

Физическая аппаратура

УДК 543.42

О подходе к автоматическому дешифрированию многоспектральных изображений

А. Б. Бельский, В. Б. Фофанов

ОАО "Красногорский завод им. С. А. Зверева", г. Красногорск Московской обл., Россия

Описан подход к автоматическому дешифрированию нескольких разнеспектральных изображений одной сцены, где основными компонентами приняты сама сцена, признаки объектов, зона интереса объекта и сегментация зоны интереса. Примененные понятия зоны интереса и расстояния, заданного в признаковом пространстве, позволяют выстроить методы совместного дешифрирования произвольного числа изображений. Все изложенные результаты иллюстрируются на примерах дешифрирования конкретных трех спектральных изображений.

Известно, что для распознавания объектов, образующих сцену, необходимо располагать индивидуальными наборами характерных признаков, позволяющих отличать их друг от друга. Если исходной информацией о сцене служат ее оптические изображения, то эффективность их дешифрирования зависит в значительной степени от объема представленных на них признаков объектов, которые образуют сцену, т. е. от информативности исходных данных. Один из возможных путей повышения информативности исходных данных состоит в регистрации в каждом пикселе одновременно нескольких признаков. Результатом такой съемки будет набор взаимодополняющих друг друга изображений сцены.

Практической реализации совместного дешифрирования набора изображений в настоящее время нет. Во-первых, предварительные эксперименты показывают, что зрительный анализатор человека практически не приспособлен к анализу информации такого вида. Во-вторых, непосредственный перенос известных формальных методов, применяемых при дешифрировании одиночных изображений на общий случай, когда сцена описывается набором изображений, также пока не реализован. В настоящей работе для сравнения объектов сцены, описываемой набором ее изображений, используется расстояние, определяемое тем или иным способом в многомерном признаковом пространстве, обеспечивающее построение методов дешифрирования независимо от количества используемых изображений.

Модель сцены

В качестве сцены рассмотрим участки местности с различной растительностью и различными

типовыми транспортными средствами. Габаритные размеры и форма транспортных средств предполагаются известными. Исходной информацией о сцене служат наборы оптических изображений, сформированных одновременно в нескольких спектральных зонах оптико-электронными системами, расположенными на летательных (воздушных или космических) аппаратах. Предполагается, что направление съемки, задаваемое оптической осью, близко к нормали относительно поверхности участка сцены. Кроме того, предполагается, что известен масштаб съемки (размер пикселя на местности), гарантирующий присутствие на изображении геометрических признаков заданных объектов. Целью дешифрирования оптических изображений является выявление (поиск) на местности объектов определенных типов с указанием их координат, например, в системе координат, заданной на изображении.

Далее рассмотрим реальную сцену как совокупность неделимых элементов — пикселей. Каждый пиксель характеризуется индивидуальными целочисленными координатами $z = (z_1, z_2)$, заданными на двухмерной целочисленной решетке

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

и векторной (ν -мерной) случайной величиной $\xi_z = (\xi_z^j)_{1 \leq j \leq \nu}$. Каждая ее компонента ξ_z^j , $1 \leq j \leq \nu$ является скалярной случайной величиной со значением из конечного множества $Y = \{0, 1, \dots, |Y| - 1\}$, $|Y| > 1$. Она описывает j -свойства пикселя, значение которого становится известным только после его измерения (съемки). Предполагается, что все скалярные случайные величины определены на одном и том же вероятностном пространстве (Ω, A, P) . Счетное семейство $\xi = (\xi_z)_{z \in Z^2}$ векторных

случайных величин называется, как известно, векторным случайным полем на Z^2 .

Пусть ω — некоторое элементарное событие из Ω . Для каждого пикселя сцены с координатами $z \in Z^2$ и каждого j , $1 \leq j \leq v$ можно вычислить выборочное значение $x_z^j = \xi_z^j(\omega)$ скалярной случайной величины ξ_z^j . Для каждого j , $1 \leq j \leq v$ отображение вида $x^j = (x_z^j)_{z \in Z^2}$ будет называться j -м изображением сцены. В общем случае результатом съемки является набор $x = (x^j)_{1 \leq j \leq v}$ из v таких изображений. Основное внимание в настоящей работе уделяется описанию таких методов, которые имеют место при любых $v \geq 1$.

Как правило, интерес представляют не отдельные пиксели, а их конечные совокупности, так называемые элементы сцены или объекты. Каждый объект определяется подмножеством A точек из Z^2 , являющихся координатами его пикселей, и семейством $\xi_A = (\xi)_{\alpha \in A}$ из $|A|$ векторных случайных величин. Подмножество A назовем проекцией объекта. Всегда будет предполагаться, что все v -мерные случайные величины, образующие объект ξ_A , имеют одно и то же распределение с вектором средних значений $m_A = (m_A^j)_{1 \leq j \leq v}$ и вектором дисперсий $\sigma_A^2 = ((\sigma_A^j)^2)_{1 \leq j \leq v}$.

Признаки

Наиболее часто в качестве исходной информации о сцене используются оптические изображения, на которых представлены значения энергетических яркостей, измеренных в каждом пикселе сцены в заранее выбранных спектральных зонах. Хорошо известно, что энергетические яркости сильно зависят от метеоусловий, состояния объекта и некоторых других условий съемки. Из-за их изменчивости подготовить заранее описания заданных объектов в терминах энергетических яркостей с необходимой точностью не представляется возможным. Поэтому в данной работе распределения вероятностей образующих сцену объектов, их векторы средних значений и векторы дисперсий считаются неизвестными.

При дешифрировании будем пользоваться геометрическими признаками: формой, габаритными размерами, площадью и некоторыми другими данными, вычисляемыми по проекции объекта. Они более устойчивы по отношению к условиям съемки и позволяют заранее подготовить описания объектов. Для описания геометрических свойств объектов применяется Евклидово расстояние d , определенное на двумерной решетке Z^2 .

В качестве площади объекта ξ_A используется количество $|A|$ точек решетки Z^2 , образующих его проекцию. Габаритными размерами объекта называются длины сторон прямоугольника наи-

меньшей площади, описанного вокруг его проекции. Наибольшая сторона — длина объекта, а наименьшая — его ширина. Так как ориентация объектов заранее неизвестна, то используем их диаметр. По определению, диаметром объекта ξ_A будет называться диаметр его проекции $d(A)$, определяемый обычным образом

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

Если d — Евклидово расстояние, то для любого $t \in Z^2$

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

Пиксели с координатами z и t , для которых $d(z, t) = 1$, будут называться “соседями”. При использовании Евклидова расстояния каждый пиксель имеет четыре “соседа”. В общем случае количество “соседей” определяется выбранным расстоянием.

Подмножество A на Z^2 будет называться связным, если для любых двух пикселей с координатами z и t из A существует цепочка из $n \geq 2$ пикселей $(w_j)_{1 \leq j \leq n}$, принадлежащих A и таких, что $z = w_1$, $t = w_n$ и $d(w_j, w_{j+1}) = 1$, $1 \leq j \leq n$. Предполагается, что проекции рассматриваемых в настоящей работе объектов являются связными.

При крупномасштабной съемке характерными примерами объектов, проекции которых являются связными множествами, могут служить транспортные средства. В некоторых случаях интерес представляют группы таких объектов, например походные колонны или боевые порядки. Очевидно, что проекция группы не является связным множеством.

Понятие соседства, определенное для пикселей, обобщается естественным образом для произвольных объектов. Пусть $A \subset Z^2$ и $B \subset Z^2$ — проекции разных объектов. Известно, что расстояние $d(A, B)$ между множествами $d(A, B)$ определяется равенством

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

Поэтому объекты ξ_A и ξ_B сцены, для которых $d(A, B) = 1$, будут называться “соседями”. Предполагается, что у “соседних” объектов сцены разные векторы средних значений. Другими словами, пусть ξ_A и ξ_B — соседние объекты, а $m_A = (m_A^j)_{1 \leq j \leq v}$ и $m_B = (m_B^j)_{1 \leq j \leq v}$ — их векторы средних значений. Тогда должно выполняться неравенство $m_A \neq m_B$ и эквивалентное ему $d(m_A, m_B) > 0$. Отметим, что условие $d(m_A, m_B) > 0$ является формализацией известного правила, которому стараются следовать при выборе состава спектральных зон. Оно означает, что среди рабочих спектральных зон

оптико-электронной системы всегда найдется такая зона, в которой средняя яркость заданного объекта отличается от средней яркости его окрестности (фона).

Использование геометрических признаков предъявляет повышенные требования к масштабу съемки. Практически это означает, что для измерения размеров или определения формы объекта с необходимой точностью требуется, чтобы его изображение содержало от нескольких десятков до нескольких сотен пикселей. Кроме того, вычисление геометрических признаков требует значительных вычислительных ресурсов. Таким образом, повышенные требования к масштабу съемки и вычислительным ресурсам должны учитываться и выполняться при отсутствии сведений о спектральных признаках (энергетических яркостях) заданных объектов, соответствующих условиям съемки. С другой стороны, при наличии такой информации можно вести распознавание даже точечных объектов непосредственно в ходе формирования изображения. При этом значительно снижаются требования к вычислительным ресурсам.

Поиск зон интереса

В ходе поиска заданных объектов приходится на практике выполнять съемку больших по площади участков местности. Тщательное дешифрирование полученных изображений также предполагает наличие значительных вычислительных ресурсов. Требования к ним можно существенно снизить, если принять во внимание, что суммарная площадь, занимаемая всеми представляющими интерес объектами (например транспортными средствами), составляет, как правило, лишь небольшую часть от площади всего участка. Указанное обстоятельство позволяет организовать дешифрирование в два этапа. На первом этапе за счет простых и поэтому быстрых операций выполняется поиск по всему изображению фрагментов сцены, содержащих заданные объекты. Такие фрагменты называют зонами интереса. Возможность решения этой задачи следует из предположения о том, что свойства изображений зон интереса отличаются от свойств изображений фрагментов такой же формы и размеров, но не содержащих объектов. На втором этапе проводится анализ изображений только выявленных зон интереса.

Рассмотрим формальное определение зоны интереса. Пусть A — проекция с диаметром $d(A)$ заданного элемента сцены. Очевидно, что для любого $l \geq d(A) + 2$ существует квадрат $C \times Z^2$ со стороной, равной l такой, что $A \subset (C/Fr(C))$. Будем называть семейство в общем случае v -мерных случайных величин вида $\xi_C = (\xi_c)_{c \in C}$ зоной интереса

для элемента сцены ξ_A , если все случайные величины $\xi_c, c \in C/A$ имеют одно и то же распределение вероятностей с вектором средних значений $m_{C/A} = (m^j_{C/A})_{1 \leq j \leq v}$ и вектором дисперсий $\sigma^2_{C/A} = ((\sigma^j_{C/A})^2)_{1 \leq j \leq v}$. Из этого определения следует, что других объектов в зоне интереса нет. Набор пикселей $\xi_{C/A} = (\xi_c)_{c \in C/A}$ зоны интереса будет называться окрестностью элемента ξ_A .

Очевидно, что размер l зоны интереса конкретного объекта ξ_A ограничен снизу его диаметром $d(A)$, а сверху — расстоянием до ближайшего объекта этого же класса. Если r — минимальное расстояние между объектами класса, то в данном случае $l, d(A)$ и r связаны между собой неравенством вида

$$d(A) + 2 \leq l \leq r / \sqrt{2}.$$

Пусть z_0 — произвольный узел Z^2 , $e_1 = (1, 0)$ и $e_2 = (0, 1)$ — единичные векторы, а i и j — целые числа. Зададим на Z^2 семейство $C(l, z)$ квадратов со стороной l и левой верхней вершиной

$$z = z_0 + i(l - d(A))e_1 + j(l - d(A))e_2.$$

Можно показать, что всегда существует квадрат C из семейства $C(l, z)$ такой, что $A \subset (C/Fr(C))$.

Установим два свойства зоны интереса, необходимые для ее поиска. Пусть ξ_A — объект, ξ_C — его зона интереса, а $x_C = (x^j_C)_{1 \leq j \leq v}$ — набор ее изображений. Если $B(a, r) \subset A$ и $B(z, r) \subset (C/A)$ — круги из $n = |B(a, r)|$ узлов каждый, а

$$\bar{x}_a = \frac{1}{n} \sum_{t \in B(a,r)} x_t \quad \text{и} \quad \bar{x}_z = \frac{1}{n} \sum_{t \in B(z,r)} x_t$$

оценки для неизвестных координат векторов m_A и $m_{C/A}$ средних значений объекта и фона, соответственно, то при больших n справедливы, как известно, следующие равенства :

$$\bar{x}_a \approx m_A, \quad \bar{x}_z \approx m_{C/A} \quad \text{и} \quad d(\bar{x}_a, \bar{x}_z) \approx d(m_A, m_{C/A}) > 0.$$

С другой стороны, пусть C — квадрат на Z^2 , $\xi_C = (\xi_c)_{c \in C}$ — семейство одинаково распределенных случайных величин с неизвестным вектором средних значений $m_C = (m^j_C)_{1 \leq j \leq v}$ и вектором дисперсий $\sigma^2_C = ((\sigma^j_C)^2)_{1 \leq j \leq v}$, а $x_C = (x^j_C)_{1 \leq j \leq v}$ — набор изображений. Если $B(z, r)$ и $B(t, r)$ — непересекающиеся круги из C и

$$\bar{x}_z = \frac{1}{n} \sum_{t \in B(z,r)} x_t \quad \text{и} \quad \bar{x}_t = \frac{1}{n} \sum_{z \in B(t,r)} x_z$$

оценки неизвестных координат вектора средних значений, то при больших n справедливы следующие равенства:

$$\bar{x}_z \approx m_{C/A}, \quad \bar{x}_t \approx m_{C/A} \quad \text{и} \quad d(\bar{x}_t, \bar{x}_z) \approx d(m_{C/A}, m_{C/A}) = 0.$$

Пусть $B(a, r) \subset A$ — круг с центром $a \in A$ и радиусом r , лежащий в проекции объекта и \bar{x}_a — оценка неизвестного вектора $\bar{x}_z \approx m_A$ средних значений объекта. Построим на Z^2 квадрат C с центром в точке a и стороной $2d(A) + 2$. Очевидно, что его граница $Fr(C)$ не имеет общих точек с проекцией A . Разобьем $Fr(C)$ на s частей Fr_j , $1 \leq j \leq s$ по n пикселей в каждой. Если квадрат C является зоной интереса, то

$$\bar{x}_{Fr} = \frac{1}{|Fr(C)|} \sum_{t \in Fr(C)} x_t \quad \text{и} \quad \bar{x}_j = \frac{1}{n} \sum_{z \in Fr_j} x_z, \quad 1 \leq j \leq s$$

являются оценками неизвестного вектора $m_{C/A}$ средних значений фона. Поэтому естественно ожидать, что

$$d(\bar{x}_a, \bar{x}_{Fr}) > d(\bar{x}_j, \bar{x}_{Fr}) \approx 0, \quad 1 \leq j \leq s.$$

Теперь рассмотрим сам поиск зон интереса для объекта ξ_A с использованием набора изображений сцены. Вначале по известной проекции заранее вычисляем диаметр $d(A)$ объекта, который не зависит от его ориентации. Затем для каждого квадрата со стороной, равной $2d(A) + 2$, проверяем последнее неравенство. Если оно выполняется, то квадрат считается зоной интереса.

Сегментация в зоне интереса

Для измерения геометрических признаков объектов необходимо знать их проекции. Построение проекций образующих сцену объектов называется, как известно, ее сегментацией. После проведения поиска зон интереса выполнять сегментацию потребуется только в них.

По определению зоны интереса, каждый пиксель, описываемый векторной случайной величиной, принадлежит или объекту, или его окружению (фону). Исходной информацией для классификации пикселей зоны интереса на два указанных класса служит набор ее изображений. Кроме того, из определения зоны интереса следует, что пиксели, образующие ее границу, относятся к фону. Для построения решающего правила классификации пикселей зоны интереса воспользуемся свойствами ее пикселей, установленными выше.

Пусть z — координаты пикселя из $C/Fr(C)$. Используя набор изображений, вычислим оценку \bar{x}_z неизвестного вектора средних значений случайной величины ξ_z . Если

$$d(\bar{x}_z, \bar{x}_{Fr}) > d(\bar{x}_j, \bar{x}_{Fr}) \approx 0, \quad 1 \leq j \leq s,$$

то пиксель с координатами z следует считать принадлежащим объекту, в противном случае — фону.

Компьютерные эксперименты (моделирование)

Для проведения компьютерных экспериментов использовались изображения участка сцены квадратной формы со стороной, равной 256 пикселям. Размер пикселя на местности составил 0,3 м. Поверхность выбранного участка была покрыта травой, редким кустарником и отдельными деревьями. На ней были заметны следы транспортных средств. Съемка проводилась летом с летательного аппарата в трех спектральных зонах (0,7—1,1, 3—5 и 8—12 мкм), соответствующих окнам прозрачности атмосферы.

При выбранном разрешении элементы конструкций транспортных средств в указанных спектральных зонах практически не просматриваются. Поэтому в экспериментах изображения шести реальных объектов были заменены прямоугольниками со сторонами 10×20 пикселей (3×6 м). Средние значения и дисперсии яркостей прямоугольников соответствуют средним значениям и дисперсиям яркостей реальных объектов, зарегистрированных на момент съемки. Набор описанных изображений представлен на рис. 1. Их сравнение позволяет заметить небольшие различия в линейных размерах элементов фотоприемных устройств. Отметим, что на одном изображении (в средневолновом ИК-диапазоне) два объекта из шести почти не отличаются от окружающего их фона.

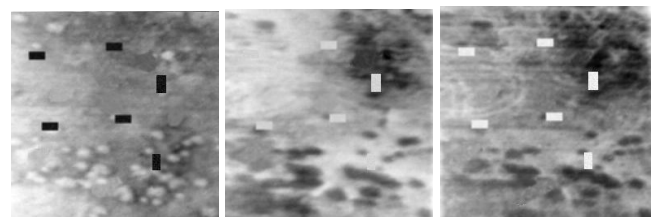


Рис. 1. Исходные изображения

На рис. 2 представлены результаты дешифрирования каждого из трех изображений без учета остальных двух. Дешифрирование включает поиск зон интереса, сегментацию зон, вычисление признаков и принятие решения о присутствии (отсутствии) объекта в зоне. В результате дешифрирования первого изображения обнаружено восемь зон интереса. В ходе последующего анализа в шести из них распознаны объекты. Две зоны интереса признаны ложными на этапе сегментации. При дешифрировании второго изображения найдено также восемь зон интереса. Последующий анализ показал, что три зоны содержат заданные объекты. Одна зона признана ложной на этапе сегментации,

а четыре зоны — на этапе принятия решения. Отметим, что при дешифрировании второго изображения не были найдены три зоны. В результате дешифрирования третьего изображения обнаружено десять зон интереса. В ходе последующего анализа в шести из них распознаны объекты. Четыре зоны были отбракованы на этапе принятия решения.

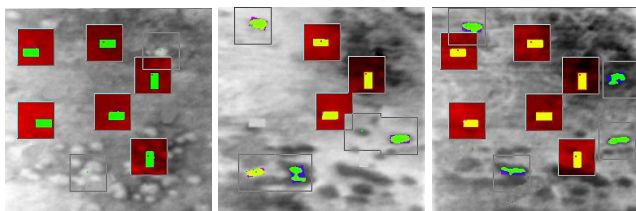


Рис. 2. Дешифрирования отдельных изображений

Из трех изображений можно создать три набора по два изображения и один — из трех изображений. Результаты дешифрирования трех наборов из четырех возможных представлены на рис. 3. Слева представлены результаты совместного дешифрирования первого и второго изображений, содержащие один ложный объект; в центре — результаты совместного дешифрирования второго и третьего изображений; пропущены два объекта. На рис. 3 справа представлены результаты совместного дешифрирования всех трех изображений; все присутствующие объекты обнаружены; ложных объектов нет.

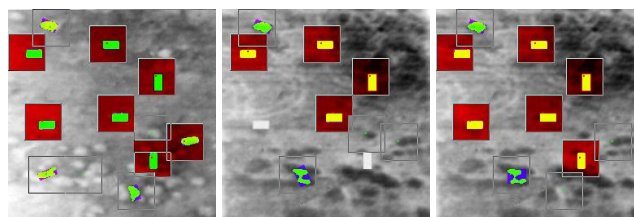


Рис. 3. Результаты совместного дешифрирования изображения

Заключение

Предложен подход к организации совместного дешифрирования произвольного количества оптических изображений сцены. Цель такого подхода — поиск на сложной сцене (реальной местности с различной растительностью) заданных объектов (например транспортных средств) по их геометрическим признакам. Основными составными частями при этом служат формальные описания сцены, признаков объектов, зоны интереса и сегментации зоны интереса. Все приводимые результаты иллюстрируются на примерах дешифрирования конкретных изображений. Очевидно, что применимость предложенного метода не ограничивается только транспортными средствами.

Авторы признательны д-ру техн. наук Р. М. Алееву (г. Казань) за полезное обсуждение подходов к данной проблеме, в результате которого работа приобрела практическую направленность.

Статья поступила в редакцию 5 октября 2006 г.

Approach to automated decoding of multispectral images

A. B. Belsky, V. B. Fofanov

JSC “Krasnogorsky Zavod im. S. A. Zvereva”, Krasnogorsk, Russia

Approach to automated decoding of several multispectral images of one scene where the main components are the scene itself, features of objects, the object of interest area and object of interest area segmentation is described. Concepts used for area of interest and distance given in feature space to build the methods of combined decoding of arbitrary number of images. All shown results are illustrated using the examples of decoding of three specific spectral images.