

ИНФОРМАТИКА

УДК 681.3

КОМБИНИРОВАНИЕ БЛОЧНЫХ АЛГОРИТМОВ ВЫЧИСЛЕНИЯ ОПТИЧЕСКОГО ПОТОКА ДЛЯ ОБНАРУЖЕНИЯ И СОПРОВОЖДЕНИЯ ДВИЖУЩИХСЯ ОБЪЕКТОВ НА ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЯХ

*канд. техн. наук, доц. Р.П. БОГУШ, В.Ю. ЛЫСЕНКО, Г.А. САМОЩЕНКОВ
(Полоцкий государственный университет)*

Рассматривается комбинированный алгоритм обнаружения и сопровождения движущихся объектов на видеопоследовательностях с использованием блочного метода оценки движения. Для сокращения временных затрат используется пирамидальное представление кадра и шаблонный поиск на этапе построения предварительной карты векторов движения. Уточнение векторного поля кадра выполняется с использованием набора векторов-кандидатов. Для повышения точности локализации объектов применяется мультипликативная минимаксная функция схожести при анализе векторов движения. Сопровождение объектов осуществляется на основе модифицированного алгоритма трассировки с использованием аддитивной минимаксной функции схожести. Представлены результаты экспериментов по оценке временных затрат на основных этапах алгоритма для видеопоследовательностей различного разрешения.

Введение. Интеллектуальные системы технического зрения достаточно перспективны для решения различных прикладных задач в производстве, медицине, робототехнике и т.д. [1 – 3]. Построение таких систем – сложная комплексная задача, которая предполагает получение цифрового видеоряда и его обработку с целью выделения и последующего анализа необходимой информации [1; 2].

К основным этапам работы интеллектуальных систем технического зрения относят: автоматическое обнаружение и сопровождение движущихся объектов в условиях воздействия различного рода помех и возмущений; распознавание и описание действий объектов, представляющих интерес [1].

Построение алгоритмов обнаружения и сопровождения движущихся объектов, инвариантных к качественным характеристикам видеопотока является очень сложной задачей вследствие влияния таких факторов, как: изменение освещенности динамической сцены, шум камеры, изменение формы объекта, одновременное присутствие нескольких объектов со схожими характеристиками и т.д. Для решения данной задачи необходимо развитие методов обработки, среди которых перспективным является метод на основе вычисления оптического потока динамических изображений [1; 4].

Под оптическим потоком понимают дискретную аппроксимацию движения в трехмерной сцене, осуществляемую путем проектирования скоростей трехмерных поверхностей на плоскость изображения или визуального сенсора [5]. Двумерное поле скоростей, которым является оптический поток, используется для описания движения в сцене.

Чтобы определить оптический поток, как правило, применяют два подхода [1]: дифференциальный анализ (градационные методы), позволяющий связывать временные и пространственные рассогласования; блочный метод, предполагающий поиск наилучшего согласования блоков фиксированного размера, относящихся к различным кадрам.

Вычисление оптического потока также выполняется на основе частотного и фазового методов.

Современные дифференциальные алгоритмы слежения за особенностями в видеопотоке опираются на работы Lucasa – Kanade [6], Tomasi – Kanade и Shi – Tomasi [7; 8]. Недостатками дифференциального метода вычисления оптического потока является значительная ресурсоемкость и сложность реализации алгоритмов, а также сложность точного численного дифференцирования из-за наличия шума камеры, малой частоты смены кадров и других возмущающих воздействий [1].

Частотный и фазовый методы вычисления оптического потока требуют еще больших временных затрат. Поэтому на практике чаще всего используется блочный метод вычисления оптического потока. При этом предполагается, что все пиксели блока кадра претерпевают одинаковое перемещение и им соответствует один и тот же вектор движения. Однако и в этом случае актуальными являются проблемы точности вычисления оптического потока и быстродействия.

В данной работе предлагается модификация алгоритмов обнаружения и сопровождения объектов на основе блочного метода вычисления оптического потока с целью обеспечения приемлемых вычислительных затрат при выполнении процедур обнаружения и сопровождения движущихся объектов на видеопоследовательностях.

1. Обнаружение движения на основе блочного метода вычисления оптического потока

Реализация блочного алгоритма вычисления оптического потока предполагает разбиение кадра $F = f_{xy}$ размером $M_1 \times M_2$ ($x \in 0, \dots, M_1 - 1$; $y \in 0, \dots, M_2 - 1$) на прямоугольные блоки одинакового размера b_{ij}^t и поиск соответствующего блока в предыдущем кадре b_{ij}^{t-1} с использованием корреляционной функции схожести [9; 10]. Таким образом, задача вычисления оптического потока сводится к задаче поиска вектора движения v_{ij} для каждого блока b_{ij} . При этом векторы v_{ij} определяются как

$$v_{ij} = \max_{v_{i,j} \in N} (Corr(t, i, j, v_{i,j})),$$

где $N \times N$ – область поиска векторов движения; $Corr(t, i, j, v_{i,j})$ – корреляционная функция схожести блоков.

Существует ряд модификаций блочного метода вычисления оптического потока [9]: метод полного перебора, шаблонные методы, методы иерархического поиска и т.д.

Известно, что основным достоинством метода полного перебора является гарантированное нахождение максимальной функции схожести для каждого блока, а основным недостатком является огромная вычислительная сложность. Развитием метода полного перебора является шаблонный поиск, который представляет собой дискретный вариант покоординатного спуска и основывается на предположении, что функция схожести достаточно гладкая для каждого блока, без локальных максимумов [9]. При этом под шаблоном понимают набор координат точек, причем координаты точек отсчитываются от центра шаблона. Такой подход использует итеративный алгоритм, который предполагает определение координат центра и всех ключевых точек шаблона, вычисление функции схожести в каждой из точек шаблона и выбор в качестве центра шаблона следующей итерации точки, которая характеризуется максимальным значением функции схожести. Далее проверяется условие прекращения процедуры поиска, и в зависимости от результата производится переход к следующей итерации или завершение поиска вектора в целом. Однако, корреляционные характеристики видеоизображений, как правило, далеки от идеальных, то есть характеризуются значительным уровнем боковых лепестков и размытостью основного пика, что приводит либо к ложным опознаниям блока, либо к неточному позиционированию блока в кадре.

Иерархический поиск предполагает пирамидальное представление кадров с N уровнями, т.е. формирование изображений кадра в различных пространственных масштабах [9; 11]. Тогда вычисление оптического потока состоит из N итераций, на каждой из которых обрабатываются изображения соответствующих уровней пары кадров. При этом в качестве стартовой точки на каждой итерации выбирается векторное поле, полученное с предыдущей итерации, т.е. каждая очередная итерация производит уточнение векторов, вычисленных на предыдущей итерации. Основными достоинствами такого подхода являются: сокращение времени обработки, улучшенная устойчивость к шуму. Однако высока вероятность неправильного определения движения объектов небольших размеров из-за использования на первом этапе кадров уменьшенного размера.

В основе методов, использующих векторы-кандидаты, лежит утверждение о том, что если соседние блоки принадлежат одному движущемуся объекту, то их вектора движения схожи [9].

Перед вычислением информации о движении для текущего блока формируется набор, состоящий из уже найденных векторов движения соседних блоков. Сформированный набор векторов называют набором векторов-кандидатов. В качестве вектора движения в каждом блоке выбирается лучший вектор из набора векторов-кандидатов. Методы данного класса имеют низкую вычислительную сложность [9].

В целом блочные алгоритмы весьма перспективны для вычисления оптического потока, а комбинирование приемов из данного класса алгоритмов позволяет строить достаточно эффективные процедуры обнаружения и сопровождения движущихся объектов.

2. Обнаружение и локализация динамических объектов

Синтезирован комбинированный алгоритм обнаружения динамических объектов на основе блочного метода оценки движения, который требует выполнения следующих основных шагов:

- 1) захват двух соседних кадров размером $M_1 \times M_2$ F^{t-1} и F^t видеоряда;
- 2) получение N уровней пирамидального представления кадров, при котором в основе пирамиды лежит его поэлементное описание, а каждое последующее описание формируется из предыдущего путем прореживания в два раза:

$$f_{ij}^t = \frac{\sum_{x=i-n}^{i+n} \sum_{y=j-n}^{j+n} f_{xy}^t}{n^2},$$

где $x \in M_1, y \in M_2$; f_{xy}^t – яркость пикселя исходного изображения; f_{ij}^t – яркость пикселя уменьшенного изображения; n – коэффициент уменьшения.

Количество уровней разложения определяется минимальными размерами движущихся объектов;

3) *вычисление предварительной карты векторов движения* для изображения верхнего уровня пирамиды с применением шаблонного поиска. В качестве функции схожести используется минимаксная аддитивная функция [12], которая в общем случае для изображения $A = a_{ij}$ и изображения $B = b_{ij}$ размером $N_1 \times N_2$ определяется как

$$R^S = \frac{1}{N_1 N_2} \sum_{i=0}^{N_1-1} \sum_{j=0}^{N_2-1} \frac{\min a_{ij}, b_{ij}}{\max a_{ij}, b_{ij}};$$

4) *уточнение предварительной карты векторов движения* с использованием набора векторов-кандидатов для вектора v_{ij} [13]: левый вектор $v_{i-1,j}$, верхний вектор $v_{i,j-1}$, правый верхний вектор $v_{i+1,j-1}$; среднее значение данных векторов; вектор движения, полученный из предыдущего кадра. Полученная карта векторов используется в качестве исходных данных для более низкого уровня пирамиды;

5) *выполнение рекурсивной медианной фильтрации векторного поля в пределах одного кадра* для удаления ложных векторов движения. Однако следует отметить, что такая процедура требует значительных временных затрат, поэтому ее следует применять при реализации алгоритма лишь на аппаратных специализированных вычислительных средствах;

6) *локализация объектов*. Результатом вычисления оптического потока является карта векторов движения, которая отображает движущиеся области. Для того чтобы соотнести области движения с объектами на кадре необходимо выполнить процедуру локализации путем определения характеристик векторов на основе анализа векторного поля кадра.

Известно, что векторы, относящиеся к одному объекту, обладают следующими характеристиками [2; 13]: располагаются в одной области кадра и образуют связную группу; являются сонаправленными и имеют схожие величины смещения по горизонтали и вертикали.

Для оценки сонаправленности векторов определяется угол между векторами движения:

$$\alpha = \arccos \left(\frac{dx_1 dx_2 + dy_1 dy_2}{\sqrt{(dx_1^2 + dy_1^2)(dx_2^2 + dy_2^2)}} \right),$$

где dx_1, dy_1, dx_2, dy_2 – проекция векторов 1 и 2 на оси координат.

Сходство величины смещения векторов вычисляется с использованием мультипликативной минимаксной функции схожести, обладающей высокой контрастностью:

$$R^M = \frac{\min(dx_1, dx_2) \cdot \min(dy_1, dy_2)}{\max(dx_1, dx_2) \cdot \max(dy_1, dy_2)}.$$

Для учета связности векторов рассматривается восемь соседних векторов.

Для каждого соседнего вектора проверяется угол отклонения между векторами α , который не должен превышать 90 градусов и величины $R^M < T$, где T – пороговый уровень. Если проверка выполнена удачно, то текущий вектор относится к объекту.

3. Сопровождение движущихся объектов

Сопровождение (трассировка, трекинг, слежение) объекта предполагает установление соответствий между различными объектами или их частями в видеоряде и нахождение траектории движения, а также определение других динамических характеристик объекта или определение движения объекта по данной последовательности изображений [1; 14]. Процедура сопровождения осуществляется после обнаружения и локализации движущихся объектов. Для этого вычисляются следующие характеристики объекта $O = o_{ij}$ размером $W \times W$ [15]:

1) *центр тяжести объекта* (i_0, j_0) , под которым понимают центр энергии светового изображения объекта:

$$i_0 = \frac{\sum_{i=0}^{W-1} \sum_{j=0}^{W-1} [o_{ij} i]}{\sum_{i=0}^{W-1} \sum_{j=0}^{W-1} o_{ij}}; \quad j_0 = \frac{\sum_{i=0}^{W-1} \sum_{j=0}^{W-1} [o_{ij} j]}{\sum_{i=0}^{W-1} \sum_{j=0}^{W-1} o_{ij}}.$$

Распределение яркости по полю изображения объекта дает дополнительную информацию о его местоположении и позволяет увеличить точность определения координат по сравнению с точечным объектом, координаты которого определяются по единственному отсчету видеосигнала с точностью до элемента разложения [15];

2) *площадь объекта*:

$$S = \sum_{i=0}^{W-1} \sum_{j=0}^{W-1} o_{ij}^b,$$

где $o_{ij}^b = \begin{cases} 1 & \text{при } i, j \in W; \\ 0 & \text{при } i, j \notin W; \end{cases}$

3) *линейные размеры по горизонтали и по вертикали* соответственно:

$$l = \sum_{i=0}^{N-1} o_i^b, \quad J = \sum_{j=0}^{N-1} o_j^b.$$

Рассчитанные признаки локализованных объектов используются для их обнаружения на последующем кадре. Каждый объект, обнаруженный и локализованный с использованием блочного метода вычисления оптического потока на первом кадре, либо на последующих кадрах впервые, относится к движущимся. Объекты, локализованные на текущем кадре и обнаруженные на следующем с использованием вероятностного подхода, определяются как сопровождаемые, их характеристики обновляются, и выполняется построение траектории движения таких объектов с применением фильтра Кальмана [14]. В качестве функции схожести на этапе сопровождения используется аддитивная минимаксная функция [12]. Если объект с предыдущего кадра не обнаружен на текущем кадре, то он относится к потерянным. Однако с целью возможности его обнаружения на последующих кадрах в случае кратковременной потери оптической связи с таким объектом алгоритмом предусмотрено хранение его характеристик заданное время. Для поиска объектов на следующем кадре относительно предыдущего используется подход, представленный в [16].

4. Результаты исследований

Для исследования представленных алгоритмов обнаружения и сопровождения движущихся объектов на видеопоследовательностях разработано программное обеспечение на языке C++. Программная реализация алгоритмов на языке C++ позволяет переносить их логику работы на различные аппаратные решения для комплексных интеллектуальных систем видеонаблюдения.

Для оценки временных затрат работы алгоритмов проведены экспериментальные исследования на основе разработанного программного обеспечения с использованием ПЭВМ с основными параметрами AMD Athlon (tm) 64 2.21 ГГц, 960 МГБ ОЗУ.

Исходными данными при выполнении экспериментов были полутоновые и цветные статические и динамические изображения, полученные в различных условиях съемки и с различными качественными характеристиками.

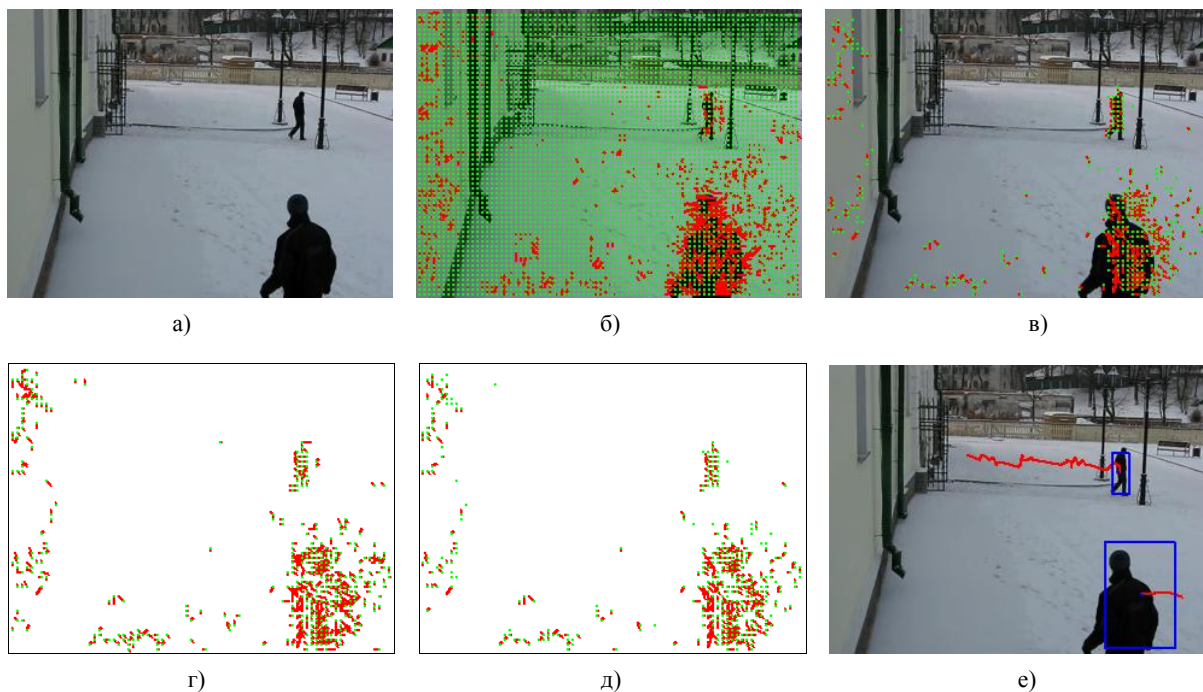
Распределение временных затрат комбинированного алгоритма обнаружения и сопровождения движущихся объектов по его основным этапам при обработке одного кадра видеоряда и тотальные временные затраты приведены в таблице. Следует отметить, что временные затраты этапа сопровождения в основном определяются количеством трассируемых объектов, поэтому использовались динамические изображения различного разрешения, но с равным числом движущихся объектов.

Распределение временных затрат алгоритма по его основным этапам при обработке одного кадра видеоряда

Размер кадра	Временные затраты по основным этапам обработки, мс			Тотальная временная сложность, мс
	построение карты векторов оптического потока	локализация объектов	сопровождение объектов	
320 × 240	15	8	10	33
336 × 272	25	10	10	45
640 × 480	91	39	10	190
720 × 576	139	59	10	208

Из таблицы видно, что формирование карты векторов оптического потока характеризуется максимальными вычислительными затратами по сравнению с другими итерациями алгоритма.

На рисунке представлены результаты основных этапов обработки кадра видеоряда при обнаружении и трассировке движущихся объектов.



Основные этапы обработки кадра видеоряда:

- а – исходное изображение; б – изображение кадра с предварительной картой векторов движения на нижнем уровне пирамиды; в – изображение кадра с уточненной картой векторов движения; г – векторное поле кадра; д – векторное поле кадра после его медианной фильтрации; е – результат сопровождения с построением траектории движения

Заключение. В результате проведенного исследования синтезирован алгоритм обнаружения динамических объектов на основе блочного метода оценки движения. Алгоритм предполагает пирамидальное представление кадра для сокращения временных затрат. Для повышения вероятности правильного обнаружения объектов применяется уточнение предварительной карты векторов движения на основе полученных векторов движения на предыдущем шаге. Отличительной особенностью алгоритма является также применение мультипликативной минимаксной функции схожести при анализе векторов движения для повышения точности локализации объектов. Для сопровождения объектов при использовании метода оптического потока представлен модифицированный алгоритм трассировки. В качестве функции схожести на этапе сопровождения используется аддитивная минимаксная функция. Алгоритмом предусмотрено хранение основных характеристик движущегося объекта в случае кратковременной потери оптической связи с ним с целью возможности его обнаружения на последующих кадрах. Проведен ряд экспериментов, которые подтвердили эффективность предложенного подхода и возможность его использования в интеллектуальных системах технического зрения. Дальнейшее развитие алгоритма предполагается путем сокращения вычислительных затрат на этапе построения карты векторов оптического потока.

ЛИТЕРАТУРА

1. Лукьяница, А.А. Цифровая обработка видеоизображений / А.А. Лукьяница, А.Г. Шишкин. – М.: «Ай-Эс-Эс Пресс», 2009. – 518 с.
2. Методы автоматического обнаружения и сопровождения объектов. Обработка изображений в управлении / Б.А. Алпатов [и др.]. – М.: Радиотехника, 2008. – 176 с.
3. Птицын, Н.В. Аппаратная видеоаналитика для охраны стратегических объектов / Н.В. Птицын, А.С. Чижов [Электронный ресурс]. – Режим доступа: <http://synesis.ru/ru/surveillance/articles/embedded-va-slides>. – Дата доступа: 24.09.2010.

4. Обухова, Н.А. Сегментация объектов интереса на основе признака движения в видеокомпьютерных системах / Н.А. Обухова // Инфокоммуникационные технологии. – 2007. – Т. 5, № 1. – С. 77 – 84.
5. Садыхов, Р.Х. Модификация алгоритма Хона – Шунка для выделения в видеопотоке движущихся людей / Р.Х. Садыхов, П.Г. Мельников // Весці НАН Беларусі. Сер. фіз.-тэхн. навук. – 2006. – № 4. – С. 93 – 97.
6. Lucas, B. An Iterative Image Registration Technique with an Application to Stereo Vision / B. Lucas, T. Kanade // Int. Joint Conf. on Artificial Intelligence, 1981. – P. 674 – 679.
7. Tomasi, C. Detection and Tracking of Point Features. Carnegie Mellon University Technical Report CMU-CS-91-132 / C. Tomasi, T. Kanade, 1991. – 22 p.
8. Shi, J. Good Features to Track / J. Shi, C. Tomasi // IEEE Conference on Computer Vision and Pattern Recognition. – 1994. – P. 593 – 600.
9. Обзор блочных методов оценки движения в цифровых видеосигналах / С.В. Гришин [и др.] // Программные системы и инструменты: темат. сб. – 2008. – № 9. – С. 50 – 62.
10. Садыхов, Р.Х. Инструментальная система для обработки видеоинформации / Р.Х. Садыхов, Д.В. Ламовский // Доклады БГУИР. – 2007. – № 4(20). – С. 175 – 180.
11. Выделение и межкадровое прослеживание движущихся объектов при регистрации изображений сложных пространственных сцен произвольно движущимися двумерными сенсорами / Ю.В. Визильтер [и др.] // Вестн. компьютерных и информационных технологий. – 2006. – № 3. – С. 34 – 39.
12. Bogush, R. New families similarity function for images processing / R. Bogush, S. Maltsev // Proc. of the 9 Int. Conf. PRIP'2007, Minsk, May 22 – 24, 2007 / UIIP of NASB. – Minsk, 2007. – Vol. 1. – P. 40 – 44.
13. A Fast Adaptive Motion Estimation Algorithm / A. Ishfaq [and others] // IEEE Transactions on Circuits and Systems for Video Technology. – 2006. – Vol. 16, № 3. – P. 420 – 438.
14. Форсайт, Д. Компьютерное зрение. Современный подход / Д. Форсайт, Ж. Понс; пер. с англ. – М.: Издат. Дом «Вильямс», 2004. – 928 с.
15. Коротаев, В.В. Телевизионные измерительные системы / В.В. Коротаев, А.В. Краснящих. – СПб.: СПбГУ ИТМО, 2008. – 108 с.
16. Narayana, M. Automatic Tracking of Moving Objects in Video for Surveillance Applications / M. Narayana [Electronic resource]. – 2007. – Mode of access: <http://www.citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.142.2246...pdf>. – Date of access: 20.06.10.

Поступила 09.02.2011

BLOCKS-BASED ALGORITHMS COMBINATION FOR OPTICAL FLOW CALCULATION FOR MOVING OBJECTS DETECTION AND TRACKING IN VIDEO SEQUENCES

R. BOGUSH, V. LYSENKO, G. SAMOSCHENKOV

The combined algorithm moving objects detection and tracking in video sequences with use of blocks-based approach for motion estimation is considered. For time acceleration we use multilevel successive elimination algorithm and diamond search for an interim maps of motion vector construction. We use multiplicative minimax similarity function for accuracy improvement of object localization at motion vector analysis. Since interim map not necessarily leads to optimal motion vectors, we use an additional post processing for motion field based on candidate vectors. We apply minimax similarity function for updating algorithm of moving object tracking. In paper results of experiments are presented according to time expenses of the moving object detection and tracking.