

Е. А. Толкачев, А. В. Лубочкин
(УО «ГТУ им. Ф. Скорины», Гомель)

РАСПОЗНАВАНИЕ РЕЧИ МЕТОДОМ МЕЛ-ЧАСТОТНЫХ КЕПСТРАЛЬНЫХ КОЭФФИЦИЕНТОВ

Метод мел-частотных кепстральных коэффициентов имеет множество направлений для применения. Чаще всего данный метод используется в системах распознавании речи и аутентификации по голосу: для характеристики речевых сигналов, необходимой для дальнейшего распознавания [1]. Кроме этих направлений, метод мел-частотных кепстральных коэффициентов может быть применен во множестве сфер: синтез речи, медицинская диагностика, неразрушающий контроль и т. д. Мел – единица высоты звука, основанная на восприятии звука органами чувств человека.

Характеристики сигнала вычисляются путем разложения его спектра на мел-шкалу соответствующим фильтром, представляющим

собой треугольную оконную функцию. Оценка получается для определенного диапазона частот одного фрейма продолжительностью 15-20 мс. Предварительно запись речи следует разбить на отдельные слова и слова на отдельные, вышеупомянутые, фреймы. Спектр сигнала рассчитывается дискретным преобразованием Фурье. Таким образом, для каждого слова получается небольшой набор признаков, по которым будет проводиться кластеризация. Данный подход существенно упрощает задачу распознавания, исключая необходимость работы с множеством отсчетов речевого сигнала. Кластеризация осуществляется методом скрытых марковских моделей или нейронными сетями. В случае работы с небольшим словарем можно использовать более примитивные методы кластеризации [2].

Приложение разработано на C/C++. Для математического моделирования в процессе разработке использовался пакет MATLAB.

Литература

- 1 Тампель, И. Б. Автоматическое распознавание речи / И. Б. Тампель, А. А. Карпов. – СПб. : Университет ИТМО, 2016. – 138 с.
- 2 Маковкин, К. А. Гибридные модели: скрытые марковские модели и нейронные сети, их применение в системах распознавания речи / К. А. Маковкин. – М. : Вычислительный центр им. А. А. Дородницына РАН, 2006. – 201 с.