

## АЛГОРИТМ ИДЕНТИФИКАЦИИ ДИКТОРА ПО ГОЛОСУ НА ОСНОВЕ СТАТИСТИЧЕСКОГО МЕТОДА

Белорусский государственный университет информатики и радиоэлектроники  
г. Минск, Республика Беларусь

Акунович А.А.

Карпушкин Э.М. – к.т.н., доцент

В данной работе исследуется эффективность алгоритма идентификации диктора по голосу на основе статистического метода. Эффективность алгоритма оценивается на основе собранной базы голосов *voxforge*.

Идентификация диктора по голосу – это вид аутентификации, который позволяет установить личность человека по его голосу. Аутентификация становится возможной благодаря тому, что каждый голос обладает своими уникальными характеристиками. В данной работе мы проведем исследование и оценим эффективность решения данной задачи с помощью статистического метода. Статистический метод основан на вычислении спектров входного и эталонного сигналов и последующего сравнения этих сигналов с помощью коэффициента взаимной корреляции.

Протестируем эффективность данного метода на базе голосов *voxforge*, которая находится в свободном доступе. База состоит из аудиозаписей 219 голосов длиной 5 – 7 сек., принадлежащих 11 дикторам (7 мужских голосов и 4 женских). Каждому диктору принадлежит по 20 аудиозаписей. Для начала попробуем протестировать алгоритм текстозависимой классификации диктора на основе статистического метода. Для этого из базы данных голосов выберем 110 голосов (по 10 для каждого из 11 дикторов). Эти голоса будут использоваться в качестве эталонных. Затем на основе этих сигналов создадим смесь полезного сигнала и шума (отношение С/Ш = 65 дБ) в качестве тестовой выборки. В качестве вектора признаков использовался спектр сигнала, который вычислялся с помощью оконного преобразования Фурье (ширина окна – 1024). В результате моделирования точность классификации составила 100 %. Исследуем теперь эффективность статистического метода для решения задачи идентификации диктора под воздействием шумов. В таблице 1 представлена зависимость точности классификации от отношения сигнал/шум.

Таблица 1 – Зависимость точности классификации от отношения сигнал/шум

С/Ш (дБ)	Точность классификации (%)
65	100
50	98,2
35	91,8
30	90,0
20	84,54
10	78,2

Как видно из таблицы, точность классификации падает вместе с уменьшением отношения сигнал/шум, чего и следовало ожидать.

Таким образом, можно сделать вывод о том, что статистический метод может применяться в качестве простого и не требующего серьезных вычислительных затрат для решения задачи идентификации диктора по голосу. Однако у данного метода есть ряд недостатков, таких как восприимчивость к шуму и невозможность использования для текстонезависимой классификации.

Список использованных источников:

1. Huang X., Acero A., Hon H.-W. (2001). Spoken Language Processing: a Guide to Theory, Algorithm, and System Development. Prentice-Hall, New Jersey.
2. Free speech recognition [электронный ресурс] – режим доступа: <http://www.voxforge.org/>