

УДК 656.2-50:519.8

МОНИТОРИНГ ДВИЖУЩЕГОСЯ ОБЪЕКТА С ИСПОЛЬЗОВАНИЕМ АДАПТИВНОГО КОРРЕЛЯЦИОННОГО ФИЛЬТРА

А.М. ПОЛОНЕВИЧ, М.П. РЕВОТЮК

Закрытое акционерное общество «ЦНИП»
Смоленская, 15А, Минск, 220088, Беларусь

Белорусский государственный университет информатики и радиоэлектроники
П. Бровка, 6, Минск, 220013, Беларусь

Поступила в редакцию 27 октября 2015

Рассматривается задача локализации движущегося объекта по результатам видеонаблюдения в реальном времени на изменяющемся фоне в условиях полного или частичного перекрытия объекта другими объектами. Предлагается алгоритм локализации объектов на основе адаптивного корреляционного фильтра. Алгоритм не требует ввода дополнительных параметров или настройки в процессе работы.

Ключевые слова: автосопровождение объектов, адаптивный корреляционный фильтр.

Введение

Область применения алгоритмов локализации движущегося объекта в видеопотоке (или трекинга объекта) очень широка, начиная от слежения за игроком в игровых программах и заканчивая военными операциями с использованием автоматизированных оборонительных комплексов. На сегодняшний день существует немало методов трекинга и приемов повышения качества их работы, однако проблема остается актуальной и о революционном прорыве говорить нельзя.

Типичные сцены с целями для трекинга представлены на рис. 1. В подобных ситуациях невозможно применить вычитание фона и основанные на этом методы [1]. Камера, посредством которой ведется наблюдение за объектом, не всегда фиксирована стационарно, но может поворачиваться (не обязательно синхронно) в направлении к цели по результатам работы алгоритма локализации. Практически не удастся воспользоваться такими методами как SIFT (Scale-invariant Feature Transform) [2], SURF (Speeded Up Robust Features) [3] или RIFF (Rotation-Invariant Fast Features) [4], основанными на анализе особых характерных точек объекта. Объект, подлежащий локализации, может не иметь ярко выраженной текстуры и быть сильно размытым (рис. 1, в).

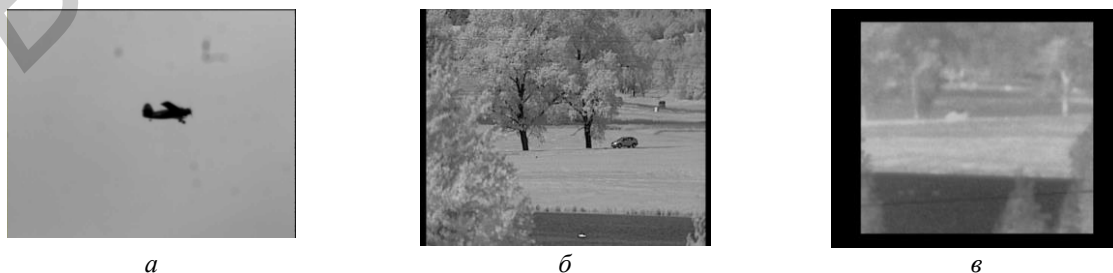


Рис. 1. Кадры видеоряда с движущимися объектами: *а* – самолет на фоне ясного неба; *б* – автомобиль на сложном, неоднородном фоне; *в* – изображение автомобиля через тепловизор

В работах [5, 6] для решения задачи локализации рассматривается метод на основе процедуры «сдвига среднего» (mean-shift). Алгоритм легко реализуется, опирается на анализ распределения цветовой гаммы изображения, но требует предварительной настройки и корректировки параметров в процессе трекинга.

В [7] описывается метод сравнения областей исходного кадра со скользящим по матрице шаблоном (template matching). Такой подход к решению задач мониторинга движущегося объекта зарекомендовал себя с лучшей стороны.

Идея поиска объекта на изображении путем корреляции с шаблоном не нова и является распространенной методикой. Взаимная корреляция часто применяется в обработке сигналов для получения вектора чисел, показывающих степень выраженности образца в сигнале. Классический пример задач обработки видео – распознавание лиц [8]. По новому позволяет взглянуть на проблему локализации движущегося объекта подход, предложенный в [9, 10]. Его основная идея – построение усредненного фильтра по набору изображений из обучающего множества, что существенно повышает стабильность выходного результата и позволяет продолжать трекинг даже при значительных вариациях обрабатываемых данных.

Методика локализации объекта

Взаимнокорреляционная функция – стандартный метод определения соответствия двух последовательностей. Для одномерного случая, дискретная взаимная корреляция функций $h[x]$ и $f[x]$ определяется следующим образом:

$$(h \otimes f)[x] = \sum_{dx=-\infty}^{\infty} h[dx]f[x + dx]. \quad (1)$$

Максимум функции (1) находится в той области, где сигналы совпадают больше всего. В рассматриваемом случае задач мониторинга выбранного на кадре объекта представим содержимое кадра и изображение цели в виде двух дискретных функций яркости $h[x, y]$ и $f[x, y]$, а затем определим двумерную взаимнокорреляционную функцию

$$(h \otimes f)[x, y] = \sum_{dx=-\infty}^{\infty} \sum_{dy=-\infty}^{\infty} h[dx, dy]f[x + dx, y + dy]. \quad (2)$$

Использование функции (2) дает хороший результат для локализации априорно заданного объекта. В частности, упомянутый метод template matching, реализованный в открытой библиотеке компьютерного зрения OpenCV [11], опирается на взаимнокорреляционную функцию. Результат работы алгоритма локализации представлен на рис. 2.

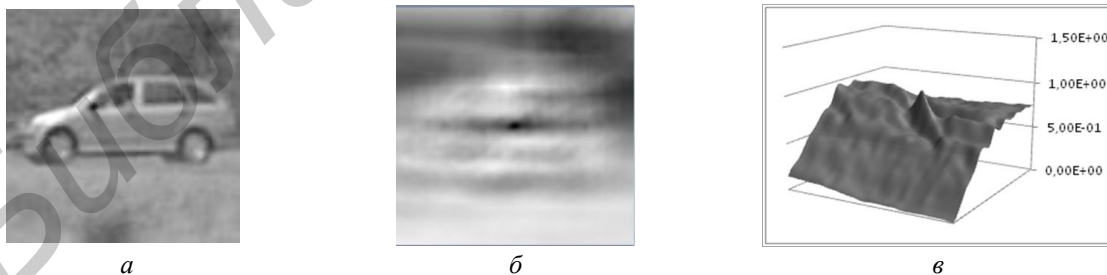


Рис. 2. Локализация объекта на основе построения взаимнокорреляционной функции: *a* – фрагмент исходного кадра, содержащего цель локализации; *б* – результат поиска по шаблону с использованием взаимнокорреляционной функции; *в* – характерный пик, определяющий положение объекта

Корреляционный фильтр в контексте обсуждаемой задачи – искусственно создаваемый шаблон, корреляция которого с исходным кадром позволит получить заранее заданную желаемую реакцию при обнаружении искомого объекта. Под желаемой реакцией будем понимать некую легко узнаваемую функцию, которая в общем случае может быть любой. Удобно воспользоваться двумерной функцией Гаусса:

$$g_i(x, y) = e^{-\frac{(x-x_i)^2+(y-y_i)^2}{\delta^2}}.$$

Таким образом, по определению, фильтр $h[x, y]$ и дискретная функция $f[x, y]$ яркости точек эталонного изображения связаны выражением $g=h \circ f$.

Для нахождения h воспользуемся тем, что по аналогии с теоремой свертки взаимная корреляция удовлетворяет $\mathcal{F}[h \circ f] = (\mathcal{F}[h])^* \cdot (\mathcal{F}[f])$. Обозначив функции, претерпевшие преобразование Фурье, соответствующими заглавными буквами, получаем

$$G = H^* \cdot F. \quad (3)$$

Использование (3) вместе с алгоритмами быстрого преобразования Фурье позволяет эффективно вычислять величины взаимной корреляции, а также легко выразить h в домене Фурье:

$$H^* = \frac{G}{F}. \quad (4)$$

Определенный в (4) фильтр является частным корреляционным фильтром, однозначно сопоставляющим идеальную реакцию на выходе эталонному сигналу на входе алгоритма. Трудности его практического использования заключаются в том, что при отклонениях исследуемого сигнала от эталонного реакция системы становится трудно узнаваемой.

Адаптивный корреляционный фильтр, синтезированный на базе множества частных, позволяет многократно повысить вероятность локализации движущегося объекта в видеопотоке и лишен недостатков, присущих точному корреляционному фильтру. Немного изменяя эталонный сигнал (например, фотографируя автомобиль под разными углами), вычислим N точных корреляционных фильтров H_i^* , на основе которых определим

$$H_{\mu}^* = \frac{1}{N} \sum_{i=1}^N H_i^*. \quad (5)$$

Справедливость выражения (5) обоснована в [12], где каждый из частных фильтров рассматривается в качестве слабого классификатора, а способность фильтра к обнаружению целевого сигнала повышается усреднением большого количества классификаторов. Результат работы алгоритма локализации с использованием адаптивного корреляционного фильтра показан на рис. 3.

Недостаток известного адаптивного фильтра ASEF (Average of Synthetic Exact Filters), определяемого (5) – существенные временные затраты на его первоначальную тренировку. Для сокращения времени первоначальной настройки в [13] формулируется задача минимизации суммы квадратов ошибки между полученным и желаемым результатами на обучающем множестве сигналов.

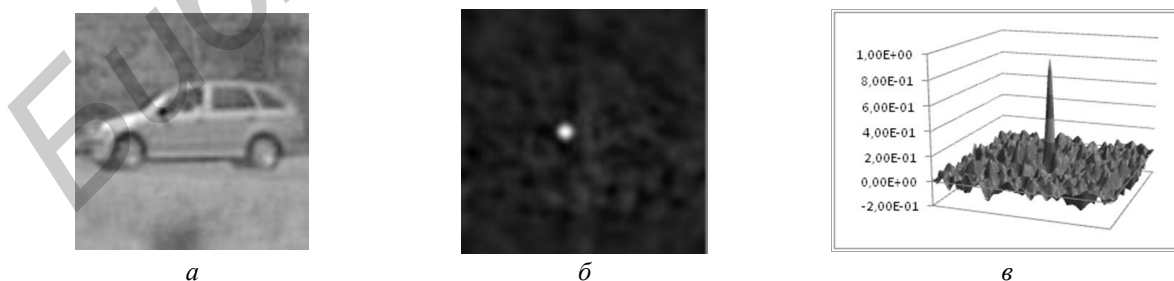


Рис. 3. Локализация объекта на основе использования адаптивного корреляционного фильтра (АКФ): a – фрагмент исходного кадра, содержащего цель локализации; b – результат локализации на основе использования АКФ; c – характерный пик, определяющий положение объекта

Оптимизация процесса адаптации корреляционного фильтра

Идея минимизации суммы квадратов ошибки не нова. Формально, для рассматриваемого случая, постановка задачи будет выглядеть так:

$$H = \min_H \sum_i |F_i H^* - G_i|^2. \quad (6)$$

Принимая во внимание тот факт, что операции производятся в домене Фурье и каждый элемент фильтра H может быть оптимизирован независимо,

$$H_{wv} = \min_{H_{wv}} \sum_i |F_{iwv} H_{wv}^* - G_{iwv}|^2. \quad (7)$$

Задача (6) практически совпадает с описанной в [14] и [15] с той лишь разницей, что в [14, 15] цель точно центрована и, следовательно, для всего обучающего множества выходной результат практически одинаковый.

Процедура нахождения минимума (7) почти не отличается от стандартной. Приравняем частную производную по H_{wv}^* к нулю, считая H независимой переменной:

$$\frac{d}{dH_{wv}^*} \sum_i |F_{iwv} H_{wv}^* - G_{iwv}|^2 = 0. \quad (8)$$

В результате стандартных преобразований определяется окончательный вид фильтра:

$$\begin{aligned} \frac{d}{dH_{wv}^*} \sum_i (F_{iwv} H_{wv}^* - G_{iwv})(F_{iwv} H_{wv}^* - G_{iwv})^* &= 0, \\ \frac{d}{dH_{wv}^*} \sum_i [(F_{iwv} H_{wv}^*)(F_{iwv} H_{wv}^*)^* - (F_{iwv} H_{wv}^*)G_{iwv}^* - G_{iwv} (F_{iwv} H_{wv}^*)^* + G_{iwv} G_{iwv}^*] &= 0, \\ \frac{d}{dH_{wv}^*} \sum_i F_{iwv} F_{iwv}^* H_{wv} H_{wv}^* - F_{iwv} G_{iwv}^* H_{wv}^* - F_{iwv}^* G_{iwv} H_{wv} + G_{iwv} G_{iwv}^* &= 0, \\ \sum_i [F_{iwv} F_{iwv}^* H_{wv} - F_{iwv} G_{iwv}^*] &= 0, \\ H_{wv} &= \frac{\sum_i F_{iwv} G_{iwv}^*}{\sum_i F_{iwv} F_{iwv}^*}, \\ H &= \frac{\sum_i F_i \times G_i^*}{\sum_i F_i \times F_i^*}. \end{aligned} \quad (9)$$

Фильтр (9) также известен, как MOSSE (Minimizing the Output Sum of Squared Error) фильтр [13]. Такой фильтр, подобно фильтру ASEF (Average of Synthetic Exact Filter), не требует большого количества эталонных сигналов для предварительной адаптации.

Результаты и их обсуждение

Рассмотренный фильтр является основным элементом системы мониторинга, формирующим данные о текущем положении наблюдаемого объекта.

Достижение высокого качества выходного результата – ключевая подзадача трекинга. Под «качеством» подразумевается отсутствие в результате ложных пиков и большая разница между значениями пика и фонового шума. В ситуации, когда срыв цели распознается поздно или ложно, значения, отправляемые на вход предиктивного фильтра, способствуют его зашумлению и препятствуют продолжению корректного функционирования. К сожалению, избавиться от срывов и ложных пиков в реальных условиях не всегда представляется возможным. Однако существенная разница между результатами анализа случайных выходных матриц алгоритма адаптивного корреляционного фильтра с алгоритмом метода template matching демонстрирует значительное превосходство первого над вторым (рис. 4).

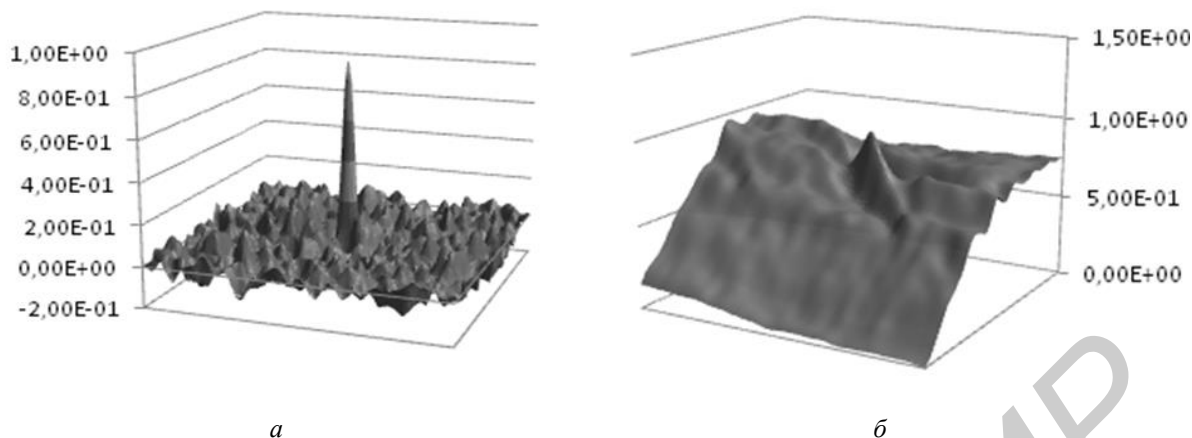


Рис. 4. Визуализация результирующих матриц: *a* – применение адаптивного корреляционного фильтра; *б* – применение метода template matching

Нетрудно заметить, что пик, полученный в результате применения метода адаптивного корреляционного фильтра, гораздо четче. Это обстоятельство, в свою очередь, позволяет использовать отношение значения пика к боковым лепесткам PSR (Peak to Sidelobe Ratio) для определения достоверности результата работы адаптивного корреляционного фильтра и свидетельствует о высоком качестве трекинга в целом.

Заключение

Предложен алгоритм локализации единичного движущегося объекта на основе адаптивного корреляционного фильтра. Представлены математические модели применения фильтра в задаче мониторинга позиции объекта. Предложено решение задачи оптимизации временных затрат на процесс предварительной адаптации корреляционного фильтра. Метод может применяться в стационарной или динамической сцене вне зависимости от типа освещения. Его реализация проведена средствами языка программирования C++ с использованием компонент библиотеки Open CV для создания системы автоматического сопровождения выделяемого движущегося объекта в реальном времени.

Экспериментальные результаты работы предлагаемого алгоритма подтверждают его эффективность и надежность в ситуациях с изменением фона, постепенным поворотом и изменением размера наблюдаемого объекта, его кратковременным пропаданием или полным/частичным перекрытием другими объектами.

REAL TIME OBJECT MONITORING USING ADAPTIVE CORRELATION FILTER

A.M. POLONEVITCH, M.P. REVOTJUK

Abstract

Real time localization of a moving or stationary object in a video stream is one of the most complex tasks in video processing due to changing background conditions or full / partial object overlaps is considered. For present moment, a number of techniques and solution proposed, nevertheless the task is still actual and no one can speak about revolutionary breakthrough. Present article describes real-time object tracking algorithm based on adaptive correlation filter usage. The algorithm is autonomous and doesn't need parameters adjustments from the beginning till the end of its functioning.

Список литературы

1. *Kaufmann M.* Machine Vision: Theory, Algorithms, Practicalities. San Francisco, 2004.
2. *David G. Lowe* // Proceedings of the International Conference on Computer Vision. Kerkyra, 20–27 September 1999. P. 1150–1157.
3. *Herbert B., Ess A., Tuytelaars T. et. al* // Computer Vision and Image Understanding (CVIU). 2008. Vol. 110, № 3. P. 346–359.
4. Unified Real-Time Tracking and Recognition with Rotation-Invariant Fast Features. [Электронный ресурс]. – Режим доступа: <http://vijaychan.github.io/Publications/2010 - Unified Real-Time Tracking and Recognition with Rotation-Invariant Fast Features.pdf>. – Дата доступа: 23.10.2015.
5. *Cheng Yizong* // IEEE Transactions on Pattern Analysis and Machine Intelligence (IEEE). 1995. P. 790–799.
6. *Comaniciu D., Meer P.* // IEEE Transactions on Pattern Analysis and Machine Intelligence (IEEE). 2002, P. 603–619.
7. *Brunelli R.* Template Matching Techniques in Computer Vision: Theory and Practice. Hoboken, 2009.
8. *Duda R., Hart P.* Template Matching. Hoboken, 1973.
9. Average of Synthetic Exact Filters. [Электронный ресурс]. – Режим доступа: <http://www.cs.colostate.edu/~vision/publications/Bolme2009Asef.pdf>. – Дата доступа: 23.10.2015.
10. Introduction to Biometric Recognition Technologies and Applications. [Электронный ресурс]. – Режим доступа: http://users.ece.cmu.edu/~jzhu/class/18200/F05/Lecture06_Marios_Lecture.pdf. – Дата доступа: 23.10.2015.
11. *Bradski G., Kaehler A.* Learning OpenCV Computer Vision with the OpenCV Library. Sebastopol, 2008.
12. *Breiman L.* Bagging Predictors. Machine Learning. New Mexico, 1996.
13. *Bolme D.* Theory and Applications of Optimized Correlation Output Filters: Дис. ... канд. физ-мат. наук. Колорадо, 2011.
14. *Kumar B., Mahalanobis A., Song S., et. al.* // Optical Engineering. 1992. Vol. 31. P. 915–922.
15. *Mahalanobis A., Kumar B., Song S., et. al.* // Applied Optics, 1994. Vol. 33. P. 3751–3759.