

Алгоритм построения траектории движения объектов в видеопотоке на основе оптического потока

В.А. Франц, В.В. Воронин, В.И. Марчук, А.В. Фисунов, М.М. Письменскова

Системы компьютерного зрения находят применение во многих областях науки и техники. В том числе: в системах безопасности, робототехнике, системах контроля качества продукции, медицине, системах связи, системах взаимодействия человек-компьютер. Часто для решения задачи не достаточно обработки статических изображений и требуется работа с видеоинформацией. Это влечет за собой высокие требования к производительности используемой вычислительной системы. В то же время, во многих случаях требуется работа в реальном времени (например, в случае контроля состояния потенциально-опасных промышленных объектов или управления мобильным роботом). В связи с этим актуальна задача разработки вычислительно-эффективных алгоритмов обработки видеоинформации [1].

С точки зрения уменьшения объема вычислений перспективным является подход, при котором вычислительно-сложная операция (например, детектирование объекта) выполняется как можно реже, а для поиска положения объекта используются алгоритмы слежения. Таким образом, детектирование объекта выполняется реже (например, только для первого кадра видеопоследовательности), а положение объекта на каждом последующем кадре определяется на основе информации о прежнем положении объекта и информации текущего кадра. Как правило, данный подход приводит к существенному снижению общих вычислительных затрат [2].

Таким образом важной частью систем компьютерного зрения является трекер – алгоритм отслеживания положения объекта на видеопоследовательности с течением времени.

Целью данной работы является разработка эффективного, с точки зрения вычислительной сложности, алгоритма отслеживания объектов на видеопоследовательности.

Задача трекера состоит в оценке траектории движения объекта. Простейшим способом отслеживания положения объекта на последовательности кадров является использование поиска по шаблону. В простейшем случае в качестве шаблона выступает область изображения, занимаемая объектом на текущем кадре [3]. Поиск положения объекта на последующем кадре выполняется в скользящем окне, и использованием некоторой меры схожести, например евклидоваго расстояния.

Более эффективными являются алгоритмы, принадлежащие к семейству алгоритмов слежения, основанных на использовании ядра [4]. Здесь используется итерационная процедура, максимизирующая некоторую меру схожести (например, коэффициент Бхаттачарая).

Как показал анализ литературных источников, существующие методы сопровождения объектов обладают рядом недостатков. Одним из важных недостатков является невысокая точность определения положения объекта [5].

На вход трекера поступает видеопоследовательность и некоторое начальное положение объекта или объектов, которое может быть оценено автоматически либо указано пользователем. Трекер позволяет в автоматическом режиме получить оценку положения объекта на всех последующих кадрах (при условии, что он на них присутствует) [6].

Методы трекинга можно разделить на два класса:

- Одногипотезные методы – позволяют отслеживать только одну возможную траекторию;
- Многогипотезные методы – позволяют отслеживать несколько траекторий, и выбирать из них лучшую.

Положение объекта может быть описано различными способами. Будем использовать описание положения объекта в виде минимального

ограничивающего прямоугольника (bounding box). Из всей информации об объекте нас будет интересовать только набор значений (x, y, w, h) .

Здесь: x -координата левого-верхнего угла целевого объекта; y -координата левого-верхнего угла целевого объекта; w – ширина целевого объекта; h – высота целевого объекта. Пример прямоугольной области, соответствующей отслеживаемому объекту представлен на рис. 1.



Рис. 1. – Пример отслеживаемого объекта

Здесь отслеживаемый объект заключен в прямоугольную рамку и четко выделяется на окружающем фоне. Но, к сожалению, задача отслеживания положения объекта может оказаться сложной в силу следующих причин:

- Нестационарный фон: подвижными могут быть не только объекты переднего плана, но так же и элементы заднего плана. Простейшим примером является движение деревьев на ветру на заднем плане сцены;

- Медленное движение объектов переднего плана. Если объект переднего плана движется медленно или неподвижен, он не может быть отличен от фона;

- Непостоянство условий освещения: освещение сцены часто меняется в зависимости от времени суток и погоды;

- Движение объектов сцены не может быть предсказано заранее. Кроме того, положение камеры относительно сцены может меняться от кадра к кадру.

Таким образом, при разработке конкретного подхода к слежению за объектом на видеопоследовательности необходимо сделать ряд предположений:

- Положение объекта от кадра к кадру меняется незначительно;

– Возможно движением камеры, но только параллельно плоскости сенсора;

– Движущиеся объекты изменяются в размере не значительно;

К однопотенциальным относятся алгоритм сдвига среднего mean shift [7].

Структурная схема алгоритма сопровождения объекта на базе mean shift представлена на рис. 2.

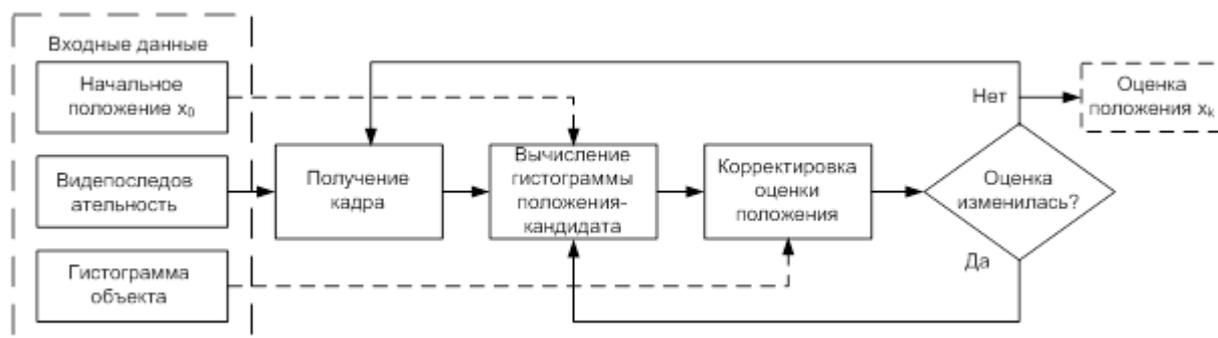


Рис. 2. – Структурная схема алгоритма meanshift

Принцип действия алгоритма mean shift основан на поиске максимума плотности вероятности некоторой функции, описывающей дискретные данные изображения. При оценке функции плотности вероятностей используется ядро, обозначаемое $K(x_i - x)$. Ядро определяет вес различных точек при оценке среднего. Обычно используется ядро гаусса $K(x_i - x) = e^{-c\|x_i - x\|^2}$. Взвешенное среднее вычисляется по формуле:

$$m(x) = \frac{\sum_{x_i \in N(x)} K(x_i - x)x_i}{\sum_{x_i \in N(x)} K(x_i - x)}$$

здесь $N(x)$ – окрестность точки x , набор точек для которых $K(x) \neq 0$.

Алгоритм производит оценку для точек $x \leftarrow m(x)$, и повторяет оценку до тех пор, пока $m(x)$ не перестанет изменяться.

Преимуществом данного подхода является высокая вычислительная эффективность. Но, к сожалению, он проигрывает более сложным многопотенциальным методам по точности отслеживания. Тем не менее, результаты данного подхода могут быть улучшены за счет использования дополнительной информации, такой как оптический поток – оценки мгновенной скорости для каждого пикселя кадра.

Методы, основанные на оптическом потоке, вычисляют движение между двумя кадрами, взятыми в момент времени t и, в каждом пикселе. Эти методы называются дифференциальными, так как они основаны на приближении сигнала отрезком ряда Тейлора; таким образом, они используют частные производные по времени и пространственным координатам.

В случае размерности $2D+t$ (случаи большей размерности аналогичны) пиксель в позиции (x, y, t) с интенсивностью $I(x, y, t)$ за один кадр будет перемещен на $\delta x, \delta y, \delta t$ и можно записать следующее уравнение:

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t)$$

Для оценки скорости движения пикселя может использоваться следующее уравнение $\nabla I^T \cdot \vec{V} = -I_t$.

Данное уравнение содержит две неизвестных и не может быть однозначно разрешено. Данное обстоятельство известно как проблема апертюры. Задачу решает наложение дополнительных ограничений — регуляризация. Существуют алгоритмы, позволяющие быстро вычислять оптический поток для каждого кадра видеопоследовательности.

Структурная схема системы отслеживания положения объекта на основе оптического потока показана на рис. 3.



Рис. 3. – Блок схема предлагаемого подхода

Блок схема содержит следующие основные блоки:

1. Блок вычисления оптического потока;
2. Блок вычисления низкоуровневых признаков;

3. Блок инициализации;
4. Блок корректировки положения объекта.

На этапе предварительной обработки производится вычисление оптических потоков для каждого кадра видеоряда, а так же низкоуровневых признаков (таких, как представление каждого пикселя в цветовом пространстве YIU). Для подвижных объектов оптический поток имеет большее значение, чем для неподвижного фона. Оптический поток позволяет нам найти объекты переднего плана на видеопоследовательности.

В первую очередь производится выделение объектов переднего плана. На практике используется несколько методов вычитания фона (background subtraction) [8]. Методы вычитания фона – это методы, которые строят модель заднего плана сцены и на каждом кадре производят попиксельное сравнение текущего значения в каждом пикселе с моделью и в соответствии с результатами сравнения принимают решение, принадлежит пиксель к заднему плану или к переднему.

Выделение объектов переднего плана включает следующие шаги:

1. Сохранение первого изображения видеопоследовательности. Обозначим его B .

2. Для каждого следующего кадра (обозначим текущий кадр C) если $|B_{ij} - C_{ij}| \leq \delta$, $i=1..w$, $j=1..h$, где w и h это ширина и высота картинки соответственно, то пиксель (i,j) принадлежит к переднему плану, иначе он принадлежит к заднему плану.

3. Для устранения шума к полученному бинарному изображению можно применить медианный фильтр или операции мат. морфологии.

Предлагается использовать вычисленный оптический поток в составе функции-ядра, при отслеживании с помощью алгоритма mean-shift.

Функция ядра примет вид $K(x_i - x) = e^{-c\|x_i - x\|^2} \cdot V_x$.

Таким образом, мы ввели учет скорости движения точки в ее общий вес, при вычислении нового положения. Траектория получения в результате работы алгоритма MeanShift не обладает достаточной гладкостью,

предлагается использовать аппроксимацию значений полученной траектории гладкой функцией. Предложенный метод реализован на персональном компьютере с использованием языка программирования C++ и библиотеки OpenCV [7]. Для исследования разработанного алгоритма использовался набор тестовых видеопоследовательностей MOTINAS [9].

Для проведения сравнения были выбраны следующие алгоритмы: Lukas-Kanade, Lukas-Kanade с использованием пирамид изображений, алгоритм Shi-Tomasi, классический MeanShift.

Для оценки качества работы алгоритма используем критерий среднеквадратичного отклонения:

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2},$$

где в качестве \bar{x} выступает действительное положение объекта.

Для сравнения результатов обработки из выбранного набора было выбрано две видеопоследовательности: `motias_emilio_webcam` и `motias_toni`.

На рис. 4 показан пример работы разработанного алгоритма для видеопоследовательности `motias_toni`.



Рис. 4. – Пример работы разработанного алгоритма.

Все, исследуемые алгоритмы, в целом справились с отслеживанием объекта, ни в одном случае он не был утерян. Отличия состоят лишь в точности определения положения объекта.

В таблице №1 приведены значения среднеквадратичной ошибки для результатов работы различных алгоритмов.

Значение среднеквадратичной ошибки

Номер испытания	Среднеквадратичное отклонение	
	motinas_toni	motinas_emilio_webcam
LK	29.8	36.7
LK-pyr	47.6	53.1
Shi-Tomasi	20.1	29.4
MeanShift	47.9	54.1
Разработанный алгоритм	17.2	23.8

Анализ результатов показывает, что предложенный алгоритм обеспечивает меньше значение среднеквадратической ошибки, в среднем на 14%.

На основе результатов эксперимента можно сделать вывод, что использование информации об оптическом потоке в составе функции-ядра при использовании алгоритма mean shift позволяет повысить точность отслеживания объектов на видеопоследовательности.

Работа выполнена в рамках федеральной целевой программы ФЦП "Научные и научно-педагогические кадры инновационной России" на 2009-2013 годы.

Литература:

1. Szeliski R. Computer Vision. Algorithms and Applications. Washington, Springer, 793 pages, 2011.

2. Костоглотов А.А., Костоглотов А.И., Лазаренко С.В., Андрашитов Д.С. Многопараметрическая идентификация конструктивных параметров методом объединенного принципа максимума [Электронный ресурс] // Инженерный вестник Дона, №1 – Ростов-на-Дону: Изд-во Северо-Кавказский научный центр высшей школы Южного федерального университета, 2011. – Режим доступа:

<http://ivdon.ru/magazine/archive/n1y2011/348> (доступ свободный) – Загл. с экрана. – Яз. рус.

3. Maggio E., Cavallaro A. Video tracking. Theory and practice. Wiley, 292 pages, 2011.

4. Shi J., Tomasi C. Good features to track. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, USA, PP. 593–600, 1994.

5. Целигоров Н.А., Целигорова Е.Н., Мафура Г.В. Математические модели неопределённостей систем управления и методы, используемые для их исследования [Электронный ресурс] // Инженерный вестник Дона, №4 – Ростов-на-Дону: Изд-во Северо-Кавказский научный центр высшей школы Южного федерального университета, 2012. – Режим доступа: <http://ivdon.ru/magazine/archive/n4p2y2012/1340> (доступ свободный) – Загл. с экрана. – Яз. рус.

6. Воронин В.В., Марчук В.И., Франц В.А., Гапон Н.В. Реконструкция видеопоследовательностей с использованием модели фона [Текст]// Нелинейный мир, №11 – Москва: Изд-во Радиотехника, 2012. – С. 777 – 781.

7. Cavallaro A., Steiger O., Ebrahimi T. Tracking video objects in cluttered background. IEEE Transactions on Circuits and Systems for Video Technology, 15(4):575–584, 2005.

8. Воронин В.В., Франц В.А., Фисунов А.В. Использование карты глубины в задаче слежения за подвижными объектами [Текст]// Успехи современной радиоэлектроники, №6 – Москва: Изд-во Радиотехника, 2012. – С. 57– 61.

9. Bradsky G., Kaehler A. Learning OpenCV. O'Reilly, 580 pages, 2008.

10. Nawaz T., A. Cavallaro, PFT: A protocol for evaluating video trackers, in Proc. of IEEE Int. Conf. on Image Processing, Brussels, PP. 11-14, 2011.