

УДК 004.93'1

Амосов Олег Семенович

ФГБОУ ВПО «Комсомольский-на-Амуре государственный технический университет»

Россия, Комсомольск-на-Амуре

Зав. кафедрой «Промышленная электроника»

Доктор технических наук, профессор

E-Mail: osa18@yandex.ru

Иванов Юрий Сергеевич

ООО МИП «Позитрон-ДВ»

при ФГБОУ ВПО «Комсомольский-на-Амуре государственный технический университет

Россия, Комсомольск-на-Амуре¹

Ведущий инженер-разработчик

E-Mail: ivanov_ys@garmvis.ru

Модифицированный алгоритм детекции лиц в видеопотоке и его программная реализация

Аннотация: Основной задачей видеоаналитики является детекция подвижных объектов в видеопотоке, таких как люди, транспортные средства, животные. Задача обнаружения лица на изображении является одной из важнейших задач, решаемых системами компьютерного зрения. Однако при попытке построения автоматической системы обнаружения лиц приходится сталкиваться с факторами, затрудняющими процесс детекции.

В данной статье предложен комплекс показателей для предварительной оценки классификатора с использованием его обучающей выборки. Предложен модифицированный алгоритм детекции подвижных объектов в видеопотоке, особенностью которого является адаптация классификатора под изменяющиеся характеристики сцены и новые образы объектов, и разработано программное обеспечение для детекции лиц людей.

Научная новизна состоит в применении классификатора OISVM и его онлайн обучении для детекции динамических объектов в видеопотоке, причем предварительное обучение происходит в автоматическом режиме, на основании данных полученных от автоматической внешней калибровки камеры.

Экспериментально доказано, что применение предложенного нами подхода увеличивает показатель правильных детекций объектов до 92,5% на зашумленных и искаженных изображениях, что позволяет применять предложенный алгоритм в реальных системах детекции образов.

С использованием тестовой выборки было доказано, что, в отличие от ранее используемых методов, предложенный алгоритм адаптируется к изменчивым условиям съемки и качеству поступающих образов.

Ключевые слова: детекция лиц; классификация; метод Виолы-Джонса; метод опорных векторов; онлайн обучение; автоматическая калибровка камеры.

Идентификационный номер статьи в журнале 23TVN314

¹ ООО МИП «Позитрон-ДВ», 681021, Хабаровский край, Комсомольск-на-Амуре, Октябрьский пр-т, 29, 38.

Введение

Основной задачей видеоаналитики является детекция подвижных объектов в видеопотоке, таких как люди, транспортные средства, животные. Задача обнаружения (детекции) лица на изображении (face detection) является одной из важнейших задач, решаемых системами компьютерного зрения. Однако при попытке построения автоматической системы обнаружения лиц приходится сталкиваться со следующими факторами, затрудняющими процесс детекции:

- индивидуальные особенности черт лица различных людей осложняют выбор признаков объектов детекции;
- небольшое изменение ориентации лица относительно камеры влечет за собой серьезное изменение изображения лица;
- часть лица может быть невидима (закрыта другими предметами) на изображении;
- условия съемки (освещение, цветовой баланс камеры, искажения изображения, качество изображения) в значительной степени влияют на изображение.

По указанным причинам в задаче обнаружения лиц существует потребность продолжать обучение на новых образах объектов. Такой процесс называется динамическим обучением (онлайн обучением – on-line learning). Главным преимуществом систем технического зрения, использующих такого рода обучение, является повышение эффективности детекции, так как с течением времени алгоритм совершает меньше ошибок на новых объектах. После детекции очередного объекта алгоритму передаётся метка класса этого объекта, которую он должен учесть, т. е. «дообучиться».

Учитывая все сказанное, можно утверждать, что задача разработки алгоритмов детекции лиц и других объектов, способных дообучению на новых данных, остается **актуальной**.

Известен метод Виолы-Джонса для детекции лица в кадре видеопотока [21]. В настоящей работе предложен способ улучшения качества работы алгоритмов детекции подвижных объектов в видеопотоке путем адаптации их под изменяющиеся характеристики сцены и новые образы объектов. Алгоритм основан на композиции метода Виолы-Джонса и последовательного обучения в методе опорных векторов (Incremental Learning Support Vector Machine).

Постановка и решение задачи детекции лица в кадре видеопотока традиционным методом

Пусть имеется видеопоследовательность, представляющая собой последовательность кадров. Каждый кадр является цифровым изображением.

В памяти компьютера изображение представляется двумерной матрицей пикселей I_t размером $weight \times height$, в которой каждый пиксель представлен некоторым неотрицательным числом от 0 до 255, если изображение черно-белое; t – номер кадра в видеопоследовательности;

Работа с видеопоследовательностью, в отличие от статичных изображений, позволяет выделить область движения, тем самым снижая размер области поиска.

После выделения области движения мы получаем матрицу C_t , путем вырезания ее из матрицы I_t . Размер полученной матрицы равен $w_c \times h_c$, где $w_c \leq weight$ и $h_c \leq height$.

По области движения требуется определить, если он есть, список прямоугольников, содержащих лицо человека:

$$rectangle_i = \{x_i, y_i, w_i, h_i\}, \quad (1)$$

где x_i, y_i – координаты центра i -го прямоугольника, w_i – ширина, h_i – высота.

Поиск объекта будет выполняться «сканирующим окном», проходящим последовательно по изображению с шагом 1 пиксель (слева-направо, сверху вниз) с изменением масштаба при каждом проходе (Рис. 1). Размер окна определяется его шириной W и высотой H .



Рис. 1. Прохождение сканирующего окна по изображению

Библиотека компьютерного зрения OpenCV [Электронный ресурс]. – Режим доступа: <http://docs.opencv.org>

При прохождении сканирующего окна по изображению, для каждого положения окна необходимо определить, содержится ли в окне лицо или нет. Это может быть выполнено путем отнесения анализируемой части изображения к одному из двух классов. Первый класс – наличие лица, второй класс – его отсутствие. Тогда задача детекции лица сводится к выбору эффективного классификатора, который был бы устойчив к перечисленным ранее факторам, затрудняющим его работу.



Рис. 2. Алгоритм детекции лица в видеопотоке.

Составлено (разработано) автором

Так как существует всего два класса {«не-лицо»; «лицо»} = {-1; +1}, то задача детекции лица в кадре видеопотока сводится к задаче двухклассовой классификации. Приведем традиционную постановку этой задачи.

Дано множество X – множество, в котором хранится описание объектов, Y – конечное множество классов. Функция $F: X \rightarrow Y$ называется решающей функцией (классификатор). Признак f объекта a – это отображение $f: X \rightarrow D_f$, где D_f – множество допустимых значений признака f . Если задан набор признаков f_1, \dots, f_n для некоторого объекта a , то вектор признаков x объекта $a \in X$ может быть определен так [21]:

$$x = (f_1(a), \dots, f_n(a)) \quad (2)$$

При этом вектор признаков допустимо отождествлять с самими объектами.

Классификатор F должен быть способен классифицировать произвольный объект $a \in X$. Оптимальным считается классификатор, который дает наименьшую вероятность ошибки при всех допустимых значениях x .

Для обучения классификатора F используется обучающая выборка, заданная множеством $D = \{(x_1, y_1), \dots, (x_m, y_m)\}$, где $y_i \in Y = \{-1; +1\}$. Общий алгоритм детекции лица в видеопотоке изображен на рисунке 2.

Недостатки классического алгоритма детекции лиц

Основной проблемой при детекции лиц является присутствие на анализируемом изображении большого количества мешающей либо несущественной шумовой информации. Для решения этой проблемы был предложен алгоритм двухпроходной классификации, особенностью которого являлась параллельная работа 2-х классификаторов [23]. Алгоритм обладал устойчивостью к изменчивым условиям съемки и качеству кадра, за счет чего показал высокие результаты при детекции номерных знаков транспортных средств.

Лицо человека может отличаться по таким критериям, как условия съёмки (расстояние до камеры, ракурс, освещённость), эмоциональное состояние (выражение лица – нейтральное, напряжённое, весёлое, грустное и т.д.), параметры оцифровки (разрешение, количество уровней яркости, контраст, насыщенность (интенсивность)), шум (вносимый как во время фотопроцесса, так и на стадии оцифровки изображения). Все эти факторы снижают процент правильной классификации.

Классический алгоритм детекции лица предполагает наличие предварительно обученного классификатора, а эффективность классификации будет зависеть от его обобщающей способности, под которой понимают способность классификатора выдавать правильные результаты не только для образов, участвовавших в процессе обучения, но и для любых новых, которые отсутствуют в обучающей выборке. Чем ближе вектор образа a к одному из векторов обучающей выборки, тем выше вероятность правильной классификации. Способность к обобщению классификатор может приобрести только за счет большого числа разнообразных комбинаций входных и целевых значений в примерах обучающего множества. При этом необходимо, чтобы выборка была репрезентативной относительно поступающих изображений с конкретной камеры наблюдения, то есть учитывала множество вариаций изменения признаков объекта a в зависимости от влияющих на него факторов. Очевидно, что размер выборки, являющейся репрезентативной для любых условий съемки может быть бесконечно огромным.

Так как формирование такой обучающей выборки, которая содержала бы все возможные вариации объекта невозможно, то для соблюдения вышеуказанного условия возникает потребность формировать обучающее множество с учетом конкретных условий съемки. Несмотря на то, что такой подход позволит получить наиболее эффективный классификатор, он влечет следующие трудности: 1) необходимо заранее знать условия съемки и характеристики объектов, что не всегда бывает выполнимо; 2) зная характеристики объектов,

необходимо сформировать обучающую выборку под конкретную камеру наблюдения, что является довольно трудоемким процессом; 3) на сформированной выборке необходимо обучить классификатор, при этом процесс обучения может занимать до нескольких недель.

Применение такого подхода в системе охранного видеонаблюдения не только не удобно, но и неприемлемо, так как конечный пользователь системы не станет измерять углы наклона камеры или освещенность, а также не будет формировать обучающую выборку, а тем более обучать классификатор.

Нами предлагается совместное использование двух классификаторов: первый из них построен на основе метода Виолы-Джонса, предварительно обучен и хорошо зарекомендовал себя в большинстве программных продуктов [21]; второй классификатор должен поддерживать онлайн обучение. Для его обучения предварительно формируется база эталонов, содержащая лица людей, снятых при различных условиях и углах наклона. При этом каждое изображение должно быть промаркировано, то есть содержать информацию об условиях съемки. Это необходимо для извлечения из базы только необходимых нам изображений.

Обучение второго классификатора будет происходить следующим образом:

1. Выполняется внешняя калибровка камеры, которая позволяет определить такие параметры как: углы наклона камеры, высоту установки, условия освещения;
2. Если параметры изображений соответствуют или близки идеальным, то есть углы наклона незначительны, освещение яркое, то необходимость в использовании второго классификатора отсутствует. Детекция выполняется только классификатором, построенном на основе метода Виолы-Джонса;
3. Если условия съемки отличны от идеальных, то на основании параметров, из предварительно сформированной базы эталонов выбираются изображения наиболее близкие к ним, таким образом формируется обучающая выборка с учетом условий съемки;
4. Полученная выборка используется для обучения второго классификатора;
5. В процессе работы алгоритма, все образы, полученные вторым классификатором также используются для его дообучения классификатора.

Постановка и решение задачи детекции лица в кадре видеопотока с использованием модифицированного алгоритма

Рассмотрим материал в такой последовательности:

1. Формирование обучающей выборки с учетом условий съемки для доказательства обоснованности предложенного подхода. Покажем зависимость между условиями съемки, обобщающей способностью и репрезентативностью выборки;
2. Опишем способы определения условий съемки;
3. Опишем метод опорных векторов, который будет использоваться в качестве второго классификатора;
4. Проанализируем методы последовательного обучения классификатора, основанного на методе опорных векторов (SVM);
5. Опишем синтезированный алгоритм детекции объектов с использованием метода Виолы-Джонса и последовательного обучения SVM.

Формирование обучающей выборки с учетом условий съемки

Изображение, поступающее на классификатор, может характеризоваться следующими признаками, влияющими на качество работы алгоритма:

- углы наклона и повороты лица;
- высота установки камеры и угол наклона относительно плоскости пола;
- уровень освещения;
- масштаб лица относительно размера изображения;
- зашумленность изображения.

Введем следующие обозначения для ряда показателей обрабатываемого изображения: α_h – боковой наклон лица относительно вертикальной оси, α_v – поворот лица относительно вертикальной оси, α_z – фронтальный угол наклона головы (зависит от высоты установки и угла наклона камеры), Lux – уровень освещённости сцены, $Noise$ – уровень зашумленности изображения (0 – идеальное изображение, 1 – бесконечно испорченное изображение).

Для удобства расчетов перейдем к относительным величинам $\{0;1\}$, для этого необходимо установить пороговые значения для каждой из характеристик. Учитывая, что объект детекции обладает симметрией, для параметров, характеризующих наклоны и повороты лица, значения допустимо брать по модулю. Детекция лица является первым шагом при решении более сложной задачи – идентификации (распознавания) лица. При этом пороговые значения должны быть установлены таким образом, чтобы детектированное лицо могло быть в дальнейшем распознано. Нами предлагаются значения, указанные в табл. 1.

Таблица 1

Параметр	Идеальное значение	Пороговое значение
α_h	0	$\pm 90^\circ$
α_v	0	$\pm 70^\circ$
α_z	0	$\pm 40^\circ$
Lux	>500 люкс	50 люкс
$Noise$	0	0.4

Составлено (разработано) автором

Воздействие внешних условий на объект a выражается следующим образом: $Q(a) = (\alpha_h, \alpha_v, \alpha_z, Lux, Noise)$. Изображение (рис. 3а), снятое в солнечный день, на котором лицо смотрит прямо на камеру, отсутствуют шумы, будет иметь следующие параметры $Q(a) = (0, 0, 0, 0, 0)$, такие изображения можем назвать идеальными. Изображение на рисунке 3б имеет параметры $Q(a) = (0.1, 0.57, 0.62, 0.2, 0)$, при этом стандартный классификатор Виолы-Джонса лица на этом изображении не обнаружил.



а)



б)

Рис. 3. Примеры изображений, снятых при разных условиях; а) – идеальное изображение; б) – изображение близкое к реальным условиям

Составлено (разработано) автором

Обучающая выборка представлена множеством объектов с метками их классов, тогда любую обучающую/тестирующую выборку, состоящую из n изображений, можно описать следующим вектором $\mathbf{D} = \{\mathbf{Q}(a_1), \mathbf{Q}(a_2), \dots, \mathbf{Q}(a_n)\} = \{(\alpha_{h1}, \alpha_{v1}, \alpha_{z1}, Lux_1, Noise_1), \dots, (\alpha_{hn}, \alpha_{vn}, \alpha_{zn}, Lux_n, Noise_n)\}$.

Для параметрической оценки обучающей выборки перейдем к следующей записи:

$$\mathbf{D} = \{(\alpha_{h1}, \dots, \alpha_{hn}), (\alpha_{v1}, \dots, \alpha_{vn}), (\alpha_{z1}, \dots, \alpha_{zn}), (Lux_1, \dots, Lux_n), (Noise_1, \dots, Noise_n)\}. \quad (3)$$

Для каждой из характеристик мы можем построить распределение количества изображений по значению параметра (рис. 4). Значения, встречающиеся наиболее часто будем называть модой. В результате вычисления моды по каждому из параметров, мы можем описать обучающую выборку так:

$$\mathbf{M_D} = \{M_{ah}, M_{av}, M_{az}, M_{Lux}, M_{Noise}\}, \quad (4)$$

где $M_{ah}, M_{av}, M_{az}, M_{Lux}, M_{Noise}$ мода по каждому из параметров

Вектор $\mathbf{M_D}$ будет описывать наиболее часто встречающиеся параметры изображений в выборке.

Минимальные и максимальные значения параметров изображений в выборке определим следующим образом:

$$\mathbf{D}_{min}^{max} = \{(\min_{\alpha_h}, \max_{\alpha_h}), (\min_{\alpha_v}, \max_{\alpha_v}), (\min_{\alpha_z}, \max_{\alpha_z}), (\min_{Lux}, \max_{Lux}), (\min_{Noise}, \max_{Noise})\} \quad (5)$$

Одним из основополагающих методов детекции лица человека является метод Виолы-Джонса. Данный метод основан на следующих принципах [21]: 1) используются изображения в интегральном представлении; 2) используются признаки Хаара; 3) используется бустинг [8] для выбора наиболее подходящих признаков для искомого объекта на данной части изображения; 4) все признаки поступают на вход классификатора, который даёт результат «верно» или «ложь»; 5) используются каскады признаков для быстрого отбрасывания окон, где не найдено лицо.

Классический алгоритм Виолы-Джонса для детекции лиц представлен в виде каскадного классификатора в виде XML файла, входящего в набор библиотеки компьютерного зрения OpenCV [25]. Классификатор был обучен с использованием базы лиц FERET [7]. С учетом предложенного подхода оценим обучающую выборку классификатора haarcascade_frontalface_alt2.xml.

Известны характеристики изображений базы FERET [7].

Оценим обучающую выборку согласно предложенным параметрам (рис. 4). Тогда:

$$\mathbf{D}_{min}^{max} = \{(0, 0.65), (0, 0.4), (0, 0.25), (0, 0.7), (0, 0.25)\}. \quad (6)$$

$$\mathbf{M_D} = (0, 0, 0, 0, 0) \quad (7)$$



Рис. 4. График распределения изображений по характеристике α_v базы FERET, S – область допустимых значений α_v , где $S_{av} < 1$; а) – количественная оценка б) – относительные величины.

Составлено (разработано) автором

Тестирующие изображения, параметры которых совпадают с параметрами *MDF* обучающей выборки, будем называть идеальными для данного обученного классификатора (Рис. 3а). Очевидно, что результат работы классификатора на идеальных изображениях, при условии, что они не входили в обучающую выборку, будет максимальным.

Протестируем классификатор на идеальных изображениях для классификатора *haarcascade_frontalface_alt2*. Уровень правильных детекций равен 96% (рис. 11)

Опишем зависимость точности классификации от характеристик тестирующей выборки $S_p(k) = [0, 1]$, где $p = \{\alpha_h, \alpha_v, \alpha_z, Lux, Noise\}$ один из параметров, k – значение этого параметра. Если $S_p(k) = 0$, значит уровень детекции от данного параметра не зависит, если $S_p(k) = 1$, значит при значении k уровень правильных детекций меньше либо равен 50%. Так для классификатора Виолы-Джонса $S_{av}(0.4) = 1$, тогда при значении параметра $\alpha_v \geq 0.4$, вероятность правильной классификации будет меньше 50%. Область допустимых значений α_v изображена заштрихованной областью (Рис. 5), в которой $S_{av} < 1$.

В ходе следующего эксперимента будем последовательно подавать тестирующие выборки, состоящие из изображений следующего вида:

$$\begin{aligned}
 M_D &= \{m, M_{av}, M_{az}, M_{Lux}, M_{Noise}\} \\
 M_D &= \{M_{ah}, m, M_{az}, M_{Lux}, M_{Noise}\} \\
 M_D &= \{M_{ah}, M_{av}, m, M_{Lux}, M_{Noise}\} \\
 M_D &= \{M_{ah}, M_{av}, M_{az}, m, M_{Noise}\} \\
 M_D &= \{M_{ah}, M_{av}, M_{az}, M_{Lux}, m\}
 \end{aligned} \tag{8}$$

где m будет изменяться от 0 до 1 с шагом 0,1. В результате серии экспериментов (по 11 на каждый параметр), мы получили область допустимых значений каждого параметра, где $S_p < 1$.

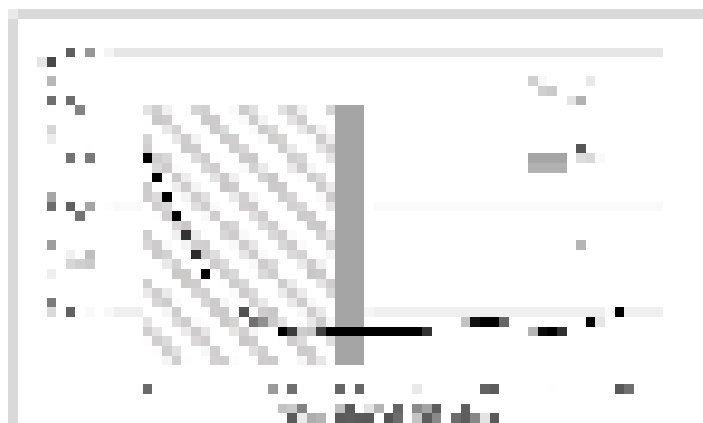


Рис. 5. Интерпретация ошибки классификатора

Составлено (разработано) автором

Объединив значения для всех параметров мы можем построить график характеризующий обучающую выборку и относительную эффективность классификатора для изображений обладающих определёнными параметрами (Рис 6.).

Также можно сделать вывод, что расстояние между значениями S_p и значениями вектора максимальных значений обучающей выборки будет характеризовать обобщающую способность классификатора.

Таким образом, мы можем сделать вывод, что для успешной работы классификатора необходимо, чтобы характеристики тестирующих изображений находились в пределах D_{min}^{max} . При этом чем ближе график M_D тестирующей выборки к графику M_D обучающей выборки, тем выше результат работы классификатора. Другими словами, обучающая выборка должна содержать наиболее часто встречающиеся вариации изображений, получаемых с конкретной камеры наблюдения, на которой классификатор будет работать, что на практике трудновыполнимо.

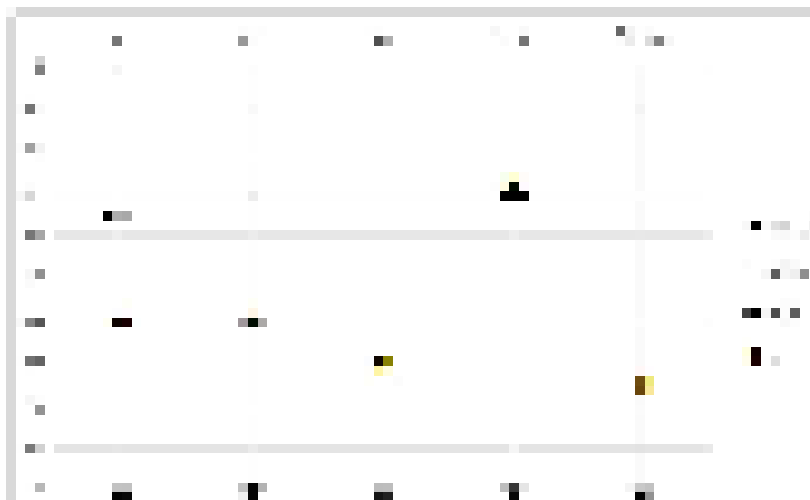


Рис. 6. Параметрическая оценка обучающей выборки

Составлено (разработано) автором

Решением данной проблемы может стать использование второго классификатора с другой область допустимых значений параметров, совпадающей с параметрами изображений, поступающих с камеры наблюдения. Проблемой становится отсутствие информации о

параметрах поступающих изображений. Использование n классификаторов, где $n > 2$, не является целесообразным, так как повысит требования к вычислительной мощности.

Выход заключается в применении последовательного дообучения второго классификатора путем добавления к нему новых образов детектируемых объектов. При таком подходе вектор M_{D2} будет постепенно приближаться к вектору значений параметров поступающих изображений Q .



Рис. 7. Увеличение области допустимых значений параметра α_v за счет использования 2-го классификатора

Составлено (разработано) автором

Классификатор Виолы-Джонса не поддерживает дообучение, значит, необходимо использовать второй классификатор, который не будет значительно уступать качеству работы метода Виолы-Джонса и при этом будет поддерживать онлайн обучение. Таким классификатором является метод опорных векторов [2]. Основной проблемой является необходимость предварительного обучения классификатора SVM на небольшом количестве образов. Причем эти образы должны соответствовать по параметрам изображению сцены. На основании информации об условиях съемки, мы можем сформировать обучающую выборку, используя предварительно сформированную базу изображений, о которой говорилось ранее. Таким образом ставится задача определения параметров изображения или внешней калибровки.

Способы определения условий съемки

Способ определения условий освещенности был изложен в [24]. В его основе лежат зависимости между типом освещения и цветовыми моделями для определения типа источника освещения с использованием интеллектуального видеоанализа.

Для определения углов наклона и поворотов камеры необходимо выполнить ее внешнюю калибровку. Под внешней калибровкой камеры понимают задачу определения ориентации и расположения камеры в пространстве по изображению сцены.

Процесс калибровки базируется на информации об объектах, которые снимаются камерой – их пропорций, ориентации, пространственного положения.

Как известно, изображение – это отображение сцены на плоскости. Модель зрительной системы камеры наблюдения (Рис. 8) состоит из:

1. прямоугольной системы координат изображения с целочисленными осями R и C , где координаты по R и C - это номера строк и столбцов пикселей на изображении;

2. прямоугольной системы координат камеры с вещественными осями X_c, Y_c, Z_c ;
3. прямоугольной мировой системы координат с вещественными осями X_w, Y_w, Z_w .

Оптический центр в системе камеры имеет координаты $(0, -f, 0)$. Плоскость изображения RC , моделирующая матрицу камеры, лежит в плоскости $X_c Y_c$, а ось Z_c является оптической осью камеры.



Рис. 8. Модель зрительной системы камеры наблюдения

Библиотека компьютерного зрения *OpenCV* [Электронный ресурс]. – Режим доступа: <http://docs.opencv.org>

Задача калибровки камеры заключается в том, чтобы определить, как расположены пиксели на плоскости изображения данной камеры относительно точек трёхмерного пространства, изображения которых и требуется получать с помощью камеры. Эта задача решается с помощью определения модели камеры C_w^I , которая позволяет сопоставить любой точки пространства (x^w, y^w, z^w) , заданной в мировой системе координат, точку на изображении (r, c) при произвольном масштабирующем элементе s :

$$\begin{bmatrix} s \cdot r \\ s \cdot c \\ s \end{bmatrix} = C_w^I \begin{bmatrix} x^w \\ y^w \\ z^w \\ 1 \end{bmatrix} \quad (9)$$

Внешние параметры описывают местоположение и ориентацию системы координат камеры в мировой системе координат. К этим параметрам относятся:

- параметры переноса $t = [t_x, t_y, t_z]^T$
- параметры поворота:

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} & 0 \\ r_{21} & r_{22} & r_{23} & 0 \\ r_{31} & r_{32} & r_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (10)$$

Параметры переноса описывают местоположение камеры в мировой системе координат, а параметры поворота – её ориентацию.

Матрицу R можно представить в виде произведения трёх матриц поворота

$$R = R_{ox} R_{oy} R_{oz}$$

$$R_{ox} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha & 0 \\ 0 & \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad R_{oy} = \begin{bmatrix} \cos \beta & 0 & \sin \beta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \beta & 0 & \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad R_{oz} = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 & 0 \\ \sin \gamma & \cos \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (11)$$

где R_{ox} – матрица поворота вокруг оси OX на угол α , R_{oy} – матрица поворота вокруг оси OY на угол β , R_{oz} – матрица поворота вокруг оси OZ на угол γ :

Таким образом, на самом деле независимых параметра поворота только 3: α, β, γ .

Преобразование T_w^C из мировой системы координат в систему камеры зависит только от внешних параметров и имеет вид:

$$T_w^C = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} x^c \\ y^c \\ z^c \\ 1 \end{bmatrix} = T_w^C \begin{bmatrix} x^w \\ y^w \\ z^w \\ 1 \end{bmatrix} \quad (12)$$

Преобразование T_w^C обратимо: $T_C^W = (T_w^C)^{-1}$

Композиция преобразований T_C^I и T_w^C является моделью камеры:

$$C_w^I = T_C^I T_w^C$$

Для автоматизации определения внешних параметров используется специальная маркерная фигура или шаблон в виде шахматной сетки (рис 9). Такой подход был предложен и подробно описан в [22]. В данной работе мы изложим основные положения на которых базируется предложенный метод.

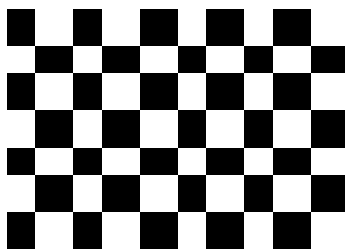


Рис. 9. Шаблон для автоматической внешней калибровки камеры

Составлено (разработано) автором

1. На принтере распечатывается шаблон для калибровки;
2. Фиксируется несколько стоп-кадров (чем количество, тем выше точность), на которых присутствует калибровочный объект;
3. На каждом изображении выполняется поиск калибровочного объекта с помощью детектора Харриса [10];
4. Выполняется перспективное преобразование плоскости калибровочного объекта в плоскость камеры (гомография).



Рис. 10. Дополненная модель зрительной системы из одной камеры: в снимаемую сцену поместили шаблон – рисунок с калибровочным узором.

Библиотека компьютерного зрения OpenCV [Электронный ресурс]. – Режим доступа: <http://docs.opencv.org>

Был проведен эксперимент, состоящий из двух этапов. В первом случае на поверхности пола поместили шаблон, таким образом, чтобы он попадал в поле зрения камеры. В качестве начала координат использовался один из углов шаблона. В процессе съемке ориентация и положение камеры относительно мировой системы координат изменялись (Рис. 11).

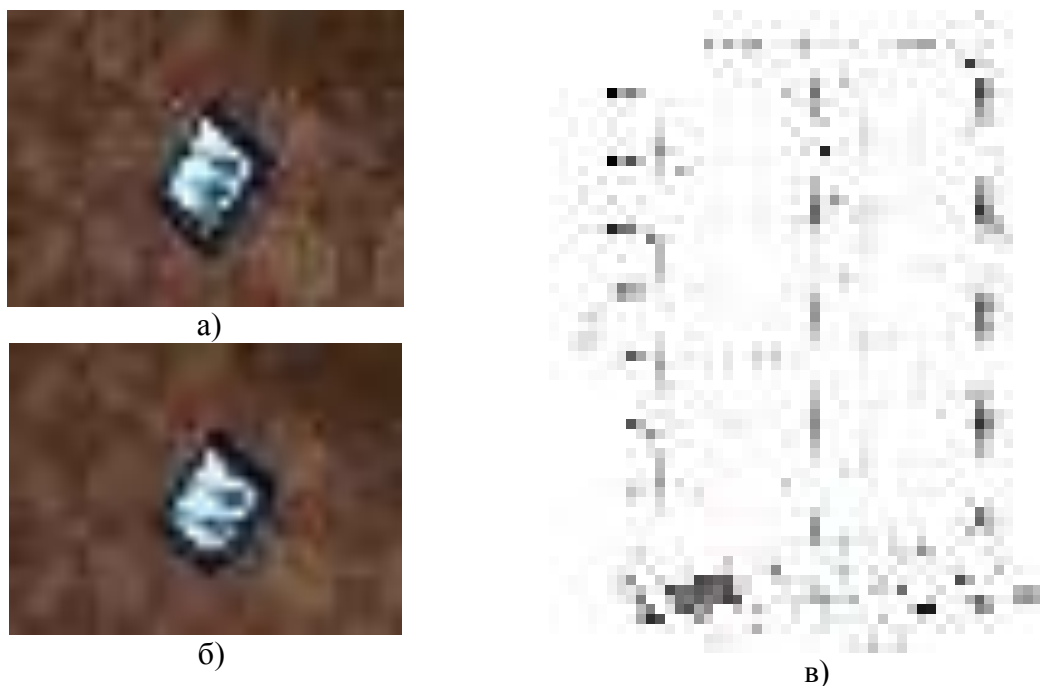


Рис. 11. Внешняя калибровка при неподвижном объекте; а), б) – изображения сцены, в) – смоделированное изображение
Составлено (разработано) автором

Во втором случае в качестве устройства видео съёмки была использована камера наблюдения, фиксированная в мировой системе координат, при этом за точку отсчета принят оптический центр камеры. В процессе съёмки изменялось положение шаблона и углы наклона шаблона (Рис. 12).

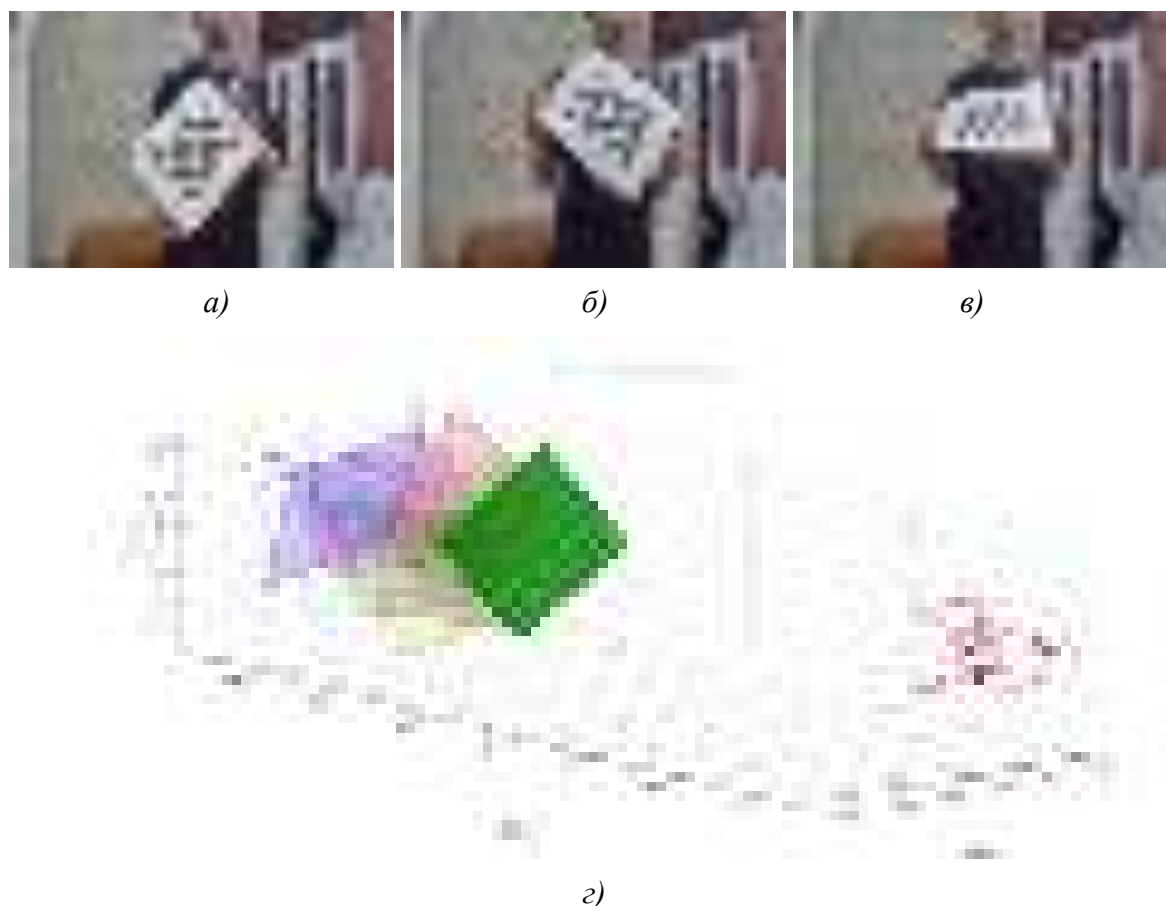


Рис. 12. Внешняя калибровка при неподвижной камере; а), б), в) – изображения сцены, г) – смоделированное изображение

Составлено (разработано) автором

С помощью описанного алгоритма были определены внешние параметры камеры и вычислено положение и наклон шаблона относительно камеры. Работа алгоритма была оценена с помощью визуального оценивания смоделированного изображения.

Нами предлагается следующая последовательность действий при начальной инициализации камеры наблюдения:

1. Определение условий освещенности и зашумленности;
2. Печать калибровочного объекта;
3. Оператор проносит калибровочный объект по наиболее вероятной траектории движения людей в кадре. Причем калибровочный объект рекомендуется держать в руках на уровне груди, сопоставив плоскость шаблона с положением головы;
4. На основе полученных кадров выполняется внешняя калибровка камеры и рассчитываются вероятные углы наклона и повороты лиц;

Таким образом, определив все параметры поступающих изображений, мы можем сформировать предварительную обучающую выборку для классификатора SVM.

Метод опорных векторов

Цель тренировки большинства классификаторов - минимизировать ошибку классификации на тренировочном наборе (называемую эмпирическим риском). В отличие от них, с помощью метода опорных векторов, можно построить классификатор, минимизирующий верхнюю оценку ожидаемой ошибки классификации (в том числе и для неизвестных объектов, не входивших в тренировочный набор). Применение метода опорных векторов к задаче обнаружения лица заключается в поиске гиперплоскости в признаковом пространстве, отделяющей класс изображений «лиц» от изображений «не-лиц».

Основная идея метода опорных векторов [2] — перевод исходных векторов в пространство более высокой размерности и поиск разделяющей гиперплоскости с максимальным зазором в этом пространстве. Две параллельных гиперплоскости строятся по обеим сторонам гиперплоскости, разделяющей наши классы. Разделяющей гиперплоскостью будет гиперплоскость, максимизирующая расстояние до двух параллельных гиперплоскостей (рис. 13). Алгоритм работает в предположении, что чем больше разница или расстояние между этими параллельными гиперплоскостями, тем меньше будет средняя ошибка классификатора. Это будет соответствовать наилучшей классификации.

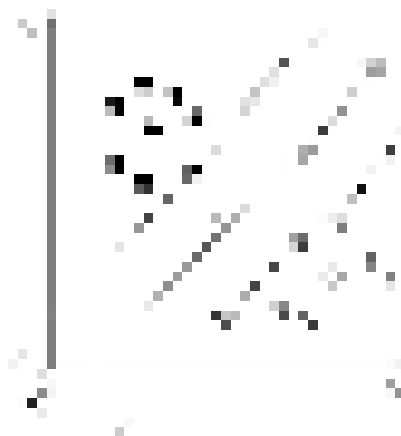


Рис. 13. Геометрическая интерпретация метода опорных векторов

Burges C. A Tutorial on Support Vector Machines for Pattern Recognition // Data Mining and Knowledge Discovery. – 1998. – Vol. 2 – P. 121-167

Рассмотрим задачу классификации на два непересекающихся класса, в которой объекты описываются n -мерными вещественными векторами: пусть $X = R^n$ — пространство объектов, $Y = \{-1; 1\}$ — множество классов. Дана обучающая выборка:

$$\{(x_1, y_1), \dots, (x_m, y_m)\}. \quad (13)$$

Функция (классификатор) $F: X \rightarrow Y$, сопоставляющая класс y произвольному признаковому описанию объекта x будет иметь следующий вид [2]:

$$F(x) = \text{sign}(\langle w, x \rangle + b), \quad (14)$$

где w — перпендикуляр к разделяющей гиперплоскости, параметр b равен по модулю расстоянию от гиперплоскости до начала координат. Если параметр b равен нулю, гиперплоскость проходит через начало координат, что ограничивает решение.

Последовательное обучение

Классический метод опорных векторов допускает предварительное (пакетное) обучение классификатора.

Задаче онлайн обучения классификатора SVM были посвящены работы [3, 13, 16, 18, 19]. Наиболее эффективное решение было предложено в работе [20]. Такой классификатор получил название ISVM. Недостатком данного метода является необходимость ручного вмешательства для остановки процесса обучения, а также отсутствует сравнение поступающих образов с уже присутствующими в базе, что может привести к переобучению классификатора.

Описанная проблема была решена в работе [17]. Далее опишем основную идею предложенного метода, который получил название OISVM.

Для классификации линейно неразделимых классов используется аппарат ядерных функций [16] для неявного проецирования векторов-признаков в пространство потенциально намного более высокой размерности, в котором классы могут оказаться линейно разделимы [4, 14]. Классификатор выглядит следующим образом [5]:

$$F(\mathbf{x}) = \sum_{i=1}^m \alpha_i y_i K(x, x_i) + b, \quad (15)$$

где $b \in \mathbb{R}$, $K(x_1, x_2) = \Phi(x_1) \cdot \Phi(x_2)$ функция ядра α_i — производные лагранжиана:

$$\min_{w, b} \left(\frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \xi_i^p \right), \quad (16)$$

при условии:

$$\begin{aligned} y_i (w \cdot x_i + b) &\geq 1 - \xi_i \\ \xi_i &\geq 0 \end{aligned}$$

где w — перпендикуляр к разделяющей гиперплоскости в пространстве где существует Φ , $\xi_i \in \mathbb{R}$ — дополнительные переменные, $C \in \mathbb{R}^+$ — параметр настройки метода, который позволяет регулировать отношение между максимизацией ширины разделяющей полосы и минимизацией суммарной ошибки (коэффициент штрафа), $p = \{1; 2\}$. В большинстве случаев, после тренировки, α_i равна 0, вектора для которых α_i отлична от нуля, являются опорными векторами. Ядро K может быть определено следующим образом $K_{ij} = K(x_i, x_j)$, где $i, j = 1, \dots, m$.

При последовательно обучении важно, чтобы добавляемые образы были линейно независимы от образов, испытываемых для построение функции классификации (15). Более подробно метод описан в [6]. Приведем основную идею алгоритма.

Обозначим индексы векторов в текущем базисе, после m тренировок, как β и новый образец как x_{m+1} . Алгоритм работает следующим образом:

1. проверить x_{m+1} линейную независимость базиса от базиса в пространстве признаков, и добавить его в β , в противном случае оставить β без изменений;
2. последовательно переобучить классификатор.

Для обучения используется известный метод [11, 12]. адаптированный для обучения в режиме онлайн. Алгоритм непосредственно минимизирует проблему (16), в отличие от стандартного способа решения двойственной задачи. Такой подход позволяет явно выбрать базисные векторы для использования. Используя выражение (16), предположим, что $p=2$ и необходимо преобразовать его для безусловной задачи. Пусть $D \subset \{1, \dots, m\}$, тогда безусловная задача формулируется следующим образом:

$$\min_{\beta} \left(\frac{1}{2} \beta^T K_{DD} \beta + \frac{1}{2} C \sum_{i=1}^m \max(0, 1 - y_i K_{iD} \beta)^2 \right), \quad (17)$$

где β есть вектор Лагранжевых коэффициентов, используемых в $F(x)$, соответствует α_i в первоначальной формулировке. Метод Ньютона с изменениями, внесенными [11, 12] может быть использован для решения (15) после добавления нового образа. Когда новый образ x_{m+1} добавляется, алгоритм выглядит следующим образом:

1. пусть $\phi = \{ i : 1 - y_i o_i > 0 \}$, где $o_i = K_{i\beta} \beta$ и β - вектор оптимальных коэффициентов с m учебными образами; если ϕ не изменилась, остановиться;
2. в противном случае, обновляем β так: $\beta - \gamma \mathbf{P}^{-1} g$, где $\mathbf{P}^{-1} = K_{\beta\beta} + CK_{\beta\phi} K_{\beta\phi}^T$ и $g = K_{\beta\beta} \beta - CK_{\beta\phi} (y_{\phi} - o_{\phi})$;
3. вернуться к шагу 1.

После шага 2, γ устанавливается в единицу.

Следует отметить, что существенным ограничением для применимости вышеупомянутого алгоритма обучения в режиме реального времени является необходимость наличия предварительно обученного классификатора, который впоследствии адаптируется. Как было указано ранее, такое требование бывает не всегда выполнимо.

Нами предлагается более эффективный подход, основанный на автоматическом процессе онлайн обучения классификатора OISVM без ручного вмешательства человека.

Синтезированный алгоритм детекции объектов с использованием метода Виолы-Джонса и последовательного обучения SVM

Предлагаемый модифицированный подход состоит из следующих этапов (рис. 14):

1. Определение параметров изображения. Если параметры не соответствуют идеальным, то сформировать из базы эталонов обучающую выборку, близкую по параметрам к изображениям сцены.
2. Обнаружение движения;
3. Предварительная обработка изображения;
4. Поиск лиц;
 - a. Если параметры изображения идеальные – методом Виолы-Джонса;
 - b. Если параметры изображения не идеальные – методом опорных векторов;
5. Дообучение классификатора OISVM с использованием нового образа.

Выполняется первоначальная настройка системы. Для этого необходимо выполнить внешнюю калибровку камеры и определить параметры изображений сцены. Если параметры изображения сцены не соответствуют идеальным, то необходимо из предварительно сформированной базы эталонов выбрать обучающие образы, подходящие под данные условия. База эталонов представляет собой предварительно промаркированную базу изображений лиц людей, содержащую различные вариации освещения, углов поворотов и наклонов. Из данной базы будут выбираться только необходимые изображения. Так, с помощью данного подхода, мы формируем обучающую выборку для классификатора OISVM, которая будет являться идеальной для наблюдаемой сцены. После предварительного обучения второго классификатора система работает в штатном режиме.

На вход подается кадр видеопотока, который анализируется детектором движения. Поиск движения может выполняться методом выделения оптического потока [1], методом сравнения обновляемого базового кадра с текущим кадром на предмет блобов [26] (массивов двоичных данных - Binary Large Object) переднего плана. При моделировании базового кадра возможно использование вероятностных моделей для возможности работы с динамическим фоном (колыхающиеся листья деревьев и т.д.) [9].

После обнаружения движения выделенная область анализируется на предмет присутствия объекта (лица человека). Поиск осуществляется одним из двух способов: классификатором Виола-Джонса или методом опорных векторов. В случае идеальных условий сцены, работа алгоритма не отличается от классического подхода. В противном случае, детекция выполняется автоматически обученным классификатором OISVM. Также в процессе работы выполняется дообучение данного классификатора.

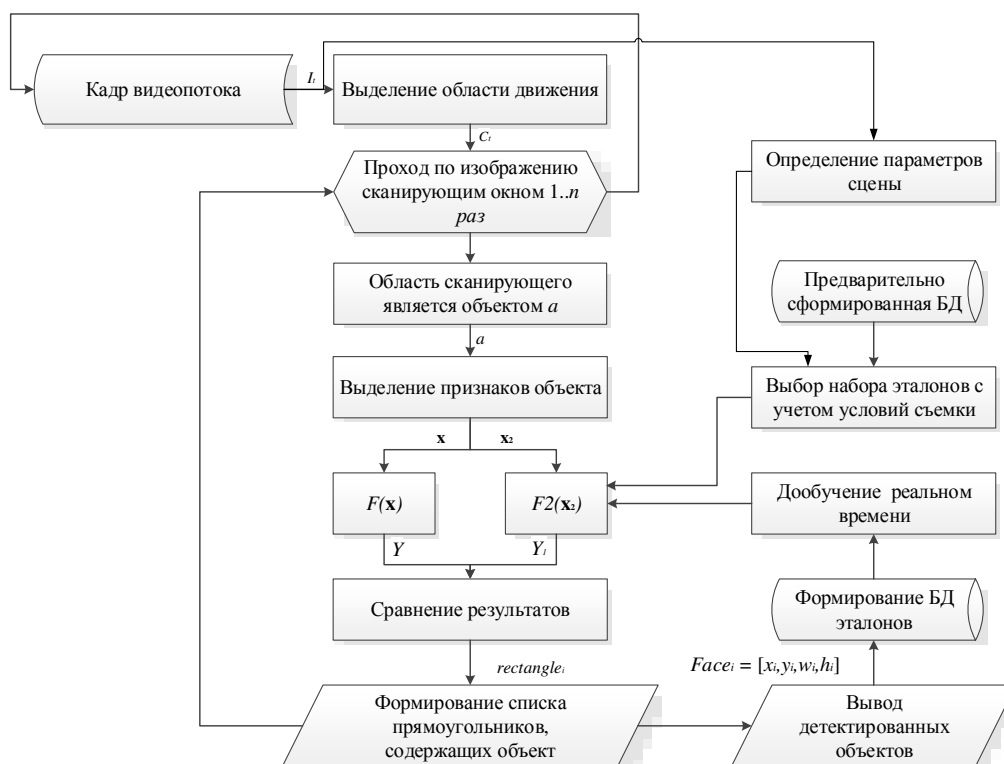


Рис. 14. Синтезированный алгоритм детекции объектов с использованием метода Виолы-Джонса и последовательного обучения OISVM (OISVM+Виола-Джонс)

Составлено (разработано) автором

Для обучения классификатора в реальном времени используются образы, найденные им самим, в случае если они отсутствовали в обучающей выборке. Нами не исключается подход, основанный на работе двух классификаторов одновременно, так как это увеличит надежность системы и обобщающую способность алгоритма. При этом положительные образы, которые определил классификатор Виолы-Джонса также могут быть использованы для дообучения второго классификатора.

Результаты эксперимента

Предложенный нами алгоритм был апробирован экспериментальным путем, с использованием изображений, получаемых с камеры наблюдения.

Нами проведено сравнение следующих методов: метод Виолы-Джонса и предложенный синтезированный алгоритм детекции объектов с использованием метода Виолы-Джонса и последовательного обучения SVM. Для сравнения использовалась база лиц, полученная с камер наблюдения и состоящая из 2 частей: обучающий набор и тестирующий набор. Обучающий набор состоит 9115 изображений (6782 лиц и 2334 «не-лиц»), а тестирующий набор состоит из 2348 изображений (1379 лиц и 969 «не-лиц»).

Эксперимент проходил в 2 этапа: на первом этапе использовался обучающий набор. Классификация выполнялась методом Виолы-Джонса.

На втором этапе использовалась тестирующая выборка. Детекция лиц выполнялась классификатором Виолы-Джонса и классификатором OISVM+Виола-Джонс. Предложенный нами подход позволил повысить эффективность предложенного алгоритма с увеличением количества поступающих образов. Результаты сравнения алгоритмов представлены на графике (рис. 15).

Сравнение классификаторов было выполнено с помощью построения ROC-кривой, выражающей соотношение уровня верных и ложных обнаружений. При этом рассчитывался численный показатель площади под ROC-кривой — Area Under Curve (AUC).

Для эксперимента использовался ПК с характеристиками: ЦПУ Pentium Dual-Core 3,2 ГГц, ОЗУ 3 ГБ.

Предложенный алгоритм был реализован в виде ПО на языке объектно-ориентированного программирования C# (Visual Studio 2010) с использованием библиотеки AForge.NET и библиотеки Matlab Runtime.

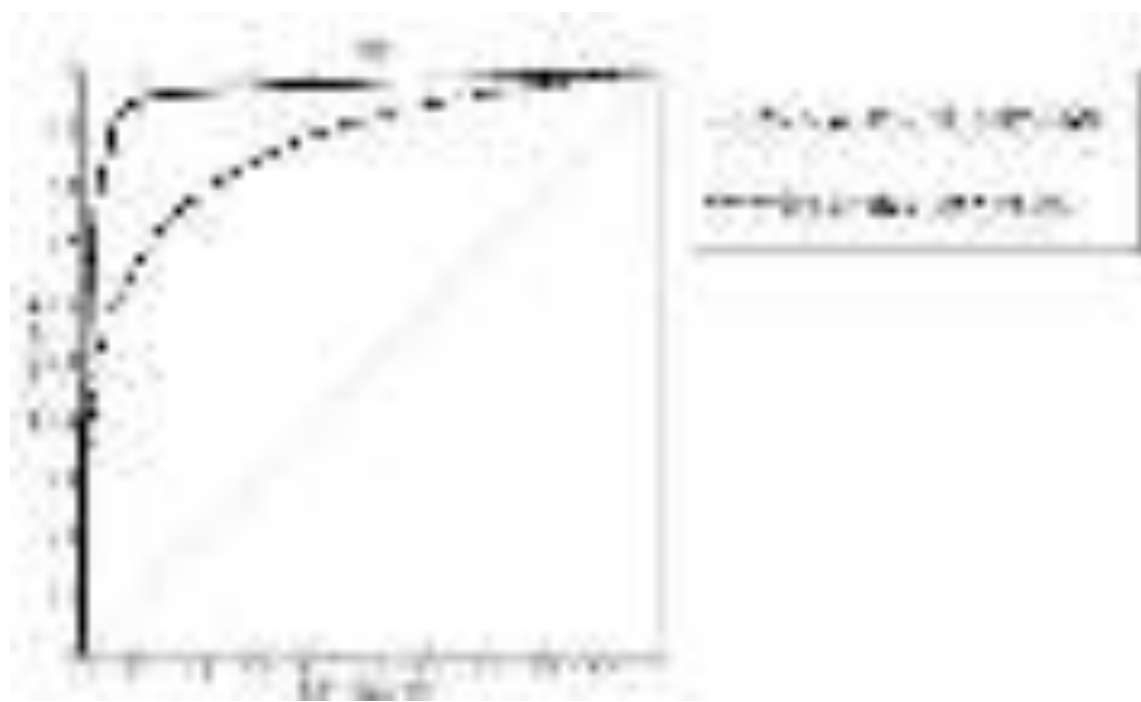


Рис 15. ROC-кривые построенные для классификатора:
Виолы-Джонса на базе FERET; Виолы-Джонса на обучающей выборке с КН;

Составлено (разработано) автором



Рис 16. ROC-кривые построенные для классификатора: Виолы-Джонса на базе FERET; Виолы-Джонса на тестирующей выборке с КН; OISVM+Виола-Джонс на тестирующей выборке с КН

Составлено (разработано) автором

Обсуждение результатов моделирования

В результате моделирования было выяснено, что эффективность классификатора Виолы-Джонса на базе лиц FERET составляет 96% и является максимальным для данного классификатора, так как база FERET использовалась при его обучении. При этом AUC классификатора Виолы-Джонса на базе полученной с камер наблюдения равен 86%, что существенно ниже (Рис. 16). Классификатором Виолы-Джонса было задетектировано 5745 положительных образов «лиц».

Результаты второго эксперимента показали следующее: на тестирующей выборке реальной БД AUC для классификатора Виолы-Джонса остался на значении 86%, с отклонениями в сотые доли. Разница может быть объяснена меньшим количеством изображений в тестирующей выборке. Большой интерес вызывают результаты работы классификатора OISVM+Виола-Джонс, AUC которого составил 92,5%, что намного выше значения полученного первым классификатором. Данный факт позволяет сделать вывод о эффективности применения предложенного алгоритма в реальных системах распознавания образов. Это подтверждает рекомендацию, сделанную нами ранее о повышении эффективности за счет постоянной работы двух классификаторов. Возможность онлайн обучения без вмешательства человека позволяет адаптировать алгоритм под изменяющиеся условия съемки и качество поступающих образов.

Недостатком предложенного алгоритма является увеличение вычислительных затрат при работе классификатора OISVM, так как данный классификатор более требователен к ресурсам. При работе с видеопотоком этот недостаток компенсируется использованием детектора движения. Тогда поиск образа выполняется не во всем изображении, а только в области движения, что снижает вычислительную нагрузку.

В настоящей работе объектом детекции являлось лицо человека, при этом применение предложенного подхода возможно для обнаружения любых других объектов. Для этого потребуется замена классификатора Хаара и замена предварительно сформированной базы

эталонов. Адаптация классификатора OISVM произойдет в автоматическом режиме при поступлении новых образов.

Заключение

В настоящей статье предложен синтезированный алгоритм детекции объектов с использованием параллельной работы 2х классификаторов: метода Виолы-Джонса и последовательного обучения OISVM. Особенностью предложенного алгоритма является онлайн обучение классификатора, основанного на методе опорных векторов.

Научная новизна состоит в применении классификатора OISVM и его онлайн обучении для детекции динамических объектов в видеопотоке, причем предварительное обучение происходит в автоматическом режиме, на основании данных полученных от внешней калибровки камеры.

Экспериментально доказано, что применение предложенного нами подхода увеличивает показатель правильных детекций объектов до 92,5%, что позволяет применять предложенный алгоритм в реальных системах детекции образов.

С использованием тестовой выборки было доказано, что, в отличие от ранее используемых методов, предложенный алгоритм адаптируется к изменчивым условиям съемки и качеству поступающих образов.

Кроме того, выявлены и недостатки, присущие предложенному алгоритму:

1. Требуется достаточно большие объемы новых образов для последовательного обучения OISVM, прежде чем его эффективность превысит уровень детекции классификатора Виолы-Джонса. Данный фактор может стать серьезной проблемой при работе на мобильных устройствах с микропроцессорной архитектурой ARM (микропроцессорная архитектура с сокращённым набором команд – Advanced RISC Machine), где уровень вычислительных ресурсов сильно ограничен.

2. Требуется добавление блока автоматического выбора используемых признаков для возможности замены объекта детекции без прекращения работы алгоритма.

3. Требуется добавление блока автоматического определения условий съемки для полной автоматизации работы алгоритма.

Однако указанные недостатки не снижают практической значимости предложенного подхода, так как архитектура ARM практически не используется при построении систем охранного видеонаблюдения, а смена объекта детекции может быть вызвана только реорганизацией системы в целом или изменением расположения камеры наблюдения.

Разработана и протестирована программа, реализующая синтезированный алгоритм детекции объектов с использованием метода Виолы-Джонса и последовательного обучения OISVM для детекции лиц, на языке объектно-ориентированного программирования C# с использованием библиотек Matlab Runtime.

ЛИТЕРАТУРА

1. Beauchemin S. S., Barron J. L. The computation of optical flow // ACM New York, USA, 1995.
2. Burges C. A Tutorial on Support Vector Machines for Pattern Recognition // Data Mining and Knowledge Discovery. – 1998. – Vol. 2 – P. 121-167.
3. Cauwenberghs G., Poggio T. Incremental and decremental support vectormachine learning // Neural Information Processing Systems. – 2000. – P. 209-218.
4. Chapelle O., Haffiier P., Vapnik V. Support vector for histogram-based image classification // IEEE transactions on Neural Networks. – 1999. – Vol. 10(5) – P. 1055-1065.
5. Cristianini N., Shawe-Taylor J., An Introduction to Support Vector Machines (and Other Kernel-Based Learning Methods). CUP, 2000.
6. Downs T., Gates K. E., Masters A., Exact simplification of support vectors solutions // Journal of Machine Learning Research – 2001. – Vol. 2 – P. 293-297.
7. FERET Face database [Электронный ресурс]. – Режим доступа: <http://itl.nist.gov/iad/humanid/feret/>
8. Freund Y., Schapire R. E. A Short Introduction to Boosting // Shannon Laboratory – USA, 1999, – P. 771-780.
9. Haritaoglu I., Harwood D., Davis L. S. Real-Time Surveillance of People and Their Activities // IEEE Transactions on pattern analysis and machine intelligence. – 2000. – Vol. 22 (8) – P. 809-830.
10. Harris C. G., Stephens M. J. Combined corner and edge detector // Proc. Fourth Alvey Vision Conference. – 1988. – P. 147–151.
11. Keerthi S. S., Chapelle O., DeCoste D., Building support vector machines with reduced classifier complexity.// Journal of Machine Learning Research – 2006. – Vol. 8 – P.1-22.
12. Keerthi S. S., DeCoste D., A modified finite newton method for fast solution of large scalelinear SVMs. // Journal of Machine Learning Research. – 2005. – Vol. 6. – P. 341-361.
13. Kivinen J., Smola A., Williamson R. Online learning with kernels //Advances in Neural Inf. Proc. Systems. – 2001. – Vol. 1 – P. 785-792.
14. Maydt J., Lienhart R. Face Detection with Support Vector Machines and a Very Large Set of Linear Features // IEEE ICME. – 2002, – P. 81-89.
15. MIT CBCL Face Data Set [Электронный ресурс]. – Режим доступа: <http://cbcl.mit.edu/software-datasets/FaceData2.html>.
16. Muller K.-R., Mika S., Ratsch G., Tsuda K., and Scholkopf B. An introduction to kernel-based learning algorithms // IEEE Transactions on Neural Networks. – 2001. – Vol. 12(2), P. 181-201.
17. Orabona F., Castellini C., Caputo B., Jie L., Sandini G., On-line Independent Support Vector Machines // Pattern Recognition – 2010. – Vol. 43(4) – P. 1402-1412.
18. Ralaivola L., d'Alcher Buc F. Incremental Support Vector Machine Learning: A Local Approach // Lecture Notes in Computer Science. – 2001. – Vol. 2130 – P. 322-329.

19. Ruping S. Incremental learning with support vector machines // Techn. Report TR-18. - 2002. – P. 641-642.
20. Syed N. A., Liu H., Sung K. K. Incremental learning with support vector machines // International Joint Conference on Artificial Intelligence (IJCAI). – 1999. – P. 876-892.
21. Viola P., Jones M. Robust Real-Time Face Detection // International Journal of Computer Vision. – 2004. – Vol. 57(2) – P. 137-154.
22. Zhang Z., A flexible new technique for camera calibration, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.22, No.11, pages 1330—1334, 2000
23. Амосов О.С., Иванов Ю.С. Модель, алгоритмы и аппаратно-программный комплекс для управления освещением на основе системы охранного телевидения // «Информатика и системы управления», № 1 (35), 2013. С. 156-166.
24. Амосов О.С., Иванов Ю.С., Модифицированный алгоритм локализации номерных знаков транспортных средств на основе метода Виолы-Джонса // Информатика и системы управления. – 2014. – № 1(39). – С.127-140.
25. Библиотека компьютерного зрения OpenCV [Электронный ресурс]. – Режим доступа:
http://docs.opencv.org/trunk/doc/py_tutorials/py_objdetect/py_face_detection/py_face_detection.html.
26. Гонсалес Р., Вудс Р., Цифровая обработка изображений. – М.: Техносфера, 2005.

Рецензент: Иванов Сергей Николаевич, доцент, д.т.н., ФГБОУ ВПО «Комсомольский-на-Амуре государственный технический университет».

Oleg Amosov

Komsomolsk-on-Amur State Technical University
Russia, Komsomolsk-on-Amur
E-Mail: osa18@yandex.ru

Yuriy Ivanov

Positron-DV, ltd.
Russia, Komsomolsk-on-Amur
E-Mail: ivanov_ys@garmvis.ru

Face detection modified algorithm in the video stream and its programming implementation

Abstract: The main problem in the video analytics is the detection of moving objects in a video stream, such as people, vehicles, animals. One of the most important problems, trying to be solved by the computer vision systems is the face detection in the image. However, in attempts to build an automatic system of face detection one faces with factors that complicate the process of detection.

The set of indicators for a preliminary assessment of the classifier using its training sample along with a modified algorithm for the face detection of moving objects in video stream, the main feature of which is an adaptation of the classifier to the changing characteristics of the scene and new images of objects have been suggested into this article. Also the software has been developed for the face detection of human faces.

Scientific novelty of the article consists in applying a classifier OISVM and its online training for the detection of dynamic objects in the video stream, moreover its pre-learning occurs automatically, based on the data, received from the external automatic camera calibration.

Experimentally proved that the application of the our suggested approach increases the rate of correct detections of objects up to 92.5% on the noisy and distorted images, which allows to apply the algorithm in real systems for detecting images.

Using the test sample it has been proved that, in contrast to the previously used methods, the proposed algorithm adapts to changing shooting conditions and quality of incoming images.

Keywords: face detection; classification; Viola-Jones method; support vector machine; online learning; automatic camera calibration.

Identification number of article 23TVN314

REFERENCES

1. Beauchemin S. S., Barron J. L. The computation of optical flow // ACM New York, USA, 1995.
2. Burges C. A Tutorial on Support Vector Machines for Pattern Recognition // Data Mining and Knowledge Discovery. – 1998. – Vol. 2 – P. 121-167.
3. Cauwenberghs G., Poggio T. Incremental and decremental support vectormachine learning // Neural Information Processing Systems. – 2000. – P. 209-218.
4. Chapelle O., Haffiier P., Vapnik V. Support vector for histogram-based image classification // IEEE transactions on Neural Networks. – 1999. – Vol. 10(5) – P. 1055-1065.
5. Cristianini N., Shawe-Taylor J., An Introduction to Support Vector Machines (and Other Kernel-Based Learning Methods). CUP, 2000.
6. Downs T., Gates K. E., Masters A., Exact simplification of support vectors solutions // Journal of Machine Learning Research – 2001. – Vol. 2 – P. 293-297.
7. FERET Face database [Jelektronnyj resurs]. – Rezhim dostupa: <http://itl.nist.gov/iad/humanid/feret/>
8. Freund Y., Schapire R. E. A Short Introduction to Boosting // Shannon Laboratory – USA, 1999, – P. 771-780.
9. Haritaoglu I., Harwood D., Davis L. S. Real-Time Surveillance of People and Their Activities // IEEE Transactions on pattern analysis and machine intelligence. – 2000. – Vol. 22 (8) – P. 809-830.
10. Harris C. G., Stephens M. J. Combined corner and edge detector // Proc. Fourth Alvey Vision Conference. – 1988. – P. 147–151.
11. Keerthi S. S., Chapelle O., DeCoste D., Building support vector machines with reduced classifier complexity.// Journal of Machine Learning Research – 2006. – Vol. 8 – P.1-22.
12. Keerthi S. S., DeCoste D., A modified finite newton method for fast solution of large scalelinear SVMs. // Journal of Machine Learning Research. – 2005. – Vol. 6. – P. 341-361.
13. Kivinen J., Smola A., Williamson R. Online learning with kernels //Advances in Neural Inf. Proc. Systems. – 2001. – Vol. 1 – P. 785-792.
14. Maydt J., Lienhart R. Face Detection with Support Vector Machines and a Very Large Set of Linear Features // IEEE ICME. – 2002, – P. 81-89.
15. MIT CBCL Face Data Set [Jelektronnyj resurs]. – Rezhim dostupa: <http://cbcl.mit.edu/software-datasets/FaceData2.html>.
16. Muller K.-R., Mika S., Ratsch G., Tsuda K., and Scholkopf B. An introduction to kernel-based learning algorithms // IEEE Transactions on Neural Networks. – 2001. – Vol. 12(2), P. 181-201.
17. Orabona F., Castellini C., Caputo B., Jie L., Sandini G., On-line Independent Support Vector Machines // Pattern Recognition – 2010. – Vol. 43(4) – P. 1402-1412.
18. Ralaivola L., d’Alcher Buc F. Incremental Support Vector Machine Learning: A Local Approach // Lecture Notes in Computer Science. – 2001. – Vol. 2130 – P. 322-329.

19. Ruping S. Incremental learning with support vector machines // Techn. Report TR-18. - 2002. – P. 641-642.
20. Syed N. A., Liu H., Sung K. K. Incremental learning with support vector machines // International Joint Conference on Artificial Intelligence (IJCAI). – 1999. – P. 876-892.
21. Viola P., Jones M. Robust Real-Time Face Detection // International Journal of Computer Vision. – 2004. – Vol. 57(2) – P. 137-154.
22. Zhang Z., A flexible new technique for camera calibration, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.22, No.11, pages 1330—1334, 2000
23. Amosov O.S., Ivanov Ju.S. Model', algoritmy i apparatno-programmnyj kompleks dlja upravlenija osveshheniem na osnove sistemy ohrannogo televidenija // «Informatika i sistemy upravlenija», № 1 (35), 2013. S. 156-166.
24. Amosov O.S., Ivanov Ju.S., Modificirovannyj algoritm lokalizacii nomernyh znakov transportnyh sredstv na osnove metoda Violy-Dzhonsa // Informatika i sistemy upravlenija. – 2014. – № 1(39). – S.127-140.
25. Библиотека комп'ютерного зренія OpenCV [Електронний ресурс]. – Режим доступа: http://docs.opencv.org/trunk/doc/py_tutorial/py_objdetect/py_face_detection/py_face_detection.html.
26. Gonsales R., Vuds R., Cifrovaja obrabotka izobrazhenij. – М.: Tehnosfera, 2005.