

1 СЕКЦИЯ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

УДК 004.89, 004.93

ИСПОЛЬЗОВАНИЕ СТРУКТУРНОГО ОПИСАНИЯ ИЗОБРАЖЕНИЙ ДЛЯ ДЕТЕКТИРОВАНИЯ И ТРЕКИНГА ДВИЖУЩИХСЯ ОБЪЕКТОВ

А.В. Агарков

Институт проблем искусственного интеллекта МОН Украины и НАН Украины,
Киев, Украина

В данной работе предложено решение задачи детектирования и трекинга объектов в видеоряде с неподвижным фоном на основе анализа движения в кадре и представления изображения и объектов как множества структурных элементов. Описание объектов имеет двухуровневую иерархию, что позволяет гибко адаптировать его при изменении объекта. Аддитивное описание объекта в виде множества структурных элементов позволяет осуществлять трекинг в условиях частичного перекрытия.

Введение

Трекинг (отслеживание) объектов в видеопотоке — определение их расположения на каждом кадре в видеопотоке. Для систем реального времени необходимо локализовать объект, имея информацию о расположении на предыдущем кадре. Трекинг нашел применение в системах видеонаблюдения для автоматизации распознавания тревожных ситуаций, в системах видеоаналитики для управления бизнес-процессами, в системах распознавания целевых объектов и т. д. К настоящему времени разработано множество методов и подходов для решения данной задачи. Актуальность решения задачи трекинга не падает — с развитием методов распознавания зрительных образов для её решения применяются все более новые методики.

В данной работе предложено решение задачи выделения и отслеживания объектов в видеоряде с неподвижным фоном на основе анализа движения в кадре и представления изображения и объектов как множества структурных элементов. Описание объектов имеет двухуровневую иерархию, что позволяет гибко адаптировать его при изменении объекта. Аддитивное описание-дескриптор объекта в виде множества структурных элементов позволяет осуществлять трекинг в условиях частичного перекрытия.

1. Структурные элементы изображений

Изображение I рассматривается как множество структурных элементов p - $I = \{p_i\}$. Видеопоследовательность - это последовательность статических изображений. Два последовательных изображения видеоряда отличаются слабо вследствие небольшой разности во времени, расположения камеры и объектов при их получении. Соответственно, видеопоследовательность V - это множество структурных элементов -

$$V = \{I_j\} = \bigcup_j \{p_i | p_i \in I_j\},$$

где j - номер кадра. Для каждого p_i заданы их наличие $b(p_i, j)$ и положение $\bar{r}(p_i, j)$ для каждого кадра I_j видеопоследовательности.

Отдельным объектам соответствуют подмножества отдельных структурных элементов (ОСЭ). При движении объекта в кадре их взаимное расположение меняется слабо, по крайней мере для близко расположенных ОСЭ. Как следствие — для соседних ОСЭ одного объекта межкадровое смещение

$$\bar{s}(p_i, j) = \bar{r}(p_i, j) - \bar{r}(p_i, j-1)$$

будет примерно совпадать. Таким образом, отдельным объектам в кадре соответствуют подмножества структурных элементов, для каждого из которых выполняется условие совпадения межкадрового смещения для соседей

$$\bar{s}(p_i, j) \approx \bar{s}(p_l, j) : (p_i, p_l) \in N_j, \quad (1)$$

где N_j - множество всех пар ОСЭ, которые являются соседними друг другу на j -м кадре. Разумеется, если два объекта в кадре находятся рядом и одинаково смещаются, то они будут отслеживаться как один объект до тех пор, пока они не разойдутся. Однако при сближении разных объектов их слияние происходить не должно.

2. Сегментация движения

При движении объекта в кадре все его части смещаются согласованно. Даже если отдельные его части описывают разные траектории (например, при ходьбе у человека руки и ноги двигаются относительно туловища), близко расположенные точки смещаются от кадра к кадру примерно одинаково, иначе нарушается целостность объекта. Таким образом, отдельному объекту на изображении соответствует подмножество структурных элементов изображения, для каждой пары соседей из которого выполняется условие подобия межкадрового смещения.

Из этого следует, что для выделения отдельного объекта на основе использования множества структурных элементов и их межкадрового смещения необходимо выделить подмножества ОСЭ, согласно условию (1). Данные подмножества не являются классами эквивалентности, поскольку условие транзитивности для подобия соседних ОСЭ в общем случае не выполняется. Эта процедура не является достаточной для выделения отдельных объектов, поскольку возможен случай, когда разные объекты при движении находятся рядом. Для их разделения необходимо применение дополнительных методов из области детектирования и распознавания объектов на изображении.

В данной работе для поиска данных подмножеств ОСЭ используется метод степного пожара. В простейшем случае межкадровое смещение двух соседних структурных элементов считается подобным, если модуль их разности меньше порога h

$$|\bar{s}(p_i, j) - \bar{s}(p_l, j)| < h : (p_i, p_l) \in N_j. \quad (2)$$

Однако результаты выделения движущихся объектов при использовании данного условия являются неудовлетворительными — при медленном движении объектов они сливаются с фоном.

Для более качественного результата выделения движущихся объектов условие (1) было модифицировано. В данной работе для определения подобия межкадрового смещения для двух соседних ОСЭ используется адаптивный по-

рог, который вычисляется отдельно для каждого ОСЭ. Для каждого ОСЭ порог формируется на основе среднего отклонения по соседям, среднеквадратических отклонений модуля межкадрового смещения и его значений x и y . Межкадровые смещения двух соседних ОСЭ считаются подобными, если выполняется условие

$$f(i, j) \vee f_x(i, j) \vee f_y(i, j) \vee f(j, i) \vee f_x(j, i) \vee f_y(j, i),$$

где

$$\begin{aligned} f(i, j) &= |\bar{s}(p_i, j) - \langle \bar{s} \rangle(p_i, j)| < \langle |\bar{s}(p_i, j) - \langle \bar{s} \rangle(p_i, j)| \rangle : (p_i, p_l) \in N_j, \\ f_x(i, j) &= |x_{\bar{s}}(p_i, j) - x_{\langle \bar{s} \rangle}(p_i, j)| < \langle |x_{\bar{s}}(p_i, j) - x_{\langle \bar{s} \rangle}(p_i, j)| \rangle : (p_i, p_l) \in N_j, \\ f_y(i, j) &= |y_{\bar{s}}(p_i, j) - y_{\langle \bar{s} \rangle}(p_i, j)| < \langle |y_{\bar{s}}(p_i, j) - y_{\langle \bar{s} \rangle}(p_i, j)| \rangle : (p_i, p_l) \in N_j, \end{aligned}$$

$\bar{s}(p, j)$ - межкадровое смещение для структурного элемента p на j -м кадре, $\langle \bar{s} \rangle(p, j)$ - средний по соседним вектор смещения структурного элемента p на j -м кадре, $x_{\bar{s}}$, $y_{\bar{s}}$ - компоненты вектора \bar{s} .

3. Формирование описания объектов трекинга

Описание объекта строится на основе использования множества структурных элементов, координаты которых от кадра к кадру изменяются. От кадра к кадру также изменяется множество ОСЭ, которое соответствует отдельному объекту, вследствие удаления и рождения части ОСЭ.

Несмотря на то, что соседние точки объекта смещаются похоже, характер движения его разных частей может отличаться. Это приводит к разбиению множества ОСЭ отдельного объекта на подмножества, каждому из которых соответствует часть (сегмент) объекта. Характер движения каждого из данных сегментов отличается от соседних. Таким образом, описание объекта представляет собой множество сегментов, каждый из которых представлен множеством ОСЭ.

$$O_k(j) = \{S_{k,l}^j\} = \bigcup_l \{p : p \in S_{k,l}^j\},$$

где $O_k(j)$ - описание k -го объекта на j -м кадре, $S_{k,l}^j$ - сегмент, принадлежащий $O_k(j)$, p - структурный объект.

Несмотря, на разный характер движения, сегменты одного объекта от кадра к кадру остаются соседями. Данное постоянство позволяет выделять сегменты, относящиеся к одному объекту, как связанное подмножество множества ОСЭ.

Заключение

В данной работе предложено решение задачи выделения и отслеживания объектов в видеоряде с неподвижным фоном на основе анализа движения в кадре и представления изображения и объектов как множества структурных элементов. Для этого видеоряд рассматривается как множество отдельных структурных элементов, для каждого из которых определены положение и наличие на каждом кадре.

Разработана методика детектирования объектов на основе анализа межкадрового смещения отдельных структурных элементов. Данная методика позволяет сегментировать изображение видеоряда на области с различным характером движения без использования информации о фоне.

Разработана методика формирования иерархического описания объектов. Данная иерархия имеет двухуровневую структуру - 1) объект описывается как множество сегментов, 2) сегмент описывается как множество отдельных структурных элементов. Таким образом, трекинг объектов при используемом подходе представляет собой выделение связных областей из структурных единиц (сегментов и ОСЭ).

Экспериментальная оценка качества предложенного метода трекинга объектов показала его устойчивость к изменению объекта, стабильное удержание подвижных и неподвижных объектов. При перекрытии объектов наблюдается неустойчивая работа - слияние объектов, потеря одного из объектов, что, видимо, связано с несовершенством процедуры сегментации по движению.

В данной работе предполагается, что фон является неподвижным, но данный подход позволяет отслеживать изменения фона, выделяя его как наибольший сегмент (при условии, что площадь объектов меньше, чем фон, что выполняется в подавляющем большинстве случаев).

Список литературы

1. Fernando de la Torre Frade. Moving object detection and tracking system: a real-time implementation / Fernando de la Torre Frade, Elisa Martínez Marroquín, Eugenia Santamaría Pérez, Jose Antonio Morán Moreno // *Seizième colloque gretsi* — 15-19 septembre, 1997. — P. 375-378.
2. Yilmaz A. Object Tracking: A Survey / Alper Yilmaz, Omar Javed, Mubarak Shah // *ACM Computing Surveys*. – 2006. – Vol. 38. – No. 4. – Article 13. – P.1-45.
3. Vasilis Papadourakis. Multiple objects tracking in the presence of long-term occlusions / Vasilis Papadourakis, Antonis Argyros // *Computer Vision and Image Understanding*. — 2010 — vol.114. – P. 835–846.
4. Andres Alarcon Ramirez and Mohamed Chouikha. A New Algorithm for Tracking Objects in Videos of Cluttered Scenes / Andres Alarcon Ramirez, Mohamed Chouikha // *International Journal of Information Technology, Modeling and Computing (IJITMC)* – May, 2013. - vol.1, No.2, – P.72-83
5. Robert T. Collins. On-Line Selection of Discriminative Tracking Features / Robert T. Collins, Yanxi Liu // *Pattern Analysis and Machine Intelligence*. – 2005. – vol. 27, issue 10. – P. 1631 – 1643.
6. D. S. Bolme. Visual Object Tracking using Adaptive Correlation Filters / D. S. Bolme, J. R. Beveridge, B. A. Draper, Y. M. Lui // *Computer Vision and Pattern Recognition*. — June, 2010. — P. 2544 – 2550.
7. Martin Godec. Hough-based Tracking of Non-rigid Objects / Martin Godec, Peter M. Roth, Horst Bischof // *In Proc. International Conference on Computer Vision (ICCV)*. — 2011. – P. 81 – 88.
8. Zdenek Kalal. Online Learning of Robust Object Detectors During Unstable Tracking / Zdenek Kalal, Jiri Matas, Krystian Mikolajczyk // *In International Conference on Computer Vision*. – 2009.
9. Kalal Z. P-N learning: Bootstrapping binary classifiers by structural constraints / Kalal Z., Matas J., Mikolajczyk K. // *Computer Vision and Pattern Recognition (CVPR)*. – 2010. – P. 49 – 56.
10. Boris Babenko. Robust Object Tracking with Online Multiple Instance Learning / Boris Babenko, Ming-Hsuan Yang, Serge Belongie // *Pattern Analysis and Machine Intelligence*. – 2010. – vol. 33, issue 8. – P. 1619 – 1632.
11. David A. Ross. Incremental Learning for Robust Visual Tracking / David A. Ross, Jongwoo Lim, Ruei-Sung Lin, Ming-Hsuan Yang // *International Journal of Computer Vision*. – 2008. – vol. 77, issue 1-3. – P. 125-141.