

## ТРЕБОВАНИЯ К ДЕТЕКТОРУ ИДЕНТИФИКАЦИИ ДИКТОРА ПО КЛЮЧЕВЫМ СЛОВАМ

М.О. Аль-Хатми, М.Ш. МАХМУД, Д.В. АНАНИЧ, К.А. ОРЛОВСКИЙ

При построении детекторов речи важно учитывать статистические характеристики сигнала (критерий стационарности спектра, энергетический критерий), и разборчивость речевого сигнала на линии. Анализ статистических характеристик спектра речи показал, что имеются различия в спектрах в зависимости от языка [1–4]. В связи с этим была проведена работа по изучению различий в спектре речи различных языков.

При определении критерия стационарности спектра, для короткого временного интервала (порядка 10–20 мс) высчитывалась спектральная мощность. Затем спектр делился на октавные полосы, и для каждой полосы высчитывалась средняя спектральная мощность. Сравнивая среднюю мощность соответствующих октавных полос для соседних временных интервалов, можно оценивать стационарность спектра, а также, устанавливая определенный критерий, принять решение о наличии речи во временном интервале. Важным фактором для данного критерия является выбор октавных полос, которые влияют на разборчивость речи.

При использовании энергетического критерия детектирования речи, задается порог, обычно на уровне порога слышимости, относительно которого определяется наличие речи во входном сигнале. Также при использовании данного критерия порог может задаваться как уровень мощности фонового шума [5]. Для энергетического критерия в детекторах речи существенным фактором является низкочастотная фильтрация сигналов, которая связана со статистическими характеристиками речи, на предварительном этапе обработки сигнала.

Для определения разборчивости речевого сигнала на конкретной линии использовался тональный метод измерения разборчивости речи [6].

В соответствии с методикой была собрана схема для телефонных аппаратов ЦБ. На микрофон подавались сигналы определенных частот с определенными звуковыми давлениями. На приемном конце оператор плавно вводил затухание до исчезновения слышимости звука, затем затухание убавлялось до момента появления звука в телефоне. Среднее значение затухания из этих двух измерений является уровнем ощущения отдельной частотной составляющей спектра речи. Затем при помощи ГОСТ 8031-78 определялась формантная разборчивость речи для каждой частоты.

Исходя из полученных данных, принимается решение о выборе оптимального диапазона (диапазонов) частот, которые будут анализироваться детектором в системе

идентификации личности говорящего по ключевым словам и, соответственно, о выборе ключевых слов.

### **Литература**

1. *Железняк В.К., Макаров Ю.К., Хорев А.А.* // Специальная техника. 2000. № 4. С. 39–45.
2. *Покровский Н.Б.* Расчет и измерение разборчивости речи. М., 1962.
3. ISO/TR 4870: 1991(E) Acoustics – The construction and calibration of speech intelligibility test. P. 22.
4. Pat. US № 5459814. Voice activity detector for speech signals in variable background noise.
5. Pat. US № 6061647. Voice activity detector.
6. ГОСТ 8031-78. Тональный метод измерения разборчивости речи.