

Метод ідентифікації людини на основі індивідуального мовного коду

Бондаренко М. Ф., Работягов А. В.

Харківський національний університет радіоелектроніки, м. Харків, Україна
software@kture.kharkov.ua

Анотація

У статті йде мова про новий метод фоноскопічної ідентифікації людини. Основна увага приділяється розгляду етапів евристичної адаптивної математико-логічної моделі методу ідентифікації. Новизну методу характеризує та обставина, що ідентифікація проводиться на основі *індивідуального мовного коду*, цифрова кодова комбінація якого являє собою набір *оптимальних* інформативних ознак, які відображують індивідуальні мовні особливості людини.

1. Вступ

Серед актуальних проблем сучасної науки одне з найбільш важливих місць із загального визнання багатьох вчених займає проблема автоматизованої обробки мовної інформації [1], зокрема, проблема розробки ефективних методів ідентифікації людини на основі аналізу акустико-фонетичних ознак. За останні 15 років (1988-2003 р.) простежується тенденція значного росту числа мовних досліджень. В основі більшості досліджень лежить *спектральна* ідеологія [1]. Однак усупереч очікуванням прогрес у даній галузі виявився не таким швидким, як того хотілося б. У чому ж причини такого положення справ? Відносно стану сучасних мовних досліджень відомий вчений-лінгвіст Р. К. Потапова пише: "Основні труднощі як для розпізнавання, так і для розуміння мови полягають у виділенні надійних ознак на акустико-фонетичному рівні" [1, стор.295].

2. Основний зміст

У спеціальній літературі в галузі прикладної лінгвістики підкреслюється, що проблема призначення інформативних ознак виникає кожний раз, коли досліджуються "важкі" мовні сигнали при проведенні, наприклад, ідентифікаційного дослідження. До такої категорії мовних об'єктів (сигналів) відносяться об'єкти, які характеризуються, зокрема, мінімальним обсягом мовної інформації (наприклад, обсягом 1-3 сек.), різним ступенем спотворення (наприклад, об'єкти записані з різних передаточних каналів). Неспроможність "традиційної" *спектральної* ознакової бази обумовлює пошук нових, нетривіальних ознак.

У чому складається новизна та особливість наших ознак? За своїм якісним змістом вони характеризуються нами як багаторівневі взаємозалежні ознаки, набір яких складає т.зв. (за прийнятою нами термінологією) *індивідуальний мовний код* (ІМК). В поняття ІМК нами вкладається наступний зміст: ІМК – це код, цифрова кодова комбінація якого являє собою набір інформативних ознак, які відображують індивідуальні мовні особливості людини; ІМК – це код, на основі якого

здійснюється класифікація мовних еталонних об'єктів (ЕО) і проводиться фоноскопічна ідентифікація; ІМК – це код, визначена цифрова кодова комбінація якого присвоюється визначеному ЕО в процесі виконання процедур адаптації, оптимізації і селекції; ІМК – це код, у якому кожний розряд цифрової кодової комбінації є *функціонально* значимим елементом.

Функція елемента (розряду) кодової комбінації ІМК – це "дія (операція, правило), що виконується в залежності від значення розряду кодової комбінації ІМК". З фактуальної сторони функції розрядів ІМК реалізовані у вигляді виконання окремих *підпрограм*, формальну математико-логічну основу яких складають прості математичні рівняння і логічні функції (предикати) математичної логіки.

Відповідно до класифікації, яка прийнята в теорії кодування, ІМК характеризується як 19-ти розрядний рівномірний код (див. табл.1).

Таблиця 1: Функціональне призначення розрядів ІМК

Простір	Розряд	Функція розряду
Vr	0	положення конструктивної лінії k1
	1	положення конструктивної лінії k2
	2	положення конструктивної лінії k3
	3	положення конструктивної лінії h1
	4	положення конструктивної лінії h2
	5	положення конструктивної лінії h3
Vc	6	кількість вікон конструктивної частини
	7	виключення конструктивної частини
Vm	8	виключення ознак конструкт. частини
	9	перетворення ознак
Vd	10	правило перетворення
	11	правило нормування
	12	значення коефіцієнта нормування "l"
Vw	13	значення коефіцієнта нормування "h"
	14	правило перетворення ознак (віднімання min, ранжирування)
Z1	15	призначення ділянки звуку для порівняння
	16	значення міри близькості ϵ^1
Z2	17	значення міри близькості ϵ^2
Z3	18	значення міри близькості ϵ^3

З метою експериментальної перевірки спроможності висунутих гіпотез ми ставимо перед собою таку задачу: в умовах невизначеності, коли ідентифікаційні ознаки відомі а ргіогі не точно, а лише в деяких межах, розробити адаптивну модель, яка б могла знайти *оптимальні* ідентифікаційні ознаки і правила, що забезпечують

класифікацію множини ЕО на скінченне число класів і проведення ідентифікації.

Формалізація задачі адаптивного моделювання полягає у математико-логічному опису та об'єднанні в єдине системне ціле етапів моделі на основі теорії розпізнавання образів, евристичного методу дослідження і деяких положень нейрофізіології. Модель складається з двох основних етапів: 1) етапу класифікації еталонних об'єктів і 2) етапу ідентифікації невідомого та еталонного об'єктів. Кожний з етапів складається з підетапів, механізм яких реалізовано у вигляді виконання визначених підпрограм (див. рис.1).

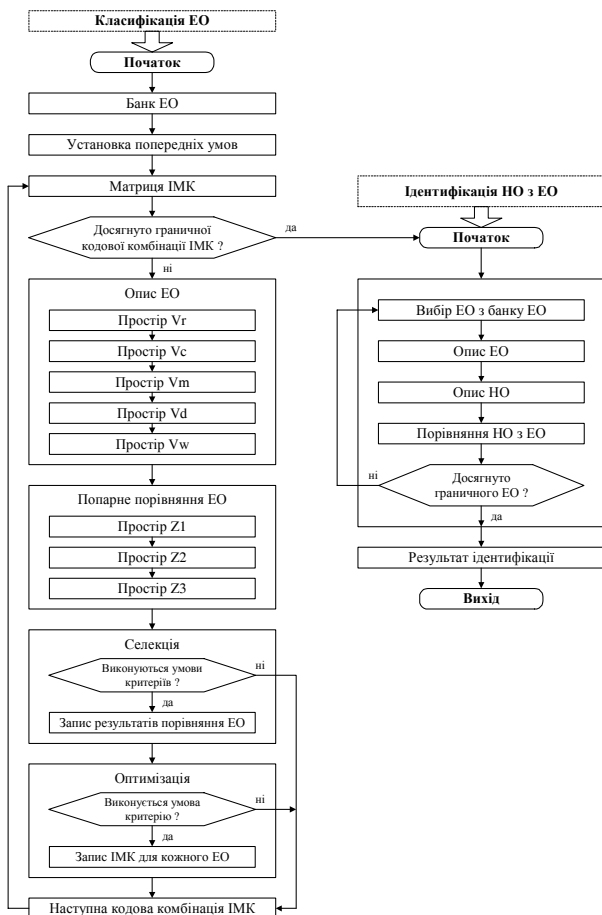


Рисунок 1: Структурно-функціональна схема моделі.

Мета етапу "Класифікація ЕО" полягає у визначенні оптимальних ідентифікаційних ознак і правил, що забезпечують класифікацію деякої множини ЕО. Об'єкт називається *еталонним* унаслідок того, що досліднику і, відповідно, комп'ютерної моделі заздалегідь відома приналежність об'єкта до визначеного класу об'єктів. Кожний ЕО являє собою звуковий файл спеціального формату, у якому міститься один голосний звук мови визначеного диктора.

Для практичної реалізації адаптивно-кодової концепції у моделі вводиться *матриця ІМК*. Матриця коду являє собою таблицю розрядів і можливих цифрових значень розрядів кодових комбінацій, з якої в процесі адаптивного пошуку відбирається визначена кодова комбінація ІМК.

Розряди кодової комбінації ІМК мають умовну позначку $FG_r(x)$, де нижній індекс r позначає порядковий номер розряду, x – цифрове значення розряду. Наприклад, $FG_2(3)$ означає: "значення 2-го розряду дорівнює числу 3". Повне число кодових комбінацій N визначається по формулі: $N = \prod n(FG_r)$, де $n(FG_r)$ – основа r -го розряду.

На етапі "Опис ЕО" формуються образи усіх ЕО. Опис кожного ЕО виконується по однотипному алгоритму в декілька етапів (у т.зв. просторах Vg, Vc, Vm, Vd, Vw), у яких реалізуються функції розрядів ІМК (див. табл.1).

У просторі Vg об'єктами первинного опису є *елементарні сегменти голосних звуків мови* (ЕСГЗМ) [2]. ЕСГЗМ уявляється у вигляді *геометричного* об'єкта (див. рис.2). На акустичному рівні як первинні ознаки нами прийняті геометричні елементи аналітичної геометрії: пряма (хорда) і складна плоска фігура. Таке уявлення (свого роду "геометрія мови") і призначення ознак апелює до інтуїції самого дослідника і спирається на систему знань не тільки в галузі лінгвістики, але і в інших галузях науки, наприклад, нейрофізіології. Формування первинного образу зводиться до виміру значень *довжини і площі* зазначених геометричних елементів, що у своїй сукупності складають *вектор* значень ознак.

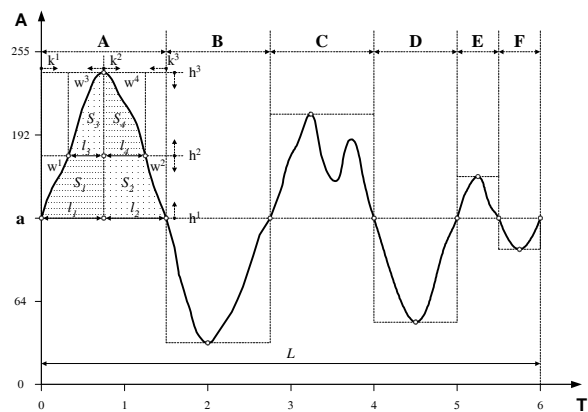


Рисунок 2: Первинний опис ЕСГЗМ у моделі.

На осі абсцис (T) розташована шкала часу (розмірність одиниць виміру – мсек). На осі ординат (A) – шкала амплітуд цифрового мовного сигналу. Тривалість ЕСГЗМ (L) дорівнює 6 мсек.

Як видно з рис.2, структура ЕСГЗМ складається з 6-ти конструктивних частин: A, B, C, D, E і F. Кожна з частин "поділяється" на декілька простіших у конструктивному відношенні елементів. Як "розподілювальний" елемент обрано прямокутне вікно, габарити якого обмежені горизонтальними h і вертикальними k конструктивними лініями. У кожному з вікон виділяються два геометричних структурних елемента: складна фігура і пряма горизонтальна лінія (хорда), які утворені габаритами вікна і формою ЕСГЗМ. Наприклад, у конструктивній частині A виділені чотири вікна: w_1, w_2, w_3 і w_4 ; у вікні w_1 визначаються площа S_1 складної фігури і довжина l_1 хорди.

У просторі Vg механізм адаптації реалізується за рахунок зміни положення конструктивних ліній l і h ,

розташування яких безпосередньо впливає на габарити вікон i , відповідно, на значення первинних ознак [3]. "Керують" положенням конструктивних ліній визначені розряди ІМК. Зі змістовної сторони "керування" здійснюється за рахунок того, що визначені значення розрядів ІМК (див. табл.1) "включають" відповідні підпрограми, які забезпечують адаптивні властивості комп'ютерної моделі. Так, наприклад, FG_3 виконує адаптивну функцію "положення конструктивної лінії hl ", його конкретне цифрове значення визначає конкретне положення цієї лінії для відповідних вікон усіх конструктивних частин ЕСГЗМ: при $FG_3(1)$ конструктивна лінія hl переміщується на 2-і одиниці нагору. Значення "середньої лінії" (а) обчислюється як \max число переходів мовного сигналу через "0" на "шумній" ділянці.

Оскільки кожний ЕО представлено одним голосним звуком мови, який утворено послідовним рядом ЕСГЗМ, то опис у цілому ЕО зводиться до опису усіх ЕСГЗМ. У результаті первинного опису формується квадратна матриця, у відповідних рядках якої розташовані вектори ЕСГЗМ, у стовпцях – координати векторів. З урахуванням значень FG_6 розмірність вектора кожного ЕСГЗМ лежить у межах від 24 до 96.

На наступних етапах побудови моделі виникає необхідність у виконанні деяких математико-логічних операцій над вже отриманими значеннями первинних ознак в залежності від відповідних значень розрядів ІМК (див. табл.1):

- у кожному векторі простору V_c виключаються визначені конструктивні частини ЕСГЗМ і координати вектора простору V_r ;
- у просторі V_m за визначеним правилом виконується віднімання (ділення) мінімального значення координати стовпця (рядка) матриці з (на) всіх однотипних координат, розташованих у стовпці (рядку) матриці простору V_c ;
- у просторі V_d за визначеними правилами виконується нормування векторів простору V_m ;
- у просторі V_w за визначеними правилами в кожному векторі простору V_d віднімається мінімальне значення координати відповідно до приналежності до конструктивної частини ЕСГЗМ і типу структурних ознак ("S" і "I"); значення координат ранжируються.

Класифікація ЕО на основі вирішальних правил виконується в т.зв. *просторах* Z_1 , Z_2 і Z_3 . Класифікація здійснюється за рахунок попарного порівняння усіх ЕО за правилом "кожний з кожним" на основі обчисленого в нормованому метричному просторі відстані d^z і значень мір близькостей. Відстань d^z між двома ЕО, які порівнюють, обчислюється за наступним алгоритмом:

1) кожний окремо узятий вектор ЕО більшої *потужності* по черзі порівнюється з усіма векторами ЕО меншої *потужності*. У процесі попарного порівняння ЕО порогово співставляється кожна пара координат: якщо відстань між координатами менше значення міри близькості \mathcal{E}^1 , то результату порівняння присвоюється логічне значення "1", якщо більше – "0". У результаті одного порівняння формується вектор Z^1 , значення координат якого належать бінарній множині $\{0, 1\}$. У результаті порівняння одного вектора ЕО більшої

потужності з усіма векторами ЕО меншої *потужності* знаходиться вектор Z_{\max} (вектор Z^1 с \max числом "1"); у результаті порівняння всіх векторів – формується множина векторів Z_{\max} .

2) На основі числа координат векторів Z_{\max} , рівних "1", розмірності і числа векторів Z_{\max} формується один вектор Z , значення координат якого належать бінарній множині $\{0, 1\}$. Значення координат вектора Z визначаються в результаті порогового співставлення числа координат векторів Z_{\max} , що дорівнюють "1", і розмірності вектора Z_{\max} : якщо відстань між ними менше значення міри близькості \mathcal{E}^2 , то результату порівняння (якій-небудь координаті вектора Z) присвоюється логічне значення "1", якщо більше – "0". Розмірність вектора Z дорівнює числу векторів Z_{\max} .

3) Відстань d^z між двома ЕО, які порівнюють, обчислюється в результаті порогового співставлення суми координат вектора Z , рівних "1", і розмірності вектора Z : якщо відстань між ними менше значення міри близькості \mathcal{E}^3 , то результату порівняння (відстані d^z) присвоюється логічне значення "1", якщо більше – "0".

Два ЕО, які порівнюють, вважаються *тотожними*, тобто приналежними до одного класу ЕО, якщо відстань d^z між ними дорівнює "1", якщо "0" – *нетотожними*.

Мета етапу "Селекція" складається з відбору *селективних* кодових комбінацій, тобто таких кодових комбінацій ІМК, при яких відсутні помилки класифікації: помилки I-го і II-го роду [2] (в іншій інтерпретації "пропуск цілі" і "захоплення помилкової цілі"). Якщо в результаті класифікації ЕО встановлені тільки лише тотожності між ЕО, то модель запише результати попарних порівнянь ЕО в спеціальну таблицю.

Мета етапу "Оптимізація" складається з відбору *оптимальних* кодових комбінацій ІМК із множини селективних кодових комбінацій і їх присвоєння ЕО. *Оптимальними* вважаються такі кодові комбінації, при яких значення критерію оптимізації для кожної пари ЕО мінімальні.

У моделі передбачені різні варіанти рішення задачі класифікації ЕО. Адаптивний пошук оптимальних рішень виконується в автоматичному режимі за рахунок послідовної, крок за кроком, зміни значень кодових комбінацій ІМК і аналізу поточного стану моделі на основі критеріїв селекції та оптимізації. На етапі "Класифікація ЕО" може бути виконано $N = 124.002.900.000$ ітерацій. Після виконання усіх ітерацій у результаті адаптивного пошуку визначається остаточний вигляд *оптимальної* кодової комбінації для кожного ЕО, яка складає *індивідуальний мовний код* (ІМК) ЕО.

У моделі умовна позначка ЕО, який характеризується своєю індивідуальною кодовою комбінацією ІМК, подається у наступному вигляді: $O_i^k(FG_0(x), FG_1(x), \dots, FG_{18}(x))$ або $O_i^k(x_0, x_1, \dots, x_{18})$, де O – умовна позначка об'єкта, i – порядковий номер ЕО,

k – номер класу ЕО, x_r – цифрове значення r -го розряду.

Наприклад, $O_{29}^6(1, 0, 2, 2, 0, 1, 0, 5, 1, 0, 0, 3, 3, 1, 1, 3, 2, 1)$.

У моделі ідентифікація виконується за "класичною" схемою – за правилом "один з усіма", згідно з яким невідомий об'єкт (НО) по черзі порівнюється з кожним ЕО. Процес ідентифікації складається з трьох основних етапів: 1) етапу опису ЕО, 2) етапу опису НО і 3) етапу порівняння НО з ЕО за правилом "один з усіма". Механізми опису і порівняння об'єктів етапу "Ідентифікація" аналогічні розглянутим вище механізмам етапу "Класифікація ЕО" з тією лише різницею, що опис, порівняння і встановлення тотожності (нетотожності) між НО і ЕО здійснюється на основі правил, які виконуються в залежності від значень розрядів кодової комбінації ІМК еталонних об'єктів.

Для експериментальної перевірки моделі були відібрані 190 ЕО банку ЕО: кожний ЕО представлено тільки лише одним голосним звуком мови, наприклад, голосним звуком *о* у фонетичному оточенні *вк* у слові *вокзал*; кожний четвертий ЕО вимовлений у навмисно спотвореній манері; кожні чотири ЕО записані з різних передаточних каналів і за різними умовами. В ході перевірки моделі досліднику був заздалегідь відомий результат ідентифікації: як "невідомий" об'єкт призначено відомий (еталонний) об'єкт, який у моделі називається *перевірочним об'єктом* (ПО). Один перевірочний клас об'єктів був представлений 8-ю ПО.

Результати ідентифікації: модель розпізнала перевірочні об'єкти як об'єкти, що належать до класу ЕО, з якого відбиралися ПО. При встановленні тотожності (нетотожності) між ПО і ЕО на усій множині ЕО не було отримано помилок II-го роду. Витрати часу на проведення в цілому ідентифікації склали 470 годин: на виконання класифікації ЕО – 462 години, ідентифікації ПО з ЕО – 8 годин.

Обмеження методу: сегментація голосних звуків на ЕСГЗМ виконується органолептично (вручну); невелика кількість об'єктів експериментальної перевірки; великі витрати часу на проведення в цілому ідентифікації.

3. Висновки

У роботі наведено нове рішення наукової задачі, яка полягає у розробці нового методу фоноскопічної ідентифікації людини. Рішення даної задачі базується на нових теоретичних положеннях і нетрадиційних поглядах у питаннях формального перетворення мовних сигналів і класифікації мовних об'єктів.

Застосування методу ідентифікації на основі "кодової" концепції дає можливість проводити ідентифікаційні дослідження, по-перше, на вкрай обмеженому за обсягом і спотвореному мовному матеріалі, по-друге, на множині об'єктів. Це приводить не тільки до розширення галузі застосування методу, але і до підвищення надійності і якості ідентифікації.

4. Література

- [1] Потапова Р. К. Речь: коммуникация, информация, кибернетика. Учебное пособие. Изд. 3-е, стереотипное. – М.: Едиториал УРСС, 2003. – 568 с.
- [2] Бондаренко М. Ф., Дрюченко А. Я., Шабанов-Кушнаренко Ю. П. Гласные звуки в теории и эксперименте. – Харьков: ХНУРЭ, 2002. – 348 с.

- [3] Работягов А. В. Решение задачи идентификации человека по параметрам речи на основе принципов адаптации и оптимизации // Автоматизированные системы управления и приборы автоматики. Всеукраинский межведомственный научно-технический сборник. Вып. 121. – Харьков: ХНУРЭ, 2002. – С. 80-87.