



УДК 004.896

## ВЫБОР И ОБОСНОВАНИЕ ПАРАМЕТРОВ РЕШАЮЩЕЙ ФУНКЦИИ ДЛЯ ОБЕСПЕЧЕНИЯ МАКСИМАЛЬНОГО РАЗЛИЧИЯ ИНФОРМАЦИОННЫХ И НЕИНФОРМАЦИОННЫХ КОМПОНЕНТ РЕЧЕВЫХ СИГНАЛОВ В ЧАСТОТНОЙ ОБЛАСТИ

**А. В. ГЛУШАК**  
**С. П. БЕЛОВ**  
**А. С. БЕЛОВ**  
**А. В. КУРЛОВ**

*Белгородский  
государственный  
университет*

*e-mail: Belov@bsu.edu.ru*

*e-mail: Belov\_AS@bsu.edu.ru*

В статье обосновывается выбор параметров решающей функции на основе вычислительных экспериментов, которые при заданной вероятности ошибок первого рода обеспечивают максимальную вероятность селекции информационных и неинформационных компонент речевых сигналов, что позволяет с высокой степенью точности обнаруживать отсутствие в анализируемом отрезке речевого сигнала энергий, обусловленных наличием речевых компонент (шумы регистрирующих приборов в паузах).

Ключевые слова: частотный анализ, кратковременный спектр, речевой сигнал, решающая функция.

### Введение

Одним из путей минимизации требуемых ресурсов каналов радиосвязи при передаче речевых сигналов является уменьшение объема их битового представления за счет обнаружения пауз с последующим кодированием. Эффективность этого подхода в значительной степени зависит от точности определения отсутствия в анализируемом отрезке речевого сигнала энергий, обусловленных наличием речевых компонент (шумы регистрирующих приборов в паузах). В [1] предлагается осуществлять реализацию указанной процедуры, используя знания об особенностях распределения энергии информационных (содержащих полезную информацию) и неинформационных (обусловленных шумами) компонент анализируемых речевых сигналов в частотной области. Однако, в зависимости от длительности обрабатываемого отрезка сигнала и выбранного количества частотных интервалов, в которых осуществляется вычисления значений нормированных частей распределения его энергии в этих интервалах, правильность принятия решения о принадлежности анализируемого участка к содержащему полезную информацию или порожденному шумами может существенно меняться.

### Выбор и обоснование параметров решающей функции

В связи с этим в статье с использованием в качестве модели решающей функции (РФ) при селекции информационных и неинформационных компонент характеристики вида определяются

$$Z_N = \max P_{rN} / S_{rN}, 1 \leq r \leq R, \quad (1)$$

где  $R$  – выбранное количество частотных интервалов анализа;  $N$  – количество отсчетов в анализируемом отрезке речевого сигнала;  $P_{rN}$  – часть энергии анализируемого отрезка речевого сигнала в  $r$ -м частотном интервале, длительность которого равна  $(v_{r+1} - v_r)$ ;  $S_{rN}$  – оценка математического ожидания соответствующей части энергии отрезка сигнала в  $r$ -м частотном интервале, обусловленном только шумами;



были проведены детальные исследования выбора наиболее подходящих компромиссных сочетаний значений длин обрабатываемых отрезков и количества используемых при этом частотных интервалов.

Ясно, что такой подход адекватен ситуации стационарного шума, когда возможность оценивания знаменателей  $S_{rN}$  в модели вида (1) имеется. Эти оценки получались усреднением по большому количеству отрезков записанного сигнала, которые заведомо не содержали информационных компонент в виде звуков речи.

Прежде всего, осуществлялось оценивание относительного количества пересечения правой частью (1) некоторых порогов, что при отсутствии информационных компонент соответствует оцениванию вероятностей ошибок первого рода (ложных тревог). В результате выполнения большого объема вычислительных экспериментов было показано, что можно указать такую величину порога, частота пересечений которого при отсутствии информационных компонент (оценка вероятности ложных тревог) стабильна и может быть достаточно мала при различных сочетаниях (с выполнением неравенства (1)) длительностей отрезков шумов и количества частотных интервалов.

С другой стороны, очень важным является отсутствие искажений в управляющих речевых сигналах, обусловленных принятием информационной компоненты за неинформационную, ошибки второго рода (пропуск сигнала). Ясно, что вероятности ошибок второго рода будут зависеть от чувствительности РФ к наличию информационных компонент.

Таким образом, становится важным установление потенциальных возможностей использования характеристики (1) в качестве модели решающей функции (РФ) при селекции информационных компонент. В данной статье, как указано выше, были проведены детальные исследования проблемы выбора наиболее подходящих компромиссных сочетаний значений длин обрабатываемых отрезков и количества используемых при этом частотных интервалов.

Проблема поиска компромисса возникает в силу того, что для минимизации искажений из-за неточности определения границ пауза/звук необходимо использовать как можно меньшие по длительности отрезки для анализа.

В свою очередь, требование достижения высокой детальности в анализе распределения частей энергии по частотным интервалам приводит к необходимости использования как можно большего их количества.

Вместе с тем, величина собственных чисел, индексы которых превосходят значение

$$J_r = \left[ \frac{N^* (\nu_{r+1} - \nu_r)}{\pi} \right] + 2, \quad (2)$$

(здесь выражение в скобках – целая часть числа), пренебрежимо мала по сравнению с единицей, что позволяет предложить упрощенные процедуры вычисления частей энергий исходных отрезков звуковых сигналов. Однако, необходимо отметить, что выражение (2) говорит о том, что имеется нижняя граница отношения длительности анализируемого отрезка к количеству частотных интервалов, которую необходимо соблюдать.

С целью поиска компромисса в статье для всех звуков русской речи были проведены вычислительные эксперименты по оценке вероятности обнаружения на основе модели (1) границ пауза/звук.

Для целей этого исследования и дальнейших применений, предлагаемых РФ была разработана итерационная процедура обработки неинформационных данных с целью установления значения порогов частоты пересечений, которые близки к задаваемым вероятностям ошибок первого рода при обнаружении пауз (ложных тревог).



В результате вычислительных экспериментов показано, что для всех звуков русской речи, записанной на диктофон в процессе чтения в аудитории лекции, наблюдалось уверенное пересечение порогов, определённых из условия частоты ложных тревог менее 0,005. На помещённом ниже рис. 1 приведен график поведения характеристики (1) как при отсутствии информационных компонент, так и при наличии их.

Этот график иллюстрирует справедливость вывода о высокой чувствительности правой части (1) к возникновению информационной компоненты, тогда как при её отсутствии она ведёт себя стабильно ниже порога.

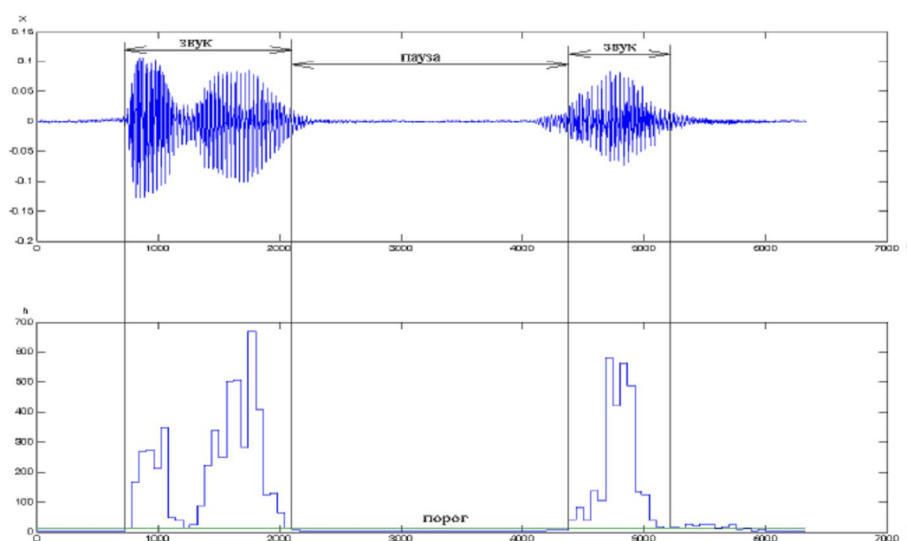


Рис. 1. Определение границ пауза/звук при длине отрезка сигнала  $N=60$ ;  $R=10$

В табл. 1-3 приведены результаты вычислительных экспериментов по обработке речевых сигналов, зарегистрированных для 3 дикторов в различной шумовой обстановке. Данные таблиц получены при обработке выборки объёмом 10000. Используемые обозначения:  $h_n$  – величина порога, обеспечивающего заданный уровень вероятности ошибок первого рода;  $\alpha_{no}$  и  $\alpha_{lo}$  – оценки относительных количеств превышений порога решающей функцией при наличии и отсутствии информационной компоненты соответственно.

Они иллюстрируют вывод о высоких потенциальных возможностях селекции информационных и неинформационных компонент на основе модели РФ вида (1) при длительностях обрабатываемых отрезков 60 отсчётов и 10 частотных интервалов для анализа. Это соответствует длительности анализируемого отрезка сигнала порядка 8 миллисекунд и ширине частотных интервалов 400 Герц, что представляется разумным компромиссом. Поэтому такое сочетание этих параметров модели рекомендуется к использованию при обработке речевых сигналов.

### Заключение

В результате проведенных исследований с использованием в качестве модели решающей функции (РФ) характеристики вида (1) были показаны ее высокие потенциальные возможности при селекции информационных и неинформационных компонент. Было показано, что наиболее подходящим компромиссным вариантом меж-



ду требованиями минимальности длительности обрабатываемого отрезка сигнала и выбранного количества частотных интервалов может служить 60 отсчётов и 10 частотных интервалов.

Таблица 1

## Диктор 1

| Номер эксперимента | N   | R  | h <sub>а</sub> | $\alpha_{no}$ | $\alpha_{ло}$ |
|--------------------|-----|----|----------------|---------------|---------------|
| 1                  | 20  | 2  | 6,9            | 0,98          | 0,01          |
| 2                  | 20  | 5  | 8,9            | 0,98          | 0,01          |
| 3                  | 20  | 10 | 10,6           | 0,98          | 0,01          |
| 4                  | 60  | 2  | 5,1            | 0,98          | 0,01          |
| 5                  | 60  | 6  | 7,1            | 0,98          | 0,01          |
| 6                  | 60  | 10 | 8,3            | 0,99          | 0,01          |
| 7                  | 60  | 15 | 11,5           | 0,99          | 0,01          |
| 8                  | 60  | 30 | 30             | 0,98          | 0,01          |
| 9                  | 200 | 2  | 5,6            | 0,99          | 0,01          |
| 10                 | 200 | 20 | 12,8           | 0,99          | 0,01          |

Таблица 2

## Диктор 2

| Номер эксперимента | N   | R  | h <sub>а</sub> | $\alpha_{no}$ | $\alpha_{ло}$ |
|--------------------|-----|----|----------------|---------------|---------------|
| 1                  | 20  | 2  | 14,53          | 0,88          | 0,01          |
| 2                  | 20  | 5  | 14,74          | 0,94          | 0,01          |
| 3                  | 20  | 10 | 16,28          | 0,98          | 0,01          |
| 4                  | 60  | 2  | 12,20          | 0,93          | 0,01          |
| 5                  | 60  | 6  | 12,39          | 0,95          | 0,01          |
| 6                  | 60  | 10 | 12,57          | 0,99          | 0,01          |
| 7                  | 60  | 15 | 12,75          | 0,99          | 0,01          |
| 8                  | 60  | 30 | 12,89          | 0,99          | 0,01          |
| 9                  | 200 | 2  | 8,08           | 0,99          | 0,01          |
| 10                 | 200 | 20 | 8,51           | 0,99          | 0,01          |

Таблица 3

## Диктор 3

| Номер эксперимента | N   | R  | h <sub>а</sub> | $\alpha_{no}$ | $\alpha_{ло}$ |
|--------------------|-----|----|----------------|---------------|---------------|
| 1                  | 20  | 2  | 4,03           | 0,96          | 0,01          |
| 2                  | 20  | 5  | 5,17           | 0,98          | 0,01          |
| 3                  | 20  | 10 | 8,09           | 0,99          | 0,01          |
| 4                  | 60  | 2  | 2,16           | 0,99          | 0,01          |
| 5                  | 60  | 6  | 6,91           | 0,98          | 0,02          |
| 6                  | 60  | 10 | 6,70           | 0,99          | 0,02          |
| 7                  | 60  | 15 | 8,19           | 0,99          | 0,02          |
| 8                  | 60  | 30 | 10,42          | 0,99          | 0,02          |
| 9                  | 200 | 2  | 1,27           | 0,99          | 0,01          |
| 10                 | 200 | 20 | 7,48           | 0,99          | 0,01          |

## Литература

1. Белов С. П., Белов А. С. О потенциальных возможностях селекции информационных и неинформационных компонент речевых сигналов // Тезисы докладов Междунар. науч.-практ. конф. (Винница, Украина 19-21 мая 2010 года).



2. Жилияков, Е.Г. Уменьшение объема битового представления речевых данных на основе нового метода удаления пауз / Е.Г. Жилияков, С.П. Белов, Е.И. Прохоренко // Вопросы радиоэлектроники. Серия «Электронная вычислительная техника (ЭВТ)». – М., 2007. – Вып. 2. – С. 82-92.

3. Жилияков, Е.Г. Вариационные методы частотного анализа звуковых сигналов [Текст] / Е.Г. Жилияков, С.П.Белов, Е.И. Прохоренко // Труды учебных заведений связи. – СПб., 2006. вып. 174. – С. 163-172.

4. Белов С. П., Белов А. С., Жилияков Е.Г., Прохоренко Е.И. Способ обнаружения пауз в речевых сигналах и устройство его реализующее. – Патент России № 2317595 от 20 февраля 2008 года.

*Исследования выполнены при поддержке ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009-2013 годы, Государственный контракт № П964 от 27 мая 2010 г., проект «Разработка методов и алгоритмов передачи речевой информации с минимизацией требуемых ресурсов каналов радиосвязи».*

## **CHOICE AND SUBSTANTIATION OF PARAMETERS OF DECISION FUNCTION FOR MAINTENANCE OF THE MAXIMUM DISTINCTION INFORMATION AND NOT INFORMATION A COMPONENT OF SPEECH SIGNALS IN FREQUENCY AREA**

**A. V. GLUSHAK  
S. P. BELOV  
A. S. BELOV  
A. V. CURLOV**

*Belgorod State University*

*e-mail: Belov@bsu.edu.ru*

*e-mail: Belov\_AS@bsu.edu.ru*

In article on the basis of computing experiments the choice of parameters of decision function which at the set probability of errors of the first sort provide the maximum probability of selection information and not information a component of speech signals that allows with a high degree of accuracy to detect the absence of a test segment of speech signal energy due to the presence of speech components (noise recording devices in the pauses).

Key words: the frequency analysis, a short-term spectrum, a speech signal, decision function.