

ВІДГУК

офіційного опонента

на дисертаційну роботу Белозьорової Яни Андріївни на тему «Метод застосування вейвлет аналізу в задачах ідентифікації мовної інформації», подану до захисту на здобуття наукового ступеня кандидата технічних наук за спеціальністю 01.05.03 – математичне та програмне забезпечення обчислювальних машин і систем

1. Актуальність дослідження.

У дисертаційній роботі Белозьорової Яни Андріївни розглядається задача розробки методів та алгоритмів, які дають можливість підвищити точність роботи систем ідентифікації мовної інформації.

Системи мовної ідентифікації особи та методи визначення мовних характеристик особи необхідні для розв'язання ряду задач, пов'язаних з ідентифікацією і верифікацією. Особливий інтерес представляють системи, що дозволяють проводити розпізнавання особи, використовуючи біометричні властивості мови. Це пов'язано з тим, що мова крім інформації, що передається мовним проголошенням, включає в себе також характеристики, що відображають індивідуальні особливості мовця.

Незважаючи на очевидні переваги використання мови, як біометричної ознаки, що визначає індивідуальні особливості особи, застосування таких програмних систем мовної ідентифікації особи на практиці може бути обмежена рядом факторів. До числа цих факторів відноситься висока трудомісткість, відсутність ефективного математичного інструменту для опису мовної інформації, неможливість описати таким чином описати мовний сигнал, щоб визначити його характеристики для однозначної ідентифікації

особи. Також слід зазначити, що такі властивості мовного сигналу, як висока варіативність, наявність спотворень з боку каналу передачі, неминуче призводять до зростання помилок розпізнавання мовної інформації.

Таким чином, представлені обставини необхідно враховувати при розробці та використанні систем ідентифікації та верифікації. Стає актуальною задача удосконалення алгоритмів та підвищення точності роботи систем ідентифікації мовної інформації.

Все це зумовлює актуальність теми дисертаційної роботи Белозьорової Я.А. проблематика якої відповідає напрямку 1.2.4.5. «Розроблення алгоритмів і програмно-апаратних засобів для систем комп'ютерного аналізу, розпізнавання та синтезу мовних і зорових образів» найважливіших проблем фундаментальних досліджень у галузі природничих, технічних, суспільних і гуманітарних наук Національної академії наук України на 2019–2023 роки, які затверджені постановою Президії НАНУ від 30.01.2019 р. № 30. Отримані в ході виконання роботи над дисертацією результати знайшли ефективне застосування на практиці для вирішення задач ідентифікації мовної інформації.

2. Достовірність і наукова новизна

За результатами аналізу роботи слід зазначити нові наукові результати, отримані автором, що представляють теоретичний і прикладний інтерес.

1) Уперше в якості набору ознак для мовної ідентифікації особи запропоновано самоподібні структури, що дозволило сформулювати новий спосіб представлення індивідуальних ознак та зменшити ймовірності виникнення помилок.

2) Уперше запропоновано метод підвищення інформативності основного тону для мовної ідентифікації особи в якому використано значення коефіцієнтів вейвлет-перетворення на відрізках мовного сигналу де спостерігаються екстремуми кореляційних функцій частоти основного тону.

3) Удосконалено знаходження границь вимовлених слів за рахунок використання структурної класифікації, що дає змогу виділення вокалізованих фрагментів та слів з мовного сигналу.

4) Одержали подальший розвиток методи ідентифікації особи, побудовані на частотно-часовому аналізі, що враховують індивідуальні характеристики особи.

Обґрунтованість і достовірність отриманих автором результатів роботи підтверджується наявністю актів їх впровадження в навчальний процес кафедри інженерії програмного при викладанні дисциплін «Мультимедійні, ігрові та розважальні системи», «Інтелектуальні системи» та «Людино-машинна взаємодія» зі спеціальності 121 «Інженерія програмного забезпечення», а також під час виконання двох НДР: «Онтології у інженерії програмного забезпечення» та «Методологія підвищення ефективності процесів життєвого циклу розробки програмного забезпечення у гнучких підходах його розробки» про що є відповідні акти. Основні положення дисертаційної роботи були опубліковані в журналах з переліку рецензованих наукових видань, і пройшли апробацію на ряді міжнародних конференцій.

3. Практична цінність результатів, отриманих автором дисертації

Основними результатами дисертаційної роботи є розробка методу сегментації, алгоритму виділення характеристик самоподібних структур та методу підвищення інформативності частоти основного тону і дозволяють їх використовувати для побудови систем ідентифікації мовного сигналу та створення інтелектуальних систем взаємодії користувача і комп'ютера. Розроблені в рамках дисертаційного дослідження методи та алгоритми впроваджені в виробничий процес ТОВ «ІСТ ТАУЕР» та ТОВ «КіберБіонік Систематікс», що підтверджено актами впровадження, наявними в додатках дисертаційної роботи.

4. Рекомендації щодо використання результатів

Отримані в дисертації результати можуть бути використані при розробці та впровадженні в експлуатацію систем ідентифікації мовної інформації, структурної розмітки мовних баз даних, для вирішення завдань, пов'язаних з ідентифікацією особи, в системах, пов'язаних з обробкою, аналізом і порівнянням мовної інформації.

5. Оцінка змісту дисертаційної роботи, її завершеність

Дисертаційна робота складається зі вступу, чотирьох розділів, висновків, переліку літературних джерел та додатків, і є завершеною науковою працею.

У вступі обґрунтовано актуальність теми дисертаційної роботи, сформульовано мету і виділено завдання дослідження, визначено об'єкт, предмет, методи дослідження, наукову новизну і практичне значення одержаних результатів, представлено загальну характеристику роботи, структуру та обсяг дисертації. Наведено відомості про впровадження результатів роботи, апробацію, особистий внесок автора, а також публікації за темою дисертації.

Перший розділ присвячено аналітичному огляду предметної області, проведено аналіз методів параметризації мовного сигналу, методів попередньої обробки мовного сигналу та його сегментації, розглянуто основні підходи до ідентифікації мовного сигналу, які застосовуються в існуючих системах розпізнавання як складових частинах інтелектуальних систем. На основі проведеного аналізу зроблено висновок про використання в якості основного інформаційного критерію ідентифікації особи частоти основного тону та формантних частот. Здійснено постановку задачі, сформульовано мету та завдання дослідження.

У другому розділі наведено модифіковану модель слухової системи людини; узагальнену послідовність математичних перетворень інформаційного каналу мовного спілкування, на основі якої виділено ознаки для сегментації мовного сигналу на вокалізовані фрагменти та паузи, та

принципи визначення характерних ознак мовного сигналу, що можуть бути використані при ідентифікації особи.

Автор наводить структурну схему узагальненої моделі слухової системи, яка складається з блоку попередньої обробки сигналу, блоків фрактального та спектрального аналізу та блоку сприйняття.

В цьому розділі автор розглядає задачу ідентифікації мовного сигналу особи в рамках контекстно-незалежної мови з необмеженим мовним словником і довільними характеристиками мови та обґрунтовує використання теорії фракталів для оцінки ступеню мінливості мовного сигналу на основі аналізу динаміки його поведінки.

В четвертому підрозділі побудовано модель представлення мовного сигналу на основі вейвлет аналізу та проаналізовано характеристики його самоподібних структур. На основі цієї моделі та узагальненої послідовності перетворень інформаційного каналу мовлення автором запропоновано підхід до виділення особливостей самоподібних структур.

У третьому розділі наведено принципи використання фрактального та вейвлет аналізу в задачі ідентифікації особи, розроблено метод сегментації мовного сигналу на вокалізовані фрагменти та паузи, а також алгоритм виділення характеристик самоподібних структур. Перший підрозділ цього розділу присвячено вибору методу розрахунку фрактальної розмірності. Автором здійснено порівняльний аналіз методів обчислення фрактальної розмірності та визначено, що для задач сегментації мовного сигналу на вокалізовані та невокалізовані фрагменти, найбільш придатним є метод обчислення фрактальної розмірності за Хаусдорфом.

Другий підрозділ присвячено розробленню методу сегментації на основі фрактальних характеристик сигналу, який полягає у розбитті сигналу на часові фрагменти, попередній його апроксимації; визначенні фрактальної розмірності з заданим часовим вікном в кожному часовому фрагменті з наступним розподілом часових фрагментів на вокалізовані та невокалізовані фрагменти згідно встановленого порогу величини фрактальної розмірності.

У третьому підрозділі створено та описано алгоритм виділення характеристик самоподібних структур в якості ідентифікаційних ознак мовного сигналу. Основою алгоритму є використання комплексного вейвлету Морле для опису мовного сигналу. На основі проведених досліджень автор пропонує метод підвищення інформативності частоти основного тону, що враховує характеристики самоподібних структур. Як підсумок, на основі розроблених методів та алгоритму, розроблено метод мовної ідентифікації особи, основними кроками якого є: сегментація досліджуваного мовного файлу на вокалізовані та невокалізовані фрагменти; виділення вокалізованих фрагментів мовного файлу; отримання з цих фрагментів характерних ознак особи; порівняння методом ранжування характерних ознак особи з наборами інших осіб; відображення найбільш близького набору ознак особи у вигляді ймовірнісної відповідності до набору ознак особи, що перевіряється.

Четвертий підрозділ присвячено визначенню частотних характеристик формант. Наведені основні методи визначення цих характеристик, та використання їх в основних формант-трекерах. В результаті автор пропонує власний алгоритм визначення формантних частот. В п'ятому підрозділі наведено порівняння характерних ознак мовного сигналу та запропоновано міру подібності статистичних характеристик мовних сигналів для прийняття рішення при ідентифікації. Автор показала, що використання в ролі статистичних характеристик густини ймовірності критеріїв ідентифікації мовного сигналу дає можливість виконати порівняння мовних сигналів.

Четвертий розділ присвячено експериментальному дослідженню інформативності алгоритмів опрацювання мовного сигналу, ознак ідентифікації особи та методів прийняття рішень, що завершується проектуванням системи ідентифікації особи за мовним сигналом та оцінюванням достовірності розпізнавання особи на основі запропонованого методу системи ідентифікації особи.

Дослідження розроблених алгоритмів та методів дали авторіві можливість встановити стійкі характеристики підвищення значення

модифікованої фрактальної розмірності для фрагментів мовних сигналів, що містять мову, показано, що фрактальна розмірність для пауз в 99 % є в межах 1,04..1,45, тоді як фрактальна розмірність фрагментів мови є не меншою за 1,55 для вікна розміром 20 мс. Автор провела оцінку ефективності алгоритму визначення формантих частот в порівнянні з існуючими формант-трекерами. Співмірна або краща точність була досягнута у випадку трекера DEEP, який, проте, був натренований з використанням набору даних, що використовувався в цій роботі. Як наслідок, проведений аналіз показав високу точність в поєднанні з низькою обчислювальною складністю запропонованого алгоритму. Порівняння точності мовної ідентифікації розробленої програмної системи та двох комерційних програмних систем (ТРАЛІ-М та ГОЛОС) показали високу ефективність розробленої системи, яка базується на моделях, методах та алгоритмах, розроблених в дисертаційній роботі.

Викладення матеріалу дисертації є достатньо послідовним та чітким, хоча подекуди бракує посилань на наступні розділи, що дещо утруднює сприйняття матеріалу. Наукові положення та висновки, сформульовані в дисертації є достатньо обґрунтованими та достовірними.

6. Зауваження до дисертаційної роботи

1) У формулюванні наукової новизни (стор. 15 дисертаційної роботи) та у Розділі 4 роботи сказано, що розроблений метод підвищення інформативності частоти основного тону дає можливість «досягти ймовірності виникнення похибок першого роду 2,12 % та ймовірності виникнення похибок другого роду 4,59 % при розпізнаванні особи за цією ознакою». Проте з роботи не зрозуміло, які були значення цих похибок без використання розробленого методу, що своєю чергою, не дає можливості оцінити ефект, який дає використання даного методу.

2) На стор. 33 у підсумку до підрозділу 1.4 сказано, що «точність методів визначення частоти ОТ можна вважати недостатньою». Проте в самому тексті підрозділу, незважаючи на детальний опис методів визначення частоти

основного тону, їх переваг та недоліків, нічого не сказано власне про *точність* її визначення.

3) На стор. 74 роботи автор вводить «мінімальний фрактальний масштаб $k \geq 3$ ». Не зрозуміло з якою метою і на підставі яких міркувань вводиться цей фрактальний масштаб, і чому його значення повинно бути не меншим саме за вказане значення.

4) На стор. 114 сказано, що «при визначенні порогового значення Δ_{OTlim} ... оцінювались параметри компонентів розробленого методу підвищення інформативності частоти ОТ та було обране порогове значення $\Delta_{OTlim} = 0.516$ », проте не наведено значень таких параметрів, критерію визначення порогового значення тощо, що не дає можливості оцінити достовірність визначеного порогового значення.

5) У дисертаційній роботі не розглядаються межі застосування запропонованого методу мовної ідентифікації особи, а також робота методу при різних рівнях шумів.

6) Не зважаючи на порівняння ефективності пропонувананих методів та алгоритмів з існуючими рішеннями, залишається не зрозумілим, наскільки запропонований методи і алгоритми будуть стійкі при низьких значеннях сигнал-шум.

7) В роботі не наведено параметрів аудіо файлів, що використовуються у дослідженні, наприклад при проведенні оцінки ефективності алгоритму визначення формантних частот.

8) В роботі наявні також деякі граматичні та стилістичні помилки, описки тощо. Наприклад: «являється» – «є», «доповнення» – «додатки», «зштовхається» – «виштовхується», «неволізовані» – «невокалізовані», «мовотворчій системи» – «мовотворчої системи», «об способі» – «про спосіб», «більш найкраща», «не можливо» – «неможливо», «не один з» – «жоден з», «в частоті» – «зокрема», «гласним» – «голосним», «дозвіл по частоті» – «роздільна здатність», «ударних голосних» – «наголошених», відсутній рисунок 3.4 тощо.

Наведені зауваження, проте, не впливають на загальну позитивну оцінку роботи в цілому.

7. Повнота викладення результатів в опублікованих матеріалах.

Результати дисертаційного дослідження апробовано належним чином. Основні результати дисертації опубліковано у двадцяти двох наукових роботах, тринадцять з яких у періодичних виданнях, з яких вісім у наукових фахових виданнях України, п'ять статей у закордонних виданнях, один патент України, а також вісім публікацій у збірниках тез та матеріалів доповідей наукових конференцій. Вісім опублікованих праць є одноосібними. Особистий внесок здобувача в сумісних публікаціях є підтвердженим.

В опублікованих працях викладено в повному обсязі основні отримані результати. Рівень та кількість публікацій, рівень апробації відповідають вимогам Міністерства освіти і науки України щодо оприлюднення наукових результатів дисертацій.

Автореферат відображає суть основних наукових положень, практичну значущість та висновки. Дисертаційна робота та автореферат оформлені у відповідності з вимогами, що ставляться до кандидатських дисертацій в Україні, хоча по тексту дисертації зустрічаються неточності, невідповідності у посиланні на рисунки, граматичні та стилістичні помилки.

Загальні висновки.

Дисертація Белозьорової Я.А. є завершеним науковим дослідженням, у якому отримано нові обґрунтовані результати, що вирішують важливу науково-прикладну задачу удосконалення алгоритмів та підвищення точності роботи систем ідентифікації мовної інформації. Результати роботи є суттєвими для розвитку таких напрямків галузі технічних наук як людино-машина взаємодія, методи ідентифікації особи, сегментація, обробка та опис характеристик мовного сигналу.

Основні результати дисертації відповідають вимогам паспорту наукової спеціальності 01.05.03 – математичне та програмне забезпечення обчислювальних машин і систем, зокрема наступним пунктам розділу II: мовні процесори; моделі програм і систем; засоби опису та аналізу предметних областей; інтелектуальні прикладні програмні системи.

Дисертант продемонстрував високу кваліфікацію, вільне володіння сучасними технологіями програмування та проектування програмного продукту. Автореферат і дисертація та їх оформлення, кількість публікації та повнота відображення результатів дисертаційних досліджень відповідають вимогам п.п. 9, 11, і 12 «Порядку присудження наукових ступенів» щодо кандидатських дисертацій.

Дисертаційна робота «Метод застосування вейвлет аналізу в задачах ідентифікації мовної інформації» відповідає вимогам Міністерства освіти і науки України до кандидатських дисертацій, а її автор Белозьорова Яна Андріївна заслуговує на присудження наукового ступеня кандидата технічних наук за спеціальністю 01.05.03 – математичне та програмне забезпечення обчислювальних машин і систем.

ОФІЦІЙНИЙ ОПОНЕНТ:

Доктор технічних наук, професор
професор кафедри систем штучного
інтелекту Національного університету
«Львівська політехніка»



Яковина Віталій Степанович

Підпис Яковини В.С. підтверджую

Вчений секретар
Національного університету
«Львівська Політехніка»



Брилинський Р.Б.