

УДК 004.932.72'1

С. Г. Антощук, д-р техн. наук,  
Н. В. Коваленко, Н. А. Годовиченко

## ОТСЛЕЖИВАНИЕ ОБЪЕКТОВ ИНТЕРЕСА ПРИ ПОСТРОЕНИИ АВТОМАТИЗИРОВАННЫХ СИСТЕМ ВИДЕОНАБЛЮДЕНИЯ ЗА ЛЮДЬМИ

*Аннотация.* Проведен анализ наиболее часто встречающихся методов отслеживания движущихся объектов, которые используются для построения системы автоматизированного наблюдения за людьми. Проведено тестирование и оценка работы на тестовом наборе видеопоследовательностей.

*Ключевые слова:* отслеживание, видеонаблюдение.

S. G. Antoshchuk, ScD.,  
N. V. Kovalenko, N. A. Godovichenko

## USING OBJECT OF INTERESTS TRACKING IN THE DEVELOPMENT OF HUMAN VIDEO SURVEILLANCE SYSTEMS

*Abstract.* The analysis of the most common moving object tracking methods used for the development of automated human surveillance systems was conducted. The testing and performance evaluation of said methods using a set of benchmark videos was held.

*Keywords:* object tracking, surveillance.

С. Г. Антощук, д-р техн. наук,  
М. В. Коваленко, М.А. Годовиченко

## ВІДСТЕЖЕННЯ ОБ'ЄКТІВ ІНТЕРЕСУ ПРИ ПОБУДОВІ АВТОМАТИЗОВАНИХ СИСТЕМ ВИДЕОСПОСТЕРЕЖЕННЯ ЗА ЛЮДЬМИ

*Анотація.* Проведено аналіз найпоширеніших методів відстеження об'єктів, що рухаються, які використовуються для побудови систем автоматизованого спостереження за людьми. Проведено тестування та оцінку роботи на тестовому наборі відеопослідовностей.

*Ключові слова:* відстеження, спостереження.

**Актуальность проблемы.** В последнее время широкое распространение получили системы видеонаблюдения, которые позволяют улучшить работу охранных структур, обеспечить более эффективное функционирование ряда производственных процессов. Однако наряду с неоспоримыми достоинствами системы видеонаблюдения имеют существенный недостаток – необходимость большого количества людских ресурсов для мониторинга и анализа видео, поступающего с камер.

Эффективность систем видеонаблюдения может значительно возрасти при наличии возможности детектирования определенных событий и выполнения, требуемых действий в реальном времени (например, привлечение внимания оператора к происходящему) [1,2].

Поэтому в настоящее время интенсивно развиваются автоматизированные

системы видеонаблюдения, успешность построения которых зависит от решения нескольких ключевых задач, одной из них является отслеживание (трекинг) объектов интереса.

Существует множество методов решения задачи отслеживания. Среди наиболее часто встречающихся методов можно выделить вычитание фона, метод сдвига среднего, метод оптического потока и метод активных контуров [3].

Выбор наиболее эффективного метода зависит от задачи, решаемой при мониторинге, от требований к конкретной системе. В данной статье авторами рассматривалось использование различных методов отслеживания применительно к системам видеонаблюдения за людьми. Подобного рода системы являются наиболее распространенными и устанавливаются в различных охранных системах, в системах наблюдения в местах скопления людей, в торговых центрах, вокзалах и т.д.

© Антощук С.Г., Коваленко Н.В.,  
Годовиченко Н.А., 2012

### Анализ алгоритмов отслеживания

*Метод вычитания фона.* Данный метод основан на отделении объектов переднего плана от статического фона.

Основная задача при использовании данного метода – получить модель статического фона.

После получения модели фона, необходимо сравнить попиксельно текущий кадр с модельным. Области текущего кадра, которые значительно отличаются от модельного, считаются областями объектов переднего плана

$$|I_t(x, y) - B_t(x, y)| > \tau, \quad (1)$$

где  $I_t(x, y)$  и  $B_t(x, y)$  означают интенсивность пикселя на текущем кадре и его значение в модели статического фона в кадре  $t$ .

В самом простом случае модель статического фона можно получить с помощью сравнения разности двух кадров с неким порогом

$$|frame_i - frame_{i-1}| > Th. \quad (2)$$

Можно использовать более сложные методы получения фона, например, использование среднего значения, полученного из нескольких кадров. Такой подход сильно зависит от выбора порога и плохо работает при изменении освещенности, наличии фонового движения (ветки деревьев, морские волны), движении камеры.

Более перспективным является построение модели фона с помощью вероятностных моделей. Такие алгоритмы используют нормальное распределение или смесь нормальных распределений для моделирования пикселя фона. Это позволяет отделять объекты переднего плана более точно и справляться с изменением освещения, динамическим фоном и т.д.

Общим недостатком для всех алгоритмов вычитания фона является предположение о том, что камера неподвижна. Любое движение камеры приводит к изменению сцены и необходимости заново строить модель фона.

*Метод сдвига среднего.* Этот подход основан на поиске кластеров в объединенной области пространства и цвета

$[l, u, v, x, y]$ , где  $[l, u, v]$  представляет цвет, а  $[x, y]$  представляет пространственные координаты.

Алгоритм инициализируется большим количеством гипотетических кластеров с центрами, выбранными случайным образом. Затем каждый кластер сдвигается на среднее значение данных, лежащих внутри многомерного эллипсоида, центр которого совпадает с центром кластера. При сдвигах некоторые кластеры могут объединяться.

Использование метода сдвига среднего требует тщательной настройки различных параметров, например, выбор цвета и ширины ядра, минимальный размер кластера и т.д.

*Метод активных контуров.* При использовании данного метода отслеживания объектов осуществляется путем эволюционирования некоторого закрытого контура в сторону границ объекта так, чтобы контур плотно охватывал область объекта. Эволюция контура регулируется функционалом энергии, который определяет то, насколько контур соответствует гипотетической области объекта. Функционал энергии определяется

$$E(\Gamma) = \int_0^1 (E_{int}(v(s)) + E_{im}(v(s)) + E_{ext}(v(s))) ds, \quad (3)$$

где  $s$  – длина контура  $\Gamma$ ,  $E_{int}$  включает ограничения регуляризации,  $E_{im}$  включает энергию, основанную на внешних признаках, и  $E_{ext}$  определяет дополнительные ограничения.  $E_{int}$  обычно включает терм кривизны.

Энергия изображения  $E_{im}$  может вычисляться локально или глобально, в первом случае с использованием градиента изображения, а во втором – с включением признаков цвета и текстуры.

Метод активных контуров имеет множество достоинств – он робастен к шуму, контрасту и позволяет отслеживать изменение формы объекта, например, если необходимо определять позу человека. К недостаткам можно отнести то, что объект интереса должен быть достаточно большого размера.

*Метод оптического потока.*

Оптический поток – это поле плотности смещения векторов, которое определяет перемещение пикселя в области. Оно вычисляется с использованием предположения о постоянстве яркости, которое предполагает постоянство яркости соответствующих пикселей в соседних кадрах.

Метод оптического потока позволяет отслеживать объекты на сложном фоне и при движениях камеры. Недостатками данного метода является невысокая точность и необходимость большого количества вычислений.

**Сравнение методов и оценка их работы**

Для сравнения эффективности алгоритмов отслеживания необходимо было решить две задачи: выбор контрольных данных и показателей эффективности.

Анализ показал, что выбор контрольных данных сильно влияет на интерпретацию результатов оценки эффективности. Существует несколько широко используемых в различных исследованиях тестовых наборов данных. В данной статье в качестве тестовых клипов для оценки работы разных методов отслеживания объектов были выбраны три клипа из базы CAVIAR [4].

Оценка эффективности алгоритмов отслеживания может проводиться с использованием контрольной выборки либо без. Во втором случае производится оценка признаков цвета и движения, либо статистически определенной модели.

При использовании контрольной выборки для количественной оценки эффективности могут использоваться показатели, основанные на анализе кадра или основанные на анализе объекта. Показатели на анализе кадра основываются на сравнении пикселей контрольного изображения и кадра, содержащего движущийся объект. Они используются для измерения эффективности всей системы в целом.

В данной статье предлагается для количественного сравнения эффективности методов отслеживания движущихся объектов

использовать показатели полноты и точности [5]. Рис. 1 иллюстрирует основную идею данных показателей. Здесь регион, очерченный сплошной линией, – это мааркированные вручную контрольные данные, а пунктирная линия, в свою очередь, показывает результат работы алгоритма.

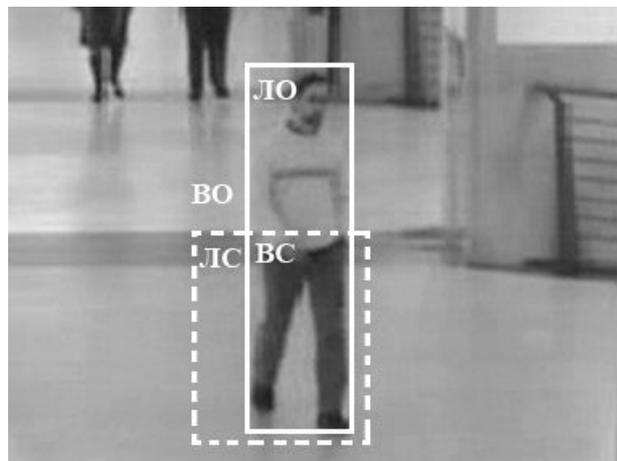


Рис. 1. Разметка изображений для расчета показателей точности и полноты

Полнота для кадра  $t$  определялась по следующей формуле:

$$\text{Полнота}(t) = \frac{|BC|}{|BC| + |ЛО|}, \quad (4)$$

где  $|BC|$  – количество пикселей, которые совпали при верном срабатывании алгоритма отслеживания, когда выделенный алгоритмом объект совпал с объектом в контрольных данных;  $|ЛО|$  – количество не совпавших пикселей при ложном отказе, ситуации, когда объект присутствует в кадре, а метод отслеживания либо не распознал его, либо выделил другую область в кадре;

Точность для кадра  $t$  определялась по следующей формуле:

$$\text{Точность}(t) = \frac{|BC|}{|BC| + |ЛС|}, \quad (5)$$

где  $|ЛС|$  – количество пикселей в случае, когда метод отслеживания выделил объект, которого нет согласно контрольным данным.

### Результаты моделирования

Авторами разработан прототип подсистемы отслеживания объектов, который позволяет определять траекторию движения человека и отслеживать его силуэт с использованием приведенных выше алгоритмов.

Для моделирования были взяты три типа наиболее характерных тестовых видеоролика из базы CAVIAR:

«EnterExitCrossingPaths1» – на котором два человека в разное время проходят перед двумя другими идущими рядом людьми (присутствует перекрытие объектов);

«OneLeaveShop1» – на котором четко видно, как один человек выходит из магазина в коридор, проходя мимо камеры (перекрытие объектов отсутствует);

«WalkByShop1» – на котором один человек появляется из-за нижнего края кадра, а затем постепенно удаляется от камеры.

Результаты отслеживания объектов для каждого тестового видеоролика (рис. 2, а), сравнивались с эталонными, т.е. маркированными вручную на контрольной выборке базы CAVIAR (рис. 2, б). Результаты сравнения использовались для вычисления показателей точности и полноты (4) и (5).



а



б

Рис. 2. Результаты моделирования (а) и эталонный образец (б)

На рисунках 3–5 для трех типов видеороликов приведены расчетные показатели эффективности – зависимости показателя полноты от показателя точности. Исследование проводилось для всех рассмотренных выше алгоритмов отслеживания: вычитания фона, метода сдвига среднего, метода оптического потока и метода активных контуров.

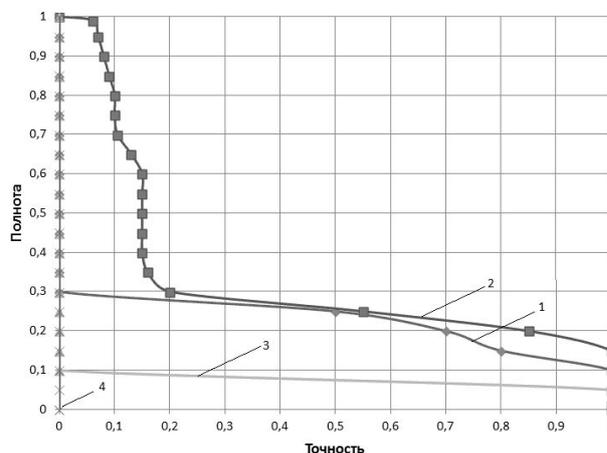


Рис. 3. Показатели эффективности для видеоролика «EnterExitCrossingPaths1»: 1 – вычитание фона, 2 – сдвиг среднего, 3 – оптический поток, 4 – активный контур

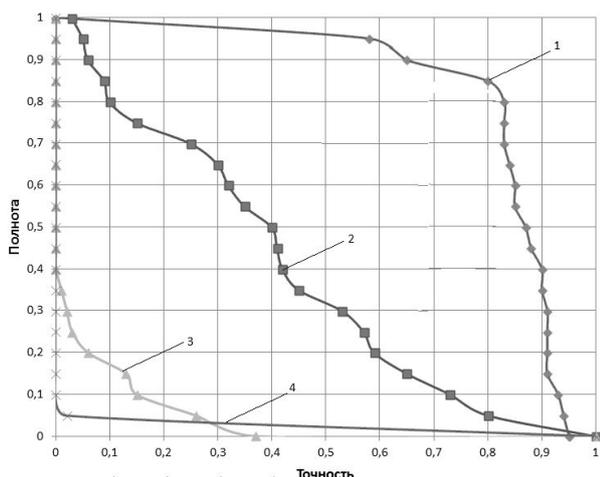


Рис. 4. показатели эффективности для видеоролика «OneLeaveShop1»: 1 – вычитание фона, 2 – сдвиг среднего, 3 – оптический поток, 4 – активный контур

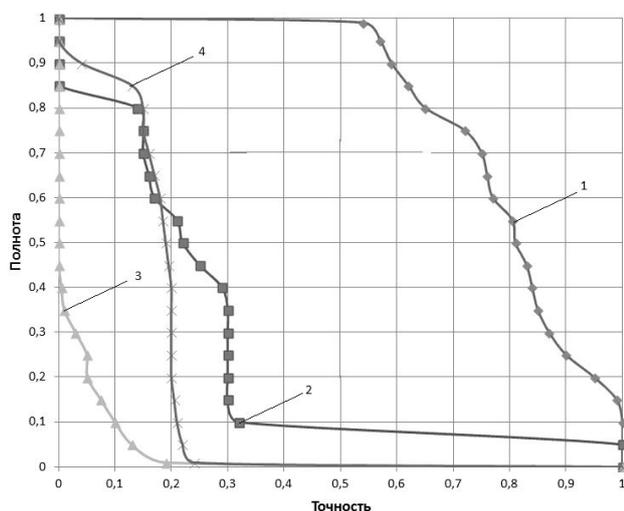


Рис. 5. Сравнение производительности для видеоролика «WalkByShop1»: 1 – вычитание фона, 2 – сдвиг среднего, 3 – оптический поток, 4 – активный контур

Из полученных результатов хорошо видно, что метод активных контуров показывал преимущественно близкие к нулевым показатели точности и полноты. Это особенно характерно, если размер эталонного объекта в кадре меньше, чем минимальный размер объекта, допустимый для метода (на всех текстовых видеороликах объекты были преимущественно небольшими и находились от камеры на достаточном отдалении).

Методы вычитания фона и сдвига среднего гораздо меньше зависят от разме-

ров объектов, и, как видно из рис. 3–5, показывают, в общем, лучшие результаты, чем метод оптического потока и метод активных контуров.

В результате проведенных экспериментов было установлено, что на тестовых клипах более предпочтительным является использование метода вычитания фона, который показал наилучшее соотношение между полнотой и точностью во втором и третьем тестовом клипе.

### Выводы

Были рассмотрены наиболее распространенные методы отслеживания движущихся объектов применительно к построению систем автоматизированного видеонаблюдения за людьми. Проведено моделирование этих методов с использованием тестовых видеороликов и контрольных данных, в результате которого было выяснено, что при построении систем автоматизированного наблюдения за людьми наиболее рациональным является использование метода вычитания фона.

Данный метод обладает высокой точностью детектирования движущихся объектов. Множество алгоритмов, реализующих данный метод, позволяет настроить систему под конкретные условия внешней среды. Данный метод имеет свои ограничения: необходимость стационарных камер наблюдения, отсутствие резких перепадов освещения.

### Список использованной литературы

1. Dick, A. Issues In Automated Visual Surveillance / A. Dick, M. Brooks // Proceedings of the VIIth Biennial Australian Pattern Recognition Society Conference – DICTA 2003. – 2003. – № 1. – P. 195–204.
2. Моделирование алгоритмов анализа аварийных ситуаций при видеонаблюдении за дорожным движением / С. Г. Антощук, А. Е. Колесников, А. А. Нутович, В. В. Прокопович // *Электротехнические и компьютерные системы*. – 2011. – № 2. – С. 76–81.
3. Javed, O. Tracking and Object Classification for Automated Surveillance / O. Javed, M. Shah // ECCV '02 Proceedings of

the 7th European Conference on Computer Vision. – 2002. – № 4. – P. 343–357.

4. CAVIAR. Context aware vision using image-based active recognition. – <http://homepages.inf.ed.ac.uk/rbf/CAVIAR> . – 2009.

5. Baumann, A. A Review and Comparison of Measures for Automatic Video Surveillance Systems / A. Baumann, M. Boltz, J. Ebling et al. // *EURASIP Journal on Image and Video Processing*. . – 2008. – Vol. 2008.

Получено 29.09.2012

#### References

1. Dick, A. Issues in Automated Visual Surveillance / A. M. Dick, M. Brooks // *Proceedings of the VIIth Biennial Australian Pattern Recognition Society Conference – DICTA 2003*. – 2003. – № 1. – P. 195–204 [in English].

2. Modeling of algorithms of accidents in traffic surveillance / S. G. Antoschuk, A. E. Kolesnikov, A. A. Nutovich, V. V. Prokopovich // *Electrotechnical and computer systems*. – 2011. – № 2. – P. 76–81 [in Russian].

3. Javed O. Tracking and Object Classification for Automated Surveillance / O. Javed, M. Shah // *ECCV '02 Proceedings of the 7th European Conference on Computer Vision*. – 2002. – № 4. – P. 343–357 [in English].

4. CAVIAR. Context aware vision using image-based active recognition. – <http://homepages.inf.ed.ac.uk/rbf/CAVIAR>. – 2009 [in English].

5. Baumann, A. Review and Comparison of Measures for Automatic Video Surveillance Systems / A. Baumann, M. Boltz, J. A. Ebling // *EURASIP Journal on Image and Video Processing* vol. – 2008 [in English].



Антошук  
Светлана Григорьевна,  
д.т.н., проф., зав. каф.  
Информационных  
систем Одесского нац.  
политехн. ун-та,  
+38048-7348-584,  
e-mail: asg@ics.opu.ua



Коваленко  
Никита Владиславович,  
аспирант Одесского нац.  
политехн. ун-та,  
м/т: +3(050)4951713



Годовиченко  
Николай Анатольевич,  
аспирант Одесского нац.  
поли-техн. ун-та,  
м/т: +3(068)2619923