

УДК. 621.391

МЕТОД ВЫДЕЛЕНИЯ ДВИЖУЩИХСЯ ОБЪЕКТОВ В ВИДЕОПОТОКЕ И ОЦЕНКА ТОЧНОСТИ ОПРЕДЕЛЕНИЯ ИХ КООРДИНАТ

Медведева Е.В., д.т.н., профессор кафедры радиоэлектронных средств Вятского государственного университета, e-mail: emedv@mail.ru;

Карлушин К.А., аспирант Вятского государственного университета, e-mail: konstantin.karlushin@gmail.com.

Ключевые слова: выделение движущихся объектов, оценка точности определения координат, изображения, многомерные цепи Маркова.

Введение

В некоторых системах видеонаблюдения необходимо выделить движущиеся объекты на сравнительно неподвижном фоне. Как правило, такие системы должны с большой вероятностью обнаруживать объекты интереса. При этом диапазон варьирования размеров объектов интереса может быть большим, а количество движущихся объектов – априорно неизвестным. Кроме того, обработка данных должна проводиться в реальном масштабе времени, что неизбежно накладывает ограничения на сложность используемых алгоритмов.

Для решения задачи выделения объектов существует много различных методов (разностный, корреляционный, на основе байесовской сегментации и др.) [1–4], но одни из них могут приводить к существенным погрешностям в оценке координат, а другие требуют относительно большого объема вычислительных ресурсов, особенно, когда в видеокадре необходимо выделить несколько объектов.

Поэтому разработка эффективных методов выделения объектов движения в видеопотоке, позволяющих в реальном масштабе времени выделять объекты интереса с высокой точностью, является актуальной задачей.

Поскольку последовательности видеокадров представляют собой трехмерные сигналы, для выделения движущихся объектов необходимо использовать пространственно-временную математическую модель изображения. Пространственно-временная модель должна быть адекватна реальным последовательностям изображений и обладать высокой вычислительной эффективностью, чтобы алгоритмы обработки изображений, синтезируемые на ее основе, требовали небольших вычислительных затрат.

Такой математической моделью является модель на основе трехмерных цепей Маркова.

Предполагается, что цифровые полутоновые изображения (ЦПИ) в видеопоследовательности представлены g -разрядными двоичными числами. Это позволяет разбить последовательность ЦПИ на последовательности g разрядных двоичных изображений (РДИ), каждое из которых является суперпозицией трех простых цепей Маркова с двумя равновероятными дискретными состояниями $M_1^{(l)}$ и $M_2^{(l)}$ и матрицами вероятностей переходов по горизонтали

${}^1\Pi = \|\pi_{ij}^{(l)}\|_{2 \times 2}$, вертикали ${}^2\Pi = \|\pi_{ij}^{(l)}\|_{2 \times 2}$ и времени

Предложен метод выделения движущихся объектов в видеопотоке. Метод основан на представлении последовательности изображений трехмерной цепью Маркова и выделении контуров движущихся объектов по вычисленной величине количества информации в элементах изображения. Проведена оценка точности определения координат движущихся объектов. Разработанный метод требует небольших вычислительных ресурсов, что делает возможным его применение при обработке данных в реальном масштабе времени.

$${}^4\Pi = \|\pi_{ij}^{(l)}\|_{2 \times 2} \quad (l = \overline{1, g}, i, j = \overline{1, 2}).$$

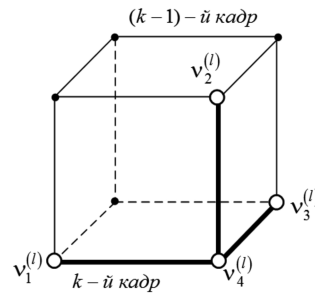


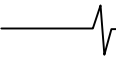
Рис. 1. Фрагмент последовательности РДИ

На приведенном фрагменте последовательности РДИ (рис. 1), элемент изображения $v_4^{(l)}$ зависит от состояний трех окрестных элементов $\Lambda_{i,j,k}^{(l)} = \{v_1^{(l)}, v_2^{(l)}, v_3^{(l)}\}$, расположенных в k -м и $(k-1)$ -м кадрах.

Количество информации в элементе $v_4^{(l)}$ относительно состояний элементов окрестности $\Lambda_{i,j,k}^{(l)} = \{v_1^{(l)}, v_2^{(l)}, v_3^{(l)}\}$ можно определить как разность собственной информации в элементе $v_4^{(l)}$ и взаимной информации, полученной от элементов окрестности $\Lambda_{i,j,k}^{(l)}$ [5–7]:

$$I(v_4^{(l)} | v_1^{(l)}, v_2^{(l)}, v_3^{(l)}) = I(v_4^{(l)}) - I(v_1^{(l)}, v_2^{(l)}, v_3^{(l)}, v_4^{(l)}) = - \left[\log p(v_4^{(l)}) + \log \frac{\prod_4 p(v_i^{(l)}, v_j^{(l)}) p(v_1^{(l)}, v_2^{(l)}, v_4^{(l)}, v_4^{(l)})}{\prod_{i=1}^4 p(v_i^{(l)}) \prod p(v_i^{(l)}, v_j^{(l)}, v_k^{(l)})} \right] = - \log \frac{w(v_4^{(l)} | v_1^{(l)}) w(v_4^{(l)} | v_2^{(l)}) w(v_4^{(l)} | v_3^{(l)}) w(v_4^{(l)} | v_1^{(l)}, v_2^{(l)}, v_3^{(l)})}{w(v_4^{(l)} | v_1^{(l)}, v_2^{(l)}) w(v_4^{(l)} | v_1^{(l)}, v_3^{(l)}) w(v_4^{(l)} | v_2^{(l)}, v_3^{(l)})}, \quad (1)$$

где произведения, обозначенные символом \prod вычисляются по всем возможным несовпадающим комбинациям различных подстрочных индексов трехмерного слу-



чайного поля; $p(v_1^{(l)}, v_2^{(l)}, v_4^{(l)}, v_4^{(l)})$, $p(v_i^{(l)}, v_j^{(l)}, v_k^{(l)})$, $p(v_i^{(l)}, v_j^{(l)})$, $i = j = k = \overline{1,4}$; $i \neq j \neq k$ - совместные плотности вероятностей значений элементов; $p(v_i^{(l)})$, $i = \overline{1,4}$ - априорные плотности вероятностей значений элементов; $w(v_4^{(l)} | v_i^{(l)})$, $i = \overline{1,3}$ - одномерные плотности вероятностей переходов; $w(v_4^{(l)} | v_i^{(l)}, v_j^{(l)})$, $i = j = \overline{1,3}$; $i \neq j$ - плотности вероятностей переходов в двумерной цепи Маркова; $w(v_4^{(l)} | v_1^{(l)}, v_2^{(l)}, v_3^{(l)})$ - плотность вероятности переходов в трехмерной цепи Маркова.

Плотность вероятности переходов в трехмерной двоячной цепи Маркова $w(v_4^{(l)} | \Lambda_{i,j,k}^{(l)})$ можно представить в форме

$$w(v_4^{(l)} | \Lambda_{i,j,k}^{(l)}) = \sum_{i, \dots, r=1}^2 \pi \left(v_4^{(l)} = M_i^{(l)} | v_1^{(l)} = M_j^{(l)}; v_2^{(l)} = M_k^{(l)}; v_3^{(l)} = M_r^{(l)} \right) \times \delta(v_1^{(l)} - M_j^{(l)}) \delta(v_2^{(l)} - M_k^{(l)}) \delta(v_3^{(l)} - M_r^{(l)}), \quad (2)$$

где $\delta(\cdot)$ - дельта функция.

В трехмерной математической модели вероятности смены состояний элемента $v_4^{(l)}$ зависят от различных сочетаний состояний элементов окрестности $\Lambda_{i,j,k}^{(l)}$ и их корреляции с элементом $v_4^{(l)}$. Все возможные значения этих вероятностей для различных комбинаций состояний элементов окрестности $\Lambda_{i,j,k}^{(l)}$ могут быть вычислены в соответствии с аргументом выражения (1) и сведены в матрицу Π вида:

$$\Pi = \begin{pmatrix} \pi_{iii}^{(l)} & \pi_{ijj}^{(l)} \\ \pi_{iji}^{(l)} & \pi_{ijj}^{(l)} \\ \vdots & \vdots \\ \pi_{jji}^{(l)} & \pi_{jjj}^{(l)} \end{pmatrix} = \begin{pmatrix} \alpha_1^{(l)} & \alpha_1^{(l)} \\ \alpha_2^{(l)} & \alpha_2^{(l)} \\ \vdots & \vdots \\ \alpha_8^{(l)} & \alpha_8^{(l)} \end{pmatrix}, \quad i, j = \overline{1,2}; \quad i \neq j, \quad (3)$$

элементы которой симметричны и удовлетворяют условию нормировки $\alpha_q + \alpha_q^{(l)} = 1$.

Например, выражения для вычисления элементов первой строки матрицы Π , имеют вид:

$$\alpha_1^{(l)} = \pi_{iii}^{(l)} = \frac{1 \pi_{ii}^{(l)} \cdot 2 \pi_{ii}^{(l)} \cdot 4 \pi_{ii}^{(l)} \cdot 7 \pi_{ii}^{(l)}}{3 \pi_{ii}^{(l)} \cdot 5 \pi_{ii}^{(l)} \cdot 6 \pi_{ii}^{(l)}}, \quad (4)$$

$$\alpha_1^{(l)} = \pi_{ijj}^{(l)} = \frac{1 \pi_{ij}^{(l)} \cdot 2 \pi_{ij}^{(l)} \cdot 4 \pi_{ij}^{(l)} \cdot 7 \pi_{ii}^{(l)}}{3 \pi_{ii}^{(l)} \cdot 5 \pi_{ii}^{(l)} \cdot 6 \pi_{ii}^{(l)}},$$

где ${}^r \pi_{ii}^{(l)}$ ($r = \overline{1,7}$) - элементы матриц вероятностей переходов в одномерных цепях Маркова с двумя состояниями по горизонтали ${}^1 \Pi^{(l)}$, вертикали ${}^2 \Pi^{(l)}$, времени ${}^4 \Pi^{(l)}$ и четырех сопутствующих матриц вероятностей переходов:

$$\begin{aligned} {}^3 \Pi^{(l)} &= {}^1 \Pi^{(l)} \times {}^2 \Pi^{(l)}; \\ {}^5 \Pi^{(l)} &= {}^1 \Pi^{(l)} \times {}^4 \Pi^{(l)}; \\ {}^6 \Pi^{(l)} &= {}^2 \Pi^{(l)} \times {}^4 \Pi^{(l)}; \\ {}^7 \Pi^{(l)} &= {}^3 \Pi^{(l)} \times {}^4 \Pi^{(l)}. \end{aligned} \quad (5)$$

Вычисление остальных элементов матрицы $\Pi^{(l)}$ осуществляется в соответствии с состоянием элементов окрестности $\Lambda_{i,j,k}^{(l)}$.

На рис. 2 показаны зависимости вероятностей переходов между состояниями элементов в последовательности РДИ по горизонтали, вертикали и между кадрами от номера разряда ЦПИ. Графики получены усреднением статистических характеристик по большому количеству реальных изображений, полученных стационарной видеокамерой, аналогичных рис. 3.



Рис. 2. Зависимости вероятностей переходов между состояниями элементов в последовательности РДИ (8 - старший разряд; 1 - младший разряд)

Из приведенного графика следует, что наибольшей корреляцией обладают элементы, расположенные в старших разрядах ЦПИ. Поэтому основные детальные области можно выделить по РДИ старших разрядов ЦПИ. А РДИ младших разрядов (при $\pi_{ii} \rightarrow 0,5$) будут составлять фон изображения в виде двумерного шума.

Для выделения движущихся объектов в каждом элементе РДИ в соответствии с состояниями элементов окрестности вычисляют величину количества информации (1). Для видеопоследовательностей, полученных стационарной видеокамерой при неизменном уровне освещенности, будет характерно незначительное изменение вероятностей переходов между элементами изображения от кадра к кадру в течение длительного промежутка времени. Поэтому предполагалось, что вероятности переходов между элементами видеопоследовательности постоянны и априорно известны, причем корреляция элементов между кадрами (${}^4 \pi_{ii}$) больше, чем внутри кадра (${}^1 \pi_{ii}$, ${}^2 \pi_{ii}$) (рис. 2).

При отсутствии объекта (контрастной области), состояния окрестных элементов $\Lambda_{i,j,k}^{(l)}$ будут совпадать с $v_4^{(l)}$, и количество информации в элементе $v_4^{(l)}$ будет минимально.

При наличии неподвижного объекта, на его границе, состояние элемента $v_4^{(l)}$ будет совпадать с состоянием элемента $v_4^{(l)}$ в $(k-1)$ -м кадре и не совпадать с одним или двумя окрестными элементами в k -м кадре, и количество информации в $v_4^{(l)}$ возрастет.

При смещении объекта элемент $v_4^{(l)}$ будет иметь различные состояния с элементом $v_3^{(l)}$ в $(k-1)$ -м кадре и с одним или двумя окрестными элементами в k -м кадре,

и количество информации в элементе $v_4^{(l)}$ еще более увеличится. Причем, чем больше корреляция между соседними кадрами, что характерно для видеопоследовательностей, полученных неподвижным датчиком, тем больше информации будет в элементе $v_4^{(l)}$ при смещении объектов.

Сравнивая значения вычисленной величины количества информации в элементе изображения с порогом, определяем принадлежность данной точки контуру. Значение порога h определяют по вычисленной величине количества информации в элементе $v_4^{(l)}$ l -го РДИ:

$$h = \frac{I(v_4^{(l)} = M_i^{(l)} | v_1^{(l)} = M_i^{(l)}, v_2^{(l)} = M_j^{(l)}, v_3^{(l)} = M_j^{(l)})}{2} + \frac{I(v_4^{(l)} = M_i^{(l)} | v_1^{(l)} = M_j^{(l)}, v_2^{(l)} = M_j^{(l)}, v_3^{(l)} = M_i^{(l)})}{2} \quad (6)$$

Предложенный метод определения контуров движущихся объектов требует небольших вычислительных ресурсов, определяемых операциями сравнения с тремя окрестными элементами.

Результатом такой обработки изображений является совокупность контурных точек, в которых происходит движение одного или нескольких объектов интереса, а также точек, представляющих собой шум, вызванный, например, движением листьев на деревьях, бликами и т.п. Следует отметить, что контурные точки, принадлежащие объекту интереса, являются связными и имеют большую протяженность, чем контурные точки, представляющие собой шум.

При решении задачи выделения объекта интереса необходимо выделить точки, которые наилучшим образом представляют объект. Для этого изображение разбивается на блоки и одновременно с определением контурных точек выполняется анализ указанных блоков на наличие в них связанных пикселей, составляющих часть предполагаемого контура движущего объекта. Размер блока выбирается в соответствии с данными о максимальной скорости движения и минимальной площади

объекта. Решение о присутствии объекта интереса на изображении принимается по наличию связанных пикселей в блоках.

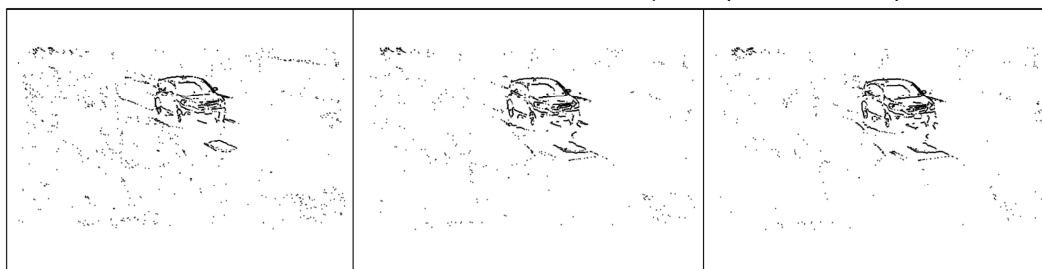
Координаты движущегося объекта соответствуют центру прямоугольного окна, в которое вписывается объект интереса. Последовательность координат движущихся объектов будет представлять собой траекторию движения. При выполнении исследования эффективности разработанного метода использовались несколько видеопоследовательностей средней длительностью 300 кадров.

Для оценки качества выделения движущегося объекта вычислялась вероятность правильного выделения объекта интереса и среднеквадратическая ошибка (СКО) определения координат объекта интереса. На рис. 3 показан фрагмент видеопоследовательности, в котором три соседних кадра совмещены для визуальной оценки передвижения объекта. Ниже на рис. 4 приводятся контурные (а) и выделенные прямоугольным окном объекты (б).



Рис. 3. Совмещенные кадры видеопоследовательности

В результате выделения объектов интереса возникают ошибки двух типов: на контурном изображении объект интереса определен (как совокупность точек шума), в то время как на реальном изображении он отсутствует; на контурном изображении объект не выделен (например, в результате частичного заслонения другим объектом), а на реальном изображении он имеется.



а) Контурные изображения



б) Изображения с выделенными объектами

Рис. 4. Кадры видеопоследовательности с выделенными объектами

Вероятность правильного выделения объекта интереса определялась как отношение количества верно выделенных объектов (N_B) к общему количеству движущихся объектов (N_O) в видеопоследовательности:

$$P_{np} = \frac{N_B}{N_O} \quad (7)$$

Для приведенной видеопоследовательности общая вероятность правильного определения объекта интереса составила 0,75, что вызвано скрытием объекта интереса препятствием. Размеры объектов интереса варьировались от 30x30 до 120x120 пикселей. Количество одновременно присутствующих объектов в кадре менялось на протяжении всей видеопоследовательности, но не превышало двух. Зависимость вероятности правильного выделения объекта в тестовой видеопоследовательности от номера кадра представлена на рис. 5.

Кадры, в которых вероятность правильного выделения объекта интереса равна нулю, соответствуют ситуации с заслонением объекта интереса. Для других видеопоследовательностей, в которых не происходит скрытие объектов интереса, величина правильного выделения достигает 0,91±0,95.

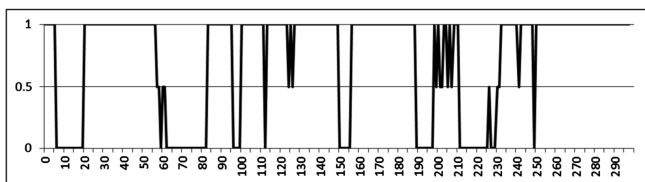


Рис. 5. Зависимость вероятности правильного выделения объекта от номера кадра

Также для оценки качества работы алгоритма вычислялась нормированная среднеквадратическая ошибка (СКО) оценки координат движущегося объекта:

$$СКО = \frac{1}{L} \sqrt{\sigma_x^2 + \sigma_y^2}, \quad (8)$$

где σ_x , σ_y — смещение координат центра движущегося объекта по горизонтали и вертикали от их истинного значения, L — половина диагонали прямоугольного окна, обрамляющего объект интереса в идеальном случае.

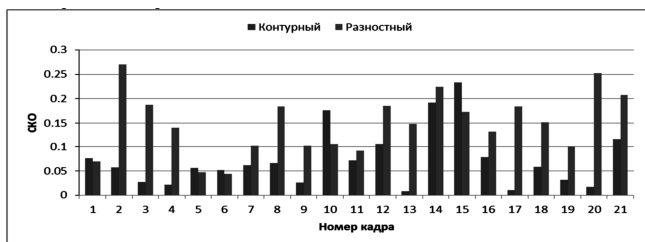


Рис. 6. Сравнение СКО определения координат движущихся объектов разными методами

При выполнении исследования сравнивались результаты СКО оценки координат разработанным и известным разностным методом [1, 3]. Результаты сравнения СКО двумя методами на нескольких кадрах последовательности (рис. 3) представлены на рис. 6. При оценке всей видеопоследовательности из 300 кадров СКО разработанного контурного метода меньше разностного в 2 раза. Для большинства видеопоследователь-

ностей выигрыш в СКО для разработанного метода по сравнению с разностным составляет 1,5÷2,5.

Из приведенных результатов следует, что предложенный метод выделения движущихся объектов в видеопотоке, основанный на вычислении количества информации в элементах изображения, позволяет определить координаты движущегося объекта с точностью, превышающей разностный метод. Разработанный метод требует небольших вычислительных ресурсов, что делает возможным его применение при обработке данных в реальном масштабе времени.

Литература

1. Яне Б. Цифровая обработка изображений. — М.: Техносфера, 2007. — 584 с.
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. — 176 с.
3. Богословский А.В. Обработка многомерных сигналов. Линейная многомерная дискретная обработка сигналов. Методы анализа и синтеза. / А.В. Богословский, Е.А. Богословский, И.В. Жигулина, В.А. Яковлев. — М.: Радиотехника, 2013. — 168 с.
4. Трифонов А.П. Обработка динамических изображений. Обнаружение и оценка параметров движения. / А.П. Трифонов, Р.В. Куцов — LAP LAMBER Academic Publishing, Germany, 2011. — 156 с.
5. Петров Е.П. Метод синтеза математических моделей видеоизображений на основе многомерных цепей Маркова / Е.П. Петров, Е. В. Медведева, А.П. Метелев // Нелинейный мир, № 4, 2011. — С.213-231.
6. Карлушин К.А. Метод выделения движущихся объектов в видеоизображениях на основе трехмерных цепей Маркова / К.А. Карлушин, Е. В. Медведева // Т-Comm. Телекоммуникации и транспорт. № 9, 2013. — С.94-97.
7. Карлушин К.А. Метод выделения движущихся объектов на основе пространственно-временной модели изображения / К.А. Карлушин, Е.В. Медведева // Труды 69-й межд. конф., посвященной Дню радио «Радиоэлектронные устройства и системы для инфокоммуникационных технологий»: — М.: 2014 — С.136-140.

A METHOD OF MOVING OBJECTS DETECTION IN VIDEO STREAM AND ESTIMATION ACCURACY OF OBJECTS COORDINATES DETERMINATION

Medvedeva E.V., Karlushin K.A.

A method of moving objects detection in video stream is offered. The method is based on the representation of a sequence of images by three-dimensional Markov chain and moving objects contours extraction by calculating the magnitude of information amount in the image elements. The accuracy of moving objects coordinates determination is estimated. The developed method requires low computational resources, which makes possible its use in the data processing in real time.