ОБРАБОТКА И РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ

Леонид Моисеевич Местецкий

профессор

кафедра математических методов прогнозирования ВМК МГУ

кафедра интеллектуальных систем МФТИ

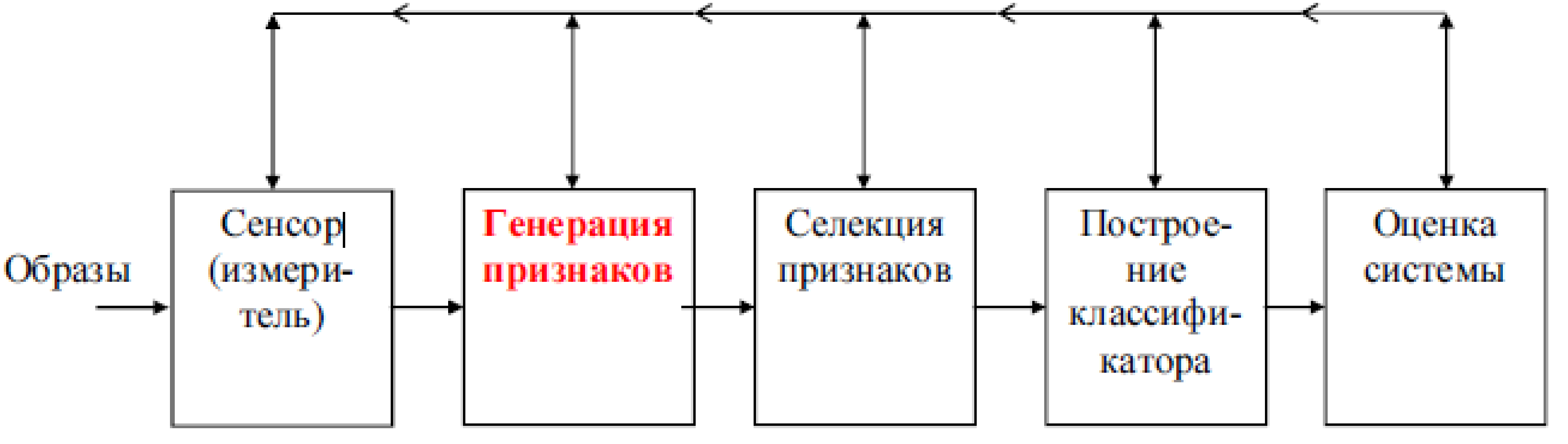
Генерация признаков на основе линейных преобразований

1. Генерация признаков осуществляется через преобразования исходных измерений образов (в частности, изображений).
2. Целью такой генерации признаков является сокращение информации до “значимой” путём преобразования исходного множества измерений в новое множество признаков.
3. Обычно задача состоит в выделении низкочастотных компонент, содержащих основную информацию.

Место генерации признаков в распознавания образов

Цель – генерация признаков через линейные преобразования исходных векторов измерений

Базовая идея – преобразовать данное множество измерений в новое множество признаков



Идея генерации признаков на основе линейных преобразований

1) Представить исходный образ в виде линейной комбинации базисных образов

2) Составить вектор признаков из коэффициентов разложения ($C_0, \ldots, C_{N-1}$)

$$\boxed{\text{Исходный образ}} = \sum C_i \times \boxed{i\text{-й базисный образ}}$$

Вектор измерений

– множество исходных измерений,

–

Эрмитов оператор

Матрицей, эрмитово сопряжённой к данной, называют матрицу , получаемую из исходной матрицы путем её транспонирования и перехода к комплексно сопряжённой, то есть .

Матрицу, равную своему эрмитовому сопряжению, называют эрмитовой, или самосопряжённой:

Унитарные матрицы

Для действительной матрицы условие унитарности обозначает, что матрица   ортогональная, т.е.

Для комплексной матрицы условие унитарности обозначает, что ,

где матрица   - транспонированная и сопряженная

Базисные вектора

$A = [a(0), a(1), \dots, a(N-1)]$ - 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 $a(i)$.

Разложение по базисным векторам

Вектор измерений x можно представить в виде линейной комбинации базисных векторов

$$x = (AA^{-1})x = (AA^{H})x = AA^{H}x = Ay = \sum_{i=0}^{N-1} y(i)a_i .$$

Поскольку $A \cdot A^H = E$ или $a_i^H \cdot a_j = (a_i, a_j) = \delta_{ij}$, получаем

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

Таким образом, в силу ортогональности векторов $a(i)$ между собой, $y(i)$ – это проекция вектора x на базисные вектора:

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

Цель разложения по базисным векторам

Мы хотим использовать подмножество компонент вектора y в качестве признакового описания исходного образа x.

Разложение по базисным матрицам

В задачах анализа изображений множество образов представляется двумерными массивами

определёнными, как матрицы, а не вектора.

Обработка матриц как векторов размерности не эффективна.

Альтернативная возможность – преобразовать матрицу через множество базисных матриц.

Базисные матрицы

Пусть U, V – унитарные $(N \times N)$ -матрицы.
u_i – вектора-столбцы матрицы U,
v_j – вектора-столбцы матрицы V.
Определим преобразование:

$$Y = U^H \cdot X \cdot V$$

и обратное

$$X = U \cdot Y \cdot V^H.$$

$$X = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} Y(i,j) \cdot u_i v_j^H$$

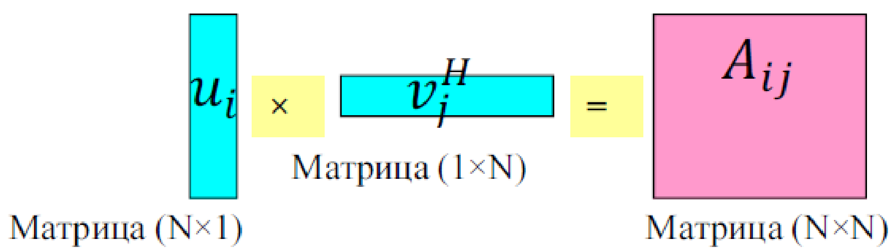
$$u_i v_j^H = \begin{bmatrix} u_{i0} \cdot v_{j0}^* & \cdots & u_{i0} \cdot v_{j\,N-1}^* \\ \vdots & \ddots & \vdots \\ u_{i\,N-1} \cdot v_{j0}^* & \cdots & u_{i\,N-1} \cdot v_{j\,N-1}^* \end{bmatrix} = A_{ij}$$

Представление через базисные матрицы

$$X = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} Y(i,j) \cdot u_i v_j^H = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} Y(i,j) \cdot A_{ij}$$

Всего получается N^2 базисных матриц.

Но если получится сделать такое разложение, чтобы матрица Y была диагональной, то количество базисных матриц сокращается до N.



Преобразование Карунена-Лоева

Пусть – вектор измерений образа. Рассматриваем его как вектор случайных величин.

Тогда полученный вектор признаков также является случайным.

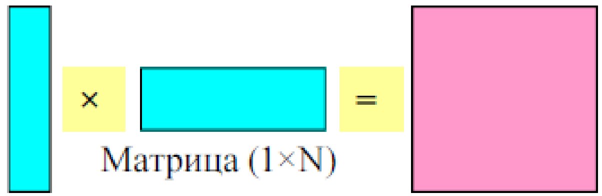
Целью преобразования является такой выбор матрицы , чтобы для вектора признаков имело место свойство: для всех , т.е. чтобы признаки были взаимно некоррелированные.

 – математическое ожидание.

Такое свойство признаков характеризует их «минимальность», поскольку нет «избыточности описания»: каждый признак не дублирует остальные признаки.

Корреляционная матрица вектора признаков

Обозначим $R_y = E[y \cdot y^T]$.



Матрица (1×N)

Матрица (N×1) Матрица (N×N)

$$R_y = \begin{pmatrix} E[y(0)^2] & E[y(0)y(1)] & \cdots & E[y(0)y(N-1)] \\ E[y(1)y(0)] & E[y(1)^2] & \cdots & E[y(1)y(N-1)] \\ \vdots & \vdots & \ddots & \vdots \\ E[y(N-1)y(0)] & E[y(N-1)y(1)] & \cdots & E[y(N-1)^2] \end{pmatrix}$$

По главной диагонали стоят дисперсии, а остальные элементы – парные корреляции компонент вектора признаков.

Корреляционная матрица вектора измерений

Поскольку $y = A^T x$ и, следовательно, $y^T = x^T A$ получаем:

$$R_y = E[y \cdot y^T] = E[A^T x \cdot x^T A] = A^T E[x \cdot x^T]A = A^T R_x A,$$

где

$R_x = E[x \cdot x^T]$ - корреляционная матрица вектора измерений x.

Теперь задача состоит в том, чтобы выбрать такую матрицу A, при которой матрица R_y будет диагональной.

Матрица преобразования

Выберем в качестве базисных векторов a_i собственные вектора корреляционной матрицы R_x.

1) Матрица R_x положительно полуопределённая, её собственные значения $\lambda_i \geq 0$.
2) Матрица R_x симметрическая, её собственные вектора ортогональны между собой.
3) Матрица $A = [a(0), a(1), \ldots, a(N-1)]$, составленная из собственных векторов матрицы R_x, приводит её к диагональной:

$$A^T R_x A = \Lambda = \begin{pmatrix} \lambda_0 & 0 & \vdots & 0 \\ 0 & \lambda_1 & \vdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ 0 & 0 & \vdots & \lambda_{N-1} \end{pmatrix}.$$

Преобразование Карунена-Лоева

Описанное преобразование называется преобразованием Карунена-Лоева. Оно имеет фундаментальное значение, т.к. оно приводит к построению некоррелированных признаков.

Другое название:

Метод главных компонент (Пирсон, 1901 г.)

Свойства преобразования Карунена-Лоева

Разложение по базисным векторам:

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

Определим новый m-мерный вектор ($m < N$):

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

где $\hat{x}$ – проекция вектора x на подпространство, соответствующее первым m компонентам этого вектора.

Вектор $\hat{x}$ можно рассматривать, как аппроксимацию вектора x с помощью $\hat{x}$.

Ошибка аппроксимации

$$E\left\|x-\hat{x}\right\|^2 = E\left[\left\|\sum_{i=m}^{N-1} y(i)a_i\right\|^2\right] = E\left[\sum_i \sum_j \left(y(i)a_i^T\right)(y(j)a_i)\right] =$$

$$= \sum_{i=m}^{N-1} E\left[y^2(i)\right] = \sum_{i=m}^{N-1} a_i^T E\left[xx^T\right]a_i = \sum_{i=m}^{N-1} a_i^T \lambda_i a_i = \sum_{i=m}^{N-1} \lambda_i .$$

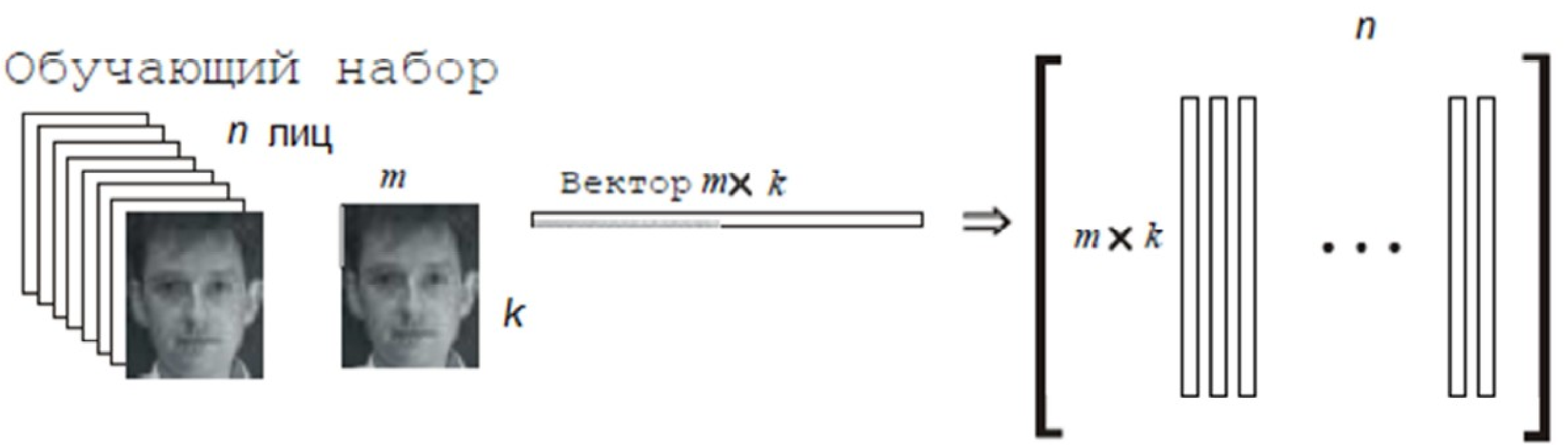
Из этого следует, что выбирать нужно m базисных векторов с максимальными собственными значениями.

В преобразовании Карунера-Лоева в качестве критерия выступает наилучшее приближение исходных измерений.

Задача распознавания лиц

Метод «Собственных лиц» (Eigenfaces) – основан на преобразовании Карунена-Лоева (методе главных компонент)

Преобразование обучающего набора лиц в одну общую матрицу



Вычисление «собственных лиц»

Дано: M изображений лиц размером $h \times w$.

1. Каждое изображение преобразуется в вектор размера $D = hw$ и помещается в множество векторов $\{\Gamma_1, \Gamma_2, \cdots, \Gamma_M\}$. Лица приводятся к одинаковому масштабу, фон (волосы, шея) должен быть одинаков или удалён.
2. Вычисляется «среднее лицо»: $\Psi = \frac{1}{M}\sum_{i=1}^{M}\Gamma_i$ и отличия от среднего $\Phi_i = \Gamma_i - \Psi$.
3. Вычисляется матрица ковариации $C = \frac{1}{M}\sum_{i=1}^{M}\Phi_i\Phi_i^T = AA^T$, где $A = \{\Phi_1, \Phi_2, \cdots, \Phi_M\} \in R^{D \times M}$.
4. Вычисляются собственные вектора матрицы C на основе матрицы $A^T A$ размерности $M \times M$.

Собственные лица


Среднее лицо

Два первых собственных лица

Три последних собственных лица

Вывод

Целью метода главных компонент является значительное уменьшение размерности пространства признаков таким образом, чтобы оно как можно лучше описывало «типичные» образы, принадлежащие множеству лиц. Используя этот метод можно выявить пределы изменчивости в обучающей выборке изображений лиц и описать эту изменчивость в базисе нескольких ортогональных векторов, которые называются собственными (eigenface).