

Е.Е. Федоров, д.т.н., доц.,
И.В. Ярош, ст. преп.
Т.А. Черняк, асс.

Донецкий национальный технический университет
fedorovee75@mail.ru

Метод классификации вокальных звуков речи на основе саундлетов и модифицированной сети Хемминга

В статье предложен метод классификации вокальных звуков речи, который базируется на саундлетах и отображениях между ними, а также на модифицированной нейронной сети Хемминга, и позволяет учитывать структуру квазипериодического сигнала и сопоставлять образцы вокальных звуков речи разной длины. На основе саундлетов и отображений между ними разработаны метод создания образцов и метод формирования опорных образцов, который также использует комбинированную кластеризацию (сочетание итеративной и иерархической кластеризации). Предложенные методы могут использоваться в интеллектуальных компьютерных системах для повышения качества распознавания речи и идентификации диктора.

Ключевые слова: классификации вокальных звуков речи, сеть Хемминга, саундлеты, комбинированная кластеризация, метод создания образцов, метод формирования опорных образцов.

Введение

На сегодняшний день актуальной является разработка программных компонент предназначенных для распознавания речи человека, идентификации диктора, синтеза речи человека и др., и используемых в интеллектуальных компьютерных системах.

В основе данной задачи лежит проблема построения эффективных методов, обеспечивающих высокую скорость обучения модели распознавания, а также высокую вероятность, адекватность и скорость распознавания речевых сигналов.

Современные системы распознавания речевых образов используют следующие подходы:

- логический;
- метрический;
- байесовский;
- нейросетевой;
- структурный.

Существующие методы и модели распознавания речевых образов обычно основаны на:

- скрытых марковских моделях (СММ) [1-2],
- КДП-подходе [3-4],
- искусственных нейронных сетях (ИНС) [5-10]

Эти методы и модели обладают следующими недостатками [4]:

- хранение большого количества опорных образцов звуков или слов, а также весовых коэффициентов;

- значительное время обучения, обычно составляющее несколько месяцев;

- большое время распознавания;
- вероятность распознавания меньше 95 %;
- наличие сотен тысяч обучающих образцов.

Целью исследования является повышение эффективности классификации вокальных звуков речи для интеллектуальных компьютерных систем.

Задачей исследования является разработка метода классификации вокальных звуков речи на основе саундлетов и отображений между ними, комбинированной кластеризации, а также и модифицированной нейронной сети Хемминга.

Для достижения поставленной цели необходимо:

1. Разработать метод создания образцов вокальных звуков.
2. Разработать метод формирования опорных образцов на основе семейства дискретных саундлетов и саундлетных отображений.
3. Разработать модель классификации вокальных звуков на основе модифицированной нейронной сети Хемминга.
4. Создать критерии оценки эффективности модели.
5. Формализовать условия классификации вокального звука по опорным образцам на основе семейства саундлетов и саундлетных отображений для оценивания результатов классификации.
6. Разработать логико-формальные правила для оценивания результатов классификации по модели

Метод создания образцов вокальных звуков

Согласно [11], образцом вокального звука речи назовем участок вокального звука в речевом сигнале, расположенный между соседними пиковыми значениями и имеющий длину соответствующую квазипериоду.

При формировании образца в режиме обучения экспертом вводятся левая и правая границы N^l, N^r вокального звука в сигнале f , а в режиме классификации автоматически определяется (на основе энергий последовательно идущих участков сигнала равной длины) левая и правая границы N^l, N^r вокальной части сигнала f .

После задания или вычисления границ N^l, N^r на множестве $\{N^l, \dots, N^r\}$ сигнала f вычисляется функции автокорреляции, с помощью которой определяется длина периода основного тона N^{FT} вокального звука.

Для формирования образца как структурообразующего элемента вокального звука множество $\{N^l, \dots, N^r\}$ сигнала f разбивается на участки на основе вычисленной длины периода основного тона N^{FT} согласно следующему правилу

$$\begin{aligned} N_0^{\max} &= \arg \max_n f(n), \\ n &\in \{N^l - 0.1 \cdot N_0^{FT}, \dots, N^l + 0.1 \cdot N_0^{FT}\}, \\ N_0^{FT} &= N^{FT}, \\ N_{i-1}^{\max} &\leq N^r \Rightarrow (N_i^{\min} = N_{i-1}^{\max}) \wedge \\ &\wedge (N_i^{\max} = \arg \max_n f(n)) \wedge \\ &\wedge (N_i^{FT} = N_i^{\max} - N_i^{\min}), \\ n &\in \{N_i^{\min} + 0.9 \cdot N_{i-1}^{FT}, \dots, N_i^{\min} + 1.1 \cdot N_{i-1}^{FT}\}. \end{aligned}$$

На основе этого разбиения формируется конечная совокупность образцов, описываемых множеством вещественнозначных ограниченных финитных дискретных функций $\{x_i | i \in \{1, \dots, I\}\}$ в виде

$$\begin{aligned} x_i(n) &= \begin{cases} f(n), & n \in \{N_i^{\min}, \dots, N_i^{\max}\} \\ 0, & n \notin \{N_i^{\min}, \dots, N_i^{\max}\} \end{cases}, \\ & i \in \{1, \dots, I\}, \\ x_i(n) &= \begin{cases} f(n), & n \in \{N_i^{\min}, \dots, N_i^{\max}\} \\ 0, & n \notin \{N_i^{\min}, \dots, N_i^{\max}\} \end{cases}, \\ & A_i^{\min} = \min_n f(n), \\ & n \in \{N_i^{\min}, \dots, N_i^{\max}\}, i \in \{1, \dots, I\}, \\ & A_i^{\max} = \max_n f(n), \end{aligned} \quad ($$

$$n \in \{N_i^{\min}, \dots, N_i^{\max}\}, i \in \{1, \dots, I\}.$$

На основе введенных в работе [11] семейств непрерывных и дискретных саундлетов и саундлетных отображений сформируем опорные образцы вокальных звуков речи.

Метод формирования опорных образцов

Пусть дана конечная совокупность обучающих образцов вокального звука, которая описывается множеством целочисленных ограниченных финитных дискретных функций

$$X = \{x_i | i \in \{1, \dots, I\}\},$$

причем A_i^{\min}, A_i^{\max} – минимальное и максимальное значение функции x_i на компакте $\{N_i^{\min}, \dots, N_i^{\max}\}$.

Для сопоставления элементов множества X между собой посредством преобразования F , предложенного в работе [11], семейство X отображается в семейство дочерних дискретных саундлетов образцов вокального звука

$$S^c = \{s^c\},$$

такое, что

$$\forall x_i \in X \exists S^c = \{s_r^c | r \in \{1, \dots, I\}\} : s_r^c = Fx_r,$$

причем все функции S^c имеют одинаковый компактный носитель $\{1, \dots, \tilde{N}^{av} + 1\}$ и одинаковые минимальные и максимальные значения $0, \tilde{A}^{av}$ на нем, причем

$$\begin{aligned} \tilde{N}^{av} &= \frac{1}{I} \sum_{i=1}^I (N_i^{\max} - N_i^{\min}), \\ \tilde{A}^{av} &= \frac{1}{I} \sum_{i=1}^I (A_i^{\max} - A_i^{\min}). \end{aligned}$$

Вычисляется нормированное расстояние между функцией, описывающей обучающий образец, и дочерним дискретным саундлетом в виде

$$\begin{aligned} \forall i, r \in \{1, \dots, I\} \quad d_{ir} &= \frac{\rho_p(s_i^c, s_r^c)}{\tilde{A}^{av} \sqrt{\tilde{N}^{av} + 1}}, \\ \rho_p(s_i^c, s_r^c) &= p \sqrt{\sum_{m \in \mathbb{Z}} |s_i^c(m) - s_r^c(m)|^p}. \end{aligned}$$

Осуществляется выбор множества функций H , описывающих опорные образцы, из множества функций S^c , описывающих обучающие образцы, на основе матрицы нормированных расстояний $[d_{ir}]$, используя комбинированный алгоритм кластеризации (сочетание алгоритма средних и иерархической кластеризации).

Каждая функция h_k множества H преобразуется к N -мерному вектору \bar{v}_k множества V , причем все вектора \bar{v}_k одинаковые минимальные и максимальные значения 0 и A .

$$\forall h_k \in H \exists \bar{v}_k : v_{kn} = (Fh_k)(n-1), n \in \{1, \dots, N\}.$$

На основе введенных в работе [11] семейств саундлетов и саундлетных отображений и сформированного множества опорных образцов создадим модель классификации вокальных звуков

Модель классификации вокальных звуков на основе модифицированной сети Хемминга

На рис. 1 приведена модифицированная нейронная сеть Хемминга, которая является рекуррентной двухслойной ИНС и в отличие от аналога [12-14] позволяет работать с вещественными данными [15, 16]. Выходной слой соответствует выходному слою MaxNet.

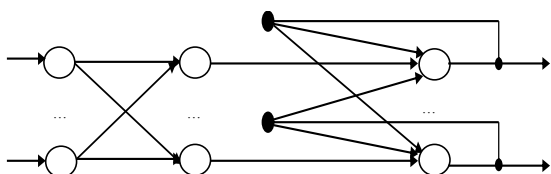


Рисунок 1 - Модифицированная нейронная сеть Хемминга

Емкость памяти нейронной сети Хемминга составляет $2^{\beta N^{(0)}}$, $\beta \leq 1$, что больше емкости нейронной сети Хопфилда или машины Больцмана [17], где $N^{(0)}$ – количество нейронов входного и выходного слоя сети.

Суть работы модифицированной нейронной сети Хемминга заключается в следующем. После начальной оценки расстояния Хемминга между входным вектором и запомненными при обучении образцами, сеть начинает циклические вычисления с использованием обратных связей, во время которых выходы сети, соответствующие классам, пытаются взаимно подавить конкурирующие выходные сигналы. В результате этого процесса активизируется выход нейрона-победителя, который соответствует запомненному образцу, ближайшему к входному вектору. Номер выигравшего нейрона определяет номер запомненного образца. Роль первого слоя весьма условна: сеть использует его веса только один раз, после чего этот слой в вычислениях не участвует.

Обучение модифицированной нейронной сети Хемминга (фаза запоминания)

Задается обучающее множество $V = \{\bar{v}_i | \bar{v}_i \in (0,1)^{N^{(0)}}\}$, $i \in \overline{1, N^{(1)}}$. Производится инициализация смещения (порога) $b_j^{(1)} = N^{(0)} / 2$, $j \in \overline{1, N^{(1)}}$, весов связей первого слоя $w_{ij}^{(1)} = \frac{v_{ij}}{A}$, $i \in \overline{1, N^{(0)}}$, $j \in \overline{1, N^{(1)}}$.

Функционирование модифицированной нейронной сети Хемминга (фаза восстановления)

1. $y_j^{(1)} = \sum_{i=1}^{N^{(0)}} (w_{ij}^{(1)} - z_i)^2$
2. $y_j^{(2)}(1) = y_j^{(1)}$, $n = 1$.
3. $y_j^{(2)}(n+1) = f(y_j^{(2)}(n) - \varepsilon \sum_{i=1}^{N^{(2)}} y_i^{(2)}(n))$,

$$f(s) = \frac{1}{1 + \exp(-\alpha s)}, j \in \overline{1, N^{(2)}}, 0 < \varepsilon < 1.$$

4. Если $\sum_{j=1}^{N^{(2)}} |y_j^{(2)}(n+1) - y_j^{(2)}(n)| > \delta$, то $n = n + 1$, переход на 3, иначе выдать номер класса $j^* = \arg \max_j y_j^{(2)}(n+1)$.

Для оценивания результатов классификации вокальных звуков необходимо сформулировать критерии оценки эффективности модели.

Критерии оценки эффективности модели

1. *Критерий скорости классификации* означает выбор из заданного набора метрик такой метрики, которая на стадии обучения модели требует наименьшего количества опорных образцов

$$F = T \rightarrow \min_p$$

2. *Критерий оценки пороговой вероятности классификации* означает выбор такого множества опорных образцов на стадии опытной эксплуатации модели, чтобы для тестового образца номер звука, вычисленный по модели, совпал с тестовым номером звука этого тестового образца

$$F = \frac{1}{I} \sum_{i=1}^I \phi(y_i^{model}, y_i^{test}) \rightarrow \max_{\{H_j\}}$$

$$\phi(a, b) = \begin{cases} 1, & a = b \\ 0, & a \neq b \end{cases}$$

где y_i^{test} – тестовый номер звука для i -го тестового образца,

y_i^{model} – номер звука для i -го тестового образца, полученный по модели,

I – количество тестовых образцов.

3. Для оценки готовности модели к эксплуатации используется критерий адекватности модели, основанный на минимуме среднеквадратичной ошибки

$$F = \frac{1}{I} \sum_{i=1}^I (\tilde{y}_i^{model} - \tilde{y}_i^{test})^2 \rightarrow \min_{\{H_j\}}$$

где \tilde{y}_i^{test} – тестовое состояние выходного нейрона-победителя для i -го тестового образца,

\tilde{y}_i^{model} – состояние выходного нейрона-победителя для i -го тестового образца, полученное по модели.

Для оценивания результатов классификации вокальных звуков необходимо сформулировать условия их классификации.

Условия классификации тестового образца вокального звука по опорным образцам

Необходимое условие классификации тестового образца. Тестовый образец классифицирован, если

$$\forall n \in \{1, \dots, J\} \quad \forall m \in \{1, \dots, J\} \\ (y^{model} = n) \wedge (\tilde{y}^{model} = m) \rightarrow, \\ \rightarrow (n = m) \wedge (\tilde{y}^{model} > \tilde{\epsilon}), \quad j \in \{1, \dots, J\},$$

где $\tilde{\epsilon}$ – заданный порог, $0 < \tilde{\epsilon} \leq 1$.

Достаточное условие классификации тестового образца. Тестовый образец классифицирован, если

$$\forall n \in \{1, \dots, J\} \quad \forall m \in \{1, \dots, J\} \\ (y^{model} = n) \wedge (\tilde{y}^{model} = m) \rightarrow, \\ \rightarrow (n = m) \wedge (\tilde{y}^{model} = 1), \quad j \in \{1, \dots, J\}.$$

На основе полученных условий возможно сформировать логико-формальные правила оценивания результатов классификации.

Логико-формальные правила оценивания результата классификации

Для оценивания результатов классификации формируются следующие логико-формальные правила

Если $\tilde{y}^{model} > \tilde{\epsilon}$, то $q = y$,

Если $\tilde{y}^{model} \leq \tilde{\epsilon}$, то $q = 0$,

где q – номер звука,

$\tilde{\epsilon}$ – это численно вычисленный порог для множества векторов тестовых образцов невокальных (непериодических) звуков.

Численное исследование метрического метода классификации вокальных звуков

В табл. 1 приведено сравнение предложенного метода и существующих нейросетевых методов на основе базы данных ТИМІТ. Классификации подлежали все вокальные звуки.

В неавторских методах в качестве образцов брались вектора мел-частотные кепстральные коэффициенты (MFCC), вычисленные на участках

равной длины, т.е. фреймах. Ошибка классификации представляет собой отношение количества правильно классифицированных образцов, содержащих вокальные звуки, к их общему количеству в процентах, при этом образцы, содержащие конец первого вокального звука и начало второго вокального звука, не учитывались. Приведенные в табл.1 стандартные нейросетевые методы были реализованы автором статьи, посредством пакета Matlab. Исследование позволяет сделать вывод, что авторский метод обеспечивает высокую вероятность классификации.

Таблица 1. Оценка нейросетевых методов классификации

Искусственные нейронные сети	Ошибка классификации (%)
трехслойный многослойный персептрон	0.80
нейронная сеть с радиально-базисными функциями	0.81
машина опорных векторов	0.82
вероятностная нейронная сеть	0.84
нейронная сеть Хопфилда	0.86
машина Больцмана	0.87
нейронная сеть с задержкой по времени	0.88
трехслойный рекурсивный многослойный персептрон	0.90
модифицированная сеть Хемминга	0.95

Для выбора наиболее эффективной метрики, используемой в методе формирования опорных образцов вокальных звуков, было проведено численное исследование, результаты которого приведены в табл.2.

Таблица 2. Количественные оценки метрик для формирования опорных образцов

Метрика	Среднее значение I / K
ρ_1	3.0
ρ_2	4.0
ρ_3	3.0
ρ_4	3.4

Отношение I / K представляет собой отношение общего количества обучающих образцов, содержащих вокальные звуки, к количеству эталонных образцов, при этом образцы, содержащие конец первого вокального звука и начало второго вокального звука, не учитывались. Как видно из табл.2., наиболее эффективной оказалась метрика ρ_2 .

Заключення

Научная новизна. В работе впервые предложено использовать саундлеты и саундлетные отображения применительно к искусственным нейронным сетям. Усовершенствован подход к классификации вокальных звуков, который отличается тем, что позволяет учитывать квазипериодическую структуру вокальных звуков и обобщать образцы одного звука различной длины и различным размахом амплитуд, что повышает эффективность классификации вокальных звуков речи. Получил дальнейшее развитие метод создания множества опорных образцов, который отличается тем, что основан на семейства саундлетов и саундлетных отображений, что повышает эффективность процедуры формирования опорных образцов. В рамках предложенных саундлетов и саундлетных отображений становится возможным использовать модифицированную сеть Хемминга применительно к структурным элементам вокального звука речи и использовать адаптивный нормиро-

ванный порог в логико-формальных правилах, что повышает вероятность классификации полезных звуков.

Практическая значимость Разработан метод построения модели классификации вокальных звуков на основе саундлетов и модифицированной сети Хемминга, что позволяет сократить количество опорных образцов, используемых для обучения. Предложен адаптивный нормированный порог для логико-формальных правил оценивания классификации речевых сигналов, который позволяет с большей вероятностью выделять полезные звуки. В результате численного исследования было установлено, что алгоритм классификации вокальных звуков на основе саундлетов и модифицированной сети Хемминга дает вероятность классификации 0.95. Созданные алгоритмы могут использоваться для решения задач, связанных с распознаванием речи оператора, идентификацией диктора, синтезом речи, анализом вибрационного сигнала.

Список литературы

1. Rabiner L.R. Fundamentals of speech recognition / L.R. Rabiner, B.H. Jang. – Englewood Cliffs, NJ: Prentice Hall PTR, 1993. – 507 p.
2. Потапова Р.К. Речь: коммуникация, информация, кибернетика / Р.К. Потапова. М.: Радио и Связь, 1997. 528 с.
3. Винцюк Т.К. Анализ, распознавание и интерпретация речевых сигналов / Т.К. Винцюк. – К.: Наук. думка, 1987. – 261 с.
4. Федоров Е.Е. Методология создания мультиагентной системы речевого управления: монография / Е.Е. Федоров. – Донецк: изд-во «Ноулидж», 2011. – 356 с.
5. Хайкин С. Нейронные сети: полный курс / С. Хайкин. – М.: Издательский дом «Вильямс», 2006. – 1104 с.
6. Галушкин А.И. Теория нейронных сетей / А.И. Галушкин. – М.: ИПРЖР, 2000. – 416 с.
7. Каллан Р. Основные концепции нейронных сетей / Р. Каллан. – М.: Издательский дом «Вильямс», 2001. – 288 с.
8. Sivanandam S.N. Introduction to Neural Networks using Matlab 6.0/ S.N. Sivanandam, S. Sumathi, S.N. Deera – New Delhi: The McGraw-Hill Comp., Inc., 2006. – 660 p.
9. Руденко О.Г. Штучні нейронні мережі: Навчальний посібник / О.Г. Руденко, С.В. Бодянский. – Харків: ТОВ Компанія СМІТ, 2006. – 404 с.
10. Галушкин А.И. Нейронные сети: история развития теории: Учеб. пособие для вузов. / А.И. Галушкин, Я.З. Цыпкин. – М.: ИПРЖР, 2001. – 840 с.
11. Федоров Е.Е. Метод классификации вокальных звуков речи на основе саундлетной байесовской нейронной сети // Управляющие системы и машины. – №6. – 2015. – С. 78-83.
12. Lippmann R.P. An introduction to computing with neural nets / R.P. Lippmann // IEEE Acoustics, Speech and Signal Processing Magazine. – 1987. – April. – P. 4-22.
13. Осовский С. Нейронные сети для обработки информации / С. Осовский. – М.: Финансы и статистика, 2002. – 344 с.
14. Комарцова Л.Г. Нейрокомпьютеры / Л.Г. Комарцова, А.В. Максимов. – М.: Изд-во МГТУ им. Н.Э. Батмана, 2002. – 320 с.
15. Куравский Л.С. Нейросетевая технология диагностики патологических состояний по аномалиям электроэнцефалограмм / Л.С. Куравский, С.Н. Баранов, О.Е. Буланова, Т.Е. Кравчук // Нейрокомпьютеры: разработка и применение. 2007. № 4. С. 4–14.
16. Федоров Е.Е. Искусственные нейронные сети: монография / Е.Е. Федоров. – Красноармейск: ДВНЗ «ДонНТУ», 2016. – 338 с.
17. Ackley D.H. A learning algorithm for Boltzmann machines / D.H. Ackley, G.E. Hinton, T.J. Sejnowski // Cognitive Science. – 1985. – Vol.9. – P. 147-169.

Надійшла до редакції 15.03.2014

Є.Є. ФЕДОРОВ, І.В.ЯРОШ, Т.О. ЧЕРНЯК

Донецький національний технічний університету

МЕТОД КЛАСИФІКАЦІЇ ВОКАЛЬНИХ ЗВУКІВ МОВЛЕННЯ НА ОСНОВІ САУНДЛЕТІВ І МОДИФІКОВАНОЇ МЕРЕЖІ ХЕМІНГА

У статті запропоновано метод класифікації вокальних звуків мовлення, який базується на саундлетах і відображеннях між ними, а також на модифікованій нейронній мережі Хеммінга, і дозволяє враховувати структуру квазіперіодичного сигналу і зіставляти зразки вокальних звуків мовлення різної довжини. На основі саундлетів і відображень між ними розроблено метод створення зразків і метод формування опорних зразків, який також використовує комбіновану кластеризацію (поєднання ітеративної і ієрархічної кластеризації). Запропоновані методи можуть використовуватися в інтелектуальних комп'ютерних системах для підвищення якості розпізнавання мовлення і ідентифікації диктора.

Ключові слова: класифікації вокальних звуків мови, мережа Хеммінга, саундлети, комбінована кластеризація, метод створення зразків, метод формування опорних зразків

E. FEDOROV, I. YAROSH, T. CHERNYAK

Donetsk National Technical University

METHOD FOR CLASSIFICATION OF VOCAL SPEECH SOUNDS BASED ON SAUNDLETS AND MODIFIED HAMMING NETWORK

Development of software components intended for human speech recognition, speaker identification, synthesis of human speech, et al., and used in the intelligent computer systems is actual problem now. This problem is based on building of effective methods, providing high speed of learning recognition model, as well as a high probability, the adequacy and speed of recognition of speech signals. Modern speech pattern recognition system use the following approaches: logic; metric; Bayesian; neural network; structural. Existing methods for pattern recognition and speech models are typically based on hidden Markov models (HMM); CDP-approach; artificial neural networks (ANN). These methods and models have the following disadvantages: storage of a large number of reference patterns of words or sounds, and the weighting coefficients; the training for several months; big time recognition; recognition probability is less than 95%; the presence of hundreds of thousands of training patterns. The aim is to develop a method for the classification of vocal sounds based on speech saundlets and mappings between them, a combination of clustering, as well as a modified Hamming neural network. For the first time it proposed to use saundlets and saundlets map in relation to artificial neural networks. Approach to the classification of vocal sounds is improved. New approach is characterized in that it takes into account the quasi-periodic structure of vocal sounds and compiles patterns of the sound of different lengths and different swing amplitudes, which increases the efficiency of the classification of vocal sounds of speech. Method of creating a plurality of reference patterns is a further developed. This method is characterized by the fact that based on the family saundlets and saundlets maps, which increases the efficiency of procedure for generating reference patterns. Through the proposed saundlets and saundlets maps it is possible to use a modified Hamming network with respect to the structural elements of speech and vocal sound to use normalized adaptive threshold logic-formal rules that increases the probability of classification of useful sounds. A method for building a classification model based on vocal sounds saundlets and modified Hamming network, thereby reducing the number of reference patterns used for training. The adaptive threshold for the normalized logical-formal rules of evaluation classification of speech signals, which allows you to select useful sounds more likely. The numerical research has found that the classification algorithm based on vocal sounds saundlets and modified Hamming network gives the probability of classification 0.95. The algorithm can be used to solve the problems associated with the recognition of operator speech, speech synthesis, analysis of the vibration signal.

Keywords: classification of vocal sounds of speech, Hamming network, saundlets, combined clustering, method creating pattern, method formation of the reference patterns

REFERENCES

1. Rabiner, L.R., Jang B.H. (1993). *Fundamentals of speech recognition*, Prentice Hall PTR, Englewood Cliffs, NJ, 507 p.
2. Potapova, R.K. (1997). *Speech: communication, information, cybernetics* [Rech: kommunikatsiya, informatsiya, kibernetika], Radio i Svyaz, Moscow, 528 p.
3. Vintsyuk, T.K. (1987). *Analysis, recognition and interpretation of speech signals* [Analiz, raspoznavanie i interpretatsiya rechevyih signalov], Nauk. dumka, Kiev, 261 p.
4. Fedorov, E.E. (2011) *Methodology for creating a multi-agent speech control system: monograph* [Metodologiya sozdaniya multiagentnoy sistemyi rechevogo upravleniya: monografiya], izd-vo «Noulidzh», Donetsk, 356 p.
5. Haykin, S. (2006). *Neural networks: full course* [Neyronnyie seti: polnyiy kurs] Izdatelskiy dom «Vilyams», Moscow, 1104 p.

-
6. Galushkin, A.I. (2000). *The theory of neural networks* [Teoriya neyronnykh setey], IPRZhR, Moscow, 416 p.
 7. Kallan, R. (2001). *Basic concepts of neural networks* [Osnovnyie kontseptsii neyronnykh setey], Izdatelskiy dom «Vilyams», Moscow, 288 p.
 8. Sivanandam, S.N., Sumathi S., Deepa S.N. (2006). *Introduction to Neural Networks using Matlab 6.0*, The McGraw-Hill Comp., Inc., New Delhi, 660 p.
 9. Rudenko, O.H., Bodianskyi, Ye.V. (2006). *Artificial neural networks: Textbook* [Shtuchni neironni me-rezhi: Navchalnyi posibnyk], TOV Kompaniia SMIT, Kharkov, 404 p.
 10. Galushkin, A.I., Tsyipkin, Ya.Z. (2001). *Neural networks: the history of the theory: Proc. Manual for high schools* [Neyronnyie seti: istoriya razvitiya teorii: Ucheb. posobie dlya vuzov], IPRZhR, Moscow, 840 p.
 11. Fedorov, E.E. (2015). The method of classification of vocal sounds of speech on the basis of sound-let's Bayesian neural network [Metod klassifikatsii vokalnykh zvukov rechi na osnove saundletnoy bayesovskoy neyronnoy seti], *Upravlyayushchie sistemy i mashiny*, № 6, pp. 78-83.
 12. Lippmann, R.P. (1987). An introduction to computing with neural nets, *IEEE Acoustics, Speech and Signal Processing Magazine*, April, pp. 4-22.
 13. Osovskiy, S. (2002). *Neural networks for information processing* [Neyronnyie seti dlya obrabotki in-formatsii], *Finansyi i statistika*, Moscow, 344 p.
 14. Komartsova, L.G., Maksimov, A.V. (2002). *Neurocomputers* [Neyrokompyuteryi], Izd-vo MGTU im. N.E. Batmana, Moscow, 320 p.
 15. Kuravskiy, L.S., Baranov, S.N., Bulanova, O.E., Kravchuk, T.E. (2007) Neural network technology for the diagnosis of pathological conditions by anomalies of electroencephalograms [Neyrosetevaya tehnologiya diagnostiki patologicheskikh sostoyaniy po anomaliiyam elektroentsefalogramm], *Neyrokompyuteryi: razrabotka i primeneniye*, № 4, pp. 4-14.
 16. Fedorov, E.E. (2016). *Artificial neural networks: monograph* [Iskusstvennyie neyronnyie seti: mono-grafiya], DVNZ «DonNTU», Krasnoarmeysk, 338 p.
 17. Ackley, D.H., Hinton, G.E., Sejnowski T.J. (1985). A learning algorithm for Boltzmann machines, *Cognitive Science*, Vol.9, pp. 147-169.