

ДЕКОМПОЗИТИВНЕ РОЗПІЗНАВАННЯ СИМВОЛЬНОЇ ІНФОРМАЦІЇ З КРЕСЛЕНЬ ІЗ
ВИКОРИСТАННЯМ ТЕХНОЛОГІЙ ШТУЧНОГО ІНТЕЛЕКТУ

Розроблено новий метод декомпозитивного розпізнавання креслень шляхом поетапного відокремлення з них і розпізнавання однотипних образів. Описано технологію розділення креслення на блоки графічної та символної інформації з використанням нейронних схем. Проведено аналіз перспективних технологій штучного інтелекту з метою визначення оптимального методу розпізнавання символної інформації на графічних зображеннях.

It is designed new method of decompositive recognitions of the drawings by way of the phased branch from them and recognitions of its parts. Described technology of division of the drawing on blocks of graphic and symbol information with use neural schemes. The organized analysis of perspective technologies of the artificial intelligence for the reason determinations of the optimum method of the recognition to symbol information on graphic images.

Ключові слова: розпізнавання, розпізнавання креслень, штучний інтелект, нейронні мережі, нейронні схеми.

Keywords: recognitions, recognitions of the drawings, artificial intelligence, neural network, neural schemes.

Вступ.

У сучасному світі щодня переводиться з паперу в електронну форму велика кількість різних документів: друковані тексти, платіжні доручення, митні або податкові декларації, бюлетені для голосування, різні анкети й ін. Активно використовуються тисячі різних систем електронного документообігу практично у всіх сферах діяльності. При сучасних обсягах потоків документів подібні операції неможливі без автоматизованої обробки.

Використання паперових креслень є одним із факторів, що ускладнюють застосування систем автоматизованого проектування (САПР). Розробка технології перетворення паперових креслень у електронні аналоги допоможе зняти існуючі базові протиріччя й перевести процедури зберігання, модифікації та зберігання креслень на єдину базу.

Паперові креслення різних видів мають велику кількість недоліків. Наприклад, паперові креслення старіють і виходять з ладу, потребують місце й спеціальні умови для зберігання, одночасне використання одного креслення кількома людьми або відділами завжди важке, а його копіювання найчастіше веде до погіршення якості. Значними є й вторинні втрати, зокрема втрати часу, коли висококваліфіковані інженери шукають необхідні креслення; або збитки від документів, що губляться. Навіть ті компанії, які успішно впроваджують системи автоматизованого проектування, як і раніше оточені багатьма паперовими документами. Це і старі архіви самої організації, і паперові креслення, отримані від підрядників; і „тверді копії” електронних документів.

Ситуація ускладнюється ще й тим, що лише деякі нові проекти розробляються „з нуля”, без використання попередніх розробок підприємства, що зберігаються на папері або плівці. Тому щораз, коли приходить час випуску модернізованого виробу, організації електронних архівів, впровадження систем документообігу або провадження реконструкції приміщення чи установки, виникає проблема спільного використання старих паперових креслень і сучасних програмних засобів САПР.

На сьогоднішній найшвидшим способом перетворення паперового документу в електронний файл є *сканування*. Відскановані зображення можна відразу включати в систему електронного документообігу й використовувати як довідкову документацію, однак для редагування растрових зображень і їхнього повноцінного використання в САПР необхідне спеціалізоване програмне забезпечення. Це обумовлено принциповими розходженнями між растровою графікою, яка одержується при скануванні, і векторною графікою, котра створюється й використовується в автоматизованих системах креслення і проектування. Для вирішення проблеми з маніпуляцією відсканованими растровими зображеннями креслень, застосовується ряд методів щодо їхньої сегментації, *розпізнавання* і подальшого застосування в САПР.

Тому було поставлене завдання розробити такий метод розпізнавання креслень, який дозволить розпізнавати структурні елементи креслень із растрових зображень.

Основна частина.

Найбільш ефективним інструментом розпізнавання креслень є використання технологій штучного інтелекту для розпізнавання образів із застосуванням ефективних методів нейрообробки інформації. Відомі схеми аналізу зображень й оцінки даних малоєфективні, тому що вони негнучкі й прив'язані до певного неадаптивного алгоритму. З метою визначення найбільш ефективного методу розпізнавання графічної інформації, було виконано *аналіз технологій штучного інтелекту*, серед яких найбільш перспективними є наступні:

- Одношарові штучні нейронні мережі.
- Нейронні мережі із зворотним поширенням помилок.

- Нейронні мережі Хеба.
- Нейронні мережі Хопфілда.
- Нейронні схеми.

Одношарові штучні нейронні мережі [1] складаються із групи нейронів, що утворюють шар. В них кожен елемент із безлічі входів окремою вагою з'єднаний з кожним штучним нейроном, а кожен нейрон видає зважену суму входів у мережу. Можуть мати місце також з'єднання між виходами й входами елементів у шарі. Ваги всіх елементів матриці можна позначити через W . Матриця має m рядків й n стовпців, де m – число входів, а n – число нейронів. Наприклад, $w_{2,3}$ – це вага, що з'єднує третій вхід із другим нейроном. Таким чином, обчислення вихідного вектора N , компонентами якого є виходи нейронів, зводиться до матричного множення $N = XW$, де N і X – вектори рядка.

Коли в мережі тільки один шар нейронів, алгоритм її навчання із учителем очевидний, тому що правильні вихідні стани нейронів єдино шару наперед відомі, і підстроювання синоптичних зв'язків йдуть у напрямку, який мінімізує помилку на виході мережі. На цьому принципі будується, наприклад, алгоритм навчання одношарового перцептрона. У багатошарових мережах оптимальні вихідні значення нейронів всіх шарів, крім останнього, як правило, не відомі, і двох або більше шарів перцептрона неможливо навчити, керуючись тільки величинами помилок на виходах штучних нейронних мереж. У мереж, розглянутих вище, немає зворотних зв'язків, тобто з'єднань, що йдуть від виходів деякого шару до входів цього ж шару або попередніх шарів. Це спеціальний клас мереж, названий мережами без зворотних зв'язків, або мережами прямого поширення. У мережах без зворотних зв'язків немає пам'яті, їхній вихід повністю визначається поточними входами й значеннями ваг.

Один із варіантів рішення даної проблеми – розробка наборів вихідних сигналів, відповідних вхідним, для кожного шару штучних нейронних мереж, що як правило є дуже трудомісткою операцією й не завжди здійснюється. Другий варіант – динамічне підстроювання ваг коефіцієнтів синапсів, у ході якої вибираються, як правило, найбільш слабкі зв'язки й змінюються на малу величину в ту або іншу сторону, а зберігаються тільки ті зміни, які спричинили зменшення помилки на виході всієї мережі. Третій варіант – поширення сигналів помилки від виходів штучних нейронних мереж до її входів, у напрямку, зворотньому прямому поширенню сигналів у звичайному режимі роботи. Цей алгоритм навчання одержав назву *процедури зворотнього поширення* [2].

Нейронні мережі Хеба базуються на концепції, що нервові клітини мозку з'єднані одна з одною великою кількістю прямих і зворотних збудливих зв'язків й утворюють нейронну мережу. Кожен нейрон здійснює просторово-тимчасову сумачію вхідних сигналів від збуджених нейронів, визначаючи потенціал на своїй мембрані. Коли потенціал на мембрані перевищує граничне значення, нейрон збуджується. Нейрон володіє рефрактерністю й втомою. Ефективність зв'язків може змінюватися в процесі функціонування мережі, підвищуючись між одночасно збудженими нейронами. Це приводить до об'єднання нейронів у клітинні ансамблі – групи кліток, які частіше усього збуджувалися разом, і до відокремлення ансамблів один від іншого. При збудженні достатньої частини ансамблю, він збуджується цілком. Різні ансамблі можуть перетинатися: один і той самий нейрон може входити в різні ансамблі. Електрична активність мозку обумовлена послідовним порушенням окремих ансамблів.

Ідеї Хеба стали основою для створення нейронних моделей довгострокової пам'яті. Ансамблеву нейронну мережу можна розглядати як структуру, що реалізує функції розподіленої асоціативної пам'яті. Формування ансамблів у такій мережі відповідає запам'ятовуванню образів (ознак, об'єктів, подій, понять), закодованих патерном активності нейронів, а сформовані ансамблі є їх внутрішнім відтворенням. Процес порушення всього ансамблю при активації частини його нейронів можна інтерпретувати як витяг запам'ятованої інформації з її частини – ключу пам'яті. Модель пам'яті на основі ансамблевої нейронної мережі має деякі властивості, властиві біологічній пам'яті, такі, як асоціативність, розподіленість, паралельність, стійкість до шумів або збоїв і надійність. Проводяться також структурні аналогії між ансамблевими моделями нейронних мереж й будовою кори головного мозку. Є експериментальні дані про синаптичні пластичності, постульованої Хебом [3].

Серед різних конфігурацій штучних нейронних мереж зустрічаються такі, при класифікації яких за принципом навчання не підходять ні навчання з учителем, ні навчання без учителя. У таких мережах вагові коефіцієнти синапсів розраховуються тільки один раз перед початком функціонування мережі на основі інформації про оброблювані дані, і все навчання мережі зводиться саме до цього розрахунку. З одного боку, представлення апріорної інформації можна розцінювати, як допомога вчителя, але з іншого боку – мережа фактично просто запам'ятовує зразки до того, як на її вхід надходять реальні дані, і не може змінювати своє поведіння, тому говорити про ланку зворотного зв'язку з учителем не коректно. З мереж із подібною логікою роботи найбільш відомі мережа Хопфілда й мережа Хемінга, які звичайно використовуються для організації асоціативної пам'яті.

Мережа Хопфілда складається з єдиного шару нейронів, число яких є одночасно числом входів й виходів мережі [4]. Кожен нейрон зв'язаний синапсами з усіма іншими нейронами, а також має один вхідний синапс, через який здійснюється введення сигналу. Вихідні сигнали утворюються на аксонах.

Серед властивостей штучних нейронних мереж основною є їхня здатність до навчання, хоча навчаються вони найрізноманітнішими методами. Більшість методів навчання виходить із загальних передумов, і має багато ідентичних характеристик. Їх навчання нагадує процес інтелектуального розвитку людської особистості.

Можливості навчання штучних нейронних мереж обмежені. Проте, вже отримано переконливі досягнення і має місце багато практичних застосувань [5, 6]. Мережа навчається, щоб для деякої кількості входів давати необхідну кількість виходів. Кожна така вхідна (або вихідна) кількість розглядається як вектор. Навчання здійснюється шляхом послідовного пред'явлення вхідних векторів з одночасним підстроюванням ваг у відповідності з певною процедурою. У процесі навчання ваги мережі поступово стають такими, щоб кожен вхідний вектор виробляв вихідний вектор.

Основним недоліком нейронних мереж є локальність застосування їх окремих конструкцій, а тому за останні роки значні зусилля в області розробки систем ШІ були перенаправлені на створення *нейронних схем* [7]. Теорія нейросхем розглядає будь-яку ділянку біологічної нервової системи як друковану плату, де ролі мікросхем виконують тіла нейронів, а електричними провідниками є аксони й дендрити, підключені один до одного в певній послідовності (аксон завжди передає, а дендрит завжди приймає сигнал). Відповідно, нервовий патерн (сигнал, що передає інформацію між нейронами) розглядається як команда мікропроцесора, передана по одиничному провіднику у вигляді бітової послідовності. Нейронні схеми базуються на використанні базових класів компонентів-примітивів, причому всі складні елементи утворюються комбінацією більш простих компонентів за єдиним принципом. Як правило, у головному типі реалізовані всі функції основних схемних взаємодій мікросхем у нейронних схемах.

Перевагами нейронних схем над іншими технологіями штучного інтелекту є:

- 1.) Модульність: кожний компонент нейросхеми складається з компонентів більш низького рівня, створених окремо, що підвищує зручність і ефективність їх використання.
- 2.) Об'єктно-орієнтований підхід: використання компонентів як класів базового типу, що значно економить час при конструюванні нейросхем.
- 3.) Фрактальність: єдність принципів створення й застосування компонентів різних рівнів зменшує інформаційне й функціональне навантаження при роботі з нейронними схемами.
- 4.) Універсальність: з нейросхемних компонентів можуть бути спроектовані нейронні схеми для вирішення будь-яких задач (від пошуку рішення математичних рівнянь до аналізу складних багатомірних об'єктів).
- 5.) Функціональність: компоненти, спроектовані для вирішення певних задач, можуть бути використані в подальшому для виконання аналогічних функцій; створена нейронна схема може бути модифікована шляхом внесення змін як в архітектуру нейронної схеми, так і в механізм роботи окремих компонентів.

Розроблений метод декомпозитивного розпізнавання креслень дозволяє трансформувати відскановані растрові креслення у цифрові формати шляхом поетапного відокремлення з них однотипних образів і їх подальшого розпізнавання. Головним етапом його застосування є *відділення й розпізнавання символічної інформації на графічних зображеннях*. Виходячи з наведених переваг нейронних схем, даний інструмент був обраний для реалізації технології розпізнавання символічної інформації на графічних зображеннях.

Відділенню й розпізнаванню символічної інформації передують сегментація зображення, яка визначає окремі елементи на кресленні й дає можливість аналізувати їх диференційовано.

Сегментація зображення складається із двох основних етапів:

- бінаризація зображення;
- розділення зображення на окремі складові.

Бінаризація зображення полягає в зведенні кольорової палітри до двох кольорів і проводиться шляхом збільшення *контрастності* чи адаптивно. Авторами було розроблено новий *метод автоматизованого адаптивного аналізу спектрального розподілу*, що виконується шляхом ітераційного аналізу поступаючої від нейросхеми інформації в процесі розподілення кольорового балансу [8]. Автоматизований аналіз спектрального розподілу виконується шляхом ітераційного аналізу поступаючої від нейросхеми інформації в процесі розподілення кольорового балансу, й модель з найбільшим показником ефективності розпізнавання вважається за оптимальну. Таким чином, реалізується еквівалент роботи генетичного алгоритму при визначенні найбільш ефективної моделі. Результатом бінаризації зображення є зведення кольорової палітри до кольору образу та кольору фону.

Для **кластеризації** (розділення зображення на окремі складові) було розроблено новий метод сегментного заповнення. В процесі його використання проводиться поступовий аналіз пікселів робочої області зображення (Рисунок 1) для пошуку кольору образу. При його знаходженні виконується наступна послідовність:

1. Копіювання знайденого пікселя образу в окрему тестову область відповідно до його координат.
2. Присвоєння даному пікселю кольору фону на робочій області.
3. Аналіз кольору всіх суміжних пікселів – у випадку, якщо піксель має колір образу, його координати заносяться у динамічний масив.
4. Вибір з масиву координат останнього пікселю образу й виконання пункту 1.

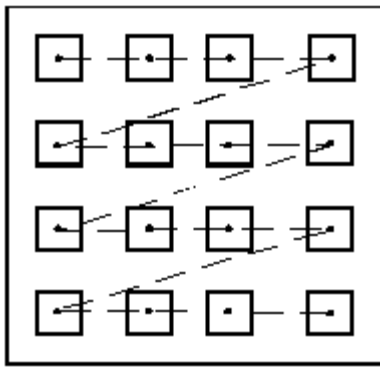


Рисунок 1. Схема шляху аналізу робочої області

Результатом роботи алгоритму є повне визначення знайденого сегменту (Рисунок 2) та його переміщення з робочої області в тестову із запам'ятовуванням координат образу на робочій області. Після чого проводиться пошук наступного образу, доки всі образи не будуть зафіксовані. У випадку, коли поданий на матрицю зображення образ буде розпізнаний із прийнятним рівнем вірогідності, його штамп на робочій області в наступному буде замінено на штамп фону. Перевагою розробленого методу є роздільний аналіз вкладених та суміжних об'єктів розпізнавання.

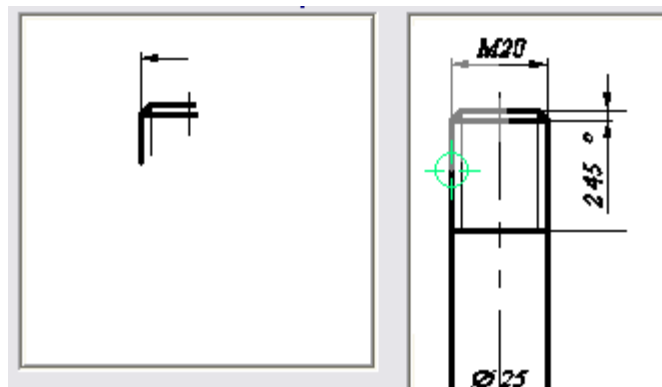


Рисунок 2. Визначення сегменту креслення

Наступним етапом є тестування кожного із знайдених образів розпізнаванням з метою встановлення, чи є даний образ символом. Критерієм приналежності образу до символічної множини є високий відсоток вірогідності **розпізнавання** його як символу. При цьому проводиться розпізнавання по наступних категоріях:

1. Друковані букви та цифри (відносяться до символічної інформації).
2. Окремі діакричні знаки – наприклад, верхні елементи українських літер „і”, „ї”, „й” (відносяться до символічної інформації за результатом подальшого аналізу).
3. Знаки пунктуації (відносяться до символічної інформації за результатом подальшого аналізу).
4. Друковані символи – наприклад, позначення діаметру (відносяться до символічної інформації).
5. Дрібні елементи креслення та розмітки – наприклад, креслення різьби (будуть віднесені до креслення).

Для базового очищення образу від шумів та спотворень використовується оптимізація параметрів образу (Рис. 3), зокрема **фасеткова згортка зображення** (Рис. 4). Після чого проводиться розпізнавання образу.

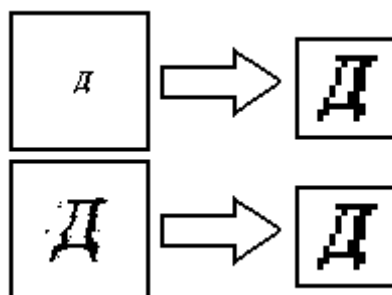


Рисунок 3. Оптимізація параметрів образу для формування матриці зображення

Для досягнення поставленого завдання розпізнавання символів за допомогою комплексної нейросхемної моделі, що використовує ймовірносно-статистичний аналіз растрових масивів, була побудована гібридна програмна модель, випробувана на прикладі розпізнавання рукописних текстів, але не прив'язана ні до якого конкретного набору символів і яка може бути застосована для будь-яких алфавітів і контурних зображень [7].

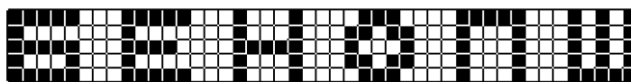


Рисунок 4. Фасеткові моделі образів

Точність й ефективність розпізнавання символів безпосередньо залежить від добре підібраних і відкаліброваних еталонних представників. Внесення в еталонну БД погано детермінованих й сильно перекручених об'єктів може викликати труднощі при класифікації образів, що істотно знизить відсоток розпізнавання [9].

Будь-який друкований текст має первинну властивість – шрифт, яким він надрукований. Із цього погляду існують два класи алгоритмів розпізнавання друкованих символів: шрифтовий й безшрифтовий [10].

Шрифтові або шрифтозалежні алгоритми використовують апріорну інформацію про шрифт, яким надруковані букви. Це означає, що програмі повинна бути пред'явлена повноцінна вибірка тексту, надрукованого даним шрифтом. Програма вимірює й аналізує різні характеристики шрифту й заносить їх у певну базу. По закінченні цього процесу шрифтова програма оптичного розпізнавання символів готова до розпізнавання даного конкретного шрифту. Цей процес умовно можна назвати навчанням програми. Далі навчання повторюється для деякої множини шрифтів, що залежить від області застосування програми.

Другий клас алгоритмів – безшрифтові, або шрифтонезалежні, алгоритми, що не мають апріорних знань про символи, які надходять до них на вхід. Ці алгоритми вимірюють й аналізують різні характеристики (ознаки), властиві буквам незалежно від шрифту й розміру, яким вони надруковані. У граничному випадку для шрифтонезалежного алгоритму процес навчання може бути відсутнім. У цьому випадку характеристики символів вимірює, кодує й поміщає в базу програми людина. Однак на практиці випадки, коли такий шлях вичерпно вирішує поставлене завдання, зустрічаються рідко. Більш загальний шлях створення бази характеристик полягає в навчанні програми на вибірці реальних символів.

Умовою для якісного розпізнавання є достатній набір введених еталонів, використовуючи які можна проводити аналітичні звірення й далі застосовувати алгоритми нейромережевої ідентифікації. Для самодостатності еталонів представники кожного символу БД повинні мати наступні характеристики: бути класифікованими, не бути сильно схожими між собою й не містити сильно деформовані елементи всередині класу.

У процесі асоціювання нових зображень із об'єктами з БД, можуть виникати проблеми, пов'язані з образотворчими невизначеностями, коли важко або навіть неможливо людині чи машині дати однозначну відповідь, до якого типу образу належить подане зображення.

З метою розпізнавання символів, що розміщені під кутом, проводиться циклічний процес розпізнавання для усіх кутів образу із кроком 10° . Критерієм вибору вірного кута розташування образу є максимальний відсоток вірогідності розпізнавання його як символу.

Після розпізнавання, одержується растрове креслення, позбавлене символічних надписів, і масив розпізнаних символів, що містить наступні дані:

- а) координати символу на робочій області;
- б) кут символу;
- в) назву символу;
- г) шрифт і розмір символу.

Наведений список даних символів є достатнім для виконання зворотнього складання растрового креслення або імпорту даних до спеціалізованих форматів цифрових креслень.

Даний етап є найскладнішим кроком в розпізнаванні растрових креслень, адже він дозволяє розділити креслення на масиви графічної та символічної інформації, які потребують відмінні алгоритми для свого розпізнавання [11].

Висновки.

Розроблено новий метод декомпаративного розпізнавання креслень, що заснований на поетапному відокремленні з креслень і розпізнаванні однотипних образів із використанням нейронних схем. Описано нові методи кластеризації та автоматизованого аналізу спектрального розподілу. Проведено аналіз перспективних технологій штучного інтелекту з метою визначення оптимального методу розпізнавання символічної інформації на графічних зображеннях. Розроблено програму, що використовує розроблену методологічну базу й проводить розпізнавання креслень із класифікацією розпізнаних елементів.

Список використаних джерел

1. Rosenblatt F. 1962. Principles of Neurodynamics. New York: Spartan Books. (Русский перевод: Розенблатт Ф. Принципы нейродинамики. – М: Мир. – 1965.)
2. Qrossberg S. 1969. Some networks that can learn, remember and reproduce any number of complicated space-time patterns. Journal of Mathematics and Mechanics, 19:53-91.
3. Hebb D. O. 1961. Organization of behavior. New York: Science Edition.
4. Сотник С. Л, Конспект лекций по курсу "основы проектирования систем искусственного интеллекта", Москва, 1998.
5. Sejnowski T. J., and Rosenberg C. R. 1987. Parallel Networks that learn to pronounce English text. Complex Systems 3:145–68.
6. DeSieno D. 1988. Adding a conscience to competitive learning Proceedings of the IEEE International Conference on Neural Networks, pp. 117-24. San Diego, CA: SOS Printing.
7. Мазурець О.В., Ковальчук С.С. Застосування багатосарових нейронних схем для вирішення складних технологічних задач // Збірник наукових праць за матеріалами другої всеукраїнської науково-технічної конференції «Актуальні проблеми комп'ютерних технологій 2008» – Хмельницький – ХНУ, 2008. – Т.2 – С.22–27.
8. Кондратюк А.В., Мазурець О.В. Розробка системи адаптивного аналізу спектральної інформації для оптимізації розпізнавання зашумлених образів за допомогою нейронних схем // Збірник наукових праць за матеріалами другої всеукраїнської науково-технічної конференції «Актуальні проблеми комп'ютерних технологій 2008» – Хмельницький – ХНУ, 2008. – Т.1 – С.127–135.
9. Ковальчук С.С., Рыбак Л.П., Мазурець А.В. Создание системы для распознавания рукописных текстов на базе нейронных схем // Сборник трудов Международной научной конференции «Нейросетевые технологии и их применение». Краматорск – 2004. – С.89–103.
10. Арлазаров В.Л., Троянker В.В., Котович Н. В. Адаптивное распознавание символов // «Интеллектуальные технологии ввода и обработки информации», М.: Эдиториал УРСС, 1998.
11. Кубик О.О., Мазурець О.В. Технологія декомпаративного розпізнавання символічної інформації з креслень на базі нейросхемних алгоритмів // Збірник наукових праць за матеріалами другої всеукраїнської науково-технічної конференції «Актуальні проблеми комп'ютерних технологій 2008» – Хмельницький – ХНУ, 2008. – Т.1 – С.151–161.