



И.Д. Иванова, Н.И. Цупрев, С.А. Шлапаков

Дифференцирование как способ проявления информации в речевом сигнале

Исследования в области систем речевого общения (СРО) начаты в индустриально развитых странах в конце 50-х – начале 60-х годов. Промышленные СРО появились в конце 70-х годов. Первые проекты были ориентированы на разработку методов распознавания отдельных слов, в последующем – анализу подверглась непрерывная речь [1].

В настоящее время наиболее качественное распознавание речи обеспечивают устройства, работающие с ограниченным словарем, ориентированные на одного диктора. Анализ речевых сигналов в большинстве СРО осуществляется с помощью спектрально-полосных анализаторов, выделяющих локальные максимумы в спектре. В некоторых системах анализ речевых сигналов осуществляется путем измерения корреляционных функций, форматных параметров, плотности нулей клиппированного сигнала [2, 3].

Основой для разработки современных СРО являются лингвоакустическая и информационная теории речеобразования и восприятия речи. Лингвистическая теория рассматривает фонетические и просодические характеристики речи, акустическая – акустические признаки фонем и просодем, информационная – структуру речевого сигнала [4].

Известно, что при генерации речевого сигнала реализуются два физических явления: резонанса – в основном для гласных и сонант, и модуляции – для предшествующих гласным согласных фонем (за основу принята "открытость" слогов в русском языке). Экспериментальные данные показывают, что речевой аппарат при его кажущейся "свободе" в изменениях посредством модуляции, настроен на спектр гармонического источника и может изменяться только вместе с ним.

Целесообразно выделить в отдельный класс модуляцию длительности звуков, которая характеризует личность диктора, несет информацию о последнем и для анализа информативности сигнала значения не имеет.

Модуляция переносчиков (контекстное влияние согласных на гласную), как предельный случай амплитудной модуляции, представляет интерес с точки зрения информативности сигнала, характеризуется дискретным процессом включения или выключения ее источника [2].

На рис. 1 представлен исходный сигнал фонемы [а] и слова "Да", произнесенные мужским голосом.

Задача состоит в нахождении способа, позволяющего упростить речевой сигнал, сохраняя его информативность. Для этого было предложено дифференцирование.

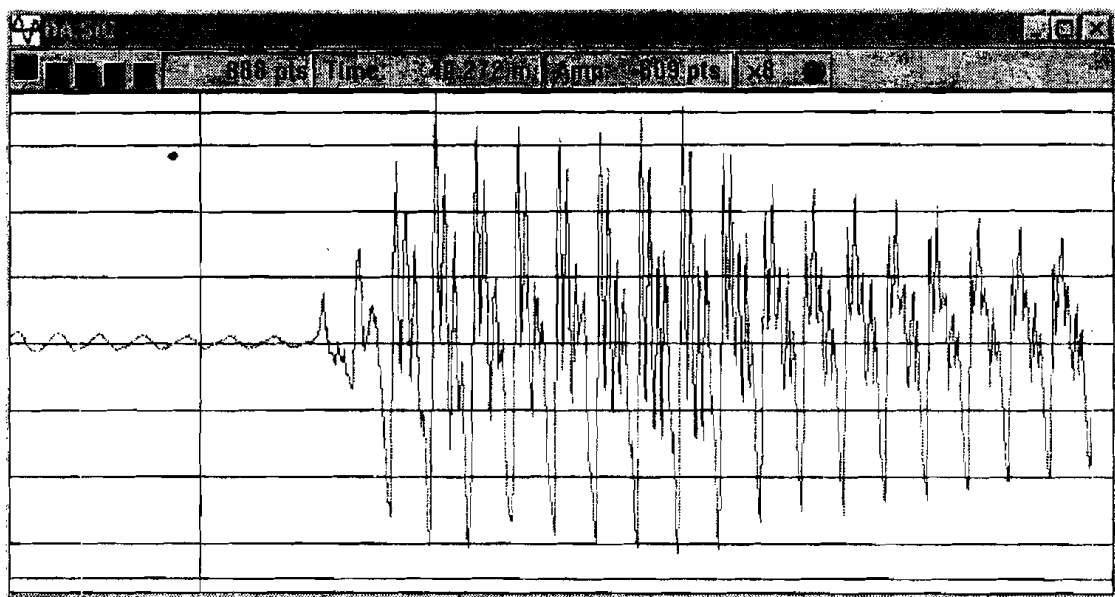
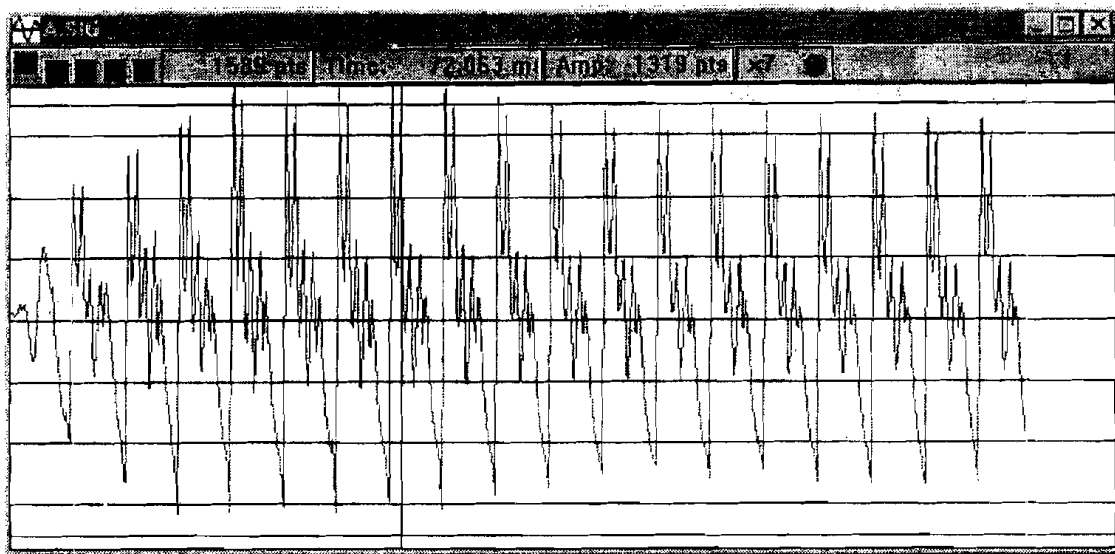


Рис.1. Исходный сигнал фонемы [a] и слова "Да"

Известно, что дифференцирование сигнала изменяет соотношение значений составляющих в спектре по следующему закону:

$$\frac{d^n y(t)}{dt^n} \Rightarrow (j\omega_a)^n \cdot Y(\omega),$$

где $y(t)$ – временное представление сигнала; $Y(\omega)$ – спектральное пред-

ставление сигнала; $\omega_a = 2\pi f_a = 2\pi \frac{1}{t_a}$ – частота первой гармоники, опре-

деляемая из условия согласования интервала анализа t_a параметрам гармонического источника; n – порядок производной.

На рис. 2 представлены вторые производные исходных сигналов фонемы [a] и слова "Да".

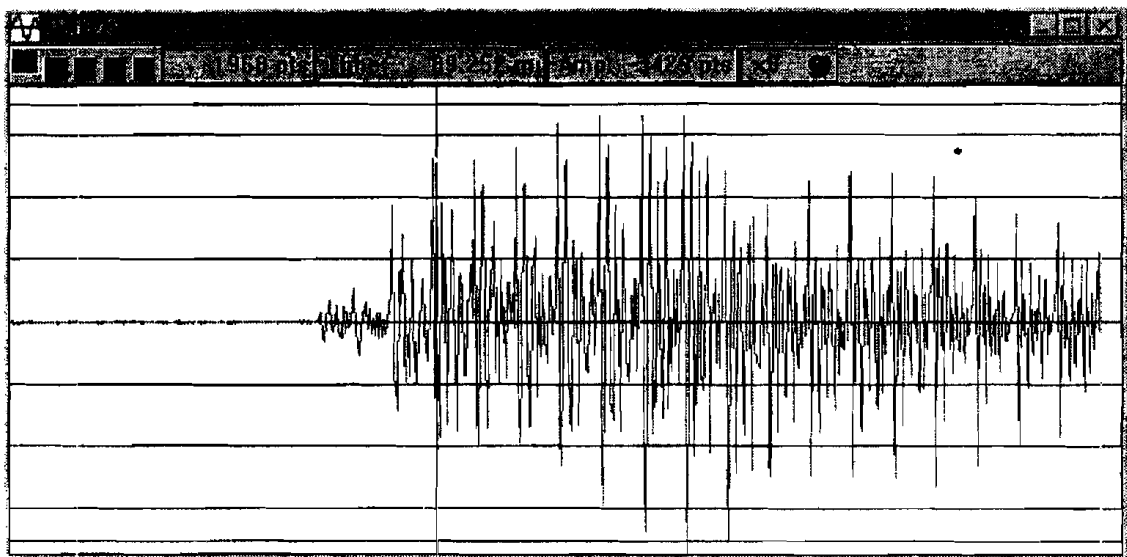
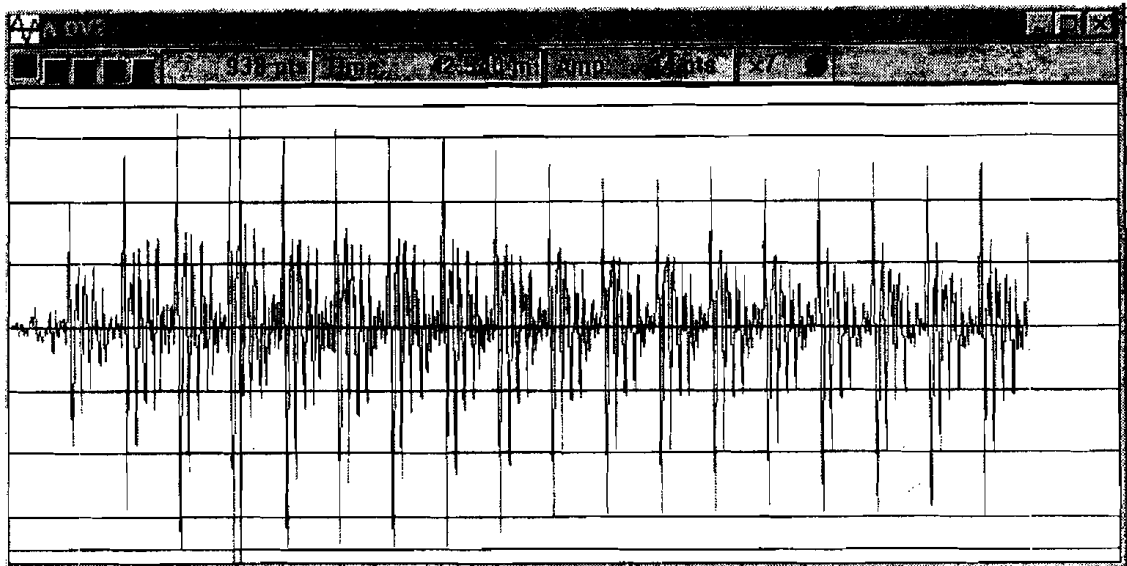
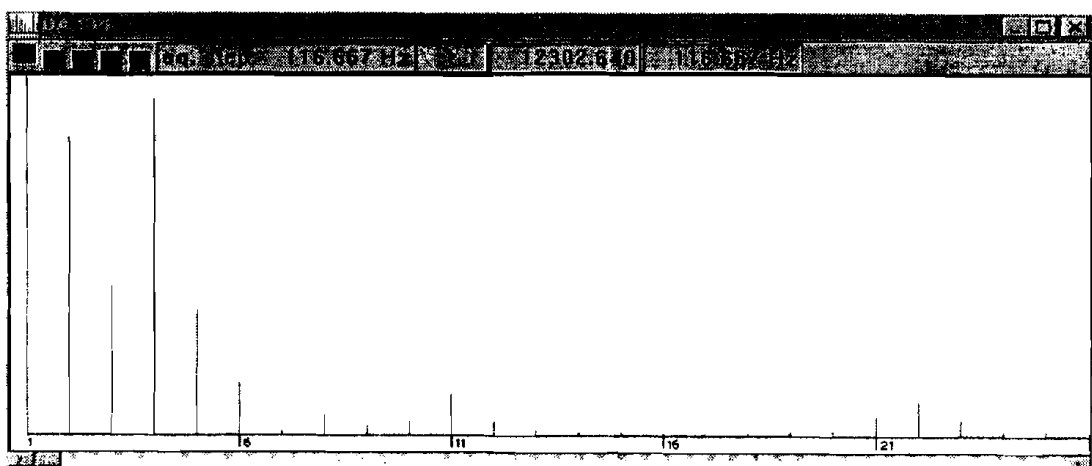


Рис. 2. Вторая производная сигналов фонемы [а] и слова "Да"

На рис. 3 и 4 представлены спектры исходных и продифференцированных сигналов фонемы [а] и слова "Да".



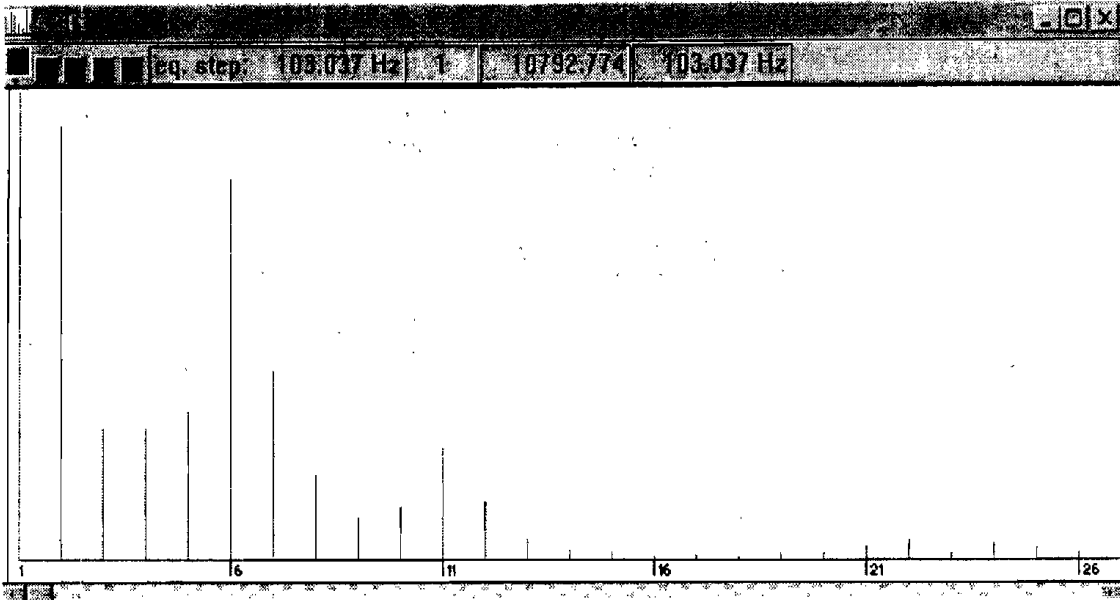


Рис. 3. Спектры исходных сигналов

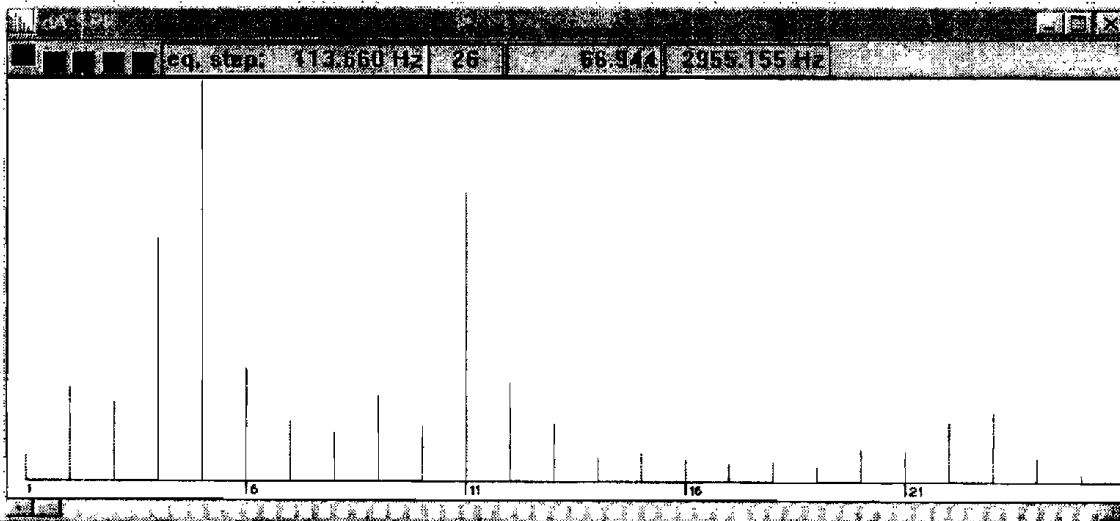
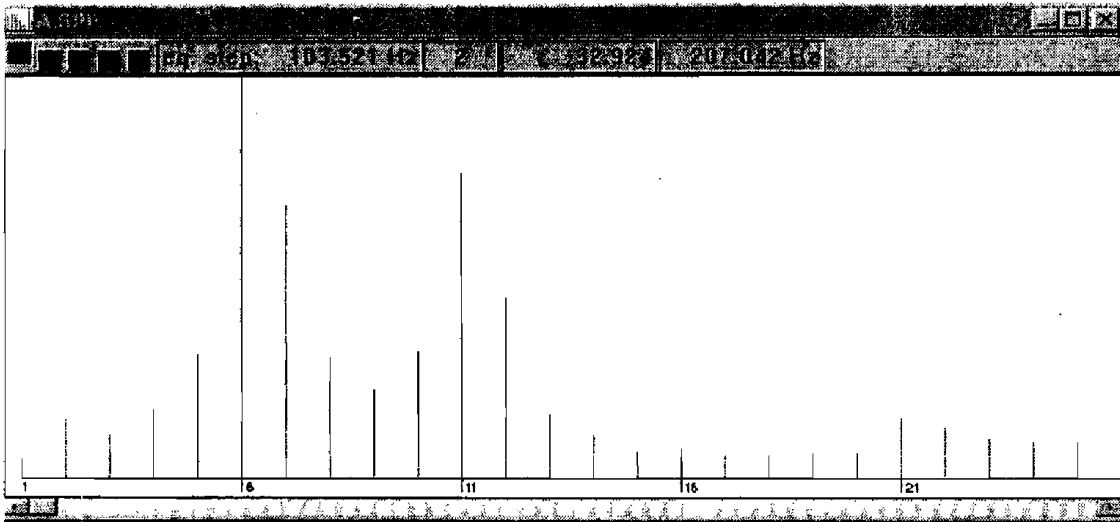


Рис. 4. Спектры проинтегрированных сигналов

Контрольное прослушивание фонемы [а] и слова "Да" до и после дифференцирования показывает полную сохранность информации о произносимом, хотя формы спектральных функций изменились. При этом в сигнале, соответствующем слову "Да", прослушивание участка, который в исходном сигнале определял фонему [а], показывает увеличение влияния фонемы [d] на структуру фонемы [а]. Сам участок предистории (фонема [d]) трансформируется в "нулевой" сигнал [5].

Так как гласные наиболее ярко выражены в сигнале, то нивелированием согласных путем дифференцирования, с одной стороны, осуществляется сегментация сигнала по слогам, с другой стороны, определяется "ядро" анализа, динамика изменения спектра в котором позволяет не только определить и описать гласную, но также и модуляционное воздействие предшествующей согласной.

На основании изложенного выше можно сделать вывод, что с помощью дифференцирования, а именно вычисляя вторую производную, можно решить проблему сегментации речевого сигнала, исследования контекста в речевом сигнале и, в конечном итоге, решить глобальную задачу распознавания речи.

ЛИТЕРАТУРА

1. **Уинстон П.** Искусственный интеллект: Пер. с англ. В.Л.Стафанюка // Под ред. Д.А.Поспелова. М.: Мир, 1980. - 520с.
2. Искусственный интеллект: В 3 кн. Кн. 1. Системы общения и экспертные системы // Справочник / Под ред. **Э.В.Попова**. М.: Радио и связь, 1990. С. 95-106.: ил.
3. **Гольденберг Л.М.** Цифровая обработка сигналов: Учеб. пособие для вузов / **Л.М. Гольденберг, Б.Д. Матюшкин, М.Н. Поляк**. 2-изд., перераб. и доп. М.: Радио и связь, 1990. - 256 с.
4. Фонологические аспекты распознавания речи. **Дж.Е. Шуп** // Методы автоматического распознавания речи. В 2 кн.: Пер. с англ. / Под ред. **У.Ли**. М.: Мир, 1983. Кн.1. - 326 с.
5. **Довнар Д.В., Предко К.Г.** Коррекция оптического изображения с помощью его предварительного дифференцирования // Доклады АН БССР. 1981. Т. XXV. N 2. С. 128-131.

S U M M A R Y

The article deals with the problem of a discernment of speech signals. The main idea is do use a differential as a way to display information.