



<http://dx.doi.org/10.35596/1729-7648-2022-20-1-73-82>

Оригинальная статья / Original paper

УДК 004.934.2+534.784

СИСТЕМА АНАЛИЗА И КЛАССИФИКАЦИИ ГОЛОСОВОГО СИГНАЛА НА ОСНОВЕ ПЕРТУРБАЦИОННЫХ ПАРАМЕТРОВ И КЕПСТРАЛЬНОГО ПРЕДСТАВЛЕНИЯ В ПСИХОАКУСТИЧЕСКИХ ШКАЛАХ

М.И. ВАШКЕВИЧ, Д.С. ЛИХАЧЁВ, И.С. АЗАРОВ

*Белорусский государственный университет информатики и радиоэлектроники
(г. Минск, Республика Беларусь)*

Поступила в редакцию 15 декабря 2021

© Белорусский государственный университет информатики и радиоэлектроники, 2022

Аннотация. Описан подход к построению системы анализа и классификации голосового сигнала на основе пертурбационных параметров и кепстрального представления. Рассмотрены два варианта кепстрального представления голосового сигнала: при помощи мел-частотных кепстральных коэффициентов (МЧКК) и при помощи барк-частотных кепстральных коэффициентов (БЧКК). В работе использовался общепринятый подход к вычислению МЧКК на основе частотно-временного анализа методом дискретного преобразования Фурье (ДПФ) с объединением энергии в субполосах. Данный метод аппроксимирует частотное разрешение слуха человека, но имеет фиксированное временное разрешение. В качестве альтернативы предложен вариант кепстрального представления на основе БЧКК. При расчете БЧКК использовался неравнополосный ДПФ-модулированный банк фильтров, аппроксимирующий частотную и временную разрешающую способность слуха. Целью работы ставилось сравнение эффективности применения признаков на основе МЧКК и БЧКК для построения систем анализа и классификации голосового сигнала. Результаты эксперимента показали, что в случае использования акустических признаков на основе МЧКК можно получить систему классификации голоса со средней полнотой классификации 80,6 %, а в случае использования признаков на основе БЧКК этот показатель равен 83,7 %. При дополнении набора МЧКК признаков пертурбационными параметрами голоса средняя полнота классификации повысилась до 94,1 %, при аналогичном дополнении набора БЧКК признаков средняя полнота классификации увеличилась до 96,7 %.

Ключевые слова: голосовой сигнал, МЧКК, БЧКК, патология голоса.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Для цитирования. Вашкевич М.И., Лихачёв Д.С., Азаров И.С. Система анализа и классификации голосового сигнала на основе пертурбационных параметров и кепстрального представления в психоакустических шкалах. Доклады БГУИР. 2022; 20(1): 73-82.

VOICE ANALYSIS AND CLASSIFICATION SYSTEM BASED ON PERTURBATION PARAMETERS AND CEPSTRAL PRESENTATION IN PSYCHOACOUSTIC SCALES

MAXIM I. VASHKEVICH, DENIS S. LIKHACHOV, ELIAS S. AZAROV

Belarusian State University of Informatics and Radioelectronics (Minsk, Republic of Belarus)

Submitted 15 December 2021

© Belarusian State University of Informatics and Radioelectronics, 2022

Abstract. The paper describes an approach to design a system for analyzing and classification of a voice signal based on perturbation parameters and cepstral representation. Two variants of the cepstral representation of the voice signal are considered: based on mel-frequency cepstral coefficients (MFCC) and based on bark-frequency cepstral coefficients (BFCC). The work used a generally accepted approach to calculating the MFCC based on the time-frequency analysis by the method of discrete Fourier transform (DFT) with summation of energy in subbands. This method approximates the frequency resolution of human hearing, but has a fixed temporal resolution. As an alternative, a variant of the cepstral representation based on the BFCC has been proposed. When calculating the BFCC, a warped DFT-modulated filter bank was used, which approximates the frequency and temporal resolution of hearing. The aim of the work was to compare the effectiveness of the use of features based on the MFCC and BFCC for the designing systems for the analysis and classification of the voice signal. The results of the experiment showed that in the case when using acoustic features based on the MFCC, it is possible to obtain a voice classification system with an average recall of 80.6 %, and in the case when using features based on the BFCC, this metric is 83.7 %. With the addition of the set of MFCC features with perturbation parameters of the voice, the average recall of the classification increased to 94.1 %, with a similar addition to the set of BFCC features, the average recall of the classification increased up to 96.7 %.

Keywords: voice signal, MFCC, BFCC, vocal pathology.

Conflict of interests. The authors declare no conflict of interests.

For citation. Vashkevich M.I., Likhachov D.S., Azarov E.S. Voice Analysis and Classification System Based on Perturbation Parameters and Cepstral Presentation in Psychoacoustic Scales. Doklady BGUIR. 2022; 20(1): 73-82.

Введение

Патологические изменения в голосе могут возникать в результате неврологических заболеваний, инфекций, паралича мышечных тканей гортани и др. В настоящее время для выявления патологии голосовой функции используют субъективное и объективное оценивание. Под субъективной оценкой понимается суждение, выносимое врачом-специалистом в результате слухового восприятия и визуального исследования состояния голосовых складок. Объективное оценивание основывается на компьютерном анализе акустического голосового сигнала с целью выявления патологии в голосе, которая может даже быть не слышна для человека [1]. Данный тип оценивания состояния голоса в меньшей степени подвержен влиянию человеческого фактора. Кроме того, голос легко может быть записан при помощи смартфона, что позволяет использовать его в качестве платформы для построения системы анализа и классификации голосового сигнала [2].

Компьютерный анализ голосового сигнала подразумевает извлечение из сигнала акустических признаков, которые условно можно разделить на пертурбационные (т. е. описывающие возмущения сигнала на уровне отдельных периодов основного тона) и спектральные/кепстральные признаки [2–5]. Особенно широко используется кепстральное представление голосового сигнала [1, 3, 5]. Чаще всего применяются мел-частотные кепстральные коэффициенты (МЧКК), которые описывают структуру и динамику спектральных огибающих сигнала. В процессе вычисления МЧКК используется метод частотного-временного анализа сигнала на основе дискретного преобразования Фурье (ДПФ)

с объединением энергий в субполосах. Данный метод аппроксимирует частотное разрешение слуха человека, но имеет фиксированное временное разрешение. В настоящей работе предлагается альтернативный вариант кепстрального представления голосового сигнала при помощи барк-частотных кепстральных коэффициентов (БЧКК), при расчете которых используется неравнополосный ДПФ-модулированный банк фильтров, аппроксимирующий частотную и временную разрешающую способность слуха. Целью работы ставилось сравнение эффективности применения признаков на основе МЧКК и БЧКК для построения систем анализа и классификации голосового сигнала.

Пертурбационные параметры голоса

Пертурбационные параметры описывают возмущения временной формы голосового сигнала, который относится к классу квазипериодических сигналов. В группу пертурбационных параметров в первую очередь включают джиттер и шиммер. При помощи джиттера оценивают степень относительного изменения длительности следующих друг за другом периодов основного тона (ОТ). Существует несколько вариантов расчета параметра джиттер, которые оценивают либо краткосрочные, либо долгосрочные изменения периода основного тона. Имеется также параметр частотной пертурбации, учитывающий только знаки изменения следующих друг за другом периодов ОТ. Это так называемый направленный коэффициент возмущений (англ. *DPF – directional perturbation factor*). Параметр шиммер оценивает степень относительных изменений амплитуды следующих друг за другом периодов ОТ. Для него также существует несколько вариантов расчета.

Пертурбационные параметры также рассчитываются исходя из контура частоты основного тона (ЧОТ). К этой группе относят: 1) частотный диапазон фонации (англ. *PFR – phonatory frequency range*); 2) среднеквадратичное отклонение ЧОТ – SD_{f_0} ; 3) энтропию периодов ОТ (англ. *PPE – pitch period entropy*) [4]; 4) индекс патологичности вибрато (англ. *PVI – pathology vibrato index*). Используемые в работе пертурбационные параметры приведены в табл. 1. Более подробное их описание можно найти в работе [5].

Таблица 1. Пертурбационные параметры голоса
Table 1. Perturbation voice parameters

Группа параметров	Число параметров	Названия параметров
Частотная пертурбация	5	$J_{loc}, J_{ppq3}, J_{ppq5}, J_{ppq55}$
Амплитудная пертурбация	5	$S_{loc}, S_{apq3}, S_{apq5}, S_{apq11}, S_{apq55}$
Пертурбация контура ЧОТ	4	SD_{f_0}, PFR, PPE, PVI
Всего	14	–

Кепстральное представление голоса в психоакустических шкалах

В данном разделе рассматривается кепстральное представление голосового сигнала, получаемое на основе спектрального анализа сигнала в психоакустически мотивированной частотной шкале. Анализируется широко применяемое для описания голосового сигнала мел-частотное кепстральное представление [6], которое сравнивается с предлагаемым в работе барк-частотным кепстральным представлением, получаемым на основе неравнополосного ДПФ-модулированного банка фильтров.

Расчет МЧКК относится к методам кратковременного анализа голосового сигнала, которые предполагают разбиение сигнала на кадры анализа. Как правило, в интервале от 10 до 30 мс голосовой сигнал можно считать стационарным. Схема вычисления МЧКК показана на рис. 1.

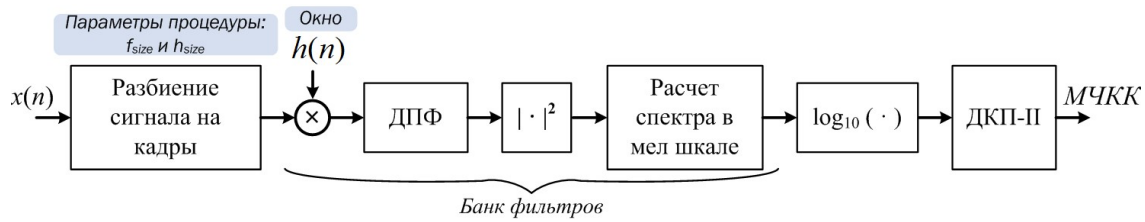


Рис. 1. Схема вычисления мел-частотных кепстральных коэффициентов (МЧКК)
Fig. 1. Scheme for calculating mel-frequency cepstral coefficients (MFCC)

В соответствии со схемой на рис. 1 при вычислении МЧКК входной сигнал $x(n)$ разбивается на кадры размера f_{size} с перекрытием h_{size} отсчетов. Каждый кадр умножается на оконную функцию Хэмминга $h(n)$, и к нему применяется ДПФ. Далее рассчитывается энергетический спектр сигнала, который затем переводится в мел-шкалу. Перевод в мел-шкалу осуществляется путем взвешивания энергетического спектра сигнала на треугольные окна, имеющие равную ширину в шкале мелов. От полученного мел-спектра вычисляется логарифм и применяется дискретное косинусное преобразование 2-го типа (ДКП-II).

При расчете МЧКК используются следующие значения настроечных параметров. Нижняя и верхняя границы частотного диапазона, покрываемого треугольными фильтрами, равны 50 Гц и 6400 Гц соответственно. Число полос в шкале мелов (равное числу треугольных фильтров, используемых для взвешивания энергетического спектра сигнала) $M = 20$. Число МЧКК, получающееся в результате расчета, равно $K = M - 1 = 19$ (выход ДКП-II с нулевым индексом не используется).

Заметим, что в процессе вычисления МЧКК моделируется работа банка слуховых фильтров. Имеется в виду то, что совместное действие умножения сигнала на оконную функцию с последующим вычислением ДПФ и взвешиванием энергий спектральных отсчетов можно интерпретировать, как работу неравнополосного банка фильтров, который выполняет декомпозицию энергии сигнала в соответствии с психоакустической шкалой мелов. Частотно-временное разрешение такого анализа зависит от размера анализируемого кадра f_{size} . В данной работе рассматривались несколько вариантов МЧКК, различающиеся между собой размером кадра анализа (f_{size} выбирался равным 10, 20 и 40 мс), при этом размер перекрытия всегда выбирался равным $h_{size} = f_{size} / 16$. На рис. 2 представлены амплитудно-частотные характеристики (АЧХ) фильтров, соответствующие различным вариантам расчета МЧКК. Вид АЧХ фильтров зависит от используемого для спектрального анализа временного окна (в данном случае это окно Хэмминга), а также от треугольных окон, которые используются для взвешивания энергий в субполосах.

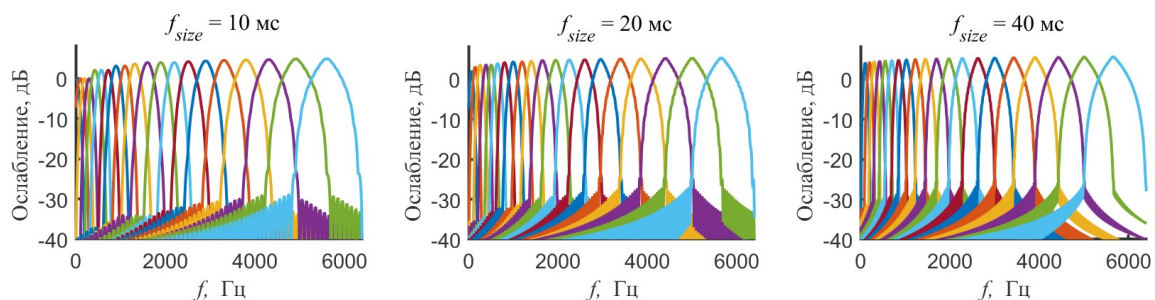


Рис. 2. АЧХ банков фильтров, использующихся при расчете МЧКК
Fig. 2. Magnitude responses of the filter banks used for MFCC calculation

В работе предлагается барк-частотное кепстральное представление сигнала, которое в отличие от мел-частотного использует неравнополосный ДПФ-модулированный банк фильтров [7], аппроксимирующий шкалу барков в качестве средства частотно-временного анализа сигнала. Неравнополосный ДПФ-модулированный банк фильтров на основе фазового преобразования имеет частотно-временное покрытие неравномерное как по частоте, так и по времени и, следовательно, может более точно моделировать слуховое восприятие человека. Схема вычисления БЧКК показана на рис. 3.

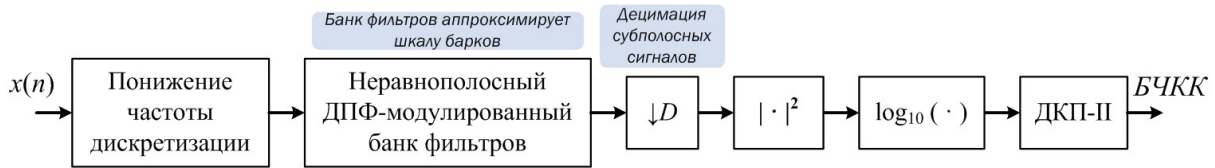


Рис. 3. Схема вычисления барк-частотных кепстральных коэффициентов (БЧКК)

Fig. 3. Scheme for calculating bark-frequency cepstral coefficients (BFCC)

В соответствии со схемой на рис. 3 вычисление БЧКК начинается с блока понижения частоты дискретизации. Он необходим, поскольку исходный сигнал, как правило, имеет высокую частоту дискретизации (более 40 кГц), а анализируемая банком фильтров полоса чаще всего не превышает 7–8 кГц. Неравнополосный ДПФ-модулированный банк фильтров выполняет декомпозицию сигнала на субполосы, имеющие равную ширину относительно шкалы барков. В банке фильтров используется низкочастотный (НЧ) фильтр-прототип, длину его импульсной характеристики можно соотнести с размером кадра анализа f_{size} в схеме вычисления МЧКК. После банка фильтров в схеме следует блок децимации субканальных сигналов. Коэффициент децимации D можно соотнести с величиной перекрытия кадров h_{size} в схеме вычисления МЧКК. Остальные блоки имеют тот же смысл, что и в схеме вычисления МЧКК. Рассматривались три варианта БЧКК, имеющие отличия в банке фильтров. Использовались три банка фильтров, у которых длина импульсной характеристики самого НЧ фильтра равнялась 10, 20 и 40 мс соответственно. Их АЧХ приведены на рис. 4.

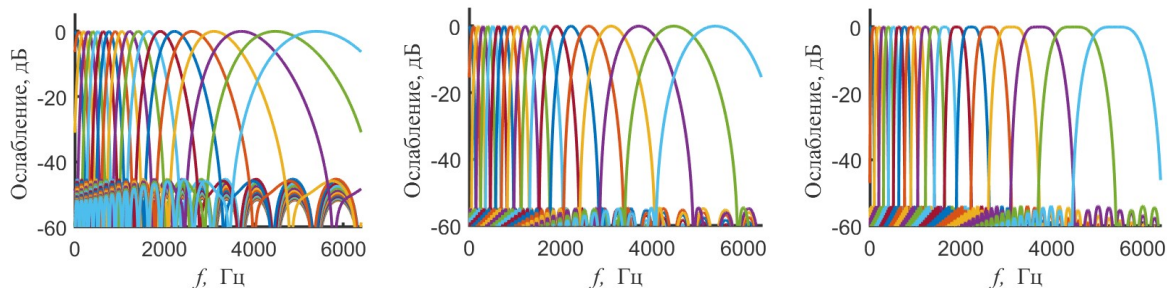


Рис. 4. АЧХ банков фильтров, использующихся при расчете БЧКК

Fig. 4. Magnitude responses of the filter banks used for BFCC calculation

Надсегментные кепстральные признаки

В результате кратковременного анализа формируется большой набор МЧКК/БЧКК, которые описывают локальную структуру спектра сигнала. Однако для построения эффективной системы анализа и классификации голосовых сигналов требуется уменьшить объем данных, получаемых в результате первичного расчета. Для этого получающиеся МЧКК/БЧКК предлагается усреднять для формирования надсегментного вектора кепстральных признаков.

На рис. 5 представлен процесс формирования надсегментных кепстральных признаков. Исходный сигнал разбивается на длинные сегменты (до 1 с), для которых затем производится покадровая обработка, в результате которой формировалась последовательность МЧКК/БЧКК и их конечных разностей (дельт). Затем для последовательностей МЧКК/БЧКК рассчитываются их средние значения и СКО. Данные параметры и являются надсегментными МЧКК/БЧКК признаками.

Длина сегмента L_{size} для получения надсегментных кепстральных признаков выбиралась равной 800 мс (перекрытие между соседними сегментами 75 %). В соответствии со схемой, показанной на рис. 5, размерность вектора надсегментных кепстральных признаков равна $1 \times 4 \cdot (M - 1)$, т. е., учитывая, что $M = 20$, она равна 1×76 . Общее число петрубационных параметров 14. Таким образом, общее число признаков, извлекаемых из сигнала, равно 90.

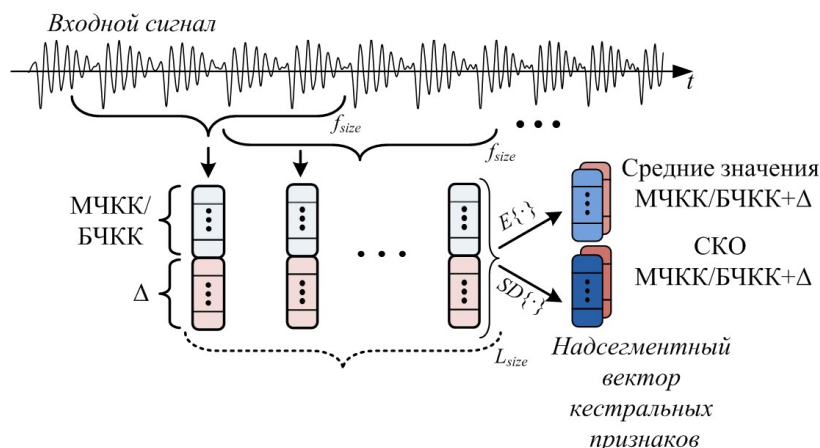


Рис. 5. Процесс формирования надсегментных признаков
Fig. 5. The process of forming suprasegment features

Система анализа и классификации голосового сигнала

В качестве примера рассматривается система, выполняющая задачу бинарной классификации, т. е. отнесения образца голосового сигнала к норме или патологии. Общая схема системы состоит из трех базовых блоков (см. рис. 6): предобработки, извлечения признаков и классификатора и принятия решения.

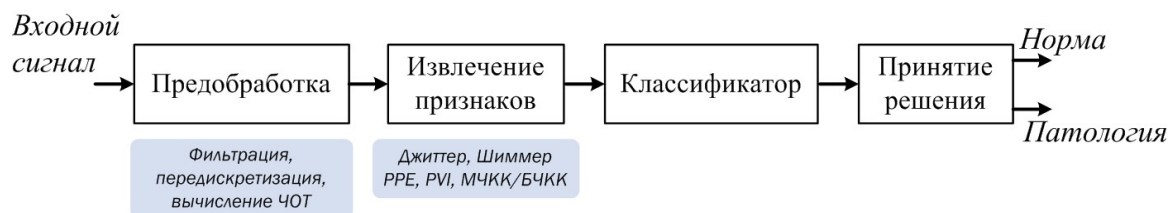


Рис. 6. Системы анализа и классификации голосового сигнала
Fig. 6. The system of analysis and classification of voice signal

В соответствии со схемой на рис. 6 на первом этапе, названном предобработкой, выполняется фильтрация сигнала, вычисление контура ЧОТ, а также, при необходимости, передискретизация сигнала. На втором этапе происходит извлечение признаков, рассчитываются пертурбационные параметры и надсегментные кепстральные признаки. Далее полученный вектор признаков попадает в блок классификации, который относит образец к одному из возможных классов. Для решения задачи классификации применялся метод линейного дискриминантного анализа (англ. LDA – *linear discriminant analysis*) [8].

Для проведения экспериментов использовалась база, содержащая 54 образца голосовых сигналов (протяжный звук /a/), описанная в работе [9]. База содержит записи 39 здоровых человек (23 мужчины, 16 женщин) и 15 больных боковым амиотрофическим склерозом (БАС) (6 мужчин, 9 женщин). В результате перевода голосовой базы в кепстральные параметры получена выборка из 877 обучающих векторов (650 – группа здоровых, 227 – группа больных).

При планировании эксперимента ставилась задача получить ответы на следующие вопросы:

1. Какое качество классификации можно получить, используя надсегментные МЧКК/БЧКК признаки?
2. Как изменится качество классификации, если к надсегментным МЧКК/БЧКК признакам добавить пертурбационные параметры сигнала?

Поскольку рассматриваемая база не является сбалансированной, то принято решение использовать среднюю полноту классификации как меру качества классификации.

Результаты и их обсуждение

Эксперимент проводился в три этапа: 1) подготовка обучающей выборки; 2) ранжирование признаков; 3) обучение и тестирование классификатора с использованием различного числа признаков.

На первом этапе исходная база звуковых файлов переводилась в надсегментные МЧКК/БЧКК признаки. Поскольку сигнал анализировался 800 мс сегментами, а средняя длительность записей превышала 4 с, то для каждой записи получалось несколько выходных значений надсегментных признаков. В случае вычисления пертурбационных параметров, каждой записи ставился в соответствие только один параметр. На втором этапе для полученной обучающей выборки выполнялось ранжирование признаков методом LASSO [8]. На третьем этапе выполнялось обучение и тестирование классификатора методом перекрестной проверки по $K = 4$ блокам [8]. Причем разбиение на блоки выполнялось на уровне дикторов, и, таким образом, тестовый и обучающий наборы содержали вектора, относящиеся к голосам разных дикторов. Процедура перекрестной проверки выполнялась $J = 76$ раз для того, чтобы проверить работу классификатора с различным числом признаков (от 1 до J). При этом на p -м запуске для классификатора выбирались признаки, имеющие ранг с 1 по p . По результатам перекрестных проверок строился график, который показывал зависимость качества классификации от числа признаков.

В исследовании использовались следующие наборы признаков (см. табл. 2), которые составлялись для получения ответов на представленные выше вопросы.

Таблица 2. Наборы признаков, использовавшиеся в экспериментах

Table 2. Feature sets used in experiments

Название набора признаков	Размерность	Описание
Petrurb	14	Пертурбационные параметры из табл. 1
MFCC-10	76	Надсегментные МЧКК признаки, кадр анализа 10 мс
MFCC-20	76	Надсегментные МЧКК признаки, кадр анализа 20 мс
MFCC-40	76	Надсегментные МЧКК признаки, кадр анализа 40 мс
BFCC-10	76	Надсегментные БЧКК признаки, длина НЧ фильтра 10 мс
BFCC-20	76	Надсегментные БЧКК признаки, длина НЧ фильтра 20 мс
BFCC-40	76	Надсегментные БЧКК признаки, длина НЧ фильтра 40 мс

В первом эксперименте проверялось качество классификации на основе надсегментных признаков MFCC-10/20/40 и BFCC-10/20/40. Для примера на рис. 7 представлены результаты оценки качества классификации при использовании признаков MFCC-40 и BFCC-40.

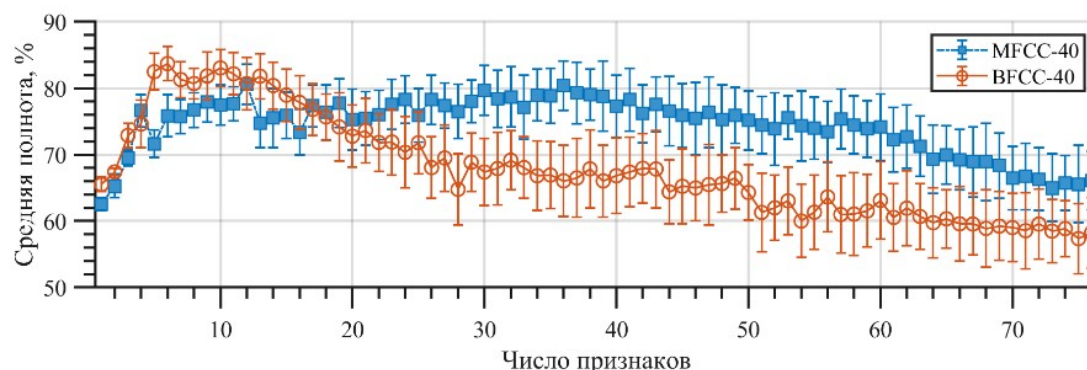


Рис. 7. Качество классификации при использовании признаков MFCC-40 и BFCC-40

Fig. 7. The quality of classification using features MFCC-40 and BFCC-40

Полученные в эксперименте результаты суммируются в табл. 3, в которой приведена оценка средней полноты классификации в зависимости от набора кепстральных признаков,

при ограничении на число признаков. Такое ограничение обусловлено тем, что для классификации всегда желательно иметь меньшее число признаков, но обеспечивающих лучшее качество. Лучший результат достигается при использовании набора признаков BFCC-40 (см. табл. 3).

Таблица 3. Максимальная средняя полнота классификации при ограничении числа признаков p
Table 3. The maximum average recall of the classification achieved when the number of features p is limited

Набор признаков	Средняя полнота, % (в скобках указано число признаков, при котором достигается средняя полнота)					
	$p \leq 10$	$p \leq 20$	$p \leq 30$	$p \leq 40$	$p \leq 50$	$p \leq 60$
MFCC-10	76,9 (5)	77,2 (15)	–	–	–	–
MFCC-20	61,9 (9)	62,4 (17)	67,6 (30)	70,1 (40)	74,0 (48)	–
MFCC-40	78,0 (9)	80,6 (12)	–	–	–	–
BFCC-10	72,2 (10)	73,5 (11)	–	–	–	–
BFCC-20	77,1 (10)	77,4 (13)	80,0 (28)	–	–	–
BFCC-40	83,7 (6)	–	–	–	–	–

Во втором эксперименте проверялось качество классификации на основе наборов надсегментных признаков, к которым добавлялись пертурбационные параметры голоса. Результаты эксперимента приведены в табл. 4. Можно заметить, что добавление к МЧКК / БЧКК признакам пертурбационных параметров позволяет существенно улучшить качество классификации.

Таблица 4. Максимальная средняя полнота классификации при ограничении числа признаков p
Table 4. The maximum average recall of the classification achieved when the number of features p is limited

Набор признаков	Средняя полнота, % (в скобках указано число признаков, при котором достигается средняя полнота)					
	$p \leq 10$	$p \leq 20$	$p \leq 30$	$p \leq 40$	$p \leq 50$	$p \leq 60$
MFCC-10 + Petrurb	84,2 (8)	89,3 (20)	93,5 (29)	94,1 (31)	–	–
MFCC-20 + Petrurb	84,7 (2)	90,3 (13)	–	–	90,9(49)	93,5 (54)
MFCC-40 + Petrurb	82,2 (5)	85,0 (14)	–	–	87,1 (46)	–
BFCC-10 + Petrurb	87,8 (8)	91,5 (18)	–	–	96,7 (45)	–
BFCC-10 + Petrurb	78,1 (9)	79,7 (18)	88,0 (22)	–	–	–
BFCC-10 + Petrurb	86,5 (10)	–	87,2 (25)	–	–	–

Среди всех наборов признаков, основанных на МЧКК, наилучший результат достигается при использовании 31 наиболее значимых признаков из набора MFCC-10+Petrurb. При этом средняя полнота классификации достигает 94,1 %. Среди всех наборов признаков, основанных на БЧКК, наилучший результат дает использование 45 наиболее значимых признаков из набора BFCC-10+Petrurb. Достижимая при этом средняя полнота классификации равна 96,7 %. Этот результат подтверждает обоснованность применения метода частотно-временного анализа сигнала на основе неравнополосного ДПФ-модулированного банка фильтров на основе фазового преобразования при расчете кепстральных коэффициентов.

Заключение

В работе предложен метод вычисления барк-частотных кепстральных коэффициентов, основанный на использовании неравнополосного ДПФ-модулированного банка фильтров, аппроксимирующего частотно-временное разрешение слуха человека. Произведено сравнение предложенных БЧКК с широко распространенными МЧКК в отношении эффективности построения на их основе системы анализа и классификации голосовых сигналов. Проведенные эксперименты по построению системы классификации голосов пациентов с боковым амиотрофическим склерозом показали эффективность применения надсегментных БЧКК

признаков. Среди классификаторов, использующих набор кепстральных признаков, лучший результат (средняя полнота 83,7 %) достигнут LDA-классификатором, использующим 6 надсегментных БЧКК признаков, отобранных методом LASSO. Среди классификаторов, использующих набор кепстральных признаков, объединенных с пертурбационными параметрами голоса, лучший результат (средняя полнота 96,7 %) достигнут LDA-классификатором, использующим 45 надсегментных БЧКК признаков, отобранных методом LASSO.

Список литературы / References

1. Harar P., Galaz Z., Alonso-Hernandez J.B., Mekyska J., Burget R., Smekal Z. Towards robust voice pathology detection. *Neural Computing and Applications*. 2020;32(20): 15747-15757.
2. Likhachov D., Vashkevich M., Azarov E., Malhina K., Rushkevich Y. A mobile application for detection of amyotrophic lateral sclerosis via voice analysis. *International Conference on Speech and Computer, 2021*. Springer, Cham; 2021:372-383.
3. Benba A., Jilbab A., Hammouch A. Discriminating between patients with Parkinson's and neurological diseases using cepstral analysis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2016;24(10):1100–1108.
4. Tsanas A., Little M.A., McSharry P.E., Spielman J., Ramig L.O. Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease. *IEEE Transactions on Biomedical Engineering*. 2012;59(5):1264-1271.
5. Vashkevich M., Rushkevich Y. Classification of ALS patients based on acoustic analysis of sustained vowel phonations. *Biomedical Signal Processing and Control*. 2021;65:1-14.
6. Huang X., Acero A., Hon H.-W. *Spoken language processing: A guide to theory, algorithm, and system development*. Prentice hall PTR; 2001: 980.
7. Bielawski K., Petrovsky A. Proposition of minimum bands multirate noise reduction system which exploits properties of the human auditory system and all-pass transformed filter bank. *IEEE Workshop Signal Processing*. 2001:65-70.
8. Gareth J., Daniela W., Trevor H., Robert T. *An introduction to statistical learning with applications in R*. NewYork: Springer; 2013.
9. Vashkevich M., Petrovsky A. Rushkevich Y. Bulbar ALS detection based on analysis of voice perturbation and vibrato. *IEEE International Conference on Signal Processing: Algorithms, Architectures, Arrangements, and Applications*. 2019: 267-272.

Вклад авторов

Вашкевич М.И. определил цель и задачи исследования, предложил идею барк-частотного кепстрального представления голосового сигнала, выполнил программную реализацию расчета БЧКК, принимал участие в подготовке текста статьи и интерпретации результатов экспериментов.

Лихачёв Д.С. выполнил программную реализацию расчета МЧКК, участвовал в подготовке программной базы для эксперимента.

Азаров И.С. предложил идею совместного использования кепстральных признаков и пертурбационных параметров, принимал участие в подготовке текста статьи и интерпретации результатов экспериментов.

Authors' contribution

Vashkevich M.I. determined the purpose and objectives of the study, proposed the idea of the bark-frequency cepstral representation of the voice signal, carried out the software implementation of the BFCC calculation, took part in the preparation of the text of the article and the interpretation of the experimental results.

Likhachov D.S. carried out the software implementation of the calculation of the MFCC, participated in the preparation of the software tools for the experiment.

Azarov E.S. proposed the idea of the joint use of cepstral features with perturbation parameters, took part in the preparation of the text of the article and interpretation of the experimental results.

Сведения об авторах

Вашкевич М.И., к.т.н., доцент кафедры электронных вычислительных средств Белорусского государственного университета информатики и радиоэлектроники.

Лихачёв Д.С., к.т.н., доцент кафедры электронных вычислительных средств Белорусского государственного университета информатики и радиоэлектроники.

Азаров И.С., д.т.н., заведующий кафедрой вычислительных средств Белорусского государственного университета информатики и радиоэлектроники.

Адрес для корреспонденции

220013, Республика Беларусь,
г. Минск, ул. П. Бровки, 6,
Белорусский государственный университет
информатики и радиоэлектроники;
тел. +375-17-293-84-78;
e-mail: vashkevich@bsuir.by
Вашкевич Максим Иосифович

Information about the authors

Vashkevich M.I., Cand. of Sci., Associate Professor at the Computer Engineering Department of the Belarusian State University of Informatics and Radioelectronics.

Likhachov D.S., Cand. of Sci., Associate Professor at the Computer Engineering Department of the Belarusian State University of Informatics and Radioelectronics.

Azarov E.S., Dr. of Sci., Head of the Computer Engineering Department of the Belarusian State University of Informatics and Radioelectronics.

Address for correspondence

220013, Republic of Belarus,
Minsk, P. Brovki st., 6,
Belarusian State University
of Informatics and Radioelectronics;
tel. +375-17-293-84-78;
e-mail: vashkevich@bsuir.by
Vashkevich Maksim Iosifovich