

# Old School Computer Vision

Калинов Алексей, ПМИ152

# Содержание

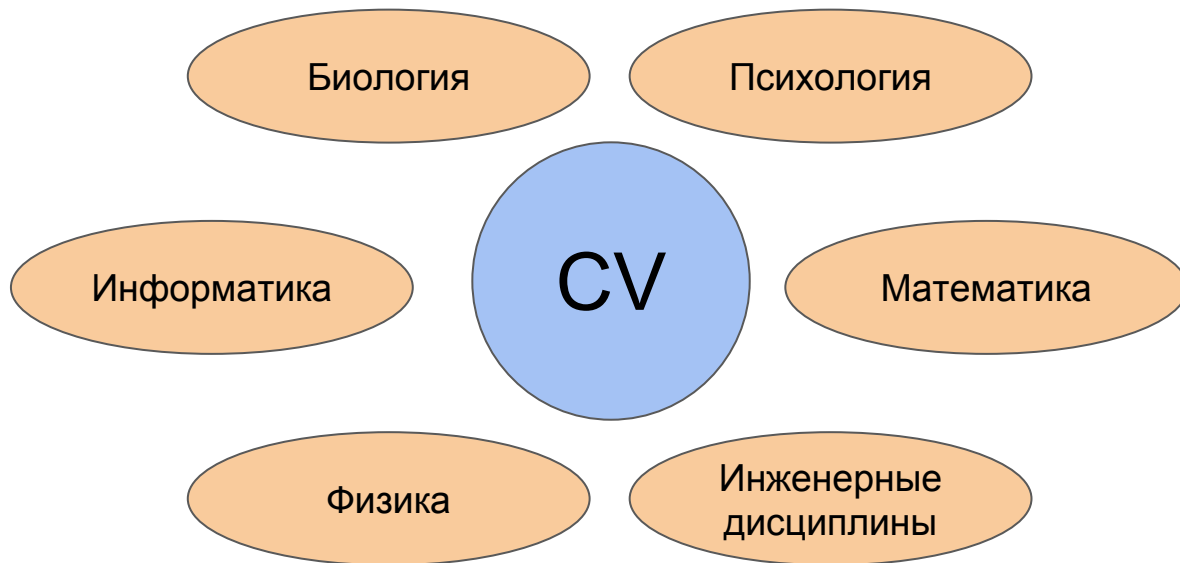
1. Определение и задачи
2. Сложности
3. Датасеты
4. Методы решения задач
  - a. Предобработка изображений
  - b. Выделение признаков

# Содержание

1. Определение и задачи
2. Сложности
3. Датасеты
4. Методы решения задач
  - a. Предобработка изображений
  - b. Выделение признаков

# Что такое компьютерное зрение?

- Междисциплинарная область науки, занимающаяся автоматическим получением информации из изображений и видео.



# Задачи: Распознавание



Изображение находится в общественном достоянии.

Обнаружение  
объекта



Сегментация  
изображения



Распознавание  
объектов

Q534 PES

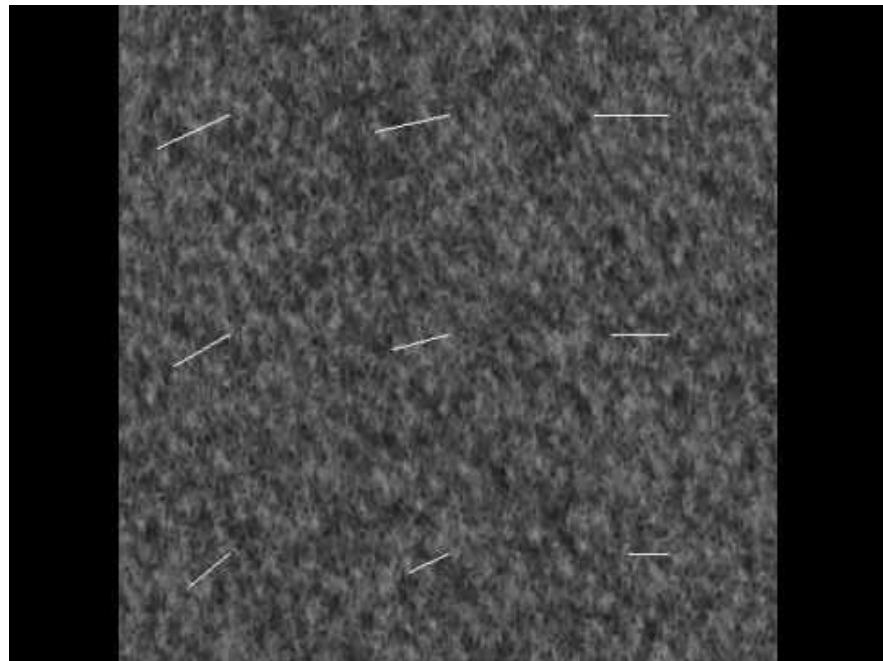
# Задачи: Анализ движения

## Слежение за объектами



Видео создано [Thijs Zumbrink](#) и распространяется под лицензией [CC BY 3.0](#).

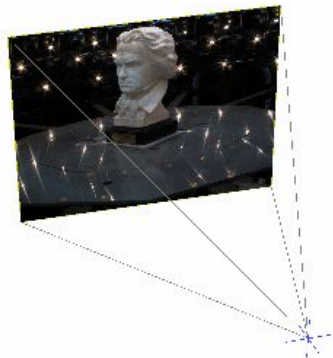
## Поиск оптического потока



Видео создано [Tomas Novak](#) и распространяется под лицензией [CC BY 3.0](#).

# Задачи: Восстановление и реконструкция

Реконструирование 3D объектов и сцен



Изображение создано [Martin Oswald](#), [Maria Klodt](#), [Jörg Stückler](#), [Prof. Dr. Daniel Cremers](#).

Восстановление изображения



Изображение взято из [1] под лицензией [ACM Copyright License](#).

# Содержание

1. Определение и задачи
2. Сложности
3. Датасеты
4. Методы решения задач
  - а. Предобработка изображений
  - б. Выделение признаков



# Сложности: Семантическая разница



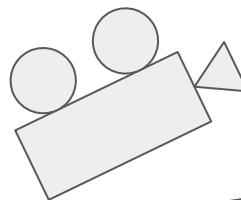
Люди воспринимают объекты

```
image =  
[  
  [57, 35, ... 26],  
  [26, 11, ... 83],  
  ...  
  [109, 47, ... 83],  
  ]
```

В компьютере цифровые сигналы  
представлены числами

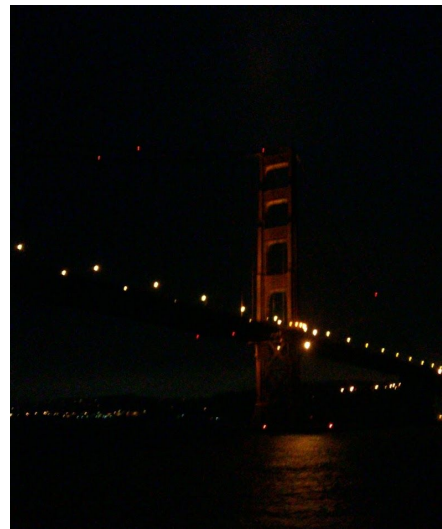
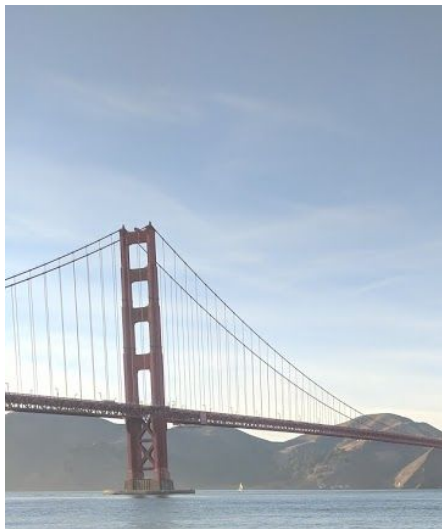
# Сложности: Вариативность ракурса

Все пиксели меняют своё значение.



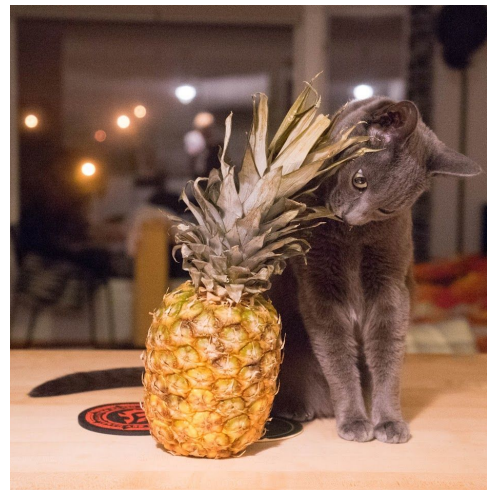
# Сложности: Вариативность освещения

Все пиксели меняют своё значение.



# Сложности: Вариативность формы объекта

Все пиксели меняют своё значение.



# Сложности: Окклюзия

Часть объекта может быть скрыта.



Изображение находится в общественном достоянии.

# Содержание

1. Определение и задачи
2. Сложности
3. Датасеты
4. Методы решения задач
  - а. Предобработка изображений
  - б. Выделение признаков

# Датасеты: MNIST



[LeCun et al., 1998a]

Рукописные цифры.

60 000 изображений для обучения, 10 000 изображений для теста.



# Датасеты: ImageNet



Изображение взято из [блога](#) Andrej Karpathy "What I learned from competing against a ConvNet on ImageNet".

Фотографии объектов на разные темы.  
14 197 122 изображений, ~22 000 классов.



# Датасеты: YouTube 8M

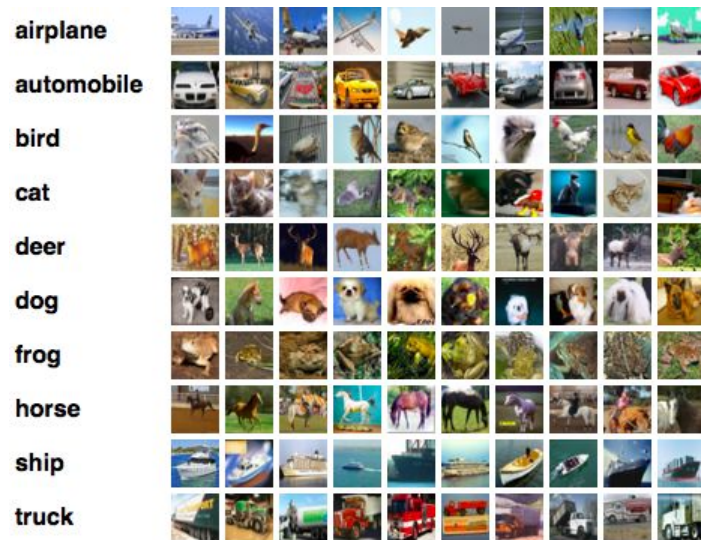


Изображение взято из [блога](#) Google Research Blog "An updated YouTube-8M, a video understanding challenge, and a CVPR workshop. Oh my!".

Видео на разные темы.

3,2 млрд признаков для 7 млн видео, 4716 классов из Knowledge Graph.

# Датасеты: CIFAR



Изображение взято из [4].

60 000 изображений 32x32.

**CIFAR-10**: 10 классов, **CIFAR-100**: 100 классов.

# Содержание

1. Определение и задачи
2. Сложности
3. Датасеты
4. Методы решения задач
  - a. Предобработка изображений
  - b. Выделение признаков

# Решение задач в общих чертах.

1. Предобработка изображений.
2. Выделение признаков.
3. Применение систем, получающих из признаков ответ.

# Предобработка изображений: ЦВЕТОВЫЕ СХЕМЫ

## Grayscale



Изображение находится в общественном достоянии.

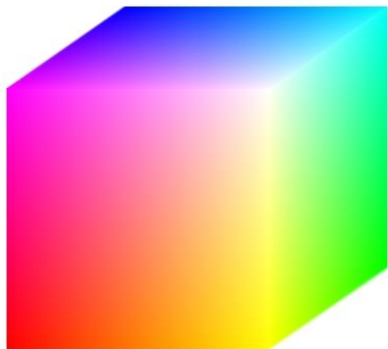


Существуют разные способы перевести цветную картинку в чёрно-белую.

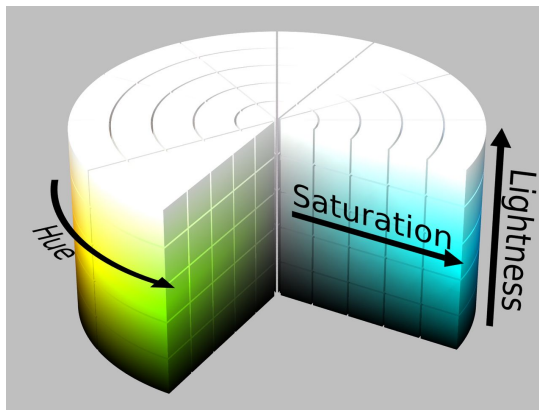
$$\text{HDTV: } Y = 0.2126R + 0.7152G + 0.0722B$$

# Предобработка изображений: ЦВЕТОВЫЕ СХЕМЫ

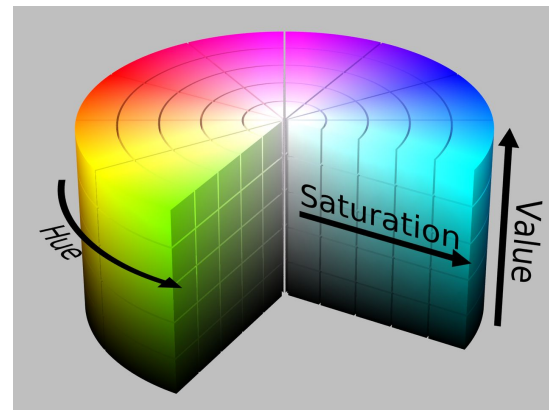
HSL, HSV



RGB



HSL



HSV

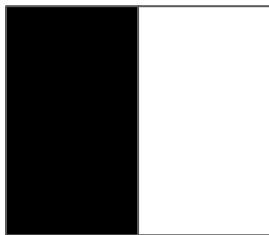
[Изображение](#) создано SharkD и распространяется под CC BY-SA 3.0. [Изображение](#) создано SharkD и распространяется под CC BY-SA 3.0.

# Предобработка изображений: нормализация

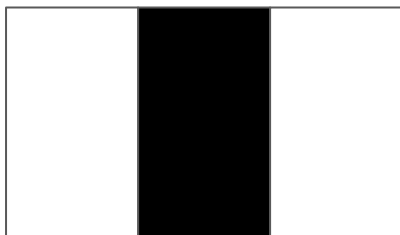
Переход из одного цветового пространства в другое посредством сдвига и растяжения.

Очень часто переходят к отрезку  $[-1, 1]$ .

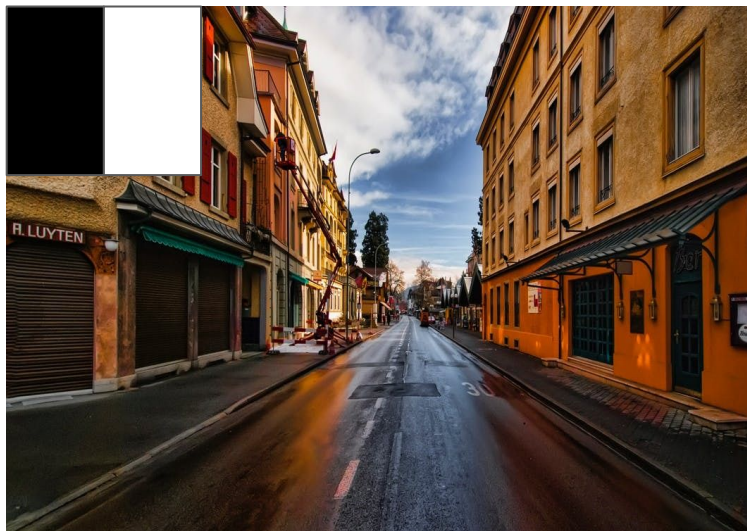
# Признаки: Хаара



2-прямоугольные



3-прямоугольные



Разность сумм пикселей в  
прямоугольниках

[  
[57, 35, ... 26],  
[26, 11, ... 83],  
...  
[109, 47, ... 83],  
]



# Признаки: Хаара (продолжение)

## **Плюсы:**

1. Быстрота вычислений.

## **Минусы:**

1. Слабый признак.
2. Медленное вычисление для повернутых прямоугольных областей.

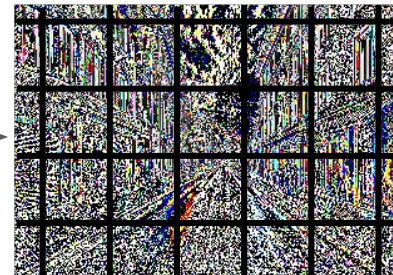
# Признаки: Гистограмма направленных градиентов



Исходное изображение

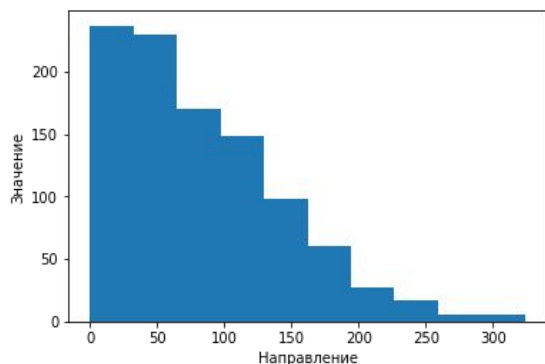


Горизонтальный и  
вертикальный градиенты

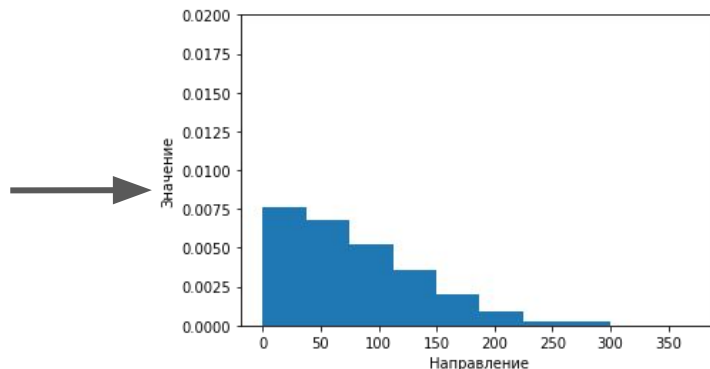


Разбиение на ячейки

# Признаки: Гистограмма направленных градиентов (продолжение)



Гистограмма  
градиентов для  
каждой ячейки



Нормализованная гистограмма  
по блоку, состоящему из  
нескольких ячеек

A light gray box containing a vertical list of values representing a concatenated vector of normalized histograms. The values are: 0.57, 0.35, an ellipsis (...), 0.13, and 0.26. An arrow points from the normalized histogram to this box.

[0.57,  
0.35,  
...  
0.13,  
0.26]

Вектор  
сконкатенированных  
гистограмм

# Признаки: Гистограмма направленных градиентов (продолжение)

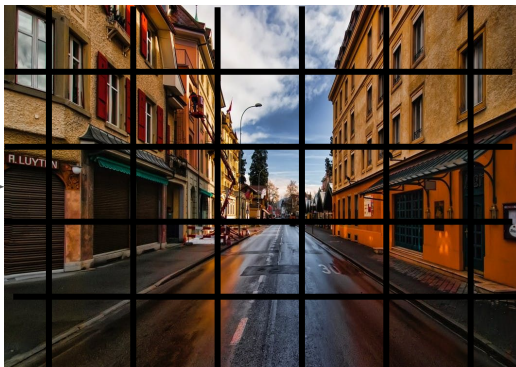
## Плюсы:

1. Локальность и инвариантность относительно геометрических преобразований.
2. Относительная быстрота вычислений.

## Минусы:

1. Не инвариантен относительно ориентации объекта.

# Признаки: Локальные бинарные шаблоны (LBP)



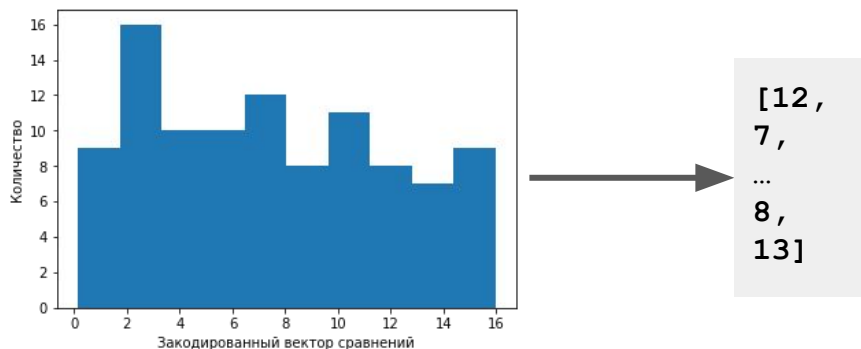
```
[  
[0, 1, 0, 0],  
[0, 1, 1, 0],  
...  
[1, 1, 0, 1]  
]
```

Исходное изображение

Разбиение на ячейки

Сравнение с соседними  
пикселями для каждого  
пикселя ячейки

# Признаки: Локальные бинарные шаблоны (LBP) (продолжение)



Гистограмма векторов  
сравнений по ячейке

Вектор сконкатенированных  
гистограмм

# Признаки: Локальные бинарные шаблоны (LBP) (продолжение)

## Плюсы:

1. Локальность и инвариантность относительно геометрических преобразований.
2. Относительная быстрота вычислений.

## Минусы:

1. Не инвариантен относительно ориентации объекта.

# Заключение

- Компьютерное зрение строится на результатах нескольких областей науки и пытается решить различные задачи получения информации из изображений и видео.
- Из-за большой семантической разницы задачи компьютерного зрения становятся сложными.
- Для многих задач собраны и обработаны данные, что облегчает тестирование методов.
- Предобработка важна; большинство методов извлечения признаков стараются использовать локальность для увеличения инвариантности.



# Библиография

- [1] Bertalmio, M., Sapiro, G., Ballester, C., and Caselles, V. 2000. Image inpainting. In Proceedings of ACM SIGGRAPH 2000, 417– 424.
- [2] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. "Gradient-based learning applied to document recognition." Proceedings of the IEEE, 86(11):2278-2324, November 1998.
- [3] Olga Russakovsky\*, Jia Deng\*, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg and Li Fei-Fei. (\* = equal contribution) ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision, 2015.

# Библиография (продолжение)

[4] Krizhevsky, Alex. (2012). Learning Multiple Layers of Features from Tiny Images. University of Toronto.

[5] Dalal, N and Triggs, B, Histograms of Oriented Gradients for Human Detection, IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2005

# Бонусные слайды

# Датасеты: Playboy Lena



Изображения взяты с [официальной страницы](#) истории изображения.

“Стандарт”, используемый для оценки обработки изображений.

# На самом деле, у всех есть мечта!



Let's Enhance: <https://www.youtube.com/watch?v=Vxq9yj2pVWk>