

УДК 004.93'1  
DOI: 10.15827/0236-235X.122.368-373

Дата подачи статьи: 07.09.17  
2018. Т. 31. № 2. С. 368–373

## АЛГОРИТМ РАСПОЗНАВАНИЯ СИТУАЦИЙ В РАСПРЕДЕЛЕННОЙ СИСТЕМЕ ВИДЕОНАБЛЮДЕНИЯ

А.Ю. Кручинин<sup>1</sup>, к.т.н., доцент, *kruchinin-al@mail.ru*

Д.В. Колмыков<sup>1</sup>, магистр, *malin.chyn@gmail.com*

Р.Р. Галимов<sup>1</sup>, к.т.н., доцент, *rin-galimov@yandex.ru*

<sup>1</sup> *Оренбургский государственный университет, просп. Победы, 13, г. Оренбург, 460018, Россия*

Системы видеонаблюдения являются важнейшим средством для предотвращения нештатных ситуаций, таких как преступления, аварийные ситуации. Большое количество камер и значительная площадь зоны контроля обуславливают необходимость внедрения видеоаналитики для распознавания опасных ситуаций. При этом нужно учитывать данные с множества видеокамер как для детектирования траектории движения распознаваемого объекта, так и для повышения достоверности распознавания.

В статье предлагается алгоритм распознавания нештатных ситуаций для распределенной системы видеонаблюдения, основанной на стохастических грамматиках. Распознавание ситуации происходит на трех уровнях: нижнем – распознаются образы, среднем – события и верхнем – ситуации. Для снижения времени отклика системы предлагается использовать многоагентную архитектуру, позволяющую распределять нагрузку между интеллектуальными камерами. Уменьшение сетевого трафика достигается тем, что обмен данными происходит только между близлежащими узлами.

Использование большого количества видеокамер предполагает наличие зон, контролируемых несколькими узлами. Совмещение результатов детектирования нескольких камер позволяет повысить оценку достоверности, но для этого требуется знать взаимное расположение камер и углов их поворотов. В статье предложены методы для автоматической калибровки камер распределенной системы видеонаблюдения, способы совмещения образов на разных камерах, в частности, на основе векторов скорости движения объектов.

С учетом определенных особенностей распределенной системы видеонаблюдения разработан алгоритм распознавания нештатных ситуаций для интеллектуальной камеры видеонаблюдения. Каждая камера генерирует вероятные ситуации на основе ранее распознанных событий. При превышении порогового значения вероятностной оценки результата детектирования осуществляется его уточнение в процессе взаимодействия с соседними узлами.

**Ключевые слова:** *распределенная система видеонаблюдения, распознавание нештатных ситуаций, многоагентная система.*

Современные системы видеонаблюдения часто охватывают большие территории с большим количеством видеокамер, за которыми операторам невозможно уследить. Поэтому для детектирования, слежения и идентификации объектов применяются алгоритмы компьютерного зрения [1, 2]. Однако в современных системах видеонаблюдения недостаточно только распознавать графические образы – требуется детектировать опасные ситуации, происходящие с этими образами. В данной работе принимается, что объект – это графический образ, событие – мгновенное совершение действия объектом или без объекта, ситуация – последовательность событий, то есть распознавание ситуаций – верхний уровень алгоритмов распознавания, которые оперируют с результатами распознавания графических образов. В простейшем случае ситуации можно разделить на два типа: нормальные и опасные. При возникновении опасной ситуации необходимо оповещать о ней заинтересованных лиц.

Распределенная система видеонаблюдения характеризуется большим количеством камер и значительной территорией контроля, в которой объекты могут перемещаться из зоны наблюдения одной видеокамеры в другую. В связи с этим существует необходимость в разработке новых алго-

ритмов распознавания опасных ситуаций, позволяющих учитывать данные с множества камер. В настоящей статье предлагается алгоритм распознавания ситуаций в условиях объединения результатов с нескольких камер.

При распознавании графического образа можно идентифицировать не только сам объект, но и совершаемое им действие. В результате распознавания последовательности кадров будет распознана последовательность образов:  $\Omega = (\omega_1, \omega_2, \omega_3, \dots, \omega_N)$ , где  $N$  – размер последовательности образов, каждый из которых соответствует одному из эталонных классов  $\Omega_1, \Omega_2, \Omega_3, \dots, \Omega_M$ , где  $M$  – общее количество классов образов. Из этой последовательности образов формируется последовательность событий:  $S = (s_1, s_2, s_3, \dots, s_K)$ , где  $K \leq N$  – размер последовательности событий.

Известны структурные методы распознавания, применяемые в случае, когда объект распознавания сложен и для него можно составить грамматику для его описания. Если каждое событие обозначить символом грамматики и использовать вероятность того или иного события, можно построить и применить стохастическую грамматику. Сам процесс распознавания ситуаций аналогичен процессу поиска подстрок (шаблонов) в строке (цепочке событий) в структурных методах распознавания [3].

Для распознавания ситуации в распределенной системе видеонаблюдения часто необходима информация о событиях, которые фиксируются в разных зонах, подконтрольных разным камерам. Например, человек что-то сделал в одной зоне, затем перешел в другую и сделал что-то там. События должны складываться, а общая последовательность событий анализироваться. С другой стороны, событие может происходить в области зрения одновременно нескольких камер, что может повысить достоверность распознавания за счет объединения результатов детектирования с разных точек наблюдения. В результате для распределенной системы видеонаблюдения можно выделить две основные задачи:

- повышение достоверности распознавания события за счет совмещения результатов с нескольких камер;
- распознавание ситуаций на основе событий, распределенных во времени и по местоположению.

Решение этих задач позволит повысить достоверность распознавания ситуаций. При этом нужно понимать, что события, принадлежащие к определенной ситуации, происходят в одной или близлежащих зонах за небольшой период времени. В результате получается последовательность событий во времени с небольшими интервалами.

Распределенные системы видеонаблюдения могут быть реализованы в виде двух основных архитектур: централизованной (с единым центром) и многоагентной (без единого центра).

В централизованной архитектуре (рис. 1а) все камеры передают видеоданные в центр обработки для распознавания. Большое количество камер может привести к увеличению времени отклика системы из-за большой нагрузки на центр обработки данных. Скорость передачи видеоизображения может зависеть от расположения камеры. Недостатком подобной архитектуры является ее низкая надежность, так как выход из строя видеосервера приведет к отказу всей системы [4].

В многоагентной архитектуре (рис. 1б) каждая камера самостоятельно выполняет функции распознавания ситуации. Это возможно, поскольку мно-

жество современных IP-камер обладают функциональностью обычного компьютера под управлением Linux [5, 6]. К подобным камерам относятся камеры компании Axis Communications. В данной архитектуре они осуществляют взаимодействие с другими близлежащими камерами:

- передают соседям предварительно распознанные образы с целью совместного распознавания; если несколько камер распознают событие с низкой достоверностью, совмещение результатов их работы повышает достоверность распознавания; например, если среднее значение ошибки распознавания некоторого образа равно 0.4, то при совместной работе двух камер можно допустить, что применяется правило перемножения вероятностей для независимых образов, и вероятность ошибки составит 0.16;
- передают распознанные цепочки событий соседним камерам для их объединения с цепочками с разных камер и распознавания ситуации в целом.

Преимуществом данной архитектуры является высокая степень масштабируемости, надежности и оперативности. Недостаток подобного решения в необходимости использования высокопроизводительных камер с встроенной видеоаналитикой.

При проектировании алгоритма распознавания ситуаций необходимо исходить из того, что в распределенной системе видеонаблюдения существуют несколько вариантов анализа ситуации:

- распознавание в одной камере;
- распознавание в камерах с пересекающимися областями зрения;
- распознавание в камерах с известными координатами, но без пересечения области зрения, на основе объединения цепочек событий.

Очевидно, что для реализации этих вариантов все камеры необходимо откалибровать с точки зрения координат камер и их углов поворота в пространстве. Автоматическая предварительная калибровка может осуществляться с помощью метода, описанного в работе [7]. Идентифицироваться объекты могут по уникальным признакам (особые точки, гистограммы яркости, скелет или контур объекта и др.), однако из-за низкого каче-

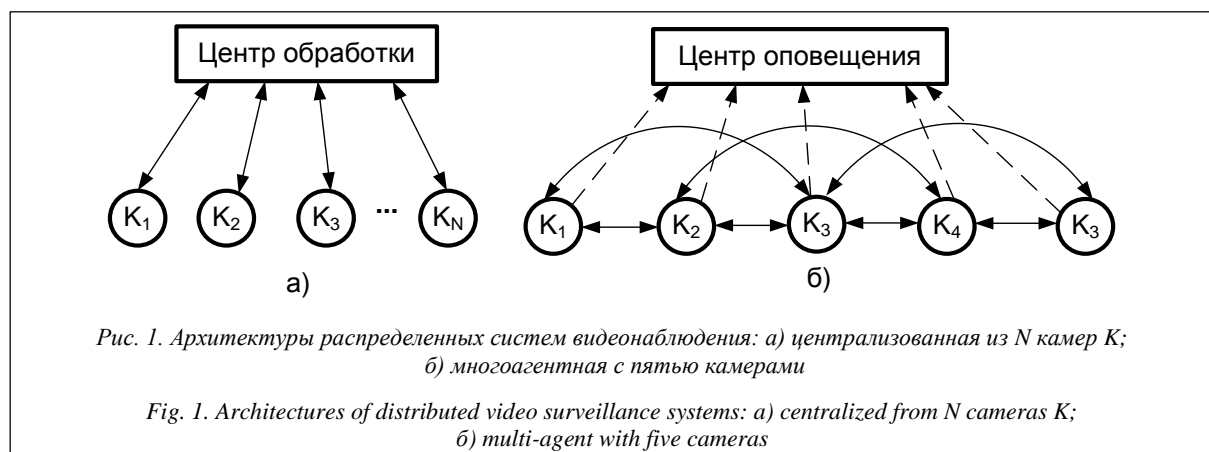


Рис. 1. Архитектуры распределенных систем видеонаблюдения: а) централизованная из N камер K; б) многоагентная с пятью камерами

Fig. 1. Architectures of distributed video surveillance systems: а) centralized from N cameras K; б) multi-agent with five cameras

ства изображения и малых размеров объектов сделать это не всегда возможно. Поэтому имеются некоторые контуры объектов, которые можно соотнести друг с другом в разных камерах, например по координатам. При невозможности определить местоположение объектов используются векторы движения распознанных объектов в кадре [8]. Образы с двух камер соотносятся друг с другом следующим образом:

$$b_{ij} = \begin{cases} 0, & \text{если } \arg \min(|v_i - v_j|) \neq i \text{ или } \min(|v_i - v_j|) > thresh, \\ 1, & \text{если } \arg \min(|v_i - v_j|) = i \text{ и } \min(|v_i - v_j|) < thresh, \end{cases}$$

где  $b$  – двумерный массив булевых переменных;

$v_i$  –  $i$ -й вектор скорости, распознанный в первой камере;  $v_j$  –  $j$ -й вектор скорости, распознанный во второй камере;  $thresh$  – некоторое заранее заданное пороговое значение.

Если  $b_{ij}$  равно 1, то  $i$ -й объект первой камеры является объектом  $j$  второй камеры. Если разность нескольких векторов движения из одной камеры мало отличается от разности векторов движения из другой, то образы дополнительно соотносятся по цвету с помощью алгоритма сегментации watershed [9] и HOG-дескрипторов.

На рисунке 2 показана схема алгоритма распознавания ситуаций и объединения результатов рас-

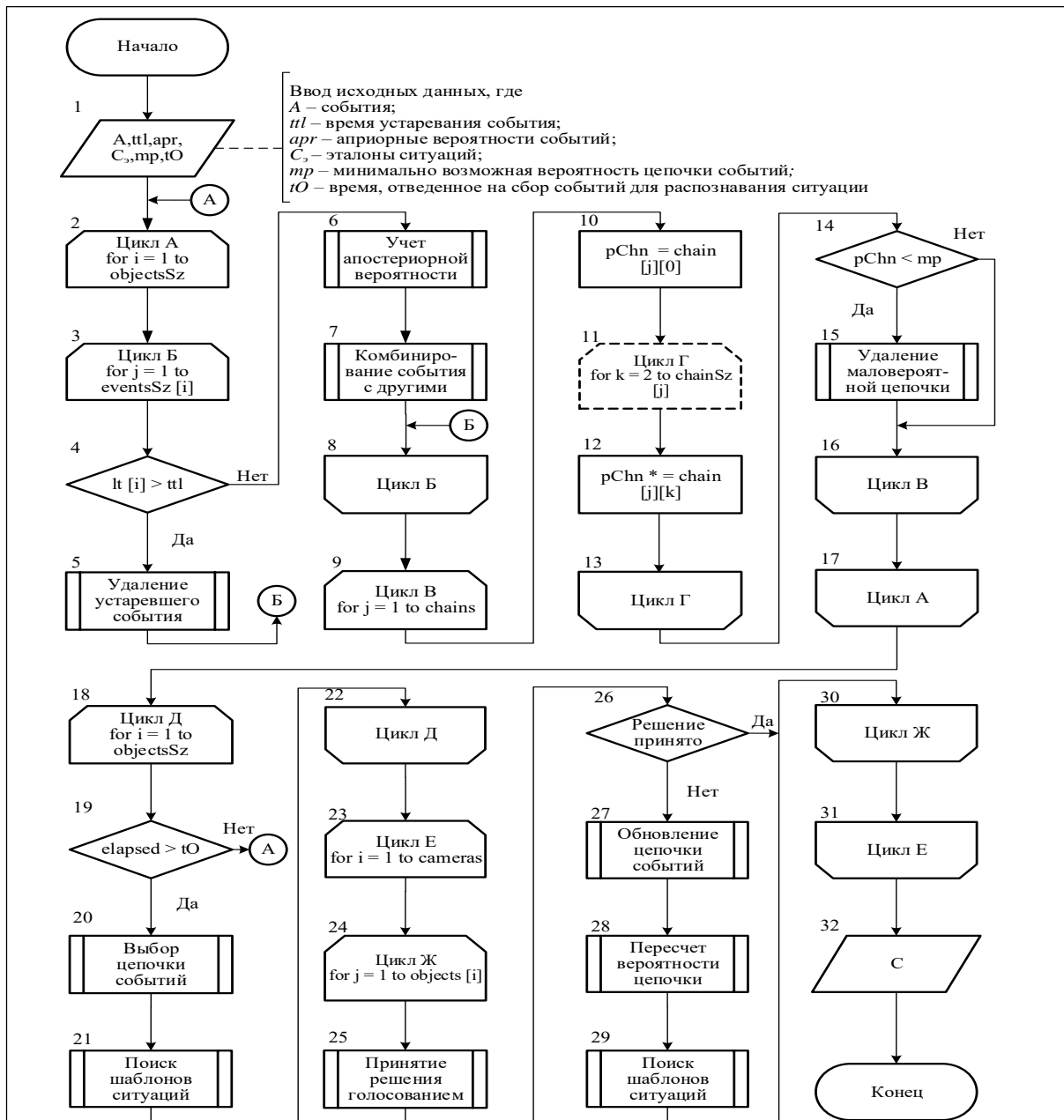


Рис. 2. Алгоритм распознавания ситуаций и объединения результатов распознавания

Fig. 2. The algorithm for recognizing situations and combining recognition results

познавания с разных камер. В блоке 1 происходит ввод исходных данных: распознанные события, время устаревания событий, априорные вероятности событий, эталоны ситуаций, минимально возможная вероятность цепочки событий, время на сбор событий для распознавания ситуации.

В блоках 2–17 осуществляется подготовка потока событий к распознаванию ситуаций. В связи с конечностью числа цепочки событий в распознаваемой ситуации и снижением требований к вычислительным ресурсам средств обработки данных устаревшие события удаляются из памяти. Если событие не устарело, в блоке 6 будет произведен пересчет вероятности события по формуле Байеса с учетом априорной и апостериорной вероятности этого события и результатов его распознавания с других камер. Сопоставление объектов в одной и другой камерах производится при помощи идентификаторов, в качестве которых выступают HOG-дескрипторы объектов, векторы скорости, а также цвет объекта. При превышении оценкой достоверности распознавания события в данной камере заданного порогового значения результаты текущего распознавания событий передаются близлежащим камерам. Это позволит формировать цепочки событий и у тех камер, в области зрения которых объект еще не появлялся.

В блоке 7 происходит комбинирование событий во все возможные цепочки. Для каждого объекта формируется свое множество цепочек. В блоках 9–16 происходит вычисление вероятности цепочки событий в результате перемножения вероятности всех событий, входящих в нее. Если вероятность цепочки меньше минимально допустимой, эта цепочка будет удалена в блоке 15.

В блоке 19 проверяется, вышло ли время, отведенное на сбор событий для распознавания ситуации. Если время не вышло, цепочка событий продолжает формироваться, если время вышло, в блоке 20 выбирается наиболее вероятная цепочка событий. Затем в блоке 21 происходит распознавание ситуации для объекта в отдельной камере. В блоках 23–31 объединяются результаты распознавания с разных камер.

В блоке 25 делается первоначальная попытка быстро принять решение. Вначале данные запрашиваются со всех интересующих камер, затем сопоставляются объекты одной и другой камер (сравнение идентификаторов). Далее осуществляется попытка принятия решения через взвешенное умножение вероятностей распознанных ситуаций на разных камерах для одного объекта в некоторый момент времени. Значимость камеры зависит от ее характеристик, размещения и общего числа камер. Здесь же учитывается число камер, в которых была распознана одна ситуация, а также учитывается, что вероятность цепочек в разных камерах не должна превышать некоторое пороговое значение. Если полученный результат не меньше требуемого

(то есть достоверен), то решение принимается, иначе выполняются блоки 27–29. В блоке 27 происходит пересоздание цепочки событий, но в новую цепочку входят все события, распознанные на всех камерах в заданный промежуток времени. В блоке 28 идет пересчет вероятности цепочки событий, а уже в блоке 29 выполняется распознавание новой, расширенной цепочки событий. В блоке 32 выводятся распознанные ситуации.

На рисунке 3 показано возможное применение разработанного алгоритма – перемещение человека в области зрения двух камер. Тестовые видеоданные взяты из [10]. Измеряются траектория движения, вектор движения в каждый момент времени и идентификационные характеристики объекта.

Алгоритм реализован в виде тестовой утилиты на языке программирования C++ в среде разработки Visual Studio 2013 и на первых этапах работает не с камерами, а с видеофайлами. В качестве методов распознавания низкого уровня использовались методы из библиотеки OpenCV (<http://opencv.org/>). Утилита включает в себя следующие основные модули:

- main (обработка аргументов программы и вызов функций распознавания);
- video (получение и обработка видеопотока от устройства);
- detection (детектирование графических объектов в кадре);
- obj\_recog (распознавание графических объектов);
- event\_recog (распознавание событий с графическими объектами);
- probability (пересчет вероятностей);
- situation (распознавание ситуаций).

Выходные результаты работы утилиты формируются в видеофайл, на рисунке 4 представлен пример изображения из него. Изображение взято из видеофайлов [11]. На рисунке вверху показан исходный кадр 326 из видеопотока, справа вверху – маска переднего плана, слева внизу – результаты распознавания без учета апостериорной информации, справа внизу – результаты распознавания с учетом апостериорной вероятности. В правом верхнем углу нижних частей изображения показано событие с наибольшей вероятностью. Тестирование утилиты проводилось в операционной системе Windows на различных вычислительных машинах. Результаты эксперимента показали увеличение достоверности распознавания при использовании данного алгоритма на 40 % по сравнению с вариантом, учитывающим только независимые события.

Так как данный алгоритм предполагает одновременное формирование цепочек событий камерами, которые связаны с возможной ситуацией, это позволит снизить время отклика системы. Алгоритм дает возможность распознавать ситуации на высоком уровне, оперируя с событиями, как с



Рис. 3. Человек в области зрения двух камер

Fig. 3. A person in the field of view of 2 chambers

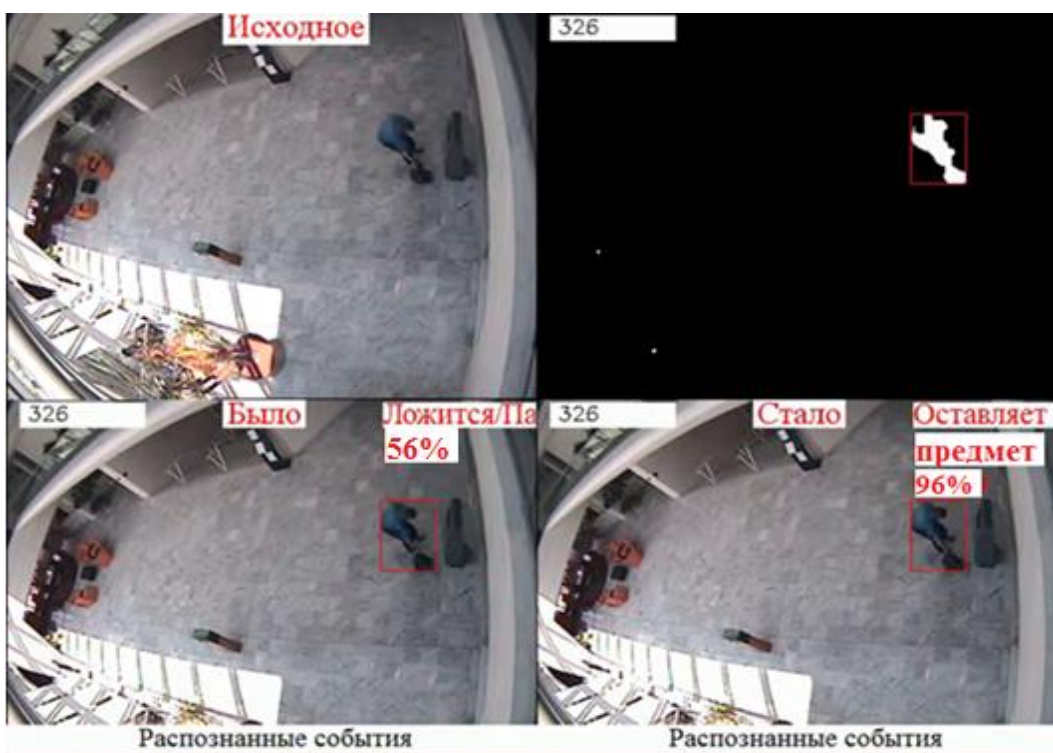


Рис. 4. Кадр из формируемого утилитой видеофайла

Fig. 4. A frame from the video file formed by a utility

символами грамматики, и используя данные множества камер. Преимуществом данного подхода является повышение достоверности распознавания ситуации за счет совмещения результатов работы нескольких узлов. Дальнейшая работа направлена на исследование стохастических грамматик, построенных по событиям в области зрения системы видеонаблюдения.

*Исследование выполнено при финансовой поддержке РФФИ и Министерства образования Оренбургской области в рамках научного проекта № 17-47-560368 р\_а.*

#### Литература

1. Скрипкина А.А. Обзор методов обнаружения движущегося объекта по видеонаблюдениям // Перспективы развития информационных технологий. 2011. № 3-1. С. 126–129.
2. Обухова Н.А. Обнаружение и сопровождение движущихся объектов методом сопоставления блоков // Информационно-управляющие системы. 2004. № 1. С. 30–35.
3. Фу К. Структурные методы в распознавании образов; [пер. с англ. З.В. Завалишина, С.В. Петрова, Р.Л. Шейнина; под ред. М.А. Айзермана]. М.: Мир, 1977. 319 с.
4. Степин Д. Некоторые аспекты проектирования IP-систем видеонаблюдения // Алгоритм безопасности. 2014. № 6. С. 34–37.
5. Бурков А.В. IP-видеонаблюдение глазами разработчика. Ч. 1: Базовые понятия и IP-камера // Алгоритм безопасности. 2016. № 6. С. 76–78.

6. Портнов Д. Современные тенденции встроенной видеоаналитики видеокamer // Алгоритм безопасности. 2016. № 6. С. 20–21.

7. Кручинин А.Ю. Автоматическая внешняя калибровка камер на основе анализа траекторий движений объектов // Информационные технологии в науке, образовании и производстве (ИТНОП): сб. тр. V Междунар. науч.-технич. конф. 2012. С. 1–6.

8. Колмыков Д.В., Кручинин А.Ю. Распознавание ситуаций в распределенной системе видеонаблюдения без единого

центра // Прикладная математика и информатика: современные исследования в области естественных и технических наук: сб. тр. III науч.-практич. Всерос. конф. 2017. С. 276–280.

9. Meyer F. Colour image segmentation. Proc. IEE Int. Conf. Image Processing and Its Applications, 1992, pp. 303–306.

10. Computer Vision Laboratory CVLab. URL: <http://cvlab.epfl.ch/data/pom> (дата обращения: 20.08.2017).

11. CAVIAR Test Case Scenarios URL: <http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/> (дата обращения: 20.08.2017).

Software & Systems

DOI: 10.15827/0236-235X.122.368-373

Received 07.09.17

2018, vol. 31, no. 2, pp. 368–373

### THE ALGORITHM FOR RECOGNIZING SITUATIONS IN A DISTRIBUTED VIDEO SURVEILLANCE SYSTEM

A.Yu. Kruchinin<sup>1</sup>, Ph.D. (Engineering), Associate Professor, [kruchinin-al@mail.ru](mailto:kruchinin-al@mail.ru)

D.V. Kolmykov<sup>1</sup>, Master of Science, [malin.chyn@gmail.com](mailto:malin.chyn@gmail.com)

R.R. Galimov<sup>1</sup>, Candidate Ph.D. (Engineering), Associate Professor, [rin-galimov@yandex.ru](mailto:rin-galimov@yandex.ru)

<sup>1</sup> Orenburg State University, Pobedy Ave. 13, Orenburg, 460018, Russian Federation

**Abstract.** CCTV systems are the most important means for preventing and timely handling of contingencies, such as crimes, emergency situations. A big number of cameras and a large control area makes it necessary to introduce video analytics to recognize dangerous situations. In this case it is necessary to take into account the data from a number of video cameras both for detecting a motion path of a recognized object and for increasing recognition reliability.

The article proposes an algorithm for recognizing emergencies for a distributed video surveillance system based on stochastic grammars. Recognition of a situation occurs at 3 levels: images are recognized at the lower level, events are recognized at the average level, and situations are recognized at the top. To reduce the system response time, it is proposed to use a multi-agent architecture that allows distributing the load between intelligent cameras. Data exchange occurs only between nearby nodes, so network traffic reduces.

The use of a large number of cameras involves zones controlled by several nodes. Combination of detection results from several cameras makes it possible to increase the estimate reliability. However, it is required to know the mutual arrangement of chambers and the angles of their turns. The article suggests some methods for automatic calibration of cameras in a distributed video surveillance system, ways of combining images from different cameras, in particular, based on speed rate vectors of objects.

Taking into account certain features of a distributed video surveillance system, there is a developed algorithm for recognizing emergencies for an intelligent surveillance camera. Each camera generates probable situations based on previously recognized events. When a threshold value of probabilistic evaluation of the detection result is exceeded, its refinement is carried out in the process of interaction with neighboring nodes.

**Keywords:** distributed video surveillance system, emergency recognition, multi-agent system.

**Acknowledgements.** The study has been financially supported by RFBR and the Ministry of Education of the Orenburg region within the framework of the scientific project no. 17-47-560368.

#### References

1. Skripkina A.A. A review of methods for detecting a moving object from video images. *Perspektivy razvitiya informatsionnykh tekhnology* [Prospects for IT Development]. 2011, no. 3-1, pp. 126–129 (in Russ.).

2. Obukhova N.A. Detection and tracking moving objects based on block matching algorithm. *Informatsionno-upravlyayushchie sistemy* [Information and Control Systems]. 2004, no. 1, pp. 30–35 (in Russ.).

3. Fu K.S. *Structural Methods in Pattern Recognition*. Academic Press, 1969, 227 p. (Russ. ed.: Z.V. Zavalishin, S.V. Petrov, R.L. Sheinin, M.A. Aizerman]. Moscow, Mir Publ., 1977, 319 p.

4. Stepin D. Some aspects of the design of IP-based video surveillance systems. *Algoritm bezopasnosti* [A Security Algorithm]. 2014, no. 6, pp. 34–37 (in Russ.).

5. Burkov A.V. IP-video surveillance through developer's eyes. Part 1: Basic Concepts and an IP Camera. *Algoritm bezopasnosti* [A Security Algorithm]. 2016, no. 6, pp. 76–78 (in Russ.).

6. Portnov D. Modern trends of built-in video analytics of video cameras. *Algoritm bezopasnosti* [A Security Algorithm]. 2016, no. 6, pp. 20–21 (in Russ.).

7. Kruchinin A.Yu. Automatic external calibration of cameras based on the analysis of object movement trajectories. *ITNOP: sb. tr. V Mezhdunar. nauch.-tekhnich. konf.* [Proc. V Int. Sci. and Tech. Conf. ITNOP]. Orel, 2012, pp. 1–6 (in Russ.).

8. Kolmykov D.V., Kruchinin A.Yu. Recognition of situations in a distributed video surveillance system without a unified center. *Prikladnaya matematika i informatika: sovremennye issledovaniya v oblasti estestvennykh i tekhnicheskikh nauk: sb. tr. III nauch.-praktich. Vseros. konf.* [Applied Mathematics and Informatics: Modern Research in the Field of Natural and Engineering Sciences: Proc. III Sci. and Pract. All-Russian Conf.]. Togliatti, 2017, pp. 276–280 (in Russ.).

9. Meyer F. Color image segmentation. *Proc. IEE Int. Conf. on Image Processing and its Applications*. Maastricht, Netherlands, 1992, pp. 303–306.

10. *Computer vision laboratory CVLab*. Available at: <http://cvlab.epfl.ch/data/pom> (accessed August 8, 2017).

11. *CAVIAR Test Case Scenarios*. Available at: <http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/> (accessed August 8, 2017).