

УДК 004

Е.Е. Федоров (канд. техн. наук, доц.)Донецкая академия автомобильного транспорта, г. Донецк
кафедра специализированных компьютерных систем,
E-mail: fee75@mail.ru**МЕТОД ИДЕНТИФИКАЦИИ ДИКТОРА НА ОСНОВЕ КОМПОЗИЦИИ
DTW-КЛАССИФИКАТОРОВ**

Для повышения точности, надежности и быстродействия принимаемых решений при управлении движущимся транспортом в окружающей среде с высокой динамикой изменения характеристик в статье предлагается метод идентификации диктора на основе композиции классификаторов, использующих алгоритм Dynamic Time Warping (DTW), базирующийся на формировании системы векторов признаков слов, эталонов слов, моделей DTW-классификаторов слов, композиции DTW-классификаторов, правил принятия решения. Для предложенного метода проводится численное исследование.

Ключевые слова: идентификация диктора, алгоритм DTW, композиции классификаторов, решающее правило, адаптивный нормированный порог.

Общая постановка проблемы

В настоящее время актуальной становится разработка интеллектуальных систем поддержки принятия решений (ИСППР) для решения проблем повышения точности, надежности и быстродействия принимаемых решений при управлении движущимся транспортом в окружающей среде с высокой динамикой изменения характеристик. При создании таких ИСППР возникает задача создания естественно-языкового интерфейса и, в частности, идентификация диктора. Для идентификации диктора необходимо выбрать систему векторов признаков, модель для классификации дикторов и решающее правило.

Анализ исследований

Методы идентификации диктора [1] базируются на метрическом подходе [2] (чаще всего используются расстояния Манхеттена, Евклида, Махаланобиса), нейросетях [3] и динамическом программировании [4]. Возможно также сочетание нейросетей и динамического программирования [5]. Недостатком первого подхода является отсутствие возможности сопоставления целого ключевого слова и эталона [6]. Второй подход также плохо приспособлен для сопоставления целого ключевого слова и эталона, требует длительного обучения и большого количества обучающих данных, дает невысокую вероятность идентификации [1,3]. Поэтому обычно используется третий подход. В рамках этого подхода лучшие результаты по точности и надежности были получены с помощью непрерывных скрытых марковских моделей (НСММ) [7] и метода КДП [8], основанного на композиции (К) эталонов с помощью автоматных порождающих грамматик, и динамическом программировании (ДП). Недостатком НСММ является то, что они требуют большого объема обучающих данных и длительного обучения для настройки параметров [7]. Недостатком КДП является сложность точного определения максимального числа повторений элементов эталона для автоматной грамматики, порождающей эталоны.

Постановка задач исследования

Разработать метод идентификации диктора на основе композиции классификаторов, использующих алгоритм динамического программирования Dynamic Time Warping (DTW).

Решение задач и результаты исследований

Создание метода идентификации диктора для вербальной системы управления транспортом предусматривает формирование:

- системы векторов признаков слов;
- эталонов ключевых слов, соответствующих дикторам;
- моделей DTW-классификаторов;
- композиции DTW-классификаторов.

1. Формирование системы векторов признаков

В режиме обучения для фиксации речевых особенностей дикторов формируется система векторов признаков речевого сигнала следующих типов [9]:

1. Коэффициенты ЛП (линейного прогнозирования).
2. Коэффициенты отражения ЛП.
3. Автокорреляция ЛП.
4. Кепстр ЛП.
5. Площади поперечных сечений акустической трубы.
6. Нормированная автокорреляция.
7. Нормированный энергетический спектр ЛП.
8. Меры контрастности ЛП.
9. Нормированный энергетический спектр ДПФ.
10. Меры контрастности ДПФ (дискретного преобразования Фурье).
11. Меры контрастности ДВП (дискретного вейвлет-преобразования).
12. Меры контрастности НВП (непрерывного вейвлет-преобразования).
13. Нормированное количество импульсов равной длины.
14. MFCC (мел-частотные кепстральные коэффициенты).

Сформированная структура векторов признаков, которые отражают частотные, автокорреляционные и энергетические характеристики речи дикторов, его речевого тракта и слухового аппарата, используется при создании эталонов ключевых слов.

2. Формирование эталонов ключевых слов

Для каждого ключевого слова, характеризующего определенного диктора, записывается обучающий набор реализаций речевого сигнала, причем каждая реализация разбивается на фреймы (участки равной длины). Для каждого j -го фрейма создается вектор признаков u -го типа

$$\mathbf{x}_{uj} = (x_{uj1}, \dots, x_{ujk}, \dots, x_{ujN_u}), u \in \overline{1,14}, j \in \overline{1,J}, \quad (1)$$

который затем нормируется в виде

$$\tilde{\mathbf{e}}_{uj} = (\tilde{e}_{uj1}, \dots, \tilde{e}_{ujk}, \dots, \tilde{e}_{ujN_u}), u \in \overline{1,14}, j \in \overline{1,J}, \quad (2)$$

$$\tilde{e}_{ujk} = \frac{x_{ujk} - x1_{uk}}{x2_{uk} - x1_{uk}},$$

где \tilde{e}_{ujk} — нормированное численное значение k -й компоненты вектора признаков u -го типа j -го фрейма, $\tilde{e}_{ujk} \in [0,1]$, $x1_{uk}$ — минимальное значение k -й компоненты вектора признаков u -го типа, $x2_{uk}$ — максимальное значение k -й компоненты вектора признаков u -го типа, N_u — длина вектора признаков u -го типа, J — количество фреймов обучающей реализации ключевого слова.

Эталон обучающей реализации ключевого слова образован набором нормированных векторов фреймов

$$\tilde{\mathbf{e}}_u = (\tilde{\mathbf{e}}_{u1}, \dots, \tilde{\mathbf{e}}_{uJ}), u \in \overline{1,14}. \quad (3)$$

Модели DTW-классификаторов дикторов используют сформированные эталоны ключевых слов в качестве векторов параметров.

3. Формирование моделей DTW-классификаторов

В статье использованы четыре типа DTW-классификаторов.

Модель DTW-классификатора первого типа (детерминированного, симметричного, с направлением движения из точек $(i-1, j), (i-1, j-1), (i, j-1)$ в точку (i, j) , с евклидовым локальным расстоянием, с усреднением глобального расстояния, без взвешивания) выдает результат $DTW_1(\tilde{\mathbf{x}}_u, \tilde{\mathbf{e}}_u)$, определяемый в виде

$$\begin{aligned}
 Q(1,1) &= \|\tilde{\mathbf{x}}_{u1} - \tilde{\mathbf{e}}_{u1}\|, l(1,1) = 0, \\
 Q(i,1) &= \|\tilde{\mathbf{x}}_{ui} - \tilde{\mathbf{e}}_{u1}\| + Q(i-1,1), l(i,1) = 1, i \in \overline{2, H}, \\
 Q(1, j) &= \|\tilde{\mathbf{x}}_{u1} - \tilde{\mathbf{e}}_{uj}\| + Q(1, j-1), l(1, j) = 1, j \in \overline{2, H}, \\
 MIN &= \min(Q(i-1, j), Q(i-1, j-1), Q(i, j-1)), i \in \overline{2, H}, j \in \overline{2, H}, \\
 Q(i, j) &= \|\tilde{\mathbf{x}}_{ui} - \tilde{\mathbf{e}}_{uj}\| + MIN, i \in \overline{2, H}, j \in \overline{2, H}, \\
 l(i, j) &= \begin{cases} l(i-1, j) + 1, & MIN = Q(i-1, j) \\ l(i-1, j-1) + 1, & MIN = Q(i-1, j-1), i \in \overline{2, H}, j \in \overline{2, H}, \\ l(i, j-1) + 1, & MIN = Q(i, j-1) \end{cases} \\
 DTW_1(\tilde{\mathbf{x}}_u, \tilde{\mathbf{e}}_u) &= \frac{Q(H, H)}{l(H, H)}, \\
 \|\tilde{\mathbf{x}}_{ui} - \tilde{\mathbf{e}}_{uj}\| &= \sqrt{\sum_{k=1}^{N_u} (\tilde{x}_{uik} - \tilde{e}_{ujk})^2},
 \end{aligned}$$

где $l(i, j)$ — текущая длина маршрута, H — максимальное количество фреймов распознаваемого сигнала и эталона.

Модель DTW-классификатора второго типа (детерминированного, асимметричного, с направлением движения из точек $(i-1, j), (i-1, j-1), (i, j-1)$ в точку (i, j) , с евклидовым локальным расстоянием, с усреднением глобального расстояния, без взвешивания) выдает результат $DTW_2(\tilde{\mathbf{x}}_u, \tilde{\mathbf{e}}_u)$, определяемый в виде

$$\begin{aligned}
 Q(1,1) &= \|\tilde{\mathbf{x}}_{u1} - \tilde{\mathbf{e}}_{u1}\|, l(1,1) = 0, \\
 Q(i,1) &= \|\tilde{\mathbf{x}}_{ui} - \tilde{\mathbf{e}}_{u1}\| + Q(i-1,1), l(i,1) = 1, i \in \overline{2, I}, \\
 Q(1, j) &= \|\tilde{\mathbf{x}}_{u1} - \tilde{\mathbf{e}}_{uj}\| + Q(1, j-1), l(1, j) = 1, j \in \overline{2, J}, \\
 MIN &= \min(Q(i-1, j), Q(i-1, j-1), Q(i, j-1)), i \in \overline{2, I}, j \in \overline{2, J}, \\
 Q(i, j) &= \|\tilde{\mathbf{x}}_{ui} - \tilde{\mathbf{e}}_{uj}\| + MIN, i \in \overline{2, I}, j \in \overline{2, J}, \\
 l(i, j) &= \begin{cases} l(i-1, j) + 1, & MIN = Q(i-1, j) \\ l(i-1, j-1) + 1, & MIN = Q(i-1, j-1), i \in \overline{2, I}, j \in \overline{2, J}, \\ l(i, j-1) + 1, & MIN = Q(i, j-1) \end{cases} \\
 DTW_2(\tilde{\mathbf{x}}_u, \tilde{\mathbf{e}}_u) &= \frac{Q(I, J)}{l(I, J)}, \\
 \|\tilde{\mathbf{x}}_{ui} - \tilde{\mathbf{e}}_{uj}\| &= \sqrt{\sum_{k=1}^{N_u} (\tilde{x}_{uik} - \tilde{e}_{ujk})^2},
 \end{aligned}$$

где I, J — количество фреймов распознаваемого сигнала и эталона соответственно.

Модель DTW-классификатора третьего типа (стохастического, асимметричного, с направлением движения из точек $(i-1, j), (i-1, j-1), (i, j-1)$ в точку (i, j) , с усреднением

глобального расстояния, без взвешивания) выдает результат $DTW_3(\tilde{\mathbf{x}}_u, \tilde{\mathbf{e}}_u)$, определяемый в виде

$$\begin{aligned}
 Q(1,1) &= \ln p_{11}, l(1,1) = 0, \\
 Q(i,1) &= \ln p_{i1} + Q(i-1,1) + \ln P1_1, l(i,1) = 1, i \in \overline{2, I}, \\
 Q(1, j) &= \ln p_{1j} + Q(i, j-1) + \ln P3_j, l(1, j) = 1, j \in \overline{2, J}, \\
 MAX &= \max(Q(i-1, j) + \ln P1_j, Q(i-1, j-1) + \ln P2_j, Q(i, j-1) + \ln P3_j), i \in \overline{2, I}, j \in \overline{2, J}, \\
 Q(i, j) &= \ln p_{ij} + MAX, i \in \overline{2, I}, j \in \overline{2, J}, \\
 l(i, j) &= \begin{cases} l(i-1, j) + 1, & MAX = Q(i-1, j) \\ l(i-1, j-1) + 1, & MAX = Q(i-1, j-1), i \in \overline{2, I}, j \in \overline{2, J}, \\ l(i, j-1) + 1, & MAX = Q(i, j-1) \end{cases} \\
 DTW_3(\tilde{\mathbf{x}}_u, \tilde{\mathbf{e}}_u) &= \frac{Q(I, J)}{l(I, J)},
 \end{aligned}$$

где p_{ij} — плотность нормального распределения вектора $\tilde{\mathbf{x}}_{ui}$ i -го фрейма распознаваемого сигнала при условии, что он соответствует вектору $\tilde{\mathbf{e}}_{uj}$ j -го фрейма эталона, $P1_j$ — априорная вероятность перехода из точки $(i-1, j)$ в точку (i, j) , $P2_j$ — априорная вероятность перехода из точки $(i-1, j-1)$ в точку (i, j) , $P3_j$ — априорная вероятность перехода из точки $(i, j-1)$ в точку (i, j) .

Модель DTW-классификатора четвертого типа (стохастического, асимметричного, с направлением движения из точек $(i-1, j), (i-1, j-1), (i, j-1)$ в точку (i, j) , с усреднением глобального расстояния, без взвешивания), предложенная автором, выдает результат $DTW_4(\tilde{\mathbf{x}}_u, \tilde{\mathbf{e}}_u)$, определяемый в виде

$$\begin{aligned}
 Q(1,1) &= p_{ij}, l(1,1) = 0, \\
 Q(i,1) &= p_{ij} + Q(i-1,1)P2_1, l(i,1) = 1, i \in \overline{2, I}, \\
 Q(1, j) &= p_{ij} + Q(1, j-1)P3_j, l(1, j) = 1, j \in \overline{2, J}, \\
 MAX &= \max(Q(i-1, j)P1_j, Q(i-1, j-1)P2_j, Q(i, j-1)P3_j), i \in \overline{2, I}, j \in \overline{2, J}, \\
 Q(i, j) &= p_{ij} + MAX, i \in \overline{2, I}, j \in \overline{2, J}, \\
 l(i, j) &= \begin{cases} l(i-1, j) + 1, & MAX = Q(i-1, j) \\ l(i-1, j-1) + 1, & MAX = Q(i-1, j-1), i \in \overline{2, I}, j \in \overline{2, J}, \\ l(i, j-1) + 1, & MAX = Q(i, j-1) \end{cases} \\
 DTW_4(\tilde{\mathbf{x}}_u, \tilde{\mathbf{e}}_u) &= \frac{Q(I, J)}{l(I, J)}.
 \end{aligned}$$

На основе сформированных четырех моделей DTW-классификаторов создается четыре композиции DTW-классификаторов.

4. Формирование композиции DTW-классификаторов

Композиция DTW-классификаторов, представленная на рис. 1, является неполносвязной сетью и осуществляет параллельное сопоставление распознаваемого сигнала, представляющего диктора, с эталонами, которые соответствуют обучающим реализациям. Предложенная композиция формируется по следующему принципу. Компонентам нормированных векторов признаков речевого сигнала $\tilde{\mathbf{x}}_u$ соответствуют

элементы входного слоя; эталонам обучающих реализаций ключевых слов соответствуют элементы скрытого слоя; номерам ключевых слов соответствуют элементы выходного слоя.

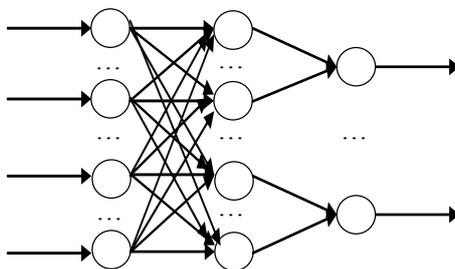


Рисунок 1 — композиция DTW-классификаторов

В скрытом слое происходит распознавание ключевого слова посредством DTW-классификаторов v -го типа с эталонами u -го типа, т.е. его сопоставление со всеми z -ми обучающими реализациями всех ключевых слов и получение результата $DTW_v(\tilde{\mathbf{x}}_u, \tilde{\mathbf{e}}_{uz})$, $z \in \overline{1, N^{(1)}}$, где $N^{(1)}$ — количество эталонов (обучающих реализаций) всех ключевых слов.

В выходном слое для всех обучающих реализаций l -го ключевого слова происходит усреднение результатов классификации, что обеспечивает обобщающую способность композиции DTW-классификаторов

$$f_{vl}(\tilde{\mathbf{x}}_u) = \frac{1}{n_l} \sum_{z=1}^{N^{(1)}} w_{zl} DTW_v(\tilde{\mathbf{x}}_u, \tilde{\mathbf{e}}_{uz}), v \in \overline{1,4}, u \in \overline{1,14}, \quad (4)$$

где n_l — количество обучающих реализаций l -го ключевого слова, $w_{zl} \in \{0,1\}$ — весовые коэффициенты.

Принятие решений о номере распознанного ключевого слова (идентифицированного диктора) осуществляется в соответствии с решающим правилом для:

– детерминированного DTW

$$\min_{l \in \overline{1, N^{(2)}}} f_{vl}(\tilde{\mathbf{x}}_u) < \varepsilon \rightarrow result_{uv} = \arg \min_{l \in \overline{1, N^{(2)}}} f_{vl}(\tilde{\mathbf{x}}_u), v \in \overline{1,2}, u \in \overline{1,14}, \quad (5)$$

$$\min_{l \in \overline{1, N^{(2)}}} f_{vl}(\tilde{\mathbf{x}}_u) \geq \varepsilon \rightarrow result_{uv} = 0, v \in \overline{1,2}, u \in \overline{1,14}; \quad (6)$$

– стохастического DTW

$$\max_{l \in \overline{1, N^{(2)}}} f_{vl}(\tilde{\mathbf{x}}_u) > \varepsilon \rightarrow result_{uv} = \arg \max_{l \in \overline{1, N^{(2)}}} f_{vl}(\tilde{\mathbf{x}}_u), v \in \overline{3,4}, u \in \overline{1,14}, \quad (7)$$

$$\max_{l \in \overline{1, N^{(2)}}} f_{vl}(\tilde{\mathbf{x}}_u) \leq \varepsilon \rightarrow result_{uv} = 0, v \in \overline{3,4}, u \in \overline{1,14}. \quad (8)$$

5. Численное исследование

Для выбора лучших DTW-классификаторов и векторов признаков было проведено численное исследование для 100 дикторов с разными голосовыми данными. Длина тестовой выборки для каждого диктора определялась как $M=10$.

Структура композиции DTW-классификаторов для каждого диктора определена следующим образом:

– количество элементов в скрытом слое (длина обучающей выборки) — $N^{(1)} = 1000$;

– количество элементов (слов) в выходном слое — $N^{(2)} = 100$.

Для стохастических DTW-классификаторов (третий и четвертый) в силу возможного отсутствия большого количества обучающих реализаций и возможного большого разброса значений компонент векторов признаков было принято, что:

– плотность нормального распределения

$$p_{ij} = \exp\left(-\frac{1}{2} \|\tilde{\mathbf{x}}_{ui} - \tilde{\mathbf{e}}_{uj}\|^2\right) = \exp\left(-\frac{1}{2} \sum_{k=1}^{N_u} (\tilde{x}_{uik} - \tilde{e}_{ujk})^2\right);$$

– априорная вероятность перехода

$$P1_j = P2_j = P3_j = \frac{1}{3}.$$

Вероятность классификации (идентификации) оценивалась для

– детерминированного DTW в виде

$$F = \frac{1}{M} \sum_{m=1}^M I(\arg \min_{l \in 1, N^{(2)}} f_{vl}(\tilde{\mathbf{x}}_{um}), test_m), v \in \overline{1,2}, u \in \overline{1,14};$$

– стохастического DTW в виде

$$F = \frac{1}{M} \sum_{m=1}^M I(\arg \max_{l \in 1, N^{(2)}} f_{vl}(\tilde{\mathbf{x}}_{um}), test_m), v \in \overline{3,4}, u \in \overline{1,14},$$

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

где $test_m$ — номер m -го тестового ключевого слова.

Адаптивный нормированный порог для решающего правила оценивался для:

– детерминированного DTW в виде

$$\varepsilon = \max_{m \in 1, M} \min_{l \in 1, N^{(2)}} f_{vl}(\tilde{\mathbf{x}}_{um}), v \in \overline{1,2}, u \in \overline{1,14} \text{ при } I(\arg \min_{l \in 1, N^{(2)}} f_{vl}(\tilde{\mathbf{x}}_{um}), test_m) = 1;$$

– стохастического DTW в виде

$$\varepsilon = \min_{m \in 1, M} \max_{l \in 1, N^{(2)}} f_{vl}(\tilde{\mathbf{x}}_{um}), v \in \overline{3,4}, u \in \overline{1,14} \text{ при } I(\arg \max_{l \in 1, N^{(2)}} f_{vl}(\tilde{\mathbf{x}}_{um}), test_m) = 1.$$

Результаты численного исследования распознавания ключевых слов (идентификации дикторов) для четырех DTW-классификаторов и четырнадцати типов векторов признаков приведены в табл.1-2.

Таблица 1 — Вероятности идентификации дикторов композициями DTW-классификаторов четырех типов, использующих систему векторов признаков (maximum)

Тип вектора признаков	Вероятность идентификации композицией DTW-классификаторов			
	типа 1	типа 2	типа 3	типа 4
коэффициенты ЛП	0.28	0.72	0.68	0.72
коэффициенты отражения ЛП	0.84	0.96	0.94	0.94
автокорреляция ЛП	0.28	0.72	0.68	0.72
кепстр ЛП	0.72	0.86	0.84	0.9
площади поперечных сечений акустической трубы ЛП	0.62	0.58	0.62	0.62
нормированная автокорреляция	0.74	0.7	0.72	0.74
нормированный энергетический спектр ЛП	0.16	0.28	0.08	0.04
меры контрастности ЛП	0.38	0.5	0.38	0.4
нормированный энергетический спектр ДПФ	0.34	0.34	0.38	0.46
меры контрастности ДПФ	0.76	0.86	0.7	0.7
меры контрастности ДВП	0.78	0.72	0.6	0.6
меры контрастности НВП	0.8	0.94	0.9	0.9
нормированное количество импульсов равной длины	0.72	0.96	0.96	0.96
MFCC	0.82	0.99	0.99	0.98

Таблица 2— Вероятности идентификации дикторов композициями DTW-классификаторов четырех типов, использующих систему векторов признаков

Тип вектора признаков	Вероятность идентификации композицией DTW-классификаторов			
	типа 1	типа 2	типа 3	типа 4
коэффициенты ЛП	0.26	0.62	0.58	0.62
коэффициенты отражения ЛП	0.78	0.92	0.94	0.94
автокорреляция ЛП	0.26	0.62	0.58	0.62
кепстр ЛП	0.62	0.74	0.72	0.78
площади поперечных сечений акустической трубы ЛП	0.48	0.48	0.44	0.44
нормированная автокорреляция	0.64	0.62	0.58	0.66
нормированный энергетический спектр ЛП	0.1	0.28	0.06	0.1
меры контрастности ЛП	0.32	0.48	0.4	0.4
нормированный энергетический спектр ДПФ	0.1	0.26	0.2	0.36
меры контрастности ДПФ	0.68	0.82	0.68	0.7
меры контрастности ДВП	0.76	0.76	0.58	0.6
меры контрастности НВП	0.26	0.9	0.76	0.78
нормированное количество импульсов равной длины	0.58	0.92	0.86	0.86
MFCC	0.82	0.96	0.94	0.92

Как следует из табл. 1, наибольшую вероятность идентификации по наибольшему количеству типов векторов признаков (половина типов векторов признаков дают вероятность идентификации больше или равную 0.98) дает композиция на основе асимметричных классификаторов (типы 2, 3 и 4). Поскольку классификаторы этих типов сопоставимы по вероятности идентификации, из них рекомендуется выбрать классификатор 4-го типа (авторский), основным достоинством которого является вероятностная оценка результатов классификации, которая для правил принятия решений позволяет выбрать единый порог для дикторов с различными голосовыми данными.

Выводы

Новизна. В статье предложен метод идентификации диктора на основе композиции DTW-классификаторов, который включает в себя формирование системы векторов признаков слов, эталонов ключевых слов, моделей DTW-классификаторов, композиции DTW-классификаторов. Предложенный метод обеспечивает:

1. Высокое быстродействие обучения и адаптации за счет отсутствия длительной настройки эталонов ключевых слов.
2. Высокое быстродействие принятия решения за счет параллельной обработки информации.
3. Высокую точность и надежность принимаемых решений за счет выбора типа:
 - вектора признаков, обеспечивающего высокую вероятность идентификации;
 - композиции DTW-классификаторов, обладающей обобщающей способностью, дающей вероятностную оценку результатов классификации, учитывающей большой разброс значений компонент векторов признаков;
 - решающих правил, использующих единый порог для дикторов с различными голосовыми данными, чтобы достичь для всех дикторов требуемого качества обучения и отвергнуть ключевые слова, значительно отличающиеся от эталонов (например, записанные в другой окружающей среде).

Практическое значение. Основные положения данной работы предназначены для реализации в интеллектуальных системах поддержки принятия решений.

Список использованной литературы

1. Федоров Е.Е. Методология создания мультиагентной системы речевого управления: монография / Е.Е. Федоров. — Донецк: Ноулидж, 2011. — 356 с.

2. Атал Б.С. Автоматическое опознавание дикторов по голосам / Б.С. Атал // ТИИЭР. — 1976. — Т. 64, №4. — С. 48–66.
3. Хайкин С. Нейронные сети: полный курс / С. Хайкин. — М.: Издательский дом Вильямс, 2006. — 1104 с.
4. Rabiner L.R. Fundamentals of speech recognition / L.R. Rabiner, B.H. Jang. — New Jersey: Prentice Hall PTR, Englewood Cliffs, 1993. — P. 507.
5. Bourlard H. Connectionist Speech Recognition: A Hybrid Approach / H. Bourlard, N. Morgan. — Boston: Kluwer Academic Publishers, 1994. — 312 p.
6. Рабинер Л.Р. Цифровая обработка речевых сигналов / Л.Р. Рабинер, Р.В. Шафер. — М.: Радио и связь, 1981. — 495 с.
7. Kasuriya S. Comparative Study of Continuous Hidden Markov Models (CHMM) and Artificial Neural Network (ANN) on Speaker Identification System / S. Kasuriya, C. Wutiw WATCHAI, C. Tanprasert // Nectec Technical Journal. — 2001. — Vol.3, №12. — P. 200–205.
8. Винцюк Т.К. Анализ, распознавание и интерпретация речевых сигналов / Т.К. Винцюк. — К.: Наук. думка, 1987. — 261 с.
9. Федоров Е.Е. Методика формирования акустических характеристик эталонов речи / Е.Е. Федоров. — Донецк: Вебер, 2008. — 282 с.
10. Nakagawa S. Speaker-Independent English consonant and Japanese word recognition by a Stochastic Dynamic Time Warping method / S. Nakagawa, H. Nakanishi // Journal of Institution of Electronics and Telecommunication Engineers. — 1988. — Vol. 34, №1. — P. 233–241.

Надійшла до редакції:
18.02.2012 р.

Рецензент:
д-р техн. наук, проф. Чичикало Н.І.

E.E. Fedorov. Method of speaker identification on the basis of a composition of DTW-qualifiers. For increase of accuracy, reliability and speed of accepted decisions at management of moving transport in environment with high dynamics of change of characteristics in article the method of speaker identification on the basis of a composition of the qualifiers using algorithm Dynamic Time Warping (DTW), based on formation of features vectors system of words, pattern of words, models of DTW-qualifiers of words is offered, to a composition of DTW-qualifiers, rules decision-making. For the offered method numerical research is carried out.

***Keywords:* speaker identification, algorithm DTW, compositions of the qualifiers, a solving rule, an adaptive normalized threshold.**

Є.Є. Федоров. Метод ідентифікації диктора на основі композиції DTW-класифікаторів. Для підвищення точності, надійності й швидкодії прийнятих рішень при керуванні транспортом, що рухається, у навколишнім середовищі з високою динамікою зміни характеристик у статті пропонується метод ідентифікації диктора на основі композиції класифікаторів, що використовують алгоритм Dynamic Time Warping (DTW), що базується на формуванні системи векторів ознак слів, еталонів слів, моделей DTW-класифікаторів слів, композиції DTW-класифікаторів, правил ухвалення рішення. Для запропонованого методу проводиться чисельне дослідження.

***Ключові слова:* ідентифікація диктора, алгоритм DTW, композиції класифікаторів, що вирішує правило, адаптивний нормований поріг.**

© Федоров Е.Е., 2012