

ОПРЕДЕЛИТЕЛЬ РЕЧЕВОЙ АКТИВНОСТИ

Волченков В.А.

Научный руководитель: д-р техн. наук, проф. Витязев В.В.
Рязанский государственный радиотехнический университет, Россия
E-mail: volchenkov.rzn@yandex.ru

Аннотация — Рассмотрен метод определения речевой активности с использованием вспомогательного сигнала при наличии транспортного шума с ОСШ 15 дБ. Приведены результаты детектирования тестового речевого сигнала.

1. Введение

В процессе диалога речевая активность абонента обычно составляет около 50 % времени от длительности обмена. Таким образом, за счет формирования и передачи речевых пакетов (РП), соответствующих только активному состоянию абонента, можно снизить речевую нагрузку, вводимую в сеть, примерно в два раза. Для эффективного решения этой задачи, прежде всего, требуется разделение речевых сигналов на периоды активной речи и паузы. Существующие на сегодняшний день способы определения активности речи позволяют выявить паузы, длительность которых значительно превышает 40 мс [1]. Однако, если проанализировать общую статистику для пауз, то наибольшей плотностью вероятности распределения пауз по длительности оказывается в интервале (5 ... 50) мс и вне этого интервала быстро убывает [2].

2. Основная часть

Для повышения вероятности правильного разделения речевых сигналов на периоды активной речи и паузы предлагается использовать следующий алгоритм (рис. 1) [3].

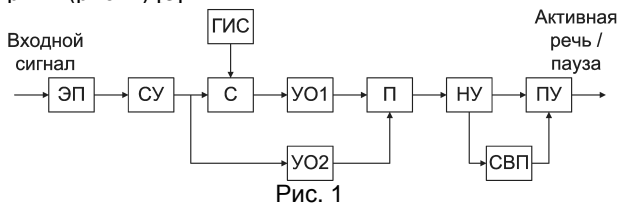


Рис. 1

Речевой сигнал с выхода электроакустического преобразователя (ЭП) усиливается селективным усилителем (СУ) и подается на вход сумматора (С). На второй вход С подается сигнал с выхода генератора измерительного сигнала (ГИС). Суммарный сигнал с выхода С поступает на вход усилителя-ограничителя 1 (УО1), где происходит усиление сигнала, а затем ограничение по амплитуде. Аналогичная операция проводится над сигналом, поступающим с выхода СУ на вход усилителя-ограничителя 2 (УО2). Сигнал с выхода УО1 подается на первый вход перемножителя (П). На второй вход П подается сигнал с выхода УО2. Сигнал с выхода П поступает на вход накопителя-усреднителя (НУ) (интервал накопления 10 мс), где происходит вычисление сигнала, по амплитуде которого принимают решение о наличии периода активного речевого сигнала или паузы в пороговом устройстве (ПУ). Значение порога вычисляется в схеме вычисления порога (СВП) путем анализа первых 150 мс, т.к. на этом интервале речь обычно отсутствует.

Для проведения исследования была выбрана тестовая фраза: «Продолжение отладки устройства». На рис. 2 представлена осциллограмма данной фразы с добавлением транспортного шума (ОСШ 15 дБ), чистого речевого сигнала и сигнала с выхода обнаружителя.

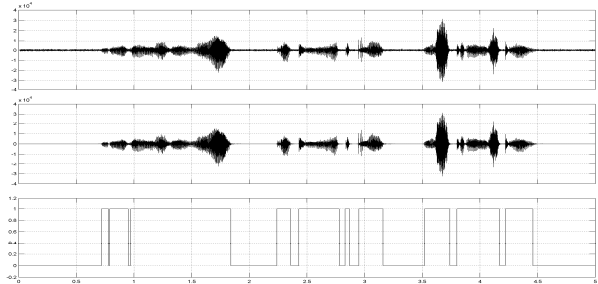


Рис. 2

Общее время записи речевого сигнала указанной фразы составило 5 с, а суммарное время активных периодов — 2,64 с, что составляет 52,8 % времени выбранного речевого сигнала. Количество ошибок обнаружения активной речи составило 3,56 %, количество ошибок обнаружения речевых пауз — 1,61 %.

Для проверки восприятия речи на слух после обработки была произведена запись речевого сигнала с помощью блока аудиозаписи, управляемого командами с обнаружителя пауз. Интервалы, в которых присутствовала речь, оставались без изменения. В интервалы, в которых находились паузы, записывался нуль. Качество полученного речевого сигнала практически не отличалось от исходного. Слова были хорошо различимы, речь легко воспринималась на слух.

3. Заключение

Таким образом, можно сделать вывод о том, что разработан новый способ обнаружения пауз в речи, который позволяет существенно повысить точность разделения на активные участки речи и паузы. При этом качество восстанавливаемой речи остается практически на том же уровне.

Предметом последующего исследования является сравнительная оценка количества ошибок и вероятности правильного обнаружения участков активной речи и паузы при различных уровнях ОСШ.

4. Список литературы

- [1] Вахитов Ш.Я. Акустика: Учебник для вузов / Ш.Я. Вахитов., Ю.А. Ковалгин, А.А. Фадеев, Ю.П. Щевьев; Под ред. профессора Ю.А. Ковалгина. — М.: Горячая линия—Телеком, 2009. — 660 с.
- [2] Шелухин О.И. Цифровая обработка и передача речи / О.И. Шелухин, Н.Ф. Лукьянцев. — М.: Радио и связь, 2000. — 456 с.
- [3] Пат. 2436173 Российская Федерация, МПК G10L 15/00, G10L 11/02, Способ обнаружения пауз в речевых сигналах и устройство его реализующее / Витязев В.В., Розов В.И., Волченков В.А.; заявитель и патентообладатель Рязанский государственный радиотехнический университет. — № 2010124342/08, заяв. 15.06.10; опубл. 10.12.11, Бюл. 34.

VOICE ACTIVITY DETECTOR

Volchenkov V.A.

Scientific adviser: Vityazev V.V.

Ryazan State Radio Engineering University, Russia

Abstract — The new method of the voice activity detection are considered with the use of a special pilot signal. The speech has been mixed with a vehicle noise of 15 dB SNR. The results of the detection are presented.