

УДК 621.391

## ОБНАРУЖЕНИЕ РЕЧЕВОЙ АКТИВНОСТИ НА ФОНЕ ПОМЕХ ПО ИЗМЕНЕНИЮ КВАЗИПЕРИОДА В ДВУХКАНАЛЬНОЙ СХЕМЕ РЕГИСТРАЦИИ ЗВУКОВОГО СИГНАЛА

© 2012 В.Р. Крашенинников, А.П. Ерофеев, В.К. Капырин

Ульяновский государственный технический университет

Поступила в редакцию 02.11.2012

Исследуется способ обнаружения речевой активности по разности длин квазипериодов в двух каналах регистрации звукового сигнала. В одном из каналов регистрируется только шум, а в другом возможно наличие речевого сигнала. Анализируется эффективность данного метода при различных алгоритмах оценивания квазипериода.

Ключевые слова: речевая активность, каналы регистрации, звуковой сигнал

В настоящее время увеличивается интерес к речевым информационно-управляющим системам, в которых запрос на выдачу информации или управление исполнительными механизмами осуществляется по голосовым командам оператора. Основной проблемой, возникающей при разработке этих систем, является распознавание речевых команд (РК) оператора, которое особенно затруднено наличием сильных акустических помех. По этой проблеме имеется очень много работ, например [1-2]. На качество распознавания РК очень сильно влияет точность определения границ РК, так как при неточном определении этих границ теряется часть команды или же к ней добавляется посторонний отрезок шума. Нахождение границ РК является частным случаем задачи обнаружения сигнала на фоне помех, решению которой посвящено огромное количество работ.

В [3] был предложен следующий метод обнаружения речевой активности по двухканальной схеме. С первого микрофона (первый канал) принимается сигнал  $Z = \{z_1, z_2, \dots, z_N\}$ , содержащий шум  $\Theta = \{\theta_1, \theta_2, \dots, \theta_N\}$  и, возможно, речевой сигнал (РС), положение которого неизвестно. Со второго микрофона (второй канал) принимается сигнал, содержащий только шум. При этом шумовая составляющая сигнала  $Z$  в силу различных условий регистрации (местоположение микрофонов, различие их характеристик и т.д.) может отличаться от шума  $\Theta$ . Однако шумовые составляющие в обоих каналах происходят от одного источника (например, от

двигателя самолёта), поэтому они могут иметь некоторые близкие характеристики, когда РС в первом канале отсутствует. При наличии РС можно ожидать возникновения разницы этих характеристик в первом канале по отношению ко второму. Наличие значимой разницы является признаком речевой активности.

В качестве одной из таких характеристик в [3] использована длина квазипериода сигнала в скользящем окне. Во многих практических случаях акустические шумы на небольших отрезках времени являются квазипериодическими сигналами. То же самое относится и к РС, при этом длины квазипериодов шума и РС обычно не совпадают, поэтому при возникновении речевой активности в первом канале может заметно изменить квазипериод суммарного сигнала. Это позволяет использовать разницу между длинами квазипериодов сигналов в двух каналах как признак наличия речевой активности. Оценка квазипериода в [3] находилась по минимуму суммы модулей разностей отсчётов сигнала со сдвигом на пробный период, то есть в качестве оценки  $\hat{T}$  квазипериода принималась точка минимума статистики

$$S_1(T) = \sum_{n=1}^N |z_n - z_{n+T}|, \quad (1)$$

где  $T$  – пробный период,  $N$  – длина скользящего окна. Решение о наличии речевой активности в скользящем окне принималось, когда разница между оценками квазипериодов сигналов в двух каналах превышала установленный порог.

Кроме оценки квазипериода по статистике (1), существует ряд других методов [4]. В настоящей работе исследуется эффективность обнаружения речевой активности по разнице квазипериодов при использовании различных мето-

*Крашенинников Виктор Ростиславович, доктор технических наук, профессор, заведующий кафедрой «Прикладная математика и информатика».*

*E-mail: kvr@ulstu.ru*

*Ерофеев Александр Павлович, аспирант.*

*E-mail: a.p.erofeev@gmail.com*

*Капырин Владимир Константинович, аспирант.*

*E-mail: kapirin@inbox.ru*

**Таблица.** СКО ошибки обнаружения начала и конца речевых команд

Соотношение сигнал/шум	Метод	Минимум суммы модулей разности	Максимум ковариации	Пик спектра	Преобразование Брукса
3	Начало команды	2074	1978	2981	2303
	Конец команды	4241	2822	3305	3620
1	Начало команды	5937	5380	8856	6548
	Конец команды	9907	9103	9000	9478

дов оценивания квазипериода.

Для оценивания длины квазипериода можно использовать автоковариационную функцию

$$S_2(T) = \sum_{n=0}^{N-1-T} z_n z_{n+T} \quad (2)$$

случайного процесса, являющуюся оценкой ковариации между отсчетами процесса, сдвинутыми на  $T$ . Отправной идеей при использовании этого метода служит предположение о высокой корреляции между отсчетами квазипериодического сигнала, отстоящими друг от друга на квазипериод. Поэтому точки максимума статистики (2) будут кратны квазипериоду сигнала.

Если исходить из предположения, что частоты вблизи частоты основного тона сигнала несут наибольшую энергию, то частота основного тона соответствует пиковому значению спектра сигнала. Поскольку точный спектр сигнала неизвестен, он оценивается по реализации сигнала (по наблюдениям  $Z$ ), поэтому положение пика оказывается нестабильным. Для улучшения оценки частоты основного тона применяется сглаживание спектра, например, вычисление суммарной мощности в скользящем окне. Для оценки спектра сигнала может быть применено непосредственно его преобразование Фурье или БПФ.

Метод оценивания квазипериода, использующий преобразование Брукса, состоит в усреднении исходного сигнала и дальнейшего рекурсивного применения к процессу  $z'_n$  разностного преобразования

$$z_n^{(m)} = \frac{1}{2} (z_n^{(m-1)} - z_{n+T}^{(m-1)}), \quad (3)$$

частотная характеристика которого есть

$$R(\omega) \approx \frac{\sin^m \lambda \omega}{\lambda \omega}, \quad (4)$$

где  $\lambda = T/2N$ . Задавая различные значения  $T$ , можно выделять частоты в узком диапазоне

(увеличивая  $m$ ), то есть синтезировать узкополосные фильтры.

Для оценки эффективности обнаружения речевой активности при перечисленных методах оценивания квазипериода были проведены эксперименты, в которых использовано 50 записей с двух микрофонов, в которых содержались РК авиационной тематики. В первый микрофон поступал РС, смешанный с шумом, во второй – только шум. В качестве источника шума был использован пылесос. Акустические сигналы оцифровывались с частотой 44кГц. В дальнейшем аудио-визуальным способом определялись истинные границы РК. Полученный материал подвергался обработке программами, реализующими описанные методы, и определялась ошибка обнаружения начала и конца каждой РК.

В таблице приведены СКО ошибок обнаружения начала и конца команд в зависимости от метода оценивания квазипериода и отношения сигнал/шум.

Результаты экспериментов показывают, что наиболее точное обнаружение границ РК получается при оценивании квазипериода по максимуму корреляции. Концы команд обнаруживаются значительно хуже, чем их начала. Это объясняется тем, что люди произносят начала слов более чётко и громко, чем их окончания.

### СПИСОК ЛИТЕРАТУРЫ

1. *Потапова Р.К.* Речевое управление роботом. М.: Радио и связь, 1989. 248с.
2. *Сапожков М.А.* Речевой сигнал в кибернетике и связи. М.: Связьиздат, 1963. 452 с.
3. *Крашенинников В.Р., Хвостов А.В.* Обнаружение речевой активности в двухканальной системе распознавания речевых команд // Труды Российского научно-технического общества радиотехники, электроники и связи имени А.С. Попова. Серия: научная сессия, посвященная дню радио. Выпуск LXIII. Москва, 2008. С. 121-123.
4. *Серебренников М.Г., Первозванский А.А.* скрытых периодичностей. М.: Наука, 1965. 244 с.

**TWO-CHANNEL VOICE ACTIVITY DETECTION  
IN NOISE BASED ON VARIATION OF QUASIPERIOD**

© 2012 V.R. Krasheninnikov, A.P. Erofeev, V.K. Kapirin

Ulianovsk State Technical University

Method of speech activity detection based on quasiperiod differences in two-channel sound records is examined. One channel contains only noise, while another can contain speech signal. Efficiency of this method given different quasiperiod length estimating algorithms was analyzed.

Keywords: speech activity, channels recording, the sound

---

*Victor Krasheninnikov, Doctor of Technical Sciences,  
Professor, Head at the Applied Mathematics and Computer  
Science Department. E-mail: kvr@ulstu.ru  
Alexander Erofeev, Post-Graduate Student.  
E-mail: a.p.erofeev@gmail.com  
Vladimir Kapirin, Post-Graduate Student.  
E-mail: kapirin@inbox.ru*