РОССИЙСКАЯ АКАДЕМИЯ НАУК

СИБИРСКОЕ ОТДЕЛЕНИЕ

АВТОМЕТРИЯ

2003, том 39, № 6

УДК 519.681

В. Б. Фофанов

(Казань)

О ТЕОРЕТИКО-ВЕРОЯТНОСТНОЙ ФОРМАЛИЗАЦИИ ЗАДАЧИ ДЕШИФРИРОВАНИЯ АЭРОКОСМИЧЕСКИХ ИЗОБРАЖЕНИЙ *

Автоматизация дешифрирования монохроматических изображений независимо от их содержания требует обязательного решения ряда задач. Главными из них являются описание реальной сцены, определение объема информации, которую можно получить о сцене по ее изображению, выбор признаков и сегментация. В предлагаемой работе в рамках теоретико-вероятностного подхода предпринята попытка исследования математических и методических вопросов, возникающих при решении перечисленных задач.

Введение. Изображения являются источником информации о реальной сцене. Основным методом их дешифрирования (анализа) до сих пор продолжает оставаться визуальный метод. Исследования по его автоматизации интенсивно проводятся во многих странах. В случае успеха полученные результаты смогут найти применение в самых разных предметных областях.

К сожалению, общепринятой теории дешифрирования изображений, которая может служить основой для решения прикладных задач из разных предметных областей, пока не существует. Успехи в распознавании печатных символов и идентификации по папиллярным узорам достигнуты благодаря использованию специфических методов, а не глубоких теоретических результатов. Это можно оправдать огромным разнообразием свойств реальных сцен, затрудняющих построение единой теории, и относительно короткой историей существования данного научно-технического направления. Тем не менее можно ожидать появления результатов, которые будут иметь принципиальное значение как для самой теории, так и ее приложений.

В литературе привлекаются различные теоретико-вероятностные понятия для описания отдельных задач, возникающих в ходе дешифрирования изображений. В предлагаемой работе делается попытка обобщения известных и оригинальных результатов исследований по теоретико-вероятностной формализации задачи дешифрирования монохроматических изображений

* Работа выполнена при финансовой поддержке научно-производственной фирмы «ОптоОйл».

реальных сцен, формируемых в оптическом диапазоне из атмосферы или космического пространства. При этом основное внимание уделяется описанию с единых позиций всего комплекса взаимосвязанных проблем, составляющих единую задачу дешифрирования изображений. Основными ориентирами в этом направлении будут служить описание реальной сцены, возможности дешифрирования, определение признаков, а также задача сегментации.

При решении задачи дешифрирования реальную сцену удобно рассматривать как объединение некоторых элементов. Состав этих элементов зависит от конкретной решаемой задачи. Например, в военных приложениях обязательными элементами сцены должны быть военная техника и вооружение, а также другие объекты искусственного и естественного происхождения, от которых зависит превосходство в вооруженном конфликте. При обновлении топографических карт в качестве элементов сцены принято рассматривать компоненты рельефа, растительного покрова, гидросферы, населенные пункты, стационарные инженерные сооружения и другие природные и социально-экономические объекты местности. В рамках вероятностного подхода каждый элемент сцены описывается конечным набором случайных величин. Их совместное распределение определяется свойствами соответствующего элемента сцены и должно отличаться от распределений других ее элементов. В качестве описания (модели) сцены естественно рассматривать множество случайных величин, задающих все ее элементы. Обычно это множество является бесконечным, т. е. случайным полем. В соответствии с теоремой Колмогорова случайное поле задается при помощи семейства согласованных конечномерных распределений, построение которого в общем случае является довольно сложной задачей. Однако если отказаться от зависимости между случайными величинами, описывающими разные элементы сцены, то построение такого семейства упрощается. Этот вопрос рассмотрен в разд. 1 предлагаемой работы.

При решении конкретных задач, как правило, возникает вопрос о том, на какую информацию о сцене можно рассчитывать в результате дешифрирования ее изображения. Очевидно, что ответ на него можно дать только при наличии математического описания сцены. С обсуждения этого вопроса начинается разд. 2 данной работы. В нем же формулируются достаточные условия, при выполнении которых отображение (признак), определенное на множестве всех изображений, является случайной величиной. Необходимость в подобном критерии очевидна, поскольку при вероятностном подходе полагают, что признаки являются случайными величинами. С другой стороны, хорошо известно, что не всякое отображение является случайной величиной. Сформулированный критерий применяется далее для доказательства того, что некоторые часто используемые признаки на самом деле являются случайными величинами.

Несмотря на большое количество работ, посвященных проблеме сегментации (см., например, [1–3]), в ней все еще остаются неисследованные вопросы. В разд. 3 для элементов сцен с известными площадью и диаметром формально определяется очень важное понятие – зона интереса. Оно позволяет формулировать и решать задачу сегментации как задачу классификации с учителем.

1. **Описание сцены.** Рассмотрим реальную сцену как совокупность неделимых элементов, называемых далее пикселами. Каждый пиксел характе-

ризуется индивидуальными целочисленными координатами $z=(z_1, z_2)$ на двумерной целочисленной решетке

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

и неотрицательной случайной величиной ξ_z, принимающей значения из конечного множества $Y=\{0,1,\dots,|Y|-1\}$, $|Y|>1$. Она описывает некоторое свойство пиксела, точное значение которого становится известным только после его измерения (съемки). Очевидно, что все случайные величины должны задаваться на одном и том же вероятностном пространстве (Ω, A, P). Другими словами, сцена является случайным полем $\xi=(\xi_z)_{z\in Z^2}$.

Пусть ω – некоторое элементарное событие из Ω. Если каждую случайную величину ξ_z, $z\in Z^2$, заменить ее выборочным значением $x_z=\xi_z(\omega)$, то совокупность $x=(x_z)_{z\in Z^2}$ всех выборочных значений, соответствующих ω, естественно рассматривать как одно из возможных изображений сцены.

Как правило, интерес представляют не отдельные пикселы, а их совокупности, которые далее будут называться элементами сцены. Каждый такой элемент определяется подмножеством A точек из Z^2, являющихся координатами его пиксел, и $|A|$-мерной случайной величиной вида $\xi_A=(\xi_a)_{a\in A}$. Поэтому все свойства элемента сцены описываются распределением вероятностей P_{Y^A} на множестве всех подмножеств из Y^A, которое однозначно задается $|Y|^{|A|}$ совместными вероятностями $p_{Y^A}(x_A)$, $x_A=(x_a)_{a\in A}\in Y^A$, вида

$$p_{Y^A}(x_A)=P(\xi_A=x_A).$$

Очевидно, что

$$p_{Y^A}(x_A)\ge 0,\quad x_A\in Y^A;\qquad \sum_{x_A\in Y^A} p_{Y^A}(x_A)=1.$$

Кроме того, если A и B – конечные подмножества на Z^2 и $A\subset B$, то для любого $x_A\in Y^A$ должно выполняться равенство

$$p_{Y^A}(x_A)=\sum_{x_{B\setminus A}\in Y^{B\setminus A}} p_{Y^B}(x_A,\ x_{B\setminus A}),$$

называемое условием согласованности. Далее подмножество A будет называться проекцией элемента сцены, а сужение $x_A=(x_a)_{a\in A}$ изображения $x=(x_z)_{z\in Z^2}$ всей сцены на проекцию A – его изображением. В то же время в соответствии с теоремой Колмогорова для задания случайного поля на Z^2 достаточно построить на ней семейство конечномерных распределений. Покажем, как это можно сделать в предположении, что проекции A и B разных элементов сцены не пересекаются и для любых $a\in A$ и $b\in B$ случайные величины ξ_a и ξ_b независимы. Другими словами, справедливо следующее утверждение.

Теорема 1. Пусть Z^2 разбита на конечные попарно не пересекающиеся подмножества – проекции элементов сцены, и пусть каждой проекции $A \subset Z^2$ поставлено в соответствие распределение вероятностей P_{Y^A} на Y^A. Тогда существует случайное поле $\xi = (\xi_z)_{z \in Z^2}$ на Z^2 такое, что

$$P(\xi_A = x_A) = p_{Y^A}(x_A)$$

для любого $x_A \in Y^A$ и любой проекции A. Кроме того, если A и B – проекции разных элементов сцены, то для любых $a \in A$ и $b \in B$ случайные величины ξ_a и ξ_b независимы.

Доказательство теоремы приведено в разд. 4. Из теоремы следует, что совместные распределения P_{Y^A} вероятностей, описывающие отдельные элементы сцены, можно задавать независимо друг от друга и что к ним не предъявляется никаких дополнительных требований. Их свойства определяются имеющейся информацией о соответствующих элементах сцены. Например, для описания лесного участка с проекцией A можно применить набор случайных величин с марковским распределением. Для описания элемента сцены с недостаточно изученными свойствами можно ограничиться набором случайных величин, задаваемых на уровне числовых характеристик.

Следует отметить, что наиболее часто в качестве описания реальных сцен применяются однородные в широком смысле (далее просто однородные) и марковские случайные поля. Однородные случайные поля являются естественным обобщением стационарных в широком смысле случайных процессов. Впервые полное изложение этой теории, включая спектральное разложение, линейный прогноз и линейную фильтрацию, было дано, по-видимому, Ягломом в [4]. К сожалению, применение однородных случайных полей для описания реальных сцен наталкивается, по меньшей мере, на два серьезных препятствия. Они следуют из определения однородных случайных полей и поэтому являются неустранимыми. Во-первых, среднее значение такого поля должно быть одним и тем же для каждого пиксела. Это требование противоречит свойствам большинства реальных сцен, которые удобно рассматривать как объединение различных элементов. Во-вторых, однородное случайное поле задается средним значением и ковариационной функцией. Это делает бессмысленным вычисление вероятностей событий, описываемых в терминах случайных величин, включая, разумеется, и признаки. Тем не менее однородные случайные поля находят широкое применение как в дешифрировании, так и в моделировании изображений.

В 1968 г. Добрушиным в работе [5] был описан класс случайных полей, названных впоследствии марковскими. Однородное марковское случайное поле определяется конечным числом условных вероятностей специального вида и позволяет в отличие от однородного в широком смысле случайного поля вычислять вероятности событий, описываемых в терминах случайных величин. Отметим, что требование однородности для марковского случайного поля не является обязательным. Отказ от него приводит лишь к увеличению числа марковских вероятностей, необходимых для задания поля. Марковские вероятности предоставляют естественный способ описания локальных свойств реальных сцен, которые обычно называют текстурными признаками. По-видимому, первые попытки построения текстурных изображений на основе случайных полей марковского типа описаны в работах [6, 7]. В 1970 г. Аверинцевым в [8] было показано, что для определения положительных марковских вероятностей можно применить потенциалы, широко используемые в статистической физике для описания гиббсовских состояний.

Устойчивый интерес к указанным классам случайных полей при решении прикладных задач объясняется в значительной степени относительно небольшим объемом информации, который необходим для их определения. В связи с этим можно отметить работу [9], в которой исследуются другие подходы к формализации доступной информации о сцене при помощи конечного числа условных вероятностей. Помимо доказательства теорем существования и единственности в ней приводятся примеры применений предлагаемых классов случайных полей для решения некоторых прикладных задач, включая моделирование изображений.

Пусть A – проекция некоторого элемента сцены. Используя распределение вероятностей $P_A = (p_A(y))_{y \in Y}$ на Y, можно определить при помощи равенства

$$p_{Y^A}(x_A) = \prod_{a \in A} p_A(x_a), \qquad x_A \in Y^A,$$

распределение вероятностей P_{Y^A} на Y^A. Легко убедиться, что случайные величины $\xi_A = (\xi_a)_{a \in A}$, описываемые этим распределением, независимы и имеют одно и то же распределение. Сцены, состоящие из таких элементов, будут называться далее бернуллиевскими. Любой элемент такой сцены однозначно определяется проекцией A и распределением P_A на Y.

Пусть d_E: $Z^2 \times Z^2 \to R$ – евклидово расстояние на Z^2. Очевидно, что для любого $t \in Z^2$

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

Поэтому пикселы с координатами z и t, для которых $d_E(z, t) = 1$, будут далее называться соседями. Легко подсчитать, что каждый пиксел имеет четырех соседей. Определение соседства для пиксел обобщается естественным образом на произвольные элементы сцены. В самом деле, пусть $A \subset Z^2$ и $B \subset Z^2$ – проекции разных элементов. Будем называть их соседями, если $d_E(A, B) = 1$. Далее предположим, что два любых соседних элемента сцены с проекциями A и B имеют разные распределения, т. е. $P_{Y^A} \neq P_{Y^B}$.

Подмножество A на Z^2 будет называться связным, если для любых двух пиксел с координатами z и t из A можно указать $n \geq 2$ пиксел $(w_j)_{1 \leq j \leq n}$, принадлежащих A, и таких, что $z = w_1$, $t = w_n$ и $d_E(w_j, w_{j+1}) = 1$, $1 \leq j \leq n$. При крупномасштабной съемке характерными примерами элементов сцен, проекции которых являются связными множествами, могут служить отдельно стоящие строения и транспортные средства. В некоторых случаях интерес представляют группы элементов, например: населенные пункты, походные колонны, боевые порядки. Очевидно, что проекция группы не является связным множеством.

Требование связности может предъявляться также к дополнению $Z^2 \setminus A$ проекции A. При визуальном восприятии это означает, что сама проекция A не содержит «дыр».

2. **Дешифрирование и признаки.** Получение информации о сцене по ее изображению называется, как известно, его дешифрированием. При решении прикладных задач всегда возникает вопрос о том, какие сведения о сцене можно получить из ее изображения. В рамках вероятностного подхода дешифрирование естественно считать полным, если по изображению удалось восстановить семейство конечномерных распределений случайного поля. В соответствии с теоремой 1 это эквивалентно определению для каждого элемента сцены его проекции A и совместного распределения вероятностей P_{Y^A} на Y^A. Для бернуллиевских случайных полей вместо P_{Y^A} достаточно определить соответствующее распределение вероятностей P_A на Y.

Довольно часто решение конкретной задачи не требует проведения полного дешифрирования. Изображение может использоваться, например, для выяснения принадлежности каждого элемента сцены к одному из заранее определенных и непересекающихся классов. При таком подходе дешифрирование, казалось бы, можно интерпретировать как хорошо известную задачу, называемую классификацией [10]. Однако, несмотря на очевидное сходство, между ними имеется важное отличие. Оно заключается в отсутствии для элементов сцен конечных наборов характеристических свойств, образующих их векторы признаков. Поэтому получение (вычисление) их по имеющемуся изображению составляет задачу первого этапа дешифрирования. Рассмотрим этот вопрос более подробно.

Для выявления присутствующего на сцене элемента с проекцией A необходимо, чтобы свойства его изображения $x_A = (x_a)_{a \in A}$ отличались от свойств изображений его соседей. При дешифрировании, как правило, сравниваются не сами изображения элементов сцены, а вычисленные по ее изображению $x = (x_z)_{z \in Z^2}$ значения некоторых заранее выбранных отображений. Они называются признаками и описывают характерные свойства элементов сцены. Очевидно, что признаки должны обладать определенными полезными свойствами. При вероятностном подходе обычно полагают, что признаки являются случайными величинами на (Ω, A, P) – исходном вероятностном пространстве. По определению каждое изображение $x = (x_z)_{z \in Z^2}$ является реализацией случайного поля $\xi = (\xi_z)_{z \in Z^2}$, описывающего сцену. Множество $\xi(\Omega)$ всех изображений образует подмножество множества Y^{Z^2} всех отображений Z^2 в Y. Если f – некоторый признак, то он должен быть определен на $\xi(\Omega)$. В этом случае суперпозиция $f \circ \xi$, определяемая равенством $f \circ \xi(\omega) = f(\xi(\omega))$, является отображением вида $\Omega \to R$. Сформулируем достаточные условия, при выполнении которых $f \circ \xi$ будет случайной величиной. Для этого ограничимся только такими признаками, при вычислении значений которых используется конечное число пиксел из счетного множества всех пиксел, образующих сцену. Если A, $A \subset Z^2$, – координаты таких пиксел, то для любых двух изображений x' и x'', совпадающих на A: $x'_A = x''_A$, значения такого признака f_A должны совпадать:

$$f_A(x') = f_A(x'').$$

Теорема 2. Пусть $A=\{a_1,\ldots,a_{|A|}\}$ – конечное подмножество из Z^2; f_A – отображение вида $Y^{Z^2}\to R$, значения которого определяются пикселами с координатами из A. Если существует борелевское отображение $f: R^{|A|}\to R$ такое, что

$$f_A(x)=f(x_{a_1},\ldots,x_{a_{|A|}}),$$

то суперпозиция $f_A\circ\xi$ является случайной величиной на (Ω, A, P).

Доказательство теоремы приведено в разд. 4. С помощью теоремы можно показать, что класс отображений вида $Y^{Z^2}\to R$, удовлетворяющих указанным требованиям, достаточно широк. В самом деле, каждое непрерывное отображение является борелевским. Поэтому борелевскими являются, в частности, тождественное отображение $x\to x$, сумма $(x_1,x_2)\to x_1+x_2$, произведение $(x_1,x_2)\to x_1x_2$, частное $x\to 1/x$, а также отображение вида $x\to c$, где $c\in R$ – константа. На этом основании все три нижеприведенных отображения являются примерами признаков:

$$f_{\{z\}}(\xi_z)=\xi_z,\qquad f_A(\xi_A)=\frac{1}{|A|}\sum_{z\in A}\xi_z,\qquad f_A(\xi_A)=|A|.$$

Значение первого признака для каждого узла $z\in Z^2$ измеряется видеодатчиком в ходе выполнения съемки, а семейство значений, соответствующих всем узлам Z^2, образует изображение сцены. Значением второго признака является интенсивность элемента сцены, осредненная по площади его проекции. В последнем случае значение признака равняется площади проекции элемента сцены. Отметим, что значение этого признака не зависит от интенсивностей пиксел. Такие признаки называются геометрическими.

Рассмотрим менее очевидную ситуацию, которая возникает при сравнении случайных величин, описывающих свойства различных пиксел. Пусть $B(z,r)$ – круг на Z^2 с центром в точке $z\in Z^2$ и радиусом r. Предположим, что принятие некоторого решения зависит от результатов сравнения x_z с каждым из оставшихся значений x_t, $t\in B(z,r)\setminus\{z\}$. Поэтому в качестве значения признака удобно рассматривать количество точек из множества $B(z,r)\setminus\{z\}$, для которых выполняется неравенство $x_z-x_t\geq 0$. Легко убедиться в том, что индикатор (характеристическая функция) $E_{[0,+\infty[}$ множества $[0,+\infty[$ является борелевским отображением вида $R\to R$. Тогда суперпозиция разности $(x,y)\to(x-y)$ и индикатора $E_{[0,+\infty[}$ тоже будет борелевским отображением вида $R^2\to R$, что позволяет определить для каждой точки $z\in Z^2$ признак при помощи равенства

$$f_{B(z,r)}(\xi_{B(z,r)})=\sum_{t\in B(z,r)}E_{[0,+\infty[}(\xi_z-\xi_t).$$

Рассмотренные примеры позволяют надеяться, что предлагаемое определение признака окажется достаточным для удовлетворения потребностей самых разных приложений.

Для определения принадлежности элемента сцены к тому или иному классу необходимо вычислить значения одного или нескольких признаков, образующих его вектор признаков. Кроме того, необходимо иметь описания самих классов. Чаще всего для этого используют обучающую выборку или указывают для каждого класса распределение вероятностей на множестве признаков и априорную вероятность [11]. Очевидно, что для описания элементов сцены и их классов должны применяться одни те же признаки. Как свидетельствует практика, получаемые в ходе съемки значения x_z, $z \in Z^2$, образующие изображение $x = (x_z)_{z \in Z^2}$ сцены, достаточно сильно зависят от условий съемки: сезона, времени суток, метеоусловий, угла визирования, масштаба и других. К настоящему времени свойства этой зависимости полностью не выяснены. Поэтому сформулировать заранее описания классов в терминах измеряемого признака с необходимой точностью, как правило, не представляется возможным. По этой причине для классификации элементов сцены на втором этапе дешифрирования обычно применяют геометрические признаки, описывающие свойства проекций элементов сцены. Они более устойчивы по отношению к метеоусловиям, времени суток, сезону и некоторым другим условиям съемки. А их зависимость от угла визирования и ориентации элемента сцены во многих случаях удается рассчитать. Наиболее часто используются форма, площадь, периметр и некоторые другие признаки, но их окончательный состав можно определить только после объявления полного списка подлежащих выявлению элементов сцены.

3. **Сегментация.** Вычисление геометрических признаков элемента сцены возможно только после определения его проекции. Разбиение Z^2 на проекции образующих сцену элементов принято называть сегментацией. В соответствии с вышеприведенным предположением проекции разных элементов сцены не пересекаются. Кроме того, их объединение совпадает с Z^2. Если считать проекции элементов сцены классами, то сегментацию можно рассматривать как классификацию узлов решетки Z^2. К сожалению, отсутствие описаний классов в терминах регистрируемых признаков, а также сведений об их числе не позволяет в общем случае воспользоваться известной теорией. Однако, учитывая специфику дешифрирования, можно сделать несколько дополнительных предположений о свойствах элементов сцен, выполнение которых допускает формулировку задачи классификации (сегментации) в байесовской форме.

Предположим, во-первых, что дешифрирование проводится с целью выявления элементов сцены, имеющих известные площадь и диаметр проекции. Пусть A – проекция одного из таких элементов; $S(A)$ – ее площадь; $d_E(A)$ – диаметр в пикселах. В этом случае для любого l, $l \geq d_E(A)+2$, независимо от ориентации элемента существует квадрат C на Z^2 со стороной, равной l пикселам и такой, что $A \subset C \setminus Fr(C)$. Во-вторых, предположим, что евклидово расстояние r между проекциями элементов не меньше чем $(d_E(A)+2)\sqrt{2}$. Из этого следует существование хотя бы одного квадрата C со стороной l, $d_E(A)+2 \leq l \leq r/\sqrt{2}$, содержащего A и не содержащего проекций других, подлежащих выявлению, элементов. Наконец, будем считать, что все пикселы, образующие множество $C \setminus A$, принадлежат проекции одного и того же элемента сцены. Далее семейство случайных величин $\xi_C = (\xi_c)_{c \in C}$ с проекцией C, удовлетворяющей трем сформулированным выше условиям, будет называться зоной интереса для элемента сцены с проекцией A.

Теорема 3. Пусть $\xi = (\xi_z)_{z \in Z^2}$ – бернуллиевское случайное поле; $\xi_A = (\xi_a)_{a \in A}$ – элемент сцены; $\xi_C = (\xi_c)_{c \in C}$ – его зона интереса; $P_A = (p_A(y))_{y \in Y}$ и $P_{C \setminus A} = (p_{C \setminus A}(y))_{y \in Y}$ – распределения вероятностей на Y, описывающие свойства элемента ξ_A и его окрестности $\xi_{C \setminus A}$ соответственно. Если Y_A и $Y_{C \setminus A}$ – подмножества Y, определенные при помощи равенств

$$Y_A = \{y \in Y: S(A) p_A(y) \geq S(C \setminus A) p_{C \setminus A}(y)\}, \quad Y_{C \setminus A} = Y \setminus Y_A,$$

то решающее правило $h^*\colon Y \to \{0,1\}$ вида

$$h^*(y) = \begin{cases} 1, & y \in Y_A; \\ 0, & y \in Y_{C \setminus A} \end{cases}$$

классификации узлов из C будет байесовским с вероятностью $e(h^*)$ ошибки классификации, вычисляемой по формуле

$$e(h^*) = \frac{S(C \setminus A)}{S(C)} \sum_{y \in Y_A} p_{C \setminus A}(y) + \frac{S(A)}{S(C)} \sum_{y \in Y_{C \setminus A}} p_A(y).$$

Доказательство теоремы приведено в разд. 4.

В ходе конкретного дешифрирования для классификации пиксел зоны интереса вместо неизвестных вероятностей $p_A(y)$ и $p_{C \setminus A}(y)$, $y \in Y$, можно использовать их оценки $\tilde{p}_A(y)$ и $\tilde{p}_{C \setminus A}(y)$, вычисленные по ее изображению x_C. В самом деле, пусть $P_A = (p_A(y))_{y \in Y}$, $P_{C \setminus A} = (p_{C \setminus A}(y))_{y \in Y}$ и $P_C = (p_C(y))_{y \in Y}$ – неизвестные распределения вероятностей на Y, описывающие свойства элемента сцены, его окрестности и квадрата C соответственно. Очевидно, что они связаны равенством

$$p_C(y) = \frac{S(A)}{S(C)} p_A(y) + \frac{S(C \setminus A)}{S(C)} p_{C \setminus A}(y), \quad y \in Y.$$

Заменим в нем неизвестные вероятности $p_C(y)$ и $p_{C \setminus A}(y)$, $y \in Y$, их оценками $\tilde{p}_C(y)$ и $\tilde{p}_{C \setminus A}(y)$, вычисленными по изображению x_C всей зоны интереса и ее границы $x_{Fr(C)}$. Тогда оценки для $p_A(y)$, $y \in Y$, можно получить, решая систему из $|Y|$ уравнений и $|Y|$ неравенств:

$$\tilde{p}_C(y) = \frac{S(A)}{S(C)} \tilde{p}_A(y) + \frac{S(C \setminus A)}{S(C)} \tilde{p}_{C \setminus A}(y), \quad y \in Y;$$

$$\tilde{p}_A(y) \geq 0, \quad y \in Y.$$

Ясно, что применять описанный метод сегментации можно только при условии, что зона интереса уже построена. Поэтому сошлемся на работу [12], в которой предлагается один из возможных методов поиска зон интереса для объектов, названных Куком и Розенфельдом в [13] пятнами. Заметим, что вместо поиска зон интереса можно организовать сегментацию каждого квадратного участка сцены, но такой перебор потребует гораздо больше вычислительных ресурсов, чем поиск зон с последующей сегментацией только в обнаруженных зонах.

4. **Доказательства теорем.**

Доказательство теоремы 1. Пусть B – проекция некоторого элемента сцены, а $P_{Y^B} = (p_{Y^B}(x_B))_{x_B \in Y^B}$ – сопоставленное ему распределение. Очевидно, что для любого $A \subset B$ семейство $P_{Y^A} = (p_{Y^A}(x_A))_{x_A \in Y^A}$, определенное при помощи равенства

$$p_{Y^A}(x_A) = \sum_{x_{B\setminus A} \in Y^{B\setminus A}} p_{Y^B}(x_A, x_{B\setminus A}),$$

также является распределением на Y^A.

Воспользуемся этим результатом для задания на Z^2 семейства конечномерных распределений. Пусть D – конечное подмножество из Z^2. Тогда существует конечное семейство проекций $(B_j)_{j \in J_D}$, позволяющее представить D в виде суммы непустых непересекающихся слагаемых

$$D = \sum_{j \in J_D} D \cap B_j.$$

Для каждого $x_D \in Y^D$ определим $p_{Y^D}(x_D)$ при помощи равенства

$$p_{Y^D}(x_D) = p_{Y^D}(x_{D\cap B_j},\ j \in J_D) = \prod_{j \in J_D} p_{Y^{D\cap B_j}}(x_{D\cap B_j}).$$

Так как

$$\sum_{x_D \in Y^D} p_{Y^D}(x_D) = \sum_{x_D \in Y^D} \prod_{j \in J_D} p_{Y^{D\cap B_j}}(x_{D\cap B_j}) =$$

$$= \sum_{k \in J_D} \sum_{x_{D\cap B_k} \in Y^{D\cap B_k}} \prod_{j \in J_D} p_{Y^{D\cap B_j}}(x_{D\cap B_j}) = 1,$$

то семейство $P_{Y^D} = (p_{Y^D}(x_D))_{x_D \in Y^D}$ является распределением вероятностей на Y^D. Покажем, что для него выполняются условия согласованности.

Пусть C и D – конечные непересекающиеся подмножества из Z^2; $(B_j)_{j \in J_C}$ и $(B_j)_{j \in J_D}$ – семейства проекций такие, что

$$C = \sum_{j \in J_C} C \cap B_j, \qquad D = \sum_{j \in J_D} D \cap B_j.$$

Покажем, что для любого $x_D \in Y^D$ выполняется равенство

$$p_{Y^D}(x_D) = \sum_{x_C \in Y^C} p_{Y^{D+C}}(x_D, x_C).$$

Для этого определим $J = J_D \cup J_C$. Тогда

$$C = \sum_{j \in J} C \cap B_j, \quad D = \sum_{j \in J} D \cap B_j, \quad C + D = \sum_{j \in J} (C + D) \cap B_j.$$

Это позволяет выполнить следующие преобразования:

$$\sum_{x_C \in Y^C} p_{Y^{D+C}}(x_D, x_C) = \sum_{x_C \in Y^C} \prod_{j \in J} p_{Y^{(D+C)\cap B_j}}(x_{(D+C)\cap B_j}) =$$

$$= \sum_{k \in J} \sum_{x_{C\cap B_k} \in Y^{C\cap B_k}} \prod_{j \in J} p_{Y^{(D+C)\cap B_j}}(x_{D\cap B_j}, x_{C\cap B_j}) =$$

$$= \prod_{j \in J} p_{Y^{D\cap B_j}}(x_{D\cap B_j}) = p_{Y^D}(x_D),$$

подтверждающие выполнение условия согласованности.

Таким образом, если каждому подмножеству B из заданного разбиения решетки сопоставлено распределение вероятностей $P_{Y^B} = (p_{Y^B}(x_B))_{x_B \in Y^B}$, то существует случайное поле $(\xi_z)_{z \in Z^2}$ такое, что

$$P(\xi_B = x_B) = p_{Y^B}(x_B).$$

Независимость случайных величин ξ_a и ξ_b, принадлежащих разным проекциям A и B, следует непосредственно из определения конечномерных распределений при $D = \{a, b\}$.

Доказательство теоремы 2. Пусть A – конечное подмножество Z^2. Если f_A – отображение вида $Y^{Z^2} \to R$, зависящее от пиксел с координатами из A, то суперпозиция $f_A \circ \xi$, определяемая равенством $f_A \circ \xi(\omega) = f_A(\xi(\omega))$, является отображением вида $\Omega \to R$. Для получения нужного результата воспользуемся хорошо известным фактом из анализа (см., например, [11]). Он утверждает, что если $(\xi_j)_{1 \le j \le n}$ – конечное семейство случайных величин и $f: R^n \to R$ – борелевское отображение, то суперпозиция вида $f(\xi_1, \ldots, \xi_n)$ будет здесь случайной величиной.

Доказательство теоремы 3. Для доказательства используются свойства зоны интереса и бернуллиевских случайных полей. Из определения зоны интереса следует, что ее проекция C состоит из узлов, принадлежащих одному из двух классов: A или $C \setminus A$. Долю каждого из них легко подсчитать при помощи следующих равенств:

$$P(A) = \frac{S(A)}{S(C)}, \qquad P(C \setminus A) = \frac{S(C \setminus A)}{S(C)}.$$

Очевидно, что $P(A)$ и $P(C \setminus A)$ неотрицательны, а их сумма равна единице. Поэтому их можно интерпретировать как априорные вероятности соответствующих классов. В качестве распределений признака внутри классов естественно рассматривать распределения $P_A = (p_A(y))_{y \in Y}$ и $P_{C\setminus A} = (p_{C\setminus A}(y))_{y \in Y}$, описывающие свойства элемента сцены и его окрестности. Задание априорных вероятностей и распределений признаков для каждого класса позволяет легко (см., например, [10]) построить оптимальное байесовское правило h^* и вычислить вероятность $e(h^*)$ его ошибки.

Заключение. Предложена формализация задачи дешифрирования аэрокосмических изображений в рамках теоретико-вероятностного подхода. В случае, когда зависимостью между значениями признаков разных элементов сцены можно пренебречь, определена структура случайного поля $\xi = (\xi_z)_{z \in Z^2}$, являющегося описанием сцены.

В рамках выбранной формализации перечислены возможные цели дешифрирования, т. е. та информация, которая может быть получена о сцене по ее изображению.

Сформулированы достаточные условия, при выполнении которых отображение (признак), заданное на множестве всех изображений, будет случайной величиной.

Дано определение зоны интереса. Показана возможность ее применения для сегментации изображений на основе байесовской классификации.

СПИСОК ЛИТЕРАТУРЫ

1. **Денисов Д. А., Низовкин В. А.** Сегментация изображений на ЭВМ // Зарубеж. радиоэлектрон. 1985. № 10. С. 5.
2. **Бакут П. А., Колмогоров Г. С., Ворновицкий И. Э.** Сегментация изображений: методы пороговой обработки // Зарубеж. радиоэлектрон. 1987. № 10. С. 6.
3. **Бакут П. А., Колмогоров Г. С.** Сегментация изображений: методы выделения границ областей // Там же. С. 25.
4. **Яглом А. М.** Введение в теорию стационарных случайных функций // Успехи мат. наук. 1952. **VII**, № 5. С. 3.
5. **Добрушин Р. Л.** Описание случайного поля при помощи условных вероятностей и условия его регулярности // Теория вероятностей и ее применения. 1968. **XIII**, № 2. С. 201.
6. **Фофанов В. Б., Скворцов В. В.** О построении L-марковских случайных полей на ЭЦВМ // Материалы всесоюз. симп. по статистике случайных процессов. Киев, 1973. С. 193.
7. **Фофанов В. Б.** Об одном алгоритме статистического моделирования дискретизированных случайных полей // Проблемы разработки и гидродинамики нефтяных месторождений. Казань: Изд-во КГУ, 1975. С. 221.
8. **Аверинцев М. Б.** Об одном способе описания случайных полей с дискретным аргументом // Проблемы передачи информации. 1970. **VI**, № 2. С. 100.
9. **Фофанов В. Б.** Описание случайных полей условными вероятностями при постановке и решении некоторых задач стохастической оптимизации: Автореф. дис. ... канд. техн. наук. Казань, 1977.
10. **Дуда Р., Харт П.** Распознавание образов и анализ сцен. М.: Мир, 1976.
11. **Лоэв М.** Теория вероятностей. М.: ИЛ, 1962.
12. **Краснова Ф. С., Фофанов В. Б.** Об автоматическом дешифрировании аэрокосмических изображений (Ч. II) // Автометрия. 1994. № 1. С. 56.
13. **Cook G. M., Rosenfeld A.** Size detectors // Proc. IEEE. Lett. 1970. **58**, N 12. P. 1956.

Казанский государственный университет,
E-mail: Viatcheslav.Fofanov@ksu.ru

Поступила в редакцию
10 июня 2002 г.