

Н.И.ЦУПРЕВ

Государственное учреждение высшего профессионального образования  
«БЕЛОРУССКО-РОССИЙСКИЙ УНИВЕРСИТЕТ»  
Могилев, Беларусь

Проблема анализа и распознавания речи как часть общей задачи ввода информации в ЭВМ на естественном языке, является одной из приоритетных проблем, связанных с созданием систем с искусственным интеллектом. Достаточно отметить, что задача создания систем данного класса на ближайшее десятилетие является одной из основных для таких разработчиков и производителей вычислительных средств и информационных технологий как IBM и Microsoft.

Основными направлениями исследований в области речевого общения являются:

- распознавание речевых образов и их интерпретация;
- идентификация и верификация диктора;
- синтез речевых сообщений.

В настоящее время наиболее качественное распознавание речи обеспечивают устройства, работающие с ограниченным словарем, ориентированные на одного диктора.

Основными производителями систем распознавания речи являются фирмы IBM, Centigram Corp., Interstate Electronics, Threshold Technology, Verbex Corp., Votan Inc.(США), NEC (Япония), Marconi, Secure Radio Systems (Великобритания).

Исследования в области синтеза и распознавания речи ведутся в 12 университетах Западной Европы и на них ассигновано более 570 млн долларов США ежегодно.

В большинстве систем речевого общения анализ речевых сигналов с целью их описания осуществляется с помощью оценки фактических характеристик сегментов текущей реализации, расчет которых производится либо с помощью интегральных преобразований (Фурье, Лапласа, Гилберта), либо итерационными методами, либо на основе корреляционных функций. Для устранения временной вариативности речевой реализации применяются методы динамического программирования.

Для анализа и распознавания речевого сигнала в настоящее время используют два основных метода:

- дескрипторный метод - метод сравнения с эталоном, причем эталон строится на основе статистической обработки большого числа реализаций;

– структурный метод - анализ составляющих сигнала в динамике изменения параметров и их взаимодействии.

При этом используются как детерминированные, так и вероятностные характеристики анализируемых сигналов.

Однако ни один из методов анализа речевых сигналов не предлагает способа однозначного и конечного кодового описания таких элементов речи как, например, фонема.

Сигнал по определению - это процесс, отслеживающий изменение некоторого параметра (в других источниках, характеристики) явления реального мира во времени.

Так как явления реального мира, как правило, многофакторные, то сигнал в общем случае может обладать достаточно сложной структурой, которую нельзя описать аналитически. Следовательно, для анализа сигналов используются методы преобразования пространства существования. Например, пространство «значение-время» преобразуется в пространство «значение-аргумент базисной функции».

До настоящего времени, выделенные характеристики использовались в качестве признаков и сравнивались с заранее сформированным «усредненным» описанием.

Предлагается СПОСОБ АНАЛИЗА И МЕТОД ОПИСАНИЯ сигнала путем сравнения его характеристик до и после некоторого преобразования (метод анализа путем оценки реакции характеристик сигнала на воздействие).

Для этого, оценивается воздействие различных аналитических преобразований на сигнал с известными свойствами.

Анализ представленных результатов по оценке реакции на аналитическое воздействие (дифференцирование и усреднение по изменяемому интервалу) позволяет сделать следующие выводы:

– щелевые не реагируют на дифференцирование при частоте дискретизации 22050 Гц.

– на усреднение по минимальному интервалу (два отсчета) – не реагируют гласные и сонанты.

Таким образом, в зависимости от характера генерируемый сигнал реагирует на преобразования строго определенным образом.