

ИССЛЕДОВАНИЕ АЛГОРИТМА СЕГМЕНТАЦИИ РЕЧЕВОГО СИГНАЛА МЕТОДОМ ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ

О. Б. Мишунин, А. В. Швец, А. П. Савинов, Е. В. Михалёва, Т. С. Петровская
Национальный исследовательский Томский политехнический университет,
Россия, г. Томск, пр. Ленина, 30, 634050
E-mail: rjawiogvozd@gmail.com

SPEECH SIGNAL SEGMENTATION ALGORITHM BASED ON DYNAMIC TIME WARPING

O. B. Mishunin, A. V. Shvets, A. P. Savinov, E. V. Mikhalyova, T. S. Petrovskaya
Tomsk Polytechnic University, Russia, Tomsk, Lenin str., 30, 634050
E-mail: rjawiogvozd@gmail.com

***Annotation.** Speech signal segmentation algorithm which uses previously known information about signal content was explored. Speech synthesizer was used to produce segmented signal that contains same speech as the signal we need to process, and compare this synthesized signal with the natural one using Dynamic Time Warping. This allows us to move segment markers from synthesized signal to natural signal. To compare signals we use feature vectors consisting of spectrum we got by using Bark scale based set of filters, pitch frequency and their finite differences.*

Сегментация речевого сигнала – очень важная задача при построении любых систем обработки речи. Данный процесс позволяет выделить минимальные смысловоразличительные единицы языка – фонемы, которые затем можно использовать для последующего анализа. Авторами было принято решение повторить эксперимент Давыдова А.Г., к.т.н. Объединенного института проблем информатики Национальной академии наук республики Беларусь, по сегментации речевого сигнала с помощью метода динамического программирования.[1,2]

Были выделены следующие этапы обработки сигнала:

1. Нормализация сигнала. Осуществляется делением сигнала на его максимальное по модулю значение.
2. Расстановка маркеров периода основного тона и расчет частоты основного тона. Для расстановки маркеров основного тона использовался собственный алгоритм нахождения периодов основного тона на основе многоуровневого анализа локальных экстремумов сигнала.
3. Расчет спектра сигнала с помощью гребенки из двадцати полосовых фильтров, соответствующей Барк шкале.[3] Полученные после фильтрации сигналы делятся на отрезки длиной 25 мс (величина может варьироваться), и для каждого из них вычисляется значение спектра как логарифм квадратного корня из среднего квадрата амплитуды.
4. Сигнум-нормализация спектра.[1] Рассчитывается как средняя разность текущего элемента спектра с несколькими соседними элементами.
5. Составление векторов признаков сигнала из нормализованного спектра, его конечных разностей, нормализованных частоты основного тона и ее конечных разностей, умноженных на

соответствующие весовые коэффициенты.

6. Сопоставление векторов признаков синтезированного и естественного сигналов с помощью метода динамического программирования.

7. Перенос границ сегментов звукового сигнала с синтезированного сигнала на естественный.

8. Коррекция границ сегментов в соответствии с найденными маркерами периодов основного тона.

Шаги 1-5 выполняются как на синтезированном, так и на естественном сигналах.

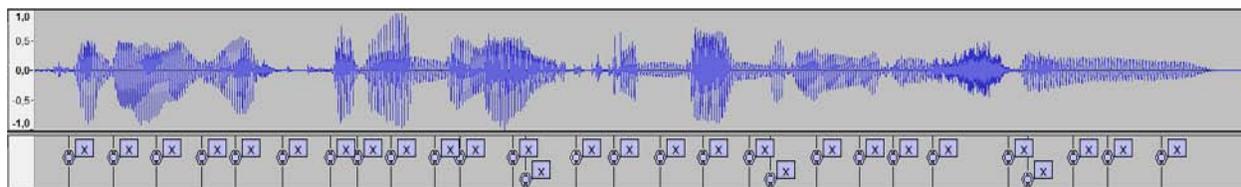


Рис. 1. Сигнал, сегментированный с использованием разработанного алгоритма



Рис. 2. Сигнал, сегментированный разработанным ПО ОИПО НАН Беларуси

Как видно из представленных рисунков, сегментация сигнала проходит практически идентично. Для более точной сегментации сигнала необходимо экспериментальным путем подбирать начальные параметры для каждого из шагов алгоритма.

СПИСОК ЛИТЕРАТУРЫ

1. Давыдов А.Г. Алгоритмы автоматической сегментации речевых сигналов в задачах распознавания и синтеза речи: к.т.н. – Минск, 2008. — 146 с.
2. Кормен, Т., Лейзерсон, Ч., Ривест, Р., Штайн, К. Глава 15. Динамическое программирование // Алгоритмы: построение и анализ / Под ред. И. В. Красикова. — 2-е изд. — М.: Вильямс, 2005. — 1296 с. — ISBN 5-8459-0857-4
3. Кавальчук, А.Н. (2011), "Формула для перехода из области частот к шкале барков и обратно". А.Н. Кавальчук, А.А. Петровский // Информатика, 2011, 4(32), стр. 71-81
4. Самарский А.А., Николаев Е.С. Методы решения сеточных уравнений. — Москва: Наука, 1978. — 592 с.