

УДК 004.932:528.85

В. А. Сметанин, аспирант, Физико-технический институт УрО РАН, Ижевск

В. Н. Милич, кандидат технических наук, Физико-технический институт УрО РАН, Ижевск

ИССЛЕДОВАНИЕ БИСПЕКТРАЛЬНЫХ МАТРИЦ РАССЕЯНИЯ ПРИ АНАЛИЗЕ МАТЕРИАЛОВ МНОГОЗОНАЛЬНОЙ СЪЕМКИ

Исследована система текстурных признаков изображений, построенная на основе биспектральных матриц рассеяния. Для принятия решения по признакам использовалась нейронная сеть. Приведены результаты экспериментов на реальных данных многозональной съемки. Представлены установленные зависимости качества обучения сети и распознавания от продолжительности и степени обучения.

Ключевые слова: многозональные изображения, нейронные сети, текстурные признаки изображений, распознавание образов.

Дистанционное зондирование Земли позволяет получать информацию о земной поверхности без непосредственного контакта с исследуемыми объектами. В последние годы возрастает роль много- и гиперзональных съемочных систем. Однако их главное преимущество – большой объем содержащейся информации – может быть использовано только при наличии эффективных методик извлечения этой информации. Поэтому разработка новых методов обработки многозональных спутниковых изображений является актуальной в настоящее время.

В настоящей статье представлены результаты исследования классификации фрагментов изображения многозональных космических снимков с целью поиска эффективных признаков изображений. Выбираемый для классификации образец представляет собой фрагмент изображения размером $N_x \times N_y$. Для описания изображений используется обобщение матриц рассеяния (co-occurrence matrices), предложенное ранее авторами [1]. Обобщение распространяет метод построения матрицы рассеяния для панхроматического изображения [2] на «перекрестные» биспектральные матрицы [1, 4]. Для конкретного образца для каждой пары каналов вычисляется матрица рассеяния [1] размером $N \times N$, где N – количество градаций яркости в одном канале. Элемент матрицы $P(i, j, k, l, d, \alpha)$ – относительная частота появления на фрагменте пары пикселей на расстоянии d в направлении α таких, что первый из них имеет в канале k яркость i , а второй – в канале l яркость j . Матрица строится для каждой пары каналов (k, l) , для каждого взаимного расположения пикселей в паре (d, α) . Элемент матрицы $P(i, j, k, l, d, \alpha)$ для краткости обозначается P_{ij}^{kl} . Помимо матрицы рассеяния для характеристики используется $f_l(x, y)$ – функция значений яркости пикселя с координатами x, y на анализируемом фрагменте в канале l , а также функция $F_l(n)$, характеризующая количество пикселей с яркостью n на фрагменте в канале l .

В работе используются следующие текстурные признаки [1, 2, 3].

- Энергия $E_{lk} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (P_{i,j}^{l,k})^2$. (1)

- Энтропия $\varepsilon_{lk} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P_{i,j}^{l,k} \ln P_{i,j}^{l,k}$. (2)

- Максимум матрицы $m_{lk} = \max_{i,j} P_{i,j}^{l,k}$. (3)

- Контраст $C_{lk} = \sum_{n=1}^{N-1} n^2 \left(\sum_{i=1}^{N-1} \sum_{j=1}^{N-1} P_{i,j}^{l,k} \right)_{|i-j|=n}$. (4)

- Однородность $O_{l,k} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{P_{i,j}^{l,k}}{1+|i-j|}$. (5)

- Начальный момент $\mu_l = \sum_{x=0}^{N_x-1} \sum_{y=0}^{N_y-1} \frac{f_l(x,y)}{N_x N_y}$. (6)

- Вариация $V_l = \sum_{g=0}^{N-1} (g - \mu_l)^2 F_l(g)$. (7)

Для принятия решения использовалась нейронная сеть с архитектурой, аналогичной сети в статье [3]. Нейронная сеть представляет собой [5] набор связанных вычислительных узлов – нейронов, имеющих n входов – синапсов, и один выход. Каждый синапс характеризуется весовым коэффициентом или синаптическим весом $w_{ij}, i = 0, \dots, n, j$ – номер нейрона в слое. Коэффициент w_0 задает пороговое значение. На вход нейрона подается n -мерный вектор $X = (x_1, \dots, x_n)$. При обработке вычисляется взвешенная сумма входных сигналов

$$s = \sum_{i=0}^n x_i w_{ij}, \quad (8)$$

причем $x_0 \equiv 1$. Затем вычисляется активационная функция

$$f(s) = \alpha \tanh(\beta s), \quad (9)$$

где α, β – параметры. Значение активационной функции подается на выход нейрона. Нейроны объединены в слои. Выходные значения предыдущего слоя являются входными значениями следующего

слоя. На вход первого слоя подаются нормированные текстурные признаки $x_i \in [-1, 1]$, для нейронов этого слоя n (число входов для каждого нейрона) равно количеству признаков. На выходе последнего слоя – вектор решения. Количество нейронов в последнем слое соответствует количеству классов в рассматриваемой модели. Последний слой называется выходным, остальные – скрытыми.

Выбрана следующая архитектура нейронной сети: многослойный персептрон с функцией активации гиперболического тангенса с двумя скрытыми слоями. Нейронная сеть содержит три слоя, каждый вход нейрона следующего слоя связан со всеми выходами предыдущего слоя. Параметры активационной функции имеют значения $\alpha = 1,7$; $\beta = 2/3$ [5].

Используется обучение «с учителем». Обучение происходит по алгоритму обратного распространения ошибки [5]. Процесс обучения является итерационным. Изначально весовые коэффициенты задаются случайными, близкими к нулю значениями. Затем выполняется процедура, включающая вычисление выходного вектора сети по поданным на вход значениям признаков обучающего образца, вычисление локального градиента для всех нейронов на основании ошибок выхода, полученных при сопоставлении с вектором ожидаемого отклика и модификацию синоптических весов для всех нейронов. Вектор ожидаемого отклика – набор априорных вероятностей принадлежности образца к существующим классам. Так, вектор (1, 0, 0, 0) соответствует предположению о принадлежности классу 1 из четырех предъявленных. Процедура итеративно повторяется либо заданное заранее количество эпох, либо до достижения достаточной близости выходных значений к вектору ожидаемого отклика. Обучающая эпоха – это цикл, в котором все образцы всех классов последовательно предъявляются для обучения сети один раз, после каждого предъявления корректируются значения весов.

Продолжительность обучения зависит от количества входных параметров (признаков), количества классов, степени обучения (количества образцов каждого класса, используемых при обучении), а также от перекрываемости значений признаков разных классов.

В качестве показателя, характеризующего качество обучения сети на отобранных классах объектов, предлагается использовать разность между компонентой выходного вектора, соответствующей правильному решению, и максимальной из оставшихся компонент. Далее будем называть эту величину уровнем правдоподобия (УП). Для оценки качества распознавания используется вероятность правильного распознавания тестовых образцов.

Исходные данные

В экспериментах с реализованной нейронной сетью использовались три группы исходных данных. В первую группу вошли модельные данные: 10 классов по 10 реализаций, заданных 10 признаками. Данные формировались так, чтобы по каждому признаку

классы отличались. Среднее значение i -го признака k -го класса выражается как $z(i, k) = 0,1 \times ((i + k) \bmod 10) - 0,05$. Значения признаков реализаций этих классов генерировались с равномерным распределением в окрестности среднего $\pm 0,02$.

Аналогично строятся исходные данные второй группы. Однако максимальное отклонение от среднего составляет $\pm 0,07$, что обеспечивает пересечение значений по некоторым признакам и усложняет процесс обучения нейронной сети.

В третью группу вошли реальные данные снимка окрестностей г. Камбарка со спутника Quickbird (разрешение 2,8 м). Система регистрации включает 4 канала: синий (B), зеленый (G), красный (R) и ближний инфракрасный (I) спектральные диапазоны. На снимке задано 7 классов природных и антропогенных объектов. В набор анализируемых классов объектов вошли: водная поверхность, пашня, болото, хвойный лес, лиственный лес, вырубки и огороды. Для каждого класса выбрано по 10 обучающих и 10 тестовых образцов.

По этим образцам вычислены признаки по всем парам каналов (всего 88 признаков [1]). Для расчета признаков использовались две биспектральные матрицы рассеяния по 5 признаков от каждой (1–5), и по два признака, вычисляемых непосредственно по фрагменту (6–7) с двух отдельных каналов. Таким образом, в каждом эксперименте на вход подавалось 14 признаков, которые выбирались из исходных 88 в зависимости от рассматриваемой комбинации каналов.

Для описания содержания входных данных будем использовать аббревиатуру, содержащую информацию о матрицах и каналах. Так, $RR + GR + B + I$ будет обозначать, что использовалась матрица рассеяния для красного канала, биспектральная матрица между зеленым и красным каналами и признаки по фрагменту в синем и инфракрасном канале.

Полученные результаты

На рис. 1 показаны графики зависимости УП от количества эпох обучения для модельных данных.

По этим графикам видно, что количество эпох, необходимых при обучении для достижения достаточного УП, существенно зависит от различимости входных данных. Для достижения этого уровня на реальных данных потребуется значительно более продолжительное обучение.

Для реальных данных исследовались следующие комбинации каналов: $BI + GR + B + G$, $GG + GR + G + R$, $RR + BI + B + R$, $BB + II + B + I$. Использовались 10 обучающих образцов каждого класса.

На рис. 2 представлены графики, показывающие качество распознавания 7 классов по 10 образцам каждого класса реальных данных.

Графики показывают, что за 100000 эпох все наборы данных обучают сеть до УП большего 0,9. Исключением является набор $BB + II + B + I$, с которым УП достигает только 0,8. Отметим, что этот набор отличается от других тем, что в нем не используются признаки биспектральных матриц рассеяния.

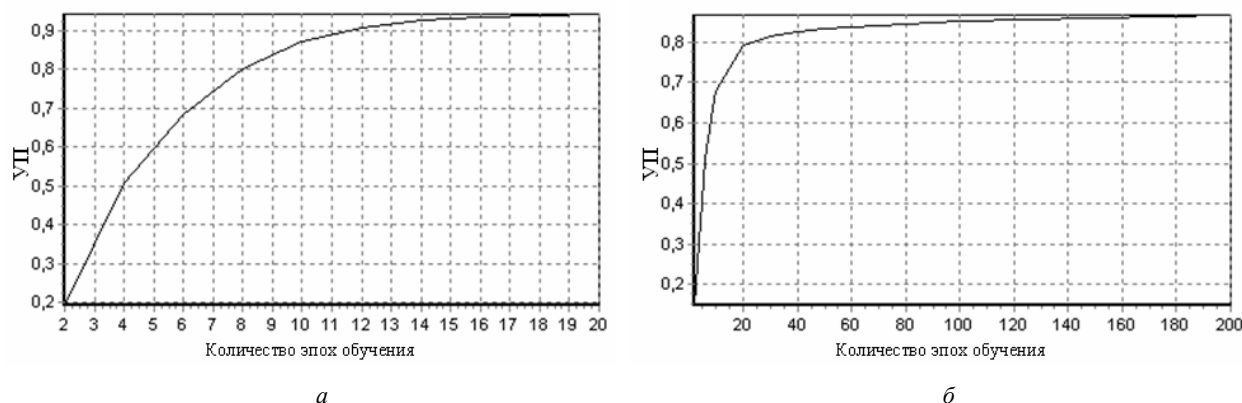


Рис. 1. Идеально различимые (а) и частично перекрывающиеся (б) данные

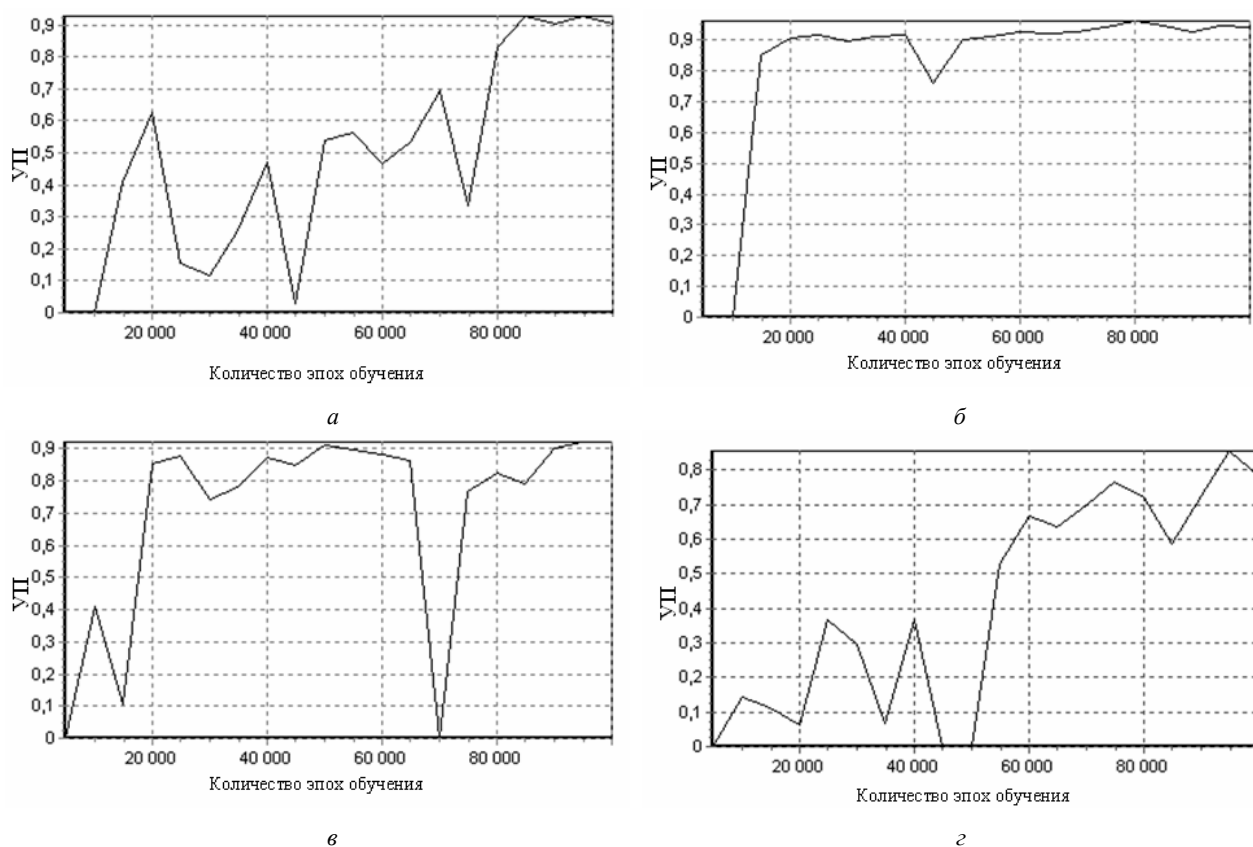


Рис. 2. Зависимости УП от продолжительности обучения для комбинаций каналов:
 а – $BI + GR + B + G$; б – $GG + GR + G + R$; в – $RR + BI + R + B$; г – $BB + II + B + I$

Немонотонное возрастание УП объясняется значительным разнообразием образцов изображений объектов внутри класса и наличием межклассовых пересечений по некоторым признакам. В этом отношении преимущество имеет набор $GG + GR + G + R$, не имеющий сильного выброса графика зависимости вниз.

При тестировании на образцах, не использованных при обучении, рассматривались два критерия принятия решения:

- строгое распознавание (СР) – отклонение от вектора ожидаемого отклика не более чем на 0,1 по норме максимума [5, 7] (УП не меньше 0,8);

- нестрогое распознавание (НР) – компонента выходного вектора, имеющая номер класса, должна превышать 0,5, в то время как остальные компоненты должны быть меньше 0,5 (положительный УП).

Для анализа зависимости результатов распознавания от степени обучения проведен следующий эксперимент. Сеть обучалась на образцах 7 классов в течение 100000 эпох с разной степенью обучения (2, 5, 7, 10). Для тестирования использовались по 10 образцов, не участвовавших в обучении. В таблице приведены вероятности правильного распознавания (в процентах) по двум предложенным критериям.

Зависимость вероятности распознавания от степени обучения

Степень обучения	Используемые комбинации каналов							
	$BB + II + B + I$		$BI + GR + B + G$		$GG + GR + G + R$		$RR + BI + R + B$	
	СП	НР	СП	НР	СП	НР	СП	НР
2	32,8	77,1	41,4	77,1	51,4	90,0	51,4	75,7
5	45,7	74,3	40,0	78,6	68,6	80,0	45,7	75,7
7	55,7	80,0	71,4	81,4	72,5	84,3	53,0	84,3
10	22,8	85,7	44,2	78,6	62,5	97,1	64,3	85,7

Из таблицы видно, что увеличение степени обучения повышает вероятность распознавания. Набор входных данных $GG + GR + G + R$ показывает лучший результат.

Таким образом, разработанная система признаков на основе биспектральных матрицы рассеяния показала свою высокую работоспособность на реальных данных многозональной съемки. При достаточной степени обучения вероятность распознавания на семи классах составляет не менее 0,8.

Качество обучения хорошо описывается показателем уровня правдоподобия. Критерием остановки обучения является $УП > 0,9$. В приведенных примерах необходимый уровень на обучающих примерах был достигнут за 100000 эпох. Для оценки качества распознавания работы системы целесообразно использовать вероятность правильного распознавания. Показано, что с увеличением степени обучения она достигает значений больше 0,8 для правила нестрогого распознавания. То есть большая степень обучения позволяет достичь большей обобщающей способности нейронной сети.

Установлено также, что межканальная информация о текстуре, содержащаяся в биспектральных матрицах, дает дополнительную различительную

информацию, что подтверждается худшими показателями по набору с матрицами по отдельным каналам ($BB + II$).

Библиографические ссылки

1. Сметанин В. А., Милич В. Н. Возможности использования текстурных признаков изображений для анализа материалов многозональной съемки // Материалы VII Всерос. науч.-техн. конф. «Приборостроение в XXI веке. Интеграция науки, образования и производства». – Ижевск : Изд-во ИжГТУ, 2012. – С. 484–489.
2. Haralick R. M., Shanmugam K., Dinstein It'sHak. Textural Features for Image Classification // IEEE Transactions on systems, man and cybernetics. – 1973. – Vol. SMC_3. – № 6. – P. 610–621.
3. Астафуров В. Г., Скороходов А. В. Классификация облаков по спутниковым снимкам на основе топологии нейронных сетей // Современные проблемы дистанционного зондирования Земли из космоса. – 2011. – Т. 8. – № 1. – С. 65–72.
4. Generalization of the Cooccurrence Matrix for Colour Images: Application to Colour Texture Classification / V. Arvis [et al.] // Image Analysis and Stereology. – 2004. – Vol. 23. – P. 63–72.
5. Хайкин С. Нейронные сети: полный курс : пер. с англ. – 2-е издание // М. : Вильямс, 2006. – 1104 с.

V. A. Smetanin, Post-graduate, Physical-Technical Institute of the Ural Branch of RAS, Izhevsk

V. N. Milich, PhD in Engineering, Physical-Technical Institute of the Ural Branch of RAS, Izhevsk

Investigation of Bispectral Co-Occurrence Matrices for Multispectral Data Analysis

A system of image texture features developed on the basis of bispectral co-occurrence matrices is explored. A neural network is used for classifying decision over the features. Results of experiments with the real multispectral data are showed. The paper presents the established dependencies of quality for both the network learning and the recognition on the learning duration and the training extent.

Key words: multispectral images, neural network, image texture features, pattern recognition.