

МОДЕЛИ ПРЕДСТАВЛЕНИЯ КЛЮЧЕВЫХ ЭЛЕМЕНТОВ В ЗАДАЧЕ ПОКОМПОНЕНТНОГО РАСПОЗНАВАНИЯ СИМВОЛОВ ТЕКСТА

А.А. КАРГИН, Е.Е. ПЯТИКОП

Рассмотрен этап формирования ключевых элементов (узлов) для решения задачи покомпонентного распознавания символов текста. Образование узлов основывается на нечетких моделях первичной зрительной коры. Введено математическое описание узла в виде нечеткой модели и степень нечеткой похожести узлов. Приведены примеры сравнения узлов с прототипами узлов.

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, поэтому 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 и 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 подетального анализа. Способ рассмотрения знакомых объектов как конфигурацию простых компонентов в своей работе предложил Д. Марр [9]. Также известна теория Бидермана — 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 три характерных узла. Простейшие детали очень похожи на результаты работы клеток первичной зрительной коры.

При цьому схеми комбінацій деталей дозволяють визначити зв'язи між деталями, які більш важливі, і відкинути менш суттєві. Так всі букви, показані на рис. 1, є буквою «А».



Рис. 1. Приклади написання букви «А»

Таким чином, піддетальний аналіз включає в себе такі процеси:

- розпізнавання простіших деталей, які складають геон (вузол);
- розпізнавання ключевих вузлів;
- ідентифікація на основі сопоставлення отриманого набору ключевих вузлів з описом прототипа з бази даних.

Перший етап — представлення об'єкта на рівні простіших деталей за допомогою моделей кліток первинної зрительної кори (ПЗК) розглянуто в роботі [10]. В даній роботі розглядається другий етап — формування і розпізнавання ключевих вузлів.

ПОСТАНОВКА ЗАДАЧИ

Результатом роботи першого етапу є нечітке представлення зображення в формі $M_{\omega}^{(3)} = \langle \omega, \{C^{\alpha}, \forall \alpha \in [0^{\circ}, 10^{\circ}, 20^{\circ} \dots 170^{\circ}]\} \rangle$, де $C^{\alpha} = \{\{G_j^{\alpha, \omega, \ell_i}\}_{\ell_i}\}$ — орієнтаційна колонка, $G^{\alpha, \omega, \ell}$ — клітка первинної зрительної кори, яка реагує на лінійні сегменти шириною ω , орієнтовані певним чином і певної довжини ℓ , α — кут орієнтації. Модель цієї клітки є виразом:

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

де $K^{\alpha, r} = \langle k^{\alpha, r+2Rk}, k^{\alpha, r+4Rk} \dots k^{\alpha, r+\ell^* \cdot 2Rk} \rangle$ — сукупність суміжних активних гангліозних кліток сітчатки, на основі якої формується детектор $D^{\alpha} = \langle K^{\alpha, r}, \alpha, \omega, \ell^*, \delta(D^{\alpha, \omega}) \rangle$, а $\delta(G^{\alpha, \omega, \ell})$ — функція впевненості наявності детектора. Для виразу (1) $\delta(G^{\alpha, \omega, \ell}) = \delta(D^{\alpha, \omega})$.

Таким чином, як базове множинство простіших деталей прийняті елементи, отримані на етапі обробки зображення за моделлю кліток ПЗК. На другому етапі, який розглядається в даній роботі, необхідно виконати представлення зображення символом набором геонів, тобто його ключевих вузлів, які формуються з простіших деталей, і зв'язями між цими геонами.

Ціль роботи — розробка моделі покомпонентного представлення і розпізнавання символів на основі нечіткого представлення, отриманого в результаті обробки зображення за моделлю кліток первинної зрительної кори.

ПРЕДСТАВЛЕНИЕ ИЗОБРАЖЕНИЯ СИМВОЛА НАБОРОМ КЛЮЧЕВЫХ УЗЛОВ

Из анализа символов латинского и русского алфавитов [11] автором было выбрано 9 типов ключевых геонов (далее узлов). Это множество базовых узлов выбрано с учетом входных данных — моделей клеток зрительной коры и является достаточным для формирования произвольных символов. Поскольку клетки зрительной коры имеют определенную ориентацию, то элементы узлов тоже ориентированы. На рис. 2 приведены предлагаемые типы узлов.

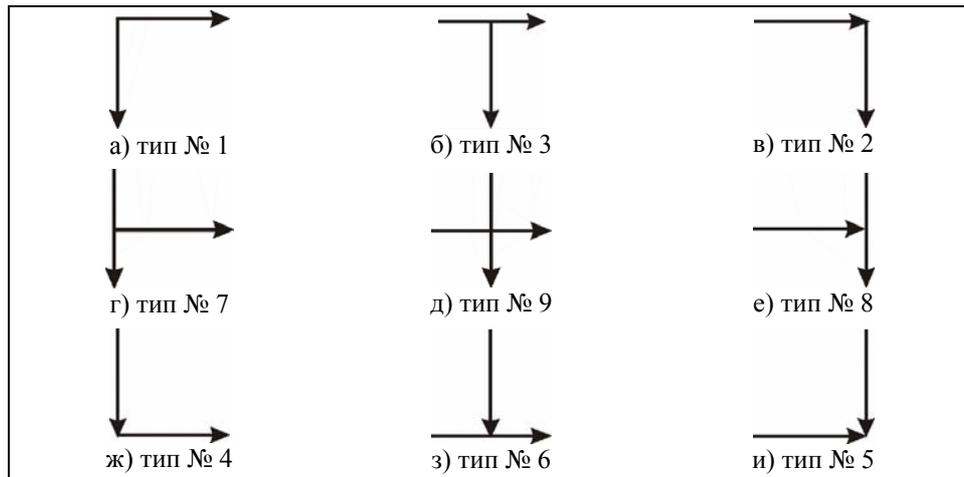


Рис. 2. Типы узлов

В базе данных хранятся прототипы символов, которые выражены через прототипы приведенных узлов. Идентификация изображения проанализированной области сводится к построению множества возможных моделей узлов и сопоставление их с прототипами узла. Прототип — это обобщенное представление, в котором воспроизведен набор общих и детализированных признаков типичного объекта (действия), и которое выступает в качестве основы для идентификации любого нового понятия. Введем модель узла-прототипа.

МОДЕЛЬ УЗЛА-ПРОТОТИПА

Узел формируется из двух гиперсложных клеток G' и G'' различной ориентации, которые определяются, как было описано ранее, на множестве $K = \{k_z\}$ активных сопряженных ганглиозных клеток. Если некоторая ганглиозная клетка k_z влияет и на формирование гиперсложной клетки G' , и на формирование гиперсложной клетки G'' , то можно предположить о некотором пересечении гиперсложных клеток. Множество таких общих клеток опишем выражением:

$$Kd = \{k_z\}_{n_{Kd}}, \quad (2)$$

где $k_z \in G'$ и $k_z \in G''$.

В модели вместо точных координат пиксельной плоскости пересечение гиперсложных клеток задается размытой областью. Для этого гиперсложная клетка разбита на три окрестности (области):

- окрестность O' с центром (x', y') в начале G ;
- окрестность O'' с центром (x'', y'') в середине G ;
- окрестность O''' с центром (x''', y''') в конце G ,

где началом клетки G принимается центр ганглиозной клетки k_0 или k_ℓ в зависимости от того, от какой из клеток меньше евклидово расстояние от центра до точки пиксельной плоскости $(0,0)$. Тогда центр другой клетки принимается как конец клетки G . Середина определяется по формуле середины отрезка.

Радиус каждой окрестности определяется:

$$R^o = \frac{\ell}{4}, \quad (3)$$

где ℓ — длина гиперсложной клетки G , равная количеству активизировавшихся ее ганглиозных клеток k_z .

Для каждого элемента k_z множества Kd необходимо определить функцию принадлежности к каждой из окрестностей O', O'', O''' относительно каждой клетки G' и G'' . Для этого введем функцию $\eta(K_z, O_G)$, вычисляемую по формуле:

$$\eta(k_z, O_G) = \begin{cases} 0 & \text{если } d(k_z, O_G) > R^o, \\ \frac{R^o - d(k_z, O_G)}{R^o}, & \text{если } d(k_z, O_G) \leq R^o, \end{cases} \quad (4)$$

где $d(k_z, O_G)$ — евклидово расстояние между центром k_z клетки и некоторой окрестностью $O_G \in \{O', O'', O'''\}_{G \in \{G', G''\}}$.

Характеристика пересечения гиперсложных клеток в модели узла-прототипа есть пара:

$$Kd = \langle \{k_z\}, \{\eta(K_z, O_G)\} \rangle, \quad (5)$$

где $k_z \in G'$ и $k_z \in G''$, $\eta(k_z, O_G)$ вычисляется по формуле (4), а

$$\begin{aligned} \{\eta(K_z, O_G)\} &= \{\eta(k_z, O_{G'}), \eta(k_z, O_{G''}), \eta(k_z, O_{G'''}), \\ &\eta(k_z, O_{G''''}), \eta(k_z, O_{G'''''}), \eta(k_z, O_{G''''''})\}. \end{aligned}$$

Модель узла-прототипа есть выражение:

$$UP = \langle \{G', G''\}, Kd, \beta \rangle, \quad (6)$$

где $G', G'' \in \{G^{\alpha, \ell}\}$ и клетка G' та, у которой угол α ближе к 90° ; Kd — множество ганглиозных клеток, которые являются общими для гиперсложных клеток G' и G'' ; $\beta = |\alpha_{G'} - \alpha_{G''}|$ — угол пересечения между гиперсложными клетками, формирующими узел.

Угол β при описании прототипа не является конкретным значением, т.к. согласно теории Бидермана, узлы не подвержены искажениям. Поэтому к типу № 1 также будут относиться и узлы, показанные на рис. 3.



Рис. 3. Множество узлов типа № 1

Аналогично, к типу № 7 будут относиться такие варианты, как показано на рис. 4.



Рис. 4. Множество узлов типа № 7

Поскольку типов узлов всего 9 будем иметь множество прототипов-узлов $\{UP_j\} \quad j \in 1, \dots, 9$. Описание узлов-прототипов приведено в таблице.

Т а б л и ц а . Описание прототипов узлов

№	Тип	$\eta(K_z, O_G)$					
		область O' (начало)		область O'' (середина)		область O''' (конец)	
		G'	G''	G'	G''	G'	G''
1		1	1	0	0	0	0
2		1	0	0	0	1	0
3		1	0	0	1	0	0
4		0	1	0	0	1	0
5		0	0	0	0	1	1
6		0	0	0	1	1	0
7		0	1	1	0	0	0
8		0	0	1	0	0	1
9		0	0	1	1	0	0

МОДЕЛЬ УЗЛА

Модель произвольного узла U по аналогии с моделью узла-прототипа описывается выражением:

$$U = \langle \{G', G''\}, Kd, \beta \rangle. \quad (7)$$

Кроме явного пересечения гиперсложных клеток при воздействии помех возможны варианты условного пересечения, показанные на рис. 5, когда общей для G' и G'' клетки k_z нет, но есть некоторое множество клеток k_z , которые попадают в окрестности O', G'', G''' клеток G' и G'' .

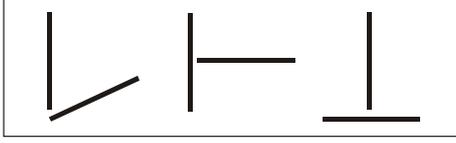


Рис. 5. Пример условного пересечения

Для учета ситуаций, показанных

в примерах на рис. 5, введено множество Kd' :

$$Kd' = \langle \{k_z\}, \{\eta(k_z, O_G)\} \rangle, \quad (8)$$

где $k_z \notin G'$ и $k_z \in G''$, но $k_z \in O_{G'}$; $\eta(k_z, O_G)$ вычисляется по формуле (4).

Тогда внесем изменения в выражение (7) и модель узла будет формально иметь описание:

$$U = \langle \{G', G''\}, Kd, Kd', \beta \rangle. \quad (9)$$

На рис. 6 приведен пример узла U_i , разбиение на окрестности гиперсложных клеток G' и G'' и функции принадлежности $\eta(K_z, O_G)$ общей ганглиозной клетки k_z к каждой из окрестностей.

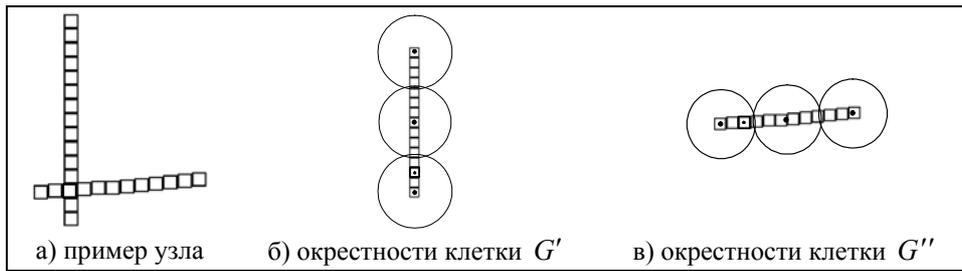


Рис. 6. Пример узла с разбиением на окрестности и значениями функции принадлежности к каждой из окрестностей O', O'', O''' относительно каждой клетки G' и G'' , полученными по формулам (4)

МЕТОД РАСПОЗНАВАНИЯ КЛЮЧЕВЫХ УЗЛОВ

Пусть для данного фрагмента изображения на множестве выделенных простейших деталей (клеток зрительной коры $\{\{G_j^{\alpha, \omega, \ell_i}\}_{\ell_i}\}_{\alpha}$) сформировано множество узлов $U = \{U_i\}$, $i \in 1, \dots, n_u$, по формуле (9).

Для сравнения узла-прототипа UP_j с узлами U_i , найденным на изображении, введена степень нечеткой похожести $\xi(UP_j, U_i)$, которая вычисляется по формуле:

$$\xi(UP_j, U_i) = \begin{cases} 0, & \text{если } Kd_{U_i} = \emptyset \text{ и } Kd'_{U_i} = \emptyset, \\ \min \left\{ \frac{1}{2} (1 + \eta_{Kd_{U_i}}(k_z, O)), \eta_{Kd_{UP_j}}(k_z, O) \right\}, & \text{если } Kd_{U_i} \neq \emptyset, \\ \min \{ \eta_{Kd'_{U_i}}(k_z, O), \eta_{Kd_{UP_j}}(k_z, O) \}, & \text{если } Kd_{U_i} = \emptyset \text{ и } Kd'_{U_i} \neq \emptyset. \end{cases} \quad (10)$$

Для каждого U_i узла вычисляется множество функций $\{\xi(UP_j, U_i)\}_j$, $j \in 1, \dots, 9$ т.е. для каждого узла-прототипа. Например, на рис. 7, а показано изображение буквы «М», а на рис. 7, б — полученные для нее детекторы.

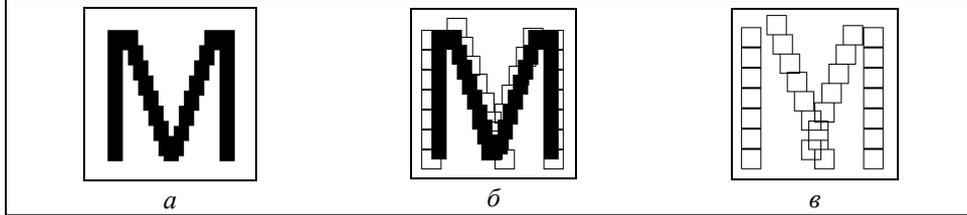


Рис. 7. Изображение буквы «М»

Тогда эти детекторы образуют узлы, показанные на рис. 8.

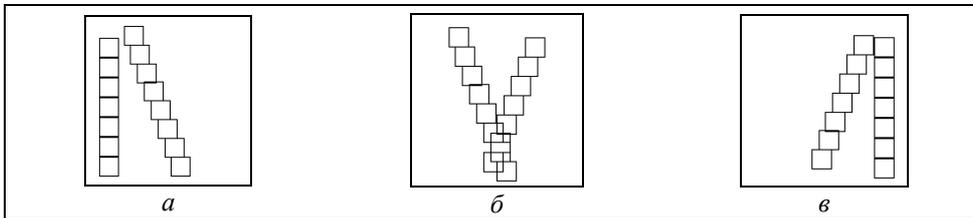


Рис. 8. Узлы для буквы «М»

Для узла показанного на рис. 8, а модель узла имеет вид:

$$U_1 = \langle \{G' = G^{90^\circ}, G'' = G^{70^\circ}\}, Kd \neq \emptyset, Kd' = \emptyset, 20^\circ \rangle,$$

где $Kd = \langle k_z, \{\eta(k_z, O_G)\} = \{1, 0, 0, 1, 0, 0\} \rangle$.

Модель узла показанного на рис. 8, б имеет вид:

$$U_2 = \langle \{G' = G^{110^\circ}, G'' = G^{70^\circ}\}, Kd \neq \emptyset, Kd' = \emptyset, 40^\circ \rangle,$$

где $Kd = \langle k_z, \{\eta(k_z, O_G)\} = \{0, 0, 0.82, 0, 0, 0.81\} \rangle$.

Модель узла показанного на рис. 8, в имеет вид:

$$U_3 = \langle \{G' = G^{90^\circ}, G'' = G^{110^\circ}\}, Kd = \emptyset, Kd' \neq \emptyset, 20^\circ \rangle,$$

где $Kd' = \langle k_z, \{\eta(k_z, O_G)\} = \{0.5, 0, 0, 0, 0, 0.5\} \rangle$.

Тогда в результате сравнения с множеством прототипов $\{UP_j\}$ $j \in 1, \dots, 9$ получено множество $\{\xi(UP_j, U_i)\}_j$, $j \in 1, \dots, 9$:

- $\{\xi(UP_j, U_1)\}_j = \{1, 0, 0, 0, 0, 0, 0, 0\}$,
- $\{\xi(UP_j, U_2)\}_j = \{0, 0, 0, 0, 0.81, 0, 0, 0\}$,
- $\{\xi(UP_j, U_3)\}_j = \{0, 0.5, 0, 0, 0, 0, 0, 0\}$.

Таким образом, для буквы «М», показанной на рис. 7, а, на основе значений множества $\{\xi(UP_j, U_i)\}_j$, приведенных выше, получаем, что первый узел близок к типу № 1 со степенью схожести равной 1, второй узел со степенью схожести равной 0,81 сходен на узел типа № 5, а третий узел со степенью схожести равной 0,5 — на тип № 2. Из этого следует

$$\begin{aligned} & \{\{\xi(UP_j, U_i)\}_j\}_{U_i} = \\ & = \{\{1, 0, 0, 0, 0, 0, 0, 0\}, \{0, 0, 0, 0, 0.81, 0, 0, 0\}, \{0, 0.5, 0, 0, 0, 0, 0, 0\}\}. \end{aligned}$$

Поскольку представление изображения символа выполнено в виде множества ключевых узлов, то можно перейти к процессу их объединения.

ВЫВОДЫ

Формализована идея представления символа множеством узлов и отношениями между узлами. Представлены девять типов узлов, введена модель узла-прототипа с основными параметрами, модель узла и степень нечеткой похожести, разработан метод распознавания ключевых узлов. Показано, что множество узлов-прототипов достаточно для представления произвольного символа из русского и латинского алфавитов. Эксперименты показали, что на этом наборе символы устойчиво различимы.

На основе описанных моделей и метода авторами представлен символ двухуровневой моделью: множеством узлов и отношениями между узлами на верхнем уровне; множеством гиперсложных клеток и отношениями между ними на нижнем уровне. Эксперименты 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 и выполнять шрифтонезависимое распознавание. Таким образом, показано использование результатов формализации когнитивных функций по обработке визуальной информации на разных уровнях.

ЛИТЕРАТУРА

1. Арлазаров В.Л. Адаптивное распознавание символов. — <http://www.cuneiform.ru/tech/adaptive.html>.
2. Васюра А.С. Моделирование нейросети для решения задачи идентификации символов // Наук. пр. ВНТУ. — 2007. — № 1. — www.nbu.gov.ua/e-journals/VNTU/2007-1/ru/07vasosi_ru.pdf.
3. Рецидив графического спама. Спамтест. — <http://www.spamtest.ru/news?id=207509462>.
4. Яфизов Р. Проактивные антиспам-технологии // Information Security «Информационная безопасность». — 2008. — № 4. — http://www.itsec.ru/articles2/Oborandteh/proaktivn_antispam_tehnologii.
5. Академики обсудили новые методы борьбы со спамом // Бизнес. — <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 // Вестн. ВГУ. Сер. Системный анализ и информационные технологии. — 2008. — № 1. — С. 74–78. — 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. — СПб.: Питер, 2003. — 928с. — (Мастера психологии).
8. Андерсон Дж. Когнитивная психология / Пер. с англ. С. Комаров. — 5-е изд. — СПб.: Питер, 2002. — 496 с. — (Мастера психологии).
9. Марр Д. Зрение. Информационный подход к изучению представления и обработки зрительных образов / Пер. с англ. Н.Г. Гуревич. — М.: Радио и связь, 1987. — 440 с.
10. Каргин А.А., Пятикоп Е.Е. Применение нечетких моделей когнитивных знаний восприятия изображения в задаче локализации строк текста // Искусственный интеллект. — 2009. — № 2. — Р. 161–167.
11. Гордон Ю. Книга про буквы от Аа до Яя. — М.: Студия «Артемия Лебедева», 2006. — 384 с.

Поступила 01.06.2009