

ГЛОССАРИЙ ПО ЦИФРОВОЙ ОБРАБОТКЕ СИГНАЛОВ

ПРЕДВАРИТЕЛЬНАЯ ВЕРСИЯ

Автор: Грибунин В.Г.

Электронная версия книги подготовлена фирмой АВТЭКС Санкт-Петербург,
<http://www.autex.spb.ru>, E-mail: info@autex.spb.ru

ВВЕДЕНИЕ

Бурное развитие цифровой обработки сигналов в последние годы привело к появлению большого количества новых терминов. По понятным причинам, большинство из них – англоязычные. На русский язык они, как правило, не переводятся, а в текстах используется калька с английского. Написание слов бывает самым разнообразным, что приводит к путанице. Еще большую путаницу вносит нередко встречающийся некорректный перевод терминов. Поэтому, мною была предпринята попытка некоторой унификации терминологии, о которой, как известно, не спорят – о ней договариваются.

Конечная цель этого словаря состоит в том, чтобы помочь вам читать англоязычные статьи, диссертации по цифровой обработке сигналов. С другой стороны его можно рассматривать как краткий справочник по темам, представленным ниже. Поэтому, в статьях словаря приводятся и рисунки, и формулы. При этом у меня нет особых претензий на соответствие приведенных терминов формулировкам каких-либо нормативных документов. Кроме того, все формулировки относятся только к цифровым сигналам.

Жду Ваших замечаний, предложений по содержанию глоссария.

Вадим Грибунин.

СОДЕРЖАНИЕ

1. Преобразования сигналов
2. Фильтры
3. Банки фильтров
4. Вейвлеты
5. Стеганография
6. Сжатие изображение

1. ПРЕОБРАЗОВАНИЯ СИГНАЛОВ

1.1. Сигнал (Signal) – то, что математики обычно называют функцией. Под сигналом можно понимать какое-то упорядоченное множество чисел, несущих информацию о некотором процессе. Обычно описывается двумя (одномерный сигнал), тремя (двумерный сигнал) или более (многомерный сигнал) параметрами. Представляется в виде конечной или бесконечной совокупности точек. Одним из параметров для всех типов сигналов является значение уровня сигнала (его энергии) во всех точках. В качестве других параметров обычно выступают время (одномерный сигнал), пространственные координаты (двумерный и многомерные сигналы). Значения всех параметров могут быть непрерывными или дискретными. Таким образом, для каждой размерности можно различать 2^L сигналов, где L – размерность сигнала. Например, в одномерном случае существуют сигналы, непрерывные по времени и уровню, дискретные по времени и уровню, дискретные по времени и непрерывные по уровню, непрерывные по времени и дискретные по уровню. Преобразование непрерывного по времени и уровню сигнала в дискретный по времени и уровню сигнал называется аналого-цифровым преобразованием (АЦП), обратное преобразование – ЦАП. Так как эти вопросы выходят за рамки глоссария, в дальнейшем все сигналы считаются дискретными по времени и уровню, то есть цифровыми.

1.2. Преобразования сигналов (Signal Transform) – понятие, вообще говоря, очень широкое и включает в себя любую операцию, производимую над сигналом. В рамках словаря под преобразованием мы будем понимать лишь обобщенное преобразование Фурье (см. Преобразование Фурье). Преобразование переводит сигнал из одной области представления в другую. Прямое преобразование $F \rightarrow Tf$ переводит сигнал из временной (пространственной) области в область спектра, которая еще называется трансформантой. Обратное преобразование $f \rightarrow T^{-1}F$ переводит сигнал из области трансформанты во временную (пространственную) область.

Преобразования сигналов используются для разных целей: его сжатия, анализа и т.д. Анализ заключается в выполнении каких-нибудь действий над сигналом и формулировке выводов на основе полученных данных.

1.3. Спектр (Spectr) – представление сигнала в виде конечной или бесконечной суммы некоторых элементарных сигналов, умноженных на некоторые числа. В качестве элементарных сигналов обычно выступают ортогональные функции, такие как Фурье, Уолша, Хаара, Адамара, Хартли, вейвлеты и т.д. Поэтому говорят о спектре Фурье, Уолша и т.д. Числа, на которые умножаются элементарные сигналы называются спектральными коэффициентами (коэффициентами трансформанты). Часто просто говорят о коэффициентах Фурье, Уолша и т.д.

1.4. Базис (Basis) – совокупность векторов пространства $\mathcal{V}$, обладающих следующим свойством: любой вектор пространства может быть представлен единственным образом как их линейная комбинация $\sum_k b_k \phi_k$. Число базисных векторов равно размерности пространства. Аналогично и последовательность функций $\phi_k(t)$ является базисом функционального пространства, если любая функция из этого пространства может быть представлена единственным образом как $\sum_k b_k \phi_k(t)$. Числа b_k называются коэффициентами разложения.

1.5. Биортогональный базис (дуальный базис). Дуальный базис w_l биортогонален исходному базису: $\langle \phi_k, w_l \rangle = \delta_{kl}$. Тогда коэффициенты разложения $b_k = \langle \phi, w_k \rangle$ (находятся на этапе анализа) и $\sum_k \langle \phi, w_k \rangle \phi_k$ (вычисляется при синтезе).

1.6. Ортогональный базис – частный случай биортогонального базиса, когда базисные векторы биортогональны самим себе. В этом случае дуальный базис идентичен исходному базису и $\langle \phi, \phi_k \rangle = \delta_k$.

1.7. Линейное преобразование (Linear transform) – преобразование, удовлетворяющее условию: преобразование суммы сигналов равно сумме преобразований каждого сигнала. Большинство рассматриваемых в ЦОС преобразований – линейно.

1.8. Унитарное преобразование (Unitary transform) – преобразование, сохраняющее норму сигнала. Вещественное унитарное преобразование называется ортогональным. Его базисные функции ортогональны между собой. Важным свойством ортогонального преобразования является простая формула вычисления обратного преобразования. Все упомянутые выше преобразования являются унитарными.

1.9. Преобразование Фурье – (Fourier transform).

- непрерывное: $\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt$ и $f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega) e^{j\omega t} d\omega$;
- дискретное: $\hat{x}[k] = \sum_{l=0}^{N-1} W^{-kl} x[l]$ и $x[l] = \sum_{k=0}^{N-1} W^{kl} \hat{x}[k]$, где $W = e^{2\pi j/N}$ - поворачивающий множитель;
- ряды Фурье: $c_n = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-jnt} dt$ и $f(t) = \sum_{n=-\infty}^{\infty} c_n e^{jnt}$.

1.10. Оконное (Short-Time) преобразование Фурье (STFT) - (ОПФ) - определяется как преобразование Фурье сигнала, умноженного на некоторую взвешивающую функцию w (окно). Так как окно локализовано по времени, ОПФ применяется для получения частотно-временного представления сигнала. Выражение для ОПФ сигнала $f(t)$ может быть записано в виде

$$F(\omega, t) = \int_{-\infty}^{\infty} w(t-\tau) f(\tau) e^{-j\omega \tau} d\tau$$
, где величина $g(\omega, t) = w(t-\tau) e^{-j\omega \tau}$ называется ядром преобразования. В настоящее время предложено большое число различных окон: прямоугольное, треугольное, Ханна, Хэмминга и др. Простейшим из них является прямоугольное, соответствующее случаю умножения сигнала на константу. Так как преобразованием Фурье прямоугольника служит осциллирующая sinc-функция, то применять такое окно не рекомендуется.

1.11 Преобразование Габора – оконное преобразование Фурье с $g(t) = g(0) e^{-\alpha t^2}$.

1.12. Преобразование Карунена-Лозва (ПКЛ) – базисные функции есть собственные векторы ковариационной матрицы входного сигнала. Является оптимальным по критерию достижения декорреляции входного сигнала. Энергия входного сигнала максимально перераспределяется в коэффициентах ПКЛ. Гарантируется, что процентное содержание энергии входного сигнала в данном количестве наибольших коэффициентов будет не меньше, чем в том же числе коэффициентов любого другого преобразования. Вычислительно трудоемко - $O(N^4)$, кроме того, для вычисления обратного преобразования декодеру надо передавать базисные функции, найденные кодером. Поэтому, на практике не применяется. Все остальные преобразования часто сравниваются с ПКЛ.

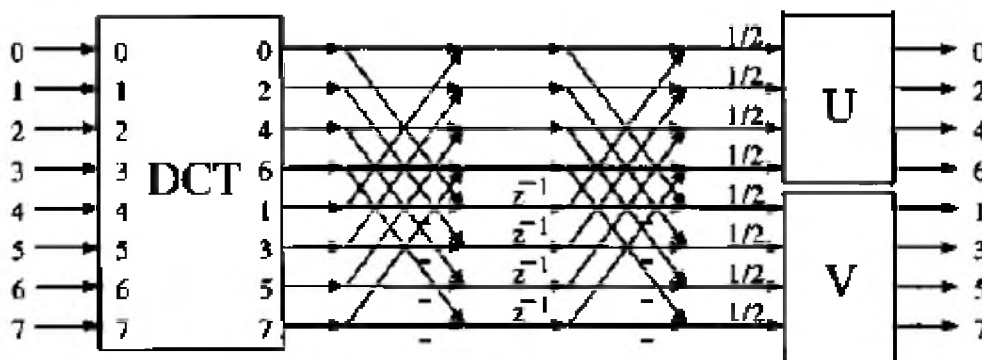
1.13. Дискретное преобразование Хаара (Haar transform) – разложение сигналов по ортонормальному базису вейвлетов Хаара.

1.14. Дискретное косинусное преобразование - ДКП (Discrete cosine transform - DCT). Существует четыре основных типа ДКП:

$$\begin{aligned}
 \text{I} \quad & \sqrt{\frac{2}{M}} c_i c_k \cos \frac{\pi}{M} i k, \quad i, k = 0, \dots, M \\
 \text{II} \quad & \sqrt{\frac{2}{M}} c_i \cos \frac{\pi}{M} i k + \frac{1}{2} \sqrt{\frac{2}{M}}, \quad i, k = 0, \dots, M-1 \\
 \text{III} \quad & \sqrt{\frac{2}{M}} c_k \cos \frac{\pi}{M} i k + \frac{1}{2} \sqrt{\frac{2}{M}}, \quad i, k = 0, \dots, M-1 \\
 \text{IV} \quad & \sqrt{\frac{2}{M}} \cos \frac{\pi}{M} i + \frac{1}{2} \sqrt{\frac{2}{M}} k + \frac{1}{2} \sqrt{\frac{2}{M}}, \quad i, k = 0, \dots, M-1,
 \end{aligned}$$

где $c_i = 1$, кроме $c_0 = 1/\sqrt{2}$. Аналогично определяется и дискретное синусное преобразование (косинусы заменяются на синусы). Оба преобразования являются ортогональными.

1.15. Перекрывающееся ортогональное преобразование (ПОП) (Lapped Orthogonal Transform - LOT). Схема вычисления ПОП изображена на рисунке. Здесь DCT – блок дискретного косинусного преобразования, ортогональные матрицы U и V размером $M/2$ параметризуются множеством поворачивающих углов. При ПОП почти отсутствуют эффекты блочности, присущие ДКП.



1.16. Обобщенное перекрывающееся ортогональное преобразование (GenLOT) – результат ПОП умножается на дополнительные матрицы. И ДКП, и ПОП являются частными случаями этого преобразования.

1.17. Преобразование Хартли (Hartley) – одно альтернатив ДПФ, использует вещественные функции. Элементы матрицы ПХ имеют вид

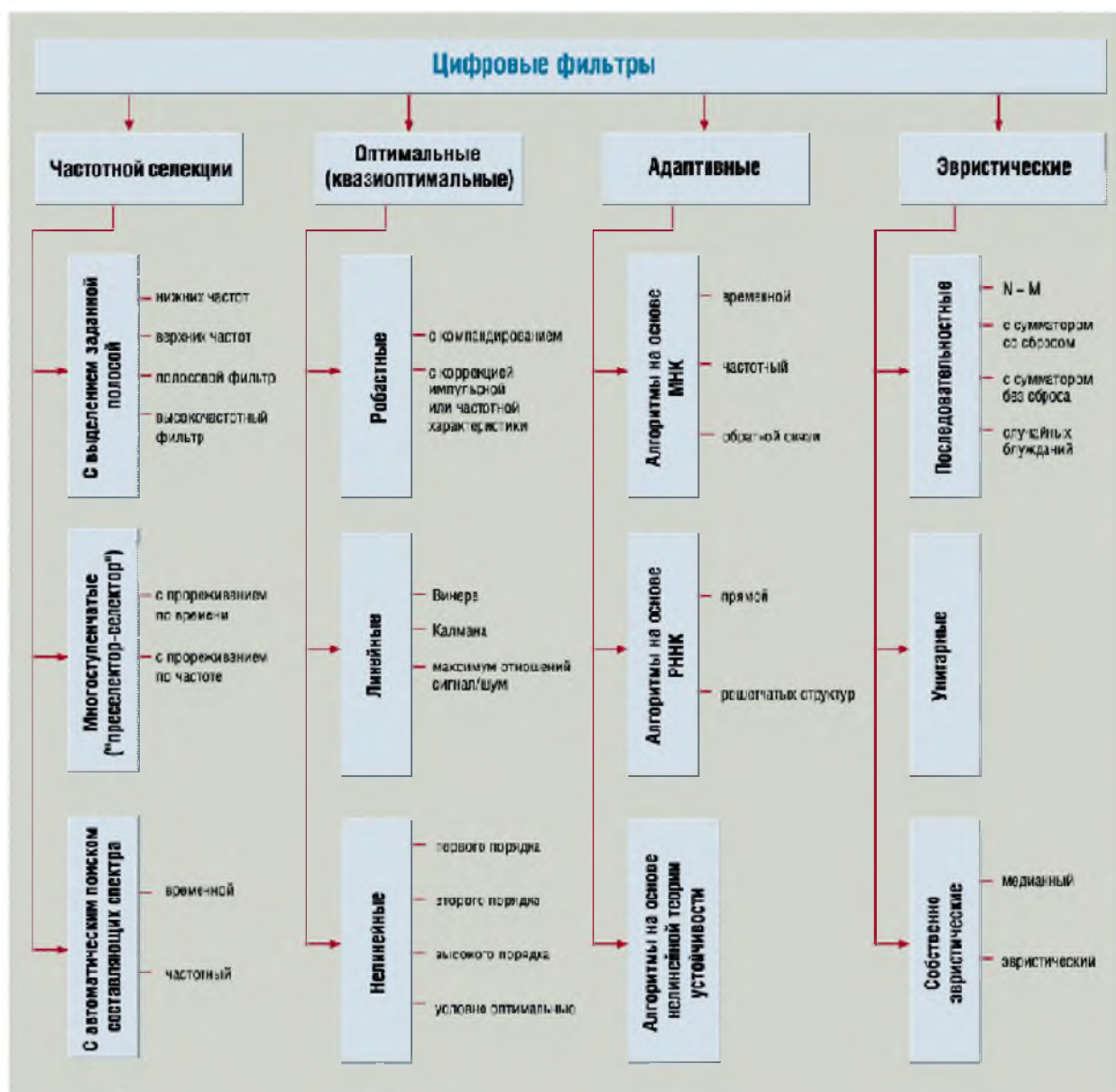
$H_{kn} = \cos \frac{2\pi kn}{L} - \sin \frac{2\pi kn}{L}$. Одно время рассматривалось как панацея ЦОС, но потом оказалось в тени вейвлет-преобразования.

1.18. Свертка (Convolution). Сверткой двух векторов $h * x$ называется вектор, n -й элемент которого равен $\sum_k h[k]x[n-k] = \sum_n h[n-k]x[k]$.

1.19. Модуляционная матрица $H_m(z)$ такая, что ее элементы $H_k(w)$.
Матрица синтеза $F_m(z) = H_m^T(z)$.

2. ФИЛЬТРЫ

Фильтр (Filter) – линейная стационарная система, то есть свойства фильтра не зависят от времени. Независимость свойств фильтра от времени означает, что задержка входа приводит к такой же задержке выхода. К основным характеристикам фильтров относятся импульсная характеристика $h[n]$, передаточная характеристика $H(z) \leftarrow h[n]z^{-n}$, частотная характеристика $H(e^{j\omega}) \leftarrow h[n]e^{j\omega n}$, порядок фильтра. Примерная классификация фильтров приведена на рисунке. Ниже рассмотрены только некоторые типы фильтров.



2.1. Каузальный (Causal) фильтр - $h[k] = 0$ при $k < 0$, то есть реакция фильтра не может предшествовать приложенному воздействию. Система каузальна, если выходной сигнал $y[n]$ зависит от входного сигнала только в моменты времени n и раньше.

2.2. Стабильный фильтр (физически реализуемый) – сумма коэффициентов импульсной характеристики бесконечна. $\sum |h[n]| < \infty$. У каузального и стабильного фильтра все нули передаточной функции находятся внутри единичного круга.

2.3. М-полосный фильтр (фильтр Найквиста) - один из наиболее часто используемых типов фильтров. Находит применение при обработке сигналов, в связи, в многоскоростных системах, при построении сигма-дельта АЦП, фильтров децимации и интерполяции. Некаузальный М-полосный фильтр имеет импульсную характеристику вида $h[n] = M \delta[n]$. Можно показать, что $|H(e^{j\omega})| = 1$. Симметричный М-полосный НЧ фильтр должен иметь нечетную длину и $\omega_p, \omega_s = 2\pi/M$, $\omega_p, \omega_s = \pi/M$. Здесь ω_p, ω_s - граничные частоты полос пропускания и задержки, ω_p, ω_s - максимальные ошибки (отклонения от требуемых величин затухания) в этих полосах.

2.4. Все пропускающий (Allpass) фильтр – фильтр, частотная характеристика которого вещественна, а ее модуль равен 1: $|A(e^{j\omega})| = 1$.

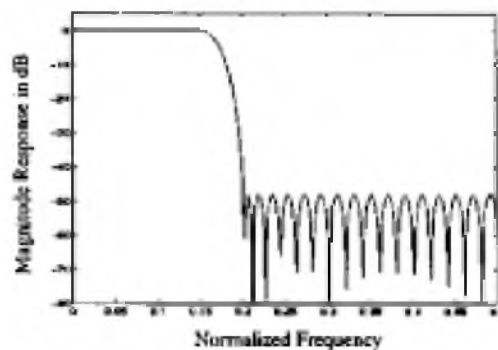
2.5. Фильтр Баттерворта (максимально плоский фильтр) – в полосе пропускания практически нет колебаний. Это является преимуществом для многих приложений, где требуется постоянство коэффициента ослабления фильтра для всех частот полосы пропускания. Недостатком такого фильтра является невысокая крутизна полосы перехода.

2.6. Дополнительные по мощности (Power Complementary) фильтры - $\sum_k |H_k(e^{j\omega})|^2 = \text{const}$. Примерами могут служить фильтры параунитарного банка фильтров.

2.7. Квадратурно-зеркальный фильтр – КЗФ (Quadrature Mirror Filter - QMF) – пара фильтров, удовлетворяющих условию $H_1(z) = H_0^*(z^*)$, то есть ВЧ и НЧ характеристики фильтров симметричны относительно средней частоты $\pi/2$. Эта частота называется квадратурной, отсюда и название фильтров. КЗФ могут быть как с КИХ, так и с БИХ. Банк фильтров КЗФ может быть М-полосным (см. также п.3.10).

2.8. Полуполосный (HalfBand) фильтр – центрированный полуполосный фильтр имеет коэффициент $h[0] = 0,5$ и $h[2k] = 0$. Нецентрированный полуполосный фильтр имеет $h[2k] = 1/2$, $h[2k+1] = 0,5$. Частотная характеристика фильтра нечетна относительно точки $\omega = \pi/2$ и $H(e^{j\omega}) = H(e^{j(\pi-\omega)})e^{j\omega}$.

2.9. Равноволновой (Equiripple) фильтр (чебышевский фильтр) – фильтр, минимизирующий максимальную ошибку в полосах задерживания и пропускания (то есть оптимальный в минимаксном смысле). Известно, что необходимым и достаточным условием этого является равенство всплесков частотной характеристики (см. рисунок). Равноволновой фильтр строится с использованием алгоритма Park-McCellan.



2.10. Линейнофазовый фильтр - фаза ЧХ $H(e^{j\omega})$ линейна: $\arg H(e^{j\omega}) = -\omega N/2$ или $\arg H(e^{j\omega}) = -\omega N/2 + \pi/2$. Групповое время задерживания постоянно. Значения коэффициентов фильтров симметричны или антисимметричны относительно отсчета $n = N/2$, то есть $h[n] = h[N-n]$ или $h[n] = -h[N-n]$.

3. БАНКИ ФИЛЬТРОВ

Банк фильтров (Filter Bank) – совокупность фильтров и следующих за ними дециматоров или следующих перед ними интерполяторов. Под дециматором понимается устройство, осуществляющее децимацию (прореживание) сигнала. Интерполятор выполняет интерполяцию сигнала.

Банки фильтров бывают равномерные и неравномерные, ортогональные, биортогональные, двухканальные и многоканальные и т.д. Каждый фильтр банка фильтров образует канал. Поэтому говорят об M -канальном банке фильтров. Сигнал в канале называется субполосой. Отсюда название субполосная фильтрация (субполосное кодирование).

В случае если число каналов равно коэффициенту децимации (интерполяции) говорят о максимально (критически) децимированном банке фильтров. Быстрый алгоритм вычисления вейвлет-преобразования строится именно на основе таких банков фильтров.

Равномерный банк фильтров – децимация в каждом канале одинаковая. В противном случае – неравномерный банк фильтров. Частный случай неравномерного банка фильтров – древовидный банк фильтров.

Обозначения:

НЧ фильтр H : $X_1 \dots X_0$

НЧ канал C : $c_k \sqrt{2} h_k$

Матрица $M \sqrt{2} L \times 2H$: $M_{jk} = 2h_j \otimes k$

3.1. Децимация (прореживание) (Decimation) – операция, заключающаяся в выбрасывании отсчетов, чей порядковый номер кратен определенному числу. Например, при децимации в два раза выбрасывается каждый второй отсчет, при прореживании в три раза – каждый третий и т.д. Децимация в M раз обозначается обычно как $y[n] \rightarrow y[n/M]$. В частотной

области это запишется как $Y(e^{i\omega}) = \frac{1}{M} \sum_{k=0}^{M-1} X(e^{i(\omega - 2\pi k/M)})$, то есть спектр

выходного сигнала операции децимации содержит M копий «расширенного» в M раз спектра входного сигнала, как это показано на рисунке. Как видно из

рисунка, если сигнал не ограничен полосой частот $[-\frac{\pi}{M}, \frac{\pi}{M}]$, то

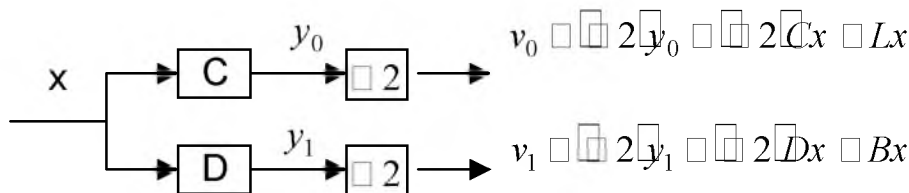
происходит наложение спектров копий, то есть элайзинг. Поэтому в банке фильтров перед децимацией выполняется НЧ фильтрация.

Совокупность фильтра и дециматора называется фильтром-дециматором.

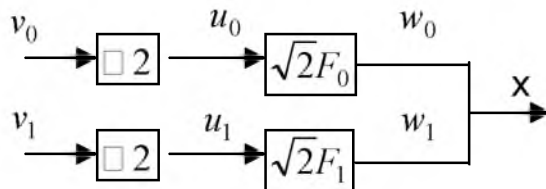
3.2. Интерполяция (Interpolation) – операция, заключающаяся во встраивании между отсчетами, чей порядковый номер кратен определенному числу, некоторой константы (обычно нуля). Интерполяция в M раз обычно обозначается как $[M]$: $u[n/M]$, n делится на M , 0 , в других случаях.

В частотной области это запишется как $U(e^{j\omega}) = X(e^{jM\omega})$, то есть спектр выходного сигнала операции интерполяции содержит M копий «сжатого» в M раз спектра входного сигнала. Эти копии повторяются через $\omega_k = \frac{2\pi k}{M}$. Для их устранения после интерполятора ставится фильтр. Совокупность интерполятора и фильтра называется фильтром-интерполятором.

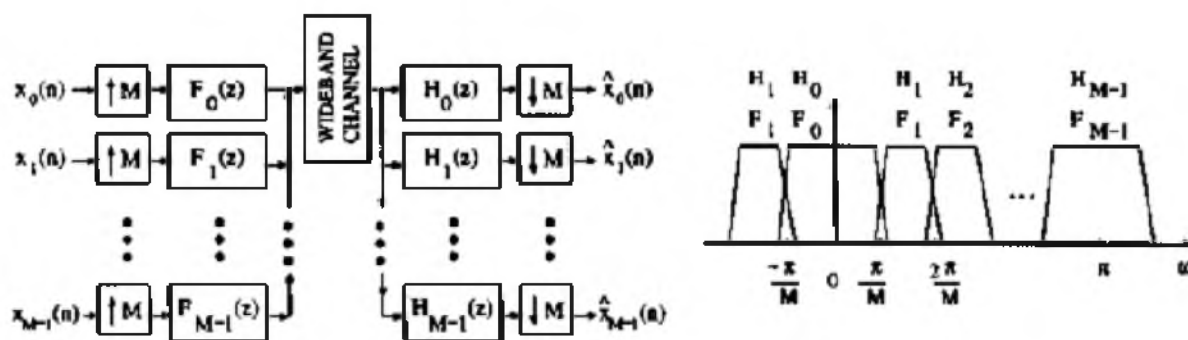
3.3. Банк фильтров анализа (Analysis Filter Bank):



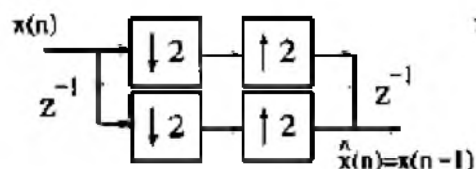
3.4. Банк фильтров синтеза (Synthesis Filter Bank) :



3.5. Система анализа-синтеза (А-С) – совокупность банка фильтров анализа и банка фильтров синтеза. Сигнал вначале декомпозируется на субполосы, в каждой из них выполняется некоторая его обработка и, затем, выполняется синтез сигнала (реконструкция). Надо отметить, что иногда нужен обратный порядок операций: сначала синтез, потом анализ. Такая последовательность действий встречается при использовании банков фильтров в связи. При этом НЧ сигналы от разных источников интерполируются, фильтруются, объединяются и передаются по каналу связи. На приеме групповой сигнал подается на схему анализа для выделения сигналов отдельных абонентов. Такая схема, показанная на рисунке, называется трансмультиплексором. На рисунке показаны также спектры сигналов в отдельных каналах. Видно, что трансмультиплексор есть разновидность схемы с частотным разделением каналов (ЧРК). Выполняя оптимизацию фильтров можно достичь серьезных выигрышей по полосе пропускания по сравнению с классическими схемами с ЧРК.



3.6. Lazy («ленивый») банк фильтров – система А-С без фильтров, как это показано на рисунке. Применяется при доказательстве теорем в банках фильтров. В частности, мы наглядно видим возможность достижения полного восстановления - выходной сигнал Lazy банка фильтров равен входному с точностью до задержки.



3.7. Два замечательных равенства (Noble Identity) позволяют уменьшить число вычислений в банке фильтров. Также широко применяются для вывода различных соотношений.



3.8. Связь между фильтрами в банке фильтров. Все фильтры в банке фильтров зачастую получаются один из другого при помощи двух операций: переворота вектора коэффициентов («flip») и чередования знака (alternating sign).

- Ordinary flip: $F_1(z) = z^{MN} H_0^*(z)$, то есть $h_1[N], \dots, h_1[0]$. Эта операция приводит к транспонированию H , задержка введена для обеспечения каузальности.

-Alternating flip: $H_1(z) = z^{MN} H_0^*(z)$, то есть $h_1[N], h_1[N-1], \dots, h_1[0]$. Таким образом получаются ортогональные фильтры (Смита-Барнуэла).

-Alternating signs: $F_1(z) = H_0(z)$ и $F_0(z) = -H_1(z)$, то есть $h_0[0], h_0[1], h_0[2], \dots, h_0[N]$. Таким образом устраняется элайзинг.

3.9. Полное восстановление (Perfect Reconstruction - PR). Сигнал, прошедший через схему анализа-синтеза идентичен входному с точностью до задержки. Для этого фильтры синтеза должны подавлять элайзинг и устранять амплитудные и фазовые искажения:

$$\begin{aligned} F_0(z)H_0(z) + F_1(z)H_1(z) &= z^{-l}, \quad l \bmod 2 = 0 \\ F_0(z)H_0(z) - F_1(z)H_1(z) &= 0. \end{aligned}$$

Модуляционная матрица должна удовлетворять соотношению

$$F_m(z)H_m(z) \neq 0 \quad \forall m$$

3.10. Двухканальный банк фильтров КЗФ (Джонстона): $H_1(z) = H_0(-z)$. Пусть $E_k(z)$ полифазные компоненты $H_0(z)$. Тогда

$$\begin{bmatrix} H_0(z) \\ H_1(z) \end{bmatrix} = \begin{bmatrix} E_0(z^2) & E_1(z^2) \\ E_0(z^2) & -E_1(z^2) \end{bmatrix} \begin{bmatrix} 1 \\ z^{-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} E_0(z^2) & 0 \\ 0 & E_1(z^2) \end{bmatrix} \begin{bmatrix} 1 \\ z^{-1} \end{bmatrix}$$

Фильтры синтеза удовлетворяют равенствам:

$$\begin{bmatrix} F_0(z) & F_1(z) \end{bmatrix} = \begin{bmatrix} H_1(-z) & -H_0(-z) \end{bmatrix} = \begin{bmatrix} z^{-1} & 1 \end{bmatrix} \begin{bmatrix} E_1(z^2) & 0 \\ 0 & E_0(z^2) \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Банк КЗФ не является банком фильтров с полным восстановлением. Суммарное искажение зависит от значений полифазных компонент.

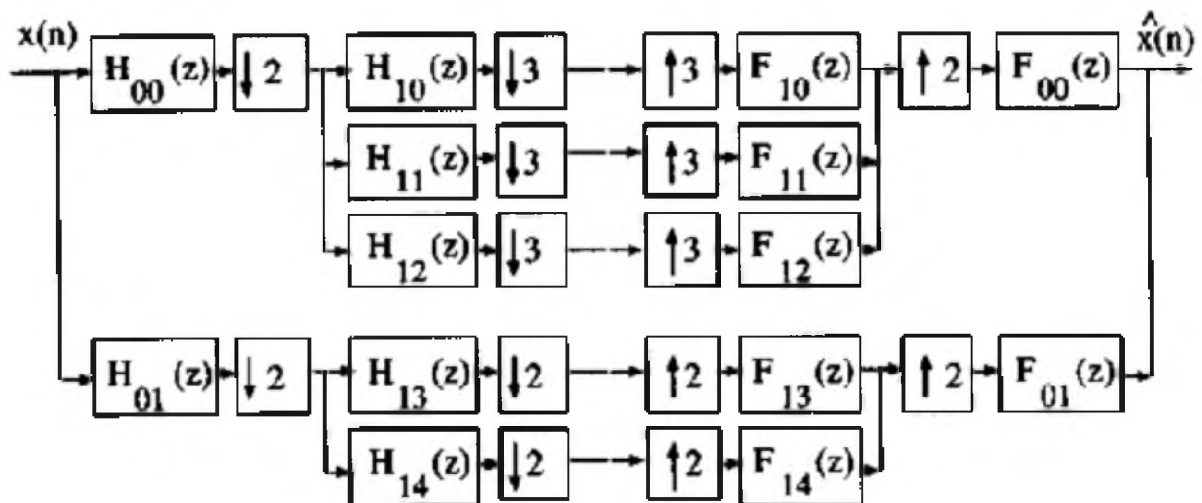
3.11. Ортогональный (параунитарный) банк фильтров – фильтры синтеза есть транспонированные версии фильтров анализа: $F_p(z) = H_p^T(z)$, или $f_k(n) = h_k(N-n)$

$$\begin{aligned} L^T B^T &= I, \quad L B^T = I, \\ L L^T &= I, \quad B B^T = I. \end{aligned}$$

3.12. Элайзинг (Aliasing) – наложение спектра, то есть два сигнала накладываются один на другой. Пусть X_1 и X_0 . Тогда после децимации в два раза их спектры будут перекрываться.

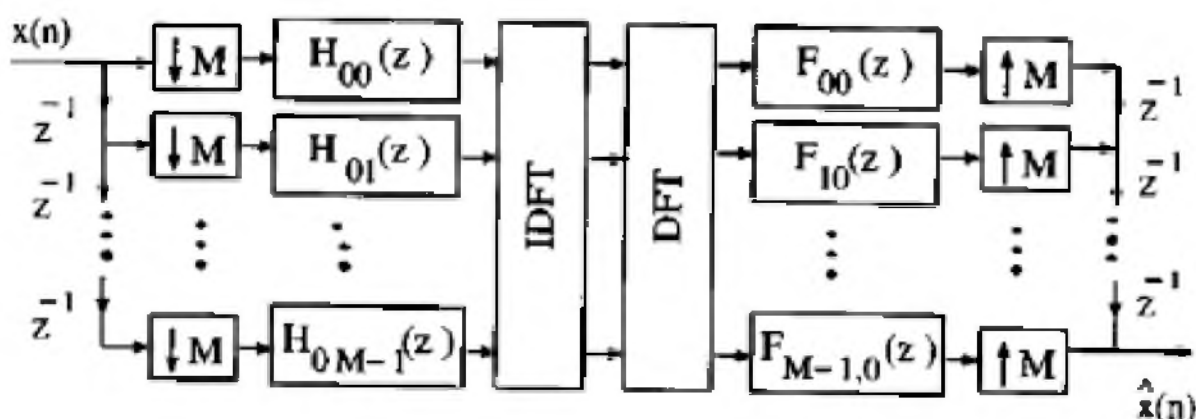
3.13. Устранение элайзинга (Alias Cancellation). Выходной сигнал системы А-С банка фильтров $\hat{x}[n]$ будет свободен от элайзинговой составляющей в случае, если произведение модуляционных матриц $F_m[z]$ и $H_m[z]$ диагонально. Для $M=2$ каналов это означает, что $F_0[z]H_0[z] = F_1[z]H_1[z] = 0$. После устранения элайзинга в сигнале на выходе системы А-С остаются только амплитудные и фазовые искажения.

3.14. Древовидный банк фильтров показан на рисунке. «Дерево» анализа строится так, что его корневой вершиной является исходный сигнал. От корня отходят две ветви, сигналы в которых являются вершинами для трех и двух ветвей. Дерево синтеза строится в обратном порядке.



Используя замечательные равенства, мы можем переместить первые дециматоры в два раза вправо и объединить их со вторыми дециматорами. Общий коэффициент децимации в каналах будет равен (6,6,6,4,4). Таким образом, древовидный банк фильтров позволяет достичь неравномерной скорости в субканалах. Основным недостатком является большая задержка (зависит от высоты дерева).

3.15. Банк фильтров равномерного ДПФ показан на рисунке. Все фильтры анализа $H_k[z]$ и синтеза $F_m[z]$ являются равномерно сдвинутыми версиями $[z] & zW^k$ НЧ фильтров. Полное восстановление возможно лишь при длине фильтров, равной M .



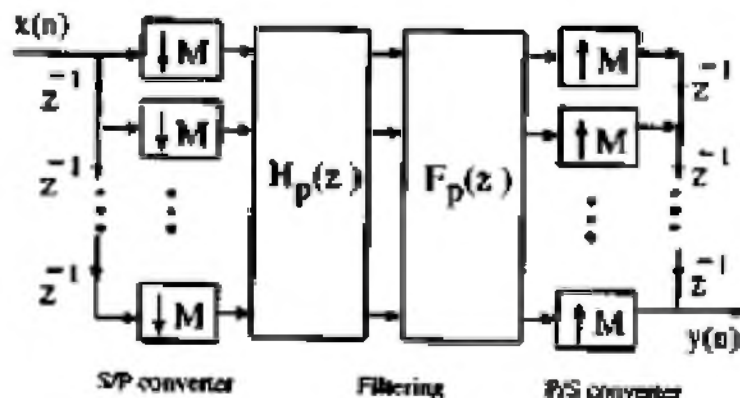
3.16. Полифазное представление (Тип I и Тип II). Последовательность $h[n]$ может быть представлена в виде совокупности M последовательностей («фаз») $e_k[n]$ или $r_k[n]$:

$$\begin{aligned} e_k[n] &= h[Mn + k] & \text{Тип I} & \quad 0 \leq k \leq M-1, \\ r_k[n] &= h[Mn + M-1-k] & \text{Тип II} & \quad 0 \leq n \leq \frac{N}{M}-1 \end{aligned}$$

Такие представления бывают очень полезными для сокращения вычислений в банках фильтров (один из примеров - лифтинговая схема вычисления вейвлет-преобразования), а также для доказательства многих теорем о банках фильтров. В области Z -преобразования можем записать

$$H(z) = \sum_{k=0}^{M-1} z^{Mk} E_k(z^M) = \sum_{k=0}^{M-1} z^{Mk} R_k(z^M) \quad \text{и} \quad R_k(z) = E_{M-1-k}(z^M)$$

Основная схема, позволяющая увеличить скорость вычислений, представлена на рисунке. Как показано, фильтрация осуществляется на скорости в M раз меньшей, чем скорость выходного сигнала.



4. ВЕЙВЛЕТЫ И ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЕ.

Вейвлеты - это математические функции, обладающие некоторыми свойствами. В научном сообществе до сих пор не решен вопрос, какие функции относить к вейвлетам. В узком смысле это семейство функций, получающихся путем масштабирования и сдвигов одной, материнской функции. Именно за счет изменения масштабов вейвлеты способны выявить особенности сигнала на разных шкалах, а за счет сдвигов они способны проанализировать сигнал во всех точках. В широком смысле вейвлеты - это функции, обладающие хорошей частотно-временной локализацией, чье среднее значение равно нулю. При этом они могут вовсе не иметь функции-прототипа (например, вейвлеты второго поколения, названные так В.Свелденсом).

В математических кругах вейвлеты назывались одно время всплесками, но сейчас этот термин встречается редко.

В приложениях чаще всего используют дискретные вейвлеты, так как непрерывные вейвлеты не образуют ортонормированного базиса и бесконечно протяжены (имеют экспоненциальный спад). С другой стороны, применение непрерывного вейвлет-анализа позволяет лучше визуализировать результаты анализа (получить более «красивые картинки») и, возможно, выявить какие-то скрытые от других видов анализа свойства сигнала. Хотя я бы лично рекомендовал использовать для анализа сигналов частотно-временные разложения Вигнера и им подобные - там, где скорость анализа не имеет первостепенной важности.

Рассмотрим основные термины, используемые в этой области. При этом надо учесть, что если англоязычные термины являются устоявшимися, то этого нельзя сказать о русских терминах. Даже перевод и издание «библии вейвлетов» - книги И.Добеши «Десять лекций по вейвлетам» - не исправил до конца положения. Так, по мнению одного из авторитетов в этой области - Л.Левковича-Маслюка - перевод терминологии (в редакции проф. А.Петухова) выполнен неудачно. Что ж, сколько людей - столько мнений.

4.1. Непрерывное (Continuos) вейвлет-преобразование (CWT). Под CWT функции $f(t) \in L^2(\mathbb{R})$ понимается ее скалярное произведение с базисными

функциями $\left(\frac{1}{\sqrt{a}} \phi\left(\frac{t-b}{a}\right) \right)_{a>0, b \in \mathbb{R}}$, $a \in \mathbb{R}^+, b \in \mathbb{R}$, то есть разложение по всем

возможным сдвигам и сжатиям некоторой функции (порождающего вейвлета).

Если для порождающего вейвлета $\phi(t)$ выполняется равенство

$C_\phi = \int_{-\infty}^{\infty} |\phi(t)|^2 dt < \infty$, то возможно обратное преобразование

$$f(t) = \frac{1}{C_\phi} \int_0^\infty \int_{-\infty}^\infty \text{CWT}_f(a, b) \frac{1}{\sqrt{a}} \phi\left(\frac{t-b}{a}\right) \frac{db da}{a^2}.$$

4.2. Вейвлет-ряды (Ортогональное дискретное (диадное) вейвлет-преобразование). В этом случае рассматриваются не все сдвиги и растяжения базисной функции, но только взятые на некоторой дискретной сетке (обычно логарифмической). Заметьте, что если сигнал остается непрерывным, то называть это преобразование дискретным неверно и приводит к путанице. Иногда его называют диадным. Мне кажется (и не только мне), что лучше называть это преобразование рядами вейвлетов непрерывного времени (CTWS) по аналогии с рядами Фурье.

$$\text{CTWS}_f = \sum_{m,n} a_0^{m/2} (a_0^m t - n) f\left(\frac{t-n}{a_0^m}\right),$$

$$a = a_0^m; b = nb_0 a_0^m, m, n \in \mathbb{Z}, a_0 \neq 1.$$

Если же сигнал - дискретный, то аналогичное преобразование правильно называть дискретным вейвлет-преобразованием. Формулами вычисления прямого ДВП является, по сути, масштабирующее уравнение (см. ниже) для вейвлет-функции и для масштабирующей функции. Обратное ДВП всегда существует.

4.3. Кратномасштабный (многомасштабный) анализ (КМА) - Multiresolution Analysis (MRA) - математическая конструкция, схема представления сигналов. Эта конструкция позволяет «с другой стороны» зайти к построению вейвлет-рядов и заключается в представлении пространства (например, $L^2(\mathbb{R})$) в виде бесконечной последовательности вложенных подпространств, являющихся отмасштабированными версиями друг друга и связанных определенными свойствами. Различают ортогональный и биортогональный КМА. Значение КМА заключается также и в том, что при его помощи мы можем дать более точные определения масштабирующей и вейвлет функциям, а также найти связь между ними.

4.4. Избыточное (redundancy) и безизбыточное преобразования. Преобразование считается безизбыточным, если в коэффициентах содержится ровно столько информации, чтобы можно было совершить обратное преобразование. При анализе сигнала этой информации нам может показаться мало, а обратное преобразование и не потребоваться. В этом случае прибегают к избыточному преобразованию. Например, в случае вейвлет-фреймов (см. ниже) число коэффициентов на каждом уровне может равняться числу отсчетов в исходном сигнале, а число уровней быть весьма значительным (например, 64, 128).

4.5. Масштабирующая (скейлинг) функция (Scaling Function) - называется еще «отцовским вейвлетом» - определяется как решение уравнения

$$\sum_{k \in \mathbb{Z}} h_k \phi\left(\frac{t}{2} - k\right) = \phi(t)$$

Это уравнение носит различные названия - refinement equation, dilation equation, two-scale difference equation. Наиболее удачным переводом на русский является, по-моему, масштабирующее уравнение.

Масштабирующая функция вместе с материнским вейвлетом порождают ортогональный кратномасштабный анализ.

4.6. Материнский (базисный) вейвлет - можно построить по известной масштабирующей функции: $\phi(t) = \sum_{k=0}^{2M-1} h_k 2^{-j/2} \phi(2^j t - k)$, где

$h_k = \int_{-\infty}^{\infty} \phi(t) 2^{j/2} \phi(2^j t - k) dt$. Материнский вейвлет порождает семейство вейвлет-функций $\phi_{j,k}(t) = 2^{j/2} \phi(2^j t - k)$, где j - параметр масштаба, k - сдвига.

4.7. Компактный носитель масштабирующей функции или вейвлета. Носитель - часть области определения функции, где она не равна нулю. Свойство компактности функций означает, что соответствующие им фильтры будут фильтрами с конечной импульсной характеристикой, то есть иметь конечное число коэффициентов.

4.8. Ортогональное дополнение и прямая (direct) сумма. Эти термины надо знать для понимания сущности КМА. Пусть V_0 - подпространство в V_1 . Ортогональное дополнение W_0 содержит все векторы из V_1 , ортогональные каждому вектору из V_0 . Тогда каждый вектор в V_1 может быть представлен в виде суммы двух проекций - на подпространства V_0 и W_0 . Прямая сумма вычислима в случае, если объединение базисов V_0 и W_0 дает базис пространства V_1 и не подразумевает ортогональности этих подпространств. Прямая сумма обозначается знаком $\oplus$.

4.9. Биортогональные масштабирующие функции и вейвлеты:

$$\int_{-\infty}^{\infty} \phi(t) \tilde{\phi}(t - k) dt = \delta(k) \quad \text{and} \quad \int_{-\infty}^{\infty} w(t) 2^{j/2} \tilde{w}(2^j t - k) dt = \delta(j) \delta(k)$$

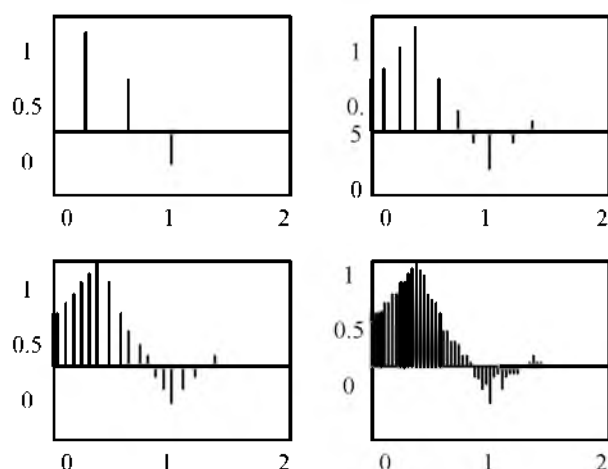
$$\int_{-\infty}^{\infty} \phi(t) \tilde{w}(t - k) dt = 0 \quad \text{and} \quad \int_{-\infty}^{\infty} \tilde{\phi}(t) w(t - k) dt = 0$$

4.10. Ортогональные масштабирующие функции и вейвлеты - частный случай биортогональных, когда $\tilde{\phi} = \phi^*$ и $\tilde{w} = w^*$.

4.11. Каскадный (Cascade) алгоритм - алгоритм последовательного построения приближений к непрерывной функции согласно выражению

$$f(t) = \sum_{k=0}^{2M-1} h_k 2^{-k} f(2t - k)$$

Четыре шага приближения (рис.).



4.12. Сжатие (dilation): $f(at)$ Тогда $\hat{f}\left(\frac{\omega}{a}\right) = \frac{1}{|a|} \hat{f}\left(\frac{\omega}{a}\right)$.

Сдвиг (translation, shift, delay, advance): $f(t-k)$ Тогда $\hat{f}(\omega) e^{-i\omega k} \hat{f}(\omega)$

4.13. Быстрое вейвлет-преобразование (Fast Wavelet Transform) - пирамида Малла (Mallat). Входной сигнал подается на пару фильтров-дециматоров. Выход ВЧ фильтра-дециматора считается коэффициентами, а выход НЧ фильтра поступает на точно такую же схему (то есть в алгоритме выполняется итерация по НЧ каналу). Сложность - не более $2LN$, где L - длина фильтра, N - длина сигнала. В самом деле, на первом уровне сложность выполнения фильтрации LN , на втором - $LN/2$ и т.д.; сумма последовательности $1+1/2+1/4+\dots$ сходится к 2.

Схема прямого и обратного вейвлет-преобразования показана на рисунке.

Прямое преобразование:

$$\begin{array}{ccccccc} x & \rightarrow & Lx & \rightarrow & LLx & \rightarrow & \dots \rightarrow LL \dots Lx \\ & \searrow & & \searrow & & \searrow & \\ & Bx & & BLx & & \dots & BL \dots Lx \end{array}$$

Обратное преобразование:

$$\begin{array}{ccccccc} LL \dots Lx & \rightarrow & \dots & \rightarrow & LLx & \rightarrow & Lx \rightarrow x \\ & \nearrow & & \nearrow & & \nearrow & \\ BL \dots Lx & & \dots & & BLx & & Bx \end{array}$$

4.14. Маска (Mask) - в литературе по аппроксимации так называется вектор, в ЦОС называющийся вектором коэффициентов фильтра.

4.15. Масштабограмма (Scalogram) - квадрат амплитуды коэффициентов непрерывного вейвлет-преобразования. Изображается обычно на плоскости. По оси абсцисс откладывается, например, время, а по оси ординат - масштабы, на которых сигнал анализировался. Масштабограмма аналогична спектрограмме, построенной при спектральном анализе с постоянной относительной полосой (constant-Q). Под спектрограммой понимается квадрат амплитуды коэффициентов оконного преобразования Фурье.

4.16. Гладкость (Smoothness) - параметр, ограничивающий сверху число производных функции. Различные вейвлет-фильтры обладают различной гладкостью (то есть, конечно, гладкостью обладают функции, получаемые в пределе при бесконечном числе итераций пирамиды Малла). Например, считается что сигнал изображения имеет гладкость 1,5-2. Поэтому, для его анализа, сжатия можно применять вейвлет-фильтры соответствующей или большей гладкости. Как правило, чем больше коэффициентов у вейвлет-фильтра, тем больше гладкость соответствующей ему функции.

4.17. Симметрия (антисимметрия) базисных функций - симметрия означает, что если перенести центр функции в начало координат, то функция будет четной. Антисимметрия означает, что функция после переноса будет нечетной.

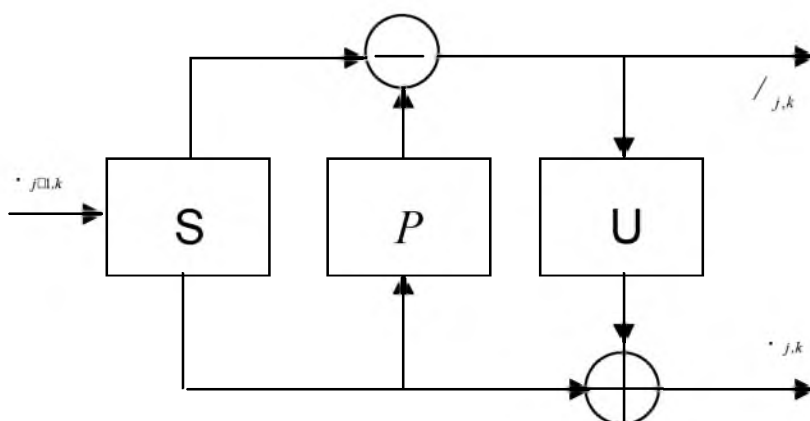
4.18. Принцип неопределенности (Uncertainty Principle) Гейзенберга - в рамках рассматриваемой темы означает, что невозможно одновременно произвольно точно зафиксировать частоту сигнала и время ее возникновения. То и другое фиксируется с ошибкой, то есть истинное значение параметров находится внутри некоторого окна. Если считать это окно прямоугольным, то его площадь будет равна произведению частоты и времени. Площадь окна величина постоянная. Именно поэтому улучшению разрешения по частоте соответствует ухудшение разрешения по времени и наоборот.

4.19. Вейвлет-пакеты (Wavelet Packet). Если в пирамиде Малла выполнять разложения не только по НЧ, но и по ВЧ каналу, то в конечном счете мы получим дерево базисов, аналогичное получаемому при выполнении быстрого преобразования Фурье. Это множество базисов и называется вейвлет-пакетами. Функции вейвлет-пакетов ортогональны между собой на каждом уровне, а также и между уровнями.

4.20. Алгоритм нахождения наилучшего базиса (Best basis). Вначале строится полное дерево вейвлет-пакетов, затем по нему ищется оптимальный в каком-то смысле путь. Для поиска оптимального пути каждому узлу дерева сопоставляется некая стоимость. Алгоритм стремится максимизировать суммарную стоимость. Он обходит дерево снизу вверх и на каждом уровне принимает решение: что выгоднее - оставить два узла-потомка или же их

родительский узел. Если стоимость узла-родителя больше суммы стоимостей потомков, то принимается решение об обрезании дерева в этом месте.

4.21. Лифтинговая (Lifting) схема вычисления вейвлет-преобразования. Альтернативный пирамиде Малла способ быстрого вычисления ВП. По некоторым оценкам снижает вычислительную стоимость в два раза, позволяет экономить память и конструировать вейвлеты, которые другим способом не построить (вейвлеты второго поколения). Лифтинговая схема состоит из трех звеньев: разбиения, предсказания и обновления (S,P,U на рисунке).



4.22. Целочисленное (Integer) вейвлет-преобразование - применяется в двух смыслах: 1)использование целочисленных фильтров; 2)получение целых значений коэффициентов преобразования целочисленного сигнала. В первом случае повышается скорость вычислений, во втором случае появляется возможность выполнения преобразования без потерь.

4.23. Мультивейвлеты получают за счет отказа от стационарности характеристик фильтров. Они могут применяться для многоканальных сигналов, причем каждый канал обрабатывается «своим» фильтром. Мультивейвлеты могут применяться и для обычных сигналов после их полифазной декомпозиции. Интересной особенностью мультивейвлетов является возможность строить ортогональные симметричные базисы, что было невозможным в случае обычных вейвлетов (кроме фильтров Хаара).

4.24. М-полосные вейвлеты. В этом случае имеется одна масштабирующая функция и несколько вейвлет-функций на каждом уровне анализа. В пирамиде Малла сигнал на каждом уровне пропускается через М фильтров.

4.25. Вейвлет-фрейм (Frame) - преобразование, сочетающее в себе свойства вейвлет-рядов и непрерывного вейвлет-преобразования. Вейвлеты соседних уровней являются сжатыми копиями друг друга (как у вейвлет-рядов), а преобразование ведется для всех сдвигов вейвлета (как при непрерывном ВП).

5. СТЕГАНОГРАФИЯ

Стеганография – наука о скрытии самого факта наличия информации. В этом ее отличие от криптографии, которая занимается скрытием содержания информации. В связи с появлением компьютеров в последние годы получила определенное развитие компьютерная стеганография. Особенно активно ведутся исследования в области цифровой стеганографии, или встраивания одних данных в другие с применением методов цифровой обработки сигналов.

5.1. Цифровая стеганография. Исследования ведутся по следующим направлениям: скрытая передача данных, внедрение цифровых водяных знаков (Watermarking), встраивание заголовков (Captioning), идентификационных номеров (Fingerprinting). Несмотря на полное различие областей применения, а также предъявляемых с их стороны требований, используемые методы встраивания во многом идентичны.

5.2. Контейнер – сообщение, в которое внедряется скрытая информация. Обычно представляет собой речь, изображение, видео, аудиосигнал. Различают потоковый (формируемый в реальном времени) и фиксированный (например, изображение) контейнеры. Контейнеры могут быть случайными, выбранными из некоторого множества, навязанными, а также являться функцией скрытого сообщения. Ясно, что в зависимости от типа контейнера, появляются различные особенности встраивания скрытых сообщений.

5.3. Стего – контейнер, заполненный скрытым сообщением.

5.4. Стегосообщение – скрываемое сообщение.

5.5. Стегосистема – система, выполняющая встраивание сообщений в контейнер, а также (как правило) их обнаружение или выделение. Состоит из предварительного кодера, стежокдера, стежокдетектора (стежокдекодера). Предварительный кодер преобразует стегосообщение к виду, удобному для внедрения. Стежокдер осуществляет это внедрение с использованием некоторого ключа. Стежокдетектор осуществляет определение наличия стегосообщения в сигнале. Стежокдекодер выделяет это сообщение. Стегосистема считается успешной в случае, если нарушитель никаким образом не может определить факт наличия стегосообщения в сигнале. Различают следующие виды стегосистем:

		Что требуется детектору		Выход детектора	
		Исходное изображение	Исходный ЦВЗ	Да/Нет	ЦВЗ
Закрытые	Тип I	+	+	+	-
	Тип II	+	-	-	+
Полузакрытые		-	+	+	-
Открытые		-	-	-	+

5.6. Цифровой водяной знак (ЦВЗ) (Digital Watermark) – внедренное в мультимедийный сигнал стегосообщение, назначением которого может быть аутентификация контента, доказательство прав его собственника, охрана прав собственников и т.д. ЦВЗ бывают видимыми и невидимыми. Наибольший интерес представляют невидимые ЦВЗ. Они делятся на три типа по выполняемым функциям: робастные, хрупкие, полухрупкие.

5.7. Робастные (robust) ЦВЗ – ЦВЗ, устойчивые к различным воздействиям (атакам). К числу возможных атак на ЦВЗ относятся, например, сжатие и другие операции ЦОС контейнера. Под устойчивостью понимается следующее: изменение контейнера, не приводящее к его существенным искажениям, не должно приводить к необнаружению ЦВЗ. Например, после выполнения фильтрации изображение может быть еще вполне приемлемым для пользователя, а нестойкий ЦВЗ будет разрушен. Робастный ЦВЗ должен быть устойчив и к атаке, направленной на «отказ в обслуживании». Например, ЦВЗ выполняет функцию защиты изображения от копирования. Повернем изображение на 90 градусов, и детектор не сможет его обнаружить (хотя он не искажен). После копирования можем повернуть изображение обратно. Для достижения робастности ЦВЗ должен встраиваться в важные детали контейнера, например, в значимые коэффициенты вейвлет-преобразования. Это противоречит требованию малозаметности ЦВЗ. Можно сказать, что робастность ЦВЗ обратно пропорциональна пропускной способности, т.е. числу встраиваемых бит.

5.8. Хрупкие ЦВЗ – ЦВЗ, разрушающиеся при изменении контента. Применяются для аутентификации, аналогично электронной цифровой подписи. Отличие заключается в том, что ЭЦП плохо подходит для защиты мультимедийного контента: незначительные изменения его приведут к отказу от приема законного сообщения. Хрупкие ЦВЗ допускают небольшую модификацию данных, не приводящих к изменению их смыслового содержания. Например, законный пользователь может пожелать сжать изображение; при передаче на него могут воздействовать шумы и т.д. Хрупкие ЦВЗ встраиваются в незначительные детали контейнера.

5.9. Полухрупкие ЦВЗ – ЦВЗ, разрушающиеся при одних типах воздействия и сохраняющиеся при других типах. Например, небольшое зашумление изображения не приведет к ликвидации ЦВЗ, а вырезка какого-то его фрагмента – приведет. Полухрупкие ЦВЗ позволяют не только определить факт вмешательства в контент, но и определить местоположение этого вмешательства и его тип. Для решения этой задачи как нельзя лучше подходят вейвлеты, в силу хорошей частотно-временной локализации.

5.10. Встраивание идентификационных номеров (Fingerprinting) – выполняет собственник контента для отслеживания его распространения. Каждая копия сообщения снабжается уникальным идентификационным номером, который не может быть удален из нее. Таким образом, с точки зрения методов встраивания, здесь много сходства с робастными ЦВЗ. Отличие заключается лишь в том, что идентификационные номера занимают небольшой объем памяти, и задача их встраивания решается легче.

5.11. Встраивание заголовков (Captioning) – может применяться, например, для аннотации медицинских снимков, что позволяет хранить изображение и его описание совместно. Пожалуй, это единственное приложение стеганографии, не носящее антагонистического характера и, поэтому, не представляющее особого интереса.

5.12. Стеганографические методы. Информация может встраиваться в области исходного сигнала (временной, пространственной) или в области преобразования. Для применения в стеганографии предлагались, наверное, все известные линейные преобразования сигналов. Наиболее часто предлагается использовать преобразование Фурье, дискретное косинусное преобразование, преобразование Хаара, вейвлет-преобразование. Некоторые авторы рекомендуют использовать для встраивания информации те же преобразования, что и при сжатии. Другие авторы, наоборот, указывают, что использование стегоалгоритмами и алгоритмами сжатия одних и тех же “дыр” неэффективно. Методы встраивания можно разделить на линейные и нелинейные. Линейные методы: В качестве приемника используется, как правило, корреляционный приемник. При этом в качестве помехи рассматривается сигнал контейнера. Нелинейные методы основаны на квантовании. Например, использованию одной кодовой книги при квантовании отсчета может соответствовать встраиванию «1», другой кодовой книги – встраиванию «0» и т.д..

6. СЖАТИЕ ИЗОБРАЖЕНИЙ

Сжатие изображений (Image compression) – сокращение объема их цифрового представления. Под изображением понимается прямоугольная матрица, элементы которой обычно принимают значения $(0..2^8-1)$ (полутоновые изображения), $(0..2^{24}-1)$ (цветные изображения), $(0,1)$ (бинарные изображения). Цветные изображения могут быть в RGB-формате (по 8 бит на канал), YCN-формате (яркостная и две цветоразностные составляющие), а также и в других форматах. Как правило, большинство алгоритмов сжатия разрабатываются для полутоновых изображений (256 градаций серого). Это связано с тем, что такие алгоритмы легко обобщить на случай цветного изображения. Различают сжатие без потерь и с потерями.

6.1. Сжатие без потерь (lossless compression) – восстановленное после сжатия изображение полностью идентично исходному («бит в бит»). То есть при сжатии используется только статистическая избыточность, имеющаяся в изображении. Ее используют так называемые энтропийные кодеры (Хаффмана, арифметический). Понятно, что таким образом большого коэффициента сжатия достичь нельзя (обычно, коэффициент сжатия не более 1.5-2). Сжатие без потерь находит ограниченное применение в некоторых областях, например, медицине или астрономии. Алгоритмы сжатия JPEG, JPEG2000 имеют режимы работы без потерь. Иногда рассматривается сжатие «почти без потерь», то есть с незаметными для человека потерями. При этом достигаются коэффициенты сжатия порядка (6..10).

Для сжатия изображений без потерь находит применение целочисленное вейвлет-преобразование. Его преимуществом является то, что коэффициенты преобразования - целые числа, поэтому после обратного преобразования можно восстановить изображение «бит в бит». Целочисленное вейвлет-преобразование выполняется для перераспределения энергии исходного изображения, так как в этом случае можно построить более эффективные энтропийные кодеры

6.2. Сжатие с потерями (Lossy Compression) - восстановленное изображение визуально мало отличается от исходного. Это различие называется искажением. Говорят, что сжатие с потерями вносит искажение. Мерой искажения может быть визуальное отличие двух изображений. Однако, эта характеристика плохо описывается на языке формул. Поэтому, на практике в качестве меры искажения используется среднеквадратическая ошибка. В среде исследователей не прекращаются споры об адекватности этой меры, однако ничего лучшего по сей день не придумано.

Сжатие с потерями выполняется, как правило, в три этапа: линейное преобразование (ДКП, вейвлет), квантование, энтропийное кодирование. Целей первого этапа несколько: перераспределение энергии между отсчетами, декорреляция отсчетов и придание коэффициентам удобной для квантования структуры (например, нульдеревя). Квантование бывает скалярное (каждый отсчет в отдельности) и векторное. В качестве энтропийного кодера обычно используется арифметический (более эффективный, но и более вычислительно сложный) или кодер Хаффмана.

6.3. Перераспределение энергии - одним из предназначений линейного преобразования изображения при его сжатии является перераспределение энергии. Оно заключается в том, что большая часть энергии изображения после преобразования оказывается сосредоточенной в малой части коэффициентов (низкочастотных). Поэтому многие коэффициенты можно «обнулить», не передавать приемнику и, тем самым, достичь сжатия. Кроме того известно, что квантование неравномерно распределенной случайной величины более эффективно, чем равномерно распределенной.

6.4. Декорреляция - одним из предназначений линейного преобразования изображения при его сжатии является декорреляция коэффициентов преобразования. Отсчеты исходного изображения коррелированы друг с другом (коэффициент корреляции достигает 0.95). Это означает, что любой можно довольно точно восстановить по соседним отсчетам, то есть имеется избыточность. Оптимальным декоррелирующим преобразованием (для гауссовского процесса) является преобразование Карунена-Лоэва. Все остальные преобразования только приближаются к нему.

6.5. Средняя квадратическая ошибка – основная мера искажения, возникающего при сжатии изображений. Определяется как сумма квадратов разностей значений пикселей исходного и восстановленного изображений. Некоторые исследователи считают, что СКО слабо коррелировано с субъективным восприятием искажений изображения. Однако, ничего лучшего не придумано. Вариацией является взвешенное СКО, когда измеренное в субполосах преобразования СКО умножается на весовой коэффициент с учетом свойств системы человеческого зрения.

6.6. Нульдерев – одна из наиболее эффективных структур, упорядочивающих множество вейвлет-коэффициентов изображения для целей их сжатия. Заключается в следующем: вейвлет-коэффициенты более НЧ областей «отвечают» за четыре коэффициента более ВЧ областей. Тем, в свою очередь, соответствуют 16 коэффициентов еще более ВЧ областей и т.д. Таким образом, получаем дерево коэффициентов с вершиной в самой НЧ области. Выберем некоторый порог и назовем меньшие этого порога (по абсолютной величине) коэффициенты незначимыми.

Начиная с самой НЧ области проверяем значимость коэффициентов относительно текущего порога. Если коэффициент незначимый, проверяем его потомков по дереву. Если все потомки незначимые, то на месте коэффициента ставится признак нульдерев. Таким образом одним символом обозначаем большое множество коэффициентов. После «обхода» всех коэффициентов порог уменьшается в два раза, и процесс повторяется.

Сжатие на основе нульдерев было впервые предложено Льюисом и Ноулесом (1992) и улучшено в дальнейшем Шапиро (1993) и Саидом-Перельманом (1995).

6.7. Скалярное квантование – квантование каждого отсчета в отдельности. Под квантованием цифрового сигнала понимается процесс замены значения отсчета из одного счетного множества значением из другого счетного множества меньшей мощности. За счет этого достигается уменьшение числа требуемых для кодирования отсчета бит, но возникает ошибка кодирования. Расстояния между соседними возможными значениями квантованных отсчетов называют интервалами квантования. Различают равномерное квантование, логарифмическое квантование, оптимальное квантование. Оптимальные интервалы квантования находятся с применением процедуры Ллойда-Макса (квантователь Макса).

6.8. Векторное квантование – квантуется сразу некоторое множество отсчетов (вектор). Если скалярное квантование одномерный процесс, то векторное – многомерный, то есть гораздо более сложный. Построение оптимального векторного квантователя основано также на применении алгоритма Ллойда-Макса. Сложность кодирования и декодирования намного выше, чем при скалярном квантовании. Поэтому, обычно применяются структурированные методы векторного квантования (то есть кодовой книге придается структура): древовидное, решетчатое, пирамидальное ВК.

6.9. Древовидный векторный квантователь – кодовые слова упорядочены в виде дерева. Таким образом, сложность поиска по дереву высотой N будет $N \log N$ вместо N^N , как было бы при не структурированной кодовой книге. Проигрыш заключается в неоптимальности найденных кодовых слов. Разновидностью является усеченный древовидный векторный квантователь (PTSVQ), суть которого заключается в следующем. Вначале строится полное дерево, затем оно усекается (аналогично, как в алгоритме вейвлет-пакетов). Таким образом получается адаптивный квантователь.

6.10. Решетчатый квантователь – кодовые слова книги расположены в узлах решетки. Каких-то преимуществ по сравнению с древовидным я не знаю. По поводу теории решеток смотрите замечательную книгу Конвэя и Слоэна.

6.11. Стандарты на сжатие изображений – сжатие бинарных изображений – JBIG, сжатие полутоновых и цветных – JPEG, JPEG2000. В стандарте JPEG в качестве линейного преобразования используется ДКП, в JPEG2000 – вейвлет-преобразование. На роль стандарта де-факто для сжатия смешанных текстографических изображений претендует алгоритм DjVu («дежавю»), основанный на распознавании образов и последующем применении различных алгоритмов к различным типам изображений.

6.12. Стандарты на сжатие видео – MPEG-2, MPEG-4 (ДКП и вейвлеты, соответственно), рекомендация МСЭ H.263, H.263+ и др. Сжатие видео состоит из двух частей: сжатие неподвижного кадра (во многом аналогично сжатию изображений) и межкадровое кодирование (в основном применяют что-то типа ДИКМ).