

УДК 004.932

А. П. Кирпичников, С. А. Ляшева, М. П. Шлеймович

ОБНАРУЖЕНИЕ И СОПРОВОЖДЕНИЕ ОБЪЕКТОВ В БОРТОВЫХ СИСТЕМАХ ОБРАБОТКИ ИЗОБРАЖЕНИЙ

Ключевые слова: обнаружение объектов на изображении, сопровождение объектов на изображениях, локальные бинарные шаблоны, сопоставление гистограмм, алгоритм сдвига среднего, фильтр частиц.

В работе рассмотрены и проанализированы методы обнаружения и сопровождения объектов, которые можно эффективно реализовать в реальном времени в составе бортовых систем мобильных объектов.

Keywords: detect objects on images, tracking objects on images, local binary patterns, histogram matching, mean shift algorithm, particle filters.

In this paper discusses and analyzes the methods of detection and tracking of objects that can be implemented efficiently in real-time as part of on-board systems of mobile objects.

Бортовые системы современной движущейся техники (мобильных объектов) представляют собой сложные интегрированные комплексы, в составе которых, как правило, присутствуют средства машинного зрения, основанные на применении технологий обработки и анализа изображений различной природы, например, оптических, инфракрасных и радиолокационных [1].

Среди задач, решаемых бортовыми средствами машинного зрения, можно выделить обнаружение и сопровождение объектов на изображениях. В этом случае предполагается получение и обработка последовательности кадров видеоизображения. Примерами объектов, которые требуется обнаруживать и сопровождать, могут служить автомашины, люди, дома, маяки, дорожная разметка [2].

Условно обработку видеоизображений в системах обнаружения и сопровождения объектов можно разделить на следующие этапы:

1. Обнаружение объектов на изображении;
2. Отслеживание траектории объектов;
3. Анализ траектории объектов.

При этом необходимо обеспечить реализацию эффективных алгоритмов обработки изображений, ориентированных на использование в реальном времени [3].

При проектировании систем на базе средств машинного зрения, возникает ряд трудностей, связанных с тем, что нет универсальных алгоритмов решения задач обработки изображений. Кроме того, имеются ограничения, определяемые конструктивными недостатками аппаратных устройств, таких, как ограничения динамических диапазонов свето- и цветопередачи, дискретность фотоприёмников, астигматизм оптического канала, ошибки калибровочных характеристик и временных привязок потока видеоданных и т.д. [1, 2].

Преодоление указанных трудностей возможно только при условии разработки новых эффективных математических методов и

алгоритмов обработки и анализа изображений в процессе решения конкретных задач управления мобильными объектами.

Таким образом, разработка методов и средств машинного зрения, используемых в системах управления мобильными объектами, является актуальной научно-исследовательской и практической задачей.

Рассмотрим методы, применяемые для обнаружения и сопровождения объектов.

Под обнаружением объекта будем понимать определение некоторой области изображения, ограничивающей данный объект. На выходе процедуры обнаружения должно быть получено описание данной области, включающее ее положение на изображении, размеры и характерные особенности. Обнаружение может быть ручным или автоматическим. Ручное обнаружение объекта интереса заключается в просмотре кадров человеком-оператором и указании им области интереса. Автоматическое обнаружение подразумевает поиск области, содержащей объект, в соответствии с заданным критерием, например, допустимой степенью сходства с заданным шаблоном (изображение объекта или совокупность его характеристик).

Один из возможных подходов к решению задачи обнаружения состоит в использовании методов машинного обучения для построения моделей классов объектов [4]. Данные методы эффективно используются, например, для обнаружения лиц, пешеходов, транспортных средств. Однако необходимо отметить, что в этом случае часто необходимо предварительно подготовить обучающую выборку – множество примеров объектов классов.

Другой подход к обнаружению объектов основан на применении методов поиска по шаблону [4]. В этом случае выделяются и сравниваются однотипные признаки изображения и шаблона. В простейшем случае в качестве шаблона может выступать изображение объекта. Более сложные

методы предполагают использование характерных черт (дескрипторов) изображения объекта, которые можно разделить на цветовые, текстурные, геометрические и структурные. В процессе поиска осуществляется перемещение по изображению окна заданных размеров и сравнение описания области, соответствующей текущему положению окна и шаблонного изображения по некоторой метрике.

Традиционная техника сравнения текущего изображения с шаблоном основывается на рассмотрении изображений как двумерных функций яркости или цвета. При этом часто используется корреляционная метрика, основанная на нормированном коэффициенте корреляции.

Часто применяются методы поиска с использованием пирамиды изображений, позволяющие значительно сократить время обработки.

Эффективны методы, основанные на применении алгоритмов детектирования точечных особенностей (ключевые точки, особые точки), например, детектор углов Харриса, SIFT, SURF.

Один из развивающихся подходов основан на вейвлет-преобразовании. В этом случае используются сигнатуры – укороченные квантованные версии вейвлет-разложений. Для определения степени сходства изображений вычисляется количество их общих значимых вейвлет-коэффициентов или характеристик на их основе.

Результаты исследований показали, что эффективным соотношением между точностью, временем и ресурсоемкостью для бортовых систем мобильных объектов обладают методы обнаружения на основе локальных бинарных шаблонов [5].

Локальные бинарные шаблоны используются для описания текстурных характеристик изображений. Локальный бинарный шаблон (ЛБШ) представляет собой описание окрестности точки изображения в двоичной форме. Оператор ЛБШ, который применяется к пикселу изображения, использует N пикселей окрестности, принимая центральный пиксел в качестве порога. Пиксели окрестности, интенсивности которых больше или равны интенсивности центрального пиксела принимают значения 1, а остальные – значения 0. Таким образом, получается N -разрядный бинарный код, который описывает окрестность пиксела:

$$LBP(x_c, y_c) = \sum_{i=0}^{N-1} s(g_i - g_c) 2^i, \quad (1)$$

где x_c, y_c – координаты пиксела, g_c – интенсивность пиксела, g_i – интенсивности пикселей окрестности, $i = 0, \dots, N-1$, s – пороговая функция, определяемая следующим образом:

$$s(x) = \begin{cases} 1, & x \geq 0, \\ 0, & x < 0. \end{cases} \quad (2)$$

Обычно используются восьмиразрядные ЛБШ.

Локальный бинарный шаблон называется равномерным, если он содержит не более трех серий 0 и 1 (например, 00000000, 001110000 и 11100001). Равномерные ЛБШ определяют только наиболее

важные локальные особенности изображения и обеспечивают существенное сокращение множества различных кодов. Например, на рис. 1 показаны коды для описания следующих элементов изображения: пятно (а), пятно/фон (б), конец линии (в), грань (г), угол (д).

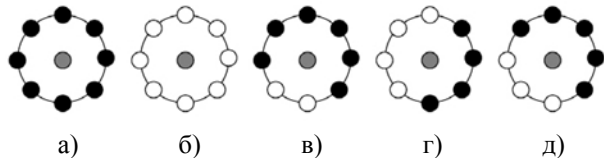


Рис. 1 - Равномерные ЛБШ

Дескриптором области изображения на основе ЛБШ является гистограмма кодов. При сравнении гистограмм H_1 и H_2 , содержащих K ячеек, могут быть использованы различные метрики, например [6]:

- корреляционное расстояние:

$$d(H_1, H_2) = \frac{\sum_{k=1}^K H'_1(k) H'_2(k)}{\sqrt{\sum_{k=1}^K H_1'^2(k) H_2'^2(k)}}, \quad (3)$$

где $H'_i(k) = H_i(k) - \frac{1}{K} \sum_{l=1}^K H_i(l)$, $i = 1, 2$,

- расстояние Хи-квадрат:

$$d(H_1, H_2) = \sum_{k=1}^K \frac{(H_1(k) - H_2(k))^2}{H_1(k) + H_2(k)}, \quad (4)$$

- пересечение:

$$d(H_1, H_2) = \sum_{k=1}^K \min(H_1(k), H_2(k)), \quad (5)$$

- расстояние Бхатачария:

$$d(H_1, H_2) = \sqrt{1 - \frac{\sum_{k=1}^K \sqrt{H_1(k) H_2(k)}}{\sqrt{\sum_{k=1}^K H_1(k) \sum_{k=1}^K H_2(k)}}}, \quad (6)$$

- логарифмическое правдоподобие:

$$d(H_1, H_2) = - \sum_{k=1}^K H_1(k) \log H_2(k). \quad (7)$$

Поскольку вычисление расстояния пересечения гистограмм требует наименьшего времени, то именно его можно рекомендовать для применения в бортовых системах мобильных объектов, работающих в реальном времени.

Алгоритмы на основе локальных бинарных шаблонов активно применяются для обнаружения и сопровождения лиц, транспортных средств, объектов естественной природы и решении других задач.

После обнаружения объекта на одном из кадров видеоизображения необходимо выполнить отслеживание траектории его движения, т.е. определить его положение на следующих кадрах.

Одним из подходов здесь является корреляционное слежение за объектами, которое осуществляется путем сравнения изображения объекта, полученного с одного из предыдущих

кадров (или некоторого базового шаблона объекта) с последующими кадрами. При этом считается, что максимуму корреляционной функции соответствует местоположение искомого объекта на новом кадре. Последовательное сканирование изображения с вычислением значения корреляционной функции для каждой его области в этом случае требует значительного времени.

Менее ресурсоемким и при этом эффективным алгоритмом является алгоритм сдвига среднего, который принято называть MeanShift [7]. Идея алгоритма состоит в том, что для каждого объекта выбирается окно поиска и вычисляется центр масс распределения его характеристик, который задает положение объекта на текущем кадре. Это же положение является начальным для поиска на следующем кадре. Далее подробно рассмотрим подход, на котором базируется этот алгоритм.

В основе метода MeanShift лежит ядерная оценка плотности распределения:

$$f(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n K(\mathbf{x} - \mathbf{x}_i) = \frac{1}{n} \sum_{i=1}^n c k \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right), \quad (8)$$

где $\mathbf{x}_i$ – вектора признаков точек распределения; $i = 1, \dots, n$; K – ядро распределения; k – ядерная функция (или парзеновское окно); h – размер парзеновского окна.

Градиент функции $f(\mathbf{x})$ имеет следующий вид:

$$\begin{aligned} \nabla f(\mathbf{x}) &= \frac{1}{n} \sum_{i=1}^n \nabla K(\mathbf{x} - \mathbf{x}_i) = \\ &= \frac{c}{n} \left[\sum_{i=1}^n g \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right) \right] \left[\frac{\sum_{i=1}^n \mathbf{x}_i g \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right)}{\sum_{i=1}^n g \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right)} - \mathbf{x} \right], \quad (9) \end{aligned}$$

где $g(r) = -k'(r)$ – функция первой производной ядерной функции.

Выражение (9) в качестве множителя содержит вектор сдвига среднего с ядром

$$\begin{aligned} G(\mathbf{x}) &= c g \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right) : \\ \mathbf{m}(\mathbf{x}) &= \frac{\sum_{i=1}^n \mathbf{x}_i g \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right)}{\sum_{i=1}^n g \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right)} - \mathbf{x}. \quad (10) \end{aligned}$$

Из выражения (10) видно, что сдвиг среднего представляет собой нормализованный градиент плотности, задаваемой ядром G .

В общем виде алгоритм сдвига среднего можно сформулировать следующим образом:

1. Определить положение и характеристики центра масс области с заданным объектом на первом кадре последовательности;
2. Для каждого последующего кадра:

2.1. Вычислить вектор сдвига среднего относительно предыдущего центра масс;

2.2. Переместить центр масс в соответствии в новое положение;

2.3. Если центр масс не изменился, то перейти к анализу следующего кадра, иначе – перейти к шагу 2.1.

Повысить устойчивость алгоритма можно за счет уточнения положения предполагаемой области изображения с сопровождаемым объектом на основе сравнения гистограмм используемых характеристик. Например, в литературе описаны алгоритмы, базирующиеся на применении расстояния Бхатачария.

В качестве модификации указанных алгоритмов можно предложить использование гистограмм кодов, полученных с помощью оператора локальных бинарных шаблонов.

Сопровождение объектов на изображении можно рассматривать как задачу теории управления, которая заключается в оценке состояния системы на основе последовательности зашумленных измерений. В этом случае строится модель наблюдения и модель состояния. При обработке кадров видеоизображения последовательно выполняется предсказание и коррекция данных моделей [4].

Одними из часто применяемых алгоритмов данного класса являются алгоритмы, базирующиеся на фильтре Кальмана [4, 7]. Они строятся в предположении, что движение является линейным, а шумы описываются функцией гауссова распределения с нулевым математическим ожиданием. Однако на практике данные условия часто не выполняются. Поэтому алгоритмы сопровождения на основе фильтра Кальмана могут давать значительные ошибки.

Более общим подходом к решению задач сопровождения являются фильтры частиц, основанные на методах Монте-Карло [4, 8, 9].

Данный подход предполагает создание и анализ набора частиц (выборок). Каждая частица предсказывает положение отслеживаемого объекта. Действительным положением объекта на кадре считается математическое ожидание положений, задаваемых частицами.

Пусть $\{\mathbf{x}_k^i\}_{i=1}^N$ – множество положений центров частиц на k -м кадре видеоизображения, а $\{\mathbf{z}_k\}$ – множество наблюдений, сделанных на кадрах видеоизображения. Тогда условная плотность распределения вероятности положения объекта на k -м кадре в зависимости от сделанных до этого наблюдений может быть аппроксимирована на основе метода Монте-Карло следующим образом:

$$p(\mathbf{x}_k | \mathbf{z}_{1:k}) \approx \sum_{i=1}^N w_k^i \delta(\mathbf{x}_k - \mathbf{x}_k^i), \quad (11)$$

где $\{w_k^i\}_{i=1}^N$ – множество весов частиц на k -м кадре, δ – дельта-функция Дирака, $\mathbf{z}_{1:k} = \mathbf{z}_1 \mathbf{z}_2 \dots \mathbf{z}_k$.

Весы, связанные с частицами, определяются по формуле:

$$w_k^i = \frac{p(\mathbf{x}_k | \mathbf{z}_{1:k})}{q(\mathbf{x}_k | \mathbf{z}_{1:k})}, \quad (12)$$

где q – функция важности, обеспечивающая генерацию множества частиц. Например, в алгоритме воспроизведения условной плотности CONDENSATION используется функция важности:

$$q(\mathbf{x}_k | \mathbf{z}_{1:k}) = p(\mathbf{x}_k | \mathbf{z}_{1:k-1}), \quad (13)$$

которая позволяет определить веса следующим образом:

$$w_k^i = p(\mathbf{z}_k | \mathbf{x}_k^i). \quad (14)$$

Для вычисления условной плотности $p(\mathbf{z}_k | \mathbf{x}_k^i)$ при расчете весов по формуле (14) часто применяются функции, основанные на расстоянии между характеристиками области на кадре, соответствующей частице, и характеристиками области объекта. Например, часто используется гауссова функция.

Вычислив веса и определив положение центров частиц, можно вычислить оценку положения объекта на k -м кадре:

$$\hat{\mathbf{x}}_k \approx \sum_{i=1}^N w_k^i \mathbf{x}_k^i. \quad (15)$$

Двумя важными этапами алгоритмов на основе фильтра частиц являются:

- 1) начальные распределения положений частиц и их весов,
- 2) создание нового множества частиц на основе текущего.

Например, начальное распределение положений частиц можно задать как смещенное случайным образом положение объекта на первом кадре, а начальные веса сделать равными $\frac{1}{N}$, где

N – число частиц. Для создания нового множества частиц вычисляется оценка значимости частиц текущего множества, удаляются незначимые частицы, а вместо них формируются новые, исходя из положения объекта на текущем кадре.

Отметим также, что часто применяются техники, основанные на уточнении значений весов.

Для повышения эффективности алгоритмов на основе фильтров частиц в настоящее время разрабатываются их модификации, использующие гистограммные характеристики и процедуры сдвига среднего [9].

Достоинством фильтров частиц является возможность естественного распараллеливания, что позволяет значительно повысить скорость отслеживания объектов на изображениях.

Последний этап обобщенной схемы обработки видеоизображений в системах обнаружения и сопровождения объектов заключается в анализе траектории движения объекта. На данном этапе определяются характеристики движения высокого уровня, например, связанные с возможностью выезда на полосу встречного движения или столкновения с находящимся впереди транспортным средством. Точность и оперативность решения задач на этом этапе в значительной степени определяются алгоритмами, описанными выше.

Таким образом, применение алгоритмов анализа текстурных характеристик на основе локальных бинарных шаблонов, обнаружения и сопровождения объектов с помощью алгоритмов сдвига среднего, фильтров частиц и их комбинации позволяет обеспечить эффективное принятие решений в системах управления мобильными объектами.

Литература

1. Техническое зрение в системах управления мобильными объектами-2010: Труды научно-технической конференции-семинара. Вып. 4 / Под ред. Р.Р.Назирова. – М.: КДУ, 2011.
2. Методы автоматического обнаружения и сопровождения объектов. Обработка изображений и управление / Б.А.Алпатов, П.В.Бабаян, О.Е.Балашов, А.И.Степашкин. – М.: Радиотехника, 2008.
3. Аун С., Шарнин Л.М., Кирпичников А.П. Информационно-измерительная система слежения за движущимися объектами // Вестник Казан. технол. ун-та. 2011. – Т. 14. №16. – С. 224-232.
4. Разработка мультимедийных приложений с использованием библиотек OpenCV и IPP [Электронный ресурс] //Национальный Открытый Университет «ИНТУИТ» [Офиц. сайт]. URL: <http://www.intuit.ru/studies/courses/10622/1106/info>.
5. Matti Pietikinen, Abdenour Hadid, Guoying Zhao, Timo Ahonen Computer Vision Using Local Binary Patterns. – Springer; 2011.
6. Gary Bradsky, Adrian Kaehler Learning OpenCV Computer Vision with the OpenCV Library – O'Reilly, 2008.
7. Форсайт Д.А., Понс Ж. Компьютерное зрение. Современный подход: Пер. с англ. – М.: Издательский дом «Вильямс», 2004.
8. Мокшин В.В., Кирпичников А.П., Шарнин Л.М. Отслеживание объектов в видеопотоке по значимым признакам на основе фильтрации частиц // Вестник Казан. технол. ун-та. 2013. – Т. 16. №18. – С. 297-303.
9. Emilio Maggio, Andrea Cavallaro Video tracking: theory and practice / Andrea Cavallaro, Emilio Maggio. - John Wiley & Sons, Ltd.

© А. П. Кирпичников – д. ф.-м. н., зав. каф. интеллектуальных систем и управления информационными ресурсами КНИТУ, kirpichnikov@kstu.ru; С. А. Ляшева - к.т.н., доц. каф. прикладной математики и информатики КНИТУ-КАИ; М. П. Шлеймович - к.т.н., доц. каф. автоматизированных систем обработки информации и управления КНИТУ-КАИ, shlch@mail.ru.

© А. P. Kirpichnikov - Dr. Sci, Head of the Department of Intelligent Systems and Information Systems Control, KNRTU, kirpichnikov@kstu.ru; S. A. Lyasheva - PhD, Associate Professor of the Department of Applied Mathematics & Informatics, KNRTU-KAI; M. P. Shleymovich - PhD, Associate Professor of the Department of Automated Information Processing Systems & Control, KNRTU-KAI, shlch@mail.ru.