

УДК 519.681

О БАЙЕСОВСКОЙ СЕГМЕНТАЦИИ ЗОН ИНТЕРЕСА СКАЛЯРНЫХ СЦЕН

В.Б. Фофанов, Р.Г. Сабиров

Аннотация

Рассматривается задача о сегментации зон интереса, которые являются фрагментами сцены, содержащими заданный объект и его некоторую окрестность. Она формулируется как задача байесовской классификации пикселей зоны интереса на два класса. Показывается, что для оценки неизвестных априорных вероятностей и условных распределений классов достаточно знать площадь объекта и изображение зоны интереса. Приводятся результаты сегментации предлагаемым методом модельных и реальных сцен

Введение

Сегментацией принято называть задачу о разбиении сцены на составляющие ее объекты по одному или нескольким изображениям. Отметим, что сегментация может рассматриваться как в качестве основной цели дешифрирования (например, в задаче об обновлении топографических карт), так и в качестве промежуточной (при поиске на сцене заданных объектов по их геометрическим признакам). Исследованию различных подходов к формализации и решению задачи сегментации посвящено чрезвычайно большое количество работ (см., например, обзоры [1–4]). Однако эта тема не закрыта до сих пор [5–9]. Не ослабевающий интерес к задаче сегментации объясняется, по-видимому, отсутствием эффективных методов ее решения.

В настоящей работе предлагается математическая модель и методы решения задачи сегментации, возникающей в ходе поиска на сцене заданных объектов по их геометрическим признакам. Специфика рассматриваемого случая заключается в том, что сегментации подвергается не вся сцена, а только те ее фрагменты, которые содержат заданный объект и некоторое его окружение. Они называются зонами интереса. При сегментации зон интереса удается построить обучающую выборку, вычислить по ней оценки для вероятностей условных распределений классов и воспользоваться байесовским решающим правилом. Показывается, что эффективность сегментации, описываемая вероятностью ошибки, зависит от качества изображения, выражаемого отношением сигнал/шум.

Исходной информацией о сцене в настоящей работе служит ее единственное изображение. Такие изображения будут называться далее скалярными. Набор из нескольких одновременно сформированных скалярных изображений будет называться векторным изображением сцены. Предполагается, что на каждом скалярном изображении задана система координат, и координаты любого пикселя сцены на этих изображениях одни и те же.

1. Скалярные сцены и скалярные изображения

Будем рассматривать сцену как совокупность неделимых элементов, называемых пикселями. Каждый пиксель характеризуется индивидуальными целочисленными координатами $z = (z_1, z_2)$, заданными на двумерной целочисленной решетке

$$Z^2 = \{z = (z_1, z_2) : z_1 \in Z, z_2 \in Z\},$$

и скалярной случайной величиной ξ_z со значениями из конечного множества $Y = \{0, 1, \dots, 2^k - 1\}$, содержащего $|Y| = 2^k$ элементов. Здесь и в дальнейшем под модулем конечного множества понимается число элементов в нем. Случайная величина ξ_z описывает свойство пикселя, значение которого становится известно только после его измерения (съемки). Предполагается, что все случайные величины определены на одном и том же вероятностном пространстве (Ω, A, P) . Счетное семейство $(\xi_z)_{z \in Z^2}$ случайных величин называется, как известно, скалярным случайнym полем на Z^2 . В настоящей работе оно будет называться также скалярной сценой.

Пусть ω – некоторое элементарное событие из Ω . Для каждого пикселя сцены с координатами $z \in Z^2$ и случайной величины ξ_z можно вычислить выборочное значение $x_z = \xi_z(\omega)$. Назовем скалярным изображением сцены отображение $x : Z^2 \rightarrow Y$, определяемое равенством $x(z) = x_z$.

Как правило, интерес представляют не отдельные пиксели, а их конечные совокупности, которые будут называться элементами сцены или объектами. Каждый объект определяется подмножеством A точек из Z^2 , являющихся координатами его пикселей, и семейством $\xi_A = (\xi_a)_{a \in A}$ из $|A|$ скалярных случайных величин. Подмножество A будет называться далее проекцией объекта. Предполагается, что проекции A и B разных объектов не пересекаются. Если x – изображение сцены, то изображением x_A объекта с проекцией A естественно считать сужение x на A , то есть $x_A = (x_a)_{a \in A}$.

В работе [10] доказано, что в случае независимости случайных величин, принадлежащих разным объектам сцены, для ее задания достаточно разделить Z^2 на конечные непересекающиеся подмножества – проекции объектов и определить для каждой проекции A распределение вероятностей $P_{Y^A} = (p_{Y^A}(x_A))_{x_A \in Y^A}$ на множестве всех подмножеств из Y^A .

Далее под площадью объекта ξ_A подразумевается количество $|A|$ пикселей, образующих его проекцию A . В качестве расстояния в Z^2 применяется евклидово расстояние d . Габаритными размерами объекта называются длины сторон прямоугольника наименьшей площади, описанного вокруг его проекции, а диаметром – диаметр его проекции A , определяемый обычным образом:

$$d(A) = \max_{a \in A, b \in A} d(a, b).$$

Легко видеть, что для любого $t \in Z^2$

$$\inf_{z \in Z^2 \setminus \{t\}} d(t, z) = 1.$$

Поэтому пиксели с координатами z и t , для которых $d(t, z) = 1$, называются соседями. Очевидно, что каждый пиксель имеет четырех соседей.

Для множества $A \subset Z^2$ его границей $Fr(A)$ будет называться совокупность точек из A , каждая из которых имеет хотя бы одного соседа из $Z^2 \setminus A$. Подмножество A из Z^2 , $|A| \geq 2$, будет называться связным, если для любых двух пикселей с координатами z и t из A таких, что $t \neq z$, существует последовательность из

$n \geq 2$ пикселей $(w_j)_{1 \leq j \leq n}$ из A , для которой $z = w_1$, $t = w_n$ и $d(w_j, w_{j+1}) = 1$, $1 \leq j \leq n$. Предполагается, что проекции рассматриваемых в настоящей работе объектов, являются связными.

Понятие соседства, определенное для пикселей, переносится естественным образом на объекты. Если ξ_A и ξ_B – объекты, то расстоянием между ними будет называться расстояние $d(A, B)$ между их проекциями:

$$d(A, B) = \min_{a \in A, b \in B} d(a, b) \geq 1.$$

Объекты ξ_A и ξ_B назовем соседями, если $d(A, B) = 1$.

При поиске на сцене заданных объектов по геометрическим признакам необходимо знать их проекции. Для нахождения проекций в настоящей работе предлагается выполнять сегментацию не всей сцены, а только тех ее фрагментов, которые, кроме самого объекта, содержат его некоторое окружение и называются зонами интереса. С формальной точки зрения, определение зоны интереса выглядит следующим образом. Пусть A – связное подмножество Z^2 с диаметром $d(A)$, $\xi_A = (\xi_a)_{a \in A}$ – один из заданных (искомых) объектов. Очевидно, что для любого $l \geq d(A) + 2$ существует квадрат C на Z^2 со стороной, равной l , такой, что $A \subset (C \setminus Fr(C))$. Будем называть семейство случайных величин $\xi_C = (\xi_c)_{c \in C}$ зоной интереса объекта ξ_A , если существует такой объект ξ_B , что $(C \setminus A) \subset B$. Далее семейство случайных величин $\xi_{C \setminus A} = (\xi_z)_{z \in C \setminus A}$ будет называться окрестностью объекта ξ_A .

Перечислим свойства сцены, наличие которых далее будет предполагаться.

Во-первых, будем считать, что случайные величины, образующие сцену, взаимно независимы, а случайные величины, образующие один и то же объект, кроме того, имеют одно и то же распределение. Если A – проекция некоторого объекта и $a \in A$, то для распределения случайной величины ξ_a будет использоваться обозначение $P_A = (p_A(y))_{y \in Y}$. Не предполагается наличие каких-либо сведений о распределении P_A . Однако у соседних объектов ξ_A и ξ_B их распределения P_A и P_B должны отличаться.

Во-вторых, предполагается, что известны геометрические признаки заданных (искомых) объектов: площадь, габаритные размеры, диаметр, а также минимальное расстояние d_{\min} между ними. Кроме того, для каждого заданного объекта ξ_A существует зона интереса ξ_C . Из ее определения и первого предположения следует, что случайные величины, образующие окрестность $\xi_{C \setminus A}$ объекта, должны иметь одно и тоже распределение, которое будет обозначаться символом $P_{C \setminus A} = (p_{C \setminus A}(y))_{y \in Y}$.

2. Сегментация зоны интереса

Очевидно, что сторона l зоны интереса заданного объекта ξ_A ограничена как снизу, так и сверху. В самом деле, если $d(A)$ – диаметр этого объекта, то нижняя граница l_1 принимает вид $l_1 = d(A) + 2$.

Для вычисления верхней границы l_2 предположим сначала, что проекция объекта имеет форму круга. Тогда сторона l зоны интереса должна удовлетворять неравенству

$$l^2 \leq (d(A) + d_{\min})^2,$$

вытекающему из теоремы Пифагора. Оно означает, что при сделанном предположении о форме проекции верхняя оценка имеет вид

$$l_2 = \sqrt{2} \left(d_{\min} + \frac{d(A)}{2} \right).$$

В общем случае предыдущие рассуждения также проходят, если учесть, что вокруг каждой проекции можно описать круг.

Так как $l_1 \leq l_2$, то

$$d_{\min} \geq \frac{d(A)(2 - \sqrt(2)) + 4}{2\sqrt{2}}.$$

Пусть z_0 – произвольный узел Z^2 , $e_1 = (1, 0)$ и $e_2 = (0, 1)$ – единичные векторы, а i и j – целые числа. Для произвольного объекта ξ_A выберем l из $[l_1, l_2]$ и построим на Z^2 семейство квадратов со стороной l и левой верхней вершиной z , определяемой равенством

$$z = z_0 + i\Delta e_1 + j\Delta e_2, \quad i \in Z, \quad j \in Z, \quad 1 \leq \Delta \leq l - d(A) - 1.$$

В этом случае справедливо следующее утверждение.

Теорема 1. *Пусть A – проекция объекта сцены с диаметром $d(A)$, неизвестными координатами и неизвестной ориентацией. Тогда существует квадрат C из описанного выше семейства такой, что $A \subset (C \setminus Fr(C))$.*

Доказательство теоремы приводится в пункте 5.

Если w и h – ширина и высота прямоугольного фрагмента сцены, а l – сторона зоны, то максимальное количество квадратов $n(w, h)$ из этого фрагмента, содержащихся в определенном выше семействе, равняется

$$n(w, h) = \left(\left[\frac{w-l}{\Delta}\right] + 1\right) \left(\left[\frac{h-l}{\Delta}\right] + 1\right).$$

Сформулированное утверждение подсказывает способ поиска объектов. Во-первых, по диаметру $d(A)$ каждого объекта и величине d_{\min} – минимальному из расстояний между всевозможными парами объектов – необходимо выбрать длину l стороны зоны интереса и величину Δ . Во-вторых, для каждого квадрата из полученного семейства требуется выполнить сегментацию, затем вычислить по полученной проекции значения геометрических признаков и на их основе принять решение о наличии объекта в рассматриваемом квадрате.

С другой стороны, так как зоны интереса отличаются от пустых квадратов (в которых нет объектов) таких же размеров, то зоны интереса можно рассматривать в качестве вспомогательных элементов сцены и заменить их перебор поиском. В этом случае второй операцией будет классификация квадратов указанного семейства на зоны интереса и пустые квадраты. Сегментацию и все оставшиеся операции требуется выполнять только в зонах интереса. Предполагая, что зоны интереса найдены (см., например, [11]), перейдем к решению задачи их сегментации.

При известном изображении $x_C = (x_c)_{c \in C}$ зоны интереса задачу сегментации можно сформулировать как задачу классификации ее пикселей. В самом деле, в качестве признака пикселя с координатами $z \in C$ рассматривается зарегистрированное в ходе съемки значение $x_z \in Y$. Из определения зоны интереса следует, что любой ее пиксель принадлежит к одному из двух классов: объекту ξ_A (класс 1) или его окрестности $\xi_{C \setminus A}$ (класс 2). Если Y_A – некоторое подмножество признакового пространства Y , а $Y_{C \setminus A} = Y - Y_A$, то отображение $h : Y \rightarrow \{1, 2\}$, определяемое равенством вида

$$h(y) = \begin{cases} 1, & y \in Y_A, \\ 2, & y \in Y_{C \setminus A}, \end{cases}$$

является решающим правилом для классификации пикселей зоны интереса на два класса. Оно предписывает отнести пиксель с координатами $z \in C$ к объекту ξ_A , если $x_z \in Y_A$, и к его окрестности $\xi_{C \setminus A}$ – в противном случае.

Наибольший интерес представляет байесовское решающее правило h^* , для которого вероятность $e(h^*)$ ошибки классификации является минимальной. Как известно [12, с. 42], для его применения требуется знать априорные вероятности $P(A)$ и $P(C \setminus A)$ принадлежности пикселя зоны интереса к объекту или к его окрестности, а также условные распределения $P_A = (p_A(y))_{y \in Y}$ и $P_{C \setminus A} = (p_{C \setminus A}(y))_{y \in Y}$, описывающие изменчивость признаков объекта и его окрестности. К сожалению, в явном виде эта информация, как правило, отсутствует. Однако при сделанных предположениях о свойствах сцены ее можно получить из изображения x_C зоны интереса, которую в этом случае следует рассматривать в качестве выборки объема $|C|$. Покажем, как это можно сделать.

Во-первых, по предположению, площадь $|A|$ объекта считается известной. Площадь $|C|$ зоны интереса выбирается на этапе дешифрирования по $d(A)$ и минимальному расстоянию между объектами d_{\min} . Поэтому для определения априорных вероятностей $P(A)$ и $P(C \setminus A)$ классов можно воспользоваться следующими равенствами

$$P(A) = \frac{|A|}{|C|}, \quad P(C \setminus A) = \frac{|C \setminus A|}{|C|}. \quad (1)$$

Пусть $n_A(y)$, $n_{C \setminus A}(y)$ и $n_C(y)$ – количество пикселей объекта, его окрестности и всей зоны интереса соответственно, со значением $y \in Y$, присутствующих на изображении x_C зоны интереса. Очевидно, что $n_C(y)$, $n_A(y)$ и $n_{C \setminus A}(y)$ связаны уравнением

$$n_C(y) = n_A(y) + n_{C \setminus A}(y). \quad (2)$$

Для любого $y \in Y$ определим $\bar{p}_A(y)$, $\bar{p}_{C \setminus A}(y)$ и $\bar{p}_C(y)$ равенствами

$$\bar{p}_A(y) = \frac{n_A(y)}{|A|}, \quad \bar{p}_{C \setminus A}(y) = \frac{n_{C \setminus A}(y)}{|C \setminus A|}, \quad \bar{p}_C(y) = \frac{n_C(y)}{|C|}.$$

Тогда уравнение (2) можно переписать в виде

$$\bar{p}_C(y) = P(A)\bar{p}_A(y) + P(C \setminus A)\bar{p}_{C \setminus A}(y), \quad y \in Y. \quad (3)$$

Известно [13, с. 221], что относительные частоты $\bar{p}_A(y)$ и $\bar{p}_{C \setminus A}(y)$ являются несмешенными и состоятельными оценками неизвестных вероятностей $p_A(y)$ и $p_{C \setminus A}(y)$, образующих условные распределения P_A и $P_{C \setminus A}$ классов.

Из определения зоны интереса следует, что ее граница является частью окрестности объекта. Поэтому ее изображение x_{Fr} также можно использовать для получения оценок

$$\bar{p}_{Fr}(y) = \frac{n_{Fr}(y)}{|Fr|}, \quad y \in Y,$$

неизвестных вероятностей распределения $P_{C \setminus A}$. После замены в уравнении (3) неизвестной оценки $\bar{p}_{C \setminus A}(y)$ на оценку $\bar{p}_{Fr}(y)$, вычисленную по изображению границы, оно превращается в уравнение с одним неизвестным. Для каждого $y \in Y$ его решением является оценка $\bar{p}_A(y)$ неизвестной вероятности $p_A(y)$. Таким образом, вместо неизвестных условных распределений P_A и $P_{C \setminus A}$ для классификации пикселей зоны интереса можно использовать

$$\bar{P}_A(y) = (\bar{p}_A(y))_{y \in Y}, \quad \bar{P}_{C \setminus A}(y) = (\bar{p}_{C \setminus A}(y))_{y \in Y}.$$

Сказанное позволяет предложить для классификации пикселей зоны интереса решающее правило $\bar{h}^* : Y \rightarrow \{1, 2\}$, вид которого определяется следующим утверждением.

Теорема 2. Пусть априорные вероятности объекта и его окрестности заданы по формулам (1),

$$\bar{P}_A = (\bar{p}_A(y))_{y \in Y}, \quad \bar{P}_{C \setminus A} = (\bar{p}_{C \setminus A}(y))_{y \in Y}$$

есть условные распределения признаков, описывающих объект и его окрестность, а Y_A и $Y_{C \setminus A}$ – подмножества из Y , определяемые равенствами вида

$$Y_A = \{y \in Y : |A| \bar{p}_A(y) \geq |C \setminus A| \bar{p}_{C \setminus A}(y)\},$$

$$Y_{C \setminus A} = Y \setminus Y_A.$$

Тогда решающее правило $\bar{h}^* : Y \rightarrow \{1, 2\}$ вида

$$\bar{h}^*(y) = \begin{cases} 1, & y \in Y_A, \\ 2, & y \in Y_{C \setminus A} \end{cases}$$

является байесовским с вероятностью ошибки $e(\bar{h}^*)$, определяемой равенством

$$e(\bar{h}^*) = \frac{|A|}{|C|} \sum_{y \in Y_{C \setminus A}} \bar{p}_A(y) + \frac{|C \setminus A|}{|C|} \sum_{y \in Y_A} \bar{p}_{C \setminus A}(y).$$

Сформулированное утверждение следует из теоремы Байеса с учетом вида априорных вероятностей и условных распределений обоих классов.

3. Сегментация модельных сцен

Для тестирования предлагаемых в работе методов и иллюстрации их возможностей используется фрагмент сцены, проекцией которого является квадрат со стороной, равной 256 пикселям. Четыре объекта фрагмента с номерами 1, 2, 3, 4 имеют прямоугольные проекции со сторонами 30 и 40 пикселей, параллельными осям координат. Все оставшиеся пиксели образуют объект сцены с номером нуль. По традиции он называется фоном. Яркость каждого пикселя сцены является случайной величиной с нормальным распределением. Средние значения объектов и фона приведены в табл. 1. Дисперсии всех элементов сцены одинаковы и равны 225 единицам. В этой же таблице для каждого объекта сцены приведено значение k_j , определяемое равенством

$$k_j = \frac{|m_0 - m_j|}{\sigma_0 + \sigma_j}, \quad j = 1, 2, 3, 4, \tag{4}$$

и называемое отношением сигнал/шум. Ниже показано, что вероятность ошибки сегментации с ростом k_j быстро уменьшается. Для сцены с описанными свойствами выполняются все предположения математической модели. Поэтому она будет называться далее модельной.

Для сравнения фактических результатов сегментации с теоретическими для каждого объекта построена прямоугольная зона интереса площадью 2400 пикселей. Легко видеть, что в этом случае $P(A) = P(C \setminus A) = 0.5$ и вероятность $e(h^*)$ принимает вид

$$e(h^*) = \frac{1}{\sqrt{2\pi}} \int_{|m_0 - m_j|/2\sigma}^{+\infty} \exp\left(-\frac{t^2}{2}\right) dt, \quad j = 1, 2, 3, 4.$$

Табл. 1

Свойства модельной сцены

	Элементы сцены				
	№ 0	№ 1	№ 2	№ 3	№ 4
m_j	165	150	135	120	105
k_j	—	0.5	1.0	1.5	2

Очевидно, что результат сегментации определяется в основном двумя обстоятельствами. Прежде всего это степень сходства между объектом и фоном. При полном сходстве, которое описывается равенством $P_A = P_{C \setminus A}$, вероятность ошибки $e(h^*)$ определяется только априорными вероятностями классов

$$e(h^*) = \min(P(A), P(C \setminus A)) \leq 1/2.$$

Этот случай можно рассматривать как неудачный выбор признака (или спектральной зоны при использовании оптических изображений). Вторым обстоятельством, от которого также зависит результат сегментации, является абсолютная погрешность оценок, образующих распределение $\bar{P}_{Fr} = (\bar{p}_{Fr}(y))_{y \in Y}$. Так как

$$E\bar{p}_{Fr}(y) = p_{C \setminus A}(y), \quad D\bar{p}_{Fr}(y) = \frac{p_{C \setminus A}(y)(1 - p_{C \setminus A}(y))}{|Fr(C)|} \leq \frac{1}{4|Fr(C)|},$$

то из неравенства Чебышева следует, что

$$P(|\bar{p}_{Fr}(y) - p_{C \setminus A}(y)| \geq \varepsilon) \leq \frac{D\bar{p}_{Fr}(y)}{\varepsilon^2} = \frac{p_{C \setminus A}(y)(1 - p_{C \setminus A}(y))}{\varepsilon^2 |Fr(C)|}.$$

Изменить вероятности условного распределения $P_{C \setminus A}$ на этапе дешифрирования невозможно. Поэтому уменьшить погрешность ε оценки можно только за счет увеличения объема выборки $|Fr(C)|$. Из последнего неравенства следует, что для уменьшения погрешности ε в t раз объем $|Fr(C)|$ выборки требуется увеличить в t^2 раз.

Табл. 2

Сегментация модельной сцены

	Элементы сцены			
	№ 1	№ 2	№ 3	№ 4
$e(h)$	0.308	0.159	0.067	0.023
$n = 196$	0.395	0.220	0.146	0.070
$n = 5440$	0.355	0.161	0.076	0.027

Для иллюстрации сказанного была проведена сегментация зон интереса модельной сцены с использованием выборок разного объема. В табл. 2 для каждой из четырех зон интереса указаны байесовская вероятность ошибки $e(h^*)$ и относительные частоты ошибок классификации (сегментации), подсчитанные для двух выборок.

В качестве малой выборки использовалось изображение x_{Fr} границы зоны интереса, состоящей из $n = 196$ пикселей. Большая выборка состояла из $n = 5440$ пикселей. Кроме пикселей границы, она включала дополнительные пиксели, расположенные по соседству с зоной интереса. Из представленных результатов следует, что использование большой выборки позволило значительно снизить относительную частоту ошибок классификации в каждой зоне интереса, приблизив ее к теоретической. На рис. 1 приведено изображение фрагмента модельной сцены, а на

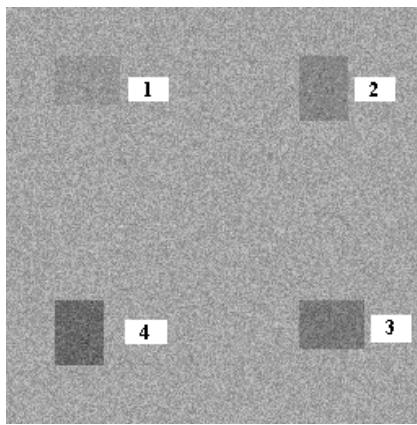


Рис. 1. Изображение модельной сцены

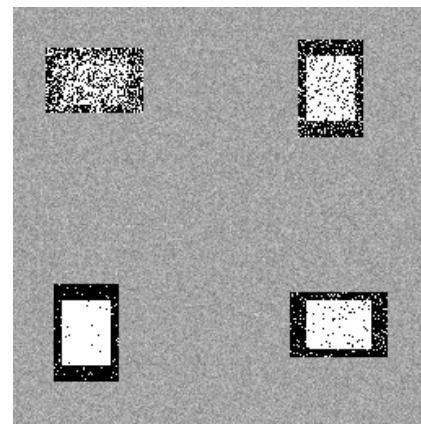


Рис. 2. Результаты сегментации модельной сцены

рис. 2 – результаты сегментации зон интереса, соответствующие большой выборке. На нем пиксели, отнесенные к первому классу (объект), закрашены максимальной яркостью. Пиксели, отнесенные ко второму классу (фон), закрашены темным цветом.

4. Сегментация реальных сцен

Очевидно, что предположения о независимости и однородности признаков, использованные при построении модели зоны интереса, для реальных сцен в полном объеме не выполняются. Поэтому для проверки предлагаемого метода сегментации в условиях, близких к реальным, в качестве исходной информации использовались изображения участков земной поверхности, полученные в различных спектральных зонах оптического диапазона. Участки поросли травой, редким кустарником и отдельными деревьями. Кроме того, на них присутствуют следы колесных и гусеничных транспортных средств. По условиям съемки сторона квадратного пикселя на местности должна равняться 0.3 м.

При таком масштабе индивидуальные особенности конструкций транспортных средств на изображениях отсутствуют, а форма их проекций приближается к прямоугольной. Последнее обстоятельство позволяет размещать (дорисовать) на изображениях фрагментов сцены прямоугольники нужных размеров и рассматривать их как изображения транспортных средств. Координаты и ориентация таких прямоугольников выбираются случайным образом. Их размеры совпадают с размерами реальных транспортных средств, присутствующих на сцене, а яркости являются реализациями случайных величин с нормальным законом распределения. В качестве его среднего значения и дисперсии используются оценки соответствующих параметров, вычисленные по изображениям реальных объектов. Далее такие сцены будут называться реальными. В качестве примера на рис. 3 представлено изображение участка земной поверхности с проекцией, имеющей форму квадрата со стороной, равной 256 пикселям. Оно сформировано оптико-электронной системой в спектральной зоне [0.7, 1.1] мкм инфракрасного диапазона, а затем на нем были дорисованы четыре прямоугольника с номерами 1, 2, 3 и 4.

Вначале по изображению сцены был проведен автоматический поиск зон интереса. Было обнаружено восемь зон, включая четыре ложных. Их границы на рис. 4 выделены сплошными линиями.

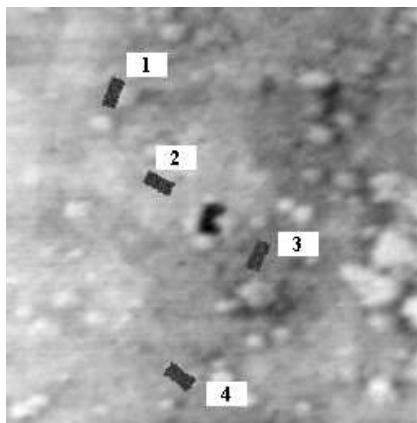


Рис. 3. Изображение реальной сцены

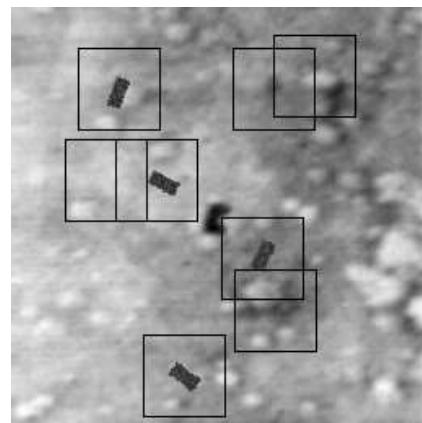


Рис. 4. Результаты поиска зон интереса

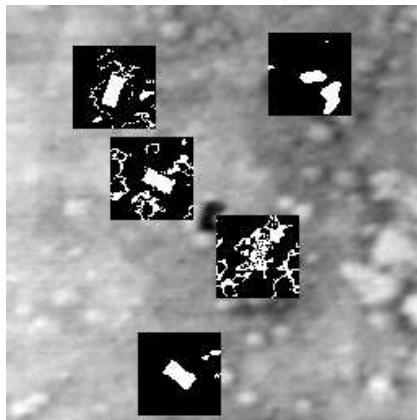


Рис. 5. Результаты сегментации реальной сцены

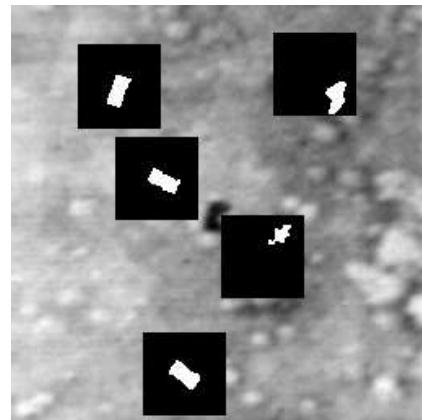


Рис. 6. Результаты коррекции

Далее в четырех зонах, содержащих объекты, и в одной ложной, взятой для примера, была проведена сегментация. Ее результаты представлены на рис. 5. Белым цветом закрашены пиксели, отнесенные к объекту, черным – отнесенные к фону. Как и следовало ожидать, в каждой зоне интереса оказалось несколько изолированных компонент. Поэтому перед измерением габаритных размеров была проведена коррекция результатов сегментации. Она включала удаление линий шириной в один пиксель, изолированных компонент и компонент, имеющих общие пиксели с границей зоны интереса. Окончательные результаты, полученные после выполнения операции коррекции, представлены на рис. 6. Из них следует, что в третьей зоне построить проекцию объекта не удалось.

Для сведения в табл. 3 для каждого дорисованного объекта приведено значение отношения сигнал/шум и результаты измерения габаритных размеров. Длина и ширина «объекта» из ложной зоны составили 18 и 12 пикселей соответственно. Плановые значения габаритных размеров равняются 19 и 9 условным единицам.

Эксперименты с реальными сценами показали, что при значительных отклонениях статистических свойств границы от свойств остальной части окрестности вероятность ошибки сегментации может стать недопустимо большой (зона № 3).

Табл. 3

Сегментация реальной сцены

	Элементы сцены			
	№ 1	№ 2	№ 3	№ 4
k_j	2.9	4.6	2.3	2.8
Длина	18	17	12	18
Ширина	8	8	11	9

5. Доказательство теоремы 1

Пусть a_1 – горизонтальная координата самого левого пикселя, a_2 – вертикальная координата самого верхнего пикселя проекции A . Существует квадрат из определенного выше семейства, у которого координаты $c = (c_1, c_2)$ левого верхнего угла удовлетворяют следующим неравенствам

$$c_1 < a_1 \leq c_1 + \Delta, \quad c_2 < a_2 \leq c_2 + \Delta.$$

В этом случае горизонтальная координата $a_1 + d(A)$ самого правого пикселя объекта ξ_A удовлетворяет неравенству

$$a_1 + d(A) \leq c_1 + l - d(A) - 1 - d(A) = c_1 + l - 1,$$

а вертикальная координата $a_2 + d(A)$ самого нижнего пикселя объекта – неравенству

$$a_2 + d(A) \leq c_2 + l - d(A) - 1 - d(A) = c_2 + l - 1,$$

Это означает, что $A \subset (C \setminus Fr(A))$. Теорема доказана.

Заключение

В работе предложена формализация и решение задачи сегментации зон интереса скалярных сцен на основе байесовской классификации. Приведены результаты компьютерных экспериментов с использованием изображений модельных и реальных сцен, подтверждающие зависимость вероятности ошибки сегментации от качества изображения, описываемого отношением сигнал/шум, и точности оценки неизвестных вероятностей. Изложенные результаты могут использоваться при решении прикладных задач, связанных с автоматическим поиском объектов по скалярному изображению сцены.

Авторы приносят благодарность директору – главному конструктору ЗАО «Научно-производственная фирма Оптоойл» Р.М. Алееву за моральную и финансовую поддержку.

Summary

V.B. Fofanov, R.G. Sabirov. On Bayesian segmentation of zone of interests of scalar scenes.

We consider a problem on the segmentation of zones of interest which represent scene fragments, containing a given object together with its certain neighborhood. We state the problem as that of the Bayesian classification of pixels of a zone of interest into two classes. We prove that for the estimation of unknown a priori probabilities and conditional distributions of classes it is sufficient to know the area of the object and to have an image of the zone of interest. The results of the segmentation of zones of interest in model and real scenes by the proposed method are offered.

Литература

1. *Бакут П.А., Колмогоров Г.С.* Сегментация изображений: методы выделения границ областей // Зарубежная радиоэлектроника. – 1987. – № 10. – С. 25–43.
2. *Бакут П.А., Колмогоров Г.С., Ворновицкий И.Э.* Сегментация изображений: методы прогрессивной обработки // Зарубежная радиоэлектроника. – 1987. – № 10. – С. 6–24.
3. *Борисенко В.И., Златопольский А.А., Мучник И.Б.* Сегментация изображений (составление, проблемы) // Автоматика и телемеханика. – 1987. – № 7. – С. 3–56.
4. *Денисов Д.А., Низовкин В.А.* Сегментация изображений на ЭВМ // Зарубежная радиоэлектроника. – 1985. – № 10. – С. 5–30.
5. *Bourylin S.A., Fofanov V.B.* Generalized Spot Criterion as Applied for Image Deciphering // Pattern Recognition and Image Analysis. – 2004. – No 2. – C. 243–248.
6. *Верденская Н.В.* Сегментация изображений — статистические модели и методы // Успехи современной радиоэлектроники. – 2002. – № 12. – С. 33–47.
7. *Грузман И.С., Новиков К.В.* Сегментация анизотропных изображений на основе локальных спектральных характеристик // Автометрия. – 2004. – № 4. – С. 26–32.
8. *Райфельд М.А.* Ранговая сегментация цветных изображений // Автометрия. – 2001. – № 1. – С. 21–26.
9. *Щербаков А.П.* Быстро действующий алгоритм сегментации изображений // Автометрия. – 2005. – № 2. – С. 59–67.
10. *Фофанов В.Б.* О теоретико-вероятностной формализации задачи дешифрирования аэрокосмических изображений // Автометрия. – 2003. – № 6. – С. 107–118.
11. *Фофанов В.Б., Демченко А.В.* Об одном подходе к поиску зон интереса // Материалы IX междунар. конф. «Интеллектуальные системы и компьютерные науки», 23–27 окт. 2006 г. – М.: Изд-во мех.-матем. фак. Моск. гос. ун-та, 2006. – Т. 2, ч. 2. – С. 291–292.
12. *Дуда Р., Харт П.* Распознавание образов и анализ сцен. – М.: Мир, 1976. – 512 с.
13. *Крамер Г.* Математические методы статистики. – М.: ИЛ, 1948. – 648 с.

Поступила в редакцию
17.08.07

Сабиров Раиль Гарифзянович – аспирант кафедры экономической кибернетики Казанского государственного университета.

Фофанов Вячеслав Борисович – кандидат технических наук, доцент кафедры экономической кибернетики Казанского государственного университета.

E-mail: *Viatcheslav.Fofanov@ksu.ru*