

# ОБРАБОТКА И РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ

Леонид Моисеевич Местецкий  
профессор

кафедра математических методов  
прогнозирования ВМК МГУ  
кафедра интеллектуальных систем МФТИ

# Идея генерации признаков на основе линейных преобразований

- 1) Представить исходный образ в виде линейной комбинации базисных образов
- 2) Составить вектор признаков из коэффициентов разложения  $(y_0, \dots, y_{N-1})$

$$\boxed{\begin{array}{c} \text{Исходный} \\ \text{образ} \end{array}} = \sum y_i \times \boxed{\begin{array}{c} i\text{-й базисный} \\ \text{образ} \end{array}}$$

# Представление базисных

## ВЕКТОРОВ

$A = [a(0), a(1), \dots, a(N-1)]$  - представление матрицы в виде столбцов.

Вектора-столбцы  $a(i)$  называются *базисными векторами*.

$$y = A^H x = \begin{pmatrix} a_0^H \\ a_1^H \\ \vdots \\ a_{N-1}^H \end{pmatrix} \cdot x$$

где  $a^H(0), a^H(1), \dots, a^H(N-1)$  – строки из транспонированных столбцов  $a(i)$ .

$$x = \sum_{i=0}^{N-1} y(i) \cdot a_i$$

# Преобразование Карунена- Лоева

1. Рассматриваем вектор измерений  $x$  как вектор случайных величин
2. Тогда полученный вектор признаков также является случайным.
3. Целью преобразования  $y = A^T x$  является такой выбор матрицы базисных векторов  $A$ , чтобы вектор  $y$  имел некоррелированные компоненты: математическое ожидание  $E[y(i) \cdot y(j)] = 0$  для всех  $i \neq j$ .
4. Такое свойство признаков характеризует их «минимальность», поскольку нет «избыточности описания»: каждый признак не дублирует остальные признаки.

# Альтернатива – генерация признаков на основе преобразования Фурье

1. Преобразование Карунена-Лоева требует построения вероятностной модели для описания вектора измерений.
2. Вычисление базисных векторов требует больших вычислительных затрат.
3. Преобразование Фурье – это разложение по универсальному базису. Оно позволяет обойтись без вероятностной модели и снизить вычислительные затраты.

# Непрерывное преобразование Фурье (примерно 1807-1822)

Жан Батист Жозеф Фурье (Jean Baptiste Joseph Fourier) 1768-1830, Париж, французский математик и физик.



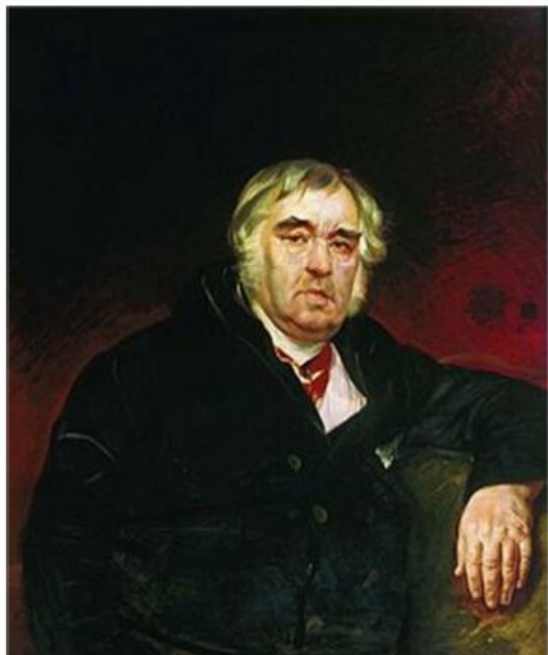
Прямое преобразование

$$X(f) = \int_{-\infty}^{+\infty} x(t) \exp(-j 2 \pi f t) dt \equiv \mathcal{F} \{x(t)\}$$

Обратное преобразование

$$x(t) = \int_{-\infty}^{+\infty} X(f) \exp(j 2 \pi f t) df \equiv \mathcal{F}^{-1} \{X(f)\}$$

# А что в России в это время?



**Иван Андреевич Крылов (1769 – 1844)**

**Водолазы (басня 1813 год)**

Какой-то древний царь впал в страшное  
сомнение:

Не более ль вреда, чем пользы, от наук?

Не расслабляет ли сердце и рук ученье?

И не разумнее ль поступит он,

Когда ученых всех из царства вышлет вон?

.....

Хотя в ученье зрим мы многих благ  
причину,

Но дерзкий ум находит в нем пучину

И свой погибельный конец,

Лишь с разницею тою,

Что часто в гибель он других влечет с  
собою.

# Одномерное ДПФ

$x(0), x(1), \dots, x(N-1)$  – вектор исходных измерений.

ДПФ определяется следующим образом:

$$y(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x(n) \cdot \exp\left(-i \frac{2\pi}{N} kn\right)$$

$$k = 0, 1, \dots, N-1$$

$$\exp\{\alpha\} = \cos(\alpha) + i \cdot \sin(\alpha),$$

$i$  - мнимая единица.

Обратное преобразование есть:

$$x(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} y(k) \cdot \exp\left(i \frac{2\pi}{N} kn\right)$$

$$n = 0, 1, \dots, N-1$$



# Базисные вектора ДПФ

Определим

$$W_N = \exp\left(-i \cdot \frac{2\pi}{N}\right) = \cos\left(\frac{2\pi}{N}\right) - i \cdot \sin\left(\frac{2\pi}{N}\right).$$

Тогда

$$\exp\left(-i \cdot \frac{2\pi}{N} \cdot kn\right) = \left(\exp\left(-i \cdot \frac{2\pi}{N} \cdot \cdot\right)\right)^{kn} = W_N^{kn}.$$

Пусть

$$y = W^H x,$$

тогда

$$x = W y$$

Матрица базисных векторов:

$$W^H = \left\| \frac{1}{\sqrt{N}} W_N^{kn} \right\|, \quad k, n = 0, \dots, N - 1$$

# Матрица базисных векторов

$$W^H = \frac{1}{\sqrt{N}} \cdot \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & W_N & \dots & W_N^{N-1} \\ 1 & W_N^2 & \dots & W_N^{2(N-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & W_N^{N-1} & \dots & W_N^{(N-1)(N-1)} \end{pmatrix}.$$

Утверждается, что  $W$  – унитарная симметрическая матрица. Пусть  $W^*$  – сопряженная матрица:  $W^* = W^H = W^{-1}$ . Тогда базисные вектора – это столбцы матрицы  $W$ .

Таким образом, имеет место разложение  $x = \sum_{i=0}^{N-1} y(i) \cdot w_i$  по базису  $w_0, w_1, \dots, w_{N-1}$  векторов-столбцов матрицы  $W$ .

# Пример

$$N = 4 \quad \rightarrow \quad W_N = \exp\left(-i\frac{2\pi}{N}\right)$$

$$W_4 = \exp\left(-i\frac{2\pi}{4}\right) = \cos\frac{\pi}{2} - i \cdot \sin\frac{\pi}{2} = -i$$

$$W_4^2 = \exp(-i \cdot \pi) = \cos(-\pi) - i \cdot \sin(-\pi) = -1$$

$$W_4^3 = \exp\left(-i\frac{2\pi}{4} \cdot 3\right) = \cos\left(-\frac{3\pi}{2}\right) + i \cdot \sin\left(-\frac{3\pi}{2}\right) = i$$

$$W_4^4 = \exp(-i \cdot 2\pi) = 1$$

# Базисные вектора

$$W^H = \frac{1}{2} \cdot \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix}$$

$$W = \frac{1}{2} \cdot \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}$$

$$w_0 = \frac{1}{2} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$w_1 = \frac{1}{2} \cdot \begin{pmatrix} 1 \\ i \\ -1 \\ -i \end{pmatrix}$$

$$w_2 = \frac{1}{2} \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

$$w_3 = \frac{1}{2} \cdot \begin{pmatrix} 1 \\ -i \\ -1 \\ i \end{pmatrix}$$

$W = (w_0, w_1, w_2, w_3)$  - матрица из базисных векторов

# Унитарная матрица

$$W^H \cdot W = \frac{1}{4} \cdot \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix} \times \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix} =$$
$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = W \cdot W^{-1}$$

# ДПФ – разложение по базисным последовательностям

$x(n)$  – последовательность,

$h_k(n)$  – множество базисных последовательностей.

$$x(n) = \sum_{k=0}^{N-1} y(k) \cdot h_k(n)$$

$$h_k(n) = \begin{cases} \frac{1}{\sqrt{N}} \exp\left(i \cdot \frac{2\pi}{N} \cdot k \cdot n\right), & n = 1, \dots, N - 1 \\ 0, & \text{иначе.} \end{cases}$$

$$\left( h_k(n), h_l(n) \right) = \sum_n h_k(n) \cdot h_l(n) = \delta_{kl} \text{ – ортонормальные последовательности}$$

# Двумерное ДПФ

Пусть  $X(m, n)$ ,  $m, n = 0, 1, \dots, N-1$  – двумерные измерения. Тогда двумерное ДПФ есть:

$$Y(k, l) = \frac{1}{N} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} X(m, n) W_N^{k \times m} W_N^{l \times n}$$

Обратное преобразование:

$$X(m, n) = \frac{1}{N} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} Y(k, l) W_N^{-k \times m} W_N^{-l \times n}$$

Данную запись компактно можно переписать в следующем виде:

$$Y = W^H X W^H, \quad X = W Y W$$



# Дискретное косинусное преобразование

$$y(k) = \alpha(k) \sum_{n=0}^{N-1} x(n) \cdot \cos\left(\frac{\pi(2n+1)k}{2N}\right), \quad k = 0, 1, \dots, N-1,$$

$$x(n) = \sum_{k=0}^{N-1} \alpha(k) \cdot y(k) \cdot \cos\left(\frac{\pi(2n+1)k}{2N}\right), \quad n = 0, 1, \dots, N-1$$

$$\alpha(k) = \begin{cases} \frac{1}{\sqrt{N}}, & k = 0 \\ \sqrt{\frac{2}{N}}, & k \neq 0 \end{cases}$$

$$y = C^T x,$$

$C$  – ортогональная матрица,  $C^{-1} = C^T$ .

$$C(n, k) = \frac{1}{\sqrt{N}}, \quad k = 0, \quad 0 \leq n \leq N-1,$$

$$C(n, k) = \sqrt{\frac{2}{N}} \cos\left(\frac{\pi(2n+1)k}{2N}\right), \quad 1 \leq k \leq N-1, \quad 0 \leq n \leq N-1.$$



# Двумерное косинусное преобразование

$X(m, n)$ ,  $m, n = 0, 1, \dots, N - 1$  – двумерные измерения

$$Y = C^T X C$$

$$X = C Y C^T$$

# Дискретное синусное преобразование

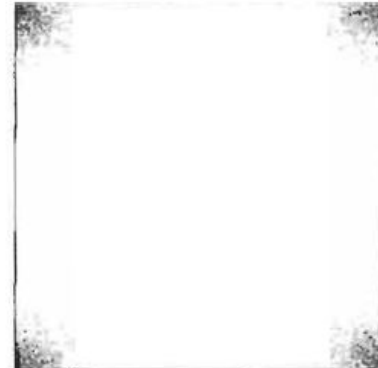
$$y = S^T x$$

$$S(k, n) = \sqrt{\frac{2}{N+1}} \sin\left(\frac{\pi(n+1)(k+1)}{N+1}\right), \quad k, n = 0, 1, \dots, N - 1.$$

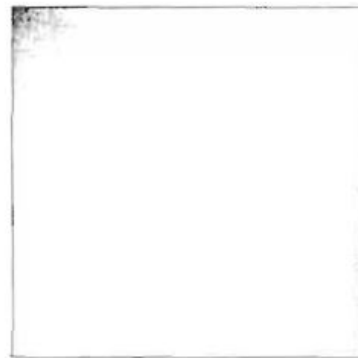
# Пример ДПФ, ДКП, ДСП преобразований



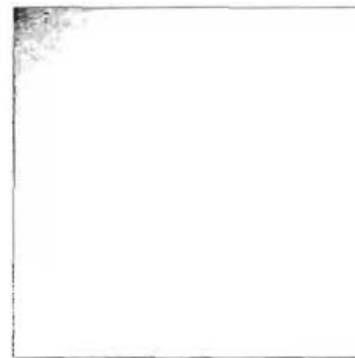
Исходное изображение



Коэффициенты ДПФ



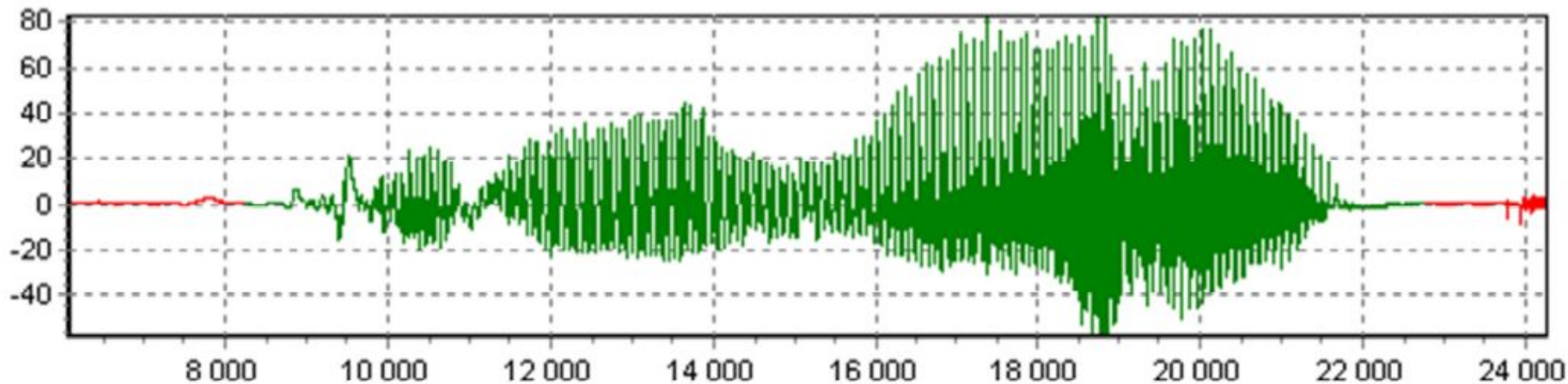
Коэффициенты ДСП



Коэффициенты ДКП

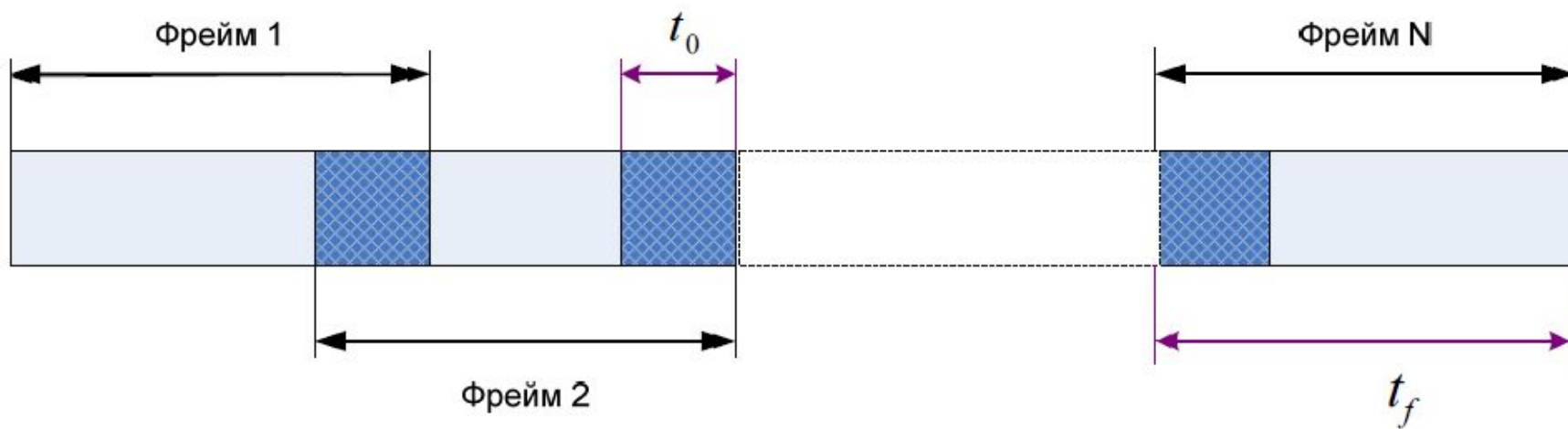
Яркость точек: чем больше модуль коэффициента, тем темнее точка

# Задача сравнения речевых команд



Звуковой сигнал с микрофона для слова  
«привет»

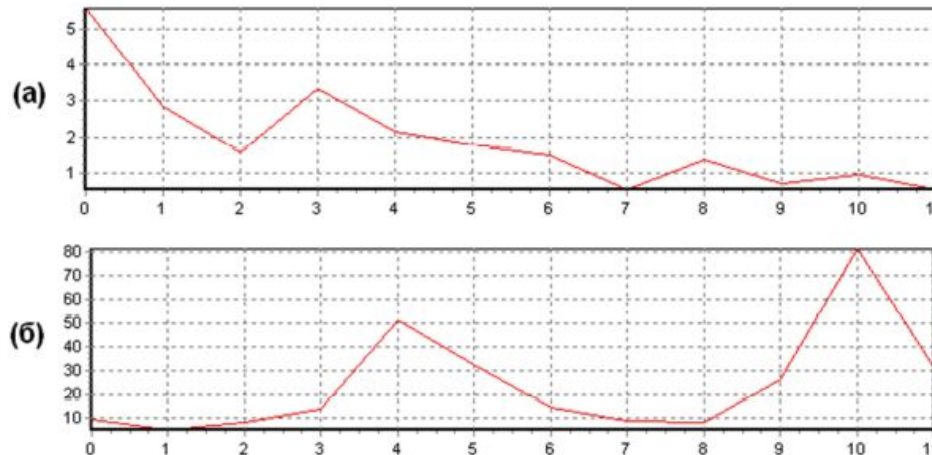
# Разделение сигнала на фреймы с перекрытием



# Спектральные коэффициенты

В качестве вектора признаков для отдельного фрейма выбираются коэффициенты ДПФ для сигнала  $x_k(n)$ ,  $n = 0, \dots, t_f - 1$  этого фрейма:

$$y_k(m) = \frac{1}{\sqrt{t_f}} \sum_{n=0}^{t_f-1} x_k(n) \exp(-i \frac{2\pi}{t_f} mn), \quad m = 0, \dots, t_f.$$



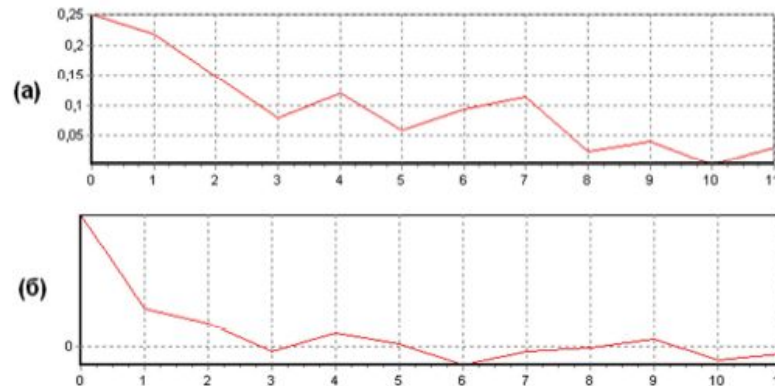
Модули спектральных коэффициентов для (а) 1-ого и (б) 30-ого фрейма

# Кепстральные коэффициенты

Второй вариант признакового описания фрейма – с помощью кепстральных коэффициентов, вычисляемых по формуле:

$$y_k(n) = \operatorname{Re} \left[ \frac{1}{t_f} \sum_{l=0}^{t_f-1} \exp(i \frac{2\pi l n}{t_f}) \ln \left\{ \left| \sum_{j=0}^{t_f-1} x_k(j) \exp(-i \frac{2\pi l j}{t_f}) \right| + 1 \right\} \right],$$

$$n = 0, \dots, t_f.$$



Кепстральные коэффициенты для (а) 1-ого и (б) 30-ого фрейма.

# Метрика в пространстве признаков

Расстояние между  $i$ -ым и  $j$ -ым фреймами, описываемых векторами признаков

$$\vec{y}_i = (y_i(1), \dots, y_i(12))$$

и

$$\vec{y}_j = (y_j(1), \dots, y_j(12))$$

соответственно, определяется как обычное Евклидово расстояние:

$$d(i, j) = \sum_{k=1}^{12} (y_i(k) - y_j(k))^2 .$$



# Мера различия речевых команд

$r_i, i = 0, \dots, p$  – фреймы в команде  $R = \{r_0, r_1, \dots, r_p\}$

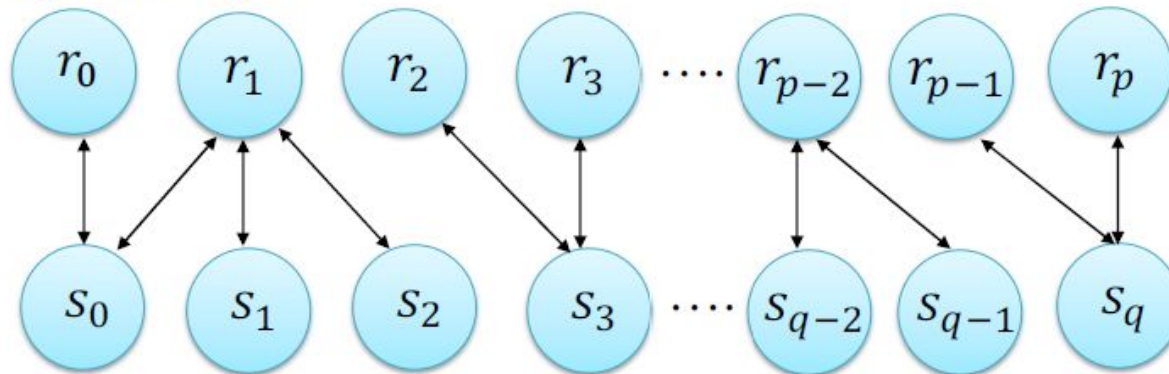
$s_j, j = 0, \dots, q$  – фреймы в команде  $S = \{s_0, s_1, \dots, s_q\}$

$D(R, S)$  – мера различия команд  $R$  и  $S$

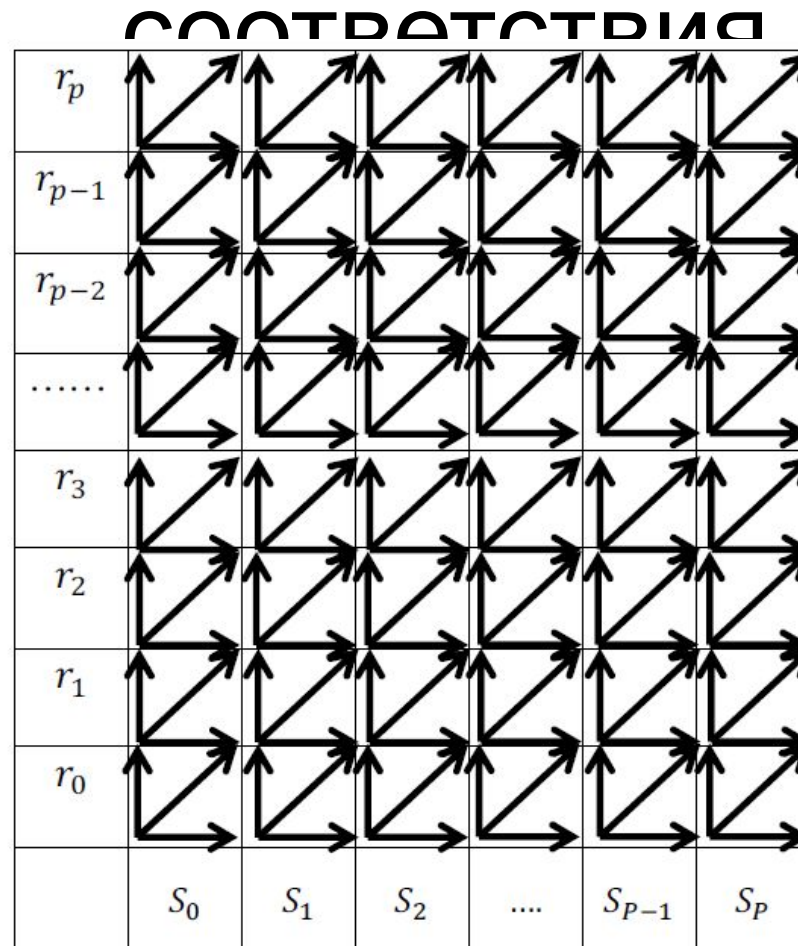
$$D(R, S) = \min_Z \sum_{(i,j) \in Z} d(r_i, s_j)$$

$Z$  – множество вариантов соответствия фреймов:

- каждому фрейму из  $R$  соответствует хотя бы один фрейм из  $S$
- каждому фрейму из  $S$  соответствует хотя бы один фрейм из  $R$
- соответствие монотонное: если есть пары  $(r_i, s_j)$  и  $(r_k, s_m)$  и  $k \geq i$ , то  $m \geq j$ .



# Выбор оптимального



Задача построения кратчайшего пути на плоском графе.

Решается алгоритмом Дейкстры