МОСКОВСКИЙ АВИАЦИОННЫЙ ИНСТИТУТ
(национальный исследовательский университет)

Элементы поиска нечетких дубликатов видео



Никитин Илья Константинович,
асп. каф. 806 МАИ

twitter: @w_495
почта: w@w-495.ru
nikitin.i@tvzavr.ru
w495@yandex-team.ru

Кто я

аспирант, преподаватель



разработчик



разработчик

Яндекс

Что такое «нечеткие дубликаты»



оригинал

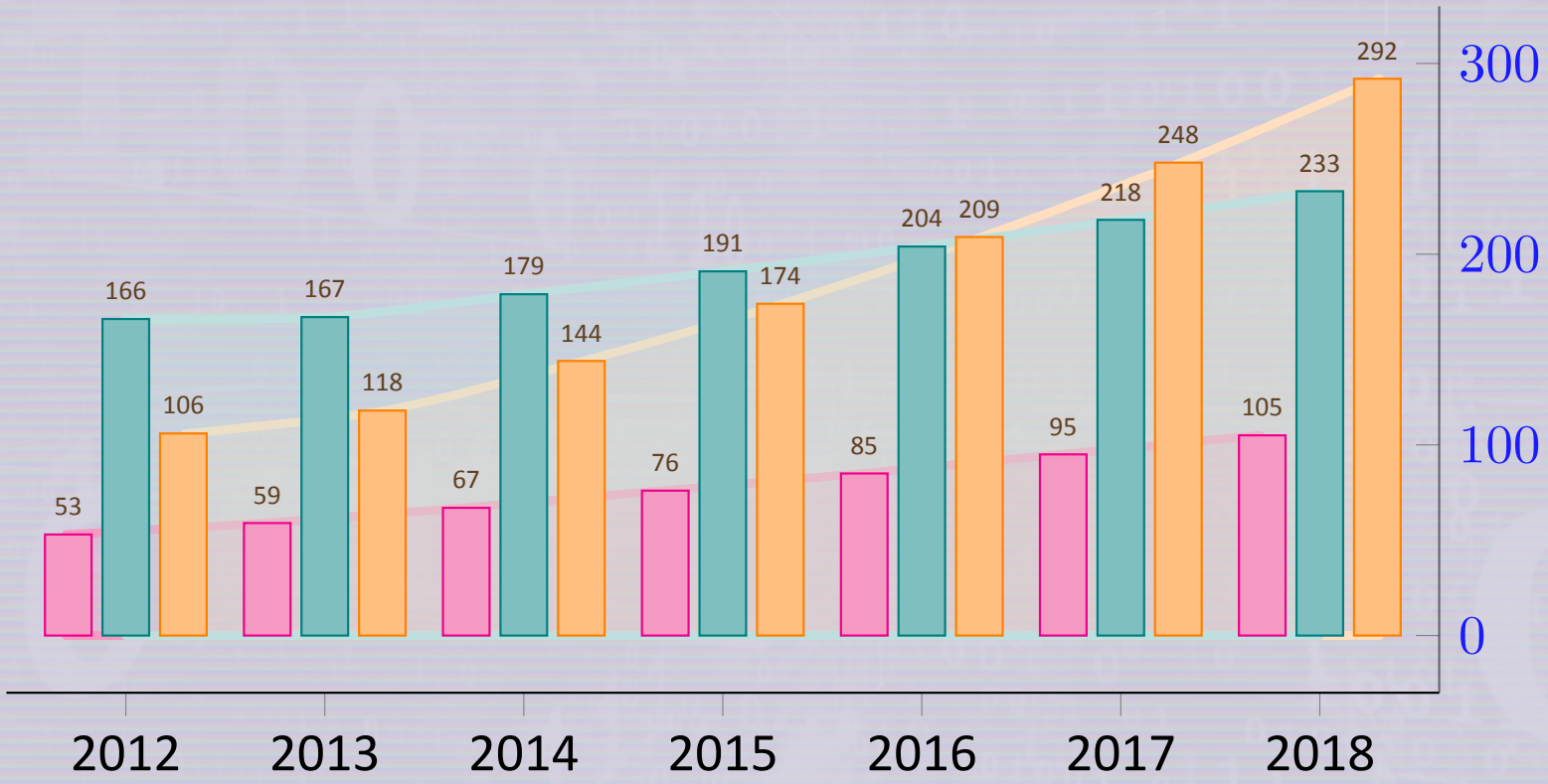


естественный дубликат

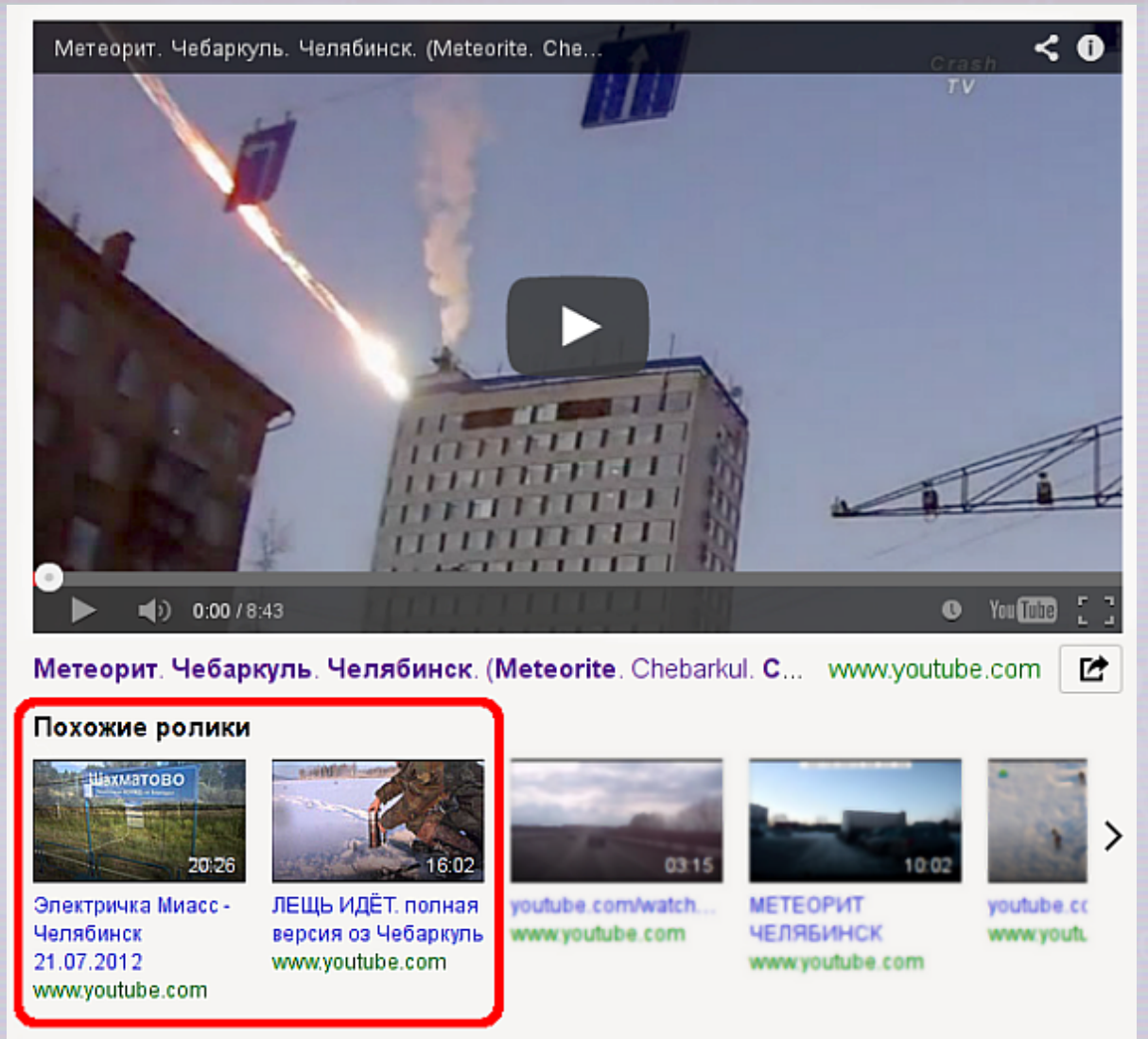


искусственный дубликат

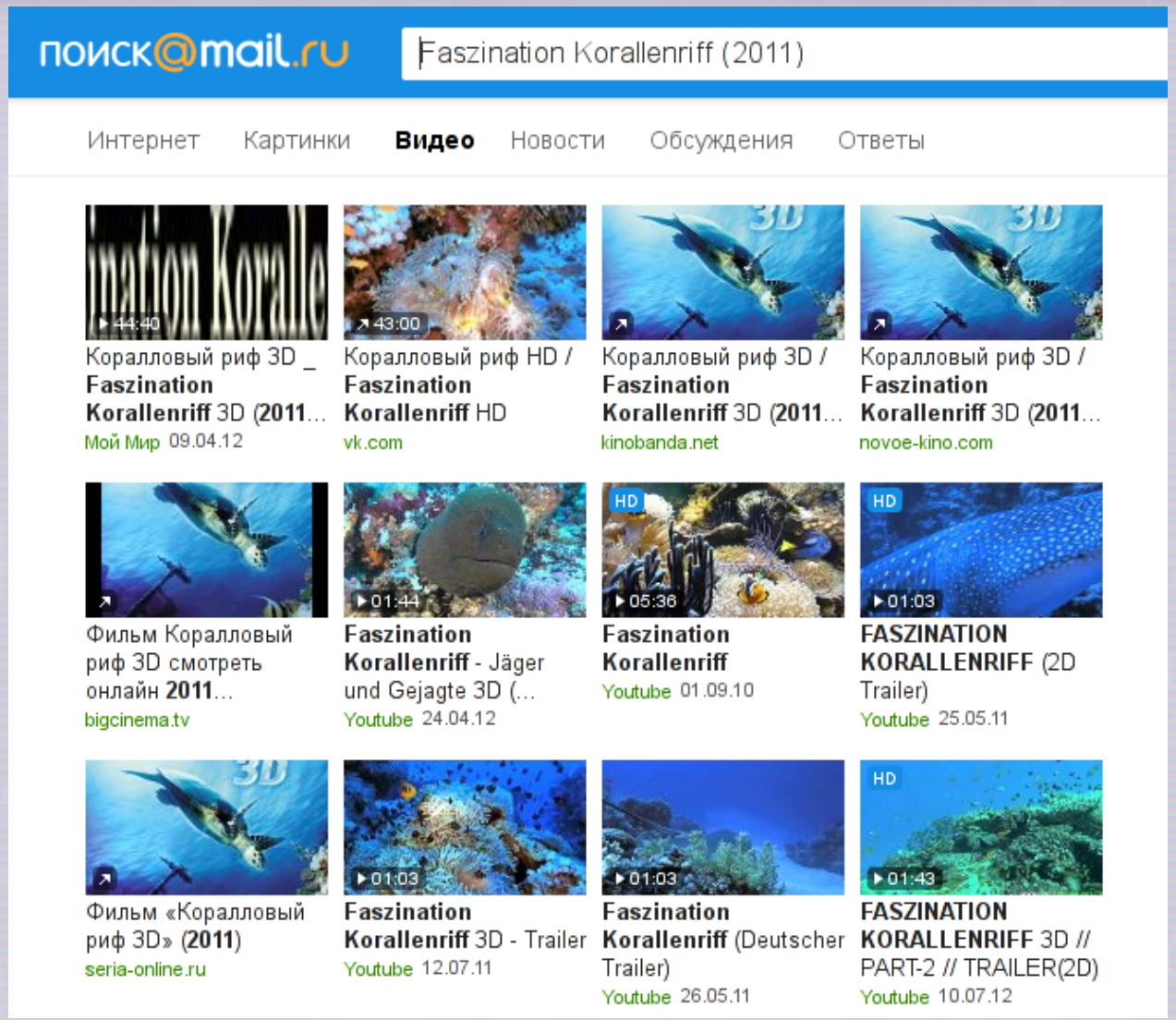
Российский рынок онлайн-видео в 2012—2018 годах



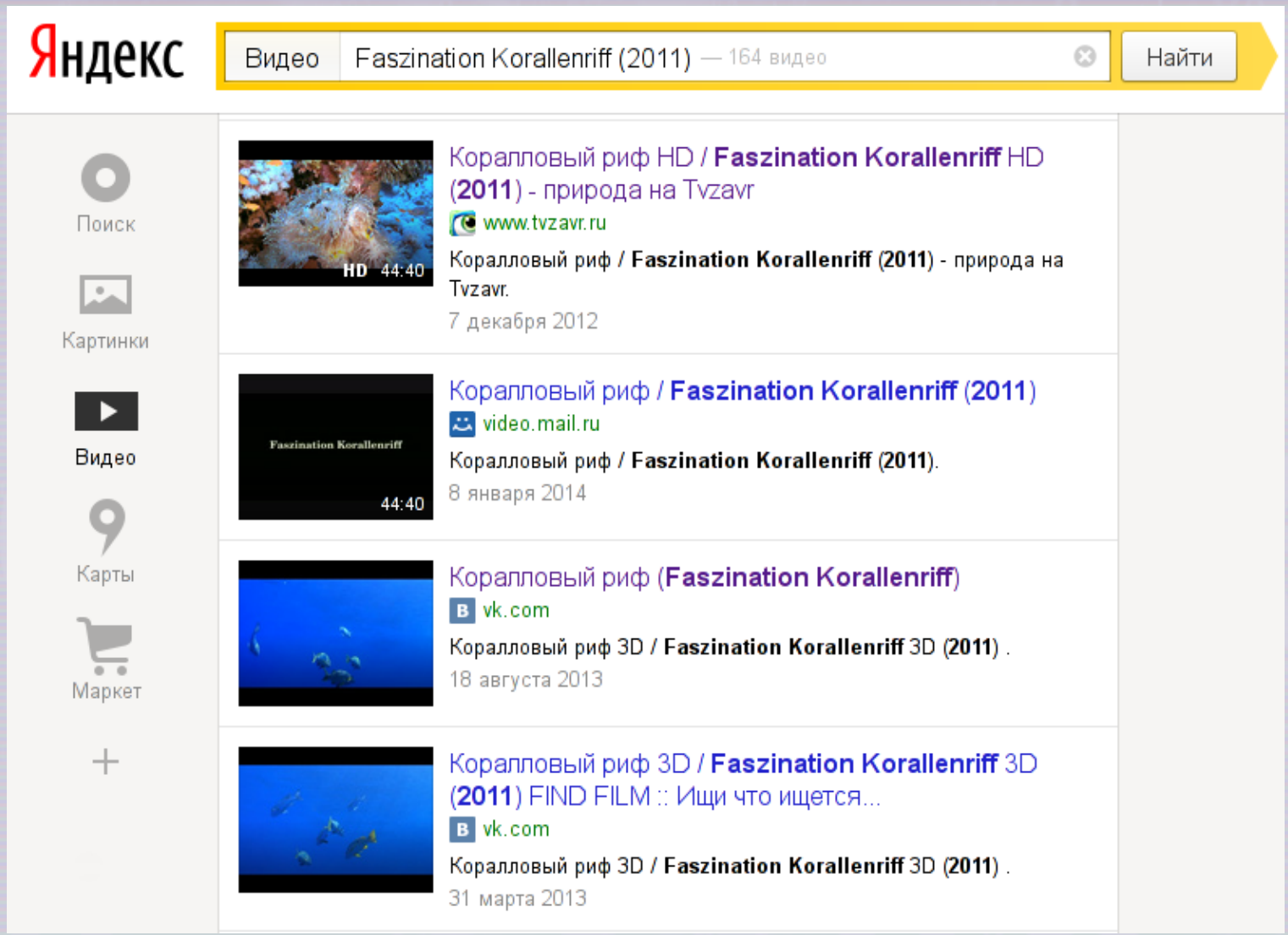
Пример ошибочной рекомендации видео



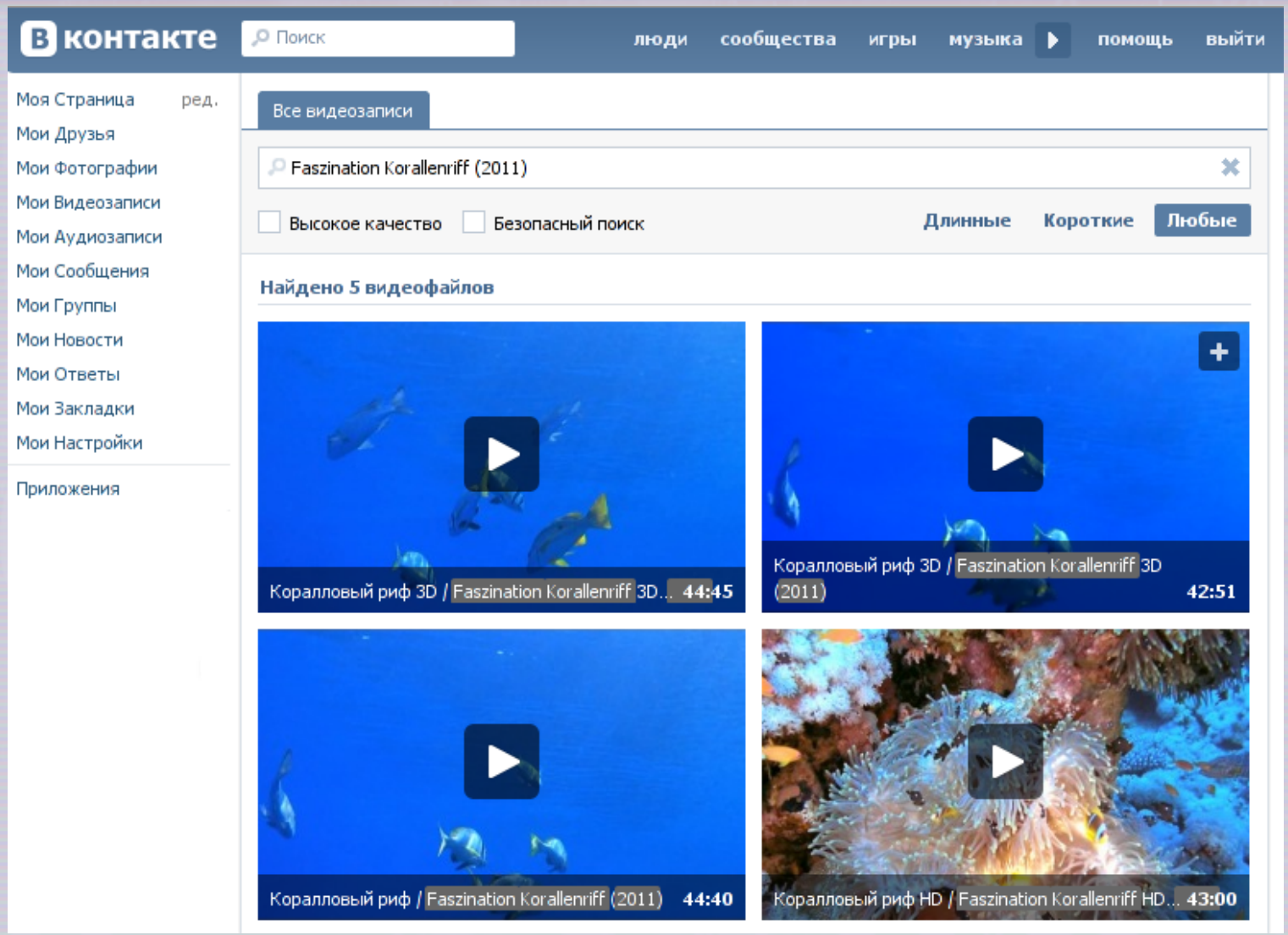
«Faszination Korallenriff (2011)». Сервис Mail.ru



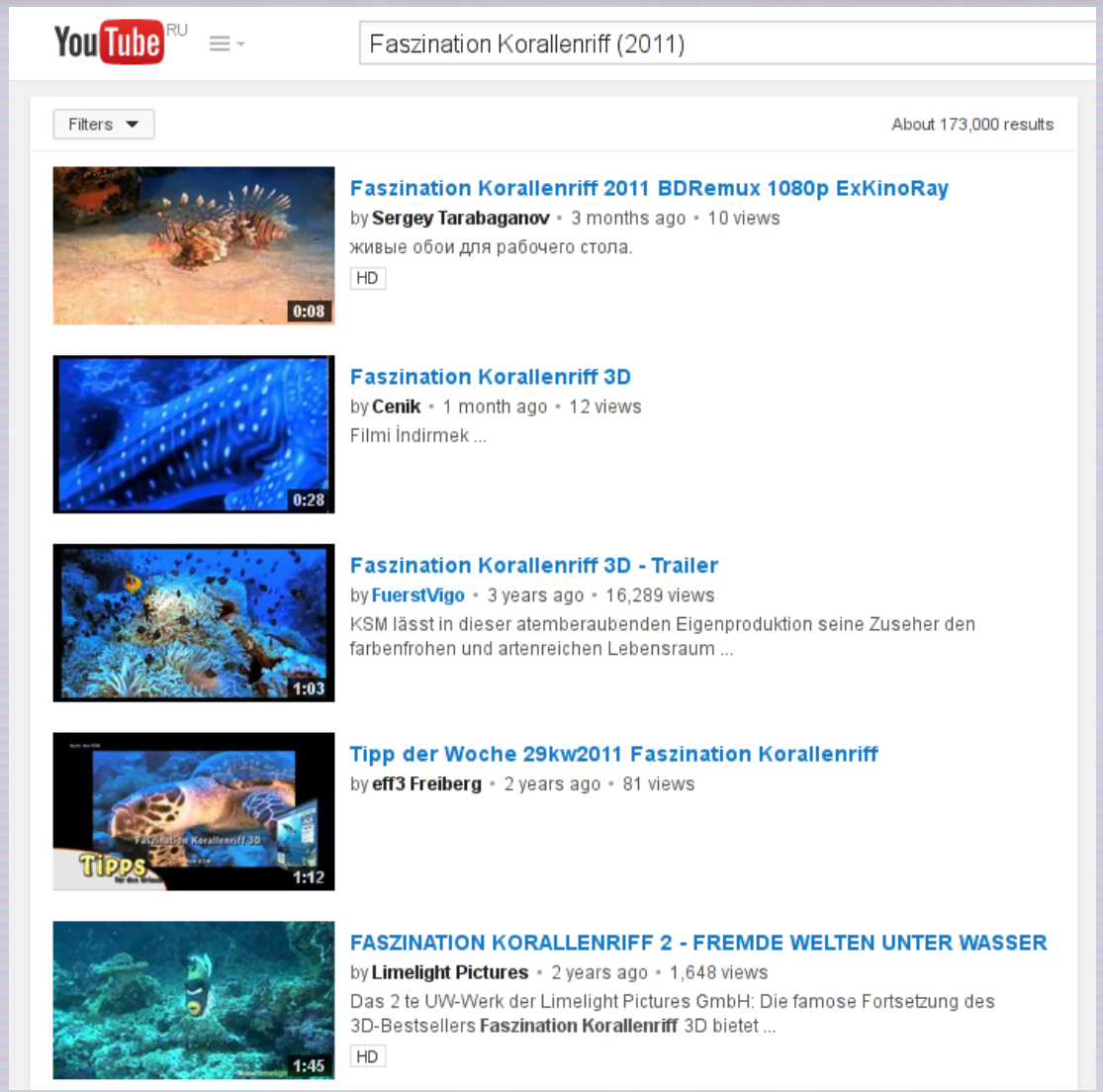
«Faszination Korallenriff (2011)». Сервис Яндекс.Видео



«Faszination Korallenriff (2011)». Сервис vk.com



«Faszination Korallenriff (2011)». Сервис YouTube.



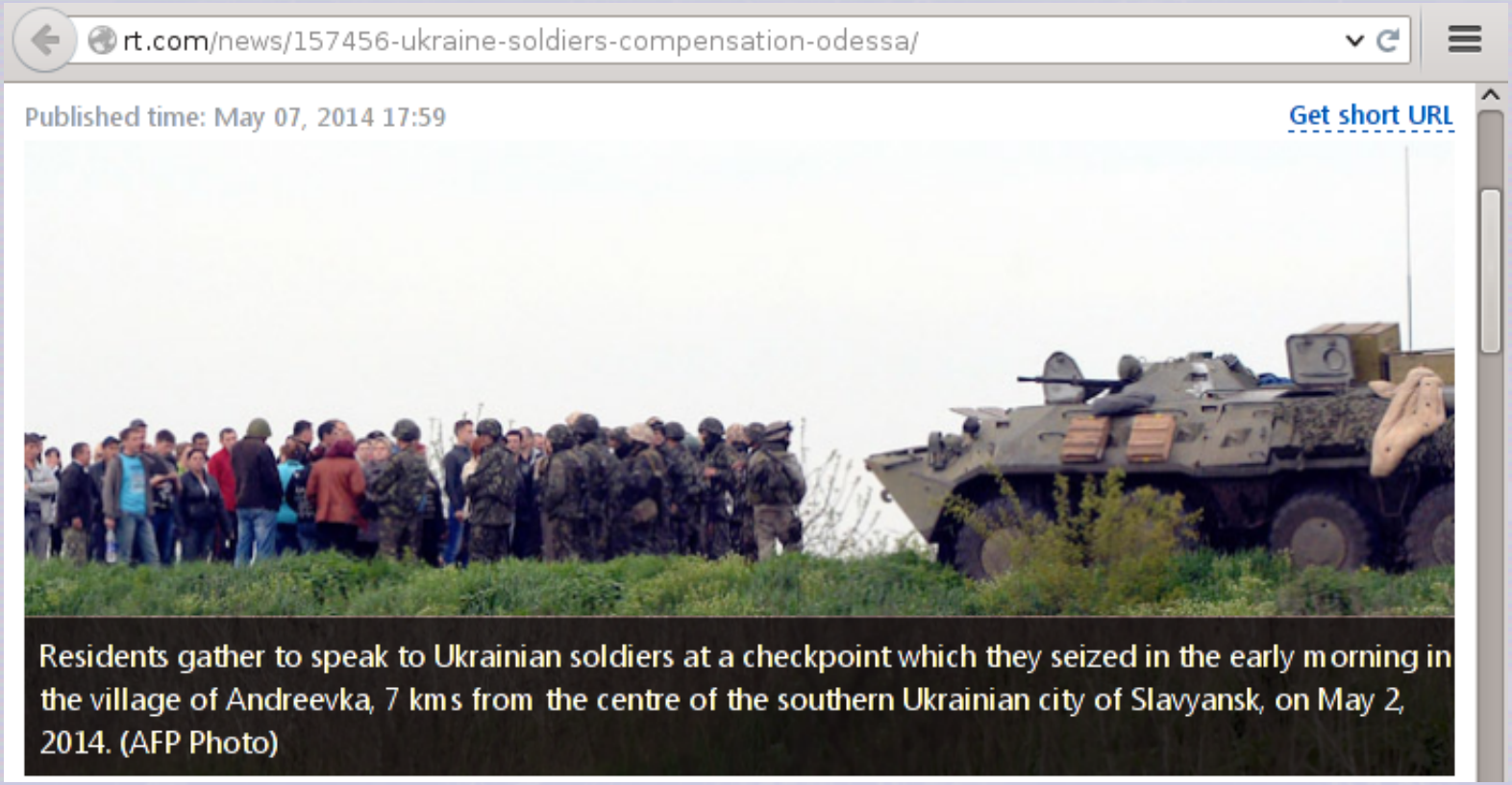
Группировка новостей. Пример с eyeopening.info



Перевод подписи

Безоружные мирные граждане встали на пути бронетехники и пытаются остановить украинских военных при входе в Славянск.

Группировка новостей. Пример с rt.com



Перевод подписи

Местные жители собираются, чтобы поговорить с украинскими военными на блокпосту, который те захватили утром в селе Андреевка, 7 км от центра Славянска, 2 мая 2014 года.

Группировка новостей. Пример с ewn.co.za



Перевод подписи

Во вторник украинские вооруженные силы нанесли удар по базам пророссийских сепаратистов в восточных регионах страны.

Нечеткие дубликаты видео по By

Копии друг друга, отличаются набором модификаций



$\simeq$ дубликат



$\simeq$ дубликат

$\not\simeq$ уже не дубликат





Нечеткие дубликаты видео по Шену

Общая сцена, разные условия съемки и варианты монтажа



$\simeq$ дубликат



$\simeq$ дубликат

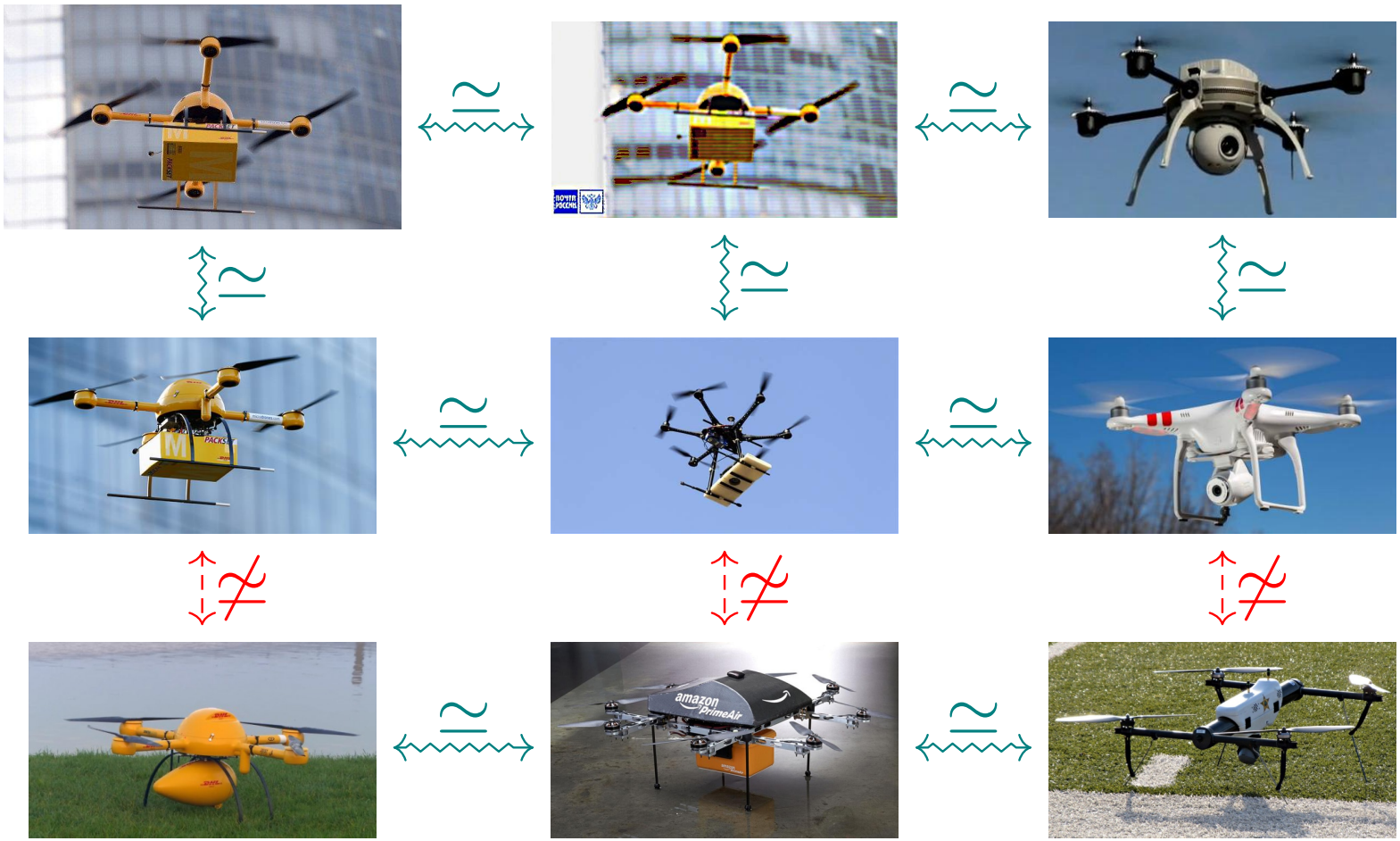
$\not\simeq$ уже не дубликат





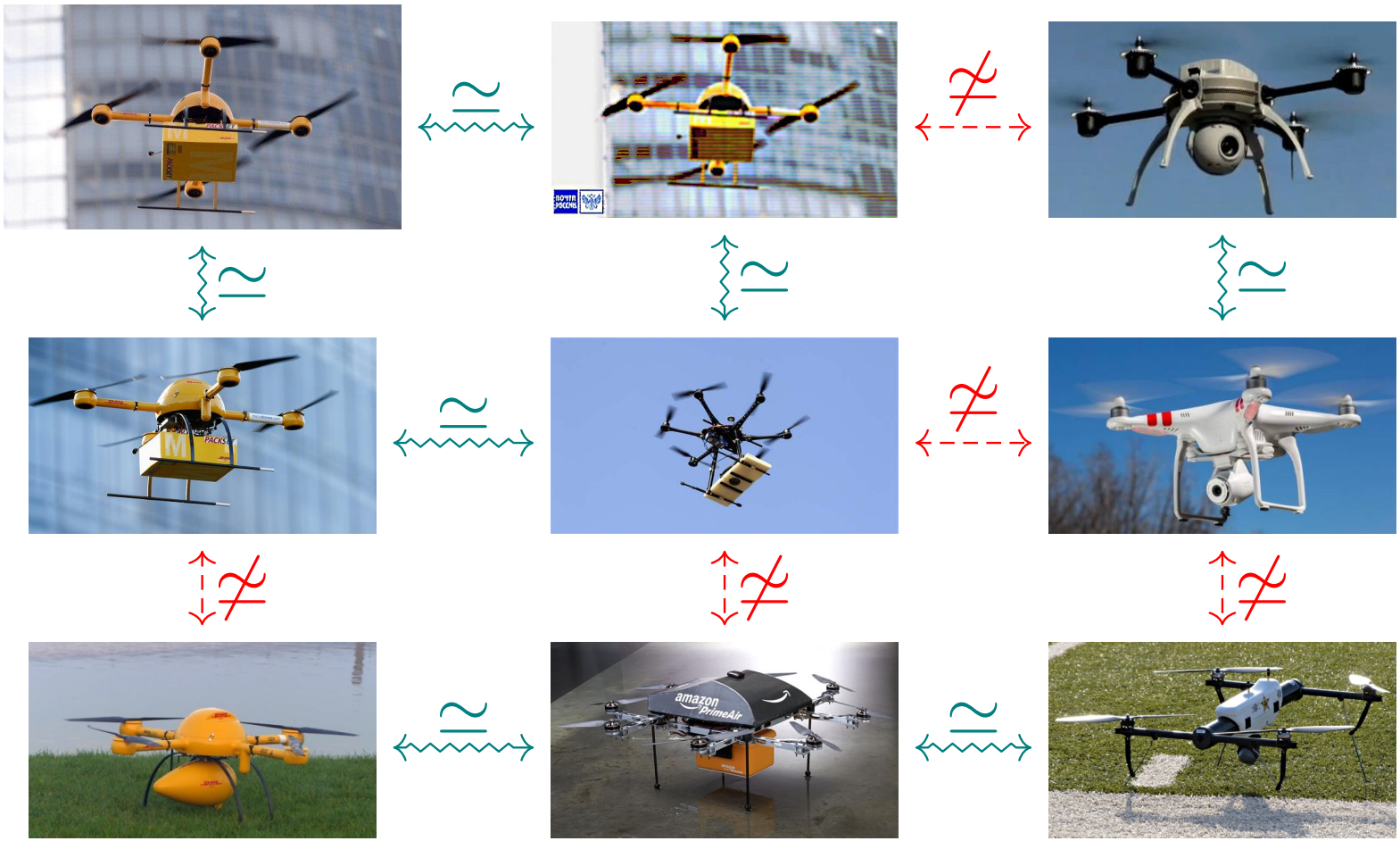
Нечеткие дубликаты видео по Башарату

Общая смысловая концепция, разные варианты представления

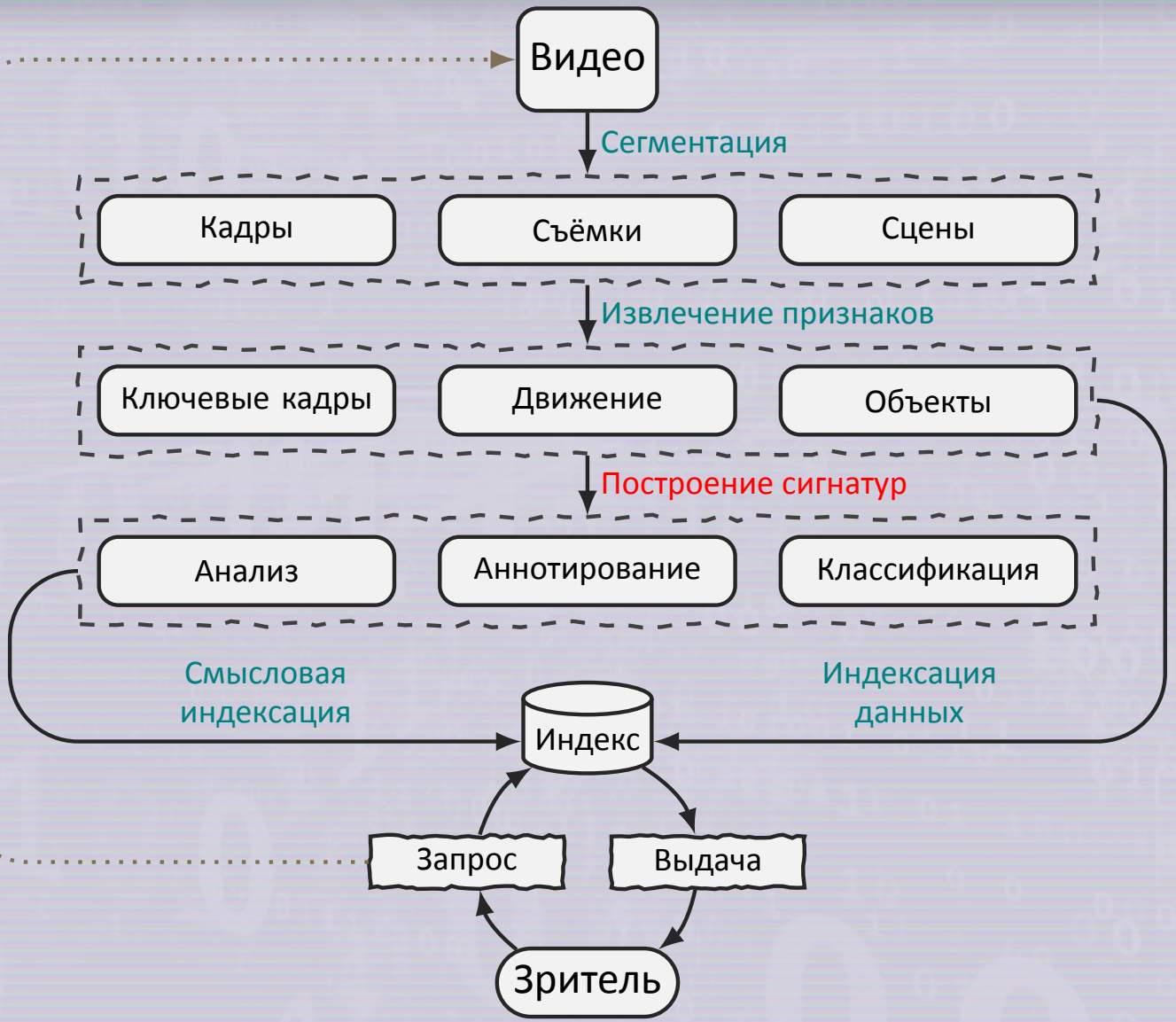


Нечеткие дубликаты видео по Керубини

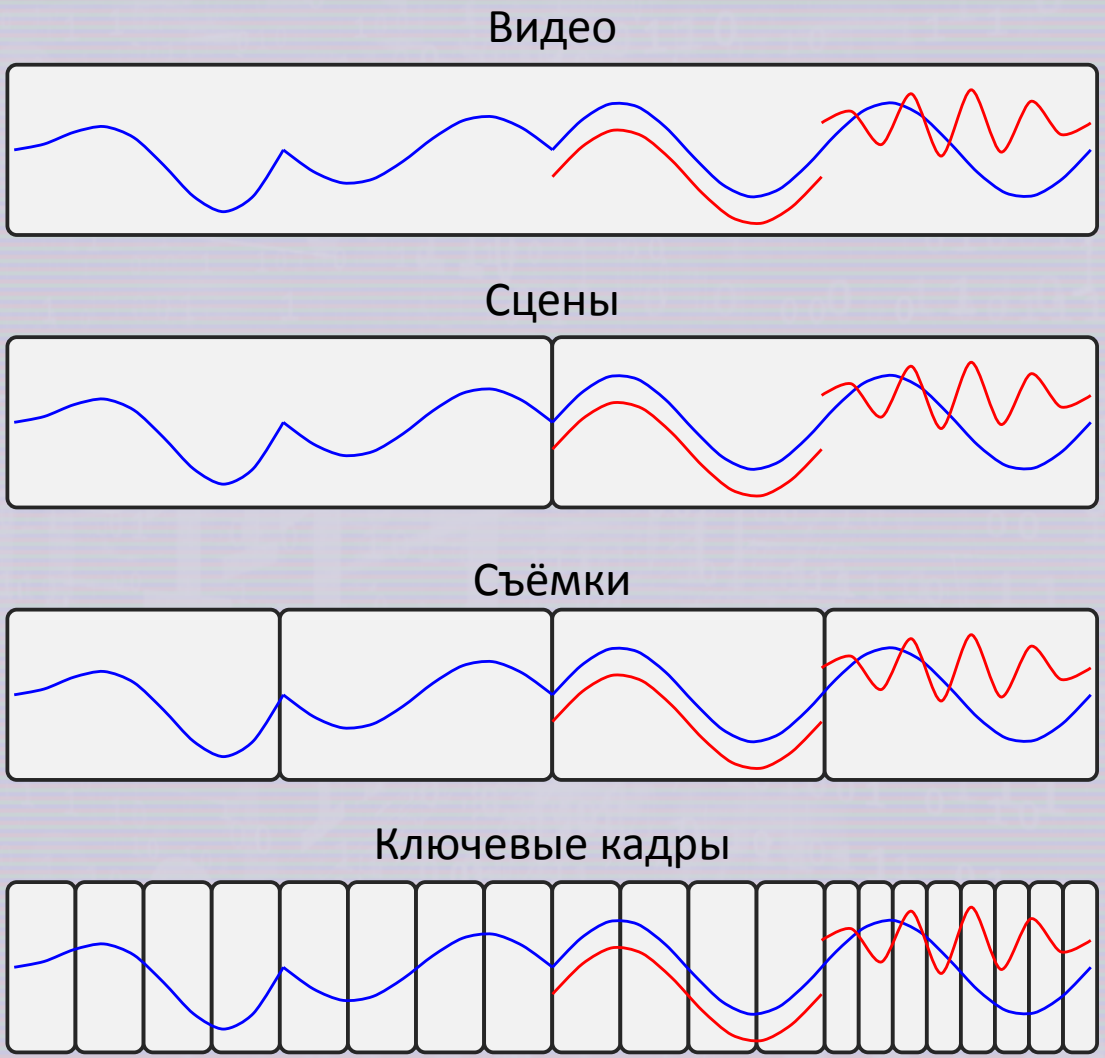
Общая семантика и нет дополнительной информации



Общая схема поиска по видео



Сегментация видео




Признаки
Звук
Шумы
Средняя частота
Музыка
Речь
Кадры
Края
Контуры
Цвет
Текстура
Движение
Камеры
Объектов
Объекты
Цвет
Размер
Текстура

Типы видео-сигнатур

Локальные

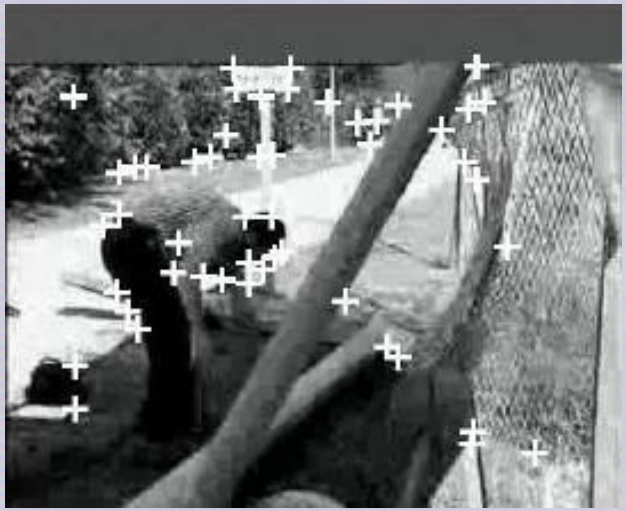
- выражают локальную информацию кадра,
- напрямую работают с дескрипторами особых точек.

Глобальные

- выражают одной сигнатурой целый кадр;
- выражают одной сигнатурой все видео целиком;
- выражают изменении кадров во времени.

Локальные сигнатуры

Видео кандидат



Плюсами показаны особые точки (запрос)

Оригинальное видео



Прямоугольниками показаны области, где могут находиться особые точки

Локальные сигнатуры: оценка

Пример

Пусть дано некоторое видео:

- разрешение 800×600 пикселей;
- в кадре может быть найдено около 1000 точек;

Если таких видео много:

- 100000 видео-файлов;
- по 5 ключевых кадров в каждом виде;

$\Rightarrow$ 500000000 ключевых точек;

Каждая точка представлена вектором:

- SIFT (128 компонент)
- или PCA-SIFT (64 компоненты);

Глобальные сигнатуры

- одна сигнатура — целый кадр:
  - средние цветовые моменты кадра;
  - цветовые моменты областей кадра (дескриптор GIST);
  - мешок визуальных слов;
  - линеаризация, видео как ДНК;
- одна сигнатура — видео целиком:
  - средний цветовой момент видео;
  - ссылочная видео-гистограмма — набор опорных кадров;
  - кластеризация кадров.
- пространственно-временные сигнатуры:
  - отслеживание траекторий;
  - линеаризация, порядковая сигнатура;
  - последовательность монтажных склеек;
  - дерево сцен.

Глобальные сигнатуры: видео как ДНК



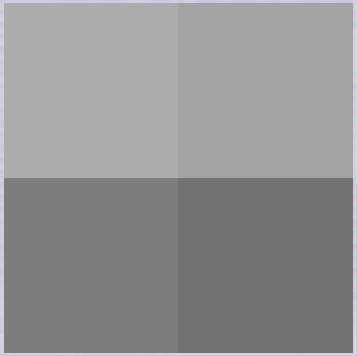
Исходное (RGB)

$\Rightarrow$



Яркость

$\Rightarrow$



2×2

$\Rightarrow$

$\Rightarrow$ $\begin{pmatrix} 172 & 164 \\ 125 & 114 \end{pmatrix}$

Матрица яркости

$\Rightarrow$ $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$

Порядок блоков

$\Rightarrow$

$\mathbf{1,2,3,4} \Rightarrow \boldsymbol{A}$

$\ldots$

$4,3,1,2 \Rightarrow W$

$4,3,2,1 \Rightarrow X$

Символ

Глобальные: порядковая, сравнение по времени



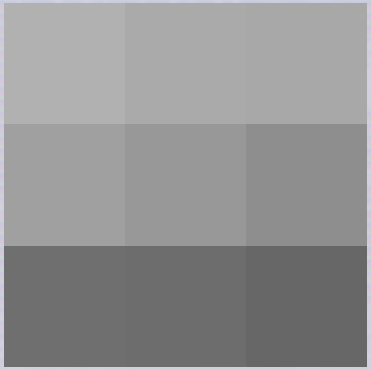
Исходное (RGB)

$\Rightarrow$



Яркость

$\Rightarrow$



3×3

$\Rightarrow$

$\Rightarrow$

$$\begin{pmatrix} 177 & 170 & 168 \\ 160 & 152 & 142 \\ 111 & 109 & 103 \end{pmatrix}$$

Матрица яркости

$\Rightarrow$

$$\begin{pmatrix} 8 & 7 & 6 \\ 5 & 4 & 3 \\ 2 & 1 & 0 \end{pmatrix}$$

Порядок блоков

Что такое видео?



Что такое видео? — Последовательность ...

Видео — последовательность фактов, развивающихся во времени.

- Свойства фактов — *пространственная* характеристика видео,
- продолжительность и порядок фактов — *временная*.

Что есть факты?

- простейший вариант — сцены-съёмки:
  - ищем точки смены съёмок (монтажных склеек).
- пространственная характеристика:
  - начальный и конечный кадр («мешок слов», GIST)
- временная характеристика:
  - отношения длин съёмок к длинам соседних съёмок
  - $+$ алгоритмы выравнивания последовательностей

Результаты и перспективы

Предложен подход

- относительные длины,
- выравнивания,

$\Rightarrow$ дескриптор ~~сцен~~ *фактов в видео*;

Проведены эксперименты (17 тыс. фильмов)

- точность $= 0.8$;
- полнота $= 0.7$;
- ложноотрицательные оценки;
- требуется более детальное практическое сравнение с существующими методами.

Приложение: алгоритм поиска нечетких дубликатов

- ν — новое видео;
- $\Pi = \{\pi_1, \pi_2, \ldots, \pi_n\}$ — исходные видео:
  - $\leftarrow$ Π может быть пустым;
  - $\leftarrow$ для непустого Π вычислены дескрипторы сцен элементов.

1. Сравниваем дескриптор каждой сцены $\sigma_{\nu,i}$ из ν с дескриптором каждой сцены $\sigma_{\pi_k,j}$ из π_k в L_2.
2. Если дескрипторы совпали ν с дескрипторами π_k. на некотором временном промежутке , то считаем эту часть ν — дубликатом π_k,

   несовпавшие части ν помещаем в Π.
3. Если дескрипторы не совпали, то считаем ν уникальным и добавляем в Π.

Приложение: относительные длины

Первые кадры сцен некоторого видео



- длина всех отрезков относительно всех, для видео это будет представлять матрицу

$$\begin{pmatrix} 1.0000 & 0.3171 & 0.3071 & 0.6000 \\ 3.1538 & 1.0000 & 0.9685 & 1.8923 \\ 3.2564 & 1.0325 & 1.0000 & 1.9538 \\ 1.6667 & 0.5285 & 0.5118 & 1.0000 \end{pmatrix};$$

- длина отрезков относительно некоторых:
  - например 3 предыдущих,
  - такой вариант применим для реального времени.

Приложение: временные отметки перемены сцен

Временные отметки перемены сцен для видео закодированного различными кодеками.

Отметки в секундах			
n	**cinepak**	**indeo5**	**h264**
1	0.133333	0.133333	0.133333
2	11.3333	—	—
3	74	74	74
4	78.9333	—	—
5	87.9333	—	87.9333
6	88.2667	88.2667	88.2667
7	88.3333	—	—
8	94.5333	94.5333	94.5333
9	—	101.133	101.133
10	101.4	—	101.4
11	—	—	112

Приложение: выравнивания последовательностей

Время	1 сек	2 сек	3 сек	4 сек	5 сек	6 сек	7 сек
Видео v_1	$\sigma_{1,1}$		$\sigma_{1,2}$				$\sigma_{1,3}$
Видео v_2	$\sigma_{2,1}$						$\sigma_{2,2}$

Алгоритм Гейла-Черча для выравниваня длин предложений параллельных корпусов на разных языках

- требуется установить, что v_1 и v_2, «переводы» друг друга;
- когда лучше выравнивать, *до* или *после* перехода к относительным длинам:

  до: перевычислять относительные длины,

  после: учитывать масштаб относительных длин;

- вычислительные затраты.