

Харківський національний університет радіоелектроніки

Работягов Андрій Валентинович

УДК 004.934 + 004.89 + 519.863

**МЕТОД ІДЕНТИФІКАЦІЇ ЛЮДИНИ НА ОСНОВІ
ІНДИВІДУАЛЬНОГО МОВНОГО КОДУ**

05.13.23 – системи та засоби штучного інтелекту

Автореферат дисертації на здобуття наукового ступеня
кандидата технічних наук

Харків – 2006

Дисертацією є рукопис.

Робота виконана

в Харківському національному університеті радіоелектроніки, Міністерство освіти і науки України.

Науковий керівник:

доктор технічних наук, професор Бондаренко Михайло Федорович, Харківський національний університет радіоелектроніки, заслужений діяч науки і техніки України, завідувач кафедри програмного забезпечення ЕОМ, ректор.

Офіційні опоненти:

- доктор технічних наук, старший науковий співробітник Єрохін Віктор Федорович, начальник управління в/ч А1906 (м. Київ);
- кандидат технічних наук, доцент Омельченко Анатолій Васильович, Харківський національний університет радіоелектроніки, доцент кафедри мереж зв'язку.

Провідна установа: Інститут кібернетики ім. В.М. Глушкова, відділ теорії цифрових математичних машин та систем, Національна академія наук України, м. Київ.

Захист відбудеться "14" червня 2006 р. о 13:00 годині на засіданні спеціалізованої вченої ради Д 64.052.01 в Харківському національному університеті радіоелектроніки за адресою: 61166, м. Харків, просп. Леніна, 14; тел. (0572) 702-14-46.

З дисертацією можна ознайомитись у бібліотеці Харківського національного університету радіоелектроніки за адресою: 61166, м. Харків, просп. Леніна, 14.

Автореферат розісланий "11" травня 2006 р.

Учений секретар
спеціалізованої вченої ради

С.Ф. Чалий

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність роботи. Серед багатьох актуальних проблем сучасної науки одне з найважливіших місць займає проблема автоматизованої обробки мовної інформації. Сучасний підхід до рішення задач в галузі мовних технологій і, зокрема, в галузі фоноскопичної ідентифікації людини характеризується в цілому як *спектральний*. Однак на сьогоднішній час відзначаються деякі недоліки такого підходу, в основі якого лежить *спектральна* теорія мовоутворення і мовосприйняття. По-перше, вузьким місцем в галузі прикладної лінгвістики є, насамперед, відсутність всебічної фонетичної теорії, що допомогла б зрозуміти мовні процеси. По-друге, основні труднощі для розпізнавання мови полягають у виділенні надійних ознак на акустико-фонетичному рівні. Ці недоліки гальмують подальший розвиток мовних технологій. Так, зокрема, у даний час відсутні прийнятні і коректні методи фоноскопичної ідентифікації людини на вкрай обмеженому за обсягом (за фразами тривалістю до 3 секунд) і спотвореному мовному матеріалі. Актуальність цієї проблеми зумовлена, зокрема, тим, що зросло число телефонних повідомлень про нібито закладену бомбу з хуліганських спонукань, що приводить до порушення роботи аеропортів, залізничних вокзалів, навчальних і адміністративних установ, підприємств України, до значного матеріального і морального збитку.

Для вирішення зазначеної проблеми необхідно проводити подальші дослідження лінгвістичної і акустико-фонетичної інформації, властивостей мовних сигналів і способів їхнього перетворення, розробляти нові методи фоноскопичної ідентифікації людини. Один з таких методів подається у даній дисертаційній роботі – метод ідентифікації людини на основі *індивідуального мовного коду* (ІМК), впровадження якого сприятиме роботі, зокрема, правоохоронних органів України.

Слід зазначити, що значний внесок у розвиток теоретичної і прикладної лінгвістики внесли такі зарубіжні і українські вчені, як: Л.В. Щерба, Г. Фант, Дж. Фланаган, Н. Хомський, Л.Р. Зіндер, Л.О. Чистович, Л.В. Бондарко, В.І. Галунов, М.А. Сапожков, Р.К. Потапова, Т.К. Вінцюк, М.Ф. Бондаренко, Ю.П. Шабанов-Кушнарєнко, О.Я. Дрюченко та ін.

Зв'язок роботи з науковими програмами, планами, темами. Робота виконувалася на кафедрі програмного забезпечення ЕОМ Харківського національного університету радіоелектроніки відповідно до плану науково-дослідної роботи

за темою № 159-2 "Ідентифікація мовника на обмеженому за обсягом мовному матеріалі" у рамках держбюджетної комплексної теми № 159 "Моделі механізмів інтелекту людини та їх застосування в інформаційних системах зі штучним інтелектом" (№ ДР 0103U001545), по якій автор дисертації є відповідальним виконавцем.

Мета і задачі дослідження. Метою роботи є розширення можливості проведення фоноскопичної ідентифікації людини скороченням обсягу і аналізом спотвореного мовного матеріалу за новим методом ідентифікації на основі індивідуального мовного коду.

Для досягнення мети в роботі ставляться такі задачі дослідження:

- 1) провести порівняльний аналіз сучасних методів фоноскопичної ідентифікації людини;
- 2) провести дослідження індивідуального мовного коду як системи інформативних диференційних ознак мовних сигналів, що визначають індивідуальні особливості мови людини;
- 3) розробити математико-логічну модель фоноскопичної ідентифікації людини;
- 4) провести експериментальну перевірку моделі фоноскопичної ідентифікації людини;
- 5) розробити метод ідентифікації людини на вкрай обмеженому за обсягом і спотвореному мовному матеріалі на основі індивідуального мовного коду.

Об'єктом дослідження є процес фоноскопичної ідентифікації людини.

Предметом дослідження є індивідуальний мовний код.

Методи дослідження. При вирішенні поставлених задач у роботі застосовуються гіпотетико-дедуктивний метод, методи штучного інтелекту, методи розпізнавання образів, емпіричний метод дослідження.

Наукова новизна отриманих результатів полягає в такому:

- 1) вперше одержано метод ідентифікації людини на основі нової системи інформативних ознак мовних сигналів – індивідуального мовного коду, в якому, на відміну від існуючих методів, для ототожнення мовних об'єктів на вкрай обмеженому за обсягом і спотвореному мовному матеріалі використовуються індивідуальні для кожного об'єкта функції (правила) обробки мовного сигналу, що дозволило повніше відобразити індивідуальні особливості мови людини;

2) удосконалено модель фоноскопичної ідентифікації людини, на відміну від існуючих моделей, введенням етапу класифікації мовних об'єктів і виділенням оптимальних ідентифікаційних ознак для кожного об'єкта, що дало змогу підвищити здатність ознак і вирішальних правил розділяти об'єкти на класи та рівень інтелектуалізації рішення задачі ідентифікації;

3) набув подальший розвиток структурний (геометричний) метод дослідження мовних сигналів, де первинні структурні ознаки елементарних сегментів голосних звуків мови найбільше відбивають закономірні особливості мови людини. Це дозволило застосувати опис елементарних площин під кривою сигналу, що найповніше відповідає механізму обробки сигналу слуховим аналізатором.

Практичне значення отриманих результатів дисертації полягає в тому, що вони є науковою базою для розробки нового методу ідентифікації людини на основі індивідуального мовного коду. У порівнянні з існуючими методами фоноскопичної ідентифікації практичне застосування розробленого методу дозволяє проводити ідентифікацію людини:

- 1) на вкрай обмеженому за обсягом мовному матеріалі (до 1 звуку мови);
- 2) на спотвореному мовному матеріалі (до мовного матеріалу не пред'являється вимога порівнянності, тобто мовні сигнали можуть бути отримані в різних умовах і різними пристроями запису; можуть бути досліджені мовні сигнали зі спотвореннями, що внесені зміною фізичного і емоційного стану людини).

Результатом застосування розробленого методу ідентифікації з'явилося:

- розширення області його застосування,
- надання діючої допомоги правоохоронним органам при встановленні особистості за короткими "тривожними" повідомленнями (наприклад, "Вокзал заміновано", "У школі закладено бомбу" та ін.), що надійшли на службу "02".

Апробація методу ідентифікації людини на основі індивідуального мовного коду проводилася в Харківському НДІ судових експертиз ім. Засл. проф. М.С. Бокаріуса (акт апробації від 2.06.2004 р.) і в ХМУ УМВС України в Харківській області (акт апробації від 31.03.2005 р.).

Результати дисертації впроваджені в науково-виробничій фірмі "Інтел-Медіа Сервіс" (м. Харків) для виконання теми "Пошук дикторів у мовній базі даних" (акт впровадження від 18.05.2005 р.).

Особистий внесок здобувача. Всі результати дисертації отримані автором самостійно. У роботах [1, 8] (співавтор М.Ф. Бондаренко) автором особисто пропонується новий метод фоноскопичної ідентифікації людини на основі ІМК; у роботі [2] надано модель ідентифікації людини за параметрами мови з використанням принципів нейронної сигналізації; у роботі [3] розглядається питання ідентифікації людини за параметрами мови на обмеженому за обсягом мовному матеріалі; у роботі [4] наведено рішення задачі ідентифікації людини за параметрами мови на основі принципів адаптації і оптимізації; у роботі [5] досліджено вплив емоційного стану людини на акустико-фонетичні ознаки; у роботі [6] (співавтор О.Я. Дрюченко) автором особисто запропоновані основні етапи методу фоноскопичної ідентифікації, який заснован на аналізі структурних характеристик елементарних сегментів голосних звуків мови; у роботі [7] розглянуто деякі методичні аспекти рішення задачі фоноскопичної ідентифікації людини.

Апробація результатів дисертації. Про результати дисертаційної роботи зроблено доповіді і вони обговорювалися: на 10-й Міжнародній конференції "Теорія і техніка передачі, прийому й обробки інформації" (Харків-Туапсе, Україна, 2004 р.), на 5-й Міжнародній науково-практичній конференції "Штучний інтелект – 2004. Інтелектуальні і багатопроекторні системи" (Крим, Україна, 2004 р.), на 7-й Всеукраїнській Міжнародній конференції "УКРОБРАЗ'2004" (Київ, Україна, 2004 р.), на 11-й Міжнародній конференції "Knowledge-Dialogue-Solution" (Варна, Болгарія, 2005 р.).

Публікації. Результати дисертації опубліковано у 8 роботах, з яких 4 – у виданнях, затверджених ВАК України.

Структура та обсяг дисертації. Дисертація складається зі вступу, чотирьох розділів і висновків, має загальний обсяг 128 сторінок основного тексту, містить 13 рисунків, з яких 6 виконані на окремих сторінках, 13 таблиць, список використаних джерел з 149 найменуваннями на 12 сторінках, 7 додатків на 22 сторінках.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** обґрунтовано актуальність теми дисертації, сформульовані мета і задачі дослідження, наведені відомості про зв'язок дослідження з планами організації, де виконана робота, про наукове і практичне значення та апробацію отриманих результатів дисертації.

У **першому розділі** досліджено передумови проведення роботи. Виділено, що основними причинами, які перешкоджають впровадженню мовних технологій у практику, є незадовільний стан спектральної теорії та відсутність надійних акустико-фонетичних ознак. Розглянуто сучасні методи фоноскопичної ідентифікації людини, питання, які на сьогоднішній день ще

невирішені. Визначено місце дисертації у вирішенні проблеми в галузі фоноскопичної ідентифікації людини.

У **другому розділі** розглянуто об'єкт і предмет дослідження. Наведені деякі теоретичні положення розробки методу ідентифікації людини на основі ІМК. Розглянуто деякі питання когнітивної лінгвістики. Викладено деякі дані з галузі нейрофізіології, які мають відношення до кодового принципу уявлення інформації, який реалізується, наприклад, у системі пам'яті людини.

Розкрито сутність ІМК, який покладено в основу методу ідентифікації:

- 1) ІМК – це система інформативних диференційних ознак, що являють собою інформаційну структуру, комплементарну властивостям мовного сигналу, і визначають індивідуальні особливості мови людини;
- 2) ІМК – це код, на основі якого проводиться класифікація мовних еталонних об'єктів (ЕО) і здійснюється фоноскопична ідентифікація людини;
- 3) ІМК – це код, цифрова десяткова кодова комбінація якого привласнюється ЕО в процесі виконання процедур оптимізації та селекції;
- 4) ІМК – це код, у якому кожний розряд цифрової десятикової кодової комбінації являє собою *показник* типу функцій (правил) обробки мовного сигналу і умов їхнього виконання. Типи функцій і умови їх виконання реалізовано у виді окремих *підпрограм*, основу яких складають прості математичні рівняння та логічні функції (предикати) математичної логіки.

Новизна ІМК полягає в тому, що для ототожнення мовних об'єктів як ідентифікаційні ознаки використовуються індивідуальні для кожного об'єкта *функції (правила)* обробки мовного сигналу, що ставляться у відповідність до об'єкта в процесі пошуку, тобто ідентифікація виконується не за первинними ознаками, як це прийнято в сучасних методах, а за індивідуальними для кожного об'єкта наборами *оптимальних функцій (правил)*.

Викладено суть ефекту згладжування у слусі. Розглянуто експериментальний метод дослідження мовних сигналів, яким є *графічний* метод. Акцентовано, що графічний метод широко використовував видатний російський вчений-лінгвіст Л.В. Щерба. Обґрунтовано вибір напрямку дослідження, наведені методи та загальна методика проведення дослідження.

У **третьому розділі** представлено модель фоноскопичної ідентифікації людини, яка складає математико-логічну базу методу ідентифікації на основі ІМК. Поставлено таку задачу: в умовах невизначеності, коли ідентифікаційні ознаки мовних сигналів відомі *a priori* не точно, а лише в деяких межах, розробити модель, яка б могла знайти оптимальні ідентифікаційні ознаки (ІМК), що забезпечують класифікацію множини ЕО на кінцеве число класів і проведення ідентифікації.

Основою формалізації моделювання є теорія розпізнавання образів.

Об'єкт називається *еталонним* унаслідок того, що досліднику і, відповідно, моделі заздалегідь відома приналежність об'єкта до класу ЕО. Кожний ЕО являє собою звуковий файл спеціального формату, у якому міститься один голосний звук мови людини.

Модель ідентифікації складається з двох етапів: 1) етапу "Класифікація ЕО", 2) етапу "Ідентифікація невідомого об'єкта (НО) з ЕО" (рис.1).

Метою етапу "Класифікація ЕО" є визначення оптимальних ідентифікаційних ознак (ІМК), що забезпечують класифікацію множини ЕО на кінцеве число класів. Поставлена мета досягається таким чином:

- 1) складається таблиця (матриця) кодових комбінацій ІМК при використанні коду на всі можливі сполучення, з якої на етапі "Класифікація ЕО" відбирається одна кодова комбінація ІМК;
- 2) на основі однієї обраної кодової комбінації ІМК здійснюється опис і класифікація всіх ЕО;
- 3) на основі результатів класифікації ЕО з усіх кодових комбінацій ІМК відбирається оптимальна кодова комбінація ІМК для кожного ЕО.

Матриця ІМК являє собою таблицю розрядів і можливих цифрових значень розрядів кодових комбінацій, з якої відбирається одна кодова комбінація ІМК (табл.1). Відповідно до класифікації, яка прийнята в теорії кодування, ІМК характеризується як 13-розрядний рівномірний код; як значення розрядів коду використовуються десяткові цифри. Розряди кодової комбінації ІМК мають умовну позначку:

$$FG_r(x),$$

де індекс r – номер розряду, x – цифрове значення розряду. Наприклад, $FG_2(3)$ позначає: "значення 2-го розряду дорівнює десятковому числу 3".

Рис.1. Структурно-функціональна схема моделі фоноскопичної ідентифікації

Кодові комбінації ІМК при використанні коду на всі можливі сполучення утворюються таким чином: у першій колонці таблиці записуються номери всіх можливих комбінацій, починаючи з другої колонки – самі кодові комбінації. При цьому отримані кодові комбінації відрізняються одна від одної не менше ніж одним цифровим значенням у будь-якому розряді кодової комбінації (табл.2).

Повне число кодових комбінацій N визначається за формулою:

$$N = \prod_{r=0}^i \left(\sum_{j=0}^m FG_r(x_j) \right),$$

де i – номер останнього розряду, m – номер останнього значення розряду, вираз в дужках – основа r -го розряду.

На етапі "Опис ЕО" формуються *образи* усіх ЕО. Опис кожного ЕО виконується за однотипним алгоритмом в декілька етапів (у просторах V_r, V_c, V_m, V_d, V_w), у яких реалізуються функції розрядів ІМК.

Варто звернути увагу, що в моделі ідентифікації застосовуються дві системи ознак: першу систему ознак складає ІМК (FG_r), другу – структурні ознаки (vr_0, vr_1, \dots, vr_n) простору V_r , над якими виконуються математико-логічні операції (функції) у просторах V_c, V_m, V_d, V_w (табл.3).

Виконання функції залежить від розряду кодової комбінації ІМК, який являє собою *показник* типу функцій обробки мовного сигналу:

$$O_i^k(FG_0^j, FG_1^j, \dots, FG_n^j) \Rightarrow O_i^k(vr_0, vr_1, \dots, vr_n),$$

де O – умовна позначка об'єкта, i – номер ЕО, k – номер класу ЕО.

Умовна позначка розряду кодової комбінації ІМК, що виконує визначену функцію, має такий вид:

$$FG_r(x) = p,$$

де індекс r – номер розряду, x – цифрове значення розряду, p – функція, яку виконує розряд. Наприклад, 1-й розряд ІМК виконує функцію $FG_1(x) =$ "кількість вертикальних ліній розбивки v "; значення розряду вказує кількість ліній розбивки v : $FG_1(1) = "1"$ кількість ліній розбивки v дорівнює 1 (табл.3).

У просторі V_r об'єктами первинного опису є елементарні сегменти голосних звуків мови (ЕСГЗМ). ЕСГЗМ – це акустична мікроподія, що є відгуком резонатора мовоутворення на один імпульс збудження від голосових зв'язок (рис.2). Сегментація голосних звуків на ЕСГЗМ виконується вручну.

Рис.2. Елементарні сегменти (ЕС) голосного звуку o

Голосний звук *о* представлено 12-тьма ЕС. Ось абсцис (Т) – шкала часу (розмірність одиниць виміру – мсек). Ось ординат (А) – шкала амплітуд цифрового мовного сигналу. Тривалість звуку (L) дорівнює 84 мсек. (а) - "нульова" лінія

У моделі фоноскопичної ідентифікації ЕСГЗМ представлено у виді геометричного об'єкта (рис.3) відповідно до структурного (геометричного) методу дослідження мовних сигналів.

Рис.3. Первинний опис елементарного сегменту голосного звуку мови

Тривалість ЕСГЗМ (L) дорівнює 6 мсек

Для первинного опису об'єкта використовуються геометричні елементи аналітичної геометрії: пряма (хорда) і складна плоска фігура. Формування первинного образу об'єкта у виді вектора значень первинних структурних ознак зводиться до виміру значень довжини хорди і площі складної плоскої фігури (рис.3).

Структура ЕСГЗМ складається з 6-ти конструктивних частин: А, В, С, D, Е і F (рис.3). Кожна з частин "розбивається" на декілька структурних вікон. Як "розділюючий" елемент використовуються горизонтальні *g* і вертикальні *v* лінії розбивки. У кожному вікні виділяються два геометричних структурних елементи: складна фігура і пряма горизонтальна лінія (хорда), утворені габаритами вікна і формою ЕСГЗМ, і вимірюються значення довжини хорди і площі складної плоскої фігури. Наприклад, у конструктивній частині А виділені чотири вікна: *w1*, *w2*, *w3* і *w4*; у вікні *w1* визначаються значення площі S_l складної фігури і довжини l_l хорди. Обчислення довжин і площ проводиться за такими формулами: довжина l хорди визначається як число семплів $n(sp)$, що утворюють хорду; площа S фігури визначається як добуток суми значень семплів $[sp_i]$, що утворюють складну фігуру, і періоду дискретизації T_d :

$$l = n(sp), \quad S = T_d \cdot \sum_{i=1}^l [sp_i].$$

У просторі V_r кількість і розташування ліній розбивки *g* і *v* безпосередньо впливає на габарити вікна і, відповідно, на значення первинних структурних ознак. "Керують" завданням числа горизонтальних *g* і вертикальних *v* ліній розбивки розряди ІМК: FG_0 і FG_1 . "Керування" здійснюється тим, що значення

розрядів ІМК "включають" відповідні підпрограми моделі фоноскопичної ідентифікації (табл.3). Для всіх конструктивних частин ЕСГЗМ положення h_g ліній розбивки g визначається за такою формулою:

де h_i – висота i -тої конструктивної частини ЕСГЗМ, x_0 – цифрове значення розряду FG_0 . Положення l_v ліній розбивки v визначається положенням точки завдання висоти i -тої конструктивної частини ЕСГЗМ.

Значення "нульової" лінії (рис.2-3) визначається автоматично значенням рівня, при якому число переходів мовного сигналу через "0" на "шумній" ділянці досягає максимального значення.

У залежності від FG_2 первинний образ об'єкта формується з урахуванням типу структурних ознак. Наприклад, для $FG_2(1)$ образ формується тільки із значень площ S складних плоских фігур структурних вікон (табл.3).

У просторі V_r модель формує з урахуванням значень розрядів ІМК первинний опис ЕСГЗМ у виді вектора в n -мірному просторі ознак:

Для кожної з конструктивних частин ЕСГЗМ складається типовий опис. Оскільки кожний ЕО представлено одним голосним звуком, який утворено ланцюжком ЕСГЗМ, то опис у цілому ЕО зводиться до опису усіх ЕСГЗМ ланцюжка. У результаті первинного опису формується квадратна матриця, у відповідних рядках якої розташовані вектори ЕСГЗМ, у стовпцях – координати векторів. З урахуванням значень FG_0 - FG_2 розмірність вектора кожного ЕСГЗМ лежить у межах 12–144.

На наступних етапах моделі фоноскопичної ідентифікації виконують математико-логічні операції над первинними ознаками ЕО в залежності від значень розрядів ІМК (табл.3).

У кожному векторі простору V_c в залежності від FG_3 виключаються ознаки конструктивної частини ЕСГЗМ. Наприклад, для $FG_3(5)$ виключення ознак здійснюється тільки для конструктивної частини Е ЕСГЗМ. У результаті таких перетворень розмірність векторів V_c скорочується і дорівнює 10-144.

У просторі V_m в залежності від FG_4 - FG_5 виконується віднімання (ділення) мінімального значення координати стовпця (рядка) матриці від (на) усіх однотипних координат, розташованих у стовпці (рядку) матриці простору V_c :

де n – номер стовпця матриці, m – номер рядка матриці; $\min(vc)$ – мінімальне значення координати вектора Vc .

У просторі Vd в залежності від FG_6 - FG_8 виконується нормування векторів простору Vm :

де n – номер стовпця матриці, m – номер рядка матриці; $k(l)$ і $k(h)$ – коефіцієнти нормування, значення яких залежать від FG_7 - FG_8 ; $\max(l)$ – тривалість ЕСГЗМ, $\min(l)$ – \min із значень довжин конструктивних частин ЕСГЗМ; $\max(h)/\min(h)$ – \max/\min із значень висот конструктивних частин ЕСГЗМ.

У просторі Vw в залежності від FG_9 в кожному векторі простору Vd віднімається мінімальне значення координати відповідно до приналежності до конструктивної частини ЕСГЗМ і типу структурних ознак ("S", "I"); значення координат ранжируються:

де n^* – номер стовпця матриці відповідно до приналежності до конструктивної частини ЕС і типу ознак; m – номер рядка матриці.

У моделі фоноскопичної ідентифікації класифікація ЕО на основі вирішальних правил виконується в просторах $Z1$, $Z2$ і $Z3$. Класифікація здійснюється попарним порівнянням усіх ЕО за правилом "кожний з кожним" на основі обчисленої в метричному просторі відстані $Z3$ і значень мір близькостей ε^1 , ε^2 , ε^3 , які залежать від FG_{10} - FG_{12} , за таким алгоритмом:

1) кожний вектор ЕО більшої потужності (v) по черзі порівнюється з усіма векторами ЕО меншої потужності (w). У процесі попарного порівняння ЕО порогово співставляється кожна пара координат: якщо відстань між координатами γ^1 менше значення міри близькості ε^1 , то результату порівняння присвоюється логічне значення "1", якщо більше – "0". У результаті одного порівняння формується вектор $z1$, значення координат якого належать бінарній множині $\{0, 1\}$. У результаті порівняння одного вектора ЕО більшої потужності з усіма векторами ЕО меншої потужності знаходиться вектор $z1_{\max}$ (вектор $z1$ с \max числом "1"); у результаті порівняння всіх векторів формується множина векторів $z1_{\max}$:

2) На основі числа координат векторів $\max\{n(1)\}$, рівних "1", розмірності $n(Vw)$ і числа векторів $m(Z1)$ формується один вектор $z2$, значення координат якого належать бінарній множині $\{0, 1\}$. Значення координат вектора $z2$ визначаються в результаті порогового співставлення числа координат векторів $\max\{n(1)\}$, що дорівнюють "1", і розмірності $n(Vw)$ вектора (v): якщо відстань γ^2 між ними менше значення міри близькості ε^2 , то результату

порівняння (координаті вектора $\mathbf{z2}$) присвоюється логічне значення "1", якщо більше – "0". Розмірність вектора $\mathbf{z2}$ дорівнює числу векторів $m(\mathbf{Z1})$:

$$\gamma_i^2 = \frac{n(\bar{Vw}) - \max\{n(1)\}_i}{m(\mathbf{Z1})} \cdot 100, \quad z_{2i} = \begin{cases} 1, & \text{якщо } \gamma_i^2 \leq \varepsilon^2, \\ 0, & \text{якщо } \gamma_i^2 > \varepsilon^2, \end{cases} \quad \varepsilon^2 = f(FG_{11}),$$

$$\bar{z2} = (z_{21}, z_{22}, \dots, z_{2m}), \quad m = m(\mathbf{Z1}), \quad \mathbf{Z2} = \{\bar{z2}\}.$$

3) Відстань $\mathbf{z3}$ між двома ЕО, які порівнюють, обчислюється в результаті порогового співставлення суми координат вектора $\mathbf{n(1)}$, рівних "1", і розмірності вектора $\mathbf{m(v)}$ більшої потужності: якщо відстань γ^3 між ними менше значення міри близькості ε^3 , то результату порівняння (відстані $\mathbf{z3}$) присвоюється логічне значення "1", якщо більше – "0":

$$n(1) = \sum_{i=1}^m z_{2i}(1), \quad \gamma^3 = \frac{m(v) - n(1)}{m(v)} \cdot 100, \quad z_3 = \begin{cases} 1, & \text{якщо } \gamma^3 \leq \varepsilon^3, \\ 0, & \text{якщо } \gamma^3 > \varepsilon^3, \end{cases} \quad \varepsilon^3 = f(FG_{12}),$$

$$\bar{z3} = (z_3), \quad \mathbf{Z3} = \{\bar{z3}\}.$$

4) Два ЕО, які порівнюють, вважаються *тотожними*, тобто приналежними до одного класу ЕО, якщо відстань $\mathbf{z3}$ між ними дорівнює "1", якщо "0" – *нетотожними*:

$$z_3 = \begin{cases} 1, & \text{об'єкти тотожні,} \\ 0, & \text{об'єкти нетотожні.} \end{cases}$$

Метою етапу "Селекція" є відбір *селективних* кодових комбінацій – кодових комбінацій ІМК, при яких відсутні помилки класифікації: помилки І-го і ІІ-го роду ("пропуск цілі" і "захоплення помилкової цілі"). Якщо в результаті класифікації ЕО встановлені тільки лише тотожності між ЕО, то модель ідентифікації записує результати попарних порівнянь ЕО.

Метою етапу "Оптимізація" є відбір *оптимальних* кодових комбінацій ІМК із множини *селективних* кодових комбінацій і їх присвоєння ЕО. *Оптимальними* вважаються такі кодові комбінації, при яких значення критерію оптимізації Q_{opt} для кожної пари (i) ЕО є мінімальним.

$$Q_{opt} = \frac{\sum_{i=1}^{m(v)} (\max\{n(1)\}_i \cdot z_{2i})}{m(v) \cdot n(Vw)} \cdot 100, \quad opt(Q_{opt}(i)) = \min\{Q_{opt}(i)\}, \quad opt(FG_i) = FG_i.$$

У моделі фоноскопичної ідентифікації передбачені різні варіанти рішення задачі класифікації ЕО. Пошук оптимальних рішень виконується в автоматичному режимі послідовною зміною значень кодових комбінацій ІМК і аналізу поточного стану моделі на основі критеріїв селекції та оптимізації. На етапі "Класифікація ЕО" може бути виконано $N = 104.186.250$ ітерацій. Після виконання усіх ітерацій у результаті пошуку визначається остаточний вигляд *оптимальної* кодової комбінації ІМК для кожного ЕО, яка складає *індивідуальний мовний код* ЕО. ІМК привласнюється кожному ЕО окремо. Таким чином, кожний ЕО характеризується своїм *індивідуальним мовним кодом*, у якому в "стислому" виді відбита специфічна *індивідуальна* мовна інформація:

$$O_i^k(FG_0(x), FG_1(x), \dots, FG_{12}(x)) \quad \text{або} \quad O_i^k(x_0, x_1, \dots, x_{12}),$$

де O – умовна позначка ЕО, i – номер ЕО, k – номер класу ЕО, x – цифрове значення розряду ІМК. Наприклад, $O_{29}^k(1, 0, 2, 3, 0, 1, 2, 5, 4, 0, 0, 3)$.

У моделі ідентифікація виконується за правилом "один з усіма", згідно з яким НО (позначка: $O^?$) по черзі порівнюється з кожним ЕО. Процес ідентифікації складається з трьох етапів: 1) "Опис ЕО", 2) "Опис НО", 3) "Порівняння НО з ЕО". Механізми опису і порівняння об'єктів етапу "Ідентифікація" аналогічні механізмам етапу "Класифікація ЕО" з тією лише різницею, що опис, порівняння і встановлення тотожності між НО і ЕО здійснюється на основі функцій (правил), які виконуються в залежності від значень розрядів кодової комбінації ІМК ЕО:

$$O_i^k(FG_0(x), FG_1(x), \dots, FG_{12}(x)) \Rightarrow O^?(FG_0(x), FG_1(x), \dots, FG_{12}(x)).$$

Прийняття рішення про тотожність (нетотожність) двох об'єктів, НО і ЕО, проводиться на підставі такого правила (табл.4):

$$z_i(O^?, O_i^{k(i)}) = Z3_i = \begin{cases} 1, & \text{об'єкти } O_i^k \text{ і } O^? \text{ тотожні,} \\ 0, & \text{об'єкти } O_i^k \text{ і } O^? \text{ нетотожні.} \end{cases}$$

У **четвертому розділі** виконано експериментальну перевірку моделі фоноскопичної ідентифікації людини. Для перевірки моделі застосовуються *перевірочні* об'єкти, що відбираються з банку ЕО. Так усувається фактор непередбачуваності і випадковості, оскільки номер класу перевірконого ("невідомого") об'єкта заздалегідь вірогідно відомий дослідникові.

Варто звернути увагу, що ідентифікація проводилася на варіативному, на вкрай обмеженому за обсягом мовному матеріалі: мінімальний обсяг мовної інформації складав один голосний звук мови (наприклад, звук *о* у фонетичному оточенні *вк* у слові *вокзал* фрази *вокзал замінено*); досліджувався спотворений мовний матеріал (деякі зразки мови вимовлені у навмисно спотвореній манері; мовний матеріал записаний з різних передатних каналів і в різних умовах).

Розрахована потенційна інформативна здатність моделі фоноскопичної ідентифікації. Досліджена інформативна здатність ІМК. Розроблена методика застосування методу фоноскопичної ідентифікації людини на основі ІМК. Розрахована імовірність ідентифікації людини за допомогою розробленого методу, що складає 80-85 %.

Зазначено, що за однаковими показниками надійності ідентифікації обсяг мовного матеріалу для проведення ідентифікації розробленим методом у 10 разів менше обсягу, що використовується існуючими методами.

Визначені обмеження методу: сегментація голосних звуків виконується вручну; великі витрати часу на проведення в цілому ідентифікації.

У **висновках** надано основні положення, які виносяться на захист.

У додатках містяться опис та структура банку ЕО, результати експериментальної перевірки моделі фоноскопичної ідентифікації, акти апробації методу ідентифікації та впровадження результатів дисертації.

ВИСНОВКИ

У дисертації відбито нове рішення наукової задачі, що складається в розробці нового методу фоноскопичної ідентифікації людини на основі індивідуального мовного коду. У порівнянні з існуючими методами ідентифікації застосування розробленого методу дає можливість проводити ідентифікацію людини на вкрай обмеженому за обсягом і спотвореному мовному матеріалі, і на множині об'єктів одночасно. Особливості методу забезпечують розширення області його застосування.

1) Проведено порівняльний аналіз сучасних методів фоноскопичної ідентифікації людини. Розглянуті питання, які на сьогоднішній день ще не вирішені. Обґрунтовано необхідність пошуку і дослідження нових ідентифікаційних ознак мовних сигналів, що визначають індивідуальні особливості мови людини, розробки нового методу фоноскопичної ідентифікації людини.

2) Вперше розроблено метод ідентифікації людини на основі нової системи інформативних диференційних ознак мовних сигналів – індивідуального мовного коду. Новизна індивідуального мовного коду, на відміну від існуючих систем ознак, полягає в тому, що для ототожнення мовних об'єктів використовуються індивідуальні для кожного об'єкта функції (правила) обробки мовного сигналу, що дозволяє повніше відобразити індивідуальні особливості мови людини. Теоретичною базою нової системи ознак є: поняття комплементарності і комплементарних пар у мові; той факт, що обробка мовних сигналів не може проводитися за однорідною схемою. Під індивідуальним мовним кодом розуміємо інформаційну структуру, комплементарну властивостям мовного сигналу, які визначають індивідуальні особливості мови людини. Новизною розробленого методу в порівнянні з існуючими методами є те, що ідентифікація людини здійснюється на вкрай обмеженому за обсягом і спотвореному мовному матеріалі. Проведено емпіричне дослідження індивідуальних особливостей мови людини і розробку комплементарних до них способів аналізу.

Розроблена методика застосування методу. Розраховано імовірність ідентифікації людини.

3) Запропоновано удосконалену модель фоноскопичної ідентифікації людини. По-перше, введено етап класифікації мовних об'єктів. Особливість моделі полягає в тому, що перед здійсненням ідентифікації, на відміну від існуючих моделей, проводиться класифікація об'єктів: на етапі класифікації формується система диференційних ознак об'єктів і вирішальних правил, на основі яких проводиться ідентифікація. Цим підвищується здатність ознак і вирішальних правил розділяти об'єкти на класи. По-друге, розроблена модель, на відміну від існуючих моделей, виділяє оптимальні ознаки для кожного об'єкта, тим самим підвищується рівень інтелектуалізації рішення задачі ідентифікації. Виконана експериментальна перевірка моделі.

4) Набув подальший розвиток структурний (геометричний) метод в галузі прикладної лінгвістики. По-перше, для здійснення ідентифікації в мовному сигналі виділяють релевантні ділянки, елементарні сегменти голосних звуків мови, що є вкрай мінімальними функціональними складовими одиницями голосних звуків, об'єктивними і закономірними акустичними мікроподіями. По-друге, як первинні інформативні ознаки прийняті структурні ознаки елементарних сегментів голосних звуків, які якнайбільше відбивають закономірні індивідуальні особливості мови людини. Вибір структурних ознак спирається на ефект згладжування у слусі (закон Тальбота, сформульований спочатку для зору). Це дозволяє відмовитися від методів спектрального опису мовних сигналів і застосувати методи опису елементарних площин під кривою сигналу, що найповніше відповідає механізму їх обробки слуховим аналізатором.

5) Апробація методу ідентифікації людини проводилася в Харківському НДІ судових експертиз ім. Засл. проф. М.С. Бокаріуса і в ХМУ УМВС України в Харківській області. Результати дисертації впроваджені в науково-виробничій фірмі "Інтел-Медіа Сервіс" (м. Харків).

СПИСОК ОПУБЛІКОВАНИХ РОБІТ

1. *Бондаренко М.Ф., Работягов А.В.* Метод идентификации человека на основе индивидуального речевого кода // Искусственный интеллект. – 2004. – № 3. – С. 617-625.

2. *Работягов А.В.* Моделирование системы идентификации человека по параметрам речи с использованием принципов нейронной сигнализации // Проблемы бионики. – Харьков: ХНУРЭ, 2002. – Вып. 56. – С. 60-64.
3. *Работягов А.В.* К вопросу об идентификации человека по параметрам речи на ограниченном объеме речевого материала // Проблемы бионики. – Харьков: ХНУРЭ, 2002. – Вып. 57. – С. 7-12.
4. *Работягов А.В.* Решение задачи идентификации человека по параметрам речи на основе принципов адаптации и оптимизации // Автоматизированные системы управления и приборы автоматики. – Харьков: ХНУРЭ, 2002. – Вып. 121. – С. 80-87.
5. *Работягов А.В.* Устная речь как отражение эмоционального состояния человека // Вісник Луганського інституту внутрішніх справ МВС України. – Луганськ: ЛІВС, 2000. – Вип. 2. – С. 164-172.
6. *Дрюченко А.Я., Работягов А.В.* Метод судебной фоноскопической идентификационной экспертизы, основанный на анализе структурных характеристик элементарных сегментов гласных звуков речи // Теорія та практика судової експертизи і криміналістики. – Харків: Право, 2001. – С. 197-205.
7. *Работягов А.В.* Некоторые методические аспекты проведения судебной фоноскопической экспертизы // Теорія та практика судової експертизи і криміналістики. – Харків: Право, 2004. – Вип. 4. – С. 219-227.
8. *Бондаренко М.Ф., Работягов А.В.* Об одном эвристическом методе идентификации говорящего // Тр. 10-й Междунар. научн. конф. "Теория и техника передачи, приема и обработки информации". – Харьков-Туапсе: ХНУРЭ, 2004. – С. 251-252.

АНОТАЦІЯ

Работягов А.В. Метод ідентифікації людини на основі індивідуального мовного коду. – Рукопис.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.23 – системи та засоби штучного інтелекту. – Харківський національний університет радіоелектроніки, Харків, 2006.

У роботі одержано новий метод фоноскопичної ідентифікації людини на основі індивідуального мовного коду (ІМК). Ідентифікація за новим методом здійснюється на вкрай обмеженому за обсягом (до 1 звуку мови) і спотвореному мовному матеріалі. Новизна ІМК полягає в тому, що для ототожнення мовних об'єктів як диференційні ідентифікаційні ознаки використовуються індивідуальні для кожного об'єкта

функції обробки мовного сигналу, які ставляться у відповідність об'єкту в процесі селекції і оптимізації. Удосконалено модель фоноскопичної ідентифікації людини введенням етапу класифікації мовних об'єктів і виділенням оптимальних ознак для кожного об'єкта. Розвинуто структурний (геометричний) метод дослідження мовних сигналів, що дозволило відмовитися від методів спектрального опису і застосувати методи опису елементарних площин під кривою сигналу.

Ключові слова: фоноскопична ідентифікація, індивідуальний мовний код, структурний метод дослідження мовних сигналів, селекція, оптимізація.

АННОТАЦИЯ

Работягов А.В. Метод идентификации человека на основе индивидуального речевого кода. – Рукопись.

Диссертация на соискание ученой степени кандидата технических наук по специальности 05.13.23 – системы и средства искусственного интеллекта. – Харьковский национальный университет радиоэлектроники, Харьков, 2006.

Диссертация посвящена решению одной из приоритетных задач прикладной лингвистики – фоноскопической идентификации человека на предельно ограниченном по объему и искаженном речевом материале.

В диссертации проведен анализ современного состояния в области теоретической и прикладной лингвистики, современных методов фоноскопической идентификации человека. Рассмотрены предпосылки проведения работы, вопросы, которые на сегодняшний день являются нерешенными. Определено место диссертационного исследования в решении проблемы в области фоноскопической идентификации человека.

Даны определения объекту и предмету исследования. Объяснен механизм эффекта сглаживания в слухе. Изложены основные положения экспериментального метода исследования речевых сигналов. Разработана общая методика проведения диссертационного исследования.

В работе впервые разработан метод фоноскопической идентификации человека на основе новой системы информативных дифференциальных признаков речевых сигналов – индивидуального речевого кода (ИРК). Теоретической основой новой системы признаков, ИРК, выступают: понятия комплементарности и комплементарных пар в речи; тот факт, что обработка речевых сигналов не может опираться на однородную схему. В работе проводится эмпирическое исследование коммуникативных свойств речевых сигналов и разработка на этой основе комплементарных к ним способов анализа. Под ИРК понимается информационная структура,

комплементарная свойствам речевого сигнала, которые определяют индивидуальные особенности речи человека.

Новизна ИРК состоит в том, что для отождествления речевых объектов в качестве дифференциальных идентификационных признаков используются индивидуальные для каждого речевого объекта функции (правила) обработки речевого сигнала, которые ставятся в соответствие речевому объекту в процессе селекции и оптимизации, т.е. идентификация выполняется не по первичным признакам, как это принято в существующих на сегодняшний день методах идентификации, а по индивидуальным для каждого объекта наборам оптимальных функций (правил).

Применение метода идентификации дает возможность осуществлять идентификационные исследования на предельно ограниченном по объему (до 1 звука речи) и искаженном речевом материале, и на множестве объектов одновременно. Это позволяет расширить область применения метода.

Усовершенствована модель фоноскопической идентификации человека. Во-первых, введен этап классификации речевых объектов. Эта особенность модели заключается в том, что перед осуществлением идентификации проводится классификация речевых объектов: на этапе классификации формируется система дифференциальных признаков объектов и решающих правил, на основе которых проводится идентификация. Этим повышается способность признаков и решающих правил разделять объекты на классы. Во-вторых, модель выделяет оптимальные признаки для каждого речевого объекта, что дает возможность **повысить уровень интеллектуализации решения задачи идентификации**. Проведена экспериментальная проверка модели идентификации. Рассчитана вероятность идентификации человека. Исследована информативная способность ИРК.

Получил дальнейшее развитие структурный (геометрический) метод в области прикладной лингвистики: для осуществления идентификации в речевом сигнале выделяются особые релевантные участки, элементарные сегменты гласных звуков речи (ЭСГЗР), которые являются предельно минимальными функциональными составными единицами гласных звуков речи, объективными и закономерными акустическими микрособытиями; в качестве первичных информативных признаков приняты структурные признаки ЭСГЗР, которые в наибольшей мере отражают закономерные индивидуальные особенности речи человека. Выбор структурных признаков опирается на эффект сглаживания в слухе (закон Тальбота, сформулированный первоначально для зрения). Это позволяет отказаться от методов спектрального описания речевых сигналов и применить методы описания элементарных площадей под кривой сигнала, наиболее полно соответствующих механизму их обработки слуховым анализатором.

Апробация метода фоноскопической идентификации человека проводилась в Харьковском НИИ судебных экспертиз им. Засл. проф. Н.С. Бокариуса и в ХГУ УМВД Украины в Харьковской области. Результаты диссертации внедрены в научно-производственной фирме "Интел-Медиа Сервис".

Ключевые слова: фоноскопическая идентификация, индивидуальный речевой код, структурный метод исследования речевых сигналов, селекция, оптимизация.

ABSTRACT

Rabotyagov A.V. A Method of Identification of a Person on the Basis of the Individual Speech Code. – Manuscript.

Thesis for a scientific degree of Cand. Tech. Sci. on a specialty 05.13.23 – Systems and Aids of Artificial Intellect. – The Kharkov National University of Radioelectronics, Kharkov, 2006.

A new method of phonoscopic identification of a person has been obtained on the basis of the individual speech code (ISC). The identification by the new method is carried out using an extremely limited, by the length (as short as one sound of the language) distorted speech material. The novelty of ISC consists in that for identification of speech objects as differentiating identification signs the functions individual for each object are used for processing speech signals which are brought into line with the object in the course of selection and optimization. The model of phonoscopic identification of a person was improved by introducing a new stage of classification of speech objects and selection of optimal signs for each object. A structural (geometric) method of studying the speech signals was developed that allowed to reject the methods of spectral description and to use the methods of description of elementary areas under the signal curve.

Key words: phonoscopic identification, individual speech code, structural method of research of speech signals, selection, optimization.

Работягов Андрій Валентинович

Метод ідентифікації людини на основі
індивідуального мовного коду

Автореферат дисертації

Відповідальний за випуск

В.П. Машталір.

Підп. до друку 10.05.2006 р. Формат 60×84¹/₁₆. Спосіб друку – ризографія.
Умов. друк. арк. 1,2. Облік. вид. арк. 1,0. Тираж 100 прим. Зам. № 2-352.

Надруковано в навчально-науковому видавничо-поліграфічному центрі ХНУРЕ.
Україна, 61166, м. Харків, просп. Леніна, 14.