

Нижегородский
Государственный
Технический Университет
им. Р.Е. Алексеева

МОДЕЛЬ И АЛГОРИТМЫ ОБНАРУЖЕНИЯ ОБЪЕКТОВ НА ИЗОБРАЖЕНИИ С ИСПОЛЬЗОВАНИЕМ ГЛОБАЛЬНОГО ПРИЗНАКОВОГО ОПИСАНИЯ

Выполнил:

студент группы М17-ИВТ-3

Смирнов Александр Вадимович

Научный руководитель:

Гай Василий Евгеньевич

Нижний Новгород, 2019г.

Цель работы и задачи исследования

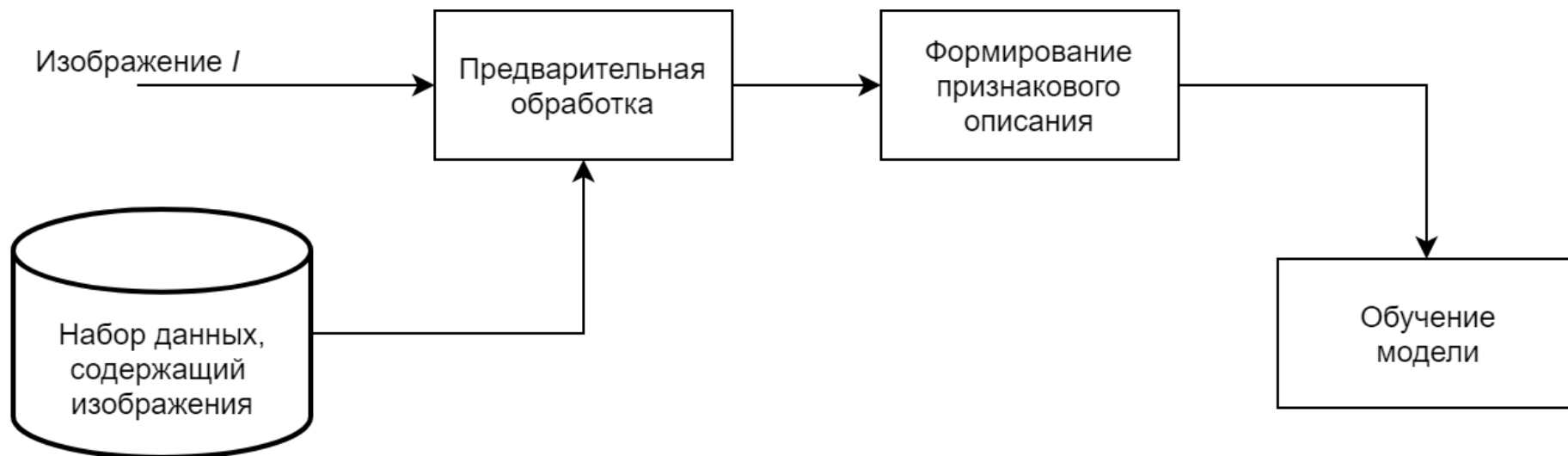
- Цель: разработка и исследование новых моделей и алгоритмов решения задачи обнаружения объектов на изображении с использованием глобального признакового описания.
- Задачи:
 - Обзор и анализ существующих известных методов решения задачи обнаружения объектов с использованием глобального признакового описания
 - Создание информационной модели описания объекта с использованием глобального признакового описания.
 - Создание нового алгоритма формирования глобального признакового описания изображения
 - Проведение исследования с целью выявления наилучшей комбинации параметров разрабатываемой системы
 - Проведение вычислительного эксперимента для установления корректности работы созданных моделей и алгоритмов.

Этапы решения задачи

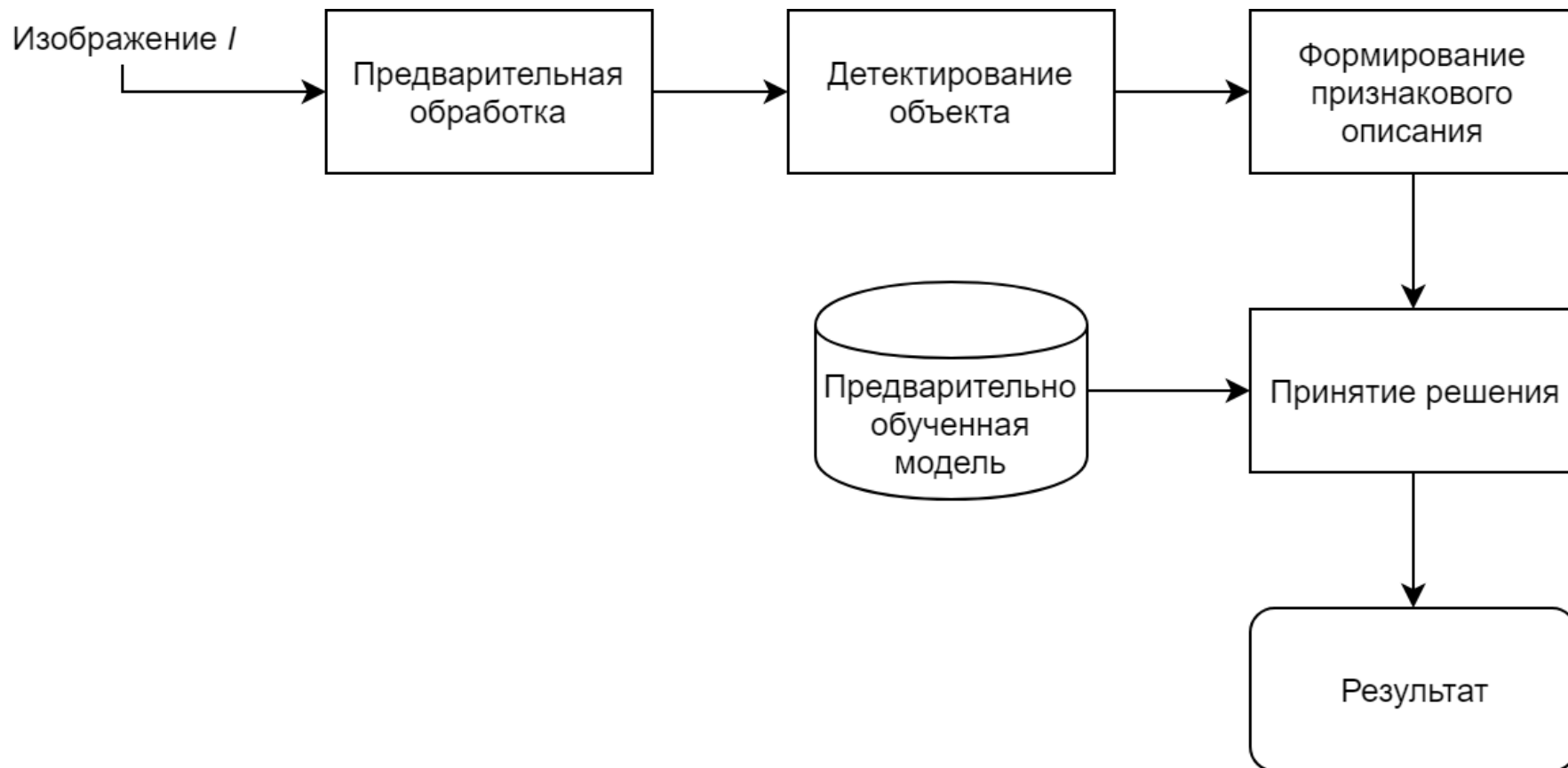
- Предварительная обработка изображения объекта
- Формирование глобального признакового описания объекта
- Локализация объекта на изображении (принятие решения)



Информационная модель этапа обучения



Информационная модель этапа применения



Набор данных

Шахуро В., Конушин А. Российская база изображений автодорожных знаков.
Компьютерная оптика, 2016

Данная база содержит:

- 179138 размеченных кадров с авторегистратора, содержащих дорожные знаки
- 156 типов дорожных знаков
- 104358 изображений знаков (полученных путем кадрирования)

Пример изображения объекта



1



2



3



4

1. Изображение объекта при благоприятной погоде и хорошем освещении
2. Изображение объекта при неблагоприятной погоде и хорошем освещении
3. Изображение объекта при благоприятной погоде и плохом освещении
4. Изображение объекта при неблагоприятной погоде и плохом освещении

Предварительная обработка

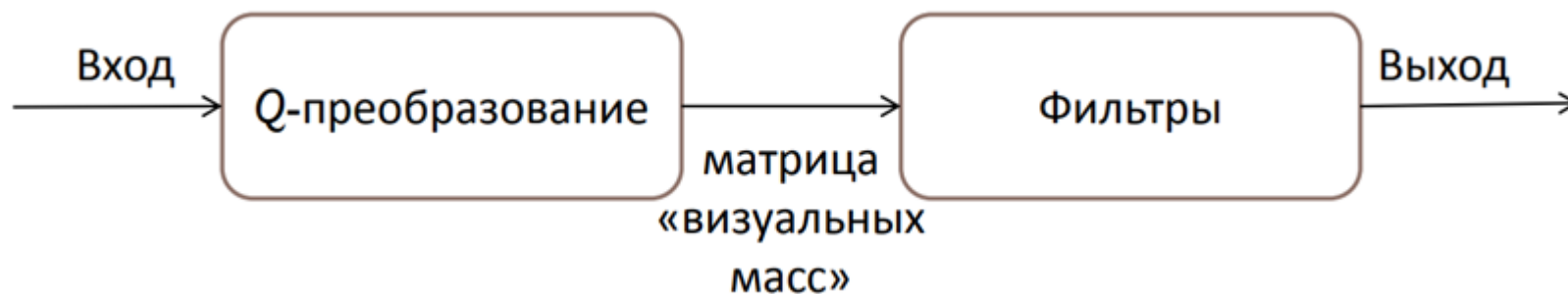
- Изображения представлены в формате RGB
- Преобразование изображения в функцию яркости:

$$I(x, y) = \frac{(R_{xy} + G_{xy} + B_{xy})}{3}, x \in X, y \in Y$$

- Применение фильтра Гаусса:

$$G(u, v) = \frac{1}{2\pi r^2} e^{-\frac{(x^2 + y^2)}{2\sigma^2}}$$

Элементы ТАВ



0.1	0.1	0	0.3	0.1	0.1	0	0
0.2	0.2	0	0	0.2	0.2	0	0.7
0	0	0.8	0.7	0	0.2	0.8	0.7
0.1	0	0.9	1	0	0	0.9	1
0.1	0.1	0	0	0.1	0.1	0	0
0.2	0.2	0	0	0.2	0.2	0	0
0	0	0.8	0.7	0	0	0.8	0.7
0	0.9	0.9	1	0	0	0.9	1

Матрица I
размером 8×8 ($N \times M$) пикселей

Q-преобр.
→

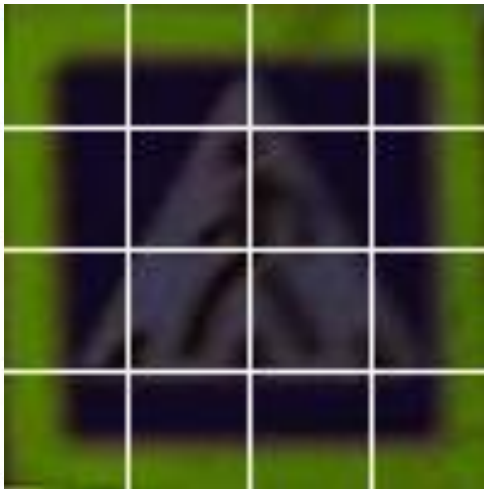
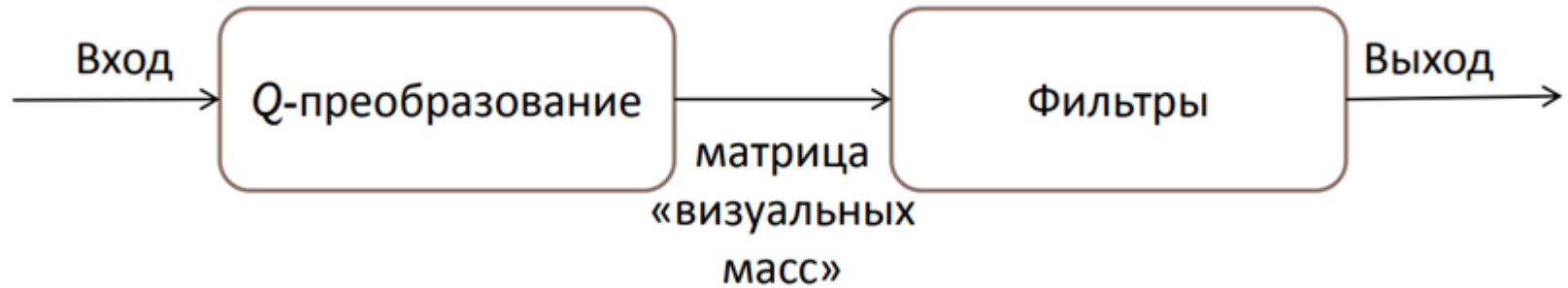
$$B_{kl} = \sum_{i=1}^2 \sum_{j=1}^2 A_{kl}[i, j]$$

$k = \overline{1,4} \quad l = \overline{1,4}$

0.6	0.3	0.6	0.7
0.1	3.4	0.2	3.4
0.6	0	0.6	0
0.9	3.4	0	3.4

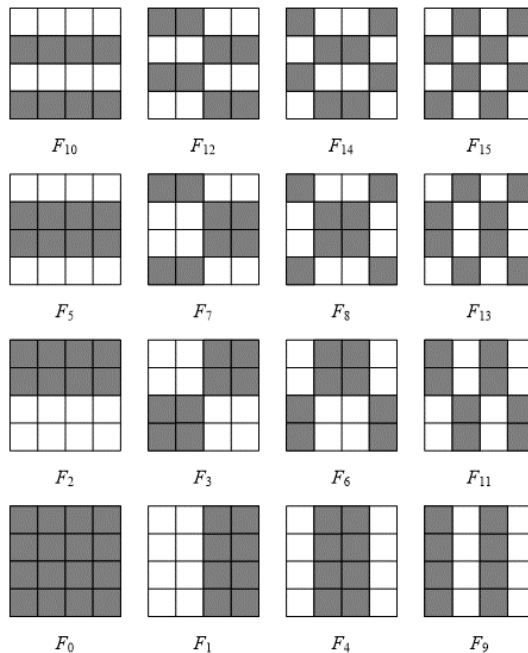
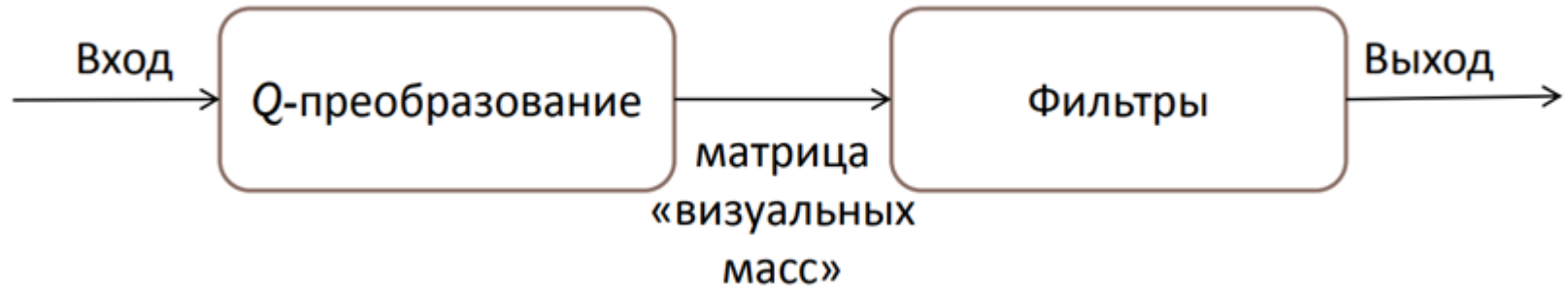
Матрица «визуальных масс»
размером 4×4 пикселей

Элементы ТАВ



357.794	265.373	243.915	335.485
239.140	151.990	120.429	222.775
279.018	220.485	214.542	239.990
347.242	236.747	228.757	308.971

Элементы ТАВ



Характеристики

Количество – 16;

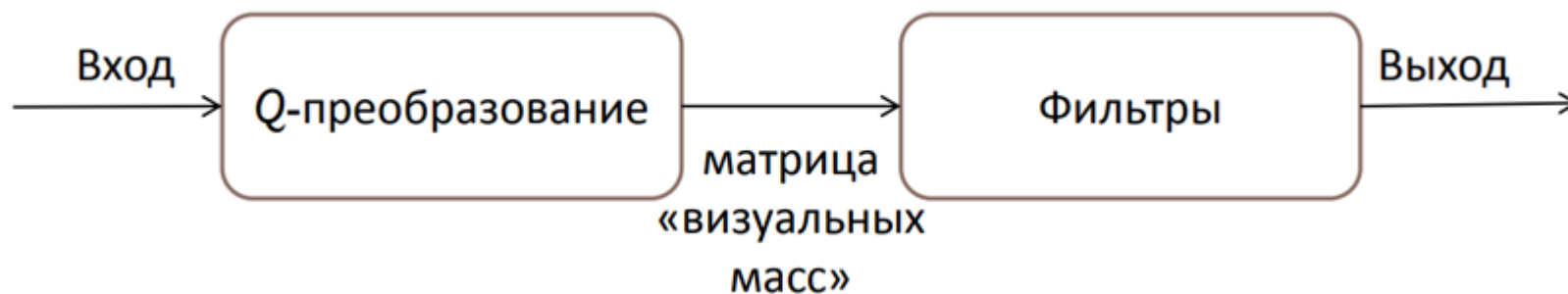
Темный элемент – значение – 1

Светлый элемент – значение +1

-1	-1	+1	+1
-1	-1	+1	+1
-1	-1	+1	+1
-1	-1	+1	+1

F_1

Элементы ТАВ



$$\Sigma \left[\begin{array}{c} \text{Матрица} \\ \text{«визуальных масс»} \end{array} \times \begin{array}{c} \text{Фильтр } F_1 \end{array} \right] = -0.4$$

0.6	0.3	0.6	0.7
0.1	3.4	0.2	3.4
0.6	0	0.6	0
0.9	3.4	0	3.4

-1	-1	+1	+1
-1	-1	+1	+1
-1	-1	+1	+1
-1	-1	+1	+1

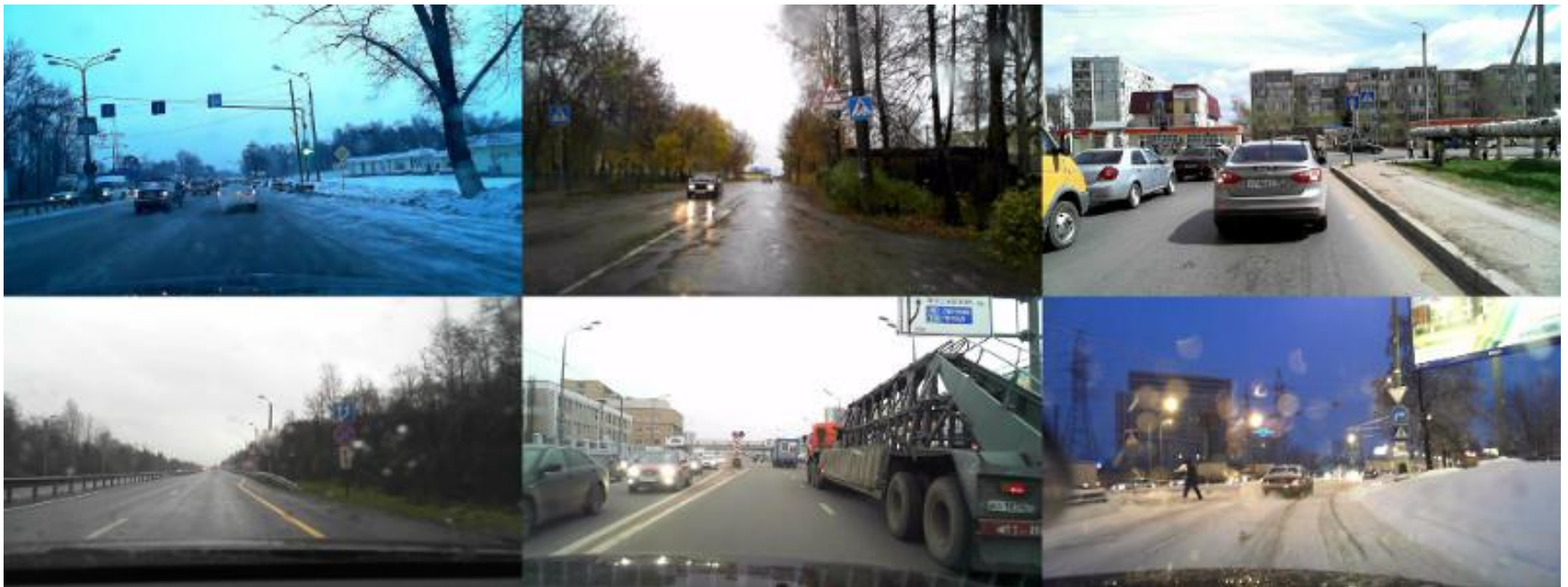
$$\mu =$$

F_0	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	F_{14}	F_{15}
18.2	-0.4	0.4	1.4	-1.2	-1.6	0.6	0.6	1.4	-11	11.4	-1.6	-1.2	0.4	-0.4	-13.8

Обучение модели

- Каждому объекту присваивается класс (наименование знака)
- Все классы кодируются числовыми значениями
- Признаковое описание сохраняется в базе

Набор входных изображений

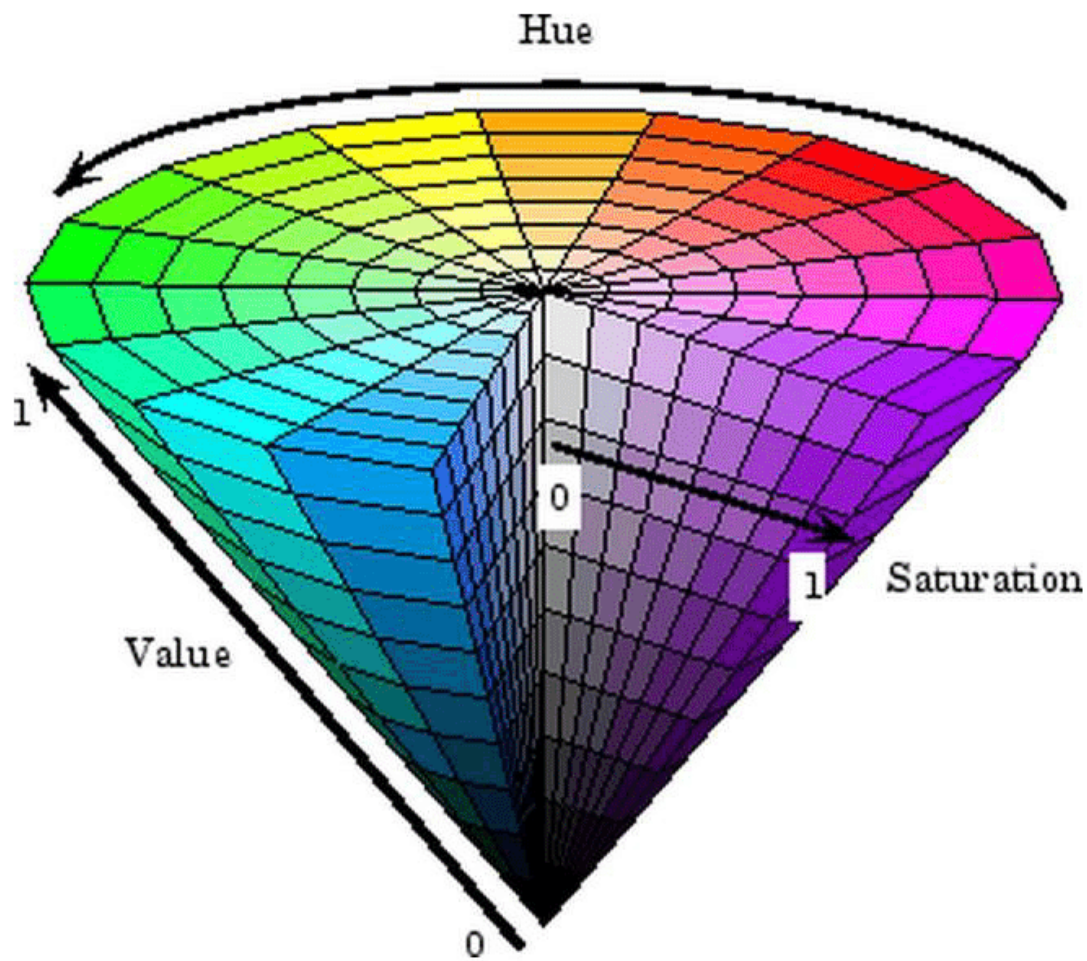


Предварительная обработка

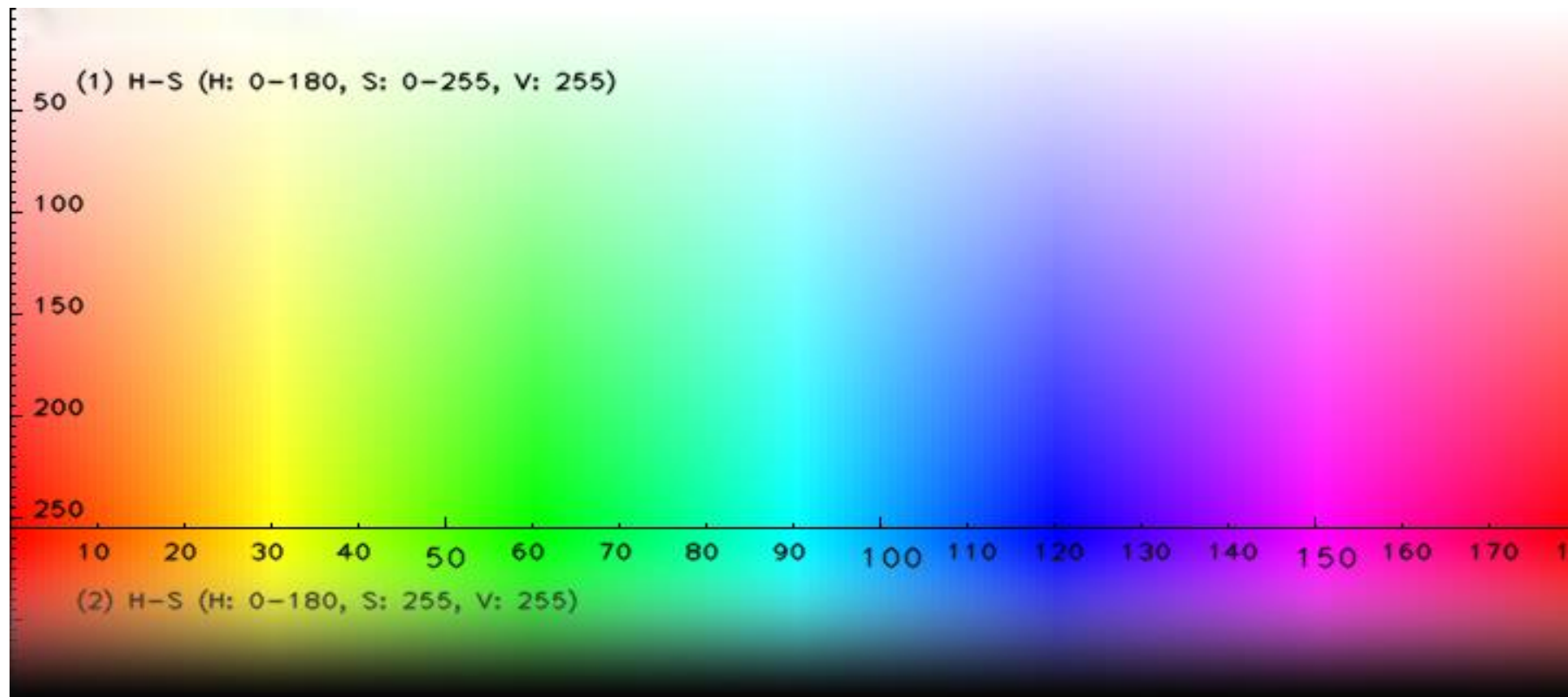
- Конвертация изображения в цветовое пространство HSV
- Применение фильтра Гаусса:

$$G(u, v) = \frac{1}{2\pi r^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}$$

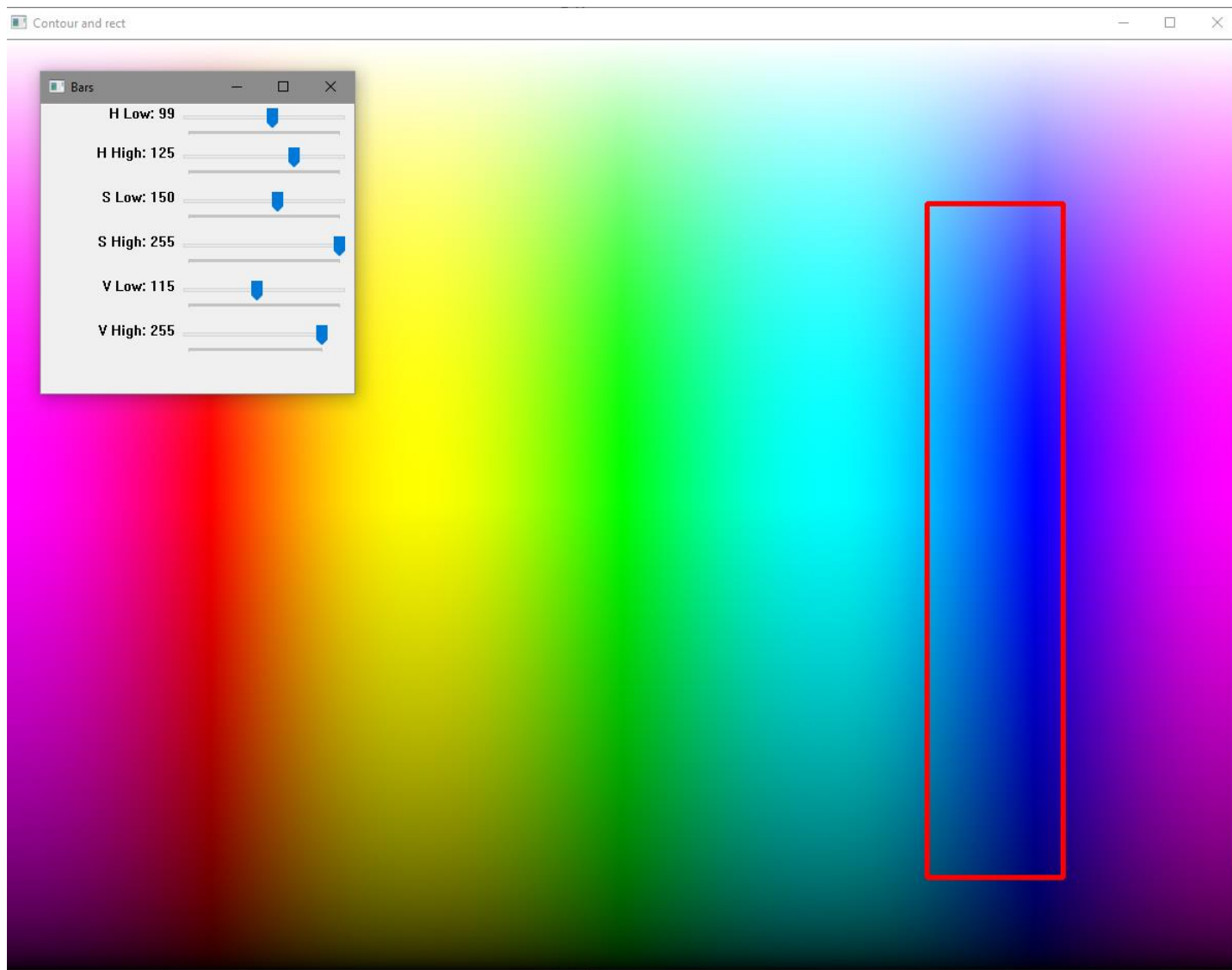
Цветовое пространство HSV



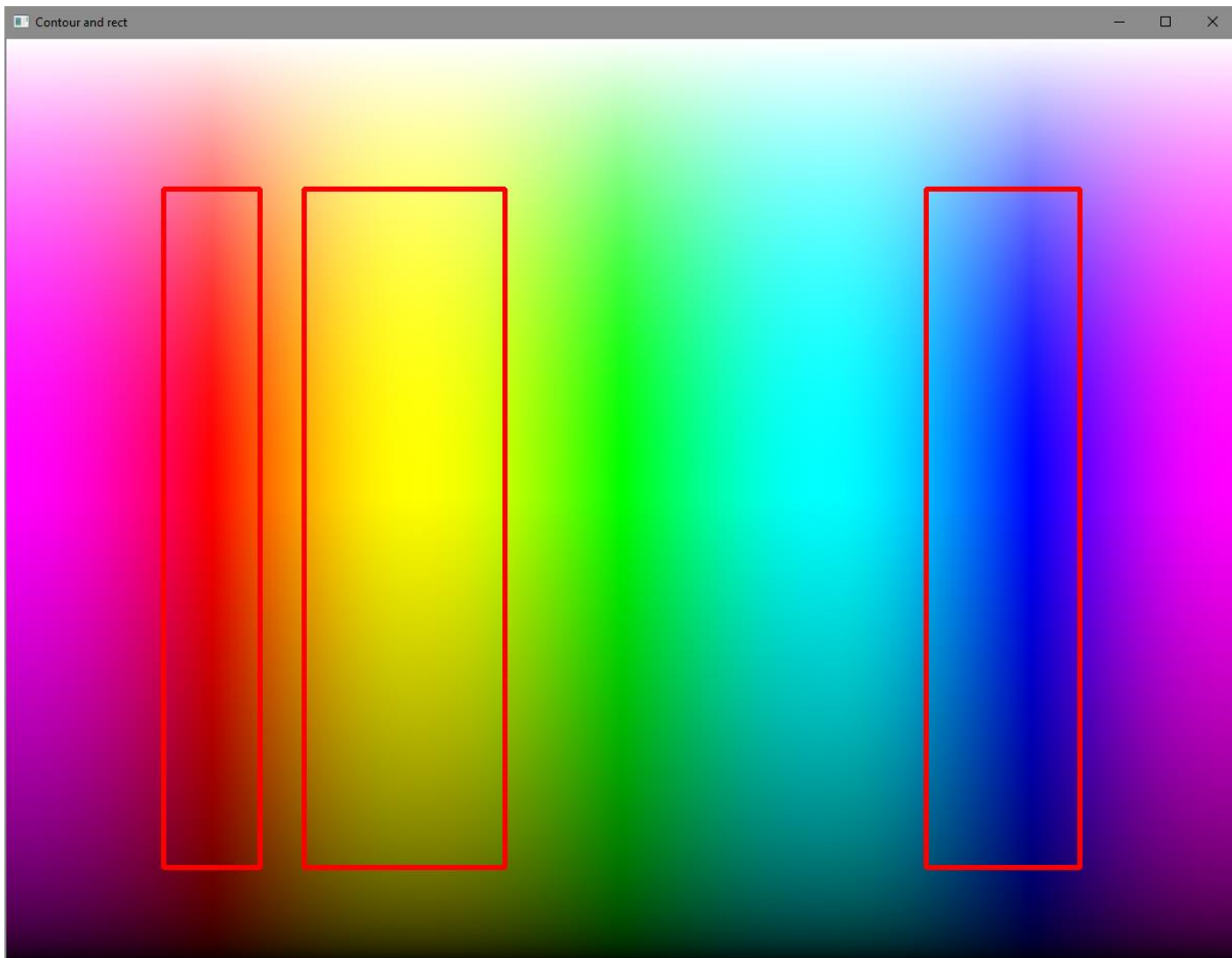
Калибровка цветового диапазона



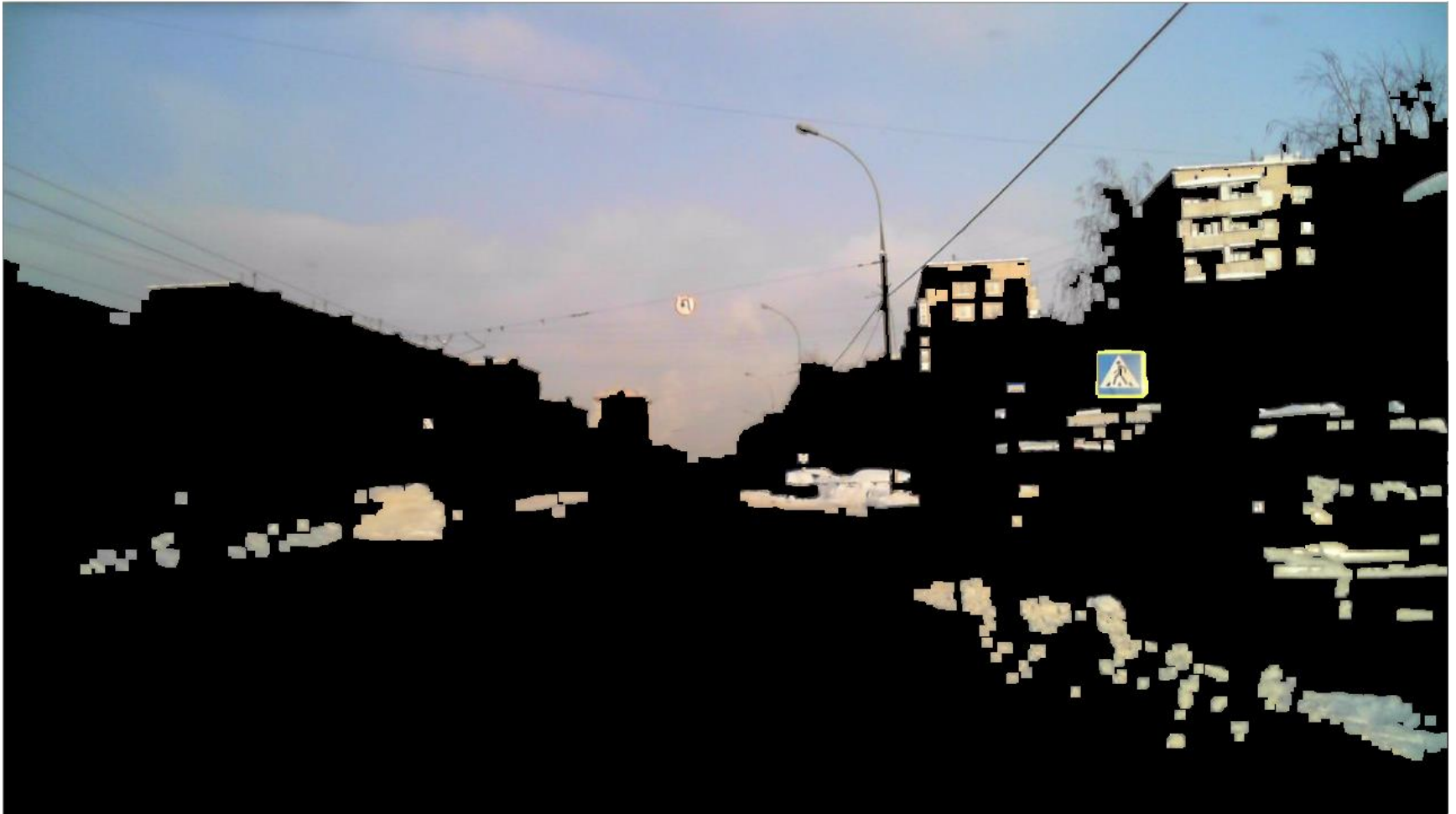
Калибровка цветового диапазона



Калибровка цветового диапазона



Выделение маски



Морфологические преобразования маски

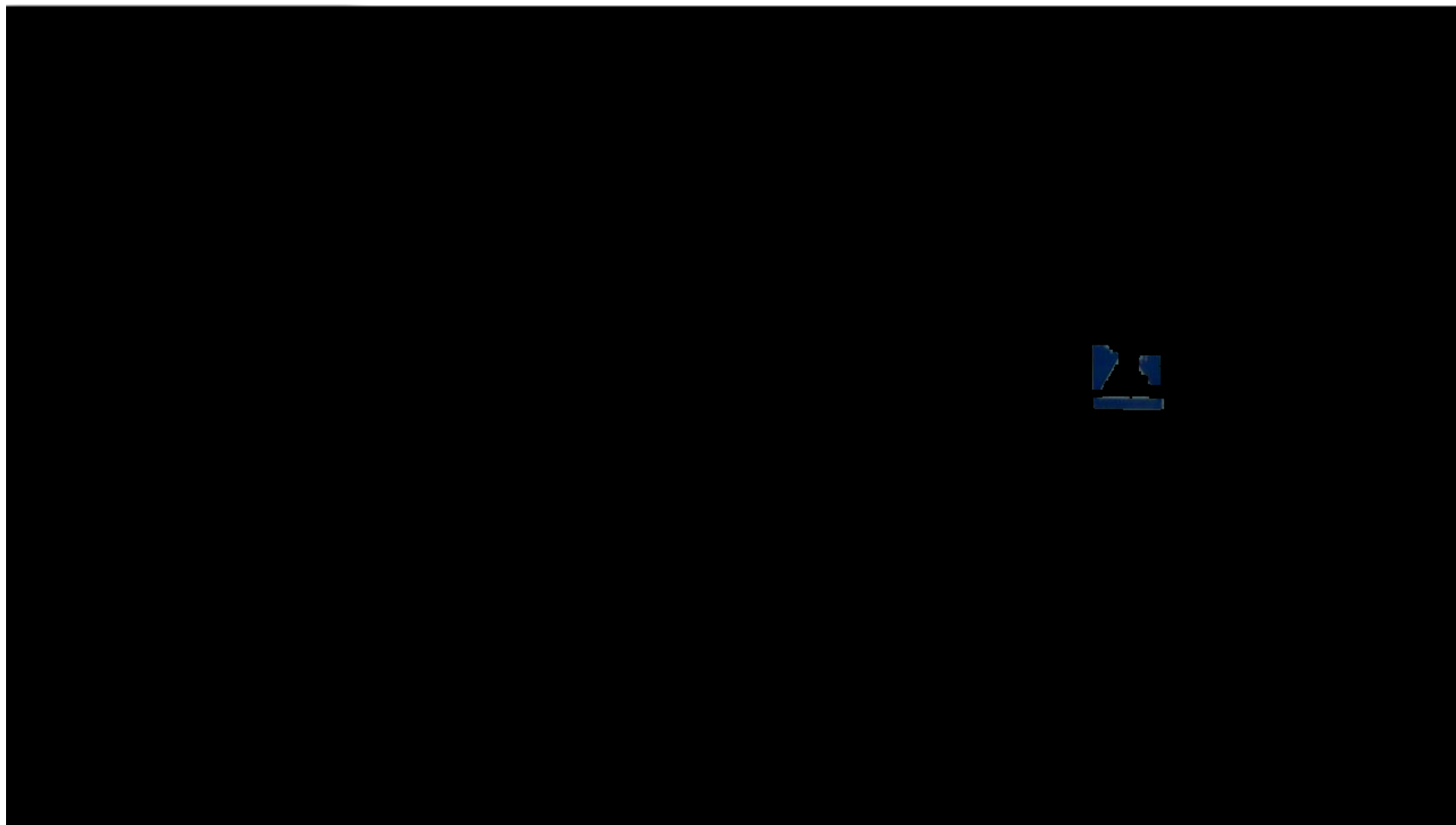


(a)

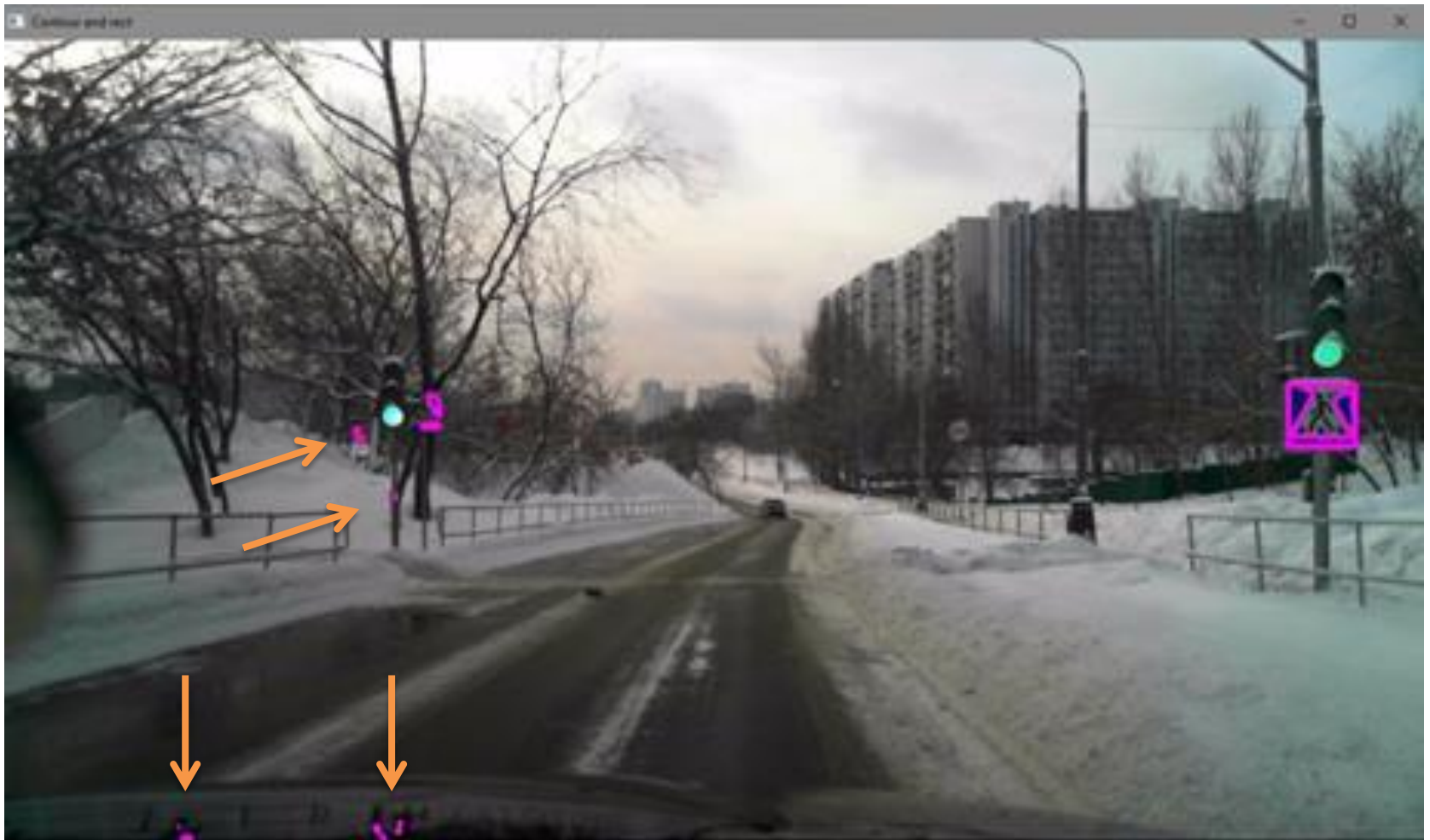
(б)

(в)

Результирующая маска

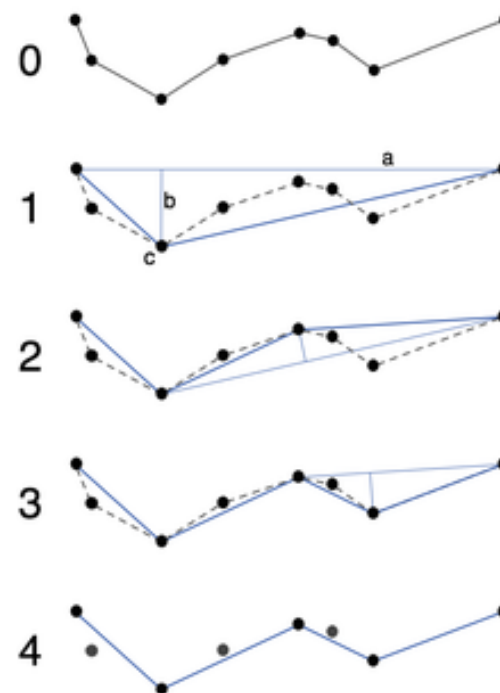


Построение контуров



Фильтрация контуров

- Иерархия контуров – объединение по старшинству
- Отсечение по пороговой площади (мелкие помехи)
- Отсечение по форме (крупные помехи):
 - Аппроксимация контуров (алгоритм Рамера-Дугласа-Пекера)
 - Подсчет ребер и анализ результатов:
 - 3 ребра – треугольник
 - 4 ребра – квадрат или прямоугольник (соотношение сторон)
 - 6 и более ребер – круг
 - Фильтрация контуров по целевой форме



Принятие решения

- Поиск ближайшего признакового описания эталонного объекта по отношению к данному объекту
- Мера близости – евклидово расстояние:

$$(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} = \sqrt{\sum_{k=1}^n (p_k - q_k)^2}$$

Результат

- Определяется изображение эталонного объекта
- Определяется класс объекта (название знака)
- Режим отладки:
 - Выводится анализируемое изображение
 - Изображение знака и его название выводятся в кадре поверх анализируемого изображения
- Режим оценки точности:
 - Положение знака в кадре и его класс сверяются с истинными значениями
 - Выводится оценка точности

Вычислительный эксперимент:

подготовка данных

1. Выборка изображений, сделанных в нормальных условиях
2. Выборка изображений с добавлением шума различной интенсивности:
 - ОСШ = 20дБ
 - ОСШ = 10дБ
 - ОСШ = 0дБ
3. Выборка изображений с поворотом искомых объектов:
 - Угол поворота = 15 градусов
 - Угол поворота = 30 градусов
 - Угол поворота = 45 градусов

Вычислительный эксперимент: локализатор

№	Синий цвет			Желтый цвет			Красный цвет		
	Н	S	V	Н	S	V	Н	S	V
1	95-130	140-255	100-255	19-37	130-255	110-255	160-20	135-255	115-255
2	99-125	155-255	115-255	23-34	141-255	124-255	170-10	141-255	124-255
3	105-115	165-255	125-255	25-31	150-255	136-255	175-5	152-255	135-255

Вычислительный эксперимент: локализатор, первая выборка

№	Итераций Erosion	Итераций Dilation	Точность
1	2	2	76.43%
1	3	4	79.64%
1	4	1	74.12%
2	3	4	89.78%
2	4	3	90.33%
2	4	2	94.71%
2	2	2	86.52%
3	2	3	83.46%
3	4	4	79.12%
3	3	4	81.83%

Вычислительный эксперимент: локализатор, вторая выборка

№	Итераций Erosion	Итераций Dilation	Точность		
			ОСШ = 20дБ	ОСШ = 10дБ	ОСШ = 0дБ
1	4	2	90.11%	85.21%	70.41%
2	2	2	89.81%	86.52%	71.90%
2	5	3	91.26%	87.46%	72.91%
3	4	4	88.28%	81.12%	72.14%
3	3	4	87.92%	84.61%	70.76%

Вычислительный эксперимент: локализатор, третья выборка

№	Итераций Erosion	Итераций Dilation	Точность		
			УП = 15°	УП = 30°	УП = 45°
2	4	3	90.33%	90.33%	90.33%
2	3	4	89.78%	89.78%	89.78%
2	4	2	94.71%	94.71%	94.71%
1	3	4	79.64%	79.64%	79.64%
3	2	3	83.46%	83.46%	83.46%

Вычислительный эксперимент: классификатор, первая выборка

АРДП	Точность	Время
Да	98.53%	0.34 сек.
Нет	96.07%	0.26 сек.

Вычислительный эксперимент: классификатор, вторая выборка

АРДП	ОСШ	Точность	Время
Да	20дБ	91.67%	0.38 сек.
	10дБ	87.32%	0.38 сек
	0дБ	81.12%	0.4 сек.
Нет	20дБ	89.71%	0.27 сек.
	10дБ	84.59%	0.28 сек.
	0дБ	78.97%	0.31 сек.

Вычислительный эксперимент: классификатор, третья выборка

АРДП	Угол поворота	Точность	Время
Да	15°	97.87%	0.35 сек.
	30°	97.34%	0.34 сек.
	45°	96.91%	0.36 сек.
Нет	15°	95.17%	0.27 сек.
	30°	94.76%	0.24 сек.
	45°	94.23%	0.26 сек.

Точность аналогов

Метод	Точность
GIST	93-96%
HOG	89-94%
CoHOG	95-98%
Моменты Ху	94-98%

Заключение

- Выполнен обзор и анализ существующих методов получения глобального признакового описания
- Выполнен обзор и анализ существующих методов классификации признаковых описаний
- Разработана информационная модель
- Разработана алгоритмическая и программная реализация системы
- Приведены результаты проведённых экспериментов, которые показывают работоспособность предлагаемого метода.
- Полученные результаты демонстрируют, что разработанный метод устойчив к масштабированию, перевороту и изменению углу наблюдения объекта на изображении.

Публикации

- А.В. Смирнов, В.Е. Гай - Модель и алгоритмы обнаружения объектов на изображении с использованием глобального признакового описания // Материалы XXV международной научно-технической конференции «Информационные системы и технологии - 2019», ИСТ-2019, Россия, Н. Новгород, 2019г.

Демонстрация

Спасибо за внимание!

Обзор альтернативных методов

- Инварианты Ху
- GIST
- HOG
- CoHOG

Инварианты Ху

$$M_{pq} = \iint_{-\infty}^{\infty} x^p y^q f(x, y) dx dy$$

$$M_{ij} = \sum_x \sum_y x^i y^j I(x, y)$$

$$\mu_{pq} = \iint_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy$$

$$\bar{x} = \frac{M_{10}}{M_{00}}, \bar{y} = \frac{M_{01}}{M_{00}}$$

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y)$$

$$\mu_{00} = M_{00}$$

$$\mu_{01} = 0$$

$$\mu_{02} = 0$$

$$\mu_{11} = M_{11} - \bar{x}M_{01} = M_{11} - \bar{y}M_{10}$$

$$\mu_{20} = M_{20} - \bar{x}M_{10}$$

$$\mu_{02} = M_{02} - \bar{y}M_{01}$$

$$\mu_{21} = M_{21} - 2\bar{x}M_{11} - \bar{y}M_{20} + 2\bar{x}^2 M_{01}$$

$$\mu_{12} = M_{12} - 2\bar{y}M_{11} - \bar{x}M_{02} + 2\bar{y}^2 M_{10}$$

$$\mu_{30} = M_{30} - 3\bar{x}M_{20} + 2\bar{x}^2 M_{10}$$

$$\mu_{03} = M_{03} - 3\bar{y}M_{02} + 2\bar{y}^2 M_{01}$$

$$\mu_{pq} = \sum_m^p \sum_n^q \binom{p}{m} \binom{q}{n} (-\bar{x})^{(p-m)} (-\bar{y})^{(q-n)} M_{mn}$$

Инварианты Ху

$$\eta_{ij} = \frac{\mu_{ij}}{\mu_{00}^{(1+\frac{i+j}{2})}}$$

$$I_1 = \eta_{20} + \eta_{02}$$

$$I_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$I_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$I_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$I_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

$$I_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$

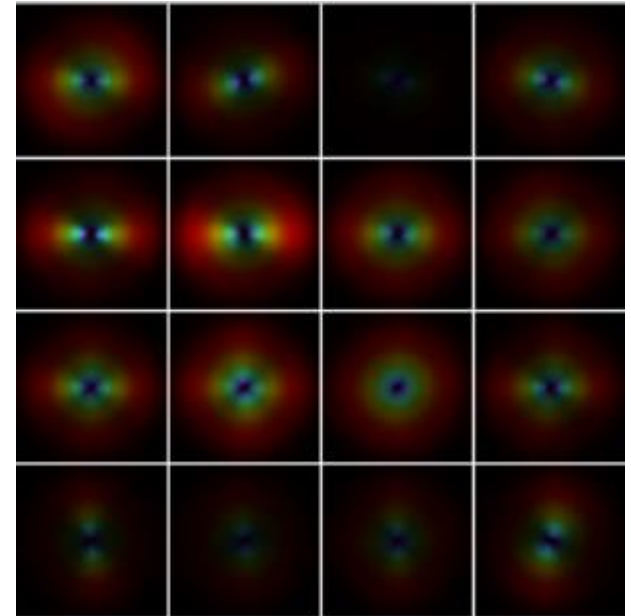
$$I_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2].$$

GIST

Входное изображение

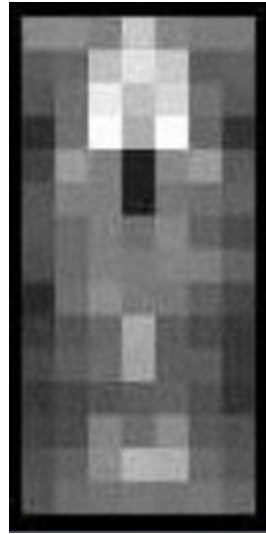


Дескриптор GIST



- GIST суммирует градиентную информацию (масштабы и ориентацию) для разных частей изображения. Вычисления происходят следующим образом:
1. Свертка изображения, используя 32 фильтра Габора в 4 масштабах, 8 ориентациях.
 2. Разделение каждой карты на 16 областей и усреднение значений в каждой области.
 3. Конкатенация всех усредненных значений.

HOG



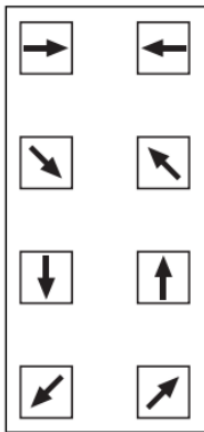
В основе метода HOG лежит предположение, что вид распределения градиентов интенсивности изображения позволяет достаточно точно определить наличие и форму присутствующих на нем объектов. При описании изображения оно разбивается на несколько небольших участков, далее называемых ячейками. В ячейках вычисляются гистограммы h :

$$h_{L_2} = \frac{h}{\sqrt{|h|_2^2 + \varepsilon}}, h_{L_1} = \frac{h}{\sqrt{|h|_2 + \varepsilon}}, h_{\sqrt{L_1}} = \sqrt{h_{L_1}}$$

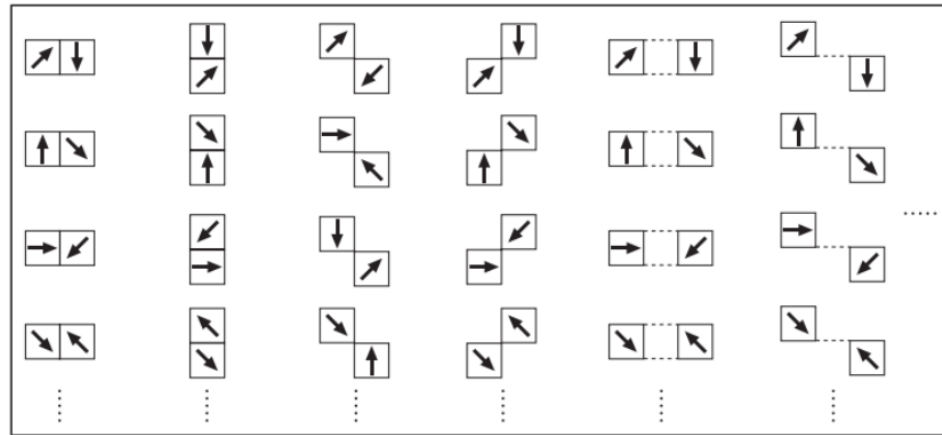
CoHOG

CoHOG - это пары градиентных ориентаций.

В HOG для одного градиента ориентация имеет только восемь разновидностей, но пара из них имеет гораздо больше возможных разновидностей. Таким образом, CoHOG может выражать формы более подробно, чем HOG, который использует ориентацию с одним градиентом. Как можно заметить, одиночный набор ориентаций градиентов (a) представляет собой набор из восьми доступных положений, в то время как набор пар, составленный из их комбинаций (b) имеет намного больше доступных положений.



(a)



(b)

Набор допустимых ориентаций градиентов для дескрипторов HOG (a) и CoHOG (b)