

Моделирование знаков произвольной природы на дискретных множествах с целью их дальнейшего опознавания

Шевчук Е.В., Котвинский В.В
Донецкий национальный университет
Донецк, Украина

АННОТАЦИЯ

На сегодняшний день проблема автоматического распознавания знаков произвольной природы, примерами которых могут служить рукописные символы, геометрические фигуры и т.д., 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, на практике это трудноразрешимая проблема. Например, Бонгард в работе «Проблемы узнавания» [1] отмечает, что в большинстве практических задач распознавания определение полного набора различительных признаков и построение модели опознаваемого объекта оказывается исключительно трудным делом.

Как следует из многочисленных работ, посвященных проектированию систем распознавания зрительных образов, большинство методов распознавания основано на предварительном формировании классов образов, заданных некоторыми характеристическими векторами признаков. Процесс распознавания, таким образом, заключается в отнесении распознаваемого объекта к одному из классов.

Однако, например, попытки автоматизации распознавания написанных рукой человека букв и цифр выявили ряд проблем. Говоря о знаках произвольной природы, невозможно сформировать конечное множество эталонов, поскольку приходится иметь дело с бесконечным разнообразием их начертаний. 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.

Основная идея метода заключается в том, что модель знака произвольной природы предлагается формировать в виде совокупности локально-глобальных направлений движения, воспроизводящих его начертание. Направление движения при начертании двух сходных вариантов одного и того же знака совпадают, что позволяет создать модель, инвариантную к его изменениям.

Ключевые слова: локально-глобальное направление, ориентированное множество, стратегия, многоуровневая модель.

1. ВВЕДЕНИЕ

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.

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

Для исследования свойств знака произвольной природы на дискретном множестве необходимо исследовать свойства подмножеств, в совокупности составляющих этот знак.

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

2. ОСНОВНЫЕ ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ

Поместим в компьютер при помощи некоторого устройства ввода бинарное изображение. При преобразовании исходного изображения в «машинный» вид происходит так называемая «оцифровка» изображения. Изображение, прошедшее процесс оцифровки, моделируется конечным множеством атомарных элементов [2] $A = \{\alpha_n\}$, где $\alpha_n = \alpha(i_n, j_n)$, $i_n \in [1, I]$, $j_n \in [1, J]$, $h \in [1, H]$.

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* (примеры связанных атомарных элементов приведены на рис.1), дадим определение знака.

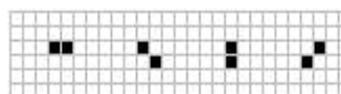


Рисунок 1. Примеры парно связанных между собой атомарных элементов

Пусть A_a – множество активных атомарных элементов, представимое в виде совокупности подмножеств связанных

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, связанных с данным атомарным элементом. Максимальное значение функции равно восьми, минимальное – нулю.

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 $V(\alpha_r) \leq 2$.

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 подмножеств атомарных элементов: $S = \bigcup_{c=1}^{C_1} S_c, c \in [1, C_1]$.

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

Примеры скоплений приведены на рисунке 2.

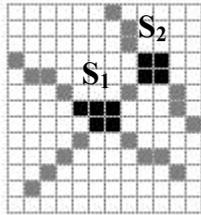


Рисунок 2. Пример множества S , состоящего из двух скоплений – S_1 и S_2

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

$\forall t \neq p, t, p \in [1, C_2]$ (см. рис. 3).

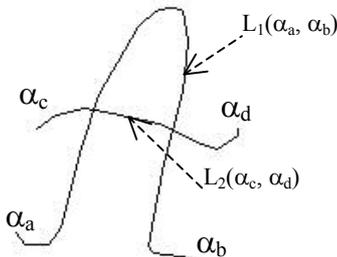


Рисунок 3. Буква «А» представлена объединением двух путей: $L_1(\alpha_a, \alpha_b)$ и $L_2(\alpha_c, \alpha_d)$.

Однако на практике часто приходится сталкиваться с «зашумленными» знаками (Рис. 4.а), содержащими в себе атомарные элементы с весом, превышающим 2, не принадлежащие *пересечению путей* [3]. Предварительная обработка таких знаков предполагает необходимость избавления от указанных атомарных элементов (Рис.4.б).

Пусть $S_c = \{\alpha_{r_1}\}_{r_1=1}^{R_1}$ и существует подмножество атомарных элементов $S'_c \subset S_c, S'_c = \{\alpha_{r_2}\}_{r_2=1}^{R_2}, R_2 \leq R_1,$ такое, что $\forall r_2 \in [1, R_2]$ выполняется: $\alpha_{r_2} \in Z_k$ и

$\alpha_{r_2} \notin \bigcap_{n=1}^{C_2} L_n,$ тогда атомарные элементы $\alpha_{r_2} \in S'_c,$

$\forall r_2 \in [1, R_2]$ - будем считать *шумовыми*.

Представление знака в виде совокупности путей позволяет избавиться от необходимости учета шумовых атомарных элементов на этапе моделирования знаков, то

есть исключить из множества Z_k атомарные элементы со значением весовой функции, превышающим 2, не принадлежащие множеству путей знака Z_k .

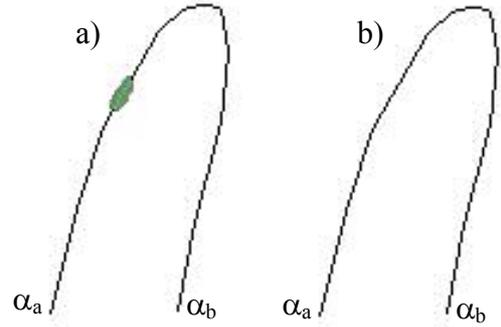


Рисунок 4. а) пример «зашумленного» знака; б) пример знака, представленного совокупностью путей (в данном случае – путем $L_1(\alpha_a, \alpha_b)$, при выделении которого были отброшены

Для исследования свойств знака произвольной природы на дискретном множестве необходимо изучить свойства подмножеств атомарных элементов, в совокупности составляющих данный знак.

Для связанных элементов, изображенных на рис. 1, введены следующие обозначения [3]: s_1 – горизонтальная связка; s_2 – вертикальная связка; s_3, s_4 – диагональные связки. На множестве пар связанных атомарных элементов определена функция η [3] следующим образом: $\eta(\alpha_a, \alpha_b) =$

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

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 $R' \subset Z_k, R' = \{(\alpha_{i_h}, \alpha_{i_{h+1}})_{m}\}_{h=1}^{N_1}, i_h \in [1, I], j_h \in [1, J], m \in \{1, 2, 3, 4\}$, 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 связок, преобладающих в данном ориентированном множестве. То есть, если в данном ориентированном множестве преобладают связки типа s_1 , то множество будем считать *горизонтально-ориентированным*. Если же преобладают связки типа s_2 , то – *вертикально-ориентированным*. Если s_3, s_4 – *лево- и право-ориентированным* соответственно.

Множество, состоящее из связок 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 (см. рис. 5).

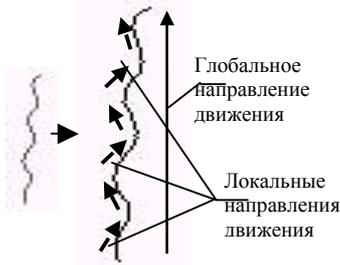


Рисунок 5. Пример локальных и глобальных движений на знаке

Однако выявление всех локальных движений может приводить к излишней детализации.

Например, знак, изображенный на рис. 5, представим в виде совокупности локальных движений, но всегда ли необходима такая подробная детализация? Очевидно, что нет. В случае знаков произвольной природы таким образом может быть представлен как знак, составленный из совокупности локальных движений, так и знак, начертанный одним движением, но, скажем, дрожащей рукой. В последнем случае локальные движения нужно не выявлять, а наоборот – считать избыточной информацией.

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.

3. МОДЕЛИРОВАНИЕ ЗНАКА ЭЛЕМЕНТАМИ Р-ГО УРОВНЯ ПРЕДСТАВЛЕНИЯ

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

Множество атомарных элементов A будем полагать атомарным уровнем и обозначать U_0 .

Обозначим через U_p – p -й уровень представления совокупности атомарных элементов, $p \in N$.

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 6. Она являет собой границу квадрата γ_p составленную из атомарных элементов: $\lambda_p = \gamma_p \setminus \gamma_{p-1}$.

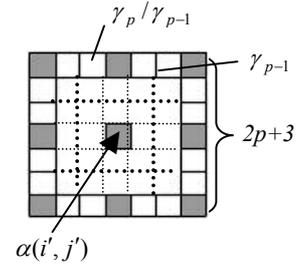


Рисунок 6. На рисунке представлен элемент 2-го уровня представления $\beta_2(\alpha(i', j'))$ с центром в атомарном элементе $\alpha(i', j')$

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 обоим 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 элемента p -го уровня представления $\beta_p(\alpha(i', j'))$ назовем элемент $\alpha(i', j')$.

Представим границу λ_p в виде совокупности секторов $G_1(\alpha(i', j')), \dots, G_g(\alpha(i', j'))$ так, как показано на рисунке 7.

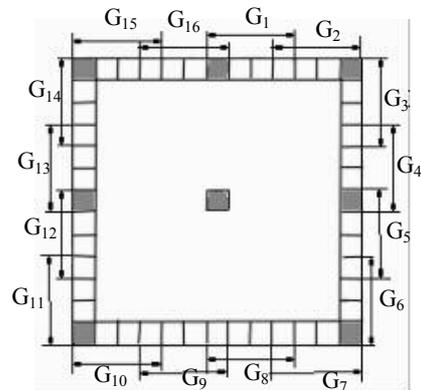


Рисунок 7. Пример разбиения граней квадрата на сектора G_1, \dots, G_{16} при $p=5$ и $q=2$

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. Обозначим Q_p множество размеров секторов, допустимых для уровня p :

$$Q_p = \{\vartheta \in [1, p+2] \mid \vartheta = [(p+2)/q] + \text{Rem}((p+2)/q);$$

$$\forall q \in [1, p+1]: \text{Rem}((p+2)/q) = 0 \text{ или } \text{Rem}((p+2)/q) = 1\},$$

где значение функции $\text{Rem}(x/y)$ равно остатку от деления некоторого целого числа x на некоторое целое число y .

Сектора G_1 и G_9 , G_2 и G_{10} , G_3 и G_{11} и т.д., приведенные на рисунке 7 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 представления 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 $G'_e(\alpha(i^t, j^t))$: $\alpha' \in G'_e(\alpha(i^t, j^t))$, $t \in [1, T]$ назовем *входным сектором*, а сектор $G''_v(\alpha(i^t, j^t))$: $\alpha'' \in G''_v(\alpha(i^t, j^t))$ - *выходным сектором* структуры $\beta_{pq}^t(\alpha(i^t, j^t))$ при $t \in [1, T]$, $e, v \in [1, \delta q]$.

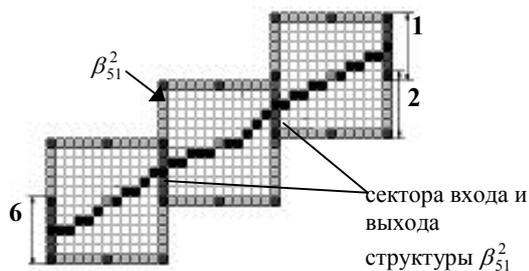


Рисунок 8. Пример элементарной стратегии, покрытой элементами представления 5-го уровня. 1,2,6 – некоторые номера секторов

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: $G'_{e_1}(\alpha(i^t, j^t)) \equiv G''_{v_1}(\alpha(i^t, j^t))$, где $e_1, v_1 \in [1, g]$, $t \in [1, T_1]$ 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 ϑ и уровень представления p , мы можем огрублять модель рассматриваемого знака, избавляясь таким образом от избыточной информации, или, наоборот, детализировать.

В рамках построения многоуровневой модели некоторого знака произвольной природы Z_k , параллельно на всех уровнях осуществляются покрытия ориентированных подмножеств атомарных элементов знака. Обозначим через M_k многоуровневую модель знака Z_k . 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: $Z_k \rightarrow M_k \rightarrow Z'_k \rightarrow M'_k$.

При восстановлении знака Z'_k по модели M_k , каждой элементарной стратегии ставится в соответствие отрезок «машинной» прямой [5], соединяющий начало и конец данной стратегии. Это делается с целью моделирования глобальных движений на знаке и абстрагирования от избыточной информации.

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

Пример на рисунке 9 иллюстрирует приведенный критерий корректности моделей.

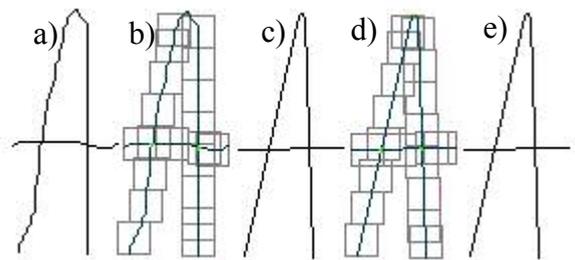


Рисунок 9. Пример корректно построенной модели:

- a) исходный знак - Z_k ;
- b) покрытие знака, построение модели M_k ;
- c) знак, восстановленный по модели - Z'_k ;
- d) покрытие знака, построение модели M'_k ;
- e) знак, восстановленный по модели M'_k

Нетрудно видеть, что знак, восстановленный по модели M_k , и знак, восстановленный по модели M'_k - тождественно совпадают. Это подтверждает, что их модели эквивалентны.

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.

Особенно важным является то, что приведенный метод построения моделей не использует никаких заранее

заданных ограничений, констант или эталонов. Множество прообразов концептов знака, сформированное посредством восстановления знака по модели, представляет собой широкий диапазон всевозможных его начертаний, причем модели всех знаков из этого множества будут эквивалентны. Любому другому знаку, модель которого будет также эквивалентна данной, будет поставлено в соответствие имя, установленное для исходного знака.

5. ЗАКЛЮЧЕНИЕ

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.

[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 conference on Optoelectronic Information-Energy Technologies – Винница, 2001.

[5] Шикин Е.В., Боресков А.В. *Компьютерная графика*. М: Диалог-МФТИ, 1995. – С.112-113.

Об авторе

Шевчук Елена Викторовна, аспирант Донецкого Национального Университета. E-mail: sssh@skif.net.

СПИСОК ЛИТЕРАТУРЫ