

# ОБНАРУЖЕНИЕ ПАТОЛОГИЙ ГОЛОСОВОГО АППАРАТА С ПОМОЩЬЮ МЕТОДОВ АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ ДИКТОРОВ

С. В. Демидчик

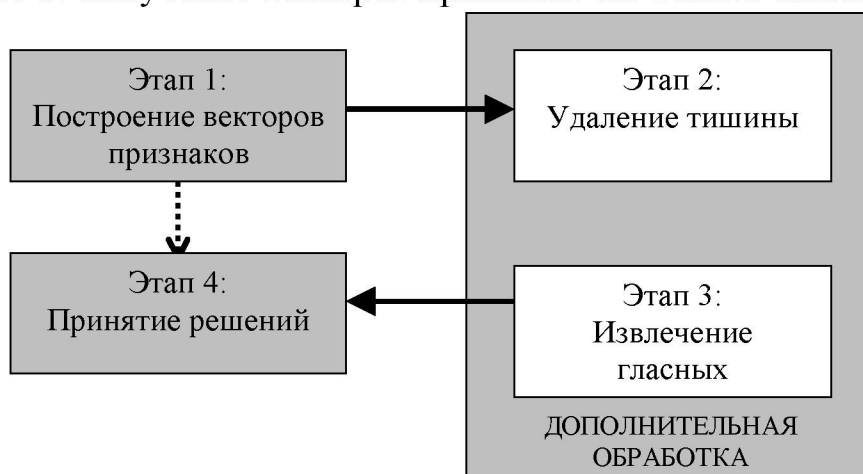
## 1. ВВЕДЕНИЕ

В настоящее время существует достаточно большое количество профессиональных методов диагностики и выявления различных заболеваний и патологий голосового аппарата (ларингоскопия, стробоскопия и др.). Однако данным методам присущ ряд определенных недостатков. В связи с этим в последнее время стали интенсивно развиваться неинвазивные методы диагностики речевых заболеваний, основным из которых является акустический анализ. Данный метод обладает рядом преимуществ, которые дают возможности для его широкого применения. Среди преимуществ данного средства диагностики заболеваний следует отметить достаточно высокую скорость проведения обследования, а также отсутствие какого-либо дискомфорта для пациента, поскольку в данном случае анализируется лишь записанный речевой сигнал, произнесенный человеком.

## 2. ОБЩАЯ СХЕМА СИСТЕМЫ ОБНАРУЖЕНИЯ БОЛЕЗНЕЙ

Как было показано в [1], из всего многообразия звуков наилучшим индикатором состояния голосового аппарата являются ударные гласные фонемы. Поэтому предложенная схема системы обнаружения патологий основана на анализе отдельно взятых гласных звуков. На *рис.1* приведена общая структура предложенной системы.

*Этап 1:* получение векторов-признаков на основе мел-кепстральных



*Рис.1.* Структурная схема системы классификации

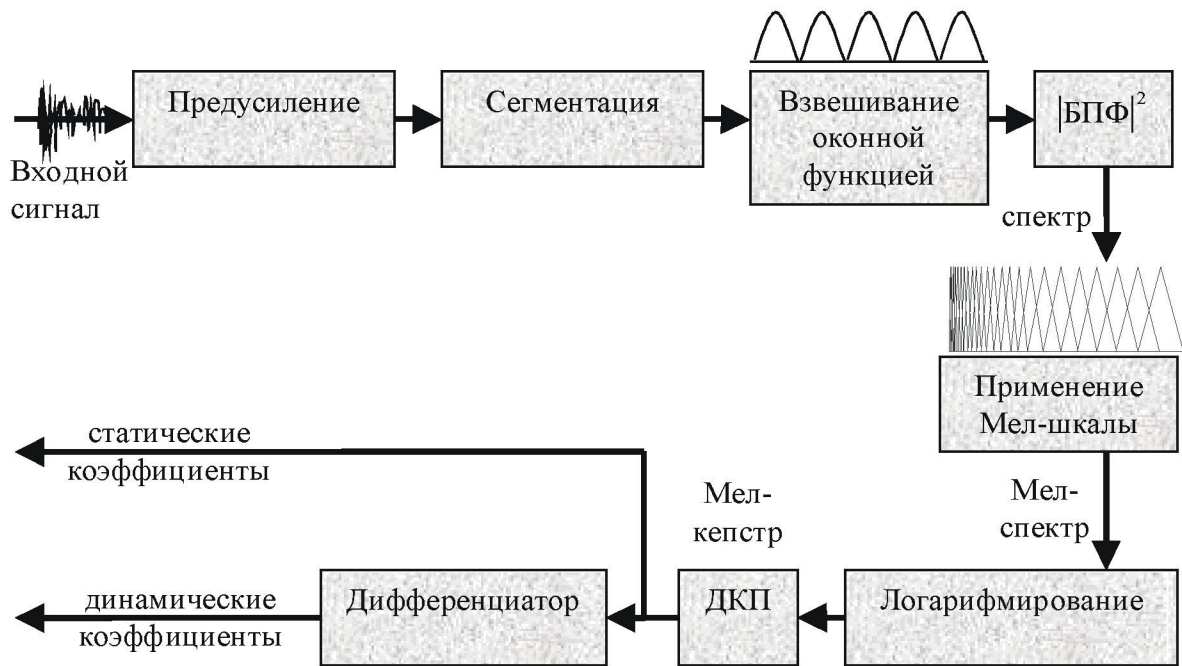


Рис. 2. Схема получения мел-кепстральных коэффициентов [2]. Существует несколько методов получения мел-кепстральных коэффициентов. На рис. 2 приведены основные шаги, необходимые для их получения.

На основе данного алгоритма были получены векторы признаков, состоящие из 12 (MFCC12) и 24 (MFCC24) мел-кепстральных коэффициентов, а также векторы признаков, включающие значение энергии сигнала (MFCC+E).

*Этап 2:* на данном шаге из сигнала удалялись фреймы с тишиной. Определение фреймов с тишиной производилось при помощи сравнения порога ( $T_{TFE}$ ) и общей энергии фрагмента.  $T_{TFE}$  вычисляется адаптивно, чтобы учесть эффект естественного изменения громкости. В начале для всего аудио-фрагмента вычисляются минимальное ( $E_{min}$ ), максимальное ( $E_{max}$ ), и среднее ( $E_{\mu}$ ) значения энергии. Затем проверяются 2 условия:

1.  $E_{max} >$  минимального уровня энергии текущего фрейма;
2.  $E_{max} \gg E_{min}$ .

Если условия не выполняются, то считается, что фрагмент содержит только тишину и дальнейшая обработка прекращается. Иначе вычисляется  $T_{TFE}$ :

$$T_{TFE} = E_{min} + \lambda_s \psi(E_m - E_{min}) \quad (1)$$

где  $0 < \lambda_s \leq 1$ .  $\lambda_s$  – коэффициент тишины, который определяет пороговое значение между  $E_{min}$  и  $E_{\mu}$ , соответствующее тишине. Если общая энергия

фрагмента меньше  $T_{TFE}$ , то он классифицируется как тишина, иначе – как не-тишина.

*Этап 3:* поскольку гласные звуки являются хорошим индикатором состояния голосовых связок, на данном этапе производилось извлечение только гласных звуков из аудиосигнала. Их извлечению производилось путем оценки наибольшего значения энергии частотного диапазона (dominant band energy ratio – DBER) и оценки основной частоты [3]. Гласные звуки представляют собой последовательность гармонических тонов и их спектральная энергия достаточно сильно локализована. В согласных же звуках, в основном, преобладают шумовые компоненты, вследствие чего их спектральная энергия значительно рассеивается по частотному диапазону.

*Этап 4:* используя обработанные на предыдущих этапах данные, классификатор на основе статистик второго порядка [4] распределяет входные данные по трем классам: класс здоровых голосов, класс голосов с патологиями и неопределенный класс. В неопределенный класс попадали те записи, которые классификатор не смог отнести к первым двум классам.

Классификатор предварительно проходил процедуру обучения. Все записи произвольным образом были разбиты на два набора: обучающий набор и тестовый набор. Обучающий набор (1/3 всех записей) использовался для создания базы данных моделей здоровых голосов и голосов с патологиями. После обучения на классификатор подавались записи из тестового набора. С помощью различных мер расстояний классификатор принимал решение о принадлежности входной записи к одному из вышеупомянутых классов. Для каждого вида меры расстояния и типа вектора признаков эксперимент повторялся 20 раз. Файлы из базы случайным образом разбивались на эталонное и тестовое множества в пропорции 1:2.

### **3. ПОЛУЧЕННЫЕ РЕЗУЛЬТАТЫ**

Для проведения эксперимента совместно с фониатрами Республиканского центра патологий слуха голоса и речи была создана база записей речевых сигналов, где каждой записи сопоставлен диагноз врача-фониатра. В настоящее время база насчитывает свыше 300 записей общей длительностью более 20 часов.

В таблице приведены значения полученной точности классификации для различных мер подобия  $d_i$  [4], которые применялись в качестве классификатора при исключении блоков предварительной обработки сигнала (Этапы 2 и 3, *рис. 1*).

Таблица 1

**Зависимость точности  
классификации от типа  
векторов признаков**

Мера по- добия	Величина ошибки, %		
	MFCC12	MFCC+E	MFCC24
<b>d1</b>	23	32	24
<b>d2</b>	16	28	17
<b>d3</b>	24	41	22
<b>d4</b>	41	53	39
<b>d11</b>	22	31	21
<b>d12</b>	46	55	45
<b>d14</b>	34	43	35
<b>d15</b>	11	23	12
<b>d16</b>	41	54	42
<b>d17</b>	25	36	26
<b>d19</b>	21	32	22
<b>d23</b>	19	29	18

#### 4. ЗАКЛЮЧЕНИЕ

В результате проведенного эксперимента была продемонстрирована возможность создания системы автоматического обнаружения патологий голосового аппарата. Были определены оптимальные параметры работы классификатора на основе статистик второго порядка. Получено значение точности классификации порядка 89% для предварительно необработанных фрагментов речи. При включении блока дополнительной обработки (Этапы 2 и 3, *рис. 1*) была достигнута точность 99% для некоторых случаев тяжелых патологий.

#### Литература

1. Демидчик С. В. Обработка биомедицинских сигналов для задач диагностики патологий речевого аппарата // Сб. работ 62 научн. конференции студентов и магистрантов БГУ, Ч.1, Минск, 2005, с. 122–126
2. Godino-Llorente J. I., Gomez-Vilda P. Automatic Detection of Voice Impairments by Means of Short-Term Cepstral Parameters and Neural Network Based Detectors // IEEE Trans. Biomedical Engineering, 51(2), 2004, 380–384.
3. Kiranyaz S., Aubazac M., Gabbouj M. Unsupervised Segmentation and Classification over MP3 and AAC Audio Bit-streams // WIAMIS Workshop, 338-345, London, 2003.
4. Bimbot F., Magrin-Chagnolleau I., Mathan L. Second Order Statistical Measures for Text Independent Speaker Identification // Speech Communication, 17:177–192, 1995.

## ВЛИЯНИЕ НИЗКОИНТЕНСИВНОГО ЛАЗЕРНОГО ИЗЛУЧЕНИЯ НА ТЕМПЕРАТУРУ ОБЛУЧАЕМОЙ ТКАНИ

**Т. А. Железнякова**

### ВВЕДЕНИЕ

Повышение эффективности лечения заболеваний кожи является важной задачей дерматологии. Одним из путей ее решения является разработка новых или совершенствование известных способов взамен традиционным (инъекции с помощью шприца, таблетки и микстуры, мази, компрессы и аппликации, лекарственный электро- и фонофорез).

Одним из перспективных методов является лекарственный лазерофорез (ЛЛФ), суть которого заключается в сочетанном воздействии на ор-