

КРАТКИЕ СООБЩЕНИЯ

УДК 621.396.96

Р.Х. Садыхов, С.А. Кучук

АЛГОРИТМ УДАЛЕНИЯ ФОНА
В ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЯХ

Предлагается способ улучшения метода удаления фона Ж.К.С. Жака, К.Р. Юнга, С.Р. Муса в полутоновых изображениях видеопоследовательностей, позволяющий удалять фон для объектов различного размера и монотонных поверхностей.

Введение

Удаление фона – одна из составных задач анализа последовательности видеоизображений [1–8], которая заключается в определении наличия движения либо нахождения человека в тревожной зоне, подсчете количества людей и направлений их движения, выделении траекторий их движения (через ограду, вверх по пожарной лестнице дома), распознавании ситуаций неадекватного поведения людей (драки; упавшего, бегущего, остановившегося человека), выделении и идентификации лиц, детекции оставленных, удаленных или переброшенных предметов.

В статье предложен способ улучшения метода удаления фона Ж.К.С. Жака, К.Р. Юнга, С.Р. Муса [1]. Постановка задачи исходного алгоритма заключалась в удалении фона внутри и вне помещений для полутоновых изображений на стационарных камерах видеонаблюдения.

1. Описание метода

В методе Ж.К.С. Жака, К.Р. Юнга, С.Р. Муса используется модель фона, которая создается на основе 8–40-секундного фрагмента видеопоследовательности.

Модель фона содержит следующие три матрицы:

$$\begin{bmatrix} m(i, j) \\ n(i, j) \\ d(i, j) \end{bmatrix} = \begin{bmatrix} \min_z V^z(i, j) \\ \max_z V^z(i, j) \\ \max_z |V^z(i, j) - V^{z-1}(i, j)| \end{bmatrix}, \quad (1)$$

где z – номер кадра; $V^z(i, j)$ – интенсивность пиксела кадра z видеопоследовательности.

Кадры z , используемые для инициализации модели фона, должны удовлетворять выражению

$$|V^z(i, j) - \lambda(i, j)| \leq 2\sigma(i, j), \quad (2)$$

где $\lambda(i, j)$ – медианное значение интенсивностей пикселей по z кадрам; $\sigma(i, j)$ – дисперсия интенсивностей пикселей по z кадрам.

Пиксеты, принадлежащие фону кадра V видеопоследовательности, определяются по условию

$$m(i, j) - kd(i, j) \leq V(i, j) \leq n(i, j) + kd(i, j), \quad (3)$$

где k – настраиваемый коэффициент, который авторы исходного метода предлагают брать равным двум.

После удаления фона происходит устранение теней. Поскольку устранение теней не является предметом улучшения, описание этого процесса опущено.

На финальной стадии изображение обрабатывается маской 5x5 в форме ромба операциями морфологической обработки: сначала операцией закрытия, затем открытия.

2. Улучшение метода

Улучшение метода осуществляется путем использования коэффициента k при инициализации модели фона, изменения расчетов $\lambda(i, j)$ на вычисление среднего значения яркости и $d(i, j)$ для близких к нулю значений и нахождения зависимости диаметра маски ромба от размера искомого объектов.

Для отбора кадров инициализации модели фона в выражении (2) используется коэффициент k следующим образом:

$$|V^z(i, j) - \lambda(i, j)| \leq k \sigma(i, j). \quad (4)$$

Этот коэффициент – параметр алгоритма – подбирается экспериментальным путем для определенного класса задач, в дальнейшем не меняется. Авторы исходного метода предлагают брать его равным двум.

На экспериментальных данных при расчете $\lambda(i, j)$ применение среднего значения вместо медианного позволило более качественно удалять фон согласно выражению (4) за счет уменьшения числа ложноположительных результатов на ярких (снег, стекла) и однотонных (дорога) объектах, на видеоизображениях которых особенно заметны шумы.

Один из недостатков определения фона по условию (3) заключается в том, что при близких к нулю значениях $d(i, j)$ это выражение работает неэффективно. Поэтому предлагается в (3) вместо $kd(i, j)$ использовать функцию от $d(i, j)$, которая бы усиливала на порядок близкие к нулю значения $d(i, j)$:

$$m(i, j) - (c + kd(i, j)) \leq V(i, j) \leq n(i, j) + (c + kd(i, j)). \quad (5)$$

Наилучшие результаты детекции фона по выражению (5) достигаются при c от 1 до 3 (рис. 1).

Подобного результата можно добиться добавлением 2 %-го импульсного шума вида «соль и перец» к изображениям инициализации модели фона, при котором меняется интенсивность случайных пикселей изображения на максимальное или минимальное значение. Результат обусловлен тем, что после добавления шума уменьшается среднее значение интенсивности пиксела $\lambda(i, j)$, увеличивается дисперсия $\sigma(i, j)$, так как при ее расчете будут обработаны в том числе пикселы, которым присвоили крайние значения интенсивностей. Это приводит к достаточному увеличению $d(i, j)$, чтобы выражение (3) работало эффективно. Однако в сравнении с использованием формулы (5) добавление шума не позволяет ощутимо уменьшить количество кадров, требуемых для инициализации модели фона.

Эксперименты показали, что маска размером 5x5 не подходит для всех размеров искомого объектов на изображении. Наилучшие результаты обработки достигаются, когда размер маски-ромба зависит от наименьшей окружности движущихся фрагментов объекта в пикселах, за которыми просматривается фон. Поскольку на теле человека такими объектами являются конечности и шея, то предлагается определять размер маски-ромба следующим образом:

$$L_d = \max(3; \min(l_{\min}, n_{\min}) / 2), \quad (6)$$

где l_{\min} – диаметр окружности голени искомого объекта в пикселах; n_{\min} – диаметр окружности шеи искомого объекта в пикселах. Оба параметра, предполагается, будут настраиваться интеграторами при инсталляции системы охранного видеонаблюдения для каждой камеры вручную по двум кадрам, в которых вырезанный трафарет человека имеет минимальную и максимальную высоты в пикселах соответственно.

В случае, когда минимальный размер искомого объектов велик, предлагается уменьшать размер обрабатываемых данных.

Для расположенных фронтально или с небольшим наклоном относительно горизонтали камер высота объектов на кадре меняется незначительно, стало быть, выражение (6) можно упростить. Эксперименты показали, что диагональ маски ромба L_d составляет приблизительно

2,0–4,65 % (в среднем на 3,16 %) от высоты искомого объекта на кадре, что позволяет формулу (6) сделать зависимой от аргументов, получаемых более простым путем:

$$L_d = \max(3; 0,02 * h_{\min}), \quad (7)$$

где h_{\min} – минимальная высота искомого объекта в пикселах.

Таким образом, полученная зависимость диаметра маски-ромба от минимальной высоты человека (7) позволяет удалять шумы более эффективно, чем в случае, когда этот диаметр задан жестко.

3. Результаты обработки видеоизображений

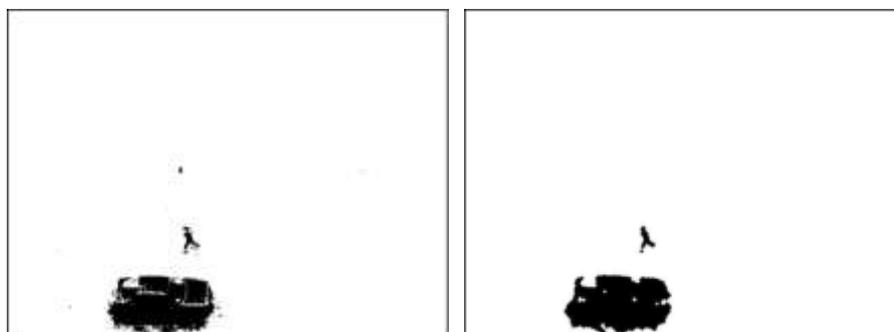
Система тестировалась на четырех видеоизображениях размером 640x480 пикселей. Для первого – третьего видеоизображений объем данных для инициализации фона составил 600 кадров, а для четвертого – 40 кадров. Первое и второе видеоизображения использовались для детекции объектов высотой 33 и 13 пикселей соответственно при $k = 3$ (рис. 1 и 2).



а)



б)

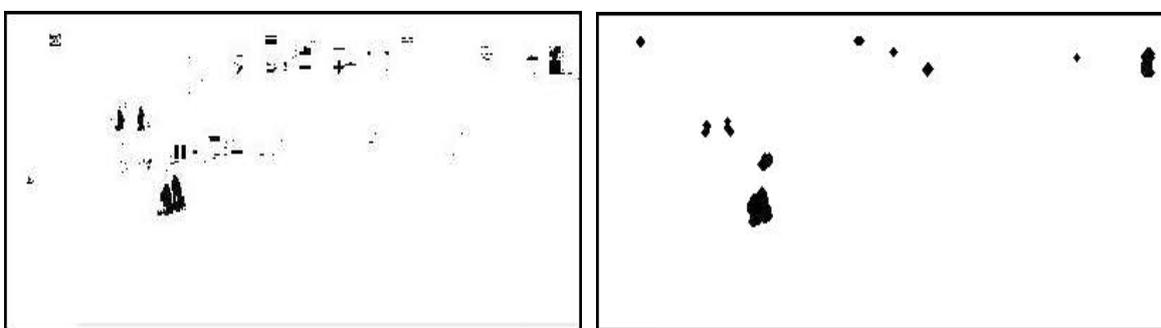


в)

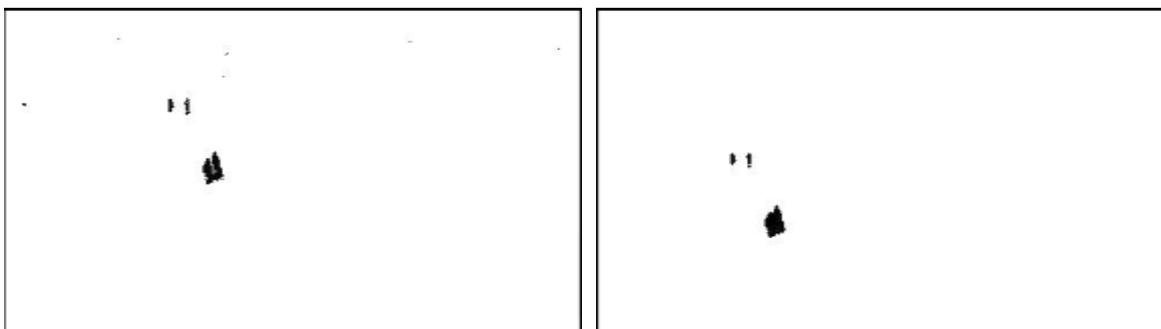
Рис. 1. Результаты обработки кадра из первой видеопоследовательности: а) исходное изображение; б) вычисленный фон до и после применения морфологических операций в исходном методе; в) вычисленный фон до и после применения морфологических операций в предлагаемом методе



а)



б)



в)

Рис. 2. Результаты обработки кадра из второй видеопоследовательности: а) исходное изображение; б) вычисленный фон до и после применения морфологических операций в исходном методе; в) вычисленный фон до и после применения морфологических операций в предлагаемом методе

На рис. 1 и 2 видно, что после проведенных улучшений метод намного эффективнее работает на объектах с вариациями сравнительно небольшой высоты. Рассмотрим, как он поведет себя при удалении фона внутри помещений, где размеры объектов велики. На третьей видеопоследовательности (i-LIDS AVSS 2007 Abandoned baggage detection, easy level) представлены результаты проверки работы алгоритма для поиска объектов высотой 160 пикселей при $k = 2$ (рис. 3).



а)



б)



в)

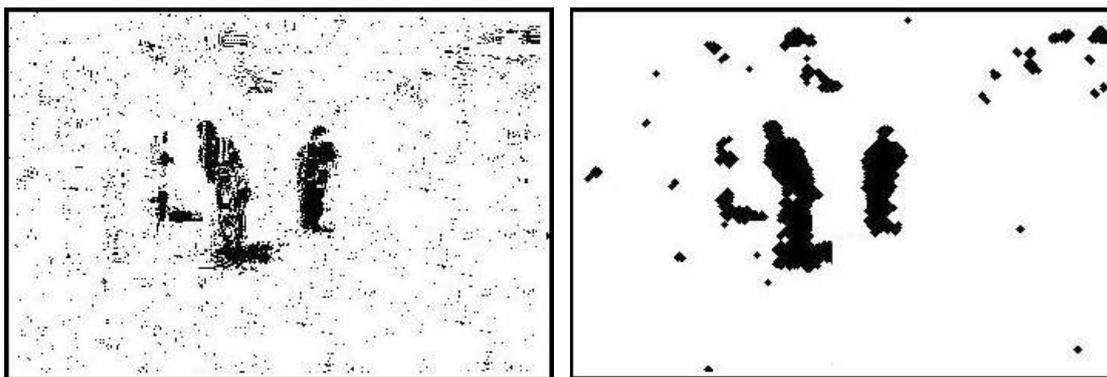
Рис. 3. Результаты обработки кадра из третьей видеопоследовательности: а) исходное изображение; б) вычисленный фон до и после применения морфологических операций в исходном методе; в) вычисленный фон до и после применения морфологических операций в предлагаемом методе

Рассмотрим, как проявят себя улучшения на четвертой видеопоследовательности (hall_monitor.mpg) с ограниченным количеством кадров (40 вместо 600) для настройки модели

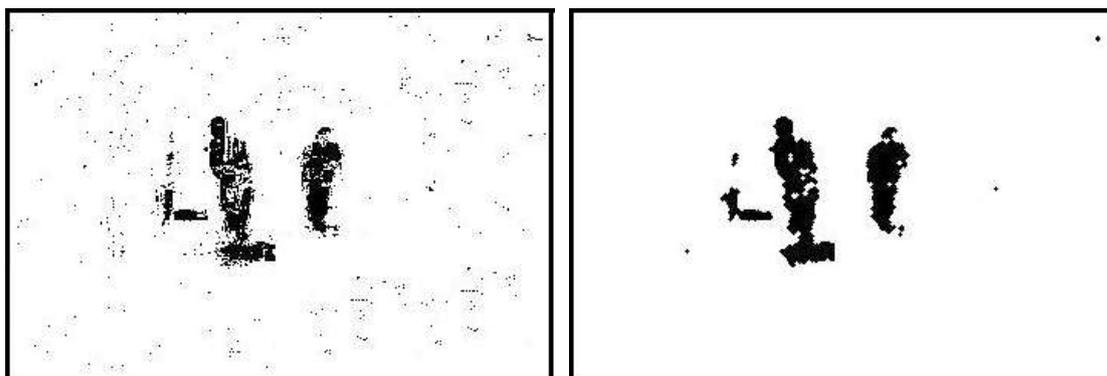
фона. Эта видеопоследовательность использовалась для поиска объектов шириной от 30 до 65 пикселей и высотой от 70 до 110 пикселей (рис. 4).



а)



б)



в)

Рис. 4. Результаты обработки кадра из четвертой видеопоследовательности: а) исходное изображение; б) вычисленный фон до и после применения морфологических операций в исходном методе; в) вычисленный фон до и после применения морфологических операций в предлагаемом методе

Как видно из рис. 1–4, предложенные способы улучшения метода позволили снизить количество ложноположительных результатов на различных тестовых данных. Кроме того, площади

ложноположительных результатов стали меньше, чем площади искомых объектов. Эксперименты показали, что наилучшие результаты достигаются при $k = 3$ для сцен на улице и $k = 2$ для сцен внутри помещений.

Заключение

Представленные изменения в методе Ж.К.С. Жака, К.Р. Юнга, С.Р. Муса повысили качество удаления фона: уменьшилось количество ложноположительных результатов (размеры таких результатов стали меньше размеров искомых объектов), улучшилось качество удаления фона на видеоизображениях ярких и однотонных поверхностей. Кроме того, появилась возможность работать с объектами разных габаритных размеров.

Список литературы

1. Jacques, J.C.S. Background Subtraction and Shadow Detection in Grayscale Video Sequences / J.C.S. Jacques, C.R. Jung, S.R. Musse // Computer Graphics and Image Proc. 18th Brazilian Symposium. – Brazilian, 2005. – P. 189–196.
2. Ермолаев, Е. Комплексные решения для контроля и управления высотными зданиями / Е. Ермолаев // Высотные здания. – 2008. – Июнь-июль – С. 118–121.
3. Ерошин, Е.В. Новые рынки для сетевого видеонаблюдения / Е.В. Ерошин // Каталог CCTV [Электронный ресурс]. – 2010. – Режим доступа : <http://ip-kamera.ru/articles.phtml?a=813>. – Дата доступа : 20.11.2011.
4. Ерошин, Е.В. Системы видеонаблюдения на транспорте в вопросах и ответах / Е.В. Ерошин, Д.А. Горнушенков, С.В. Кучумаров // Системы безопасности. – 2010. – № 3 (93). – С. 136–142.
5. Андрианов, Е.Ю. Системы видеонаблюдения с функциями видеоанализа для удаленных объектов / Е.Ю. Андрианов, С.Ю. Исправников, В.В. Старцев // Системы безопасности. – 2011. – № 2 (98). – С. 88–93.
6. Чижов, А.С. Видеоанализ в регионах / А.С. Чижов // Системы безопасности. – 2011. – № 3 (99). – С. 76–77.
7. Шумейко, М. Передовые технологии для систем видеонаблюдения / М. Шумейко // F+S: технологии безопасности и противопожарной защиты. – 2009. – № 4 (40). – С. 4–6.
8. Птицын, Н.В. Мегапиксельная видеоаналитика для сложных систем видеонаблюдения / Н.В. Птицын, В.Б. Булычева // Системы безопасности. – 2010. – № 4 (94). – С. 66–68.

Поступила 20.01.2012

*Белорусский государственный университет
информатики и радиоэлектроники,
Минск, П. Бровки, 6
e-mail: rsadykhov@bsuir.by,
Cuchuk.Sergey@gmail.com*

R.Kh. Sadykhov, S.A. Kuchuk

VIDEO SEQUENCES BACKGROUND REMOVAL ALGORHYTHM

An improvement of the method by J.C.S. Jacques, C.R. Jung and S.R. Musse for background removal and video sequences shadow detection is proposed.