

## ОЦЕНИВАНИЕ ПАРАМЕТРОВ СМЕЩЕНИЯ ИЗОБРАЖЕНИЯ ПРИ ВЫДЕЛЕНИИ ДВИЖУЩИХСЯ ОБЪЕКТОВ

Бабаян П.В., Стротов В.В.

Рязанская государственная радиотехническая академия

Автоматическое выделение объектов – это одна из важнейших задач анализа последовательностей изображений. Под выделением объектов понимается классификация точек изображения на точки, принадлежащие объекту, и точки, принадлежащие фону. Часто выделение объектов является промежуточным этапом обработки изображений на пути к дальнейшему обнаружению и распознаванию объектов.

Известно большое количество методов выделения объектов. Среди них можно выделить класс методов, основанных на пространственной обработке отдельных изображений [1], и класс методов, основанных на временной обработке последовательностей изображений [2]. Методы, относящиеся к первому классу, предназначены для выделения движущихся и неподвижных объектов, имеющих статистические, яркостные или текстурные характеристики, значительно отличающиеся от соответствующих характеристик фона. Второй класс методов главным образом ориентирован на выделение подвижных объектов и основан на выявлении таких яркостных изменений изображения во времени, причиной которых является движение объектов.

Существенным недостатком методов выделения объектов, основанных на временной обработке, в отличие от пространственных методов, является необходимость учёта геометрических деформаций фонового изображения. Геометрические деформации фонового изображения можно разделить на две группы. К первой группе относятся искажения, связанные с движением датчика в пространстве. В простейших случаях эти искажения являются параметрическими: сдвиг, поворот, евклидовы, аффинные, проективные преобразования. Ко второй группе относятся случайные геометрические искажения, вызванные, прежде всего, преломлением световых лучей в атмосфере.

Важными проблемами при реализации временных методов выделения объектов являются, во-первых, оценивание параметров геометрических преобразований, вызванных движением датчика, а, во-вторых, учёт случайных геометрических искажений фона. Второй проблеме посвящена, в частности, работа [2]. В настоящем докладе рассмотрена первая проблема – оценивание параметров геометрических преобразований изображения с целью дальнейшего выделения движущихся объектов. При этом мы ограничиваемся одним практически важным типом геометрических преобразований фонового изображения – смещением. Этот тип преобразований имеет место в двух основных случаях: во-первых, когда датчик изображения поворачивается на небольшой угол вокруг прямой, перпендикулярной к оптической оси датчика, и, во-вторых, когда датчик перемещается в плоскости, перпендикулярной к своей оптической оси.

Существует большое количество методов оценивания параметров геометрических преобразований изображений. Основными группами этих методов являются корреляционные, спектральные и структурные [3]. **Корреляционные методы** основаны на поиске точки максимума некоторой функции сходства, вычисленной при различных значениях параметров преобразования. При использовании **спектральных методов** оценки параметров геометрических преобразований каждое изображение предварительно преобразуется в частотную область, и дальнейшие вычисления производятся с использованием Фурье - спектров изображений. **Структурные методы** обычно включают в себя три этапа. На первом этапе производится поиск на изображениях характерных структурных элементов. В качестве таких элементов могут выступать точечные, линейные или площадные объекты, которые могут быть хорошо идентифицированы. На втором этапе устанавливаются типы взаимосвязей между структурными элементами различных изображений. На третьем этапе координаты структурных элементов и связи между ними используются для составления системы уравнений, решив которую, можно найти искомые параметры преобразований.

Перед тем как сформулировать решаемую в настоящей работе задачу оценивания параметров смещения, рассмотрим модель формирования наблюдаемого датчиком изображения. Пусть датчик формирует дискретную последовательность изображений  $l_n(x, y)$ , где  $n$  – номер кадра,  $(x, y) \in R^2$ . Пусть  $g(x, y)$  – неизменное во времени изображение сцены в отсутствие искажений и движущихся объектов. Пусть  $q_n(x, y)$  – изображение фона с присутствующими на нём объектами на кадре с номером  $n$ . Связь между  $g(x, y)$  и  $q_n(x, y)$  описывается моделью заслона:

$$q_n(x, y) = g(x, y)[1 - r_n(x, y)] + h_n(x, y)r_n(x, y), \quad (1)$$

где  $r_n(x, y)$  – бинарное изображение, единичные значения которого задают размещение объектов на кадре  $n$ ,  $h_n(x, y)$  – яркостное изображение всех движущихся объектов на кадре  $n$ .

Связь между  $q_n(x, y)$  и  $l_n(x, y)$  задаётся выражением:

$$l_n(x, y) = q_n(x - \alpha_n - u_n(x, y), y - \beta_n - v_n(x, y)) + \xi_n(x, y), \quad (2)$$

где  $(\alpha_n, \beta_n)$  – вектор смещения фонового изображения, который равен нулю в первоначальный момент наблюдения:  $\alpha_1 = 0, \beta_1 = 0$ ,  $(u_n(x, y), v_n(x, y))$  – векторное поле деформаций изображения, обусловленных случайными геометрическими искажениями,  $\xi_n(x, y)$  – аддитивный шум датчика. Задача оценивания параметров преобразования смещения состоит в нахождении оценок вектора  $(\alpha_n, \beta_n)$  на основе анализа последовательности наблюдаемых изображений  $l_k(x, y)$ ,  $k = \overline{1, n}$ . Особенностью задачи, по сравнению с классической задачей поиска взаимного смещения изображений, является присутствие в поле зрения объектов, которые двигаются не согласованно с движением фона. Заметим, что в рассматриваемой задаче суммарная площадь движущихся объектов невелика и составляет не более 5-10% площади кадра.

Одним из наиболее эффективных методов оценивания параметров смещения является корреляционный подход [4], который заключается в поиске точки максимума критериальной функции  $\rho(s_n(x, y), l_n(x - \alpha, y - \beta))$  при различных значениях  $(\alpha, \beta)$ . Здесь  $s_n(x, y)$  – эталонное изображение, сформированное к кадру с номером  $n$ . Существует несколько разновидностей функций  $\rho$ : разностные, корреляционные, фазовые критериальные функции. Известны различные стратегии формирования эталонного изображения  $s_n(x, y)$ , в частности, межкадровое сглаживание, полная замена эталонного изображения и другие.

Недостатком корреляционного подхода в рассматриваемых условиях наблюдения является большая вычислительная сложность. Поскольку размер обрабатываемых изображений может составлять 256x256 пикселей и более, а частота их поступления – 25 кадров в секунду, реализация корреляционного алгоритма оценивания смещений в реальном времени затруднена даже при использовании современных вычислительных средств.

Одним из путей снижения вычислительной сложности при реализации корреляционных методов является использование подходов, основанных на анализе лишь определённых участков наблюдаемого и эталонного изображения, являющихся наиболее подходящими с точки зрения повышения точности оценивания параметров смещения [5]. Очевидно, что выбранные участки должны обладать хорошим отношением сигнал/шум и не должны располагаться вблизи движущихся объектов. При этом возникают две важные задачи. Во-первых, задача поиска этих участков, а, во-вторых, задача объединения результатов оценивания смещения различных опорных участков, с тем, чтобы вычислить оценку параметров смещения для всего изображения.

Предлагаемый в настоящей работе алгоритм основан на одновременном корреляционном отслеживании смещения нескольких участков изображения. В процессе слежения некоторые опорные участки отбрасываются, и в процесс обработки вовлекаются новые опорные участки. Количество опорных участков обозначим  $M$ . Номер конкретного опорного участка обозначим  $m$ ,  $m = \overline{1, M}$ . Поиск новых участков происходит с помощью так называемого поискового участка. Среди опорных участков один всегда является поисковым. Обозначим номер поискового участка на  $n$ -ом кадре как  $\mu_n$ . На первом кадре видеопоследовательности поисковым является последний участок, то есть  $\mu_1 = M$ .

При реализации корреляционного слежения за опорными участками изображения в данной работе использовалась разностная критериальная функция, дискретная форма которой имеет вид:

$$\rho(a, b) = - \sum_{i=-H}^H \sum_{j=-H}^H |l_n(i - a, j - b) - s_n(i, j)|, \quad (3)$$

где  $s_n(i, j)$  – эталонное изображение, сформированное к кадру с номером  $n$ ,  $H$  – размер эталонного изображения,  $a, b$  – предполагаемое смещение опорного участка.

Положение опорных участков изображения на первом кадре видеопоследовательности выбирается случайным образом.

При слежении за смещениями  $M$  опорных участков, на каждом кадре формируются  $m$  пар оценок, по одной на каждый опорный участок:  $(\hat{\alpha}_{m,n}, \hat{\beta}_{m,n})$ . На основе полученных оценок формируется общая оценка смещения изображения. При этом учитываются все полученные оценки, кроме пары оценок, соответствующих поисковому участку:

$$\hat{\alpha}_n = \underset{m \neq \mu_n}{\text{med}} \{ \hat{\alpha}_{m,n} \}, \quad \hat{\beta}_n = \underset{m \neq \mu_n}{\text{med}} \{ \hat{\beta}_{m,n} \}. \quad (4)$$

Затем находится номер участка, оценка смещения которого больше всего отличается от полученной оценки  $(\hat{\alpha}_n, \hat{\beta}_n)$ .

$$p_n = \arg \max_{m=1..M} \left\{ (\hat{\alpha}_n - \hat{\alpha}_{m,n})^2 + (\hat{\beta}_n - \hat{\beta}_{m,n})^2 \right\}. (5)$$

Поскольку опорный участок с номером  $p_n$  вероятнее всего соответствует либо движущемуся объекту, либо участку фона с малым отношением сигнал/шум, то он заменяется новым опорным участком изображения, координаты которого определяются случайным образом. Поисковый участок на  $(n+1)$ -м кадре последовательности соответствует найденному значению  $p_n$ , то есть  $\mu_{n+1} = p_n$ . Таким образом, алгоритм имеет тенденцию через некоторое количество кадров заменять ранее захваченные «некачественные» опорные участки на более подходящие. Заметим, что ещё одним событием, при котором опорный участок должен заменяться новым, является его близость к краю наблюдаемой области изображения.

Вычисленные значения  $(\hat{\alpha}_n, \hat{\beta}_n)$ , полученные в результате обработки  $n$ -го кадра, поступают на вход алгоритма выделения движущихся объектов.

Были проведены экспериментальные исследования представленного алгоритма с использованием искусственных и естественных видеопоследовательностей. При этом оценки смещения  $(\hat{\alpha}_n, \hat{\beta}_n)$  поступали на вход алгоритма выделения движущихся объектов [2]. Количество опорных областей принималось равным 5, размер каждой области – 25x25 пикселей. Результаты экспериментов показали, что эффективность алгоритма зависит от типа фонового изображения. При наблюдении воздушных объектов на облачном фоне успешное выделение объектов оказалось практически невозможным, в то время как при наблюдении объектов на фоне пересечённой местности достигнуты наилучшие результаты.

#### Литература

1. Муравьев В.С. Пространственный алгоритм обнаружения объектов на основе адаптивного порога // Проблемы передачи и обработки информации в сетях и системах телекоммуникаций. Тез. докл. 14-й междунар. научно-технич. конф. - Рязань, 2005. - С.122-123.
2. Алпатов Б.А., Бабаян П.В. Выделение движущихся объектов в условиях геометрических искажений изображения // Цифровая обработка сигналов. – 2004. – №4. – С. 9-14.
3. Zitova V., Flusser J. Image registration methods: a survey // Image and Vision Computing 21 – 2003, pp. 977–1000.
4. Баклицкий В.К., Бочкарёв А.М. Методы фильтрации сигналов в корреляционно-экстремальных системах навигации. – М.: Радио и связь, 1986. – 216 с.
5. Бачило С.А., Дзягун Д.Ю., Итенберг И.И., Овчаров А.Б., Сивцов С.А. Средства электронной стабилизации телевизионных изображений для подвижных систем наблюдения. Тез. докл. 4-й конф. DSPA. – М.: МЦНТИ 2002. – С. 252-254.



### SHIFT PARAMETERS ESTIMATION IN IMAGE SEQUENCE FOR AUTOMATIC OBJECT EXTRACTION

Babayan P., Strotov V.

Ryazan state radioengineering academy

In this work the algorithm of image shift estimation is proposed. This algorithm is aimed to be used commonly with object extraction algorithm, described in [1]. There are many approaches to image shift estimation [2]. One of the most effective image shift estimation methods is correlation algorithm. The algorithm suggested is based on simultaneously tracking on selected reference regions of image. During the tracking, some of reference regions are excluded, but some regions are included into tracking. Let the number of regions be  $M$ , the number of current region be  $m$ ,  $m = \overline{1, M}$ . The search of the new reference region is implemented by means of search region. One of the current regions is always a search region. Let the number of search region at the frame  $n$  be  $\mu_n$ . At the first frame search region is the last region:  $\mu_1 = M$ .

Correlation tracking is implemented using difference criterial function. In discrete form this function can be described as: 
$$\rho(a, b) = - \sum_{i=-H}^H \sum_{j=-H}^H |l_n(i-a, j-b) - s_n(i, j)|, \quad (1)$$

where  $l_n(x, y)$  – image from the image sensor,  $s_n(i, j)$  – template image,  $n$  – frame number,  $H$  – template image height,  $a, b$  – supposed region shift.

At the first frame the regions are situated at the image in random case.

During the tracking on  $M$  reference regions, at every frame  $m$  pairs of shift estimations are formed:  $(\hat{\alpha}_{m,n}, \hat{\beta}_{m,n})$ . The whole image shift estimation is calculated based on shift estimations for all regions, exclude search region:  $\hat{\alpha}_n = \underset{m \neq \mu_n}{med} \{\hat{\alpha}_{m,n}\}$ ,  $\hat{\beta}_n = \underset{m \neq \mu_n}{med} \{\hat{\beta}_{m,n}\}$ . (2)

Then the number of region, which has the most difference from the found estimation, is determined:

$$p_n = \arg \max_{m=1..M} \left\{ (\hat{\alpha}_n - \hat{\alpha}_{m,n})^2 + (\hat{\beta}_n - \hat{\beta}_{m,n})^2 \right\}. (3)$$

Because, most probably,  $p_n$ -th region is moving object or background with small signal-to-noise rate, we change this region by new region with random position. The search region number at  $(n+1)$ -th frame will be  $p_n$ , in other words  $\mu_{n+1} = p_n$ .

The experimental results are shown, that the effectiveness of the proposed algorithm depends upon the background type. If there are only clouds at the background, usually object extraction is impossible. If there is rugged terrain at the background, extraction results are good.

#### References

1. Alpatov B.A., Babayan P.V. Extraction of moving targets from the image sequence, distorted with geometric transformations // Digital signal processing. – 2004. – №4. – pp. 9-14.
2. Zitova B., Flusser J. Image registration methods: a survey // Image and Vision Computing 21 – 2003, pp. 977–1000.

