

Анализ изображений

Лекция#1

Введение
Цифровое изображение
Задачи компьютерного зрения

ФИВТ МФТИ

2021

Лектор

Александр Жуковский

- Заведующий лаборатории компьютерного зрения и обработки естественного языка
- Руководитель группы исследований в области компьютерного зрения, ABBYY
- mailto: zhukovsky+cv21@phystech.edu

Материалы

- Регистрация на курс:

airtable.com/shrpfmaV5zydU750p

- Слайды, материалы семинаров, работа с ДЗ:

github.com/miptcv/cv20

- Лекции:

drive.google.com/drive/u/1/folders/1ZH_t-sNHNS2yqt64OoQ315ObuGJz70qB

- Канал/чат:

t.me/miptcv

Программа курса

1. Классическое CV:

- Базовая обработка изображений
- Границы на изображении
- Локальные особенности, точки и прямые
- Цвет, сегментация и бинаризация
- Геометрия

2. ML/DL:

- Классификация и регрессия
- Детектирование объектов
- Основы нейронных сетей
- Распознавание лиц

3. ML-контекст

Литература

Основная:

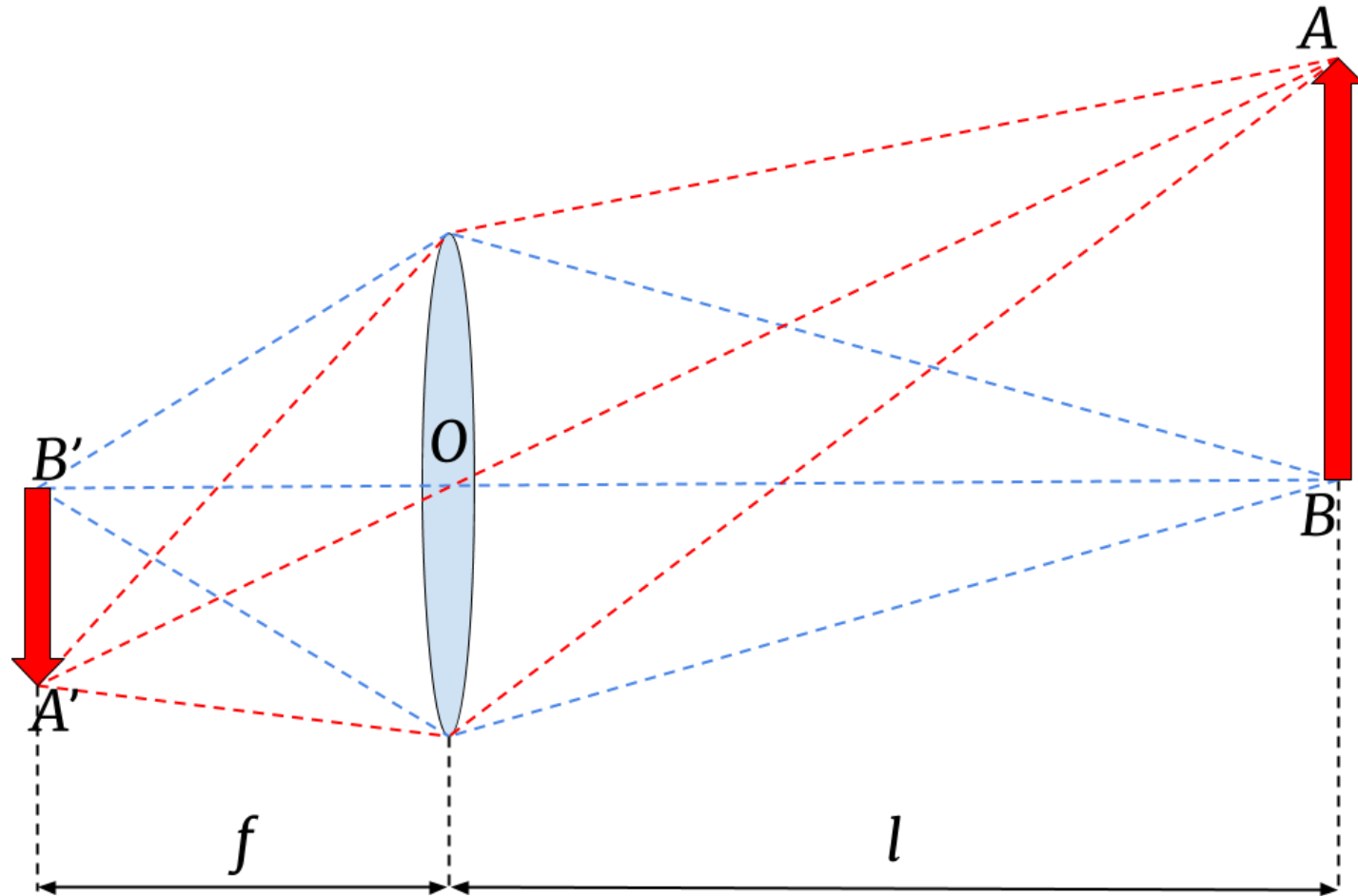
- Forsyth, Ponce – **CV. A modern approach**, 2004
- Szelinsky – **CV. Algorithms and applications**, 2011
- Prince – **CV. Models, learning and inference**, 2012
- Goodfellow, Bengio, Courville – **Deep learning**, 2016

Дополнительная:

- Horn – **Robot vision**, 1989
- Shapiro, Stockman – **CV**, 2006
- Сойфер – **Методы компьютерной обработки изображений**, 2003
- Hartley, Zisserman – **Multiple view geometry in CV**, 2003

Формирование цифрового изображения

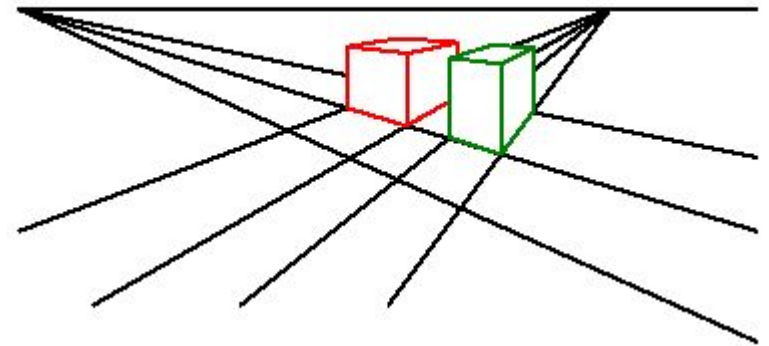
Формирование изображения. Оптика



Формирование изображения. Геометрия. Проективное преобразование

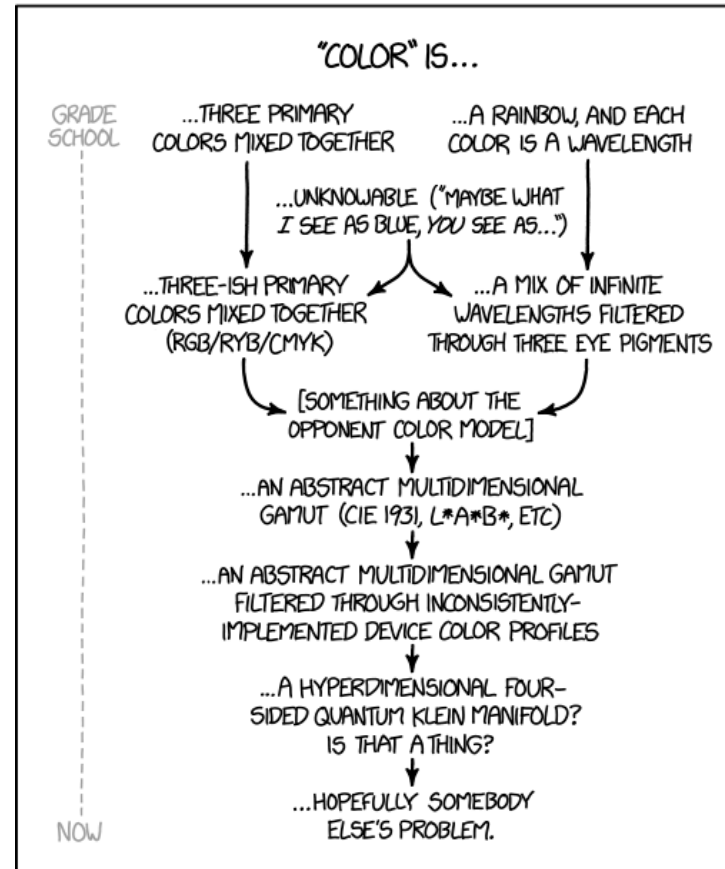
$$p' = Hp$$

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



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 — Э. Шредингер, 1920

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

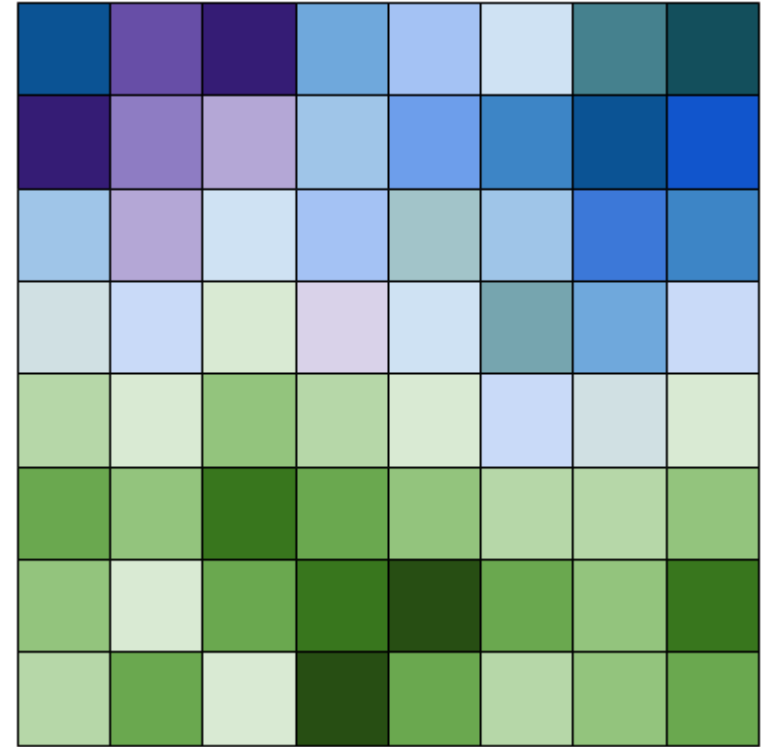
$R(\lambda)$, $G(\lambda)$, $B(\lambda)$ — функции абсолютной спектральной чувствительности колбочковых приёмников

$L(\lambda)$ — спектральный состав света

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

Цифровое изображение

- 2-мерная матрица пикселей (растр): $a(x, y)$
- RGB, 3 значения на пиксель
- обычно – 8 бит: один беззнаковый байт, значения в диапазоне $[0..255]$
- Другие форматы также допустимы, например, float32 и значения со знаком
- Векторное изображение = набор векторных примитивов



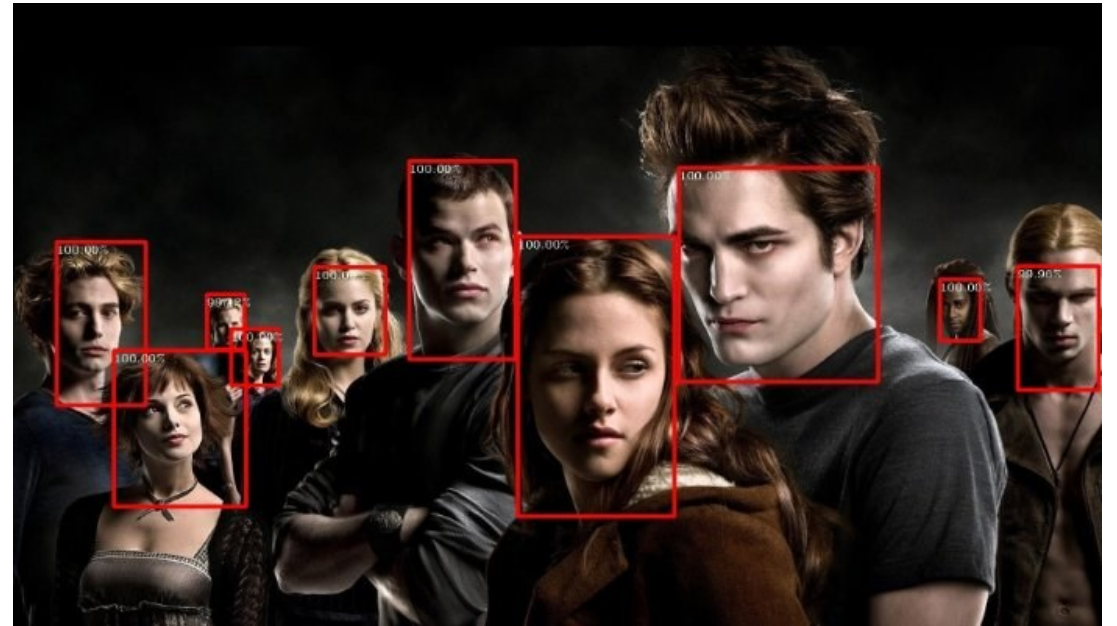
Задачи компьютерного зрения

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 со смартфона



Распознавание документов

Распознавание текста:

- Печатного
- Рукопечатного
- Рукописного
- Онлайн распознавание
- Текст в естественных условиях

Распознавание документов:

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

#4. Автономные автомобили



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, ...

