

ПРОСТРАНСТВЕННЫЙ АЛГОРИТМ ВЫДЕЛЕНИЯ ОБЪЕКТОВ НА ОСНОВЕ АДАПТИВНОЙ ПОРОГОВОЙ ОБРАБОТКИ

Алпатов Б.А., Муравьев В.С., Муравьев С.И.

Рязанская государственная радиотехническая академия

Во многих областях военной и гражданской направленности находят применение видеоинформационные системы, предназначенные для обнаружения и сопровождения движущихся объектов, которыми могут являться автомобили, корабли, вертолеты и т.п. Для успешного решения этой задачи необходимо максимально использовать априорную информацию об объекте и фоне. Эта информация может содержать параметры движения объектов, их яркостные характеристики и размеры. Существенное влияние на выбор алгоритма обработки также оказывает тип фона.

Всё многообразие методов, используемых для обнаружения и выделения объектов, можно разделить на пространственные и временные. Пространственная обработка базируется на информации, заложенной в одном телевизионном кадре, в то время как временная обработка учитывает корреляционные связи между выборками сигналов в одной и той же точке изображения в различных кадрах.

Одной из актуальных задач является обнаружение воздушных летательных аппаратов наземными средствами. Рассматриваемая задача обладает следующими основными особенностями:

- широкий диапазон размеров обнаруживаемых объектов (от 2x2 до 64x64)
- относительно ровный фон, без резких яркостных перепадов, примером которого может служить небо с равномерной облачностью;
- большая подвижность фона из-за высокой относительной скорости носителя видеодатчика и объекта обнаружения;
- возможность реализации алгоритма обнаружения в реальном масштабе времени на существующей элементной базе (ПЛИС, ЦПОС).

В качестве модели наблюдаемого изображения была использована модель заслона (1)

$$l(i, j, n) = h(i, j, n)r(i, j, n) + g(i, j, n)(1 - r(i, j, n)) + \xi(i, j, n), \quad i = \overline{1, N_x}, j = \overline{1, N_y}, \quad (1)$$

где n – номер кадра, N_x, N_y – ширина и высота кадра, $l(i, j, n)$ – наблюдаемое изображение, $g(i, j, n)$ и $h(i, j, n)$ – неизвестные яркости точек фона и объектов соответственно в кадре с номером n , $\xi(i, j, n)$ – гауссовый шум с нулевым средним. Функция $r(i, j, n)$ задаёт расположение объектов на изображении и задаётся правилом:

$$r(i, j, n) = \begin{cases} 1, & \text{если в точке } (i, j) \text{ кадра } n \text{ находится объект,} \\ 0, & \text{в противном случае} \end{cases} \quad (2)$$

Задача выделения объектов на изображении заключается в нахождении оценки $\hat{r}(i, j)$ на основе наблюдаемого изображения $l(i, j)$.

В процессе анализа поставленной задачи был разработан алгоритм выделения объектов, использующий пространственный подход к обработке последовательности изображений. Функциональная схема алгоритма, представлена на рисунке 1.

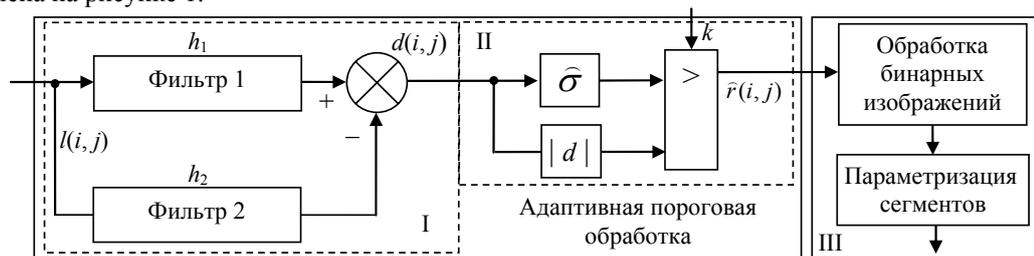


Рис. 1. Функциональная схема пространственного алгоритма выделения объектов

В блоке I осуществляется процедура компенсации фона, широко освещенная в литературе [1, 2], которая была несколько видоизменена.

Для выполнения оценки фона, исходное изображение $l(i, j)$ пропускается через фильтр 2, имеющий маску h_2 большой размерности. Параллельно $l(i, j)$ сглаживается маской h_1 меньшего размера, после чего из результата фильтрации вычитается оценка фонового изображения. Подобная структура блока I позволяет распараллелить операции свертки. Такой подход увеличит вычислительную эффективность алгоритма, которая при аппаратной реализации на ПЛИС во многом определяется числом операций обращения к памяти.

При большом диапазоне варьирования размеров объектов можно использовать линейные усредняющие фильтры, которые могут быть вычислительно эффективно реализованы. Изображения, полученные на выходе фильтров 1 и 2, обозначенные $f_1(i, j)$ и $f_2(i, j)$, можно описать выражениями (3), (4)

$$f_1(i, j) = \sum_{m_x=-q_1}^{q_1} \sum_{m_y=-q_1}^{q_1} h_1(m_x, m_y)l(i - m_x, j - m_y), \quad (3)$$

$$f_2(i, j) = \sum_{m'_x=-q_2}^{q_2} \sum_{m'_y=-q_2}^{q_2} h_2(m'_x, m'_y)f(i - m'_x, j - m'_y), \quad (4)$$

h_1 и h_2 – маски фильтров размерностями $(2q_1 + 1) \times (2q_1 + 1)$ и $(2q_2 + 1) \times (2q_2 + 1)$ соответственно, причем $q_1 < q_2$, N_x, N_y – ширина и высота кадра. Предложено использовать маски следующего вида:

$$h_1(m_x, m_y) = 1/(2q_1 + 1)^2, \quad m_x, m_y = \overline{-q_1, q_1};$$

$$h_2(m'_x, m'_y) = \begin{cases} 0, & \text{при } m'_x, m'_y = \overline{-q_1, q_1} \\ 1/\sqrt{(2q_2 + 1)^2 - (2q_1 + 1)^2}, & \text{иначе.} \end{cases} \quad (5)$$

После фильтрации вычисляется разностное изображение:

$$d(i, j) = f_1(i, j) - f_2(i, j), \quad i = \overline{1, N_x}, j = \overline{1, N_y}, \quad (6)$$

Введение дополнительного НЧ фильтра (фильтра 1) призвано усреднить значение анализируемого пикселя со своими ближайшими соседями до операции вычитания фона, что позволяет более точно оценить яркость объекта и снизить влияние шумов. Теоретически, эта идея весьма близка оптимальному корреляционному приемнику с заданным размером объекта. Наличие нулевых значений в центральной части маски h_2 позволяет более точно оценить фоновую составляющую, что в конечном итоге благоприятно сказывается на точности выделения объектов.

В блоке II происходит вычисление адаптивного порога на основе оценки дисперсии шума и остаточного фона по всему кадру (7), которое сводится к нахождению второго начального момента [3].

$$\hat{\sigma}^2 = \frac{1}{N_x \cdot N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} d^2(i, j) \quad (7)$$

После этого применяется решающее правило (6).

$$\hat{r}(i, j) = \begin{cases} 1, & |d(i, j)| > k\hat{\sigma}, \\ 0, & \text{иначе,} \end{cases} \quad (8)$$

где k – полуширина доверительного интервала.

Эксперименты показывают, что остаточный фон и шум, присутствующие на разностном изображении d многих видеосюжетов, имеют выраженное гауссово распределение [2, 3] с математическим ожиданием близким к нулю. Моды, порожденные объектами, дистанцированы и могут быть выделены согласно критерию $k \cdot \hat{\sigma}$.

Блок III осуществляет пост-обработку бинарных изображений, основанную на логической либо морфологической фильтрации, что приводит к удалению точечного шума и уменьшению фрагментации сегментов. В процессе параметризации сегментов измеряются координаты центров сегментов (i_u, j_u) , их габаритные размеры, площадь и средняя яркость.

Экспериментальная проверка алгоритма, в основном, производилась с использованием характерных естественных видеопоследовательностей продолжительностью 200-300 кадров, снятых в ТВ и ИК диапазонах.

Для оценки точности выделения объектов строились кривые зависимости вероятности правильного выделения P_{ns} от вероятности ложного выделения P_{ln} , называемые рабочими характеристиками выделения. Если $\hat{r}(i, j, n)$ – оценка бинарного изображения, полученная на n -ом кадре, а $r(i, j, n)$ – эталонная бинарная маска объекта, то оценки вероятностей ложного $\hat{P}_{\bar{e}a}$ и правильного $\hat{P}_{\bar{i}a}$ выделения можно найти по формулам (9):

$$\hat{P}_{\bar{e}a} = E \left[\frac{\sum_{i=1}^N \sum_{j=1}^M \hat{r}(i, j, n)(1 - r(i, j, n))}{\sum_{i=1}^N \sum_{j=1}^M (1 - r(i, j, n))} \right], \quad \hat{P}_{\bar{i}a} = E \left[\frac{\sum_{i=1}^N \sum_{j=1}^M \hat{r}(i, j, n)r(i, j, n)}{\sum_{i=1}^N \sum_{j=1}^M r(i, j, n)} \right]. \quad (9)$$

где E – символ математического ожидания. В качестве эталонного бралось изображение, созданное в графическом редакторе, усреднение проводилось по 6–8 кадрам. Варьируя параметры алгоритма, можно получить совокупность кривых. Чем выше будет лежать кривая, т.е., чем больше, \hat{P}_{ia} при фиксированном значении \hat{P}_{ea} , тем больше будет точность алгоритма. Таким образом, можно подобрать параметры алгоритма, дающие наилучшие результаты. Пример графиков кривых приведен на рисунке 2.

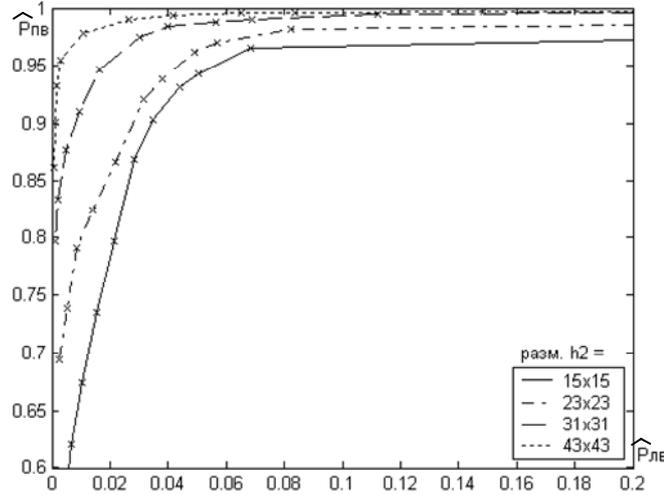


Рис. 2. Характеристики выделения при размерности маски $h_1 = 5 \times 5$

В процессе исследования работоспособности алгоритма было найдено предельное значение отношения сигнал/шум, которое для малоразмерных объектов (размерами от 2×2 до 8×8) составило приблизительно 4, для крупноразмерных – около 3. Под предельным понималось такое отношение сигнал/шум, при котором объект не обнаруживался в течение 5 кадров подряд. Также исследовались возможности применения медианной фильтрации и вычисление оценки дисперсии в скользящем окне, но полученные результаты были неудовлетворительными. Возможным способом улучшения качества выделения является построение двух и более каналов на базе разработанного алгоритма, но возникают вопросы, связанные с объединением результатов обработки в единое бинарное изображение.

Литература

1. Li Biao, Shen ZhenKang, Li JiCheng. Automatic target detection and tracking system using infrared imagery // Proc. of SPIE, Vol. 3069, Automatic Target Recognition VII, 1997, pp. 534 – 540.
2. Dennis Braunreiter, Nam Banh. Detection of moving subpixel targets in infrared clutter with space - time filtering // Proc. of SPIE, Vol. 1481, Signal and Data Processing of Small Targets. – 1991, pp. 73 – 83.
3. Алпатов Б.А., Муравьев В.С, Муравьев С.И. Алгоритм предварительного обнаружения объектов на основе адаптивной пороговой обработки // Сети и системы связи. Мат-лы всерос. научн. практ. сем-ра. – Рязань, РВВКУС, 2005.– С.195-197.

THE SPATIAL ALGORITHM FOR OBJECT EXTRACTION BASED ON ADAPTIVE THRESHOLDING

Alpatov B., Muraviev V., Muraviev S.

Ryazan state radioengineering academy

In this work we propose the spatial algorithm for object extraction in images, based on adaptive thresholding. The functional scheme of this algorithm is presented on figure 1.

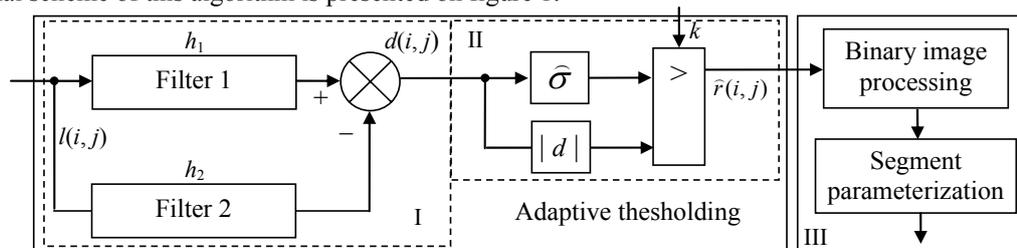


Fig. 1. The functional scheme of object extraction algorithm

At the first stage the background estimation is evaluated, for that an observed image $l(i, j)$ passes through the filter 2 with the high dimension mask h_2 . Simultaneously $l(i, j)$ is smoothed with h_1 of smaller size, after that the background estimation is subtracted from filter 1 output.

Masks h_1 and h_2 with dimensions $(2q_1+1) \times (2q_1+1)$ and $(2q_2+1) \times (2q_2+1)$ respectively, where $q_1 < q_2$, have the next appearance (1)

$$h_1(m_x, m_y) = 1 / (2q_1 + 1)^2, \quad m_x, m_y = \overline{-q_1, q_1};$$

$$h_2(m'_x, m'_y) = \begin{cases} 0, & \text{when } m'_x, m'_y = \overline{-q_1, q_1}, \\ 1 / \left((2q_2 + 1)^2 - (2q_1 + 1)^2 \right), & \text{otherwise.} \end{cases} \quad (1)$$

After obtaining the difference image $d(i, j)$ the variance estimate $\hat{\sigma}$ can be found, which is used for thresholding later. Logical or morphological processing can be performed over output binary image. The segment parameterization procedure is used for object coordinates measurement.

Basically, the algorithm has been tested on typical natural video sequences of 200 – 300 frames length, in both TV and IR subbands. For object extraction accuracy estimation in this work we have plotted the dependence of probability of object detection P_d from probability of false alarm P_{fa} , which is called segmentation operating curve (SOC). Measured minimal signal-to-noise ratio is about 4 for small objects (from 2x2 to 10x10 pixels) and 3 for large objects.

Much attention has been focused on real time working in modern hardware realizations (FPGA, DSP). First of all the proposed algorithm is designed for aircraft detection observed on background with uniform cloudiness.

References

1. Li Biao, Shen ZhenKang, Li JiCheng. Automatic target detection and tracking system using infrared imagery // Proc. of SPIE, Vol. 3069, Automatic Target Recognition VII, 1997, pp. 534 – 540.
2. Dennis Braunreiter, Nam Banh. Detection of moving subpixel targets in infrared clutter with space - time filtering // Proc. of SPIE, Vol. 1481, Signal and Data Processing of Small Targets. – 1991, pp. 73 – 83.
3. Алпатов Б.А., Муравьев В.С, Муравьев С.И. Алгоритм предварительного обнаружения объектов на основе адаптивной пороговой обработки // Сети и системы связи. Мат-лы всерос. научн. практ. сем-ра. – Рязань, РВВКУС, 2005.– С.195-197.

