

Курс “Анализ изображений”

Лекция#1.

Введение. Задачи компьютерного зрения.
Формирование цифрового изображения

Как выглядит лектор?

Александр Жуковский

- CTO, NVI Solutions
- TeamLead, Smart Engines
- mailto: zhukovsky@phystech.edu



Где лежит всё по курсу?

Основной ресурс: github.com/miptcv/cv19

- Слайды: [/lectures](#)
- Код с семинаров: [/code](#)
- Заготовки для домашних заданий: [/homework](#)
- Задачи и сроки: [/issues](#)
- Рефераты: github.com/miptcv/wiki_cv

Какого цвета учебник?

Основная:

- Forsyth, Ponce – **CV. A modern approach**, 2004
- Horn – **Robot vision**, 1989
- Shapiro, Stockman – **CV**, 2006
- Goodfellow, Bengio, Courville – **Deep Learning**, 2016

Еще:

- Szelinsky – **CV. Algorithms and applications**, 2011
- Prince – **CV. Models, learning and inference**, 2012
- Hartley, Zisserman – **Multiple view geometry in CV**, 2003
- Сойфер – **Методы компьютерной обработки изображений**, 2003

Программа курса

1. Классическое CV:

- Базовая обработка изображений
- Локальные особенности
- Геометрия
- Цвет и сегментация
- Сжатие изображений
- Работа с видео

2. ML/DL:

- Классификация и регрессия
- Детектирование объектов
- Сегментация
- Style transfer
- Распознавание лиц

3. Kaggle ML контекст

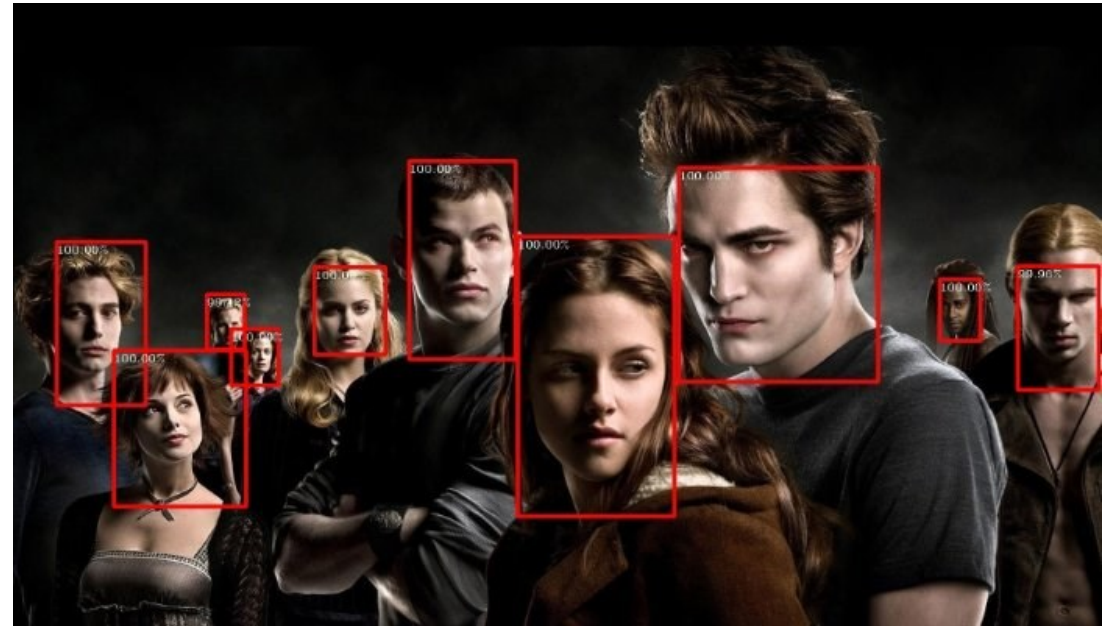
Задачи компьютерного зрения

Задачи компьютерного зрения

- Распознавание текста
- Распознавание документов
- Обнаружение лиц
- Распознавание лиц
- Картография
- Медицина
- Автономные автомобили
- Навигация роботов
- Камеры наблюдения
- Поиск фото/видео
- Виртуальная реальность
- Дополненная реальность
- Интеллектуальная обработка фотографий
- ...

#1. Лица

- Обнаружение
- Отслеживание
- Сравнение двух на совпадение (= распознавание)
- Определение эмоций
- Сегментация
- Фильтрация



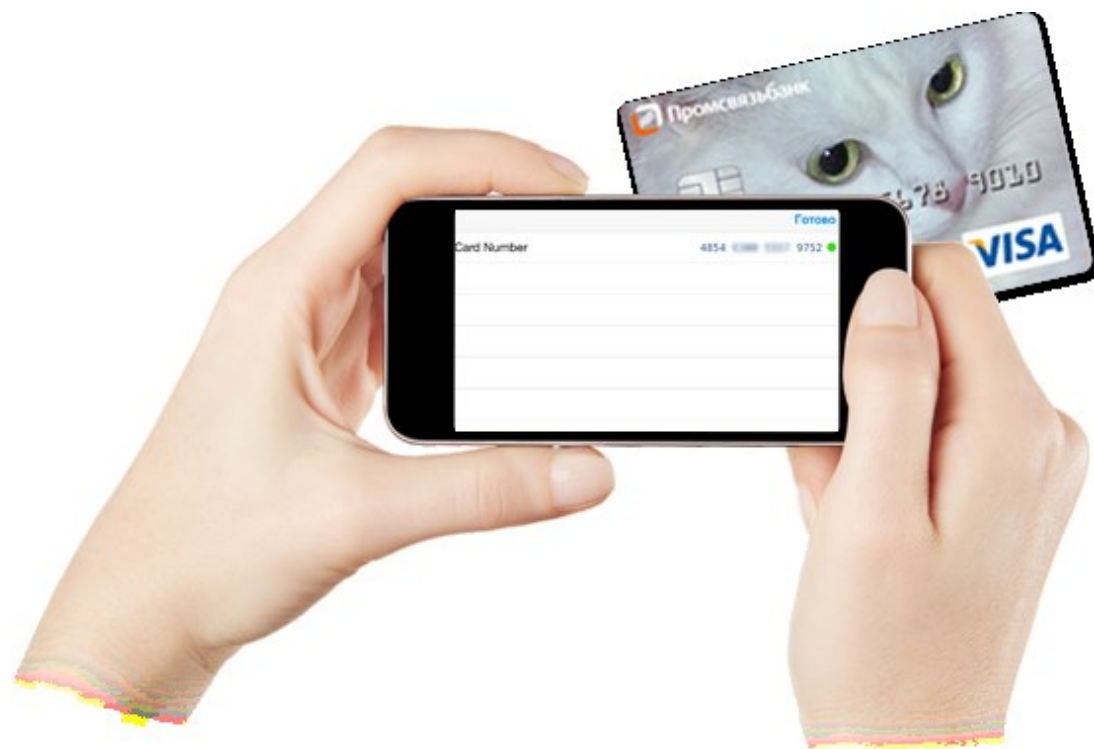
Сегментация



#2. Перенос художественного стиля (Artistic Style Transfer)



#3. Распознавание документов



Распознавание документов

Распознавание текста:

- Печатного
- Рукопечатного
- Рукописного
- Онлайн распознавание
- Текст в сцене

Распознавание документов:

- Жесткие формы
- Гибкие формы
- Паспорта
- Водительские права
- Банковских карты
- Сравнение двух договоров

#4. Автономные автомобили



#5. DeepFake



Анализ дорожной сцены

- Пешеходы
 - Автомобили
 - Велосипедисты
 - Мотоциклисты
 - Животные
 - Препятствия
 - Ямы
 - Дорожное полотно
 - Линии разметки
 - Светофоры
 - Дорожные знаки
 - Сигналы автомобилей
 - Номера автомобилей
 - Согласованность камер с радарам, GPS, ...
- Кто? Tesla, Google, Uber, Mercedes, Mobileye, ...

Формирование цифрового изображения

Цифровое изображение



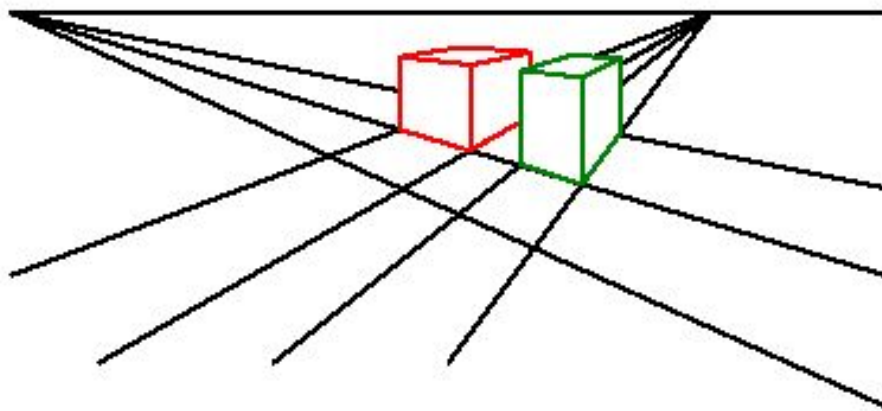
- 3-х мерный тензор $W \times H \times C$
- C – количество каналов (3 для RGB, 1 для полутоновых, 4 – с альфа каналом)
- Тип: для хранения обычно байтовый [0..255]
- для обработки часто float32

Формирование изображения. Геометрия

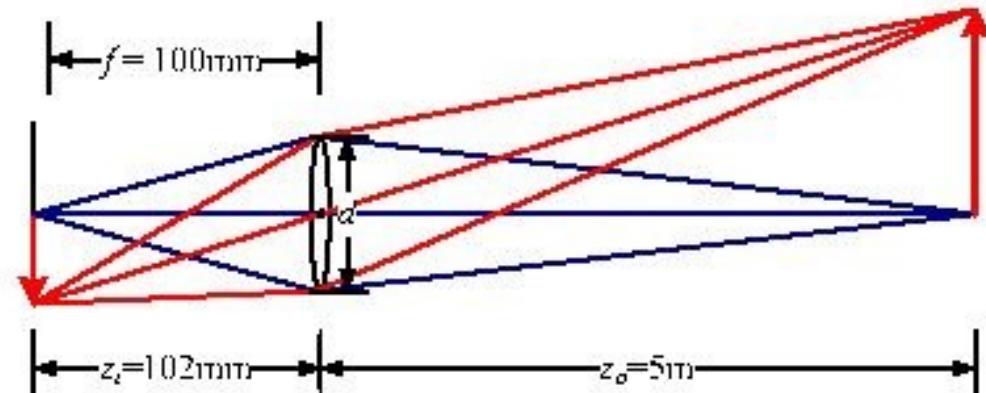
Проективное преобразование:

$$p' = Hp$$

$$H = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{pmatrix}; \quad p = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

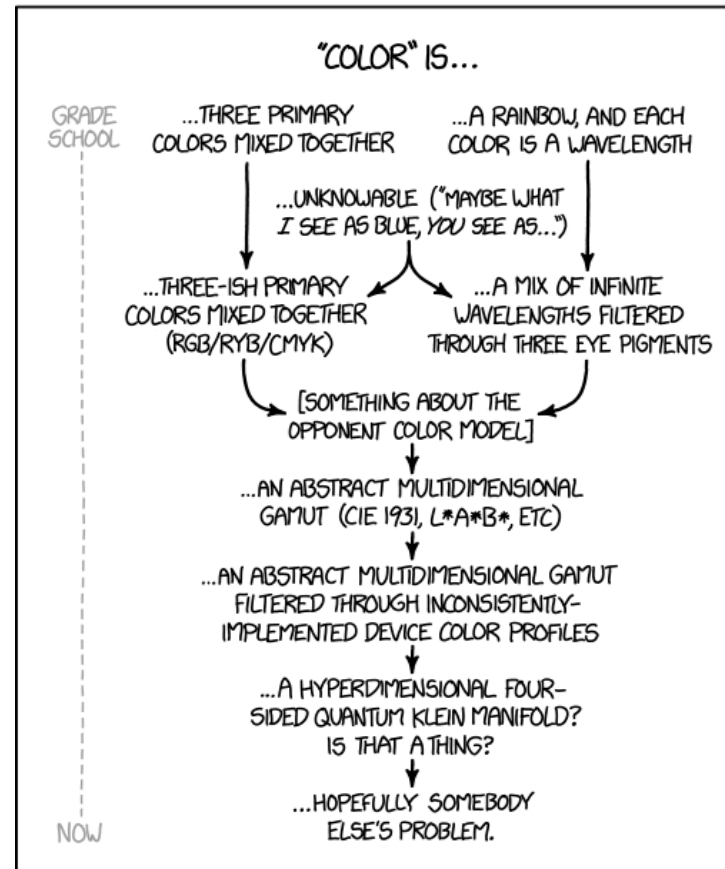


Из оптики: преобразование центральной проекции



Формирование изображения. Цвет

EVOLUTION OF MY UNDERSTANDING OF COLOR OVER TIME:



Формирование изображения. Цвет

Цвет — это общее свойство излучений разного спектрального состава, неразличимых для человека — *Эрвин Шредингер*

$$a_r = \int R(\lambda)L(\lambda)d\lambda \quad a_g = \int G(\lambda)L(\lambda)d\lambda \quad a_b = \int B(\lambda)L(\lambda)d\lambda$$

$R(\lambda)$, $G(\lambda)$, $B(\lambda)$ — функции абсолютной спектральной чувствительности колбочковых приёмников

$L(\lambda)$ — спектральный состав света

$a_{r,g,b}$ — составляющие совокупного действия света на все приёмники глаза

