

## ОСОБЛИВОСТІ СИСТЕМ ТА МЕТОДІВ СТРУКТУРНОГО РОЗПІЗНАВАННЯ ТЕКСТОВИХ ЗОБРАЖЕНЬ

Б.М. Гавриш<sup>1</sup>, к.т.н., ст. викладач

О.В. Тимченко<sup>1,2</sup>, д.т.н., професор

Ю.О. Борзов<sup>3</sup>, к.т.н. ст. викладач

О.О. Тимченко<sup>1</sup>, аспірант

<sup>1</sup>Українська академія друкарства

<sup>2</sup>Uniwersytet Warmińsko-Mazurski w Olsztynie, Polska

<sup>3</sup>Львівський Державний університет безпеки життєдіяльності

***Анотація.** Розглядаються особливості побудови систем оптичного розпізнавання текстів, методи формування ознак зображень та метод побудови граматики для структурного розпізнавання текстових зображень.*

***Ключові слова:** систем оптичного розпізнавання текстів, структурне розпізнавання*

### **Вступ**

Оптичне розпізнавання образів є одним з завдань, які людина вирішує неперервно. При цьому з органів зору надходить інформація, що обробляється мозком, який забезпечує прийняття рішення про об'єкт розпізнавання. Однією з важливих областей застосування розпізнавання, чи класифікації, є розпізнавання друкованого або рукописного текстів. З розвитком інформаційних технологій стало можливим полегшити, прискорити та підвищити якість розпізнавання. Першим елементом системи розпізнавання літер є сканерами, що вводять в комп'ютер зображення текстів. Для створення текстового документу, необхідно розпізнати на цьому зображенні окремі символи. Існує цілий ряд програмних продуктів, які практично автоматизували процеси розпізнавання текстів. Проте забезпечити задовільний результат у випадку спотворень зображень текстів різного типу (геометричних, шумів та подібних) вдається не завжди. Проте необхідно зазначити, що здатність людини читати друкований чи рукописний текст низької якості перевершує можливості комп'ютерних систем [1].

Оптичне розпізнавання тексту (англ. optical character recognition, OCR) — це електронне переведення зображень рукописного або друкованого тексту в послідовність кодів, що використовуються для представлення в текстовому редакторі.

Проблема ефективного розпізнавання тексту посідає важливе місце в сферах

інформатизації різних процесів людської діяльності. Текстове представлення інформації, порівняно із графічним, дозволяє істотно скоротити витрати на зберігання та передачу інформації, а також дозволяє реалізувати всі методи використання та аналізу електронних документів. Тому найбільший інтерес з практичної точки зору представляє саме перетворення інформації з паперових носіїв в текстовий електронний документ [1].

### Загальна характеристика систем розпізнавання текстів

Система OCR (наприклад, FineReader, OmniPage, ReadIris чи інші) отримує цифрове зображення сканованого або сфотографованого документа і формує текст, який містить це зображення, у одному з форматів електронних текстових документів. Загальна структурна системи показана на рис. 1 [2].

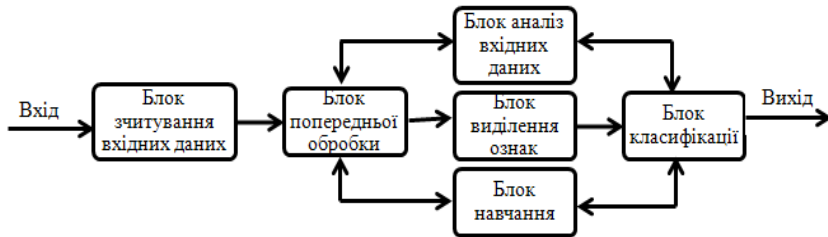


Рис.1. Узагальнена структурна схема системи розпізнавання символів

Призначення блоків системи наступне. Блок зчитування вхідних даних (сканування, фотографування) проводить перетворення енергії світлового сигналу в електричний. Розмір матриці сканування залежить від точності процесу отримання зображення та кольору зображення або його фрагментів [3].

Блок попередньої обробки переводить зображення у чорно біле з метою значно зменшують розміри матриці ознак, економії пам'яті і зменшення витрат на обробку інформації. Також відбувається зменшення рівня шумів у вигляді дрібних ізольованих крапок, ліквідації так званої бахроми, дрібних прогалин (порожнеч) всередині ліній і, якщо необхідно, операції потоншення ліній. Також тут виконуються функції виправлення геометричних спотворень зображень, отриманих в процесі, наприклад, фотографування.

Блок виділення ознак може мати різну складність в залежності від природи зображення, яке розпізнається і використаних методів. Кожен підданий попередній обробці об'єкт повинен бути представлений у вигляді структури мовного типу, наприклад ланцюжка. Цей процес складається з двох підпроцесів: з сегментації і з виділення непохідних елементів. Основною метою процесу сегментації є виділення з зображення окремих елементів для виділення певних ознак або порівняти відмінності зі стандартними зразками в пам'яті. Тому при розпізнаванні друкованого тексту, в якому символи відокремлені один від іншого,

сегментація дає результати, які дозволяють в результаті розпізнавання отримати текст, практично повністю ідентичний оригінальному. Проблеми з точністю процесу сегментації виникають і тоді, коли символи пов'язані з графіками, таблицями чи малюнками.

Блок класифікації формує ознаки відповідності елементів зображення тим чи іншим еталонним зразкам. Оцінка ефективності класифікації залежить від якості виконання попередніх етапів і оцінюється кількістю ознак, за якими буде встановлена ця відповідність. Тому процес класифікації досить часто вимагає повторних циклів аналізу подання вхідного зображення.

Для розпізнавання символів у системах розпізнавання текстів використовуються такі класифікатори: растровий, ознаковий, структурний та їх поєднання. Використовуються вони в залежності від змісту документа, вхідних параметрів зображення і задачі розпізнавання [4].

Растровий класифікатор порівнює символ з набором еталонів, по черзі накладаючи зображення одне на одне. Еталонами виступають спеціально підготовлені зображення, кожне з них об'єднує в собі контури багатьох варіантів написання того або іншого символу.

Ознаковий класифікатор висуває гіпотези, виходячи із ступеня схожості параметрів символу з еталонними значеннями. Оперує певними числовими ознаками, такими, наприклад, як довжина периметра, кількість чорних точок в різних областях або уздовж різних напрямів, тощо.

Структурний класифікатор проводить структурний аналіз символу, розкладаючи останній на елементарні складові (відрізки, дуги, кола, точки) і формуючи точну схему аналізованого знаку. Потім отримана схема (структурний опис букви) порівнюється з еталоном. Цей класифікатор працює повільніше растрового і ознакового, зате відрізняється високою точністю.

Блок навчання модифікує представлення елементів таким чином, щоб якнайшвидше класифікувати елементи зображення на наступних етапах. Він надає можливість збереження системою методу, стилю та інших результатів класифікації, які після успішного завершення процесу виділення ознак можуть бути використані для більш точного і швидкого розпізнавання зображень.

**Мета роботи:** визначити особливості систем оптичного розпізнавання текстів та дослідити метод структурного розпізнавання текстових зображень і побудови відповідної граматики.

### **Методи розпізнавання зображень текстів**

Можливість розпізнавання спирається на схожість однотипних об'єктів. Незважаючи на те, що в загальному всі об'єкти унікальні, між деякими з них завжди можна знайти подібності за тією або іншою ознакою. Звідси виникає можливість класифікації - розбиття всієї множини об'єктів на непересічні

підмножини, що носять назву класів, елементи яких мають схожі властивості, що відрізняють їх від елементів інших класів.

Розглянемо наступні методи розпізнавання зображень і класифікації текстів.

У методі порівняння проводиться порівняння виділеного елемента з базою даних, де для кожного представлені об'єкти під різними кутами, масштабами, зміщеннями, деформаціями, тощо. Для букв в базі даних слід зазначити також шрифт, властивості шрифту, і т.п.

Другий підхід - здійснюється глибокий аналіз характеристик образу. У разі оптичного розпізнавання тексту це може бути визначення геометричних характеристик окремих символів.

Новітні методи розпізнавання використовують штучні нейронні мережі. Хоча методи вимагають великої кількості прикладів при навчанні, але мають більш високу ефективність і продуктивність. У цих методах зображення символу приводиться до деякого стандартного розміру, наприклад, 16x16 пікселів. Значення яскравості у вузлах нормалізованого растра використовуються, як вхідні параметри нейронної мережі. Число вихідних параметрів нейронної мережі дорівнює числу розпізнаваних символів. Результатом розпізнавання є символ, якому відповідає найбільше зі значень вихідного вектора нейронної мережі.

Традиційно для опису зображень використовувалися розкладання в ряди за ортогональними функціями (ряди Фур'є, Чебишева, Карунена- Лоєва та ін.).

Структурний опис на відміну від розкладань:

- органічно зрозумілий для людини, що вирішує задачу розпізнавання об'єкта;
- прийнятний для комп'ютерної реалізації при розпізнаванні;
- вільний від трудомісткості обчислень і втрат інформації, властивих розкладанням.

Структурні ознаки, які використовують в структурному описі, це непохідні (тобто елементарні унікальні) елементи (символи), примітиви зображення об'єкта розпізнавання. Поява структурних ознак зобов'язана виникненню проблеми розпізнавання зображень з її специфічними особливостями.

Структурні методи розпізнавання символів зберігають інформацію не про точкове написання символу, а про його топологію. Тобто, еталон містить інформацію про взаємне розміщення окремих складових частин символу [5, 6]. Ясно, що при цьому стає неважливим розмір розпізнаваної букви і навіть шрифт, яким вона надрукована. Але основною проблемою структурних методів розпізнавання залишається ідентифікація знаків, які мають дефекти (наприклад, розрив лінії або злиття сусідніх ліній). Розглянемо особливості розпізнавання на конкретному прикладі [5].

Символ, що розпізнається, приводиться до деякого зразка, що має "еталонні" розміри і положення, згладжується, тобто проводиться виключення з зображення всіх елементів, які привели до його спотворення, шляхом заміни групи елементів зображення (як правило, сусідніх) на один еквівалентний їм елемент, кодується за допомогою послідовностей Фрімана, в результаті чого створюються неперервні лінії-послідовності (рис.2).

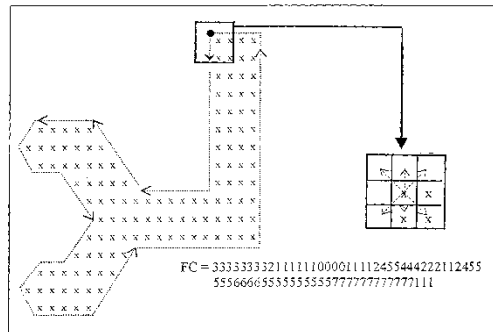


Рис.2. Кодування послідовності Фрімана для геометричного зображення символу

Отриманий образ піддається процедурі скелетизації (зменшення) [6]. Отриманий контур скелетного представлення описується у вигляді послідовного набору особливих точок і ланцюгового коду обходу контура за годинниковою стрілкою, що складається з точки прив'язки, числа кодів і масиву напрямків з чергової точки на наступну точку. Особливі точки - це кінцеві точки і точки розгалуження (тріюди), тобто точки, сусіди яких утворюють не менше трьох зв'язкових областей. Потім за допомогою перенумерування особливих точок і зміни початку контуру робиться спроба ототожнення контуру з одним з основних типів. Оперуючи обмеженим числом атомарних (непохідних) елементів (примітивів), можна одержати опис різноманітних об'єктів. Тобто, для об'єктів, що відрізняються, можна мати набір однакових непохідних елементів. Але для того, щоб опис можна було виконати, поряд з визначенням непохідних елементів повинні вводитися правила комбінування, що визначають способи побудови об'єкта із непохідних елементів. У результаті два однакових непохідних елементи різних об'єктів можуть бути з'єднані один з одним за різними правилами. Це і буде відрізнити об'єкти між собою.

У цілому для опису довільного об'єкта непохідні елементи поєднуються в ланцюжки (речення) за своїм, характерним тільки для цього об'єкта, набором правил.

У результаті зв'язків з непохідних елементів (структурних ознак) утвориться об'єкт, аналогічно до того, як речення мови будуються шляхом з'єднання слів, які в свою чергу складаються з букв. У цьому структурні методи мають аналогію із

синтаксисом природної мови. Звідси структурні ознаки носять ще назву лінгвістичних чи синтаксичних.

Таким чином, описом кожного класу повинен бути набір ланцюжків з непохідних елементів із правилами з'єднання. Кожний з них – характеристика структурних особливостей об'єктів цього класу.

Наведемо формальне визначення класу: класом називається сукупність об'єктів, пов'язаних між собою деяким відношенням еквівалентності, або, в крайньому разі, толерантності. Відношенням еквівалентності буде відношення, яке є симетричним, рефлексивним і транзитивним (наприклад, відношення “дорівнювати”). Для відношення ж толерантності властивість транзитивності в цілому не виконується (наприклад, відношення “бути схожим”).

Структурні або лінгвістичні системи для розпізнавання використовують речення, кожне з яких описує структуру (будову) об'єкта з непохідних (елементарних) елементів. Ці речення складають спеціальну мову. Класифікація об'єкта виконується шляхом порівняння речення невідомого об'єкта з еталонними реченнями класів.

### **Характеристика зображення текстових документів**

Особливістю цих зображень є те [5], що вони складаються з великої кількості взаємозалежних частин. Тому аналіз таких зображень доцільно проводити за допомогою методів структурного розпізнавання. Результатом такого розпізнавання є не віднесення зображення символу до прототипу, а перелік символів і відношень між ними. Складність, характерна для розпізнавання саме таких зображень, полягає у тому, що символ не визначається однозначно своїм зображенням. Ім'я зображення символу залежить не тільки від зображення цього фрагмента, а і від його місця, оточення, тобто контексту.

Ряд методів такого розпізнавання базується на евристичних прийомах, тобто розпізнавання за контекстом. Як правило, вказані методи залежать від предметної області і досвіду. Програмні продукти цього типу дуже ефективно розв'язують конкретну задачу. Таким чином, в блоці навчання (рис.1) доцільно реалізовувати можливість збереження системою методу, стилю та інших результатів класифікації.

Процедури розпізнавання, тобто функціонування вже навченої системи, є, як правило, достатньо простими і простими в реалізації. Тому, коли структурне розпізнавання застосовується до розв'язання проблем розпізнавання рукописного тексту, актуальність процедур навчання різко зростає. Однак проблеми навчання в структурному розпізнаванні значно складніші і значно менш досліджені порівняно з навчанням в неструктурному розпізнаванні.

Структурне розпізнавання символів в умовах випадкових шумів зводиться до пошуку оптимального зображення символів, які розпізнаються. Функція якості

полягає у відшукуванні найімовірнішої сукупності прихованих параметрів зображення. Наприклад, при розпізнаванні рядка тексту вимога знайти найімовірнішу послідовність літер рівнозначна вимозі мінімізації кількості неправильно розпізнаних символів.

Зображення які містять тексти, таблиці, рисунки, створюються і читаються згідно з певними правилами, набір яких можна формалізувати як певну граматику. Очевидно, що алгоритми розпізнавання таких зображень повинні базуватись на використанні правил цієї граматики.

Методи створення таких граматик наступні:

- використання теорії графів. Зображення тексту представлено у вигляді розміченого графа. Задачі розпізнавання ставляться як задачі знаходження ізоморфізму еталонного та вхідного графів, або ізоморфізму їх підграфів.
- методи теорії формальних мов та граматик. Зображення розглядається як слово у деякій формальній мові, яка задається за допомогою конструкцій, що є узагальненнями граматик Хомського. Розпізнавання полягає у відшукуванні найкращого в певному значенні виводу зображення у заданій граматиці [5].

Розглянемо методику створення відповідної граматики. Для цих граматик характерним є розгляд зображення як об'єкта, що складається за певними правилами з великої кількості елементарних частин. Ці частини і правила можуть значно відрізнятись між собою (наприклад, розпізнавання нот [9] або розпізнавання математичних формул [10]). В роботі [7] введено формалізм, названий загальною структурною конструкцією. Окрім власне визначення загальної структурної конструкції, в [7, 8] наведено загальну постановку та алгоритм розв'язання задачі розпізнавання в рамках введеного формалізму.

Нехай  $T$  – певна прямокутна підмножина двовимірної цілочисельної решітки:  $T = \{(i, j) \mid 0 \leq i < H, 0 \leq j < W\}$ . Множину  $T$  називатимемо полем зору, числа  $H$  та  $W$  – його висотою та шириною відповідно, а елементи поля зору називатимемо пікселями. Кольори пікселів вибираються з певної скінченної множини  $Y$ . Відображення  $x: T \rightarrow Y$  називатимемо зображенням, величина  $x(t)$  визначає колір зображення  $x$  в пікселі  $t \in T$ .

Двовимірною контекстно-вільною граматикою буде  $G = \langle E, V, P, \varepsilon \rangle$ , де  $E$  – множина термінальних зображень, що містить однопіксельні зображення (зображення, визначені на полі зору розміром 1 піксель) всіх кольорів з множини  $Y$ ,  $V$  – множина нетермінальних імен (метасимволи), що присвоюються частинам зображення в процесі його генерування за допомогою граматики,  $\varepsilon \in V$  – аксіома, використовується для іменування всього, повністю згенерованого зображення,  $P$  – множина правил виводу, містить правила трьох типів: правила горизонтальної конкатенації, вертикальної конкатенації (об'єднання блоків зображення) та

правила заміни. Кожну окрему множину цих правил позначатимемо відповідно  $P_h$ ,  $P_v$ ,  $P_s$ .

Правила заміни  $P_s$  мають вигляд  $v \rightarrow e$ , де  $v \in V$  – деяке нетермінальне ім'я, а  $e \in E$  – термінальне зображення.

Множина правил горизонтальної конкатенації  $P_h$  містить трійки нетермінальних імен вигляду  $v \rightarrow v_l/v_r$ . Наявність такого правила вказує на те, що будь-який прямокутний фрагмент зображення може отримати ім'я  $v$ , якщо його можна розбити вертикальною лінією на такі два прямокутні фрагменти, що лівий вже має ім'я  $v_l$ , а правий –  $v_r$ . Символ  $|$  використовується для розділення пари нетермінальних імен у правилах горизонтальної конкатенації, а не для альтернативного запису кількох правил, як це прийнято у формальних граматиках.

Аналогічно множина правил вертикальної  $P_v$  конкатенації містить трійки нетермінальних імен вигляду  $v \rightarrow \frac{v_t}{v_b}$ .

Наявність такого правила вказує на те, що будь-який прямокутний фрагмент зображення може отримати ім'я  $v$ , якщо його можна розбити горизонтальною лінією на такі два прямокутні фрагменти, що верхній вже має ім'я  $v_t$ , а нижній –  $v_b$ .

Мова граматики  $G$  складається з таких зображень, яким може бути присвоєне ім'я  $\varepsilon$ . Послідовність правил, застосованих до зображення  $x$ , результатом чого є присвоєння імені  $\varepsilon$  всьому зображенню, називається виводом зображення  $x$  в граматиці  $G$ .

Природнім чином для двовимірних контекстно-вільних граматик формулюється задача на точний збіг: для вхідного зображення  $x$  та граматики  $G$  визначити, чи належить  $x$  до мови граматики  $G$ , тобто чи існує вивід зображення  $x$  і якщо так, то необхідно знайти цей вивід.

Алгоритм (див. [8]) розв'язання задачі є прямим узагальненням алгоритму Кока-Янгера-Касамі [11] для визначення належності слова мові певної контекстно-вільної граматики Хомського. Він полягає в тому, що при послідовному перегляді всіх прямокутних фрагментів зображення для кожного з них визначається, які імена можуть бути присвоєні йому в даній граматиці. При цьому для перегляду фрагменти впорядковуються за розмірами.

Введене означення двовимірної контекстно-вільної граматики та постановка задачі на точний збіг, вочевидь мають ряд недоліків, що суттєво обмежують їх практичне застосування:

- з точки зору розпізнавання зображень формалізм двовимірних контекстно-вільних граматик не завжди дозволяє знайти реальне зображення, яке можна розділити на прямокутні фрагменти і при цьому дані фрагменти не перетинаються;



- надмірна деталізація правил граматики: для кожного зображення слід вказувати, яким чином воно складається з менших частин аж до рівня окремих пікселів. Очевидно слід оперувати більшими фрагментами зображення без вказання, яким чином вони можуть бути виведені з пікселів, наприклад, фрагментами, що відповідають окремим літерам тексту;

- постановка задачі на точний збіг вимагає точного збігу значень кольорів окремих пікселів зображення з кольорами термінальних зображень. За наявності шуму, що є характерним для розпізнавання, ця умова не виконується.

Базова контекстно-вільна конструкція, введена в [5] дозволяє усунути вказані недоліки, оскільки:

- вона описує зображення як таке, що складається з фрагментів довільної, а не лише прямокутної форми;

- процес побудови зображень задається до рівня довільного набору термінальних фрагментів, розміри яких можуть значно перевищувати розмір пікселя;

- визначається штраф за присвоєння імен фрагментам зображення, який для заданого фрагмента дорівнює сумі штрафів фрагментів, з яких його було утворено;

- штраф за термінальні фрагменти визначається довільною функцією, яка не пов'язана з формалізмом контекстно-вільних граматики. Це дозволяє використовувати у її побудові евристичні міркування. Наприклад, в задачі розпізнавання текстів нею може бути будь-яка функція, яка визначає схожість довільної літери та фрагмента зображення.

**Висновки.** Як зауважено в [5, 8] алгоритм розв'язання задачі синтаксичного аналізу в такій граматиці є узагальненням відповідного алгоритму для двовимірних контекстно-вільних граматики і полягає в послідовному обчисленні для кожного фрагмента штрафу за присвоєння йому кожного імені.

Спільним недоліком таких алгоритмів є їх часова та просторова складність, яка, наприклад, у випадку двовимірних контекстно-вільних граматики становить  $O(H^2W^2(H + W))$  та  $O(H^2W^2)$  відповідно. Ця складність хоч і є поліноміальною від розмірів  $H$  та  $W$  зображення, та все ж обмежує застосування алгоритмів на практиці.

Часова та просторова складність цих алгоритмів визначаються перш за все кількістю фрагментів, які переглядаються в ході їх роботи. Тому зменшення цієї кількості є основним способом зменшення складності алгоритмів розпізнавання.

### Список використаних джерел

1. Анисимов Б. В. Распознавание и цифровая обработка изображений: Учебное пособие для вузов / Б. В. Анисимов, В. Д. Курганов, В. К. Злобин. – М.: Высшая

- школа, 1983. – 295 с.
2. Тимченко О.В. Алгоритми та функції інформаційної системи розпізнавання символів на основі методів поліпшення зображень / Тимченко О.В., Кульчицька І.О., Тимченко О.О. // Моделювання та інформаційні технології. Зб. наук. пр. ПІМЕ НАН України. – Вип.69. – К.: 2013. – С.167-173.
  3. Гавриш Б.М., Дурняк Б.В., Тимченко О.В., Ющик О.В. Відтворення зображень растровими скануючими пристроями / Львів: Вид. УАД, 2016. – 180 с.
  4. Патрик Э. Основы теории распознавания образов / Эдвард А. Патрик. – М.: Советское радио, 1980; (пер. с англ. под ред. Б.Р. Левина). – 864с.
  5. Савчинський Б. Д. Контекстно-вільні граматичні конструкції для розпізнавання зображень текстових та графічних документів. Дис. ... канд. техн. наук за спец. 05.13.23 - системи та засоби штучного інтелекту. Київ, 2007.
  6. Methodologies for Evaluating Thinning Algorithms for Character Recognition [Text] / R.Plamondon, C. Y. Suen, M. Bourdeau, C. Barriere // J. Pattern Recognition and Artificial Intelligence, special issue thinning algorithms. – 1993. – Vol. 7, № 5. – P. 1247–1270.
  7. Шлезингер М., Главач В. Десять лекций по статистическому и структурному распознаванию. – К.: Наук. думка, 2004. - 545 с.
  8. Фу, К. Структурные методы в распознавании образов [Текст] : пер.с англ. / К. Фу ; под ред. М. А. Айзермана. – М. : Мир, 1977. – 319 с.
  9. Корес G. E., Chou P. A., Maltz D. A. Markov source model for printed music decoding // Journal of Electronic Imaging.\_ January 1996.\_ Vol. 5, no. 1. Pp. 7–14.
  10. Garcia P., Couasnon B. Using a genetic document recognition method for mathematical formulae recognition // IAPR Intern. Workshop on Graphics Recognition / Ed. by D. Blostein, Y.-B. Kwon. LNCS 2390. Berlin, Heidelberg: Springer-Verlag, 2002.\_ Pp. 236–244.
  11. Kasami T. An efficient recognition and syntax analysis algorithm for context-free languages: Tech. Rep. AFCLR-65-758. Bedford, Mass., USA: Air Force Cambridge Research Laboratory, 1965.

### References

1. Anisimov, B.V. Recognition and digital image processing: A manual for high schools / B.V. Anisimov, V.D. Kurganov, V.K. Zlobin. - М.: Higher school, 1983. - 295 p.
2. Tymchenko O.V. Algorithms and functions of the information recognition system for symbols based on image enhancement methods / Tymchenko OV, Kulchytska I.O, Tymchenko O.O. // Modeling and Information Technologies. Coll. Sciences IEEE of NAS of Ukraine. - Vip.69. - К.: 2013. - p.167-173.
3. Gavrish B.M., Durniak B.V., Tymchenko O.V., Yuschik O.V. Image reproduction with raster scanning devices / Lviv: UAD, 2016. - 180 p.
4. Patrick E. Fundamentals of Pattern Recognition Theory / Edward A. Patrick. - М.: Soviet Radio, 1980; (Translated from English under the editorship of Boris Levin). - 864s.
5. Savchinsky B. D. Context-free grammatical constructions for recognition of images of

- text and graphic documents. Dis ... Candidate Tech. Sciences for special 05.13.23 - systems and means of artificial intelligence. Kyiv, 2007.
6. Methodologies for Evaluating Thinning Algorithms for Character Recognition [Text] / R.Plamondon, C. Y. Suen, M. Bourdeau, C. Barriere // J. Pattern Recognition and Artificial Intelligence, special issue thinning algorithms. – 1993. – Vol. 7, № 5. – P. 1247–1270.
  7. Shlesinger M., Glavach B. Ten lectures on statistical and structural recognition. - K.: Science. Opinion, 2004. - 545 pp.
  8. Fu, K. Structural Methods in Pattern Recognition [Text]: Peres Eng. / K. Fu; Ed. MA Aiserman. - M.: World, 1977. - 319 p.
  9. Kopec G. E., Chou P. A., Maltz D. A. Markov source model for printed music decoding // Journal of Electronic Imaging.\_ January 1996.\_ Vol. 5, no. 1. Pp. 7–14.
  10. Garcia P., Couasnon B. Using a genetic document recognition method for mathematical formulae recognition // IAPR Intern. Workshop on Graphics Recognition / Ed. by D. Blostein, Y.-B. Kwon. LNCS 2390. Berlin, Heidelberg: Springer-Verlag, 2002.\_ Pp. 236–244.
  11. Kasami T. An efficient recognition and syntax analysis algorithm for context-free languages: Tech. Rep. AFCLR-65-758. Bedford, Mass., USA: Air Force Cambridge Research Laboratory, 1965.

## FEATURES OF SYSTEMS AND METHODS OF STRUCTURAL RECOGNITION OF TEXT IMAGES

*B.M. Gavrish*<sup>1</sup>, Ph.D., Senior Lecturer

*O.V. Tymchenko*<sup>1,2</sup>, D.Sc., professor

*Yu.O. Borzov*<sup>2</sup>, Ph.D.

*O.O. Tymchenko*<sup>1</sup>, graduate student

<sup>1</sup>*Ukrainian Academy of Printing*

<sup>2</sup>*Uniwersytet Warmińsko-Mazurski w Olsztynie, Polska*

<sup>3</sup>*Lviv State University of Life Safety*

**Abstract.** *Features of construction of systems of optical recognition of texts, methods of formation of image signs and method of constructing grammar for structural recognition of text images are considered.*

**Key words:** *systems of optical character recognition*

**Методологія дослідження.** Досліджуються системи OCR (optical character recognition) в яких цифрове зображення сканованого або сфотографованого документа перетворюється в текст, який містить це зображення. Для аналізу процесу перетворення створена та проаналізована узагальнена структурна схема

OCR, визначені функції її основних вузлів, дослідженні структурні та лінгвістичні методи, застосовані для побудови блоків класифікації та навчання.

**Результати.** Розглянуті методи структурного та неструктурного розпізнавання текстів, проведено їх порівняння, визначені переваги структурних методів. Узагальнено методи створення граматики для структурного розпізнавання, наведено загальну постановку та алгоритм розв'язання задачі розпізнавання текстів в рамках введеного формалізму. Показаний приклад створення відповідної граматики та визначено часову та просторову складність алгоритмів структурного розпізнавання зображень текстових документів.

**Новизна.** Показано, що система OCR повинна включати блок попередньої обробки зображення для зменшення шумів та виправлення існуючих геометричних спотворень, отриманих, наприклад, в результаті фотографування текстового зображення. Розглянуті функції та принципи побудови блоків класифікації та навчання. Узагальнено принципи застосування лінгвістичних методів побудови відповідної граматики.

**Практична значущість.** Узагальнено структурну схему системи розпізнавання символів та функцій її вузлів. Докладно проаналізовано створення двовимірної контекстно-вільної граматики для структурного розпізнавання текстів, визначена часова та просторова складність процесу розпізнавання зображень текстів.