

3.1. Методы преобразования контейнера в пространство признаков

Преобразование в пространство признаков

2

□ Основные форматы данных в СВИ:

- \mathbb{X}_{\square}^m , где $\mathbb{X} \subseteq \mathbb{R}$ – множество цифровых сигналов
- \mathbb{Y}_{\square}^l , где $\mathbb{Y} \subseteq \mathbb{C}$ – множество матриц признаков

□ Порядок преобразований носителя информации

$$C \in \mathbb{X}_{\square}^m \rightarrow f \in \mathbb{Y}_{[\square]}^l \rightarrow f^W \in \mathbb{Y}_{\square}^l \rightarrow C^W \in \mathbb{X}_{\square}^m$$

□ До сих пор мы рассматривали упрощённый вариант:

$$C \in \mathbb{X}_{\square}^m \rightarrow C^W \in \mathbb{X}_{\square}^m$$

□ Формы представления встраиваемой информации (тоже преобразуются из одной в другую по мере необходимости)

- бинарный вектор $\mathbf{b}, \mathbf{b}^R \in \mathbb{B}_{[N_b]}^1$
- цифровой сигнал $W, W^R \in \mathbb{X}_{\square}^m$
- матрица признаков $\Omega, \tilde{\Omega} \in \mathbb{Y}_{\square}^l$

Способы перехода в пространство признаков

3

- 1. Пространство признаков = пространственная область (Spatial Domain)
 - ▣ Исходные отсчёты
 - ▣ Смена цветового пространства и выделение компонент для встраивания
- 2. Спектральное преобразование (Spectral Domain)
 - ▣ Дискретное преобразование Фурье
 - ▣ Дискретное косинусное преобразование
 - ▣ Дискретное вейвлет-преобразование
- 3. Трансформанты, рассчитываемые при сжатии данных
 - ▣ JPEG, JPEG 2000, WebP, JPEG-XR,...
 - ▣ MJPEG, H.263, H.264, WebM,...
 - ▣ MP3, AAC,...

Пространственная область. Встраивание в яркостную или наиболее значимую компоненту

4

□ Пространство RGB

- $f(n_1, n_2) = \max\{C_R(n_1, n_2), C_G(n_1, n_2), C_B(n_1, n_2)\}$

Далее для поэлементных формул не будем вводить координаты пикселей для лаконичности:

$$f = \max\{C_R, C_G, C_B\}$$

- $f = (C_R + C_G + C_B)/3$

- $f = C_G$

□ Другие цветовые пространства

- $f = C_{V[HSV]} = \max\{C_R, C_G, C_B\}$

- $f = C_{L[HSL]} = (\max\{C_R, C_G, C_B\} + \min\{C_R, C_G, C_B\})/2$

- $f = C_{Y[YCbCr]} = \frac{77}{256} C_R + \frac{150}{256} C_G + \frac{29}{256} C_B$

Пространственная область. Встраивание в наименее значимую компоненту

5

- Пространство RGB

- $f = C_B$

- Другие цветовые пространства

- $f = C_{S[HSV]}$

- $f = C_{S[HSL]}$

- $f = C_{Cb[YCbCr]}$ или $f = C_{Cr[YCbCr]}$ или $f = (C_{Cb[YCbCr]}, C_{Cr[YCbCr]})$

Развёртки двумерной области в вектор

6

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

Построчная

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

Серпантинная

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

Гильберта-Пеано

- $f(m) = C(\mu(m))$, где $\mu: \mathbb{R} \mapsto \mathbb{R}^2$
- Может быть и развёртка 3D в 2D или в 1D
- Развёртки применяются не только при встраивании в пространственной области

Встраивание в спектральной области

7

- Крупные объекты на изображении человек воспринимает лучше, чем мелкие детали
- = система человеческого зрения более чувствительна к низкочастотной информации, нежели к высокочастотной
- \Rightarrow логичный подход - переход от пространственного представления сигнала к частотному с последующей модификацией уже не пикселей изображения, а спектральных компонент
- Как правило, используются преобразования, для которых известны **быстрые прямые и обратные** алгоритмы и спектральные компоненты которых являются **слабо коррелированными**.

Встраивание в спектральной области

8

- 1. Расчёт ДОП изображения C (будем обозначать его C_{DOT})

$$C_{DOT}(m_1, m_2) = \sum_{n_1=0}^{N_1} \sum_{n_2=0}^{N_2} C(n_1, n_2) h_{m_1}(n_1) g_{m_2}(n_2),$$

- ▣ где $\{h_{m_1}(n_1)\}_{m_1=0}^{N_1-1}, n_1 = 0..N_1 - 1, \{g_{m_2}(n_2)\}_{m_2=0}^{N_2-1}, n_2 = 0..N_2 - 1$ – два семейства ортогональных базисных функций, отличающихся только периодом (то есть при $N_1 = N_2$ они идентичны)

- Сокращённо:

$$C_{DOT}(m_1, m_2) = \mathcal{F}(C(n_1, n_2))$$

Встраивание в спектральной области

9

- 2. Выбор подмножества спектральных компонент для встраивания информации (пространства признаков):

$$f(m) = C_{DOT}(\mu(m)), m = 0..M - 1,$$

- где отображение $\mu: \mathbb{R} \mapsto \mathbb{R}^2$ либо жёстко фиксировано в рамках конкретного алгоритма, либо определяется на основе секретного ключа **k**.
- Шаг 2 у многих систем отсутствует – используются все отсчёты C_{DOT} и двумерное представление данных.

Встраивание в спектральной области

10

- 3. Собственно встраивание информации в пространстве признаков:

$$f^W = \mathcal{E}(f, \Omega, \mathbf{k}).$$

- Матрицу спектральных компонент после встраивания будем обозначать $C_{DOT}^W(m_1, m_2)$, если отбор компонент (шаг 2) проводился.
- 4. Переход в пространственные координаты при помощи обратного ДОП:

$$C^W(n_1, n_2) = \sum_{m_1=0}^{N_1} \sum_{m_2=0}^{N_2} C_{DOT}^W(m_1, m_2) h_{n_1}^{-1}(m_1) g_{n_2}^{-1}(m_2),$$

- сокращённо

$$C^W(n_1, n_2) = \mathcal{F}^{-1} \left(C_{DOT}^W(m_1, m_2) \right)$$

Семейства базисных функций

11

$$\{h_m(n)\}_{m=0}^{N-1}, n = 0..N-1$$

- Базис ДПФ

$$h_m(n) = \frac{1}{\sqrt{N}} \exp \frac{2\pi i m n}{N}$$

- Базис ДП Хартли

$$h_m(n) = \frac{1}{\sqrt{N}} \left(\cos \frac{2\pi m n}{N} + \sin \frac{2\pi m n}{N} \right)$$

- Базис ДКП

$$h_m(n) = \lambda_m \cos \frac{\pi m \left(n + \frac{1}{2} \right)}{N},$$

$$\lambda_m = \begin{cases} 1/\sqrt{N}, & m \neq 0 \\ 2/\sqrt{N}, & m = 0 \end{cases}$$

Дискретное вейвлет-преобразование

12

- Это целый класс преобразований, основанных на использовании различных семейств вейвлет-функций
- Оно заключается в итеративном применении ко входному сигналу двух функций с разными масштабами и сдвигами. Первая из этих функций называется *скейлинг-функцией* (scaling function), вторая – *вейвлетом* (wavelet).
- Особенность: можно варьировать число итераций, называемое числом уровней декомпозиции. Это экономит время и позволяет не разделять низкочастотные компоненты, если это не требуется.
- Самое простое семейство – семейство Хаара. В качестве функций используются вычисление полусумм и полуразностей

Иллюстрация ДВП Хаара в одномерном случае

13

(1,2,3,4,5,6,7,8)

$$\left(\underbrace{\frac{3}{2}, \frac{7}{2}, \frac{11}{2}, \frac{15}{2}}_{L_1}, \underbrace{-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}}_{H_1} \right)$$

$$\left(\underbrace{\frac{5}{2}, \frac{13}{2}}_{L_2}, \underbrace{-1, -1}_{H_2}, \underbrace{-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}}_{H_1} \right)$$

$$\left(\underbrace{\frac{9}{2}}_{L_3}, \underbrace{-2}_{H_3}, \underbrace{-1, -1}_{H_2}, \underbrace{-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}}_{H_1} \right)$$

ДВП в двумерном случае

14

- Итерации вейвлет-декомпозиции осуществляются то по горизонтали, то по вертикали
- Различают стандартное и пирамидальное разложение

Пирамидальное разложение

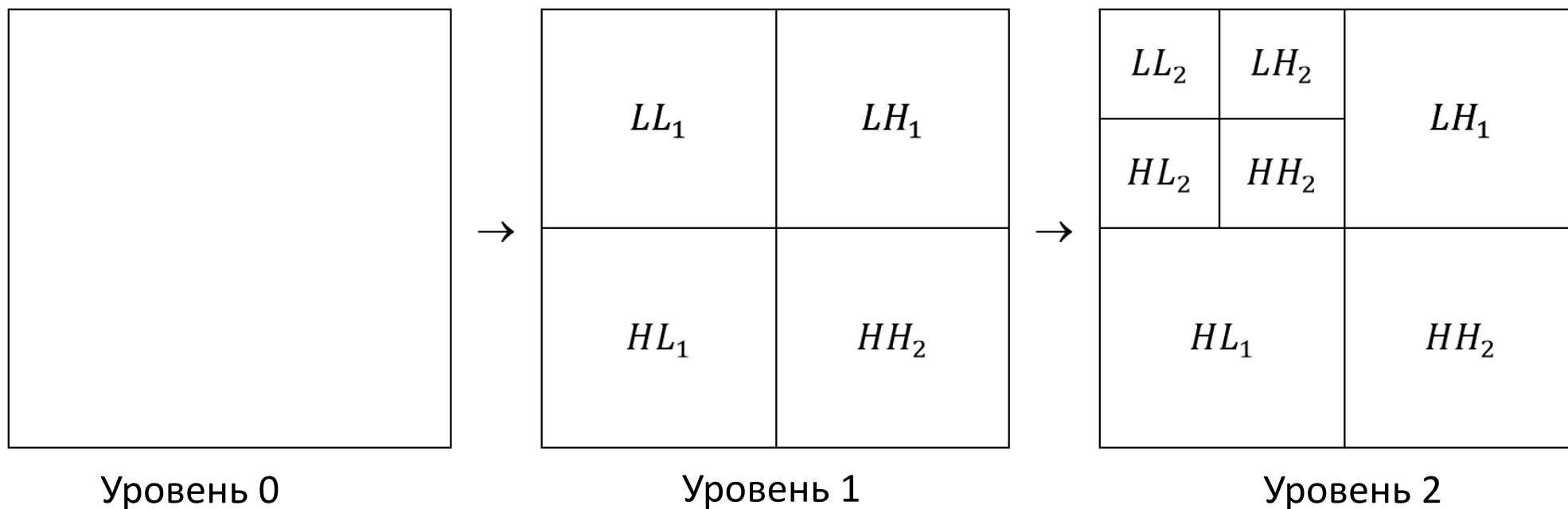
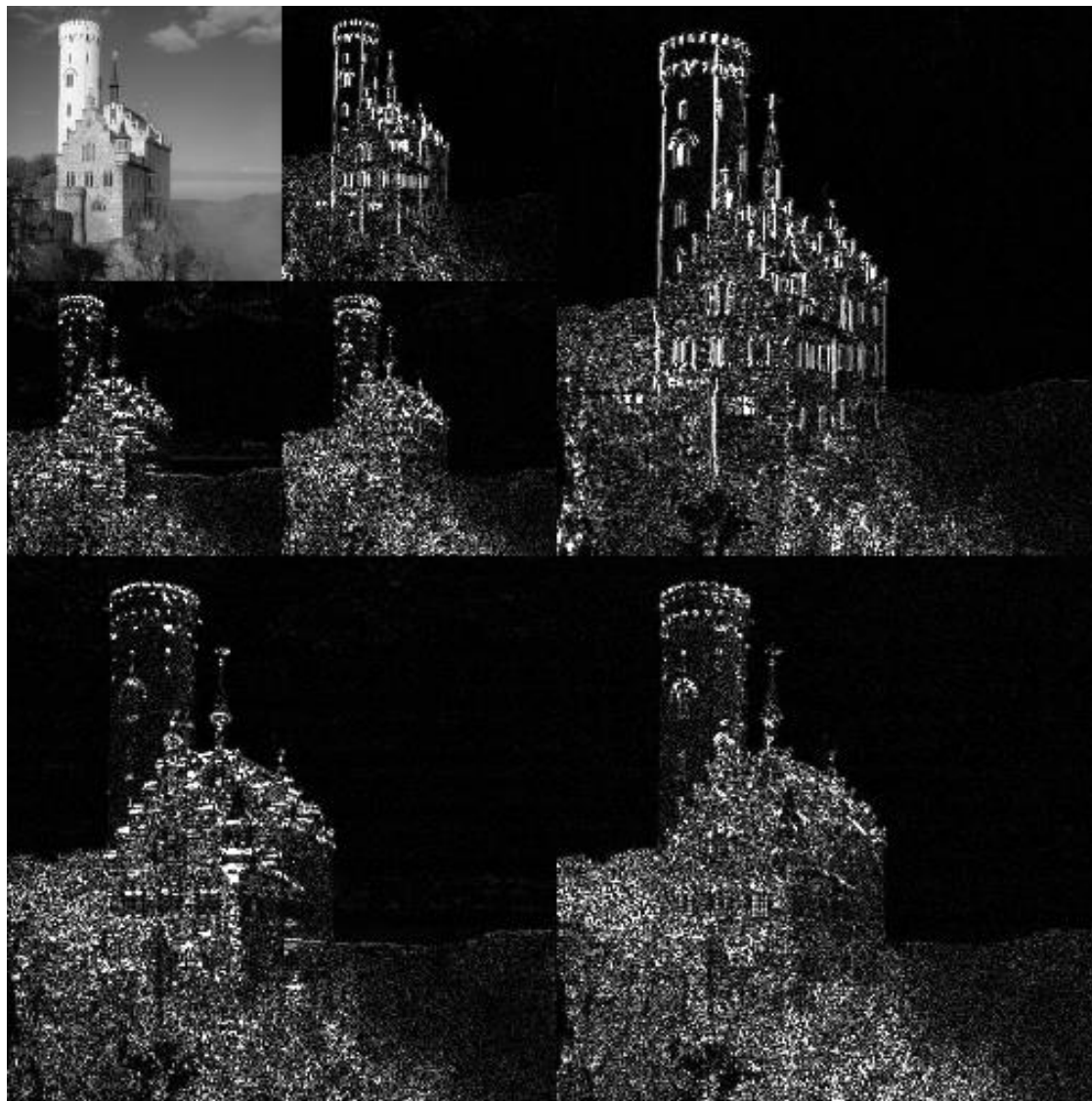


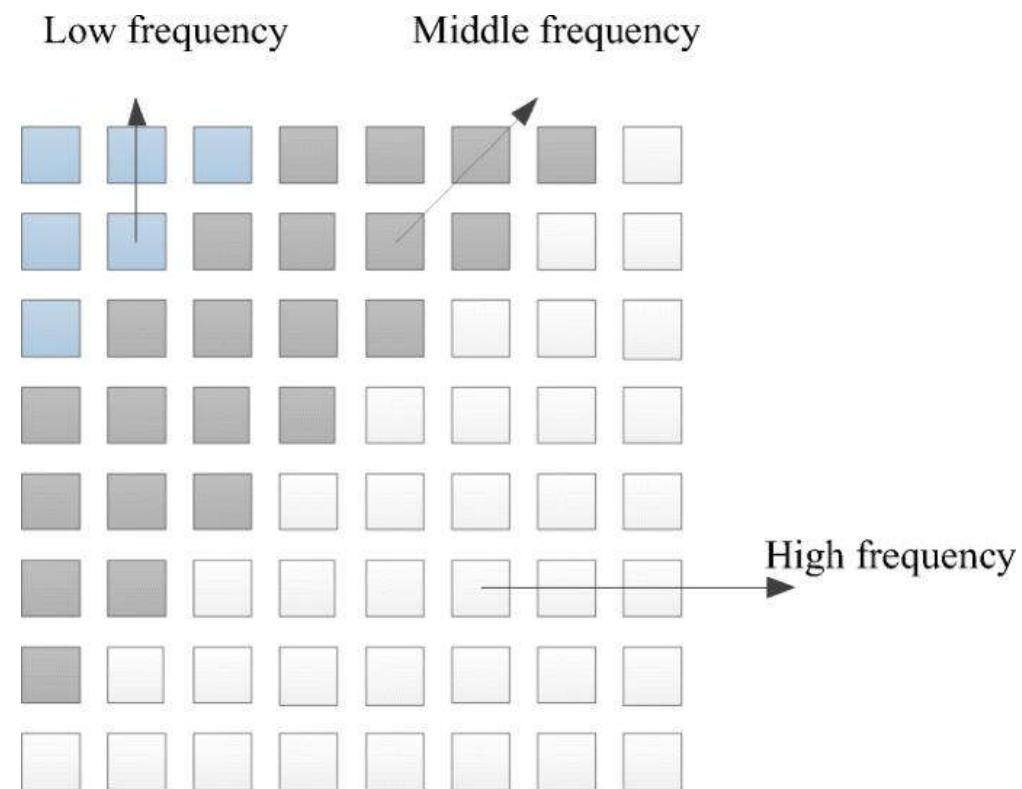
Иллюстрация ДВП

15



Частотные области в спектре ДКП

16



Разделение на частотные области

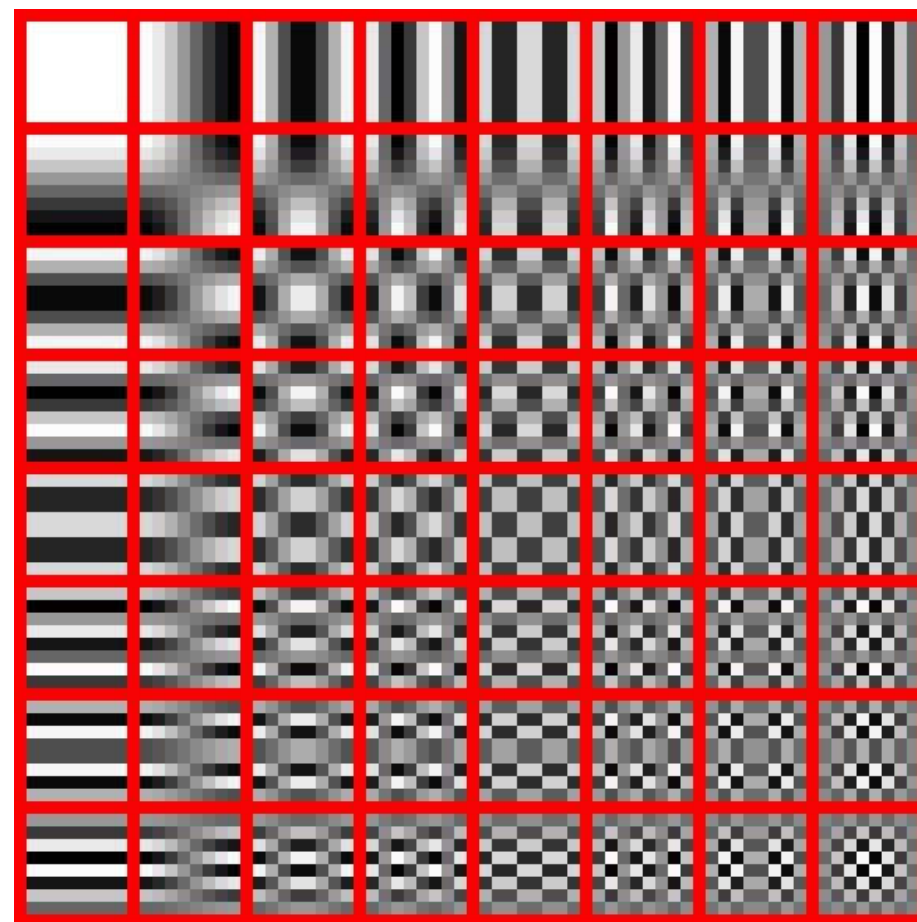
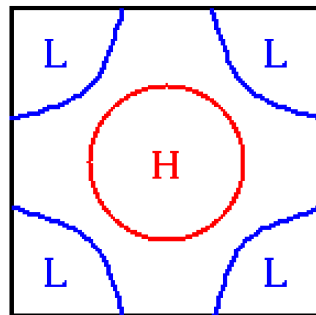


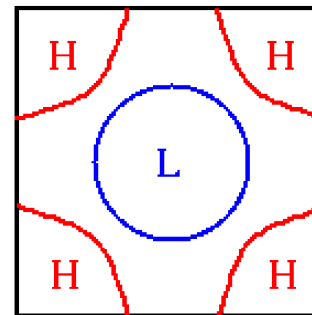
Иллюстрация базисных функций

Частотные области в спектре ДПФ

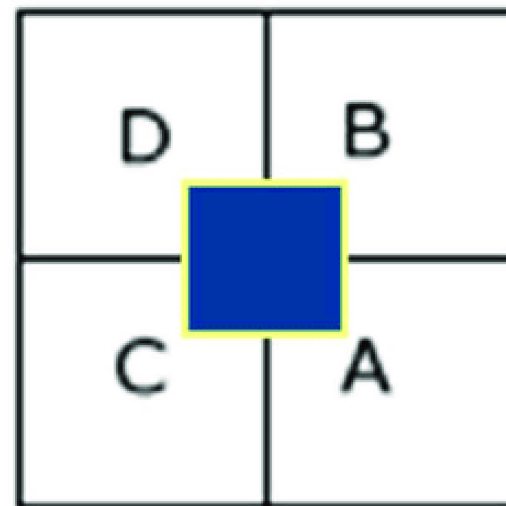
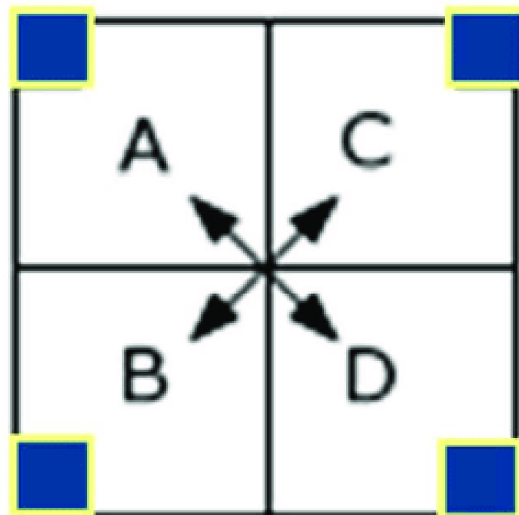
17



before
centralization

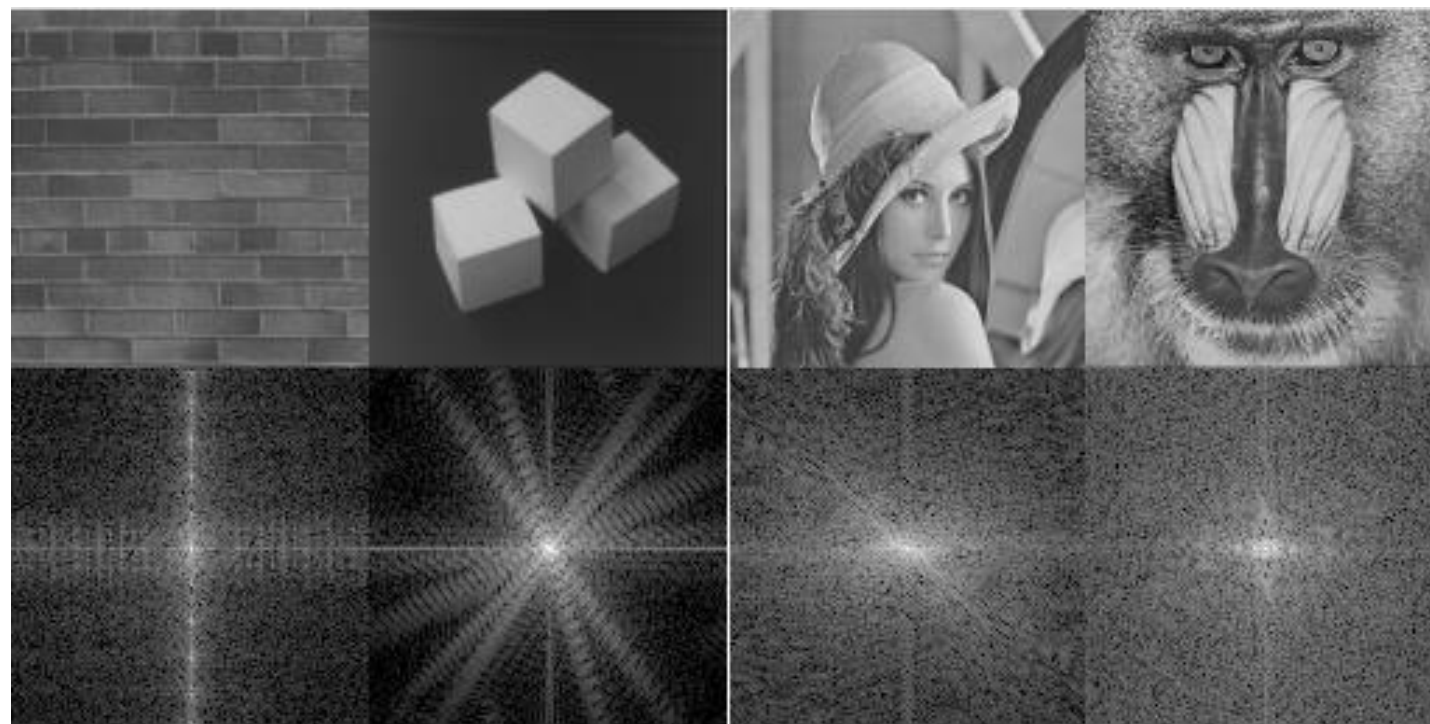
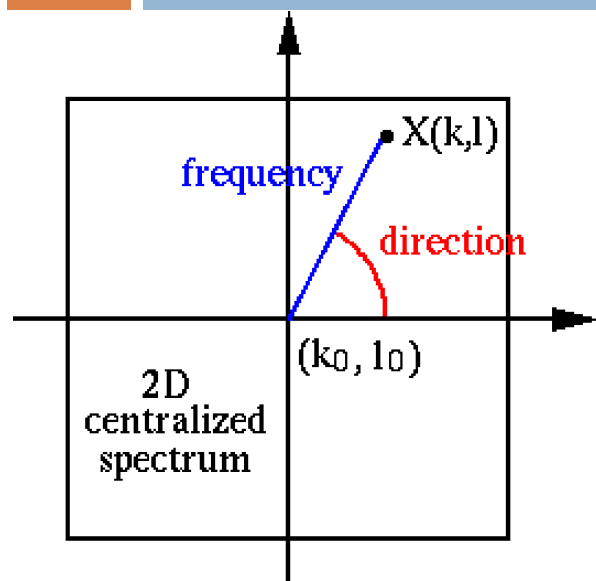


after
centralization



Как читать спектр ДПФ

18



Преобразование Фурье-Меллина

19

- Одно из важных требований к ЦВЗ – стойкость к геометрическим преобразованиям
- Оно может быть выполнено благодаря особой структуре преобразования
- Пример: для обеспечения стойкости к циклическим сдвигам в качестве пространства признаков может быть выбран модуль спектра Фурье
- Преобразование Фурье-Меллина обеспечивает стойкость:
 - ▣ к циклическому сдвигу
 - ▣ к изотропному масштабированию
 - ▣ к повороту

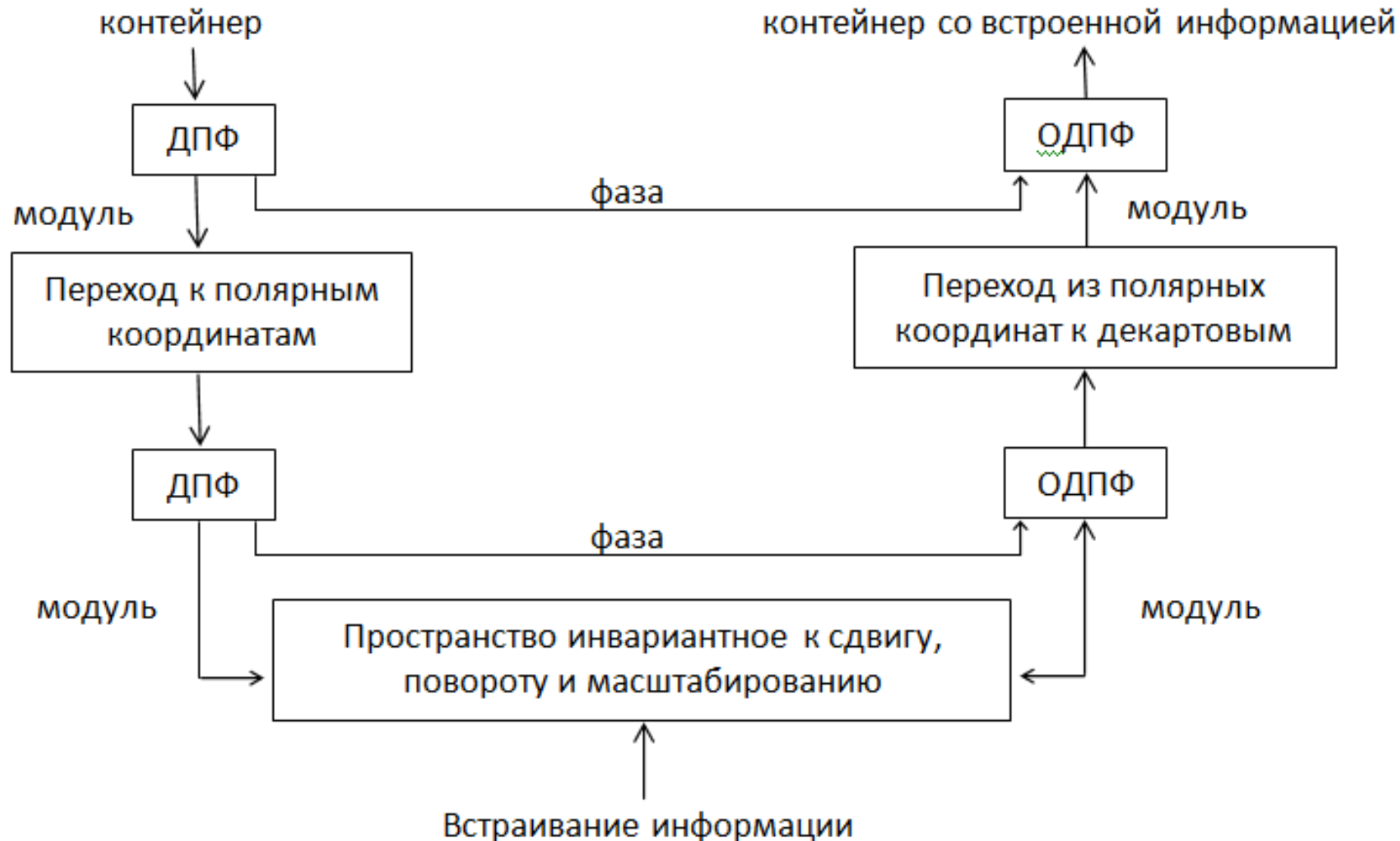
Преобразование Фурье-Меллина

20

- Масштабирование в пространственной области приводит к обратному масштабированию в области спектральной.
- Поворот в пространственной области приводит к такому же повороту в спектральной области
- Идея: масштабирование и поворот есть сдвиг в полярных координатах!
- $(m_1, m_2) \rightarrow \begin{cases} m_1 = e^\rho \cos \theta \\ m_2 = e^\rho \sin \theta \end{cases} \rightarrow (\rho, \theta)$
- Для чего e^ρ , а не ρ ?

Преобразование Фурье-Меллина. Полная схема

21



Встраивание в трансформанты JPEG

22

- В отдельном файле “JPEG Compression.pdf”

Встраивание в трансформанты MP3.

Понятие битрейта

23

- Битрейт – величина, показывающая скорость передачи аудиоданных и измеряющаяся в битах в секунду.
- Для несжатых форматов битрейт можно вычислить по следующей формуле:

$$B_r = F_S \times m \times N_{ch},$$

F_S – частота дискретизации; m – глубина квантования, или количество бит на отсчет; N_{ch} – количество каналов

- Для одноканального сигнала с частотой дискретизации 44100 Гц и глубиной квантования 16 бит на отсчет битрейт будет равен 705,6 кбит/с, а для двухканального – 1411,2 кбит/с
- Для MP3 битрейт задается при кодировании. Самые распространенные значения: 128, 192, 256 и 320 кбит/с.

Встраивание в трансформанты MP3.

Режимы сжатия

24

- CBR (Constant Bit Rate, постоянный битрейт): битрейт задается пользователем при кодировании и не меняется на протяжении всего файла;
- VBR (Variable Bit Rate, переменный битрейт): битрейт динамически изменяется при кодировании в зависимости от насыщенности аудиосигнала и установленного пользователем качества кодирования;
- ABR (Average Bit Rate, усредненный битрейт): битрейт задается пользователем, а кодек варьирует его так, чтобы соответствовать заданному.

Встраивание в трансформанты MP3.

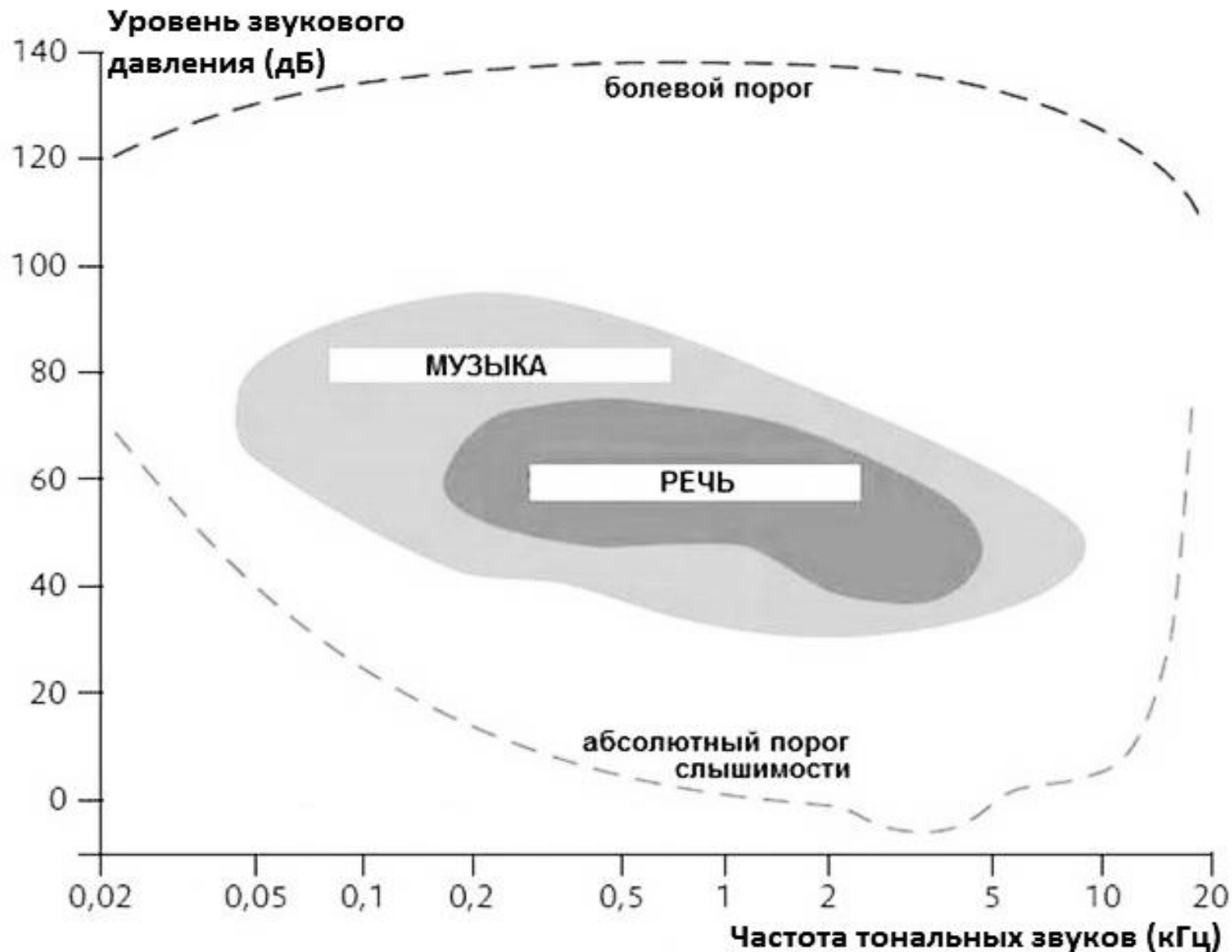
Разделение на фреймы исходного сигнала

25

- Исходный сигнал делится на кадры (frames), которые содержат 576 отсчетов при одноканальной аудиодорожке и 1152 отсчета при двухканальной.
- При частоте дискретизации 44100 Гц один кадр будет иметь длительность 26 мс
- Каждый кадр преобразуется независимо от других и после такого преобразования состоит из самих кодированных данных и дополнительной информации, которая необходима для их декодирования.
- При записи в конечный MP3-файл полученные преобразованные кадры komponуются в единый поток данных.

Встраивание в трансформанты МРЗ. Диапазон слышимых звуков с цифрами

26



Встраивание в трансформанты MP3.

Процедура сжатия кадра (режим CBR)

27

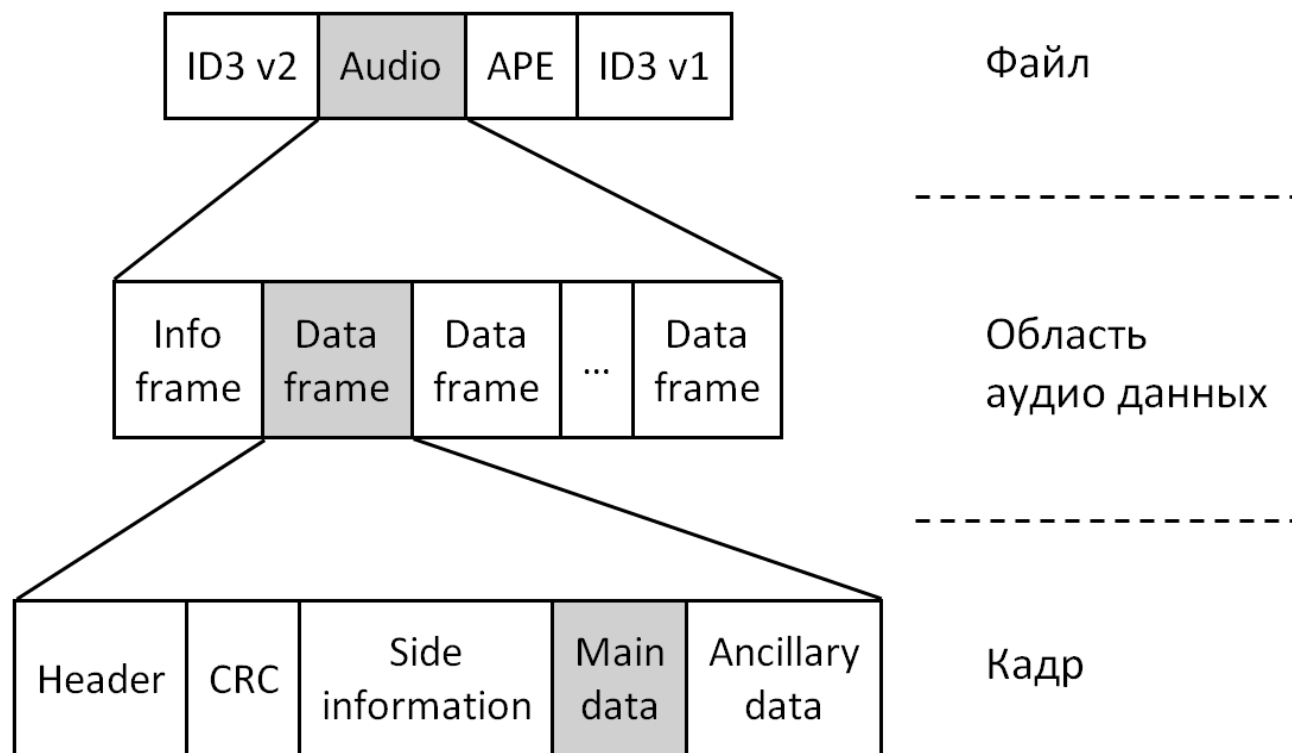
- МДКП кадра, затем БПФ
- Убираются компоненты выше 16 кГц и маскированные сигналы, согласно принятой психоакустической модели.
- Далее 3 основных операции: деление на коэффициенты по 22 полосам, квантование и кодирование Хаффмана.
- Они реализуются в двух вложенных циклах по отношению к 576 спектральным значениям за раз
- Внешний цикл – цикл контроля искажений, внутренний – цикл контроля битрейта.
- Контроль битрейта: перебор шагов квантования начиная с малого
- Контроль искажений: перебор значений масштабных коэффициентов. Цель этого цикла – держать шум квантования ниже порога маскировки

Встраивание в трансформанты MP3.

Конечные фреймы. Логическая структура файла

28

- Полученные параметры вместе с самими данными формируют кадр архивного формата
- Все кадры записываются последовательно в конечный MP3-файл и могут быть декодированы независимо от последующих
- Кадр конечного файла содержит 1-2 кадра исходного сигнала



Улучшения в формате AAC

29

- Больше двух каналов (до 48)
- Произвольная частота дискретизации от 8 до 96 кГц
- Больше в 2 раза максимальная частота дискретизации
- Не применяется ДПФ, МДКП длиной 2048
- И мн. Др.