

РЕФЕРАТ

Актуальність теми. Зараз систем розпізнавання мовлення набувають все більшої популярності та зустрічаються все частіше. Успішними прикладами використання технології розпізнавання мови в мобільних додатках є: введення адреси голосом в Яндекс.Навігатор, голосовий пошук Google Now.

Крім мобільних пристроїв, технологія розпізнавання мови знаходить широке поширення в різноманітних сферах людської діяльності:

- Телефонія: автоматизація обробки вхідних і вихідних дзвінків шляхом створення голосових систем самообслуговування зокрема для: отримання довідкової інформації та консультування, замовлення послуг, товарів, зміни параметрів чинних послуг, проведення опитувань, анкетування, збору інформації, інформування та будь-які інші сценарії;
- Рішення "Розумний будинок": голосовий інтерфейс управління системами «Розумний будинок»;
- Побутова техніка і роботи: голосовий інтерфейс електронних роботів; голосове керування побутовою технікою тощо;
- Автомобілі: голосове управління в салоні автомобіля - наприклад, навігаційною системою;
- Соціальні сервіси для людей з обмеженими можливостями;
- Комплексні системи захисту інформації. Голосова аутентифікація.
- Визначення емоційного забарвлення голосу диктора.

Об'єктом дослідження є процес розпізнавання голосових сигналів.

Предметом дослідження є методи та моделі розпізнавання голосових сигналів в умовах обмеженої обчислювальної потужності.

Мета роботи: підвищення ефективності процесу розпізнавання голосових сигналів в умовах обмеженої обчислювальної потужності.

Методи дослідження. В роботі використовуються методи математичного моделювання, методи оптимізації, методи системного аналізу, чисельні методи.

Наукова новизна роботи полягає в наступному:

1. Була модифікована модель генерації параметризованих еталонів голосових сигналів за рахунок проведення апроксимації схожих еталонів.

2. Отримав подальший розвиток метод використання алгоритму динамічної трансформації часової шкали в системах розпізнавання голосових сигналів з обмеженим словником за рахунок використання розробленої моделі генерації параметризованих еталонів.

Практична цінність отриманих в роботі результатів полягає в тому, що:

- запропонована модифікація алгоритму динамічної трансформації часової шкали;
- розроблена модель для тестування запропонованого алгоритму ;
- наведена порівняльна характеристика розробленого алгоритму та існуючих аналогів.

Апробація роботи. Основні положення і результати роботи були представлені та обговорювались на:

- науковій конференції магістрантів та аспірантів «Прикладна математика та комп'ютеринг» ПМК-2018 (Київ, 21-13 березня 2018 р.) та опубліковані у збірнику "Наукові вісті НТУУ «КПІ»" № 20, 2018р;
- міжнародній науково-практичній конференції студентів і молодих учених "Інформаційні технології в соціокультурній сфері, освіті, економіці та праві" (Київ, 18-19 квітня 2018 р.) та опубліковані у збірнику тез конференції;
- науково-практичній конференції студентів і аспірантів "Теоретичні та прикладні аспекти розробки комп'ютерних систем" (Київ, 29 березня 2018 р.) та опубліковані у збірнику тез конференції.

Структура та обсяг роботи. Магістерська дисертація складається з вступу, чотирьох розділів, висновків та додатків.

У вступі надано загальну характеристику роботи, виконано оцінку сучасного стану проблеми, обґрунтовано актуальність напрямку досліджень, сформульовано мету і задачі досліджень, показано наукову новизну отриманих результатів і практичну цінність роботи, наведено відомості про апробацію результатів і їх впровадження.

У першому розділі наведені теоретичні засади щодо розробки систем розпізнавання голосових сигналів та проаналізовані існуючі системи розпізнавання голосових сигналів.

У другому розділі наведена модель процесу розпізнавання. Детальніше описані проблеми, що виникають під час розробки систем розпізнавання зливої мови.

У третьому розділі наведена математична модель розробки систем розпізнавання голосових сигналів. Детальніше описані етапи роботи алгоритму динамічної трансформації часової шкали.

У четвертому розділі описано хід експериментів та наведено результати експериментів.

У висновках проаналізовано отримані результати роботи.

У додатках наведено алгоритм алгоритму динамічної трансформації часової шкали, часові характеристики роботи програмного модуля розпізнавання голосових сигналів та порівняльні діаграми ефективності отриманих результатів.

Робота виконана на 80 аркушах, містить 4 додатки та посилання на список використаних літературних джерел з 40 найменувань. У роботі наведено 12 рисунків та 8 таблиць.

Ключові слова: алгоритм динамічної трансформації часової шкали, розпізнавання голосових сигналів.

ABSTRACT

Theme urgency. Speech recognition systems are becoming increasingly popular and increasingly common. Successful examples of using speech recognition technology in mobile applications are: entering a voice address in Yandex.Navigator, Google Now voice search.

In addition to mobile devices, speech recognition technology is widely used in various areas of human activity:

- Telephony: automates the processing of incoming and outgoing calls by creating voice self-service systems in particular for: receiving background information and advice, ordering services, goods, changing the parameters of current services, conducting surveys, questionnaires, collecting information, informing and any other scenarios;
- “Smart House” solutions: voice interface for intelligent home systems management;
- Household appliances and work: voice interface of electronic robots; voice control of home appliances, etc .;
- Cars: voice control in the car - for example, the navigation system;
- Social services for people with disabilities;
- Comprehensive information security systems. Voice authentication.
- Determination of the emotional color of the speaker's voice.

Object of research are systems and algorithms for voice recognition.

Subject of research is a usage of dynamic time warping algorithm in speech recognition systems in the conditions of limited computing power

Research objective: development and modification of the dynamic time warping algorithm for recognizing a limited vocabulary.

Research methods. Methods of mathematical modeling, methods of optimization, methods of system analysis, numerical methods are used in this work.

Scientific novelty of the work is as follows:

1. Proposed the modification of the dynamic time warping algorithm, which is based on the approximation of similar standards, which allows to reduce the time of the algorithm without losing the accuracy of the recognition.

2. An example of the practical use of the algorithm of the dynamic time warping in the systems of recognition of voice signals of a limited dictionary and comparison of the algorithm with existing analogues is given.

Practical value obtained in the work is:

- proposed modification of the algorithm of the dynamic time warping;
- a model for testing the proposed algorithm is developed;
- comparative characteristic of the developed algorithm and existing analogues is given.

Approbation. The main positions and results of the work were presented and discussed at:

- the 2nd scientific conference of masters and postgraduates "Applied Mathematics and Computing", PMK-2018 (Kyiv, March 21-13, 2018) and published in the collection "Scientific News of NTUU" KPI "" No. 12, 2018;
- international scientific and practical conference of students and young scientists "Information Technologies in the Socio-Cultural, Education, Economics and Law" (Kyiv, April 18-19, 2018) and published in the conference abstracts;
- scientific and practical conference of students and postgraduates "Theoretical and Applied Aspects of Computer Systems Development" (Kiev, March 29, 2018) and published in the conference abstracts.

Structure and content of the thesis. The master's thesis consists of an introduction, four chapters, conclusions and appendices.

The introduction provides a general description of the work, evaluates the current state of the problem, substantiates the relevance of the research direction, formulates the purpose and objectives of the research, shows the scientific novelty of the results obtained and the practical value of the work, provides information about the approbation of the results and their implementation.

In the first chapter the theoretical principles concerning the development of voice signal recognition systems are presented and the existing systems of voice signal recognition are analyzed.

The second chapter presents the model of the recognition process. The problems that arise during the development of intelligent language recognition systems are described in more detail.

The third chapter provides a mathematical model for the development of voice recognition systems. Details of the stages of the algorithm of the dynamic time table transformation are described.

The fourth chapter describes the progress of the experiments and gives the results of the experiments.

The conclusions are analyzed the results of work.

The appendices provide algorithm for the dynamic time warping algorithm, time characteristics of the program module for recognition of voice signals and comparative diagrams of the effectiveness of the obtained results.

The work is performed on 80 sheets, contains 4 attachments and a link to the list of used literary sources of 40 titles. The paper presents 12 figures and 8 tables.

Key words: dynamic time warping algorithm, voice recognition.

РЕФЕРАТ

Актуальность темы. Сейчас системы распознавания речи приобретают все большую популярность и встречаются все чаще. Успешными примерами использования технологии распознавания речи в мобильных приложениях являются: ввод адреса голосом в Яндекс.Навигатор, голосовой поиск Google Now.

Кроме мобильных устройств, технология распознавания речи находит широкое распространение в различных сферах человеческой деятельности:

- **Телефония:** автоматизация обработки входящих и исходящих звонков путем создания голосовых систем самообслуживания в частности для: получения справочной информации и консультирование, заказ услуг, товаров, изменения параметров действующих услуг, проведения опросов, анкетирования, сбора информации, информирование и любые другие сценарии;
- **Решение "Умный дом":** голосовой интерфейс управления системами «Умный дом»;
- **Бытовая техника и работы:** голосовой интерфейс электронных роботов голосовое управление бытовой техникой и т.д.;
- **Автомобили:** голосовое управление в салоне автомобиля - например, навигационной системой;
- **Социальные сервисы для людей с ограниченными возможностями;**
- **Комплексные системы защиты информации.** Голосовая аутентификация.
- **Определение эмоциональной окраски голоса диктора.**

Объектом исследования являются системы и алгоритмы распознавания голосовых сигналов.

Предметом исследования является алгоритм динамической трансформации временной шкалы в системах распознавания голосовых сигналов в условиях ограниченной вычислительной мощности.

Цель работы: разработка и модификация алгоритма динамической трансформации временной шкалы для распознавания ограниченного словаря.

Методы исследования. В работе используются методы математического моделирования, методы оптимизации, методы системного анализа, численные методы.

Научная новизна работы заключается в следующем:

1. Предложена модификация алгоритма динамической трансформации временной шкалы, которая базируется на аппроксимации схожих эталонов, что позволяет уменьшить время работы алгоритма не теряя точности распознавания.

2. Приведен пример практического использования алгоритма динамической трансформации временной шкалы в системах распознавания голосовых сигналов ограниченного словаря и сравнения алгоритма с существующими аналогами.

Практическая ценность полученных в работе результатов заключается в том, что:

- предложена модификация алгоритма динамической трансформации временной шкалы;
- разработана модель для тестирования предложенного алгоритма;
- приведена сравнительная характеристика разработанного алгоритма и существующих аналогов.

Апробация работы. Основные положения и результаты работы были представлены и обсуждались на:

- II научной конференции магистрантов и аспирантов «Прикладная математика и компьютеринг» ПМК-2018 (Киев, 21-13 марта 2018) и опубликованы в сборнике "Научные вести НТУУ «КПИ» № 12 2018;

- международной научно-практической конференции студентов и молодых ученых "Информационные технологии в социокультурной сфере, образовании, экономике и праве" (Киев, 18-19 апреля 2018) и опубликованы в сборнике тезисов конференции;
- научно-практической конференция студентов и аспирантов "Теоретические и прикладные аспекты разработки компьютерных систем" (Киев, 29 марта 2018) и опубликованы в сборнике тезисов конференции.

Структура и объем работы. Магистерская диссертация состоит из введения, четырех глав, заключения и приложений.

Во введении приведена общая характеристика работы, выполнена оценка современного состояния проблемы, обоснована актуальность направления исследований, сформулированы цели и задачи исследований, показано научную новизну полученных результатов и практическую ценность работы, приведены сведения об апробации результатов и их внедрение.

В первом разделе приведены теоретические основы по разработке систем распознавания голосовых сигналов и проанализированы существующие системы распознавания голосовых сигналов.

Во втором разделе приведена модель процесса распознавания. Подробнее описаны проблемы, возникающие при разработке систем распознавания слитной речи.

В третьем разделе приведена математическая модель разработки систем распознавания голосовых сигналов. Подробнее описаны этапы работы алгоритма динамической трансформации временной шкалы.

В четвертом разделе описан ход экспериментов и приведены результаты экспериментов.

В выводах проанализированы полученные результаты работы.

В приложениях приведены алгоритм алгоритма динамической трансформации временной шкалы, временные характеристики работы программного модуля распознавания голосовых сигналов и сравнительные диаграммы эффективности полученных результатов.

Работа выполнена на 80 листах, содержит 4 приложения и ссылки на список использованных литературных источников из 40 наименований. В работе приведены 12 рисунков и 8 таблиц.

Ключевые слова: алгоритм динамической трансформации временной шкалы, распознавание голосовых сигналов.