РОССИЙСКАЯ АКАДЕМИЯ НАУК
СИБИРСКОЕ ОТДЕЛЕНИЕ
АВТОМЕТРИЯ
№ 6 1993

УДК 519.681

Ф. С. Краснова, В. Б. Фофанов
(Казань)

ОБ АВТОМАТИЧЕСКОМ ДЕШИФРИРОВАНИИ АЭРОКОСМИЧЕСКИХ ИЗОБРАЖЕНИЙ. Ч. I

Работа является обобщением как оригинальных, так и известных результатов, относящихся к автоматическому дешифрированию аэрокосмических изображений. В частности, предлагается определение такого понятия, как информативность набора признаков, исследуется влияние на результат дешифрирования числа используемых признаков и различных способов разбиения множества объектов на классы.

Введение. Под дешифрированием в настоящей работе будет подразумеваться выявление на местности заданных подвижных объектов по ее аэрокосмическим изображениям. Типичным примером подвижных объектов являются транспортные средства, находящиеся на дорогах или естественных фонах, летательные аппараты на стоянках или после совершения вынужденной посадки, морские и речные суда.

Своевременное выявление изменений, происходящих на местности, отсутствие систем жизнеобеспечения на носителях с датчиками видеоинформации, а также необходимость снижения степени субъективности результатов визуального дешифрирования — вот далеко не полный перечень причин, инициировавших около тридцати лет назад исследования по автоматизации дешифрирования аэрокосмических изображений. Получить представление о современном состоянии этого научного направления можно, ознакомившись с обзором [1].

Настоящая работа является обобщением как оригинальных, так и известных результатов, опубликованных в различных изданиях и связанных с решением некоторых задач, возникающих при автоматизации дешифрирования изображений. В частности, предлагается определение такого понятия, как информативность набора признаков, исследуется влияние на результат дешифрирования числа используемых признаков и различных способов разбиения множества объектов на классы. Впервые в явном виде рассмотрена схема многоэтапного дешифрирования. Приводятся результаты экспериментов по автоматическому дешифрированию реальных изображений.

Вероятностный подход. Будем считать земную поверхность плоскостью и разобьем ее на непересекающиеся квадратные участки одинаковой площади — элементы разложения. Зададим на плоскости прямоугольную систему координат, поместив ее начало в центр одного из элементов разложения и направив оси параллельно его сторонам. За единицу масштаба примем длину стороны элемента разложения. В таким образом выбранной системе координат центр z каждого элемента разложения будет иметь целочисленные координаты (z_1, z_2). Множество Z^2 таких точек $z = (z_1, z_2)$ называется двумерной целочисленной решеткой.

Аэрокосмическим изображением назовем неотрицательную функцию x, определенную на Z^2, со значениями в конечном множестве $X \subset R : x =$

$= (x_z)_{z \in Z^2}$. Для каждого $z \in Z^2$ значение x_z получается путем регистрации осредненного по площади соответствующего элемента разложения или его части электромагнитного излучения датчиком, расположенным в атмосфере или космосе, в фиксированной спектральной зоне $\lambda = [\lambda_1, \lambda_2]$ оптического диапазона.

Экспериментальное изучение поведения энергетических яркостей выявило отсутствие устойчивой функциональной связи между их значениями даже в соседних элементах разложения. Вместе с тем наблюдается устойчивость частот наблюдения того или иного значения яркости для фиксированных условий. Поэтому земную поверхность до ее съемки естественно описывать не при помощи какого-то одного изображения x, а статистическим ансамблем изображений, каждое из которых повторяется в нем с определенной частотой. Обобщая изложенное, назовем земной поверхностью семейство неотрицательных случайных величин $(\xi_z)_{z \in Z^2}$ со значениями в конечном множестве X, или, другими словами, случайное поле на Z^2. Изображение $x = (x_z)_{z \in Z^2}$ будет рассматриваться в качестве одной из выборочных поверхностей (реализаций) этого случайного поля.

Пусть d — эвклидово расстояние на Z^2. Два узла $a \in Z^2$ и $b \in Z^2$ назовем соседними, если $d(a, b) = 1$. Подмножество $A \subset Z^2$ назовем связным, если любую пару принадлежащих ему точек $a \in A$ и $b \in A$ можно соединить конечной последовательностью $(z_j)_{1 \le j \le n}$ из A, в которой любые следующие друг за другом точки являются соседними и $a = z_1$, $b = z_n$. Пусть A — конечное односвязное подмножество из Z^2, r — целое неотрицательное число. Совокупность точек $V_r(A)$ из $Z^2 \setminus A$, отстоящих от A на расстояние не более чем r, назовем r-окрестностью A.

Пусть A — конечное односвязное подмножество из Z^2, состоящее из элементов разложения, принадлежащих объекту. Назовем его проекцией объекта. Самим объектом будем называть семейство случайных величин $(\xi_a)_{a \in A}$, являющееся подмножеством $(\xi_z)_{z \in Z^2}$. Случайные величины ξ_a, образующие объект, имеют одно и то же распределение вероятностей p_0 со средним значением $E(\xi_a) = m_0$ и дисперсией $D(\xi_a) = \sigma_0^2$. Будем предполагать, что проекция объекта A имеет r-окрестность $V_r(A)$, случайные величины которой имеют одно и то же распределение вероятностей $p_ф$ со средним $m_ф$ и дисперсией $\sigma_ф^2$. Для определенности будем считать, что $m_ф < m_0$.

Предложенное определение соответствует одному из простых, но распространенных (типовых) случаев, допускающему экспериментальную проверку. Требуется, чтобы была известна масштабная единица, позволяющая определить все элементы разложения, составляющие объект, и объект был единой конструкцией. Предполагается, что отражательные и излучательные свойства объекта статистически мало изменяются при переходе от одного его элемента разложения к другому и яркость большинства его элементов разложения выше самого яркого элемента разложения местности из его окружения. Кроме того, объектам запрещается находиться рядом друг с другом. На изображениях объекты имеют разные проекции и значения энергетических яркостей.

В рамках описанного вероятностного подхода дешифрирование сводится к выяснению свойств случайного поля $(\xi_z)_{z \in Z^2}$ по его известной выборочной поверхности x, или, более конкретно, к вычислению координат проекции объекта и отнесению его к одному из заранее названных классов.

Признаки и дешифрирование. В настоящем разделе широко используются терминология и определения теории распознавания образов (см., например, [2]). Рассмотрим каждый объект как совокупность только ему одному присущих свойств (признаков), позволяющих считать его как нечто единое и отличать от других объектов. При дешифрировании объектами являются участки земной поверхности, которые могут оказаться как подвижными транспортными средствами, так и просто участками местности.

С формальной точки зрения такой подход означает, что каждому элементу ω из множества Ω всех объектов поставлен в соответствие с помощью отображения $f: \Omega \to R^{\nu}$ ν-мерный вектор $\pi = (\pi_1, ..., \pi_\nu)$ с вещественными координатами. Для каждого объекта ω вектор $\pi = f(\omega)$ называется его образом, координаты π_j, $1 \le j \le \nu$, — признаками, а R^{ν} — признаковым пространством. Подробнее об ограничениях, которым должно удовлетворять отображение, см. [3].

Пусть J — конечное множество из R, элементы которого называются индексами, а $(\Omega_j)_{j \in J}$ — совокупность непустых непересекающихся подмножеств из Ω с объединением, равным Ω. Они будут называться классами.

В процессе дешифрирования изображения $x = (x_s)_{s \in Z^2}$ для каждого его фрагмента $(x_a)_{a \in A}$ вычисляется ν признаков, каждый из которых является функцией x, и принимается решение об отнесении объекта ω с этими признаками к одному из классов Ω_j, $j \in J$. По существу определение номера класса эквивалентно вычислению значения некоторого отображения $h: R^{\nu} \to J$, называемого решающим правилом (см. [3]).

Пусть H — множество всех решающих правил. Для сравнения их между собой определим на H неотрицательную функцию e, называемую целевой или критерием эффективности. Для этого построим на Ω вероятностное пространство (Ω, A, P) таким образом, чтобы каждый класс Ω_j, $j \in J$, был в нем событием (см. [3]). Потребуем, чтобы отображения f и h были измеримыми функциями. В этом случае для любого $j \in J$ подмножество из Ω вида

$$\{\omega \in \Omega : \omega \in \Omega_j,\ h(f(\omega)) \ne j\},$$

состоящее из неправильно проклассифицированных объектов из класса Ω_j, и множество всех неправильно проклассифицированных объектов из Ω

$$\bigcup_{j \in J} \{\omega \in \Omega : \omega \in \Omega_j,\ h(f(\omega)) \ne j\}$$

будут событиями. Поэтому, если выбрать в качестве целевой функции e вероятность неправильной классификации, то ее значение для решающего правила h можно вычислить по формуле

$$e(h) = \sum_{j \in J} P(\Omega_j) P(h(f(\omega)) \ne j \mid \Omega_j). \tag{1}$$

Решающее правило h^*, удовлетворяющее неравенству $e(h^*) \le e(h)$ для любого $h \in H$, называется оптимальным в смысле выбранного критерия e.

По определению (1) целевая функция e ограничена: $0 \le |e| \le 1$. Поэтому существует $\inf\limits_H e(h)$. Он не зависит от конкретного h, а характеризует возможности используемого набора признаков, который определяется отображением f, при оптимальном решающем правиле h^*. Если назвать число $I(f) = 1 - \inf\limits_H e(h)$ информативностью набора признаков, задаваемых отображением f, то появляется возможность сравнивать между собой различные наборы признаков.

Сделанные определения позволяют ответить на ряд вопросов, возникающих при дешифрировании. Начнем с выяснения влияния размерности ν признакового пространства на качество дешифрирования. Суть проблемы состоит в том, что после добавления к уже имеющимся признакам новых качество дешифрирования может ухудшиться по сравнению с первоначальным.

Теорема 1. Пусть ν_1 и ν_2 — целые неотрицательные числа и $\nu_1 \le \nu_2$, а отображения $f_1 : \Omega \to R^{\nu_1}$ и $f_2 : \Omega \to R^{\nu_2}$ таковы, что из $f_1(\omega) = (\pi_1, ..., \pi_{\nu_1})$ следует $f_2(\omega) = (\pi_1, ..., \pi_{\nu_1}, ..., \pi_{\nu_2})$. Пусть H_1 и H_2 — множества решающих правил, определенных на R^{ν_1} и R^{ν_2} соответственно, e_1 и e_2 — целевые функции, определенные на H_1 и H_2 соответственно, а

$$\inf_{H_1} e_1(h_1) \quad \text{и} \quad \inf_{H_2} e_2(h_2)$$

— качество дешифрирования при ν_1 и ν_2 признаках соответственно. Тогда

$$\inf_{H_2} e_2(h_2) \leq \inf_{H_1} e_1(h_1). \tag{2}$$

Доказательство теоремы см. в [3].

Последнее неравенство означает, что при использовании оптимальных решающих правил увеличение размерности признакового пространства не ухудшает качество дешифрирования. Если применяемые правила не являются оптимальными, то добавление дополнительных признаков может вызвать даже ухудшение качества.

Разбить на классы множество обьектов Ω можно многими способами. Пусть $(\Omega_j)_{j \in J}$ — одно из них, которое будем называть первоначальным. Обьединяя некоторые из классов первоначального разбиения в более крупные, можно получить новое разбиение, которое естественно назвать более грубым по сравнению с первоначальным. Наша следующая задача состоит в ответе на вопрос о том, как меняется ошибка классификации при переходе к более грубому разбиению.

Пусть K — конечное множество, $2 \leq |K| < |J|$ и $(J_k)_{k \in K}$ — разбиение множества J на $|K|$ подмножеств. Для каждого $k \in K$ построим подмножество $\Omega_{J_k} \subset \Omega$ следующим образом:

$$\Omega_{J_k} = \bigcup_{j \in J_k} \Omega_j. \tag{3}$$

Очевидно, что семейство $(\Omega_{J_k})_{k \in K}$ является разбиением Ω на классы, более грубым по сравнению с первоначальным.

Пусть h_J — решающее правило для классификации обьектов из Ω на $|J|$ классов. Для каждого $k \in K$ определим подмножество $X_{J_k} \subset R^\nu$ при помощи равенства

$$X_{J_k} = \bigcup_{j \in J_k} h_J^{-1}(j).$$

Поскольку $(X_{J_k})_{k \in K}$ является разбиением R^ν, то отображение $h_K : R^\nu \to K$, определяемое соотношением

$$h_K(\pi) = \sum_{k \in K} k E_{X_{J_k}}(\pi), \tag{4}$$

где $E_{X_{J_k}}$, $k \in K$, — характеристическая функция множества X_{J_k}, будет решающим правилом для классификации обьектов из Ω на $|K|$ больших классов Ω_{J_k}, $k \in K$. Будем считать, что правило h_J порождает более грубое правило h_K.

Теорема 2. Если правило h_J порождает более грубое правило h_K, а e_J и e_K — соответствующие им критерии эффективности, то

$$e_K(h_K) \leq e_J(h_J). \tag{5}$$

Последнее неравенство означает, что при переходе от первоначального разбиения $(\Omega_j)_{j \in J}$ к более грубому $(\Omega_{J_k})_{k \in K}$ вероятность ошибки классификации не увеличивается. Доказательство см. в [3].

Наконец, рассмотрим классификацию обьектов из Ω на $|J|$ классов, проводимую в два этапа. На первом этапе с помощью грубого решающего правила

h_K выбирается один из $|K|$ больших классов Ω_{J_k}, на втором — один из $|J_k|$ исходных классов Ω_j, $j \in J_k$, образующих Ω_{J_k}.

Пусть $h_K : R^\nu \to K$ — решающее правило для классификации обьектов из Ω на $|K|$ больших классов Ω_{J_k}, $k \in K$; $h_{J_k} : R^\nu \to J_k$, $k \in K$, — решающее правило для классификации обьектов из Ω_{J_k} на $|J_k|$ исходных классов Ω_j, $j \in J_k$. Легко проверить, что отображение $h_J : R^\nu \to J$, определяемое равенством

$$h_J(\pi) = \sum_{k \in K} h_{J_K}(\pi) E_{h_K^{-1}(k)}(\pi), \tag{6}$$

будет решающим правилом для классификации обьектов из Ω на $|J|$ исходных классов.

С другой стороны, если задано решающее правило h_J для классификации обьектов из Ω на $|J|$ классов, то оно порождает более грубое правило h_K согласно (4). Кроме того, для каждого $k \in K$ можно определить отображение $h_{J_k} : R^\nu \to J_k$ при помощи равенства

$$h_{J_k}(\pi) = h_J(\pi) E_{X_{J_k}}(\pi), \tag{7}$$

которое будет решающим правилом для классификации обьектов из Ω_{J_k} на $|J_k|$ классов Ω_j, $j \in J_k$. Изложенное позволяет рассматривать равенство (6) как разложение решающего правила h_J по совокупности правил h_{J_k}, $k \in K$, а равенства (4) и (7) определяют способ их вычисления для h_J.

Если определить для каждого $k \in K$ на Ω_{J_k} соответствующим образом (см. [3]) вероятностное пространство $(\Omega_{J_k}, A_{J_k}, P_{J_k})$, то для каждого правила h_{J_k} можно подсчитать его эффективность $e_{J_k}(h_{J_k})$ по формуле (1):

$$e_{J_k}(h_{J_k}) = \sum_{j \in J_k} P_{J_k}(\Omega_j) P_{J_k}(h_{J_k}(f(\omega)) \neq j \mid \Omega_j). \tag{8}$$

Более того, имеет место следующее утверждение, связывающее эффективности e_J и e_{J_k}, $k \in K$.

Теорема 3. Пусть для любого $k \in K$ e_{J_k} — эффективность решающего правила $h_{J_k} : R^\nu \to J_k$, а e_J — эффективность решающего правила $h_J : R^\nu \to J$, определяемого равенством (6), тогда

$$e_J(h_J) = \sum_{k \in K} P(\Omega_{J_k}) e_{J_k}(h_{J_k}). \tag{9}$$

Кроме того, правило h_J будет оптимальным тогда и только тогда, когда оптимальными являются правила h_{J_k}, $k \in K$.

Последнее равенство можно рассматривать как разложение $e_J(h_J)$ по совокупности $e_{J_k}(h_{J_k})$, $k \in K$. В этом смысле оно дополняет (6).

СПИСОК ЛИТЕРАТУРЫ

1. Бурлак Л. Ф. // Радиоэлектроника за рубежом. Обзоры.—1992.—Вып. 4(60).
2. Дуда Р., Харт П. Распознавание образов и анализ сцен: Пер. с англ.—М.: Мир, 1976.
3. Краснова Ф. С., Фофанов В. Б. Об автоматическом дешифрировании аэрокосмических изображений. Ч. II // Автометрия.—1994.—№ 1.

Поступила в редакцию 11 марта 1993 г.