

К ВОПРОСУ О МОДЕЛИРОВАНИИ ЗНАКОВ ПРОИЗВОЛЬНОЙ ПРИРОДЫ НА ДИСКРЕТНЫХ МНОЖЕСТВАХ С ЦЕЛЬЮ ИХ ОПОЗНАВАНИЯ

С.В. Мышко, Е.В. Шевчук

Донецкий национальный университет.

Проблема автоматического распознавания знаков произвольной природы, примерами которых могут служить рукописные символы, геометрические фигуры и т.д., на сегодняшний день не утратила своей актуальности. При проектировании систем автоматического распознавания зрительных образов, одной из важных задач, подлежащих решению, является выделение характерных признаков и свойств объектов распознавания. Однако, как отмечается в ряде работ и, в частности в [1], на практике это трудноразрешимая проблема. Многочисленные попытки автоматизации распознавания написанных рукой человека букв и цифр выявили ряд проблем. В силу того, что приходится иметь дело с бесконечным разнообразием начертаний знаков произвольной природы, невозможно сформировать конечное множество эталонов. Значительно меняются размеры и детали структур, которые часто оказываются существенно разными даже для одинаковых символов, написанных одной и той же рукой.

В данной работе предлагается для рассмотрения конструктивный метод автоматического формирования модели знака, не предполагающий необходимости заведения меры близости и, как следствие, введения субъективно-статистически определённых констант.

В последующих разделах дается ряд основных определений и излагается метод моделирования знаков произвольной природы на дискретных множествах.

Поместим в компьютер при помощи некоторого устройства ввода бинарное изображение. В подавляющем большинстве графические устройства являются растровыми, то есть представляют изображение в виде прямоугольной матрицы пикселей. Система опознавания, после так называемой «оцифровки» изображения в качестве исходной информации получает данные о дискретном множестве элементов поверхности изображения, представленные множеством чисел. По-

этому модель опознаваемого знака предлагается генерировать на основании свойств этого множества, а не исходного объекта.

Изображение, прошедшее процесс оцифровки, моделируется конечным множеством атомарных элементов [2] $A = \{\alpha_h\}$, где $\alpha_h = \alpha(i_h, j_h)$, $i_h \in [1, I]$, $j_h \in [1, J]$, $h \in [1, H]$.

Атомарные элементы, участвующие в формировании знака, назовем активными атомарными элементами а не участвующие (составляющие фон) - пассивными атомарными элементами.

Дальнейшие рассуждения проведем для активных атомарных элементов и, говоря «атомарный элемент», будем подразумевать активный атомарный элемент.

С учетом введенного в [3] понятия связности атомарных элементов и связного множества атомарных элементов, дадим определение знака.

Пусть A_a - множество активных атомарных элементов, представимое в виде совокупности подмножеств связных атомарных элементов: $A_a = \bigcup_{k=1}^K Z_k$, $Z_t \cap Z_p = \emptyset$, $\forall t \neq p$, $t, p \in [1, K]$.

Определение 1. Связное множество активных атомарных элементов $Z_k \in A_a$ будем называть знаком.

На множестве атомарных элементов знака Z_k определим весовую функцию $V(\alpha(i, j))$, ставящую в соответствие каждому атомарному элементу $\alpha(i, j) \in Z_k$ целое число, равное количеству атомарных элементов знака, связных с данным атомарным элементом.

Из определения пути [3] на множестве атомарных элементов следует, что для любого атомарного элемента α_r такого, что $\alpha_r \in L$, где L - некоторый путь, всегда верно, что $V(\alpha_r) \leq 2$. Атомарные элементы с весом, превышающим 2, как правило принадлежат пересечению путей [3].

Пусть Z_k -знак и существует подмножество $S \subset Z_k$, значение весовой функции атомарных элементов которого превышает 2. То есть, $\exists S \subset Z_k: S = \{\alpha_r \in Z_k \mid V(\alpha_r) > 2; r = 1, 2, \dots, R, R \in \mathbb{N}\}$.

Пусть $S \subset Z_k$ представимо в виде совокупности связных подмножеств атомарных элементов: $S = \bigcup_{c=1}^{C_1} S_c, c \in [1, C_1]$.

Определение 2. Всякое подмножество $S_c \subset S, c \in [1, C_1]$ будем называть *скоплением*.

Из определения связного множества следует, что любой знак должен быть представлен в виде пути [3] или объединения путей [3], то есть, $Z_k = \bigcup_{n=1}^{C_2} L_n, L_t \cap L_p = \emptyset, \forall t \neq p, t, p \in [1, C_2]$.

Для связных элементов, введены следующие обозначения [3]: s_1 – горизонтальная связка; s_2 – вертикальная связка; s_3, s_4 – диагональные связки. На множестве пар связных атомарных элементов определена функция η [3] следующим образом:

$$\eta(\alpha_a, \alpha_b) = \begin{cases} s_1, & \text{если } i_a = i_b, |j_a - j_b| = 1; \\ s_2, & \text{если } j_a = j_b, |i_a - i_b| = 1; \\ s_3, & \text{если } |i_a - i_b| = |j_a - j_b| = 1, (i_a - i_b)(j_a - j_b) = 1; \\ s_4, & \text{если } |i_a - i_b| = |j_a - j_b| = 1, (i_a - i_b)(j_a - j_b) = -1. \end{cases}$$

Если $\eta(\alpha_a, \alpha_b) = s_m$, то при $m=1$ будем говорить, что пара (α_a, α_b) является горизонтальной связкой; при $m=2$ – вертикальной, при $m=3$ – диагональной связкой первого типа, при $m=4$ – диагональной связкой второго типа.

Зафиксируем α_a, α_b – произвольные атомарные элементы из Z_k и определим понятие ориентированного множества атомарных элементов как упорядоченной последовательности связок.

Определение 3. Конечное вполне упорядоченное подмножество $R' \subset Z_k, R' = \{(\alpha_h, \alpha_{h+1})_m\}_{h=1}^{N_1}, i_h \in [1, I], j_h \in [1, J], m \in \{1, 2, 3, 4\}$, будем называть ориентированным множеством, если в нем, наряду с вертикальными и (или) горизонтальными связками, встречаются диагональные связки только одного типа или отсутствуют вообще. То есть, не существует $t, r \in [1, N_1]$ таких, что для $(\alpha_t, \alpha_{t+1}), (\alpha_r, \alpha_{r+1}) \in R', \eta((\alpha_t, \alpha_{t+1})) = s_w, \eta((\alpha_r, \alpha_{r+1})) = s_v$, где $w, v \in \{3, 4\}$ и $w \neq v$.

Ориентацией множества будем считать тип связей, преобладающих в данном ориентированном множестве. То есть, если в данном ориентированном множестве преобладают связи типа s_1 , то множество будем считать вертикально-ориентированным. Если же преобладают связи типа s_2 , то – горизонтально-ориентированным. Если s_3, s_4 – лево- и право-ориентированным соответственно.

Множество, состоящее из связей только типа s_3 или s_4 будем называть соответственно лево- и право-диагональным.

Выявление на знаке ориентированных множеств позволяет с определенной долей уверенности моделировать породивший этот знак прообраз концепта в виде совокупности локально-глобальных направлений движения.

Однако выявление всех локальных движений может приводить к излишней детализации. Например, знак, изображенный на рис. 4, представим в виде совокупности локальных движений, но всегда ли необходима такая подробная детализация?

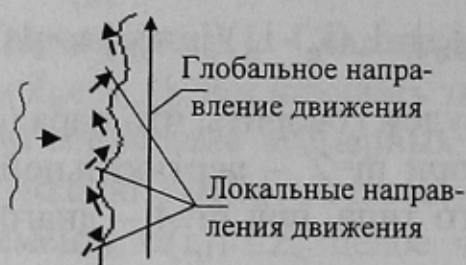


Рисунок 4 - Пример локальных и глобальных движений на знаке

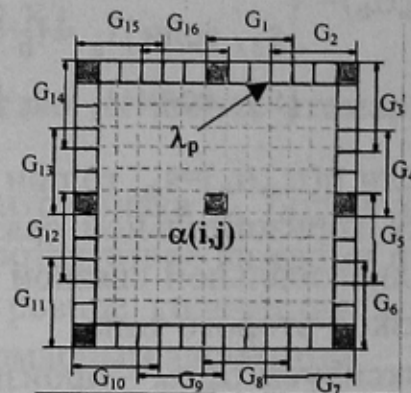


Рисунок 5 - Пример разбиения граней квадрата на сектора G_1, \dots, G_{16} при $p=5$ и $q=2$

Очевидно, что нет. Но на практике определить необходимую степень детализации для произвольного знака невозможно, поэтому для построения модели предлагается многоуровневая система представлений знаков машинных концептов с различной степенью детализации на каждом уровне. Различная детализация достигается за счет изменения от уровня к уровню параметров элементов представления.

Как указано ранее, описываемая в данной работе модель знака являет собой многоуровневую систему представлений.

Обозначим через U_p – p -й уровень представления совокупности атомарных элементов, $p \in \mathbb{N}$. Множество атомарных элементов A будем полагать атомарным уровнем и обозначать U_0 .

Для определения элементов представления и структур уровней представления U_p , введем понятие квадрата, составленного из атомарных элементов $\alpha(i, j) \in A$ с центром в атомарном элементе $\alpha(i', j') \in A$, и обозначим его γ_p :

$$\gamma_p(\alpha(i', j')) = \{\alpha(i, j) \mid i = i' + k_1, j = j' + k_2; k_1 = \overline{-r, r}, k_2 = \overline{-r, r}, r = p + 1\}$$

Рассмотрим конструкцию λ_p , представленную на рисунке 5. Она являет собой границу квадрата γ_p составленную из атомарных элементов:

$$\lambda_p = \gamma_p \setminus \gamma_{p-1}.$$

Представим границу λ_p объединением *граней* $\Gamma_1 \cup \Gamma_2 \cup \Gamma_3 \cup \Gamma_4$, где:

$$\Gamma_1 = \{\alpha(i, j) \in \lambda_p \mid i = i' + k_1, j = j' - r; k_1 = \overline{-r, r}, r = p + 1\},$$

$$\Gamma_2 = \{\alpha(i, j) \in \lambda_p \mid i = i' + r, j = j' + k_1; k_1 = \overline{-r, r}, r = p + 1\},$$

$$\Gamma_3 = \{\alpha(i, j) \in \lambda_p \mid i = i' + k_1, j = j' + r; k_1 = \overline{-r, r}, r = p + 1\},$$

$$\Gamma_4 = \{\alpha(i, j) \in \lambda_p \mid i = i' - r, j = j' + k_1; k_1 = \overline{-r, r}, r = p + 1\}.$$

Центральными элементами граней $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$ будем полагать атомарные элементы $\alpha(i', j' - (p+1))$, $\alpha(i' + (p+1), j')$, $\alpha(i', j' + (p+1))$, $\alpha(i' - (p+1), j')$ соответственно.

Центральные элементы разбивают грань на две полуграни, причем центральный элемент грани принадлежит обеим полученным полуграням.

Определение 4. Элементом p -го уровня представления совокупности атомарных элементов будем называть множество атомарных элементов $\beta_p(\alpha(i', j'))$, такое, что:

$$\beta_p(\alpha(i', j')) = \{\alpha(i, j) \mid \alpha(i, j) \in \lambda_p \cup \alpha(i', j'); i, i' \in [1, I], j, j' \in [1, J]\}.$$

Центром елемента p -го рівня представлення $\beta_p(\alpha(i', j'))$ назовем елемент $\alpha(i', j')$.

Представим границю λ_p в виде совокупности секторов $G_1(\alpha(i', j')), \dots, G_g(\alpha(i', j'))$ так, как показано на рисунке 5. При этом, каждая полугрань разбивается на q секторов. Для каждого уровня p допустимые размеры секторов варьируются. Максимальный размер сектора равен количеству атомарных элементов на полуграни, а минимальный – одному атомарному элементу.

Сектора G_1 и G_9 , G_2 и G_{10} , G_3 и G_{11} и т.д., приведенные на рисунке 5 назовем эквивалентными секторами и в дальнейшем будем обозначать эквивалентность секторов следующим образом: $G_1 \equiv G_9$.

Определение 5. Элемент $\beta_p(\alpha(i', j'))$ назовем структурой p -го уровня с q секторами на полуграни и будем обозначать $\beta_{pq}(\alpha(i', j'))$, если для его полуграни заведено q секторов.

Пусть на атомарном уровне задан произвольный знак Z_k . Выберем некоторый уровень представлення U_p , значение q и покроем его структурами $\beta_{pq}^1(\alpha(i_1, j_1)), \dots, \beta_{pq}^T(\alpha(i_T, j_T))$ так, что для всякого $\beta_{pq}^t(\alpha(i_t, j_t))$, $t \in [1, T-1]$, соответствующий ему центральный атомарный элемент принадлежит знаку Z_k : $\alpha(i_t, j_t) \in Z_k$ и $(\beta_{pq}^t \setminus \alpha(i_t, j_t)) \cap (\beta_{pq}^{t+1} \setminus \alpha(i_{t+1}, j_{t+1})) \neq \emptyset$.

Знак некоторым образом пересечет границу структуры $\beta_{pq}^t(\alpha(i_t, j_t))$, $t \in [1, T]$. Атомарные элементы $\alpha', \alpha'' \in Z_k$: $\alpha', \alpha'' \in Z_k \cap (\beta_{pq}^t(\alpha(i_t, j_t)) \setminus \alpha(i_t, j_t))$, $t \in [1, T]$ будем называть соответственно точкой входа и точкой выхода структуры $\beta_{pq}^t(\alpha(i_t, j_t))$.

Сектор, содержащий точку входа $\alpha' \in G'_e(\alpha(i_t, j_t))$, $t \in [1, T]$ назовем входным сектором, а сектор, содержащий точку выхода $\alpha'' \in G''_v(\alpha(i_t, j_t))$ – выходным сектором структуры $\beta_{pq}^t(\alpha(i_t, j_t))$ при $t \in [1, T]$, $e, v \in [1, 8q]$.

Определение 6. Элементарной стратегией S_{pq} p -го уровня с q секторами на полуграни будем называть подмножество атомарных

элементов знака Z_k , покрытое совокупностью структур $\beta_{pq}^1(\alpha(i_1, j_1), \dots, \beta_{pq}^{T_1}(\alpha(i_{T_1}, j_{T_1})))$, при $2 \leq T_1 \leq T$ такое что: $G'_{e_t}(\alpha(i_t, j_t)) \equiv G''_{v_t}(\alpha(i_t, j_t))$, где $e_t, v_t \in [1, g]$, $t \in [1, T_1]$ и $e_1 = e_2 = \dots = e_{T_1}, v_1 = v_2 = \dots = v_{T_1}$.

Точку входа структуры β_{pq}^1 будем считать началом стратегии, а точку выхода структуры $\beta_{pq}^{T_1}$ - концом стратегии S_{pq} .

Варьируя размер секторов и уровень представления p , мы можем огрублять модель рассматриваемого знака, избавляясь таким образом от избыточной информации, или, наоборот, детализировать.

Модель M_k на каждом уровне представляется как упорядоченная по определенному правилу последовательность выявленных на данном знаке элементарных стратегий.

Критерием корректности моделей служит следующее правило: по знаку Z_k строится его модель M_k , по этой модели восстанавливается знак Z'_k и по нему строится новая модель M'_k . Схематично изобразим это следующим образом: $Z_k \rightarrow M_k \rightarrow Z'_k \rightarrow M'_k$.

В том случае, если модели M_k и M'_k совпадают, моделирование знака Z_k осуществлено корректно.

Пример на рисунке 8 иллюстрирует приведенный критерий корректности моделей.

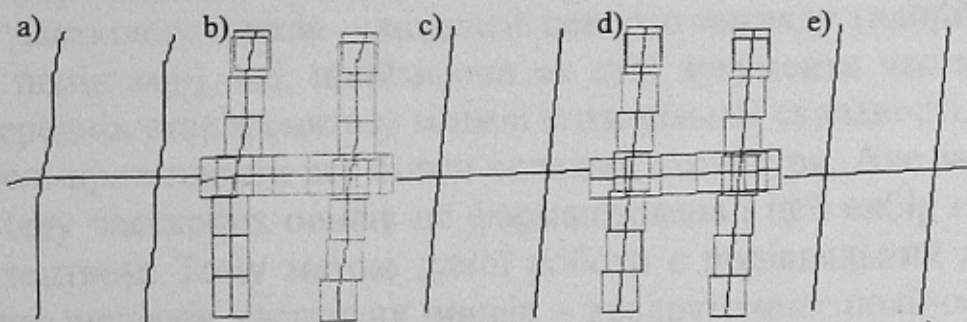


Рисунок 8 - Пример корректно построенной модели: а) исходный знак - Z_k ; б) покрытие знака, построение модели M_k ; в) знак, восстановленный по модели - Z'_k ; д) покрытие знака, построение модели M'_k ; е) знак, восстановленный по модели M'_k

Сформированные по приведенному методу модели являются конструктивными – позволяют строить по ним знаки, причем по одной модели можно построить множество знаков, считающихся «похожими», включающее исходный знак. Любому другому знаку, модель которого будет также эквивалентна данной, будет поставлено в соответствие имя, установленное для исходного знака.

В соответствии с изложенным были разработаны алгоритмы и программа автоматического формирования многоуровневых моделей знаков для дальнейшего их опознавания. Полученные при тестировании программы результаты подтверждают перспективность использования данного метода для построения систем автоматического опознавания знаков произвольной природы.

Список источников

1. Бонгард М.М. Проблемы узнавания. –М: Наука, 1967.
2. Мышко С.В., Шевчук Е.В, Шевцов Д.В. Моделирование знаков элементарными стратегиями. // Сб. докл. Первой международной научно-практической конференции «Вычислительная техника в информационных и управляющих системах» - Мариуполь, ПГТУ, 2000. -С. 79-80.
3. Мышко С.В., Шевцов Д.В. Конструктивное определение прямой в терминах свойств множеств атомарных элементов. // Праці наукової конференції Донецького національного університету за підсумками науково-дослідної роботи за період 1999-2000 рр. – Донецьк: 2001. - С. 99.
4. Мышко С.В., Шевцов Д.В. Определение прямой на множестве атомарных элементов. //International conferance on Optoelectronic Information-Energy Technologies – Винница, 2001.
5. Шикин Е.В., Боресков А.В. Компьютерная графика. М: Диалог-МФТИ, 1995. – С.112-113.