начала координат, построенные для различных метрик будут иметь вид:

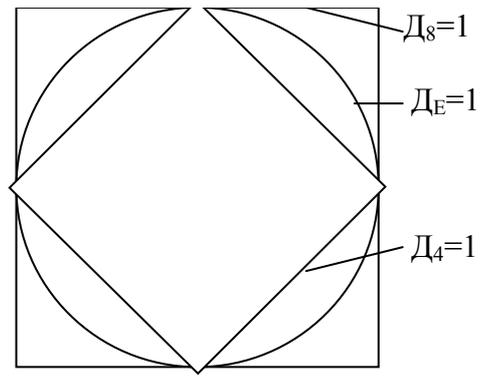


Рис. 15.1

16. МЕТОДЫ ПРОСТРАНСТВЕННОЙ И ЧАСТОТНОЙ ОБЛАСТИ ПРИМЕНИТЕЛЬНО К ПРЕДВАРИТЕЛЬНО ОБРАБОТКЕ ИЗОБРАЖЕНИЯ

К пространственной области относится совокупность пикселей, составляющих изображение. Функция предварительной обработки в пространственной области записывается в виде:

$$g(x,y) = h[f(x,y)],$$

где $f(x,y)$ - входное изображение,

$g(x,y)$ – выходное (обработанное) изображение,

h – оператор функции f , определенный в некоторой области (x,y) .

Основным подходом при определении окрестности точки (x,y) является использование квадратной или прямоугольной области части изображения с центром в точке (x,y) . Центр этой части изображения перемещается от пиксела к пикселу, начиная, например, от левого верхнего угла. При этом для получения $g(x,y)$ оператор применяется для каждого положения (x,y) . Хотя используются иногда и другие формы окрестности (например, круг), квадратные формы более предпочтительны из-за простоты их реализации.

Один из наиболее применяемых методов пространственной области основан на использовании так называемых масок свертки (шаблонов, окон, фильтров). Обычно маска представляет собой небольшую (например, размерность 3×3) двумерную систему, коэффициенты которой выбираются таким образом, чтобы обнаружить заданное свойство изображения.

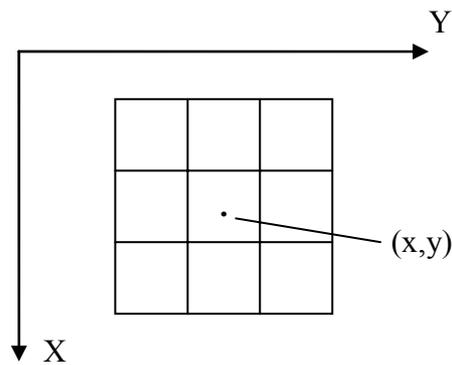


Рис. 16.1

На рис.м 16.1 представлена окрестность размерностью 3*3 точки (x,y) изображения. Предположим для начала, что дано изображение с постоянной интенсивностью, которое содержит отдельные, удаленные друг от друга пиксели с отличной от фона интенсивностью. Эти точки могут быть обнаружены маской, показанной на рис. 16.2

-1	-1	-1
-1	8	-1
-1	-1	-1

Рис. 16.2

Процесс заключается в следующем. Центр маски, помеченный цифрой 8, перемещается по изображению определенным образом. При совпадении центра маски с положением каждого пиксела производится умножение значений всех пикселов, находящихся под маской, на соответствующий коэффициент на маске, т.е. значение пиксела под центром маски умножается на 8, а значения восьми соседних пикселов умножаются на -1. Затем результаты этих умножений суммируются. Если все пиксели под маской имеют одинаковые значения (постоянный фон), то сумма будет равна нулю. Если же центр маски разместится над точкой с другой интенсивностью, сумма будет отлична от нуля. В случае размещения указанной точки вне центра, сумма также будет отлична от нуля, но на меньшую величину. Это

различие может быть устранено путем сравнения значения суммы с пороговым значением.

Если величины w_1, w_2, \dots, w_9 представляют собой коэффициенты, маски пиксела (x, y) и его восьми соседей (рис.16.3), то описанный выше алгоритм можно представить как выполнение следующей операции на окрестности 3×3 точки (x, y) .

w_1 $(x-1, y-1)$	w_2 $(x-1, y)$	w_3 $(x-1, y+1)$
w_4 $(x, y-1)$	w_5 (x, y)	w_6 $(x, y+1)$
w_7 $(x+1, y-1)$	w_8 $(x+1, y)$	w_9 $(x+1, y+1)$

Рис. 16.3

$$\begin{aligned}
 h[f(x, y)] = & w_1 f(x-1, y-1) + w_2 f(x-1, y) + w_3 f(x-1, y+1) + \\
 & + w_4 f(x, y-1) + w_5 f(x, y) + w_6 f(x, y+1) + w_7 f(x+1, y-1) + \\
 & + w_8 f(x+1, y) + w_9 f(x+1, y+1)
 \end{aligned}
 \tag{16.1}$$

Методы частотной области используют для своей реализации алгоритмы прямого и обратного двумерного преобразований.

$$\begin{aligned}
 F(u, v) &= \frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(ux+vy)/N} \\
 f(x, y) &= \sum_{U=0}^{N-1} \sum_{V=0}^{N-1} F(u, v) e^{j2\pi(ux+vy)/N}
 \end{aligned}
 \tag{16.2}$$

При вычислении прямого двумерного преобразования Фурье, сначала с использованием техники одномерного преобразования Фурье оперируют строками массива изображения, а затем одномерные преобразования осуществляются над столбцами (построчно-столбцовый алгоритм).

Операция сглаживания используется для снижения шума и других помех, которые могут появиться на изображении в результате его дискретизации и квантования.

Усреднение окрестности является прямым методом пространственной области для сглаживания изображения. Для изображения $f(x,y)$ процесс заключается в получении сглаженного изображения $g(x,y)$, интенсивность которого в каждой точке (x,y) равна усредненному значению интенсивности пикселей функции f , содержащихся в заданной окрестности точки (x,y) . Иными словами сглаженное изображение получается с использованием соотношения:

$$g(x, y) = \frac{1}{P} \sum_{n,m \in S} f(n, m)$$

для всех x,y функции $f(x,y)$, где p - общее число точек в окрестности.

Для окрестности размерностью $3*3$ указанное соотношение является частным случаем выражения (16.1) при $w_i=1/9$.

17. СЕГМЕНТАЦИЯ ИЗОБРАЖЕНИЯ ПОСРЕДСТВОМ ВЫДЕЛЕНИЯ ГРАНИЦ ОБЛАСТЕЙ

Сегментация означает выделение областей, однородных по своим яркостным, цветовым или текстурным свойствам, либо посредством выделения их границ, либо путем разметки внутренних точек. В свою очередь, можно выделить два класса, которыми определяются методы выделения границ: пространственное дифференцирование и высокочастотная фильтрация. Причем общим для указанных методов является стремление рассматривать границу (край), как область резкого перепада функции яркости изображения.

Метод пространственного дифференцирования основан на предположении о том, что граничные (краевые) точки имеют большую величину модуля градиента функции $f(x,y)$. Градиент изображения $f(x,y)$ в точке (x,y) определяется как двумерный вектор:

$$\overline{G[f(x,y)]} = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (17.1)$$

7

Из векторного анализа известно, что вектор G указывает направление максимального изменения функции f в точке (x,y) . Величина этого вектора определяется как:

$$G[f(x,y)] = [G_x^2 + G_y^2]^{1/2} = \left[\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 \right]^{1/2} \quad (17.2)$$

На практике, как правило, градиент аппроксимируется абсолютными значениями

$$G[f(x,y)] \approx |G_x| + |G_y| \quad (17.3)$$

Как видно из указанных соотношений, вычисление градиента основано на нахождении первых производных $\partial f/\partial x$, $\partial f/\partial y$.

Для цифрового изображения это можно сделать несколькими путями. Один из подходов состоит в использовании разности между соседними пикселями.

$$G_x = \partial f/\partial x = f(x,y) - f(x-1,y) \quad (17.4)$$

$$G_y = \partial f/\partial y = f(x,y) - f(x,y-1)$$

Несколько более сложный способ включает пиксели в окрестности размерностью 3*3 и более.

Структура метода пространственного дифференцирования может быть представлена следующим образом



Рис. 7

Получаемое после преобразования D поле $g(x,y)$ называется градиентным изображением или изображением с усиленными границами. Обработка градиентного изображения осуществляется с помощью порогового оператора согласно правилу:

$$b(x,y) = \begin{cases} 1 & \text{при } g(x,y) \geq T \\ 0 & \text{при } g(x,y) < T \end{cases}$$

где T – величина постоянного или меняющегося по определенному правилу порога.

Вычислительная реализация данного метода сводится к синтезу численных алгоритмов оценки частных производных в некоторой точке изображения и выполнению далее несложных арифметически операций.

Пусть через $f(m,n)$ и $g(m,n)$, где $m, n = \overline{0, N-1}$ обозначены дискретизированные исходное и градиентное изображения и в качестве дискретного аналога соотношения (17.2) выступает:

$$g(m,n) = [d_1^2(m,n) + d_2^2(m,n)]^{1/2}$$

или его упрощенный вычислительный вариант:

$$g(m,n) = |d_1(m,n)| + |d_2(m,n)|$$

В этих соотношениях $d_1(m,n)$ и $d_2(m,n)$ обозначают оценки частных производных функции яркости в точке (m,n) .

Обозначим через F_{mn} и F_{mn} векторы

$$(F_{mn})^T = (f_{mn}, f_{m,n+1}, f_{m+1,n}, f_{m+1,n+1})$$

$$(F_{mn})^T = (f_{m-1,n-1}, f_{m-1,n}, f_{m-1,n+1}, f_{m,n-1}, f_{m,n}, f_{m,n+1}, f_{m+1,n-1}, f_{m+1,n}, f_{m+1,n+1}),$$

представляющие собой упорядоченные по строкам элементы окрестности точки (m,n) , размерностью $2*2$ и $3*3$.

В качестве масок (шаблонов) можно использовать разностный оператор:

$$W_1 = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}; W_2 = \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}$$

ТОГДА:

$$d_1(m,n) = [f(m,n) - f(m,n+1)] + [f(m+1,n) - f(m+1,n+1)]$$

$$d_2(m,n) = [f(m+1,n) - f(m,n)] + [f(m+1,n+1) - f(m,n+1)]$$

Оператор Робертса

$$W_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}; W_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

$$d_1(m,n) = [f(m,n) - f(m+1,n+1)]$$

$$d_2(m,n) = [f(m+1,n) - f(m,n+1)]$$

Оператор Превитта

$$W_1 = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}; W_2 = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

$$d_1(m,n) = [f(m-1,n-1) + f(m,n-1) + f(m+1,n-1)] - [f(m-1,n+1) + f(m,n+1) + f(m+1,n+1)]$$

$$d_2(m,n) = [f(m+1,n-1) + f(m+1,n) + f(m+1,n+1)] - [f(m-1,n-1) + f(m-1,n) + f(m-1,n+1)]$$

Оператор Собеля

$$W_1 = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}; W_2 = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

$$d_1(m,n) = [f(m-1,n-1) + 2f(m,n-1) + f(m+1,n-1)] - [f(m-1,n+1) + 2f(m,n+1) + f(m+1,n+1)]$$

$$d_2(m,n) = [f(m+1,n-1) + 2f(m+1,n) + f(m+1,n+1)] - \\ - [f(m-1,n-1) + 2f(m-1,n) + f(m-1,n+1)]$$

18. ОСНОВНЫЕ СВОЙСТВА ДВУМЕРНОГО ПРЕОБРАЗОВАНИЯ ФУРЬЕ

Сдвиг

Пара преобразований Фурье обладает следующими трансляционными (сдвиговыми) свойствами

$$f(x,y)e^{j2\pi(u_0 x/M+v_0 y/N)} \Leftrightarrow F(u-u_0, v-v_0)$$

$$f(x-x_0, y-y_0) \Leftrightarrow F(u, v)e^{-j2\pi(x_0 u/M+y_0 v/N)},$$

где двойная стрелка использована для обозначения того, что соответствующие функции образуют пару Фурье-преобразований.

При $u_0=M/2$, $v_0=N/2$ будем иметь

$$e^{j2\pi(u_0 x/M+v_0 y/N)} = e^{j\pi(x+y)} = (-1)^{x+y}$$

В этом случае будем иметь

$$f(x,y)(-1)^{x+y} \Leftrightarrow F(U-M/2, V-N/2)$$

и аналогично при $x_0=M/2; y_0=N/2$

$$f(x-M/2, y-N/2) \Leftrightarrow F(u, v)(-1)^{u+v}$$

Из определения преобразования Фурье следует, что

$$F[f_1(x,y)+f_2(x,y)] = F[f_1(x,y)] + F[f_2(x,y)] \text{ и}$$

$$F[f_1(x,y)*f_2(x,y)] \neq F[f_1(x,y)]*F[f_2(x,y)]$$

Иными словами, преобразование Фурье обладает свойством дистрибутивности по отношению к сложению, но не обладает этим свойством по отношению к умножению. Указанное относится и к обратному Фурье-преобразованию.

Аналогично, если a, b – две постоянные, то

$$f(ax, by) \Leftrightarrow aF(u, v)$$

$$f(ax, by) \Leftrightarrow \frac{1}{|ab|} F(u/a, v/b), ab \neq 0$$

(изменение масштаба)

Поворот

Введем понятие координаты

$$x = r \cdot \cos \theta, \quad y = r \cdot \sin \theta$$

$$u = \omega \cdot \cos \varphi$$

$$v = \omega \cdot \sin \varphi$$

И обозначим через $f(r, \theta)$ и $F(\omega, \varphi)$ функции, рассматриваемые как функции полярных координат соответственно.

$$\text{Тогда } f(r, \theta + \theta_0) \Leftrightarrow F(\omega, \varphi + \theta_0)$$

Это выражение указывает на то, что поворот функции $f(x, y)$ на угол θ_0 приводит к повороту функции $F(u, v)$ на тот же угол.

Дискретное преобразование Фурье обладает следующими свойствами периодичности

$$F(u, v) = F(u + M, v) = F(u, v + N) = F(u + M, v + N)$$

Обратное Фурье-преобразование также периодично:

$$f(x,y)=f(x+M,y)=f(x,y+N)=f(x+M,y+N)$$

Свойство симметрии относительно операции комплексного сопряжения

$$F(u,v)=F^*(-u,-v)$$

Отсюда следует, что спектр центрально симметричен относительно начала координат

$$|F(u,v)| = |F(-u,-v)|$$

В случае двумерного преобразования спектр определяется как

$$|F(u,v)| = [R^2(u,v)+I^2(u,v)]^{1/2}$$

где $R=\text{Re}(F)$, $I=\text{Im}(F)$

Фаза определяется как $\Phi(u,v)=\text{arctg}[I(u,v)/R(u,v)]$

Энергетический спектр вычисляется как $P(u,v)=|F(u,v)|^2$, а среднее значение будет равно

$$\bar{f}(x,y) = F(0,0) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y)$$

Представление в полярных координатах имеет вид

$$F(u,v) = |F(u,v)| e^{-j\Phi(u,v)}$$

Свойство разделения переменных означает, что для вычисления двумерного преобразования Фурье необходимо вначале вычислить одномерные преобразования по каждой строке изображения, а затем вычислить одномерные преобразования по каждому столбцу полученного промежуточного результата. Изменение порядка вычислений (сначала по столбцам, затем по строкам) приводит к тому же результату.

Дискретная свертка двух функций $f(x,y)$ и $h(x,y)$ размерностью $M \times N$ обозначается как $f(x,y) * h(x,y)$ и определяется как

$$f(x,y) * h(x,y) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m,n) h(x-m, y-n)$$

Используя двумерное Фурье-представление, будем иметь

$$f(x,y) * h(x,y) \Leftrightarrow F(u,v) H(u,v)$$

$$f(x,y) h(x,y) \Leftrightarrow F(u,v) * H(u,v)$$

При вычислении двумерной свертки необходимо использовать технологию дополнения нулями.

Пусть мы имеем два изображения $f(x,y)$ и $h(x,y)$ размерами $A \times B$ и $C \times D$ соответственно. Как и в одномерном случае необходимо рассматривать эти массивы как периодические с некоторыми периодами P в направлении x и Q в направлении y . Причем выбор

$$P \geq A + C - 1$$

$$Q \geq B + D - 1$$

позволяет избежать ошибок перекрытия при вычислении двумерной свертки. При этом дополнение нулями функций $f(x,y)$, $h(x,y)$ в основном периоде осуществляется следующим образом

$$fe(x,y) = \begin{cases} f(x,y) & 0 \leq x \leq A-1 & 0 \leq Y \leq B-1 \\ 0 & A \leq x \leq P & B \leq Y \leq Q \end{cases}$$

$$he(x,y) = \begin{cases} h(x,y) & 0 \leq x \leq C-1 & 0 \leq Y \leq D-1 \\ 0 & C \leq x \leq P & D \leq Y \leq Q \end{cases}$$

Корреляционная функция двух функций $f(x,y)$ и $h(x,y)$ определяется соотношением

$$f(x,y) \circ h(x,y) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f^*(m,n)h(x+m,y+n)$$

где f^* обозначает функцию, комплексно сопряженную функции f . Обычно мы имеем дело с действительными функциями и в этом случае $f^*=f$. Корреляция от свертки отличается комплексным сопряжением (в общем случае) и заменой минусов на плюсы во всех членах под знаком суммы. В остальном вычисление корреляционной функции идентично вычислению свертки, включая необходимость дополнения нулями.

Тогда теорема корреляции

$$f(x,y) \circ h(x,y) \Leftrightarrow F^*(u,v)H(u,v)$$

означает, что корреляционная функция в пространственной области может быть получена как результат обратного Фурье-преобразования, примененного к произведению $F^*(u,v)H(u,v)$, где F^* - функция, комплексно сопряженная к F .

Аналогично вычисление корреляционной функции в частотной области сводится к умножению в пространственной области.

$$F(u,v) \circ H(u,v) \Leftrightarrow f^*(x,y)h(x,y)$$

Указанные два утверждения составляют содержание теоремы о корреляции, при чем все функции должны быть дополнены нулями.

Следует отметить, что свертка является связующим звеном между процедурами фильтрации в пространственной и частотной областях, а основная задача, для решения которой используется корреляция – это задача совмещения.

Изображение $f(x,y)$ участвующее в совмещении состоит, как правило, из некоторых объектов или областей. Если мы хотим определить, содержит ли изображение f конкретный интересующий нас объект (область), то мы формируем отдельное изображение $h(x,y)$ этого объекта (обычно называемое эталоном) и вычисляем корреляционную функцию изображения и эталона. Тогда, если данное изображение содержит объект, совмещенный с эталоном, то корреляция двух функций будет иметь максимум, расположенный там, где обнаружено соответствие между эталоном h и изображением. При решении большинства практических задач совмещения необходимы процедуры предварительной обработки, такие как утоньшение, масштабирование, выравнивание (эквализация), однако основой является вычисление корреляции. Вместо термина корреляция очень часто используется термин «взаимная корреляция» чтобы подчеркнуть, что корреляция вычисляется между различными изображениями. Если изображения сравниваемые совпадают, то

$$f(x,y) \circ f(x,y) \Leftrightarrow F^*(u,v)F(u,v) \Leftrightarrow |F(u,v)|^2$$

Это соотношение представляет собой энергетический спектр.

Аналогично

$$|f(x,y)|^2 \Leftrightarrow F(u,v) \circ F(u,v)$$

19. ОСНОВЫ ФИЛЬТРАЦИИ В ЧАСТОТНОЙ ОБЛАСТИ

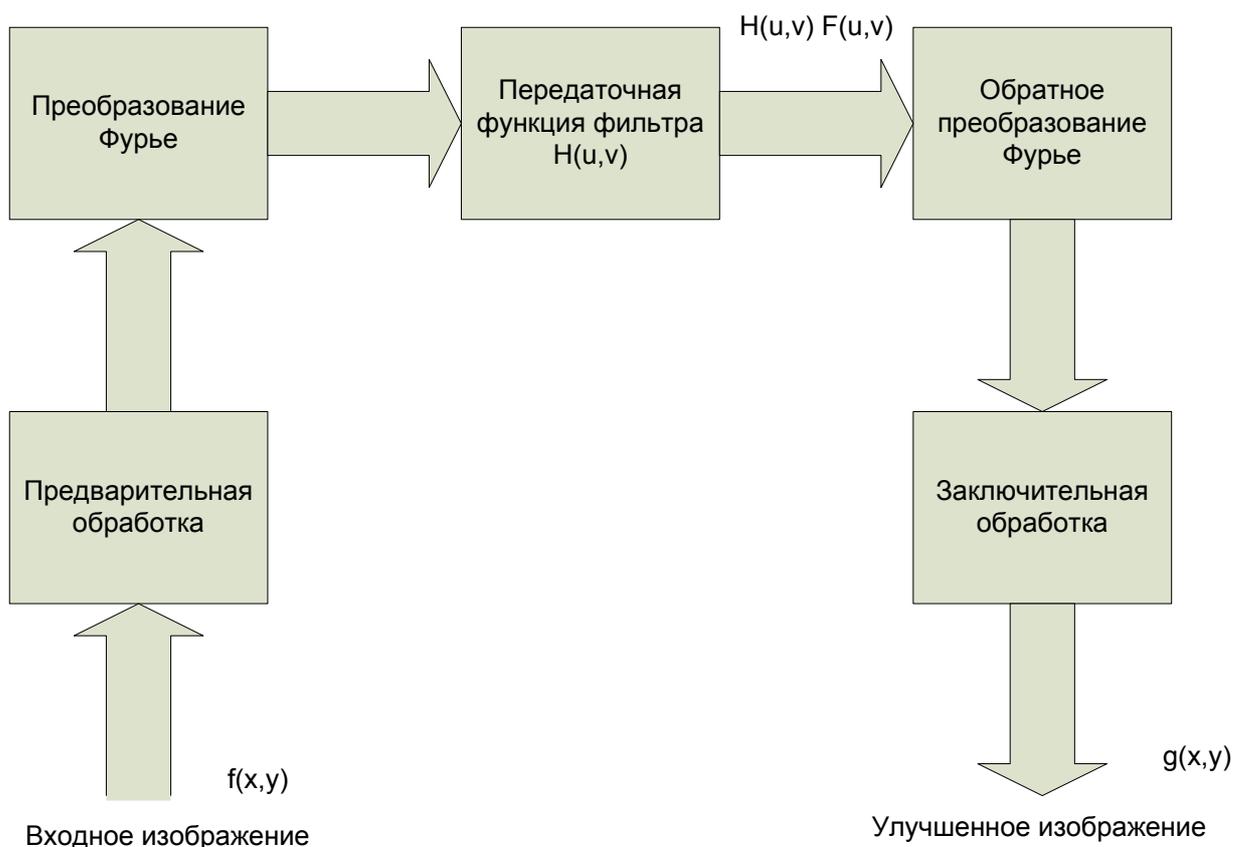


Рис. 19.1

$$F[f(x,y)(-1)^{x+y}] = F(u-M/2, v-N/2) \quad (19.1)$$

Процедура фильтрации в частотной области состоит из следующих шагов:

1. Исходное изображение умножается на $(-1)^{x+y}$, чтобы его Фурье-преобразование в соответствии с (19.1) оказалось центрированным.
2. Вычисляется прямое ДПФ $F(u,v)$ изображения, полученного после шага 1.
3. Функция $F(u,v)$ умножается на $H(u,v)$ (передаточная функция фильтра).
4. Вычисляется обратное ДПФ от результата шага 3.
5. Выделяется вещественная часть результата шага 4.

6. Результат шага 5 умножается на $(-1)^{x+y}$

Причина, по которой множитель $H(u,v)$ называется фильтром, состоит в том, что он подавляет некоторые частоты преобразования, оставляя при этом другие без изменения.

Аналогия с повседневной жизнью возникает при рассмотрении сетчатого фильтра (сито), который пропускает некоторые предметы и не пропускает другие в соответствии с их размерами.

Пусть $f(x,y)$ обозначает входное изображение после шага 1 и пусть $F(u,v)$ есть его Фурье-образ. Тогда Фурье-образ выходного изображения определяется соотношением

$$G(u,v)=H(u,v)F(u,v)$$

Умножение функций двух переменных H и F осуществляется поэлементно. Это означает, что первый элемент H умножается на первый элемент F , второй элемент H – на второй элемент F и так далее.

Фильтрованное изображение получается вычислением обратного преобразования Фурье от Фурье-образа $G(u,v)$

$$g_{\text{фильтр}}(x,y)=F^{-1}[G(u,v)]$$

Искомое изображение получается выделением действительной части из последнего результата и умножения на $(-1)^{x+y}$, чтобы скомпенсировать эффект от умножения входного изображения на ту же величину.

Обратное Фурье-преобразование в общем случае является комплексным. Однако, в случае вещественного входного изображения и вещественной передаточной функции фильтра мнимые части всех значений обратного Фурье-преобразования должны равняться нулю. Однако в связи с ошибками округления при вычислениях на практике значения обратного Фурье-преобразования, как правило, содержат паразитную мнимую

составляющую, которой необходимо при решении прикладных задач пренебречь.

Процедура фильтрации, приведенная на рис.19.1, представлена в общем виде, включая предварительную и заключительную обработку. Помимо умножения на $(-1)^{x+y}$ такая обработка может включать обрезание входного изображения так, чтобы его размеры приняли ближайшие четные значения по отношению к исходным (это необходимо для правильного центрирования Фурье-преобразования), яркостное масштабирование, преобразование форматов входных и выходных данных. Возможны многоступенчатые процедуры фильтрации, а также разнообразные операции предварительной и заключительной обработки. Существенным является то, что метод фильтрации основан на некотором изменении Фурье-образа изображения посредством передаточной функции фильтра и последующем обращении результата для получения обработанного выходного изображения.

Фильтры, основанные на порядковых статистиках

Фильтры, основанные на порядковых статистиках, относятся к классу нелинейных пространственных фильтров. Отклик фильтра определяется предварительным упорядочиванием (ранжированием), значений пикселей, покрываемых маской фильтра и последующим выбором значения, находящегося на определенной позиции упорядоченной последовательности (т. е. имеющего определенный ранг). При этом фильтрация сводится к замещению исходного значения пикселя (в центре маски) на полученное значение отклика фильтра. Из указанного класса наиболее известен медианный фильтр, который заменяет значение пикселя на значение медианы распределения яркости всех пикселей в окрестности, включая исходный. Медианные фильтры особенно эффективны при фильтрации импульсных шумов, иногда называемых шумами «соль и перец», которые выглядят как наложение на изображение случайных черных и белых точек.

Медиана набора чисел есть такое число ξ , что половина чисел из набора меньше или равны ξ , а другая половина – больше и равны ξ . Чтобы выполнить медианную фильтрацию для элемента изображения, необходимо сначала упорядочить по возрастанию значения пикселей внутри окрестности, затем найти значение медианы и наконец присвоить полученное значение обрабатываемому элементу. Так для окрестности размером 3×3 медианой будет пятое значение по величине, для окрестности 5×5 – тринадцатое значение и так далее.

Аналитически действие медианного фильтра определяется как

$$\hat{f}(x, y) = \underset{s, t \in S_{xy}}{\text{med}} \{g(s, t)\}$$

Причем при вычислении медианы значение в самой точке (т. е. в центре окружности) также учитывается.

Если несколько элементов в окрестности имеют одинаковые значения, то значения будут сгруппированы. Например, пусть в окрестности 3×3 имеют место следующие значения (10, 20, 20, 20, 15, 20, 20, 25, 100). После упорядочения они будут расположены следующим образом (10, 15, 20, 20, 20, 20, 20, 25, 100), а, следовательно, медианой будет значение 20.

Хотя медианный фильтр значительно более распространен в обработке изображений, чем остальные виды фильтров, тем не менее, он не является единственным. Медиана представляет собой 50-ый перцентиль упорядоченного набора чисел.

Широкая популярность медианных фильтров обусловлена тем, что они прекрасно приспособлены для подавления импульсных шумов и при этом приводят к меньшему размыванию по сравнению с линейными сглаживающими фильтрами.

Использование 100-го перцентиля приводит к фильтру, основанному на выборе максимального значения (фильтр максимума), который задается выражением

$$\hat{f}(x, y) = \max_{(s,t) \in S_{xy}} \{g(s, t)\}$$

Такой фильтр полезен при обнаружении наиболее ярких точек на изображении.

Использование 0-го перцентиля приводит к фильтру, основанному на выборе минимального значения (фильтр минимума)

$$\hat{f}(x, y) = \min_{(s,t) \in S_{xy}} \{g(s, t)\}$$

Такой фильтр полезен при обнаружении наиболее темных точек на изображении.

Применение фильтра срединной точки заключается в вычислении среднего между максимальным и минимальным значениями в соответствующей окрестности

$$\hat{f}(x, y) = \frac{1}{2} [\max_{(s,t) \in S_{xy}} \{g(s, t)\} + \min_{(s,t) \in S_{xy}} \{g(s, t)\}]$$

Этот фильтр объединяет методы порядковых статистик и усреднения. Этот фильтр лучше всего работает при наличии таких случайно распределенных шумов, как гауссов или равномерный.

Предположим, что мы удалили $d/2$ наименьших и $d/2$ наибольших значений яркости из множества всех значений функции $g(s, t)$ в окрестности S_{xy} . Пусть $g_r(s, t)$ представляет оставшиеся элементы изображения, количество которых равно $(mn-d)$.

Фильтр, действие которого заключается в усреднении оставшихся значений, называется фильтром усеченного среднего

$$\hat{f}(x, y) = \frac{1}{mn - d} \sum_{(s,t) \in S_{xy}} q_r(s, t),$$

где $d = \overline{0, mn - 1}$

При $d=0$ фильтр усеченного среднего сводится к среднеарифметическому фильтру. В случае $d/2=(mn-1)/2$ фильтр трансформируется в медианный.

20. СЖАТИЕ ДАННЫХ

Важной областью применения ортогональных преобразований является сжатие данных. Если дискретный сигнал содержит N отсчётов, то его можно рассматривать как точку N -мерного пространства. Тогда каждый отсчёт является координатой N -мерного вектора данных. Для более эффективного представления можно осуществить ортогональное преобразование X , что приводит к:

$$Y = T * X,$$

где Y и T – вектор коэффициентов преобразования и матрица преобразования соответственно. Целью сжатия данных является выбор подмножества M координат вектора Y , где M существенно меньше N . Остальные $(N - M)$ координат можно отбросить, не вызывая существенной ошибки при восстановлении сигнала по M координатам вектора Y . Следовательно, сравнивать ортогональные преобразования следует в соответствии с некоторым критерием ошибки. Одним из часто используемых является критерий среднеквадратической ошибки.

Поиск оптимального преобразования

Найдём ортогональное преобразование, которое с одной стороны будет обеспечивать представление сигнала, а с другой стороны будет оптимальным в смысле среднеквадратичного критерия.

Пусть T - ортогональное преобразование, заданное в виде:

$$T = [\phi_1 \ \phi_2 \ \dots \ \phi_N],$$

где ϕ – векторы размерностью $N \times 1$. Для удобства базисные векторы $\{\phi_m\}$ $m = \overline{1, N}$, будем считать вещественнозначными и ортонормированными.

Для каждого вектора X , принадлежащего данному классу векторов исходных данных получим:

$$Y = T * X, \quad (20.1)$$

где $X = [x_1 \ x_2 \ \dots \ x_N]$, $Y = [y_1 \ y_2 \ \dots \ y_N]$

Из соотношения (20.1) :

$X = T^{-1}Y = [y_1 \ y_2 \ \dots \ y_N]Y$, что можно записать как

$$X = y_1\phi_1 + y_2\phi_2 + \dots + y_N\phi_N = \sum_{i=1}^N y_i \phi_i$$

Желательно сохранить подмножество $\{y_1, y_2, \dots, y_M\}$ координат Y и при этом получить оценку X . Это может быть осуществлено заменой остальных $(N - M)$ координат Y заранее выбранными константами, что приводит к

$$\tilde{X}(M) = \sum_{i=1}^M y_i \phi_i + \sum_{i=M+1}^N b_i \phi_i,$$

где $\tilde{X}(M)$ обозначает оценку X . Ошибку, возникающую при отбрасывании $N - M$ координат можно представить в виде

$$\Delta X = X - \tilde{X}(M) \text{ т.е.}$$

$$\Delta X = X - \sum_{i=1}^M y_i \phi_i - \sum_{i=M+1}^N b_i \phi_i = \sum_{i=M+1}^N y_i \phi_i - \sum_{i=M+1}^N b_i \phi_i = \sum_{i=M+1}^N (y_i - b_i) \phi_i.$$

Таким образом, среднеквадратичная ошибка $\varepsilon(M)$ определяется в виде

$$\varepsilon(M) = \varepsilon\{\|\Delta x\|^2\} = \varepsilon\{(\Delta x)^T \Delta x\}$$

С учётом подстановки получим

$$\varepsilon(M) = \varepsilon\left\{\sum_{i=M+1}^N \sum_{j=M+1}^N (y_i - b_i)^T (y_j - b_j) \phi_i^T \phi_j\right\},$$

что в упрощённом виде можно записать как

$$\varepsilon(M) = \sum_{i=M+1}^N E\{(y_i - b_i)^2\}$$

Из последнего соотношения следует, что для каждого выбора ϕ_i и b_i получаем определённое значение $\varepsilon(M)$. Требуется определить такую комбинацию этих величин, чтобы минимизировать $\varepsilon(M)$. Процедура выбора оптимальных ϕ_i и b_i разбивается соответственно на 2 этапа.

Первый этап

Оптимальное значение b_i определяется из выражения

$$\frac{\partial}{\partial b_i} E\{(y_i - b_i)^2\} = -2[E\{y_i\} - b_i] = 0$$

Тогда $E\{y_i\} = b_i$.

В свою очередь $y_i = \phi_i \cdot x$.

Таким образом

$$b_i = \phi_i E\{x\} = \phi_i \bar{x} \quad \text{где } E\{x\} = \bar{x}$$

Так как разность $(y_i - b_i)$ является скалярной, то $\varepsilon(M)$ можно записать как

$$\varepsilon(M) = \sum_{i=M+1}^N E\{(y_i - b_i)(y_i - b)^T\} \quad (20.2)$$

Подстановка

$$y_i = \phi_i x \quad b_i = \phi_i \bar{x} \quad \text{в (20.1) приводит к}$$

$$\varepsilon(M) = \sum_{i=M+1}^N \phi_i E\{(x - \bar{x})(x - \bar{x})^T\} \phi_i^T$$

Так как $\sum_x = E\{(x - \bar{x})(x - \bar{x})^T\}$ является ковариационной матрицей x , то получим

$$\varepsilon(M) = \sum_{i=M+1}^N \phi_i \sum_x \phi_i^T \quad (20.3)$$

Второй этап

Для нахождения оптимального ϕ_i следует не только минимизировать $\varepsilon(M)$ по отношению к ϕ_i , но и удовлетворить ограничению

$$\phi_i^T \phi_i = 1$$

При этом, используя метод множителей Лагранжа, получим

$$\sum_x \phi_i^T = \phi_i^T \lambda_i \quad (20.4)$$

Подставляя (20.4) в (20.3), получим минимальное значение среднеквадратической ошибки в виде

$$E_{\min}(M) = \sum_{i=M+1}^N \lambda_i$$

Таким образом, разложение

$x = \sum_{i=1}^N y_i \phi_i$ представляет собой разложение по собственным векторам ковариационной матрицы. Это разложение называется разложением Карунена-Лоэва. В литературе по статистике задачи по минимизации $\varepsilon(M)$ называются факторным анализом или анализом главных компонент.

Таким образом можно сделать 2 вывода:

1. Преобразование Карунена-Лоэва является оптимальным для представления сигналов по отношению к критерию среднеквадратичной ошибки.
2. Поскольку $Y = TX$, то ковариационная функция в области изображения \sum_Y определяется как

$$\sum_Y = T \sum_X T^{-1} = T \sum_X T^T$$

Так как T состоит из собственных векторов \sum_X , то

$$\sum_Y = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N)$$

где λ_i - собственные значения \sum_X . Так как \sum_Y - диагональная матрица, то координаты вектора преобразованных данных y_i в выражении $y_i = \phi_i^T x$ некоррелированы.

Из выражения $E_{\min}(M) = \sum_{i=M+1}^N \lambda_i$ следует, что эффективность коэффициента преобразования y_i для представления вектора данных X определяется соответствующим ему собственным значением. Если какой-либо коэффициент y_k не учитывается, то среднеквадратичная ошибка увеличивается на соответствующее собственное значение λ_k . Таким образом,

необходимо выбрать множество y_i соответствующее M наибольшим собственным значениям, а остальные y_i отбросить, так как их можно заменить константами $b_i \quad i = \overline{M+1, N}$.

Так как $b_i = \phi_i \bar{X}$, то остальные y_i можно приравнять к нулю, если исходные данные центрированы, то есть $\bar{X} = 0$.

Так как собственные значения являются элементами \sum_Y , стоящими на главной диагонали, то они соответствуют дисперсиям коэффициентов преобразования y_i . Для всех остальных преобразований \sum_Y содержит ненулевые внедиагональные элементы. Поэтому естественным критерием при выборе множества оставляемых коэффициентов преобразования является сохранение M коэффициентов с наибольшими дисперсиями, а остальные $(N - M)$ коэффициентов можно отбросить. Приведённый выше критерий выбора коэффициентов преобразования носит название дисперсионного критерия.

Выше предполагалось, что X принадлежит одному классу сигналов, но все рассуждения легко распространяются на несколько классов. Отличие заключается в том, что базисные векторы для соответствующего преобразования Карунена-Лоэва являются собственными векторами общей матрицы \sum_X^* , а не \sum_X . Матрица \sum_X^* определяется как:

$$\sum_X^* = p_1 \sum_{X1} + p_2 \sum_{X2} + \dots + p_K \sum_{XK} ,$$

Где K – число классов сигналов,

\sum_{X_i} - ковариационная матрица i -ого класса с априорной вероятностью

p_i .



Рис. 20.1 Модель сжатия данных в соотношении M:1.

При цифровой обработке изображений каждый элемент обычно кодируется в виде слова, содержащего 6 бит. Таким образом, элемент изображения можно представить десятичным числом от 1 до 64 или от 0 до 63 (то есть 64 уровнями). Закодированные, таким образом, данные обычно обрабатываются блоками размером $N \times N$, как показано на рис.

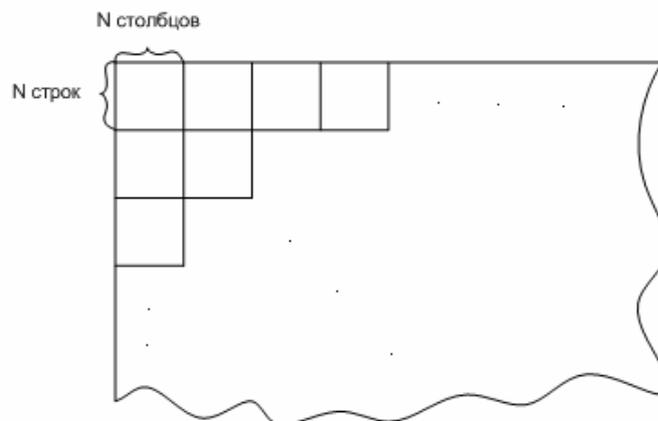


Рис. 20.2

Элемент изображения в i -ой строке и j -ом столбце можно представить в виде случайной величины $f(i, j)$, а матрицу ($N \times N$) случайных величин как $[f(i, j)]$. Тогда двумерное преобразование $[f(i, j)]$ и обратное преобразование можно записать как:

$$[F(u, v)] = T[f(i, j)]T'$$

$$[f(i, j)] = T' [F(u, v)] T,$$

где $[F(u, v)]$ - матрица коэффициентов преобразования, T – матрица преобразования.

Обозначим через $\sigma^2(i, j)$ и $\sigma^2(u, v)$ дисперсии $f(i, j)$ и $F(u, v)$ соответственно. Если функция распределения дисперсии $\sigma^2(u, v)$ неизвестна для изображения или класса изображений, подвергаемых преобразованию, её получают обычно моделированием. Один из подходов заключается в том что изображение описывается статистически как Марковский процесс первого порядка, а строки и столбцы обрабатываются независимо. Если дисперсия каждого столбца и строки равна σ^2 , то ковариационные матрицы для строк и столбцов можно записать как

$$\sum_k = \sigma^2 R_k$$

$$\text{где } R_k = \begin{vmatrix} 1 & \rho_k & \rho_k^2 & \dots & \rho_k^{N-1} \\ \rho_k & 1 & \rho_k & \rho_k^2 & \dots \\ \rho_k^2 & \rho_k & 1 & \rho_k & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \rho_k^{N-1} & \dots & \dots & \dots & 1 \end{vmatrix}$$

а ρ_1 и ρ_2 - коэффициенты корреляции для строк и столбцов случайных величин соответственно

Тогда ковариационные матрицы в области преобразований, соответствующие \sum_k запишутся как:

$$\tilde{\sum}_k = \sigma^2 [TR_k T'] \quad k=1,2$$

В свою очередь, функция распределения дисперсии $\sigma^2(u, v)$ определяется как

$$\tilde{\sigma}^2(u, v) = \tilde{\Sigma}_1(u, u) \tilde{\Sigma}_2(v, v)$$

Рассмотрим простейший способ сжатия данных, называемый зональным кодированием, который реализуется в 3 этапа:

1. Получают двумерное преобразование данного изображения, путём обработки блоками ($N \times N$).
2. Из N^2 коэффициентов преобразования сохраняют N^2/m коэффициентов, обладающих наибольшими двумерными дисперсиями $\sigma^2(u, v)$. Все остальные коэффициенты приравниваются нулю.
3. Каждый из N^2/m сохранённых коэффициентов кодируется с помощью k бит. Затем восстанавливается соответствующий ($N \times N$) блок с помощью обратного преобразования.

Так как кодируется только N^2/m коэффициентов преобразования, а не N^2 исходных элементов изображения для каждого ($N \times N$) блока, то среднее число бит на элемент восстановленного изображения равняется R/m .

Следует отметить, что при заданном m выбор типа преобразования определяется объёмом вычислений, объёмом требуемой аппаратуры и требованиями, предъявляемыми к качеству изображения.

21. ВЫБОР ПРИЗНАКОВ И РАСПОЗНАВАНИЕ ОБРАЗОВ.

ПРИНЦИП ОБУЧЕНИЯ

Задача распознавания образов заключается в классификации некоторой группы объектов на основе определённых требований.

Требования эти могут быть различными, так как в различных ситуациях возникает необходимость в определённом типе классификатора. К примеру при распознавании английских букв образуется 26 классов образов. Однако, чтобы отличить при распознавании буквы от китайских иероглифов нужны лишь два класса образов.

Проблема распознавания получила широкое распространение при решении различных задач в различных областях народного хозяйства: распознавание рукописных букв или слов, медицинская диагностика, диагностика неисправностей в технических системах, обнаружение цели, прогноз погоды, классификация сейсмических волн.

Простейший подход к распознаванию образов заключается в сопоставлении с эталоном. В этом случае некоторое множество образов по одному из каждого класса образов хранится в памяти машины. Входной, подлежащий распознаванию образ (неизвестного класса) сравнивается с эталоном каждого класса. Классификация основывается на заранее выбранном критерии соответствия или критерии подобия. Иными словами, если входной образ лучше соответствует эталону i -го класса образов, чем любому другому эталону, то входной образ классифицируется как принадлежащий i -му классу образов.

Более совершенный подход заключается в том, что вместо сравнения входного образа с эталонами, классификация основывается на некотором множестве отобранных замеров, называемых информативными признаками, которые предполагаются инвариантными или малочувствительными по отношению к изменениям и искажениям.

В этом случае распознавание образов можно рассматривать состоящим из двух задач.

Первая задача заключается в выборе подмножества признаков и их упорядочиванию в заданном множестве измерений. Вторая задача заключается в классификации (принятии решения о принадлежности входного образа тому или иному классу), которая основывается на измерениях отобранных признаков.

Упрощенная блок-схема системы распознавания приведена на рис.1.

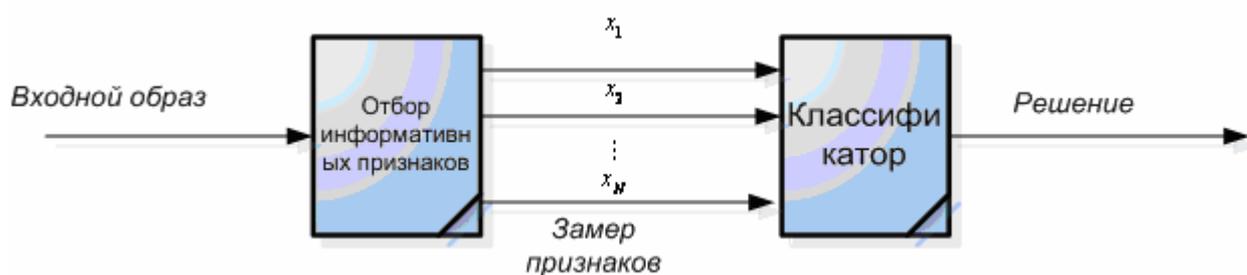


Рис. 21.1

Таким образом, подход, заключающийся в сопоставлении с эталоном, можно рассматривать как частный случай второго подхода- подхода измерения признаков, при этом эталоны хранятся в виде измеренных признаков, а в классификаторе используется специальный критерий сопоставления.

Математически задача классификации может быть сформулирована с помощью разделяющей функции.

Пусть $\omega_1, \omega_2, \dots, \omega_m$ обозначают m возможных классов образов, подлежащих распознаванию и пусть $X' = [x_1, x_2, \dots, x_n]$ вектор замеров признаков, где x_k предоставляет k -ый замер.

Тогда разделяющая функция $D_j(x)$, относящаяся к классу образов $\omega_j, j = \overline{1, m}$, такова, что если входной образ представленный вектором признаков X , принадлежит классу ω_i , то величина $D_i(x)$ должна быть

наибольшей . Пусть $X \sim \omega_i$ обозначает, что вектор признаков X входного образа принадлежит классу ω_i . Тогда можно записать , что для всех $X \sim \omega_i$

$$D_i(x) > D_j(x); i, j = \overline{1, m} \quad i \neq j \quad (21.1)$$

Таким образом, в пространстве признаков Ω_x граница разбиений , называемая решающей границей между областями , относящимися соответственно к классам ω_i, ω_j , выражается уравнением

$$D_i(x) - D_j(x) = 0$$

Общая схема классификатора , использующего критерий (21.1) и типичный двумерный пример приведены на рис. 21.2 и рис. 21.3.

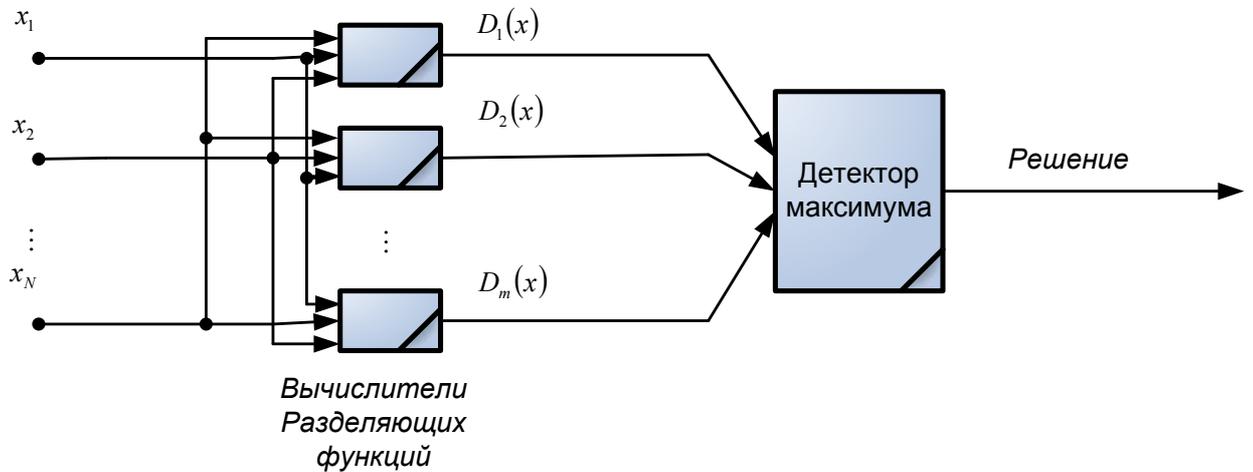


Рис. 21.2

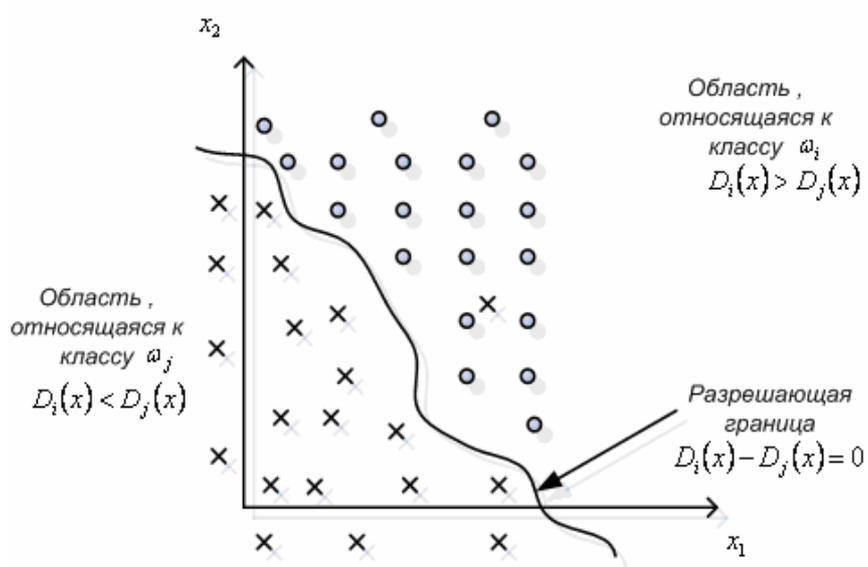


Рис. 21.3

22. ВИДЫ РАЗДЕЛЯЮЩИХ ФУНКЦИЙ. КЛАССИФИКАТОР ПО МИНИМАЛЬНОМУ РАССТОЯНИЮ

А. Линейные разделяющие функции.

В этом случае в качестве $D_i(x)$ используется линейная комбинация измеренных признаков x_1, x_2, \dots, x_n

$$D_i(x) = \sum_{k=1}^n w_{ki} x_k + w_{i,n+1}$$

Решающая граница между областями w_i и w_j в пространстве признаков Ωx имеет вид

$$D_i(x) - D_j(x) = \sum_{k=1}^n w_k x_k + w_{n+1} = 0 \quad (22.1)$$

где $w_k = w_{ki} - w_{kj}$,
 $w_{n+1} = w_{i,n+1} - w_{j,n+1}$

Уравнение (22.1) представляет собой уравнение гиперплоскости в пространстве признаков Ωx .

Общая схема вычислителя линейной разделяющей функции представлена на рис. 22.1.

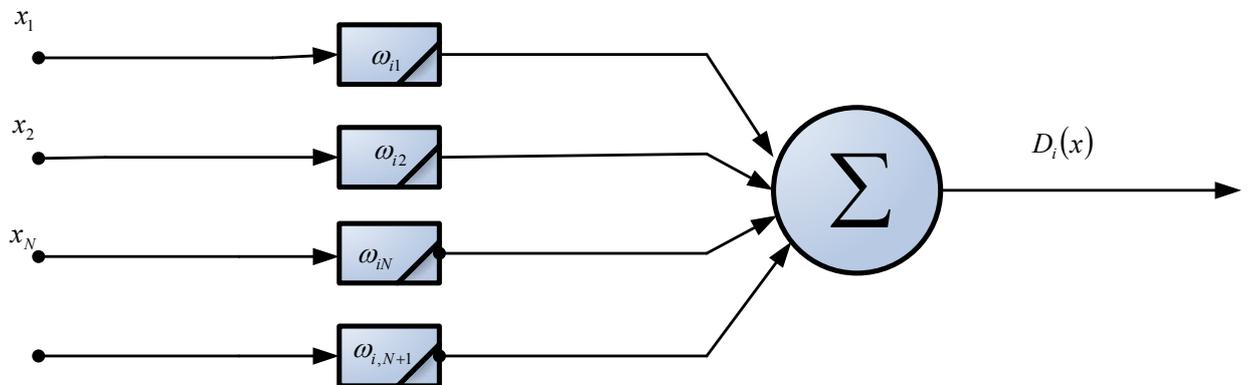


Рис. 22.1

Если $m=2$, то согласно (22.1) $i, j=1,2$ и в качестве классификатора, использующего линейную разделяющую функцию, можно применить пороговый логический элемент.

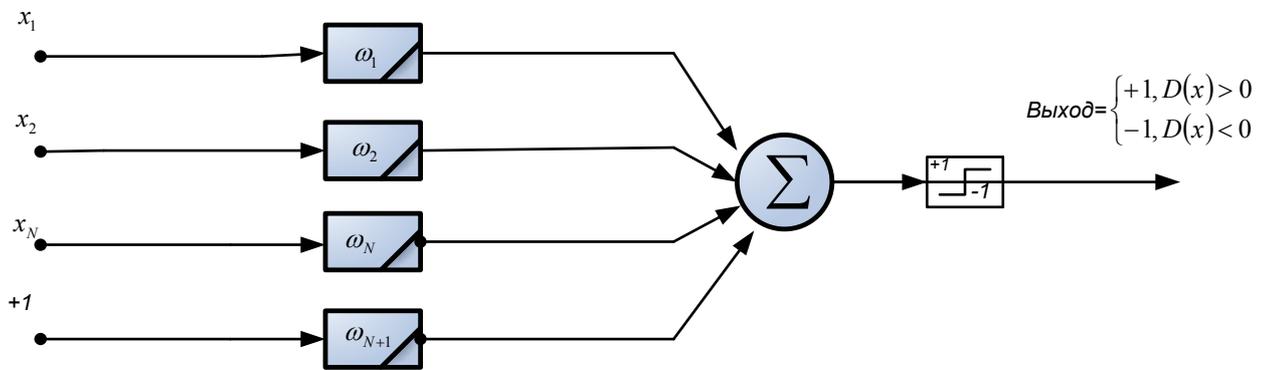


Рис. 22.2

Из этого рисунка получим, что если выходной сигнал $=+1$, т.е. при $D(x) > 0$, то $X \sim \omega_1$; если выходной сигнал $=-1$, т.е. $D(x) < 0$, то $X \sim \omega_2$.

Когда число классов образов больше двух, т.е. $m > 2$ можно применить параллельное соединение нескольких пороговых логических элементов. При этом комбинации выходных сигналов M пороговых логических элементов достаточны для различения m классов при $2^M \geq m$.

В. Классификатор по минимальному расстоянию.

Важный класс составляют линейные классификаторы, в которых в качестве критерия классификации используется расстояние между входным образом и множеством опорных векторов или эталонных точек в пространстве признаков. Предположим, что задано m опорных векторов R_1, R_2, \dots, R_m , где R_j соответствует классу образов ω_j . При классификации по минимальному расстоянию относительно R_1, R_2, \dots, R_m входной сигнал X предполагается принадлежащим ω_i , т.е.

$X \sim \omega_i$, если $|X - R_i|$ минимально, где $|X - R_i|$ есть расстояние между X и R_i .

Расстояние можно определить ,например, следующим образом

$$|X - R_i| = \left[(X - R_i)^T (X - R_i) \right]^{1/2}, \quad (22.2)$$

где индекс T определяет операцию транспонирования вектора.

Из последнего соотношения следует, что

$$|X - R_i|^2 = X^T X - X^T R_i - X R_i^T + R_i^T R_i$$

Так как $X^T X$ не зависит от i , то соответствующая разделяющая функция для классификатора по минимальному расстоянию имеет вид

$$D_i(x) = X^T R_i + X R_i^T - R_i^T R_i, \quad i = \overline{1, m}$$

Как видно из этого соотношения, классификатор по минимальному расстоянию является линейной функцией. В свою очередь, свойства классификатора по минимальному расстоянию конечно зависят от того, как выбраны опорные векторы.

С. Кусочно-линейная разделяющая функция.

Изложенная выше идея может быть распространена на классификацию по минимальным расстояниям до m множеств опорных векторов. Пусть R_1, R_2, \dots, R_m обозначают m множеств опорных векторов, относящихся соответственно к классам $\omega_1, \omega_2, \dots, \omega_m$ и пусть опорные векторы в R_j обозначены через $R_j^{(k)}$, т.е.

$$R_j^{(k)} \in R_j, \quad k = 1, \dots, U_j$$

где U_j -число опорных векторов множества R_j . Определим расстояние между вектором входных признаков и R_j следующим образом.

$$d(X, R_j) = \min_{k=1, \dots, U_j} |X - R_i^{(k)}|,$$

То есть расстояние между X и R_j равно наименьшему из расстояний между X и каждым вектором в R_j . Такой классификатор будет относить входной сигнал к классу образов, которому соответствует ближайшее множество векторов.

Если расстояние между X и $R_j^{(k)}$ определить согласно (22.1) разделяющая функция в данном случае имеет вид

$$D_i(x) = \max_{k=1, \dots, U_j} \{X^T R_i^{(k)} + X(R_i^{(k)})^T - (R_i^{(k)})^T R_i^{(k)}\} \quad i = \overline{1, m}$$

Пусть $D_i^{(k)}(x) = X^T R_i^{(k)} + X(R_i^{(k)})^T - (R_i^{(k)})^T R_i^{(k)}$,

Тогда

$$D_i(x) = \max_{k=1, \dots, U_j} \{D_i^{(k)}(x)\}$$

Следует отметить, что $D_i^{(k)}(x)$ является линейной комбинацией признаков. Поэтому указанный классификатор часто называют кусочно-линейным классификатором.

Д. Полиномиальная разделяющая функция.

Полиномиальная функция r -ой степени может быть представлена в виде

$$D_i(x) = w_{i1} f_1(x) + w_{i2} f_2(x) + \dots + w_{iL} f_L(x) + w_{i,L+1},$$

где $f_i(x)$ является формой

$$x_{k_1}^{n_1} x_{k_2}^{n_2} \dots x_{k_r}^{n_r} \quad \text{при} \quad \begin{cases} k_1, k_2, \dots, k_r = 1, \dots, N \\ n_1, n_2, \dots, n_r = 0, 1 \end{cases}$$

Решающая граница между двумя классами также имеет форму полинома r -ой степени. В частности, если $r=2$, решающая функция называется квадратичной.

В этом случае

$$f_j(x) = x_{k_1}^{n_1} x_{k_2}^{n_2} \quad \text{при} \quad \begin{cases} k_1, k_2 = 1, \dots, N \\ n_1, n_2 = 0, 1 \end{cases}$$

Разделяющая функция будет иметь вид

$$D_i(x) = \sum_{k=1}^N w_{kk} x_k^2 + \sum_{j=1}^{N-1} \sum_{k=j+1}^N w_{kj} x_j x_k + \sum_{j=1}^N w_j x_j + w_{i,L+1},$$

где $L = 1/2N(N+3)$

В общем случае границей для квадратичных разделяющих функций является гиперboloид. В частных случаях это будут гиперсфера, гиперэллипсоид и гиперэллипсоидальный цилиндр. Общая схема вычислителя квадратичного разделения показана на рис. 22.3

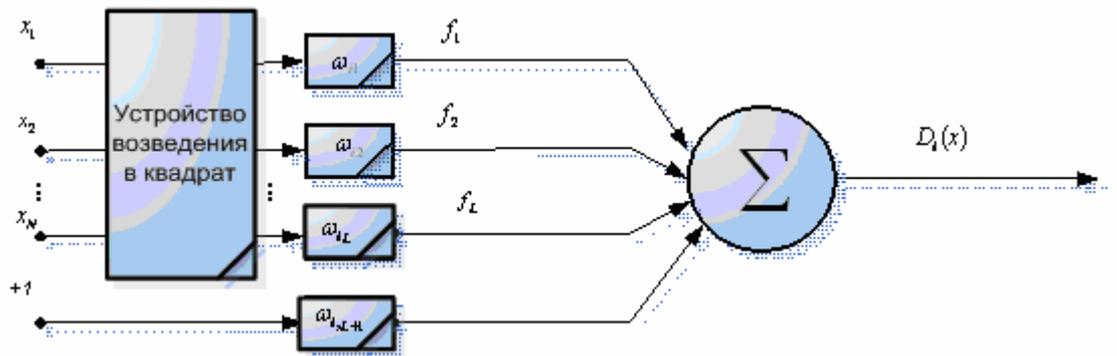


Рис. 22.3

23. СТРУКТУРА СИСТЕМЫ РАСПОЗНАВАНИЯ ОБРАЗОВ С ПРИМЕНЕНИЕМ ОРТОГОНАЛЬНЫХ ПРЕОБРАЗОВАНИЙ. ПРИНЦИП ОБУЧЕНИЯ В ДВУХКЛАССОВОЙ СИСТЕМЕ РАСПОЗНАВАНИЯ

Систему распознавания образов с применением ортогональных преобразований можно представить в виде

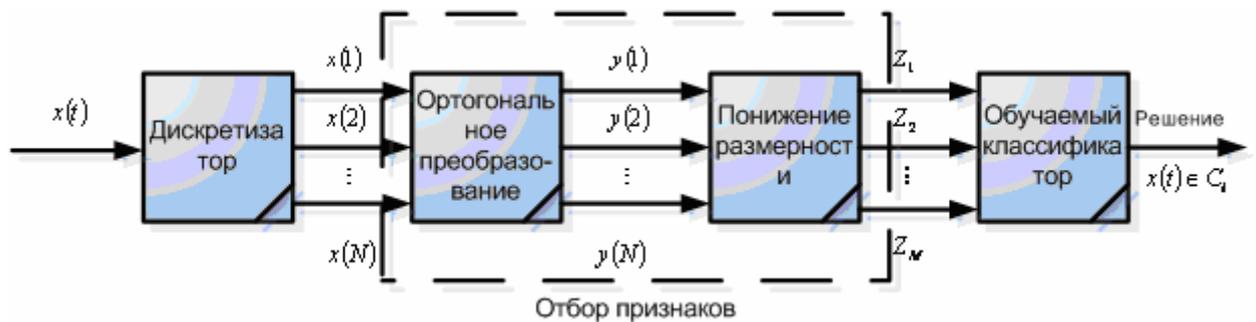


Рис. 23.1

Через $x(t)$ обозначается сигнал, принадлежащий одному из K классов C_1, C_2, \dots, C_K .

На первом этапе выбора осуществляется ортогональное преобразование. Вторым этапом выбора признака является понижение размерности, после чего получаем подмножество M признаков z_1, z_2, \dots, z_M из $\{Y(M)\} = \{Y(1), Y(2), \dots, Y(N)\}$ причём $M \ll N$.

Понижать размерность следует таким образом, чтобы сопутствующее этому увеличение ошибки классификации было относительно невелико. Классификатор, изображенный на рис. 23.1 является решающим устройством, которое обучается с целью классификации входного сигнала $x(t)$, принадлежащего к одному из K классов.

Принцип обучения наилучшим образом можно объяснить с помощью простого примера. Предположим, что мы желаем обучить классификатор автоматически классифицировать образ Z , принадлежащий либо классу C_1 ,

либо классу C_2 . Предположим также что обучающее множество (т.е. множество истинная классификация которого известна) состоит из следующего множества двумерных образов Z_{ij} , где Z_{ij} обозначает j -ый образ, принадлежащий $C_i, i=1,2$

$$C_1 : Z_{11} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}; Z_{12} = \begin{bmatrix} 6 \\ 5 \end{bmatrix}; Z_{13} = \begin{bmatrix} 6 \\ 6 \end{bmatrix}; Z_{14} = \begin{bmatrix} 6 \\ 7 \end{bmatrix}; Z_{15} = \begin{bmatrix} 7 \\ 5 \end{bmatrix}$$

$$C_2 : Z_{21} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}; Z_{22} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}; Z_{23} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}; Z_{24} = \begin{bmatrix} -3 \\ 3 \end{bmatrix}; Z_{25} = \begin{bmatrix} -4 \\ 3 \end{bmatrix}$$

Образы $\{Z_{ij}\}$, принадлежащие классам C_1 и C_2 , располагаются в двумерном пространстве признаков, как показано на рис. 23.2

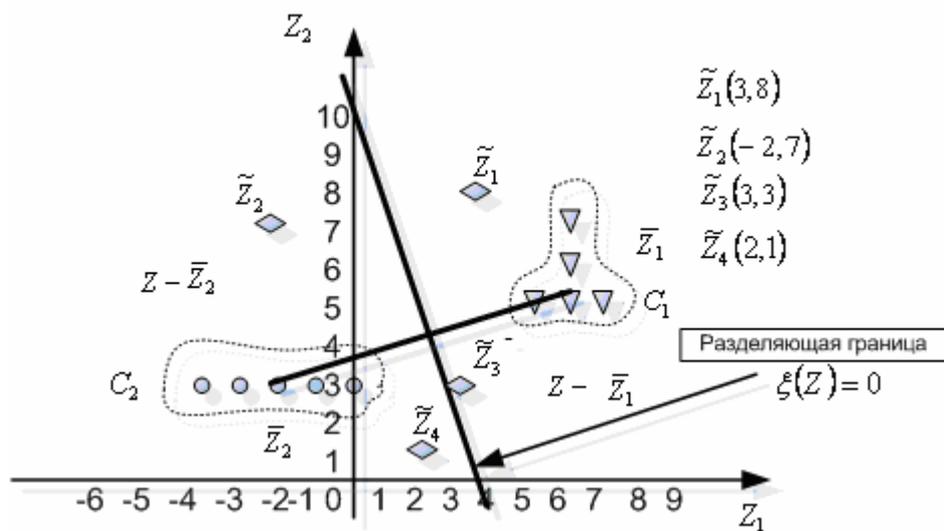


Рис. 23.2

Пусть \bar{Z}_1 и \bar{Z}_2 - средние векторы образов, связанные с C_1 и C_2 соответственно. Тогда

$$\bar{Z}_i = \frac{1}{5} \sum_{j=1}^5 Z_{ij}, i=1,2$$

$$\bar{Z}_1 = \begin{bmatrix} 6 \\ 5,6 \end{bmatrix}, \bar{Z}_2 = \begin{bmatrix} -2 \\ 3 \end{bmatrix} \quad (23.1)$$

Из рис. 23.2 видно, что наиболее целесообразной решающей границей (т.е. линией на плоскости), разделяющей классы C_1 и C_2 , является срединный перпендикуляр к прямой, соединяющей \bar{Z}_1 и \bar{Z}_2 . Следующим шагом является описание решающей границы с помощью уравнения.

Рассмотрим любую точку Z , принадлежащую решающей границе, как показано на рис. Так как решающая граница является срединным перпендикуляром к прямой, соединяющей \bar{Z}_1 и \bar{Z}_2 , то

$$\|Z - \bar{Z}_1\|^2 = \|Z - \bar{Z}_2\|^2 \text{ или в упрощенном виде}$$

$$(\bar{Z}_1 - \bar{Z}_2)' Z = \frac{1}{2} \left\{ \|\bar{Z}_1\|^2 - \|\bar{Z}_2\|^2 \right\} \quad (23.2)$$

Подставляя (23.1) в (23.2), получим уравнение для решающей границы в виде

$$8Z_1 + 2,6Z_2 = 27,18 \quad (23.3)$$

Величина $0,5 \left(\|\bar{Z}_1\|^2 - \|\bar{Z}_2\|^2 \right) = 27,18$ называется порогом классификатора.

Уравнение (2) даёт всю информацию, необходимую для создания классификатора. Основной характеристикой, представляющей классификатор, является дискриминантная функция $\xi(Z)$, которая определяется как

$$\xi(Z) = 8Z_1 + 2,6Z_2 - 27,18.$$

Возникает вопрос, что в действительности делает $\xi(Z)$? Чтобы ответить на этот вопрос, рассмотрим несколько испытательных образов (т.е. образов точная классификация которых неизвестна), которые обозначим как \tilde{Z}_k ($k = 1, 2, 3, 4$).

Вычислим теперь значение $\xi(Z)$ для каждого из испытательных образов

$$\tilde{Z}_1 : \xi(Z) = 8(3) + 2,6(8) - 27,18 > 0$$

$$\tilde{Z}_2 : \xi(Z) = 8(-2) + 2,6(7) - 27,18 < 0$$

$$\tilde{Z}_3 : \xi(Z) = 8(3) + 2,6(3) - 27,18 > 0$$

$$\tilde{Z}_4 : \xi(Z) = 8(2) + 2,6(1) - 27,18 < 0$$

Рассмотрение последнего выражения приводит к выводу, что когда образ \tilde{Z} лежит справа от решающей границы, то $\xi(Z) > 0$ и наоборот, когда образ \tilde{Z} лежит слева от решающей границы, то $\xi(Z) < 0$. Все точки лежащие справа от границы ближе к \tilde{Z}_1 , а точки, лежащие слева от границы ближе к \tilde{Z}_2 . Таким образом, благодаря дискриминантной функции $\xi(Z)$ получим следующее простое решающее правило:

если $\xi(Z) > 0$, то $Z \in C_1$

если $\xi(Z) < 0$, то $Z \in C_2$

Описанный выше классификатор, называемый элементом линейной пороговой логики, реализуется как показано на рис. 23.3.

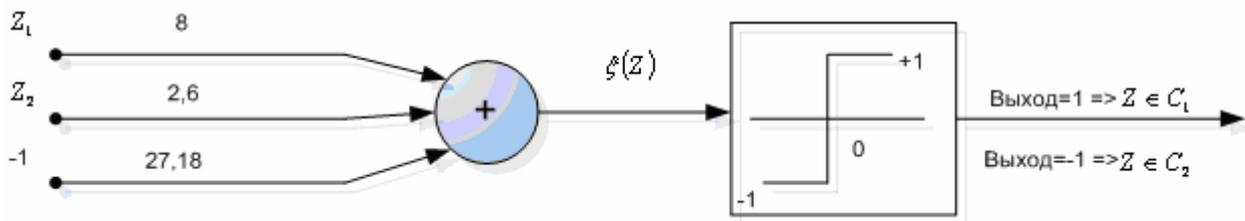


Рис. 23.3

Получив дискриминантную функцию $\xi(Z)$, говорят, что классификатор обучен, т.е. способен классифицировать образы с помощью соответствующего решающего правила. Элемент пороговой логики является классификатором, работающим по критерию минимума расстояния, так как решающее правило может быть сформулировано также следующим образом:

если Z ближе к \bar{Z}_1 , то $Z \in C_1$

и если Z ближе к $\overline{Z_2}$, то $Z \in C_2$

Рассмотрим элемент линейной пороговой логики, на который воздействуют $(d \times 1)$ образы, где $d > 2$. В этом случае $Z' = [Z_1 Z_2 \dots Z_d]$.

Соответствующая дискриминантная функция записывается как

$$\xi(Z) = w_1 Z_1 + w_2 Z_2 + \dots + w_d Z_d - \theta$$

где w_i - веса или параметры классификатора, а θ - порог. Веса w_i и θ получаются из обучающего множества. Границей в этом случае является гиперплоскость, определяемая как

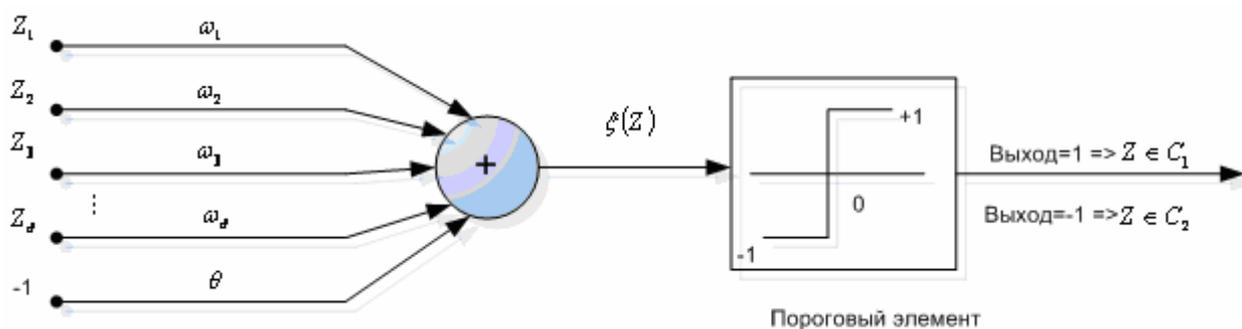


Рис. 23.4

24. ЗАДАЧА ТРЕХ КЛАССОВ ОБРАЗОВ ПО КРИТЕРИЮ МИНИМАЛЬНОГО РАССТОЯНИЯ

Ранее рассматривались аспекты так называемой задачи двух классов .
Распространим эти понятия на общий случай трёх классов: C_1, C_2, C_3 . Пусть обучающее множество записывается как

$$\begin{aligned}
 C_1 : Z_{11} &= \begin{bmatrix} 0 \\ 3 \end{bmatrix}; Z_{12} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}; Z_{13} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}; Z_{14} = \begin{bmatrix} -3 \\ 3 \end{bmatrix}; Z_{15} = \begin{bmatrix} -4 \\ 3 \end{bmatrix} \\
 C_2 : Z_{21} &= \begin{bmatrix} 5 \\ 5 \end{bmatrix}; Z_{22} = \begin{bmatrix} 6 \\ 5 \end{bmatrix}; Z_{23} = \begin{bmatrix} 6 \\ 6 \end{bmatrix}; Z_{24} = \begin{bmatrix} 6 \\ 7 \end{bmatrix}; Z_{25} = \begin{bmatrix} 7 \\ 5 \end{bmatrix} \\
 C_3 : Z_{31} &= \begin{bmatrix} 6 \\ -1 \end{bmatrix}; Z_{32} = \begin{bmatrix} 7 \\ 0 \end{bmatrix}; Z_{33} = \begin{bmatrix} 8 \\ 1 \end{bmatrix}; Z_{34} = \begin{bmatrix} 9 \\ 1 \end{bmatrix}; Z_{35} = \begin{bmatrix} 10 \\ 1 \end{bmatrix}
 \end{aligned}$$

В этом случае средние векторы образов равны

$$\bar{Z}_1 = \begin{bmatrix} -2 \\ 3 \end{bmatrix}; \bar{Z}_2 = \begin{bmatrix} 6 \\ 5,6 \end{bmatrix}; \bar{Z}_3 = \begin{bmatrix} 8 \\ 0,4 \end{bmatrix}$$

Классификатор, работающий по минимуму расстояния имеет следующее решающее правило : данный образ Z принадлежит C_i , если Z ближе всего к \bar{Z}_i , $i=1,2,3$.

Пусть D_i обозначает расстояние образа Z от \bar{Z}_i . Тогда получим

$$D_i^2 = \|Z - \bar{Z}_i\|^2 = (Z - \bar{Z}_i)^T (Z - \bar{Z}_i)$$

Упрощение D_i^2 приводит к

$$D_i^2 = Z^T Z - Z^T \bar{Z}_i - \bar{Z}_i^T Z + \bar{Z}_i^T \bar{Z}_i = \|Z\|^2 - 2 \left(\bar{Z}_i^T Z - \frac{1}{2} \|\bar{Z}_i\|^2 \right) \quad (24.1)$$

Очевидно, что D_i^2 минимально, когда величина $\left\{ \bar{Z}_i^T Z - \frac{1}{2} \|\bar{Z}_i\|^2 \right\}$ максимальна. Поэтому вместо вычисления D_i^2 по формуле (24.1) проще потребовать, чтобы в классификаторе вычислялось значение $\left\{ \bar{Z}_i^T Z - \frac{1}{2} \|\bar{Z}_i\|^2 \right\}$. Классификатор в этом случае описывается дискриминантными функциями.

$$\xi_i(Z) = \left\{ \bar{Z}_i^T Z - \frac{1}{2} \|\bar{Z}_i\|^2 \right\} \quad i = 1, 2, 3.$$

Подстановка численных значений \bar{Z}_i и $\|\bar{Z}_i\|^2$ приводит к

$$\begin{aligned} \xi_1(Z) &= -2Z_1 + 3Z_2 - 6,5 \\ \xi_2(Z) &= 6Z_1 + 5,6Z_2 - 33,68 \\ \xi_3(Z) &= 8Z_1 + 0,4Z_2 - 32,08 \end{aligned}$$

Таким образом, классификатор вычисляет три числа: $\xi_1(Z)$, $\xi_2(Z)$ и $\xi_3(Z)$, а затем сравнивает их. Классификатор относит Z к классу C_1 , если $\xi_1(Z)$ максимально, к классу C_2 , если $\xi_2(Z)$ максимально и к классу C_3 , если $\xi_3(Z)$ максимально.

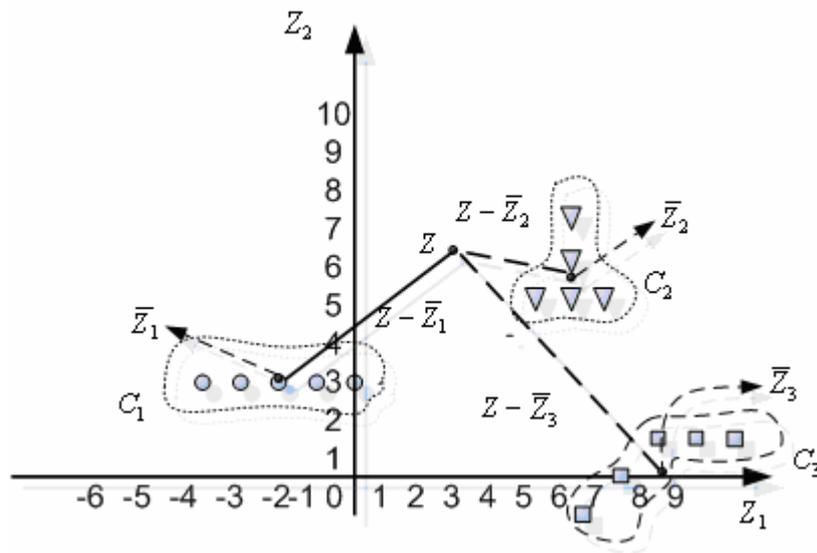


Рис. 24.1 Двумерное пространство признаков связанное с C_1, C_2, C_3 .

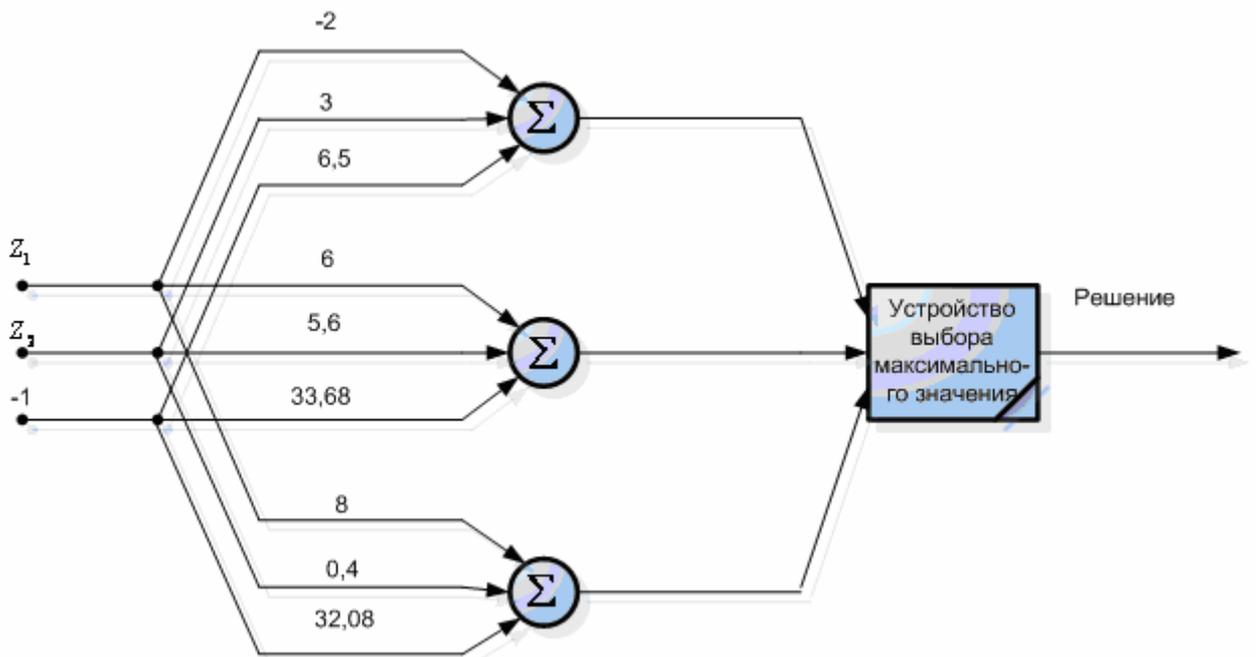


Рис. 24.2 Классификатор для трёх классов образов, работающий по критерию минимального расстояния.

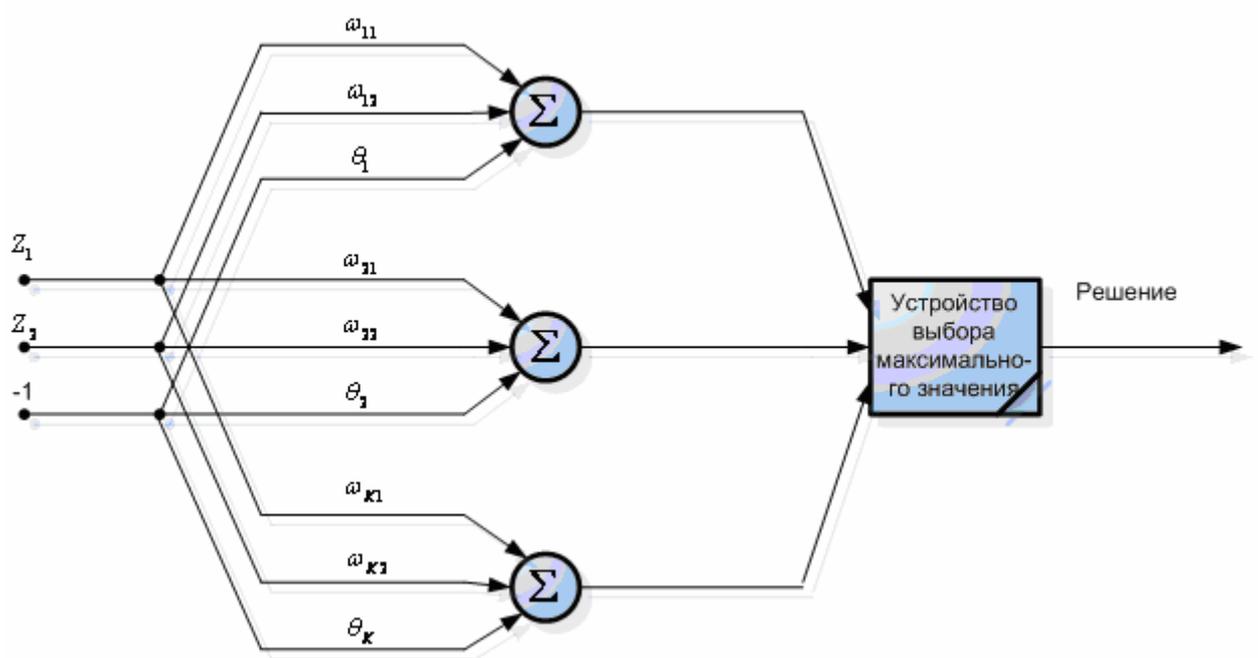


Рис. 24.3 Классификатор для K классов образов , работающий по критерию минимального расстояния.

25. МЕТОД ОТОБРАЖЕНИЯ ПО АЛГОРИТМУ НАИМЕНЬШИХ КВАДРАТОВ

При обсуждении классификаторов, работающих по критерию наименьшего расстояния, предполагалось, что классы образов в пространстве признаков группируются вокруг соответствующих им средних \bar{Z}_i ($i=1,2,\dots,K$). Однако, возможен и другой подход. При этом классификатор должен в первую очередь отображать образы в пространство решений, в котором образы, принадлежащие C_i обязательно группируются вокруг заранее выбранной точки V_i . Преобразование A , которое позволяет осуществлять это отображение из пространства признаков в пространство решений в общем случае выбирается таким, чтобы общая среднеквадратичная ошибка была минимальной. Для классификации некоторого образа этот образ сначала отображается в пространство решений, а затем классифицируется как принадлежащий C_{i_0} , если он отображён ближе к точке V_{i_0} .

Введём отображение по методу наименьших квадратов, на котором основываются классификаторы с минимальным среднеквадратичным расстоянием.

Рассмотрим множество M -мерных образов Z_{ij} , $j=1,2,\dots,N_i$, которые должны отображаться в определённую точку в K -мерном пространстве, обозначаемую $V_i = [v_1, v_2, \dots, v_k]$.

Найдём преобразование A , которое отображает $\{Z_{ij}\}$ в точку V_i , таким образом, чтобы общая среднеквадратичная ошибка, вызываемая отображением, была минимальной.

Обозначим результат отображения образа Z_{ij} через L_{ij} . Тогда соответствующий вектор ошибки будет равен

$$\varepsilon_j = L_{ij} - V_i = AZ_{ij} - V_i \quad (25.1)$$

Из выражения (25.1) следует , что общая среднеквадратичная ошибка при отображении Z_{ij} в V_i определяется как

$$\varepsilon = \frac{1}{N_i} \sum_{j=1}^{N_i} \|\varepsilon_j\|^2 \quad (25.2)$$

Подстановка (25.1) в (25.2) приводит к

$$\begin{aligned} \varepsilon &= \frac{1}{N_i} \sum_{j=1}^{N_i} \|AZ_{ij} - V_i\|^2 = \frac{1}{N_i} \sum_{j=1}^{N_i} (AZ_{ij} - V_i)' (AZ_{ij} - V_i) = \\ &= \frac{1}{N_i} \sum_{j=1}^{N_i} (AZ_{ij})' AZ_{ij} - 2(AZ_{ij})' V_i + V_i' V_i = \\ &= \frac{1}{N_i} \sum_{j=1}^{N_i} \left\{ Z_{ij}' A' AZ_{ij} - 2Z_{ij}' A' V_i + \|V_i\|^2 \right\} \end{aligned} \quad (25.3)$$

Так как A должно быть выбрано так , чтобы ε было минимальным, то оно получается в результате решения уравнения $\nabla_A \varepsilon = 0$,

что приведёт к

$$\frac{1}{N_i} \sum_{j=1}^{N_i} \nabla_A \left\{ Z_{ij}' A' AZ_{ij} \right\} - \frac{2}{N_i} \sum_{j=1}^{N_i} \nabla_A \left\{ Z_{ij}' A' V_i \right\} + \frac{1}{N_i} \sum_{j=1}^{N_i} \nabla_A \left\{ \|V_i\|^2 \right\} = 0 \quad (25.4)$$

Поскольку

$$\begin{aligned} \nabla_A \left\{ Z_{ij}' A' AZ_{ij} \right\} &= 2A \left(Z_{ij} Z_{ij}' \right), \\ \nabla_A \left\{ Z_{ij}' A' V_i \right\} &= V_i Z_{ij}', \\ \nabla_A \left\{ \|V_i\|^2 \right\} &= 0 \end{aligned}$$

то подстановка в (25.4) приведёт к

$$A \left[\frac{2}{N_i} \sum_{j=1}^{N_i} (Z_{ij} Z_{ij}') \right] = \frac{2}{N_i} \sum_{j=1}^{N_i} V_i Z_{ij}' ,$$

что позволяет определить A как

$$A = \left[\sum_{j=1}^{N_i} V_i Z'_{ij} \right] \left[\sum_{j=1}^{N_i} (Z_{ij} Z'_{ij}) \right]^{-1}$$

Рассмотрим пример . Пусть множество $\{Z_{ij}\}$ имеет вид

$$Z_{i1} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}; Z_{i2} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}; Z_{i3} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}; Z_{i4} = \begin{bmatrix} 6 \\ 6 \end{bmatrix}; Z_{i5} = \begin{bmatrix} 6 \\ 7 \end{bmatrix}$$

что соответствует $N_i = 5$. Пусть $V_i = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ тогда

$$\begin{aligned} \sum_{j=1}^5 V_i Z'_{ij} &= \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 5 & 5 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 5 & 6 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 5 & 7 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 6 & 6 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 6 & 7 \end{bmatrix} = \\ &= \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 5 & 6 \end{bmatrix} + \begin{bmatrix} 5 & 7 \\ 5 & 7 \end{bmatrix} + \begin{bmatrix} 6 & 6 \\ 6 & 6 \end{bmatrix} + \begin{bmatrix} 6 & 7 \\ 6 & 7 \end{bmatrix} = \begin{bmatrix} 27 & 31 \\ 27 & 31 \end{bmatrix} \\ \sum_{j=1}^5 Z_{ij} Z'_{ij} &= \begin{bmatrix} 25 & 25 \\ 25 & 25 \end{bmatrix} + \begin{bmatrix} 25 & 30 \\ 30 & 36 \end{bmatrix} + \begin{bmatrix} 25 & 35 \\ 35 & 49 \end{bmatrix} + \begin{bmatrix} 36 & 36 \\ 36 & 36 \end{bmatrix} + \begin{bmatrix} 36 & 42 \\ 42 & 49 \end{bmatrix} = \begin{bmatrix} 147 & 168 \\ 168 & 195 \end{bmatrix} \end{aligned}$$

Тогда

$$A = \begin{Bmatrix} 0,129 & 0,0475 \\ 0,129 & 0,0475 \end{Bmatrix}$$

Далее вычислим

$$L_{ij} = AZ_{ij}$$

$$L_{i1} = \begin{bmatrix} 0,883 \\ 0,883 \end{bmatrix}; L_{i2} = \begin{bmatrix} 0,930 \\ 0,930 \end{bmatrix}; L_{i3} = \begin{bmatrix} 0,978 \\ 0,978 \end{bmatrix}; L_{i4} = \begin{bmatrix} 1,06 \\ 1,06 \end{bmatrix}; L_{i5} = \begin{bmatrix} 1,11 \\ 1,11 \end{bmatrix}$$

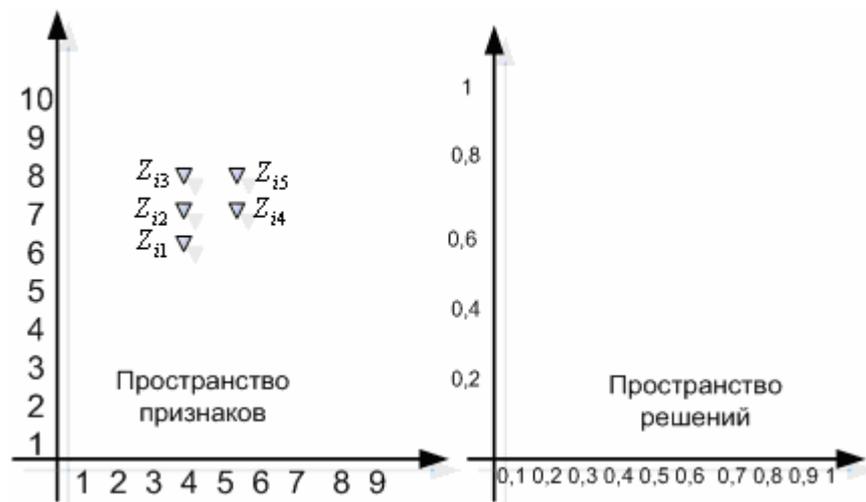


Рис. 25.1

Понятие расширенного пространства признаков необходимо при разработке классификатора с минимальным среднеквадратичным расстоянием.

Такое понятие вытекает из определения дискриминантной функции

$$\xi(Z) = w_1 Z_1 + w_2 Z_2 + \dots + w_d Z_d - \theta$$

В матричном виде это соотношение можно записать как

$$\xi(Z) = W' \hat{Z},$$

где $W' = [w_1 w_2, \dots, w_d \theta]$; $\hat{Z}' = [Z_1 Z_2, \dots, Z_d -1] = [Z' -1]$

Из последнего соотношения следует, что \hat{Z} можно получить из данного образа Z приписыванием к нему дополнительной координаты, равной -1. Пространство, состоящее из $(d+1)$ -мерных образов Z , называется расширенным пространством признаков.

26. КЛАССИФИКАТОР ДЛЯ РАСПОЗНАВАНИЯ ТРЁХ КЛАССОВ ОБРАЗОВ ПО КРИТЕРИЮ НАИМЕНЬШЕГО СРЕДНЕКВАДРАТИЧНОГО РАССТОЯНИЯ

Классификатор для распознавания трёх классов, работающий по критерию наименьшего среднеквадратичного расстояния можно представить иллюстративно следующим образом

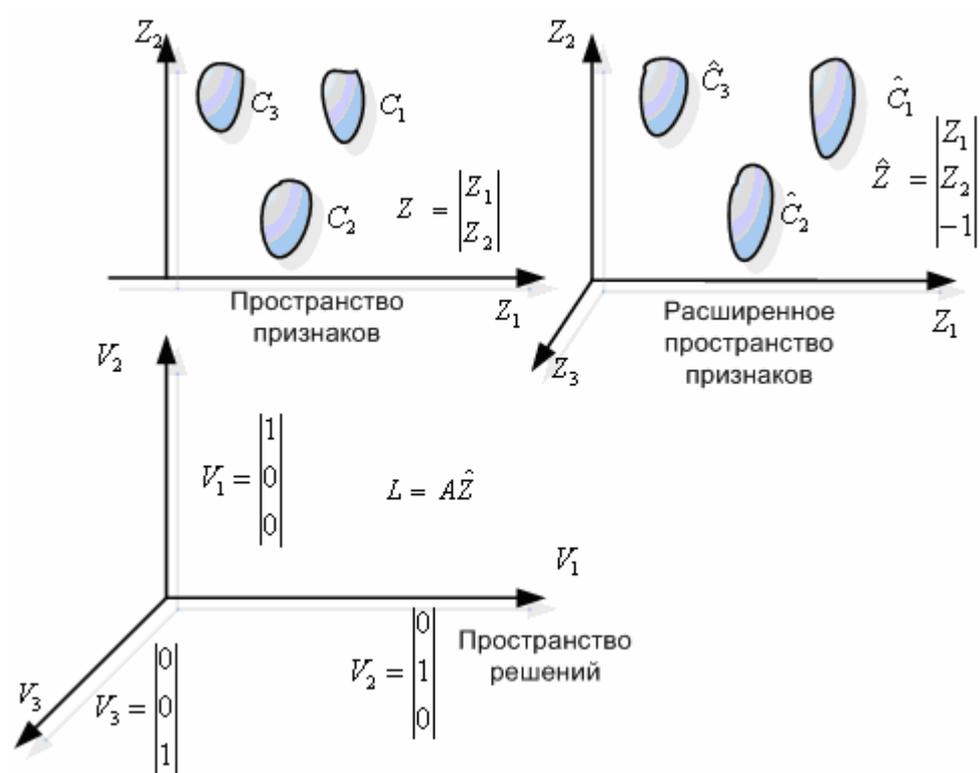


Рис. 26.1

Определяем преобразование A , отображающее $\hat{Z} \in C_k ; k = 1, 2, 3$ в V_k в смысле наименьших квадратов. Выберем в качестве вершины трёх единичных векторов, как показано на рис. Обозначим априорные вероятности классов C_k через P_k . Преобразование A можно непосредственно найти из соотношения

$$A \left[\frac{P_1}{N_1} \sum_{j=1}^{N_1} \hat{Z}_{1j} \hat{Z}'_{1j} + \frac{P_2}{N_2} \sum_{j=1}^{N_2} \hat{Z}_{2j} \hat{Z}'_{2j} + \frac{P_3}{N_3} \sum_{j=1}^{N_3} \hat{Z}_{3j} \hat{Z}'_{3j} \right] =$$

$$= \frac{P_1}{N_1} \sum_{j=1}^{N_1} V_1 \hat{Z}'_{1j} + \frac{P_2}{N_2} \sum_{j=1}^{N_2} V_2 \hat{Z}'_{2j} + \frac{P_3}{N_3} \sum_{j=1}^{N_3} V_3 \hat{Z}'_{3j}$$

Решая это уравнение относительно A , получим

$$A = S_{v\hat{z}} S_{\hat{z}\hat{z}}^{-1};$$

$$\text{где } S_{v\hat{z}} = \sum_{i=1}^3 \sum_{j=1}^{N_i} \frac{P_i}{N_i} (V_i \hat{Z}'_{ij}) = E\{V\hat{Z}\} \quad (26.1)$$

$$S_{\hat{z}\hat{z}} = \sum_{i=1}^3 \sum_{j=1}^{N_i} \frac{P_i}{N_i} (\hat{Z}_{ij} \hat{Z}'_{ij}) = E\{\hat{Z}\hat{Z}'\} \quad (26.2)$$

Очевидно, что $S_{v\hat{z}}$ и $S_{\hat{z}\hat{z}}$ являются матрицами взаимной корреляции и автокорреляции соответственно. Следует отметить, что матрица A имеет размерность $[3 \times (d+1)]$. Для классификации образа $\hat{Z}' = [Z_1 Z_2, \dots, Z_d - 1]$ классификатор вычисляет в первую очередь $L = A\hat{Z}$, а затем отображает \hat{Z} в пространство решений. При этом применяется следующее решающее правило по критерию минимума расстояния: если L наиболее близко к V_{i_0} , то Z классифицируется как принадлежащий C_{i_0} . Расстояния, которые вычисляет классификатор определяются как

$$D_i^2 = \|L - V_i\|^2, i = \overline{1,3}; \text{ т.е.}$$

$$D_i^2 = (L - V_i)'(L - V_i) = \|L\|^2 - 2V_i' L + \|V_i\|^2$$

Так как $\|V_i\|^2 = 1$, то D_i^2 будет минимально, когда $V_i' L$ максимально. Поэтому вместо D_i^2 в последнем соотношении достаточно, чтобы классификатор вычислял

$$d_i = V_i' L \quad (26.3)$$

Подставляя $V_1' = [1 \ 0 \ 0]$, $V_2' = [0 \ 1 \ 0]$, $V_3' = [0 \ 0 \ 1]$ и $L = A\hat{Z}$ в (26.3) получаем

$$\begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A\hat{Z} = A\hat{Z}$$

Матрица преобразования обозначается как

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1d} & \theta_1 \\ a_{21} & a_{22} & \dots & a_{2d} & \theta_2 \\ a_{31} & a_{32} & \dots & a_{3d} & \theta_3 \end{bmatrix},$$

где d_i представляют собой следующие дискриминантные функции, которые определяют классификатор

$$\xi(Z) = d_i = a_{i1}Z_1 + a_{i2}Z_2 + \dots + a_{id}Z_d - \theta_i.$$

Из последнего соотношения следует, что классификатор полностью определяется матрицей преобразования A , которая получается из множества обучающих образов. Для классификации данного образа классификатор вычисляет три числа $\xi_1(Z)$, $\xi_2(Z)$ и $\xi_3(Z)$.

Если $\max\{\xi_i(Z)\} = \xi_{i_0}(Z)$, то этот образ приписывается C_{i_0} .

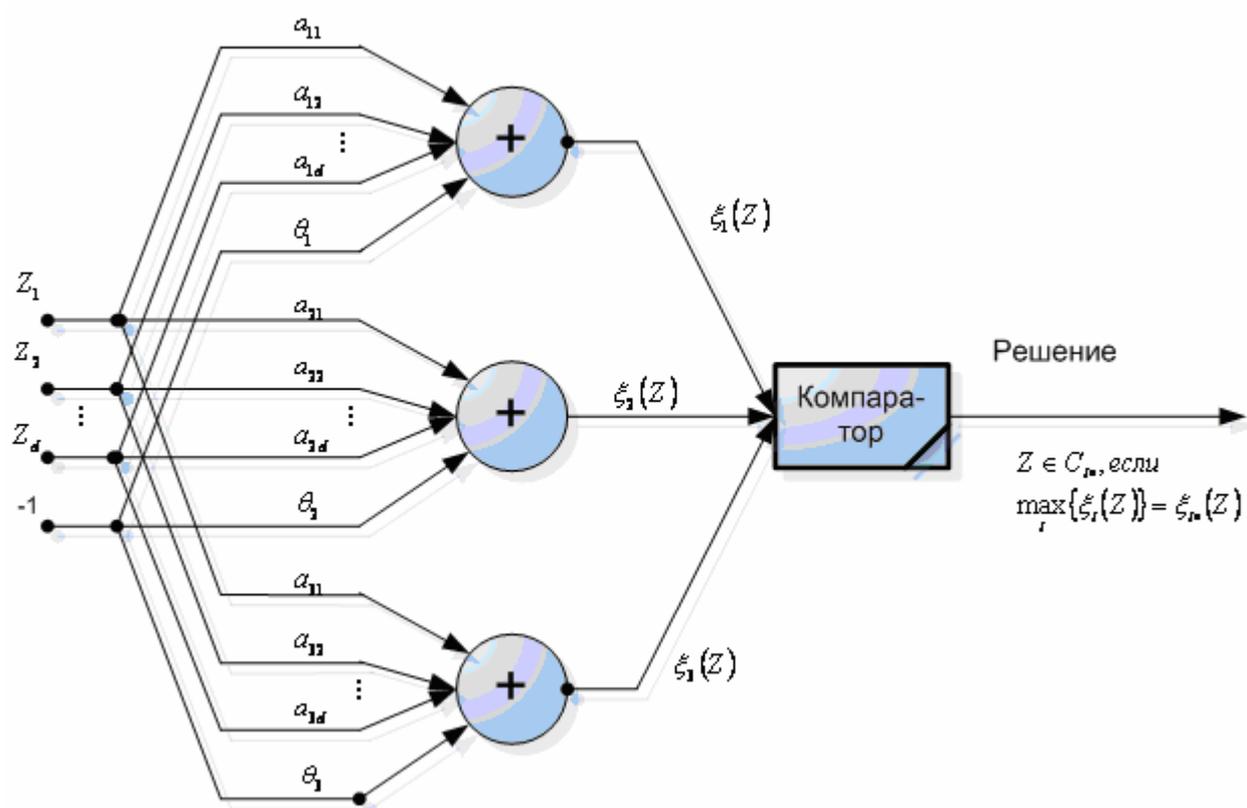


Рис. 26.2 Структура классификатора для трех классов

27. КЛАССИФИКАТОР ДЛЯ РАСПОЗНАВАНИЯ K КЛАССОВ ОБРАЗОВ ПО КРИТЕРИЮ НАИМЕНЬШЕГО СРЕДНЕКВАДРАТНОГО РАССТОЯНИЯ

Задача трёх классов может быть обобщена для случая K классов. Расширенные образы, принадлежащие C_i отображаются в вершине единичного K -вектора V_i

$$V_i = [0 \cdots 0 1 0 \cdots 0],$$

где ненулевой элемент «1» находится в i -ой позиции V_i , $i = \overline{1, K}$. Вместо матрицы преобразования A размерностью $3 \times (d+1)$ получаем матрицу размерностью $K \times (d+1)$

$$A = S_{vz} S_{zz}, \text{ где}$$

$$S_{vz} = \sum_{i=1}^K \sum_{j=1}^{N_i} P_i / N_i (V_i \hat{Z}_{ij}') = E(V \hat{Z}')$$

$$S_{zz} = \sum_{i=1}^K \sum_{j=1}^{N_i} P_i / N_i (\hat{Z}_{ij} \hat{Z}_{ij}') = E(\hat{Z} \hat{Z}')$$

Таким образом, дискриминантные функции, определяющие классификатор, имеют вид

$$\xi_i(z) = a_{i1} z_1 + a_{i2} z_2 + \cdots + a_{id} z_d - \theta_i, i = \overline{1, K}$$

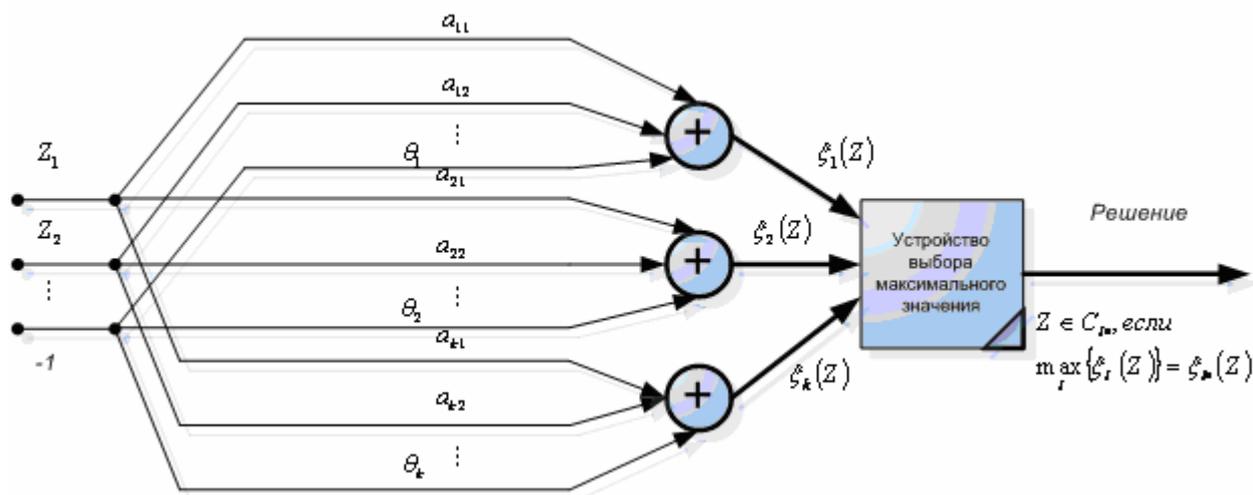


Рис. 27.1 Структура классификатора к классов

При обсуждении классификаторов по минимуму расстояния мы ограничивались рассмотрением только линейных классификаторов, т.е. классификаторов, которые обладают линейными разделяющими границами в пространстве признаков. Соответствующую процедуру обучения нетрудно получить для случая квадратичных разделяющих границ, но при этом реализация классификатора усложняется. Квадратичный классификатор определяется дискриминантными функциями вида

$$\xi_i(z) = \sum_{j=1}^d w_{ij} z_j^2 + \sum_{j=1}^{d-1} \sum_{k=j+1}^d w_{jk} z_j z_k + \sum_{j=1}^d w_j z_j - \theta_i, i = \overline{1, K} \quad (27.1)$$

Из последнего соотношения следует, что квадратичная дискриминантная функция $\xi_i(z)$ имеет $[(d+1)/(d+2)]/2$ весовых коэффициентов или параметров, а именно:

d параметров, соответствующих коэффициентам при $z_j^2 - w_{jj}$

d параметров, соответствующих коэффициентам при $z_j - w_j$

$d(d-1)/2$ параметров, соответствующих коэффициентам при $z_j z_k - w_{jk}, j \neq k$

порог- θ_i

Приведённые выше параметры определяются в процессе обучения .

Для пояснения реализации классификатора, связанного с вычислением (27.1) определим Q -мерный вектор G , координаты которого f_1, f_2, \dots, f_Q являются функциями z_i , $i = \overline{1, d}$. Первые d координат вектора G имеют вид $z_1^2, z_2^2, \dots, z_d^2$, следующие $[d(d-1)]/2$ координаты всем парам $z_1z_2, z_1z_3, \dots, z_{d-1}z_d$; последние α координат представляют собой z_1, z_2, \dots, z_d .

Тогда реализация классификатора будет иметь вид

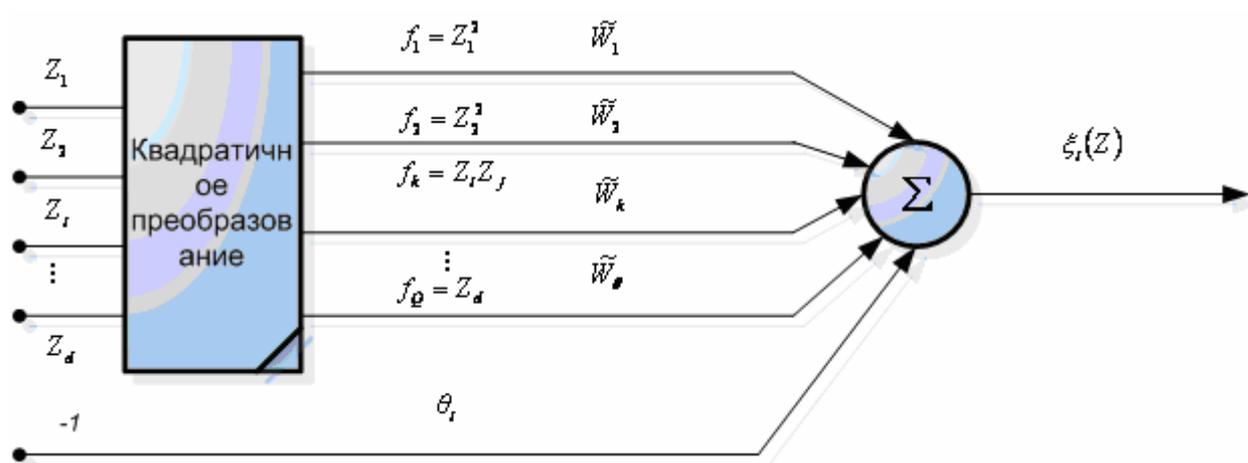
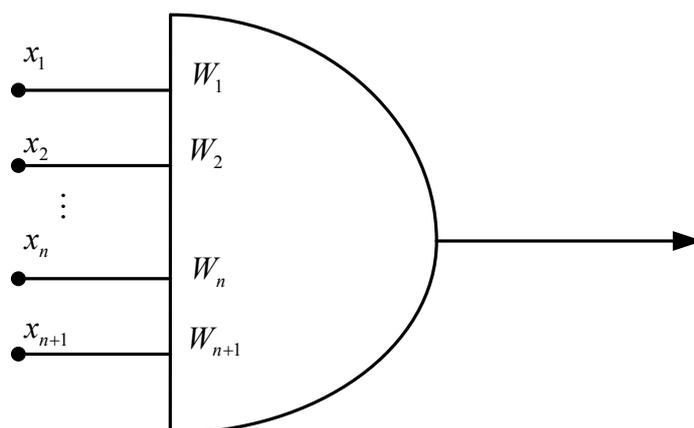


Рис. 27.2

$$\xi_i(z) = \tilde{w}_1 f_1 + \tilde{w}_2 f_2 + \dots + \tilde{w}_Q f_Q - \Theta_i$$

28. МОДЕЛЬ НЕЙРОННОЙ СЕТИ. ОБУЧАЮЩИЙ АЛГОРИТМ ДЛЯ ПЕРЦЕПТРОНА

При разбиении на два класса реализация классификатора может быть осуществлена с помощью порогового устройства



$$R = 1, \text{ если } W'X > T$$

$$R = -1, \text{ если } W'X < T$$

Рис. 28.1

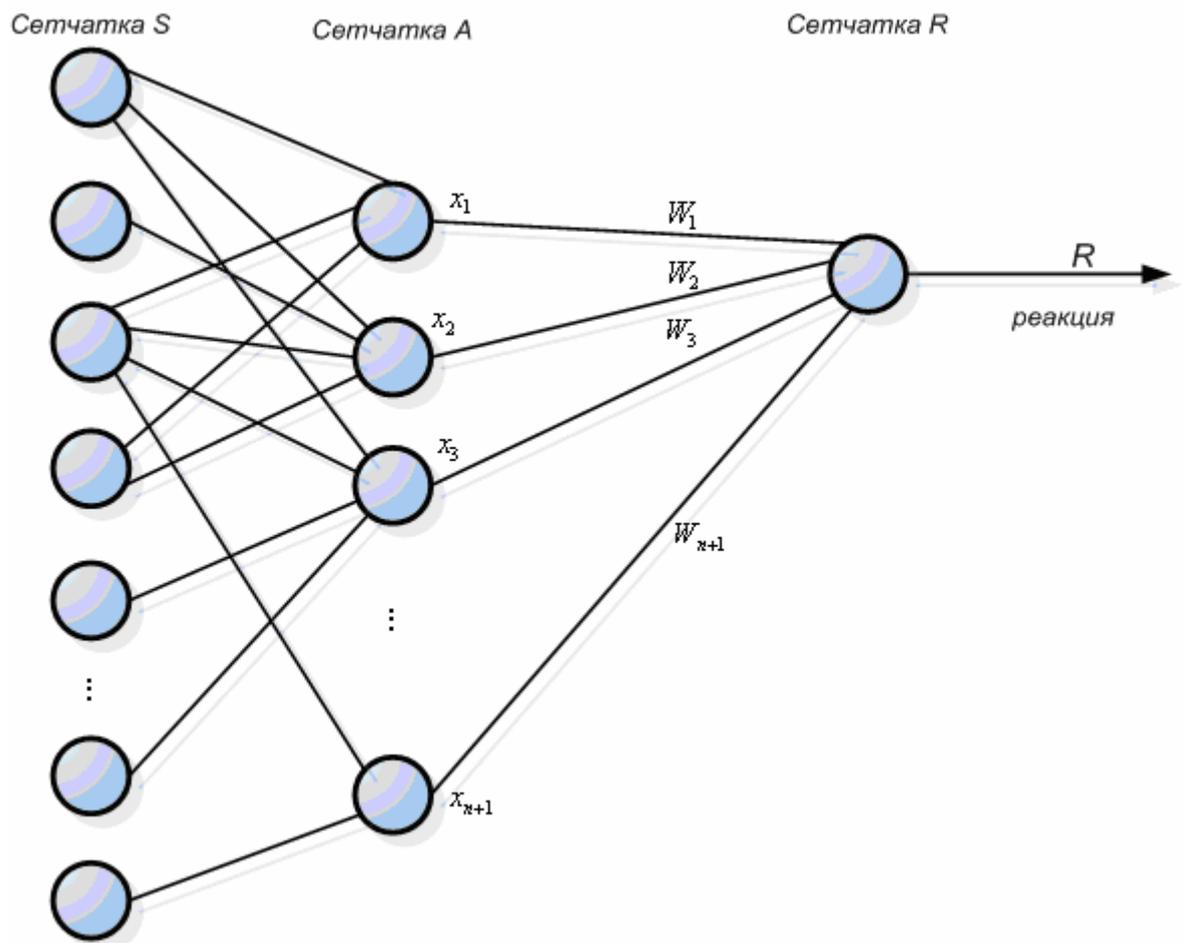
Эти две выходные величины (реакции порогового устройства) обычно обозначаются как 1 и -1 соответственно

Если заданы два множества образов, принадлежащих соответственно классам ω_1, ω_2 , то решение ищется в виде вектора весов W , обладающего тем свойством, что для всех образов класса ω_1 выполняется условие $W'X > 0$, а для всех образов класса ω_2 - условие $W'X < 0$. Если образы класса ω_2 умножить на -1 , то эквивалентное условие $W'X > 0$ становится общим для всех образов.

Причём класс устройств, предложенных в начале 80-х годов американским учёным Розенблаттом и называемых перцептронами,

представляет естественную и обладающую большими возможностями модель процесса обучения машин.

Основная модель пещетрона, обеспечивающая отнесение образа к одному из заданных классов приведена на рис. 28.2.



Устройство состоит из сетчатки S сенсорных элементов , которые случайным образом соединены с ассоциативными элементами второй сетчатки A . Каждый из элементов второй сетчатки воспроизводит выходной сигнал только в том случае ,если достаточное число сенсорных элементов, соединённых с его входом, находится в возбуждённом состоянии. Сенсорные элементы можно рассматривать в качестве устройств, с помощью которых вся система воспринимает из внешней среды стимулы ,т.е. как некие измерительные устройства ,а ассоциативные элементы –как входную часть системы.

Реакция всей системы пропорциональна сумме взятых с определёнными весами реакций элементов ассоциативной сетчатки. Таким образом, обозначив через x_i реакцию i -го ассоциативного элемента и через w_i - соответствующий вес, реакцию системы можно записать как

$$R = \sum_{i=1}^{n+1} W_i' x_i = W'X$$

Если $R > 0$, значит предъявленный системе образ принадлежит классу ω_1 , если $R < 0$, то образ относится к классу ω_2 .

Отметим, что перцептронная модель представляет собой за исключением сенсорной сетчатки не что иное, как реализацию линейной решающей функции.

Схему, приведённую на рис 28.2, легко распространить на случай разделения на несколько классов посредством увеличения числа реагирующих элементов в R - сетчатке.

Классификация проводится обычным способом: рассматриваются значения реакций R_1, R_2, \dots, R_M , где M - число классов и образ причисляется к классу ω_i , если $R_i > R_j$ для всех $i \neq j$.

Основную модель можно также легко распространить на случай нелинейных решающих функций введением соответствующих нелинейных преобразователей между сетчатками A и R .

Обучающий алгоритм для перцептрона сводится к простой схеме итеративного определения вектора весов W . Дадим краткое определение этой схемы, которую обычно называют алгоритмом перцептрона. Заданы два обучающие множества, представляющие классы ω_1 и ω_2 соответственно.

Пусть $W(1)$ начальный вектор весов, который выбирается произвольно. В таком случае k -ый шаг обучения выглядит следующим образом

Если $x(k) \in \omega_1$ и $W'(k)x(k) \leq 0$, то вектор весов $W(k)$ заменяется вектором

$$W(k+1) = W(k) + cx(k), (1) \quad (28.1)$$

где c -корректирующее приращение

Если $x(k) \in \omega_2$ и $W'(k)x(k) \geq 0$, то $W(k)$ заменяется вектором

$$W(k+1) = W(k) - cx(k), \quad (28.2)$$

в противном случае $W(k)$ не изменяется ,т.е.

$$W(k+1) = W(k)$$

Иными словами, алгоритм вносит изменения в вектор весов W в том и только том случае ,если образ предъявленный на k -м шаге был при выполнении этого шага неправильно классифицирован с помощью соответствующего вектора весов . Причём корректирующее приращение должно быть положительным .

Очевидно, что алгоритм перцептрона является процедурой типа «подкрепление - наказание» ,причём подкреплением за правильную классификацию образа служит отсутствие наказания .Иными словами, если образ классифицирован правильно, то система подкрепляется тем, что в вектор весов W не вносятся никаких изменений.

С другой стороны если образ классифицирован неправильно и произведение $W'(k)x(k)$ оказывается меньше нуля , когда оно должно быть больше нуля, система наказывается увеличением значения вектора весов $W(k)$ на величину, пропорциональную $x(k)$.Точно так же, если произведение $W'(k)x(k)$ оказывается больше нуля , когда оно должно быть меньше нуля, система наказывается противоположным образом.

Сходимость алгоритма наступает при правильной классификации всех образов с помощью некоторого вектора весов. Если заданные классы

линейно разделены, то алгоритм перцептрона сходится за конечное число итераций.

Пример: рассмотрим образы представленные на рис. 28.1 и 28.2

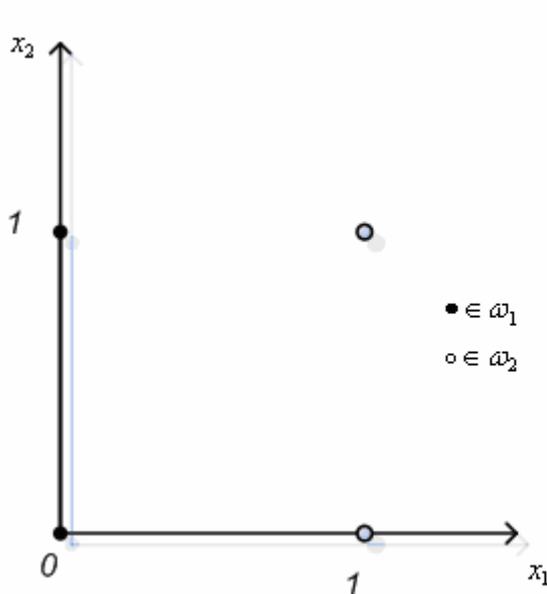


Рис. 28.1

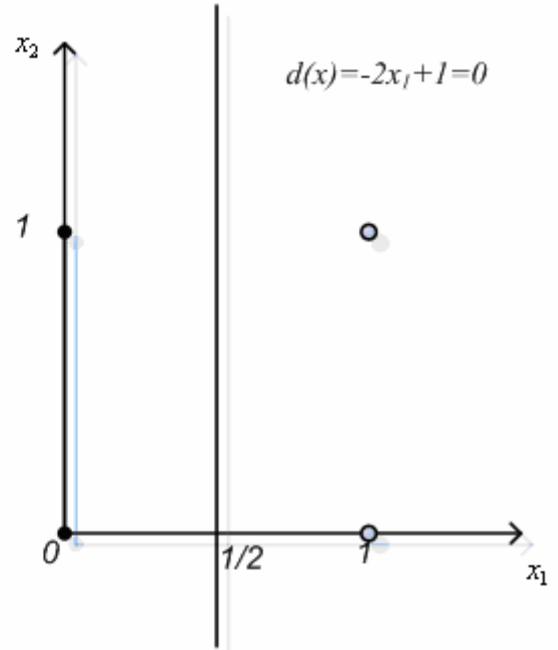


Рис. 28.2

Следует применить к этим образам алгоритм перцептрона с тем, чтобы с его помощью определить весовой вектор решения. Осмотр образов показывает, что два заданных класса линейно разделимы и следовательно применение алгоритма окажется успешным.

До начала работы алгоритма пополним все образы. При этом рассматриваемые классы обратятся

$$\begin{array}{ll} \omega_1: \{0\ 0\ 1\} & \omega_2: \{1\ 0\ 1\} \\ \omega_1: \{0\ 1\ 1\} & \omega_2: \{1\ 1\ 1\} \end{array}$$

Задав $c=1$, $W(1)=0$ и предъявив образы в указанном выше порядке, получим (по шагам)

$$W'(1)x(1) = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0$$

$$W(2) = W(1) + x(1) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$W'(2)x(2) = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = I$$

$$W(3) = W(2) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$W'(3)x(3) = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = I$$

$$W(4) = W(3) - x(3) = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$$

$$W'(4)x(4) = \begin{pmatrix} -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = -I$$

$$W(5) = W(4) = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$$

В связи с ошибками классификации коррекция вектора весов производилась на 1-м и 3-м в соответствии с формулами (28.1) и (28.2).

Так как получаемый результат можно считать искомым решением только в том случае, когда алгоритм осуществит без ошибок полный цикл

итераций по всем образам, обучающее множество следует предъявить ещё раз. Процесс обучения системы продолжается при $x_5=x_1, x_6=x_2, x_7=x_3, x_8=x_4$.

Второй цикл итерации приводит к следующим результатам:

$$W'(5)x(5) = \begin{pmatrix} -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0$$

$$W(6) = W(5) + x(5) = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

$$W'(6)x(6) = \begin{pmatrix} -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = 1$$

$$W(7) = W(6) = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

$$W'(7)x(7) = \begin{pmatrix} -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = -1 + 0 + 1 = 0$$

$$W(8) = W(7) - x(7) = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 0 \end{pmatrix}$$

$$W'(8)x(8) = \begin{pmatrix} -2 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = -2$$

$$W(9) = W(8) = \begin{pmatrix} -2 \\ 0 \\ 0 \end{pmatrix}$$

Поскольку в данном цикле итерации совершено две ошибки, все образы предъявляются ещё раз.

$$W'(9)x(9) = \begin{pmatrix} -2 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0$$

$$W(10) = W(9) + x(9) = \begin{pmatrix} -2 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$

$$W'(10)x(10) = \begin{pmatrix} -2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = -2 \times 0 + 0 \times 1 + 1 \times 1 = 1$$

$$W(11) = W(10) = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$

$$W'(11)x(11) = \begin{pmatrix} -2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = -2 \times 1 + 0 \times 0 + 1 \times 1 = -1$$

$$W(12) = W(11) = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$

$$W'(12)x(12) = \begin{pmatrix} -2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = -1$$

$$W(13) = W(12) = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$

Поскольку в данном цикле итераций совершена одна ошибка, все образы предъявляется ещё раз.

$$W'(13)x(13) = \begin{pmatrix} -2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 1$$

$$W(14)=W(13)=\begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$

$$W'(14)x(14)=(-2 \ 0 \ 1)\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}=I$$

$$W(15)=W(14)=\begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$

$$W'(15)x(15)=(-2 \ 0 \ 1)\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}=-2 \times I + I \times I = -I$$

$$W(16)=W(15)=\begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$

$$W'(16)x(16)=(-2 \ 0 \ 1)\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}=-2 \times I + I \times I = -I$$

$$W(17)=W(16)=\begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$

Итак за четыре итерации все образы классифицируются правильно.
Вектор решения имеет вид

$$W = (-2 \ 0 \ 1)$$

Соответствующей решающей функцией будет

$$d(x) = -2x_1 + I$$

Приравнивание этой функции к нулю даёт уравнение разделяющей границы.

Умножив все образы класса ω_2 на -1, алгоритм перцептрона можно записать как

$$W(k+1) = \begin{cases} W(k), & \text{если } W'(k)x(k) > 0 \\ W(k) + cx(k), & \text{если } W'(k)x(k) \leq 0 \end{cases}$$

где c -положительное корректирующее приращение.

Обобщающий алгоритм перцептрона

Рассмотрим M классов $\omega_1, \omega_2, \dots, \omega_M$.

Пусть на k -ой итерации процедуры обучения системе предъявляется образ $x(k)$, принадлежащий классу ω_i . Вычисляются значения M решающих функций

$$d_j[x(k)] = W_j'(k)x(k); \quad j = 1, 2, \dots, M.$$

Затем если выполняются условия

$$d_i[x(k)] > d_j[x(k)], \quad j = \overline{1, M} \quad i \neq j$$

то векторы весов не изменяются, т.е.

$$W_j(k+1) = W_j(k),$$

Допустим с другой стороны, что для некоторого l

$$d_l[x(k)] \leq d_l[x(k)]$$

В этом случае производятся следующие коррекции весов

$$W_i(k+1) = W_i(k) + cx(k)$$

$$W_l(k+1) = W_l(k) - cx(k)$$

$$W_j(k+1) = W_j(k)$$

$$j = 1, 2, \dots, M; j \neq i; j \neq l$$

где c -положительная константа.

Если при рассмотрении этого случая классы разделимы, то этот алгоритм сходится за конечное число итераций при произвольных начальных векторах $W_i(1)$.

29. ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ

29.1. ОБЩИЕ ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ

Приведем одно из определений искусственных нейронных сетей:

Искусственная нейронная сеть (ИНС) – это существенно параллельно распределенный процессор, который обладает способностью к сохранению и репрезентации опытного знания. Она сходна с мозгом в двух аспектах:

- знание приобретает сеть в процессе обучения;
- для сохранения знания используются силы межнейронных соединений, называемые также синаптическими соединениями.

Работа нейронной сети (НС) состоит в преобразовании входных сигналов во времени, в результате чего меняется внутреннее состояние сети и формируются выходные воздействия. Обычно НС оперирует цифровыми, а не символьными величинами. Большинство моделей НС требуют обучения. В общем случае, *обучение* — это такой выбор параметров сети, при котором сеть лучше всего справляется с поставленной проблемой. Обучение — это задача многомерной оптимизации и для ее решения существует множество алгоритмов.

Современные искусственные НС демонстрируют такие ценные свойства, как:

1. *Обучаемость*. Выбрав одну из моделей НС, создав сеть и выполнив алгоритм обучения, мы можем обучить сеть решению задачи, которая ей по силам.

2. *Способность к обобщению*. После обучения сеть становится нечувствительной к малым изменениям входных сигналов (шуму или вариациям входных образов) и дает правильный результат на выходе.

3. *Способность к абстрагированию.* Если предъявить сети несколько искаженных вариантов входного образа, то сеть сама может создать на выходе идеальный образ, с которым она никогда не встречалась.

К задачам, успешно решаемым НС на данном этапе их развития относятся:

- распознавание зрительных, слуховых образов;
- ассоциативный поиск информации и создание ассоциативных моделей, синтез речи, формирование естественного языка;
- формирование моделей и различных нелинейных и трудно описываемых математически систем, прогнозирование развития этих систем во времени;
- системы управления и регулирования с предсказанием;
- разнообразные конечные автоматы: системы массового обслуживания и коммутации, телекоммуникационные системы;
- принятие решений и диагностика, исключая логический вывод, особенно в областях, где отсутствуют четкие математические модели: в медицине, криминалистике, финансовой сфере.

На рис. 29.1 показана структура пары типичных биологических нейронов. Дендриты идут от тела нервной клетки к другим нейронам, где они принимают сигналы в точках соединения, называемых синапсами. Принятые синапсом входные сигналы подводятся к телу нейрона. Здесь они суммируются, причем одни входы стремятся возбудить нейрон, другие – воспрепятствовать его возбуждению. Когда суммарное возбуждение в теле нейрона превышает некоторый порог, нейрон возбуждается, посылая по аксону сигнал другим нейронам. У этой основной функциональной схемы много усложнений и исключений, тем не менее большинство искусственных нейронных сетей моделируют лишь эти простые свойства.

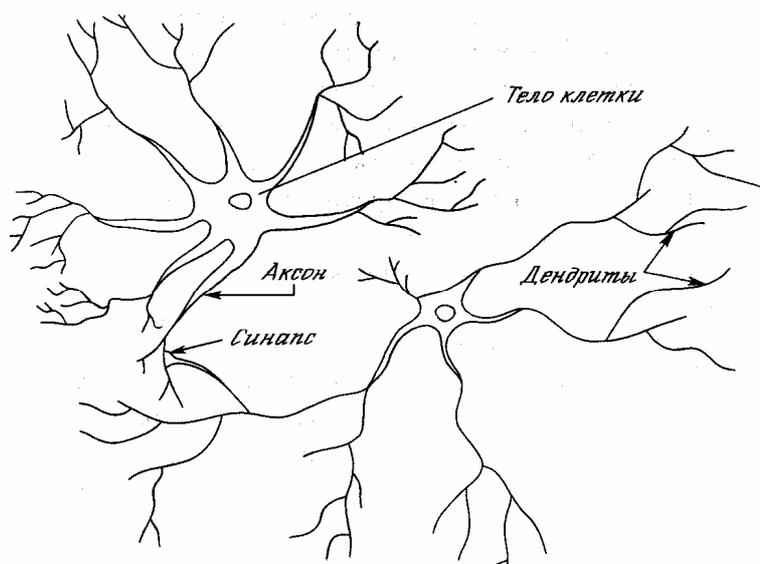


Рис. 7.1. Биологический нейрон

Основным элементом нейронной сети является **нейрон**, который осуществляет операцию нелинейного преобразования суммы произведений входных сигналов на весовые коэффициенты:

$$y = f\left(\sum_{i=1}^n w_i x_i + T\right), \quad (29.1)$$

где $X=(x_1, x_2, \dots, x_n)$ – вектор входного сигнала; $W=(w_1, w_2, \dots, w_n)$ – весовой вектор; T - порог; f – так называемая функция активации.

Схема нейронного элемента изображена на рис 29.2. Каждому i -му входу нейрона соответствует весовой коэффициент w_i (синапс), который характеризует силу синаптической связи по аналогии с биологическим нейроном.

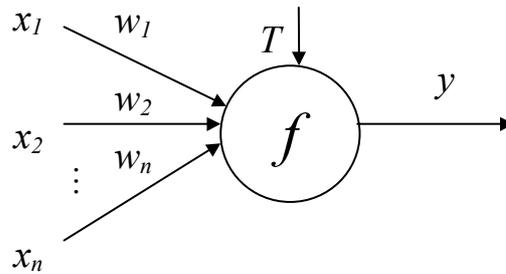


Рис. 7.2. Искусственный нейрон

Весовой вектор W , порог T и пороговая функция f определяют поведение любого нейрона – его реакцию на входные данные. Величина веса w_i определяет степень влияния входа i на выход нейрона, а знак – характер влияния. Каждый из входов связан с некоторым источником информации (рецептор, формирующий признаки, распределительная ячейка, выход другого нейрона). Положительные веса характерны для возбуждающих связей, способствующих повышению активности нейрона. Отрицательные, как правило, соответствуют тормозящим связям.

Порог, если он используется, является характеристикой, задающей начальный уровень активности (при нулевом входе) и помогающей настроить нейрон на пороговую функцию. Изменение порога эквивалентно сдвигу пороговой функции по оси абсцисс. Ряд авторов вводят дополнительный вход нейрона x_0 , всегда равный 1 , и обозначают порог как его вес w_0 . Это позволяет упростить выражение (29.1) и математическую запись некоторых алгоритмов обучения. Однако на практике при программной реализации это не приводит к экономии времени и способствует ошибкам, кроме этого, порог может настраиваться иначе, чем весовой вектор.

Функция активации используется для ограничения выхода нейрона в заданном диапазоне и для нелинейного преобразования взвешенной суммы. Последнее позволяет нейронному классификатору аппроксимировать любую нелинейную границу между классами в пространстве образов. Функция

активации выбирается для конкретной задачи и является неизменной характеристикой отдельного нейрона.

Могут использоваться следующие функции активации и их гибриды:

1) линейная функция $y = Ax$;

2) пороговая функция $y = \begin{cases} 1, x > 0 \\ 0, x \leq 0 \end{cases}$;

3) биполярная пороговая функция $y = \begin{cases} 1, x > 0 \\ -1, x \leq 0 \end{cases}$;

4) сигмоидная функция (рис. 29.2)

$$y = \frac{1}{(1 + e^{-x})}; \quad (29.2)$$

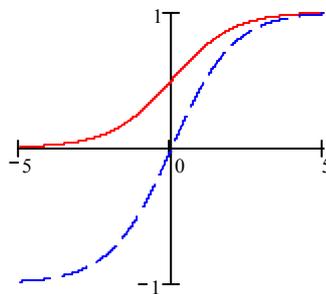


Рис. 7.2. Сигмоидные функции активации

5) биполярная сигмоидная функция (см. рис. 29.2)

$$y = \frac{2}{(1 + e^{-x})} - 1; \quad (29.3)$$

6) гиперболический тангенс (рис. 29.3)

$$y = th(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \quad (29.4)$$

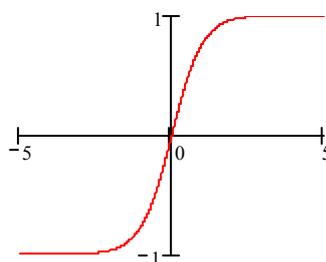


Рис. 7.3. Гиперболический тангенс

Для классификаторов чаще всего применяют функции (29.2) – (29.4), поскольку они нелинейные и хорошо дифференцируются.

Веса и порог конкретного нейрона являются настраиваемыми параметрами. Они содержат знания нейрона, определяющие его поведение. Процесс настройки этих знаний с целью получения нужного поведения называется обучением.

Нейронная сеть – совокупность нейронных элементов и связей между ними. Слоем нейронной сети называется множество нейронных элементов, на которые в каждый такт времени параллельно поступает информация от других нейронных элементов сети. Помимо слоев нейронов часто используется понятие входного распределительного слоя. Распределительный слой передает входные сигналы на первый обрабатывающий слой нейронных элементов (рис. 29.4).

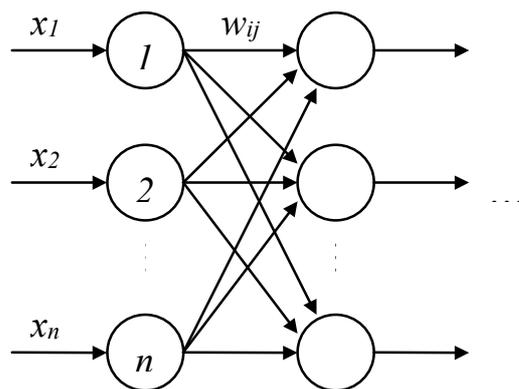


Рис. 29.4. Распределительный слой

У любой нейронной сети можно выделить 2 режима работы: обучение и воспроизведение. На этапе обучения настраиваются веса и пороги всех слоев. Воспроизведение является этапом обработки информации, следующим за обучением, при этом веса и пороги, как правило, не изменяются. Большое влияние на функционирование сети оказывает ее топология – архитектура слоев и связей между нейронами.

На способе обработки информации решающим образом сказывается наличие или отсутствие в сети обратных связей. Если обратных связей нет (каждый нейрон получает информацию только от нейронов предыдущих слоев), то обработка информации происходит в одном направлении за число тактов, равное числу слоев. При наличии обратных связей информация может проходить через сеть много раз до достижения какого-либо условия. В случае, если условие достигнуто, говорят, что сеть стабилизировалась. В общем случае сходимости не гарантируется. Тем не менее наличие обратных связей позволяет решать задачи с привлечением меньшего числа нейронов, что ускоряет процесс обучения.

Для лучшего понимания данного вопроса приведем одну из возможных классификаций НС в зависимости от различных характеристик:

1. По типу входной информации:
 - аналоговые НС (используют информацию в форме действительных чисел);
 - двоичные НС (оперируют с информацией, представленной в двоичном виде);
2. По характеру обучения:
 - с учителем (известно входное пространство решений НС);
 - без учителя (НС формирует выходное пространство решений только на основе входных воздействий – самоорганизующиеся сети);
3. По характеру настройки синапсов:
 - сети с фиксированными связями (весовые коэффициенты НС выбираются сразу, исходя из условия задачи);
 - сети с динамическими связями (в процессе обучения происходит настройка синаптических связей);
4. По методу обучения
 - НС с алгоритмом обратного распространения ошибки;
 - НС с конкурентным обучением;
 - НС, использующие правило Хебба;
 - НС с гибридным обучением, в которых используются различные алгоритмы обучения;
5. По характеру связей:
 - НС с прямыми связями;
 - НС с обратным распространением информации;
6. По архитектуре и обучению:
 - персептронные сети с прямыми связями;
 - самоорганизующиеся НС (НС Кохонена, НС адаптивного резонанса, рециркуляционные НС);

- НС с обратными связями (НС Хопфилда, НС Хэмминга, двунаправленная ассоциативная память, рекуррентные НС);
- гибридные НС (НС встречного распространения, НС с радиально-базисной функцией активации).

29.2. НЕЙРОННАЯ СЕТЬ ХОПФИЛДА

Сеть Хопфилда – однослойная, симметричная, нелинейная, автоассоциативная нейронная сеть, которая запоминает бинарные / биполярные образы. Сеть характеризуется наличием обратных связей. Топология сети Хопфилда показана на рис. 7.5. Информация с выхода каждого нейрона поступает на вход всех остальных нейронов. Образы для данной модификации сети Хопфилда кодируются биполярным вектором, состоящим из 1 и -1 .

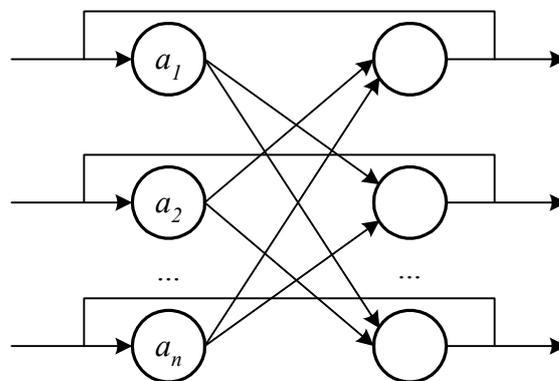


Рис. 7.5. Топология сети Хопфилда

Обучение. Обучение сети осуществляется в соответствии с соотношением

$$w_{ij} = \begin{cases} \sum_{k=1}^m a_i^k a_j^k, & i \neq j, \text{ для } i, j = \overline{1, n}, \\ 0, & i = j \end{cases} \quad (29.5)$$

где w_{ij} – вес связи от i -го нейрона к j -му;

n – количество нейронов в сети;

m – количество образов, используемых для обучения сети;

a_i^k – i -й элемент k -го образа из обучающей выборки.

Матрица весовых коэффициентов

$$W = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2n} \\ \dots & \dots & \dots & \dots \\ w_{n1} & w_{n2} & \dots & w_{nn} \end{bmatrix}. \quad (29.6)$$

В качестве матрицы весовых коэффициентов Хопфилд использовал симметричную матрицу ($w_{ij}=w_{ji}$) с нулевой главной диагональю ($w_{ii}=0$). Последнее условие соответствует отсутствию обратной связи нейронного элемента на себя. В качестве функции активации нейронных элементов может использоваться как пороговая, так и непрерывная функция, например сигмоидная или гиперболический тангенс.

Будем рассматривать нейронную сеть Хопфилда с дискретным временем. Тогда при использовании пороговой функции активации она называется нейронной сетью с дискретным состоянием и временем. Нейронная сеть с непрерывной функцией активации называется нейронной сетью с непрерывным состоянием и дискретным временем. При использовании непрерывного времени модель Хопфилда называется непрерывной.

Для описания функционирования таких сетей Хопфилд использовал аппарат статистической физики. При этом каждый нейрон имеет два состояния активности $(1, -1)$, которые аналогичны значениям спина некоторой частицы. Весовой коэффициент w_{ji} можно интерпретировать как вклад поля j – частицы в величину потенциала i – частицы. Хопфилд показал, что поведение такой сети аналогично поведению лизингового спинового стекла. При этом он ввел понятие вычислительной энергии, которую можно интерпретировать в виде ландшафта с долинами и впадинами. Структура соединений сети определяет очертания ландшафта. Нейронная сеть выполняет вычисления, следуя по пути, уменьшающему вычислительную энергию сети. Это происходит до тех пор, пока путь не приведет на дно впадины. Данный процесс аналогичен скатыванию капли жидкости по склону, когда она минимизирует свою потенциальную энергию в поле тяготения. Впадины и долины в сети Хопфилда соответствуют наборам информации, которую хранит сеть. Если процесс начинается с приближенной или неполной информации, то он следует по пути, который ведет к ближайшей впадине. Это соответствует операции ассоциативного распознавания.

Матрица весов является диагонально симметричной, причем все диагональные элементы равны 0.

Воспроизведение. Нейронная сеть Хопфилда может функционировать синхронно и асинхронно. Для воспроизведения используется соотношение

$$a_i(t+1) = f\left(\sum_{j=1}^n w_{ij} a_j(t)\right), \quad (29.7)$$

где $a_j(t)$ – выход j -го нейрона в момент времени t , а f – бинарная / биполярная функция активации;

$$f(x) = \begin{cases} 1 & x > 0, \\ -1 & x \leq 0. \end{cases} \quad (29.8)$$

При работе в синхронном режиме на один такт работы сети все нейроны одновременно меняют состояние. В случае асинхронной работы состояние меняет только один случайный нейрон. Итерации продолжаются до тех пор, пока сеть не придет в стабильное состояние.

Во время воспроизведения исходным вектором $a(0)$ является некоторый тестовый образ, не совпадающий с образами из обучающей выборки. В процессе функционирования по формуле (29.8) сеть должна прийти в состояние, соответствующее образу из обучающей выборки, наиболее похожему на тестовый.

Максимальное количество образов, которое можно запомнить в матрице W не превышает

$$m = \frac{n}{2 \ln n + \ln \ln n}, \quad (29.9)$$

где n – количество нейронов, что следует отнести к недостаткам этой сети.

29.3. МНОГОСЛОЙНЫЙ ПЕРСЕПТРОН

Многослойный персептрон является сетью с прямым распространением сигнала (без обратных связей), обучаемой с учителем. Такая сеть способна аппроксимировать любую непрерывную функцию или границу между классами со сколь угодно высокой точностью. Для этого достаточно одного скрытого слоя нейронов с сигмоидной функцией активации (29.2) – (29.4), т.е. многослойный персептрон обычно состоит из 3

слоев: первого распределительного, второго скрытого и третьего выходного (рис. 29.6).

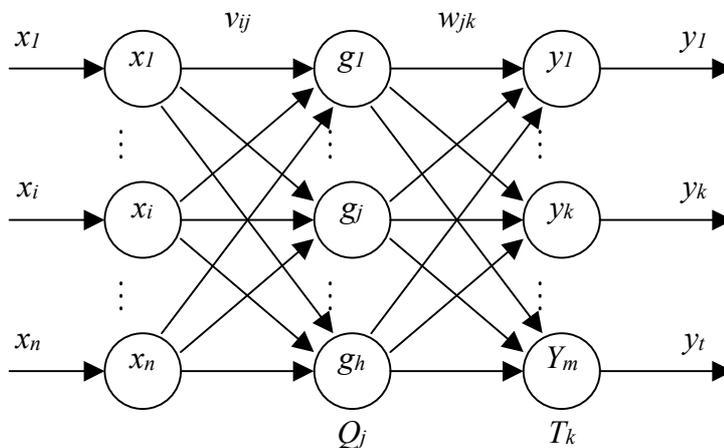


Рис. 29.6. Многослойный персептрон

Такая сеть имеет n входов и n нейронов распределительного слоя, h нейронов скрытого слоя и m выходных нейронов. Используются две матрицы весов: скрытого слоя v размером $n \times h$ и выходного слоя w размером $h \times m$. Кроме этого, с каждым слоем нейронов связан массив порогов: Q – для скрытого слоя, T – для выходного. Эти данные представляют собой знания сети, настраиваемые в процессе обучения и определяющие ее поведение. Персептрон функционирует по следующим формулам:

$$g_j = f\left(\sum_{i=1}^n v_{ij} x_i + Q_j\right), \quad (29.10)$$

$$y_k = f\left(\sum_{j=1}^h w_{jk} g_j + T_k\right). \quad (29.11)$$

В качестве функции активации используется одна из функций (29.2) – (29.4). Вид функции определяет диапазон чисел, в котором работает сеть. В дальнейшем будет использоваться сигмоидная функция (29.2), имеющая область значений от 0 до 1.

Обучение с учителем ставит перед сетью задачу обобщить p примеров, заданных парами векторов (x^r, y^r) , $r = \overline{1, p}$. Вектор $x^r = (x_1^r, x_2^r, \dots, x_n^r)$ в случае задачи классификации задает входной образ (вектор признаков), а вектор $y^r = (y_1^r, y_2^r, \dots, y_m^r)$, задающий эталонный выход, должен кодировать номер класса. При этом есть множество вариантов кодирования. Оптимальным представляется кодирование, когда номер класса определяется позицией единичной компоненты в векторе y^r , а все остальные компоненты равны 0. Каждый выходной нейрон соответствует одному классу. Такой способ позволяет при классификации определять вероятность каждого класса по величине на выходе соответствующего нейрона (чем ближе к единице, тем вероятность больше).

Обучение персептрона проводится с помощью алгоритма обратного распространения ошибки, который минимизирует среднеквадратичную ошибку нейронной сети. Для этого с целью настройки синаптических связей используется метод градиентного спуска в пространстве весовых коэффициентов и порогов нейронной сети. Рассмотрим алгоритм обратного распространения ошибки.

1. На первом этапе происходит начальная инициализация знаний сети. Простейший вариант такой инициализации – присвоить всем весам и порогам случайные значения из диапазона $[-1, 1]$.

2. Далее для каждой пары векторов (x^r, y^r) выполняется следующее:

- 2.1. Для входного вектора рассчитываются выходы нейронов скрытого слоя и выходы сети по формулам (29.10), (29.11).

2.2. Происходит коррекция знаний сети, при этом главное значение имеет отклонение реально полученного выхода сети y от идеального вектора y^r . Согласно методу градиентного спуска, изменение весовых коэффициентов и порогов нейронной сети происходит по следующим формулам:

$$w_{jk}(t+1) = w_{jk}(t) + \alpha \frac{\partial E}{\partial w_{jk}(t)}, \quad (29.12)$$

$$T_k(t+1) = T_k(t) + \alpha \frac{\partial E}{\partial T_k(t)}, \quad (29.13)$$

где E - среднеквадратичная ошибка нейронной сети для одного образа, а α – параметр, определяющий скорость обучения. Формулы записаны в терминах выходного слоя, аналогично выглядят формулы для скрытого слоя. Среднеквадратичная ошибка сети вычисляется как

$$E = \frac{1}{2} \sum_{k=1}^m (y_k^r - y_k)^2. \quad (29.14)$$

Ошибка k -го нейрона выходного слоя определяется как

$$d_k = \frac{\partial E}{\partial y_k} = y_k^r - y_k. \quad (29.15)$$

Выразим производные из формул (29.12), (29.13) через легко вычисляемые величины. Определим взвешенную сумму, аргумент функции активации как

$$S_k = \sum_{j=1}^h w_{jk} g_j + T_k. \quad (29.16)$$

$\frac{\partial E}{\partial w_{jk}}$ можно представить как:

$$\frac{\partial E}{\partial w_{jk}} = \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial S_k} \frac{\partial S_k}{\partial w_{jk}}, \quad (29.17)$$

где $\frac{\partial E}{\partial y_k} = d_k = y_k^r - y_k$ – ошибка k -го нейрона; $\frac{\partial y_k}{\partial S_k} = f'(S_k)$ – производная функции активации; $\frac{\partial S_k}{\partial w_{jk}} = g_j$ – значение j -го нейрона предыдущего слоя.

Получаем

$$\frac{\partial E}{\partial w_{jk}} = d_k f'(S_k) g_j. \quad (29.18)$$

Аналогично, с учетом того, что $\frac{\partial S_k}{\partial T_k} = 1$, получаем

$$\frac{\partial E}{\partial T_k} = \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial S_k} \frac{\partial S_k}{\partial T_k} = d_k f'(S_k). \quad (29.19)$$

Веса и пороги скрытого слоя также корректируются по формулам, аналогичным (29.12), (29.13), с учетом (29.18) и (29.19). При этом главной трудностью является определение ошибки нейрона скрытого слоя. Эту ошибку явно определить по формуле, аналогичной (29.15) нельзя, однако существует возможность рассчитать ее через ошибки нейронов выходного слоя (отсюда произошло название алгоритма обратного распространения ошибки):

$$e_j = \frac{\partial E}{\partial g_j} = \sum_{k=1}^m \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial S_k} \frac{\partial S_k}{\partial g_j} = \sum_{k=1}^m d_k f'(S_k) w_{jk}. \quad (29.20)$$

Производные от функций активации тоже легко рассчитываются, например, для сигмоидной функции получаем

$$f'(S_k) = \left(\frac{1}{1 + e^{-S}} \right)' = f(S_k)(1 - f(S_k)) = y_k(1 - y_k). \quad (29.21)$$

Аналогично:

$$(th(S_k))' = 1 - y_k^2;$$

$$\left(\frac{2}{1 + e^{-x}} - 1 \right)' = \frac{(1 - y_k^2)}{2}.$$

Таким образом, можно записать окончательные выражения (29.12), (29.13) для двух слоев, использующих сигмоидную функцию:

$$w_{jk} = w_{jk} + \alpha y_k (1 - y_k) d_k g_j, \quad (29.22)$$

$$T_k = T_k + \alpha y_k (1 - y_k) d_k, \quad (29.23)$$

$$v_{ij} = v_{ij} + \beta g_j (1 - g_j) e_j x_i, \quad (29.24)$$

$$Q_j = Q_j + \beta g_j (1 - g_j) e_j. \quad (29.25)$$

В эти формулы вводится дополнительный параметр – β – скорость обучения скрытого слоя, который может отличаться от аналогичного параметра для выходного слоя. Рекомендуется изменять скорости обучения обратно пропорционально количеству шагов 2 алгоритма обучения, однако это не всегда оправдывает себя на практике.

3. После того, как коррекция знаний произведена для каждой пары векторов, можно оценить степень успешности обучения сети для определения момента завершения алгоритма. Для этого можно использовать в качестве критерия максимальную по модулю ошибку на выходе d_k , полученную на шаге 2. Условием прекращения обучения в этом случае будет

$$\max |d_k| < D, k = \overline{1, m}, r = \overline{1, p}, \quad (29.26)$$

где D – достаточно маленькая константа – величина максимальной ошибки, которую требуется достичь в процессе обучения. Если условие (29.26) не выполняется, то шаг 2 повторяется.

Способность персептрона-классификатора разделять образы в пространстве признаков прежде всего зависит от его скрытого слоя. Именно на этот слой возлагается задача сделать множество классов линейно разделимым для успешной работы выходного слоя. Очевидно, что чем больше нейронов в скрытом слое, тем большее количество примеров этот слой может разделять. Кроме этого, увеличение числа признаков входных образов также способствует успешному их разделению в пространстве признаков. Однако увеличение этих параметров приводит к росту ошибок сети и времени обучения. Увеличение размерности входов n приводит к росту ошибки аппроксимации сети, возникающей из-за обобщения данных. Увеличение числа нейронов h скрытого слоя приводит к росту ошибки, связанной со сложностью модели. Персептрону легче провести функцию через эталонные точки, однако при этом обобщающая способность сети ухудшается. Он хуже предсказывает поведение аппроксимируемой функции

на образах, не входящих в обучающую выборку. Такое состояние сети называется переобучением.

Оптимальное соотношение между этими параметрами оценивают как

$$h \sim \sqrt{\frac{p}{n}}. \quad (29.27)$$

Эксперименты показывают, что обучение максимально успешно проходит на множестве классов, хорошо (желательно линейно) разделенных в пространстве признаков. Это достигается удачным подбором информативных признаков. Если классы в пространстве признаков хорошо кластеризуются, т.е. образы каждого класса составляют компактную группу, достаточно удаленную от других групп, то есть возможность уменьшить размер обучающей выборки p (используются только центры кластеров) и затем уменьшить число нейронов h . Это приводит к ускорению обучения и улучшает работу классификатора.

Другой проблемой является то, что алгоритм градиентного спуска не гарантирует нахождение глобального минимума среднеквадратичной ошибки сети (7.14), а гарантируется определение только локального минимума. Проблемы, возникающие в процессе градиентного спуска, можно проанализировать на примере функции ошибки, схематически изображенной на рис. 29.7.

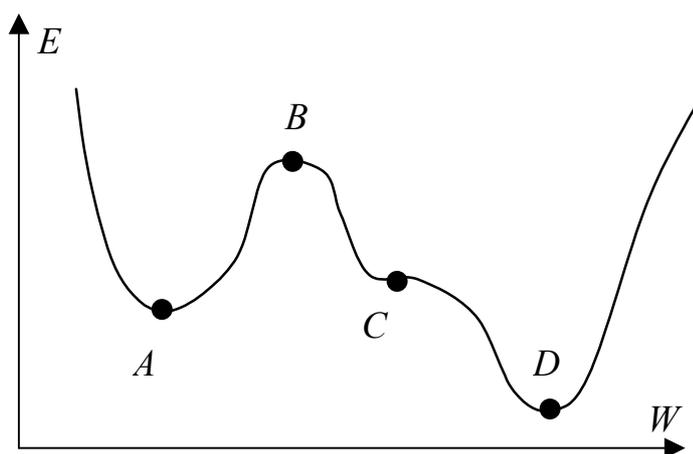


Рис. 29.7. Схематическая функция ошибки сети

На рис. 29.7 показаны четыре критические точки, производная функции ошибки в которых близка к нулю. Точка *A* соответствует локальному минимуму. Признаком достижения локального минимума в процессе обучения является полное прекращение уменьшения ошибки (29.26). В этом случае может помочь повторное обучение с другим начальным распределением знаний (тут может помочь случайная инициализация). Точка *B* – локальный максимум. В случае попадания в окрестность такой точки скорость резко падает, затем снова быстро растет. Не существует способа предсказать, в какую сторону лучше двигаться из точки с производной, близкой к нулю. Точка *C* – точка перегиба, характеризуется длительным уменьшением скорости. Точка *D* – глобальный минимум – цель алгоритма.

Существуют многочисленные способы оптимизации метода градиентного спуска, призванные улучшить поведение алгоритма в подобных критических точках. Эффективной модификацией является введение момента, накапливающего влияние градиента на веса со временем. Тогда величина Δw (29.12) в момент времени t будет вычисляться как

$$\Delta w(t) = \alpha \frac{\partial E}{\partial w} + \mu \Delta w(t-1), \quad (29.28)$$

где μ - параметр, определяющий величину влияния момента. С использованием (29.28) скорость изменения весов возрастает на участках с постоянным знаком производной. В окрестностях минимума скорость резко падает за счет колебания знака.

Достоинства алгоритма – большая скорость в точках перегиба, возможность по инерции преодолевать небольшие локальные минимумы. Недостатки – еще один параметр, величину которого следует подбирать и настраивать.

Этот и другие алгоритмы оптимизации обучения персептрона позволяют улучшить работу сети в условиях плохой сходимости. Однако они усложняют процесс обучения, не гарантируя в то же время полного успеха во всех случаях. Успех обучения классификатора зависит от самого алгоритма обучения и качества обучающей выборки.

29.4. СЕТЬ РБФ

Сеть РБФ (радиальная базисная функция) является аналогом многослойного персептрона (рис. 29.8). Скорость обучения такой сети гораздо выше, причем допускается полностью аналитический подход к расчету весовых коэффициентов. Однако эти положительные моменты сопровождаются рядом недостатков, главным из которых является ухудшение точности аппроксимации. Сеть обладает хорошей обобщающей способностью только для ограниченного класса аппроксимируемых функций. В качестве классификатора такая сеть может с успехом применяться в случае хорошей кластеризации классов в пространстве признаков.

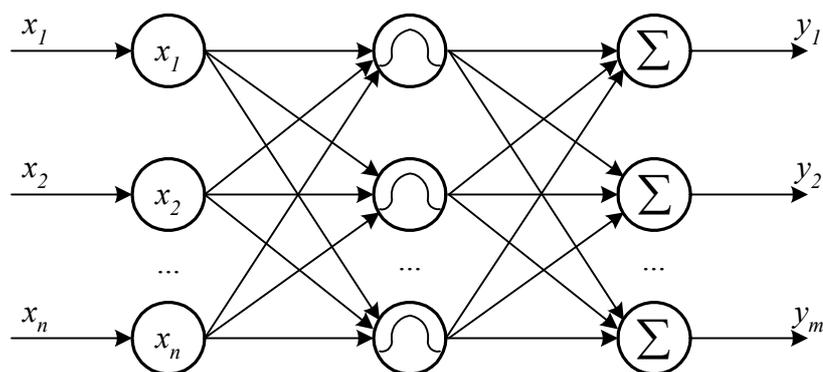


Рис. 29.8 Топология сети РБФ

Упрощение работы и обучения достигается за счет введения вместо скрытого слоя нейронов слоя РБФ ячеек. Классический закон, по которому такая ячейка функционирует, определяется формулой гауссова колокола:

$$g_j = \exp\left(\frac{-|x - t^j|^2}{\sigma_j^2}\right), \quad (29.29)$$

где x – входной вектор; t^j – вектор, определяющий математическое ожидание (центр кластера в пространстве признаков) РБФ ячейки; σ_j – среднеквадратическое отклонение или параметр, зависящий от величины разброса образов класса от его центра (рис. 29.8). В данном выражении евклидово расстояние между векторами x и t^j вычисляется как

$$|x - t^j|^2 = (x_1 - t_1^j)^2 + (x_2 - t_2^j)^2 + \dots + (x_n - t_n^j)^2.$$

Обучение. РБФ ячейки обучаются путем подбора центра и отклонения каждой из них. Для классификатора в качестве центра выбирается центр кластера в пространстве признаков, компактно содержащего образы одного и того же класса. В простейшем случае, если класс задается одним идеальным

образом, этот образ и будет являться вектором t – центром РБФ ячейки. Параметр разброса каждой ячейки выбирается в зависимости от величины радиуса кластера или расстояния до соседних центров. Ряд авторов рекомендует выбирать σ как половину расстояния до ближайшего центра ячейки, соответствующей другому классу. Количество РБФ ячеек выбирается таким образом, чтобы покрыть гауссовыми колоколами все классы.

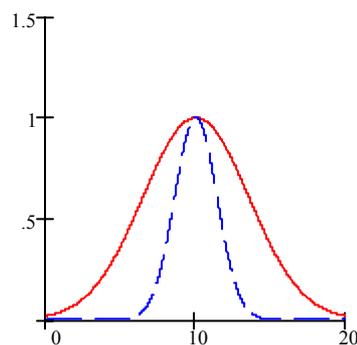


Рис. 7.9. Примеры функций РБФ ячеек с одинаковым центром и разным разбросом

Выходной слой РБФ сети обычно состоит из суммирующих ячеек

$$y_k = \sum_{j=1}^h w_{jk} g_j. \quad (29.30)$$

Это позволяет при определенных условиях использовать систему линейных уравнений для определения весов выходного слоя. В других обстоятельствах можно использовать алгоритм градиентного спуска для настройки весов выходного слоя (функция активации в данном случае линейная). С учетом того, что обучается только выходной слой нейронной сети, а скрытый уже настроен, обучение такой сети проходит на порядок быстрее, чем аналогичного многослойного персептрона.

Формула (29.22) упрощается до

$$w_{jk} = w_{jk} + \alpha d_k g_j, \quad (29.31)$$

поскольку функция активации в выходном слое сети РБФ линейная и ее производная равна 1. Для коррекции весовых коэффициентов выходных суммирующих ячеек используется только (29.31).

Контроль завершения алгоритма обучения производится аналогично.

Воспроизведение. Сеть функционирует по формулам (29.29) и (29.30).

Очевидно, что функция на выходе РБФ сети будет представлять собой суперпозицию гауссовых колоколов. В этом заключается ограничение данного класса классификаторов. Кроме этого, при неудачном выборе признаков большой проблемой является выбор количества РБФ ячеек, определение их центров и отклонения. С уменьшением числа РБФ ячеек улучшается обобщение данных сетью, но могут проявляться большие ошибки в эталонных точках.

29.5. КОНКУРЕНТНАЯ НЕЙРОННАЯ СЕТЬ

Самоорганизующиеся нейронные сети обучаются без учителя. Они способны адаптироваться к входным данным, используя содержащиеся в этих данных зависимости. Такие сети используются для нахождения более компактного описания данных (сжатия), кластеризации, выделения признаков.

Конкурентная сеть является простейшей самоорганизующейся нейронной сетью (рис. 29.10).

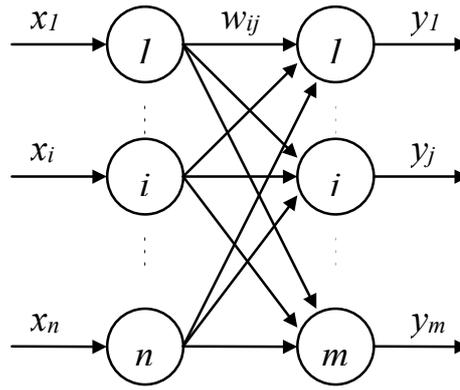


Рис. 29.10. Конкурентная нейронная сеть

Первый слой является распределительным. Нейроны второго слоя функционируют по формуле

$$y_j = \sum_{i=1}^n w_{ij} x_i = |w^j| |x| \cos \alpha, \quad (29.32)$$

где $x = (x_1, x_2, \dots, x_i, \dots, x_n)$ – входной вектор; $w^j = (w_{1j}, w_{2j}, \dots, w_{ij}, \dots, w_{nj})$ – вектор весовых коэффициентов нейрона, а $|x|$ и $|w^j|$ – их модули, α – угол между ними.

Обучение. При обучении нейронной сети при подаче каждого входного вектора определяется нейрон-победитель, для которого (29.32) максимально. Для этого нейрона синаптические связи усиливаются по формуле

$$w_{ij}(t+1) = w_{ij}(t) + \beta(x_i - w_{ij}(t)), \quad (29.33)$$

где β – скорость обучения.

Смысл этой формулы в том, что вектор весовых коэффициентов нейрона – победителя “поворачивается” в сторону входного вектора, тем самым активность нейрона усиливается. Удобно работать с нормированными

входными и весовыми векторами, когда их модуль равен 1. Нормировка уравнивает шансы в конкуренции нейронов с разным модулем вектора весовых коэффициентов. Выражение (29.32) для нормированных векторов будет выглядеть как

$$y_j = \sum_{i=1}^n w_{ij} x_i = \cos \alpha, \quad (29.34)$$

а выражение (7.33):

$$w_{ij}(t+1) = \frac{w_{ij}(t) + \beta(x_i - w_{ij}(t))}{|w^j(t) + \beta(x - w^j(t))|}. \quad (29.35)$$

Случайное начальное распределение весовых коэффициентов может привести к тому, что некоторые нейроны никогда не станут победителями, так как их весовые векторы окажутся удаленными от всех входных векторов. Существует ряд модификаций алгоритма обучения, позволяющих устранить этот недостаток. Хорошие результаты на практике показало частотно–зависимое конкурентное обучение. Согласно нему, нейрон-победитель определяется по минимуму произведения евклидова расстояния между входным и весовым вектором и количеством побед данного нейрона f_j :

$$d_v = \min_j (|x - w^j| f_j). \quad (29.36)$$

Шансы нейрона на победу уменьшаются с количеством побед, что дает преимущество другим нейронам.

Конкурентное обучение продолжается до тех пор, пока максимум евклидова расстояния между любым входным вектором и

соответствующим ему вектором весов нейрона-победителя не достигнет заданного малого значения.

Конкурентная сеть позволяет разбить входную выборку нормированных векторов на m (количество выходных нейронов сети) кластеров, расположенных на поверхности гиперсферы в пространстве признаков единичного радиуса. Входные векторы, приводящие к победе одного и того же нейрона, относят к одному кластеру.

Кохонен предложил внести в правило конкурентного обучения (29.33) информацию о расположении нейронов в выходном слое. Для этого нейроны упорядочиваются в одномерные или двухмерные решетки. Вводится функция, корректирующая изменение весов в зависимости от расстояния до нейрона-победителя $h(t,k,j)$ – сила влияния между нейроном-победителем k и нейроном j в момент времени t . Для $j=k$ эта функция всегда равна 1 и уменьшается с ростом расстояния между k и j в решетке. С течением времени радиус влияния обычно сужается. С использованием этой функции веса меняются для всех нейронов сети, а не только для нейрона-победителя:

$$w_{ij}(t+1) = w_{ij}(t) + \beta h(t,k,j)(x_i - w_{ij}(t)). \quad (29.37)$$

В качестве функции $h(t,k,j)$ может использоваться гауссовый колокол с параметром σ , зависящим от времени или функция вида “мексиканская шляпа”.

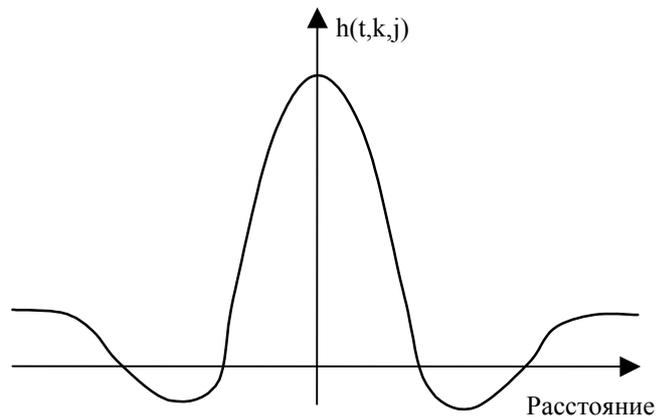


Рис. 29.11. Пример функции расстояния в сетях Кохонена

В результате модификации конкурентного обучения сеть Кохонена не только кластеризирует входные примеры, но и упорядочивает их в виде одномерной или двухмерной решетки. Это позволяет получить дополнительную информацию о близости кластеров. Если два кластера проецируются на соседние нейроны решетки, это говорит об их близости в исходном пространстве признаков. Обратное неверно. Из-за уменьшения размерности пространства отношение близости сохраняется только для ограниченного числа кластеров.

ЛИТЕРАТУРА

1. У. Прэтт. Цифровая обработка изображений. В 2-х томах. М.: Мир, 1982.
2. Гонсалес Р., Вудс Р. Цифровая обработка изображений. - М.: Техносфера, 2005.
3. Сэломон Д. Сжатие данных, изображений и звука: Учеб. пособие для вузов. - М.: Техносфера, 2004.
4. Уэлстид С. Фракталы и вейвлеты для сжатия изображений в действии: Учеб. пособие. - М.: Триумф. 2003.
5. Бутаков Е.А. и др. Обработка изображений на ЭВМ. М.: Радио и связь, 1987.
6. СБИС для распознавания образов и обработки изображений. М.: Мир, 1988.
7. К. Фу. Последовательные методы в распознавании образов и обучении машин. М.: Наука, 1971.
8. Р. Дуда, П. Харт. Распознавание образов и анализ сцен. М.: Мир, 1976.
9. Павлидис Т. Алгоритмы машинной графики и обработки изображений. М.: Радио и связь, 1986.
10. Мартинес Ф. Синтез изображений. Принципы, аппаратное и программное обеспечение. М.: Радио и связь, 1990.
11. Ярославский Л.П. Введение в цифровую обработку изображений. М.: Радио и связь, 1979.
12. Верхаген К., Дейн Р., Грун Ф. Распознавание образов: состояние и перспектива. М.: Радио и связь, 1982.
13. Горелик А.Л., Гуревич И.Б., Скрипник В.А. Современное состояние проблемы распознавания. М.: Радио и связь, 1985.
14. Дж. Бендат, А. Пирсол. Прикладной анализ случайных данных. М.: Мир, 1989.
15. Гольдберг Л.М., Матюшкин Б.Д., Поляк М.Н. Цифровая обработка сигналов. М.: Радио и связь, 1990.
16. Горелик А.Л., Скрипник В.А. Методы распознавания. М.: Высшая школа, 1989.
17. Q. Tian, Y.Faiman and S.H. Lee. Comparison of statistical pattern recognition algorithms for hybrid processing, 1 Linear mapping algorithms// Journal of the optical society of America, V.5 №10, 1988, p.1655-1669.
18. O. Tian, V. Faiman and S.Y. Lee. Comparison of statistical pattern recognition algorithms for hybrid processing. 2. Statistical pattern

- recognition algorithms// Journal of the optical society of America, v.5, №10, 1988, pp.1670-1680.
19. Кун С. Матричные методы обработки в СБИС. М. «Мир», 1990.
 20. Соколов Е.Н., Вайткявичюс Г.Г. Нейроинтеллект: от нейрона к нейрокомпьютеру. М.: Наука, 1989.
 21. Ярославский Л.П. Цифровая голография. М.: Наука, 1992.
 22. Василенко Г.И., Тараторин А.М. Восстановление изображений. М.: Радио и связь, 1986.
 23. Фу К. Структурные методы распознавания образов. М.: Мир, 1974.
 24. Фоли Дж., Дэм А. Основы интерактивной машинной графики в 2-х томах. М.: Мир, 1985.
 25. Садыхов Р.Х., Чеголин П.М., Шмерко В.П. Методы и средства обработки сигналов в дискретных базисах. Минск, Наука и техника, 1987.
 26. В.А. Головкин. Нейронные сети: обучение, организация и применение. Москва, 2001.