

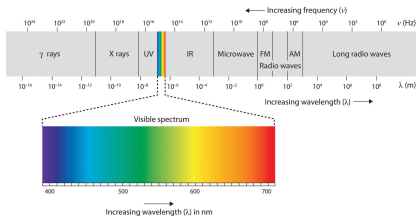
Изображения в компьютерном зрении

Основные концепции и определения

Диц Даниил Денисович

НТУ Сириус

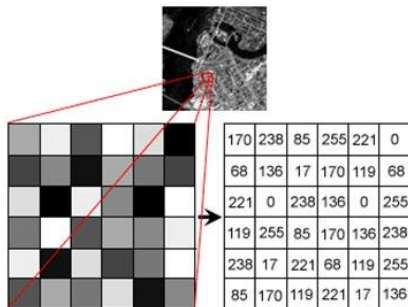
7 марта 2025 г.



Что такое изображение?

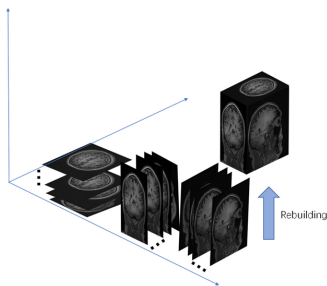
- ▶ Изображение — это визуальное представление объекта, сцены, человека или концепции.
- ▶ Оно может быть фотографией, рисунком, схемой, сканом и т.д.
- ▶ Изображение — это также функция, а именно n -мерная функция.
- ▶ В двумерном случае ($n = 2$) изображение описывается функцией $F(X, Y)$, где X и Y — пространственные координаты.

Пространственные координаты



- ▶ Амплитуда F в точке (x_i, y_i) — это интенсивность или уровень серого.
- ▶ Интенсивность определяет восприятие света и тени.
- ▶ Координаты (x_i, y_i) называются пикселями (элементами изображения).

Типы изображений



- ▶ Функция $F(X, Y)$ зависит от источника освещения и отражения объектом света
- ▶ **3D изображения:** $F(X, Y, Z)$, где (x_i, y_i, z_i) — воксели.
- ▶ Примеры 3D изображений: медицинские сканы, МРТ, микроскопия.
- ▶ **Цветовые каналы:** $F_{\text{red}}(X, Y)$, $F_{\text{green}}(X, Y)$, $F_{\text{blue}}(X, Y)$.
- ▶ Интенсивность цвета: от 0 (нет цвета) до 255 (максимальная интенсивность).

Специальные типы изображений

Base



Binary Mask



RGB Mask



- ▶ **Меченые изображения:** Пиксели имеют метки (например, 1 — передний план, 0 — фон).
- ▶ **Бинарные изображения (маски):** Только две метки.
- ▶ **4D и 5D изображения:** Добавляются время и каналы (например, 3D + время + каналы).

HSV-цветовое пространство

Определение

HSV (Hue, Saturation, Value) - альтернативное представление цвета, где:

- ▶ **Hue (Оттенок)** - основной цвет (0° - 360°)
- ▶ **Saturation (Насыщенность)** - интенсивность цвета (0-100%)
- ▶ **Value (Яркость)** - яркость цвета (0-100%)

Математическое преобразование из RGB:

$$V = \max(R, G, B)$$
$$S = \begin{cases} \frac{V - \min(R, G, B)}{V} & \text{if } V \neq 0 \\ 0 & \text{otherwise} \end{cases}$$
$$H = \begin{cases} 60^{\circ} \cdot \frac{G - B}{V - \min(R, G, B)} & \text{if } V = R \\ 60^{\circ} \cdot \left(2 + \frac{B - R}{V - \min(R, G, B)} \right) & \text{if } V = G \\ 60^{\circ} \cdot \left(4 + \frac{R - G}{V - \min(R, G, B)} \right) & \text{if } V = B \end{cases}$$

Преимущества HSV и пример кода

Преимущества перед RGB

- ▶ Отделение цветовой информации от освещения
- ▶ Более интуитивное управление цветом
- ▶ Упрощение задач цветовой сегментации

```
hsv_image = cv.cvtColor(rgb_image,  
                        cv.COLOR_RGB2HSV)
```

```
rgb_restored = cv.cvtColor(hsv_image,  
                           cv.COLOR_HSV2RGB)
```

```
lower_red = np.array([0, 100, 100])  
upper_red = np.array([10, 255, 255])  
mask = cv.inRange(hsv_image, lower_red, upper_red)
```

Изображения vs Видео

- ▶ **Изображение:** Статичное представление в один момент времени.
- ▶ **Видео:** Последовательность изображений с временной компонентой $F(X, Y, T)$.
- ▶ Частота кадров (FPS) создает иллюзию движения.

Изображения vs Табличные данные

- ▶ **Табличные данные:** Размерность определяется количеством признаков.
- ▶ **Изображения:** Размерность определяется пространственными координатами.
- ▶ **Извлечение признаков:** Для изображений — традиционная предварительная обработка или глубокое обучение.
- ▶ **Предварительная обработка:** Для изображений — изменение размера и нормализация интенсивности.

Ключевые различия между типами данных

№	Характеристика	Изображение	Видео	Аудио
Табличные данные				
Тип	Один момент времени	Последовательность изображений	Один момент времени	Структурированные данные
Представление	2D массив пикселей	3D массив кадров	1D массив сэмплов	2D массив признаков
Форматы	JPEG, PNG, RAW	MP4, AVI, MOV	WAV, MP3, FLAC	CSV, Excel
Аугментация	Поворот, кадрирование	Временные изменения	Добавление шума	SMOTE, ADASYN
Извлечение признаков	Границы, текстуры, цвета	Границы, текстуры, цвета, оптический поток	Спектрограмма, MFCC	Статистический анализ
Модели	CNN	RNN, 3D CNN	CNN, RNN	Регрессия, Деревья решений
Задачи	Классификация, сегментация	Распознавание действий, трекинг	Распознавание речи, идентификация	Регрессия, классификация
Вычислительная стоимость	Меньше	Больше	Умеренная	Меньше
Приложения	Распознавание лиц	Интерпретация языка жестов	Голосовые ассистенты	Прогнозирование, обнаружение мошенничества