



УДК 681.300

Е. Ю. Бурсиан

ПОСТРОЕНИЕ БАЗ ДАННЫХ ЭТАЛОННЫХ СИМВОЛОВ ПРИ АВТОМАТИЧЕСКОМ РАСПОЗНАВАНИИ ТЕСТОВ

Дата поступления: 27.07.2015

Решение о публикации: 08.09.2015

Цель: Исследовать процесс автоматического распознавания тестов, применяемых для самопроверки и контроля знаний студентов; определить и повысить эффективность алгоритмов обработки отсканированного изображения рукописных текстов, содержащих специальные знаки и таблицы, в тестах студентов. **Методы:** Используются методы построения скелетных представлений символов и их сравнение с эталонами. При этом базы данных скелетных представлений эталонных символов разделяются на общие и частные, дополняемыми новыми эталонами с помощью методов теории графов. Базы данных структурированы с использованием особенностей задачи. **Результаты:** В предлагаемой модели для каждого испытуемого и каждого символа построен набор скелетных представлений. Вычислена матрица подобия, элементы которой равны коэффициенту подобия между скелетными представлениями. На основе матрицы подобия построен нагруженный граф, вершины которого соответствуют распознаваемым символам, определены центры графа, радиус и степени вершин. В качестве частной базы данных скелетных графов выбрано доминирующее множество, содержащее вершины наибольшей степени и центры графа, построенного на основе матрицы подобия. В результате частные базы данных не перегружаются, так как содержат только центральные элементы. Для проверки предлагаемой модели разработана экспериментальная программа автоматической проверки тестов. **Практическая значимость:** Обращение к построенным с помощью методов теории графов частным базам данных образцов скелетных графов символов позволит повысить вероятность распознавания символов теста, а применение метрических характеристик графов в качестве параметров частных баз данных – структурировать их и оптимизировать по объёму.

Моделирование, процесс распознавания, скелетный граф, вектор информативных характеристик.

Yelena Yu. Bursian, Cand. Sci. (Eng.), associate professor, bursianeu@mail.ru (Petersburg State Transport University) BUILDING REFERENCE SYMBOL DATABASES IN AUTOMATIC RECOGNITION OF TESTS

Objective: To study the process of automatic recognition of tests used for self-examination and control of students' knowledge; to determine and to increase efficiency of algorithms for processing scanned images of hand-written texts which contain special signs and tables in students' tests. **Methods:** Methods for building skeletal representation of symbols and their comparison with reference symbols were used. In the process, databases of skeletal representations of reference symbols are divided into general and specific

ones, with the latter databases augmented by new reference symbols with the use of graph theory methods. Databases are structured in accordance with specific tasks. **Results:** In the model proposed, for each person undergoing test and for each symbol a set of skeletal representations is built. Similarity matrix is calculated, elements of which equate the coefficient of similarity between skeletal representations. Loaded graph is built on the basis of similarity matrix, with graph points corresponding to recognised symbols; graph centres, range and degrees of points established. Dominating set which includes graph points of the highest degree and centres of a graph built on the basis of similarity matrix was selected as a specific database. As a result, specific databases are not overloaded as they only contain central elements. An experimental automatic test check program was developed to check the proposed model. **Practical importance:** Application to specific databases of skeletal graphs of exemplary symbols built with help of graph theory methods will allow to increase probability of recognising test symbols, and the application of metric characterisation of graphs as parameters of specific databases will allow to structure them and to optimise their size.

Simulation, recognition process, skeletal graph, informative characteristics vector.

На современном этапе развития образования применяют разные формы контроля деятельности студентов, в частности, широко используется тест как текущая и периодическая форма проверки знаний.

Для самоконтроля и ускорения проверки тестов необходима система компьютерной поддержки обработки отсканированных тестов. Составляющей такой системы является комплекс программ автоматического распознавания отдельных рукописных символов и знаков.

Разработки комплексов программ распознавания отдельных рукописных символов основаны на методах общей теории распознавания, наиболее известны российские исследования в этой области Ю. И. Журавлёва, В. В. Рязанова, О. В. Сенько, Е. В. Дюковой [4]. К прикладным исследованиям относятся работы Л. М. Местецкого, Ю. В. Визильтера, Я. А. Фурмана [2, 5]. Среди иностранных исследований по анализу и обработке изображений, адаптивному бустингу известны работы R. C. Gonzalez, T. Y. Zhang, C. Y. Suen, H. Blum, R. O. Duda, P. E. Hart, P. Viola, M. Jones, S. Rosset [3, 7, 8].

Созданы комплексы программ распознавания отдельных рукописных символов для изображений хорошего качества, в частности, ABBYY FormReader, CuneiForm, Omnipage, ReadirisPro. При этом их применение для

распознавания тестов не всегда удобно из-за специфики учебных дисциплин и различных методов составления тестов.

При использовании данных пакетов часто возникают затруднения при автоматической настройке на индивидуальные особенности почерка тестируемых и пишущего инструмента. Таким образом, актуальна разработка программ автоматического чтения отдельных рукописных символов и специальных знаков, пригодных для автоматической проверки тестов студентов.

В данной работе рассматривается модель процесса распознавания символов, основанная на представлении символов в виде скелетных графов и их сравнении с эталонными графами [3].

Постановка задачи

Требуется построить модель процесса автоматического распознавания изображений отдельных рукописных символов и специальных знаков, применимую в условиях многократного тестирования групп испытуемых.

На основе построенной модели необходимо разработать алгоритм распознавания изображения тестов, использующий допол-

нительную информацию об индивидуальных особенностях написания знаков, полученную в ходе предыдущих обращений к системе автоматического тестирования.

Далее необходимо разработать экспериментальный комплекс программ автоматической обработки тестов. При этом предполагается, что указанный программный комплекс позволит обрабатывать изображения нечётко прорисованных знаков при многократном применении к документам с повторяющимися особенностями.

Для хранения и использования информации, полученной в ходе предыдущих обращений, надо создать инструменты систематического обновления дополнительных (частных) баз данных.

Модель процесса распознавания отдельных рукописных символов теста в условиях многократного тестирования групп испытуемых

На предварительном этапе обработки отсканированного изображения удаляют шумы и помехи с последующим созданием соответствующего бинарного изображения и с выделением на нём дискретных фигур [5].

Символы предварительно классифицируют, сравнивая скелетные графы распознаваемых дискретных фигур с элементами стандартной базы эталонных графов символов [3].

На следующем этапе строят частную базу эталонных графов, предполагая, что распознаваемый тест заполнен одним почерком. Частные базы данных эталонных графов систематически обновляют и сохраняют для повторного использования.

С помощью частной базы данных вычисляют дополнительные параметры, используемые на завершающем этапе в алгоритме сравнения скелетных графов распознаваемых областей с эталонными графами. Ведение частных баз данных эталонных графов позволяет учитывать характеристики пишущего инструмента

и индивидуальные особенности изображения символов тестируемым.

Построение дополнительных баз данных нагруженных графов для эталонных символов

Одним из основных параметров, содержащих информацию об индивидуальных особенностях написания символов, являются скелетные представления образцов символов (нагруженные графы эталонных символов). При построении системы скелетных представлений образцов символов применяются методы теории графов.

Множество ветвей каждого скелетного графа эталонного символа разбивается на два подмножества: основных и второстепенных ветвей (рис. 1).

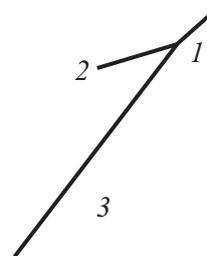


Рис. 1. Скелетный граф эталонного символа

На рис. 1 подмножество основных ветвей — $\{2, 3\}$, второстепенных — $\{1\}$. Отдельно для основных и для первоначальных скелетных представлений вычисляются топологические характеристики графов эталонных символов: $k(G)$ — число компонент связности графа G , $\nu(G) = p(G) - b(G) + k(G)$ — цикломатическое число ($p(G)$ — число вершин, $b(G)$ — число ребер), число ветвей, степени вершин.

Предполагается, что на этапе распознавания гомеоморфизм между скелетными графами распознаваемых и эталонных символов устанавливается отдельно для основных и для первоначальных скелетных представлений.

При вычислении векторов информативных характеристик нагруженные графы эталон-

ных символов приводят к единому прямоугольнику расположения. Множество ветвей скелетных графов эталонных символов упорядочивают по координатам узлов.

С каждым эталонным символом сопоставляют вектор информативных характеристик – последовательность значений кодированных углов наклонов ребер на ветвях скелетных графов, взятых через равные промежутки по длине ребер. Таким образом, для каждой ветви скелетного графа применяют цепное кодирование (вариант естественного уравнения кривой для дискретного изображения) [3].

Определение параметров множества модификаций изображения для каждого символа дополнительной базы данных

Для определения структуры множества модификаций, соответствующих различным изображениям символа, применим методы теории графов.

Множество модификаций изображения символа предлагаем представлять в виде нагруженного орграфа $G_S(V, E)$, в котором каждой вершине $v \in V$ соответствует модификация изображения символа S . Дуга $e(i, j) \in E$, где $i, j \in \{1, \dots, N\}$, N – число различных модификаций символа S , если в процессе построения частной базы данных будет установлено подобие между различными модификациями $v(i)$ и $v(j)$.

Для установления подобия для каждой модификации $v(i)$ вычислим вектор информативных характеристик – набор значений закодированных углов наклонов ветвей скелетных графов, взятых через равные расстояния на соответствующих ветвях:

$$\bar{\varphi}(i) = (\varphi_1(i), \dots, \varphi_n(i));$$

$$\varphi_k(i) \in \{0, 1, \dots, m-1\},$$

где $k \in \{1, \dots, n\}$.

В качестве вектора информативных характеристик рассматривается также вектор

коэффициентов дискретного преобразования Фурье углов наклонов ветвей скелетных графов:

$$c_l(j) = \sum_{k=0}^{n-1} \varphi_k(j) \exp\left(\frac{2\pi ikl}{n}\right),$$

где $l \in \{1, \dots, n\}$.

Для каждого графа $G_S(V, E)$ вычислим матрицу подобия $A = [a_{ij}]$. Дуга $e(i, j)$ не принадлежит E ($a_{ij} = 0$), если не существует гомеоморфизма между скелетным графом, соответствующим модификации $v(i)$ и подграфом, соответствующим модификации $v(j)$, при котором расстояния между соответствующими узлами не превышают порогового значения. При этом рассматриваются только подграфы скелетных графов, в которых последовательно удаляются второстепенные ветви.

В качестве меры подобия $v(i)$ и $v(j)$ выберем цепную взаимно-корреляционную функцию $C(i, j)$:

$$C(i, j) = \frac{1}{n} \sum_{k=1}^n \overline{\varphi_k(i) \varphi_k(j)};$$

$$\overline{\varphi_k(i) \varphi_k(j)} = \cos(\varphi_k(i) - \varphi_k(j)).$$

В альтернативном варианте как мера подобия рассматривается корреляционная функция $r(i, j)$:

$$r(i, j) = \frac{\sum_{k=1}^n \varphi_k(i) \varphi_k(j) - \varphi_0(i) \varphi_0(j)}{\sqrt{\sum_{k=1}^n (\varphi_k(i) - \varphi_0(i))^2 \sum_{k=1}^n (\varphi_k(j) - \varphi_0(j))^2}}.$$

Матрица подобия вычисляется по формуле

$$a_{ij} = \begin{cases} 1, & C(i, j) \geq C_0, (r(i, j) \geq r_0); \\ 0, & C(i, j) < C_0, (r(i, j) < r_0). \end{cases}$$

Константы C_0 (соответственно r_0) – пороговое значение цепной взаимно-корреляционной

функции (корреляционной функции). Дуга $e(i, j)$ принадлежит E при $a_{ij} = 1$, $e(i, j)$ не принадлежит E при $a_{ij} = 0$.

На следующем этапе предлагается дополнить множество дуг графа $G_S(V, E)$ с помощью евклидова расстояния между векторами информативных характеристик соответствующих вершин либо с помощью расстояния Ли [1]. Каждой дуге $e(i, j)$ приписывается метрическая характеристика – расстояние между векторами информативных характеристик модификаций $v(i)$ и $v(j)$:

$$d(i, j) = d_L(\bar{\varphi}_i, \bar{\varphi}_j).$$

Множество E дополняется дугой $e(i, j)$, если соответствующее ей расстояние Ли либо евклидово расстояние меньше порогового значения:

$$d(i, j) < d_0.$$

Расстояние Ли и нормированное расстояние Ли определим с помощью веса Ли:

$$W_L(\bar{\varphi}) = \sum_{k=1}^n |\varphi_k|,$$

где

$$|\varphi| = \begin{cases} \varphi, & \varphi \leq \frac{m}{2} \\ m - \varphi & \varphi > \frac{m}{2} \end{cases}.$$

$$d_L(\bar{\varphi}_i, \bar{\varphi}_j) = W_L(\bar{\varphi}_i - \bar{\varphi}_j);$$

$$d_L^0(\bar{\varphi}_i, \bar{\varphi}_j) = \frac{W_L(\bar{\varphi}_i - \bar{\varphi}_j)}{n}.$$

Для описания структуры множества модификаций, относящихся к различным изображениям эталонного символа, и построения частных баз данных эталонных символов вычислим метрические характеристики неорграфа $G_S^0(V, E)$, соотнесённого с оргграфом

$G_S(V, E)$: центры, диаметр и радиус связных компонент графа. При этом также определим вершины с наибольшей степенью, внешне устойчивые множества вершин и ядра неорграфа $G_S^0(V, E)$ [1].

В множестве вершин неорграфа $G_S^0(V, E)$ выберем подмножество V_0 , вершины которого соответствуют эталонным символам, сравниваемым с распознаваемыми символами. При этом на первом этапе рассматривается множество $V_0 = \{v_1\}$, где v_1 – вершина наибольшей степени. Если вершин наибольшей степени несколько, выбирается вершина, наиболее близкая к одному из центров связных компонент неорграфа $G_S^0(V, E)$.

На следующем этапе в множество V_0 добавляется вершина наибольшей степени неорграфа $G_S^0(V - v_1, E)$, т.е. подграфа, из которого исключены вершины, включённые в V_0 и т.д., пока множество V_0 не станет являться доминирующим множеством в графе $G_S^0(V, E)$.

При построении множества V_0 предпочтительнее отдаётся вершинам, наиболее близким к одному из центров связных компонент неорграфа $G_S^0(V, E)$ и несмежным с вершинами, уже включёнными в множество V_0 . Таким образом, V_0 в наилучшем случае – одно из ядер неорграфа $G_S^0(V, E)$. В остальных случаях V_0 будет внешне устойчивым (доминирующим) множеством, не являясь при этом ядром.

При распознавании символов в качестве эталонных изображений символов берем те, которые соответствуют вершинам множества V_0 в неорграфе $G_S^0(V, E)$. При этом неорграф $G_S^0(V, E)$ пополняется новыми вершинами за счёт распознаваемых символов и используется в дальнейшем для статистической обработки.

Метрические характеристики неорграфа $G_S^0(V, E)$ вычисляются с помощью евклидова расстояния либо расстояния Ли. Пороговое значение p_0 в алгоритме распознавания не превышает радиуса неорграфа $p(G_S^0(V, E))$.

Если евклидово расстояние либо расстояние Ли между векторами информативных характеристик распознаваемого символа и

эталонного символа S меньше порогового значения p_0 , распознаваемый символ идентифицируется как символ S :

$$d(\bar{\varphi}, \bar{\varphi}_S) < p_0. \quad (1)$$

При этом предполагается, что расстояния Ли между векторами информативных характеристик различных эталонных символов в случае установления гомеоморфизма между скелетными графами превышают диаметры связанных компонент неорграфа $d(G_S^0(V, E))$.

Если множество вершин в неорграфе $G_S^0(V, E)$ достаточно велико для проведения статистического анализа выборки векторов информативных характеристик, то применение формулы (1) в качестве решающего правила при распознавании символов обосновывается с помощью результатов статистического анализа.

Если есть основания предполагать, что углы, соответствующие закодированным значениям, нормально распределены и ковариационную матрицу векторов информативных характеристик можно оценить, то разделяющие функции могут быть записаны в явном виде.

В качестве разделяющих функций, согласно Байесовской теории принятия решений [3], можно взять логарифм плотности нормального распределения, умноженной на вероятность появления символа S :

$$g(\bar{\varphi}_S) = -\frac{1}{2}(\bar{\varphi} - \bar{\varphi}_S)^T \Sigma_S^{-1}(\bar{\varphi} - \bar{\varphi}_S) - \frac{n}{2} \ln 2\pi - \frac{1}{2} \ln |\Sigma_S| + \ln P(S),$$

где $P(S)$ – априорная вероятность появления символа S ; $|\Sigma_S|$ – определитель ковариационной матрицы символа.

Если вероятности появления различных символов равны, что возникает при распознавании ответов в виде целых чисел, и для гомеоморфных графов величины логарифмов модулей ковариационных матриц близки, то в качестве решающего правила применяется сравнение расстояний махаланобиса:

$$r_L(\bar{\varphi}, \bar{\varphi}_S) < r_0;$$

$$r_L(\bar{\varphi}, \bar{\varphi}_S) = (\bar{\varphi} - \bar{\varphi}_S)^T \Sigma_S^{-1}(\bar{\varphi} - \bar{\varphi}_S).$$

Если предположить, что компоненты векторов информативных характеристик имеют одинаковые дисперсии и при сравнении выбираются статистически независимо, разделяющие функции могут быть приведены к виду

$$g(\bar{\varphi}_S) = (\bar{\varphi} - \bar{\varphi}_S)^T (\bar{\varphi} - \bar{\varphi}_S) = d_L^2(\bar{\varphi}, \bar{\varphi}_S).$$

Вычисление параметров множества модификаций изображения символов при построении частной базы данных

В эксперименте рассматривали модель автоматического распознавания, в которой характеристиками изображения символа считались разности между текущими и предыдущими ненулевыми изменениями углов наклона ветвей скелетных графов:

$$\bar{\varphi} = (\varphi_1, \dots, \Delta\varphi_v, \dots, \Delta\varphi_n);$$

$$\Delta\varphi_i = \begin{cases} \varphi_i - \varphi_{i-1}, & \varphi_i - \varphi_{i-1} \neq 0; \\ \Delta\varphi_{i-1}, & \varphi_i - \varphi_{i-1} = 0. \end{cases}$$

В качестве порогового значения, при котором дуга считалась не принадлежащей орграфу $G_S(V, E)$ множества модификаций символа S , рассматривалось значение коэффициента корреляции между векторами углов наклона ветвей скелетных графов, равное $r_0 = 0,9$.

На рис. 2 представлены различные модификации изображения символа, полученные при тестировании одного из испытуемых, на рис. 3 – соответствующие скелетные графы.

В табл. 1, 2 представлены значения коэффициентов корреляции и нормированных расстояний Ли между векторами информативных характеристик для рассматриваемых модификаций символа.



Рис. 2. Модификации изображения символа



Рис. 3. Скелетные графы

По полученным значениям коэффициентов корреляции построен неорграф $G_S^0(V, E)$. $p(G_S^0(V, E)) = 0,446$, $d(G_S^0(V, E)) = 0,519$; пороговым значением, при котором изображение символа считается изображением символа S , принимается значение $p_0 = 0,25$. В частную базу данных эталонных скелетных графов для символа S вносятся нагруженные графы 7 и 3, соответствующие центрам неорграфа

$G_S^0(V, E)$, $V_0 = \{3, 7\}$ – ядро $G_S^0(V, E)$. При хорошем почерке испытуемого предлагаемый метод позволяет повысить вероятность распознавания символа до 0,95.

Заключение

Тестирование студентов необходимо проводить как на компьютере, так и на бумажном носителе информации. При этом индивидуальные особенности изображения символов испытуемыми во многих случаях затрудняют автоматическое распознавание и проверку работ непосредственно после выполнения.

В процессе учебной деятельности, как правило, один и тот же испытуемый многократно обращается к тестированию, таким образом, имеется возможность собрать необходимые

ТАБЛИЦА 1. Значения коэффициентов корреляции между информативными характеристиками различных модификаций символа

№	1	2	3	4	5	6	7
1	1	0,95	0,90	0,89	0,93	0,89	0,92
2	0,95	1	0,90	0,92	0,92	0,89	0,95
3	0,90	0,90	1	0,86	0,90	0,92	0,88
4	0,89	0,92	0,86	1	0,93	0,88	0,93
5	0,93	0,92	0,90	0,93	1	0,90	0,95
6	0,89	0,89	0,92	0,88	0,90	1	0,90
7	0,92	0,95	0,88	0,93	0,95	0,90	1

ТАБЛИЦА 2. Значения нормированных расстояний Ли между информативными характеристиками различных модификаций символа

№	1	2	3	4	5	6	7
1	0	0,205	0,258	0,286	0,231	0,282	0,254
2	0,205	0	0,279	0,263	0,265	0,289	0,213
3	0,258	0,279	0	0,279	0,244	0,211	0,276
4	0,286	0,263	0,279	0	0,240	0,307	0,238
5	0,231	0,265	0,244	0,240	0	0,258	0,202
6	0,282	0,289	0,211	0,307	0,258	0	0,265
7	0,254	0,213	0,276	0,238	0,202	0,265	0

данные и построить частные базы данных скелетных графов символов.

Экспериментальные вычисления показывают, что введение дополнительных параметров и обращение к частным базам данных образцов скелетных графов символов позволяет повысить вероятность распознавания символов теста. Применение метрических характеристик графов: центров, диаметров и радиусов графов как параметров частных баз данных – позволяет их структурировать и оптимизировать по объёму.

Библиографический список

1. Берж К. Теория графов и её применения / К. Берж. – М. : Иноиздат, 1962. – 316 с.
2. Визильтер Ю. В. Построение морфологических спектров полутоновых изображений / Ю. В. Визильтер, С. В. Сидякин // Вестн. компьютерных и информационных технологий. – 2012. – № 4. – С. 8–17.
3. Дуда Р. О. Распознавание образов и анализ сцен / Р. О. Дуда, П. Е. Харт. – М. : Мир, 1976. – 507 с.
4. Журавлев Ю. И. Распознавание. Математические методы. Программная система. Практические применения / Ю. И. Журавлев, В. В. Рязанов, О. В. Сенько. – М. : Физматлит, 2005. – 159 с.
5. Местецкий Л. М. Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры / Л. М. Местецкий. – М. : Физматлит, 2009. – 288 с.
6. Питерсон У. Коды, исправляющие ошибки / У. Питерсон, Э. Уэлдон. – М. : Мир, 1976. – 593 с.
7. Rosset S. Boosting as a Regularized Path to a Maximum Margin Classifier / S. Rosset, J. Zhu, T. Hastie // J. Mach. Learn. Res. – 2004. – N 5. – P. 941–973.

8. Viola P. Robust real-time face detection / P. Viola, M. Jones // Int. J. Comput. Vision. – 2004. – Vol. 57, N 2. – P. 137–154.

References

1. Berzh K. Teoriya grafov i yeyo primeneniya [Graph Theory and Its Applications]. Moscow, Inoizdat, 1962. 316 p.
2. Vizilter Yu. V. & Sidiyakin S. V. *Vestnik kompyuternykh i informatsionnykh tekhnologiy – Computer and Information Technologies Newsletter*, 2012, no. 4, pp. 8-17.
3. Duda R. O. & Khart P. Ye. *Raspoznavaniye obrazov i analiz stsen* [Image Recognition and Scene Analysis]. Moscow, Mir, 1976. 507 p.
4. Zhuravlev Yu. I., Ryazanov V. V. & Senko O. V. *Raspoznavaniye. Matematicheskiye metody. Programmnaya sistema. Prakticheskiye primeneniya* [Recognition. Mathematical Methods. Program System. Practical Applications]. Moscow, Fizmatlit, 2005. 159 p.
5. Mestetskoy L. M. *Nepreryvnaya morfologiya binarnykh izobrazheniy: figury, skelety, tsirkulyary* [Continuous Morphology of Binary Images: Figures, Skeletons, Circulars]. Moscow, Fizmatlit, 2009. 288 p.
6. Piterson U. & Ueldon E. *Kody, ispravlyayushchiye oshibki* [Error-Correcting Codes]. Moscow, Mir, 1976. 593 p.
7. Rosset S., Zhu J. & Hastie T. Boosting as a Regularized Path to a Maximum Margin Classifier. *J. Mach. Learn. Res.*, 2004, no. 5, pp. 941-973.
8. Viola P. & Jones M. Robust real-time face detection. *Int. J. Comput. Vision.*, 2004, Vol. 57, no. 2, pp. 137-154.

БУРСИАН Елена Юрьевна – канд. техн. наук, доцент, bursianeu@mail.ru (Петербургский государственный университет путей сообщения Императора Александра I).