

УДК 004.81+004.93

Е.Е. Пятикоп¹, А.А. Каргин²¹ ГВУЗ «Приазовский государственный технический университет», Мариуполь² Донецкий национальный государственный университет, Донецк

ПОКОМПОНЕНТНОЕ РАСПОЗНАВАНИЕ СИМВОЛОВ ТЕКСТА НА ОСНОВЕ НЕЧЕТКОГО ПРЕДСТАВЛЕНИЯ

В статье рассматривается этап идентификации символа на основе сопоставления полученного набора ключевых узлов с описанием прототипа из базы данных. Образование узлов основывается на нечетких моделях клеток первичной зрительной коры. Вводятся математические описания модели прототипа символа, модели символа и нечеткой степени похожести. Приводятся результаты экспериментов.

Ключевые слова: обработка изображения, зрительная система, распознавание, символ, узлы, модель, нечеткая степень похожести.

Введение

Одним из направлений в области создания искусственных систем обработки и анализа зрительной информации является распознавание символов. На сегодняшний день уже имеется опыт разработки методов распознавания печатных символов с применением различного вида классификаторов, нейронных сетей [1 – 2]. Но объем цифровой визуальной информации увеличивается, меняются условия распознавания, и от современных систем требуется более глубокий интеллектуальный анализ. Так, одной из актуальных задач является борьба с «графическим спамом» [3, 4], в рамках которой необходимо выполнять анализ изображения с целью обнаружения рекламного текста. Изобретательность создателей спама требует совершенствование технологий распознавания [5]. Поэтому для обработки изображения исследуется возможность применения знаний когнитивной психологии [6].

В когнитивной психологии на основе наблюдений за человеком за длительный период предложены и апробированы экспериментально модели организации практически всех когнитивных функций человека: восприятие зрительной, звуковой информации, организация памяти разных уровней [7 – 8]. Так в [8] описывается, что в основе идентификации символов лежит задача распознавания паттернов, которая решается с помощью поддетального анализа. Способ рассмотрения знакомых объектов как конфигурацию простых компонентов в своей работе предложил Д. Марр [9]. Также известна теория Бидермана, названная «Распознавание по компонентам» (по прототипу). Основное положение теории заключается в том, что любой объект может быть разложен на ряд элементарных составляющих – геометрических модулей, или компонентов, называемых геонами, комбинируя которые можно получить объект практически любой формы [8]. Геоны

(компоненты) рассматриваются как комбинации простейших деталей. Так, деталями алфавита можно считать различные виды линий (вертикальные, горизонтальные, диагональные), а геонами (компонентами) – ключевые признаки букв – узлы. Итак, заглавная буква «А» может рассматриваться как состоящая из двух линий – под углом 30° и горизонтальной линии, которые в результате пересечения образуют 3 характерных узла. Простейшие детали очень похожи на результаты работы клеток первичной зрительной коры.

При этом схемы комбинаций деталей позволяют определить связи между деталями, которые более важны, и отбросить менее существенные. Так, все буквы, показанные на рис. 1, являются буквой А.



Рис. 1. Примеры написания буквы «А»

Таким образом, поддетальный анализ включает в себя такие процессы:

- распознавание простейших деталей, которые составляют геон (узел);
- распознавание ключевых узлов;
- идентификация на основе сопоставления полученного набора ключевых узлов с описанием прототипа из базы данных.

Первый этап – представление объекта на уровне простейших деталей с помощью моделей клеток первичной зрительной коры (ПКЗ) рассмотрен в работе [10]. Второй этап описан в работе [11]. В настоящей статье рассматривается третий этап – идентификация символов на основе сопоставления полученного набора ключевых узлов с описанием прототипа из базы данных.

Постановка задачи

Результатом работы первого этапа является нечеткое представление изображения в виде

$$M_{\omega}^{(3)} = \langle \omega, \{C^{\alpha}, \forall \alpha \in [0^{\circ}, 10^{\circ}, 20^{\circ} \dots 170^{\circ}]\rangle,$$

где $C^{\alpha} = \left\{ \{G_j^{\alpha, \omega, \ell_i}\}_{\ell_i} \right\}$ – ориентационная колонка;

$G^{\alpha, \omega, \ell}$ – модель клетки первичной зрительной коры, которая реагирует на линейные сегменты шириной ω , ориентированные определенным образом и определенной длины ℓ ; α – угол ориентации. Модель этой клетки есть выражение

$$G^{\alpha, \omega, \ell} = \langle K^{\alpha, r}, \alpha, \omega, \ell, \delta(G^{\alpha, \omega, \ell}) \rangle, \quad (1)$$

где $K^{\alpha, r} = \langle k^{\alpha, r+2Rk}, k^{\alpha, r+4Rk} \dots k^{\alpha, r+\ell^* \cdot 2Rk} \rangle$ – совокупность моделей смежных активных ганглиозных клеток сетчатки, на основе которой формируется детектор $D^{\alpha} = \langle K^{\alpha, r}, \alpha, \omega, \ell^*, \delta(D^{\alpha, \omega}) \rangle$; $\delta(D^{\alpha, \omega})$ – функция уверенности наличия детектора. Для выражения (1) $\delta(G^{\alpha, \omega, \ell}) = \delta(D^{\alpha, \omega})$.

На втором этапе получены модели прототипа узла UP_j (2), модели узла U_i (3) и введена степень нечеткой похожести $\xi(UP_j, U_i)$.

Модель узла-прототипа есть выражение

$$UP = \langle \{G', G''\}, Kd, \beta \rangle, \quad (2)$$

где $G', G'' \in \{G^{\alpha, \ell}\}$, $\beta = |\alpha_{G'} - \alpha_{G''}|$ – угол пересечения между моделями гиперсложных клеток, формирующими узел; $Kd = \langle \{k_z\}, \{\eta(K_z, O_G)\} \rangle$ – характеристика пересечения моделей гиперсложных клеток, где $k_z \in G'$ и $k_z \in G''$.

$$U = \langle \{G', G''\}, Kd, Kd', \beta \rangle, \quad (3)$$

где $Kd' = \langle \{k_z\}, \{\eta(k_z, O_G)\} \rangle$, $k_z \notin G'$ и $k_z \in G''$, но $k_z \in O_{G'}$.

Таким образом, после определения набора узлов, представленных во фрагменте изображения символа, решается задача распознавания символа, составленного из этих узлов. Для решения этой задачи необходимо иметь описание классов (прототипов), а далее выполнить сопоставление набора узлов входного изображения с набором узлов каждого из прототипов. Опишем модели символа-прототипа, символа и степень их похожести.

Модель символа-прототипа

Строится на наборе узлов-прототипов и описывается выражением:

$$LP = \langle U^{LP}, Gd \rangle, \quad (4)$$

где $U^{LP} = \{UP_j\}$ определяется выражением (2), а Gd – сложная структура, описанная формулой

$$Gd = \langle \{G_k\}, \{\rho(G_k, UP_j)\} \rangle, \quad (5)$$

где $\{G_k\}$ – множество моделей гиперсложных клеток, которые являются общими для некоторых узлов множества UP ; $\{\rho(G_k, UP_j)\}$ – множество функций принадлежности модели гиперсложной клетки G_k к каждому из типов узлов $\{UP_j\}$ $j \in 1 \dots 9$.

Рассмотрим пример описания символа-прототипа «М». Дано множество узлов $U = \{U_i\}$, $i \in 1 \dots 3$:

$$U_1 = \langle \{G^{100 \pm 30^{\circ}}, G^{60 \pm 20^{\circ}}\}, Kd = \{1, 0, 0, 1, 0, 0\}, 30^{\circ} \pm 20^{\circ} \rangle,$$

$$U_2 = \langle \{G^{60 \pm 20^{\circ}}, G^{120 \pm 20^{\circ}}\}, Kd = \{0, 0, 1, 0, 0, 1\}, 60^{\circ} \pm 40^{\circ} \rangle,$$

$$U_3 = \langle \{G^{80 \pm 30^{\circ}}, G^{120 \pm 20^{\circ}}\}, Kd = \{1, 0, 0, 0, 0, 1\}, 30^{\circ} \pm 20^{\circ} \rangle.$$

Множество общих моделей клеток $Gd = \langle \{G_k\}, \{\rho(G_k, UP_j)\} \rangle$:

$$Gd_1 = \langle G^{120 \pm 20^{\circ}}, \rho(G, UP_j) = \{1, 0, 0, 0, 1, 0, 0, 0, 0\} \rangle,$$

$$Gd_2 = \langle G^{60 \pm 20^{\circ}}, \rho(G, UP_j) = \{0, 1, 0, 0, 1, 0, 0, 0, 0\} \rangle.$$

На рис. 2 приводится графическое описание символа-прототипа М, где черные круги обозначают узлы, а цифра рядом соответствует типу; линии соответствуют простейшим деталям (моделям гиперсложных клеток), при этом двойная линия обозначает, что модель клетки участвует в двух узлах.

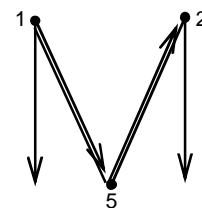


Рис. 2. Графическое представление прототипа буквы М

Модели символов-прототипов сформированы на основе изучения строения символов алфавитов, опроса экспертов и проведения компьютерных экспериментов.

Модель символа

По аналогии модели символа-прототипа LP можно описать модель символа L выражением

$$L = \langle U^L, Gd \rangle, \quad (6)$$

где $U^L = \langle \{U_i\}, \{\xi(UP_j, U_i)\} \rangle$ – двойка сложной структуры.

В выражении (6) Gd – сложная структура, описанная выражением (5), в котором

$$\rho(G_k^{U_i}, U_i) = \xi(UP_j, U_i).$$

Для изображения символа на рис. 2 описание модели будет следующим:

$$U_{\xi} = \langle U_1 = \langle \{G^{90^{\circ}}, G^{70^{\circ}}\}, Kd = \{1, 0, 0, 1, 0, 0\}, 20^{\circ} \rangle,$$

второго столбца (модели символа «Е») степень похожести с прототипом символа «Е» составляет 0,83. Следующее по убыванию значение степени похожести (0,67) относится к прототипу символа «Б», описание которого близко к прототипу символа «Е» по количеству общих узлов и общих клеток. Далее значения еще больше уменьшаются, поскольку прототипы этих символов имеют все меньше общих узлов с моделью исходного изображения.

Вторая группа экспериментов проводилась с целью идентификации символа на множестве идентичных символов с искажениями. Пример исходных данных для эксперимента показан на рис. 5. Величина степени схожести показана на рис. 5 под каждой из букв.

Таблица 1
Результаты экспериментов

Прототипы символов	Исходные символы						
	Е`	Н`	П`	Г`	Ф`	Б`	Ь`
1	2	3	4	5	6	7	8
Е	0,83	0,17	0,17	0,13	0,43	0,35	0,27
Н	0,17	1,00	0,00	0,00	0,19	0,08	0,10
П	0,08	0,00	0,88	0,19	0,13	0,08	0,00
Г	0,17	0,00	0,50	0,78	0,25	0,17	0,00
Ф	0,50	0,25	0,25	0,19	0,64	0,23	0,10
Б	0,67	0,10	0,07	0,08	0,34	0,63	0,48
Ь	0,50	0,08	0,00	0,00	0,06	0,52	0,60

Результаты данного эксперимента иллюстрируют, что незначительные искажения символа позволяют его верно идентифицировать с достаточно высокой степенью схожести. На рис. 5, а - г искажения в пределах отклонений заложенных в прототипе для этой буквы, а на рис. 5, д - искажения за пределами допустимых норм, поэтому степень сравнения значительно падает. Только изменение структуры символа (появление лишних узлов, как показано на рис. 5, е или потеря необходимых узлов, как показано на рис. 5, и значительно повышает ошибки идентификации.

Следующая группа экспериментов направлена на подтверждение факта, что нечеткая модель описания символа является шрифтонезависимой и не требует масштабируемости исходного символа под прототип.

Другими словами, необходимо было проверить, действительно ли буквы различных размеров, начертаний и шрифтов идентифицируются с высокой степенью схожести. Для этого эксперимента были выбраны изображения, показанные на рис. 6.

Числовые величины степени схожести показаны на рис. 6 под каждой из букв. Видно, что самая низкая степень схожести для символа, приближенного к рукописному.

Также были проведены эксперименты идентификации символа с учетом контекста. В условиях контекста предлагаемый метод идентификации символов дает существенное улучшение результатов (символ вне контекста получает степень схожести 0,57, а похожие на него – 0,53), в то же время с учетом контекста его степень схожести возрастает до 0,68.

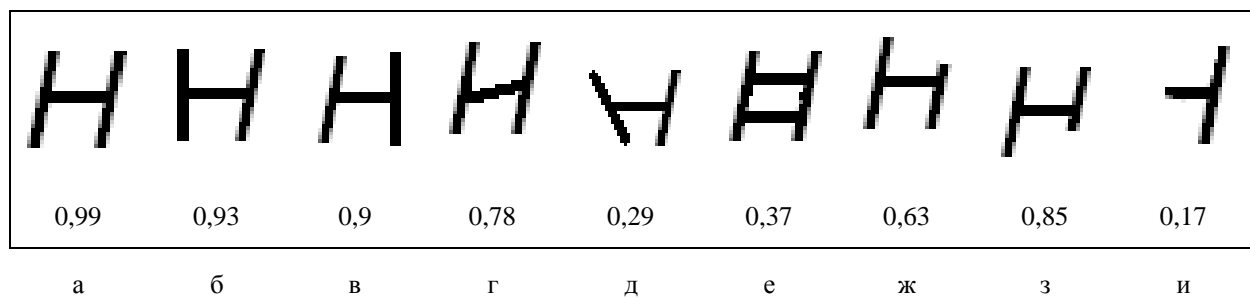


Рис. 5. Символ «Н» с искажениями

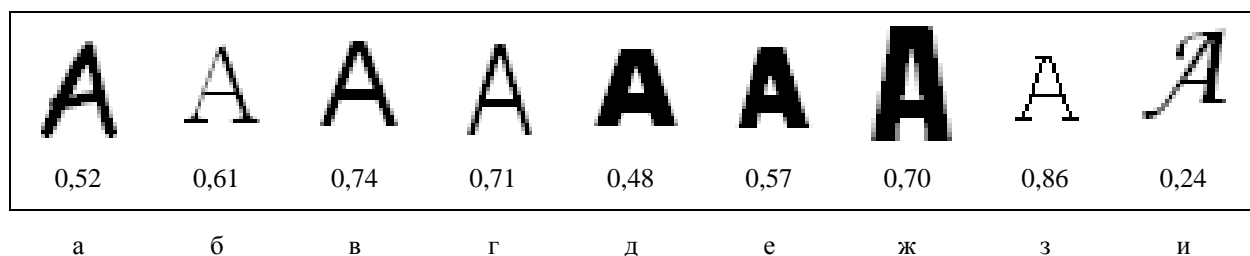


Рис. 6. Изображения одного символа с помощью различных шрифтов и размеров

Выводы

Таким образом, авторами формализована идея представления символа двухуровневой моделью: множеством узлов (геонов) и отношениями между узлами на верхнем уровне; множеством моделей гиперсложных клеток и отношениями между ними на нижнем уровне. Эксперименты по распознаванию символов подтвердили возможность интерпретации символа с помощью множества ключевых компонентов (узлов). В отличие от других методов предлагаемый подход имеет преимущества, поскольку не требует предварительного обучения и позволяет выполнять дальнейшее распознавание по ключевым узлам символов. Это позволяет исключить процесс предварительного масштабирования символов, а также выполнять шрифтонезависимое распознавание.

Список литературы

1. Арлазаров В.Л. Адаптивное распознавание символов [Электронный ресурс] / В.Л. Арлазаров // Режим доступа к статье: <http://www.cuneiform.ru/tech/adaptive.html>.
2. Васюра А.С. Моделирование нейросети для решения задачи идентификации символов [Электронный ресурс] / А.С. Васюра // Наукові праці ВНТУ. – 2007. – № 1. – Режим доступа к статье: www.nbuv.gov.ua/e-journals/VNTU/2007-1/ru/07vasosi_ru.pdf.
3. Рецидив графического спама. Спамтест [Электронный ресурс] // Режим доступа к ресурсу: <http://www.spamtest.ru/news?id=207509462>.
4. Яфизов Р.Н. Проактивные антиспам-технологии [Электронный ресурс] / Р.Н. Яфизов // "Information Security/Информационная безопасность". – 2008. – №4. – Режим доступа к статье: http://www.itsec.ru/articles2/Oborandteh/proaktivn_antispam_tehnologii.
5. Академики обсудили новые методы борьбы со спамом / Snews|Бизнес [Электронный ресурс] // Режим доступа к ресурсу: <http://www.cnews.ru/news/line/index.shtml?2007/04/05/244199>.
6. Ярмошевич Е.И. Функциональная спектральная пространственно-временная модель формирования изображений объектов зрительной системой человека [Электронный ресурс] / Е.И. Ярмошевич // Вестник ВГУ, серия: системный анализ и информационные технологии. – 2008. – №1. – С. 74-78. – Режим доступа к статье: www.vestnik.vsu.ru/pdf/analiz/2008/01/yarmoshevich.pdf.
7. Шиффман Х.Р. Ощущение и восприятие: Пер. с англ. З. Замчук; 5-е изд. / Х.Р. Шиффман. – СПб.: Питер, 2003. – 928 с., ил. – (Серия «Мастера психологии»).
8. Андерсон Дж. Когнитивная психология: пер. с англ. С. Комаров; 5-е изд. / Дж. Андерсон. – СПб.: Питер, 2002. – 496 с. – (Серия «Мастера психологии»).
9. Марр Д. Зрение. Информационный подход к изучению представления и обработки зрительных образов: пер. с англ. Н.Г. Гуревич / Д. Марр. – М.: Радио и связь, 1987. – 440 с.
10. Каргин А.А. Применение нечетких моделей когнитивных знаний восприятия изображения в задаче локализации строк текста / А.А. Каргин, Е.Е. Пятюков // Искусственный интеллект. – 2009. – №2. – С. 161-167.
11. Каргин А.А. Модели представления ключевых элементов в задаче покомпонентного распознавания символов текста / А.А. Каргин, Е.Е. Пятюков // Системні дослідження та інформаційні технології. – 2011. – №3. – С. 48-55.

Поступила в редколлегию 2.11.2011

Рецензент: д-р техн. наук, проф. Ю.А. Скобцов, Донецкий национальный технический университет, Донецк.

ПОКОМПОНЕНТНЕ РОЗПІЗНАВАННЯ СИМВОЛІВ ТЕКСТУ НА ОСНОВІ НЕЧІТКОГО ПРЕДСТАВЛЕННЯ

А.О. Каргін, О.Є. П'ятюков

У статті розглядається етап ідентифікації символу на основі зіставлення отриманого набору ключових вузлів з описом прототипу з бази даних. Формування вузлів ґрунтується на нечітких моделях клітин первинної зорової кори. Вводяться математичні описи моделі прототипу-символу, моделі символу і нечіткої ступені схожості. Наводяться результати експериментів.

Ключові слова: обробка зображення, зорова система, розпізнавання, символ, модель, вузли, нечітка ступінь схожості.

CHARACTER RECOGNITION OF TEXT BASED ON A FUZZY REPRESENTATION

A.A. Kargin, O. Ye. Piatykov

The article describes the stage of the identification character of text, which is performed by comparing the resulting set of core units with a description of the prototype of the database. The formation of the nodes is based on fuzzy models of primary visual cortex cells. We described the mathematical description of the model prototype, symbol models and the fuzzy degree of similarity. The paper presents experiments

Keywords: image processing, visual system, recognition, symbol, model, nodes, fuzzy degree of similarity.