

ПРОГРАММНОЕ СРЕДСТВО КЛАССИФИКАЦИИ РЕЧИ НА ОСНОВЕ КЕПСТРАЛЬНОГО АНАЛИЗА

Райкевич А.С., Никитенко Ю.Н, Зельманский О.Б.

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

Зельманский О.Б. – к.т.н., доцент

Предложено программное средство классификации фонетических единиц на фонемы, реализующее заключающийся в нахождении минимума расстояния между кепстрами анализируемой фонетической единицы и базы образцов алгоритм. Данное программное средство позволяет определить к какой группе фонем относится анализируемая фонетическая единица и, кроме того, какой фонеме она соответствует в большей степени.

Устная речь, производимая речевым аппаратом и передаваемая в естественных условиях посредством звуковых волн, сегодня является самым оперативным и распространенным способом передачи информации в любой сфере человеческой деятельности. Поэтому в современных компьютерных системах приветствуется, а в некоторых случаях и является крайне необходимым использование средств речевого взаимодействия с пользователем.

На этапе цифровой обработки сигнала непрерывный электрический сигнал проходит оцифровку и преобразуется в набор параметров. Основной задачей на этом этапе является получение компактных, но в то же время достаточно полно описывающих речевой сигнал характеристик, позволяющих максимизировать показатели эффективности распознавания. Для того чтобы получить векторы признаков одинаковой длины, происходит сегментация речевого сигнала на равные части, а затем выполнение преобразования внутри каждого сегмента. На выходе получается последовательность признаков речевого сигнала.

Для выполнения задачи распознавания речи наиболее эффективными являются кепстральные признаки [1-3]. Кепстральный вектор может быть получен с помощью обратного дискретного преобразования Фурье от логарифма амплитуды спектра, полученного с помощью прямого дискретного преобразования Фурье [1]. Кепстральный анализ применяется для отделения сигнала возбуждения от сигнала речевого тракта. Полученные в таком случае параметры обеспечивают качественное выполнение разделительности звуков – классификацию речи.

В качестве входных параметров программного средства классификации речи используются фонетические единицы речи. Каждая фонетическая единица представляет собой звуковой файл средней длительностью от 30 до 200 мс, записанный в формате WAV. Частота дискретизации для каждого файла 44100 Гц, амплитуда каждого отсчета характеризуется 32-мя битами.

Блок вычисления кепстральных коэффициентов сигнала выполняет расчет кепстра анализируемого сигнала, который передается в блок вычисления коэффициента различия, на второй вход которого из базы данных звуковых волн аллофонов, содержащей образцы вариантов реализации фонем, обусловленные конкретным фонетическим окружением этих фонем, последовательно поступают значения их кепстральных коэффициентов. В свою очередь блок вычисления коэффициента различия на основе корреляционной матрицы осуществляет расчет коэффициентов различия для анализируемой фонетической единицы и всех образцов, содержащихся в базе данных аллофонов, а также сопоставляет их длительности, результаты вычислений передаются в блок принятия решения. В функции последнего входит принятие решения о принадлежности анализируемой фонетической единицы к той или иной группе фонем, а также ее классификация как определенной фонемы путем нахождения образца с наиболее близкой анализируемой фонетической единице длительностью, которому соответствует наименьшее значение коэффициента различия. Классифицированная фонетическая единица сохраняется в новой базе данных аллофонов или обновляет уже существующую базу.

Таким образом, в ходе работы программного средства формируется массив фонетических единиц, разделенных по классам, которые сохраняются в базе данных звуковых волн аллофонов.

Список использованных источников:

1. Рабинер, Л.Р. Цифровая обработка речевых сигналов: пер. с англ. / Л.Р. Рабинер, Р.В. Шафер; под ред. М.В. Назарова и Ю.Н. Прохорова. – Москва: Радио и связь, 1981. – 495 с.
2. Шарий, Т.В. О проблеме параметризации речевого сигнала в современных системах распознавания речи / Т.В. Шарий // Вісник Донецького національного університету. Сер. А: Природничі науки. 2008. – № 2 [Электронный ресурс] / Національна бібліотека України ім. В.І. Вернадського. – Київ, 2008. – Режим доступа: http://www.nbuv.gov.ua/portal/Natural/VDU/a/2008_2/Control%20systems/9_Shariy.pdf. – Дата доступа: 1.11.2010.
3. Граничин, О.Н. Решение задачи автоматического распознавания отдельных слов речи при помощи рандомизированного алгоритма стохастической аппроксимации / О.Н. Граничин, Д.С. Шалымов // Нейрокомпьютеры: разработка, применение. – 2009. – № 3. – С. 58–64.