УДК 519.681

ФОРМАЛИЗАЦИЯ СЦЕНЫ В ЗАДАЧЕ ДЕШИФРИРОВАНИЯ МНОГОЗОНАЛЬНЫХ ИЗОБРАЖЕНИЙ

© 2007 г. **В. Б. Фофанов**, канд. техн. наук

ЗАО "Научно-производственная фирма Оптоойл", г. Казань

E-mail: optooil@tbit.ru

Предлагается математическая модель сцены для построения методов совместного дешифрирования нескольких ее изображений. Цель дешифрирования состоит в выявлении на сцене объектов по их геометрическим признакам. Основными элементами модели являются пиксел, объект, его зона интереса и изображение. Для сравнения пикселов, описываемых векторными случайными величинами, применяется случайное расстояние.

Коды OCIS: 100.0100.

Поступила в редакцию 07.09.2006.

Введение

Визуальное дешифрирование рассматривается сегодня некоторыми специалистами в качестве основного или даже единственно возможного метода дешифрирования. Однако против такой точки зрения можно привести ряд серьезных возражений. Во-первых, визуальное дешифрирование является принципиально субъективным. Во-вторых, участие человека в дешифрировании может оказаться невозможным, например, из-за отсутствия рядом с оптико-электронной системой средств жизнеобеспечения и канала связи. В-третьих, при скорости съемки, превышающей психофизические возможности восприятия зрительного анализатора, все преимущества визуального дешифрирования становятся бесполезными. Наконец, необходимо отметить, что эксперименты по визуальному дешифрированию нескольких одновременно формируемых изображений, изложенные в работе [1], позволяют утверждать, что зрительный анализатор человека плохо приспособлен к интегрированию подобной видеоинформации. Поэтому, несмотря на некоторое увеличение вероятностей правильного распознавания, отмеченное в ряде экспериментов с двумя изображениями, принципиальным решением задачи дешифрирования является его автоматизация, вначале частичная, а затем и полная.

К сожалению, общепринятой математической теории дешифрирования изображений, которая могла бы служить основой его автоматизации, пока не существует. В настоящей работе описывается математическая модель сцены, используемая в задаче совместного дешифрирования нескольких ее изображений. Они формируются оптико-электронной системой одновременно в нескольких спектральных диапазонах. Предполагается, что на каждом изображении имеется своя система координат и что изображения каждого пиксела сцены в этих системах имеют одни и те же координаты. Целью дешифрирования является выявление на сцене заданных объектов. Поскольку в общем случае каждый пиксел сцены описывается векторной случайной величиной, то для описания свойств объектов и построения методов применяется так называемое случайное расстояние, введенное в работе [2].

Необходимо отметить, что исследования в этом направлении были начаты в НПО "Государственный институт прикладной оптики" [3, 4] и продолжены в ЗАО "Научно-производственная фирма Оптоойл".

Скалярные и векторные сцены

Вначале будем рассматривать реальную сцену как совокупность неделимых элементов, называемых далее пикселами. Каждый пиксел характеризуется 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 скалярной случайной величиной ξ_z со значениями из конечного множества $Y = \{0, 1, \ldots, |Y| - 1\}$, $|Y| > 1$. Она описывает заданное свойство пиксела. Все случайные величины определены на некотором вероятностном пространстве (Ω, A, P). Счетное семейство $\xi = (\xi_z)_{z \in Z^2}$ будет называться далее скалярной сценой.

Для каждого $z \in Z^2$ и каждого элементарного события $\omega \in \Omega$ можно вычислить выборочное значение $x_z = \xi_z(\omega)$ случайной величины ξ_z. Скалярным изображением сцены будет называться отображение x: $Z^2 \to Y$, определяемое равенством $x = (x_z)_{z \in Z^2}$.

Как правило, интерес представляют конечные совокупности пикселов. Они будут называться объектами. Каждый объект определяется подмножеством A точек из Z^2, являющихся координатами его пикселов, и семейством $\xi_A = (\xi_a)_{a \in A}$ из $|A|$ случайных величин. Подмножество A будет называться далее проекцией объекта. Предполагается, что проекции A и B разных объектов не пересекаются.

Если в пикселе с координатами $z \in Z^2$ измеряется одновременно $\nu > 1$ признаков, то для его описания естественно использовать векторную случайную величину $\xi_z = (\xi_z^j)_{1 \le j \le \nu}$. Ее компоненты ξ_z^j, $1 \le j \le \nu$, являются скалярными случайными величинами со значениями из Y, определенными на том же самом вероятностном пространстве (Ω, A, P). Счетное семейство $(\xi_z)_{z \in Z^2}$ векторных случайных величин будет называться далее векторной сценой.

Пусть $\omega \in \Omega$, для пиксела с координатами $z \in Z^2$ и каждого j, $1 \le j \le \nu$, можно вычислить выборочное значение $x_z^j = \xi_z^j(\omega)$ скалярной случайной величины ξ_z^j. Скалярным изображением с номером j (j-й компонентой) векторной сцены будет называться отображение $x^j{:}Z^2 \to Y$, $1 \le j \le \nu$, вида $x^j = (x_z^j)_{z \in Z^2}$. Таким образом, при $\nu > 1$ результатом съемки является набор $x = (x^j)_{1 \le j \le \nu}$ из ν скалярных изображений. Если для каждого $z \in Z^2$ определить вектор $x_z = (x_z^j)_{1 \le j \le \nu}$, то x становится векторным отображением $Z^2 \to Y^\nu$ вида $x = (x_z)_{z \in Z^2}$. Оно будет называться далее векторным изображением сцены. Основное внимание в настоящей работе уделяется получению результатов, которые имеют место одновременно для скалярных ($\nu = 1$) и векторных ($\nu > 1$) сцен.

Очевидно, что объектом с проекцией A векторной сцены является семейство $\xi_A = (\xi_a)_{a \in A}$ из $|A|$ векторных случайных величин.

В настоящей работе будем считать, что случайные величины, образующие сцену, попарно независимы, а случайные величины, образующие объект, $\xi_A = (\xi_a)_{a \in A}$ имеют, кроме того, одно и то же распределение вероятностей. Для скалярных сцен его среднее значение m_A и дисперсия σ_A^2 являются скалярами. Для векторной сцены будут использоваться вектор средних значений $m_A = (m_A^j)_{1 \le j \le \nu}$ и ковариационная матрица $C_A = (c_{ij}^A)$ с элементами $c_{ij}^A = E(\xi_a^i - m_A^i)(\xi_a^j - m_A^j)$, $a \in A$.

Поскольку регистрируемые признаки (энергетические яркости) очень изменчивы, то никаких сведений о них заранее не предполагается. Все методы, используемые в ходе дешифрирования, строятся на основе предположения о том, что соседние объекты с проекциями A и B должны иметь разные векторы средних значений: $m_A \ne m_B$ или, что то же самое, $d(m_A, m_B) > 0$. С содержательной точки зрения это означает, что среди ν спектральных диапазонов оптико-электронной системы всегда найдется диапазон (хотя бы один) с номером j, $1 \le j \le \nu$, в котором $m_A^j \ne m_B^j$.

Предположение о независимости случайных величин позволяет рассматривать изображение сцены в качестве случайной выборки и использовать его для оценки их средних значений и ковариаций. На первый взгляд оно может показаться слишком оптимистичным. Однако хорошо известно [5], что оценивать средние значения можно и по выборочным поверхностям (реализациям) однородных, в широком смысле случайных, полей, состоящих из зависимых случайных величин.

Объекты и их признаки

Напомним, что целью дешифрирования изображений сцены в настоящей работе является выявление на ней заданных объектов, описываемых геометрическими признаками. Прежде всего, это площадь, габаритные размеры, форма и некоторые другие свойства, определяемые по проекции объекта.

Площадью объекта с проекцией A будет называться количество $|A|$ образующих ее пикселов. Для вычисления размеров будет применяться, как правило, евклидово расстояние d на Z^2. Габаритными размерами объекта будут называться длины сторон прямоугольника наименьшей площади, описанного вокруг его проекции.

Диаметром объекта с проекцией A будет называться диаметр $d(A)$ его проекции, определяемый обычным образом, $d(A) = \max\limits_{a \in A, b \in A} d(a, b)$. Очевидно, что для любого $t \in Z^2$ $\inf\limits_{z \in Z^2 \setminus \{t\}} d(t, z) = 1$. Поэтому пикселы с координатами z и t, для которых $d(z, t) = 1$, будут называться соседями. У каждого пиксела их ровно четыре.

Подмножество A на Z^2 будет называться связным, если для любых двух пикселов с координатами z и t из A, $z \ne t$, существует последовательность из $n \ge 2$ пикселей $(w_j)_{1 \le j \le n}$ принадлежащих A и таких, что $z = w_1$, $t = w_n$ и $d(w_j, w_{j+1}) = 1$, $1 \le j \le n$. Проекции всех рассматриваемых в работе объектов предполагаются связными.

Объекты ξ_A и ξ_B будут называться соседями, если расстояние $d(A, B)$ между их проекциями, определяемое обычным образом $d(A, B) = \inf\limits_{a \in A, b \in B} d(a, b)$, равно единице.

Значения геометрических признаков заданных объектов или их распределения вероятностей должны быть известны до начала дешифрирования.

Случайные расстояния

При переходе к описанию пикселов сцены ν-мерными случайными величинами тотчас возникает ряд новых проблем. Во-первых, в ходе съемки каждому пикселу сопоставляется не скаляр, а ν-мерный вектор признаков. К сожалению, в ν-мерном векторном пространстве R^ν отсутствует отношение порядка, сопоставимое по наглядности с таковым из R^1. Во-вторых, объем выборки, необходимой для оценки неизвестных вероятностей, описывающих изменчивость признаков, зависит экспоненциально от ν. Для решения этих проблем можно воспользоваться введенным в работе [2] случайным расстоянием. Для удобства читателя приведем определение и используемые далее свойства случайного расстояния.

Расстояние d на R^ν будет называться борелевским, если борелевским является отображение $d: R^\nu \times R^\nu \to R$. Из свойств измеримых функций следует, что расстояние Евклида и многие другие известные расстояния являются Борелевскими.

Пусть d – борелевское расстояние на R^ν, а $\xi = (\xi_j)_{1 \le j \le \nu}$ и $\eta = (\eta_j)_{1 \le j \le \nu}$ – векторные случайные величины, определенные на некотором вероятностном пространстве (Ω, A, P). Известно, что в этом случае отображение $d(\xi, \eta): \Omega \to R$, определяемое равенством $d(\xi, \eta)(\omega) = d(\xi(\omega), \eta(\omega))$, также будет случайной величиной на (Ω, A, P). Из определения расстояния следует, что при каждом $\omega \in \Omega$ будут выполняться аксиомы расстояния. Это позволяет назвать отображение d случайным расстоянием на множестве ν-мерных случайных величин, а случайную величину $d(\xi, \eta)$ – случайным расстоянием между ξ и η.

Случайное расстояние позволяет легко обобщить на векторные случайные величины понятие дисперсии. Пусть $\xi = (\xi_j)_{1 \le j \le \nu}$ – векторная случайная величина, $m_\xi^j = E\xi_j$, $1 \le j \le \nu$, и $m_\xi = (m_\xi^j)_{1 \le j \le \nu}$. Тогда дисперсией ξ естественно назвать неотрицательное число $D\xi$, определяемое равенством $D\xi = Ed^2(\xi, m_\xi)$. Это определение позволяет установить целый ряд полезных свойств векторных случайных величин, включая знаменитое неравенство Чебышева.

1. Если $\xi = (\xi_j)_{1 \le j \le \nu}$ – векторная случайная величина с конечной дисперсией $\sigma_\xi^2 = D\xi$, то при любом вещественном $k > 0$ выполняется неравенство $P(d(\xi, m_\xi) \ge k\sigma_\xi) \le 1/k^2$.

Следующее утверждение позволяет оценить вероятность отклонения случайного расстояния $d(\xi, \eta)$ между случайными величинами от расстояния $d(m_\xi, m_\eta)$ между их средними значениями.

2. Пусть ξ и η – векторные случайные величины, m_ξ и m_η – их векторы средних значений, а σ_ξ^2 и σ_η^2 – их конечные дисперсии. Тогда для любого вещественного $k > 0$ выполняется неравенство $P(|d(\xi, \eta) - d(m_\xi, m_\eta)| \ge k(\sigma_\xi + \sigma_\eta)) \le 1/k^2$.

Далее в качестве d будет использоваться евклидово расстояние. Из его определения тотчас следует, что $D\xi = D\xi_1 + \ldots + D\xi_\nu$.

3. Если ξ и η – попарно независимые векторные случайные величины, $a \in R$ и $b \in R$, то $D(a\xi + b\eta) = a^2 D\xi + b^2 D\eta$.

Пусть $\sigma_\xi^2 = D\xi$. Легко проверить непосредственно, что для любого вектора $m = (m_j)_{1 \le j \le \nu}$ из R^ν имеет место равенство $Ed^2(\xi, m) = \sigma_\xi^2 + d^2(m_\xi, m)$. Из него тотчас следует неравенство $Ed^2(\xi, m) \ge \sigma_\xi^2$, обобщающее на векторные случайные величины важное свойство дисперсии скалярных случайных величин.

Если ξ и η независимы, то легко убедиться в справедливости равенства $Ed^2(\xi, \eta) = \sigma_\xi^2 + d^2(m_\xi, m_\eta) + \sigma_\eta^2$.

Будем говорить, что последовательность $\xi_n = (\xi_{n,j})_{1 \le j \le \nu}$, $n \ge 1$, векторных случайных величин сходится по вероятности к векторной случайной величине $\xi = (\xi_j)_{1 \le j \le \nu}$, и обозначать это символом $\xi_n \to \xi$, если последовательность $(d(\xi_n, \xi))_{n \ge 1}$ скалярных случайных величин сходится по вероятности к нулю. Другими словами, если $P(d(\xi_n, \xi)) > \varepsilon) \to 0$ при $n \to +\infty$ для любого $\varepsilon > 0$, имеет место следующее свойство.

4. Пусть $(\xi_n)_{n \ge 1}$ – последовательность независимых векторных случайных величин таких, что $E\xi_n = m$ и $D\xi_n = \sigma^2$ для любого $n \ge 1$. Тогда последовательность $(\eta_n)_{n \ge 1}$ вида $\eta_n = (\xi_1 + \ldots + \xi_\nu)/n$ сходится по вероятности к m при $n \to +\infty$.

Зоны интереса

Для распознавания объектов по геометрическим признакам необходимо знать их проекции. Разбиение Z^2 на проекции образующих сцену объектов называется, как известно, сегментацией. В общем случае для решения этой задачи необходимо располагать довольно подробными сведениями о сцене, которые на этапе дешифрирования, как правило, отсутствуют. В частном случае, когда дешифрирование проводится с целью выявления заданных объектов, вместо сегментации всей сцены можно ограничиться сегментацией только таких ее участков, которые содержат заданный объект и окружающие его пикселы (окрестность). Такие участки были названы зонами интереса. Для скалярных сцен, когда $\nu = 1$, их определение дано в работе [6]. Общий случай рассматривается ниже.

Пусть $\xi_A = (\xi_a)_{a \in A}$ – заданный объект с вектором средних значений $m_A = (m_A^j)_{1 \le j \le \nu}$ и диаметром $d(A)$. Очевидно, что для любого $l \ge d(A) + 2$ существует квадрат C на Z^2 со стороной l такой, что $A \subset (C \backslash \mathrm{Fr}(C))$. Символ $\mathrm{Fr}(C)$ обозначает границу квадрата C, т. е. подмножество его точек, имеющих соседей в $Z^2 \backslash C$. Будем называть семейство ν-мерных случайных величин вида $\xi_C = (\xi_c)_{c \in C}$ зоной интереса для элемента сцены $\xi_A = (\xi_a)_{a \in A}$, если все ξ_c, $c \in C \backslash A$, имеют одно и то же распределение вероятностей с вектором $m_{C \backslash A} = (m_{C \backslash A}^j)_{1 \le j \le \nu}$ средних значений и если $d(m_A, m_{C \backslash A}^j) > 0$.

Из определения зоны интереса следует, что каждый ее пиксел может принадлежать только к объекту

ξ_A или его окрестности $\xi_{C\setminus A}$ (фону). В частном случае, когда $\nu = 1$, неравенство $d(m_A, m_{C\setminus A}) > 0$ принимает вид $|m_A, m_{C\setminus A}| > 0$, а при известном знаке контраста – $m_A > m_{C\setminus A}$ (объект светлее фона) или $m_A < m_{C\setminus A}$ (объект темнее фона).

Так как свойства зон интереса отличаются от свойств пустых квадратов таких же размеров, то зоны интереса можно рассматривать в качестве вспомогательных элементов сцены. В связи с этим целесообразно начать дешифрирование с поиска зон интереса, а сегментацию и все остальные операции выполнять только в них. При успешном решении задачи поиска зон интереса такой подход может заметно снизить потребность в вычислительных ресурсах.

Отметим, что поиск зон интереса скалярных сцен ($\nu = 1$) подробно рассмотрен в работе [6]. Показано, что дополнительная информация о знаке контраста позволяет значительно снизить вероятность ложного обнаружения зоны.

Заключение

В работе предложена модель сцены, предназначенная для автоматизации совместного дешифрирования ее изображений, цель которого состоит в выявлении заданных объектов. Все пикселы сцены описываются векторными случайными величинами. Каждый объект является семейством независимых и одинаково распределенных векторных случайных величин, определенных на конечном подмножестве целочисленной решетки Z^2. Оно называется проекцией объекта. Объекты, подлежащие выявлению, описываются геометрическими признаками или их распределениями вероятностей. Никаких данных о расположении объектов, их ориентации или спектральных признаках не предполагается. Исходной информацией о сцене служит набор ее изображений.

Рассмотренная модель сцены будет использоваться далее для решения задач поиска зон интереса, сегментации зон интереса и вычисления геометрических признаков объектов.

ЛИТЕРАТУРА

1. *Смирнов В.П., Алексеев Ю.В., Аллик Ю.А. и др.* Эффективность комплексирования разноканальных изображений при опознавании объектов // Оптический журнал. 1992. Т. 59. № 2. С. 20–24.

2. *Фофанов В.Б., Чемоданов К.Ю.* О выборе признаков для классификации пикселей цветных изображений // Актуальные проблемы математического моделирования и информатики. Материалы научной конференции. Казань: Изд-во Казанского математического общества, 2002. С. 102–105.

3. *Краснова Ф.С., Фофанов В.Б.* Об автоматическом дешифрировании аэрокосмических изображений. Ч. I // Автометрия. 1993. № 6. С. 79–83.

4. *Краснова Ф.С., Фофанов В.Б.* Об автоматическом дешифрировании аэрокосмических изображений. Ч. II // Автометрия. 1994. № 1. С. 56–59.

5. *Крамер Г., Лидбеттер М.* Стационарные случайные процессы. Свойства выборочных функций и их приложения / Пер. с англ. М.: Мир, 1969. 400 с.

6. *Bourylin S.A., Fofanov V.B.* Generalized Spot Criterion as Applied for Image Deciphering // Pattern Recognition and Image Analysis. 2004. V. 14. № 2. P. 243–248.