

СЕГМЕНТАЦИЯ ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ РЕКУРРЕНТНОЙ НЕЙРОННОЙ СЕТИ

В.Б. Немировский, А.К. Стоянов

Томский политехнический университет
E-mail: stoj-ak@ad.cctpu.edu.ru; nemir-vb@ad.cctpu.edu.ru

Рассматривается возможность применения одномерного точечного отображения для сегментации изображений. В связи с сегментацией обсуждаются вопросы расчёта параметров нейронов рекуррентной сети, реализующей такое отображения. Приводятся результаты экспериментов по многошаговой сегментации.

Ключевые слова:

Изображение, пиксель, точечные отображения, неподвижные точки, сегментация.

Key words:

Image, pixel, point map, critical point, segmentation.

Введение

Для решения различных задач обработки изображений широко используется их сегментация. Под сегментацией понимается разбиение изображения на ряд областей со схожими характеристиками пикселей. Результатом является множество сегментов (областей), полностью покрывающих исходное изображение. Предполагается, что области соответствуют реальным объектам (или их частям).

Методы сегментации можно разделить на два класса: автоматические — не требующие взаимодействия с пользователем и интерактивные — изменяющие пользовательский ввод непосредственно в процессе работы [1, 2]. В статье рассматривается разработка метода сегментации, относящегося к первому классу. Методы этого класса делятся на две категории [1–3]: методы, базирующиеся на информации либо о контурах, либо об областях.

Методы первой категории используют для сегментации априорную информацию о контурах объектов [2], которая всегда есть в задачах анализа сцен, поиска объектов на изображении и некоторых других задачах. Методы этой группы узко специализированы для каждой конкретной задачи и не являются объектом нашего интереса.

Методы второй категории делят изображение на области в соответствии с совпадением некоторых характеристик пикселей изображения (интенсивности, цветовых каналов, текстуры и т. д.) [3, 4]. При этом априорная информация о свойствах областей отсутствует, но на само разбиение изображения накладываются некоторые условия (например, однородность областей по яркости или цвету). Методы сегментации, относящиеся к этой категории, универсальны и применимы к любым изображениям. Существует большое количество методов сегментации, использующих разные подходы к формированию однородных областей [1]. Одним из распространенных подходов является применение для этих целей кластеризации (различные модификации метода k -средних [2, 5]). Отметим, что основной проблемой применения метода k -средних является необходимость наличия априорной

информации для выбора числа кластеров [2, 4, 5], что ограничивает широкое использование кластеризации для сегментации изображений.

Таким образом, существует потребность в разработке методов сегментации изображений, основанных на кластеризации, и свободных от указанного выше недостатка. Таким решением этой задачи может быть использование кластеризации, осуществляемой нейронной сетью. Использование для кластеризации рекуррентной нейронной сети [6] не требует (в отличие от k -средних) никакой предварительной информации о числе кластеров. Кроме того, достоинствами любой искусственной нейронной сети является распределенная параллельная обработка информации, способность к самонастройке и подавлению шумов.

Целью данной работы является исследование возможности применения кластеризации, выполняемой рекуррентной нейронной сетью, и разработка на основе этого исследования метода сегментации полутоновых изображений.

Кластеризация значений на одном нейроне

Для кластеризации множества M -мерных данных рекуррентной нейронной сетью в состав входного слоя сети включают M нейронов. Каждый нейрон обрабатывает все значения какого-то одного из M признаков, характеризующих вектор данных из исходного множества [6]. Отдельное значение одного такого признака является входным сигналом для нейрона. В рассматриваемой нами рекуррентной сети для каждого входного нейрона вводится локальная обратная связь, соединяющая выход нейрона с его входом. Эта обратная связь реализует процесс одномерного отображения величины x входного сигнала на активационной функции нейрона, которая описывается выражением

$$f(x) = \frac{\mu}{1 + \exp(-\alpha x + \beta)}, \quad (1)$$

где α — коэффициент наклона; $\alpha > 0$, β — величина смещения; μ — коэффициент увеличения, $\mu > 0$. Одномерное отображение величины x входного сигнала на активационной функции описывается ите-

рациональным соотношением $x_{n+1}=f(x_n)$, где $n=1,2,3,\dots$ — номер итерации. После нескольких итераций отображаемая величина сигнала достигает устойчивой неподвижной точки, значение которой x^* определяется из нелинейного уравнения $f(x^*)=x^*$. В численных расчётах значение x^* определяется с некоторой заданной точностью ε .

Динамика работы такого нейрона проанализирована в [6], и мы не будем касаться подробностей. Отметим лишь те из деталей анализа, которые проясняют механизм кластеризации для некоторого множества M -мерных данных.

Для кластеризации параметры α и β выбирают так, чтобы диапазон величин всех входных сигналов $[x_{\min}, x_{\max}]$, отображаемых на одном нейроне, располагался между значениями неустойчивой точки отображения x_1^* и одной из двух устойчивых точек x_2^* . Процесс отображения разбивает диапазон величин входных сигналов на ряд интервалов. Интервалы обладают следующими свойствами: величины, принадлежащие одному интервалу, в результате отображения сходятся к устойчивой точке x_2^* за одно и то же число итераций. Таким образом, величины входных сигналов группируются по степени их близости. Мерой близости является количество итераций, нужное для достижения устойчивой точки с заданной степенью точности ε . Точки, попавшие в один интервал, относятся, согласно выбранной мере, к одному кластеру. Таким образом, каждый из входных нейронов кластеризует множество значений одного из M признаков исходных данных. Вся же сеть в целом выделяет кластеры во всех признаках, что и даёт возможность кластеризации элементов множества M -мерных данных.

В рассматриваемой сети возможен расчёт параметров активационной функции отдельного нейрона. Он основан на том, что при отображении любого значения на некоторой функции неподвижные (устойчивые и неустойчивые) точки лежат на пересечении графика функции и биссектрисы прямого угла [7]. В нашем случае такой функцией является сигмоида, которая описывается ур. (1), и её параметры α и β можно получить, решая систему уравнений

$$x_i^* = f(x_i^*, \alpha, \beta, \mu), \quad i = 1, 2. \quad (2)$$

Значения x_i^* связаны со значениями границ диапазона $[x_{\min}, x_{\max}]$ соотношениями $x_1^* = x_{\min} - \delta_1$ и $x_2^* = x_{\max} + \delta_2$, где $|\delta_1| < 1$ и $|\delta_2| < 1$. Система уравнений (2) решается относительно α и β лишь в случае, если коэффициент увеличения μ задан заранее. В предположении известности μ решение системы (2) имеет вид:

$$\alpha = \frac{1}{x_2^* - x_1^*} \ln \left(\frac{(\mu - x_1^*)x_2^*}{(\mu - x_2^*)x_1^*} \right),$$

$$\beta = \alpha x_1^* + \ln \left(\frac{\mu - x_1^*}{x_1^*} \right). \quad (3)$$

Очевидно, что любое изменение μ управляет значениями α и β , и, следовательно, меняет форму

активационной функции, оставляя неизменными положения неустойчивой и устойчивой точек отображения. В свою очередь, параметры α и β определяют условия сходимости отображения входных значений, что, в конечном счете, и создаёт возможность управления процессом кластеризации данных.

Экспериментально было установлено [6], что наилучшая кластеризация входных значений одного нейрона наступает при значении μ , которое находится в окрестности $\max|\Delta H|$, где ΔH — изменение энтропии, соответствующее изменению μ . Энтропия отображения H подсчитывается для данного μ по формуле Шеннона

$$H = - \sum_i p_i \log_2 p_i.$$

Здесь $p_i = N_i/N$; N_i — количество значений входного сигнала, входящих в i -й кластер (т. е. множество значений соответствующего признака, сходящихся к устойчивой точке за одно и то же число итераций); N — общее количество всех значений сигнала. Суммирование ведётся по количеству кластеров, образовавшихся при отображении на активационной функции с параметрами α и β , соответствующими текущему значению μ . Смысл критерия выбора μ становится понятным, если вспомнить, что энтропия даёт нам среднее значение информации, приходящееся на один элемент изучаемой системы (в нашем случае таким элементом является любой заполненный из образовавшихся кластеров). Следовательно, $\max|\Delta H|$ показывает наибольший прирост средней информации на один кластер, получаемый за счёт изменения значения μ . Поскольку энтропия H связана со степенью хаотичности системы, то этот прирост обусловлен переходом системы из менее упорядоченного состояния в более упорядоченное. Зная значения μ , соответствующие такому переходу, можно выбрать то из них, которое приводит к большей упорядоченности отображения, то есть лучшему (при данном выборе параметров α и β) набору кластеров.

Многошаговая сегментация

Будем считать, что исходное изображение представлено в полутоновой серой шкале. Выделение однородных по яркости областей (сегментация) в этом случае фактически сводится к кластеризации значений всего одной характеристики (признака) — яркости. Это означает, что обработка всего изображения может осуществляться в рекуррентной сети с одним входным нейроном ($M=1$). В то же время обычное представление изображения — двумерная матрица. Такое представление естественным образом предполагает использование сети со многими входами ($M>1$) для распараллеливания процесса обработки. Применение такой рекуррентной сети для кластеризации в нашем случае возможно, если модифицировать сеть, рассмотренную в [6], таким образом, чтобы у всех нейронов входного слоя была бы одна и та же ак-

тивационная функция с одинаковыми значениями параметров. Тогда обработка всех яркостей изображения входным слоем сети математически будет эквивалентна их обработке на одном нейроне, что можно использовать при программном моделировании сети. Параметры α и β активационной функции этого нейрона определяются по методике, описанной в предыдущем разделе. Границам диапазона значений входных сигналов нейрона $[x_{\min}, x_{\max}]$ соответствует минимальная и максимальная яркости пикселей изображения. На основе этих граничных значений по формулам (3) вычисляются α и β .

Сам процесс сегментации, рассматриваемый в этой статье, требует некоторых дополнительных пояснений, приведённых ниже.

Пусть μ меняется дискретно от μ_0 до некоторого $\mu_{\text{конеч}}$. Из ур. (3) ясно, что $\mu_0 > x_2^* > x_1^*$. Каждому значению $\mu_i \in [\mu_0, \mu_{\text{конеч}}]$, $i=1, 2, \dots$ при отображении соответствует своя последовательность итераций, разбивающая диапазон яркостей изображения на ряд интервалов с одинаковыми свойствами (т. е. кластеров). В зависимости от распределения значений яркостей, часть этих интервалов заполнится, а часть останется пустыми. В целом, всей совокупности значений μ_i соответствует некоторое конечное множество разбиений диапазона.

Процесс кластеризации предполагает нахождение оптимального по некоторому критерию значения $\mu_{\text{опт}}$, которое обеспечивает наилучшее выделение кластеров. Пусть таким критерием является $\max[\Delta H]$ энтропии распределения, получаемого в результате отображения, разбивающего диапазон яркостей на ряд кластеров. Тогда полученное разбиение обеспечивает максимальную упорядоченность распределения яркостей изображения для данного $\mu_{\text{опт}}$ по всем выделенным кластерам. Выполним теперь для каждого заполненного кластера следующую процедуру — усредним его яркости и полученным средним заменим все значения, входящие в кластер. Энтропия распределения не изменится, а результатом будет сегментированное изображение, поскольку в нём множество разных исходных яркостей пикселей разделено на несколько подмножеств, объединяющих в себе близкие яркости пикселей.

Повторим описанный выше процесс кластеризации для нового распределения яркостей после усреднения. Для этого распределения нужно искать новое $\mu_{\text{опт}}$. Действительно, для некоторого μ_i в повторном процессе разбиения может быть кластер, включающий в себя, по крайней мере, два или более значения яркости нового распределения. Это означает, что возможно объединение яркостей для этого кластера, т. е. дальнейшее упорядочивание распределения яркостей.

На рис. 1, а, схематично представлено распределение значений яркости изображения для $\mu_{\text{опт}}$ до их усреднения в пределах кластера (верхняя часть рис. 1, а) и после усреднения с последующей заменой всех текущих значений кластера (нижняя часть рис. 1, а). Энтропия в обоих случаях неизменна и равна 1,89 бита. На рис. 1, б, представлено распределение для некоторого значения $\mu_i \neq \mu_{\text{опт}}$ при повторном процессе. Энтропия, соответственно, равна 1,24 бита. Иными словами, новое распределение более упорядочено.

Таким образом, используя последовательно в роли исходного изображения результат предыдущей его сегментации, можно получить ряд сегментированных изображений с последовательно снижающейся энтропией. Назовём такой процесс многошаговой сегментацией изображения. При этом энтропия изображения может служить критерием качества сегментации.

При реализации многошагового процесса возникает вопрос о его сходимости к устойчивому сегментированному изображению (энтропия которого не меняется). Ответ на этот вопрос дали экспериментальные исследования.

Экспериментальные результаты

Для проверки возможности получения сегментации в многошаговом процессе были проведены экспериментальные исследования на основе созданной программы, реализующей описанный выше алгоритм. Исследования проводились на полутоновых изображениях (серая шкала).

Яркосты X_i i -го пикселя изображения приводились к диапазону $[\gamma_1, \gamma_1 + \gamma_2]$ использованием линейной нормировки (числа γ_1 и γ_2 лежат в единичном интервале; яркости изображения $X_{\min}$ и $X_{\max}$ не равны):

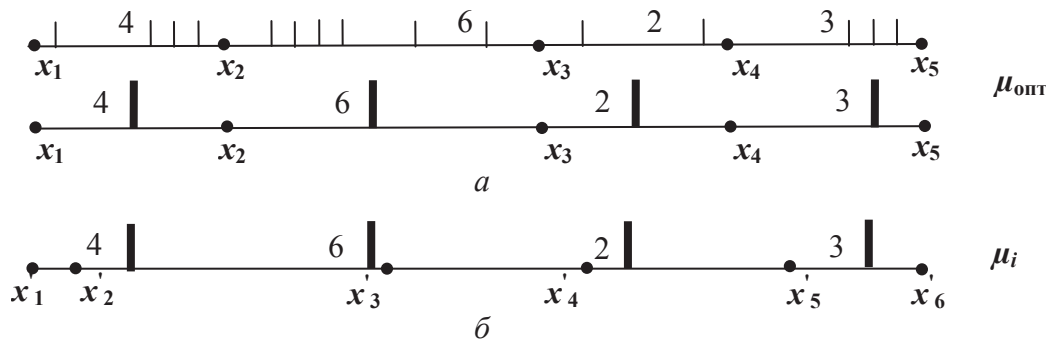


Рис. 1. Распределение значений яркостей по кластерам для разных μ . Цифрами указано количество значений, попавших в интервал $[x_k, x_{k+1}]$, $k=1, 2, \dots, 5$

$$x_i = \gamma_2 \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} + \gamma_1.$$

Здесь x_i – приведённое значение яркости пикселя. В эксперименте для нормировки использовались значения $\gamma_1=0,025$; $\gamma_2=0,20$ в одном варианте и $\gamma_1=0,01$; $\gamma_2=0,20$ – в другом. Затем в изображении отыскивались минимальное и максимальное значения нормированных яркостей пикселей, которые использовались для определения параметров α и β . Значения α и β определялись для заданного μ в соответствии с (3). Значение μ последовательно задавалось в диапазоне $x_{\max} + h < \mu \leq 1$ с постоянным шагом h . Для каждого μ подсчитывалась энтропия H распределения яркостей изображения. Таким образом, получалась последовательность значений энтропии, зависящая от величины μ , в которой затем определялся $\max|\Delta H|$. Величина μ_0 , соответствующая найденному максимуму, использовалась в дальнейшем для определения границ кластеров в яркостях пикселей изображения.

В соответствии с описанными выше особенностями процесса сегментации далее яркости пикселей изображения группировались по выделенным

кластерам, после чего близкие яркости, попавшие в один кластер, заменялись средним значением всех яркостей по этому кластеру. Таким образом, уменьшалось количество градаций яркости по сравнению с их числом в исходном изображении. Преобразованная шкала градаций использовалась для формирования матрицы пикселей нового изображения. Описанная последовательность действий выполнялась до тех пор, пока энтропия для нового изображения, полученного на очередном текущем шаге процесса, не становилась равной энтропии изображения, исходного для этого шага.

На рис. 2 и 3 приведены результаты, иллюстрирующие процесс многошаговой сегментации для разных изображений и условий нормировки яркостей пикселей.

Хорошо видно, что результаты отдельных шагов сегментации дают дополнительную информацию об объектах обрабатываемого изображения, позволяя, в зависимости от задачи, изучить структуру этих объектов с различной степенью детализации. Вид сегментированного изображения определяется тем шагом процесса, на котором оно было получено, и слабо зависит от условий нормировки ярко-

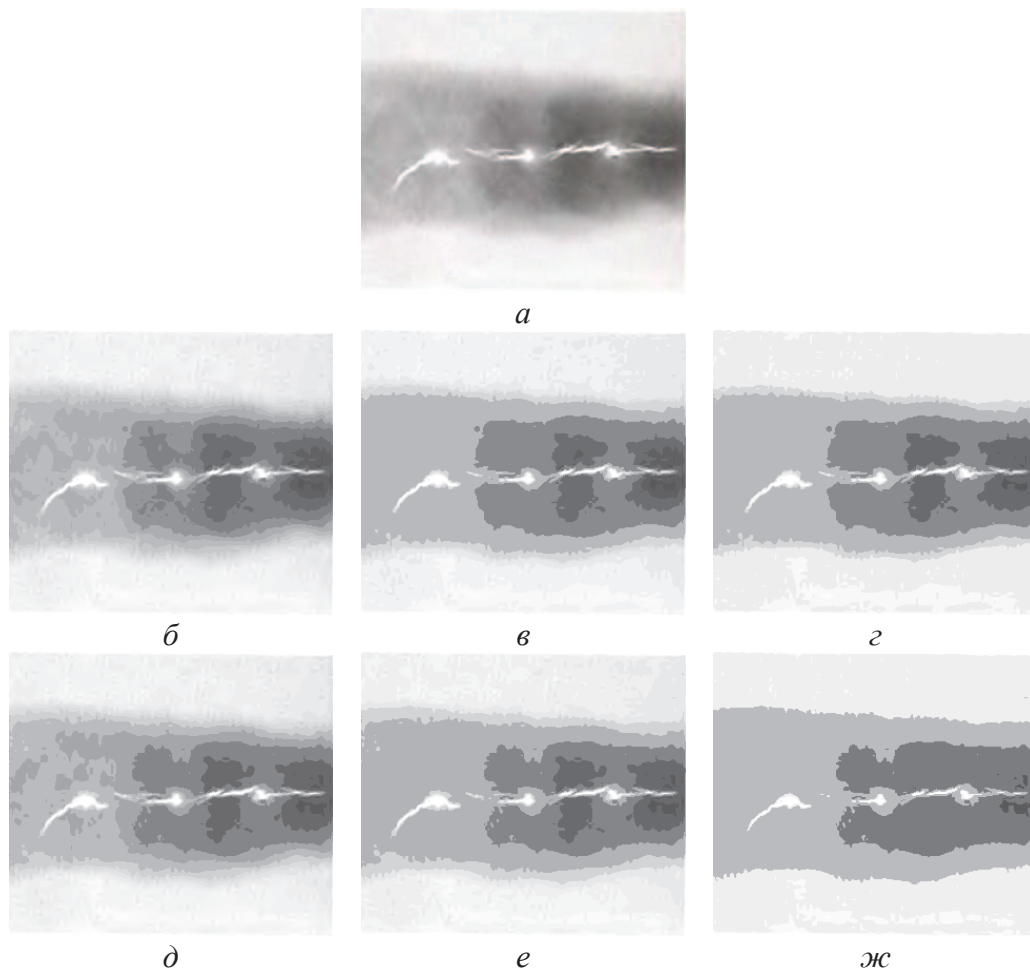


Рис. 2. Многошаговая сегментация рентгенограммы сварного шва: а) исходное изображение, взятое из [1]; б) шаг 1, $H=4,31$ бита; в) шаг 4, $H=2,92$ бита; г) шаг 6, $H=2,41$ бита; ($\gamma_1=0,200$, $\gamma_2=0,025$) д) шаг 1, $H=3,76$ бита; е) шаг 3, $H=2,92$ бита; ж) шаг 7, $H=1,86$ бита ($\gamma_1=0,20$, $\gamma_2=0,01$)

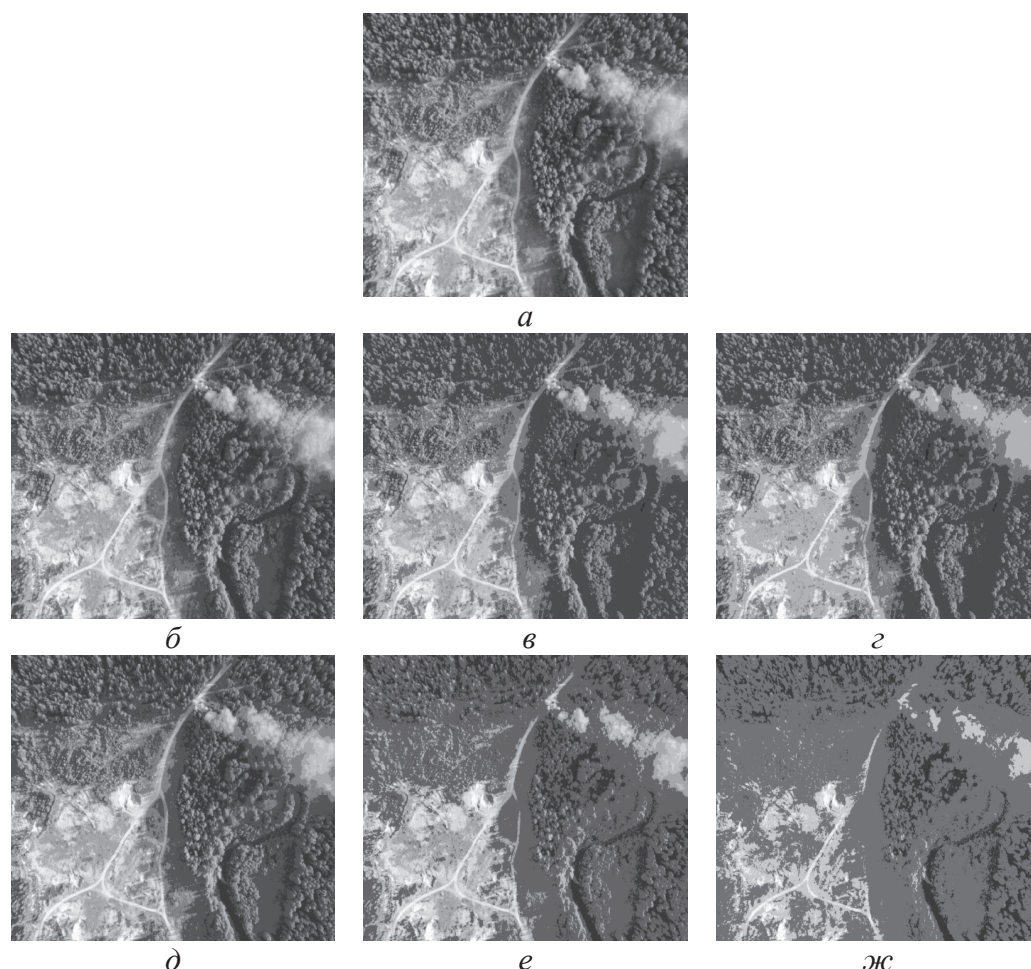


Рис. 3. Многошаговая сегментация аэрофотоснимка лесного пожара: а) исходное изображение; б) шаг 1, $H=3,05$ бита; в) шаг 2, $H=2,10$ бита; г) шаг 3, $H=1,87$ бита; ($\gamma_1=0,200$, $\gamma_2=0,025$) д) шаг 1, $H=2,53$ бита; е) шаг 2, $H=1,87$ бита; ж) шаг 3, $H=1,20$ бита ($\gamma_1=0,20$, $\gamma_2=0,01$)

стей исходного изображения. В этом легко убедиться, сравнивая изображения с одинаковой энтропией (рис. 2, в и е).

Многошаговая сегментация завершается при условии равенства энтропий изображений предыдущего и текущего шагов процесса. Шаг завершения сегментации в эксперименте был равен 8 для рентгенограммы и 6 для аэрофотоснимка. Причём на последних шагах сегментированные изображения в каждом из рассмотренных случаев визуально почти не различались, несмотря на разные значения их энтропии.

Таким образом, многошаговая сегментация может быть использована для изучения с различной степенью детализации структуры объекта, представленного на исходном изображении.

Выводы

1. Рекуррентная нейронная сеть, основанная на одномерных отображениях, позволяет выделить в изображении кластеры близких значений яркости, что даёт возможность сегментировать изображение. Используя многошаговый

процесс сегментации, можно получить сегментированное изображение с энтропией, минимально возможной для данного распределения исходных яркостей.

2. Экспериментально показано уменьшение энтропии распределения яркостей на каждом шаге сегментации и сходимости многошаговой сегментации за конечное число шагов при четко выраженном выделении сегментов изображения. Результат сегментации зависит только от параметров линейной нормировки диапазона яркостей изображения на единичный диапазон.
3. Обработанные изображения одного и того же объекта, обладающие равной энтропией при разных параметрах одномерного отображения, визуально слабо различаются.
4. Установлено, что параметры нейрона входного слоя, обеспечивающие отображение входных значений яркости и успешную их сегментацию, можно рассчитать.

Работа выполнена в рамках государственного задания «Наука».

СПИСОК ЛИТЕРАТУРЫ

1. Гонсалес Р., Вудс Р. Цифровая обработка изображений. — М.: Техносфера, 2005. — 1072 с.
2. Skarbek W., Koschan A. Colour image segmentation — a survey // Techn. Rep. 94–32. — Technical University Berlin. 1994. URL: <http://iristown.engr.utk.edu/~koschan/paper/coseg.pdf> (дата обращения: 09.04.2012).
3. Malik J., Belongie S., Leung T., Shi J. Contour and texture analysis for image segmentation // International Journal of Computer Vision. — 2001. — V. 43. — № 1. — P. 7–27.
4. Форсайт Д., Поппе Д. Компьютерное зрение. Современный подход. — М.: Вильямс, 2004. — 928 с.
5. Lucchese L., Mitra S. Color image segmentation: A state-of-the-art survey // Proc. of the Indian National Science Academy (INSA-A). — 2001. — V. 67. — № 2. — P. 207–221.
6. Стоянов А.К. Применение рекуррентной нейронной сети для решения задачи кластеризации // Известия Томского политехнического университета. — 2009. — Т. 315. — № 5. — С. 144–149.
7. Малинецкий Г.Г. Математические основы синергетики. — М.: Изд-во ЛКИ, 2007. — 312 с.

Поступила 11.04.2012 г.

УДК 004.652.8:004.825

ДИАГНОСТИКА МОДЕЛЕЙ БАЗ ДАННЫХ НА ОСНОВЕ ПРОДУКЦИОННОЙ СИСТЕМЫ

А.А. Мицель, В.Д. Сибилев, И.А. Красиков

Томский государственный университет систем управления и радиоэлектроники

E-mail: Krasikov86@sibmail.com

Представлена логическая структура и механизм работы проектируемой автоматизированной системы проверки моделей баз данных. Приведена формализация базы знаний продукционной системы искусственного интеллекта, используемая в синтаксическом анализе моделей баз данных. Рассмотрены примеры логического вывода на основе знаний, представленных в формализованной базе знаний продукционной системы искусственного интеллекта.

Ключевые слова:

База данных, синтаксические ошибки, семантические ошибки, логическая модель, метаданные, модель «сущность–связь», продукционная система, база знаний, правило-продукция.

Key words:

Database, syntax errors, semantic errors, logic model, metadata, model «entity–relationship», production system, knowledge base, rule-production.

Введение

Диагностика моделей баз данных на сегодняшний день является одной из важнейших задач, стоящих перед проектировщиком баз данных (БД), от выполнения которой зависит будущее всей разрабатываемой информационной системы. Процесс проектирования модели БД представляет собой логически сложную, трудоемкую задачу, требующую от проектировщика знания семантического моделирования БД и навыков в работе с инструментальными средствами построения моделей БД, называемых CASE-системами (Computer Aided Software/System Engineering).

В настоящее время специалистов в области проектирования БД готовят практически во всех вузах и иных учебных заведениях нашей страны, однако уровень знаний выпускаемых специалистов в данной области зачастую оставляет желать лучшего. Низкий уровень знаний проектировщиков БД проявляется в отсутствии навыка выделять объекты описываемой предметной области (ПО), формулировать законы взаимодействия объектов ПО друг с другом, строить на их основе модели БД, которые в конечном итоге сводятся к физическим БД. В таких моделях БД зачастую наблюдаются

различного рода аномалии обновления данных, их избыточность, потеря данных, либо в БД могут иметь место данные, не несущие никакой смысловой нагрузки. Перечисленные ошибки относятся к так называемым ошибкам нормализации БД, которые можно не допускать на этапе построения модели БД, владея методологией построения моделей БД и соблюдая применяемые к ним правила нормализации. В проектировании БД принят стандарт IDEF1X, который широко используется, начиная с 1993 г. и по сегодняшний день [1]. В стандарте IDEF1X четко и однозначно описаны все синтаксические правила построения моделей БД, на которых базируется большинство CASE-систем, однако знания методологии IDEF1X для успешного построения схем БД недостаточно, проектировщику необходимо также выявлять все функциональные зависимости (ФЗ) сущностей, вытекающие из делового регламента описываемой ПО. Правила нормализации БД, подробно описанные в книге [2], могли бы помочь студентам избежать большинства ошибок еще на этапе проектирования модели БД.

Данная статья является продолжением работ [3, 4], в которых: