

## Висновки

У роботі\* запропоновано метод узагальненого тригонометричного перетворення, що ґрунтується на періодичних Ateb-функціях. Розглядаються властивості ортогональних тригонометричних перетворень Фурє та Хартлі. Аналогічно доводяться властивості лінійності, симетрії та подібності для введеного ортогонального Ateb-перетворення.

\* Роботу виконано у межах договору ДЗ 465-2011 “Створення та освоєння інформаційної технології захистуцифрованих паперів на основі нових методів цифрової обробки графічної інформації” з державним номером реєстрації 0111U009029.

1. Яцимірський М.М. Швидкі алгоритми ортогональних тригонометричних перетворень. – Львів: Академічний експрес, 1997. – 219 с. 2. Добеши І. Десять лекцій по вейвлетам. Москва, “РХД”, 2001 г. 3. Сенік П. М. Про Ateb-функції / П. М. Сенік // Докл. АН УРСР, сер. А. — 1968. — № 1. — С. 23 – 27. 4. Сенік П. М. О табулюванні періодической Ateb-функції / П. М. Сенік, А. М. Возний // Докл. АН УССР. Сер. А. – 1969. – № 12. – С. 1089 – 1092. 5. Сокіл Б. І. Нелінійні коливання механічних систем і аналітичні методи їх досліджень: Автореф. дис. д-ра техн. наук: 05.02.09 / Національний ун-т “Львівська політехніка”. – Львів, 2001. – 36 с. 6. Грицик В. В. Математичні моделі алгоритмів і реалізація Ateb-функцій / В. В. Грицик, М. А. Назаркевич // Доповіді НАН України. – 2007. – № 12. – С. 37 – 43. 7. Брейсуэлл Р. Преобразование Хартли: теория и приложения / Р. Брейсуэлл ; пер. А. И. Пашков ; ред. пер. И. С. Рыжак. – М. : Мир, 1990. – 174 с.

УДК 531.36+534

М. Назаркевич, І. Вербенко

Національний університет “Львівська політехніка”,  
кафедра ІТВС

## ІДЕНТИФІКАЦІЯ ДАНИХ КОРЕЛЯЦІЙНИМИ МЕТОДАМИ НА ОСНОВІ АТЕВ-ФУНКЦІЙ

© Назаркевич М., Вербенко І., 2011

Побудовано метод ідентифікації даних на основі Ateb-функцій для встановлення відповідності між розпізнаним документом та його електронним еталоном. Ідентифікація інформації здійснювалася на основі кореляційних методів. Розроблено програмне забезпечення для здійснення ідентифікації.

**Ключові слова:** Ateb-функції, ідентифікація інформації, кореляційні методи.

**We construct a method of identification based on data Ateb-functions for mapping between the recognized document and its electronic standard. Identifying information based on correlation methods. Is developed software for identification.**

**Keywords:** Ateb-functions, identifying information, correlation methods.

### Вступ

У ХХІ ст. стрімко розвиваються такі напрями інформаційної діяльності, як телекомунікації, інтернет-технології та мультимедіа. Компанії переходять на безпаперові технології, перевагою яких є швидке зберігання документів, їх пересилання та опрацювання. Електронну інформацію зберігають у сховищах або базах даних, де існує простий та швидкий доступ до документів. Незважаючи на це, в установах, де працюють із друкованими документами, весь час виникає потреба знаходити потрібну інформацію – для внесення даних, перевірки їх правильності або

коригування. Прикладами таких документів є акти на володіння землею, податкові та митні декларації, страхові поліси, документи про освіту тощо. Більшість таких документів потребують певного захисту та ідентифікації.

Ідентифікація проводиться з метою порівняння документів, розпізнавання, встановлення відповідності між розпізнаним документом і його електронним еталоном. У процесі ідентифікації визначаються відповідності між друкованим документом і його зразком, які характеризуються сукупністю певних ознак.

Існують загальноживані методи для ідентифікації документів в процесі документообігу. В США і Канаді застосовується штрих-код *UPC*, на основі якого в Європі створено міжнародну систему товарної нумерації *EAN*. Система *EAN* належить до міжнародного стандарту ідентифікації товарів та послуг на основі штрихових кодів та електронного обміну даними *EDI (Electronic Data Interchange)*. За цим стандартом кодова величина виражається у восьми (*EAN-8*) або тринадцяти (*EAN-13*) цифрах [1]. Недоліками цього стандарту є:

- штрихові коди взагалі не піддаються зчитуванню, коли у процесі друку виникає занадто велике розтискування фарби, що призводить до розширення штрихів і до звуження проміжків між ними;
- немає встановлених правил щодо присвоєння та реєстрації коду товару чи послуги. Тобто за присвоєним кодом може надавати інформацію лише підприємство-виробник;
- неприпустимим є нанесення на штриховий код, зокрема, на зони його стабілізації, чорнильних штампів або будь-яких інших позначок;
- при скануванні штрих-кодів сканер не може визначити повторно відсканований один і той самий код.

У роботі запропоновано новий оригінальний спосіб ідентифікації, який є простим у застосуванні, не вимагає спеціального додаткового апаратного обладнання і може використовуватися з метою встановлення автентичності документів.

### Математичний апарат *Ateb*-функцій для побудови графічних примітивів

Сучасні програми для опрацювання векторної графіки застосовують графічні примітиви для побудови зображень. До графічних примітивів належать коло, еліпс, прямокутник та багатокутник. У цій роботі графіки функцій, задані аналітично, застосовуються для побудови графічних примітивів.

Введемо у розгляд математичний апарат *Ateb*-функцій, який дає можливість отримувати аналітичні розв'язки  $(x(t), y(t))$  систем диференціальних рівнянь, поданих у вигляді [2]

$$\begin{cases} \frac{dx}{dt} + by^m = 0 \\ \frac{dy}{dt} + ax^n = 0 \end{cases}, \quad (1)$$

де  $a, b$  – деякі дійсні сталі, а  $n, m$  задовольняють умови

$$n = \frac{2q'_1 + 1}{2q''_1 + 1}, \quad m = \frac{2q'_2 + 1}{2q''_2 + 1}, \quad (q'_1, q''_1, q'_2, q''_2 = 0, 1, 2, \dots) \quad (2)$$

*Ateb*-функції є оберненням до *Beta*-функцій [3].

Періодичності *Ateb*-функцій досягають, якщо параметри  $n, m$  задовольняють умову (5)

$$\frac{1}{n+1} > 1, \quad \frac{1}{m+1} > 1; \quad (3)$$

У такому випадку система диференціальних рівнянь (1), описує коливальний рух [3].

Введемо у розгляд вираз (4), позначивши  $w$  як функцію від параметрів  $n, m$

$$w = \frac{1}{2} \int_0^{-1 \leq y \leq 1} t^{-\frac{n}{n+1}} (1-t)^{-\frac{m}{m+1}} dt \quad (4)$$

Для побудови *Ateb*-функції синусу  $sa(n, m, w)$  розглядаємо обернену залежність  $y$  від  $w$ , що є одночасно функцією  $n$  і  $m$

$$y = sa(n, m, w) \quad (5)$$

Аналогічно вводимо у розгляд *Ateb*-функцію косинусу  $ca(m, n, w)$ .

З властивостей періодичності [3] було виведено співвідношення

$$ca^{m+1}(m, n, w) + sa^{n+1}(n, m, w) = 1 \quad (6)$$

З (6) можна отримати функцію косинуса  $ca(m, n, w)$ , якщо заданий синус і навпаки.

Опишемо властивості періодичних *Ateb*-функцій, які важливі для побудови графічних примітивів.

*Ateb*-синус  $y = sa(n, m, w)$ . Неперервна крива з періодом  $T = 2\Pi(m, n)$ . Перетин з віссю  $w$  в точках  $k\Pi(m, n)$ ,  $k = 0, \pm 1, \pm 2, \dots$ . У цих точках дотична має кут  $g = \pm \arctg \frac{2}{n+1}$  нахилу до осі  $w$ . Екстремумів досягають в точках  $(k + \frac{1}{2}) \cdot \Pi(m, n)$  і відповідно рівні  $(-1)^k$ .

*Ateb*-косинус  $y = ca(m, n, w)$ . Неперервна крива з періодом  $T = 2\Pi(m, n)$ . Перетин з віссю  $w$  в точках  $(k + \frac{1}{2}) \cdot \Pi(m, n)$ ,  $k = 0, \pm 1, \pm 2, \dots$  відбувається під кутом  $g = \pm \arctg \frac{2}{m+1}$ . Екстремумів досягають в точках  $k\Pi(m, n)$ ,  $k \in Z$  і відповідно рівні  $(-1)^k$ .

За допомогою математичного апарату *Ateb*-функцій створюємо бібліотеку графічних примітивів, на основі яких будемо проводити ідентифікацію друкованих документів. Різновиди еталонних графіків періодичних *Ateb*-функцій будуть залежати від різноманітності вибору значень параметрів  $m$  і  $n$ .

Обчислюємо  $m$  і  $n$  за формулами (2) з врахуванням (3), кількість яких залежить від  $q'_1, q''_1, q'_2, q''_2 \in N \cup \{0\}$ , де  $N$  – множина натуральних чисел. Таким чином можна створити як завгодно велику кількість графічних примітивів.

На основі формул (4) та (6) обчислюємо значення *Ateb*-функцій. Застосовуємо обчислення з подвійною точністю, що дає можливість стверджувати про точність обчислень порядку  $e = 10^{-10}$ . Така точність представлення цифрових файлів дасть можливість провести якісну ідентифікацію.

Підбираючи різні значення  $n$  і  $m$ , описаних вище, створюємо бібліотеку графічних примітивів. У роботі наведено шість графічних примітивів з розробленої бібліотеки, які показано на рис. 1. Поряд з графічними примітивами вказані числові значення  $w$  та  $y$  *Ateb*-функцій, які обчислені з (4). З наведених рисунків можна побачити, що форма та розмір кожного графічного примітиву є різною.

### Етапи процесу ідентифікації друкованого документу

Розглянемо технологічну послідовність запропонованого методу ідентифікації (рис. 2), який полягає у виокремленні одиничного графічного примітиву зі сканованого зображення, його нормуванні та порівнянні з певним еталонним зображенням. Для реалізації поставленої цілі необхідно побудувати еталонні графічні примітиви за допомогою *Ateb*-функцій, які будуть наноситись на документ у вигляді захисної сітки. На основі унікальних для кожного графічного примітиву параметрів буде відбуватися ідентифікація документа. Крім того необхідно спеціальним чином обробити скановане зображення, щоб виділити графічний елемент для порівняння з еталонним.

*Сканування зображення.* Процес ідентифікації починається з сканування друкованого документу. Зображення документу з захисною сіткою має бути скановане в режимі градацій сірого з роздільною здатністю 300 dpi.

Вхідне растрове зображення у цифровому вигляді підлягає попередньому опрацюванню внаслідок присутності характерних для сканованих зображень завад. Програма за певним алгоритмом автоматично перетворює вхідне зображення з режиму градацій сірого на режим чорного і білого. Результат зображено на рис. 3.

a) $m=0,01; n=0,1$			б) $m=1,3; n=1$		
	$\omega$	Y		$\omega$	Y
	0	0		0	0
	0.1	0.1822739424		0.1	0.0997829244
	0.2	0.3618369807		0.2	0.1992899384
	0.3	0.5472236808		0.3	0.2976918446
	0.4	0.7257083364		0.4	0.3945261688
	0.5	0.9057950637		0.5	0.4892490963
0.56	1	0.6	0.5811187920		
в) $m=1,7; n=1$			г) $m=1; n=1,3$		
	$\omega$	Y		$\omega$	Y
	0	0		0	0
	0.1	0.1156789041		0.05	0.0872214946
	0.2	0.2001915239		0.1	0.1561575045
	0.3	0.3052659404		0.15	0.2145016524
	0.4	0.3932688757		0.2	0.2829709021
	0.5	0.4914367401		0.25	0.3585748864
0.6	0.5937697877	0.3	0.4131626228		
д) $m=7; n=1$			е) $m=3; n=1,7$		
	$\omega$	Y		$\omega$	Y
	0	0		0	0
	0.1	0.1755735615		0.1	0.1715768421
	0.2	0.1755735615		0.2	0.3132964289
	0.3	0.2999893323		0.3	0.4517585124
	0.4	0.3820681240		0.4	0.5476376320
	0.5	0.4460062106		0.5	0.6438603865
0.6	0.5443026013	0.6	0.7228811012		

Рис. 1. Приклади графічних примітивів з бібліотеки, створених за допомогою математичного апарату Ateb-функцій



Рис. 2. Етапи ідентифікації друкованого документа, захищеного на основі Ateb-функцій

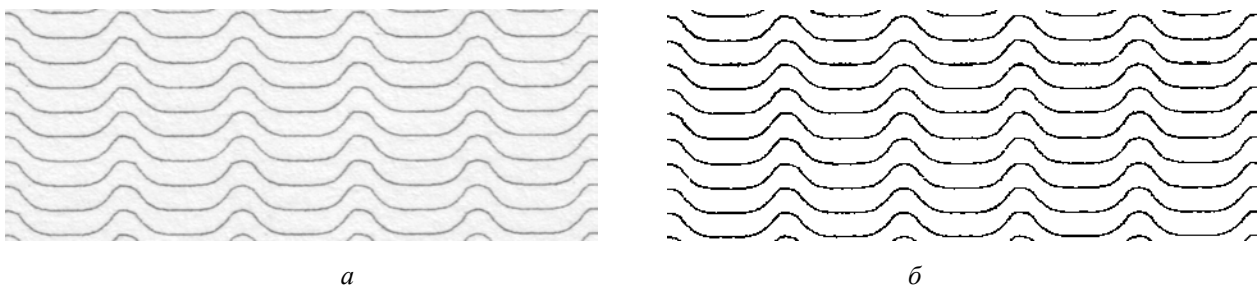


Рис. 3. Фрагмент сканованого документа у режимі Grayscale (a) та його вигляд після попереднього опрацювання (б)

*Скелетизація кривих.* Продовженням попереднього опрацювання вхідного растру є сегментація. В нашому випадку сегментація – це виділення складових частин зображення, тобто кривих. Розпізнавання проводилось також з використанням алгоритму скелетизації [6]. Для цього проведено сегментацію в окремі кластери та виділено границі кривих (рис. 4 а).

*Виокремлення періоду та амплітуди кривих.* Знаходимо період сканованої кривої  $T$  і амплітуду  $2A$  [6]. За допомогою цих двох параметрів виокремлюємо одиничний графічний примітив зображення. Результат наведено на рис. 4 б.

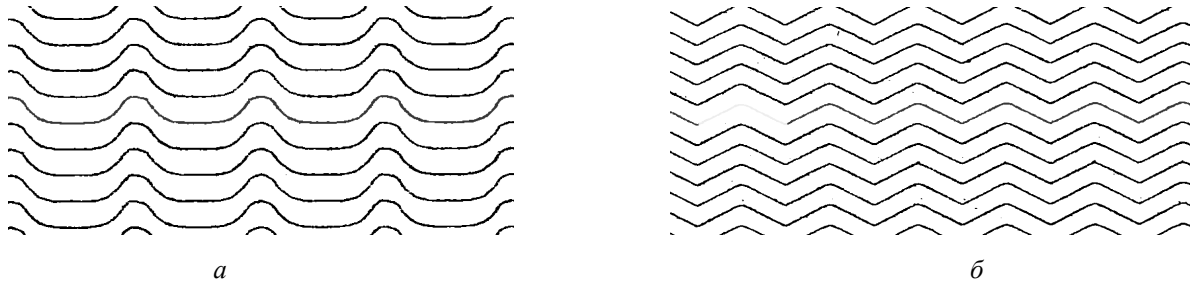


Рис. 4. Сегментація сканованого документа – виділення контурів кривих (а) та виділення сегменту з кривої (б)

*Паралельне перенесення системи координат* – формування простору ознак – паралельного перенесення системи координат у точку максимуму по осі  $X$  і точку  $0$  відносно амплітуди коливань по осі  $Y$ .

Скановане зображення на проміжку від  $0$  до  $2\pi$  розуміємо як вхідне зображення  $P_i$  – простір ознак вхідного зображення. Необхідно зазначити, що нормалізовані зображення відрізняються між собою у величинах порядку  $10^{-6}$ . Тому для якісного порівняння з еталоном потрібно обчислювати *Ateb*-функції з точністю порядку  $10^{-10}$ .

*Порівняння з еталоном за критерієм максимальної кореляції.* На етапі розпізнавання крива, яка вимагає верифікації порівнюється з етальною, побудованою у вигляді періодичних *Ateb*-функцій. Будемо вважати за розпізнавання зображення порівняння вхідного сканованого зображення кривої з отриманим раніше етальним зображенням. Визначимо міру близькості двох зображень  $P_i$   $P_j$  як метрику  $\mu(P_i, P_j)$  у просторі зображень  $G$ . Якщо позначити  $P_0$  етальне зображення і визначити значення міри  $\mu_0 = \mu(P_0, P_j)$  у випадку збігу зображень, то використаний нами кореляційний критерій означає знаходження такого перетворення  $g \in G$ , для якого  $|\mu_0, \mu_g| < \varepsilon$ . Класичною мірою ідентичності є визначення максимуму функції [6]:

$$m_g = \frac{\sum_{(x,y) \in D} P(x,y)P_0(x,y)}{\sqrt{\sum_{(x,y) \in D} P^2(x,y)} \sqrt{\sum_{(x,y) \in D} P_0^2(x,y)}} \rightarrow \max_{g \in G} \quad (7)$$

де  $D$  – поле зору,  $P(x,y)$ ,  $P_0(x,y)$  – вхідне та етальне зображення відповідно.

З (7) випливає, що  $\mu_g$  прямує до одиниці. У випадку збігу зображень  $\mu_0 = 1$ .

Результатом роботи нашої програми є міра близькості  $\varepsilon$  між значеннями  $\mu_0$  і  $\mu_g$  графічних примітивів етального і сканованого зображень, подана у відсотках. Для зручності обчислюємо відносний результат у відсотках:

$$k = \left| 1 - \frac{|\mu_0 - \mu_g|}{\mu_0} \right| \cdot 100\% \quad (8)$$

Вважаємо, що якщо відсоток збігу є більшим за 90, то документ є ідентифікований. У протилежному випадку документ вважаємо не ідентифікованим.

На рис. 5 показано зображення з захисною сіткою, яка має параметри  $ca(0,01;0,1)$ . Сканований у режимі градацій сірого документ було перетворено на режим чорного і білого з метою видалення характерних для сканованих зображень завад. За допомогою алгоритму сегментації зображень було виділено складові частини зображення – криві. Після процесу

знаходження періоду та амплітуди виділеної кривої виокремлюємо її одиничний сегмент – графічний примітив. Паралельне перенесення системи координат у точку максимуму по осі  $X$  і точку  $O$  відносно амплітуди коливань по осі  $Y$  дає можливість визначити значення координат графічного примітиву. На основі визначених значень і значень еталонного графічного примітиву відбувається ідентифікація за критерієм максимальної кореляції. Як видно з рис. 5, сканований документ розпізнано на 98,84%, а це означає, що він є ідентифікований.

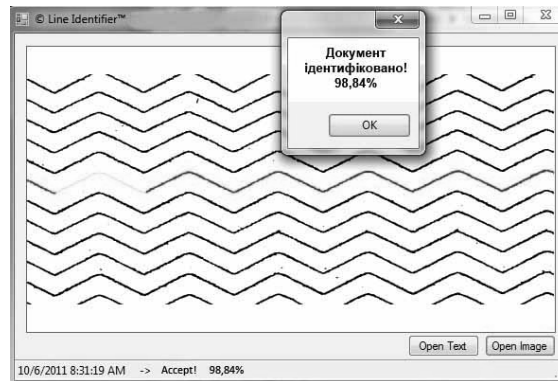


Рис. 5. Інтерфейс програми з прикладом ідентифікації документа з захисною сіткою

### Висновки

Розроблено\* новий метод ідентифікації друкованих документів, який дає змогу підтвердити ідентичність друкованого документа та еталона.

Документи створюються на основі бібліотеки графічних примітивів за допомогою математичного апарату Атеб-функцій, що дає можливість стверджувати про унікальність документів.

Ідентифікують за критерієм максимуму кореляції, у результаті чого обчислюють міру близькості між еталонним зображенням та сканованим документом. Результатом роботи є число у відсотках, яке вказує на подібність між еталоном і сканованим зображенням.

Програмний продукт можна використовувати у різноманітних сферах діяльності: наприклад, для ідентифікації оригінальності актів на володіння землею, різних договорів чи інших цінних паперів.

\*Роботу виконано у межах договору ДЗ 465-2011 “Створення та освоєння інформаційної технології захисту цінних паперів на основі нових методів цифрової обробки графічної інформації” з державним номером реєстрації 0111U009029.

1. ДСТУ 3146-95. Коди та кодування інформації. Штрихове кодування. Маркування об'єктів ідентифікації. Штрихковий позначки EAN. Вимоги до побудови [Текст]. – К.: Держспоживстандарт України, 2006. – 48 с. 2. Грицик В. В. Математичні моделі алгоритмів і реалізація Атеб-функцій / В.В. Грицик, М.А. Назаркевич // Доповіді НАН України. – 2007. – № 12. – С. 37 – 43. 3. Грыцьк В.В. Информационные технологии защиты документов средствами Атеб-функцій. Часть 1. Построение базы данных Атеб-функцій для защиты документов / В. В. Грыцьк, И. М. Дронюк, М. А. Назаркевич // Проблемы управления и информатики. – 2009. – № 2. – С. 139 – 152. 4. Грыцьк В. В. Информационные технологии защиты документов средствами Атеб-функцій. Часть 2. Об одном способе защиты электронных и напечатанных документов / В. В. Грыцьк, И. М. Дронюк, М. А. Назаркевич // Проблемы управления и информатики. – 2009. – № 3 – С. 144 – 153. 5. Пуятин Е.П., Аверин С.И. Обработка изображений в робототехнике. – М.: Машиностроение, 1990. – 320 с. 6. Любченко В.А. Применение одномерной нормализации в задачах распознавания изображений при проективных искажениях // Сб. научн. тр. по материалам 6-го Межд. молодежного форума “Радиоэлектроника и молодежь в ХНІ веке”. Ч.2. – Харьков: ХНУРЕ. – 2002. – С.388–389.