

УДК 621.397.13

## ПРЕДВАРИТЕЛЬНАЯ КЛАССИФИКАЦИЯ ИЗОБРАЖЕНИЯ В ЗАДАЧАХ СЕГМЕНТАЦИИ ОБЪЕКТОВ

**Н. А. Обухова,**

канд. техн. наук, доцент

Санкт-Петербургский государственный университет аэрокосмического приборостроения

*Рассматривается алгоритм предварительной классификации изображения с целью выделения областей, содержащих с высокой вероятностью объекты интереса. Классификация областей изображения выполняется по уровню высокочастотной энергии. Приведены результаты сравнительного анализа различных способов, подчеркивающих высокочастотные составляющие изображения, и выбран способ, обеспечивающий наиболее высокий уровень качества классификации. Подробно рассмотрена процедура пороговой обработки.*

*A method of preliminary image classification is considered. The aim of the analysis is to distinguish the areas containing, with high probability, the objects of interest. Image classification is carried out according to the level of high-frequency energy. The results of the comparative analysis for various methods emphasizing the high-frequency image components are given and the method providing the highest quality of classification is chosen. A threshold processing procedure is considered in detail.*

### Введение

Реализовать автоматический захват и сопровождение нескольких объектов, разрешить ситуацию окклюзии — перекрытия объекта интереса фоном или другим объектом, сегментировать неподвижные объекты, а также исключить потерю объекта при его остановке позволяет метод сегментации по совокупности признаков [1].

Основными процедурами метода являются:

— предварительная классификация изображения с целью выделения областей, в которых с высокой вероятностью возможно присутствие объектов интереса;

— сегментация объектов по признакам временной и пространственной корреляции [2, 3];

— сопровождение объектов путем определения вероятности принадлежности фрагмента изображения объекту  $g$  на основе математического аппарата теории нечетких множеств.

Метод разработан для сегментации и сопровождения объектов размером (в пересчете к плоскости кадра) от  $4 \times 4$  до  $100 \times 100$  пикселей, обладающих жестким движением со скоростью от 0 до 20 пикселей за кадр. Жестким (non-rigid) движением называется движение, при котором все части объекта движутся в соответствии с основным направлением движения. Наиболее часто объектами с такими характеристиками являются объекты искусственного происхождения: автомобили и другие транспортные средства, летательные аппараты, морские цели.

### Общие положения

Существенной задачей при сегментации объектов по совокупности признаков является предварительная классификация изображения. Ее цель — выделение областей, в которых возможно присутствие объектов интереса. Именно эти области подвергаются дальнейшей обработке. Соответственно, предварительная классификация, не потеряв объект интереса или какую-либо его часть, должна уменьшить объем обрабатываемого далее материала. Проведение предварительного анализа на основе таких признаков как движение, форма или яркость не решает указанной выше задачи. Признак движения исключает из дальнейшего рассмотрения все неподвижные и медленно движущиеся (маневрирующие) объекты. Признак формы требует существенных априорных знаний об объектах интереса и, соответственно, неприемлем для первоначальной классификации. При работе по яркостному признаку на объект накладывается требование однородной яркости, что достаточно часто нарушается в реальных условиях видеонаблюдения для протяженных объектов интереса.

Отличительной особенностью областей изображения, включающих описанные выше объекты интереса, является большое количество перепадов яркости (детальность): на границах объект/фон, на внутренних контурах объекта и др. Для оценки этого признака введем понятие уровня детальности

$$D(k, l) = \sum_{j=1}^M \sum_{i=1}^N PL(x_k + i, y_l + j),$$

где  $PL$  — перепад яркости в пикселе;  $x_k, y_l$  — координаты левого верхнего угла фрагмента (блока) изображения;  $M, N$  — число пикселей по вертикали и горизонтали во фрагменте. Поверхности, описывающие изменения уровней детальности по плоскости кадра, показаны на рис. 1. Области, которые должны быть оставлены для дальнейшей обработки, характеризуются высоким уровнем детальности. Этот признак и положен в основу предварительной классификации.

Для ее реализации предложен следующий алгоритм:

- обработка изображения с целью усиления перепадов яркости;
- разбиение изображения на фрагменты заданного размера (блоки  $8 \times 8$ ) и определение числа перепадов яркости в каждом фрагменте;
- пороговая обработка.

Первый шаг может быть реализован большим количеством способов: методами линейного и нелинейного контрастирования, морфологической обработкой, вейвлет-преобразованием. При этом способ обработки непосредственно определяет качество предварительной классификации.

Для оценки эффективности различных способов обработки в рамках решаемой задачи был проведен их сравнительный анализ. Из каждой группы выбран метод, максимально учитывающий особенности видеоматериала: существенная часть используемых для тестирования изображений имеет фон с выраженными горизонтальными структурами (береговая линия, линия горизонта), поэтому подчеркивание перепадов яркости по вертикали заведомо приведет к включению больших фрагментов фона в области для дальнейшей сегментации. Таким образом, рассматривались следующие методы.

Линейная обработка — пространственное дифференцирование в горизонтальном направлении:

$$D(x, y) = |L(x, y) - L(x + \Delta x, y)|,$$

где  $L$  — яркость пикселя в кадре;  $(x, y)$  — координаты пикселя;  $\Delta x$  — приращение координаты  $x$  (для блока  $8 \times 8$  равна 5).

Нелинейная обработка — выделение контуров с помощью оператора Превитта [4]:

$$O\_P = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}.$$

Морфологическая обработка — многомасштабный градиент [5]:

$$MG(L) = \frac{1}{3} \sum_{i=1}^3 [(L \oplus S_i) - (L \ominus S_i)] \ominus S_{i-1},$$

где  $L(x, y)$  — исходное изображение;  $\oplus$  и  $\ominus$  — морфологические операции наращивания и эрозии [3];  $S_i$  — квадратная группа структурных элементов. Размер  $S_i$  равен  $(2i+1)(2i+1)$  пикселей для  $0 \leq i \leq 3$ . В соответствии с этим выражением значения градиентов рассчитывают трижды с использованием структурных элементов различной размерности, а затем результаты складывают.

Вейвлет-преобразование — вейвлет Добеши Д4 [6, 7]. Для анализа используются коэффициенты вейвлет-разложения матрицы **НН**, полученные операцией свертки с коэффициентами фильтра высоких частот. Коэффициенты фильтра для вейвлет Добеши Д4

$$g_0 = -\frac{1 - \sqrt{3}}{4\sqrt{2}}; g_1 = -\frac{3 - \sqrt{3}}{4\sqrt{2}};$$

$$g_2 = \frac{3 + \sqrt{3}}{4\sqrt{2}}; g_3 = -\frac{1 + \sqrt{3}}{4\sqrt{2}}.$$

### Критерии эффективности

Для оценки эффективности различных способов обработки сформулированы следующие критерии.

*Степень выделения объектов интереса* — первый и наиболее важный критерий. Области изображения, исключенные на этапе предварительной классификации, далее вообще не рассматриваются, поэтому присутствие в них объекта интереса или его части — грубая ошибка. Значительно безопаснее включение в дальнейшую обработку фрагментов изображения, не содержащих объекты интереса, так как эти области будут отсегментированы в ходе дальнейшего анализа. Для оценки критерия предлагается использовать отношение площадей:

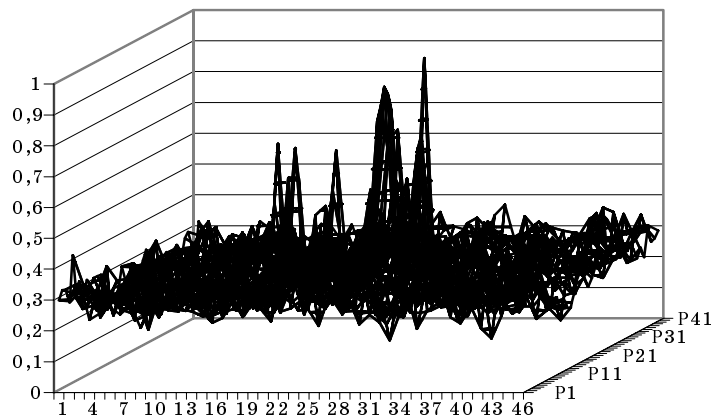
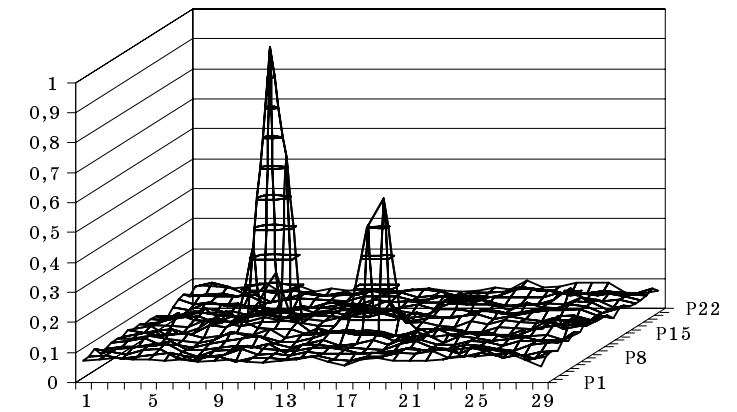
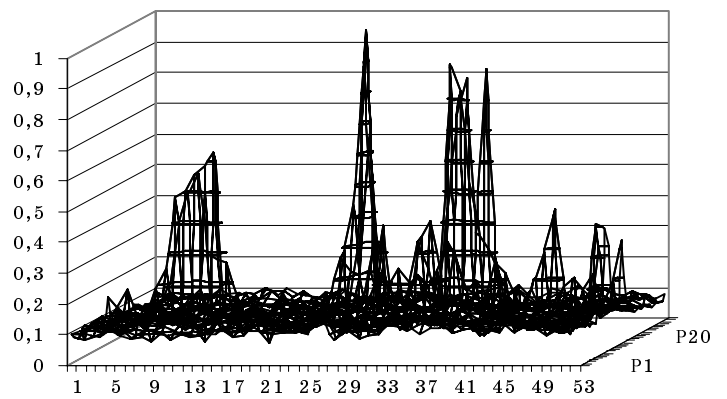
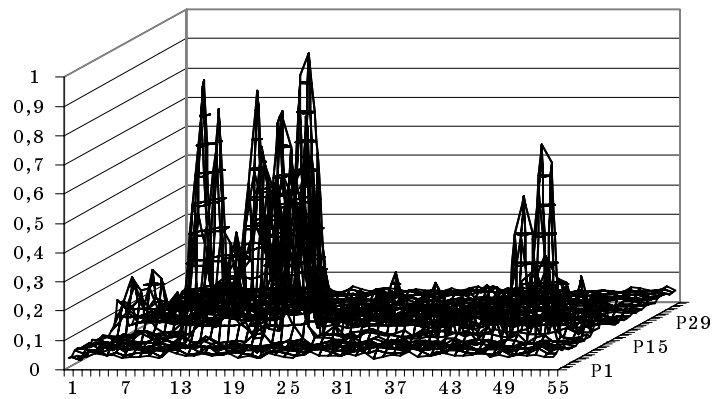
$$K_1 = \frac{S_c}{S_n},$$

где  $S_c$  — площадь объекта, сегментированного одним из методов;  $S_n$  — истинная площадь объекта на изображении. Соответственно, чем ближе коэффициент  $K_1$  к единице, тем эффективнее применяемый метод обработки.

*Степень выделения фона*. Этот критерий показывает, какова площадь изображения, выделенного для дальнейшего анализа, но не содержащая объектов интереса. Его оценка

$$K_2 = 1 - \frac{S_\phi}{S_n},$$

где  $S_\phi$  — площадь фона в выделенных для дальнейшего анализа областях изображения;  $S_n$  — площадь всего изображения. Соответственно, чем ближе коэффициент  $K_2$  к единице, тем эффективнее применяемый метод обработки.



■ Рис. 1. Поверхности, описывающие уровни детальности для тестовых последовательностей «Корабль», «Катер», «Самолеты», «Туман»

Эффективность обработки — критерий уровня полезной информации в данных, оставленных для дальнейшего анализа в результате классификации:

$$K_3 = \frac{S_c}{S_\phi}$$

Инвариантность к типу изображения  $K_4$  — критерий устойчивости значений коэффициентов  $K_1$ ,  $K_2$  и  $K_3$  для различных изображений (разная степень присутствия фона, различный диапазон перепадов яркости и др.).

### Экспериментальное исследование и его результаты

Для проведения экспериментов были выбраны четыре тестовые последовательности. Их основные характеристики сведены в табл. 1.

С помощью специально разработанного программного обеспечения были получены значения коэффициентов  $K_1$ ,  $K_2$ ,  $K_3$  для 100 кадров каждой тестовой последовательности. В полученных выборках были определены средние значения коэффициентов  $K_1$ ,  $K_2$ ,  $K_3$ . Полученные результаты представлены на рис. 2. По этим данным была проведена ранговая оценка методов (табл. 2).

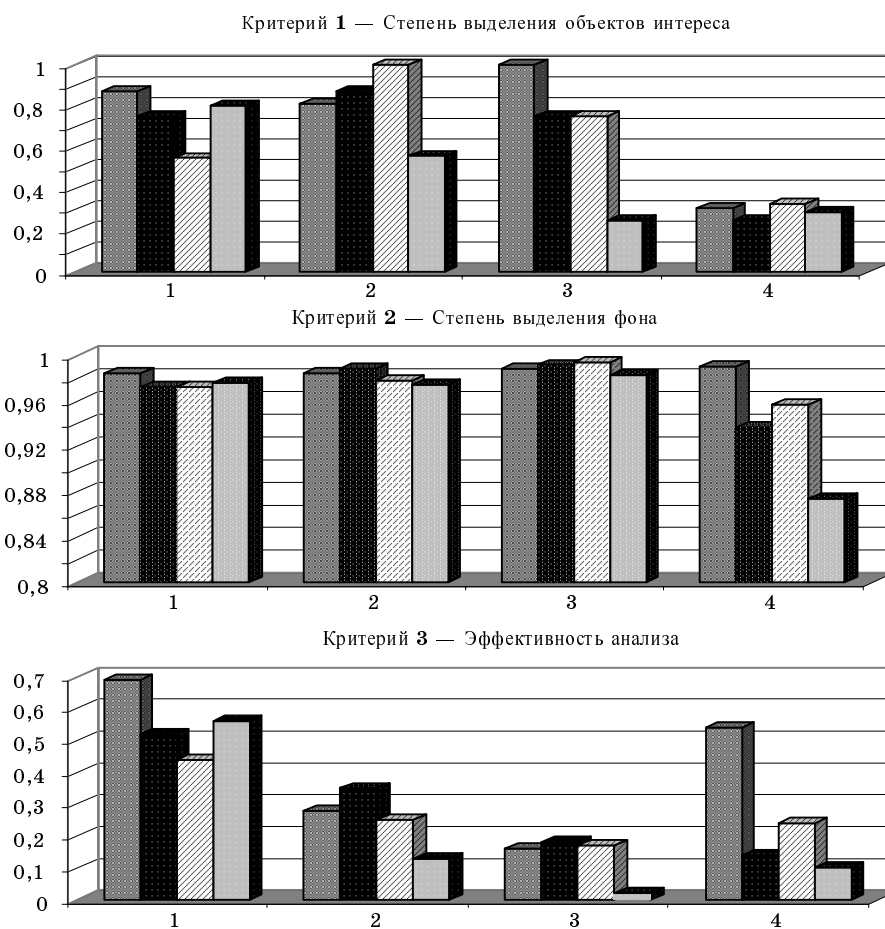
Соответствие рангов и численных значений критериев  $K_1$ ,  $K_2$ ,  $K_3$  очевидно, ранжирование ме-

тодов по критерию  $K_4$  было выполнено следующим образом. Наиболее устойчивым является метод пространственного дифференцирования. Это единственный метод, успешно справившийся со всеми тестовыми последовательностями, включая последовательность «Туман». Критерий  $K_2$ : степень выделения фона у него на всех последовательностях колеблется около 0,985. Все другие методы на последовательности «Туман» имеют значение критерия в 3–4 раза больше, что приводит к неудовлетворительному результату — существенному выделению фона. Наибольшей чувствительностью к характеру изображения обладает вейвлет-фильтр Добеши. Для последовательностей «Корабль» и «Катер» его характеристики близки к характеристикам остальных методов, а для последовательности «Самолеты» и «Туман» результат классификации практически отсутствует, что отражает и значения коэффициентов  $K_1$ ,  $K_2$ ,  $K_3$ . Соответственно, по критерию  $K_4$  ранги распределились следующим образом: 1 — пространственное дифференцирование, 2 — морфологический градиент, 3 — оператор Превитта, 4 — вейвлет-фильтр Добеши.

Таким образом, по сумме оценок наиболее эффективным способом обработки является пространственное дифференцирование. Результаты предварительной классификации на его основе представлены на рис. 3. Следует отметить, что по значениям критериев  $K_1$ ,  $K_2$ ,  $K_3$  обработка с помо-

■ Таблица 1

Название тестовой последовательности	Описание тестовой последовательности
«Корабль»	Характеристика объекта интереса: протяженный, медленно двигающийся (маневрирующий) объект Характеристика фона: средний уровень детальности — линия горизонта, небольшая рябь на воде Средняя яркость по фону — 163 Средняя яркость объекта — 103 Уровень контраста фон/объект — 0,36
«Катер»	Характеристика объекта интереса: протяженный, быстро двигающийся объект Характеристика фона: высокий уровень детальности — береговая линия с пирсом, существенная рябь на воде Средняя яркость по фону — 126 Средняя яркость объекта — 151 Уровень контраста фон/объект — 0,16
«Самолеты»	Характеристика объекта интереса: близкий к точечному, медленно двигающийся объект Характеристика фона: средний уровень детальности — облака Средняя яркость по фону — 205 Средняя яркость объекта — 98 Уровень контраста фон/объект — 0,52
«Туман»	Характеристика объекта интереса: протяженный, медленно двигающийся объект Характеристика фона: низкий уровень детальности — очень слабая рябь на воде Средняя яркость по фону — 150 Средняя яркость объекта — 147 Уровень контраста фон/объект — 0,02



■ Рис. 2. Средние значения критериев  $K_1, K_2, K_3$  для тестовых последовательностей: 1 — «Корабль», 2 — «Катер», 3 — «Самолеты», 4 — «Туман»; ■ — пространственное дифференцирование; ▨ — оператор Превитта; ■ — морфологический градиент; □ — вейвлет-фильтр Добеши

■ Таблица 2

Критерий	Метод обработки	Ранг для тестовых последовательностей				Сумма рангов	Результирующий ранг
		«Корабль»	«Катер»	«Самолеты»	«Туман»		
$K_1$	Пространственное дифференцирование	1	3	1	3	8	1
	Многомасштабный морфоградиент	2	2	2	4	10	3
	Оператор Превитта	4	1	2	2	9	2
	Вейвлет-фильтр Добеши Д4	3	4	3	1	11	4
$K_2$	Пространственное дифференцирование	1	2	1	3	7	1
	Многомасштабный морфоградиент	3	1	3	2	8	2
	Оператор Превитта	3	3	2	1	9	3
	Вейвлет-фильтр Добеши Д4	2	4	4	4	14	4
$K_3$	Пространственное дифференцирование	1	2	1	3	7	1
	Многомасштабный морфоградиент	3	1	3	1	8	2
	Оператор Превитта	4	3	2	2	11	3
	Вейвлет-фильтр Добеши Д4	2	4	4	4	14	4



■ Рис. 3. Результаты классификации изображения с помощью пространственного дифференцирования

щью пространственного дифференцирования близка к обработке на основе морфологического градиента (исключая последовательность «Туман»). Соответственно, при достаточных соотношениях контраста этот подход также обеспечивает весьма эффективное подчеркивание высокочастотной составляющей при предварительной классификации.

### Пороговая обработка

Последним шагом алгоритма предварительной классификации является пороговая обработка: на основании полученной оценки детальности блок должен быть оставлен или исключен из дальнейшего анализа.

Процедура формирования порога базируется на двух основных положениях.

1. Совокупность оценок детальности, полученных для всех блоков кадра, является выборкой случайной величины. Анализ и исследование ее гистограмм позволяет утверждать, что это случай-

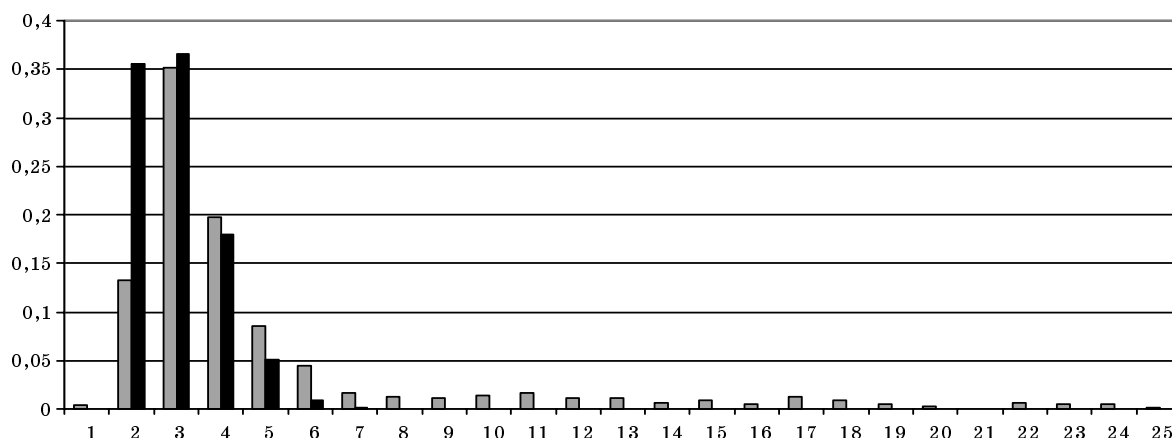
ная величина, описываемая распределением Релея:

$$w(\rho) = \frac{\rho}{\sigma^2} \exp\left(\frac{-\rho^2}{2\sigma^2}\right).$$

Гистограммы распределения оценок детальности, полученные в ходе исследования, и теоретический вид функции распределения Релея при  $\sigma = 3$  показаны на рис. 4.

2. Задача предварительной обработки изображения является задачей бинарной классификации, что позволяет использовать критерий Неймана—Пирсона.

Первый класс  $D_1$  — блоки изображения с высокой вероятностью, содержащие объекты интереса. Второй класс  $D_2$  — блоки, не подлежащие дальнейшей обработке. Ошибкой первого рода  $p_0$  в данном контексте является отнесение блока изображения, содержащего объект интереса, к классу  $D_2$ ,



■ Рис. 4. Вид гистограммы:

■ — экспериментальная гистограмма; ■ — теоретическая гистограмма

ошибкой второго рода  $p_1$  — отнесение к классу  $D_1$  блока, не подлежащего дальнейшей обработке.

Суть стратегии Неймана — Пирсона состоит в следующем: задают допустимое значение вероятности ошибки первого рода  $p_0$ , а затем классификатор строят таким образом, чтобы обеспечить минимум вероятности ошибки второго рода  $p_1$ :

$$\begin{cases} p_0 = p_0^* \\ p_1 \rightarrow \min_{D_0, D_1} \end{cases}$$

Решением является классификатор вида

$$\lambda(y) = \frac{p(y/\Omega_1)}{p(y/\Omega_0)} > \lambda \Rightarrow \begin{cases} y \in D_1 \\ y \in D_0 \end{cases}$$

значение пороговой величины  $\lambda$  определяется, исходя из условия  $p_0 = p_0^*$ ;  $p(y/\Omega_1)$  и  $p(y/\Omega_0)$  — функции правдоподобия классов  $D_1$  и  $D_2$  [4].

Соответственно, в нашем случае пороговое значение

$$\lambda = \sqrt{\left(-2 \hat{\sigma}^2 \ln(p_0)\right)},$$

где  $\hat{\sigma}$  — оценка  $\sigma$ .

Из теории известно, что максимальная плотность при распределении Релея достигается при  $\rho = \sigma$ . Эта особенность позволяет предложить быстрый и удобный для практической реализации способ оценки параметра  $\sigma$  по экспериментальной гистограмме

$$\hat{\sigma} = \min + \text{num\_max} \times \text{int} + \text{int}/2,$$

где  $\min$  — минимальное значение оценки детальности;  $\text{num\_max}$  — номер интервала гистограммы с максимальной частотой попадания случайной величины;  $\text{int}$  — величина интервала.

## Выводы

Предложенный алгоритм предварительной классификации позволяет выделить области изображения с заданной вероятностью, содержащие объекты интереса. Как показали экспериментальные исследования, это дает возможность исключить из последующей обработки до 80 % кадра, тем самым резко снизить вычислительную и алгоритмическую сложность последующих процедур. Значимым достоинством предлагаемого алгоритма является отсутствие существенных ограничений на скорость, линейные размеры и яркостные характеристики объектов интереса.

## Литература

1. Обухова Н. А., Тимофеев Б. С. Сегментация и сопровождение объектов на основе анализа видеопоследовательности // IV Междунар. конф. «Телевидение: передача и обработка изображений»: Тез. докл. / СПбГЭТУ «ЛЭТИ», 2005. С. 87–89.
2. Тимофеев Б. С. Видеокомпьютерные системы для наблюдения за движущимися объектами // Изв. вузов. Сер. Радиоэлектроника. 2003. № 4. С. 32–44.
3. Обухова Н. А. Обнаружение и сопровождение движущихся объектов методом сопоставления блоков // Информационно-управляющие системы. 2004. № 1. С. 30–37.
4. Методы компьютерной обработки изображений / Под ред. В. А. Соифера. М.: Физматлит, 2003. 784 с.
5. Maragos S. Tutorial an advances in morphological image processing and analysis // Optical Engineering. July 1987. N 26(7). P. 623–632.
6. Киселев А. Вейвлет своими руками / BaseGroup Labs, 2003. [http://www.basegroup.ru/filtration/making\\_wavelet.htm](http://www.basegroup.ru/filtration/making_wavelet.htm)
7. Уэлстид С. Фракталы и вейвлеты для сжатия изображений в действии: Учеб. пособие. М.: Триумф, 2003. 320 с.