

РАСПОЗНАВАНИЕ ДОРОЖНЫХ ЗНАКОВ С ПОМОЩЬЮ МЕТОДА ОПОРНЫХ ВЕКТОРОВ И ГИСТОГРАММ ОРИЕНТИРОВАННЫХ ГРАДИЕНТОВ

Лисицын С.О., Байда О.А.

Самарский государственный аэрокосмический университет имени академика С.П. Королёва
(национальный исследовательский университет)

Аннотация

В работе рассмотрена задача распознавания изображения дорожных знаков с помощью метода опорных векторов и признакового описания на основе гистограмм ориентированных градиентов. Для построения системы распознавания используются две ранее представленные в литературе формулировки задачи обучения классификатора на основе опорных векторов: формулировка Вестона–Уоткинса, сведённая с помощью преобразования Кеслера к более изученной задаче оптимизации, и формулировка Краммера–Зингера, допускающая применение модифицированного метода покоординатного спуска. Для построения спрямляющих пространств используются ядерные функции Йенсена–Шеннона и пересечения гистограмм. В качестве вычислительно эффективной линейной аппроксимации признакового пространства решающей функции предлагается использование отображений в пространство со скалярным произведением, соответствующим однородной аддитивной ядерной функции. Представлен сравнительный анализ рассматриваемых методов на экспериментальных данных.

Ключевые слова: распознавание образов, гистограммы ориентированных градиентов, метод опорных векторов.

Введение

Важнейшей практической проблемой распознавания образов является отбор или построение информативных признаков. Простейшим описанием изображения является вектор значений яркостей в каждом пикселе изображения. Недостатки подобного описания приводят к необходимости построения альтернативных систем признаков, обладающих некоторыми качествами, часто связанными с природой самих изображений. На данный момент в области компьютерного зрения разработано множество алгоритмов построения дескрипторов изображений, включая SIFT [1], LBP [2], HOG [3], PCA-SIFT [4].

В работе рассматривается задача распознавания изображений дорожных знаков с помощью метода опорных векторов. Для построения системы признаков используется метод, основанный на гистограммах ориентированных градиентов. Приведены результаты экспериментального анализа предложенного метода распознавания дорожных знаков.

Постановка задачи

Будем рассматривать стандартную задачу классификации. Пусть имеется некоторое пространство объектов X произвольной природы и определено функциональное отображение

$$f : X \rightarrow Y,$$

ставящее в соответствие каждому объекту некоторый «ответ» из множества Y . При этом на некотором подмножестве пространства объектов $X_i \subset X$ известны значения этого отображения f . Множество пар $\{(x, y) \mid x \in X_i, y = f(x) \in Y\}$ будем называть обучающей выборкой. Решением задачи классификации является некоторая функция f^* , аппроксимирующая отображение f на всём множестве X . При

этом в рамках теории распознавания образов решение задачи классификации должно удовлетворять минимуму эмпирического или структурного риска [5].

Метод построения признаков на основе гистограмм ориентированных градиентов был впервые представлен в работе [3]. Изначально данный метод построения признаков был предназначен для задач обнаружения фигур людей на изображениях, однако дальнейшие исследования [6,7] показали эффективность метода для множества других задач классификации. Приведём описание алгоритма построения признакового вектора с помощью гистограмм ориентированных градиентов. Построение вектора признаков может быть разделено на следующие этапы:

1. Нормализация яркости, контраста и других параметров монохромного изображения $I_{i,j}$, $i = 0, \dots, M, j = 0, \dots, N$. Данный этап позволяет достичь устойчивости системы признаков по отношению к изменениям яркости и контрастности изображения.
2. Вычисление градиента изображения по горизонтальному и вертикальному направлению. Для аппроксимации градиента изображения могут быть использованы различные дискретные дифференциальные операторы, такие как оператор Собеля, оператор Превитта или оператор Щарра [8]. Отметим, что выбор оператора, аппроксимирующего градиент изображения, оказывает влияние на качество классификации.
3. Вычисление для каждого (i, j) пикселя изображения направления $\Theta_{i,j}$ и величины $G_{i,j}$ градиента:

$$\Theta_{i,j} = \arctan \left(\frac{G_{i,j}^y}{G_{i,j}^x} \right);$$

$$G_{i,j} = \sqrt{(G_{i,j}^x)^2 + (G_{i,j}^y)^2}.$$

Пример полей изображения и соответствующих ему полей $\Theta_{i,j}$ и $G_{i,j}$ приведён на рис. 1.

4. Построение для каждой непересекающейся ячейки $C_{m,n}$ из (w,v) пикселей изображения (рис. 1) гистограммы направлений градиентов по t направлениям $(0, 2\pi/t, \dots, 2\pi)$ следующим образом:

$$H_{m,n} = \left\{ \sum_{\left| \Theta_{i,j} - \frac{2\pi k}{t} \right| \leq \frac{\pi}{t}} |G_{i,j}| \mid (i,j) \in C_{m,n}, k = 0, \dots, t \right\}.$$

Размер ячейки и количество направлений градиентов t являются важными параметрами метода, существенно влияющими на информативность признаков и их эффективность для задач классификации. Пример разбиения на ячейки и вычисления гистограмм приведён на рис. 2. Отметим, что при нормализации гистограмм в ячейках допустимо использование различных размеров ячеек среди изображений.

5. Нормализация гистограмм в блоках $B_{q,p}$, содержащих ячейки:

$$H'_{m,n} = H_{m,n} / \left\| \bigcup_{H_{m,n} \in B_{q,p}} H_{m,n} \right\|.$$

Отметим, что для блоков гистограмм допустимо пересечение между собой.

6. Формирование вектора признаков из нормализованных гистограмм в ячейках изображения:

$$HOG(I) = [H'_{0,0} \quad \dots \quad H'_{m,n} \quad \dots].$$

Приведённая схема построения признаков легко может быть обобщена для цветных изображений с K различными цветовыми каналами $I_{i,j,k}$, $i = 0, \dots, M$, $j = 0, \dots, N$, $k = 0, \dots, K$ простым объединением гистограмм ориентированных градиентов для каждого из изображений $I_{i,j,k}$, $k = 0, \dots, K$. Однако в данной работе предлагается использование более информативных поэлементных максимальных значений гистограмм среди всех цветовых каналов.

Отметим, что построение гистограмм по ячейкам постоянного размера обладает существенным недостатком: так, изображения различного разрешения должны быть приведены к некоторому общему среди выборки изображений разрешению. Это требование существенно повышает вычислительную сложность алгоритма распознавания.

В данной работе предлагается подход к распознаванию изображений, основанный на методе опорных векторов и гистограммах ориентированных градиентов с постоянным общим количеством, но разным размером ячеек.

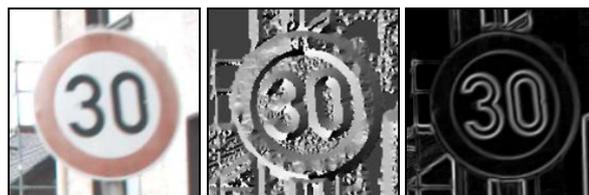


Рис. 1. Изображения и соответствующие им поля направлений и величины градиента

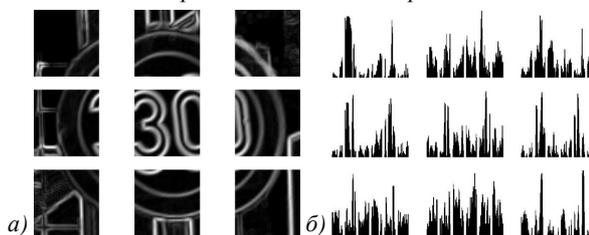


Рис. 2. Пример разбиения поля величины градиента на ячейки (а) и построения гистограмм величин по направлениям (б)

Метод опорных векторов

Метод опорных векторов является прямым развитием метода обобщённых портретов, разработанного группой математиков под руководством Владимира Вапника [5]. Приведём краткое описание метода. Для этого рассмотрим задачу бинарной классификации и зададим множество ответов $Y = \{-1, +1\}$. Для распознавания будем использовать следующее дискриминантное решающее правило¹:

$$f(x) = \text{sgn}(\langle w, x \rangle + b),$$

где разделяющая гиперплоскость определяется из решения следующей задачи оптимизации:

$$\min \left[\frac{1}{2} \|w\|^2 + C \sum_i \xi_i \right] \tag{1}$$

при следующих ограничениях:

$$f(x_i) [\langle w, x_i \rangle + b] \geq 1 - \xi_i, \quad \xi_i \geq 0.$$

Решение задачи, двойственной к задаче (1), определяет положение гиперплоскости как следующую линейную комбинацию векторов обучающей выборки:

$$f(x) = \text{sgn} \left(\sum_l \alpha_l \langle x_l, x \rangle + b \right), \tag{2}$$

где коэффициенты α_l являются множителями Лагранжа, а векторы $x_l \in X_l$ с $\alpha_l \neq 0$ называют *опорными векторами*. При этом решающее правило по-прежнему может быть сформулировано в терминах вектора нормали разделяющей гиперплоскости:

$$f(x) = \text{sgn} \left(\sum_l \alpha_l \langle x_l, x \rangle + b \right) = \text{sgn}(\langle w, x \rangle + b).$$

¹ При этом будем считать, что на векторах обучающей выборки x_i значения $f(x_i) = y_i$.

Известно, что в случае K классов может быть построено K решающих правил, разделяющих векторы каждого класса и векторы остальных классов. Формально, согласно [9], решающее правило может быть представлено в следующем виде:

$$f_m(x) = \text{sgn}(\langle w_m, x \rangle + b_m) = \text{sgn}\left(\sum_i \alpha_i^m \langle x_i^m, x \rangle + b_m\right).$$

Тогда решение о принадлежности к классу может быть найдено из следующего соотношения:

$$f(x) = \arg \max_m f_m(x),$$

а задача оптимизации, аналогичная (1), следуя [10], может быть сформулирована следующим образом:

$$\min \left[\frac{1}{2} \sum_{m=1}^k \|w_m\|^2 + C \sum_i \sum_{m \neq c_i} \xi_i^m \right] \quad (3)$$

при следующих ограничениях:

$$\langle w_{f(x_i)}, x_i \rangle + b_{f(x_i)} - (\langle w_m, x_i \rangle + b_m) \geq 1 - \xi_i^m, \\ \forall m \neq f(x_i), \xi_i^m \geq 0, \forall i, \forall m.$$

Решение задачи (3) в исходном виде, как правило, является чрезвычайно трудоёмким. Преобразовав задачу (3) к задаче:

$$\min \left[\frac{1}{2} \sum_{m=1}^K (\|w_m\|^2 + b_m^2) + C \sum_i \sum_{m \neq f(x_i)} \xi_i^m \right] \quad (4)$$

при тех же ограничениях:

$$\langle w_{f(x_i)}, x_i \rangle + b_{f(x_i)} - (\langle w_m, x_i \rangle + b_m) \geq 1 - \xi_i^m, \\ \forall m \neq f(x_i), \xi_i^m \geq 0, \forall i, \forall m,$$

используя преобразование Кеслера, можно свести её к более простой [10]. Для нахождения решения такой задачи оптимизации может быть использован модифицированный метод Митчелла-Демьянова-Малозёмова (МДМ-метод), представленный в работе [11]. Применимость метода МДМ для решения задачи оптимизации (4) изучена в [10].

Другим способом формулировки задачи обучения классификатора является метод, предложенный в работе [12]. Используя этот метод, задача оптимизации принимает следующий вид:

$$\min \left[\frac{1}{2} \sum_{m=1}^K \|w_m\|^2 + C \sum_i \xi_i \right] \quad (5)$$

при ограничениях:

$$\langle w_{f(x_i)}, x_i \rangle - \langle w_m, x_i \rangle \geq 1 - \delta_{f(x_i), m} - \xi_i.$$

В частности, при $m = f(x_i)$ условие принимает вид:

$$\xi_i \geq 0, \forall i.$$

Данный метод во многих случаях оказывается предпочтительнее в связи с тем, что он требует меньше вычислительных затрат на этапе обучения. В част-

ности, существуют эффективные реализации модифицированного метода покоординатного спуска, решающие задачу (5) в её двойственной формулировке [13].

Отметим, что порог b в исходной формулировке (5) является нулевым, что понижает качество классификации. Для решения этой проблемы задачу (5) можно переформулировать следующим образом с помощью регуляризованного порога:

$$\min \left[\frac{1}{2} \sum_{m=1}^K (\|w_m\|^2 + b_m^2) + C \sum_i \xi_i \right] \quad (6)$$

при ограничениях:

$$\langle w_{f(x_i)}, x_i \rangle + b_{f(x_i)} - \langle w_m, x_i \rangle + b_m \geq 1 - \delta_{f(x_i), m} - \xi_i.$$

Отметим, что для решения задачи (6) при этом могут быть использованы те же методы, что и для решения задачи (5).

В современной формулировке метод опорных векторов также допускает так называемый *ядерный переход* (kernel trick). Этот переход позволяет достичь линейной разделимости выборки за счёт неявного отображения в пространство большей размерности и тесно связан с методом потенциальных функций, представленных в монографии [14]. Суть метода состоит в замене всех скалярных произведений в исходной формулировке задачи (2) или (3) на некоторую симметричную и неотрицательно определённую функцию $k(x, x')$.

Ядерный переход позволяет использовать «нестандартные» меры схожести векторов признаков, обладающие необходимыми для решения задачи свойствами. Так, одной из наиболее распространённых ядерных функций является радиальная базисная функция:

$$k(x, x') = e^{-\gamma \|x - x'\|^2}, \quad x, x' \in X.$$

Однако предложенная система признаков основана на гистограммах, поэтому особый интерес представляют ядерные функции, учитывающие природу гистограмм. Естественной мерой схожести гистограмм является сумма их пересечения, поэтому в качестве меры схожести гистограмм ориентированных градиентов предложено использовать ядро пересечения гистограмм (histogram intersection kernel)

$$k_{HI}(x, x') = \sum_i \min\{|x_i|, |x'_i|\},$$

впервые введённое в работе [15] для классификации изображений с помощью цветových гистограмм. Отметим, что для корректной нормировки ядерной функции необходимо, чтобы $\forall x: \|x\|_1 = \sum_i |x_i| = 1$.

В работе [16] для задач с гистограммными признаками также показана эффективность ядерной функции, основанной на дивергенции Йенсена-Шеннона:

$$k_{JS}(x, x') = \sum_i \left[\frac{x_i}{2} \log_2 \left(\frac{x_i + x'_i}{x_i} \right) + \frac{x'_i}{2} \log_2 \left(\frac{x_i + x'_i}{x'_i} \right) \right].$$

Быстрая классификация с использованием ядерной функции пересечения гистограмм

Отметим, что ядерная функция пересечения гистограмм является *аддитивной* по отношению к компонентам вектора. В работе [17] было показано, что возможно переформулировать решающее правило (3) следующим образом:

$$\begin{aligned} f(x) &= \operatorname{sgn} \left(\sum_i \alpha_i k_{HI}(x_i, x) + b \right) = \\ &= \operatorname{sgn} \left(\sum_i \alpha_i \sum_i \min \{ |x_i^l|, |x_i| \} + b \right) = \\ &= \operatorname{sgn} \left(\sum_i \sum_i \alpha_i \min \{ |x_i^l|, |x_i| \} + b \right) = \\ &= \operatorname{sgn} \left(\sum_i h_i(x_i) + b \right), \end{aligned}$$

где функция $h_i(x_i) = \sum_i \alpha_i \min \{ |x_i^l|, |x_i| \}$ может быть выражена как

$$h_i(x_i) = \sum_{x_i^l < x_i} \alpha_i x_i^l + \left[\sum_{x_i^l \geq x_i} \alpha_i \right] x_i. \tag{7}$$

Выражение (7), приведённое в [17], позволяет вычислять точное значение функции $h_i(x_i)$ с помощью бинарного дерева сортировки за время, сопоставимое с $O(\ln L)$, где L – общее количество опорных векторов. Для этого в каждом узле дерева должны будут содержаться предварительно вычисленные $\sum_{l < k} \alpha_l x_i^l$ и $\sum_{l \geq k} \alpha_l$, что несколько увеличивает требования к памяти.

Покажем, что аналогичная техника может быть применена для задачи с несколькими классами. Для этого выразим решающее правило для класса m :

$$\begin{aligned} f^m(x) &= \sum_i \alpha_i^m k_{HI}(x_i^m, x) + b_m = \\ &= \sum_i \sum_i \alpha_i^m \min \{ |x_i^{l,m}|, |x_i| \} + b_m = \\ &= \sum_i h_i^m(x_i) + b_m. \end{aligned}$$

Объединяя вычисленные $\sum_{l < k} \alpha_l^m x_i^{l,m}$ и $\sum_{l \geq k} \alpha_l^m$ в единую структуру данных, допускающую сортировку по отношению к значениям $x_i^{l,m}$, для вычисления каждой из функций $h_i^m(x_i)$ требуется количество операций, сопоставимое с $O\left(\ln \sum_i L_i\right)$, где L_i – количество опорных векторов i -го классификатора. С помощью приведённой схемы может быть построено эффективное с вычислительной точки зрения решающее правило.

Явные отображения спрямляющих пространств однородных аддитивных ядерных функций

В статье [18] предложен эффективный метод явного построения признаков пространств со скалярным произведением, соответствующим ядерной функции в заданном пространстве признаков. Данный метод позволяет использовать линейные методы (в частности, метод опорных векторов в случае линейной ядерной функции), обладающие существенно меньшей вычислительной сложностью классификации, что позволяет улучшить общее быстродействие распознавания знаков.

Согласно [18] будем рассматривать однородные ядерные функции, удовлетворяющие равенству:

$$k(cx, cy) = c^\gamma k(x, y) \quad \forall c.$$

При этом к такому классу функций относится рассмотренная в работе ядерная функция пересечения гистограмм.

В силу произвольности c , положив $c = 1/\sqrt{xy}$, может быть получена следующая форма ядерной функции:

$$k(x, y) = (xy)^{\gamma/2} \mathcal{K}(\log y - \log x),$$

где \mathcal{K} называют сигнатурой ядерной функции. Представление однородной ядерной функции через её сигнатуру также позволяет определить её спектральную функцию с помощью теоремы Бохнера [19]. Спектр ядерной функции тогда может быть найден обратным преобразованием Фурье:

$$\kappa(\omega) = \frac{1}{2\pi} \int_{\mathbb{R}} \exp\{i\langle \omega, \lambda \rangle\} \mathcal{K}(\lambda) d\lambda.$$

Данное представление позволяет определить значения признаков в пространстве частот спектра:

$$\begin{aligned} k(x, y) &= (xy)^{\gamma/2} \int_{-\infty}^{+\infty} e^{-i\omega \lambda} \kappa(\omega) d\omega = \\ &= \int_{-\infty}^{+\infty} \left(e^{-i\omega \log x} \sqrt{x^\gamma \kappa(\omega)} \right)^* \left(e^{-i\omega \log y} \sqrt{y^\gamma \kappa(\omega)} \right) d\omega. \end{aligned}$$

Полученное выражение, очевидно, является скалярным произведением признаков:

$$\Psi_\omega(x) = e^{-i\omega \log x} \sqrt{x^\gamma \kappa(\omega)}.$$

Признаки, выраженные в таком виде, зависят от непрерывной переменной частоты спектра, поэтому для явного формирования конечного числа признаков используется аппроксимация, полученная с помощью периодизации и дискретизации спектра. Таким образом, представленный метод позволяет вычислять явные значения признаков с соответствующим ядерной функции скалярным произведением для целого класса ядерных функций, в том числе и для используемых в работе.

Описание обучающей и тестовой выборки

Для экспериментального изучения предложенного метода была использована выборка German

Traffic Sign Recognition Benchmark (GTSRB), представленная в 2011 году для сравнительного анализа методов классификации дорожных знаков. Данная выборка создана на основе реальных цифровых фотографий дорожных знаков и содержит изображения, полученные в различных условиях освещённости, углов обзора. Обучающая выборка содержит 39209 аннотированных изображений, разделённых на 43 непересекающихся класса. Тестовая выборка содержит 12630 изображений. Кроме того, для проверки точности классификации для изображений тестовой выборки доступны их истинные классы. Примеры изображений обучающей и тестовой выборки приведены на рис. 3.



Рис. 3. Примеры изображений обучающей выборки: изображение предупреждающего знака «Дети» (а); изображение знака «Дорожные работы» (б); изображение знака «Остановка» (в)

Экспериментальный анализ

Основной характеристикой качества метода классификации является точность на тестовой выборке, то есть отношение правильно распознанных изображений к общему числу изображений. Кроме того, важным свойством метода классификации является его низкая вычислительная сложность, требуемая в режиме реального времени.

В табл. 1 показана зависимость точности классификации от используемого оператора градиента. Для данного эксперимента на этапе обучения была использована модифицированная реализация алгоритма, решающего задачу в формулировке Краммера. Параметр C при этом выбран с помощью скользящего контроля на обучающей выборке.

В табл. 2 приведена зависимость точности классификации от используемой ядерной функции. Для данного эксперимента использована реализация алгоритма, решающего задачу в формулировке, предложенной Вестоном. Аналогично предыдущему эксперименту параметр регуляризации выбран с помощью скользящего контроля.

В табл. 3 приведена сравнительная характеристика процента ошибки в зависимости от используемых методов построения признаков и методов распознавания.

Все алгоритмы и их модификации, использованные в экспериментальном анализе, реализованы в библиотеке алгоритмов машинного обучения с открытым исходным кодом Shogun toolbox [20]. В частности, для обучения классификатора на основе формулировки Вестона была использована реализация General Minimal Norm Problem SVM, а для обучения классификатора, соответствующего формулировке Краммера – модифицированная реали-

зация по координатного спуска, доступная в библиотеке LibLinear.

Таблица 1. Зависимость точности распознавания от оператора градиента

| Оператор градиента | Процент ошибки |
|--------------------|----------------|
| Щарра | 2,58 |
| Собея | 2,56 |
| Превитта | 2,52 |

Таблица 2. Зависимость точности распознавания от ядерной функции

| Ядерная функция | Процент ошибки |
|------------------------|----------------|
| Линейная | 7,52 |
| Гауссова | 3,85 |
| Пересечения гистограмм | 2,12 |
| Йенсена–Шеннона | 2,34 |

Таблица 3. Зависимость процента ошибки классификации от метода классификации и используемых признаков изображений

| Метод / Признаки | Метод опорных векторов | Линейный дискриминантный анализ | Метод k ближайших соседей |
|--|------------------------|---------------------------------|-----------------------------|
| Гистограммы ориентированных градиентов с переменным размером ячеек | 2,12 | 4,39 | 8,67 |
| Гистограммы ориентированных градиентов | 2,53 | 5,13 | 9,78 |
| Вектор яркостей признаков | 7,54 | 13,45 | 23,25 |

Заключение

В работе представлен подход к распознаванию локализованных дорожных знаков с помощью метода опорных векторов и гистограмм ориентированных градиентов. Предложен способ построения гистограмм ориентированных градиентов, не зависящий от размера изображения. Для задач со множеством классов обобщено быстрое решающее правило метода опорных векторов с ядерной функцией пересечения гистограмм. На экспериментальных данных показаны преимущества предложенного метода.

Дальнейшими направлениями исследований, представленных в работе, являются эффективные методы построения информативных признаков изображений и подходы к решению задач распознавания со множеством классов.

Благодарности

Работа выполнена при поддержке Министерства образования и науки (ГК № 07.514.11.4105) и РФФИ (проект № 11-07-12051-офи-м).

Авторы выражают признательность профессору Фурсову В.А. за ряд ценных замечаний и полезных обсуждений.

Литература

1. **Lowe, D.G.** Object recognition from local scale-invariant features / D.G. Lowe // Proceedings of the Seventh IEEE International Conference on Computer Vision. – 1999. – Vol. 2. – P. 1150-1157.
2. **Ojala, T.** Performance evaluation of texture measures with classification based on Kullback discrimination of distributions / T. Ojala, M. Pietikainen, D. Harwood // Proceedings of the 12th IAPR International Conference. – 1994. – Vol. 1 – P. 582-585.
3. **Dalal, N.** Histograms of Oriented Gradients for Human Detection. / N. Dalal, W. Triggs // IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR05. – 2005. – Vol. 1(3). – P. 886-893.
4. **Ke, Y.** PCA-SIFT : A More Distinctive Representation for Local Image Descriptors / Y. Ke, R. Srikthar. – Computer Vision and Pattern Recognition, 2004.
5. **Вапник, В.Н.** Теория распознавания образов (статистические проблемы обучения) / В.Н. Вапник, А.Я. Червоненкис. – М.: Главная редакция физико-математической литературы изд-ва «Наука», 1974. – 416 с.
6. **Wang, X.** An HOG-LBP Human Detector with Partial Occlusion Handling / X. Wang, T.X. Han // IEEE Computer Engineering. – 2009. – P. 32-39.
7. **Ruta, A.** In-vehicle camera traffic sign detection and recognition / A. Ruta, F. Porikli, S. Watanabe, Y. Li // Machine Vision and Applications. – 2011. – Vol. 22(2) – P. 359-375.
8. **Шапиро, Л.** Компьютерное зрение / Л. Шапиро, Дж. Стокман, под ред. С. Соколова, пер. с англ. Андрей Богуславский. – М.: Бином. Лаборатория знаний. – 2006. – 752 с.
9. **Bordes, A.** Solving multiclass support vector machines with LaRank / A. Bordes, L. Bottou, P. Gallinari, J. Weston // Proceedings of the 24th International Conference on Machine Learning. – 2007. – P. 89-96.
10. **Franc, V.** Multi-class Support Vector Machine / V. Franc, V. Hlavac // IEEE Communication. – 2002. – Vol. 2(1) – P. 5-8.
11. **Митчелл, Б.Ф.** Нахождение ближайшей к началу координат точки многогранника / Б.Ф. Митчелл, В.Ф. Демьянов, В.Н. Малоземов // Л.: Вестник ЛГУ. – 1971. – № 19. – С. 38-45.
12. **Crammer, K.** On the Algorithmic Implementation of Multiclass Kernel-based Vector Machines / K. Crammer, Y. Singer // Journal of Machine Learning Research. – Vol. 2. – P. 265-292.
13. **Keerthi, S.S.** A sequential dual method for large scale multiclass linear SVMs / S.S. Keerthi, S. Sundararajan, K.-W. Chang, C.-J. Hsieh, C.-J. Lin // Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining KDD. – 2008. – P. 408-413.
14. **Айзерман, М.А.** Метод потенциальных функций в теории обучения машин / М.А. Айзерман, Э.М. Браверман, Л.И. Розоноэр. – М.: Главная редакция физико-математической литературы изд-ва «Наука», 1970. – 384 с.
15. **Barla, A.** Histogram intersection kernel for image classification. / A. Barla, F. Odone, A. Verri // Proceedings of International Conference on Image Processing. – 2003. – Vol. 2(1). – P. 513.
16. **Martins, A.F.T.** Nonextensive entropic kernels / A.F.T. Martins, M.A.T. Figueiredo, P.M.Q. Aguiar, N.A. Smith, E.P. Xing // Proceedings of the 25th International Conference on Machine Learning. – 2008. – P. 640-647.
17. **Maji, S.** Classification using intersection kernel support vector machines is efficient / S. Maji, A.C. Berg, J. Malik // IEEE Conference on Computer Vision and Pattern Recognition. – 2008. – P. 1-8.
18. **Vedaldi, A.** Efficient Additive Kernels via Explicit Feature Maps / A. Vedaldi, A. Zisserman // IEEE Engineering. – 2011. – Vol. 10(3) – P. 3539-3546.
19. **Иосида, К.** Функциональный анализ / К. Иосида, пер. с англ. В.М. Волосова. – М.: Мир, 1967. – 616 с.
20. **Sonnenburg, S.** The SHOGUN Machine Learning Toolbox / S. Sonnenburg, G. Raetsch, S. Henschel, C. Widmer, J. Behr, A. Zien, F. De Bona // Journal of Machine Learning Research. – 2010. – Vol. 11(5). – P. 1799-1802.

References

1. **Lowe, D.G.** Object recognition from local scale-invariant features / D.G. Lowe // Proceedings of the Seventh IEEE International Conference on Computer Vision. – 1999. – Vol. 2. – P. 1150-1157.
2. **Ojala, T.** Performance evaluation of texture measures with classification based on Kullback discrimination of distributions / T. Ojala, M. Pietikainen, D. Harwood // Proceedings of the 12th IAPR International Conference. – 1994. – Vol. 1 – P. 582-585.
3. **Dalal, N.** Histograms of Oriented Gradients for Human Detection. / N. Dalal, W. Triggs // IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR05. – 2005. – Vol. 1(3). – P. 886-893.
4. **Ke, Y.** PCA-SIFT : A More Distinctive Representation for Local Image Descriptors / Y. Ke, R. Srikthar. – Computer Vision and Pattern Recognition, 2004.
5. **Vapnik, V.N.** Pattern recognition theory / V.N. Vapnik, A.Ya. Chervonenkis // Moscow: “Nauka” Publisher, 1974. – 416 p. – (In Russian).
6. **Wang, X.** An HOG-LBP Human Detector with Partial Occlusion Handling / X. Wang, T.X. Han // IEEE Computer Engineering. – 2009. – P. 32-39.
7. **Ruta, A.** In-vehicle camera traffic sign detection and recognition / A. Ruta, F. Porikli, S. Watanabe, Y. Li // Machine Vision and Applications. – 2011. – Vol. 22(2) – P. 359-375.
8. **Shapiro, L.** Computer vision / L. Shapiro, G. Stockman // Prentice Hall, 2001. – 608 p.
9. **Bordes, A.** Solving multiclass support vector machines with LaRank / A. Bordes, L. Bottou, P. Gallinari, J. Weston // Proceedings of the 24th International Conference on Machine Learning. – 2007. – P. 89-96.
10. **Franc, V.** Multi-class Support Vector Machine / V. Franc, V. Hlavac // IEEE Communication. – 2002. – Vol. 2(1) – P. 5-8.
11. **Mitchell, B.F.** Finding point of polyhedron nearest to the origin / B.F. Mitchell, V.F. Demyanov, V.N. Malozemov // Leningrad: “Vestnik LGU”. – 1971. – N 19 – P. 38-45 – (In Russian).
12. **Crammer, K.** On the Algorithmic Implementation of Multiclass Kernel-based Vector Machines / K. Crammer, Y. Singer // Journal of Machine Learning Research. – Vol. 2. – P. 265-292.
13. **Keerthi, S.S.** A sequential dual method for large scale multiclass linear SVMs / S.S. Keerthi, S. Sundararajan, K.-W. Chang, C.-J. Hsieh, C.-J. Lin // Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining KDD. – 2008. – P. 408-413.
14. **Aizerman, M.A.** Theoretical foundations of the potential function method in pattern recognition learning / M.A. Aizerman, M.A. Braverman, L.I. Rosonoer // Moscow: “Nauka” Publisher, 1970. – 384 p. – (In Russian).

15. **Barla, A.** Histogram intersection kernel for image classification. / A. Barla, F. Odone, A. Verri // Proceedings of International Conference on Image Processing. – 2003. – Vol. 2(1). – P. 513.
16. **Martins, A.F.T.** Nonextensive entropic kernels / A.F.T. Martins, M.A.T. Figueiredo, P.M.Q. Aguiar, N.A. Smith, E.P. Xing // Proceedings of the 25th International Conference on Machine Learning. – 2008. – P. 640-647.
17. **Maji, S.** Classification using intersection kernel support vector machines is efficient / S. Maji, A.C. Berg, J. Malik // IEEE Conference on Computer Vision and Pattern Recognition. – 2008. – P. 1-8.
18. **Vedaldi, A.** Efficient Additive Kernels via Explicit Feature Maps / A. Vedaldi, A. Zisserman // IEEE Engineering. – 2011. – Vol. 10(3) – P. 3539-3546.
19. **Yosida, K.** Functional analysis / K. Yosida // Springer classic in mathematics, 1996. – 513 p.
20. **Sonnenburg, S.** The SHOGUN Machine Learning Toolbox / S. Sonnenburg, G. Raetsch, S. Henschel, C. Widmer, J. Behr, A. Zien, F. De Bona // Journal of Machine Learning Research. – 2010. – Vol. 11(5). – P. 1799-1802.

ROAD SIGN RECOGNITION USING SUPPORT VECTOR MACHINES AND HISTOGRAM OF ORIENTED GRADIENTS

S.O. Lisitsyn, O.A. Bayda

S.P. Korolyov Samara State Aerospace University (National Research University)

Abstract

In this paper we consider recognition of traffic signs using support vector machines (SVMs) and features based on histograms of oriented gradients (HOG). We approach training of classifier with two well developed multiclass support vector machine formulations proposed earlier by Weston & Watkins and Crammer & Singer. Feature space straightening is approached with Jensen-Shannon and histogram intersection kernels. Due to computational efficiency reasons we propose to make use of homogeneous kernel mapping presented recently. Comparative study based on German Road Traffic Sign Recognition Benchmark dataset shows the effectiveness of our approach.

Key words: machine learning, pattern recognition, histogram of oriented gradients, multiclass support vector machines.

Сведения об авторах



Лисицын Сергей Олегович, 1991 года рождения. Студент Самарского государственного аэрокосмического университета им. академика С.П. Королёва направления «Прикладная математика и информатика».

E-mail: lisitsyn.s.o@gmail.com.

Sergey Olegovich Lisitsyn (b. 1991) is a student of Samara State Aerospace University with major in Computer Science. His main interest is machine learning including support vector machines, nonlinear dimensionality reduction and computer vision case studies.



Байда Оксана Александровна, 1991 года рождения. Студентка Самарского государственного аэрокосмического университета им. академика С.П. Королёва направления «Прикладная математика и информатика».

E-mail: bayda.o.a@gmail.com.

Oksana Aleksandrovna Bayda (b. 1991) is a student of Samara State Aerospace University with major in Computer Science. Her interests are probabilistic pattern recognition, optimization theory and support vector machines.

Поступила в редакцию 4 марта 2012 г.