



УДК 004.932

СТРУКТУРА НЕЙРОСЕТИ ДЛЯ РАСПОЗНАВАНИЯ ОБЪЕКТОВ НА АЭРОКОСМИЧЕСКИХ ИЗОБРАЖЕНИЯХ НА ОСНОВЕ АНАЛИЗА РАСПРЕДЕЛЕНИЯ ИХ ЭНЕРГИЙ ПО ЧАСТОТНЫМ ИНТЕРВАЛАМ

Е.Г. ЖИЛЯКОВ**А.Ю. ЛИХОШЕРСТНЫЙ****В.В. КРАСИЛЬНИКОВ**

*Белгородский государственный
национальный исследовательский
университет*

*e-mail: Zhilyakov@bsu.edu.ru
ozzy.osbourne.man@gmail.com*

В статье рассматривается новый метод распознавания объектов на снимках земной поверхности на основе построения нейронной сети, в которой вычисляются доли энергии выделяемых объектов.

Ключевые слова: аэрокосмические снимки, нейронная сеть, частотное представление, доли энергии, распознавание образов.

Мониторинг состояния земной поверхности осуществляется на основе процесса дешифрирования снимков и, в частности, выделения объектов с той или иной точки зрения, составляющий единый класс. Такие процедуры естественно называть распознаванием объектов. Причем для реализации этой процедуры применяется предварительное обучение, заключающееся в описании класса искомого объекта на основе задания значений признаков из некоторого их пространства. Выбор пространства признаков является важнейшим этапом подготовки к решению задачи распознавания объектов. Другой важный аспект процедуры распознавания заключается в выборе инструмента, с помощью которого производится сопоставление значений признаков анализируемого объекта с их значениями, полученными на этапе обучения. Предполагается, что этап обучения осуществляется по объекту-образу (фрагменту изображения), который указывается оператором, в том числе непосредственно на обработанном изображении.

В каждой из задач распознавания присутствует некоторый объект, представленный значениями своих признаков, а также некоторое число классов, к одному из которых необходимо отнести данный объект. Таким образом, распознавание — это отнесение некоторого неизвестного объекта по его описанию к одному из классов.

Распознавание производится по прецедентам, классификация которых известна и которые в виде формализованных описаний могут быть предъявлены алгоритму распознавания для настройки на задачу в процессе обучения.

Одним из современных подходов распознавания объектов на аэрокосмических изображениях является построение нейронных сетей. Одной из ведущих программных систем, использующих для дешифрирования нейронные сети, является ENVI. Для автоматизированного распознавания объектов на аэрокосмических снимках эта программная система использует следующие три группы признаков распознавания: геометрические (форма, размер); яркостные (уровень яркости, цвет); структурные (текстура, структура). Недостатками таких подходов являются:

- сложность распознавания в условиях сильных помех на изображении;
- необоснованное отнесение объектов снимка в один и тот же класс (рис. 1);
- слабая адаптация существующих методов автоматизированного распознавания к изменяющимся условиям съемки и обработки изображений;
- для разных классов объектов порой приходится подбирать различные функциональные модели объектов.



Рис. 1. Результат распознавания частных домов байесовским методом в ENVI

Одной из перспективных возможностей изменения ситуации в области автоматизированного распознавания являются разработка и внедрение программно-технологических средств, которые используют новые методы обработки космофотоснимков на основе частотных представлений.

Новизна идеи состоит в использовании нового метода обнаружения и выделения объектов, где в качестве группы признаков распознавания берутся энергетические характеристики изображения (доли энергии распознаваемых объектов) и на их основе строится нейронная сеть.

Основная цель работы состоит в разработке, исследовании и программной реализации алгоритма нейросетевого распознавания объектов на аэрокосмических изображениях на основе анализа распределения их энергий по частотным интервалам.

Поставленная цель достигается решением следующих конкретных задач:

1. Вычисление долей энергии изображения, которые являются многомерным вектором признаков распознавания;
2. Построение нейронной сети для выделения и идентификации объектов на космических снимках, группами признаков распознавания для которого являются доли энергии каждого объекта изображения;
3. Оценивание работоспособности разработанных алгоритмов на основе вычислительных экспериментов с реальными изображениями.

При распознавании объектов на изображении исходными данными является массив значений интенсивностей для каждого пикселя. Очевидно, этот массив можно представить в виде многомерного вектора признаков, но распознавание в подобном пространстве признаков неосуществимо. Даже небольшая смена ракурса приведет к тому, что значения интенсивностей большинства пикселей изменятся, т.е. образы одного класса в данном пространстве занимают области, имеющие очень сложные формы. Понятно, что каждый из этих признаков (значений интенсивностей отдельных пикселей) мало информативен: даже выкинув значительную часть пикселей, человек сможет распознать объект практически также хорошо, как и по исходному изображению. С другой стороны, чтобы отличить, например, мяч от линейки, нам достаточно только одного такого признака, как форма. Это означает, что в рамках данной задачи существуют некоторые признаки, в пространстве которых классы образов будут хорошо отделимы с помощью сравнительно простых процедур



распознавания. К сожалению, такие признаки очень сложным образом зависят от исходных описаний, с чем и связана трудность проблемы выбора признаков.

Основной проблемой, связанной с представлениями на пиксельном уровне, является проблема эффективности этих представлений в целях их хранения и передачи. Представления на пиксельном уровне являются исходными для любых приложений интерпретации изображений с помощью цифровых вычислительных машин. Именно потому, что такое представление является общим для различных задач компьютерной обработки изображений, часто говорят, что «изображения – это массивы пикселей».

Представления на пиксельном уровне содержат в себе всю имеющуюся информацию о наблюдаемой сцене, но в форме, неудобной для автоматического анализа. Это и вызывает необходимость привлечения других представлений изображений с целью извлечения содержащейся в них релевантной информации.

В связи с этим возникает естественное желание представить изображение как элемент некоторого математического пространства, чтобы воспользоваться уже введенными на нем операциями. Часто отображение из пиксельного представления в выбранное математическое пространство является взаимно однозначным и непосредственно выражается через исходные значения интенсивностей, а результаты математических операций над изображениями снова представляются в виде массива пикселей.

Существующие на данный момент строгие математические модели изображений являются достаточно низкоуровневыми и имеют ограниченную область применения.

В качестве основных причин различия внешнего вида одного и того же объекта на разных изображениях можно назвать[7]:

- смену ракурса (пространственное преобразование);
- изменение освещения (преобразование интенсивности);
- смену типа сенсора;
- собственную изменчивость объекта (например, сезонные изменения на аэрокосмических фотографиях).

Следовательно, для назначения меток необходимо построить описания изображений, инвариантные перечисленным преобразованиям. К сожалению, математические модели позволяют добиться инвариантности только по отношению к весьма ограниченным классам преобразований. Изменчивость других типов, вызванная, например, сменой сенсора или собственными изменениями объекта, оказывается трудноформализуемой, так как сильно зависит от свойств объектов наблюдения.

Одна из проблем заключается в том, чтобы построить инвариантные представления, потеряв при этом как можно меньше полезной информации. Представления, решающие в той или иной степени эту задачу, обращаются к структуре изображения, т. е. к взаимосвязям между различными пикселями, поэтому такие представления часто называют структурными.

Структура возникает из-за периодического повторения значений коэффициента отражения точек некоторой физической поверхности, что вызывает периодические (но, возможно, искаженные перспективой) повторения значений интенсивностей пикселей изображения [7]. Природные объекты, однако, не обладают подобной регулярной периодической структурой. Тем не менее пространственные вариации значений коэффициента отражения у таких объектов также содержат некие закономерности, различные для разных объектов. При этом повторяются не интенсивности отдельных пикселей, а значения некоторых параметров (текстурных признаков) в локальных распределениях интенсивностей пикселей. Это позволяет выделять на изображении области, соответствующие разным объектам, что и является целью структурной сегментации изображения. Структурный анализ может использоваться не только для выделения объектов на изображении, но и непосредственно для их распознавания.

Изображение можно определить как двумерную функцию f_{ik} , где $i=1,2,\dots,M$ и $k=1,2,\dots,N$ – координаты в пространстве (конкретно, на плоскости), и значение f которой в любой точке, задаваемой парой координат (i,k) , называется интенсивностью изображения в этой точке [1,2,6].



Частотным представлением функции f_{ik} называется следующее выражение [3-5]:

$$f_{ik} = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} F(u, v) e^{ju_i} e^{jv_k} du dv \quad (1)$$

Здесь e^{ju_i} и e^{jv_k} – базисные функции (функции представляются в базисе Фурье), аргументы u и v – круговые частоты, отражающие периодичность (цикличность) изменений исходной функции f_{ik} с изменением аргументов i и k .

В качестве весовой функции $F(u, v)$ можно использовать трансформанту Фурье:

$$F(u, v) = \sum_{i=1}^M \sum_{k=1}^N f_{ik} e^{-ju(i-1)} e^{-jv(k-1)} \quad (2)$$

На основе равенства Парсеваля энергию изображения можно представить в виде суммы:

$$\sum_{i=1}^M \sum_{k=1}^N f_{ik}^2 = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |F(u, v)|^2 du dv = \sum_{r=1}^{R_x} \sum_{m=1}^{R_y} P_{\Omega_{r,m}} \quad (3)$$

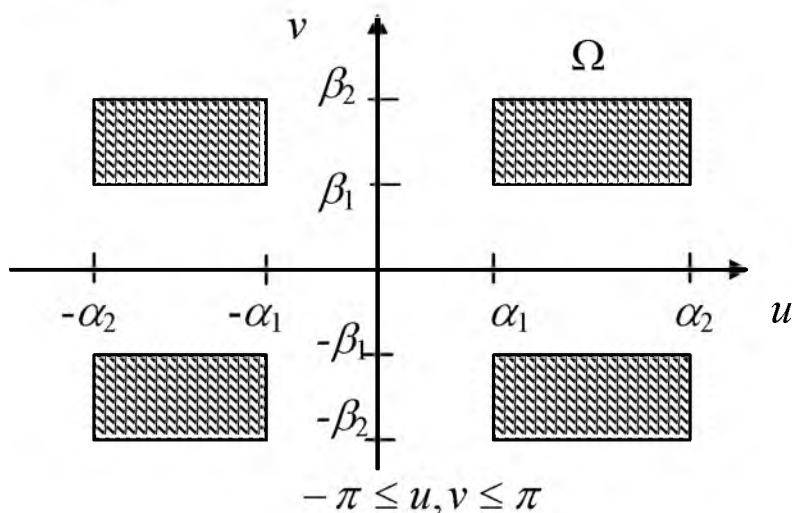
где значение доли энергии $P_{\Omega_{r,m}}$ двумерного сигнала f_{ik} в двумерной частотной области Ω_{ik} :

$$P_{\Omega_{r,m}} = \frac{1}{4\pi^2} \iint_{(u,v) \in \Omega_{ik}} |F(u, v)|^2 du dv \quad (4)$$

В качестве области Ω_{ik} рассматривается следующая центрально-симметричная область частотной плоскости [4]:

$$\Omega_{ik} : \{ \Omega_{ik}(u, v) \mid (u \in [\alpha_1, \alpha_2], v \in [\beta_1, \beta_2]) \cup (u \in [\alpha_1, \alpha_2], v \in [-\beta_2, -\beta_1]) \cup (u \in [-\alpha_2, -\alpha_1], v \in [-\beta_2, -\beta_1]) \cup (u \in [-\alpha_2, -\alpha_1], v \in [\beta_1, \beta_2]) \} \quad (5)$$

где $0 \leq \alpha_1, \alpha_2, \beta_1, \beta_2 \leq \pi$



Если в правую часть представления (4) подставить определение (2), то после преобразований можно получить соотношение [5]:

$$P_{rm} = \frac{\text{tr} (A_r^T \cdot \Phi \cdot B_m \cdot \Phi^T)}{\sum_{i=1}^M \sum_{k=1}^N f_{ik}^2} \quad (6)$$

где элементы матриц $A=(a_{i_1i_2})$ и $B=(b_{k_1k_2})$ вычисляются следующим образом:

$$a_{i_1i_2} = \begin{cases} \frac{\sin(\alpha_2(i_1 - i_2)) - \sin(\alpha_1(i_1 - i_2))}{\pi(i_1 - i_2)}, & i_1 \neq i_2, \\ \frac{\alpha_2 - \alpha_1}{\pi}, & i_1 = i_2, \end{cases}$$

$$b_{k_1k_2} = \begin{cases} \frac{\sin(\beta_2(k_1 - k_2)) - \sin(\beta_1(k_1 - k_2))}{\pi(k_1 - k_2)}, & k_1 \neq k_2, \\ \frac{\beta_2 - \beta_1}{\pi}, & k_1 = k_2. \end{cases} \quad (7)$$

Здесь Φ – исходное изображение, $\text{tr}ec$ – след матрицы.

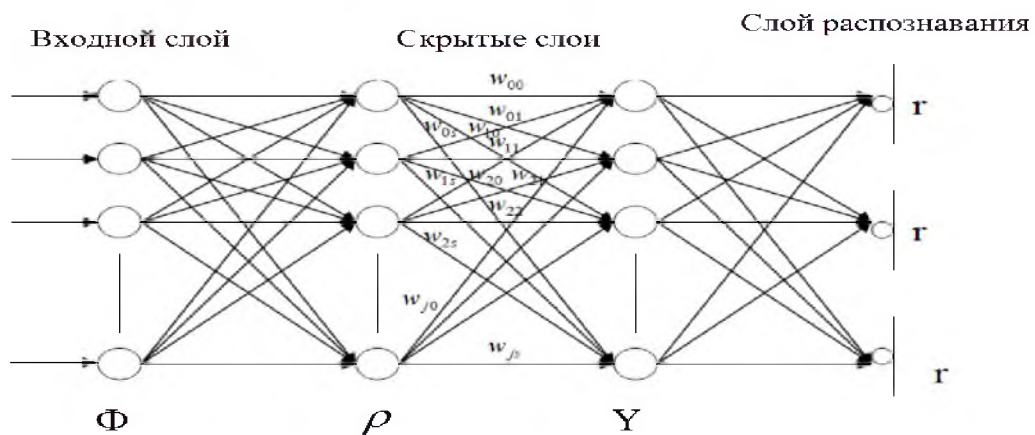
Возможность проведения анализа изображений на основе частотных представлений определяется тем, что в графических данных, зачастую, наблюдается квазипериодичность отображаемых процессов. На изображении могут присутствовать повторяющиеся объекты, которые задают некоторую периодичность изменения яркости изображения.

На первом шаге на изображении выделяется определенный объект изображения, который будет служить эталоном для распознавания других подобных объектов (рис. 2).



Рис. 2. Задание эталона для распознавания

На втором шаге строится нейронная сеть. Топология сети имеет следующий вид:





Здесь на входной слой подаются исходные данные (значения интенсивностей пикселей изображения). В скрытом слое вычисляются доли энергии, а на его выходе сумма произведений значений долей энергии и весовых коэффициентов. На слое распознавания вычисляется ошибка распознавания (относительная среднеквадратическая погрешность) относительно желаемого отклика.

Алгоритм обучения имеет следующий вид:

1. Поступление на входной слой исходных данных.
2. Инициализация весовых коэффициентов:

$$w_{sj} = \frac{1}{M * N}$$

3. Вычисление долей энергии на первом скрытом слое:

$$P_{rm}^{норм} = \frac{tr(A_r^T \cdot \Phi \cdot B_m \cdot \Phi^T)}{\sum_{i=1}^M \sum_{k=1}^N f_{ik}^2} \quad (8)$$

4. Получение одномерного вектора:

$$P_{rm}^{норм} \rightarrow \rho_j = \rho_{(r-1)R_1 + m + 1},$$

$$j = (r-1)R_1 + m + 1,$$

$$m = 1, \dots, R_2$$

5. while ($\varepsilon > r$) do:

- a) Поиск инвариантов с точностью допуска γ на втором скрытом слое:

$$Y^{(l)}_s = \sum_{j=1}^{R_1 * R_2} w^{(l)}_{js} \rho_j \quad (9)$$

- b) Вычисление относительной среднеквадратической погрешности:

$$\varepsilon^{(l)} = \sqrt{\frac{\sum_{s=1}^{R_1 * R_2} (Y^{(l)}_s - \rho_s)^2}{\sum_{s=1}^{R_1 * R_2} \rho_s^2}} \quad (10)$$

где ρ_s – желаемый отклик.

- v) Изменение весовых коэффициентов для второго скрытого слоя:

$$w_{sj}^{(l+1)} = w_{sj}^{(l)} + \mu * \varepsilon^{(l)} * \frac{\rho_j^{\min} * \rho_j}{\rho_j^{\max}} * \frac{Y_s}{Y_s^{\max}} \quad (11)$$

где $\mu = 0, \dots, 1$ – коэффициент скорости обучения.

- г) Конец цикла while.

6. Завершение обучения.

Результаты распознавания представлены на рис. 3.



Рис. 3. Результаты распознавания

Ошибки распознавания первого и второго рода:

$$Er_1 = \frac{N_{o1}}{N_{пол}} * 100\% = 0\%$$

$$Er_2 = \frac{N_{o2}}{N_{пол}} * 100\% = 0\%$$

где N_{o2} – количество объектов, ошибочно отнесенных к интересующему классу распознавания; $N_{пол}$ – количество объектов на снимке, которые относятся к классу распознавания; N_{o1} – количество объектов, принадлежащих к классу распознавания, но не отнесенных к нему.

Работа выполнена при поддержке ФЦП «Научные и научно-педагогические кадры инновационной России», на 2009-2013 годы, гос. контракт № 14.740.11.0390.

Литература

1. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2006. – 1072 с.
2. Грузман И.С. Цифровая обработка изображений в информационных системах. Учебное пособие. – Новосибирск, 2000. – 166 с.
3. Жилияков Е.Г. Вариационные методы анализа и построения функций по эмпирическим данным на основе частотных представлений. – Белгород: Изд-во БелГУ, 2007. – 160 с.
4. Жилияков Е.Г. Метод определения точных значений долей энергии изображений в заданных частотных интервалах / Е.Г. Жилияков, А.А. Черноморец, И.В. Лысенко // Вопросы радиоэлектроники. – Сер. РЛТ, 2007. – Вып. 4. – С. 115-123.
5. Жилияков Е.Г. Оптимальная фильтрация изображений на основе частотных представлений / Е. Г. Жилияков, А.А. Черноморец // Вопросы радиоэлектроники. Сер. ЭВТ. – 2008. – Вып.1. – С.118-132.



6. Ярославский Л.П. Введение в цифровую обработку изображений. – М.: Сов. радио, 1979. – 312 с.
7. Потапов А.С. Распознавание образов и машинное восприятие / А.С. Потапов. Спб: – Политехника, 2007. – 552 стр.

THE STRUCTURE OF NEURAL NETWORKS FOR OBJECT RECOGNITION IN THE AEROSPACE IMAGES BY ANALYZING THE DISTRIBUTION OF ENERGY OVER THE FREQUENCY RANGES

E.G. ZHYLYAKOV
A.U. LIKHOSHERSTNYI
V.V. KRASILNIKOV

*Belgorod National
Research University*

*e-mail: Zhilyakov@bsu.edu.ru
ozzy.osbourne.man@gmail.com*

In the article discusses a new method for recognition of objects or images the earth's surface by constructing a neural network, which calculated the proportion of energy allocated objects.

Key words: aerospace images, neural network, frequency representation, the proportion of energy pattern recognition.