



УДК 004.934

РАЗРАБОТКА СИСТЕМЫ РАСПОЗНАВАНИЯ РЕЧИ НА ОСНОВЕ СКРЫТЫХ МАРКОВСКИХ МОДЕЛЕЙ ОТДЕЛЬНЫХ СЛОВ

А. Н. Савин, Н. Е. Тимофеева, А. С. Гераськин, Ю. А. Мавлютова

Савин Александр Николаевич, кандидат физико-математических наук, доцент кафедры дискретной математики и информационных технологий, Саратовский национальный исследовательский государственный университет имени Н. Г. Чернышевского, 410012, Россия, Саратов, Астраханская, 83, savinan@info.sgu.ru

Тимофеева Надежда Евгеньевна, заведующий лабораторией теоретических проблем информатики и ее приложений кафедры дискретной математики и информационных технологий, Саратовский национальный исследовательский государственный университет имени Н. Г. Чернышевского, 410012, Россия, Саратов, Астраханская, 83, timofeevane@yandex.ru

Гераськин Алексей Сергеевич, кандидат педагогических наук, доцент кафедры теоретических основ компьютерной безопасности и криптографии, Саратовский национальный исследовательский государственный университет имени Н. Г. Чернышевского, 410012, Россия, Саратов, Астраханская, 83, gerascinas@mail.ru

Мавлютова Юлия Альбертовна, старший лаборант лаборатории теоретических проблем информатики и ее приложений кафедры дискретной математики и информационных технологий, Саратовский национальный исследовательский государственный университет имени Н. Г. Чернышевского, 410012, Россия, Саратов, Астраханская, 83, yuliyamav@yandex.ru

Приведены результаты разработки программных модулей, реализующих систему распознавания речи на основе скрытых Марковских моделей отдельных слов и использования линейного предсказания при кодировании признаков звукового сигнала. Обосновывается структура системы распознавания речи, использующая скрытые марковские модели отдельных слов, состоящая из четырех модулей: модуль выделения слов из звукового потока, модуль анализа признаков слова, модуль обучения скрытых марковских моделей и модуль распознавания слов. Приводятся алгоритмы формирования скрытых марковских моделей с лево-правой топологией для отдельных слов требуемого словаря команд системы управления объектом, основанные на кодировании признаков звукового сигнала, использующего линейные предсказания. Приведены результаты оценки достоверности последовательности наблюдений, соответствующих отдельным словам, получаемым с помощью предложенного алгоритма обработки. Разработанные программные модули позволяют эффективно подготавливать необходимые исходные данные и формировать таким образом требуемый словарь команд системы управления объектом, строить скрытые марковские модели отдельных слов, проводить их обучение с помощью алгоритма Баума – Велша. Построенные словари команд предполагается использовать в интеллектуальных системах управления различными объектами.

Ключевые слова: скрытые марковские модели, кепстральный анализ, распознавание речи, метод Баума – Велша.

DOI: 10.18500/1816-9791-2017-17-4-452-464

ВВЕДЕНИЕ

Распознавание голоса в компьютерных системах весьма распространено. Распознавание речи и, как следствие, голосовая идентификация нашли свое применение



во всех сферах человеческой деятельности. Благодаря системам распознавания речи обеспечивается безопасность от несанкционированного проникновения в защищенную зону. Такие системы содержат базу данных голосов сотрудников, имеющих доступ к защищаемой зоне, и предотвращают допуск людей, чьих голосов в ней нет [1, 2].

В настоящее время широко разрабатываются и внедряются интеллектуальные системы управления различными объектами, которые позволяют осуществлять контроль за объектами в реальном времени. Управление такими системами можно осуществлять различными способами, одним из них является метод голосовых команд. При этом защиту объекта от несанкционированного доступа можно решить, используя индивидуальные особенности голоса каждого человека.

Уровень развития современной микропроцессорной техники (например, мобильные устройства связи) позволяет использовать сложные вычислительные алгоритмы, основанные на цифровой потоковой обработке статистических данных в реальном времени. Поэтому разработка таких алгоритмов является весьма актуальной.

Одним из путей решения вышеуказанных задач является использование для распознавания фрагментов речи математического аппарата скрытых марковских моделей (СММ) [3]. Данная работа посвящена разработке алгоритма и соответствующего программного модуля, осуществляющего формирование СММ для отдельных слов требуемого словаря команд системы управления объекта, на основе кодирования признаков звукового сигнала, использующего линейные предсказания.

1. СТРУКТУРА СИСТЕМЫ РАСПОЗНАВАНИЯ РЕЧИ НА ОСНОВЕ СММ

Рассмотрим дискретную систему, имеющую конечное множество из N состояний — $S = \{s_1, \dots, s_N\}$, в каждом из которых она может принимать одно из M значений из набора наблюдаемых параметров $V = \{v_1, \dots, v_M\}$ — алфавита. Состояние системы q_t в момент времени t , принимающее одно из N значений множества S , зависит только от её состояния q_{t-1} в момент времени $t-1$, а значение наблюдаемого параметра o_t в момент времени t зависит только от состояния q_t , т. е. не зависит от времени.

Вероятности переходов между состояниями системы задаются матрицей A . Вероятности выпадения каждого из M значений наблюдаемого параметра системы в каждом из N состояний системы задаются набором векторов B . Вероятность появления некоторого начального состояния системы задаётся вектором π . При этом последовательность состояний, в которых пребывает система $Q = q_1, \dots, q_T$, внешнему наблюдателю не видна, а видит он только последовательность наблюдений $O = o_1, \dots, o_T$ (здесь T — длина последовательности), т. е. система ведёт себя как «чёрный ящик». Модель такой системы получила название СММ и в компактной записи обозначается $\lambda = (A, B, \pi)$ [2].

Для моделирования отдельного слова может быть выбрана лево-правая СММ (рис. 1) на основе предположения о том, что в каждый момент времени система переходит в новое состояние [4]. Соответственно неизвестное число скрытых состояний N в этом случае определяется



Рис. 1. Структура лево-правой СММ
Fig. 1. The structure of the left-right hidden Markov models (HMM)



длиной и количеством сегментов, на которые слово разбивается при анализе его признаков. Процесс распознавания с использованием СММ предполагает два этапа (рис. 2).

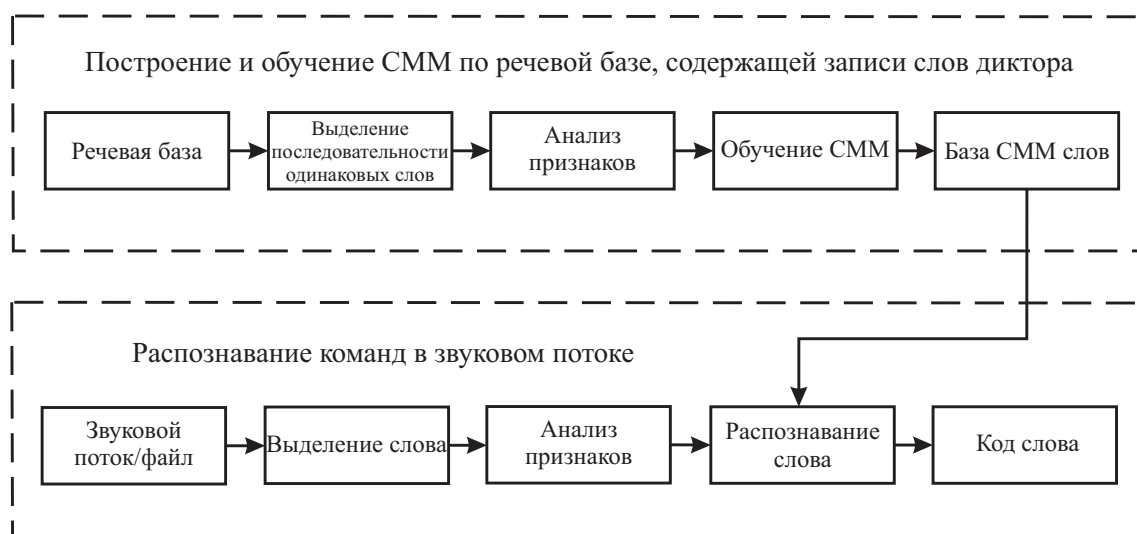


Рис. 2. Структура системы распознавания речи на основе использования СММ

Fig. 2. Structure of the speech recognition system based on the use of HMM

В режиме обучения элементы системы имеют следующее функциональное назначение:

- речевая база содержит записи слов, повторяющихся несколько раз для обеспечения адекватности получаемых СММ, которые будут доступны для распознавания;
- выделение последовательности одинаковых слов из файла речевой базы с помощью предварительной обработки (подавление шума, фильтрация и т. д.);
- анализ признаков и определение алфавита слова V , по которому формируется последовательность наблюдений O ;
- обучение СММ — подбор параметров СММ, чтобы она как можно лучше описывала реальную наблюдаемую последовательность O символов алфавита V анализируемого слова;
- сохранение СММ в базе — словаре.

В режиме распознавания:

- выделение слова из входного звукового потока с помощью предварительной обработки;
- анализ признаков распознаваемого слова и формирование соответствующей последовательности наблюдений O ;
- распознавание слова с использованием базы СММ и генерация кода распознаваемого слова.

Таким образом, для реализации данной структуры необходимо всего 4 модуля: модуль выделения слов из звукового потока, модуль анализа признаков слова, модуль обучения СММ с базой моделей, модуль распознавания слов.

2. АЛГОРИТМ ВЫДЕЛЕНИЯ ОТДЕЛЬНЫХ СЛОВ ИЗ ЕДИНОГО ЗВУКОВОГО ФАЙЛА

В режиме обучения файл должен содержать несколько раз произнесенное одним диктором требуемое слово. Это необходимо для получения достоверной последовательности наблюдений O , соответствующей данному слову. На рис. 3 приведена блок-схема алгоритма предварительной обработки звукового файла, основанного на

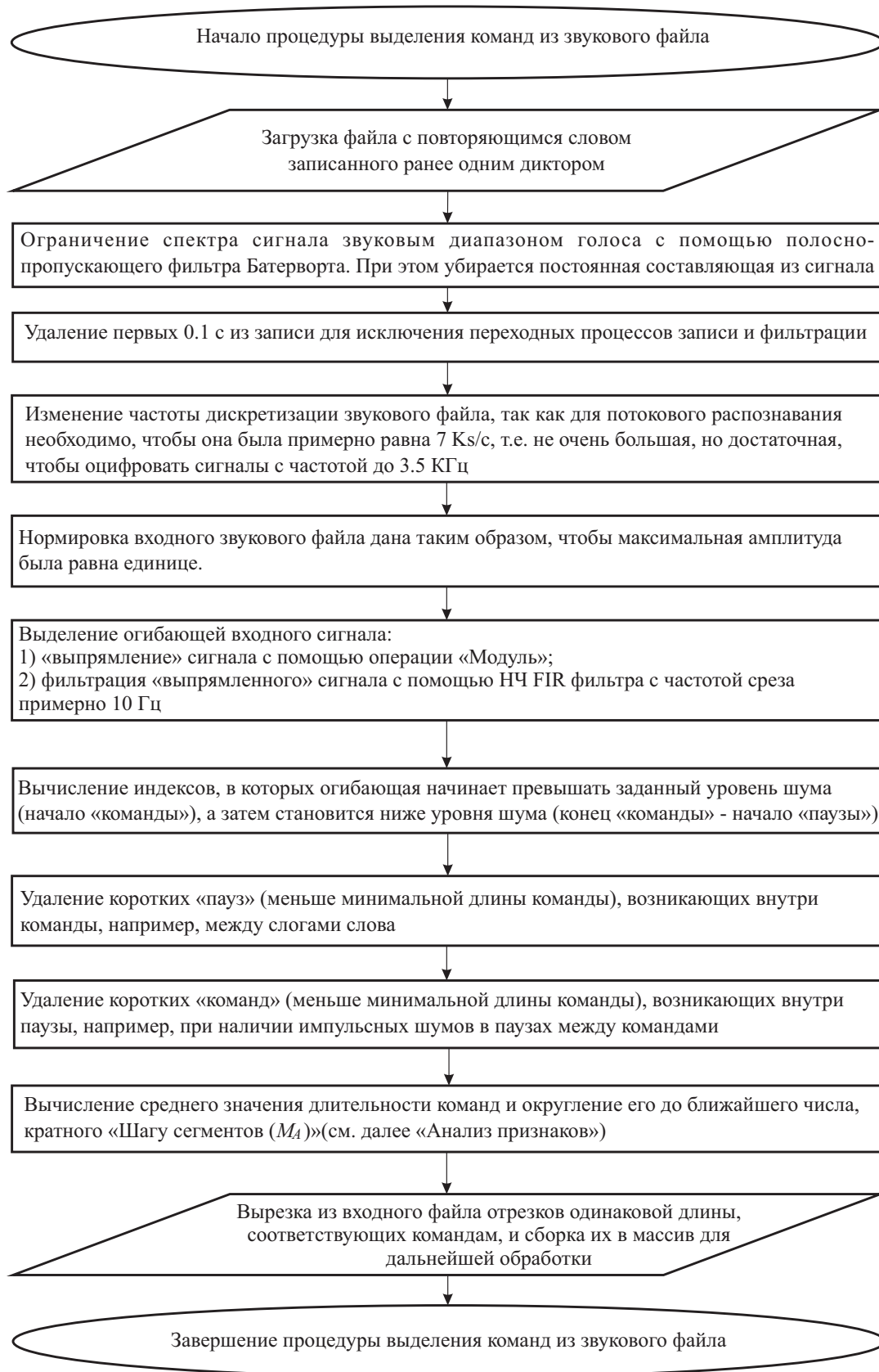


Рис. 3. Блок-схема алгоритма выделения отдельных повторяющихся слов из звукового файла

Fig. 3. A flowchart of an algorithm for selecting separate repetitive words from a sound file



вычислении огибающей и выделении на ее основе участков файла соответствующих повторяющимся словам.

При этом на выходе формируется массив отрезков звукового файла одинаковой длины, соответствующих повторяемому слову, что позволяет использовать усредненные входные данные при обучении СММ слова, делая её тем самым более адекватной.

Алгоритм выделения отдельных слов из звукового потока встроен в модуль построения СММ слов, реализованный в среде графического программирования LabVIEW компании National Instruments [5]. На рис. 4 показан процесс выделения команды из звукового файла, содержащего десять раз повторяющееся слово «Вперёд».

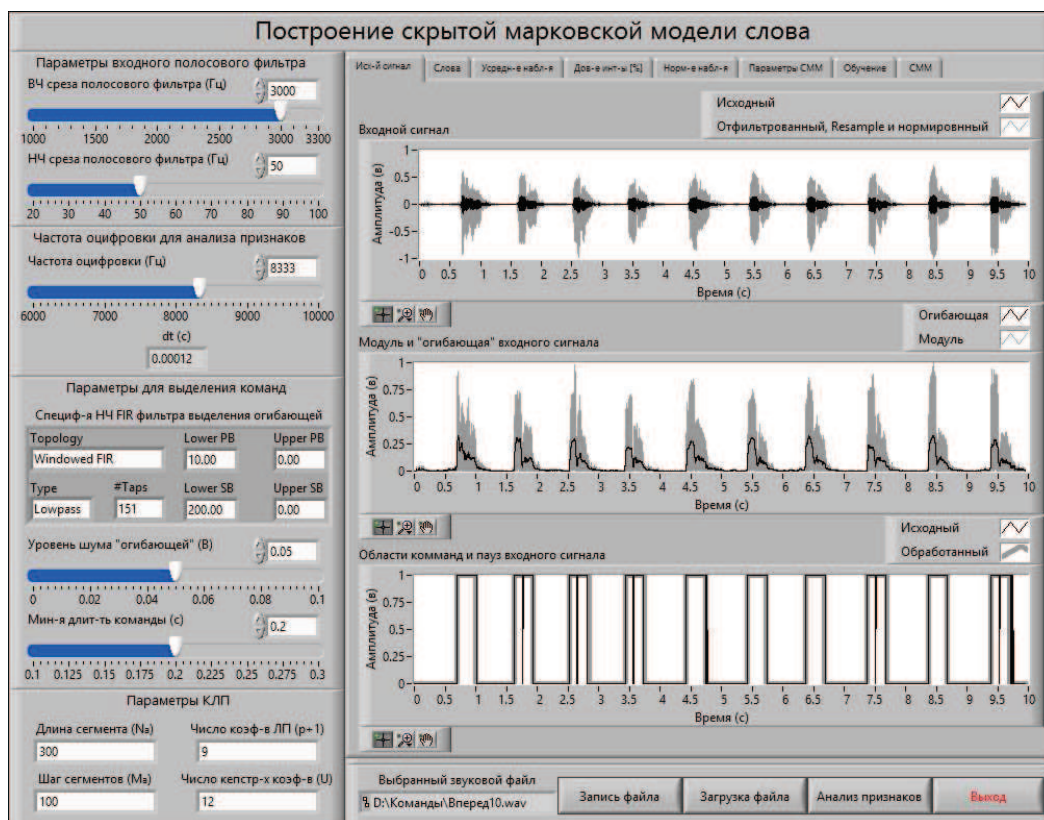


Рис. 4. Выделение повторяющихся слов из звукового файла
Fig. 4. Selecting duplicate words from a sound file

На первом графике рис. 4 чёрным изображен исходный сигнал, серым — отфильтрованный, с измененной частотой дискретизации и нормированный. Из записи убираются первые 0.1 с, соответствующие переходным процессам при включении микрофона и предварительной фильтрации.

Выделение слова осуществляется путем анализа огибающей сигнала. Индексы, в которых огибающая начинает превышать заданный в начале уровень шума, соответствуют началу команды. Индексы, в которых огибающая становится ниже уровня шума, — концу команды. Паузы внутри команды отсеиваются с помощью заданной заранее минимальной длительности команд. Таким же образом отсеиваются и лишние шумы.

Элементы управления модуля (граничные частоты среза входного полосового фильтра, частота дискретизации сигнала для анализа признаков, параметры НЧ



фильтра огибающей, уровень шума огибающей, минимальная длительность команды) позволяют подбирать требуемые параметры на этапе выделения команд для обеспечения построения адекватных СММ слов.

3. АЛГОРИТМ АНАЛИЗА ПРИЗНАКОВ СЛОВ

Для системы распознавания речи каждому слову необходимо сопоставить набор признаков. Этот процесс в [4] предложено осуществлять на основе анализа периодичности спектра фрагментов звукового сигнала (кепстральный анализ), предварительно обработанного с помощью алгоритмов линейного предсказания. Такой процесс называется кодированием на основе линейного предсказания (КЛП). Алгоритм анализа признаков на основе КЛП, используемый при распознавании отдельных слов, приведен на рис. 5, 6.

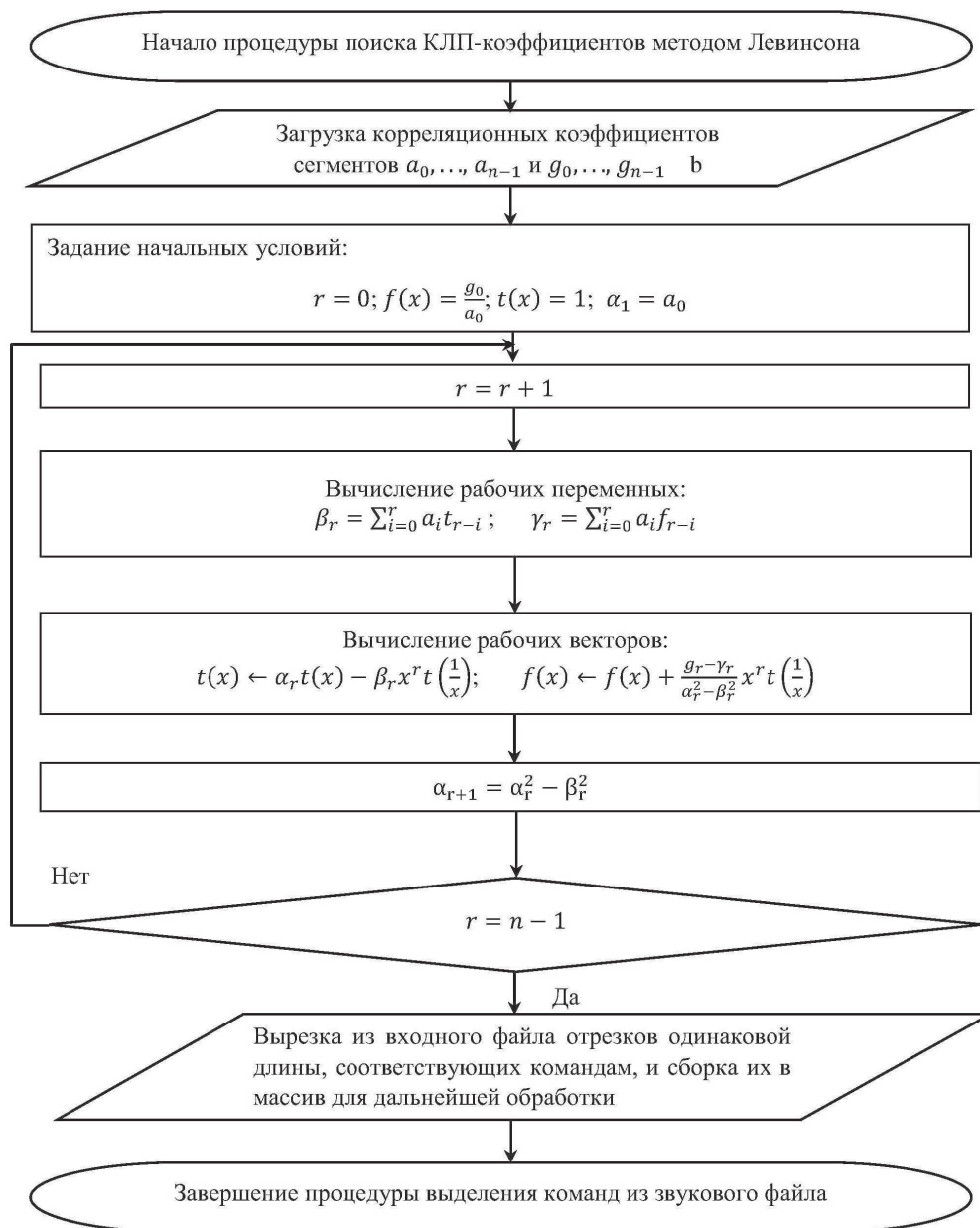
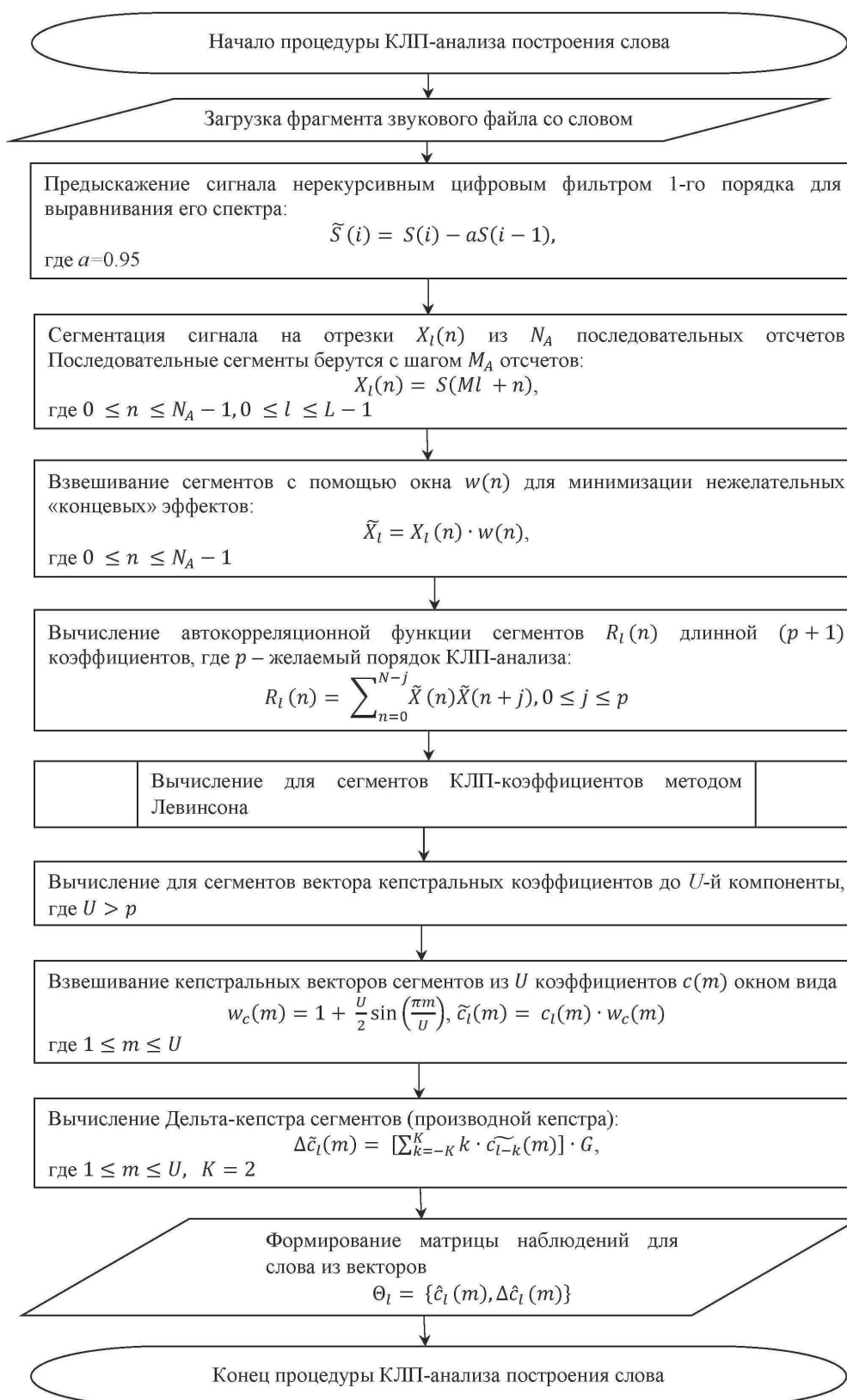


Рис. 5. Блок-схема процедуры алгоритма Левинсона вычисления КЛП-коэффициентов
 Fig. 5. Block diagram of the procedure of Levinson's algorithm for calculating linear prediction coding (LPC) coefficients



Достоинством спектральной обработки звуковых сигналов является то, что при переходе из временной области в частотную представление информации становится



более наглядным, компактным. Причем, чем более «простым» является сигнал во временной области, тем в большей степени происходит сжатие информации.

Выявление периодичности в спектре (кепстральный анализ) позволяет более достоверно и точно охарактеризовать особенности произношения дикторов. При этом спектральная информация представляется еще более компактно. Каждый гармонический ряд исходного спектра представляется в идеале всего одной составляющей в кепстре [4].

Использование линейного предсказания, основанного на автокорреляционной фильтрации, должно улучшать отношение сигнал – шум исходного сигнала и убирать из него случайные артефакты. Вычисление коэффициентов линейного предсказания осуществляется с помощью алгоритма Левинсона (см. рис. 5) [6].

В процессе анализа признаков слова каждый участок, выделенный ранее из файла и соответствующий повторяющемуся слову, разбивается на небольшие перекрывающиеся отрезки – сегменты и затем обрабатывается согласно алгоритму, приведенному на рис. 5, 6. Как видно (рис. 7, а), в результате предсказания сигнала происходит выравнивание спектра, что обеспечивает равноценность спектральных компонент при анализе признаков.

Оконное взвешивание (рис. 7, б) уменьшает сигнал на концах сегментов и увеличивает в центре, минимизируя нежелательные концевые эффекты.

Линейное предсказание (рис. 7, в) на основе алгоритма Левинсона (см. рис. 5) убирает сглаживает выбросы и случайные артефакты в анализируемом сигнале.

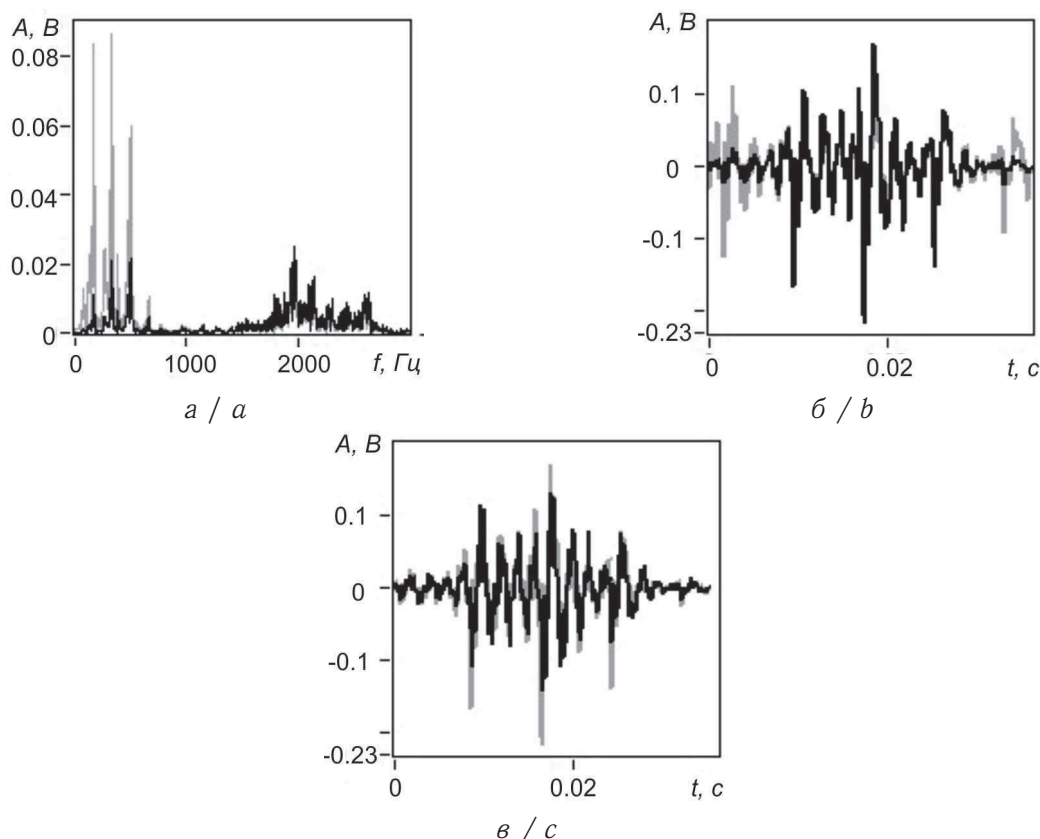


Рис. 7. Спектр исходного (---) и предсказанного (—) сигналов (а), исходный (---) и взвешенный (—) сегменты сигнала (б), исходный взвешенный (---) и предсказанный (—) сегменты сигнала (в)

Fig. 7. The spectrum of the initial (---) and pre-faded (—) signals (a), the initial (---) and weighted (—) signal segments (b), the initial weighted (---) and predicted (—) signal segments (c)



На выходе алгоритма формируется необходимая для распознавания слова матрица, строки которой образуются конкатенацией взвешенного кепстрального и соответствующего взвешенного дельта-кепстрального векторов сегментов. Каждая такая строка является набором признаков сегмента — вектором наблюдений и соответствует одному символу из алфавита V СММ слова в последовательности наблюдений O . Количество строк определяет число состояний N , в которых находилась лево-правая СММ слова.

Настройка параметров КЛП (число отсчётов в сегменте N_A , число отсчётов в смещении сегментов M_A , порядок КЛП-анализа p , число кепстральных коэффициентов U) осуществляется посредством соответствующих элементов (см. рис. 5).

4. ПОСТРОЕНИЕ СММ СЛОВА

В процессе построения СММ слова для повышения её адекватности необходимо использовать матрицу векторов наблюдений, полученную статистическим усреднением матриц векторов наблюдений повторяющихся слов. При этом усредняются вектора наблюдений, соответствующие одним и тем же моментам времени повторяющихся слов.

Оценка достоверности выборочных средних значений признаков сегментов (элементов векторов наблюдений) повторяющихся слов проводится с помощью доверительных интервалов, вычисляемых при уровне статистической значимости $\alpha = 0,05$.

Сравнение степени разброса, т.е. оценка однородности выборочных дисперсий значений элементов векторов наблюдений, вычисленных по результатам анализа признаков повторяющихся слов, осуществляется с использованием критерия Кохрена [7]

Обеспечение статистически значимых выборочных средних значений элементов векторов наблюдений и однородности их дисперсий, т.е. получение достоверной последовательности наблюдений для анализируемого слова, достигается подбором параметров обработки входного сигнала и параметров КЛП (см. п. 2, 3).

Оценку расстояния между символами алфавита V слова — усреднёнными векторами наблюдений — было предложено делать с помощью евклидовой нормы. При этом для обеспечения равнозначности признаков при вычислении расстояния проводилась их нормировка. В качестве нормирующего для каждого элемента вектора наблюдений использовался диапазон его изменения в матрице, расширенный с учётом доверительного интервала.

Формирование алфавита СММ слова осуществляется удалением повторяющихся строк из нормированной матрицы средних значений наблюдений, если такие имеются. При этом сравнивается расстояние между текущей строкой и остальными. Если оно меньше некоторой заданной величины, то строка с большим индексом удаляется, так как считаем, что эти строки соответствуют одному и тому же символу. Соответственно число строк получившейся прореженной матрицы определяет количество символов M алфавита V , индексы строк являются значениями алфавита, а сами строки — признаками символов.

Последовательность наблюдений слова определяется сравнением строк матрицы алфавита V (прореженной) с исходной нормированной матрицей средних значений наблюдений. Если расстояние между строками меньше некоторой заданной величины, использованной при построении алфавита, то индекс строки матрицы алфавита V записывается в последовательность наблюдений O . Длина последовательности наблюдений T равна числу строк исходной матрицы, а число состояний N равно T в случае лево-правой СММ слова.



На рис. 8 приведены результаты экспериментов по определению числа состояний N для слов «Вперёд» и «Стоп». На вход модуля построения СММ для каждого из этих слов подавалось тридцать звуковых файлов, содержащих по пятнадцать повторений, произнесённых одним диктором. Во всех случаях число состояний N совпадало с количеством символов M алфавита V , что соответствует предположению о лево-правой структуре СММ (см. рис. 1) для этих слов из-за отсутствия в них повторяющихся звуков.

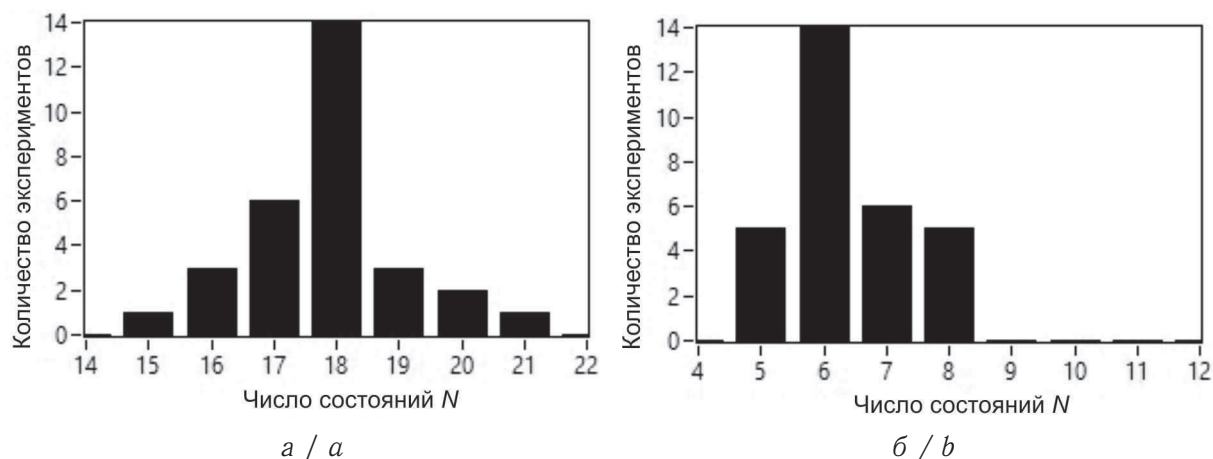


Рис. 8. Количество полученных в экспериментах состояний для слов «Вперёд» (а) и «Стоп» (б)

Fig. 8. The number of states obtained in the experiments for the words „Forward“ (a) and „Stop“ (b)

Как видно из рис. 8, закон распределения отклонений числа состояний N от средних значений близок к нормальному. Соответственно для слова «Вперёд» среднее значение N составило $18 \pm 2.6\%$ при 95% -й доверительной вероятности, а для слова «Стоп» — $6 \pm 5.7\%$.

Следовательно, данные, полученные с помощью предложенного алгоритма предварительной обработки звукового файла в совокупности с КЛП-анализом, являются достаточно надёжными и их можно использовать для построения СММ слова.

Процесс построения начинается генерированием по известному числу состояний N и количеству символов M алфавита V исходной СММ $\lambda = (A, B, \pi)$, имеющей случайные параметры. При этом матрица вероятностей переходов между состояниями A , матрица вероятностей каждого наблюдения в каждом состоянии B , а также вектор вероятностей начального состояния π должны удовлетворять стохастическим ограничениям [3].

Далее необходимо так подобрать параметры исходной СММ, чтобы вероятность соответствия последовательности наблюдений, сгенерированной этой СММ $\lambda^* = (A^*, B^*, \pi^*)$ и полученной ранее последовательности наблюдений $O = o_1, \dots, o_T$ слова, была максимально возможной. То есть исходную СММ $\lambda = (A, B, \pi)$, имеющую вероятность $p(O|\lambda)$ генерирования заданной последовательности наблюдения $O = o_1, \dots, o_T$ слова, надо обучить по этой последовательности наблюдения $O = o_1, \dots, o_T$, чтобы вероятность $p(O|\lambda^*)$ генерирования последовательности $O = o_1, \dots, o_T$, обученной СММ $\lambda^* = (A^*, B^*, \pi^*)$, была максимально возможной.

Одним из вариантов обучения СММ $\lambda = (A, B, \pi)$ по заданной последовательности наблюдений $O = o_1, \dots, o_T$ является применение алгоритма Баума – Велша [3].



Алгоритм позволяет уточнять параметры исходной СММ $\lambda = (A, B, \pi)$ таким образом, чтобы у уточнённой СММ $\lambda^* = (A^*, B^*, \pi^*)$ вероятность $p(O|\lambda^*)$ увеличилась. Итеративное применение алгоритма до схождения в одной точке позволяет максимизировать $p(O|\lambda^*)$, т. е. настроить СММ $\lambda^* = (A^*, B^*, \pi^*)$ на заданную последовательность наблюдений $O = o_1, \dots, o_T$ слова. На рис. 9 приведены зависимости изменений вероятности $p(O|\lambda^*)$ и её приращения $\Delta p(O|\lambda^*)$ на каждом итерационном шаге при настройке методом Баума – Велша СММ на слово «Стоп». Эти зависимости имеют характерный для метода Баума – Велша вид.

В начале обучения значения вероятности $p(O|\lambda^*)$ имеют, как правило, величины меньшие или сравнимые с используемой для оценки сходимости положительной величиной ε (см. рис. 9, а), но разность значений $p(O|\lambda^*)$ на каждом шаге увеличивается, т. е. приращение $\Delta p(O|\lambda^*) > 0$ растёт (см. рис. 9, б).

При завершении обучения значения вероятности $p(O|\lambda^*)$, как правило, сходятся к некоторой величине, при этом приращение $\Delta p(O|\lambda^*) > 0$, но оно начинает уменьшаться и стремиться к 0.

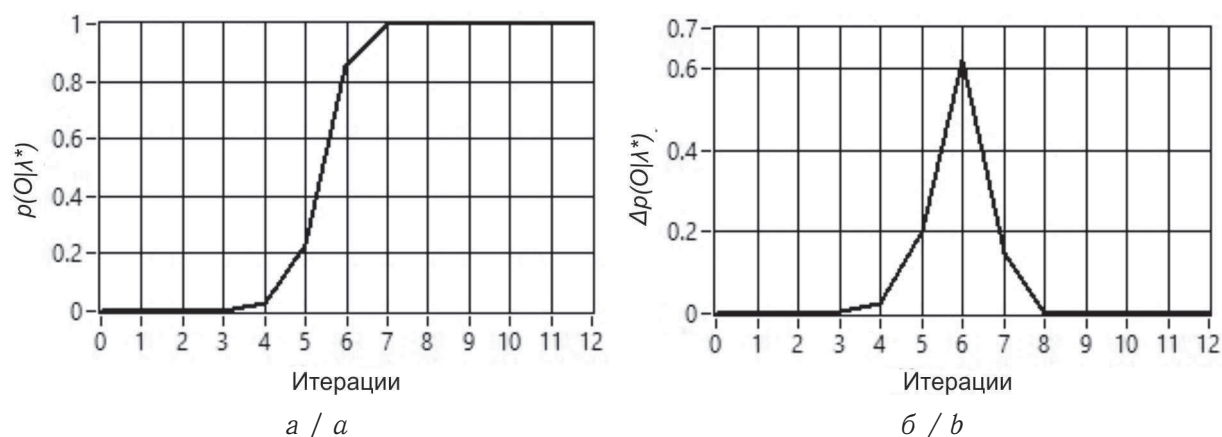


Рис. 9. Изменения вероятности СММ $p(O|\lambda^*)$: а — и её приращения $\Delta p(O|\lambda^*)$; б — при обучении методом Баума – Велша

Fig. 9. Changes in the probability of the HMM $p(O|\lambda^*)$: а — and its increment $\Delta p(O|\lambda^*)$; б — when learning by the Baum – Welsh method

Следовательно, для корректной оценки сходимости итерационного процесса обучения необходимо контролировать не только величину приращения $\Delta p(O|\lambda^*)$, но и знак его изменения, т. е. для завершения обучения должно выполняться условие $\Delta p(O|\lambda^*) \leq \varepsilon$ при уменьшении $\Delta p(O|\lambda^*)$.

В разработанном модуле построения СММ для оценки сходимости процесса обучения использовано значение $\varepsilon = 1.1 \cdot 10^{-19}$, соответствующее машинной точности.

Недостатком алгоритма Баума – Велша при обучении СММ является поиск локального максимума $p(O|\lambda^*)$, а не глобального. Поэтому для достижения хорошего результата требуется, как правило, несколько запусков при различных начальных условиях.

Таким образом, используя последовательности наблюдений и алфавит моделируемых слов, получаемые на первых этапах обработки, с помощью алгоритма Баума – Велша можно строить соответствующие адекватные СММ для систем распознавания речи.



ЗАКЛЮЧЕНИЕ

Разработанный программный модуль позволяет эффективно подготавливать необходимые исходные данные на основе кодирования признаков звукового сигнала, использующего линейные предсказания, строить СММ отдельных слов и проводить их обучение с помощью алгоритма Баума – Велша. Построенные СММ слов предполагается использовать в интеллектуальных системах управления различными объектами.

Библиографический список

1. Жилияков Е. Г., Бабаринов С. Л., Чадюк П. В. Исследование сервиса компании Google Inc. по распознаванию русской речи // Научные ведомости БелГУ. Сер. История. Политология. Экономика. Информатика. 2013. № 15(158), вып. 27/1. С. 247–255.
2. Титов Ю. Н. Современные технологии распознавания речи // Вестн. Тамб. ун-та. Сер. Естественные и технические науки. 2006. Т. 11, вып. 4. С. 571–574.
3. Савин А. Н., Тимофеева Н. Е., Гераськин А. С., Мавлютова Ю. А. Разработка компонентов программного комплекса для потоковой фильтрации аудиоконтента на основе использования скрытых марковских моделей // Изв. Саратов. ун-та. Нов. сер. Сер. Математика. Механика. Информатика. 2015. Т. 15, вып. 3. С. 340–350. DOI: 10.18500/1816-9791-2015-15-3-340-350.
4. Рабинер Л. Р. Скрытые Марковские модели и их применение в избранных приложениях при распознавании речи : Обзор // ТИИЭР. 1989. Т. 77, № 2. С. 86–120.
5. Портал компании National Instruments Russia. URL: <http://www.labview.ru> (дата обращения: 15.12.2017).
6. Блейхут Р. Быстрые алгоритмы цифровой обработки сигналов. М. : Мир, 1989. 448 с.
7. Адлер Ю. П., Маркова Е. В., Грановский Ю. В. Планирование эксперимента при поиске оптимальных условий. М. : Наука, 1976. 279 с.

Образец для цитирования:

Савин А. Н., Тимофеева Н. Е., Гераськин А. С., Мавлютова Ю. А. Разработка системы распознавания речи на основе скрытых марковских моделей отдельных слов // Изв. Саратов. ун-та. Нов. сер. Сер. Математика. Механика. Информатика. 2017. Т. 17, вып. 4. С. 452–464. DOI: 10.18500/1816-9791-2017-17-4-452-464.

Development of Speech Recognition Systems Based on Hidden Markov Models of Individual Words

A. N. Savin, N. E. Timofeeva, A. S. Geraskin, Yu. A. Mavlutova

Alexander N. Savin, orcid.org/0000-0001-5148-9166, Saratov State University, 83, Astrakhanskaya Str., Saratov, Russia, 410012, savinan@info.sgu.ru

Nadezhda E. Timofeeva, orcid.org/0000-0002-3976-3115, Saratov State University, 83, Astrakhanskaya Str., Saratov, Russia, 410012, timofeevane@yandex.ru

Aleksej S. Geraskin, orcid.org/0000-0002-3118-1022, Saratov State University, 83, Astrakhanskaya Str., Saratov, Russia, 410012, gerascinas@mail.ru

Yuliya A. Mavlutova, orcid.org/0000-0002-1190-90064, Saratov State University, 83, Astrakhanskaya Str., Saratov, Russia, 410012, yuliyamav@yandex.ru



The results of the development of software modules implementing the speech recognition system based on the hidden Markov models of individual words and the use of linear prediction in the coding of signs of an audio signal are presented. The structure of the speech recognition system is based on the hidden Markov models of individual words, consisting of four modules: a module for extracting words from the sound stream, a module for analyzing the features of a word, a module for learning the hidden Markov models, and a word recognition module. Algorithms for the formation of hidden Markov models with left-right topology for individual words of the required dictionary of commands of the object control system are based on the coding of signs of a sound signal using linear predictions. Results of an estimation of reliability of a sequence of observations corresponding to separate words obtained with the help of the proposed processing algorithm are given. The developed software modules allow to prepare efficiently the necessary initial data and thus form the required dictionary of commands of the object management system, build hidden Markov models of individual words, and conduct their training using the Baum – Welsh algorithm. The designed command dictionaries are supposed to be used in intelligent control systems for various objects.

Key words: Hidden Markov models, cepstral analysis, speech recognition, method of Baum – Welch.

References

1. Zhilyakov E. G., Babarinov S. L., Chadyuk P. V. Google Inc. Russian Speech Recognition Service Research. *Belgorod State University Scientific Bulletin. Ser. History. Political science. Economics. Information technologies*, 2013, no. 15(158), iss. 27/1, pp. 247–255 (in Russian).
2. Titov Y. N. Modern technologies of speech recognition. *Tambov University Reports. Ser. Natural and Technical Sciences*, 2006, vol. 11, iss. 4, pp. 571–574 (in Russian).
3. Savin A. N., Timofeeva N. E., Geraskin A. S., Mavlutova Yu. A. The development of software components for streaming audio content filtering through the use of hidden Markov models. *Izv. Saratov Univ. (N.S.), Ser. Math. Mech. Inform.*, 2015, vol. 15, iss. 3, pp. 340–350. DOI: 10.18500/1816-9791-2015-15-3-340-350.
4. Rabiner L. R. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 1989, vol. 77, no. 2, pp. 257–286. DOI: 10.1109/5.18626.
5. *Portal of company National Instruments Russia*. Available at: <http://www.labview.ru> (accessed 25 December, 2012).
6. Blahut R. E. *Fast Algorithms for Digital Signal Processing*. Addison-Wesley Publ. Co; Repr. with corrections edition, 1987. 441 p. (Russ. ed. : Moscow, Mir, 1989. 448 p.).
7. Adler Ju. P., Markova E. V., Granovskij Ju. V. *Planirovanie eksperimenta pri poiske optimal'nykh uslovii* [Planning an experiment searching for optimal conditions]. Moscow, Nauka, 1976. 279 p. (in Russian).

Cite this article as:

Savin A. N., Timofeeva N. E., Geraskin A. S., Mavlutova Yu. A. Development of Speech Recognition Systems Based on Hidden Markov Models of Individual Words. *Izv. Saratov Univ. (N. S.), Ser. Math. Mech. Inform.*, 2017, vol. 17, iss. 4, pp. 452–464 (in Russian). DOI: 10.18500/1816-9791-2017-17-4-452-464.
