

ДОДАТОК А
ПУБЛІКАЦІЇ

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ РАДІОЛЕКТРОНІКИ

МАТЕРІАЛИ 25-го МІЖНАРОДНОГО МОЛОДІЖНОГО ФОРУМУ

«РАДІОЕЛЕКТРОНІКА І МОЛОДЬ У ХХІ СТОЛІТТІ»

20 – 22 квітня 2021 р.

Том 4

КОНФЕРЕНЦІЯ

**«ПЕРСПЕКТИВИ РОЗВИТКУ ІНФОКОМУНІКАЦІЙ
ТА ІНФОРМАЦІЙНО-ВИМІРЮВАЛЬНИХ ТЕХНОЛОГІЙ»**

Харків 2021

ДОСЛІДЖЕННЯ АЛГОРИТМІВ РОЗПІЗНАВАННЯ ГЕНДЕРНОЇ ПРИНАЛЕЖНОСТІ ЛЮДИНИ ЗА ЇЇ ГОЛОСОМ

Єрмолаєв А.А.

Науковий керівник к.т.н., доц. Омельченко С.В.

Харківський національний університет радіоелектроніки
(61166, Харків, просп. Науки, 14, каф. ІМІ, тел. (057) 702-14-29)

e-mail: artem.yermolaiev@nure.ua

Object of research – methods of gender recognition using speech signals.

The purpose of this work is a research and implementation of methods for automatic gender identification by voice.

The main methods of gender recognition using speech signals are: method of support vector machines, method of gaussian mixture model. A comparative analysis of methods has been made. Voice features and a method of their extraction on a basis of mel-frequency cepstral coefficients were reviewed. Also algorithms of gender recognition through speech were reviewed. In the practical part the algorithm of gender recognition was implemented based on the frequency of the main tone of voice.

Мета роботи: дослідити ефективність запропонованих методів автоматичної ідентифікації статі людини за голосом. У ході роботи були розглянуті основні методи розпізнавання статі людини за голосом: метод опорних векторів, метод гаусових сумішей. Були розглянуті голосові ознаки та метод їх вилучення на основі мел-частотних кепстральних коефіцієнтів. В практичній частині було реалізовано алгоритм розпізнавання статі на основі частоти основного тону голосу та 2-ої формантної частоти.

Експериментальні дослідження алгоритму розпізнавання статі диктора з одноетапним визначенням кількості нулів в смугах формант проводилися методом статистичного випробування на вибірках 5-ти сигналів для кожного з 5-х різних дикторів чоловічої та жіночої статі. За вибірками оцінювалися параметри вирішального правила, а також контрольні вибірки реальних сигналів використовувалися для оцінювання якості розпізнавання сигналів.

Дослідження показали, що найбільший внесок в розпізнавання вносить друга форманта і показник її ексцесу, де можливе лінійне розділення класів. На рис. 1 показано отримане експериментальне зосередження пар вимірювань середнього значення другої формантної частоти основного тону та коефіцієнту ексцесу для п'яти жіночих голосів (трикутники) і п'яти чоловічих голосів (квадрати). Це дозволяє використовувати для другої форманти лінійні розділяючі межі між ними.

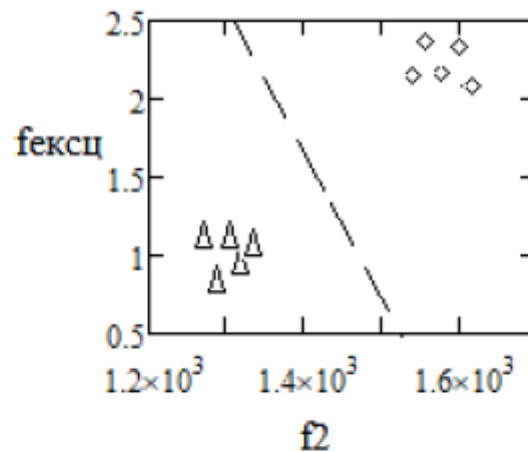


Рисунок 1 – Зосередження пар вимірювань середнього значення другої формантної частоти основного тону і коефіцієнта ексцесу для жіночих голосів (трикутники) і чоловічих голосів (квадрати)

Експериментальне дослідження показало помилку ймовірності розпізнавання, що дорівнює 0. При додатковій дії адитивної завади типу гаусів білий шум, отримана оцінка середньої ймовірності правильного розпізнавання при відношенні сигнал/шум $q = 20$, для алгоритму розпізнавання по частоті основного тону та 2-ої формантної частоти – 0.8.

Проведене дослідження підтверджує можливість використання запропонованого алгоритму розпізнавання за оцінками моментів частоти основного тону та 2-ої формантної частоти.

ПЕРЕЛІК ПОСИЛАНЬ

1. Пресняков И. Н. Автоматическое распознавание отдельных слов и фонем речи / И. Н. Пресняков, С. В. Омельченко. // Радиоэлектроника и информатика научно-технический журнал. – 2003. – №2. – С. 41–47.
2. Пресняков И. Н. Автоматическое распознавание речи в каналах передачи / И. Н. Пресняков, А. В. Омельченко, С. В. Омельченко. // Радиоэлектроника и информатика научно-технический журнал. – 2002. – №1. – С. 26–31.

АЛФАВІТНИЙ ПЕРЕЛІК

- A**
Andrii Zhuravka, 14, 16, 18
Ayodele Tega Ajadi, 46
- D**
David Ogamune, 16
Denis Zhuravka, 14, 16, 18
- E**
Ethel Chila, 18
- I**
Ikeza Obasi A. D., 44
- J**
Joel Kashaija, 66
- O**
Okwudili Gene Omukaogu, 14
- P**
Perslyn I. V., 22
- S**
Samad Habib Suhel, 8
- T**
Tresor M.A., 42
- A**
Аль-Вандави Саиф Ахмед Искандар
Исмаель, 104
Ащепков В.О., 119
- Б**
Безрученко О.Ю., 104
Бельков Е.А., 116
Беленцов А.С., 38
Білик В. О., 26
Білокурова А.О., 40
Богомазов С.А., 129
Бондаренко С.В., 159
Босенко Д.В., 169
- Бураківська А. О., 100
- В**
Варченко М.А., 123
Вервейко В.В., 110
- Г**
Герус М.А., 68
Гонтарь І. А., 78
Греков І. С., 24
- Д**
Давиденко Н.В., 112
Дікаленко Д. Д., 36
Добринін К.І., 56
Дученко П.Ю., 153
- Є**
Єрмолаєв А.А., 88
- Ж**
Жирна Г.А., 155
- З**
Запотроєв Д.І., 123
- К**
Каницька І.В., 167
Кепещук Д. Т., 141
Кононов В.Б., 149
Кононова О.А., 149
Коржов И.М., 161
Корнейцова Н.В., 32
Красюкова В.В., 70
Куліченко В.В., 121
Курлан О.О., 90
Кухарчук М.М., 30
- Л**
Ларіонов В.В., 72
Лісняк О.О., 12
Луценко М.И., 157
Лялічев В. Д., 102

ДОДАТОК Б СЛАЙДИ ПРЕЗЕНТАЦІЇ

Харківський національний університет радіоелектроніки

Кафедра Інформаційно-мережної інженерії

Кваліфікаційна робота магістра на тему:

Сумісне розпізнавання віку людини за її голосом та зображенням обличчя.

Студент:	Єрмолаєв Артем Андрійович
Група:	ІМІм-19-2
Керівник:	доц. Омельченко Сергій Васильович

Харків - 2021

1

Мета роботи

Метою даної роботи є розробка та дослідження методів сумісної ідентифікації віку людини за голосом та зображенням обличчя людини.

Для досягнення поставленої мети необхідно вирішити наступні завдання:

1. Дослідити існуючі методи вирішення проблеми ідентифікації віку людини по голосу та зображенню обличчя, оцінити якість, а також існуючі проблеми та обмеження в їх роботі.
2. Впровадити один з кращих алгоритмів ідентифікації віку людини за голосом та зображенням обличчя.
3. Провести експериментальне дослідження розробленого алгоритму.

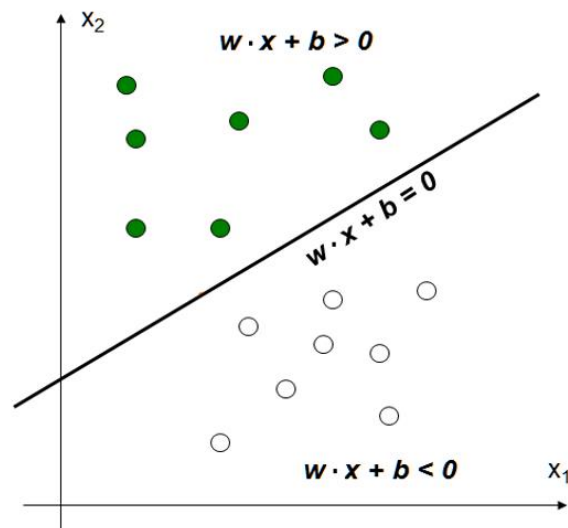
2

Розглянуті методи розпізнавання

- Метод опорних векторів
- Оцінка віку за зображеннями обличчя на основі вибору ієрархічних ознак
- Метод KNN – K-найближчих сусідів

3

Метод опорних векторів



4

База еталонів дікторів

Таблиця 1 – кількість дикторів

Акцент	Стать		
	Жіноча	Чоловіча	Всього
Британський	65	57	122
Американський	59	65	124
Всього	124	122	246

5

Спектрограма чоловічого голосу

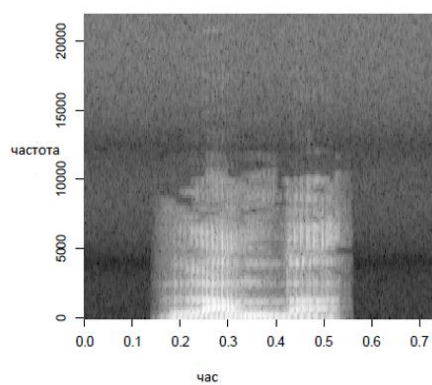


Рисунок 1 – Спектрограма чоловічого голосу

6

Спектрограма жіночого голосу

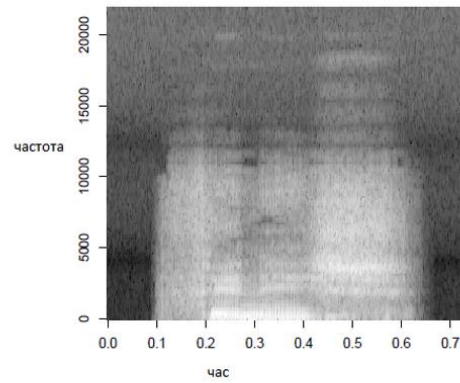


Рисунок 2 – Спектрограма жіночого голосу

7

Результати досліджень алгоритмів з застосуванням ознак MFCC

Таблиця 2 – Результати досліджень

# MFCC	ERR(РБФ)	ERR(поліном)
12	0.0748	0.0549
19	0.0232	0.0418
26	0.0072	0.0131
34	0.0023	0.0051
40	0.0019	0.0063

8

Особливості методу опорних векторів

Метод опорних векторів дає високу точність класифікації, має теоретичне обґрунтування, дозволяє застосовувати різні методи до класифікації відповідно до вибору функцій ядра.

Серед недоліків необхідно відмітити проблему вибору ядра, а також повільне навчання у випадку задачі многокласового розпізнавання.

9

Еталони зображень що застосовувались для алгоритмів розпізнавання віку за зображеннями обличчя на основі вибору ієрархічних ознак



10

Результати розпізнавання віку

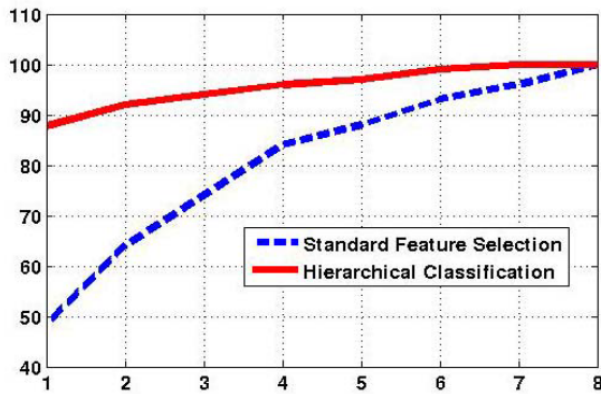


Рисунок 2 - Сукупний бал відповідності для оцінки віку

Таблиця 3 – Ієрархічна кластеризація та оцінка вікових груп

Level 1		Level 2		Level 3	
18-43	96.1%	18-32	93.3%	18-24	91.5%
		33-43		25-32	
44-100	97.2%	44-56	97.2%	33-37	92.1%
				38-43	
		44-48		90.7%	
		49-56			
		57-100		57-72	94.6%
				73-100	

Дослідження методу KNN – К-найближчих сусідів

Процент правильного розпізнавання за віком та статі лише за допомогою мел-частотних кепстральних коефіцієнтів MFCC

МФСК	Стать вірно	Вік вірно
MFCC 1 2 3	55,55%	44,44%
MFCC 4 5 6 7 8 9	66,66%	44,44%
MFCC 10 11 12	22,22%	33,33%
Усі MFCC	66,66%	44,44%

Процент правильного розпізнавання за віком та статтю разом за допомогою мел-частотних кепстральних коефіцієнтів MFCC

MFCC	Рекрівень пізнання (вік і стать разом)
MFCC 1 2 3	33,33%
MFCC 4 5 6 7 8 9	44,44%
MFCC 10 11 12	11,11%
Усі MFCC	44,44%

13

Результати розпізнавання віку та розпізнавання статі за допомогою форматних частот

Форманти	Процент правильного розпізнавання віку	Процент правильного розпізнавання статі
F3, F4	22,22%	22,22%
F3	38,88%	27,77%
F4	27,77%	44,44%
F1, F4	38,88%	55,55%
F1, F2, F3	50%	55,55%
F1, F2, F4	44,44%	50%
F1, F2	55,55%	61,11%

14

Результати розпізнавання віку та розпізнавання статі за допомогою форматних частот (продовження)

F1, F3, F4	33,33%	44,44%
F1, F3	50%	55,55%
F1	38,88%	38,88%
F2, F3	55,55%	50%
F2, F3, F4	38,88%	50%
F2, F4	50%	38,88%
F2	38,88%	38,88%
F1,F2, F3, F4	50%	66,66%

15

Результати розпізнавання статі та віку разом використовуючи лише формантні частоти

Форманти	Процент правильного розпізнавання (вік та стать разом)
F1, F2, F3, F4	50%
F1, F4	33%
F1, F3	39%
F1, F2	55%
F1	28%
F1, F2, F3	39%
F1, F2, F4	39%
F1, F3, F4	28%
F2, F3, F4	33%
F2, F4	33%
F2	33%
F4	17%
F3	22%
F2, F3	39%
F3, F4	22%

16

Висновки

- Були розглянуті методи сумісного розпізнавання віку людини за ознаками голосу та зображенню обличчя людини.
- Були розглянуті особливості вилучення ознаки віку людини.
- В практичній частині було розглянуто алгоритм KNN (К-Найближчих сусідів), а також алгоритм розпізнавання віку за зображеннями обличчя на основі вибору ієрархічних ознак.

17

Дякую за увагу!

18

