

Министерство образования Республики Беларусь
Учреждение образования «Барановичский государственный университет»
Студенческое научное общество БарГУ

СОДРУЖЕСТВО НАУК. БАРАНОВИЧИ-2016

Материалы XII Международной
научно-практической конференции
молодых исследователей

(Барановичи, 19—20 мая 2016 года)

В трёх частях

Часть 2

Барановичи
БарГУ
2016

В части 2 сборника материалов XII Международной научно-практической конференции молодых исследователей «Содружество наук. Барановичи-2016» представлены результаты исследований в области физики и математики, а также рассмотрены актуальные проблемы в области информационных систем и технологий в образовании, науке и технике. Особое внимание уделено современным тенденциям в технологиях и материалах машиностроительного и сельскохозяйственного производств, а также экономическим аспектам развития предприятия, региона.

Сборник адресован научным работникам, аспирантам, магистрантам и студентам инженерных и экономических специальностей учреждений высшего образования.

Редакционная коллегия:

А. В. Никишова (гл. ред.), Ю. Е. Горбач, В. Н. Кременевская (отв. секретари), Е. Н. Кирюхова,
О. И. Наранович, А. К. Гавриленя, М. В. Нерода, В. Н. Познякевич, Г. Я. Житкевич

Рецензент

кандидат технических наук, заведующий лабораторией механофизики гетерогенных систем
Государственного научного учреждения «Физико-технический институт
Национальной академии наук» А. М. Милюкова

Научное издание

СОДРУЖЕСТВО НАУК.
БАРАНОВИЧИ-2016

Материалы XII Международной
научно-практической конференции
молодых исследователей

(Барановичи, 19—20 мая 2016 года)

На русском, белорусском, английском языках

В трёх частях

Часть 2

Ответственный за выпуск Е. Г. Хохол
Технический редактор А. Ю. Сидоренко
Компьютерная вёрстка С. М. Глушак
Корректор Н. Н. Колодко

Подписано в печать 04.10.2016. Формат 60 × 84 ¹/₈. Бумага ксероксная.

Отпечатано на копировально-множительной технике. Усл. печ. л. 28,00. Уч.-изд. л. 25,10. Тираж 9 экз. Заказ 681.

Учреждение образования «Барановичский государственный университет».
Свидетельство о государственной регистрации издателя № 1/424 от 09.09.2016.
Ул. Войкова, 21, 225404 г. Барановичи. Тел. 8 (0163) 45 46 28, e-mail: rio@barsu.by .

СРАВНЕНИЕ МЕТОДОВ ИДЕНТИФИКАЦИИ ДИКТОРОВ ПО ГОЛОСУ

Введение. В настоящий момент область речевых технологий является одной из наиболее динамично развивающихся областей науки. Это связано с наличием большого числа актуальных задач, связанных с обработкой речи человека. Многие задачи из области речевых технологий представляют интерес как для бизнес-структур, так и для правительственных организаций. Системы автоматического выделения ключевых слов могут использоваться как для индексирования речевых баз данных, так и для поиска в телефонном канале определённых слов, фраз и т. д.

Поддержка речевых технологий из разных источников привела к развитию автоматических методов исследования. Но ввиду сложности и комплексного характера задач, возникающих в процессе обработки естественного речевого сигнала, эти задачи далеки от того, чтобы можно было считать их решёнными как в практическом, так и в научном направлении [1].

Целью данной работы является исследование методов и алгоритмов идентификации дикторов по голосу.

В соответствии с указанной целью сформулированы следующие задачи: 1) выполнить обзор методов, алгоритмов и систем идентификации дикторов по голосу; 2) провести сравнительный анализ основных методов идентификации.

Основная часть. Каждый человек имеет индивидуальные голосовые характеристики, которые определяют особенностями строения его голосовых органов. В процессе общения люди способны на подсознательном уровне различать голоса других людей, однако для вычислительной техники данная задача является нетривиальной.

Задача распознавания личности по голосу сводится к тому, чтобы выделить, классифицировать и соответствующим образом отреагировать на человеческую речь из входного звукового потока. При этом обычно выделяют две подзадачи: идентификация и верификация.

Идентификация — процесс определения личности по образцу голоса путём сравнения данного образца с шаблонами, сохранёнными в базе. Верификация — процесс, при котором с помощью сравнения представленного образца с хранимым в базе шаблоном проверяется запрошенная идентичность. Из определения следует, что при верификации вместе с образцом голоса передаётся идентификатор пользователя, зарегистрированного в системе.

Кроме того, системы распознавания могут быть разделены на текстозависимые и текстонезависимые. При текстозависимом распознавании могут использоваться как фиксированные фразы, так и фразы, сгенерированные системой и предложенные пользователю. Текстонезависимые системы предназначены обрабатывать произвольную речь [2].

Процесс определения диктора, зарегистрированного в системе, по входному речевому сигналу во всех рассматриваемых методах состоит в поиске наиболее подходящей сохранённой модели на основе каких-либо критериев.

На основании данных, полученных с помощью опытов, использующих субъективные методы, основное проявление индивидуальности речи человека следует искать в двух основных группах признаков. Они связаны с физиологическими (анатомическими) особенностями механизма речеобразования человека и уникальным характером поведения его в действии (артикуляционной деятельностью), обусловленным работой центральной нервной системы.

Первая группа признаков основывается на хорошо известной модели речевого тракта, состоящей из передаточной функции резонансной системы и генератора импульсов сигнала возбуждения. Передаточная функция практически полностью характеризует индивидуальную геометрическую форму полостей речевого аппарата: задняя глоточная полость, сужение между языком и небом, передняя полость рта, сужение между губами и т. д. Основными параметрами здесь выступают характеристики четырёх формантных областей (средняя частота, частотный диапазон, энергия), огибающая спектра, формантные траектории и производные от этих параметров. Частота импульсов возбуждения находится в прямой зависимости от колебаний голосовых связок, которые, в свою очередь, зависят от длины, толщины и натяжения последних. Основными параметрами здесь являются частота основного тона, тон-шум, звонкость, подъём основного тона и производные от этих параметров.

Для расчёта параметров, связанных с физиологическими особенностями речевого тракта, наиболее часто используются методы спектрально-временного анализа. Такие методы анализа речевого сигнала адекватны природному механизму восприятия речи, что делает понятной тенденцию многих исследователей искать индивидуальные особенности в мгновенных спектральных распределениях отдельных фонем и в распределениях текущего спектра. В основе таких методов лежит классический Фурье-анализ или параметрический авторегрессионный анализ (линейное предсказание как частный случай).

Тесно связан со спектральным представлением речевого сигнала довольно часто применяемый в последнее время гомоморфный метод. Этот метод представляет речевой сигнал в виде последовательности векторов кепстральных коэффициентов, которые требуют значительно меньшего объёма памяти для хранения эталонных образов. Небольшим количеством кепстральных коэффициентов (обычно 8 или 16) можно аппроксимировать формантный разрез, имеющий высокое спектральное разрешение. Это обеспечивает более компактное представление речевых отрезков без существенной потери основных информативных признаков (формантной структуры, огибающей, параметра тон-шум). Что касается параметров сигнала возбуждения, то они могут быть рассчитаны одним из

широко известных методов выделения частоты основного тона (например, корреляционный метод, кепстральный метод, метод Голда—Рабинера) [3].

Если первая группа признаков отражает статические свойства речеобразующего тракта, то вторая группа призвана полностью описать его поведение во времени, т. е. артикуляционную динамику речи. Согласно существующему предположению, исходным и основным этапом в организации процесса речеобразования является управляемая центральной нервной системой человека программа комплекса артикуляционных движений, соответствующая тому сообщению, передача которого планируется в данный момент времени.

Для расчёта параметров, описывающих артикуляционную динамику речи, могут быть использованы методы спектрально-временного анализа данных, описанные выше. Однако необходимо отметить такую особенность расчёта просодических параметров, как их жёсткая связь с лексическим и синтаксическим контекстом исследуемой фразы. Это требует комплексного применения как средств лингвистического анализа, так и параметрических методов обработки, что явно определяет сложность анализа данных характеристик. При этом основной задачей является установление прямой связи между деятельностью речеобразующего аппарата (динамикой его артикуляционных движений) и характеристиками спектральной картины потока речи [4].

Большинство разработанных на сегодня систем идентификации личности по голосу построены на основе однократной проверки соответствия требуемой ключевой фразы и произнесённой в первоначальный момент доступа к вычислительной системе. Данные системы поддерживают два основных режима работы: обучение системы и проверка подлинности при доступе.

Основным достоинством описанных выше систем является простота построения. Широкие возможности их реализации на основе стандартных процедур цифровой обработки сигнала и невысокие требования к вычислительным ресурсам и объёму памяти ЭВМ сделали такие системы почти хрестоматийным примером при изучении теории автоматического определения человека по голосу.

Заключение. В рамках данной работы исследовались существующие методы решения задачи автоматической идентификации диктора по голосу.

Речь представляет собой сложный сигнал, образующийся в результате преобразований, происходящих на нескольких различных уровнях: семантическом, языковом, артикуляционном (уровне голосового аппарата человека) и акустическом (уровне физических свойств звука). Различия в этих преобразованиях влекут за собой различия в свойствах речевого сигнала. При решении задачи распознавания диктора по голосу все эти различия могут быть использованы для того, чтобы выделить индивидуальные характеристики голоса каждого человека.

Список цитируемых источников

1. Симончик К. К. Методы и алгоритмы автоматической текстонезависимой верификации дикторов и их программная реализация. СПб. : КопиСервис, 2010.
2. Машонский И. Д. Автоматическая идентификация диктора по голосу. М. : МГУ им. Ломоносова, 2014.
3. Там же.
4. Идентификация пользователей по голосу [Электронный ресурс]. URL: <http://www.speech-soft.ru/user-identification-by-voice?destination=node/14> (дата обращения: 04.02.2016).

УДК 004.02

А. Ю. Бузук, А. В. Максимович

Учреждение образования «Барановичский государственный университет», Барановичи

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ОБЪЕКТНО-ОРИЕНТИРОВАННОГО ПРОГРАММИРОВАНИЯ И ПРОЦЕДУРНОГО ПРОГРАММИРОВАНИЯ НА ПРИМЕРЕ РАЗРАБОТАННОГО ПРИЛОЖЕНИЯ

Введение. Сегодня существуют различные по сложности программы, что обуславливает программистов изучать и использовать тот или иной стиль программирования для создания своего приложения.

Выбор подхода для написания программ зависит от сложности технического задания, а иногда и от подготовки и выбора технологии программирования. Но всё же хотелось бы разобраться, что целесообразнее использовать объектно ориентированное программирование (далее — ООП) или процедурное программирование (далее — ПП).

Основная часть. Метод ООП основан на том, что любая программа состоит из объектов. Объект, в свою очередь, является экземпляром определённого класса. Классы позволяют проводить конструирование из полезных компонент, обладающих простыми инструментами, что даёт возможность абстрагироваться от деталей реализации. Данные и операции вместе образуют определённую сущность и они не «размазываются» по всей программе, как это нередко бывает в случае процедурного программирования. Локализация кода и данных улучшает наглядность и удобство сопровождения программного обеспечения. Инкапсуляция информации защищает наиболее критичные данные от несанкционированного доступа [1].

Стиль ООП даёт возможность создавать расширяемые системы. Это одно из самых значительных достоинств ООП, и именно оно отличает данный подход от традиционных методов программирования. Расширяемость