

Федеральное государственное автономное образовательное учреждение высшего профессионального образования  
«БЕЛГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ НАЦИОНАЛЬНЫЙ  
ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ»  
(НИУ «БелГУ»)

Факультет компьютерных наук и телекоммуникаций  
Кафедра информационно-телекоммуникационных систем и технологий

## **КУРСОВАЯ РАБОТА**

по дисциплине «Компьютерная обработка аудиоданных»  
на тему:

**«Сегментация речевых сигналов с использованием  
кепстрального анализа»**

1403.210406.238.ПЗКР

Выполнил  
студент группы 140608

Д.В. Попов

Руководитель  
кандидат техн. наук,  
доцент кафедры ИТСиТ

Е.И. Прохоренко

# Актуальность

На всем множестве разнообразных задач обработки речи (распознавание речи, компрессия речи, создание обучающих баз, идентификация диктора по голосу и т. д.) исследователи неизбежно сталкиваются с проблемой сегментации речи. Сегментация речи - это процесс поиска границ между фразами, словами, слогами или артикуляторно-акустическими сегментами речевого сигнала. Этот термин применяется как к мыслительному процессу человека, так и к процессу автоматической сегментации, выполняемой машинами.

Целесообразность того или иного типа сегментации определяется: конкретной задачей обработки речи; моделью, выбранной для решения этой задачи; требованиями к точности и времени работы системы, реализующей модель.

В настоящее время в современных АТС все большее развитие набирает услуги в той или иной степени затрагивающие голосовое управление, это обуславливает необходимость поиска надежных методов и алгоритмов сегментации речевых сигналов как неотъемлемой составляющей распознавания речи, способных с высокой точностью выделять необходимые составляющие речевого сигнала.

Известные по литературе методы поиска границ сегментов обладают значительными недостатками. В частности, эти методы либо используют априорную информацию о содержании речевого сигнала, которая обычно недоступна, либо производят поиск границ слишком грубых элементов - слогов, слов, предложений, либо дают слишком большие погрешности. В свою очередь, описанные в литературе методы распознавания типа сегментов также обладают различными недостатками и не позволяют решать реальные речевые задачи: они не обладают достаточной точностью, не выполняют распознавания кардинальных типов речевых сегментов.

# Цели и задачи

Целью данной работы является разработка и исследование алгоритма сегментации речевых сигналов с использованием кепстрального анализа.

Задачи работы:

- Обзор методов сегментации речевых сигналов ;
- Исследование теоретических основ кепстрального анализа;
- Разработка и исследование алгоритма сегментации речевых сигналов с использованием кепстрального анализа.

## Основы кепстрального анализа

Кепстральные коэффициенты являются результатом применения обратного преобразования Фурье к логарифмированному энергетическому спектру.

$$c_j = \sum_{i=1}^M x_i \cos\left(j \cdot (i - 0.5) \cdot \frac{\pi}{M}\right), j = 1..J \quad (1)$$

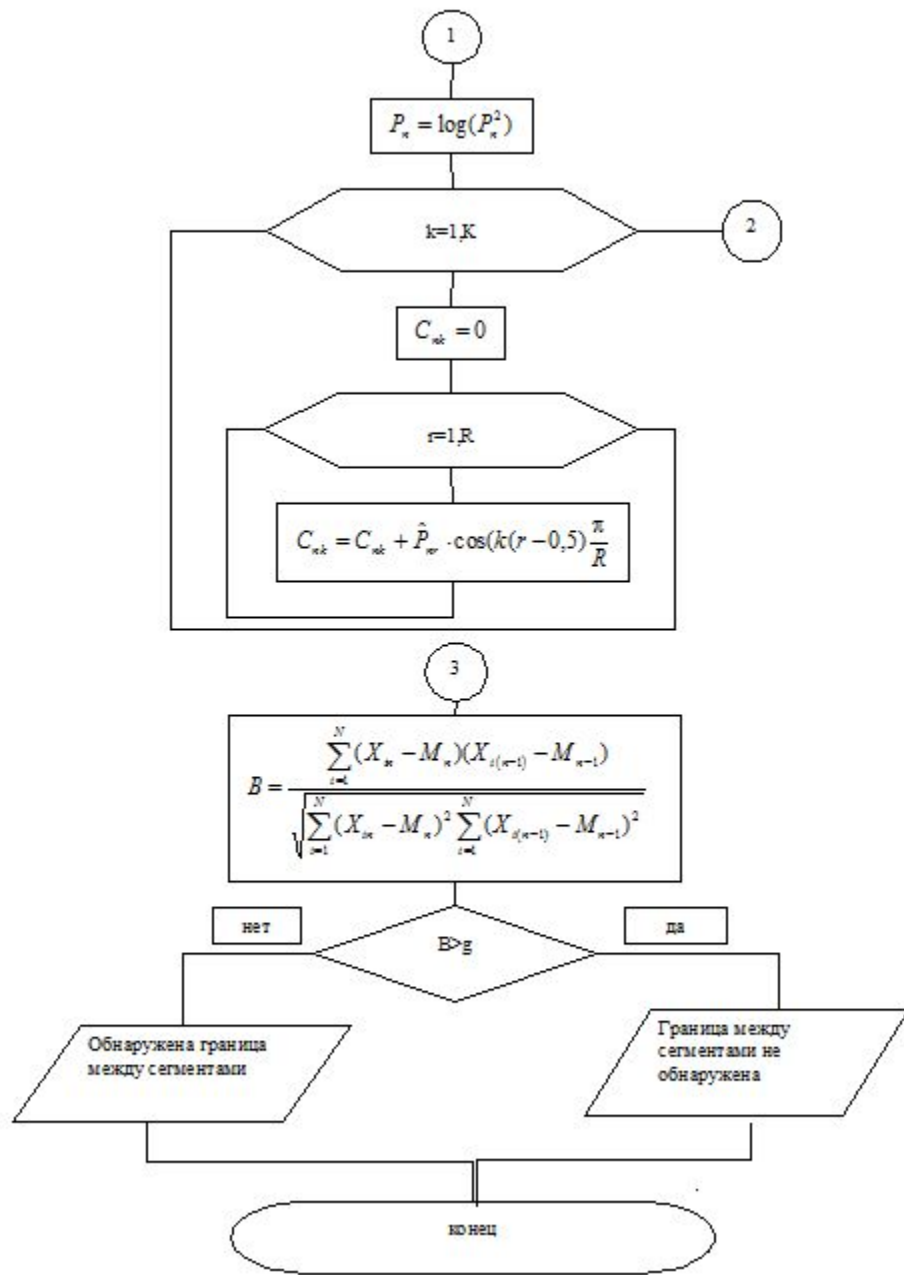
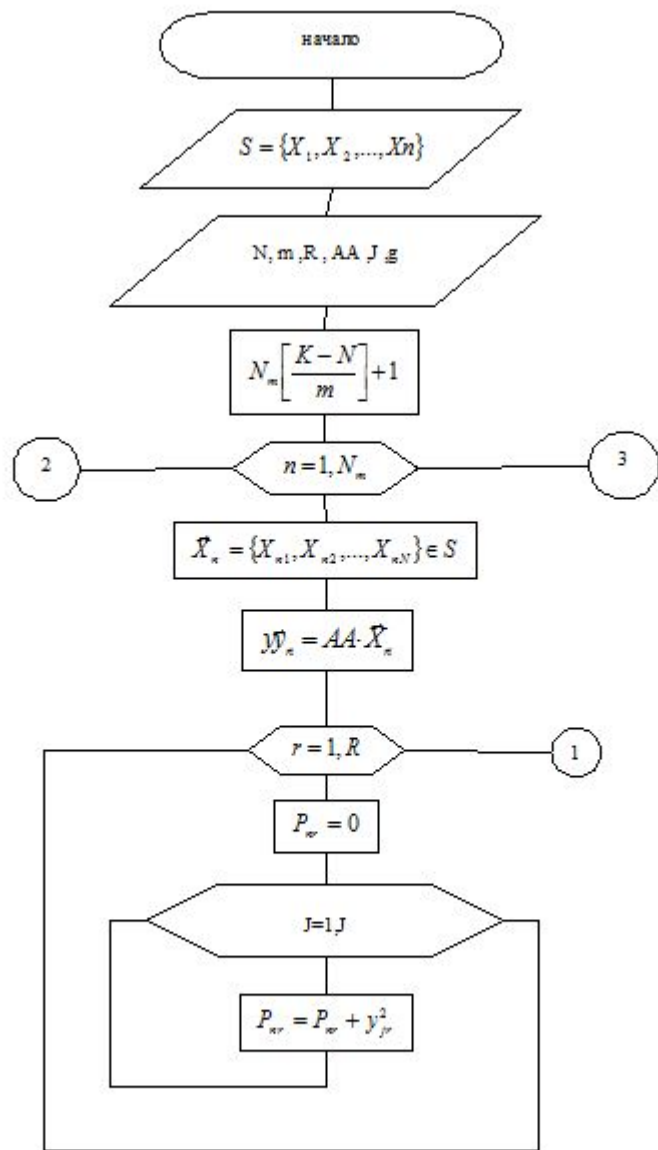
Удаление «сверточных» искажений в сигнале осуществляется путем вычитания из кепстральных характеристик их среднего значения (математического ожидания):

$$\bar{c}_t = c_t - \frac{1}{T} \sum_{i=0}^{T-1} c_i \quad (2)$$

Метод маскировки шума заключается в добавлении некоторой константы  $C$  к спектральным коэффициентам при вычислении кепстра

$$c = \text{DCT}(\log(C + x e(j\omega))) \quad (3)$$

# Блок-схема алгоритма сегментации речевых сигналов с использованием кепстрального анализа



# Результаты работы алгоритма сегментации

На рисунках 1-2 представлены результаты алгоритма сегментации для слова «характеристика»

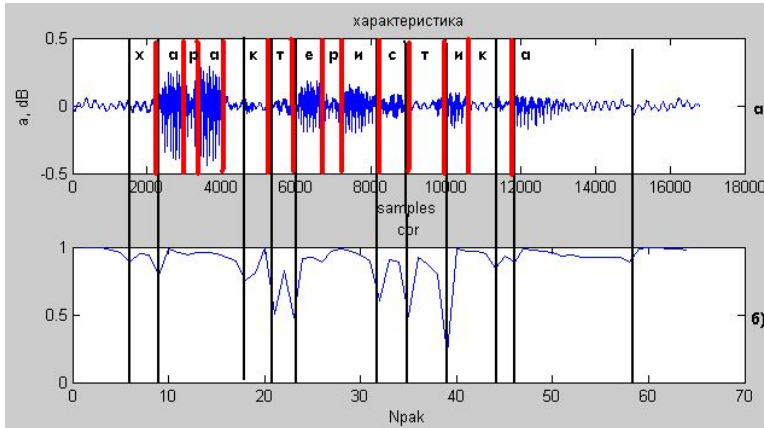


Рис.1-Результат сегментации слова «характеристика» при  $N=256$  и  $g=0.9$

Количество границ сегментов определенных «на слух»---13  
Количество правильно определенных границ сегментов---7  
количество ошибок первого рода -----4  
Количество ошибок второго рода -----6

Максимальное количество границ сегментов для слова «характеристика» определено при  $n=256$  и  $g=0.9$

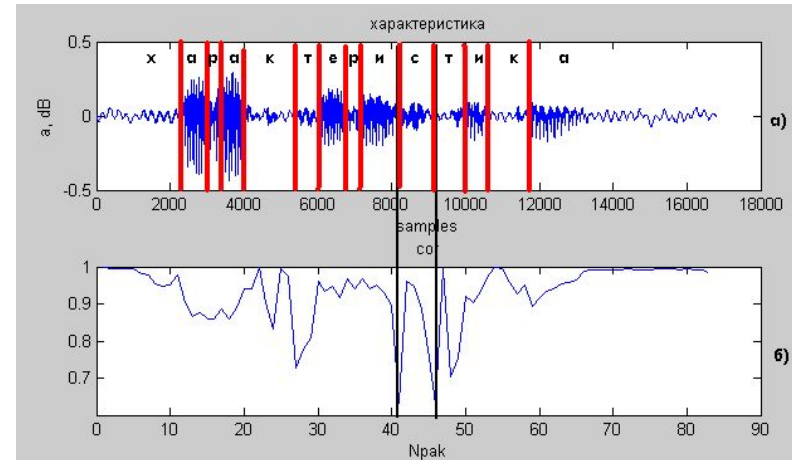


Рис.2-Результат сегментации слова «характеристика» при  $N=200$  и  $g=0.7$

Количество границ сегментов определенных «на слух»---13  
Количество правильно определенных границ сегментов---2  
количество ошибок первого рода -----0  
Количество ошибок второго рода -----11

Минимальное количество границ сегментов для слова «характеристика» определено при  $n=256$  и  $g=0.9$

## Результаты алгоритма сегментации речевых сигналов с использованием кепстрального анализа

В таблице 1 приведены вероятности обнаружения границ сегментов, а так же вероятности ошибок первого и второго родов в зависимости от значения порога корреляции и длины фреймов.

**Таблица 1-результаты сегментации**

N	100	150	200	256	N	100	150	200	256
<b>g=0.7</b>									
<b>P'</b>	0,38	0,22	0,24	0,14	<b>M'</b>	30	17	19	19
<b>P''</b>	0,14	0,13	0,004	0,19	<b>M''</b>	11	10	3	15
<b>P'''</b>	0,61	0,77	0,75	0,75	<b>M'''</b>	48	61	59	59
<b>g=0.8</b>									
<b>P'</b>	0,51	0,3	0,34	0,39	<b>M'</b>	30	24	27	31
<b>P''</b>	0,25	0,23	0,1	0,2	<b>M''</b>	20	18	8	16
<b>P'''</b>	0,47	0,84	0,65	0,59	<b>M'''</b>	37	66	51	47
<b>g=0.9</b>									
<b>P'</b>	0,62	0,49	0,54	0,54	<b>M'</b>	49	39	43	43
<b>P''</b>	0,44	0,41	0,2	0,35	<b>M''</b>	35	32	16	28
<b>P'''</b>	0,37	0,51	0,44	0,44	<b>M'''</b>	29	40	35	35

P'-вероятность правильного обнаружения границ сегментов

P''-вероятность ошибки первого рода (вероятность ложной тревоги)

P'''-вероятность ошибки второго рода (вероятность пропуска цели)

M-количество границ сегментов определенных на слух

M'-количество границ сегментов определенных правильно

M''-количество границ сегментов определенных не правильно

M'''-количество пропущенных границ сегментов

g-значение порога корреляции

N-длина фрейма

## Основные полученные результаты:

1. Проведен обзор методов сегментации речевых сигналов , который показал что несмотря на довольно большее количество методов сегментации проблема до конца не решена , так как ни один из предлагаемых методов и алгоритмов не позволяет с близкой к ста процентам вероятностью выделять сегменты речевых сигналов.
2. Рассмотрен метод кепстрального анализа , в результате чего была получена теоретическая основа для создания собственного алгоритма на основе этого анализа.
3. Разработан алгоритм сегментации речевых сигналов с использованием кепстрального анализа , было исследовано влияние на качество таких параметров как порог корреляции и длинна фрейма для данного алгоритма.
  - 3.1. С увеличением порога корреляции увеличивается количество правильно определяемых элементов , но при этом происходит выделение лишних сегментов для согласных шипящих таких как -с-, -ш- , -ф-. При этом хорошо разделяются гласные звуки, стоящие рядом -а-, -и- , -е- , разделяется сочетание -кс-.
  - 3.2. Увеличение длинны фрейма уменьшает чувствительность алгоритма к изменениям речевого сигнала. При снижении длинны происходит выделение лишних сегментов шипящих.