

УДК 004.89:004.912

В. В. Бондарчук, Е. А. Шевченко

Государственное учреждение «Институт проблем искусственного интеллекта», г. Донецк
83048, г. Донецк, ул. Артема, дом 118 б

ИССЛЕДОВАНИЕ АЛГОРИТМА ДЕТЕКЦИИ ОБЪЕКТОВ ИНТЕРЕСА В ВИДЕОПОТОКЕ

V. V. Bondarchuk, J. A. Shevchenko

Public institution «Institute of Problems of Artificial intelligence», Donetsk
83048, Donetsk, Artema st., 118b

STUDY OF ALGORITHM OF DETECTION OF OBJECTS OF INTEREST IN A VIDEO STREAM

В. В. Бондарчук, Є. О. Шевченко

Державна установа «Інститут проблем штучного інтелекту», м. Донецьк
83048, м. Донецьк, вул. Артема, буд. 118 б

ДОСЛІДЖЕННЯ АЛГОРИТМУ ДЕТЕКЦІЇ ОБ'ЄКТІВ ІНТЕРЕСУ У ВІДЕОПОТОЦІ

Разработан алгоритм детекции объектов в видеопотоке на основе видеосигналов, видеоданных полученных с помощью стационарных средств видеонаблюдения (видеокамер) с помощью использования программного пакета SPECPLUS.

Ключевые слова: детекция объектов; классификация; интегральные характеристики оптического потока, метод Виолы-Джонса; графы, метод опорных векторов; онлайн-обучение; автоматическая калибровка камеры.

The developed algorithm for detection of objects in a video stream is based on video signals, the video data obtained by stationary means (cameras) through the use of a software package SPECPLUS.

Key words: detection of objects in a video stream; classification; method of Viola-Jones; graphs; support vector machine; online learning; automatic camera calibration.

Розроблено алгоритм детекції об'єктів у відеопотоці на основі відеосигналів, відеоданих отриманих за допомогою стаціонарних засобів відеоспостереження (відеокамер) за допомогою використання програмного пакета SPECPLUS .

Ключові слова: детекція об'єктів; класифікація; інтегральні характеристики оптичного потоку, метод Віолі-Джонса ; графи, метод опорних векторів ; Онлайн-навчання; автоматичне калібрування камери.

Введение

Переход к цифровым методам передачи и коммутации помимо преимуществ, обусловленных новыми технологиями связи, вызвал множество проблем, специфика которых связана с принципами передачи сигналов в цифровом виде. В настоящее время наблюдается значительный интерес к интеллектуализации систем видеонаблюдения, автоматизации подсчета количества посетителей супермаркета, длины очередей, количества автомобилей, которые проехали в области наблюдения, детектирования направлений и интенсивности человеческих потоков, распознавания и отслеживания целевых объектов и т.п. Относительно этих направлений интеллектуализации видеонаблюдения имеется значительный спрос во всех инфраструктурах, промышленности. Задача представляет и оборонный интерес.

В данной статье проводится исследование алгоритмов детекции объектов в видеопотоке и его реализации на основе видеосигналов, видеоданных полученных с помощью стационарных (неподвижных) средств видеонаблюдения (видеокамер) с помощью использования программного пакета SPECPLUS. Новые направления исследований включают разработки новых методов решения основных задач видеонаблюдения: исключения фона, сегментация и описание сцен, выявление объектов, перспективного описания и представления объектов интереса и сцены видеонаблюдения, обучение и дообучение алгоритмов обнаружения и распознавания объектов в процессе трекинга, а также мультисервисную систему повышенной безопасности и надежности.

Постановка и решение задачи детекции объекта интереса 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 [1].

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 Ct , путем вырезания ее из матрицы I_t . Размер полученной матрицы равен $wc \times hc$, где $wc < weight$ и $hc < height$. Общий подход решения проблем систем видеонаблюдения представлен на рис. 1.

По области движения требуется определить, есть ли список прямоугольников, содержащих объект человека. Поиск объекта будет выполняться «сканирующим окном», проходящим последовательно по изображению с шагом 1 пиксель (слева направо, сверху вниз) с изменением масштаба при каждом проходе.

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

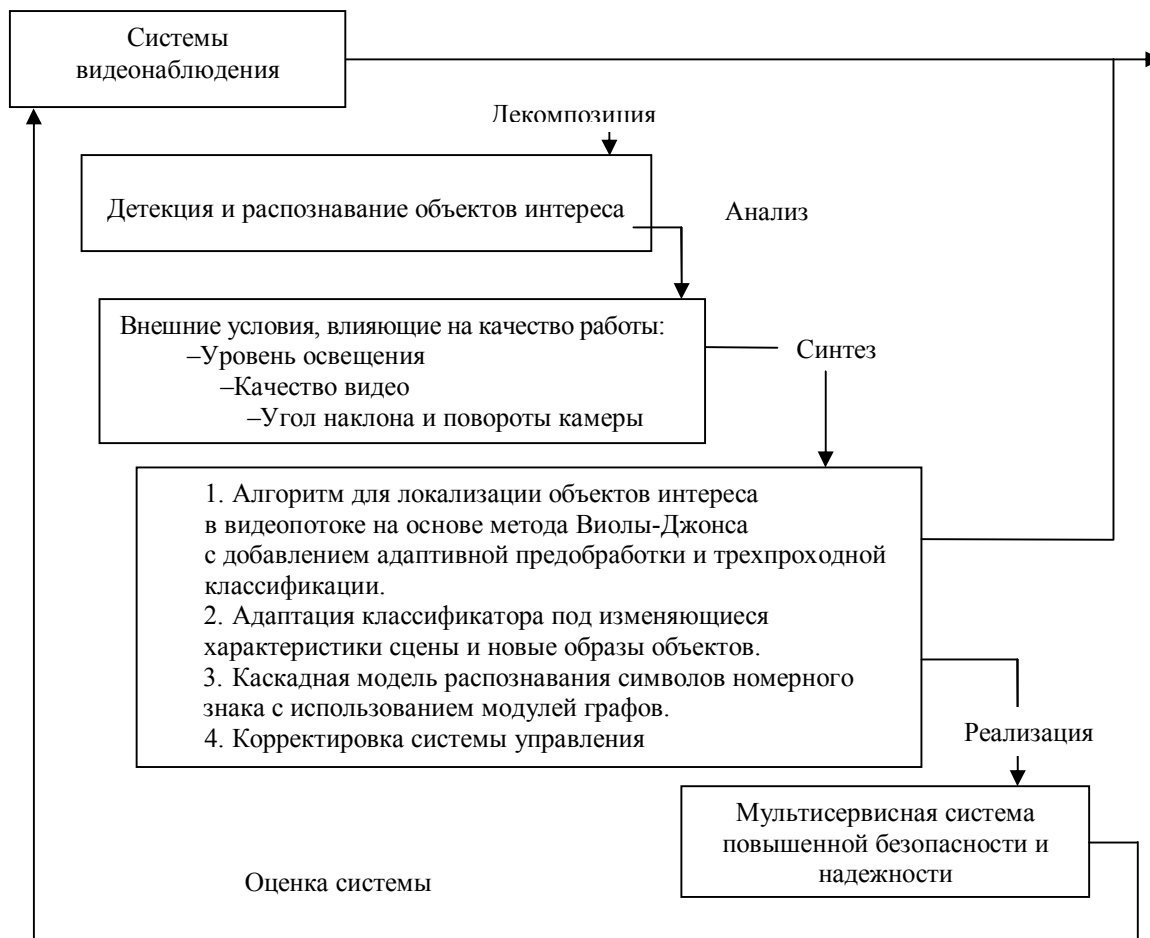


Рисунок 1 – Общий подход к решению проблем СВН

Так как существует всего два класса {«не-объект»; «объект»} = {-1; +1}, то задача детекции объекта в кадре видеопотока сводится к задаче двухклассовой классификации. При этом вектор признаков допустимо отождествлять с самими объектами. Классификатор F должен быть способен классифицировать произвольный объект a . Оптимальным считается классификатор, который дает наименьшую вероятность ошибки при всех допустимых значениях x . Для обучения классификатора F используется обучающая выборка.

Преимущества разработанного алгоритма детекции объектов интереса

Основной проблемой при детекции объектов является присутствие на анализируемом изображении посторонних объектов и большого количества шума.

Для решения этой проблемы был предложен алгоритм двухпроходной классификации, особенностью которого являлась параллельная работа 2-х классификаторов. Алгоритм обладал устойчивостью к изменчивым условиям съемки и качеству кадра, за счет чего показал высокие результаты при детекции номерных знаков транспортных средств.

Классический алгоритм детекции объекта 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, под конкретную камеру наблюдения, что является довольно трудоемким процессом. На сформированной выборке необходимо обучить классификатор, при этом процесс обучения может занимать до нескольких недель.

Применение такого подхода 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.

Предлагается совместное использование трех классификаторов: первый из них построен на основе метода Виолы-Джонса, предварительно обучен и хорошо зарекомендовал себя в большинстве программных продуктов; второй классификатор на основе метода графов – изучен и аргументировано обоснован, третий классификатор должен поддерживать онлайн-обучение [2]. Для его обучения предварительно формируется база эталонов, содержащая объекты, снятые 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. Алгоритм реализации системы с использованием классификаторов представлен на рис. 2.

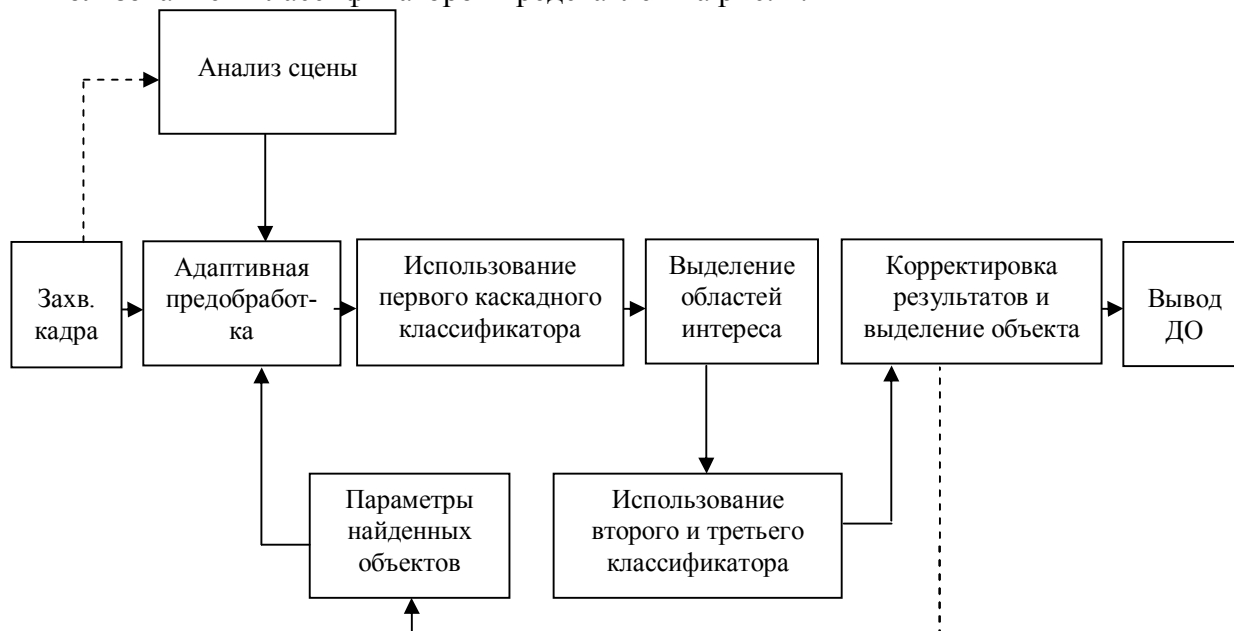


Рисунок 2 – Алгоритм реализации системы с использованием классификаторов

Обучение третьего классификатора будет происходить в несколько этапов. Сначала выполняется внешняя калибровка камеры, которая позволяет определить такие параметры, как: углы наклона камеры, высоту установки, условия освещения и т.д. Если параметры изображений соответствуют или близки идеальным, то есть углы наклона незначительны, освещение яркое, то необходимость в использовании второго классификатора отсутствует и 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. Полученная выборка используется для обучения второго классификатора. В процессе работы алгоритма, все образы, полученные вторым и третьим классификаторами также используются для его дообучения.

Алгоритм детекции объектов с использованием метода Виолы-Джонса и обучения представлен на рис. 3.

Формирование обучающей выборки с учетом условий съемки

Основными признаками, влияющими на качество работы алгоритма, являются: зашумленность изображения; углы наклона и повороты объекта интереса; уровень освещения; высота установки камеры и угол наклона относительно плоскости пола; масштаб объекта интереса относительно размера изображения;

Детекция объекта интереса является первым шагом при решении более сложной задачи – идентификации (распознавания). При этом пороговые значения должны быть установлены таким образом, чтобы детектируемый объект мог быть в дальнейшем распознан.

Одним из основополагающих методов детекции объекта интереса является метод Виолы-Джонса. Данный метод основан на следующих принципах:

- используются изображения в интегральном представлении;
- используются признаки Хаара;
- используется бустинг для выбора наиболее подходящих признаков для искомого объекта на данной части изображения;
- все признаки поступают на вход классификатора, который даёт результат «верно» или «ложь»;
- используются каскады признаков для быстрого отбрасывания окон, где не найден объект.

Для успешной работы классификатора необходимо, чтобы характеристики тестирующих изображений находились в определенных пределах. При этом, чем ближе график тестирующий выборки к графику 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. Решением данной проблемы может стать использование второго классификатора с другой областью допустимых значений параметров, совпадающей с параметрами изображений, поступающих с камеры наблюдения. Проблемой становится отсутствие информации о параметрах поступающих изображений.

Выход заключается в применении последовательного дообучения второго классификатора путем добавления к нему новых образов детектируемых объектов. При таком подходе искомый вектор будет постепенно приближаться к вектору значений параметров поступающих изображений. 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. Основной проблемой является необходимость предварительного обучения классификатора 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.

Таким образом, ставится задача определения параметров изображения или внешней калибровки.

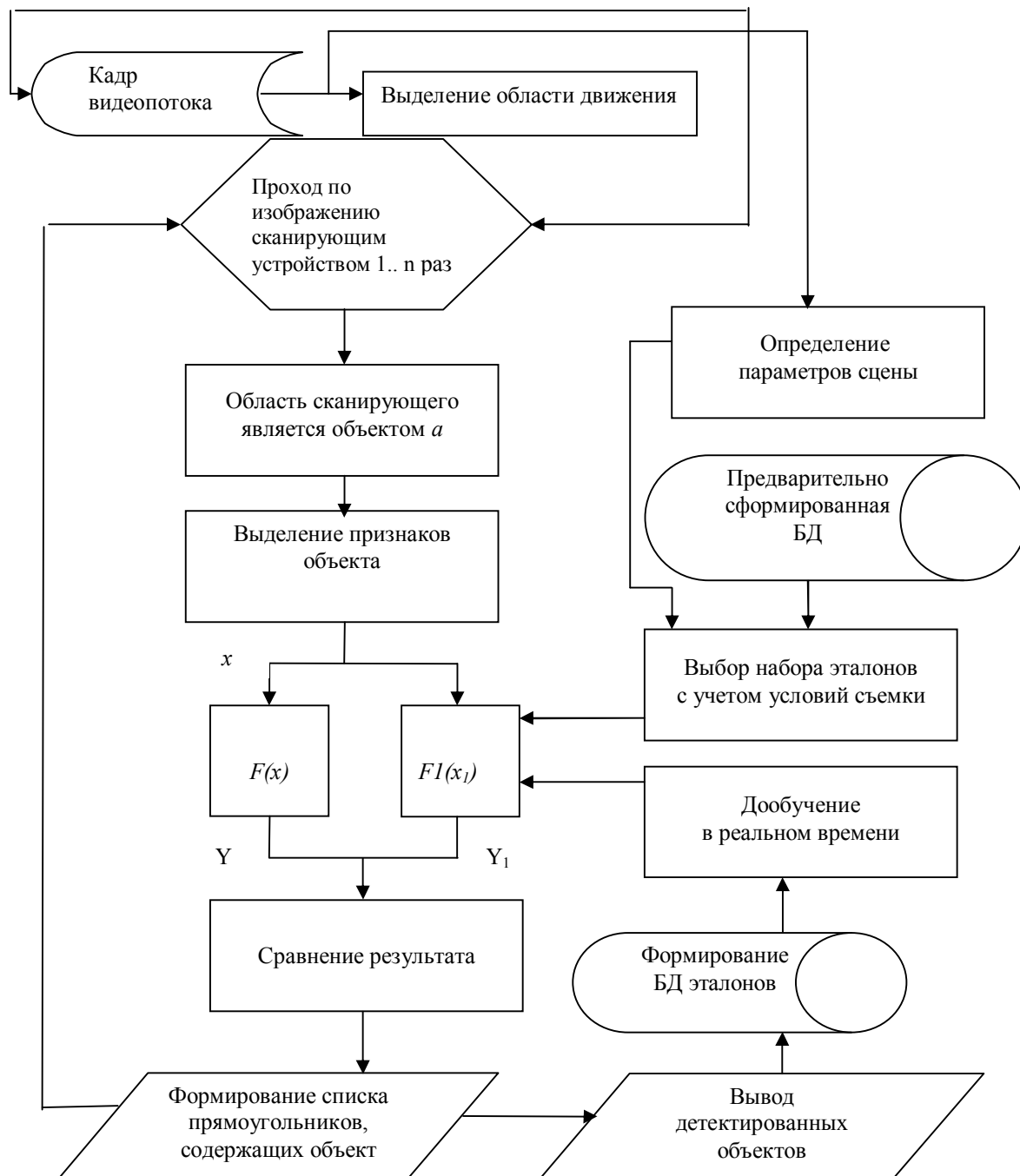


Рисунок 3 – Модифицированный алгоритм детекции объектов с использованием метода Виолы-Джонса и обучения

Способы определения условий съемки

Способ определения условий освещенности был изложен в [2]. 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. На каждом изображении выполняется поиск калибровочного объекта с помощью детектора Харриса. Предлагается следующая последовательность действий при начальной инициализации камеры наблюдения:

1. Определение условий освещенности и зашумленности.
2. Печать калибровочного объекта.
3. Оператор пронесит калибровочный объект по наиболее вероятной траектории движения объекта.
4. На основе полученных кадров выполняется внешняя калибровка камеры, и рассчитываются вероятные углы наклона и повороты объектов.

Определив все параметры поступающих изображений, можно сформировать предварительную обучающую выборку для классификатора.

Метод графов

Предложен метод поиска в графе подграфов с заданными свойствами на основе использования процесса формирования дополнительного графа-пирамиды[4], что позволило на его основе разработать алгоритмы для анализа, сравнения графов, поиска множества максимальных клик, вычислительная сложность которых линейно зависит от количества решений и которые удобны для параллельной реализации. Учёт особенностей графов, описывающих изображения, позволяет проводить их обработку и сравнение с полиномиальной сложностью.

Разработанный алгоритм детекции объектов с использованием метода Виолы-Джонса и последовательного обучения состоит из следующих этапов:

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 объекта:
 - а) Если параметры изображения идеальные – методом Виолы-Джонса;
 - б) Если параметры изображения не идеальные – методом графов.
5. Дообучение классификатора с использованием онлайн-обучения 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 объектов 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, которая будет являться идеальной для наблюдаемой сцены.

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 трех 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. Возможность онлайн-обучения без вмешательства человека позволяет адаптировать алгоритм под изменяющиеся условия съемки и качество поступающих образов.

В работе объектом детекции являлся номерной знак автомобиля, при этом применение предложенного подхода возможно для обнаружения любых других объектов. Адаптация классификатора произойдет в автоматическом режиме при поступлении новых образов. Результаты работы алгоритма приведены на рис. 4.

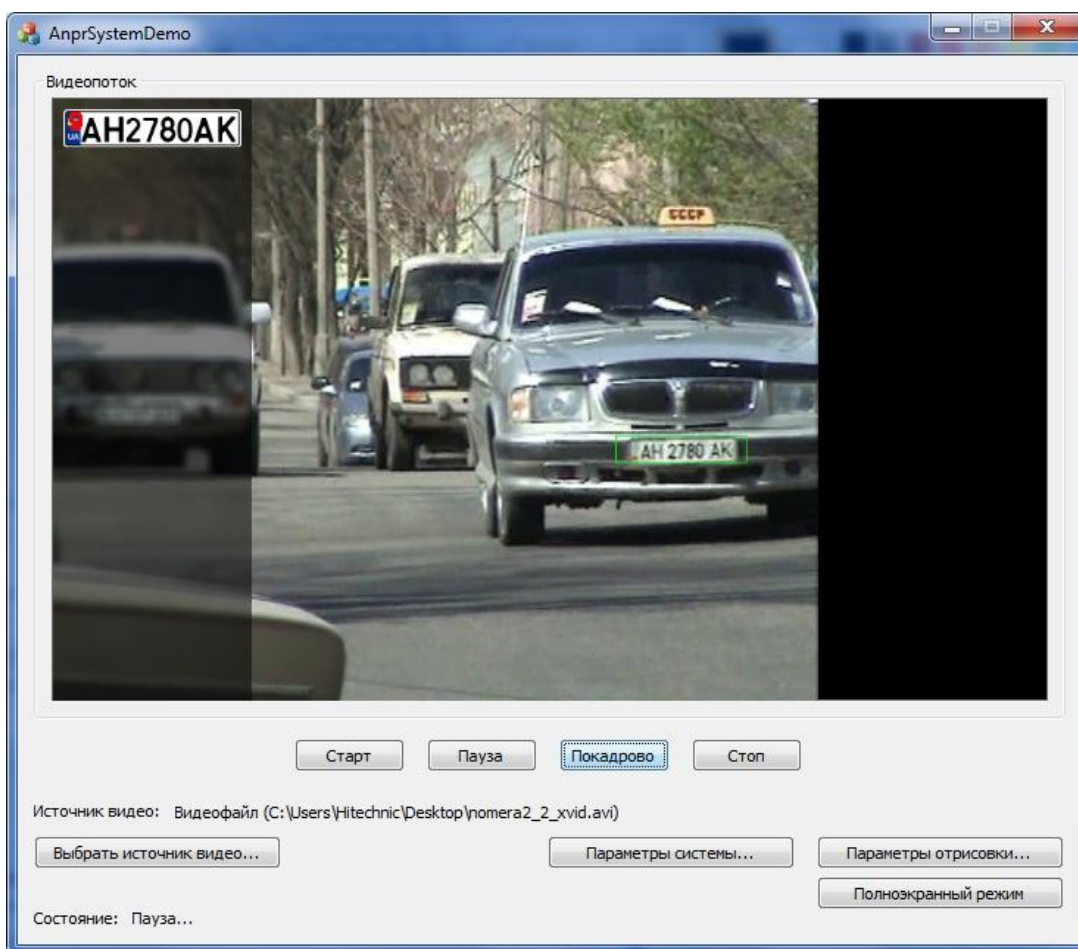


Рисунок 4 – Распознавание номерного знака автомобиля в реальном времени

Выводы

Предложен алгоритм детекции объектов с использованием параллельной работы 3-х классификаторов: метода Виолы-Джонса, метода графов и его онлайн-обучения. Научная новизна состоит в адаптации классификатора и обучении для детекции динамических объектов в видеопотоке, причем предварительное обучение происходит в автоматическом режиме, на основании данных полученных от внешней калибровки камеры.

Разработаны алгоритмы детекции объектов в видеопотоке на основе видеосигналов, видеоданных полученных с помощью стационарных средств видеонаблюдения (видеокамер) с помощью использования программного пакета SPECPLUS. Новые направления исследований включают разработки новых методов решения основных задач видеонаблюдения: исключения фона, сегментация и описание сцен, выявление объектов, перспективного описания и представления объектов интереса и сцены видеонаблюдения, обучение и дообучение алгоритмов обнаружения и распознавания объектов в процессе трекинга, а также мультисервисную систему повышенной безопасности и надежности. Экспериментально доказано, что применение предложенного подхода увеличивает показатель правильных детекций объектов до 90%, что позволяет применять предложенный алгоритм в реальных системах детекции образов.

Список литературы

1. Амосов О. С. Модель, алгоритмы и аппаратно-программный комплекс для управления освещением на основе системы охранного телевидения / О. С. Амосов, Ю. С. Иванов // Информатика и системы управления. – 2013. – № 1 (35). – С. 156-166.
2. Иванов Ю. С. Модифицированный алгоритм детекции лиц в видеопотоке и его программная реализация / О. С. Амосов, Ю. С. Иванов // Наукоедение. – 2014 – № 3 (22).
3. Viola P. Robust Real-Time Face Detection / P. Viola, M. Jones // International Journal of Computer Vision. – 2004. – Vol. 57(2). – P. 137-154.
4. Agarkov A. V. Search of the Set of Maximal Cliques Based on the Method for Constructing Complementary Graph / A. V. Agarkov // Artificial Intelligence. – 2011. – № 3. – P. 190-199.

References

1. Amosov O. S., Ivanov Yu. s. Model, algorithms and hardware-software complex for lighting control on the basis of closed circuit television // "Informatics and control systems", № 1 (35), 2013. S. 156-166.
2. Ivanov, Y. S. the Modified algorithm of face detection in a video stream and its implementation / O. S. Amosov, Yu. s. Ivanov // Science studies. – 2014 – №3 (22).
3. Viola P., Jones M. Robust Real-Time Face Detection // International Journal of Computer Vision. – 2004. – Vol. 57(2) – P. 137-154.
4. Agarkov A.V. Search of the Set of Maximal Cliques Based on the Method for Constructing Complementary Graph / A.V. Agarkov // Artificial Intelligence. – 2011. – № 3. – P. 190-199.

RESUME

V. V. Bondarchuk, J. A. Shevchenko

The Algorithm for Detecting Objects in a Video Stream and its Software Implementation

Background: the transition to digital methods of transmission and switching in addition to the advantages due to new communication technologies caused a lot of problems related to the principles of digital signal transmission. Currently there is considerable interest in the intellectualization of the systems of video surveillance, automation of counting the number of visitors to the supermarket, the length of queues, number of cars crossing the fields of surveillance, detection of directions and intensity of human flows, recognition and tracking of targets, and the like. Relative to these areas of intellectualization of the video there is a significant demand in the infrastructure industry. The problem is also of defense interest.

Materials and methods: the paper uses three classifiers: the first one is based on the method of Viola-Jones; the second classifier relies on the method of graphs; the third classifier should support online learning. This learning implies development of a database of standards containing objects, taken under various conditions and angles. Each image must be marked, i.e. it must contain information about the shooting conditions. It is important in order to extract from the database only necessary images. Learning of the third classifier has several stages. The first one is the external camera calibration, which allows us to determine such parameters: camera angles, mounting height, lighting conditions, etc. If the images match or they are almost perfect, then the detection is done by a classifier based on the method of Viola-Jones. If shooting conditions differ from the ideal ones, then on the basis of parameters obtained from a pre-developed database of standards we select images, which are the closest to these conditions, thereby we form the learning sample taking into account the shooting conditions.

Results: it is experimentally proved that the application of the proposed approach increases the correct detection of objects up to 90%, which allows applying the proposed algorithm in real systems of image detection.

Conclusion: the developed algorithms of detection of objects in a video stream based on video data obtained by stationary means (cameras) through the use of a software package SPECPLUS. New areas of research include the development of new methods of solving the basic problems of video: the exclusion of the background, segmentation and description of scenes, detection of objects, perspective descriptions and representations of objects of interest and scene of surveillance, learning algorithms of detection and recognition of objects in the tracking process, as well as a multi-service system of high security and reliability.

Статья поступила в редакцию 21.12.2015.