

Дескрипторы HOG и SIFT

Как считать градиент

- Считаем градиент изображения
 - Оператор Собеля:

$$\mathbf{G}_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} * \mathbf{A} \quad \text{and} \quad \mathbf{G}_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * \mathbf{A}$$

- Одномерная дифференцирующая маска:

$$G_x = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} * A \quad G_y = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} * A$$

- Можно перед подсчетом градиента применить фильтр Гаусса

$$G_{(x,y,\sigma)} = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}$$

$$G = \sqrt{G_x^2 + G_y^2}$$

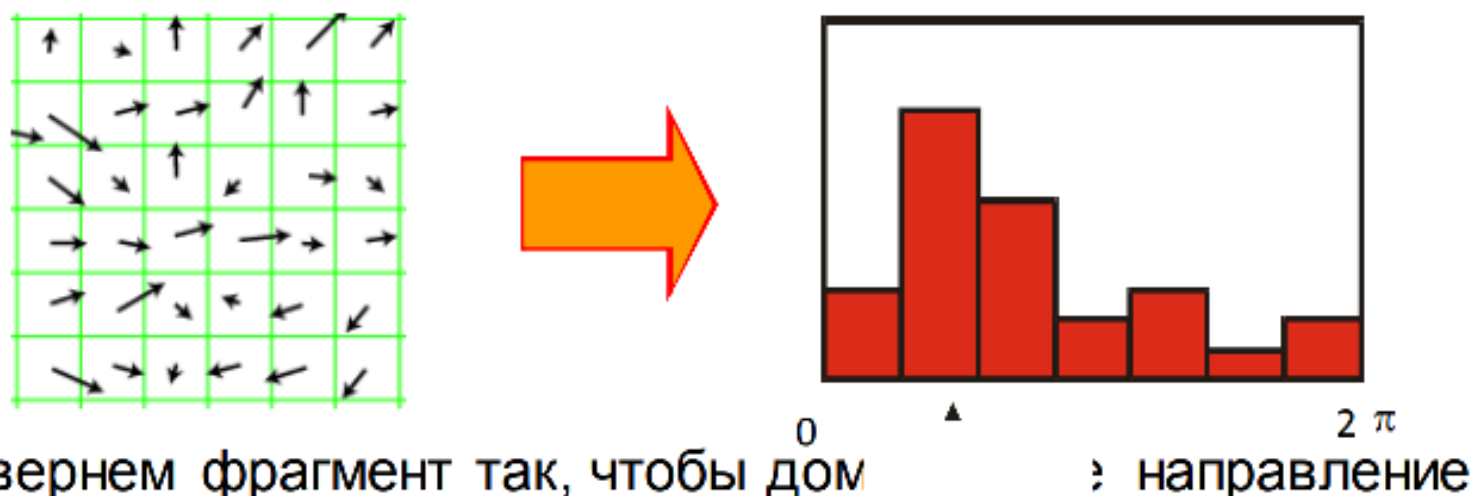
- Модуль градиента

$$\Theta = \arctan\left(\frac{G_y}{G_x}\right)$$

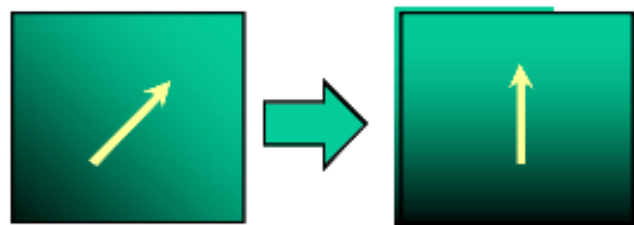
- Направление градиента

- Мы научились выделять ключевые точки.
- Если точка является ключевой, нужно вычислить её ориентацию
- Окрестность ключевой точки играет роль.
- Гистограмма состоит из 36 компонент, которые равномерно покрывают промежуток в 360 градусов, и формируется она следующим образом: каждая точка окна (x, y) вносит вклад, равный $m * G(x, y, \sigma)$, в ту компоненту гистограммы, которая покрывает промежуток, содержащий направление градиента
- $G(x, y, \sigma)$ – двумерный гауссиан, m – модуль градиента

- Идея: найти основное (доминантное) направление градиентов пикселей в окрестности точки
- Выберем в гистограмме ячейку с максимальным значением, возьмём это направление как доминирующее



- Повернем фрагмент так, чтобы дом градиента было направлен вверх



- Если локальных максимумов несколько – считаем, что несколько точек с разной ориентацией

- Как искать «доминирующее» направление?
- Берем бин с максимальным значением гистограммы
- Берем два соседних бина гистограммы и интерполируем параболу. Есть 3 точки – нужно задать параболу. После задания параболы находим ее максимум, тогда припишем ключевой точке эту характеристику – направление и величину, а так же поворачиваем окрестность ключевой точки
- Это нам дает инвариантность относительно поворотов

SIFT-дескриптор

- Слева пиксели и их градиенты
- В центре ключевая точка
- Круг – радиус, из которого берутся точки
- Чем дальше пиксель, тем меньше он делает вклад. Это достигается умножением на гауссиан

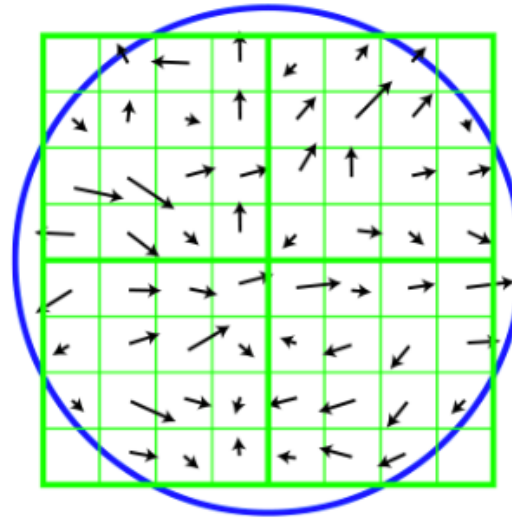
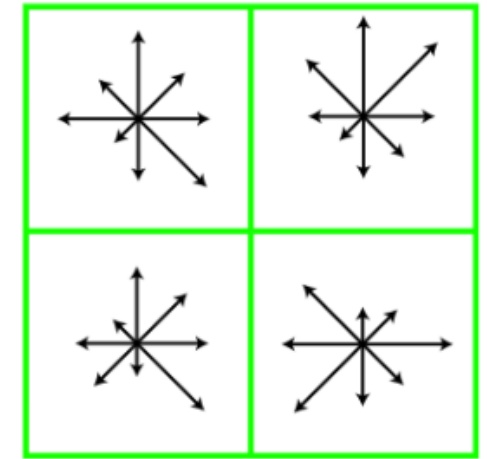


Image gradients



Keypoint descriptor

- Делим окружность на 4 части, в каждой из которой будем накапливать гистограммы на n направлений

HOG

- Сначала считаем градиент изображения одним из предыдущих методов
- Делим на ячейки
- Накапливаем гистограммы
- Объединяем ячейки в блоки
- Нормируем блоки
- Возвращаем массив чисел



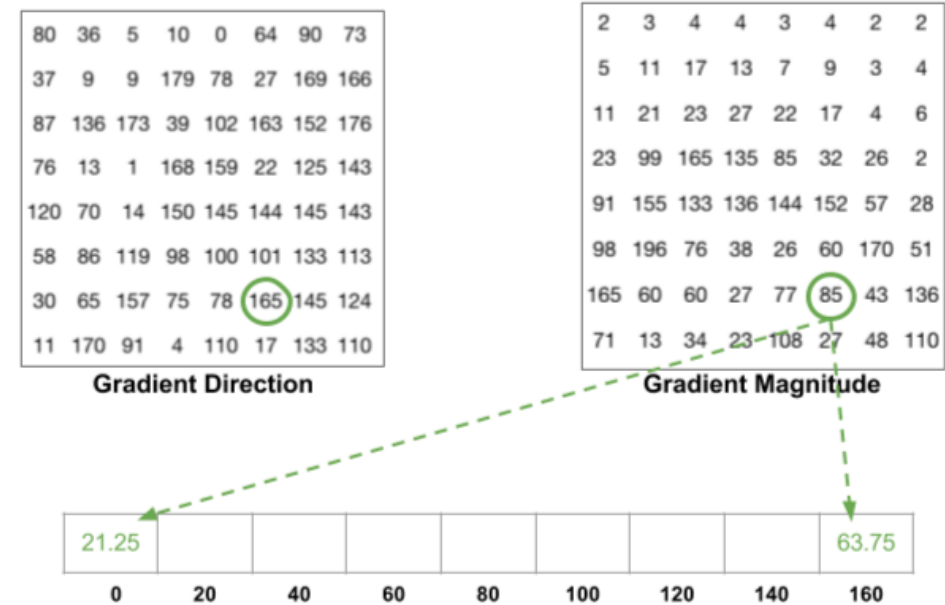
**Left : Absolute value of x-gradient. Center : Absolute value of y-gradient.
Right : Magnitude of gradient.**

- Разбиваем изображение на ячейки разм $m \times n$
- В каждой ячейке накапливаем гистограммы по направлениям градиентов



Как накапливать гистограммы?

- Идем по всем пикселям, считаем направление и модуль градиента
- Выбираем два направления(бины) гистограммы, между которыми лежит наш пиксель
- Этот пиксель добавляет модуль градиента к соседним бинам с некоторыми весами.
- Чем дальше от бина направление, тем меньше вес

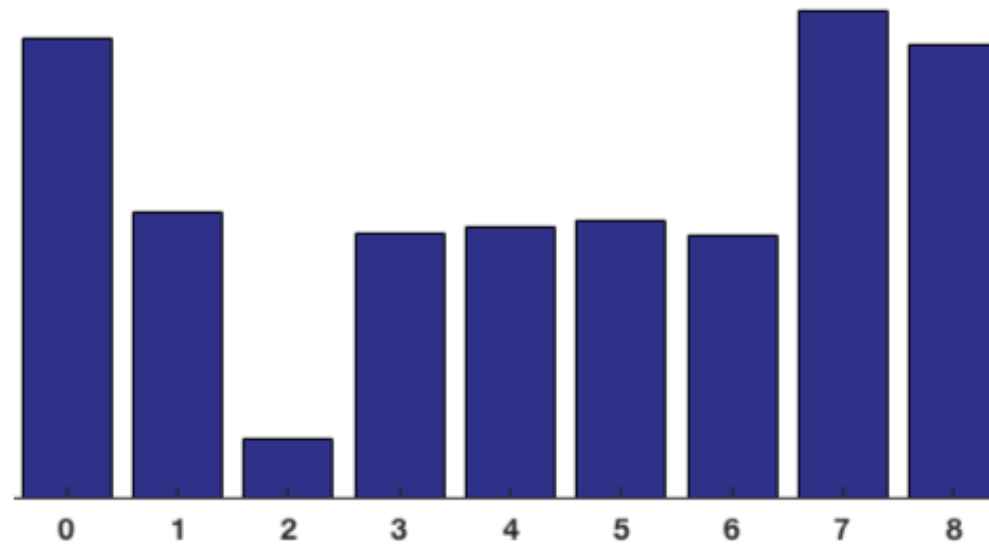


К бину с направлением 0 (или 180 – одно и то же)

добавляе $\frac{165 - 160}{20} \cdot 85$ ина

К бину с направлением 160 добавляется велич $\frac{180 - 165}{20} \cdot 85$

В итоге одна ячейка выглядит следующим образом:
массив из чисел, каждое из которых, это сумма всех
градиентов по данному направлению



В данном примере 8 направлений, которые
равномерно
разбивают промежуток $[0, 180]$ или $[0, 360]$

Деление ячеек на блоки

- Ячейки объединяются в блоки разм $k \times l$
- Блоки идут с некоторым шагом, который нужно задать
- Шаг может быть, например, размеру ячейки, тогда блоки пересекаются
- Если шаг равен размеру блока, то блоки не пересекаются

Нормализация блоков

- После объединения ячеек в блоки каждый блок описывается несколькими векторами $k \times l$, а именно k . Вытянем эти вектора в один большой вектор и нормируем его

$$f = \frac{v}{\sqrt{\|v\|_2^2 + e^2}}$$

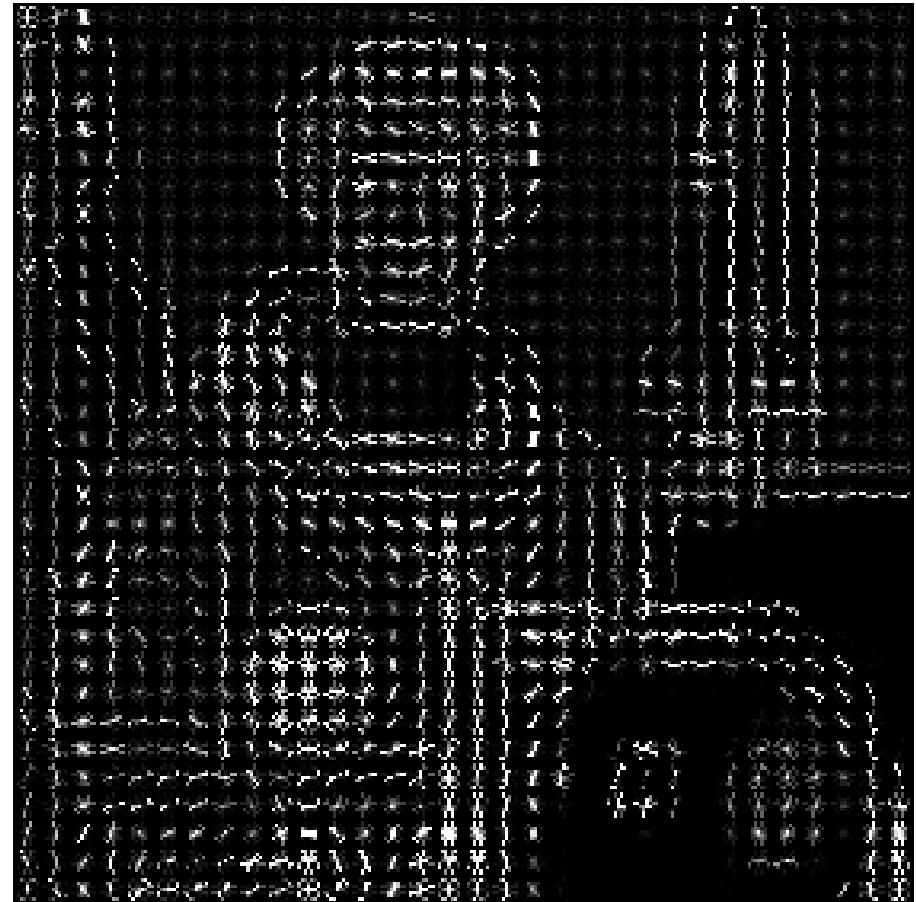
- L2 нормализация
- L2-Nys нормализация: ограничиваем L2-норму сверху некоторым числом, а потом снова нормируем

Визуализация

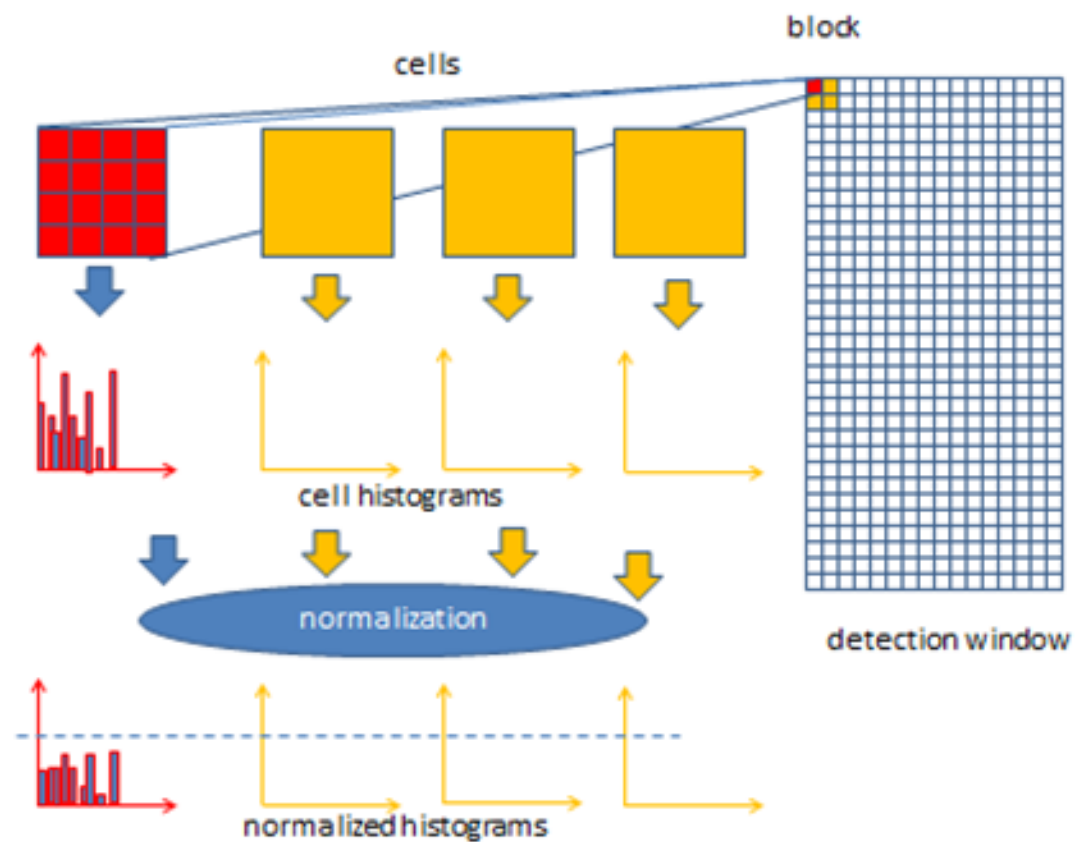
Input image



Histogram of Oriented Gradients



Весь дескриптор в одной картинке



Вывод

- Детекторы SIFT и HOG устойчивы к изменению освещения
- SIFT устойчив к поворотам
- Используются для распознавания пешеходов, машин