

1. Введение. Базовая обработка видео

Дмитрий Ватолин

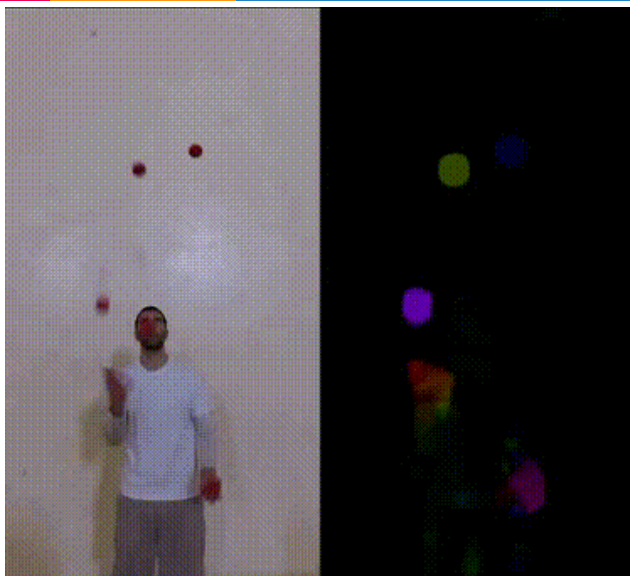
Video Group

CS MSU Graphics&Media Lab

План лекции

- **Введение, понятие видео**
- Задание «Детектор смены сцен»

Содержание курса (1)



Базовая обработка видео и принципы устройства видеокодеков

Классические и нейросетевые подходы обработки видео и устранения артефактов

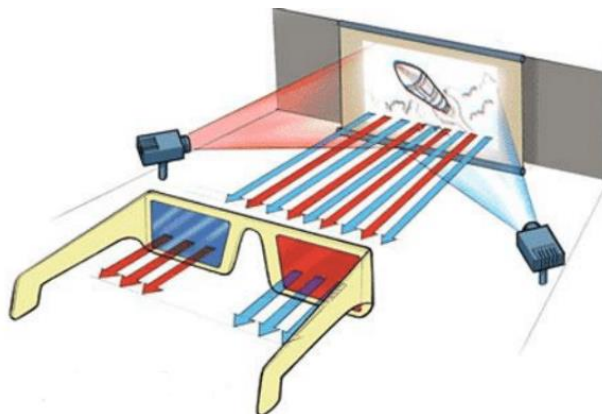
Содержание курса (2)



Обработка пленоптических
фотографий

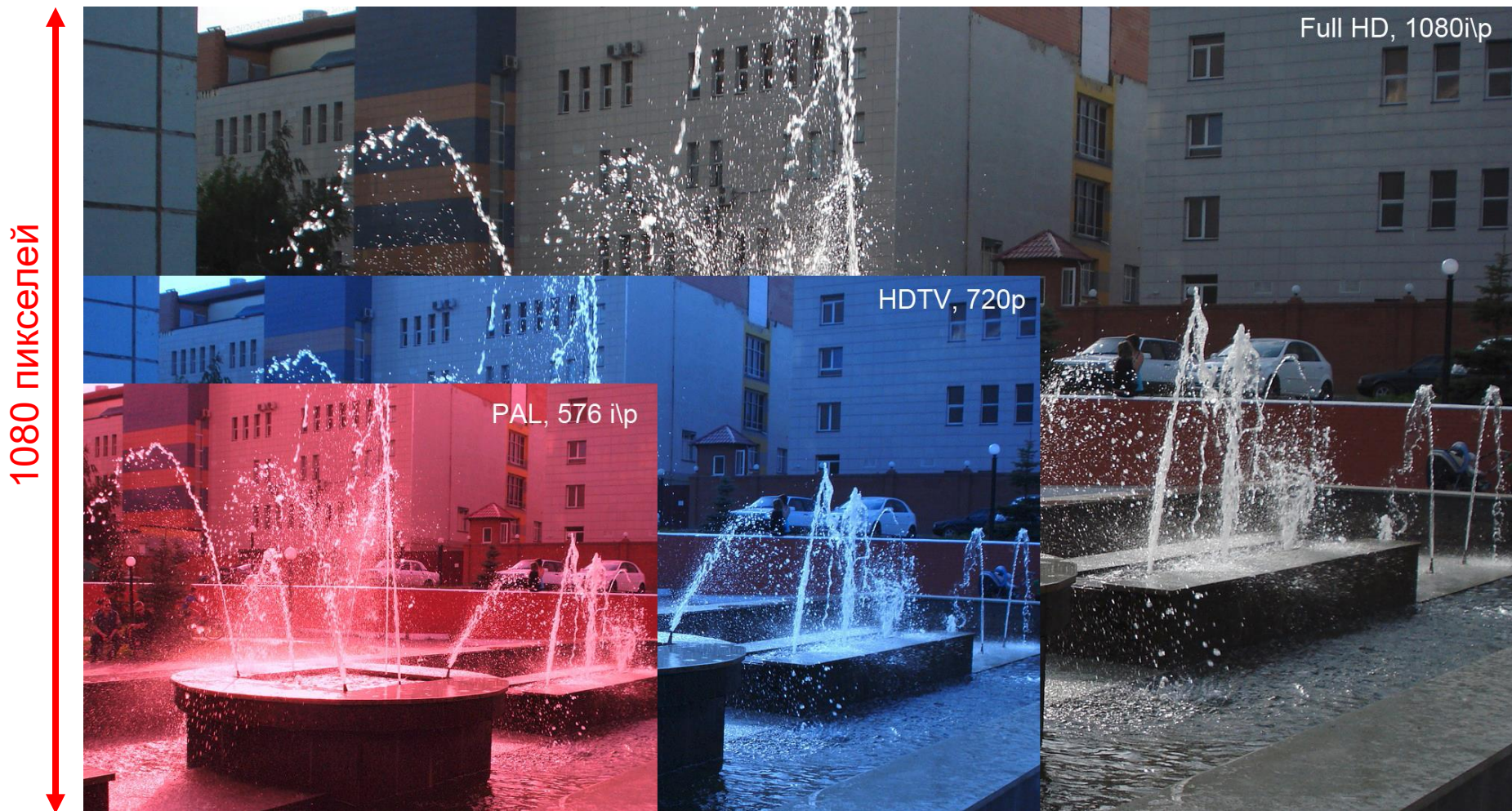


3D и 360 видео



Характеристики видео

1. Разрешение



Характеристики видео

2. Частота кадров



Normal speed



30fps



60fps

Цвет каждого пикселя изображения описывается при помощи цветового пространства

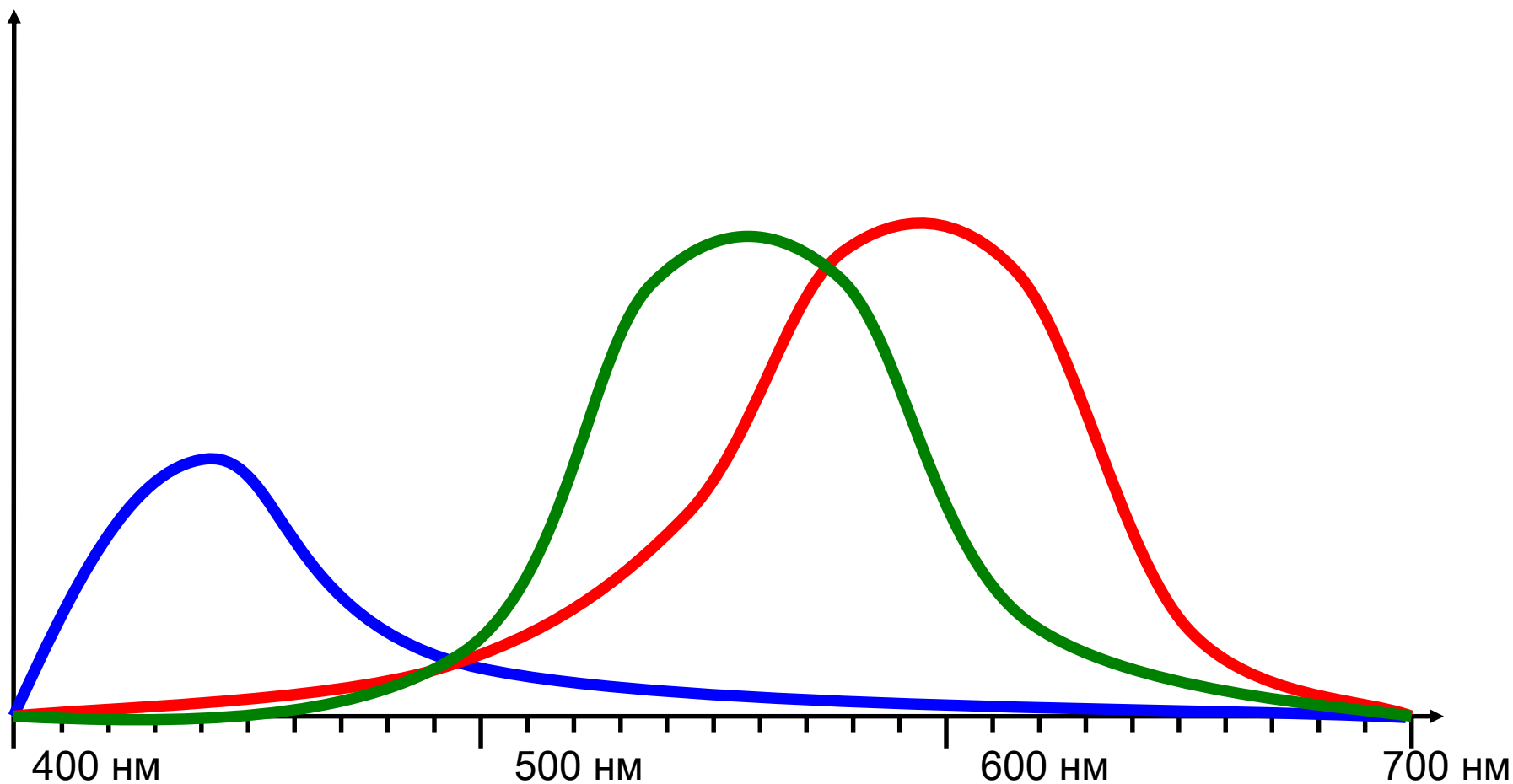
Цветовое пространство строится таким образом, чтобы любой цвет был представлен точкой, имеющей определённые координаты

Примеры цветовых пространств:

- RGB
- CMYK
- HSV
- YUV

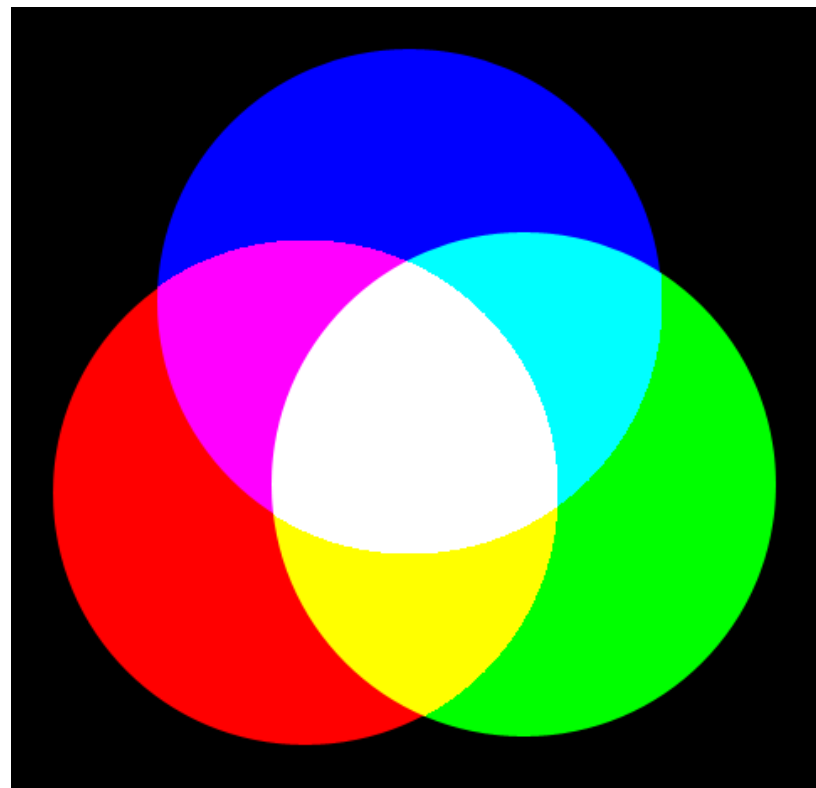
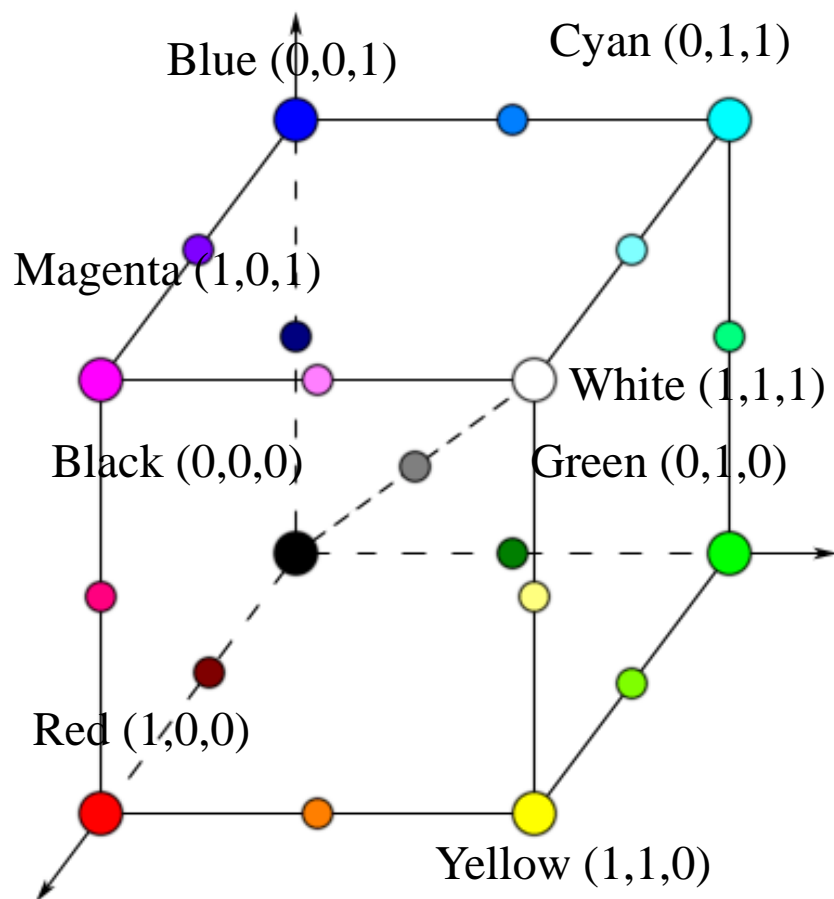
Цветовые пространства

Чувствительность



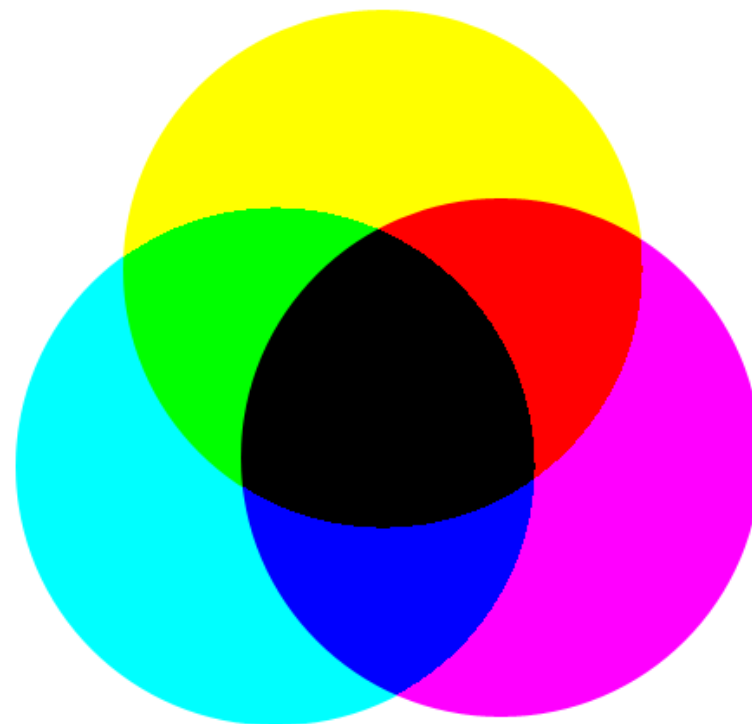
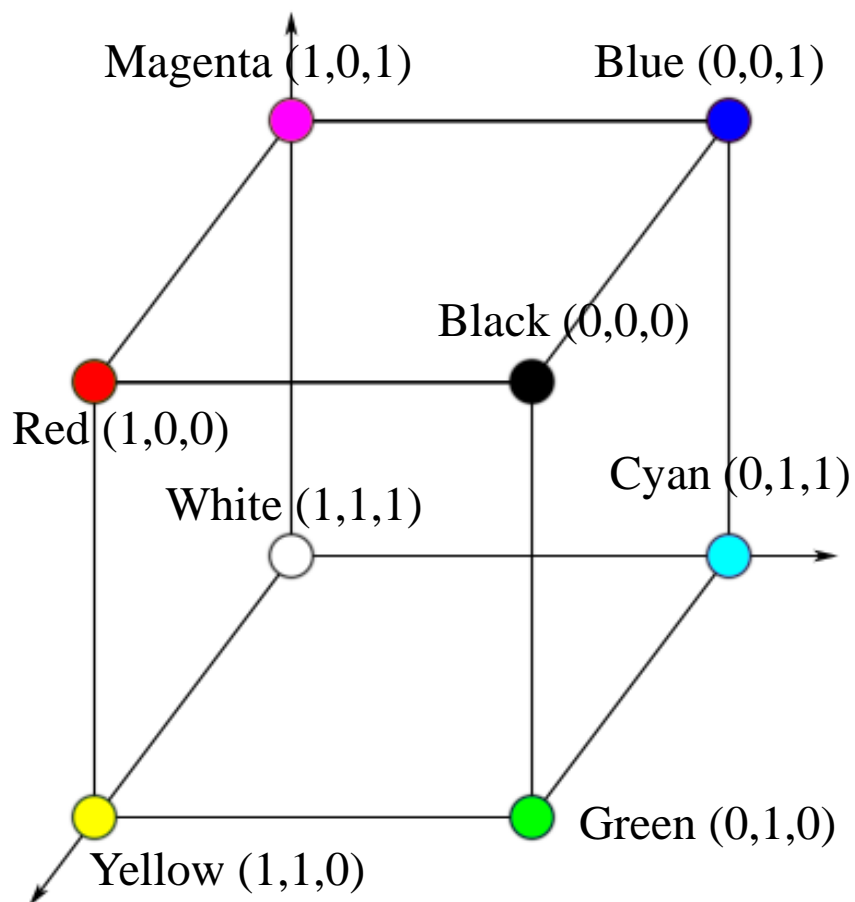
Цветовые пространства

RGB (Red, Green, Blue)



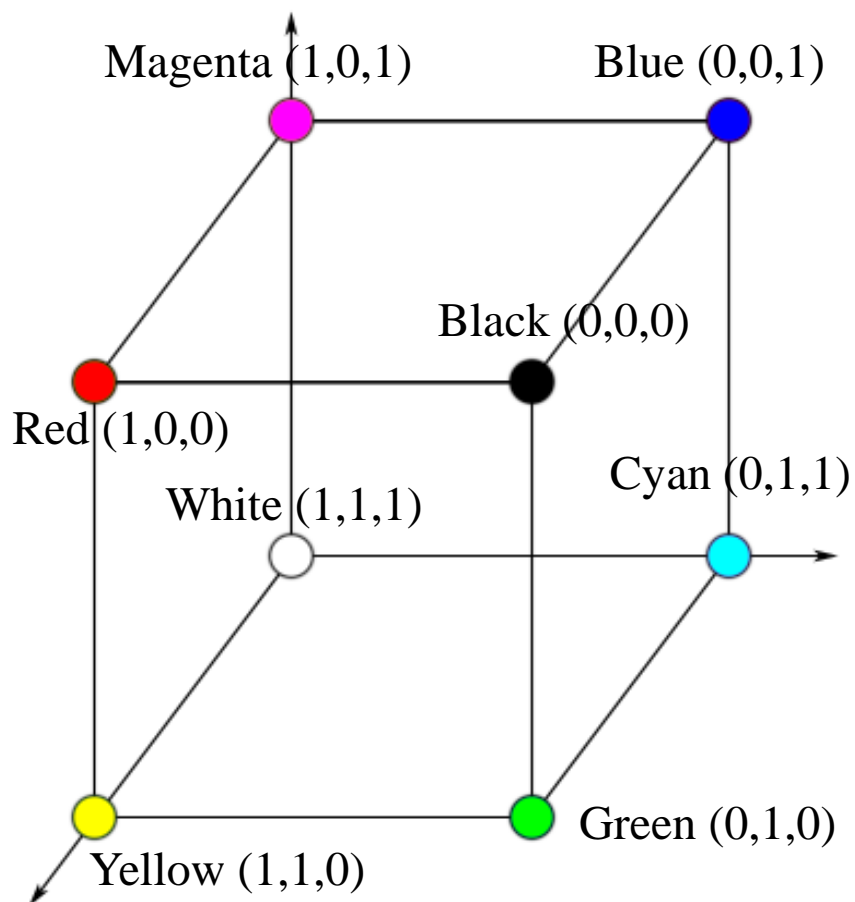
Цветовые пространства

СМΥК (Cyan, Magenta, Yellow, black)



Цветовые пространства

CMYK ↔ RGB



$$C = 255 - R$$

$$M = 255 - G$$

$$Y = 255 - B.$$

$$K = \min(C, M, Y),$$

$$C = C - K,$$

$$M = M - K,$$

$$Y = Y - K.$$

Цветовые пространства

HSV (Hue, Saturation, Value)

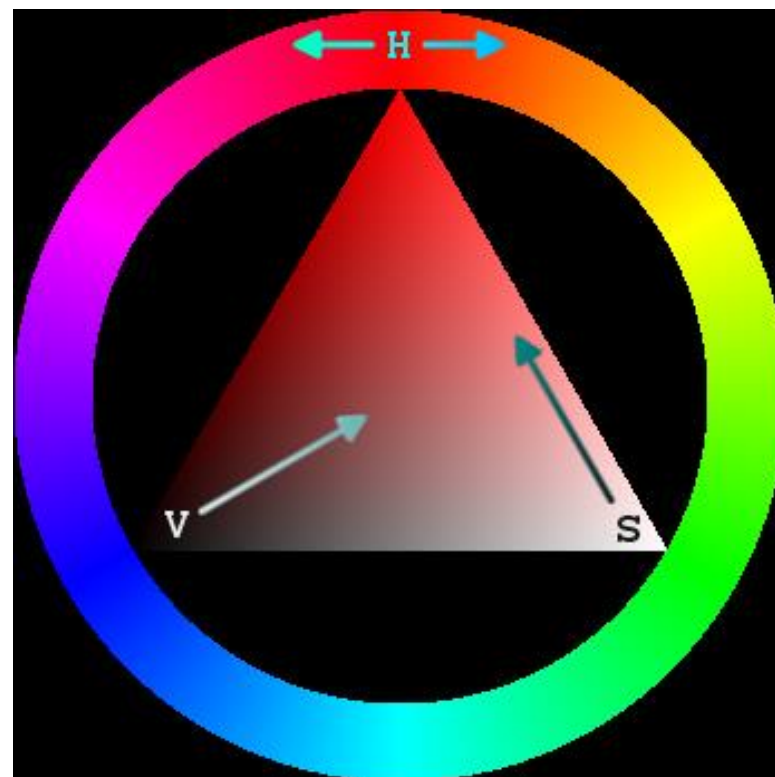
$$Max = \max(R, G, B)$$

$$Min = \min(R, G, B)$$

$$V = Max$$

$$S = \begin{cases} 0, & \text{если } Max = 0 \\ 1 - \frac{Min}{Max} & \text{иначе} \end{cases}$$

$$H = \begin{cases} 0, & \text{если } Max = Min \\ 60 \frac{G - B}{Max - Min}, & \text{если } Max = R \text{ и } G \geq B \\ 60 \frac{G - B}{Max - Min} + 360, & \text{если } Max = R \text{ и } G < B \\ 60 \frac{B - R}{Max - Min} + 120, & \text{если } Max = G \\ 60 \frac{R - G}{Max - Min} + 240, & \text{если } Max = B \end{cases}$$



Цветовые пространства

YUV



$$\begin{matrix} Y \\ U \\ V \end{matrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.189 \end{bmatrix} \begin{matrix} R \\ G \\ B \end{matrix}$$

Представление изображения в цветовом пространстве YUV



Цветовые пространства

Зачем их так много?



- RGB — телевизоры и мониторы
- YUV — совместимость чёрно-белого и цветного телевидения
- HSV — фоторедакторы: удобное представление цветового пространства на плоскости
- CMYK — печать

План лекции

- Введение, понятие видео
- **Задание «Детектор смены сцен»**

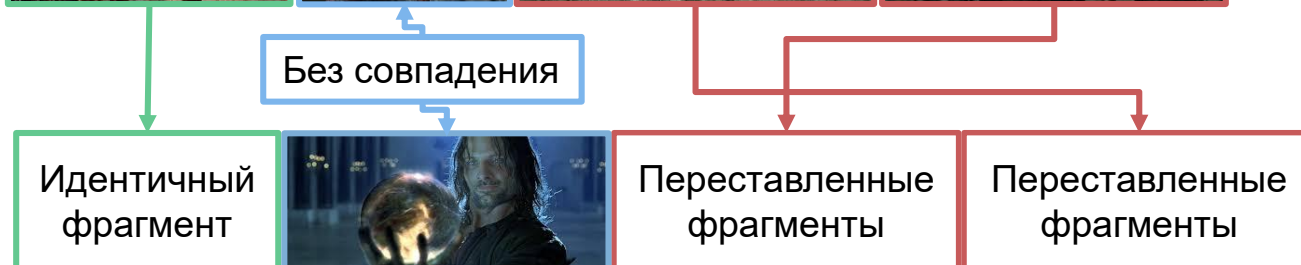
Детектор смены сцен

Задача сопоставления версий фильмов

Театральная
версия фильма



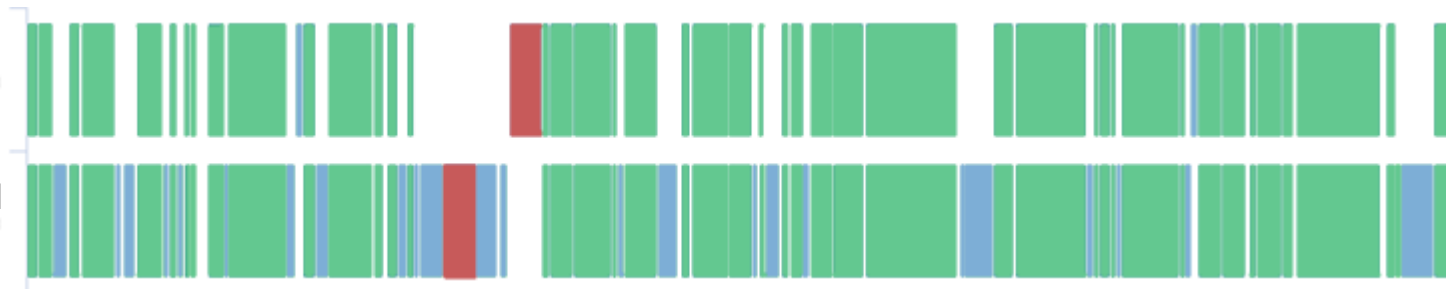
Режиссёрская
версия фильма



Пример: «Властелин колец: Две крепости» (2002),
101 переставленная сцена, почти час новых сцен

Театральная
версия

Режиссёрская
версия



Детектор смены сцен

Первое задание: постановка задачи

Важной подзадачей является разбить фильмы на **сцены** — непрерывные отрезки, снятые одной камерой без прерывания



Детектор смены сцен

Идея решения

Требуется оценить близость (похожесть) соседних кадров (могут ли они принадлежать одной сцене)



Кадры похожи →
нет смены сцен



Кадры не похожи →
есть смена сцен

Вопрос: как определить «похожесть» между кадрами?

Детектор смены сцен

Среднеквадратическое отклонение между кадрами

Подсчитаем попиксельную разницу между соседними кадрами: $diff(A, B) = \frac{1}{\# \text{пикселей}} \sum_{\forall i} \sum_{\forall j} (A_{ij} - B_{ij})^2$



Детектор смены сцен

Среднеквадратическое отклонение между кадрами

Подсчитаем попиксельную разницу между соседними кадрами:
$$diff(A, B) = \frac{1}{\# \text{пикселей}} \sum_{\forall i} \sum_{\forall j} (A_{ij} - B_{ij})^2$$



Детектор смены сцен

Среднеквадратическое отклонение между кадрами

Подсчитаем попиксельную разницу между соседними кадрами: $diff(A, B) = \frac{1}{\# \text{пикселей}} \sum_{\forall i} \sum_{\forall j} (A_{ij} - B_{ij})^2$



Разница: 33.13

Детектор смены сцен

Среднеквадратическое отклонение между кадрами



Подсчитаем попиксельную разницу между соседними

кадрами: $diff(A, B) = \frac{1}{\text{\#пикселей}} \sum_{\forall i} \sum_{\forall j} (A_{ij} - B_{ij})^2$



Детектор смены сцен

Среднеквадратическое отклонение между кадрами

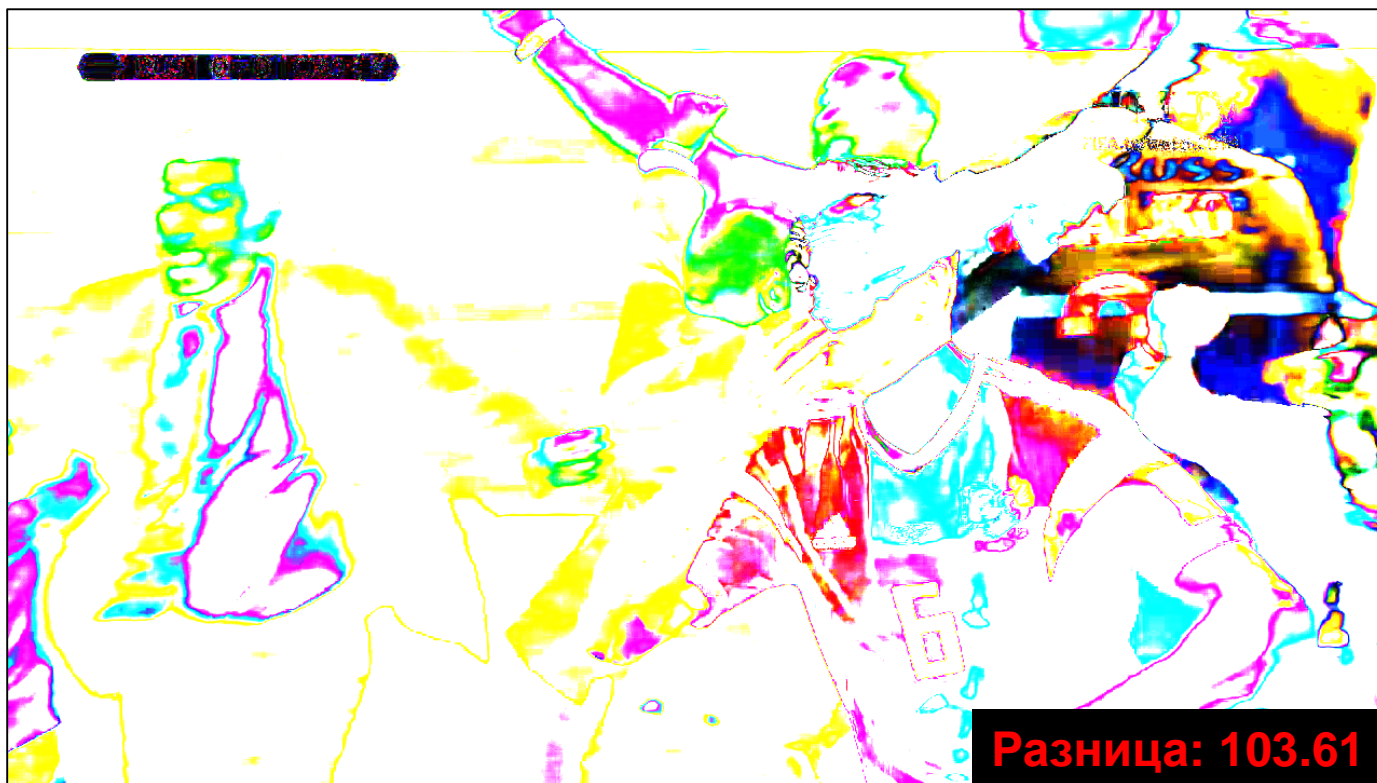
Подсчитаем попиксельную разницу между соседними кадрами: $diff(A, B) = \frac{1}{\text{\#пикселей}} \sum_{\forall i} \sum_{\forall j} (A_{ij} - B_{ij})^2$



Детектор смены сцен

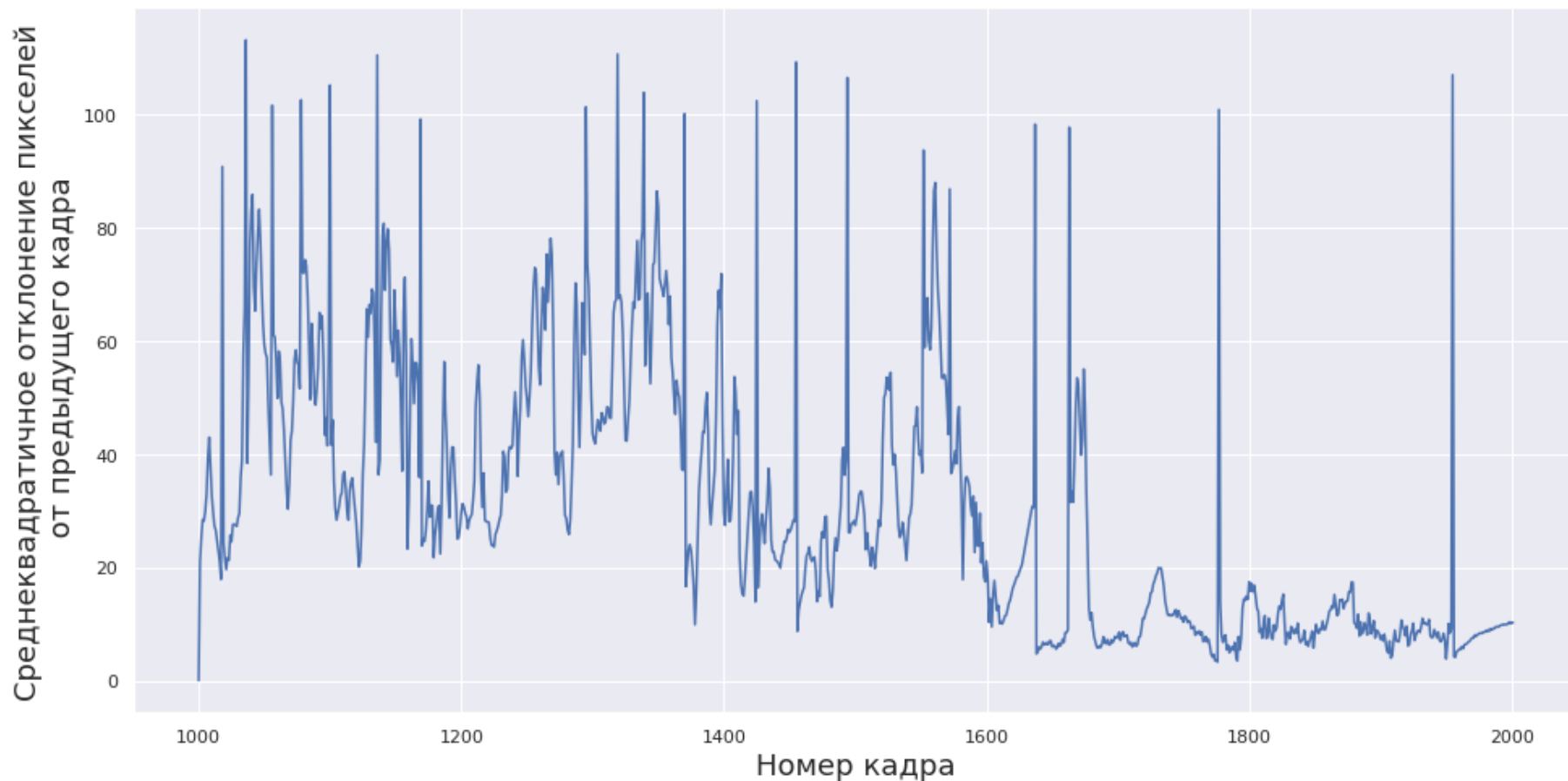
Среднеквадратическое отклонение между кадрами

Подсчитаем попиксельную разницу между соседними кадрами: $diff(A, B) = \frac{1}{\text{\#пикселей}} \sum_{\forall i} \sum_{\forall j} (A_{ij} - B_{ij})^2$



Детектор смены сцен

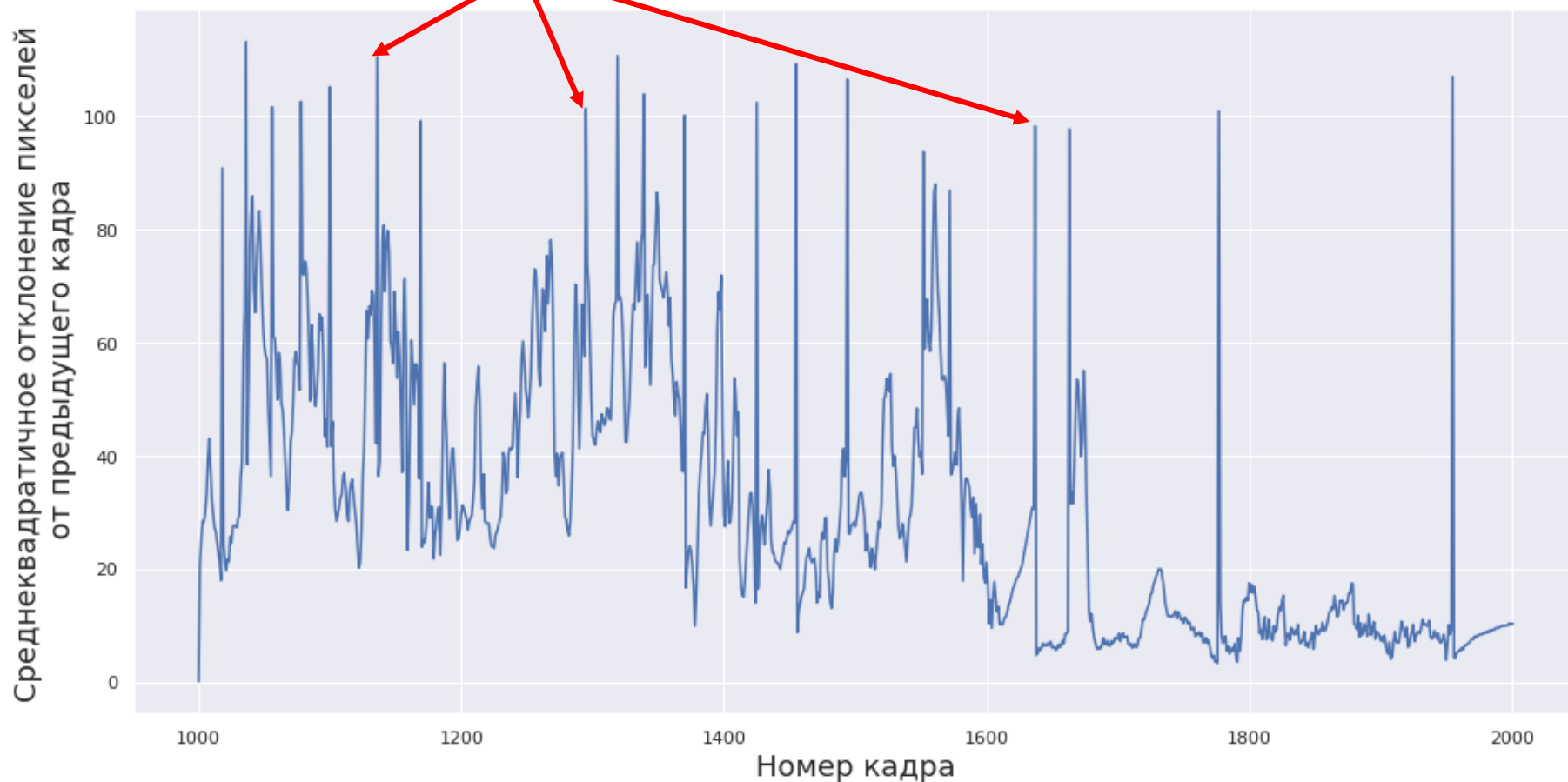
Среднеквадратическое отклонение между кадрами



Детектор смены сцен

Среднеквадратическое отклонение между кадрами

Пики соответствуют сменам сцен

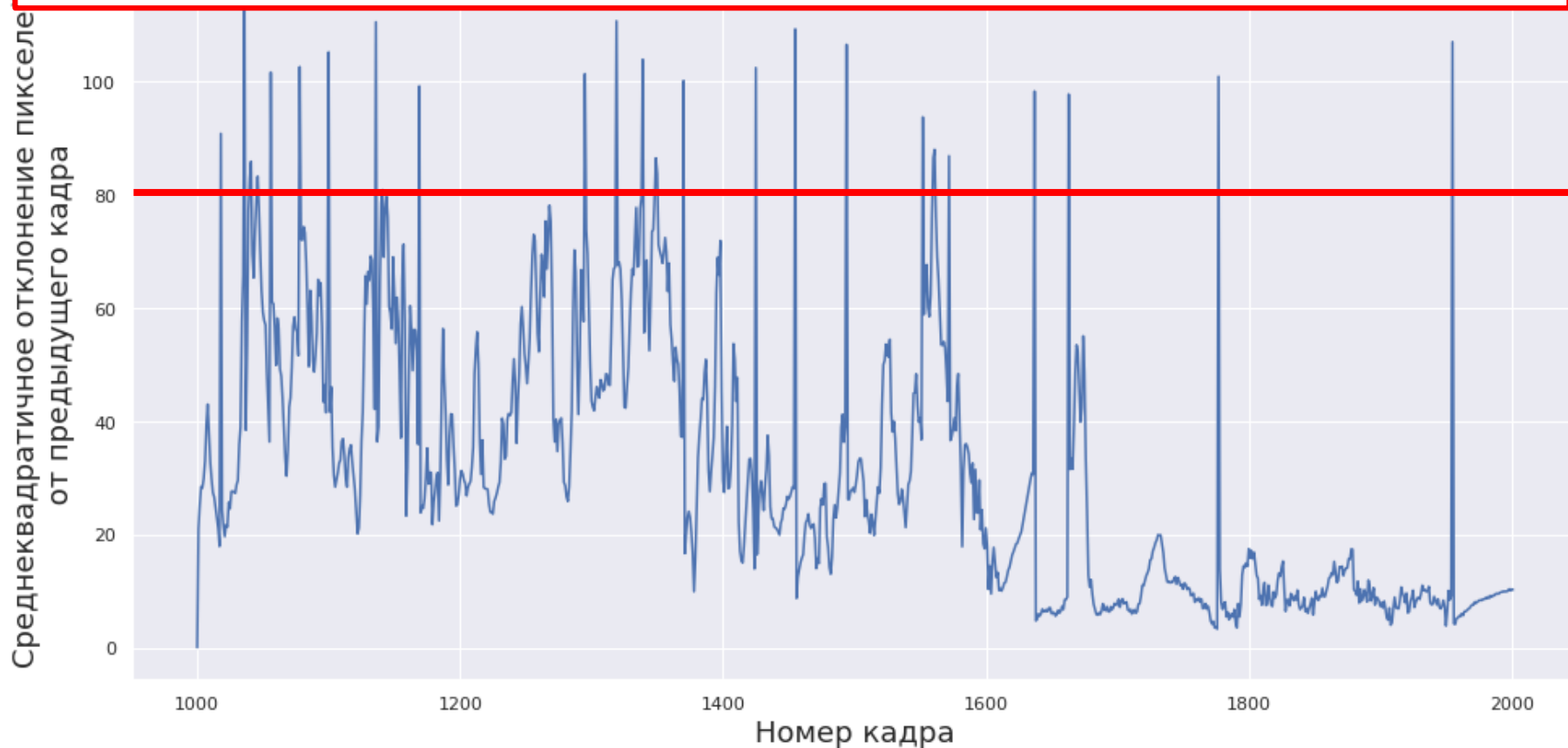


Детектор смены сцен

Среднеквадратическое отклонение между кадрами



Подбирается порог, чтобы отсечь большинство смен сцен



Детектор смены сцен

Среднеквадратическое отклонение между кадрами

На динамичных сценах, с большим движением в кадре будут ложные срабатывания:



Детектор смены сцен

Среднеквадратическое отклонение между кадрами

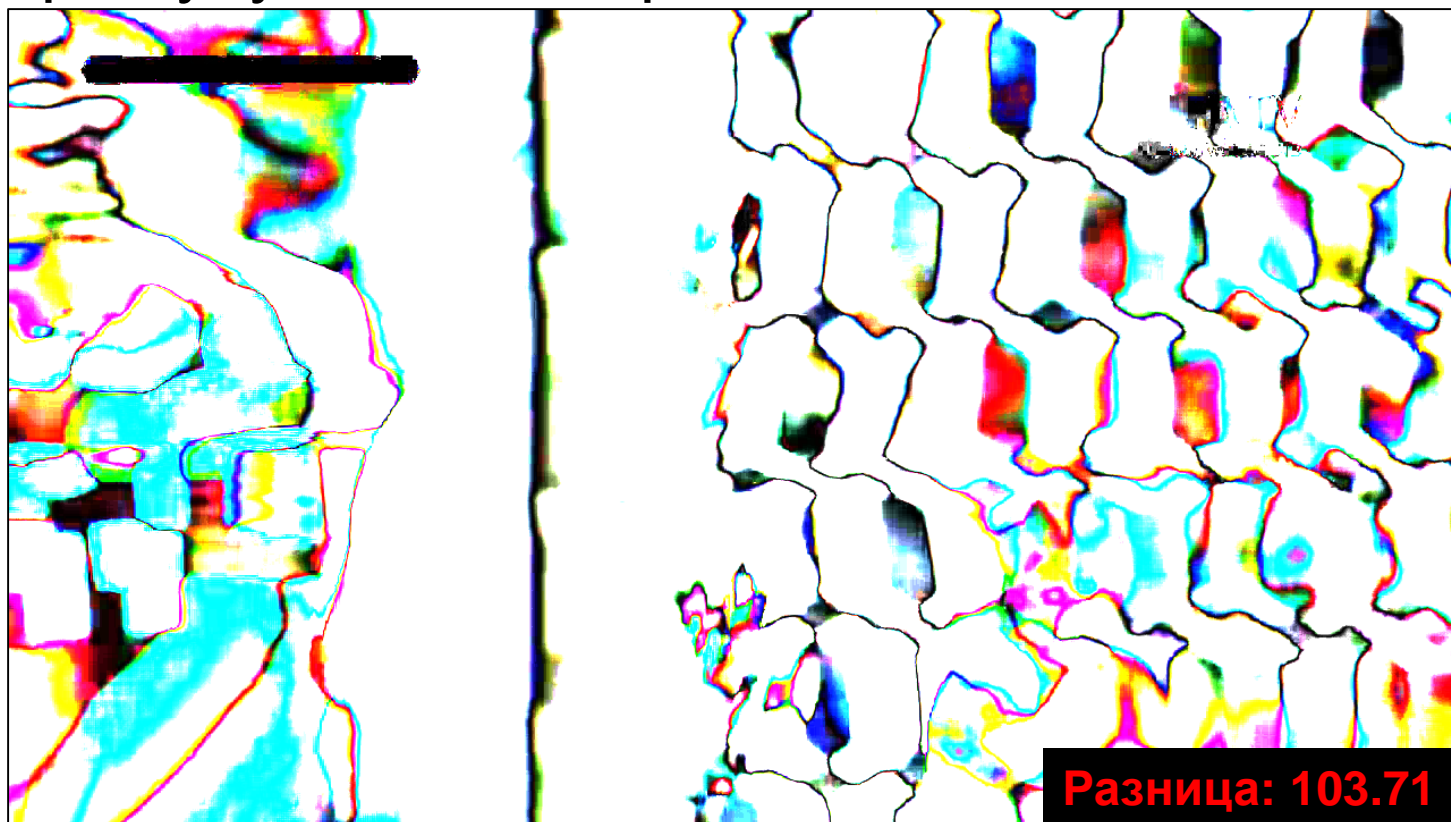
На динамичных сценах, с большим движением в кадре будут ложные срабатывания:



Детектор смены сцен

Среднеквадратическое отклонение между кадрами

На динамичных сценах, с большим движением в кадре будут ложные срабатывания:



Детектор смены сцен

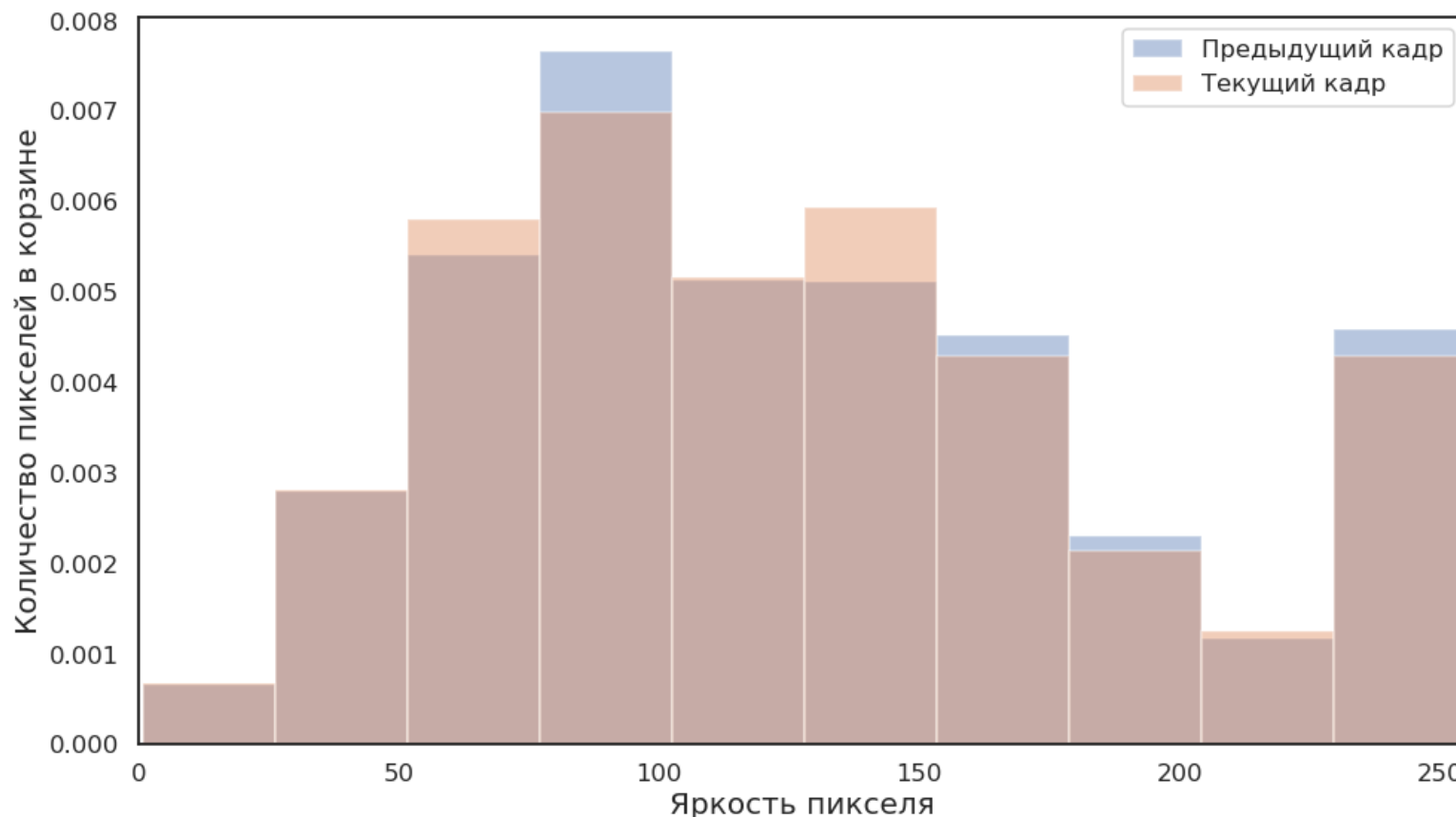
Гистограммы цветов кадров



Подсчитаем гистограмму распределения цветовой компоненты (например, яркости) для каждого из кадров

Детектор смены сцен

Гистограммы цветов кадров



Для сравнения кадров можно использовать среднюю разность между значениями в корзинах

Детектор смены сцен

Гистограммы цветов кадров



Предыдущий кадр

Детектор смены сцен

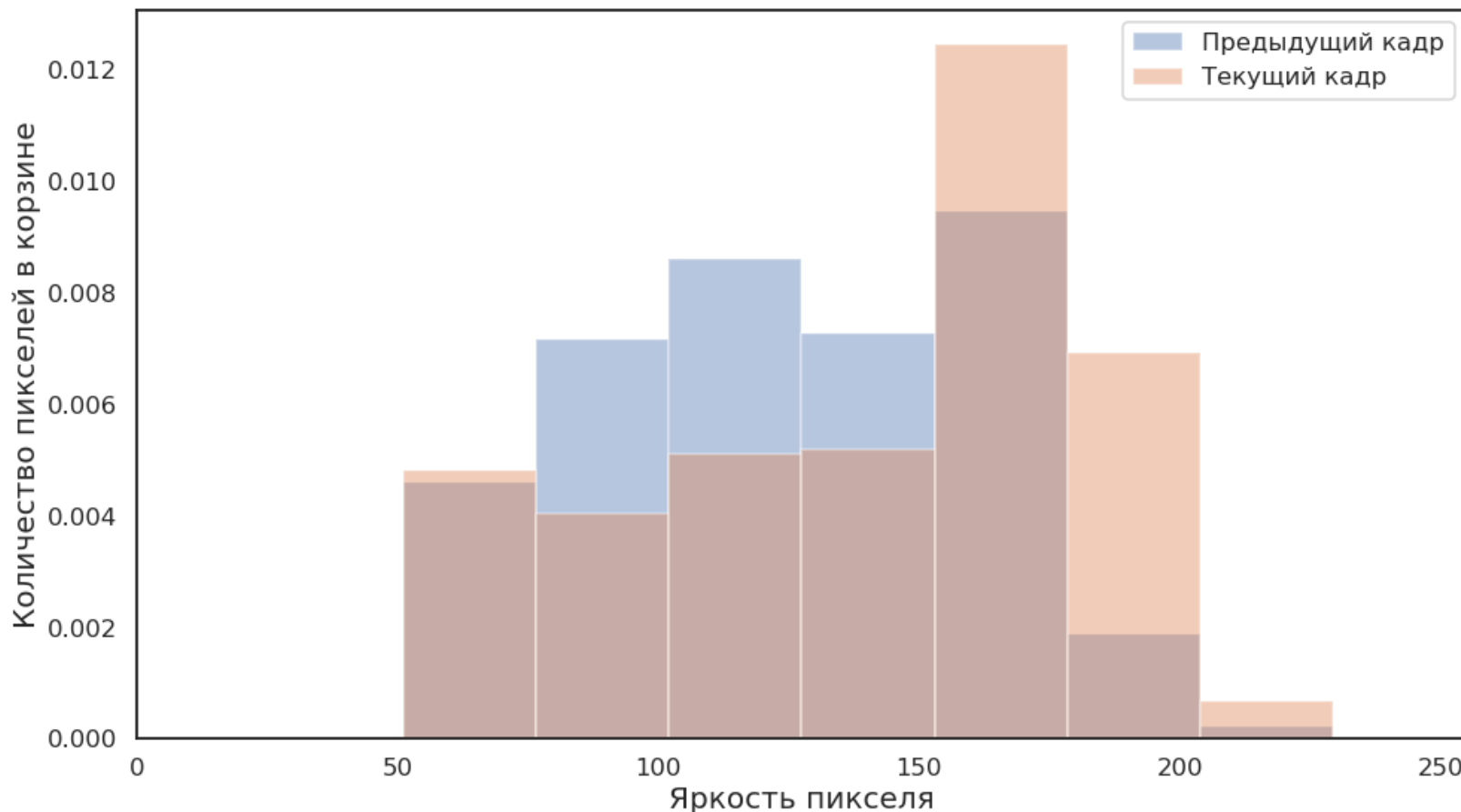
Гистограммы цветов кадров



Текущий кадр

Детектор смены сцен

Гистограммы цветов кадров



Вывод: необходимы цветонезависимые признаки

Детектор смены сцен

Цветонезависимые признаки



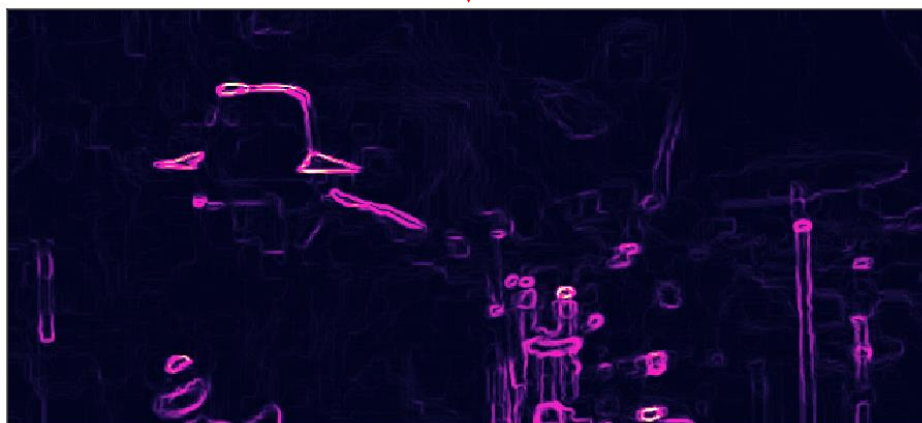
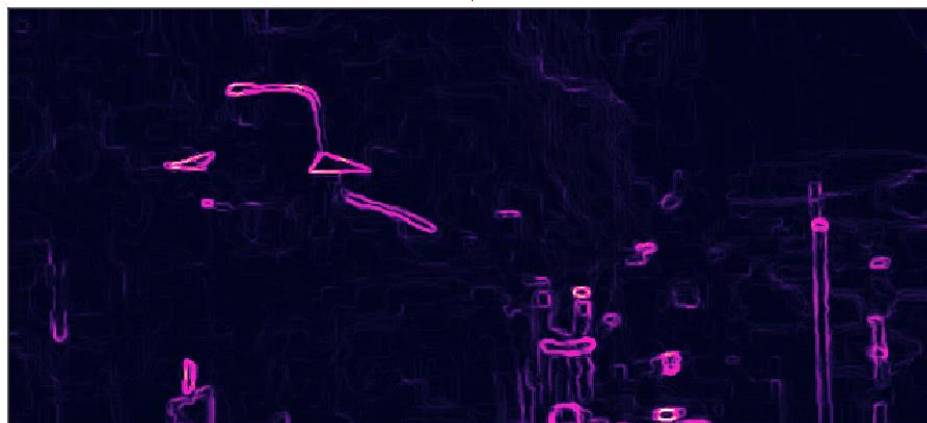
- Вычислим по кадру видео его границы — области с большими перепадами яркости
- При большом глобальном изменении яркости (вспышка камеры, лучи света, блики) локальные перепады яркости сохраняются

Детектор смены сцен

Цветонезависимые признаки



Границы на кадре



Детектор смены сцен

Обнаружение границ изображения



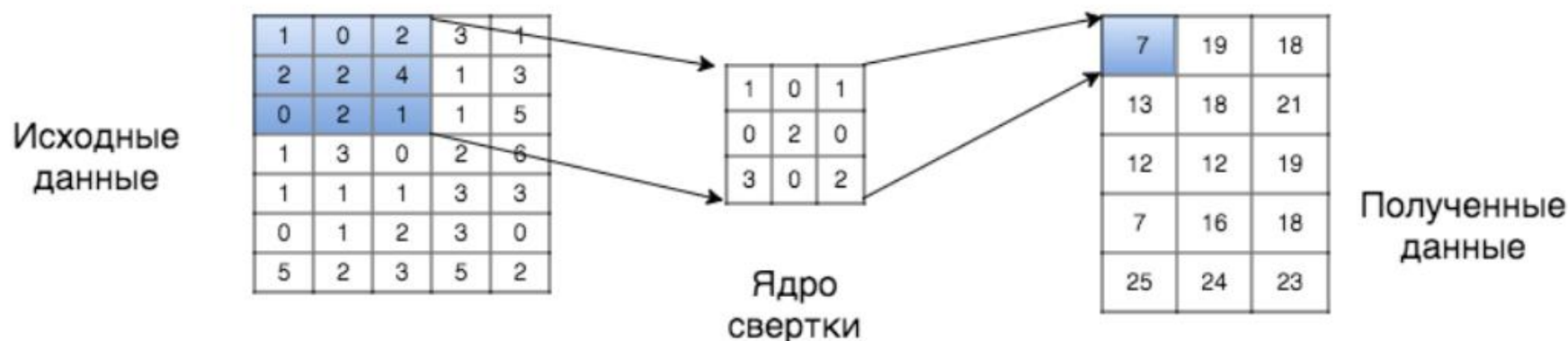
Простой способ обнаружения границ на изображении — вычисление приближённых производных в точке

$$\bullet \text{ Sobel}_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}; \text{ Sobel}_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

$$\bullet \text{ Edges}_{abs} = \sqrt{(A * \text{Sobel}_y)^2 + (A * \text{Sobel}_x)^2}$$

- A — исходное изображение
- Edges — границы
- $*$ — применение операции свёртки

Что такое свёртка



При вычислении нового значения выбранного пикселя изображения, ядро свёртки прикладывается своим центром к этому пикселю, а затем вычисляется сумма поэлементного произведения ядра свёртки и матрицы, которую накрыло данное ядро

Вычисление границ по шагам



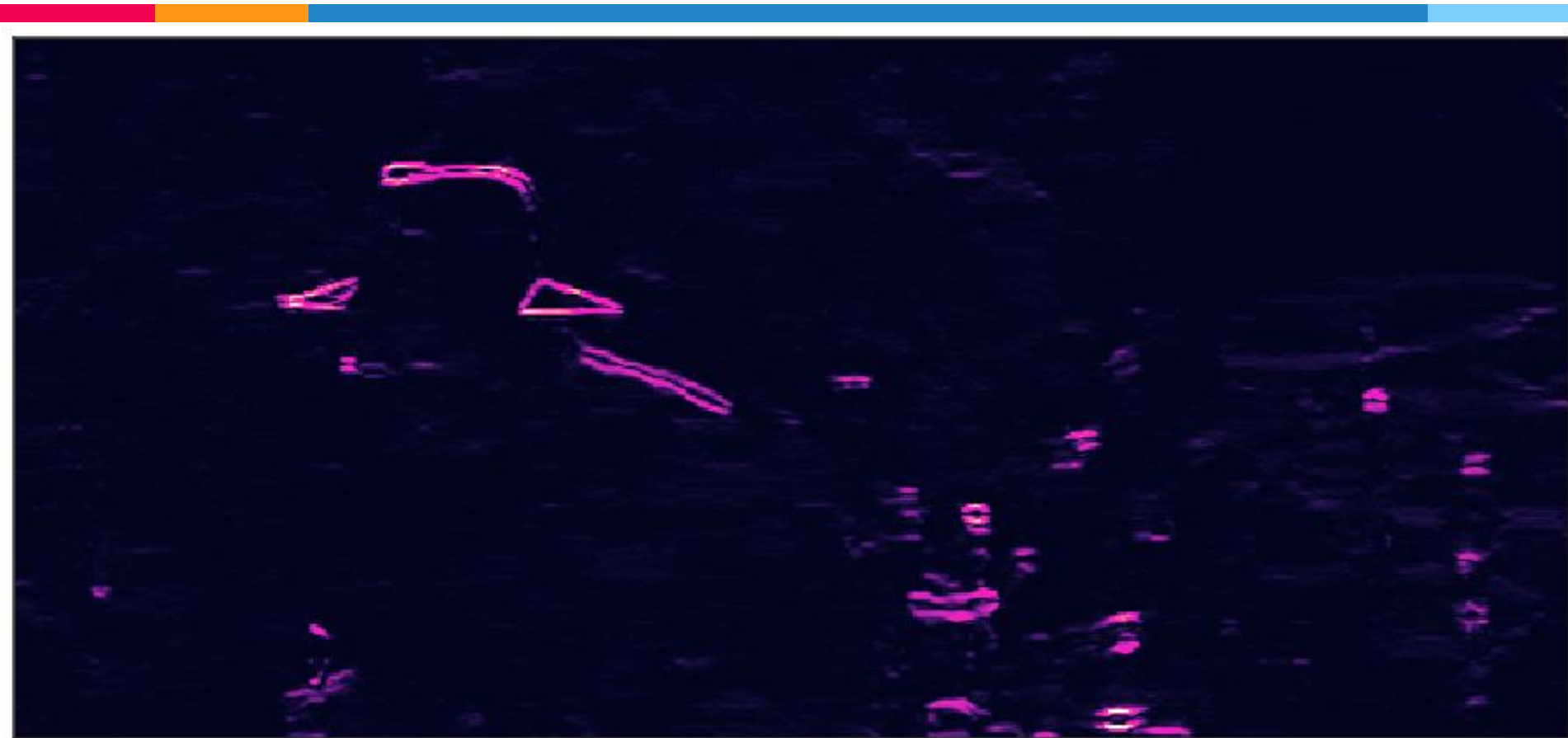
Исходное изображение

Вычисление границ по шагам



$Sobel_x * A$ — вертикальные границы

Вычисление границ по шагам



$Sobel_y * A$ — горизонтальные границы

Вычисление границ по шагам



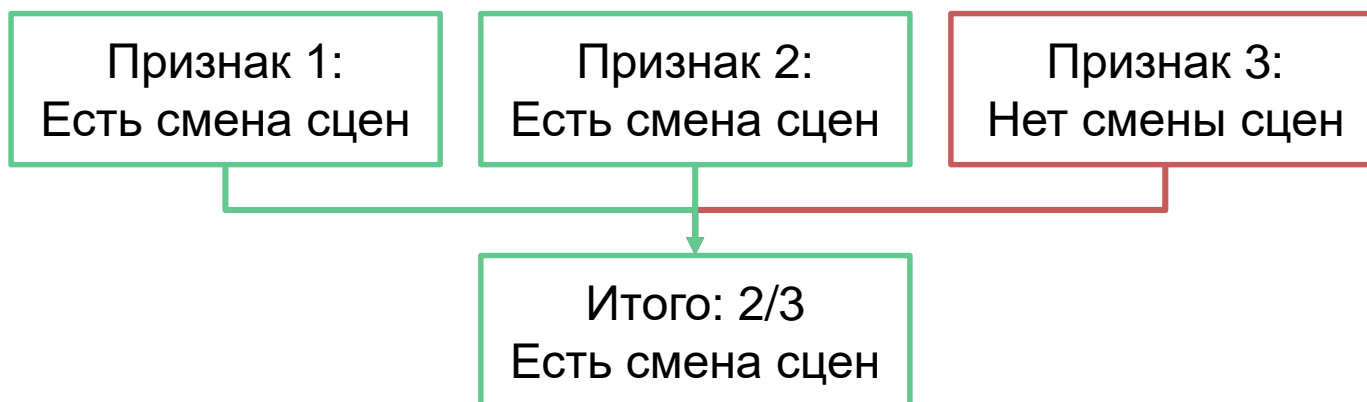
$$\sqrt{(Sobel_y * A)^2 + (Sobel_x * A)^2}$$

Детектор смены сцен

Объединение признаков

Объединить несколько признаков можно следующими способами:

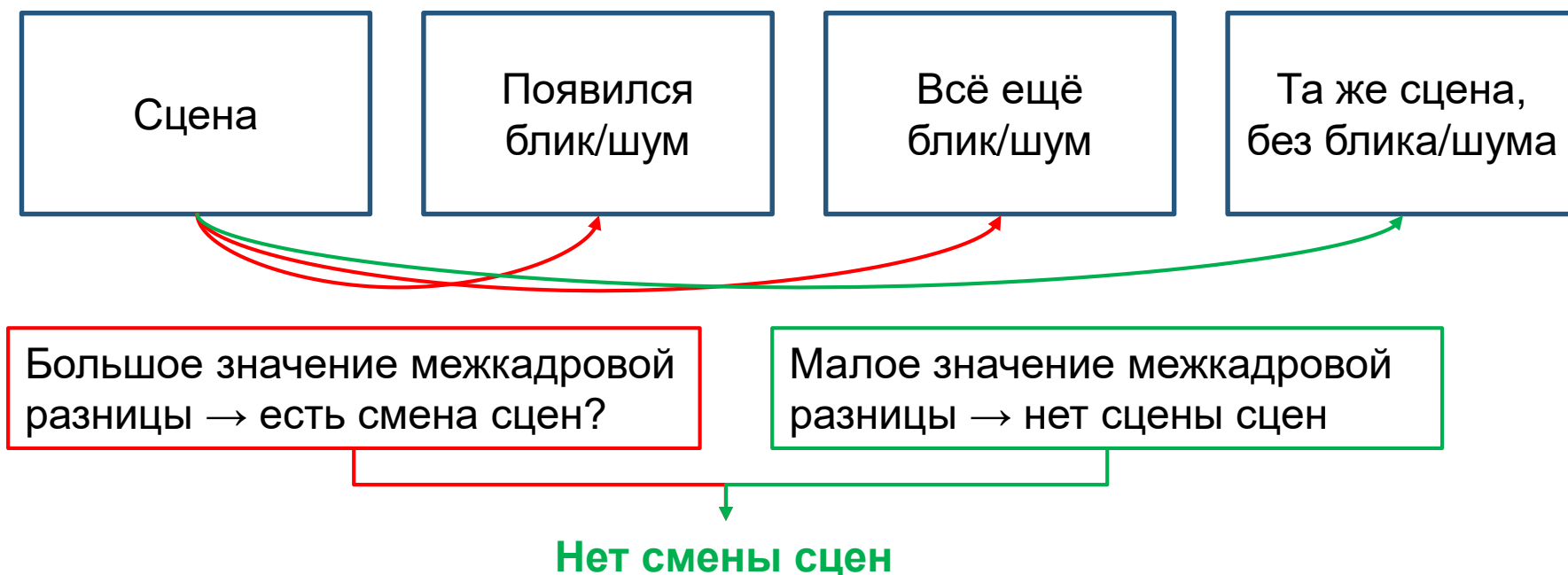
- 1) Итоговое решение детектора принимается путём голосования признаков
- 2) Совместный подбор параметров по сетке
- 3) Методы машинного обучения



Детектор смены сцен

Повышение устойчивости решения

Для повышения устойчивости решения также можно вычислять признаки на расстоянии нескольких кадров



Детектор смены сцен

Более сложные смены сцен



Cut (мгновенная смена сцен)



Dissolve (один кадр «растворяется» в другом)



Wipe (что любит Windows Movie Maker)

Детектор смены сцен

Dissolve — модель смены сцен



Введём модель смены сцен:

$$F_t = \alpha_t A_t + (1 - \alpha_t) B_t$$

где

- F_t — наблюдаемый кадр в момент времени t
- A_t — кадр первой сцены в момент времени t
- B_t — кадр второй сцены в момент времени t
- α_t — прозрачность сцены A в момент времени t (монотонно убывает)

Детектор смены сцен

Dissolve — следствия из модели



Вычислить α_t или характеристики кадров A_t , B_t не всегда возможно, однако, можно сделать несколько следствий из модели смены сцен (так как $A_t \approx A_{t+1}$, $B_t \approx B_{t+1}$, $\alpha_t - \alpha_{t+1} \approx \alpha_{t-1} - \alpha_t$:

- 1) Большинство пикселей либо монотонно возрастает либо монотонно убывает
- 2) Справедливо $F_t \approx \frac{1}{2} (F_{t+1} + F_{t-1})$

Детектор смены сцен

Dissolve — монотонные пиксели



Предыдущий кадр

Детектор смены сцен

Dissolve — монотонные пиксели



Текущий кадр

Детектор смены сцен

Dissolve — монотонные пиксели



Пиксели, являющиеся монотонными

Детектор смены сцен

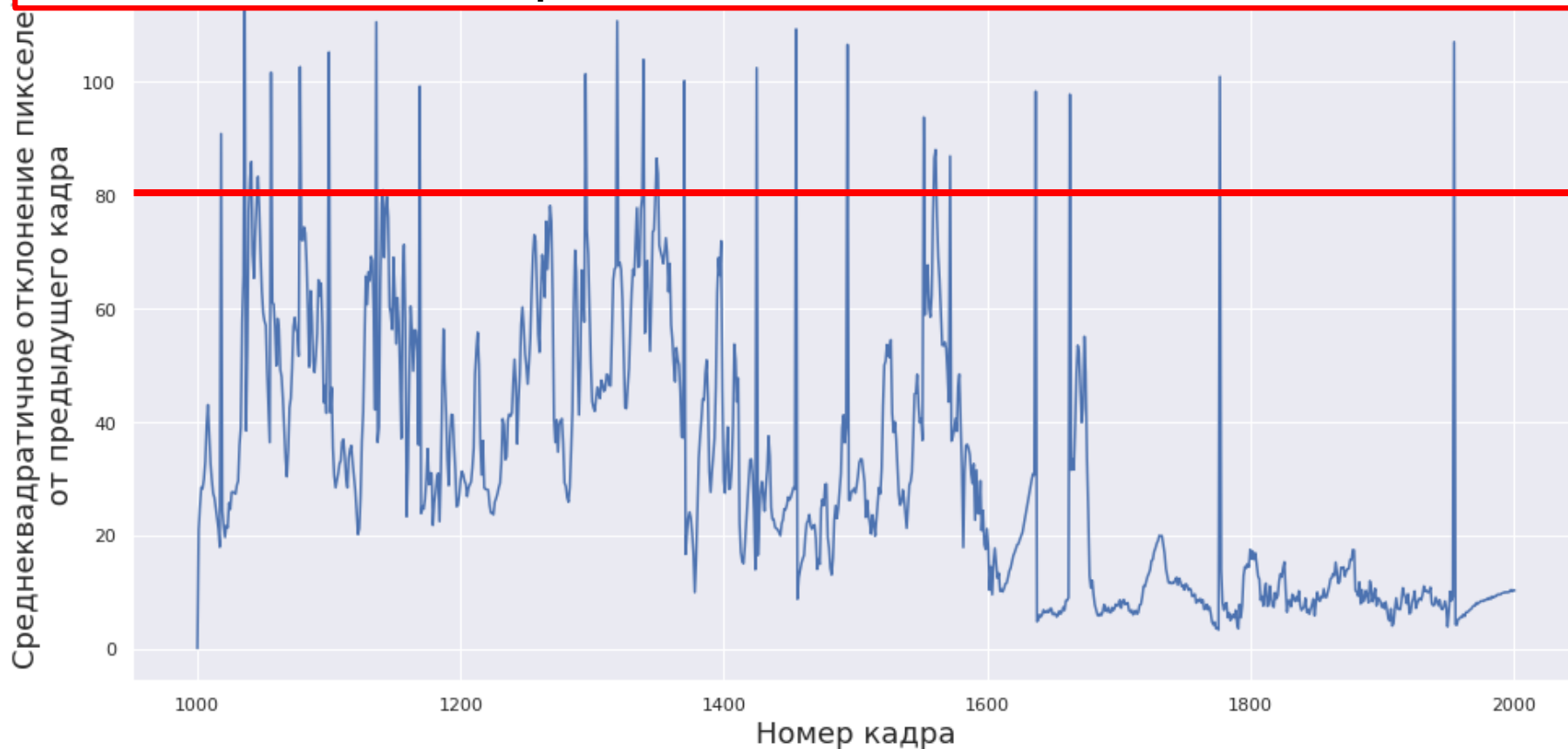
Dissolve — монотонные пиксели



Детектор смены сцен

Проблема подбора порогов

Подобрать порог так, чтобы полностью отделить все смены сцен, как правило, невозможно

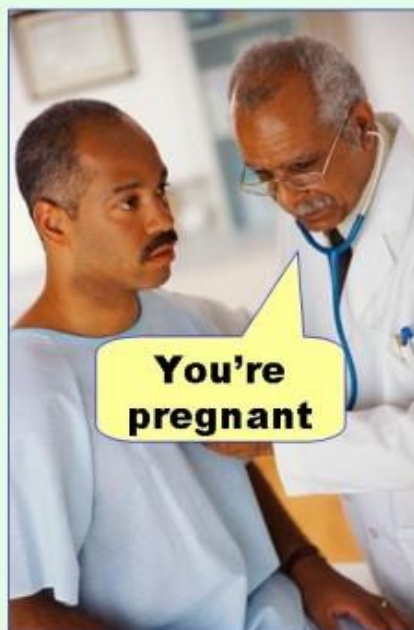


Детектор смены сцен

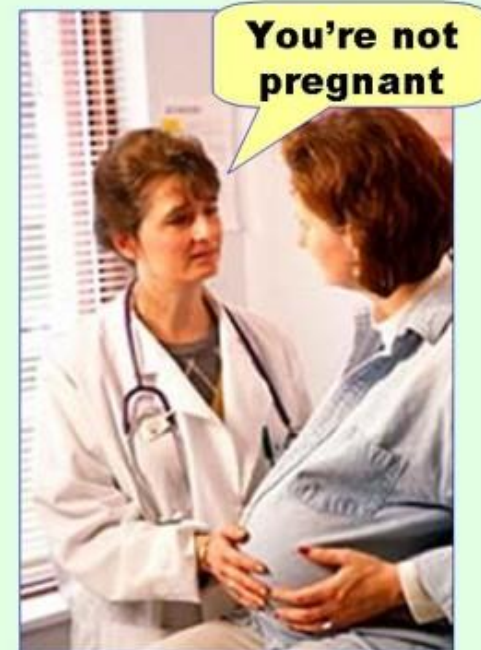
Ошибки 1 и 2 рода

		Метка	
		Да	Нет
Алгоритм	Да	True Positive	False Positive
	Нет	False Negative	True Negative

Type I error
(false positive)



Type II error
(false negative)



Детектор смены сцен

Ошибки 1 и 2 рода

Среднеквадратическое отклонение между кадрами: 87.93

Предыдущий кадр



Текущий кадр



Среднеквадратическое отклонение между кадрами: 81.41

Предыдущий кадр



Текущий кадр



Детектор смены сцен

Accuracy



		Метка	
		Да	Нет
Алгоритм	Да	True Positive	False Positive
	Нет	False Negative	True Negative

Accuracy (Точность средства измерений) вычисляется по формуле:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

Детектор смены сцен

Accuracy — недостатки



Accuracy нельзя использовать, если классы «несбалансированы»

		Метка	
		Да	Нет
Алгоритм А	Да	4	6
	Нет	6	84

$$Accuracy = 0.88$$

		Метка	
		Да	Нет
Алгоритм Б	Да	0	0
	Нет	10	90

$$Accuracy = 0.9$$

Алгоритм Б по метрике лучше А, однако, он ничего не делает!

Детектор смены сцен

Precision & Recall



		Метка	
		Да	Нет
Алгоритм	Да	True Positive	False Positive
	Нет	False Negative	True Negative

Precision (точность результата измерений) и Recall (полнота) вычисляются по формулам:

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

Детектор смены сцен

F1-score



$$F1 = \frac{2 * Precision * Recall}{(Precision + Recall)}$$

- F1-score (F-мера) применяется в случае, когда классы несбалансированны
- «Смысл» метрики — среднее геометрическое между Precision и Accuracy

Детектор смены сцен

F1-score — сравнение двух алгоритмов



F1-score для алгоритма А: 0.4

F1-score для алгоритма Б: 0

		Метка	
		Да	Нет
Алгоритм А	Да	4	6
	Нет	6	84

		Метка	
		Да	Нет
Алгоритм Б	Да	0	0
	Нет	10	90

Детектор смены сцен

Итоговый алгоритм решения



- Придумываем признаки, позволяющие отделить смены сцен от последовательных кадров
- Смен сцен гораздо меньше, чем последовательных кадров, поэтому для оценки качества используем F1-score
- Для повышения качества решения можно объединять признаки, а также можно считать признаки не только между соседними кадрами, но и через несколько кадров
- Пороги для признаков можно подобрать путём максимизации F1-score на тренировочной выборке

Детектор смены сцен

Итоговый алгоритм решения — Dissolve



- Придумываем признаки, описывающие кадр
- Проверяем насколько они удовлетворяют модели смены сцен
- Пороги для признаков можно подобрать путём максимизации F1-score на тренировочной выборке
- Не забываем фильтровать кадры — для каждой смены сцен оставляем только 1 кадр

Детектор смены сцен

Сроки сдачи задания



- Время на выполнение задания — 1 неделя
- Вопросы по заданию задавать в телеграм-чате

Полезные ссылки

- anytask.org/course/465
(ВМК МГУ/Интеллектуальные методы
обработки видео)
Инвайт: QYpg93Y
- Телеграм-канал: bit.ly/vgcourse_channel



Страница курса



Телеграм-канал