

## ІДЕНТИФІКАЦІЯ ДРУКОВАНИХ ДОКУМЕНТІВ ЗАХИЩЕНИХ НА ОСНОВІ АТЕВ-ФУНКЦІЙ

*The method for identification documents is presented. Protecting documents built using unique Ateb-functions graphic elements. Standard images were constructed for identification. Selection algorithm and normalization is applied to the input image. For comparison with the standard image used maximum correlations method.*

*Розроблено метод для ідентифікації документів. Захист документів будується за допомогою унікальних графічних елементів Атев-функцій. З метою ідентифікації побудовано еталонні зображення. До вхідного зображення застосовано алгоритм скелетизації сканованих даних, виокремлення одиничного графічного елемента та його нормування. Для порівняння з еталоном використано метод максимуму кореляцій.*

### 1. ВСТУП

Значні досягнення в області обчислювальної техніки зумовлюють постійну потребу в створенні нових методів захисту документів, а отже і в розробці способів ідентифікації та методів підтвердження оригінальності документів. Проблема ідентифікації документів, які надруковані поліграфічним способом, тісно пов'язана з задачею розпізнавання зображень.

Базовими для систем обробки графічної інформації є два основних формати представлення зображень - растровий та векторний. Розглянемо у порівнянні растровий і векторний формати представлення зображень та їх особливості при розпізнаванні.

Будь-який фізичний об'єкт в растровому зображенні відображається у вигляді множини пікселів цього растру. Оскільки пікселі не містять інформації про свою належність до певного об'єкту, доводиться опрацьовувати певний об'єм даних растру. Це веде до великого зростання часу опрацьовування при розпізнаванні. Складність растрового розпізнавання пов'язана також з тим, що побудова ознак об'єктів, які б були стійкими до афінних перетворень зображення, пов'язана зі значними труднощами.

---

<sup>1</sup> Національний університет "Львівська політехніка"

Основою векторних зображень є векторні об'єкти. Визначається певна система координат та метрика. Розпізнавання зображення у векторному представленні має такі переваги: можливість групування, сортування, відображення, аналізу об'єктів прошарками; виграш у швидкодії алгоритмів розпізнавання, що пов'язано з можливістю розпізнавання для невеликої частки об'єктів; виграш у швидкодії та достовірності розпізнавання, до яких застосовуються специфічні методи обробки, критерії, бази прототипів та ін.; векторне зображення використовує суттєво менший об'єм пам'яті ЕОМ.

## 2. ПОСТАНОВКА ПРОБЛЕМИ

Проблема захисту поліграфічної продукції від підробки є надзвичайно практично важливою. Це зумовлює актуальність розробки методів визначення аутентичності друкованих документів. Авторами раніше було запропоновано метод захисту інформації [1], зокрема друкованих документів, на основі теорії Атеб-функцій. Проте проблема ідентифікації не розглядалася. Водночас, очевидно, що разом з надійним методом захисту необхідно розробити відповідну процедуру ідентифікації. Тому була запропонована задача створення методу ідентифікації для друкованих документів, що захищені на основі теорії Атеб-функцій.

Поставлена задача вимагає подальшого наукового дослідження, а також розробки відповідного методу оцінки достовірності поліграфічного документу. Результати роботи авторів над розв'язанням цієї проблеми представлено у даній статті.

## 3. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

На сучасному етапі розвитку інформаційних технологій широке розповсюдження отримали системи введення інформації через різноманітні типи сканерів, цифрові фото- та відеокамери. Коректна робота таких систем вимагає інтелектуального аналізу зображень [2]. Аналогічні задачі виникають у процесі ідентифікації поліграфічної продукції. Розпізнавання зображень - це широкий клас задач, моделі та методи для яких останнім часом стрімко розвиваються [3].

Мета та алгоритми сегментації залежать від алгоритмів векторизації та розпізнавання, що мають бути застосовані до даного зображення. Найчастіше критерієм векторизації є колір (яскравість, або ін.) пікселів, тобто об'єктом називають зв'язану множину пікселів однакового кольору (кластерів) при тому, що суміжні з цією множиною пікселі мають інші кольори.

Від якості попередніх етапів оброблення зображення залежать хід та результат автоматичної векторизації. В ідеальному випадку множина векторних об'єктів повинна співпадати з множиною реальних об'єктів на зображенні. Існує тільки один класичний метод в теорії обробки зображень, який працює з проєктивною моделлю зору, це кореляційний метод [4]. Саме цей метод використано для ідентифікації зображення у даній статті.

#### 4. ФОРМУВАННЯ ЦІЛЕЙ СТАТТІ

Метою даної роботи є побудова методу ідентифікації для друкованих та захищених документів на основі класичного методу максимуму кореляції. Для запропонованого раніше методу захисту документів [1] на основі теорії Ateb-функцій пропонується метод ідентифікації, що полягає у виокремленні одиничного графічного елемента зі сканованого зображення, його нормуванні та порівнянні його з певним еталонним зображенням. Для реалізації поставленої цілі необхідно побудувати еталонні зображення. Крім того необхідно спеціальним чином обробити скановане зображення, щоб виділити зображення для порівняння з еталоном. Для цього застосовуємо спеціальні алгоритми з теорії розпізнавання зображень, а саме сегментації, скелетизації та нормування. На завершальному етапі необхідно провести оцінку достовірності вхідного зображення шляхом порівняння з еталоном. Розробці описаного методу присвячена дана стаття.

#### 5. ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

##### 5.1. Побудова еталонних зображень у вигляді періодичних Ateb-функцій

Періодичними Ateb-функціями є  $u = ca(m, n, \omega)$ , які введені в [5]. Доведено, що Ateb-функції мають період  $\Pi(m, n)$ . Отже при необхідності обчислення функцій достатньо провести для аргумента  $\omega$ , що змінюється на відрізьку

$$0 \leq \omega \leq \frac{1}{2} \Pi(m, n). \quad (1)$$

Використовуючи властивості періодичності, можна продовжити отримані значення на весь період. Обчислено значення Ateb – косинусу  $ca(m, n, \omega)$  і сформовано табл.1 та табл.2-3.

Обчислені значення будуть еталонами вхідного зображення (прос-тір ознак вхідного зображення  $P_0$  для задачі ідентифікації).

Таблиця 1

Значення періодів Ateb-функцій

$2\Pi(1,1) = 6,28318556$	$2\Pi\left(1, \frac{1}{5}\right) = 4,47671031$
$2\Pi\left(\frac{1}{3}, 1\right) = 4,79242125$	$2\Pi\left(3, \frac{1}{7}\right) = 8,39009293$

Таблиця 2

Значення Ateb-функцій та значення нормованої Ateb-функції  
 $ca(1,1,\omega)$  та  $ca(1,1/5,\omega)$ 

$\omega$	$ca(1,1,\omega)$	Нормована $ca(1,1,\omega)$	$ca(1,1/5,\omega)$	Нормована $ca(1,1/5,\omega)$
0	1	0	1	0
0.1	0.99500415	0,09994930	0.94140625	0,14028158
0.2	0.98006657	0,19989860	0.86328125	0,28056316
0.3	0.95533648	0,29984790	0.78515625	0,42084474
0.4	0.92106101	0,39979720	0.69921875	0,56112632
0.5	0.87758258	0,49974650	0.60546875	0,70140791
0.6	0.82533559	0,59969580	0.51171875	0,84168949
0.7	0.76484218	0,69964510	0.41796875	0,98197107
0.8	0.69670668	0,79959440	0.31640625	1,12225265
0.9	0.62160996	0,89954370	0.22265625	1,26253423
1.0	0.54030231	0,99949300	0.12109375	1,40281581
1.1	0.45359615	1,09944230	0.01953125	1,54309739
1.2	0.36235777	1,19939160	0	1,68337897
1.3	0.26749882	1,29934090		
1.4	0.16996715	1,39929020		
1.5	0.16996715	1,49923950		

Таблиця 3

Значення Ateb-функцій та значення нормованої Ateb-функції  
 $ca(1/3,1,\omega)$  та  $ca(3,1/7,\omega)$ 

$\omega$	$ca(1/3,1,\omega)$	Нормована $ca(1/3,1,\omega)$	$ca(3,1/7,\omega)$	Нормована $ca(3,1/7,\omega)$
0	1	0	1	0
0.2	0.97006226	0,26208047	0.92578125	0,14970037
0.4	0.88082886	0,52416093	0.83984375	0,29940074
0.6	0.73416138	0,78624140	0.74609375	0,44910110
0.8	0.53359985	1,04832187	0.64453125	0,59880147
1.0	0.28463745	1,31040234	0.55078125	0,74850184
1.2	0	1,57248280	0.44921875	0,89820221
1.4			0.34765625	1,04790258
1.6			0.24609375	1,19760294

1.8			0.15234375	1,34730331
2.0			0.05078125	1,49700368

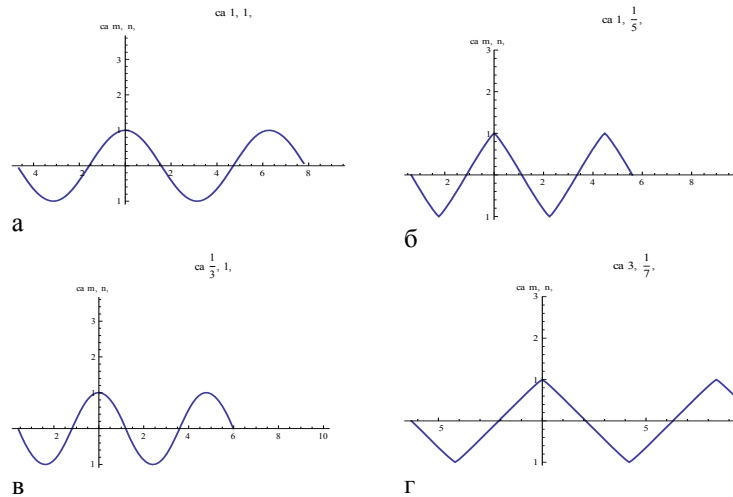


Рис. 1. Графічне представлення *Ateb*-функцій а)  $ca(1, 1, \omega)$ ; б)  $ca(1, 1/5, \omega)$ ; в)  $ca(1/3, 1, \omega)$ ; г)  $ca(3, 1/7, \omega)$

Графіки представлені на рис. 1. та цифрові еквіваленти в табл. 2-3 використані для побудови еталонних зображень процесу ідентифікації. Встановлено, що одиничні графічні елементи, які будуть використовуватися як захисні елементи у процесах додрукарської підготовки, можна представити математичними моделями, які описуються певним класом систем диференціальних рівнянь. Розроблено математичний апарат для знаходження точних розв'язків цих моделей у вигляді *Ateb*-функцій. Вперше запропоновано застосувати цей апарат до побудови ефективних методів захисту інформації на етапі процесу додрукарської підготовки.

На основі отриманих графіків було сформовано захисну сітку товщиною 0,25 pt. Сітку сформовано на основі одиничного графічного елемента *Ateb*-косинусу  $ca(1/3, 1, \omega)$  та здійснено його копіювання. Було реалізовано розмноження одиничного графічного елемента за законом паралельного перенесення та повороту. Роздруковано документ з захисною сіткою, який використано для сканування та представлено на рис. 3.

## 5.2. Етапи процесу ідентифікації друкованого документу

Розглянемо технологічну послідовність запропонованого методу ідентифікації, яка показана на рис. 2. Вхідне растрове зображення у цифровому вигляді підлягає попередньому опрацюванню. Вважаємо при цьому, що лінії захисної сітки не перетинаються і сформовані з Ateb-функцій лише за допомогою афінних перетворень.

**Сканування зображення** Процес ідентифікації починаємо з оцифрування зображення. Беремо зображення захищеного документу, яке скановане в режимі градацій сірого.

Вхідне растрове зображення у цифровому вигляді підлягає попередньому опрацюванню або сепарації внаслідок присутності характерних для сканованих зображень завад. Для цих зон проводиться подальше опрацювання. Результат представлений на рис. 3.

**Скелетизація кривих.** Продовженням попереднього опрацювання вхідного растру є сегментація. В нашому випадку сегментація - це процес формування в інформативних зонах кластерів, які відповідають кривим. Розпізнавання проводилось також з використанням алгоритму скелетизації [2]. Для цього проведено сегментацію в окремі кластери та виділено границі кривих.

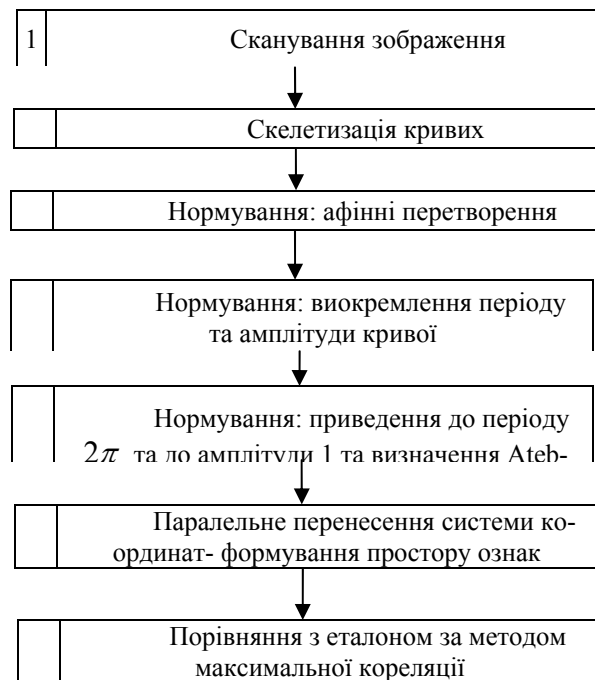


Рис. 2. Етапи ідентифікації друкованого документа, захищеного на основі Атеb-функції

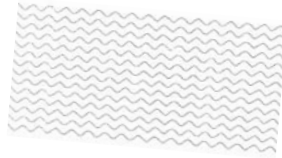


Рис. 3. Скановане зображення захисної сітки

Лінії сортуємо за довжиною, товщиною (зберігається як властивість лінії після скелетизації), орієнтацією, кутом повороту та іншими параметрами.

**Афінні перетворення.** Для подальшого розпізнавання та ідентифікації отримані зображення підлягають нормалізації. Спочатку реалізуємо афінне перетворення повороту. Розташовуємо осі координат паралельно прямій, що містить максимуми кривих.

**Виокремлення періоду та амплітуди кривої.** Виокремлюємо одиничний графічний елемент зображення. Знаходимо період сканованої кривої  $T$  та амплітуду  $2A$ . Результати представлені на рис.4.

**Приведення до періоду  $2\pi$  та до амплітуди 1 та визначення Атеb-функції.** Наступний етап: нормування. На цьому етапі сканована крива перетворюється до періоду  $2\pi$  та амплітуди 1. Результат представлено на рис.5.

**Паралельне перенесення системи координат – формування простору ознак.** Переходимо до наступного етапу – 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.

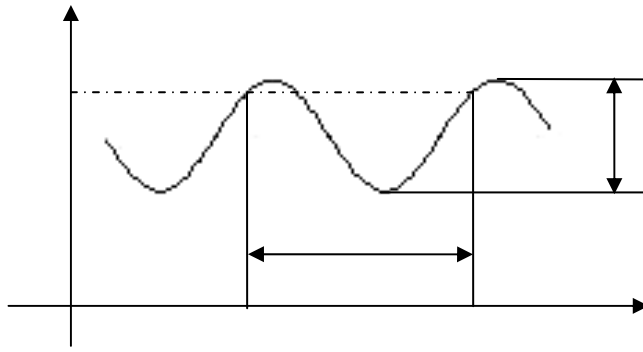


Рис. 4. Виокремлення одиничного графічного елемента

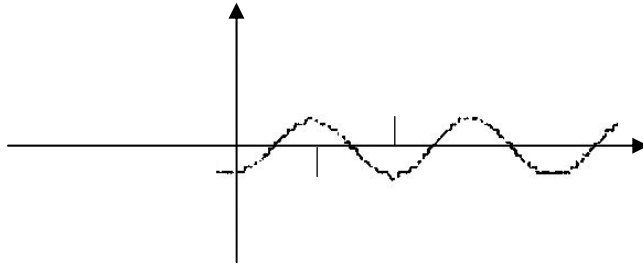


Рис. 5. Приведення кривої до періоду  $2\pi$  та до амплітуди 1

**Порівняння з еталоном за методом максимальної кореляції.** На етапі розпізнавання крива, який вимагає верифікації порівнюється з еталоною, побудованою в п.4.1. Будемо вважати за розпізнавання зображення порівняння вхідного сканованого зображення кривої з отриманими раніше еталоновими зображеннями. Визначимо міру близькості двох зображень  $P_i P_j$  як метрику  $\mu(P_i, P_j)$  у просторі зображень  $G$ . Якщо позначити  $P_0$  еталонове зображення і визначити значення міри  $\mu_0 = \mu(P_0, P_j)$  у випадку співпадіння зображень, то використаний нами кореляційний метод [2] означає знаходження такого перетворення  $g \in G$ , для якого  $|\mu_0, \mu_g| < \varepsilon$ . Класичною мірою ідентичності є визначення максимуму функції [4]

$$\mu_g = \frac{\sum_{(x,y) \in D} P(x,y)P_0(x,y)}{\sqrt{\sum_{(x,y) \in D} P^2(x,y)} \sqrt{\sum_{(x,y) \in D} P_0^2(x,y)}} \rightarrow \max_{g \in G} \quad (4)$$



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

З (4) випливає, що  $\mu_g$  прямує до одиниці. У випадку співпадіння зображень  $\mu_0 = 1$ .

## 6. ВИСНОВКИ

Здійснення ідентифікації друкованих документів проведено для векторних об'єктів. Розроблено та приведено етапи методу ідентифікації.

Побудову захисних сіток здійснено на основі Ateb-функцій, що значно розширює спектр одиничних графічних елементів. Приведено числові значення та графічне представлення Ateb-функцій, побудовано еталонні зображення.

Ідентифікацію та порівняння з еталоном здійснено методом максимальної кореляції.

1. Грицик В. В. Метод захисту та відтворення інформації засобами Ateb-функцій / В. В. Грицик, І. М. Дронюк, М. А. Назаркевич. // Доповіді НАН України. — 2008. — № 5. — С.48 – 52. 2. Пуятин Е.П., Аверин С.И. Обработка изображений в робототехнике. -М.: Машиностроение, 1990. – 320 с. 3. Shalkoff R.J. Digital image processing and computer vision. – New York-Chichester-Brisbane-Toronto-Singapore: John Wiley & Sons, Inc., 1989. — 489 p. 4. Любченко В.А. Применение одномерной нормализации в задачах распознавания изображений при проективных искажениях // Сб. научн. тр. по материалам 6-го Межд. молодежного форуму «Радиоэлектроника и молодежь в XXI веке». Ч.2. - Харьков: ХНУРЭ. - 2002. С.388-389. 5. Сенник П. М. Численное обращение одного класса неполной Beta-функции / П. М. Сенник, А. М. Возный. // Математическая физика. — 1973. — № 14. — С. 160 – 164.