

Дескриптор в аккумуляторном пространстве Хафа градиентного поля изображения для детектирования пешеходов*

Н. А. Самсонов¹, А. Н. Гнеушев²

nikita.samsonov@phystech.edu; gneushev@ccas.ru;

¹Московский физико-технический институт, Россия, г. Долгопрудный, Институтский пер., 9

²ФИЦ «Информатика и управление» РАН, Россия, г. Москва, ул. Вавилова, 44/2

Рассматривается задача выделения признаков для распознавания пешеходов на изображении. Наиболее популярный и эффективный подход к выбору признаков для распознавания объекта на изображении — использование дескриптора на основе гистограмм ориентаций градиентов (Histograms of Oriented Gradients, HOG). В работе предлагается использовать аккумуляторное пространство Хафа для обобщения HOG дескриптора путем получения проекций не только ориентаций, но и положений границ на локальном участке изображения — гистограмм аккумуляторного пространства Хафа (Hough Accumulator Histograms, HAH). Аккумуляторное пространство Хафа строится на основе лучевого преобразования Радона градиентного поля изображения. Предлагаемые методы были протестированы совместно с линейным классификатором метода опорных векторов (Support Vector Machine, SVM) на базе изображений пешеходов INRIA. Результаты экспериментов показали лучшую разделяющую способность новых дескрипторов, уменьшение ложных срабатываний детекторов в сравнении с HOG.

Ключевые слова: *histograms of oriented gradients, HOG; Hough accumulator histograms, HAH; дескриптор; пешеходы; преобразование Радона; преобразование Хафа*

DOI: 10.21469/22233792.3.3.03

1 Введение

Автоматическое детектирование объектов на изображениях и видео является одной из основных задач компьютерного зрения. Цель детектирования — определить наличие и расположение искомого объекта на видеоизображении для последующего анализа. Такие объекты на изображении, как люди, являются неотъемлемой частью дорожной сцены, играют важную роль в развитии дорожной ситуации. Задача локализации человека на видеоизображениях широко востребована в таких областях, как мониторинг и анализ дорожных ситуаций, обнаружение дорожно-транспортных происшествий, контроль за соблюдением правил дорожного движения, системы безопасности и следящие системы, беспилотные автомобили, робототехника, системы помощи водителю.

Распознавание человека на изображении — сложная задача. Ее сложность обусловлена несколькими причинами: изменение структуры изображения вследствие движения объекта, неравномерная освещенность изображения, большая вариабельность изображений человека из-за разных ракурсов (поз, размеров, углов поворота), частичные перекрытия фигуры человека другими объектами.

Как правило, множество признаков, которые используются для распознавания объекта, определяются его характерной структурой на изображении. Изображение человека

*Работа выполнена при финансовой поддержке РФФИ, грант №16-07-01171.

может быть представлено совокупностью контуров, силуэтов частей тела, которые являются контурными признаками и представляются на изображении как максимальные перепады значений яркости [1]. Однако контуры на изображении человека могут иметь достаточно сложную и разнообразную структуру, и для описания этой структуры необходимо учитывать совокупность контурных признаков — локальных дескрипторов для описывания изображения объекта в целом [2, 3]. На сегодняшний день наиболее популярны подходы, основанные на градиентах яркости в локальных областях изображения и учитывающие совокупность контурных признаков — локальных дескрипторов для описания изображения объекта в целом. Так, в подходе, описанном в [4, 5], строится каскад бинарных классификаторов, где признаки, напоминающие базисные вейвлеты Хаара, вычисляются по локальным прямоугольным областям изображения. Однако детекторы, основанные на данном методе, являются неустойчивыми по отношению к таким факторам, как неравномерное изменение освещения и вариабельность изображения объекта. Эффективная схема построения интегрального вектора признаков для задачи детектирования человека на основе локальных контурных дескрипторов — гистограмм ориентаций градиентов в локальных областях изображения — используется в методе HOG [6] и подходах, которые его развивают [7, 8]. HOG-дескриптор обладает рядом преимуществ: он устойчив к неравномерному освещению, вычислительно эффективен, показывает лучшие результаты для задачи детектирования пешеходов, однако он недостаточно полно описывает локальные контурные особенности объекта, используя лишь информацию об ориентации, вследствие чего детектор на его основе имеет неудовлетворительную для многих реальных приложений долю ложных срабатываний при приемлемом уровне пропуска. Гистограммы ориентаций градиентов можно рассматривать как аппроксимацию проекций модуля градиента на основные направления. Так, в работах [9, 10] предлагаются методы на основе преобразования Радона для детектирования сложных фигур на изображении. Описанные в работах дескрипторы строятся как проекции в пространстве преобразования Радона бинарных изображений, являются инвариантными относительно поворота, переноса и масштабирования, показывая при этом хорошие результаты, однако для полутоновых изображений результатов получено не было.

В данной работе предлагается обобщить HOG-дескриптор с помощью определения нового дескриптора — НАН — для детектирования пешеходов на изображениях. Для построения аккумуляторного пространства Хафа используется подход на основе лучевого поперечного преобразования Радона градиентного поля изображения. Использование проекций в аккумуляторном пространстве Хафа с целью описания изображения позволяет учитывать в НАН-дескрипторе не только локальное распределение ориентаций контурных признаков, но также и их положение. Таким образом, предложенный метод преодолевает недостатки [6], показывая, что использование дополнительной информации о положении локальных контурных особенностей объекта позволяет превзойти HOG по качеству детектирования и уменьшить уровень пропуска при приемлемом числе ложных срабатываний.

2 Постановка задачи

Определим задачу построения признакового пространства G как получение интегрального вектора признаков $\mathbf{g}(f) = (g_1, \dots, g_n)^T$ для входного изображения f , где n — размерность пространства признаков. Вектор признаков $\mathbf{g}(f)$ будем строить на основе общей схемы метода HOG [6], используя классификацию методом SVM при построении эффективного детектора пешехода. Критерием эффективности детектора будем считать уровень ошибок 1-го и 2-го рода на тестовой выборке пешеходов INRIA.

В соответствии со схемой НОГ [6], изображение $f(x, y)$, где x, y — координаты точки изображения размером $X \times Y$, разбивается на смежные области — квадратные ячейки размером $K \times K$ пикселей, $I = X/K$ ячеек по горизонтали и $J = Y/K$ ячеек по вертикали. В каждой ячейке (i, j) , где i, j — индексы ячейки по вертикали и горизонтали соответственно, выделяются локальные признаки, которые определяются дескриптором ячейки — вектором $\mathbf{u}_{i,j} = (u_1, \dots, u_l)^T$, где l — размерность вектора. Четыре смежные ячейки объединяются в пересекающиеся блоки. Таким образом, каждый блок содержит $4K \times 4K$ пикселей и имеет общие ячейки с соседними блоками. Дескриптор блока определяется объединением дескрипторов собственных ячеек, вектором $\mathbf{v}_{i,j} = \mathbf{u}_{i,j} \cup \mathbf{u}_{i+1,j} \cup \mathbf{u}_{i,j+1} \cup \mathbf{u}_{i+1,j+1}$, где i, j — индексы блока. Общее количество блоков — $(I - 1)(J - 1)$. Под операцией объединения \cup двух векторов \mathbf{u}_1 и \mathbf{u}_2 с размерностью l_1 и l_2 соответственно будем понимать результирующий вектор из пространства размерности $l_1 + l_2$, первые l_1 компонент которого являются компонентами вектора первого аргумента, последние l_2 — компонентами вектора второго аргумента.

Дескриптор блока нормируется с помощью одной из двух норм: L_2 нормы

$$\tilde{\mathbf{v}} = N_{L_2}(\mathbf{v}) = \frac{\mathbf{v}}{\sqrt{\|\mathbf{v}\|_2^2 + \varepsilon^2}} \tag{1}$$

либо L_2 -hys нормы [6]

$$\tilde{\mathbf{v}} = N_{L_2\text{-hys}}(\mathbf{v}) = N_{L_2}(\min(N_{L_2}(\mathbf{v}), h)), \tag{2}$$

где операция \min применяется покомпонентно к вектору первого аргумента; h — пороговое значение, которое используется для ограничения значений компонент вектора в операции \min . Данное пороговое ограничение определено с целью фильтрации шумовых всплесков, которые могут появляться в компонентах дескриптора. В данной работе, как и в работе [6], в качестве порогового значения используется $h = 0,2$.

Интегральный дескриптор определяется объединением всех блочных дескрипторов, т. е. вектором

$$\mathbf{g} = \bigcup_j^{J-1} \bigcup_i^{I-1} \tilde{\mathbf{v}}_{i,j}.$$

Интегральные дескрипторы \mathbf{g} множества изображений формируют признаковое пространство G . Для решения задачи детектирования пешеходов необходимо определить разбиение этого пространства на два непересекающихся класса: первый характеризует пешеходов; второй — фон, не содержащий пешехода. Ограничимся использованием линейного классификатора, который задается выражением:

$$s(\mathbf{g}) = \text{sign} \left(\sum_{j=1}^n w_j g_j - w_0 \right) = \text{sign} (\langle \mathbf{w}, \mathbf{g} \rangle - w_0),$$

где вектор $\mathbf{w} = (w_1, \dots, w_n)^T \in R^n$ и скалярный порог $w_0 \in R$ определяют разделяющую гиперплоскость $\langle \mathbf{w}, \mathbf{g} \rangle = w_0$ в пространстве G .

Вектор весов \mathbf{w} классификатора и скалярный порог w_0 находятся из процедуры обучения на основе метода SVM [6] по специально подготовленной обучающей выборке из базы изображений INRIA, содержащий два подмножества изображений: с положительными примерами, содержащими пешеходов, и отрицательными примерами, содержащими фон.

Критерием качества классификации $s(\mathbf{g})$ на специально подготовленной тестовой выборке из базы изображений INRIA будем считать отношение $MR = FN / (TP + FN)$ — доля неверно отвергнутых классификатором изображений (Miss Rate), к $FPPW = FP / (TN + FP)$ — доля неверно принятых изображений (False Positive Per Window), где FN — количество неверно отвергнутых классификатором положительных примеров; TP — количество верно классифицированных положительных примеров; TN — количество верно классифицированных отрицательных примеров; FP — количество неверно классифицированных отрицательных примеров.

3 Построение текстурных признаков для детектирования пешеходов в аккумуляторном пространстве Хафа

3.1 Лучевое поперечное преобразование Радона градиентного поля изображения

Рассмотрим преобразование Радона [11] функции $f \in L_2(\mathbb{R}^2)$ вдоль прямой $L : \mathbf{x} = \rho \mathbf{n} + t \mathbf{a}$, которое определяется формулой:

$$\mathcal{R}_{\alpha, \rho}(f) = \int_L f(\rho \mathbf{n} + t \mathbf{a}) dt, \quad (3)$$

где t — параметр прямой L ; $\mathbf{n} = (\cos \alpha, \sin \alpha)^T$ и $\mathbf{a} = (-\sin \alpha, \cos \alpha)^T$ — нормаль и направляющий вектор прямой; α задает ориентацию вектора нормали \mathbf{n} ; $\rho = x \cos \alpha + y \sin \alpha$ — модуль радиуса-вектора ближайшей к началу координат точки на прямой. Таким образом, α и ρ — параметры прямой L , задающие пространство Радона \mathcal{R} .

Преобразование Радона обладает свойством обратимости [11]. Таким образом, при переходе в пространство Радона информация о функции f не теряется.

На основе преобразования Радона (3) определим лучевое поперечное преобразование [11] градиента функции $\nabla f \in L_2(\mathbb{R}^2)$ (рис. 1):

$$\mathcal{R}_{\alpha, \rho}^{\perp}(\nabla f) = \int_L \langle \mathbf{n}, \nabla f(\rho \mathbf{n} + t \mathbf{a}) \rangle dt. \quad (4)$$

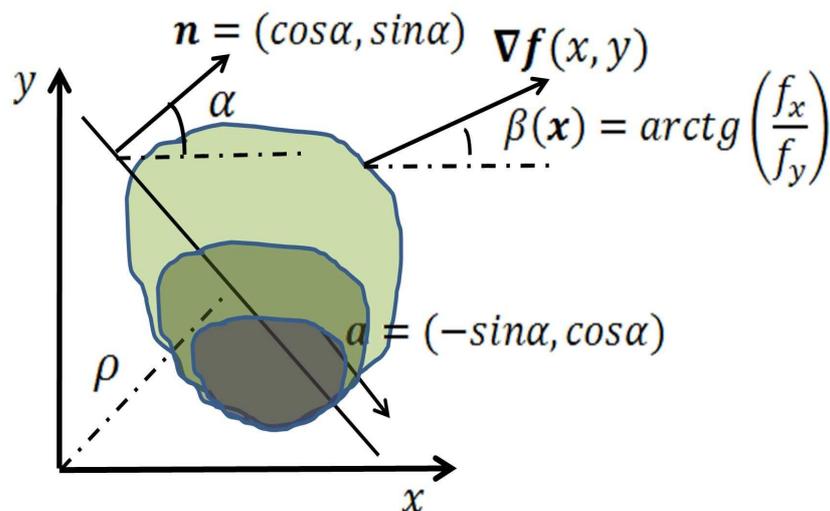


Рис. 1 Преобразование Радона градиентного поля

Выражение (4) можно переписать в виде:

$$\begin{aligned} \mathcal{R}_{\alpha,\rho}^{\perp}(\nabla f) &= \int_L \langle \mathbf{n}, \nabla f(\rho\mathbf{n} + t\mathbf{a}) \rangle dt = \int_L |\nabla f(\rho\mathbf{n} + t\mathbf{a})| \cos(\alpha - \beta(\mathbf{x})) dt = \\ &= \int_L |\nabla f(\rho\mathbf{n} + t\mathbf{a})| q(\alpha, \beta(\mathbf{x})) dt, \end{aligned} \quad (5)$$

где $\nabla f(\mathbf{x})$ — вектор градиента в точке $\mathbf{x} = (x, y)^T$; $\beta(\mathbf{x})$ — угол, задающий ориентацию градиента в точке \mathbf{x} ; $\mathbf{n} = (\cos \alpha, \sin \alpha)^T$; $\mathbf{a} = (-\sin \alpha, \cos \alpha)^T$ — нормаль и направляющий вектор прямой L соответственно; α задает ориентацию вектора нормали \mathbf{n} ; $q(\alpha, \beta(\mathbf{x})) = \cos(\alpha - \beta(\mathbf{x}))$ — функция, задающая вклад модуля градиента в точке \mathbf{x} в направлении α .

3.2 Гистограммные признаки в аккумуляторном пространстве Хафа

Рассмотрим дискретный случай лучевого поперечного преобразования Радона (5). Разобьем пространство параметров α, ρ на ячейки, для чего зададим разбиение осей на бины: $k = 1, \dots, N$ — индексы разбиения по оси α ; $l = 1, \dots, M$ — индексы разбиения по оси ρ . Получим множество ячеек $\{(\alpha_k, \rho_l)\}$, задающее аккумуляторное пространство Хафа (рис. 2), где (α_k, ρ_l) — центр ячейки с соответствующими индексами. Аккумуляторное пространство Хафа для поля градиентов ∇f строится путем сбора двумерной гистограммы. Из формулы преобразования Радона (5) для значения ячейки (α_k, ρ_l) аккумуляторного пространства получаем выражение:

$$\mathcal{H}_{\alpha_k, \rho_l}^{\perp}(\nabla f) = \sum_{x,y} |\nabla f(x, y)| q(\alpha_k, \beta(\mathbf{x})) [\rho_l = x \cos \alpha_k + y \sin \alpha_k], \quad (6)$$

где операция $[\cdot]$ возвращает индикаторную величину, равную 0, если выражение в скобках не выполняется, и 1 — в противном случае.

Аккумуляторное пространство Хафа, описываемое выражением (6), можно рассматривать в качестве пространства, характеризующего контурные признаки изображения, и строить на его основе дескрипторы. Определим локальные дескрипторы ячеек изображения в аккумуляторном пространстве Хафа. Из формулы (6) следует, что для локальной

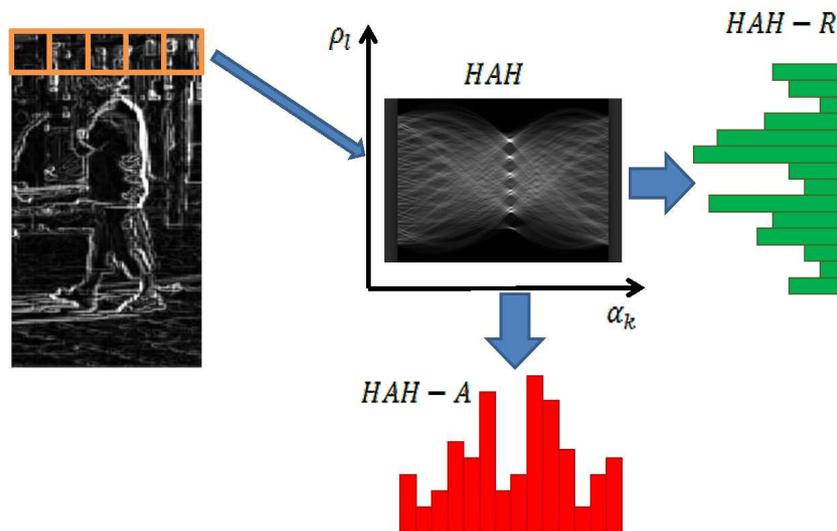


Рис. 2 Преобразование Хафа градиентного поля

ячейки (i, j) изображения значение в точке (α_k, ρ_l) признакового пространства описывается формулой:

$$\mathcal{H}_{\alpha_k, \rho_l}(i, j, \nabla f) = \sum_{(x, y) \in \text{Cell}_{i, j}} |\nabla f(x, y)| q_{\mathcal{H}^\perp}(\alpha_m, \beta(\mathbf{x})) [\rho_l = x \cos \alpha_k + y \sin \alpha_k]. \quad (7)$$

В качестве функции вклада $q_{\mathcal{H}^\perp}(\alpha_m, \beta(\mathbf{x}))$ будем использовать функцию:

$$q_{\mathcal{H}^\perp}(\alpha_m, \beta(\mathbf{x})) = |\cos(\alpha_k - \beta(\mathbf{x}))|, \quad (8)$$

где модуль берется для того, чтобы вклады противоположных по направлению градиентов яркости на изображении учитывались одинаково.

Для уменьшения размерности пространства признаков (7) построим его проекции в пространство одного из параметров.

Определение 1. НАН-А дескриптор является проекцией пространства (7) на пространство параметра α и определяется выражением:

$$\mathbf{u}_{i, j}^{\text{НАН-А}} = \{u_{i, j}^k\}_{k=1}^N, \quad u_{i, j}^k = \sum_l \mathcal{H}_{\alpha_k, \rho_l}(i, j). \quad (9)$$

Полученный дескриптор совпадает по размерности с НОГ [6], но отличается от него видом интерполяционной функции.

Утверждение 1. НАН-А дескриптор $\mathbf{u}_{i, j}^{\text{НАН-А}}$, определяемый выражением (9) с функцией вклада

$$q(\alpha_k, \beta) = \left(1 - \frac{|\alpha_k - \beta(x, y)|}{\alpha_{k+1} - \alpha_k}\right) [|\alpha_k - \beta(x, y)| < (\alpha_{k+1} - \alpha_k)], \quad (10)$$

эквивалентен локальному дескриптору ячейки НОГ [6].

Для доказательства утверждения достаточно заметить, что функция вида (10) для q в сумме (9) является весовой функцией, линейно интерполирующей модуль градиента между соседними бинами параметра α . Именно данная интерполяция используется в дескрипторе НОГ.

Определение 2. НАН-R дескриптор является проекцией пространства (7) на пространство параметра ρ и определяется выражением:

$$\mathbf{u}_{i, j}^{\text{НАН-R}} = \{u_{i, j}^l\}_{l=1}^M, \quad u_{i, j}^l = \sum_k \mathcal{H}_{\alpha_k, \rho_l}(i, j). \quad (11)$$

Данный дескриптор, в отличие от НОГ [6], учитывает только положение границ объектов внутри ячеек.

На основе дескрипторов, задаваемых выражениями (9) и (11) можно построить дескриптор, учитывающий как ориентацию, так и положение границ контуров внутри ячеек.

Определение 3. НАН-AR дескриптор является объединением обеих проекций пространства (7) и определяется выражением:

$$\mathbf{u}_{i, j}^{\text{НАН-AR}} = \mathbf{u}_{i, j}^{\text{НАН-А}} \cup \mathbf{u}_{i, j}^{\text{НАН-R}}. \quad (12)$$

3.3 Построение детектора пешеходов на основе гистограммных признаков в аккумуляторном пространство Хафа

Построение детектора пешеходов включает две подзадачи: выделение гистограммных признаков в аккумуляторном пространстве Хафа (6) для построения дескриптора и его использование для бинарной классификации (рис. 3). Для выделения признаков на изображении будем пользоваться схемой HOG [6]. В качестве локальных дескрипторов ячеек будем использовать дескрипторы (9), (11) и (12). При использовании в качестве локального дескриптора ячейки HAH-AR на этапе нормализации гистограмм блоков, где блочный дескриптор обозначается как $v_{i,j}^{HAH-AR}$, будем нормировать следующим образом: сначала отдельно нормируются (1) и (2) гистограммы:

$$\begin{aligned} v_{i,j}^{HAH-A} &= u_{i,j}^{HAH-A} \cup u_{i+1,j}^{HAH-A} \cup u_{i,j+1}^{HAH-A} \cup u_{i+1,j+1}^{HAH-A}, \\ v_{i,j}^{HAH-R} &= u_{i,j}^{HAH-R} \cup u_{i+1,j}^{HAH-R} \cup u_{i,j+1}^{HAH-R} \cup u_{i+1,j+1}^{HAH-R}. \end{aligned}$$

Затем строится нормализованная гистограмма блока:

$$v_{i,j}^{HAH-AR} = u_{i,j}^{HAH-AR} \cup u_{i+1,j}^{HAH-AR} \cup u_{i,j+1}^{HAH-AR} \cup u_{i+1,j+1}^{HAH-AR}.$$

При использовании в качестве локальных дескрипторов ячейки HAH-A или HAH-R (9) или (11) нормализация гистограммы блока осуществляется по формулам (1) либо (2).

Для задачи классификации пешеходов на изображении будем применять линейный метод SVM, как и в [6]. Будем обучать алгоритм SVM в два этапа. На первом этапе обучение производится на выборке, составленной из базы размеченных изображений пешеходов (положительные примеры) и случайной выборки областей из изображений фона (отрицательные примеры). После первого этапа полученный классификатор используется для нахождения областей на изображениях фона, на которых детектор ошибается и ложно срабатывает. Эти области — «сложные примеры» — добавляются к обучающей выборке отрицательных примеров первого этапа. На втором этапе алгоритм SVM обучается второй раз на тех же положительных примерах и расширенной выборке отрицательных примеров для получения финального вектора весов линейного классификатора.

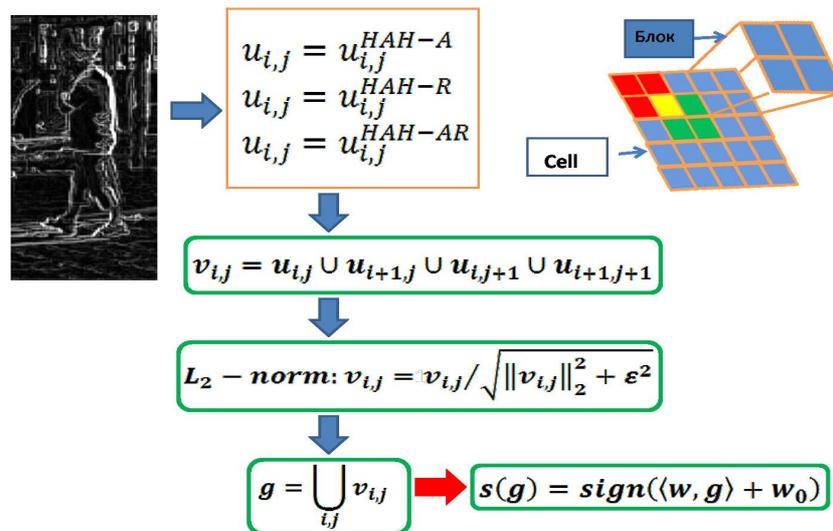


Рис. 3 Общая схема HAH-детектора

4 Результаты вычислительных экспериментов

Целью вычислительных экспериментов является проверка работы предлагаемых в статье дескрипторов на реальных данных, а также сравнение их качества детектирования и времени работы с аналогичными показателями HOG дескриптора [6].

Вычисления производились с помощью персонального компьютера с процессором Intel(R) Core(TM) i7 CPU с частотой 3,07 ГГц, 9 ГБ оперативной памяти, в системе ©Microsoft-Windows 7.

В качестве обучающей выборки была использована база изображений пешеходов INRIA [12]. База содержит 2478 изображений пешеходов размером 96×160 пикселей (рис. 4). Из каждого изображения выбирается центрированное окно размером 64 пикселей. Таким образом, 2478 окон составляют положительную часть обучающей выборки. В качестве изображений фона база данных содержит 1218 изображений (рис. 5) из которых в отрицательную часть обучающей выборки случайным образом выбирается 12 180 окон размером 64 пикселей. Число «сложных примеров» при использовании различных вариантов дескрипторов варьируется от 30 000 до 45 000.

В качестве локальных дескрипторов ячеек были использованы дескрипторы НАН-А, НАН-R и НАН-AR (9), (11) и (12), где множество параметра α аккумуляторного пространства Хафа разбивается на 9 бинов, а ρ — на 16 бинов. В зависимости от числа бинов интегральный дескриптор имеет следующую размерность: НАН-А — 3780, НАН-R — 6720 и НАН-AR — 10 500.

Для обучения классификатора была использована библиотека OpenCV [13]. Классификатор обучается с помощью линейного метода SVM [14]. В данных экспериментах в каче-



Рис. 4 Примеры позитивных изображений из обучающей выборки



Рис. 5 Примеры изображений фона из обучающей выборки, не содержащих пешеходов

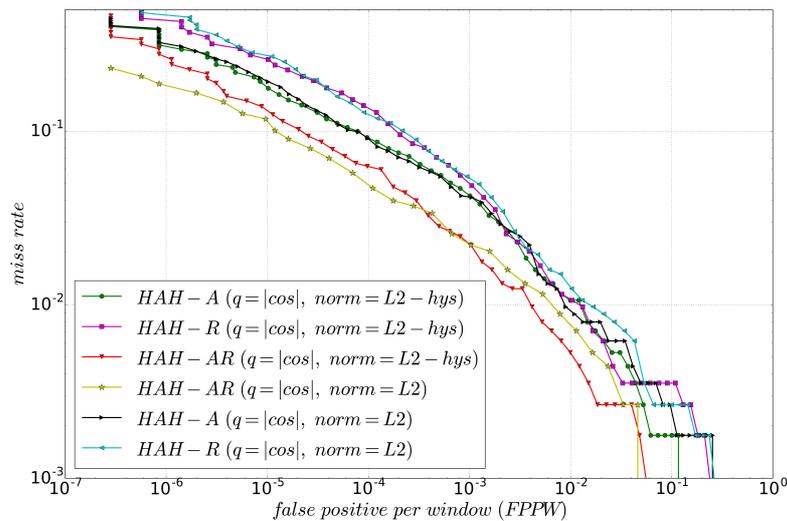


Рис. 6 Сравнение качества предлагаемых дескрипторов при различных вариантах нормализации гистограмм блоков

стве значений параметра метода C , который управляет балансом между максимизацией разделяющей полосы и суммарной ошибкой на объектах обучающей выборки, использовались значения из множества $\{0, 1, 10\}$. Выбиралось то значение параметра C , при котором достигался наилучший результат на тестовой выборке.

Полученные классификаторы были проверены на тестовой выборке изображений INRIA [12]. Тестовая выборка содержит 1132 изображения пешехода и 3 526 548 изображений фона размером 64×128 пикселей. Тестовые изображения фона были сформированы путем отбора областей из всевозможных окон с размерами 64×128 пикселей на 453 изображениях отрицательных примеров.

Для определения качества классификации были построены DET (detection error trade-off) кривые, где по оси абсцисс отложена доля неверно отвергнутых классификатором изображений (miss rate), а по оси ординат — доля неверно принятых решений (FPPW).

Для обнаружения влияния разных способов нормировки дескриптора при их обучении использовались два вида нормировки: L_2 (1) и L_2-hys (2), а в качестве функции вклада — функция (8).

На рис. 6 приведены DET кривые для различных вариантов нормализации дескрипторов блоков. Из него видно, что качество детектирования при использовании L_2 -нормы повышается на 32% при $10^{-4}FPPW$ по сравнению с L_2-hys нормализацией. Это связано с тем, что функция вклада (8) сглаживает вклад модуля градиента яркости по всем направлениям в каждой точке аккумуляторного пространства Хафа, что уменьшает вероятность всплесков. Таким образом, использование L_2 нормализации совместно с функцией вклада (8) имеет преимущества перед L_2-hys за счет того, что не теряется информативность признаков из-за порогового ограничения, определенного в функции (2) с целью фильтрации шумовых всплесков, которые могут появляться при использовании функции вклада (10).

На рис. 7 и в таблице отражены результаты для лучших вариантов предложенных в работе дескрипторов и НОГ-дескриптора [6].

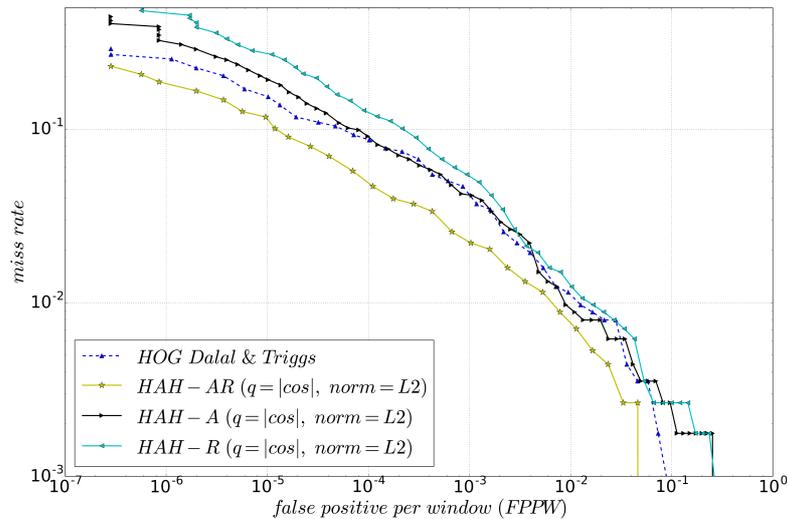


Рис. 7 Сравнение качества предлагаемых дескрипторов с HOG-дескриптором

Сравнение miss rate (MR) на уровне 10^{-4} FPPW
и скорости работы

Дескриптор	MR	t , мс
НОГ [6]	0,08	0,6
НАН-А($q = \cos $, $\text{norm} = L_2$)	0,08	1,4
НАН-Р($q = \cos $, $\text{norm} = L_2$)	0,13	1,5
НАН-АР($q = \cos $, $\text{norm} = L_2$)	0,04	2,2

На рис. 8 представлены результаты работы НАН-АР детектора на тестовых изображениях большого масштаба.

5 Заключение

В работе представлено описание метода построения текстурных признаков для обнаружения пешеходов на изображении. Описан подход построения нового дескриптора в аккумуляторном пространстве Хафа с использованием лучевого поперечного преобразования Радона градиентного поля изображения. Описаны несколько вариантов дескрипторов, построенных в рамках данного подхода. Представлены результаты вычислительного эксперимента с использованием предложенных дескрипторов и их сравнение с НОГ-дескриптором по производительности. Предложенный метод существенно уменьшает ошибки детектирования пешеходов по сравнению с НОГ-дескриптором за счет использования информации о положении контуров объекта. Основным недостатком предлагаемого подхода является довольно низкая производительность, вызванная использованием функции вклада, требующей интерполяции модуля градиента в каждой точке изображения во все бины по углам аккумуляторного пространства Хафа. Частично подобную проблему можно решить использованием других функций вклада, которые позволят уменьшить количество вычислительных операций.

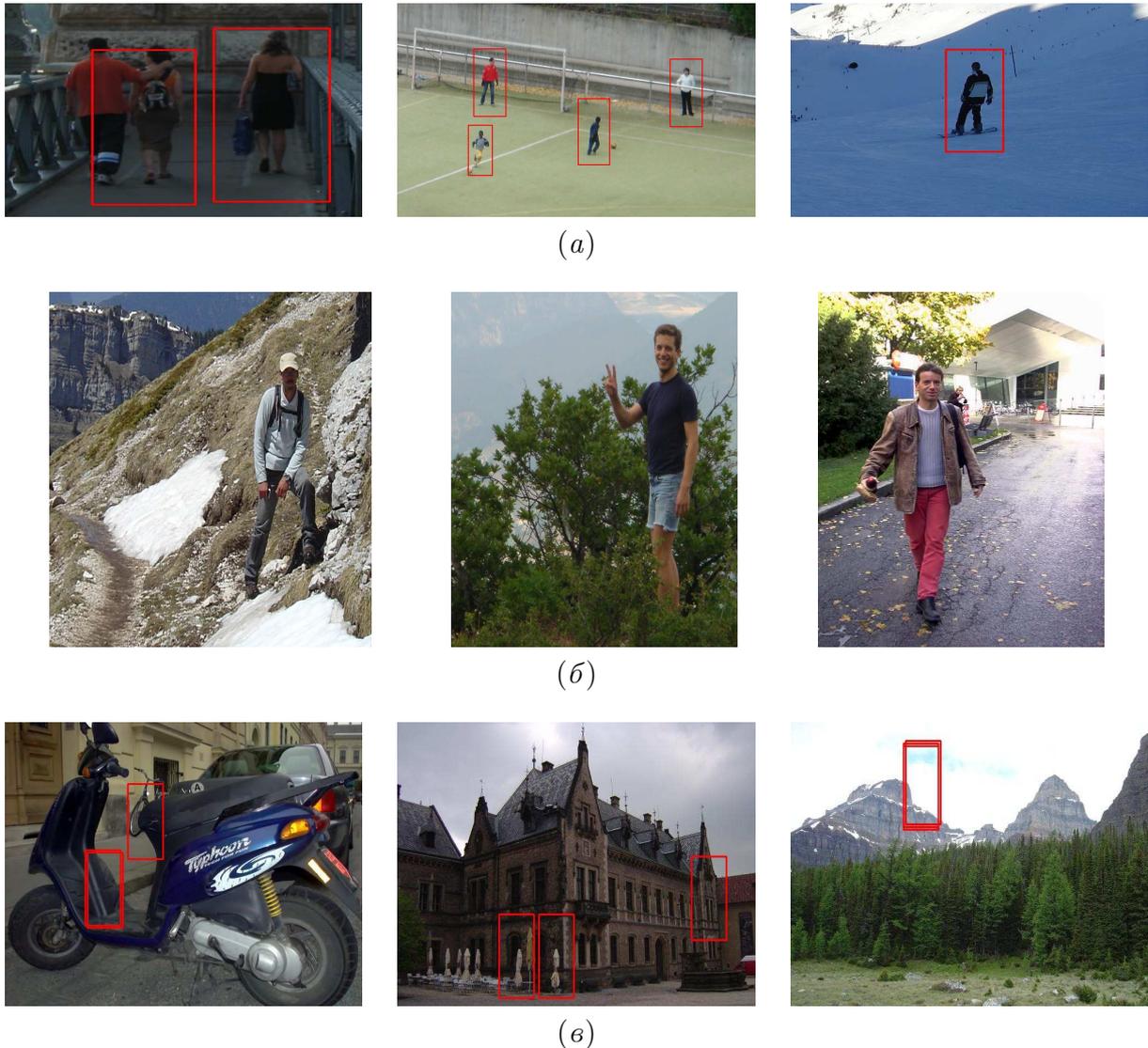


Рис. 8 Результаты положительной работы (а), пропущенные изображения (б) и ложные срабатывания (в) детектора NAN-AR при пороговом значении, соответствующем 10^{-4} FPPW

Литература

- [1] Гнеушев А. Н., Мурьшин А. Б. Адаптивный градиентный метод выделения контурных признаков объектов на изображениях реальных сцен // Изв. РАН. Теория и системы управления, 2003. № 6. С. 128–135.
- [2] Гнеушев А. Н. Построение и оптимизация текстурно-геометрической модели изображения лица в пространстве базисных функций Габора // Изв. РАН. Теория и системы управления, 2007. № 3. С. 85–96.
- [3] Гнеушев А. Н. Оптимизация текстурно-геометрической модели изображения для оценивания параметров лица // Автоматика и телемеханика, 2012. № 1. С. 159–168.
- [4] Viola P., Jones M. J. Rapid object detection using a boosted cascade of simple features // CVPR, 2001. Vol. 1. P. 511–518.
- [5] Viola P., Jones M. J., Snow D. Detecting pedestrians using patterns of motion and appearance // 9th Conference (International) on Computer Vision, 2003. Vol. 1. P. 734–741.

- [6] *Dalal N., Triggs B.* Histograms of oriented gradients for human detection // IEEE CVPR. — San Diego, CA, USA, 2005.
- [7] *Felzenszwalb P. F., Girshick B. R., McAllester D., Ramanan D.* Object detection with discriminatively trained part based models // IEEE T. Pattern Anal., 2010. Vol. 32. No. 9. P. 1627–1645.
- [8] *Sun D., Watanada J.* Detecting pedestrians and vehicles in traffic scene based on boosted HOG features and SVM // IEEE 9th Symposium (International) on Intelligent Signal Processing, 2015.
- [9] *Tabbone S., Wendling L., Salmon J. P.* A new shape descriptor defined on the Radon transform // Comput. Vis. Image Und., 2006. Vol. 102. No. 1. P. 42–51.
- [10] *Nacereddine, N., Tabbone S., Ziou D., Hamami L.* Shape-based image retrieval using a new descriptor based on the Radon and wavelet transforms // 20th Conference (International) on Pattern Recognition. — Istanbul, Turkey, 2010.
- [11] *Светов И. Е.* Формулы обращения для восстановления двумерных гармонических векторных полей по известным лучевым преобразованиям // Сиб. электрон. матем. изв., 2015. № 12. С. 436–446.
- [12] INRIA Person Dataset. <http://pascal.inrialpes.fr/data/human/>.
- [13] Open Source Computer Vision Library. <http://opencv.org/releases.html>.
- [14] *Воронцов К. В.* Лекции по методу опорных векторов. — 2007. <http://www.ccas.ru/voron/download/SVM.pdf>.

Поступила в редакцию 19.09.2017

Textural descriptor in the Hough accumulator space of the gradient field for detecting pedestrians*

N. A. Samsonov¹ and A. N. Gneushev²

nikita.samsonov@phystech.edu; gneushev@ccas.ru;

¹Moscow Institute of Physics and Technology, 9 Institutskiy per., Dolgoprudny, Moscow, Russia

²Federal Research Center “Computer Science and Control” of RAS, 44/2 Vavilova Str., Moscow, Russia

The problem of selecting features for recognizing pedestrians on an image is considered. The most popular and effective approach to selecting features is using descriptors based on Histograms of Oriented Gradients (HOG). In this paper, it is proposed to use the Hough accumulator space to generalize the HOG descriptor by obtaining projection not only of orientations, but also of the positions of the boundaries in the local area of the image Hough Accumulator Histograms (HAH). The Hough accumulator space is built on the basis of the beam Radon transform of the gradient field of the image. The proposed methods were tested together with linear support vector machine (SVM) classifier on the INRIA pedestrian database. The results of the experiment have shown the best separating ability of new descriptors and reduction of false detections in comparison with HOG.

Keywords: *HOG, histograms of oriented gradients; HAH, Hough accumulator histograms; descriptor; pedestrians; Radon field transform; Hough transform*

DOI: 10.21469/22233792.3.3.03

*The work was supported by the Russian Foundation for Basic Research, grant No. 16-07-01171.

References

- [1] Gneushev, A. N., and A. B. Murynin. 2003. Adaptive gradient method for extracting contour features of objects in images of real-world scenes. *J. Comput. Sys. Sci. Int.* 42(6):973–980.
- [2] Gneushev, A. N. 2007. Construction and optimization of a texture–geometric model of a face image in the space of basic Gabor functions. *J. Comput. Sys. Sci. Int.* 46(3):418–428.
- [3] Gneushev, A. N. 2012. Optimization of the texture–geometric image model for estimation of the face parameters. *Autom. Rem. Contr.* 73(1):144–152.
- [4] Viola, P., and M. J. Jones. 2001. Rapid object detection using a boosted cascade of simple features. *CVPR* 1:511–518.
- [5] Viola, P., M. J. Jones, and D. Snow. 2003. Detecting pedestrians using patterns of motion and appearance. *9th Conference (International) on Computer Vision*. 1:734–741.
- [6] Dalal, N., and B. Triggs. 2005. Histograms of oriented gradients for human detection. *IEEE CVPR*. San Diego, CA.
- [7] Felzenszwalb, P. F., B. R. Girshick, D. McAllester, and D. Ramanan. 2010. Object detection with discriminatively trained part based models. *IEEE T. Patt. Anal.* 32(9):1627–1645.
- [8] Sun, D., and J. Watanada, 2015. Detecting pedestrians and vehicles in traffic scene based on boosted HOG features and SVM. *IEEE 9th Symposium (International) on Intelligent Signal Processing*.
- [9] Tabbone, S., L. Wendling, and J. P. Salmon. 2006. A new shape descriptor defined on the Radon transform. *Comput. Vis. Image Und.* 102(1):42–51.
- [10] Nacereddine, N., S. Tabbone, D. Ziou, and L. Hamami. 2010. Shape-based image retrieval using a new descriptor based on the Radon and wavelet transforms. *20th Conference (International) on Pattern Recognition*. Istanbul, Turkey.
- [11] Svetov, I. 2015. Formuly obrashcheniya dlya vosstanovleniya dvumernykh garmonicheskikh vektornukh poley po izvestnym luechym preobrazovaniyam [Inversion formulas for recovering the harmonic 2D-vector fields by known ray transform]. *Sib. Electron. Math. Izv.* 12:436–446.
- [12] INRIA Person Dataset. Available at: <http://pascal.inrialpes.fr/data/human/> (accessed June 4, 2017).
- [13] Open Source Computer Vision Library. Available at: <http://opencv.org/releases.html> (accessed May 16, 2017).
- [14] Vorontsov, K. V. 2007. Lectures on support vector machine. <http://www.ccas.ru/voron/download/SVM.pdf> (accessed June 25, 2017).

Received September 19, 2017