

## Захват и отслеживание удаленных объектов в видеопотоке

© А.Н. Алфимцев, Н.А. Демин

МГТУ им. Н.Э. Баумана, Москва, 105005, Россия

*Рассмотрено пять алгоритмов захвата и отслеживания удаленных объектов в видеопотоке: Лукаса – Канаде, шаблонов движений, сдвига среднего, Виолы – Джонса, непрерывно адаптирующегося сдвига. Проведен сравнительный анализ данных алгоритмов, выявлены их достоинства и недостатки. Проверена эффективность и представлены результаты работы каждого алгоритма применительно к задаче отслеживания удаленного объекта. Проведен анализ вероятностей возникновения ошибок I и II рода при использовании алгоритмов. Рассмотрены особенности алгоритмов при обработке видеопотока. Выделены два наиболее подходящих алгоритма для захвата и отслеживания. Предложен интегральный алгоритм, использующий возможности алгоритма Лукаса – Канаде и алгоритма Виолы – Джонса для захвата и отслеживания удаленного объекта.*

**Ключевые слова:** захват объекта, отслеживание объекта, видеопоток, цифровое изображение, распознавание образов.

**Введение.** Использование электронно-вычислительных средств цифровой обработки изображений способствует расширению функциональных возможностей бытовых устройств. Стабилизация, автоматическая настройка экспозиции, отслеживание лиц и многие другие подобные задачи решаются этими устройствами в реальном времени. Увеличение вычислительных мощностей процессоров позволяет разработчику реализовывать все более сложные алгоритмы и оснащать создаваемые устройства новыми функциями, практически граничащими с ожидаемыми возможностями искусственного интеллекта. Одной из фундаментальных основ современной интеллектуальной обработки изображений является захват движения некоторого объекта в видеопотоке и последующее его отслеживание.

*Цифровое изображение* (кадр), получаемое видеокамерой в момент времени  $t$  и имеющее разрешение по вертикали  $h$  1600 пикселей, а по горизонтали  $w$  — 1200, обозначим  $I_t(h, w)$ . *Видеопоток* — это последовательность цифровых изображений (кадров)  $I_t(h, w)$ ,  $I_{t+1}(h, w)$ , ...,  $I_{t+k}(h, w)$ . Под объектом понимается прямоугольная область интересов  $Ob_t(X, Y)$ , представленная множеством пикселей цифрового изображения  $I_t(h, w)$ , очерчивающих искомый объект, содержащая  $X$  пикселей по вертикали и  $Y$  по горизонтали. *Захватом объекта* называется выделение области интересов  $Ob_t(X, Y)$  на циф-

ровом изображении  $I_t(h, w)$  в момент времени  $t$ . *Отслеживанием* называется захват  $Ob_t(X, Y)$  на последовательности цифровых изображений. Под захватом объекта *в реальном времени* понимается обработка видеопотока с частотой не менее 10 кадров в секунду. *Обучением* называется предварительная настройка и задание параметров алгоритма захвата и отслеживания для обнаружения интересующего объекта. Обучение может выполняться на основе обучающей выборки, состоящей из одного или нескольких обучающих изображений. Под обучающим цифровым изображением понимается изображение  $I_t(h, w)$  с дополнительной информацией о наличии или отсутствии на нем искомого объекта.

Для решения задач захвата и отслеживания создано множество алгоритмов, позволяющих проектировать актуальное программное обеспечение для различных сфер применения. В данной работе проведен анализ алгоритмов Лукаса – Канаде, шаблонов движений, сдвига среднего, Виолы – Джонса, непрерывно адаптирующегося сдвига применительно к решению задач захвата и отслеживания удаленных объектов в видеопотоке. Под *удаленным объектом* здесь понимается объект, находящийся на расстоянии более 100 м от пассивного сенсора — видеокамеры, не использующего оптические и программные средства увеличения изображения. Работа с данным типом объекта значительно усложняет задачу захвата и отслеживания. Не все алгоритмы успешно справляются с ее решением из-за возрастающего уровня шумов и оптических иллюзий, возникающих в потоке цифровых изображений.

**Краткий обзор алгоритмов захвата и отслеживания.** Для анализа выбраны следующие наиболее часто используемые алгоритмы для захвата и отслеживания движущихся объектов  $Ob_t(X, Y)$ , исходный код которых свободно доступен на языке программирования C++ в сети Интернет для экспериментов и модификации:

- шаблонов движений (*Motion Templates*) — основан на поиске границ объектов в каждом кадре видеопотока  $I_t(h, w)$  [1]. Смещение границы на новом кадре  $I_{t+1}(h, w)$  относительно предыдущего задает вектор движения объекта. Данный алгоритм наиболее эффективен при движении крупных объектов и часто используется для распознавания динамических жестов в человеко-машинных интерфейсах;

- сдвига среднего (*Mean-Shift*) — основан на математической модели, которая заключается в том, что вычисляется локальный экстремум плотности распределения набора характерных точек, т. е. алгоритм отслеживает смещение центра масс точек, определяющих объект слежения, получая на выходе вектор движения объекта [2]. Высокая эффективность достигается при ярко выраженном цветовом различии объекта и фона;

• непрерывно адаптирующегося сдвига (*CamShift*) — основан на алгоритме сдвига среднего, но отличается тем, что автоматически подстраивает границы и размер окна, в пределах которого расположены характерные точки [3]. Таким образом производится более точное отслеживание объекта, изменяющегося в размерах;

• Лукаса – Канаде (*Lucas – Kanade*) — основан на дифференциальном вычислении оптического потока с помощью анализа пикселей (предполагается, что оптический поток одинаков для пикселей, лежащих в окрестности центра окна слежения), при этом смещение пикселей между соседними кадрами  $I_t(h, w)$  должно быть невелико [4]. Данный алгоритм более двадцати лет активно используется в приложениях компьютерного зрения и уже доказал свою высокую эффективность для широкого круга задач;

• Виолы – Джонса (*Viola – Jones*) — основан на обнаружении в кадре  $I_t(h, w)$  наборов пикселей, совпадающих с заранее подобранными шаблонами, состоящими из белых и черных прямоугольников [5, 6]. Для распознавания различных объектов требуется свой уникальный набор шаблонов, который создается путем обучения алгоритма на конкретном объекте. Правильно обученный алгоритм работает с высокой эффективностью, однако сам процесс обучения достаточно трудоемкий, требует от разработчика-исследователя специальных знаний о данном алгоритме и хорошо организованной обучающей выборки.

**Анализ алгоритмов захвата и отслеживания объектов.** Для начала работы любого алгоритма отслеживания требуется каким-либо образом инициализировать первоначальную область слежения или массив характерных точек. Существуют алгоритмы, самостоятельно обнаруживающие искомый объект  $Ob_t(X, Y)$  в кадре  $I_t(h, w)$  при его появлении, а для остальных объектов либо указывается вручную, либо его координаты/характерные точки передаются из функции обнаружения. Кроме того, некоторые алгоритмы требуют обучения перед их использованием. Все вышеописанное определяет степень автономности, информация о которой для исследуемых алгоритмов приведена в табл. 1.

Таблица 1

Степень автономности алгоритмов слежения

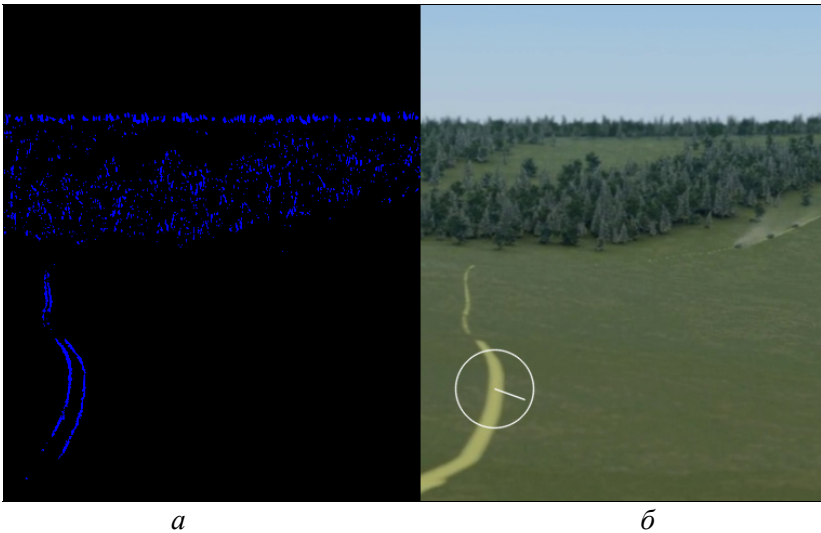
Алгоритм	Автообучение	Автоинициализация
Шаблоны движений	+	+
Сдвиг среднего	+	–
Непрерывно адаптирующийся сдвиг	+	–
Лукаса – Канаде	+	–
Виолы – Джонса	–	+

Кроме того, в процессе исследования рассмотренные алгоритмы были проверены на устойчивость под влиянием различных негативных параметров. Результаты проверки приведены в табл. 2.

Таблица 2

**Влияние негативных параметров видеопотока на работу алгоритмов слежения**

Алгоритм	Негативный параметр			
	Объект движется очень быстро	Объект мало отличим по цвету от фона	Подвижный фон	Размеры объекта меняются со временем
Шаблоны движений	–	+	–	+
Сдвиг среднего	–	–	+	–
Непрерывно адаптирующийся сдвиг	–	–	+	+
Лукаса – Канаде	–	+	+	–
Виолы – Джонса	+	+	+	+

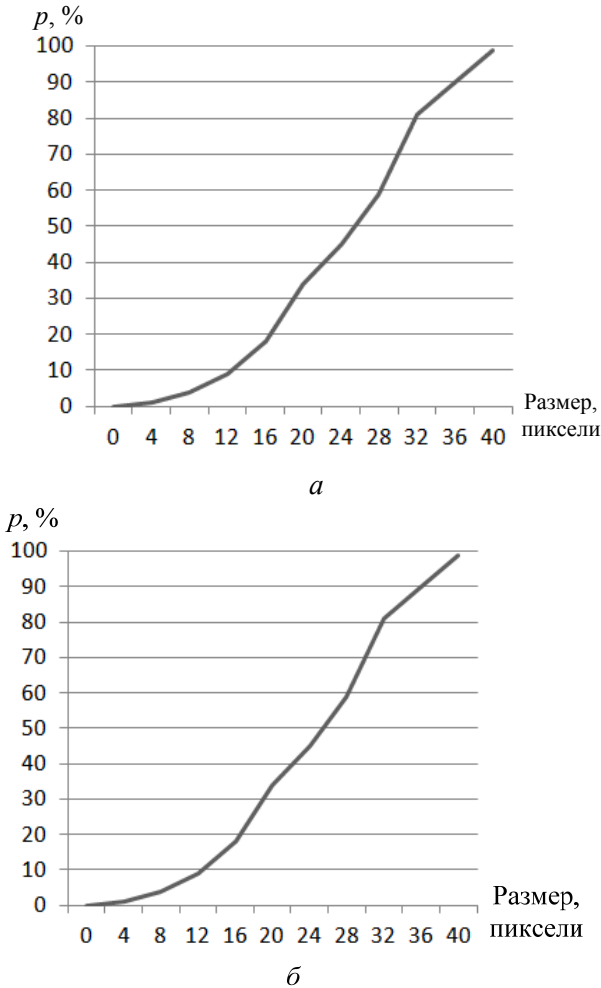


**Рис. 1.** Алгоритм шаблонов движений ошибочно принимает дорогу за движущийся объект:  
*a* — обработанное изображение; *б* — исходное изображение с выделенным объектом

Рассмотрим алгоритмы, которые позволяют автоматически распознать нужный объект  $Ob_i(X, Y)$ , — Виолы – Джонса и шаблоны движений. При автоматическом решении данной задачи существует вероятность возникновения ошибки II рода: захват участка фона вме-

сто отслеживаемого объекта  $Ob_i(X, Y)$ . Это происходит если участок имеет определенные черты, сходные с характерным набором, наличие которого позволяет алгоритму сделать вывод, что распознаваемый объект  $Ob_i(X, Y)$  найден. К примеру, алгоритм шаблонов движений опирается при захвате объекта  $Ob_i(X, Y)$  на цифровом изображении  $I_i(h, w)$  на резкие цветовые границы. Во время его работы любое движение фона, на котором имеются подобные участки, может вызвать ошибку распознавания (рис. 1).

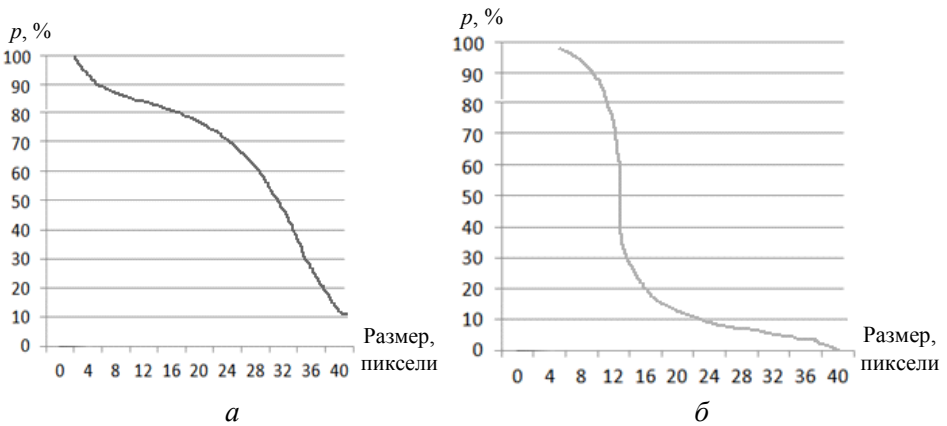
На рис. 2 приведены графики зависимости вероятности возникновения ошибки II рода от размера области с характерными свойствами для шаблона Виолы – Джонса и шаблонов движений.



**Рис. 2.** Вероятность  $p$  возникновения ошибки II рода в зависимости от размера «отвлекающих» областей на фоне для шаблонов движений (а) и Виолы – Джонса (б) (размер объекта слежения составляет 25 пикселей)

Алгоритм шаблонов движений не имеет каких-либо недостатков, делающих его априорно непригодным для решения поставленной задачи, однако при его использовании вероятность возникновения ошибки II рода очень велика: при размере «отвлекающей» области больше размера объекта  $Ob_i(X, Y)$  вероятность ошибки — от 60 до 100 % (см. рис. 2). В свою очередь, алгоритм Виолы – Джонса обладает гораздо большей точностью и является эффективным средством, однако в некоторых случаях он непригоден (к примеру, во время потоковой обработки видео), так как требует начального обучения и высоких вычислительных мощностей.

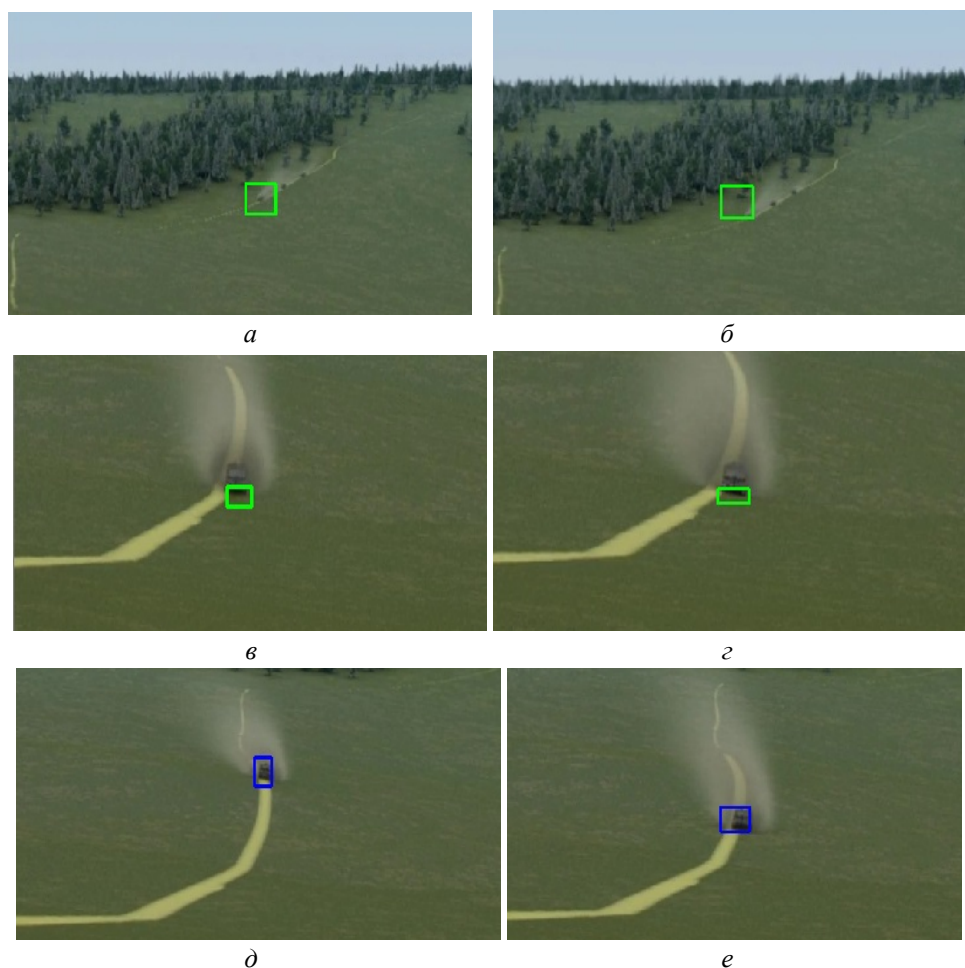
В отличие от этих алгоритмов, алгоритмы Лукаса – Канаде и непрерывно адаптирующегося сдвига не имеют описанного выше недостатка (ошибок II рода), поскольку не осуществляют непосредственно поиск движущегося объекта  $Ob_i(X, Y)$ , а лишь сопровождают уже обнаруженный другими методами объект. Однако при их использовании существует вероятность потери фокуса на объекте  $Ob_i(X, Y)$  во время отслеживания (ошибка I рода). Зависимость вероятности возникновения этой ошибки от размера объекта  $Ob_i(X, Y)$  для алгоритма Лукаса – Канаде и непрерывно адаптирующегося сдвига изображена на рис. 3, где в качестве объекта отслеживания взят автомобиль, который по цвету не сильно отличается от фона. Алгоритм сдвига среднего показал те же результаты, что и алгоритм непрерывно адаптирующегося сдвига, а так как он является предшественником непрерывно адаптирующегося сдвига, его результаты не представлены.



**Рис. 3.** Вероятность  $p$  возникновения ошибки I рода для алгоритмов Лукаса – Канаде (а) и непрерывно адаптирующегося сдвига (б) в зависимости от размера объекта

Как видно из рис. 3, в поставленном эксперименте алгоритм непрерывно адаптирующегося сдвига заметно проигрывает алгоритму Лукаса – Канаде. Это объясняется тем, что для достижения высо-

кой эффективности при использовании алгоритма непрерывно адаптирующегося сдвига отслеживаемый объект  $Ob_i(X, Y)$  должен иметь четкую цветовую границу с фоном, а в данном случае это требование не выполняется. Примеры работы алгоритма Лукаса – Канаде и непрерывно адаптирующегося сдвига показаны на рис. 4. В реальных условиях довольно часто возникает необходимость отслеживания объекта  $Ob_i(X, Y)$ , практически неотличимого от фона, для этого больше подходит алгоритм Лукаса – Канаде. Существует несколько путей использования этого алгоритма при решении задачи отслеживания удаленного объекта  $Ob_i(X, Y)$ . Рассмотрим один из них.



**Рис. 4.** Демонстрация работы алгоритмов непрерывно адаптирующегося сдвига (а–г) и Лукаса – Канаде (д, е):

а, б — алгоритм теряет удаленный объект; в, г — при приближении объекта к камере алгоритм захватывает участок объекта, сильнее всего отличающийся по цвету от фона, — черный бампер автомобиля; д, е — алгоритм корректно отслеживает удаленный объект

**Интегральный алгоритм захвата и отслеживания удаленного объекта.** Для комплексного решения задачи отслеживания удаленных объектов необходимо использовать два алгоритма: захвата объекта  $Ob_i(X, Y)$  и его отслеживания. Для этого лучше всего подходят алгоритмы Виолы – Джонса и Лукаса – Канаде. В случае отслеживания наилучший выбор — это алгоритм Лукаса – Канаде. Использование алгоритма Виолы – Джонса в этих целях нецелесообразно, поскольку в таком случае потребуются высокие вычислительные мощности, а алгоритм непрерывно адаптирующегося сдвига уступает алгоритму Лукаса – Канаде по вероятности ошибки I рода. В свою очередь, алгоритм шаблонов движений не подходит для отслеживания удаленного объекта  $Ob_i(X, Y)$ , что подтверждается полученными данными.

Приведенная ниже последовательность шагов представляет собой интегральный алгоритм захвата и отслеживания объекта  $Ob_i(X, Y)$ , эффективным образом использующий возможности алгоритмов Лукаса – Канаде и Виолы – Джонса.

*Шаг 1.* С помощью алгоритма Виолы – Джонса (или через ручной ввод, если невозможно выполнить предварительное обучение) определяются координаты объекта отслеживания  $Ob_i(X, Y)$ .

*Шаг 2.* Инициализируется массив характерных точек (дескрипторов) объекта  $Ob_i(X, Y)$ .

*Шаг 3.* Для следующего кадра  $I_t(h, w)$  видеопотока начинает выполняться алгоритм Лукаса – Канаде, который принимает на вход массив характерных точек.

*Шаг 4.* Если какая-либо из характерных точек слишком отдалена от центра масс характерных точек, она удаляется.

*Шаг 5.* Если число характерных точек стало ниже определенного порогового значения, объект  $Ob_i(X, Y)$  считается потерянным, производится переход к шагу 1.

*Шаг 6.* На экран выводится кадр  $I_t(h, w)$ , на котором отмечен объект отслеживания  $Ob_i(X, Y)$ .

*Шаг 7.* Производится переход к шагу 3.

Исходя из вычислительных оценок работы алгоритмов Лукаса – Канаде [7] и Виолы – Джонса [5], верхнюю оценку вычислительной сложности рассмотренного интегрального алгоритма можно оценить как  $O(N)$ , где  $N$  — количество пикселей исследуемого изображения  $I_t(h, w)$ . Таким образом, алгоритм имеет линейную оценку сложности, что позволяет проводить захват и отслеживание удаленного объекта  $Ob_i(X, Y)$  в реальном времени.

**Заключение.** Анализ особенностей алгоритмов при обработке видеопотока и вероятностей возникновения ошибок I и II рода рассмотренных алгоритмов позволил выделить два наиболее подходя-



щих алгоритма для захвата и отслеживания удаленного объекта: Виолы – Джонса и Лукаса – Канаде. Алгоритм Виолы – Джонса является мощным средством и способен эффективно и корректно захватывать требуемый объект слежения, а алгоритм Лукаса – Канаде — с высокой точностью отслеживать его перемещение. Совместное использование позволило объединить их достоинства и построить устойчивый механизм трекинга, который имеет возможность обрабатывать видеопотоки с плохо различимыми объектами и движущимся фоном, а также осуществлять слежение одновременно за несколькими объектами.

## ЛИТЕРАТУРА

- [1] Müller M., Röder T. Motion templates for automatic classification and retrieval of motion capture data. *Symposium on Computer Animation — SCA*. Vienna, Austria, 2006, pp. 137–146.
- [2] Comaniciu D., Meer P. Mean Shift: A Robust Approach Toward Feature Space Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence — PAMI*, 2002, vol. 24, no. 5, pp. 603–619.
- [3] Wang Z., Yang X., Xu Y., Yu S. CamShift guided particle filter for visual tracking. *Pattern Recognition Letters — PRL*, 2009, vol. 30, no. 4, pp. 407–413.
- [4] Baker S., Matthews I. Lucas – Kanade 20 Years On: A Unifying Framework. *International Journal of Computer Vision — IJCV*, 2004, vol. 56, no. 3, pp. 221–255.
- [5] Алфимцев А.Н., Лычков И.И. Метод обнаружения объекта в видеопотоке в реальном времени. *Вестник Тамбовского государственного технического университета*, 2011, т. 17, № 1, с. 44–55.
- [6] Kasinski A., Schmidt A. The architecture and performance of the face and eyes detection system based on the Haar cascade classifiers. *Pattern Analysis and Applications — PAA*, 2010, vol. 13, no. 2, pp. 197–211.
- [7] Liu H. Accuracy vs Efficiency Trade-offs in Optical Flow Algorithms. *Computer Vision and Image Understanding — CVIU*, 1998, vol. 72, no. 3, pp. 271–286.

Статья поступила в редакцию 28.06.2013

Ссылку на эту статью просим оформлять следующим образом:

Алфимцев А.Н., Демин Н.А. Захват и отслеживание удаленных объектов в видеопотоке. *Инженерный журнал: наука и инновации*, 2013, вып. 11. URL: <http://engjournal.ru/catalog/it/hidden/1049.html>

**Алфимцев Александр Николаевич** окончил МГТУ им. Н.Э. Баумана в 2005 г. Канд. техн. наук, доцент кафедры «Информационные системы и телекоммуникации» МГТУ им. Н.Э. Баумана. Автор более 70 научных публикаций, в том числе пяти охранных документов на интеллектуальную собственность. Область научных интересов: методы искусственного интеллекта, мультимодальные интерфейсы, распознавание образов. e-mail: [alfim@bmstu.ru](mailto:alfim@bmstu.ru)

**Демин Никита Александрович** — студент кафедры «Информационные системы и телекоммуникации» МГТУ им. Н.Э. Баумана. Автор двух научных публикаций. Область научных интересов: распознавание образов, сетевые интерфейсы, цифровая обработка изображений. e-mail: [dyominick@gmail.com](mailto:dyominick@gmail.com)