МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ТЕРНОПІЛЬСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ ІВАНА ПУЛЮЯ
ФАКУЛЬТЕТ КОМП'ЮТЕРНО-ІНФОРМАЦІЙНИХ СИСТЕМ
І ПРОГРАМНОЇ ІНЖЕНЕРІЇ

ЧАСНИК ВАСИЛЬ АНАТОЛІЙОВИЧ

УДК 004.522

МЕТОДИ ТА ЗАСОБИ АВТОМАТИЧНОГО РОЗПІЗНАВАННЯ МОВИ НА БАЗІ МІКРОКОНТРОЛЕРНОЇ СИСТЕМИ

123 «Комп'ютерна інженерія»

Автореферат
дипломної роботи на здобуття освітнього ступеня «магістр»

Тернопіль 2019

Роботу виконано на кафедрі комп'ютерних систем та мереж Тернопільського національного технічного університету імені Івана Пулюя Міністерства освіти і науки України

Керівник роботи: Ph.D., доцент кафедри
комп'ютерних систем та мереж
Луцик Надія Степанівна,
Тернопільський національний технічний університет
імені Івана Пулюя

Рецензент: асистент кафедри
комп'ютерних наук
Липак Галина Ігорівна,
Тернопільський національний технічний університет
імені Івана Пулюя

Захист відбудеться 26 грудня 2019 р. о $9^{.00}$ годині на засіданні екзаменаційної комісії №37 у Тернопільському національному технічному університеті імені Івана Пулюя за адресою: 46001, м. Тернопіль, вул. Руська, 56, навчальний корпус №1, ауд. 603

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми роботи. Не важко зрозуміти привабливість мовлення як засіб для взаємодії між людиною та комп'ютером. Для більшості людей мова є найшвидшим та найлегшим способом комунікації, на відміну будь-яким іншим способам (наприклад, набору або написання тексту). Таким чином, мовлення є більш природним та інтуїтивно зрозумілим для середньостатистичного користувача комп'ютера, що змушує практично всіх лідерів індустрії вважати, що забезпечення взаємодії людини з комп'ютером через мовлення є перспективним напрямком розвитку машинно-людинних інтерфейсів.

Розпізнавання мовлення, також відоме як автоматичне розпізнавання мовлення, визначається як процес перетворення мовного сигналу в набір слів певним алгоритмом, який може бути реалізований як системна програма, або процес перетворення акустичного сигналу в набір слів.

Сучасні комп’ютерні системи обробки мови досягнули великої складності та високої точності внаслідок використання складних алгоритмів обробки сигналів та появи потужних обчислювальних засобів. Тим не менш, ріст продуктивності вбудованих процесорів (які також називаються мікроконтролерами) дозволив реалізувати розпізнавання мови в споживчій електроніці, де вартість має найбільше значення.

Застосування дешевих мікроконтролерів для автоматичного розпізнавання мови накладає обмеження на складність системи внаслідок малого об’єму оперативної та постійної пам’яті і невеликої обчислювальної потужності. Тому вибір параметрів та методів її реалізації є досить складною задачею.

В опублікованих роботах, у цьому напрямі, пропонувалось застосування спеціалізованих інтегральних схем (ASIC) або окремих спеціалізованих процесорів. Недоліком таких систем є висока вартість та складність розробки. Також пропонувалось використання недорогих стандартних мікроконтролерів, однак достатню точність розпізнавання, у таких системах, досягнуто лише для дикторів, на мовленні яких системи навчено.

На основі проведеного аналізу існуючих методів та засобів автоматичного розпізнавання мови поставлено за мету створити недорогу систему розпізнавання мови на базі мікроконтролера STM32F446. Користуючись стандартними інструментами розробки і застосувавши новий метод попередньої обробки голосового сигналу для мікроконтролера та навчання класифікаційної нейронної мережі, необхідно запрограмувати мікроконтролер початкового рівня для автоматичного розпізнавання мови з достатньою точністю.

Метою роботи є дослідження існуючих методів та засобів автоматичного розпізнавання мовлення на базі мікроконтролерної системи та створення недорогої системи розпізнавання мови на базі мікроконтролера STM32F446, застосувавши новий метод попередньої обробки голосового сигналу для мікроконтролера та навчання класифікаційної нейронної мережі, для автоматичного розпізнавання мовлення з достатньою точністю.

Для досягнення вказаної мети в роботі поставлено наступні **задачі**:

- аналіз наукових публікацій щодо автоматичного розпізнавання мови, зокрема розпізнавання мови на базі мікроконтролерних систем;
- дослідження алгоритмічного та математичного забезпечення автоматичного розпізнавання мови;
- формалізація алгоритмічного та математичного забезпечення запропонованих методів та засобів автоматичного розпізнавання мови;
- побудова моделі автоматичного розпізнавання мови на базі запропонованих методів та засобів;
- порівняння існуючих систем автоматичного розпізнавання мови на базі мікроконтролерних систем з побудованою системою;
- апробація запропонованих методів та засобів автоматичного розпізнавання мови на базі мікроконтролерної системи.

Об'єкт дослідження: процес автоматичного розпізнавання мови на базі мікроконотролерної системи.

Предмет дослідження: методи і засоби автоматичного розпізнавання мови на базі мікроконотролерної системи.

Методи дослідження. Для вирішення поставлених задач використано наступні методи: аналіз та узагальнення – при проведені аналізу існуючих методів та засобів автоматичного розпізнавання мови; теорії аналізу часових рядів, теорії машинного навчання та добування даних – при розробці запропонованого методу розпізнавання мови; проектування та програмування – при побудові системи автоматичного розпізнавання мови на базі мікроконтролерної системи; експеримент та вимірювання – для апробації запропонованих методів та засобів.

Наукова новизна одержаних результатів. Запропоновано поліпшений метод виявлення ознак у мовленнєвому сигналі, який базується на використанні вейвлет-аналізу сигналу за допомогою вейвлету Добеші четвертого порядку, з розбиттям сигналу на одинадцять частотних смуг.

Для реалізації системи автоматичного розпізнавання мови використано мікроконтролерну систему на базі мікроконтролера STM32F446.

Практичне значення одержаних результатів. Можливим є впровадження запропонованих методів та засобів автоматичного розпізнавання мови на базі мікроконтролерної системи та їх реалізація у вигляді системи, яка дає змогу автоматизувати керування іншими системами за допомогою набору голосових команд.

Публікації. Результати дослідження апробовано на VIII Міжнародній науково-технічній конференції молодих учених та студентів. «Актуальні задачі сучасних технологій» (27-28 листопада 2019 р.) Тернопільського національного технічного університету імені Івана Пулюя та на VII науково-технічній конференції Тернопільського національного технічного університету імені Івана Пулюя «Інформаційні моделі, системи та технології» (11-12 грудня 2019 року) у вигляді тез конференцій.

Структура роботи. Робота складається з пояснювальної записки та графічної частини. Пояснювальна записка складається із вступу, шести розділів, висновків, переліку посилань та додатку. Обсяг роботи: пояснювальна записка – 116 арк. формату А4, графічна частина – 10 аркушів формату А1.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У вступі обґрунтовано актуальність дослідження методів і засобів автоматичного розпізнавання мови, сформульовано мету, задачі і методи дослідження, наведено наукову новизну та практичне значення одержаних результатів.

У першому розділі «Аналіз існуючих методів та засобів автоматичного розпізнавання мови» досліджено існуючі методи та засоби розпізнавання мови, досліджено класифікацію систем розпізнавання мови. Проаналізовано існуючі реалізації систем розпізнавання мови на базі мікроконтролерних систем, у результаті встановлено, що існуючі реалізації можна покращити шляхом застосування поліпшених методів та засобів автоматичного розпізнавання мови, та використання нової мікроконтролерної системи.

У другому розділі «Вибір оптимальних методів та засобів автоматичного розпізнавання мови» визначено оптимальний метод автоматичного розпізнавання мови, який здійснюється за допомогою двох блоків: front-end та back-end процесорів. Front-end процесор займається зчитуванням звукового сигналу, його попередньою. фільтрацією, розбиттям на фрейми та накладанням вікна, визначенням початку та кінця мовлення та виділенням ознак за допомогою вейвлет-перетворення. Back-end процесор займається порівнянням представлення звукового сигналу у вигляді вектора ознак, з базою шаблонів відомих представлень мовленнєвих сигналів, використовуючи повнозв’язну штучну нейронну мережу, з сигмоїдальною активаційною функцією.

У третьому розділі «Реалізація системи розпізнавання мови» за вибраними у попередньому розділі методами та засобами автоматичного розпізнавання мовлення, було змодельовано процес розпізнавання мовлення та здійснено його реалізацію на базі мікроконтролерної системи. Процес моделювання виконано засобами програмного пакету Matlab. У процесі моделювання було створено моделі попередньої фільтрації сигналу, розбиття сигналу на фрейми, виявлення голосової активації та виявлення ознак, за допомогою вейвлет-перетворення та класифікаційної штучної нейронної мережі. Отримані моделі було реалізовано на мікроконтролерній системі на базі плати розробки «STM32 Nucleo-64». Після чого було здійснено оцінку роботи отриманої системи та здійснено порівняння отриманих результатів з існуючими системами розпізнавання мовлення.

У четвертому розділі «Обґрунтування економічної ефективності» обґрунтовано економічну ефективність проведення досліджень дипломної роботи магістра шляхом проведення відповідних розрахунків, що дало змогу встановити термін окупності на рівні 1,79 років при ціні 46749,71 грн.

У п’ятому розділі «Охорона праці та безпека в надзвичайних ситуаціях» проаналізовано вимоги з охорони праці та безпеки в надзвичайних ситуаціях, що дало змогу визначити шляхи підтримання безпеки життєдіяльності в системі "людина - життєве середовище" та дослідити виробничий і побутовий травматизм, наслідки і заходи щодо його попередження.

У шостому розділі «Екологія» проаналізовано методи узагальнення екологічної інформації, та отримання енергії за рахунок альтернативних джерел.

Додатки до пояснювальної записки містять матеріали конференцій у яких опубліковано основні результати дипломної роботи магістра.

У графічній частині до дипломної роботи магістра проілюстровано основні наукові та практичні результати щодо запропонованих методів та засобів автоматичного розпізнавання мови на базі мікроконтролерної системи, наведене представлення запропонованого процесу розпізнавання мови.

ВИСНОВКИ

У дипломній роботі магістра досліджено існуючі методи та засоби розпізнавання мови та класифікацію систем розпізнавання мови. Проаналізовано існуючі реалізації систем розпізнавання мови на базі мікроконтролерних систем, у результаті встановлено, що існуючі реалізації можна покращити шляхом застосування поліпшених методів та засобів автоматичного розпізнавання мови, та застосування нової мікроконтролерної системи.

Визначено оптимальний метод автоматичного розпізнавання мови, який здійснюється за допомогою двох блоків: front-end та back-end процесорів. Front-end процесор займається зчитуванням звукового сигналу, його попередньою. фільтрацією, розбиттям на фрейми та накладанням вікна, визначенням початку та кінця мовлення та виділенням ознак за допомогою вейвлет-перетворення. Back-end процесор займається порівнянням представлення звукового сигналу у вигляді вектора ознак, з базою шаблонів відомих представлень мовленнєвих сигналів, використовуючи повнозв’язну штучну нейронну мережу.

За вибраними методами та засобами автоматичного розпізнавання мовлення, було змодельовано процес розпізнавання мовлення та здійснено його реалізацію на базі мікроконтролерної системи. Процес моделювання виконано засобами програмного пакету Matlab. У процесі моделювання було створено моделі попередньої фільтрації сигналу, розбиття сигналу на фрейми, виявлення голосової активації та виявлення ознак, за допомогою вейвлет-перетворення та класифікаційної штучної нейронної мережі. Отримані моделі було реалізовано на мікроконтролерній системі на базі плати розробки «STM32 Nucleo-64». Після чого було здійснено оцінку роботи отриманої системи та здійснено порівняння отриманих результатів з існуючими системами розпізнавання мовлення.

Обґрунтовано економічну ефективність проведення досліджень дипломної роботи магістра шляхом проведення відповідних розрахунків, що дало змогу встановити термін окупності на рівні 1,79 років при ціні 46749,71 грн.

Проаналізовано вимоги з охорони праці та безпеки в надзвичайних ситуаціях, що дало змогу визначити шляхи підтримання безпеки життєдіяльності в системі "людина-життєве середовище" та дослідити виробничий і побутовий травматизм, наслідки і заходи щодо його попередження.

Проаналізовано методи узагальнення екологічної інформації, та отримання енергії за рахунок альтернативних джерел.

СПИСОК ОПУБЛІКОВАНИХ АВТОРОМ ПРАЦЬ ЗА ТЕМОЮ РОБОТИ

1. Часник В. А., Луцик Н. С. Методи та засоби автоматичного розпізнавання мови на базі мікроконтролерної системи. *Актуальні задачі сучасних технологій*: зб. тез доповідей VIII міжн. наук.-техн. конф. Тернопіль: ТНТУ, 2019. С. 114.

2. Часник В. А., Луцик Н. С. Процес автоматичного розпізнавання мови на базі мікроконтролерної системи. *Інформаційні моделі, системи та технології*: зб. тез доповідей VII наук.-техн. конф. Тернопіль: ТНТУ, 2019. С. 143.

АНОТАЦІЯ

Часник В.А. Методи та засоби автоматичного розпізнавання мови на базі мікроконтролерної системи // Дипломна робота // Часник Василь Анатолійович // Тернопільський національний технічний університет імені Івана Пулюя, факультет комп'ютерно – інформаційних систем та програмної інженерії, група СІм – 61 // Тернопіль, 2019 // с. – 116, рис. – 23, табл. – 8, аркушів А1 – 10, додат. – 1, бібліогр. – 96.

Ключові слова: РОЗПІЗНАВАННЯ МОВИ, ВЕЙВЛЕТ-АНАЛІЗ, НЕЙРОННА МЕРЕЖА, МІКРОКОНТРОЛЕРНА СИСТЕМА, STM32.

Дипломна робота на здобуття освітнього ступеня магістра 123 – Комп'ютерна інженерія. – Тернопільський національний технічний університет імені Івана Пулюя, Тернопіль 2019.

У дипломній роботі досліджено існуючі методи та засоби автоматичного розпізнавання мови на базі мікроконтролерних систем, визначено їх переваги і недоліки та обґрунтовано актуальність розробки методів і засобів побудови таких систем.

Запропоновано використання нового методу виявлення ознак в мовленнєвому сигналі, у основі якого лежить вейвлет-аналіз попередньо обробленого сигналу. Визначено оптимальний метод автоматичного розпізнавання мови, який здійснюється за допомогою двох блоків: front-end та back-end процесорів. Front-end процесор займається зчитуванням звукового сигналу, його попередньою. фільтрацією, розбиттям на фрейми та накладанням вікна, визначенням початку та кінця мовлення та виділенням ознак за допомогою вейвлет-перетворення. Back-end процесор займається порівнянням представлення звукового сигналу у вигляді вектора ознак, з базою шаблонів відомих представлень мовленнєвих сигналів, використовуючи повнозв'язну штучну нейронну мережу.

Розроблено систему автоматичного розпізнавання мови на базі мікроконтролерної системи на базі плати розробки «STM32 Nucleo» з попереднім моделюванням процесу розпізнавання мови засобами програмного пакету Matlab.

ANNOTATION

Methods and tools of language automatic recognition based on microcontroller system // Master thesis // Chasnyk Vasyl Anatoliiovych // Ternopil Ivan Pul'uj National Technical University, Faculty of Computer Information Systems and software engineering, group CIm – 61 // Ternopil, 2019 // p. – 116, fig. – 23, tab. – 8, A1 Sheets – 10, Add. – 1, Ref. – 96.

Keywords: SPEECH RECOGNITION, WAVELET ANALYSIS, NEURAL NETWORK, MICROCONTROLER SYSTEM, STM32.

The diploma thesis for obtaining the Master's degree 123 – Computer engineering – Ternopil Ivan Pul'uj National Technical University, Ternopil 2019.

In the thesis, there were studied the current methods and means of automatic speech recognition on microcontroller systems. There were determined advantages and disadvantages of the mentioned above methods. There were justified the actuality of development of methods and means of such systems construction.

There was proposed the new method of detecting features in a speech signal, which is based on wavelet analysis of a pre-processed signal. The optimal method of automatic speech recognition was determined. This method is implemented using two blocks: front-end and back-end processors. The front-end processor reads the audio signal, filters, splits signal into frames, and overlays window, determines the beginning and end of speech, and extracts speech features using wavelet transform. The back-end processor compares the presentation of the audio signal as a feature vector with the database of patterns of known representations of speech signals, using a fully connected artificial neural network.

There was developed the system of automatic speech recognition on microcontroller system based on the development board «STM32 Nucleo» with preliminary modeling of speech recognition process.