

## Разработка алгоритма синтеза оптимальных эталонов на основе метода главных компонент<sup>1</sup>

А. В. Полиев

Московский физико-технический институт (государственный университет)  
117303, Москва, ул. Керченская, 1А, корп. 1

e-mail: poliyev@gmail.com

*Аннотация.* Работа посвящена проблеме автоматического распознавания речевых команд, используемых для аудиоинтерфейсов современных авиационных систем управления. Предлагается алгоритм оптимизации эталона, в котором существенное снижение размерности оптимизационной задачи достигается на основе применения метода главных компонент. Оптимальный эталон формируется путем разложения на главные компоненты и дальнейшей оптимизацией коэффициентов разложения на обучающей выборке с помощью метода покоординатного спуска. В работе представлено описание алгоритма вычисления главных компонент для решаемой задачи, а также алгоритм получения оптимального эталона. Приведены положительные результаты тестирования предложенного подхода на примерах распознавания слов естественного русского языка.

*Ключевые слова:* автоматическое распознавание речи, эталон, метод главных компонент, метод покоординатного спуска.

### 1. Введение

Надежная и рациональная организация человеко-машинного взаимодействия является одной из важных задач современной техники [1]. В последние годы активно повышается необходимость использования аудиоинтерфейсов во вновь разрабатываемых авиационных системах управления, так как их успешное внедрение позволит повысить безопасность полетов и снизит нагрузку на пилотов [2]. Так, на самолете Eurofighter Typhoon с 2005 г. эксплуатируется дикторозависимая система от фирмы Eurofighter GmbH, EU, основанная на сравнении с эталонами. Речевые сигналы можно использовать не только для организации человеко-машинного интерфейса в формате речевых команд, но и для оценки степени утомления оператора [3, 4] или состояния его слуха [5].

Наиболее важным требованием, предъявляемым к речевым интерфейсам авиационных систем, является высокая вероятность правильного распознавания слов

---

<sup>1</sup> Работа выполнена при поддержке Российского фонда фундаментальных исследований (РФФИ), проект 15-08-06946-а.

даже в условиях сильных шумов и помех. Также при создании бортового программного комплекса важным моментом является временная сложность алгоритма.

Для речевого управления бортовым оборудованием хорошую результативность показывает способ распознавания речевых команд на основе сравнения с эталоном [6]. В данной работе предлагается новый способ получения эталона и его улучшения с помощью метода главных компонент. В экспериментальной части приводятся результаты распознавания слов с помощью новых эталонов. В заключении, рассматривается способ упрощения алгоритма с сохранением результатов с целью уменьшения времени его работы.

## 2. Постановка задачи

Традиционно применяемый метод автоматического распознавания речевых команд с помощью эталона использует спектрально-временное преобразование записи входного слова [7]. Вариант преобразования, используемый в данной статье, заключается в получении параметрического портрета и состоит в следующем. В начале речевой сигнал разделяется на равные интервалы — временные отрезки длиной 10–30 мс, и далее каждый из интервалов разделяется на 30–40 частотных полос [8]. После этого используются такие стандартные процедуры цифровой обработки сигналов, как усиление высокочастотных составляющих сигнала, взвешивание интервалов окном Ханна, быстрое преобразование Фурье, осреднение по частотам и логарифмирование спектральных плотностей [9].

В итоге получается, что в качестве параметров речевого сигнала рассматриваются логарифмы оценок спектральных плотностей в функции дискретных значений частоты, рассчитанные на скользящем временном интервале. Эти параметры, составленные в единую матрицу, называются параметрическим портретом слова. Столбцы этого портрета характеризуют спектральный состав речевого сигнала на каждом временном интервале.

Далее из имеющихся параметрических портретов записей речевых команд составляется эталонный параметрический портрет. Самый простой способ получения эталона — это усреднение параметрических портретов нескольких схожих записей. Затем параметрический портрет распознаваемого слова сравнивается с параметрическим портретом эталона. Это сравнение можно осуществить по критерию максимума коэффициента корреляции векторов, некоторым образом полученных из исходных матриц параметрического портрета. После сравнения распознаваемого слова со всеми эталонами выбирается тот эталон, с которым коэффициент корреляции оказался максимальным. Речевая команда, соответствующая этому эталону, признается результатом распознавания данного алгоритма.

### 3. Метод главных компонент

Метод главных компонент (МГК) применяется для снижения размерности пространства наблюдаемых векторов, не приводя к существенной потере информативности. Главные компоненты представляют собой ортогональную систему координат, в которой дисперсии компонент характеризуют их статистические свойства.

Пусть дан исходный набор векторов  $X$  линейного пространства  $L^p$ . Применение метода главных компонент позволяет перейти к базису пространства  $L^{p'}$ , ( $p' \ll p$ ) такому, что первая компонента (первый вектор базиса) соответствует направлению, вдоль которого дисперсия векторов исходного набора максимальна. Направление второй компоненты (второго вектора базиса) выбрано таким образом, чтобы дисперсия исходных векторов вдоль него была максимальной при условии ортогональности первому вектору базиса. Аналогично определяются остальные векторы базиса. В результате направления векторов базиса выбраны так, чтобы максимизировать дисперсию исходного набора вдоль первых компонент, называемых главными компонентами. Получается, что основная изменчивость векторов исходного набора векторов представлена несколькими первыми компонентами и появляется возможность, отбросив оставшиеся (менее существенные) компоненты, перейти к пространству меньшей размерности [10].

Пусть имеется многомерное случайное наблюдение  $X = (x^{(1)} \dots x^{(p)})^T$ . Задача сводится к переходу от числа признаков  $p$  к  $p'$ . Эту задачу можно решить, если определить всевозможные линейные ортогональные нормированные комбинации показателей  $z^{(j)}(X) = c_{j1}(x^{(1)} - \mu^{(1)}) + \dots + c_{jp}(x^{(p)} - \mu^{(p)})$ , где  $[\mu^{(1)}, \dots, \mu^{(p)}]^T$  — вектор средних для переменной  $X$ . В качестве меры информативности  $p'$ -мерной системы показателей  $(z^{(1)}(X), \dots, z^{(p')}(X))$  принимается выражение  $I_{p'} = \frac{Dz^1 + \dots + Dz^{p'}}{Dx^1 + \dots + Dx^p}$ , где  $Dz$  — операция вычисления дисперсии случайной величины. Можно показать, что соотношения для определения всех  $p$  главных компонент матрицы  $X$  могут быть представлены в виде  $Z = LX$ , где  $Z = (z^{(1)}, \dots, z^{(p)})^T$ ,  $X = (x^{(1)}, \dots, x^{(p)})^T$ , а матрица  $L$  состоит из строк  $l_j = (l_{j1}, \dots, l_{jp})$ ,  $j = \overline{1, p}$ , являющихся собственными векторами ковариационной матрицы  $\Sigma$  случайной величины  $X$ . Этим собственным векторам соответствуют собственные числа  $\lambda_j$ ,  $j = \overline{1, p}$ .

Основные свойства главных компонент.

1. Матрица  $L$  является ортогональной, т. е.  $LL' = L'L = I$ , где  $I$  — единичная матрица.
2. Ковариационная матрица вектора главных компонент:

$$\Sigma_Z = L \Sigma L' = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_p \end{pmatrix}.$$

3. Сумма дисперсий исходных признаков равна сумме дисперсий всех главных компонент.

Критерий информативности метода:

$$I_{p'}(Z(X)) = \frac{\lambda_1 + \lambda_2 + \dots + \lambda_{p'}}{\lambda_1 + \lambda_2 + \dots + \lambda_p},$$

где  $\lambda_1, \lambda_2, \dots, \lambda_p$  — собственные числа ковариационной матрицы  $\Sigma$  вектора  $X$ , расположенные в порядке убывания. Такой критерий является основой при вынесении решения о том, сколько последних главных компонент можно без особого ущерба изъять из рассмотрения, сократив тем самым размерность исследуемого пространства.

#### 4. Разработка и проверка алгоритмов разложения спектрального портрета слова на главные компоненты

Пусть имеется  $M$  параметрических портретов [11] различных реализаций одного слова  $\{x_{ij}(k)\}$ ,  $k = \overline{1, M}$ ;  $l = \overline{1, N_l}$ ;  $j = \overline{1, N_j}$ . Преобразуем для каждого  $k$  матричный портрет в одномерный массив с числом элементов  $i = \overline{1, P}$ ,  $P = N_j \cdot N_l$ . Итак, имеем  $M$  векторов размерности  $P$  каждый  $\{x_{ik}\}$ ,  $k = \overline{1, M}$ ;  $i = \overline{1, P}$ . В явном виде

$$x_1 = \begin{bmatrix} x_{11} \\ x_{21} \\ \dots \\ x_{p1} \end{bmatrix}, x_2 = \begin{bmatrix} x_{12} \\ x_{22} \\ \dots \\ x_{p2} \end{bmatrix}, \dots, x_M = \begin{bmatrix} x_{1M} \\ x_{2M} \\ \dots \\ x_{pM} \end{bmatrix}. \quad (1)$$

Объединим эти  $M$  векторов в матрицу размерности  $p \times M$ :

$$X = [x_1 \ x_2 \ \dots \ x_M] = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1M} \\ x_{21} & x_{22} & \dots & x_{2M} \\ \dots & \dots & \dots & \dots \\ x_{p1} & x_{p2} & \dots & x_{pM} \end{bmatrix}. \quad (2)$$

Вычислим матрицу корреляционных моментов векторов  $x_1, x_2, \dots, x_M$  размерности  $M$ :

$$K_x = X^T X = \begin{bmatrix} x_1^T x_1 & x_1^T x_2 & \dots & x_1^T x_M \\ x_2^T x_1 & x_2^T x_2 & \dots & x_2^T x_M \\ \dots & \dots & \dots & \dots \\ x_M^T x_1 & x_M^T x_2 & \dots & x_M^T x_M \end{bmatrix}. \quad (3)$$

Заметим, что в матрице  $K_x$  каждый элемент есть скалярное произведение соответствующих векторов и матрица  $K_x$  симметричная. Для нее можно вычислить  $M$  собственных чисел  $\lambda_1, \lambda_2, \dots, \lambda_M$  и соответствующих собственных векторов  $l_1, l_2, \dots, l_M$  размерности  $M$ . Собственные числа можно упорядочить по убыванию  $\lambda_1 > \lambda_2 > \dots > \lambda_M$ .

Первая главная компонента  $z_1$  определяется как линейная комбинация исходных векторов  $x_1, x_2, \dots, x_M$ , взятых с коэффициентами, равными элементам собственного вектора  $l_1^T = [l_{11} \ l_{21} \ \dots \ l_{M1}]$ :

$$z_1 = l_{11}x_1 + l_{21}x_2 + \dots + l_{M1}x_M = \sum_{i=1}^M l_{i1}x_i. \quad (4)$$

Аналогично главные компоненты  $j = \overline{2, M}$  вычисляются по формуле

$$z_j = \sum_{i=1}^M l_{ij}x_i. \quad (5)$$

Согласно теории, изложенной в части 3, суммарная дисперсия главных компонент равна суммарной дисперсии исходных векторов, т. е. преобразование не изменяет энергию сигналов. Дисперсия главной компоненты  $z_j$  равна соответствующему собственному числу  $\lambda_j$ . Главные компоненты обладают свойством взаимной ортогональности. Смысл применения главных компонент состоит в том, что поведение системы определяется в основном несколькими первыми главными компонентами  $z_j, j = \overline{1, M'}$ . Это позволяет уменьшить размерность задачи и рассматривать  $M'$  главных компонент вместо  $M$  исходных векторов. Для оценки погрешности, вносимой при переходе от  $M$  исходных векторов к  $M'$  главных компонент, можно использовать величину

$$I_p = \frac{\lambda_1 + \lambda_2 + \dots + \lambda_{M'}}{\lambda_1 + \lambda_2 + \dots + \lambda_M}. \quad (6)$$

## 5. Разработка алгоритма синтеза оптимальных эталонов на основе метода главных компонент

В работе тестирование метода производится на примере простой задачи распознавания 3 слов: «пилотаж», «масштаб» и «навигация», которые произносятся 10 различными дикторами в условиях без шума и 4 дикторами в условиях с шумом в наушниках [12–14].

Шум в наушниках реализуется следующим образом. Диктор одевает наушники, в которые подается шум заданной громкости, при этом диктор плохо слышит свой голос, поскольку в наушниках не подключена обратная связь. Это используется для того, чтобы в экспериментальных условиях проверить степень изменения речи диктора в зависимости от уровня шума. Шум не записывается вместе с произнесенным словом, а только подается в наушники. Это сделано для того, чтобы оценить не влияние шума на процесс распознавания, а на изменение речи диктора в условиях шума.

Пусть эталоны для распознавания каждого из трех слов сформированы как усредненный параметрический портрет всех реализаций. Самая простая оценка качества этих эталонов есть количество ошибок распознавания на заданном множестве дикторов. В качестве непрерывной меры оценки качества распознавания можно рассмотреть нижнюю границу значения  $Z$ -преобразования Фишера коэффициента корреляции эталона с распознаваемым словом. Представим эталон как линейную комбинацию, например, первых 6 главных компонент и постоянной составляющей с некоторыми коэффициентами:

$$e_{syn} = k_0 Z_0 + k_1 Z_1 + \dots + k_6 Z_6. \quad (7)$$

Задача поиска оптимального эталона сводится к подбору коэффициентов  $k_0, \dots, k_6$ , таких, которые будут удовлетворять критерию:  $N_{ошибок} \rightarrow min$ , либо же, что эквивалентно, значение  $Z$  коэффициента корреляции  $\rightarrow max$ . Выбор коэффициентов  $k_0, \dots, k_6$  можно произвести, используя различные численные методы.

При построении 6 главных компонент используются речевой материал 10 дикторов без шума, причем это множество дикторов не пересекается с множеством 4 дикторов записей с шумом в наушниках. В качестве начального приближения коэффициентов предлагается выбрать значения коэффициентов разложения усредненного эталона (взятое как среднее по 10 выбранным эталонам) по полученным 6 главным компонентам. В качестве метода оптимизации  $Cr = f(k_0, \dots, k_6)$  будет использоваться метод покоординатного спуска [15].

Закон изменения коэффициентов разложения при  $k_j > 0.001$ :  $k_{j+1} = k_j + l\Delta k_j$ ,  
 $\Delta k_j = 0.01 |k_j|$ ,  $l = 0, \pm 1, \pm 2, \pm 3, \pm 5, \pm 10, \pm 15, \pm 25, \pm 50, \pm 100, \pm 200$ . Если

$k_j \leq 0.001$ , то используется закон изменения с постоянным шагом:

$$k_{j+1} = k_j + 0.001\Delta k_j, \quad \Delta k_j = 0, \pm 1, \pm 2, \pm 3, \dots, \pm 10.$$

Критерий остановки:

$$|Cr_{i+1} - Cr_i| \leq \varepsilon, \quad \varepsilon = 0.02 |Cr_i|. \quad (8)$$

Подбор коэффициентов производится для каждого слова в отдельности, при этом коэффициенты разложения двух других слов остаются неизменными и равными начальному приближению.

В ходе оптимизации методом покоординатного спуска вычисление осредненного значения нижней границы  $Z$  коэффициента корреляции проводится с использованием обучающей выборки — реализациях слов диктора Н-в с шумом в наушниках 80 дБ, в то время как исходное разложение на главные компоненты было получено по записям без шума.

## 6. Результаты

Для проверки работоспособности оптимизированного эталона было проведено распознавание для всех 4 дикторов с различными уровнями шума в наушниках: 80 дБ и 90 дБ. Результаты применения оптимизированного на тестовой выборке эталона показывают улучшение при распознавании реализаций не только для слов диктора Н-в, но также и для других дикторов, не входящих в обучающую выборку. Общее количество ошибок для 3 слов до оптимизации равнялось 3.15%, а после оптимизации уменьшилось до 0.89%. Результаты представлены в табл. 1 и 2.

Таблица 1. Результаты распознавания записей с шумом в наушниках 80 дБ на обычном (1) и оптимизированном (2) эталонах.

Диктор	Слово	Ошибки 1	Ошибки 2
Бурлак	пилотаж	6	0
	масштаб	0	0
	навигация	0	8
Габдрахманов	пилотаж	3	0
	масштаб	0	0
	навигация	0	0
Набатчиков	пилотаж	8	0
	масштаб	0	0
	навигация	0	0
Финаев	пилотаж	9	0
	масштаб	0	0
	навигация	0	0

Таблица 2. Результаты распознавания записей с шумом в наушниках 90 дБ на обычном (1) и оптимизированном (2) эталонах.

Диктор	Слово	Ошибки 1	Ошибки 2
Бурлак	пилотаж	7	0
	масштаб	0	0
	навигация	0	1
Габдрахманов	пилотаж	3	0
	масштаб	0	0
	навигация	0	2
Набатчиков	пилотаж	12	0
	масштаб	0	0
	навигация	0	0
Финаев	пилотаж	4	0
	масштаб	0	0
	навигация	8	4

Также было протестировано использование ограниченного числа реализаций слов, используемых при построении оптимального эталона. Кроме того, была проверена возможность использования заранее заданного количества итераций оптимизации эталона вместо использования критерия останова (8). Эксперимент проводился для одного из дикторов на обоих вариантах шума в наушниках. Результаты данных экспериментов приведены в табл. 3.

Таблица 3. Результаты распознавания записей с шумом в наушниках для различного числа используемых реализаций слова и итераций при оптимизации для 3 слов (пилотаж-масштаб-навигация)

Реализации	Итерации	Финаев 80 дБ	Финаев 90 дБ
1	0	6-0-0	1-0-2
	1	3-0-0	1-0-2
	3	2-0-0	0-0-2
	10	3-0-0	0-0-0
	30	3-0-0	0-0-0
3	0	6-0-0	1-0-2
	1	3-0-0	1-0-2
	3	2-0-0	0-0-2
	10	3-0-0	0-0-0
	30	3-0-0	0-0-0
10	0	6-0-0	1-0-2
	1	3-0-0	1-0-2
	3	2-0-0	0-0-2
	10	3-0-0	0-0-0
	30	3-0-0	0-0-0

Из результатов эксперимента можно сделать вывод, что для процедуры получения оптимального эталона достаточно 1 реализации слова. Использование большего числа реализаций не приводит к улучшению результатов, часто достаточно



провести всего 10 итераций оптимизации для получения устойчивых результатов. Важно отметить что использование большего числа реализация слов при получении оптимального эталона линейно увеличивает время работы алгоритма, поэтому полученный вывод очень значим для сокращения времени оптимизации.

## 7. Заключение

Эталон, полученный с помощью оптимизации коэффициентов при главных компонентах, показал значительно меньшее число ошибок при распознавании большинства записей. Этот метод оказался эффективным способом улучшения результатов распознавания слов [8] наряду с использованием разбиения слов на однородные части [16], которые также дают улучшение результатов для данной задачи. В отдельных случаях возросшее количество ошибок, вероятнее всего, связано с тем, что алгоритм покоординатного спуска попадает в локальный экстремум вместо глобального, поэтому целесообразно перейти к более эффективному алгоритму численной оптимизации. Также был получен вывод о том, что для получения приемлемых результатов достаточно использовать только одну реализацию слова и проводить всего 10 итераций при получении оптимального эталона, что заметно сокращает время работы программы.

## Литература

- [1] Себряков Г. Г. Проблемы проектирования полуавтоматических систем наведения летательных аппаратов // *Вестн. компьютерных и информационных технологий*. 2007. № 10. С. 2–7.
- [2] Бондарос Ю. Г., Маковкин К. А., Чучупал В. Я. Система распознавания команд речевого интерфейса пилота для интегрированной модульной авионики // *Вестн. компьютерных и информационных технологий*. 2007. № 4. С. 2–13.
- [3] Бондарос Ю. Г., Иванов А. И., Тищенко А. А. Определение степени утомления оператора по экспоненте Ляпунова для его речи // *Вестн. компьютерных и информационных технологий*. 2010. № 6. С. 22–30.
- [4] Иванов А. И. и др. Исследование характеристик речи дикторов-пилотов с нарушениями слуха в интересах создания системы речевого управления летательным аппаратом // *Медицина труда и промышленная экология*. 2014. № 11. С. 40–45.
- [5] Иванов А. И. и др. Экспериментальное исследование взаимосвязи характеристик речи и состояния слуха у летного состава с нейросенсорной тугоухостью: пилотное исследование // *Медицина труда и промышленная экология*. 2017. № 5. С. 58–62.

- [6] *Корсун О. Н., Габдрахманов А. Ш., Михайлов Е. И., Нахаев М. З., Тулекбаева А. К.* Алгоритм автоматического распознавания речевых команд, инвариантный к изменению языка // *Мехатроника, автоматизация, управление*. 2015. № 9. С. 599–604.
- [7] *Rabiner L., Juang B.H.* Fundamentals of Speech Recognition. — Englewood Cliffs, New Jersey : Prentice-Hall International, Inc. 1993. 507 p.
- [8] *Колоколов А. С., Любинский И. А.* Сравнительное исследование нескольких способов кратковременного частотного анализа речевого сигнала // *Автоматика и телемеханика*. 2015. № 10. С. 144–151.
- [9] *Колоколов А. С.* Обработка сигнала в частотной области при распознавании речи // *Проблемы управления*. 2006. № 3. С. 13–18.
- [10] *Айвазян С. А., Бухтитабер В. М., Енюков И. С., Мешалкин Л. Д.* Прикладная статистика: Классификация и снижение размерности; под ред. С .А. Айвазяна. — М. : Финансы и статистика, 1989. 607 с.
- [11] *Савченко Л. В.* Алгоритм пофонемного распознавания устной речи на основе метода нечеткого фонетического кодирования-декодирования слов // *Информационно-управляющие системы*. 2014. № 1. С. 23–31.
- [12] *Корсун О. Н., Габдрахманов А. Ш.* Помехозащищенный алгоритм речевого управления бортовым оборудованием самолета // *Вестн. компьютерных и информационных технологий*. 2012. № 4. С. 3–7.
- [13] *Корсун О. Н., Иванов А. И., Филатов В. Н., Красавин И. В., Чучупал В. Я.* Методика экспериментального исследования влияния пилотажной перегрузки на характеристики речи в целях создания речевого интерфейса бортового оборудования самолетов // *Вестн. компьютерных и информационных технологий*. 2012. № 5. С. 3–7.
- [14] *Schmidt-Nielsen A., Marsh E., Tardeli J., Gatewood P., Kreamer E., Tremain T., Cieri C., Wright J.* Speech in Noisy Environments (SPINE) Evaluation Audio. — Linguistic Data Consortium, 2000
- [15] *Алексеева Е. В., Кутненко О. А., Плясунов А. В.* Численные методы оптимизации : учеб. пособие. — Новосибирск : Новосиб. ун-т, 2008.
- [16] *Корсун О. Н., Полиев А. В.* Разработка метода анализа фонетически однородных частей слов естественного языка // В Сб. тр. конф.: Человеческий фактор в сложных технических системах и средах. — СПб. : Межрегиональная эргономическая ассоциация, 2016. С. 370–377.

**Автор:**

*Александр Владимирович Полиев* — аспирант кафедры управляющих и информационных систем, Московский физико-технический институт (государственный университет)

## Principal component analysis in the problem of an optimal word pattern synthesis

A. V. Poliyev

Moscow Institute of Physics and Technology  
Kerchenskaya str., 1A, building 1, Moscow, Russia, 117303

e-mail: poliyev@gmail.com

*Abstract.* The speech recognition techniques that are currently used have a significant number of errors, especially for samples with noise. The algorithm for creating an optimal pattern, which increases the probability of correct recognition, is developed and presented in this work. An optimal pattern is formed by the decomposition of a usual pattern to principal components. The coefficients of the first principal components are used in the further optimization on the training set with the coordinate descent method. Applied mathematical methods will be also described. The paper has a description of the pattern to principal components decomposition algorithm and the algorithm of generating of an optimal pattern. In the end of the work the results of word recognition for natural language Russian and conclusion describing the outcomes of this research are presented.

*Key words:* speech recognition, pattern, principal component analysis, coordinate descent algorithm

### References

- [1] *Sebryakov G. G. (2007) Vestn. komp'yuternykh i informatsionnykh tekhnologiy, 10:2–7.*
- [2] *Bondaros Y. G., Makovkin K. A., Chuchupal V. Y. (2007) Vestn. komp'yuternykh i informatsionnykh tekhnologiy, 4:2–13.*
- [3] *Bondaros Y. G., Ivanov A. I., Tishchenko A. A. (2010) Vestn. komp'yuternykh i informatsionnykh tekhnologiy, 6:22–30.*
- [4] *Ivanov A. I. (2014) Meditsina truda i promyshlennaya ekologiya, 11:40–45.*
- [5] *Ivanov A. I. (2017) Meditsina truda i promyshlennaya ekologiya, 5:58–62.*
- [6] *Korsun O. N., Gabdrakhmanov A. S. et al. (2015) Mekhatronika, avtomatizatsiya, upravleniye, 9:599–604.*
- [7] *Rabiner L., Juang B. H. (1993) Fundamentals of Speech Recognition. Englewood Cliffs, New Jersey: Prentice Hall International, Inc.*
- [8] *Kolokolov A. S., Lyubinskiy I. A. (2015) Automation and Remote Control, 76(10):1828–1833.*
- [9] *Kolokolov A. S. (2006) Problemy upravleniya, 3:13–18.*

- [10] Ayvazyan S. A., Bukhtstaber V. M., Yenyukov I. S., Meshalkin L. D. (1989) *Prikladnaya statistika: Klassifikatsiya i snizheniye razmernosti*; S.A. Ayvazyan (Eds.). Moscow, Finansy i statistika, 1989.
- [11] Savchenko L. V. (2014) *Informatsionno-upravlyayushchiye sistemy*, 1:23–31.
- [12] Korsun O. N., Gabdrakhmanov A. S. (2012) *Vestn. komp'yuternykh i informatsionnykh tekhnologiy*, 4:3–7.
- [13] Korsun O. N., Ivanov A. I., Filatov V. N., Krasavin I. V., Chuchupal V. Y. (2012) *Vestn. komp'yuternykh i informatsionnykh tekhnologiy*, 5:3–7.
- [14] Schmidt-Nielsen A., Marsh E., Tardeli J. et al. (2000) *Speech in Noisy Environments (SPINE) Evaluation Audio*. Linguistic Data Consortium.
- [15] Alekseyeva Ye. V., Kutnenko O. A., Plyasunov A. V. (2008) *Chislennyye metody optimizatsii*: Novosibirsk, Novosib. un-t.
- [16] Korsun O. N., Poliyev A. V. (2016) *Razrabotka metoda analiza foneticheskikh odnorodnykh chastey slov yestestvennogo yazyka*. In Proc. Conf. Chelovecheskiy faktor v slozhnykh tekhnicheskikh sistemakh i sredakh, p. 370–377.