УДК 681.007.05

А.В. Липанов, А.Ю. Михайлов

Харьковский национальный университет радиоэлектроники

АЛГОРИТМ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ ТЕКСТУР С ИСПОЛЬЗОВАНИЕМ МОМЕНТНЫХ ПРИЗНАКОВ И МЕТОДОВ НЕЙРОННЫХ СЕТЕЙ

В статье рассматривается алгоритм распознавания текстур, основанный на использовании моментных признаков, способа выбора участков изображения при извлечении признаков, применения методов системного анализа для принятия решений, параллельной нейросетевой архитектуры.

алгоритм распознавания текстур, моментные признаки, методы нейронных сетей

Введение

Среди задач распознавания изображений задачи распознавания текстур занимают особое место. При построении систем управления движением мобильными роботами, обработке изображений аэрофотосъемки, при неразрушающем контроле и при поиске изображений в базах данных возникает необходимость распознавания фрагментов изображения, т.е. обнаружения заведомо известных частей на изображении в том числе и текстур. Задача быстрого и эффективного распознавания изображений текстур является одной из самых сложных. Существуют различные подходы к решению этой задачи, однако они не совершенны, и поэтому поиск новых путей и подходов к решению задачи распознавания текстур на изображении остается актуальной проблемой.

Основной материал

Предлагаемый в статье подход определяет структуру системы распознавания текстур и алгоритм работы данной системы. При этом решаются задачи устранения перенасыщения нейронной сети при обучении и точной формальной интерпретации выходного сигнала сети в процессе классификации. Эффективное решение этих проблем является ключевым моментом для перехода к программной реализации системы на высокий качественный уровень. Во многих приложениях для интерпретации отклика нейронной сети используется аппарат функционального анализа – вычисление расстояния, представленного соответствующей нормой, например, вида

$$\| R_i \| = \sqrt{\sum_j (D(i,j) - O(i,j))^2}\ , \qquad (1)$$

где R_i – расстояние между ожидаемым и наблюдаемым откликами; $D(i,j)$ – j-й ожидаемый отклик i-й нейронной сети; $O(i,j)$ – j-й наблюдаемый отклик i-й нейронной сети.

Использование такого правила принятия решений является эффективным ввиду своей фундаментальности и простоты. При этом имеется возможность существенно расширить математический аппарат нейронных сетей посредством реализации специальной нейросетевой архитектуры, адаптированной к применению теории принятия решений из системного анализа.

К разработке описываемого алгоритма был применен подход распределения информационных потоков между несколькими нейронными сетями с целью ухода от известной проблемы перенасыщения нейронной сети при обучении нескольким различным классам входных сигналов.

Нейросетевая система включает в себя N нейронных сетей, каждая из которых выполняет задачу обработки некоторого одного класса входных сигналов (в данном случае – некоторого класса изображений текстур). Каждая из N нейронных сетей обладает P входами и M выходами (рис. 1). При этом M < P, что обусловлено ограничениями на производительность системы. Процесс обучения одному элементу из обучающей выборки выполняется только для нейронной сети, которая соответствует данному классу входных сигналов.

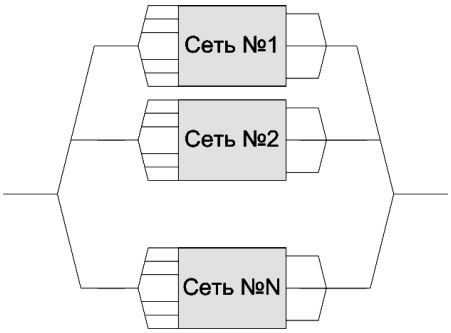


Рис. 1. Общая схема системы нейронной обработки

Для перехода к задаче системного анализа на основе полученных в процессе работы нейронных

© А.В. Липанов, А.Ю. Михайлов

сетей векторов выходных данных строится матрица принятия решений:

$$R=\begin{pmatrix} D(0,0)-O(0,0) & \dots & D(0,j)-O(0,j) & \dots & D(0,M)-O(0,M) \\ \dots & \dots & \dots & \dots & \dots \\ D(i,0)-O(i,0) & \dots & D(i,j)-O(i,j) & \dots & D(i,M)-O(i,M) \\ D(N,0)-O(N,0) & \dots & D(N,j)-O(N,j) & \dots & D(N,M)-O(N,M) \end{pmatrix} \quad (2)$$

где R – матрица расстояний между ожидаемым и наблюдаемым откликами; $D(i,j)$ – j-й ожидаемый отклик i-й нейронной сети; $O(i,j)$ – j-й наблюдаемый отклик i-й нейронной сети.

Удобным является использование матрицы рисков – матрицы абсолютных отклонений фактических значений, полученных в результате работы нейронной сети, от значений желаемого отклика. Путем поиска наиболее оптимального решения, представленного приемлемым с точки зрения выбранного правила принятия решения значением риска, легко можно произвести интерпретацию результатов обработки входного вектора нейросетевым каскадом. Таким образом, в случае обучения отдельных нейронных сетей каскада характеристикам определенных классов изображений, задача распознавания посредством нейросетевой обработки сводится к подзадаче принятия решения с минимальным риском на основе матрицы отклонений выходных сигналов нейронных сетей данного каскада.

На практике возможно применение любых методов принятия решений из системного анализа. Вопрос превосходства тех или иных методов в данной конкретной задаче является весьма важным и требует детального изучения в будущем. В построенной нами нейросетевой системе распознавания текстур качество распознавания, выражаемое в процентном соотношении правильно распознанных изображений к общему размеру выборки, существенно отличается в зависимости от выбора метода принятия решений. Опытным путем было установлено, что достаточная степень качества интерпретации отклика сети достигается при использовании критерия произведений вида

$$b_{ij} = \prod_{j=0}^{M} (D(i,j) - O(i,j)), \quad (3)$$

где b_{ij} – оценка по выбранному критерию принятия решений; $D(i,j)$ – j-й ожидаемый отклик i-й нейронной сети; $O(i,j)$ – j-й наблюдаемый отклик i-й нейронной сети.

При каждом проведенном эксперименте данный подход оказывался намного более эффективным, чем подход классификации изображений на основе сравнения норм выходных сигналов нейронной сети (являющийся на сегодняшний день наиболее распространенным), поскольку таким образом была решена задача предотвращения перенасыщения нейронной сети одновременно с задачей построения системы параллельной нейросетевой обработки.

Приводимый подход может быть использован для нейросетевой обработки различных сложных сигналов. Непосредственно же при решении задачи распознавания текстур немаловажной является подзадача специфичной реализации каждой отдельно взятой нейронной сети. На первом этапе работы системы проводится предварительная обработка изображения, в результате которой извлекаются признаки каждого отдельного фрагмента изображения. Затем эти признаки в определенном порядке передаются в нейронную сеть для обработки (обучения или распознавания). Перед подачей признаков на вход сети необходимо их обработать таким образом, чтобы этим поспособствовать наилучшей работе сети и ее обучению. Необходимым условием для корректного распознавания объектов является репрезентативность признаков объектов, используемых при работе системы.

В задаче распознавания изображений в качестве исходных данных при нейросетевой обработке могут, например, использоваться непосредственно значения яркости по базису RGB в каждой отдельно взятой точке изображения. Однако данное представление образа будет избыточным, т.е. будет содержать значительно больший объем информации, чем это фактически необходимо. С целью устранения избыточности нами применяется соответствующая процедура извлечения признаков из изображения. Среди часто используемых статистических признаков изображения можно назвать следующие: математическое ожидание яркостей по каждому цвету базиса, среднее линейное отклонение значений яркости по каждому цвету базиса RGB от математического ожидания или приближенное значение моды яркости. При использовании этих характеристик не выявляются отдельные детали изображений, и тем более не учитывается взаимное расположение точек изображения и отдельных фрагментов изображения, что приводит к потере большого объема информации, которая может использоваться для распознавания.

В качестве признаков в рассматриваемом алгоритме используются центральные моменты первого порядка [2] вида (4), вычисляемые отдельно по каждому каналу R, G и B:

$$\mu_{10} = \iint_D B_R(x,y)x\,dx\,dy\ ;\ \ \mu_{01} = \iint_D B_G(x,y)y\,dx\,dy\ ;$$

$$\mu_{11} = \iint_D B_B(x,y)xy\,dx\,dy\ , \quad (4)$$

где $B_R(x,y)$ – функция яркости канала красной составляющей; $B_G(x,y)$ – функция яркости канала

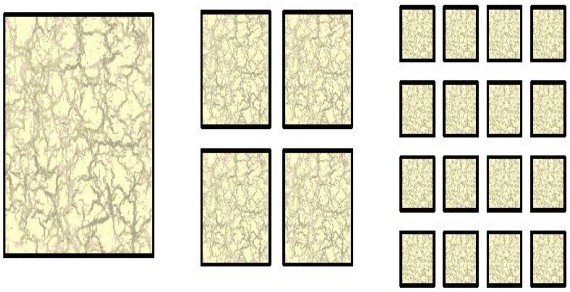
зеленой составляющей; $B_B(x, y)$ – функция яркости канала синей составляющей.

Численное вычисление интегралов в (4) производится с использованием метода, описанного в [3].

Важной особенностью данного алгоритма является то, что вычисление моментов осуществляется в определенных областях изображения. Данные области получаются путем разбиения (5) исходного изображения на ряд областей, как показано на рис. 2:

$$W_{i+1} = \{x_{i+1} \mid x_{i+1} \in W_{i+1} : x_i \bmod p = 0\};$$
$$H_{i+1} = \{y_{i+1} \mid y_{i+1} \in H_{i+1} : y_i \bmod p = 0\}, \qquad (5)$$

где p – шаг разбиения изображения; W, H – множества точек разбиения по осям X и Y соответственно; x, y – элементы множеств W и H.



а б в

Рис. 2. Схема разбиения изображения:
а исходная текстура; б – 1-й уровень разбиения;
в – 2-й уровень разбиения

В приведенном примере изображение разбито на 4 подобласти на первом уровне разбиения (рис. 2, б) и на рис. 1, 6 подобластей на втором уровне разбиения (рис. 2, в). Дальнейшее разбиение изображения будет приводить к существенному увеличению объемов вычислений и при этом – к снижению информативности признаков каждой из областей. Поэтому на втором уровне процесс разбиения изображения на области прекращается.

На каждом уровне разбиения в каждой из подобластей вычисляются моменты первого порядка по формуле (4) для каждого из цветовых каналов. В результате получаем вектор сигналов, который далее подается на вход нейронной сети для обучения или распознавания. Данный подход позволяет учесть глобальные признаки на первом уровне разбиения, а затем локальные признаки, характерные для каждой из областей, на втором уровне разбиения. Важной является возможность учета взаимного местоположения элементов изображения, объясняемая существованием привязки между подобластями изображения и синаптическими входами нейронной сети. Поскольку существует взаимно однозначное соответствие между номером подобласти изображения и номером синаптического входа (рис. 3), ассоциативная память нейронной сети автоматически учитывает взаимное расположение источников признаков. Будучи представленными отдельными независимыми модулями, нейронные сети могут использовать ту или иную архитектуру в зависимости от требуемого результата.

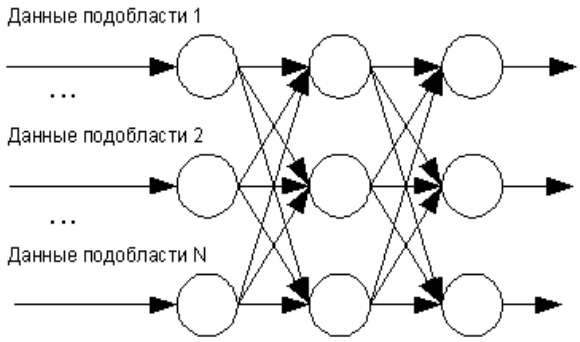


Рис. 3. Соответствие подобластей изображения
и синаптических входов

Это является важным фактором, дающим возможность изучения поведения различных нейронных сетей при решении задачи распознавания изображений, что обеспечивает возможность выбора оптимальной архитектуры каждого отдельного вычислительного элемента. В качестве вычислительного элемента использовалась полносвязная нейронная сеть с сигмоидальной функцией активации вида (6) со случайным начальным распределением синаптических весов в интервале [–0,5; 0,5]:

$$\mathrm{Sigm}(x, c) = \frac{1}{1 + e^{-cx}}, \qquad (6)$$

где x – точка вычисления функции; c – параметр, задающий крутизну сигмоиды.

Для обучения сети использовался градиентный метод оптимизации с шагом 0,01 и требуемой погрешностью 0,1. Данный выбор требуемой погрешности обусловлен необходимостью устранения перенасыщения нейронной сети при обучении большому количеству разнородных примеров. В целом, в зависимости от количества классов объектов, возможны различные пути выбора данного параметра. Использование полносвязной топологии нейронной сети обусловлено тем, что экспериментально была установлена ее повышенная надежность. Малозначимые синаптические связи при выбранной топологии отбрасываются посредством присвоения малозначимых весовых коэффициентов. Более сложные топологии нейронных сетей, как, например, БНС – топология (быстрые нейронные сети), являются вырожденным случаем полносвязной сети и, соответственно, могут быть получены путем процесса уменьшения весовых коэффициентов нейронов полносвязной сети.

Недостатком использования полносвязной сети является меньшая скорость обработки данных, которой можно пренебречь при проведении исследо-

ваний, а также с целью повышения надежности системы.

Процесс обучения отдельно взятому примеру выполняется в системе параллельно на каждом вычислительном элементе с использованием различных результирующих сигналов. В качестве входного сигнала нейронной сети используются признаки изображения, полученные посредством приведенной выше схемы. В качестве результирующего сигнала нейронной сети используется вектор величин в интервале [0, 1], характеризующий оценку вероятности принадлежности изображения к данному классу, представленному текущим вычислительным элементом.

Описанный алгоритм реализован программно с использованием Microsoft Visual C++ и технологии COM. В программной реализации системы предусмотрена возможность модификации шага разбиения изображения с целью изучения влияния способов разбиения изображения на процесс извлечения признаков. Программный комплекс реализован с учетом возможности использования различных топологий нейронных сетей в процессе работы системы. Архитектура системы распознавания текстур показана на рис. 4. Рассмотрим основные компоненты системы распознавания текстур.

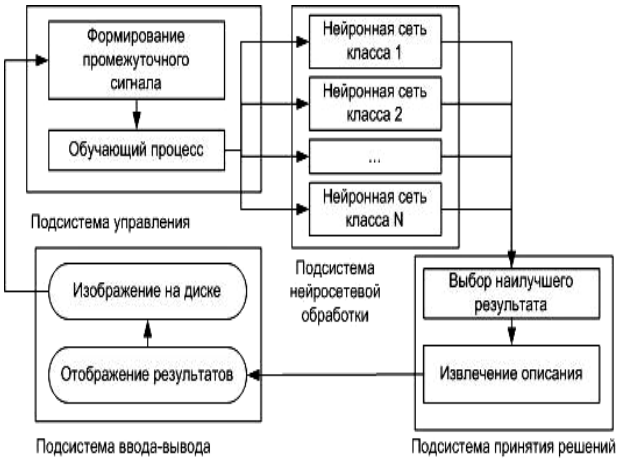


Рис. 4. Функциональная схема структурной организации системы распознавания образов

Подсистема ввода-вывода реализует хранение данных на информационных носителях и операции по файловому вводу-выводу, хранению и преобразованию изображений, отображению результатов работы пользователю, а также формирует набор входных сигналов, представляющих собой набор характеристик этого изображения, используемых затем при работе нейронной сети. Подсистема управления производит контроль над процессом обучения и формирование входных и эталонных сигналов. Подсистема нейросетевой обработки реализует все операции, относящиеся к нейросетевой обработке, включая обучение и расчет результатов обработки входного сигнала. Подсистема принятия решений осуществляет интерпретацию выходного сигнала сети в процессе распознавания и принимает решение по выбору наилучшего (наиболее вероятного с точки зрения сети) класса на основе результатов, полученных в ходе нейросетевой обработки.

Тестирование программной реализации разработанного алгоритма показало, что среднее время обучения системы одному эталонному изображению составляет 3 секунды, а среднее время распознавания составляет 0,5 секунды.

Для проведения тестирования системы была использована выборка изображений текстур, состоящая из 4000 полноцветных изображений и при этом количество правильно распознанных изображений текстур составило 93%. Эти результаты свидетельствуют о практической ценности алгоритма распознавания текстур с использованием моментных признаков и целесообразности продолжения исследований данного алгоритма.

Выводы

Разработанный алгоритм распознавания текстур является эффективным средством для решения задач распознавания изображений текстур, что достигается за счет использования: моментных признаков, способа выбора участков изображения при извлечении признаков, применения методов системного анализа для принятия решений, параллельной нейросетевой архитектуры.

Описанный алгоритм позволяет достичь следующих результатов: возможности эффективного использования наиболее простых нейросетевых архитектур, не дающих достаточного уровня точности при общепринятом подходе построения нейросистем; адаптации нейросетевой архитектуры для параллельной нейросетевой обработки, увеличивающей скорость реакции системы в режиме распознавания; упрощение формализации процесса принятия решений в нейросетевых системах.

Список литературы

1. Саймон Хайкин. Нейронные сети: полный курс, 2-е издание. – М.: Изд. дом «Вильямс», 2006. – 1104 с.

2. Путятин Е.П., Аверин С.И. Обработка изображений в робототехнике. – М.: Машиностроение, 1990. – 320 с.

3. Липанов А.В. Особенности применения численных методов при реализации алгоритмов слежения и нормализации // Системы обработки информации. – Х.: ХВУ, ХФВ «Транспорт України», 2001. – Вып. 1 (11). – С. 194-198.

Поступила в редколлегию 23.02.2007

Рецензент: д-р техн. наук, проф. В.А. Краснобаев, Харьковский национальный технический университет сельского хозяйства им. П. Василенко, Харьков.