

РОССИЙСКАЯ АКАДЕМИЯ НАУК
МЕТОДЫ ОБРАБОТКИ СИГНАЛОВ И ИЗОБРАЖЕНИЙ

УДК 681.3

В. С. Киричук, В. А. Иванов, В. Ю. Ангеров, В. В. Сяньельщиков
(Новосибирск)

**МЕТОДИКА ВЫДЕЛЕНИЯ
ПОДВИЖНЫХ ТОЧЕЧНЫХ ОБЪЕКТОВ
ПРИ РЕГИСТРАЦИИ СЦЕНЫ С ДВУХ ТОЧЕК НАБЛЮДЕНИЯ**

Исследуется методика совместной обработки последовательности стереопар для обнаружения, распознавания и классификации типов подвижных и стационарных точечных объектов. Определяются условия качественного улучшения вероятности выявления объектов. Приводятся результаты применения описанной методики в обработке модельных последовательностей при известной геометрии съемки.

Целью данной статьи является исследование методики обработки двух последовательностей изображений, получаемых с камер, регистрирующих одну и ту же сцену, для обнаружения, распознавания и классификации типов подвижных и стационарных объектов. Основная задача заключается в выявлении принципиально новых возможностей, возникающих при совместной обработке таких связанных последовательностей, для качественного улучшения вероятности обнаружения объектов и определения траектории движения в трехмерном пространстве.

Модель сигнала. Подлежащие обработке изображения и их последовательности имеют следующие характеристики: 1) изображения содержат фон (стационарную часть сигнала), случайные шумы и искомые объекты; 2) статистические характеристики фона значительно меняются по полю кадра; 3) случайный шум, сопровождающий измерения, достаточно мал и стационарен как по полю кадра, так и по времени; характер распределения шума близок к нормальному; 4) кадры одной и той же последовательности не приведены к единой системе координат.

Появление второй точки наблюдения дает принципиальную возможность определения трехмерных координат подвижных объектов (оценивания высоты положения объекта, являющейся дополнительным классификационным параметром).

Пусть анализируемая сцена в некоторой выбранной системе координат имеет рельеф $Z(X, Y)$ и яркость $D(X, Y)$. На этой сцене могут находиться точечные объекты с координатами $\{X_k(t), Y_k(t), Z_k(t)\}$ и яркостью $D_k(t)$, $k = 1, \dots, n$, где n – число таких объектов.

Изображения данной сцены регистрируются двумя камерами (условно левой (L) и правой (R)), практически неподвижными относительно анализируемой сцены.

После приведения каждой из последовательностей изображений к единой системе координат пространственно-временной сигнал, получаемый на выходе приемного устройства, допускает следующее описание:

$$D_L^t(x_i, y_j) = M_L(x_i, y_j) + \sum_{k=1}^n a_L(k, t) f_L[x_i - x_L(t), y_j - y_L(t)] + \Theta_L^t(x_i, y_j), \quad (1)$$

$$D_R^t(x_i, y_j) = M_R(x_i, y_j) + \sum_{k=1}^n a_R(k, t) f_R[x_i - x_R(t), y_j - y_R(t)] + \Theta_R^t(x_i, y_j),$$

где D_L^t, D_R^t – последовательности, получаемые с L -, R -камер; M_L, M_R – стационарный фон; a_L, a_R – амплитуды объектов; f_L, f_R – формы объектов; Θ_L, Θ_R – случайный шум, сопровождающий измерения.

Предполагая геометрию камер полностью известной, несложно получить связь между координатами XYZ и проективными координатами данной точки в изображениях L -, R -камер:

$$\begin{pmatrix} x \\ y \end{pmatrix}_L = P_L\{X, Y, Z(X, Y)\}, \quad \begin{pmatrix} x \\ y \end{pmatrix}_R = P_R\{X, Y, Z(X, Y)\}, \quad (2)$$

где P_L и P_R – операторы проектирования [1] трехмерной сцены на плоскости изображений L -, R -камер. Следовательно, в формулах (1)

$$M_L\left\{\begin{pmatrix} x \\ y \end{pmatrix}_L\right\} = F_L\{M(X, Y, Z)\}, \quad M_R\left\{\begin{pmatrix} x \\ y \end{pmatrix}_R\right\} = F_R\{M(X, Y, Z)\}, \quad (3)$$

где F_L и F_R – функции амплитудного преобразования яркости.

Таким образом, соотношения (3) устанавливают связь между стационарными составляющими в изображениях (фоне) L -, R -камер. Однако использование этих соотношений для увеличения точности оценивания фона теоретически не оправдано. Действительно, если левая и правая последовательности содержат $N \times m$ точек (N – число точек в изображении, m – число кадров последовательности), нам необходимо по результатам ($N \times m + N \times m$) измерений оценить значение стационарной составляющей в N точках левого и правого изображений ($2N$ неизвестных). Использование соотношений (3) дает дополнительно N уравнений связи и вносит дополнительно N неизвестных высот. Кроме того, вид функций амплитудного преобразования яркости F_L, F_R также не полностью известен. Тем самым использование соотношений (3) для уточнения оценок M_L, M_R невозможно в силу того, что добавляется число неизвестных, большее, чем число уравнений связи. Из этих соображений следует однозначный вывод: при отсутствии дополнительных предположений о характере поведения фоновой составляющей оценивание стационарного фона должно осуществляться в левой и правой последовательностях независимо.

Алгоритм. На первом этапе используется комплекс алгоритмов межкадровой обработки для эффективного исключения мешающего фактора – стационарного фона [2]. Данная процедура производится независимо для каждой из связанных последовательностей; ее конечным результатом является сигнал, содержащий аддитивную смесь случайного шума и объектов. Приведенная выше модель приобретает упрощенный вид:

$$\Delta D_L^t(x_i, y_j) = \sum_{k=1}^n a_L(k, t) f_L\{x_i - x_{kL}(t), y_i - y_{kL}(t)\} + \Theta_L^t(x_i, y_j), \quad (4)$$

$$\Delta D_R^t(x_i, y_j) = \sum_{k=1}^n a_R(k, t) f_R\{x_i - x_{kR}(t), y_i - y_{kR}(t)\} + \Theta_R^t(x_i, y_j),$$

где a_L, a_R – преобразованные яркости k -го объекта в L - и R -камерах;

$$\begin{pmatrix} x_{kL}(t) \\ y_{kL}(t) \end{pmatrix} = P_L\{X_k(t), Y_k(t), Z_k(t)\}, \quad \begin{pmatrix} x_{kR}(t) \\ y_{kR}(t) \end{pmatrix} = P_R\{X_k(t), Y_k(t), Z_k(t)\}. \quad (5)$$

Дальнейшее построение алгоритма зависит от вида анализируемых подвижных объектов. Целью данной работы является построение процедур поиска точечных объектов, т. е. объектов, размеры которых значительно меньше элемента разрешения считывающих камер. В силу того что объект точечный, его форма на изображениях D_L и D_R определяется используемой оптикой, типом камер и координатами проекций объектов на элементы считывания. Исходя из этих предположений, можно считать, что точечный объект вызывает реакцию на три-четыре соседних элемента разрешения по каждой оси, при этом форма объекта сильно модифицируется в зависимости от его расположения относительно элементов разрешения, а яркость изменяется в десятки раз. Поэтому для повышения вероятности определения объектов используется классический подход – согласованная фильтрация с учетом всех описанных выше ограничений [3].

Далее осуществим геометрическое преобразование так, чтобы одна из координатных осей (например, x) совпадала с базовой линией системы регистрации [4]. После проведения всех преобразований уравнения (4) приводятся к виду

$$\Delta \tilde{D}_L^t(x_i, y_j) = \sum_{k=1}^n \tilde{a}_L(k, t) \delta\{x_i - x_{kL}(t), y_i - y_k(t)\} + \tilde{\Theta}_L^t(x_i, y_j), \quad (6)$$

$$\Delta \tilde{D}_R^t(x_i, y_j) = \sum_{k=1}^n \tilde{a}_R(k, t) \delta\{x_i - x_{kR}(t), y_i - y_k(t)\} + \tilde{\Theta}_R^t(x_i, y_j),$$

где δ – символ Кронекера; $y_k(t) = y_{kL}(t) = y_{kR}(t)$, т. е. y -координаты в обоих изображениях совпадают, а величина $\Delta x_k(t) = x_{kL}(t) - x_{kR}(t)$ однозначно определяет высоту объекта. Таким образом, задача сводится к одномерной в следующей постановке.

Даны два одномерных сигнала, содержащие шум и объекты (для простоты будем рассматривать наличие одного объекта):

$$d_L(x_i) = \xi_L(x_i) + A_L \delta(x_i - x_L), \quad d_R(x_i) = \xi_R(x_i) + A_R \delta(x_i - x_R), \quad (7)$$

причем диапазон $\gamma = x_L - x_R$ известен ($\gamma < P$) и определяется возможным диапазоном высот объекта. Необходимо проверить гипотезу H_0 о наличии объекта. Если гипотеза подтверждается, то определить x_L и x_R и по этим данным рассчитать трехмерные координаты XYZ. Оптимальным критерием [5] (в предположении гауссовых шумов) в рамках метода максимального правдоподобия проверки гипотезы H_0 является следующий:

$$J(x_i, q) = d_L(x_i) + d_R(x_i + q), \quad 0 \leq q < p,$$

и если $J(x_i, q) > C$, то принимается решение о наличии объекта с координатами $x_L = x_i$; $x_R = x_i + q$.

Приведем сравнение данного подхода с результатами обработки только одной последовательности, а также с алгоритмом триангуляции – алгоритмом, когда точки, подозрительные на наличие объекта в левой и правой последовательностях, выделяются независимо, а затем осуществляется их попарное сравнение с учетом возможного диапазона высот объекта.

Для предлагаемого алгоритма вероятность пропуска объекта (в предположении, что $\sigma_L / \sigma_R = \sigma_m / \sigma_m = A$) составляет:

$$\beta_s = \Phi(\sqrt{2}A - C_s),$$

где Φ – интеграл нормального закона распределения; C_s – выбранный уровень, определяемый вероятностью появления ложной отметки $\alpha_s = \alpha P$ (P – диапазон неопределенности), и $C_s = \Phi(\alpha)$.

Для триангуляционного алгоритма

$$\alpha_T = \alpha^2 P, \quad C_T = \Phi(\alpha), \quad 1 - \sqrt{1 - \beta_T} = \Phi(A - C_T).$$

При обработке монопоследовательности

$$C_m = \Phi(\alpha_m), \quad \beta_m = \Phi(A - C_m).$$

Для количественного сравнения результатов рассчитаем амплитуду объекта (в единицах дисперсии шума) для вероятности пропуска объекта $\beta = 0,1$, вероятность появления ложной отметки 10^{-4} при зоне неопределенности $P = 1$ (высота положения объекта известна точно), $P = 10$ и $P = 100$ точек. Данные приведены в таблице.

Из анализа таблицы можно сделать следующие выводы: 1) алгоритм совместной обработки двух последовательностей дает лучшие результаты, чем

Количество точек	Стереобработка	Триангуляция	Монообработка
$P = 1$	3,44	3,97	5,03
$P = 10$	3,93	4,39	5,03
$P = 100$	4,23	4,76	5,03

алгоритмы триангуляции и монообработки; 2) с увеличением зоны неопределенности высоты объекта мощности алгоритмов сближаются и, что очевидно, при $P \rightarrow \infty$ они совпадут с алгоритмом обработки монопоследовательности.

Таким образом, при поиске малоразмерных объектов возникает следующая последовательность обработки:

- приведение левой и правой последовательностей к единой системе координат;
- компенсация фоновой составляющей (операции межкадровой обработки);
- проведение согласованной фильтрации (внутрикадровая обработка);
- геометрическое преобразование изображений для совмещения одной из координатных осей с базовой линией системы регистрации;
- совместная обработка изображений левой и правой камер для выделения точек, подозрительных на наличие объектов;
- пороговая обработка и формирование массива отметок;
- определение трехмерных координат точек, подозрительных на наличие объектов.

Экспериментальные результаты. Экспериментальная проверка разработанной методики проводилась на последовательностях стереоизображений с известной геометрией съемки. Формирование последовательностей осуществлено с помощью пакета 3D-графики RovRay. Его функции моделирования достаточно хорошо реализуют законы оптики, что дает возможность генерировать реалистичные изображения. Работа с данным пакетом показала его применимость при разработке и отладке алгоритмов обработки изображений, включая стереообработку.

Геометрия сцены, используемая для формирования стереопоследовательностей, соответствует системе наблюдения из двух камер с заданными характеристиками. Текстура данного фрагмента взята из спутниковых снимков поверхности Земли с «достаточно плотным» облачным покровом и низким разрешением. Небольшой шум от кадра к кадру вводится автоматически при формировании последовательности путем введения атмосферных эффектов.

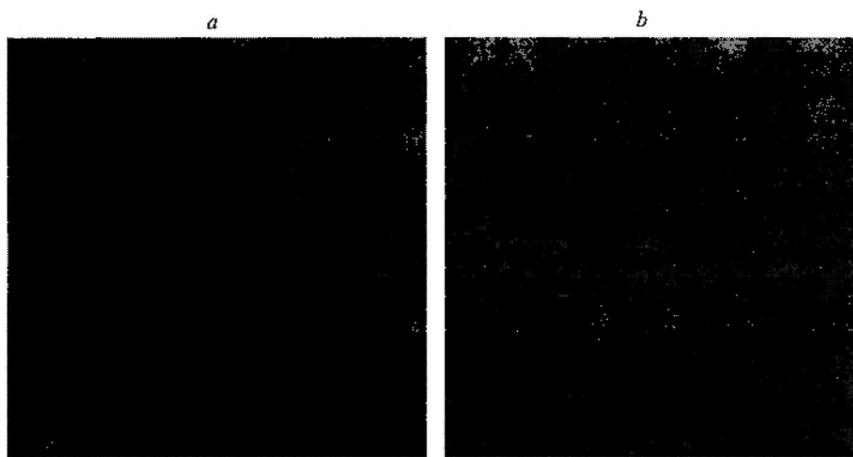


Рис. 1. Кадры исходной серии: *a* – левый, *b* – правый

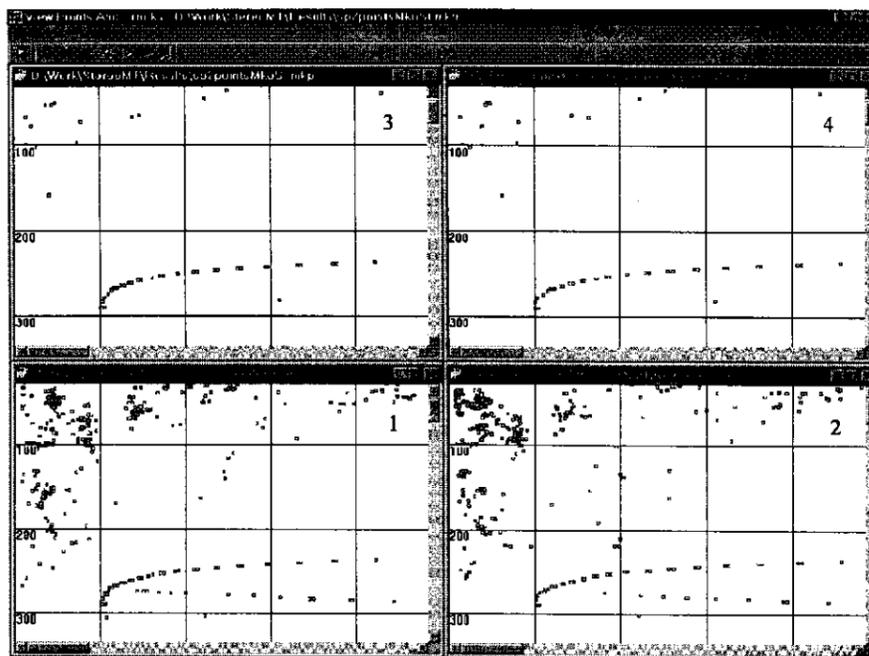


Рис. 2. Результаты обработки двух последовательностей: 1, 2 – межкадровой обработки ($N_L = 510, N_R = 524$); 3, 4 – последующей стереообработки ($N_L = 55, N_R = 52$)

Малоразмерный объект представляется сферой с заданной яркостью и траекторией движения.

На рис. 1 приведены левый и правый кадры исходной серии изображений. На рис. 2 показаны результаты обработки данной серии (N_L – число точек, подозрительных на наличие объекта и найденных на левой серии; N_R – число точек, подозрительных на наличие объекта и найденных на правой серии). Приведенные результаты показывают, что вероятность ложной тревоги ввиду использования алгоритмов стереообработки была снижена с 10^{-4} (510(524) точек из $512 \times 512 \times 19$ точек) до 10^{-5} (55(52) точек из $512 \times 512 \times 19$ точек).

ВЫВОДЫ

Эффективность совместной обработки (с точки зрения вероятности обнаружения объектов и появления ложных отметок) двух связанных последовательностей существенно превышает эффективность обработки монопоследовательности и обработки, основанной на алгоритме триангуляции.

Эффективность совместной обработки падает с увеличением диапазона неопределенности высот объектов и при полной неопределенности эквивалентна эффективности обработки монопоследовательности.

СПИСОК ЛИТЕРАТУРЫ

1. Фокс А., Пратт М. Вычислительная геометрия. Применения в проектировании и на производстве: Пер. с англ. М.: Мир, 1982.

2. Киричук В. С., Коршевер И. И., Синельщиков В. В. Анализ изображений динамических сцен: модели, алгоритмы и системы реального времени // Автометрия. 1998. № 3. С. 3.
3. Киричук В. С., Яковенко Н. С. Линейная фильтрация в задачах поиска объектов произвольной ориентации // Автометрия. 1991. № 3. С. 10.
4. Faugeras O. Three-Dimensional Computer Vision. N. Y.: MIT Press, 1993.
5. Леман Э. Проверка статистических гипотез: Пер. с англ. М.: Наука, 1964.

*Институт автоматизации и электротехники СО РАН,
E-mail: kirichuk@iae.nsk.su*

*Поступила в редакцию
26 июня 2000 г.*

Реклама продукции в нашем журнале – залог Вашего успеха!