



ПОДХОД К ОПРЕДЕЛЕНИЮ ЭФФЕКТИВНЫХ ПРИЗНАКОВ И СИНТЕЗУ ОПТИМАЛЬНОГО ПОЛОСНО-РАЗДЕЛЯЮЩЕГО КЛАССИФИКАТОРА ДЛЯ ЭЛЕМЕНТОВ ДАКТИЛЬНО-ЖЕСТОВОЙ РЕЧИ

Аннотация. Приведены новые результаты для решения задачи определения эффективных признаков и синтеза оптимального полосно-разделяющего классификатора для элементов дактильной азбуки жестового языка глухих. Рассмотрены подходы к качественной оценке разделимости элементов дактильной азбуки для различных пространств признаков. Предложен алгоритм получения гиперплоскостного классификатора, разделяющего группы дактилем в пространстве признаков.

Ключевые слова: дактильно-жестовая речь, распознавание, гиперплоскостная классификация.

ВВЕДЕНИЕ И ПОСТАНОВКА ЗАДАЧИ

Дактильная азбука жестовой речи глухих является средством, с помощью которого конфигурацией пальцев руки в виде букв алфавита отображаются собственные значения языка (фамилии, названия населенных пунктов, вещества), т.е. те значения, для которых отсутствуют соответствующие обозначения в виде жестов [1]. Развитие информационных технологий и доступность коммуникационных устройств позволяет предложить средства, которые упростили бы общение между глухими и слышащими людьми [2–4].

Данные исследования направлены на разработку подходов, которые способствуют созданию доступных средств коммуникации для глухих людей для общения их со слышащими людьми.

Для решения задачи распознавания дактильной азбуки жестового языка из потокового видеоизображения проведены исследования с целью определить эффективные характеристические признаки представления кисти руки при воспроизведении дактилем [4]. Исследовалось исходное видеоизображение, которое содержит показанные человеком знаки дактильной азбуки и преобразованные изображения, позволяющие получить на выходе идентифицированный знак (букву) украинского алфавита. Идентификация учитывала воспроизведение дактиля людьми с разными размерами рук и при фиксации данного процесса с разных фокусных расстояний. На изображении фиксировались кисть руки и ее крайние точки — сверху, слева, снизу и справа. В качестве характеристических признаков были предложены следующие группы параметров: 1) угол между векторами, проведенными из центра к крайним точкам (x_2, x_3, x_4, x_5) ; 2) нормированная длина вектора, который проведен из центра к крайним точкам (x_6, x_7, x_8, x_9) ; 3) компактность, направленность, вытянутость, отношение ширины к высоте кисти руки $(x_{10}, x_{11}, x_{12}, x_1)$ [5]; 4) дескрипторы контурного анализа $(x_{13}, x_{14}, x_{15}, x_{16})$ [6]; 5) признак динамичности дактилемы (x_0) .

Анализ дактилем с амплитудой движения (динамических дактилем), признак которых — траектории перемещения координат центра масс руки, показал устойчи-

вую идентификацию дактилем Д, З и Б. Дактилемы Й, К, Ц, Щ, Ъ, Ѓ идентифицируем как неподвижные, но при этом следует учитывать наличие траектории (это позволяет разделить одинаковые по конфигурации руки дактилемы І(Ї), І(Ї), Ш(Щ)).

Из анализа статических дактилем следует, что признак x_1 (отношение ширины к высоте) позволяет устойчиво идентифицировать подмножество дактилем {Б, Ш, К, Ц}. Признак x_{10} (компактность) позволяет однозначно определить дактилему Т, а признак x_{12} (вытянутость) — дактилему Г.

Признаки других дактилем согласно классификации оказались слаборазделимыми и требуют иных подходов для их идентификации. Отметим, что предлагаемая проблематика наиболее коррелирована с направлениями исследований распознавания образов методами, подобными Support Vector Machine [7, 8]. Однако в отличие от этих важных и хорошо известных работ в настоящей статье предлагается при синтезе систем распознавания образов использовать средства оптимального синтеза как линейных, так и нелинейных преобразований на основании результатов, полученных в теории возмущений псевдообратных и проекционных операций [9–11].

На основании полосно-разделяющей классификации предложим подход к определению эффективных признаков и синтезу оптимального классификатора для дактильной азбуки жестовой речи глухих.

ОПИСАНИЕ ТЕОРЕТИЧЕСКИХ АСПЕКТОВ ПОДХОДА

Рассмотрим классификацию данных посредством построения дискриминантных функций [9–11].

Пусть $x = (x_1, \dots, x_m)^T$ — вектор признаков, характеризующий распознаваемый сигнал. Для описания обучающей последовательности в пространстве признаков, элементы которого соответствуют различным реализациям исследуемых сигналов, формулируется задача классификации в виде поиска дискриминантной функции $y = \Phi(x)$, $y \in R^1$, $x \in R^m$, так, что при $x(j) \in \Omega_x(1)$ имеет место неравенство $y(j) = \Phi(x(j)) \geq \Delta$, а при $x(j) \in \Omega_x(2)$ имеет место неравенство $y(j) = \Phi(x(j)) \leq -\Delta$. Здесь $\Omega_x(1) \subset \Omega_x = \{x : x(1), \dots, x(n)\}$ — подмножество векторов признаков, соответствующих сигналам первого класса, $\Omega_x(2) \subset \Omega_x$ — подмножество векторов признаков, соответствующих сигналам второго класса, $\Omega_x = \Omega_x(1) \cup \Omega_x(2)$, Δ — заданный порог, определяющий характер неточности правила классификации.

В случае выбора функции $\Phi(x)$ в классе линейных функций эта задача классификации заключается в выборе как самой функции, т.е. коэффициентов линейного преобразования, так и ее значений из допустимой области. Предполагается, что в целях формирования над компонентами x однородных линейных операций x_m будет равен единице.

Это значит, что необходимо определить для функции

$$y = a^T x \quad (1)$$

такой вектор коэффициентов a и такие значения $y(j)$, $j = \overline{1, n}$, для которых имели бы место соотношения

$$a^T x(j_1) = y(j_1) \geq \Delta \quad \forall x(j_1) \in \Omega_x(1), \quad (2)$$

$$a^T x(j_2) = y(j_2) \leq -\Delta \quad \forall x(j_2) \in \Omega_x(2). \quad (3)$$

Если таких значений вектора a и величин $y(j)$, $j = \overline{1, n}$, при которых выполняются соотношения (2), (3), не существует, тогда рассматривается задача при более качественном выборе пространства признаков.

Рассмотрим алгоритм синтеза линейных систем для классификации информации. Для обучающей последовательности $x(1), \dots, x(n)$ в пространстве призна-

ков осуществляется замена знаков компонент одного из множеств:

$$x'(j_1) = x(j_1) \quad \forall x(j_1) \in \Omega_x(I) \quad (4)$$

или

$$x'(j_2) = x(j_2) \quad \forall x(j_2) \in \Omega_x(2) \quad (5)$$

и рассматривается матрица

$$X' = (x'(I) \vdots \dots \vdots x'(n)) = \begin{pmatrix} x'^T(I) \\ \dots \\ x'^T(m) \end{pmatrix}.$$

В таких обозначениях сформулированная выше задача принимает следующую интерпретацию: определить вектор $a \in R^m$ и значения $y'(j) \geq \Delta$, $j = \overline{1, n}$, при которых имеет место отношение

$$a^T x'(j) = y'(j), \quad j = \overline{1, n}, \quad (6)$$

что эквивалентно матрично-векторному виду

$$X'^T a = y', \quad y' = (y'(I), \dots, y'(n))^T. \quad (7)$$

Для рассмотренных задач распознавания принадлежности исследуемых объектов к определенному классу допускается, что известен класс, к которому относятся векторы обучающей выборки $x(j)$, $j = \overline{1, n}$, в пространстве признаков.

Алгоритм синтеза системы классификации сводится к последовательному решению следующих подзадач:

- определить необходимые и достаточные условия линейной полосной разделимости точек $x(j)$, $j = \overline{1, n}$, на два класса;
- определить оптимальную по толщине в целом нелинейную полосу разделимости точек $x(j)$, $j = \overline{1, n}$, на два класса;
- упорядочить координаты $x_i(j)$, $i = \overline{1, m-1}$, $j = \overline{1, n}$; $x_m(j) = 1$ в соответствии с их влиянием на размеры (толщину) оптимальной полосы или на критерий линейной полосной разделимости;
- определить правило замены слабовлияющих (неинформативных) координат вектора признаков на более информативные из множества конкурирующих признаков.

Для оптимизации линейной полосной разделимости в пространстве признаков предлагается следующее.

При фиксированном $\Delta > 0$ имеется необходимое и достаточное условие существования решения задачи линейной полосной разделимости точек $x(j)$, $j = \overline{1, n}$, на два класса:

$$\min_{y \in D(\Delta)} y^T Z(X) y = y_*^T(\Delta) Z(X) y_*(\Delta) = 0, \quad (8)$$

где

$$X = (x(I) \vdots \dots \vdots x(n)) = \begin{pmatrix} x^T(I) \\ \dots \\ x^T(m) \end{pmatrix}, \quad Z(X) = I_n - X^+ X, \quad (9)$$

$$D(\Delta) = \{y: y = (y_1, \dots, y_n)^T, y_j \geq \Delta, j = \overline{1, n}\}.$$

При этом искомый вектор принимает значение

$$a(\Delta) = (X^T)^+ y_*(\Delta), \quad (10)$$

а толщина полосы δ , отделяющей множество точек $x(j)$, $j = \overline{1, n}$, от начала ко-

ординат определяется как

$$y_* = \delta = \Delta / (y_*^T (\Delta) R(X) y_* (\Delta))^{1/2}, \quad (11)$$

где $R(X) = X^+ (X^T)^+$.

Поскольку справедливо равенство $y_*(k\Delta) = ky_*(\Delta)$, то без ограничения общности можно положить, что $\Delta = 1$, $y_*(1) = y_*$.

Следовательно, максимальная толщина полосы достигается при значениях

$$y_{\text{opt}} = \arg \min_{y \in D} y^T R(X) y, \quad a_{\text{opt}} = (X^T)^+ y_{\text{opt}}, \quad (12)$$

где $D = \{y: y^T Z(X) y = y_*^T Z(X) y_* = 0, e_j^T y \geq 1, j = \overline{1, n}\}$.

Используя сингулярное разложение матриц

$$X = \sum_{j=1}^r u_j v_j^T \lambda_j, \quad Z(X) = I_n - \sum_{j=1}^r v_j v_j^T, \quad u_i^T u_j = \delta_{ij}, \quad v_i^T v_j = \delta_{ij}, \\ XX^T u_j = \lambda_j^2 u_j, \quad X^T X v_j = \lambda_j^2 v_j, \quad \lambda_1^2 \geq \dots \geq \lambda_r^2, \quad i, j = \overline{1, n},$$

и учитывая, что $y^T Z(X) y = 0$ для всех $y \in D$, имеют место следующие соотношения:

$$y = \sum_{i=1}^r a_i v_i, \quad \sum_{i=1}^r a_i e_j^T v_i \geq 1 \quad \forall j = \overline{1, n}, \quad y^T R(X) y = \sum_{j=1}^r a_j^2 \lambda_j^{-2}. \quad (13)$$

Таким образом, проблема нахождения оптимальной полосной разделимости сведена к решению задачи оптимизации квадратичной функции на выпуклом множестве

$$a_{\text{opt}} = \arg \min_{a = \{a: e_j^T (v_1 \dots v_r) a \geq 1, j = \overline{1, n}\}} a^T \text{diag}(\lambda_1^{-2}, \dots, \lambda_r^{-2}) a, \\ y_{\text{opt}} = (v_1 \dots v_r) a_{\text{opt}}, \\ a_{\text{opt}} = \sum_{j=1}^r u_j v_j^T \lambda_j^{-1} \cdot (v_1 \dots v_r) a_{\text{opt}} = (\lambda_1^{-1} u_1 \dots \lambda_r^{-1} u_r) a_{\text{opt}}, \\ \delta_{\text{opt}} = (a_{\text{opt}}^T \text{diag}(\lambda_1^{-2}, \dots, \lambda_r^{-2}) a_{\text{opt}})^{-1/2}. \quad (14)$$

На основании приведенных теоретических результатов получена вычислительная схема алгоритма оптимального синтеза линейных систем распознавания образов.

Входные данные. Имеем множество точек обучающей выборки $x(j) \in R^m$, $j = \overline{1, n}$, из них точки $x(j_k)$, $k = \overline{1, n_1}$, принадлежат первому классу, а $x(j_s)$, $s = \overline{1, n_2}$, — второму классу, $n_1 + n_2 = n$.

При выполнении условия линейной полосовой разделимости

$$\min_{y \in D} y^T Z(X) y = 0, \quad D = \{y: e_{i_k}^T y \geq 1, e_{j_s}^T y \leq -1, \quad k = \overline{1, n_1}, \quad s = \overline{1, n_2}\} \quad (15)$$

строится линейная дискриминантная функция $y = a^T x$, где вектор $a \in R^m$ определяется таким образом, чтобы выполнялось условие

$$a^T x(i_k) \geq 1, \quad k = \overline{1, n_1}, \quad a^T x(j_s) \leq -1, \quad s = \overline{1, n_2}. \quad (16)$$

Алгоритм поиска значения вектора a запишем, исходя из условия максимизации толщины полосы разделения:

1) из SVD-разложения матрицы $X = \sum_{j=1}^r u_j v_j^T \lambda_j$ находятся собственные числа λ_j и собственные векторы $(v_1 \dots v_r)$;

2) на основе λ_j строится квадратичная функция относительно параметров вектора a : $y^T R(X) y = \sum_{j=1}^r a_j^2 \lambda_j^{-2}$;

3) минимизируя $y^T R(X) y$ по a , находится a_{opt} :

$$a_{\text{opt}} = \arg \min_{a \in D_1} \sum_{j=1}^r a_j^2 \lambda_j^{-2},$$

$$D \in \{a: e_{i_k}^T (v_1: \dots: v_r) a \geq 1, e_{j_s}^T (v_1: \dots: v_r) a \leq -1, k = \overline{1, n_1}, s = \overline{1, n_2}\};$$

4) вычисляется

$$\begin{aligned} y_{\text{opt}} &= (v_1: \dots: v_r) a_{\text{opt}}, \\ a_{\text{opt}} &= (u_1: \dots: u_r) \text{diag}(\lambda_1^{-1}, \dots, \lambda_r^{-1}) a_{\text{opt}}, \\ \delta_{\text{opt}} &= (a_{\text{opt}}^T \text{diag}(\lambda_1^{-2}, \dots, \lambda_r^{-2}) a_{\text{opt}})^{-1/2}. \end{aligned}$$

Заметим, что для построения системы распознавания объектов, принадлежащих нескольким различным классам, используется принцип дихотомии.

Рассмотрим алгоритм работы принципа дихотомии.

1. На основе идеальных представителей классов конкретной предметной области строится дихотомное дерево классификации образов. На этом этапе определяются характеристические признаки классификации, их количество — размерность пространства признаков.

2. На основе уже реальных представителей классов обучающей выборки для каждой вершины дерева синтезируется оптимальный полосно-разделяющий классификатор согласно информации об объектах из обучающей последовательности.

3. Построенная таким образом система распознавания может уточняться оптимизацией всех параметров в ранее организованной структуре, т.е. при этом осуществляется реализация обучения синтезированной системы распознавания образов.

АНАЛИЗ ИНФОРМАТИВНОСТИ ПРИЗНАКОВ ДАКТИЛЬНОЙ АЗБУКИ

Для множества дактилем, которые нельзя разделить простыми способами, следует провести анализ информативности характеристических признаков. Для этого по каждой дактилеме получены соответствующие наборы признаков. С помощью линейного классификатора проводится попарная классификация в рамках каждой группы получения признаков.

В качестве критерия сравнения следует для каждой пары из n букв определять относительную величину: отношение между шириной полосы w и суммой расстояний до соответствующих сторон полосы: $D = \sum_{i=1}^n d_i^j, j = 1, 2$ (рис. 1, а) (для

анализа был определен центр первого и второго множеств ($D = \sum_{i,j} d_i^j / n$) и вы-

числено расстояние от центров множеств (O_1, O_2) к соответствующим полосам, как показано на рис. 1, б).

Для корректности сравнения предварительно для каждой группы признаков осуществлено нормирование векторов признаков на промежуток $[0, 1]$. Полученные результаты представлены в виде таблиц — попарная классификация всех дактилем алфавита в рамках каждой группы. Так, отношение между парами дактилем для четвертой группы (дескрипторы контурного анализа) приведены в табл. 1, т.е. отношение между дактилемами ϵ и γ : $r(\epsilon, \gamma) = w / D = 0,94273$.

Исследования показали, что если взять минимальную величину отношения ширины полосы классификатора к сумме расстояний центров кластеров к



соответствующим сторонам, то для легко делимых дактилем эта величина приблизительно равна $r = 0,3$. При этом наиболее эффективными характеристическими признаками являются дескрипторы контурного анали-

за [10]. Для других характеристических признаков разделение букв уменьшается, т.е. $r < 0,3$.

$$r(*, *) = \frac{w}{D} \rightarrow r'(*, *) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

СИНТЕЗ ОПТИМАЛЬНОГО ПОЛОСНО-РАЗДЕЛЯЮЩЕГО КЛАССИФИКАТОРА

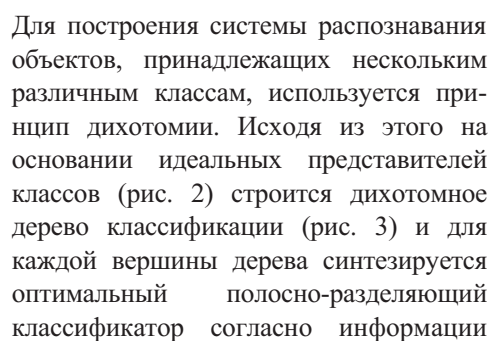


Рис. 2. Представление подмножеств слаборазделимых дактилем

об объектах из обучающей последовательности.

Из приведенного взаимного расположения (см. рис. 2) можно предложить следующее разбиение множества дактилем (см. рис. 3) сначала на два подмножества: {Е, С, Х, Ч, Я, В, Ё, Л, У} и {И, К, Н, Р, А, І, М, Ф, Ю} (рис. 4, а), а затем каждое из полученных подмножеств разбивается еще на два, т.е. на подмножества {Е, С, Х, Ч, Я} и {В, Ё, Л, У} (рис. 4, б) и на подмножества {И, К, Н, Р} и {А, І, М, Ф, Ю} (рис. 4, в).

Использование гиперплоскостной классификации требует реализации следующих шагов: 1) определение гиперплоскостей, разделяющих подмножества; 2) идентификация подмножества дактилем, к которому относится дактильный символ.

Ниже приведены алгоритмы, реализующие описанные шаги.

Для отнесения входящей дактилемы к одному из четырех подклассов дактилем выделим три гиперплоскости (см. рис. 4), разделяющие соответственно подмножества {Е, С, Х, Ч, Я, В, Ё, Л, У} и {И, К, Н, Р, А, І, М, Ф, Ю}; {Е, С, Х, Ч, Я} и {В, Ё, Л, У}, а также {И, К, Н, Р} и {А, І, М, Ф, Ю}.

Для получения указанных гиперплоскостей нужно выполнить следующие шаги:

1) на входе формируется матрица, состоящая из $k+1$ столбца (k столбцов содержат векторы одного класса, а один столбец — другого класса. Каждый столбец представляется вектором с n параметрами, значениями которых являются признаки $x_{49}, \dots, x_{52}$;

2) над элементами матрицы проводятся центрирование и нормирование на промежуток от 0 до 1;

3) дополняется каждый вектор $n+1$ компонентой, равной единице (строка свободных членов); полученную таким образом матрицу с $k+1$ столбцом и $n+1$ строкой обозначим A ;

4) находятся матрица A^T и произведение AA^T ;

5) находится обратная матрица $(AA^T)^{-1}$;

6) находится псевдообратная матрица к матрице A : $A^+ = A^T(AA^T)^{-1}$;

7) находится проекционная матрица $R = A^+ A^{+T}$;

8) находятся коэффициенты разделяющей гиперплоскости как решение уравнения $Y^T R Y \rightarrow 0$;

8.1) обозначим Y вектор размерности $k+l$, при этом k элементов этого вектора положим равными единице, а l элементов положим равными -1 ;

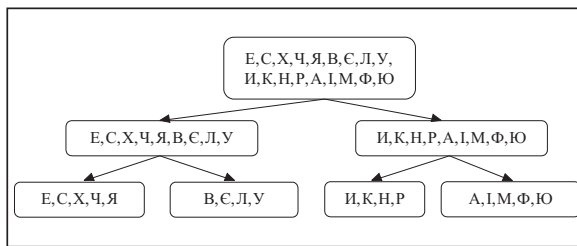


Рис. 3. Дихотомное дерево классификации

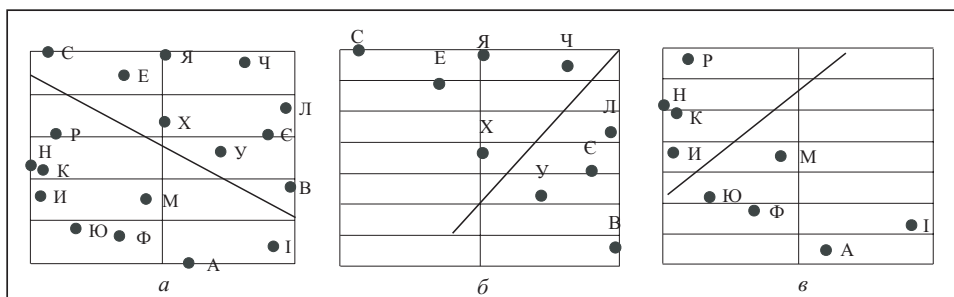


Рис. 4. Схема разбиения множеств дактилем полосно-разделяющим классификатором

8.2) найдем решение уравнения $Y^T R Y \rightarrow 0$, воспользовавшись произвольными минимаксными методами (например, поиском решения в Excel), т.е. будем искать решение (Y) такое, что минимизирует целевую функцию $Y^T R Y \rightarrow 0$ подбором компонент вектора Y со следующими ограничениями:

$$\begin{cases} y_i \geq 1, & i = \overline{1, \dots, k}, \\ y_j \leq -1, & j = \overline{k+1, \dots, k+l}; \end{cases}$$

9) для найденного Y определяются коэффициенты гиперплоскости a : $a = A^{+T} Y$.

Аналогично получаем коэффициенты трех гиперплоскостей: a_1, a_2 и a_3 .

Далее с помощью предварительно полученных коэффициентов трех гиперплоскостей ($a_i, i = \overline{1, 3}$) рассмотрим механизм отнесения модели входной дактилемы к одному из соответствующих подмножеств. Для этого последовательно определяем решение $q = a_i^T x$, где x — модель входной дактилемы ($x_{13}, \dots, x_{16}$). Если результат больше нуля, то x относится к первому классу, в противном случае — ко второму классу. Таким образом, определяем подмножество, к которому относится модель входной дактилемы.

ЗАКЛЮЧЕНИЕ

В настоящей статье предложен подход к определению эффективных характеристических признаков для распознавания элементов дактильной азбуки жестовой речи глухих. С помощью предложенного подхода произведен синтез оптимального полосно-разделяющего классификатора для элементов дактильной азбуки; значения признаков этих элементов достаточно схожи.

Дальнейшие исследования будут направлены на реализацию системы распознавания дактильной азбуки жестовой речи глухих.

СПИСОК ЛИТЕРАТУРЫ

1. Formational units in sign languages // Sign Language Typology. Ser. 3 / R. Channon, H. van der Hulst (Eds.). — Nijmegen (Netherlands); Berlin: Shara Press/Mouton de Gruyter, 2011. — 352 p.
2. Augmentative and alternative communication (AAC). — <http://www.asha.org/public/speech/disorders/AAC/>.
3. Кривонос Ю.Г., Крак Ю.В., Бармак А.В., Шкильнюк Д.В. Конструирование и идентификация элементов жестовой коммуникации // Кибернетика и системный анализ. — 2013. — № 2. — С. 3–8.
4. Кривонос Ю.Г., Крак Ю.В., Барчукова Ю.В., Троценко Б.А. Параметризация движений кисти руки человека для моделирования дактилем // Проблемы управления и информатики. — 2011. — № 6. — С. 134–143.
5. Курсы лаборатории компьютерной графики. — <http://courses.graphicon.ru/>.
6. Введение в контурный анализ / Я.А. Фурман, А.В. Кревецкий, А.К. Передрепеев и др.; под ред. Я.А. Фурмана. — М.: Физматлит, 2003. — 592 с.
7. Vapnik V.N. Statistical learning theory // Adaptive and learning systems for signal processing, communications, and control. — New York: Wiley, 1998. — 736 p.
8. Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P. Support vector machines (Sect. 16.5) / Numerical Recipes: The Art of Scientific Computing (3rd ed.). — New York: Cambridge University Press, 2007. — P. 883–888.
9. Кириченко Н.Ф. Аналитическое представление возмущений псевдообратных матриц // Кибернетика и системный анализ. — 1997. — № 2. — С. 98–107.
10. Кириченко Н.Ф., Крак Ю.В., Полищук А.А. Псевдообратные и проекционные матрицы в задачах синтеза функциональных преобразователей // Кибернетика и системный анализ. — 2004. — № 3. — С. 116–129.
11. Кириченко Н.Ф., Кудин Г.И. Анализ и синтез систем классификации сигналов средствами возмущений псевдообратных и проекционных операций // Кибернетика и системный анализ. — 2009. — № 3. — С. 47–57.

Поступила 06.11.2015