

УДК 004.056.5:35.078.3(02)

DOI: 10.18372/2310-5461.37.12444

О. К. Юдін, д-р техн. наук, проф.

Національний авіаційний університет

orcid.org/0000-0001-5098-7796

e-mail: kszi@ukr.net;

Р. В. Зюбіна, аспірант

Національний авіаційний університет

orcid.org/0000-0002-8654-6981

e-mail: kszi@ukr.net

МЕТОД ЕФЕКТИВНОЇ ШИРИНИ СПЕКТРУ**Вступ**

Розвиток інформаційних технологій потребує розробки нових систем оброблення, зберігання та захисту інформації, а належне використання новітніх інформаційних систем надання ІТ-послуг з аутентифікації та ідентифікації в галузі інформаційних технологій унеможливує процеси витоку критичної інформації, що на сьогодні стає стратегічно важливим питанням у процесах функціонування сучасного суспільства, держави та особистості.

Використання біометрії в сучасних інформаційних технологіях дало поштовх для подальших наукових досліджень, а використання унікальних особливостей будови людського тіла створило новий простір ознак для ідентифікації та аутентифікації особистості в інформаційних системах [1–4].

Аналіз останніх досліджень показав, що сучасні системи біометричної ідентифікації працюють в ідеальних умовах, а наявність завад будь-якого характеру не передбачається. На сьогодні не існує методів ідентифікації диктора в умовах наявних завад в каналі зв'язку, що робить дане дослідження досить актуальним.

Мета статті — розроблення методу ідентифікації диктора за визначеним класом біометричних ознак голосу для підвищення ефективності функціонування розподілених інформаційних систем.

Виклад основного матеріалу

Основним інструментом у системах аналізу сигналів є спектральний аналіз, який використовується для визначення характерних частот, зменшення рівня завад та ін.

Нехай існує зразок голосу людини, що підлягає ідентифікації, $S_{\text{вх}}(t)$, який залежить від таких загальновідомих параметрів t, A, φ, ω і записується як

$$S_{\text{вх}}(t) = S_{\text{вх}}(t, A, \varphi, \omega).$$

Відповідно до поставленої задачі ідентифікації, на вхід системи приходять сигнал $S_{\text{вх}}(t)$, який підлягає аналізу та визначенню його належності до класів об'єктів у базі даних $DB = \{S_1(t), S_2(t), \dots, S_n(t)\}$, тобто $S_{\text{вх}}(t) \in DB$.

Одним з класичних методів спектрального аналізу у радіотехніці став Фур'є-аналіз, або аналіз сигналів на основі перетворення Фур'є. Перетворення Фур'є — це насамперед, математичний апарат для спектрального аналізу сигналу, шляхом розкладання його на синусоїдальні коливання. Якщо існує неперервний нескінченний сигнал $S(t)$, то його можна представити у вигляді інтегрального ряду Фур'є:

$$X(k) = \int_0^T e^{i \frac{2\pi k}{T} t} x(t) dt.$$

Результатом такого перетворення будуть спектральні діаграми амплітуд A_ω і частоти ω . У спектральному аналізі більшу зацікавленість являють скінченні за часом сигнали. Мова людини являє собою змінний в часі сигнал, відповідно й спектр такого сигналу буде змінюватись залежно від часу, тому в процесі аналізу мовленнєвого сигналу зручніше використовувати короткі фрагменти аудіозапису (рис. 1). Для таких фрагментів прийнято використовувати дискретне перетворення Фур'є, яке матиме вигляд:

$$X(k) = \frac{1}{N} \sum_{n=0}^{N-1} x(n) W_N^{nk},$$

де N — кількість відліків дискретного сигналу,

$$W_N^{nk} = e^{-j \frac{2\pi nk}{N}}.$$

Виходячи з вищесказаного, для роботи з алгоритмом ідентифікації голосу людини, розкладемо всі зразки мовленнєвих сигналів, що наявні в базі даних, на спектральні діаграми (рис. 2). Надалі всі маніпуляції будуть проводитись з ними. Як приклад, проведемо аналіз зразку голосу $S_{\text{вх}}(t)$, що підлягає ідентифікації.

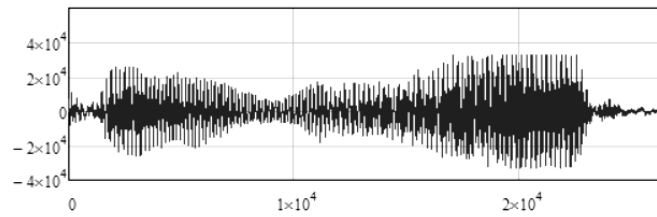


Рис. 3.1. Часове представлення біометричного ідентифікатора

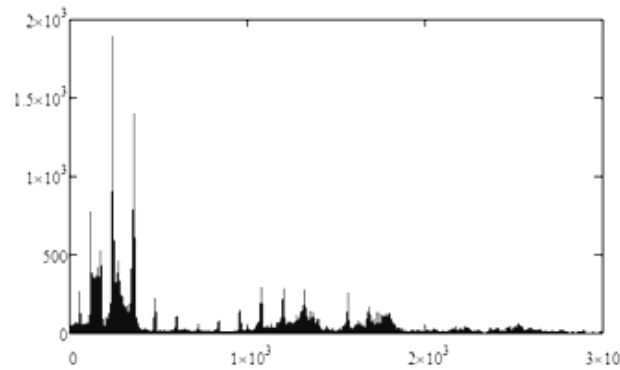


Рис. 3.2. Спектральне представлення біометричного ідентифікатора оператора БПЛА

Обмеження ряду Фур'є до N -відліків еквівалентно згортці частотної характеристики ідеального ФНЧ $H_u(e^{j\omega})$, АЧХ якої має форму прямокутника, з ядром Діріхле $W_n(e^{j\omega})$. Це означає, що коливання функції $W_n(e^{j\omega})$ потрапляють в область частотного прямокутника, у якій згортка точно відтворює ці коливання. У результаті отримуємо АЧХ реального фільтра, у якій поблизу точки розриву спостерігається два ефекти:

а) виникають помилки апроксимації у вигляді пульсацій частотної характеристики $H(e^{j\omega})$, які зумовлені бічними пелюстками функції $W_n(e^{j\omega})$.

б) утворюється згладжувальний розрив — перехідна смуга, ширина якої залежить від головної пелюстки функції $W_n(e^{j\omega})$ і приблизно дорівнює її ширині [5].

Ці помилки пов'язані з характером збіжності ряду Фур'є в точці розриву першого роду. По-перше, в точці розриву першого роду ряд Фур'є збігається до середніх граничних значень функції зліва і справа. По-друге, в точці розриву збіжність ряду Фур'є не є рівномірною і носить особливий характер, який виражається в появі пульсацій поблизу точки розриву, максимум яких ліворуч і праворуч становить 9 % від АЧХ і залишається таким, незалежно від N . Цей феномен отримав назву *явища Гіббса*. Керувати збіжністю ряду Фур'є можна за допомогою вагової послідовності кінцевої довжини $w(n)$, яка називається *вікном* або *ваговою функцією*.

Основні характеристики, що відрізняють вагові функції — розширення основної пелюстки

спектра та ступінь придушення розмиття спектра, так званих бокових пелюсток. Тут існує пряма залежність між шириною основної пелюстки та придушенням бокових пелюсток, чим більше придушені бокові, тим ширше основна пелюстка, і навпаки. Однозначного визначення кращої віконної функції немає, оскільки в процесі вирішення різних задач фільтрації необхідно знаходити компроміс для описаної залежності, тим самим жертвуючи одним із показників.

Найпростішою ваговою функцією є прямокутна (Діріхле), однак вона не забезпечує достатнього придушення бокових пелюсток, що в свою чергу, вирішується використанням спеціальних вікон.

Узагальнене вікно Хеммінга, має такий вигляд:

$$w_x(n) = \begin{cases} \alpha - (1 - \alpha) \cos\left(\frac{2\pi n}{N}\right), & 0 \leq n \leq N; \\ 0, & \text{за інших } n, \end{cases}$$

де α — лежить у межах $[0,1]$.

Випадок $\alpha = 0,5$ відповідає вікну Хеннінга, а випадок $\alpha = 0,54$ — вікну Хеммінга.

Отримати частотну характеристику представленої вагової функції можна у випадку, якщо представити її у часовій області у вигляді множника

$$w_x(n) = w_n(n) \left[\alpha - (1 - \alpha) \cos\left(\frac{2\pi n}{N}\right) \right],$$

де $w_n(n)$ — прямокутне вікно, довжина якого $N + 1$ дорівнює довжині вікна $w_x(n)$.

Шляхом використання формули Ейлера зробимо заміну косинусів експоненціальними функціями та перейдемо до частотної характеристики. Наслідком множення $x(n)$ на $e^{j\omega_0 n}$ буде зміщення в частотній області на ω_0 : $X(e^{j(\omega-\omega_0)})$.

Кінцевий вигляд частотної характеристики узагальненого вікна Хеммінга матиме такий вигляд:

$$W_x(e^{j\omega}) = \alpha W_n(e^{j\omega}) + \frac{1-\alpha}{2} W_n\left(e^{j\left(\omega-\frac{2\pi}{N}\right)}\right) + \frac{1-\alpha}{2} W_n\left(e^{j\left(\omega+\frac{2\pi}{N}\right)}\right).$$

Із сказаного вище зрозуміло, що частотна характеристика узагальненого вікна Хеммінга являє собою суму зміщених по частоті характеристик прямокутного вікна $W_n(e^{j\omega})$.

При цьому друга і третя компоненти суттєво зменшують амплітуду бічної пелюстки поблизу головної (оскільки бічні пелюстки, другий і

і третій компоненти перебувають у протифазі з бічними пелюстками першої компоненти).

Отже, рівні бічних пелюсток виявляються значно нижче, ніж у прямокутного вікна.

Однак ширина головної пелюстки узагальненого вікна Хеммінга виявляється в два рази більшою, ніж ширина головної пелюстки прямокутного вікна.

Основною відмінністю вікна Хеммінга є те, що використання $\alpha = 0,54$ дозволяє сконцентрувати 99,96 % загальної енергії спектра в головній пелюстці, при цьому максимальний рівень бокових пелюсток на 40 дБ нижче головного максимуму [6].

Порівняльні характеристики основних класичних вагових функцій представлені в таблиці. Таким чином, використання вагової функції дозволяє очистити сигнал від шуму і вирішити проблеми з помилками при збіжності ряду Фур'є.

Порівняння спектральних характеристик сигналу до використання вагової функції та після, представлені на рис. 3.

Основні параметри класичних вікон

Вагові функції		Рівень бокових пелюсток W_{\max} , дБ	Кореляція випадкових складових сигналу $C(r)$, %	
			0,75	0,5
Прямокутна		- 13,3	75,0	50,0
Трикутна		- 26,6	71,9	25,0
Хеммінга	$\alpha = 1,0$	- 23,0	75,5	31,8
	$\alpha = 2,0$	- 31,5	65,9	16,7
	$\alpha = 3,0$	- 39,3	56,7	8,5
	$\alpha = 4,0$	- 46,8	70,7	4,3
Хеммінга		- 42,7	65,9	48,6
Блекмена	Наближена	- 58,2	56,7	9,0
	Точна	- 51,0	62,7	14,0
Блекмена-Херріса	3 - люд, - 62 дБ	- 62,1	61,0	12,8
	3 - люд, - 71 дБ	- 70,9	57,2	9,5
	4 - люд, - 74 дБ	73,8	53,9	7,4
	4 - люд, - 92 дБ	92,2	46,0	3,8
Наталла		- 93,8	45,4	3,5
Блекмена- Наталла		- 98,7	47,0	4,2
Бартленна- Ханна		- 35,9	67,5	18,7
Вікно з плоскою вершиною		- 82,6	4,42	-

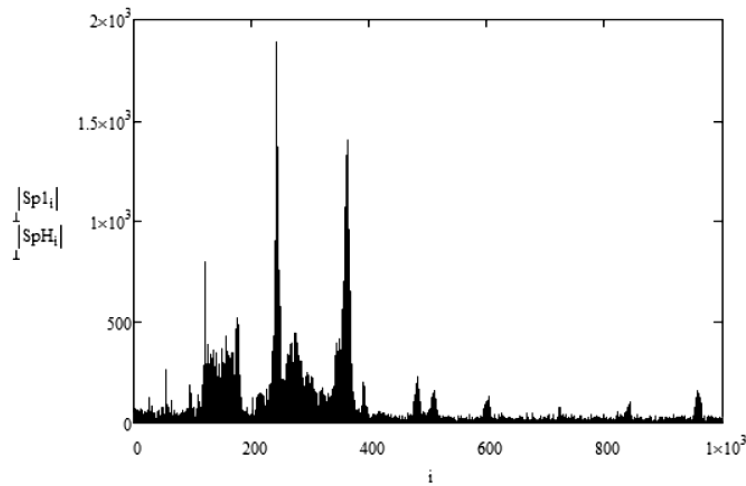


Рис. 3. Використання вікна Хеммінга в процесі ідентифікації голосу людини:

$Sp1_i$ — спектр сигналу, що підлягає ідентифікації;

SpH_i — спектр сигналу, що підлягає ідентифікації після використання вікна

Залежно від співвідношення амплітуд частотних складових спектра, звук приймає різне відтворення і сприймається системою ідентифікації, як тон або шум. Якщо розглядати спектральне представлення сигналу, можна встановити залежності у вигляді чітко виражених піків амплітуд гармонійних складових спектру, що системно повторюються з деяким інтервалом частот. Зрозуміло, що кожна така складова відтворює один тон сигналу та повинна сприйматися, як індивідуальна інформаційна характеристика повідомлення. У випадку, коли піки амплітуд не мають чітко вираженого характеру (гладка функція спектру) і визначити піки практично неможливо, тоді такий звуковий сигнал сприймається системою як шум.

У теорії і практиці ймовірно-статистичних дослідженнях фундаментальне місце займає нормальний (гаусівський) закон розподілу випадкових величин. Окрім того, що даний закон зустрічається на практиці найчастіше, його особливістю є і те, що він є граничним і всі інші закони розподілу наближаються до нього.

Функція щільності розподілу випадкових величин цього типу має вигляд:

$$f(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{(x - \mu_x)^2}{2\sigma_x^2}}$$

де основними параметрами є μ — математичне очікування, а σ — середньоквадратичне відхилення.

Прийmemo за випадкову величину спектральне представлення зразку голосу людини, який прийшов на вхід системи для ідентифікації і будемо порівнювати її зі спектральним представ-

ленням еталонного сигналу, тим самим зробивши заміну однієї з основних характеристик. У результаті таких маніпуляцій, отримаємо видозмінену функцію розподілу:

$$f(Sp_{\text{еталон}}) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{(Sp_{\text{еталон}} - Sp_{\text{вх}})^2}{2\sigma_x^2}}$$

де $Sp_{\text{еталон}}$ — спектральне представлення еталонного зразку голосу людини, а $Sp_{\text{вх}}$ — спектр сигналу, що прийшов на вхід системи для подальшої ідентифікації.

У сучасних системах статистичної теорії прийняття рішення в багатоальтернативних ситуаціях, найчастіше використовується математичний апарат Байеса. Цей спосіб дає можливість використовувати критерій та систему прийняття рішення на базі мінімально достатньої кількості інформації з визначеним порогом прийняття рішення V_k .

Результати прийняття рішень оцінюються за ступенем їх відповідності поставленій цілі. Дана відповідність, визначається за допомогою кількісної міри, що визначає виграти чи втрати від прийняття рішення. Ця міра називається функцією втрат (штрафу) або функцією виграти (цільовою функцією).

Критерій Байеса — правило, у зв'язку з яким стратегія рішень обирається таким чином, щоб забезпечити мінімум середнього ризику. Застосування критерію Байеса раціонально у випадку, коли система розпізнавання багаторазово виконує ідентифікації невідомих об'єктів чи явищ, в умовах ознакового та незмінного простору, у оптимальному та стабільному описі класів, та незмінній платіжній матриці, що в свою чергу,

задовольняє вимогам ідентифікації людини за біометричними характеристиками голосу людини.

Мінімум ризику, середній по множині альтернатив вирішення задачі ідентифікації незнайомих об'єктів, забезпечуються тоді, коли рішення про належність об'єктів класу Ω_1 чи Ω_2 приймаються відповідно з наступним правилом : якщо отримане значення ознаки у даного об'єкта розміщене в області R_1 , то об'єкт стосується класу Ω_1 , якщо в області R_2 , то об'єкт відноситься до класу Ω_2 .

Стратегію, базовану на цьому правилі, називають байєсівською стратегією, а мінімальний середній ризик — байєсівським ризиком.

Нехай у результаті дослідження буде встановлено, що значення ознаки ω у предметі, а що ідентифікують, становить величину $x = x^0$. Тоді умовна ймовірність приналежності об'єкту до класу Ω_1 (умовна ймовірність першої гіпотези відповідно до теореми гіпотез).

$$P\left(\frac{\Omega_1}{x^0}\right) = \frac{P(\Omega_1)f_1(x^0)}{f_1(x^0)},$$

а умовна ймовірність того, що об'єкт належить до класу Ω_2 (умовна ймовірність другої гіпотези)

$$P\left(\frac{\Omega_2}{x^0}\right) = \frac{P(\Omega_2)f_2(x^0)}{f_2(x^0)},$$

де $f(x^0) = P(\Omega_1)f_1(x^0) + P(\Omega_2)f_2(x^0)$ — це спільна щільність розподілу ймовірностей.

$P\left(\frac{\Omega_1}{x^0}\right)$ та $P\left(\frac{\Omega_2}{x^0}\right)$ — це апостеріорні ймовірності дійсної приналежності об'єкта що розпізнається до класів Ω_1 та Ω_2 . Існує умовний ризик, пов'язаний з рішенням $\omega \in \Omega_1$

$$P\left(\frac{\Omega_1}{x^0}\right) = c_2 P\left(\frac{\Omega_2}{x^0}\right),$$

а умовний ризик, який стосується $\omega \in \Omega_2$,

$$P\left(\frac{\Omega_2}{x^0}\right) = c_2 P\left(\frac{\Omega_1}{x^0}\right).$$

Отже, це означає, що система ідентифікації, що базується на критерії Байеса, має вирішувати завдання з найменшим умовним ризиком. Звідси маємо, що використовувати $\omega \in \Omega_1$, необхідно за умови :

$$\frac{R\left(\frac{\Omega_1}{x^0}\right)}{R\left(\frac{\Omega_2}{x^0}\right)} < 1$$

та використовувати $\omega \in \Omega_2$:

$$\frac{R\left(\frac{\Omega_2}{x^0}\right)}{R\left(\frac{\Omega_1}{x^0}\right)} < 1.$$

Використовуючи байєсівські правила прийняття рішення розрахуємо значення цільності розподілу для всіх еталонних елементів бази даних $f(Sp_{\text{еталон}_n})$ та значення щільності розподілу для вхідного сигналу, що підлягає ідентифікації $f(Sp_{\text{вх}})$ за формулою:

$$\frac{(Sp_{\text{еталон}}(S(t)_n) - Sp_{\text{вх}})^2}{2\sigma_x^2} f_n(Sp_{\text{еталон}}) = \frac{1}{\sigma_x \sqrt{2\pi}} e.$$

Отримаємо значення $f_1(Sp_{\text{еталон}_1}), f_2(Sp_{\text{еталон}_2}), \dots, f_n(Sp_{\text{еталон}_n})$ та розрахуємо апостеріорні ймовірності приналежності вхідного сигналу $S_{\text{вх}}(t)$ до класу об'єктів $DB = \{S_1(t), S_2(t), \dots, S_n(t)\}$.

Таким чином, розроблений метод дозволяє ідентифікувати людину за голосом в умовах навіть дуже високого шуму в каналі зв'язку.

Висновки

Розроблено метод ефективної ширини спектру на базі багатоальтернативних правил прийняття рішення шляхом порівняння вхідного сигналу $Sp_{\text{вх}}$ та всіх еталонних $Sp_{\text{еталон}_n}$ об'єктів бази даних за визначеним простором ознак для ідентифікації диктора за біометричними характеристиками в умовах високого рівня шуму в каналі зв'язку.

ЛІТЕРАТУРА

1. Юдін О. К., Зюбіна Р. В. Аналіз сучасних систем та методів розпізнавання аудіосигналів у задачах ідентифікації та верифікації // Проблеми інформатизації та управління: зб. наук. праць. — К. : НАУ, 2017. — Вип. 3(59). — С. 75–79.
2. Юдін О. К., Зюбіна Р. В. Визначення мінімально достатньої кількості інформації для різного класу біометричних параметрів об'єктів ідентифікації та/або аутентифікації // Вісник Інженерної академії України. — 2017. — Вип. 3. — С. 129–132.
3. Юдін О. К., Зюбіна Р. В. Класифікація методів ідентифікації частоти основного тону // Наукоємні технології. — 2017. — Т. 33. — №. 1. — С. 13–21.
4. Юдін О. К., Зюбіна Р. В. Оцінка ефективності методів ефективної ширини спектру та найбільшої інформаційної ваги основного тону в задачах ідентифікації та аутентифікації аудіосигналів // Наукоємні технології. — 2017. — Т. 35. — №. 3. — С. 209–214.
5. Дворкович В. П., Дворкович А. В. Окonné функции для гармонического анализа сигналов // М. : Техносфера, 2014. — 112 с.
6. Опенгейм А. В., Шафер Р. В. Цифровая обработка сигналов. — М.: Радио и связь, 1979. — 416 с.

Юдін О. К., Зюбіна Р. В.
МЕТОД ЕФЕКТИВНОЇ ШИРИНИ СПЕКТРУ

На основі проведеного аналізу сучасних методів ідентифікації диктора, було визначено, що на сьогодні немає систем, які б визначали особистість в умовах наявного високого рівня завад в каналі передачі даних, що робить дане дослідження досить актуальним. У роботі розроблено метод ефективної ширини спектру на базі багатоальтернативних правил прийняття рішення за визначеним класом біометричних ознак голосу для підвищення ефективності функціонування розподілених інформаційних систем. Використання розробленого методу дає можливість проводити ідентифікацію диктора навіть у випадку, коли шум перевищує інформаційний сигнал у 1,5 рази.

Ключові слова: біометрія; ідентифікація; прийняття рішення; диктор, інформаційна система.

Yudin O. K., Ziubina R. V.
METHOD OF EFFECTIVE SPECTRUM WIDTH

Based on the analysis of modern methods of speaker identification, it was determined that to date there are no systems which would define speaker identity in conditions of high level of interference in the data channel, which makes this research very relevant. In this paper, we developed the Method of Effective Spectrum Width based on of the multialternative decision rules for a particular class of biometric characteristics to improve the function efficiency of distributed information systems. Using the developed method enables identify of the speaker even in the case where the information noise exceeds the signal by 1,5 times.

Keywords: biometrics; identification; decision making; speaker, information system.

Юдин А. К., Зюбина Р. В.
МЕТОД ЭФФЕКТИВНОЙ ШИРИНЫ СПЕКТРА

На основе проведенного анализа современных методов идентификации диктора, было определено, что сегодня не существует систем, которые бы определяли личность в условиях высокого уровня помех в канале передачи данных, что делает данное исследование весьма актуальным. В работе разработан метод эффективной ширины спектра на базе многоальтернативных правил принятия решения по определенным классам биометрических признаков голоса для повышения эффективности функционирования распределенных информационных систем. Использование разработанного метода позволяет проводить идентификацию диктора даже в случае, когда шум превышает информационный сигнал в 1,5 раза.

Ключевые слова: биометрия; идентификация; принятие решения; диктор, информационная система.

Стаття надійшла до редакції 28.02.2018 р.
Прийнято до друку 28.02.2018 р.
Рецензент — д-р техн. наук, проф. Бараннік В. В.