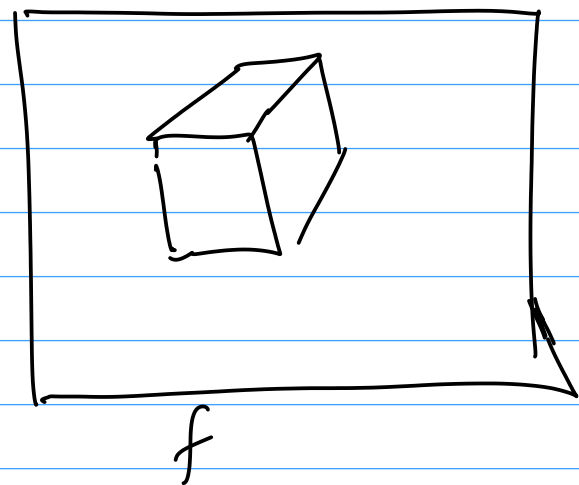


Построение формы логичного (кусочно-постоянного) изображения.

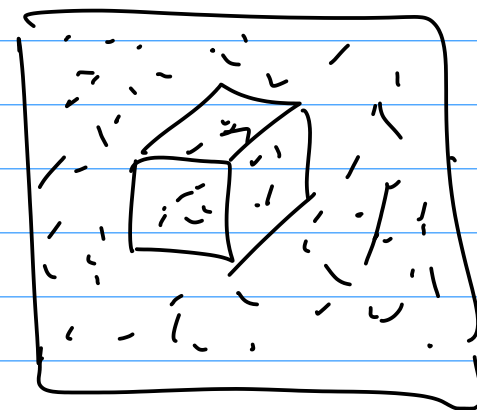


$$\vec{V}_f = \left\{ \sum_{i=1}^N c_i \chi_i, c_i \in \mathbb{R} \right\}$$

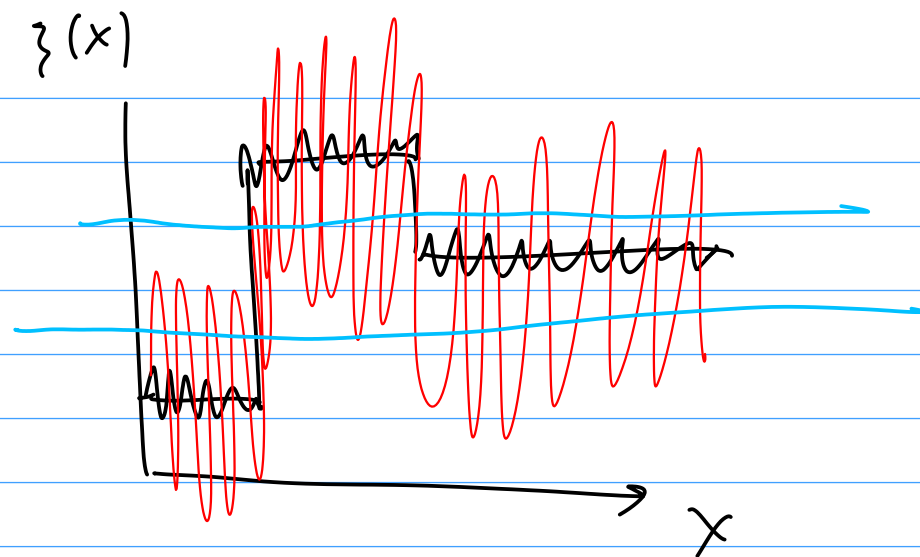
χ_i — индикаторная ф-ция i -ой области
 постоянной яркости изобр. f .

$A_i, i = 1, \dots, N$.

$$\chi_i(x) = \begin{cases} 1, & x \in A_i \\ 0, & x \notin A_i \end{cases}$$



$\vec{r} = f + \vec{v}$ — сум. изображение — шум.



Сегментация изображений — разбиение
 на время $\bar{X}$ на области, в
 пределах которых изобр. "однородно"
 в каком-либо смысле.

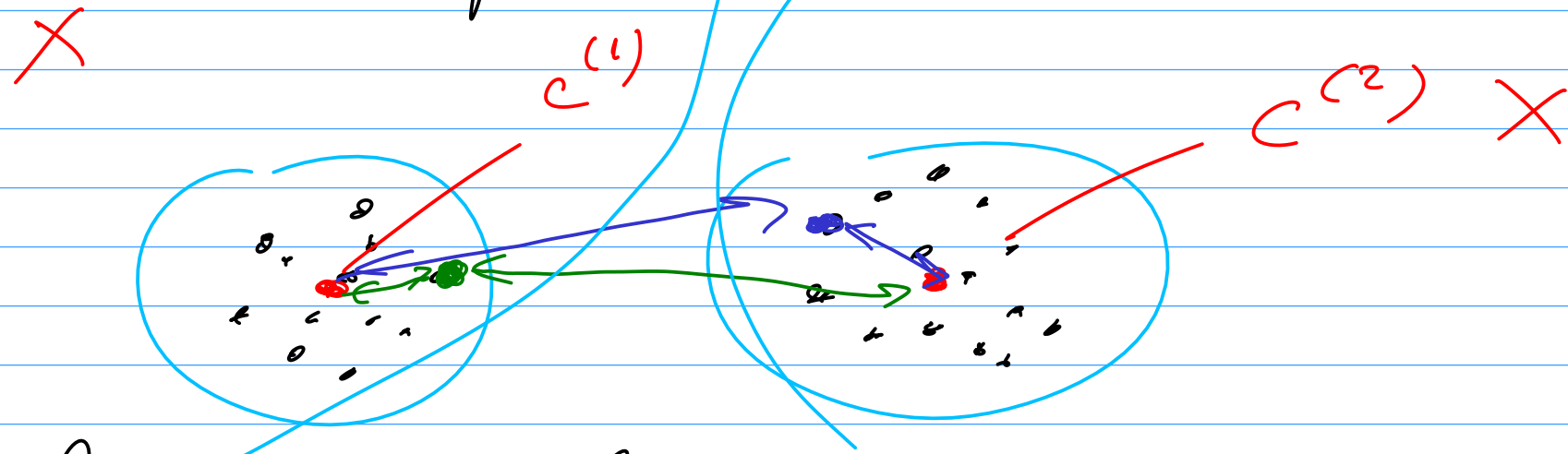
Сегментация изображений — это кластеризация пикселей.

Один из основных
 методов кластеризации —
 метод k -средних
 (k -means)

Разбить множество всех пикселей
 на "однородные" группы. — кластеры.
 по признаку в центре
 группы.

$y_1, \dots, y_m \in \mathbb{R}^d$ — вилков гр-во поглернои 2.

N — заранее заданное кол-во кластеров.



Предлагается выбрать N "центров" кластеров и разбить $\{y_1, \dots, y_m\}$ на кластеры следующим образом:

$$i\text{-ый кластер: } \{y_j, j=1, \dots, m : \|y_j - c^{(i)}\|^2 = \min_{k=1, \dots, N} \|y_j - c^{(k)}\|^2\} = C_i$$

Должен быть способ найти оптимальные
 центры кластеров $c^{(1)}, \dots, c^{(N)}$

$$\sum_{i=1}^N \frac{1}{|C_i|} \sum_{y \in C_i} \|y - c^{(i)}\|^2 \sim \min_{c^{(i)}, i=1, \dots, N}$$

используется "разброс" кластера вокруг центра.

$|C_i|$ — кол-во элементов в кластере C_i .

Кластеры зависят от центров $c^{(i)}$.
 Поэтому это не МНК.
 Задача не решается аналитически.

Убеден и проанован в SciPy, Matplotlib, ...
Функциональный метод решения задач графов:

① Инициализация: задать количество центров $c^{(1)}, \dots, c^{(N)}$
(пример — случайные)

② построить кластеры на основе выбранных центров.

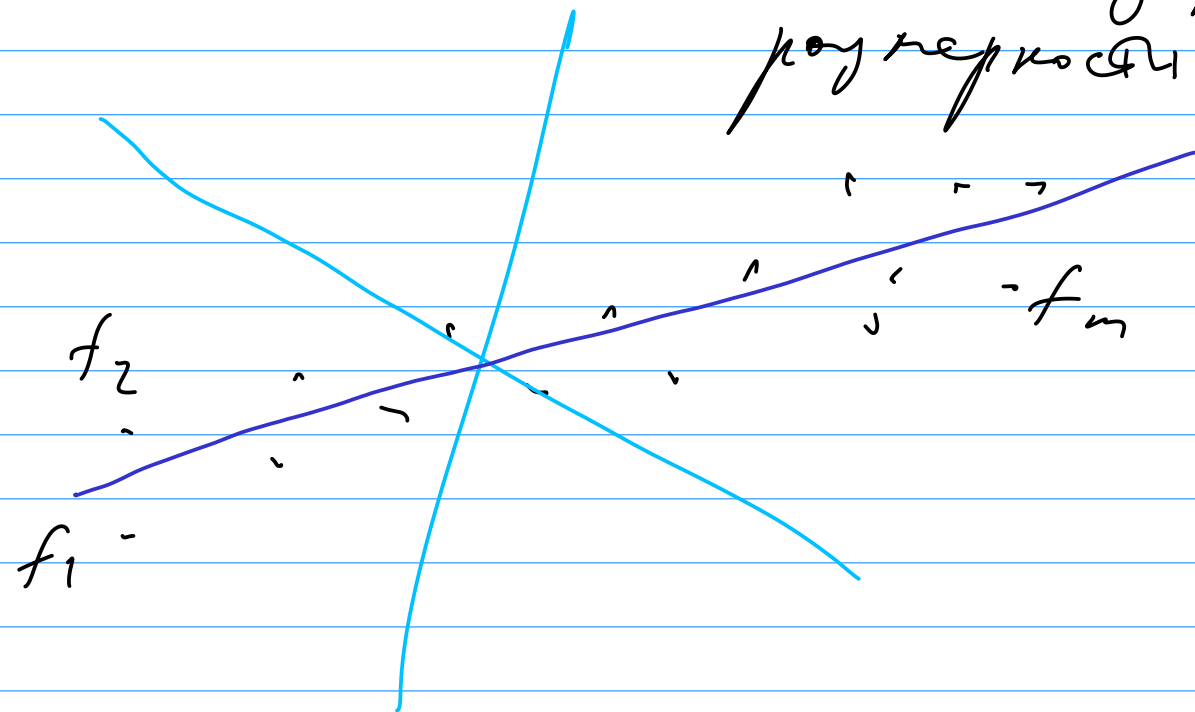
③ найти новые центры как средние по сформированным кластерам.

Построение линейного пространства по множеству
изображений,

Не кусочно-лос. изобр.

Форма изобр. канберова объекта. — лин. простран. известной
размерности.

$f_1, \dots, f_m$



Пусть $a_1, \dots, a_n$ — ОНБ пространства изображений.
 Будет искажа подпрост. размерности τ век
 $\mathcal{L}(a_1, \dots, a_\tau)$

Задача:

$$\sum_{i=1}^m \left\| f_i - \underset{\mathcal{L}(a_1, \dots, a_\tau)}{P} f_i \right\|^2 \sim \min_{\substack{a_1, \dots, a_n \text{ — ОНБ} \\ n}} \quad \uparrow \tau$$

Для широкого применения метод главных компонент $\tau < n$
 (PCA — Principal Component Analysis)

Неробусимо от τ (для $\forall \tau$) решение (т.е.
 базис $a_1, \dots, a_n$) одно и то же:

$a_1, \dots, a_n$ — собственный базис оператора $\mathcal{D}$;

$$\mathcal{D}g = \sum_{i=1}^m f_i(f_i, g)$$

Упорядочен по
 невозрастанию соб. значений.

Он же сигнатурный базис

матрица

$$F = \left(\begin{array}{c|c|c} \boxed{f_1} & \boxed{f_2} & \dots & \boxed{f_m} \end{array} \right)$$

Каждый столбец — это
 «вынесенное в столбец»
 изображение f_i

ЗАДАЧА

③ Разделение слов по форме:

— "глобальный" метод (морф. разность)

— "локальный" метод (корня морф. разности фразеологизмов, метод скользящего окна).

Орфографический словарь.

Анток

Орфограмм.

④ Морф. фильтрация сущ. аддитивного шума на
кусочно — постоянных изображениях.

Сравнить с методами скользящего среднего и медианной
фильтрации.

Рекат Газарин

⑤ Реализовать разложение убр. g на слагаемые

$$g = g_{1,0} + g_{2,0} + g_0 + g_2$$

Вадим Мухомов

(косяк преобразования)

София Киселева

Продемонстрировать на реальных изображениях.