

Національна Академія наук України
Інститут кібернетики ім. В. М. Глушкова

На правах рукопису

ЛАВРЕНЮК Алла Миколаївна

УДК 007. 001. 362

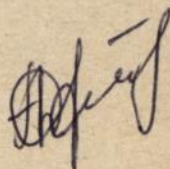
**ДОСЛІДЖЕННЯ НЕЙРОМЕРЕЖНОЇ СИСТЕМИ ДЛЯ РОЗПІЗНАВАННЯ
КРЕСЛЯРСЬКИХ СПЕЦИФІКАЦІЙ**

05.13.02 - математичне моделювання в наукових дослідженнях

Автореферат дисертації на здобуття наукового ступеня

кандидата технічних наук

Київ 1997



Дисертацією є рукопис
Роботу виконано в Інституті кібернетики ім.В.М.Глушкова НАН України

Науковий керівник: доктор технічних наук
Е. М. Куссуль

Офіційні опоненти: доктор технічних наук, проф.
В. І. Васильєв
доктор фізико-математичних наук
О.С. Макаренко

Провідна організація: Інститут проблем математичних машин та систем НАН
України

Захист відбудеться "7" квітня 1997 р. о 14 годині
на засіданні спеціалізованої вченої ради Д 01.39.05
у Інституті кібернетики ім.В.М.Глушкова НАН України
за адресою: 252187 Київ, проспект Академіка Глушкова, 40

З дисертацією можна ознайомитись у науково-технічному архіві інституту.

Автореферат розіслано "4" березня 1997 року.

Вчений секретар
спеціалізованої вченої ради
доктор біологічних наук

Л.М.КОЗАК

ЛННБ України ім.В.Стефаніка



00751998 (\$)

Актуальність. Моделювання структури та процесів обробки інформації в нейронних мережах людини та тварини дає можливість створювати технічні пристрої, що мають нові позитивні властивості: висока надійність, здатність до навчання, можливість розв'язання складних слабоформалізованих задач розпізнавання образів.

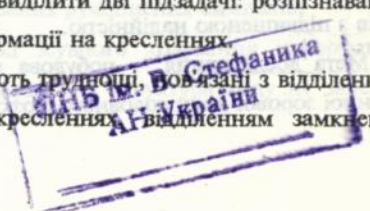
Такі пристрої широко застосовуються в різноманітних системах розпізнавання для виконання фінансових прогнозів і таке інше. Однак у таких пристроях, як правило, застосовуються нейронні мережі з так званим зосередженим або точковим кодуванням, при якому кожній ознаці або об'єкту ставиться у відповідність окремий нейрон. Реальним нейронним мережам людини та тварини властиве розподілене кодування, тобто таке кодування, при якому кожній ознаці ставиться у відповідність деяка підмножина нейронів. Воно вивчено значно гірше, а в його межах майже зовсім не досліджена можливість формування розподілених нейронних кодів, внутрішня структура яких не залежить від положення об'єкта на зображенні. Ця проблема є досить актуальною, оскільки її розв'язання з одного боку дозволило б побудувати більш адекватну математичну модель реальних нейронних мереж, а з іншого боку дало б можливість створити технічні пристрої для розв'язання складних задач штучного інтелекту.

Однією з таких задач є зберігання та обробка креслярської інформації за допомогою ЕОМ. Виділення потрібних елементів креслень та введення інформації у ЕОМ, розмноження креслень та внесення змін часто виконуються вручну, що призводить до похибок та дуже великих витрат часу. Розв'язання проблеми полягає у автоматичному розпізнаванні та перетворенні креслень.

Для введення у ЕОМ креслень використовується технологія автоматичних сканерів. При цьому важливою задачею є введення у комп'ютер не тільки образа креслень, а й його змісту, тобто перетворення графічної інформації в осмислену. Ця задача є однією із задач розпізнавання образів, якою зараз займаються багато вчених.

При розпізнаванні креслень можна виділити дві підзадачі: розпізнавання графічної та розпізнавання текстової інформації на кресленнях.

При розпізнаванні креслень виникають труднощі, пов'язані з виділенням текстової та графічної інформації на кресленнях, виділенням замкнених



контурів та з'єднувальних ліній, різними розмірами символів та щільністю написання. Можливі також пошкодження паперу та деяке спотворення зображення.

Щоб подолати ці труднощі, системи розпізнавання повинні працювати з високою точністю та надійністю, дуже чітко розпізнавати інформацію навіть при деякому ступені зашумленості, тобто моделювати функції мозку людини.

Для розв'язання подібних проблем можна використовувати нейронні мережі, які мають властивості асоціативності, паралельності обробки інформації, здатності до навчання.

Ступінь досліджуваності теми. Відомо, що в нейронних мережах людини та тварин для збереження та обробки інформації використовується розподілене кодування, при якому кожній ознаці або об'єкту ставиться у відповідність деяка підмножина нейронів, а не один нейрон. Однак у штучних нейронних мережах цей тип кодування застосовувався порівняно рідко. Хоча ідея ансамблевої організації мережі з'явилася ще в 1949 р. у Хебба, повністю реалізована вона була лише в асоціативно-проективних нейронних мережах (АПНМ). В АПНМ використовувався алгоритм локальнозв'язного кодування. Одна з переваг цього підходу полягає в тому, що коди подібних ознак або об'єктів корелюють між собою. Але недоліком алгоритму є те, що для одного й того ж об'єкту, розміщеного у різних місцях зображення, формуються різні коди. При цьому порушується схожість з біологічними нейронними мережами, що розпізнають об'єкт однаково добре незалежно від його положення.

Таким чином, у рамках розподіленого кодування майже зовсім не вивчена можливість формування розподілених нейронних кодів, внутрішня структура яких не залежить від розміщення об'єкта на зображенні. Тому для розв'язання поставлених у роботі задач пропонується метод розподіленого зсувного кодування, що дозволяє формувати код зсунутої ознаки шляхом зсуву елементів коду цієї ознаки.

Розв'язання даної проблеми для штучних нейронних мереж дозволило б досягти більшої подібності математичної моделі до біологічного прототипу та створити технічні пристрої для розв'язання складних задач розпізнавання образів з підвищеною надійністю.

Мета дослідження — побудова математичної моделі нейронної мережі первинної зорової кори головного мозку людини або тварини з використанням

розподіленого кодування інформації та її використання в технічних системах розпізнавання креслярських специфікацій.

Поставлена мета досягається шляхом розв'язання таких задач:

- вивчення проблеми кодування інформації в нейронних мережах людини і тварини, аналіз існуючих моделей нейронних мереж та підходів до кодування інформації в них;

- розробка та дослідження методу зсувного кодування інформації в моделях нейронних мереж перцептронного типу, при якому внутрішня структура коду не залежить від положення об'єкта;

- розробка алгоритмів попередньої обробки зображень символів;

- дослідження задачі виділення інформативних ознак символів на зображенні;

- створення алгоритму навчання нейронної мережі розпізнаванню креслярських специфікацій;

- розв'язання задачі розпізнавання окремих символів у креслярських специфікаціях, виділених в інтерактивному режимі;

- створення автоматичної системи розпізнавання креслярських специфікацій за допомогою нейронної мережі перцептронного типу.

Об'єктом дослідження є нейронні мережі. Предмет дослідження — моделювання функцій первинної зорової кори головного мозку для розв'язання задач розпізнавання образів.

Методи досліджень базуються на методах математичного моделювання біологічних структур та експериментальному дослідженні їх властивостей, на методах розпізнавання образів.

Застосовується також евристичний підхід до побудови алгоритмів обробки інформації.

Наукова новизна роботи полягає у тому, що

- на основі вивчення проблеми кодування інформації в нейронних мережах людини і тварини, аналізу існуючих моделей нейронних мереж та підходів до кодування інформації в них створено модель нейронної мережі для розв'язання задачі розпізнавання стилізованих рукописних символів;

- розроблено та досліджено метод зсувного кодування інформації в моделі нейронних мереж, при якому внутрішня структура коду не залежить від положення об'єкту;

- розроблено ефективні алгоритми попередньої обробки зображень символів;
- виділено інформативні ознаки символів на зображенні;
- побудовано алгоритм навчання нейронної мережі розпізнаванню креслярських специфікацій;
- розв'язано задачу розпізнавання окремих символів у креслярських специфікаціях;
- створено автоматичну систему розпізнавання креслярських специфікацій за допомогою нейронної мережі перцептронного типу.

Положення, що виносяться на захист.

1. Створена математична модель нейронної мережі перцептронного типу дозволяє будувати системи розпізнавання образів з підвищеною надійністю.
2. Розроблений метод зсувного кодування, що базується на принципі розподіленого кодування інформації в біологічних нейронних мережах, дозволяє відобразити зсув об'єкта на зображенні з високою точністю та зберегти кореляцію кодів об'єктів, що розміщені на зображенні близько один від одного.
3. Основні алгоритми обробки інформації, що використовуються в системі розпізнавання стилізованих рукописних символів, можуть бути представлені у вигляді функціонування спеціалізованих нейронних структур, що моделюють процес обробки інформації в окремих областях зорової кори головного мозку.

Особистий внесок автора у розробку наукових результатів, які виносяться на захист, полягає у тому, що:

- розроблено і досліджено модель нейронної мережі перцептронного типу та метод зсувного кодування інформації в ній, що дозволяє будувати коди ознак та об'єктів, інваріантні до зсуву об'єкта на зображенні та корельовані для близько розташованих об'єктів;
- розроблено та досліджено алгоритми попередньої обробки інформації, навчання та розпізнавання стилізованих рукописних символів;
- побудовано автоматичну систему розпізнавання креслярських специфікацій, що базується на розроблених алгоритмах.

Особистий внесок автора у роботи, виконані у співавторстві, полягає в математичному моделюванні нейронної мережі для задачі розпізнавання

образів, розробці алгоритму розпізнавання стилізованих символів та його програмній реалізації.

Теоретична та практична цінність. Розроблені математична модель нейронної мережі та алгоритми обробки інформації в ній розширюють обсяг відомостей про різні способи представлення інформації у нейронних мережах.

Теоретична цінність роботи полягає в аналізі нових властивостей нейронних кодів, що дозволяють розпізнавати символи незалежно від їх положення на сітчатці, розв'язанні проблеми кореляції таких кодів для об'єктів, зміщення яких на зображенні невелике, можливості побудови двомірних та трьохмірних зсувних кодів.

Практична цінність полягає у тому, що розроблено конкретну систему розпізнавання рукописних символів на основі нового способу кодування інформації, у якій розв'язано багато прикладних проблем: проблема нормування товщини символів, проблема виділення рядків, вирівнювання рядків та символів у рядку, розділення злитих символів, проблема навчання нейронної мережі. Ця система може бути застосована для розв'язання реальних задач розпізнавання слів у креслярських специфікаціях.

Розроблена математична модель нейронної мережі та алгоритми обробки інформації в ній можуть використовуватись при розв'язанні задач розпізнавання складних сцен, акустичної та інших видів інформації.

Рівень реалізації та впровадження результатів роботи. Результати дисертації отримано в рамках теми "Розробка принципів і методів представлення знань на різних ієрархічних рівнях нейроподібних систем" (реєстраційний номер 0194U009547) (1994-1997рр.) та проекту 6.3.3/043-92 "Нейрокомп'ютер" державної науково-технічної програми 6.3.3 "Нейрокомп'ютер" (1992-1994 рр.), що виконувалися у Інституті кібернетики ім. В.М.Глушкова НАН України. Результати дисертаційної роботи було використано у НДР "Тема" в в/ч Е6133А. Розроблені методи та алгоритми використовуються в ІК НАНУ при розробці нейрокомп'ютерних систем розпізнавання образів.

Апробація результатів роботи. Основні положення та результати дисертаційної роботи доповідались та обговорювались на Міжнародній конференції по проблемам моделювання в біоніці "Биомод-92" (Санкт-Петербург, 1992), Першій Всеукраїнській конференції "Обробка сигналів і зображень та розпізнавання образів" (Київ, 1992), Третій Українській

конференції по автоматичному керуванню "АВТОМАТИКА-96" (м. Севастополь, 1996) та щороку на Республіканському семінарі "Нейрокомп'ютери".

Публікації. За результатами роботи опубліковано 6 наукових праць.

Структура та об'єм роботи. Дисертація складається з вступу, 4 розділів, висновків, списку літератури з 120 найменувань та 1 додатку, викладена на 115 сторінках машинописного тексту; містить 33 малюнки.

ЗМІСТ РОБОТИ

В першому розділі проводиться порівняльний аналіз моделей нейронних мереж різної складності, розглядаються їх переваги та недоліки. За їх допомогою моделюється робота нейронних мереж різних відділів головного мозку людини та тварини. Застосування аналогії з біологічними нейронними мережами дозволяє створювати технічні системи для розв'язання складних слабоформалізованих задач розпізнавання образів, які успішно розв'язуються людиною або твариною. Найбільш універсальними моделями нейронних мереж є перцептрон та мережі зворотного розповсюдження помилки. Однак великий час навчання методом зворотного розповсюдження помилки та можливість попадання мережі у локальні мінімуми часто стоїть на перешкоді широкого використання цього методу у складних задачах штучного інтелекту в реальному часі.

Тому залишається актуальною проблема створення математичної моделі нейронної мережі для конкретної задачі, тобто вибору оптимальної структури нейронної мережі з точки зору швидкості роботи, складності розв'язуваної задачі та відносної простоти реалізації на ЕОМ. Враховуючи це, для розв'язання задачі розпізнавання креслярських специфікацій використовується нейронна мережа перцептронного типу, що є простою для математичного моделювання на персональній ЕОМ та показала хороші результати при розв'язанні прикладних задач.

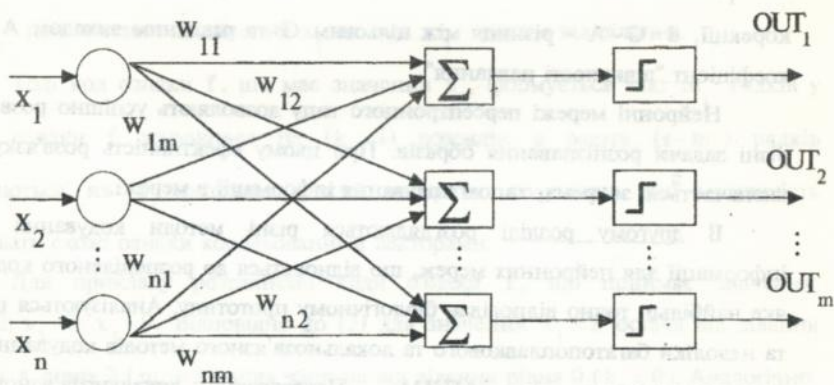
Простий перцептрон складається з одного шару штучних бінарних нейронів ("детектори ознак"), з'єднаних за допомогою вагових коефіцієнтів фіксованими зв'язками з множиною (матрицею) входів ("сенсорні нейрони"), а також бінарних нейроподібних елементів із зв'язками, що модифікуються (вирішальні елементи) (мал. 1). Кількість вирішальних елементів обирається рівною числу класів, на які необхідно розділити образи, що подаються

персептрону. Основною структурною одиницею в нейронних мережах є нейрон. Штучний нейрон, зображений на мал. 2, відображає властивості біологічного нейрона. На вхід нейрона подається вектор вхідних дій x , кожний компонент якого x_i є виходом іншого нейрона. Вхідне збудження NET нейрона визначається як зважена сума його входів.

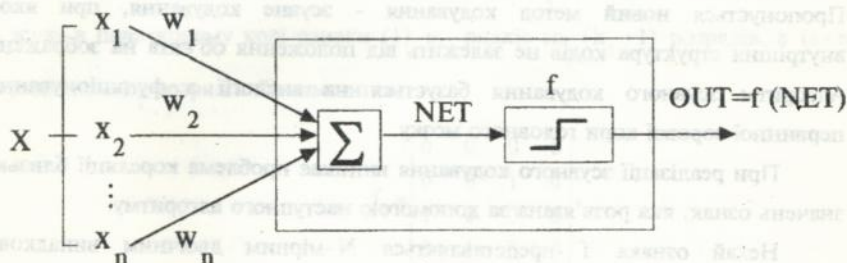
$$NET = x \cdot W,$$

де W - вектор ваги зв'язків. Вихідний сигнал нейрона визначається за допомогою активаційної функції від вхідного збудження.

$$OUT = f(NET).$$



Мал.1. Персептрон з багатьма виходами.



Мал.2. Штучний нейрон з активаційною функцією.

Активаційна функція персептрона є бінарною пороговою функцією вигляду

$$\text{OUT} = \begin{cases} 1, & \text{if } \text{NET} > T, \\ 0, & \text{if } \text{NET} \leq T, \end{cases}$$

де T - деяка стала величина.

Для навчання персептрона використовується дельта-правило. Корекція ваги зв'язків для i -го нейрона у відповідності з цим методом навчання відбувається за алгоритмом:

$$w_i(t+1) = w_i(t) + hdx_i,$$

де $w_i(t+1)$ - значення i -ї ваги після корекції, $w_i(t)$ - значення i -ї ваги до корекції, $d = G - A$ - різниця між цільовим G та реальним виходом A , а h - коефіцієнт "швидкості навчання".

Нейронні мережі персептронного типу дозволяють успішно розв'язувати різні задачі розпізнавання образів. При цьому ефективність розв'язку задачі визначається, зокрема, типом кодування інформації в мережі.

В другому розділі розглядаються різні методи кодування вхідної інформації для нейронних мереж, що відносяться до розподіленого кодування, яке найбільш точно відповідає біологічному прототипу. Аналізуються переваги та недоліки багатопоплавкового та локальнозв'язного методів кодування, котрі використовуються в АПНМ. Перевагою локальнозв'язного та багатопоплавкового методів кодування є кореляція кодів близьких ознак або об'єктів. А до їх недоліків можна віднести необхідність формування різних кодів для різних положень одного й того ж об'єкта на зображенні. Пропонується новий метод кодування - зсувне кодування, при якому внутрішня структура кодів не залежить від положення об'єкта на зображенні. Алгоритм зсувного кодування базується на аналогії з функціонуванням первинної зорової кори головного мозку.

При реалізації зсувного кодування виникає проблема кореляції близьких значень ознак, яка розв'язана за допомогою наступного алгоритму.

Нехай ознака f представляється N -мірним двоїчним випадковим вектором, у якому кількість одиничних елементів m значно менша від N . Стохастичний вектор f можна представити у вигляді двомірного масиву, що містить s рядків по g елементів. Наприклад, шістнадцятирозрядний вектор

$$f = [0 \ 0 \ 0 \ 1 \ 0 \ 1 \ 0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 0]$$

можна представити у вигляді двовірного масиву:

$$f = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \quad (1)$$

При цьому $r \cdot s = N$.

Нехай ознака f приймає цілочисельне значення x_i . Оскільки x_i - ціле число, то його можна представити у вигляді:

$$x_i = k_i \cdot s + m_i, \quad (2)$$

де k_i - ціла частина від ділення x_i на s , а m_i - остача від ділення.

Тоді код ознаки f , що має значення x_i , формується так: m_i рядків у коді ознаки f зсуваються на $(k_i + 1)$ розрядів, а решта $(s - m_i)$ рядків зсуваються на k_i розрядів. Такий алгоритм кодування дає можливість кодувати схожі ознаки корельованими векторами.

Для прикладу розглянемо коди ознаки f , що приймає значення: $x_1 = 2$; $x_2 = 3$; $x_3 = 9$. Відповідно до (2) для значення $x_1 = 2$ остача від ділення x_1 на s рівна 2 ($m_1 = 2$), ціла частина від ділення рівна 0 ($k_1 = 0$). Аналогічно

$$m_2 = 3, \quad k_2 = 0;$$

$$m_3 = 1, \quad k_3 = 2.$$

Код ознаки f , що приймає значення $x_1 = 2$, отримаємо в результаті зсуву в початковому коді ознаки (1) m_1 рядків на $(k_1 + 1)$ розрядів, а $(s - m_1)$ рядків на k_1 розрядів. Таким чином,

$$f_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

Аналогічно код ознаки f , що приймає значення $x_2 = 3$, отримаємо в результаті зсуву в початковому коді ознаки (1) m_2 рядків на $(k_2 + 1)$ розрядів,

а решту $(s - m_2)$ рядків на k_2 розрядів, тобто 3 рядки зсуваються на 1 розряд, а останній рядок не зсувається:

$$f_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Код ознаки f , що приймає значення $x_3 = 9$, отримуємо в результаті зсуву в початковому коді ознаки (1) одного рядка на 3 розряди, а решти трьох рядків - на 2 розряди:

$$f_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Представимо коди f_1, f_2, f_3 знову у вигляді векторів:

$$\begin{array}{l} x_1=2 \quad f_1 = [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad \boxed{1 \quad 0} \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0] \\ x_2=3 \quad f_2 = [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad \boxed{0 \quad 1} \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0] \\ x_3=9 \quad f_3 = [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad \boxed{0 \quad 0} \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1] \end{array}$$

Таке представлення кодів ознак визначає велику корельованість ознак, що приймають близькі значення (f_1 та f_2 , при цьому різниця у векторах показана пунктиром), і некорельованість цих ознак з f_3 .

При розв'язанні задачі розпізнавання креслярських специфікацій коди ознак розглядалися у вигляді трьохмірних двоїчних масивів з заданою ймовірністю появи одиниць у кожному елементі такого масиву $mask_1[Z][Y][X], mask_2[Z][Y][X], \dots$. Їх можна представити у вигляді паралелепіпедів розмірами X, Y, Z . Ці масиви отримали назву масок ознак. На мал. 3 зображено маску ознаки у вигляді кубу. Таке представлення маски ознаки по формі аналогічне організації клітин первинної зорової кори головного мозку в орієнтаційні колонки.

Код ознаки, що має координати (x_1, y_1) , формується за наступним алгоритмом. Маска ознаки зсувається по осі OX наступним чином: R_x шарів

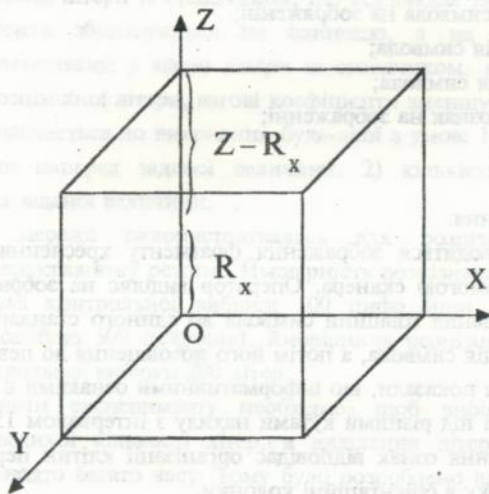
зсуваються на $(E_x + 1)$ елементів; $(Z - R_x)$ шарів зсуваються на E_x елементів, де E_x та R_x обчислюються за формулами:

$$E_x = x_1 / X,$$

$$R_x = x_1 \% X,$$

тобто E_x - ціла частина від ділення x_1 на X ; R_x - остача від ділення.

Величини E_x та R_x беруться по модулю. В залежності від знаку величини x_1 зсув виконується в додатньому чи від'ємному напрямку осі OX .



Мал.3. Маска ознаки у вигляді кубу.

Зсунута по осі OX маска ознаки аналогічним чином зсувається по осі OY : R_y шарів зсуваються на $(E_y + 1)$ елементів; $(Z - R_y)$ шарів зсуваються на E_y елементів, де E_y та R_y обчислюються за формулами:

$$E_y = y_1 / Y,$$

$$R_y = y_1 \% Y,$$

тобто E_y - ціла частина від ділення y_1 на Y ; R_y - остача від ділення.

Величини E_y та R_y беруться по модулю.

За таким же алгоритмом зсуваються всі маски ознак в залежності від їх положення на зображенні. Код об'єкта формується як порозрядна диз'юнкція кодів всіх ознак, що складають даний об'єкт.

Таким чином, якщо літера зсувається на зображенні в інше місце, то при зсувному кодуванні достатньо відповідно до цього зсунути її код і немає необхідності кодувати кожне нове положення ознаки, як при локальнозв'язному кодуванні.

У третьому розділі розглядається задача розпізнавання окремих символів (літер та цифр) у креслярських специфікаціях, що виділяються в інтерактивному режимі. Процес розв'язання задачі складається з наступних етапів:

- виділення символу на зображенні;
- скелетизація символу;
- потовщення символу;
- виділення ознак на зображенні;
- кодування;
- навчання;
- розпізнавання.

На екран виводиться зображення фрагменту креслення, введеного до комп'ютера за допомогою сканера. Оператор виділяє на зображенні окремий символ. Для приведення товщини символу до єдиного стандарту виконується спочатку скелетизація символу, а потім його потовщення до певного розміру.

Експерименти показали, що інформативними ознаками є відрізки певної довжини, розміщені під різними кутами нахилу з інтервалом 15 градусів. Така процедура формування ознак відповідає організації клітин первинної зорової кори головного мозку в орієнтаційні колонки.

Виділення ознак відбувається наступним чином. Кожна ознака послідовно накладається на всі точки зображення і, якщо ознака співпадає з точками, що належать зображенню символу, на зображенні фіксується точка, яка відповідає даній ознаці (будемо говорити, що в цій точці присутня дана ознака).

Реально отримується ціла область точок, у яких присутня дана ознака. Якщо при розпізнаванні враховувати усі ці точки, то час розпізнавання істотно збільшується. Для зменшення кількості виділених точок використовується алгоритм, котрий залишає лише частину із цих точок.

Точки, що відповідають усім виділеним ознакам, виводяться на одне і те ж зображення, в результаті чого отримується символ, складений із точок.

Виділені ознаки кодуються на основі алгоритму зсувного кодування, описаного у розділі 2. Отримані в результаті виконання такого алгоритму маски ознак об'єднуються в маску символу операцією порозрядної диз'юнкції.

Оскільки розмір ансамблю фіксований (m нейронів), а в отриманій масці може міститися більше число активних нейронів, то їх кількість нормується для підтримки фіксованого числа активних нейронів.

Для розв'язання задачі розпізнавання окремих символів використовувалась модель нейронної мережі у вигляді одношарового перцептрону розміром 8192 нейрона.

Для навчання системи пропонується використовувати наступну процедуру. На вхід мережі (див. мал. 1) послідовно подаються для розпізнавання маски літер у вигляді одномірних масивів, на вихід мережі - імена літер. Якщо літера розпізнається неправильно, то її маска подається на навчання. При цьому у матриці зв'язків на перетині рядків з одиничними елементами у масці літери із стовпчиком, що відповідає імені істинної літери, вагові коефіцієнти збільшуються на одиницю, а на перетині рядків з одиничними елементами у масці літери із стовпчиком, що відповідає імені неправильно розпізнаної літери, вагові коефіцієнти зменшуються на одиницю.

Мережа навчається до виконання будь-якої з умов: 1) кількість помилок зменшується до наперед заданої величини; 2) кількість циклів навчання досягає наперед заданої величини.

Навчена мережа використовувалась для розпізнавання символів, виділених в інтерактивному режимі. Ймовірність розпізнавання цифр складала 100% при об'ємі контрольної вибірки 100 цифр (при досягненні розміру нейронного ансамблю 500 нейронів), ймовірність розпізнавання літер - 93% при об'ємі контрольної вибірки 200 літер.

Для повноти експерименту необхідно, щоб вибірка для навчання складалася з великої кількості літер, а виділення літер в інтерактивному режимі займає надто багато часу. Тому було розроблено автоматичну систему розпізнавання креслярських специфікацій, котра розглядається у четвертому розділі.

У процесі розв'язання задачі автоматичного читання креслярських специфікацій можна виділити наступні етапи:

- виділення рядків з текстом;
- вирівнювання рядків;
- виділення окремих символів;
- розпізнавання символів;
- створення інтерфейсу користувача.

Для розв'язання задачі використовуються папки креслярської документації, написаної на стандартних бланках, розграфлених горизонтальними лініями. Отже рядок з текстом знаходиться між двома горизонтальними лініями.

При введенні зображення тексту до комп'ютера за допомогою сканера лінії, а відповідно і текст, розміщуються під невеликими кутами нахилу. Для покращення процесу розпізнавання такі рядки попередньо вирівнюються.

Для знаходження точного положення лінії та кута її нахилу будується лінійна регресійна модель за методом найменших квадратів:

$$Y = A_0 + A_1 \cdot X.$$

У ролі n вибірок експериментальних даних $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ виступали центри тяжіння кожної із n частин, на які попередньо розбивалась лінія.

Отримана пряма найбільш точно апроксимує дану лінію на зображенні. Кут нахилу прямої приймається за кут нахилу даної лінії, а відповідно і рядка з текстом, розташованого над лінією. Потім рядок вирівнюється за рахунок повороту його на знайдений кут вверх чи вниз в залежності від величини кута.

Наступним етапом розв'язання задачі автоматичного читання креслярських специфікацій є виділення літер у рядку. Для більш точного виділення літер попередньо виконувалась процедура вирівнювання літер. У креслярському шрифті літери розміщуються під певним кутом нахилу. В результаті виконання процедури вирівнювання усі літери записуються без нахилу. При розв'язанні задачі вирівнювання літер застосовувався алгоритм, аналогічний до алгоритму виділення ознак, котрий ми використовували при розв'язанні задачі розпізнавання окремих літер, виділених в інтерактивному режимі, що описаний у третьому розділі.

У задачі автоматичного розпізнавання текстів виникає проблема розділення злитих літер, для розв'язання якої застосовується наступний алгоритм. Визначаються проєкції літер на горизонтальну вісь. Якщо довжина проєкції літери більша ніж проєкція найширшої літери на зображенні, то це означає, що літери злились, і їх потрібно розділити.

У вікно розмірами $H \times W$, де H - висота вікна, рівна висоті рядка з текстом на зображенні, W - ширина вікна, рівна максимально можливій ширині літери, розміщується фрагмент рядка з текстом від лівої границі проєкції літери до правої. Цей фрагмент подається на розпізнавання. В результаті розпізнавання визначається певна літера. По середній ширині цієї літери розділяємо літери у рядку. Під час навчання мережі формуються три масиви. В один записується для кожної i -ї літери кількість екземплярів N_i , котрі зустрілися при навчанні. В другий масив записуються сумарні ширини цих екземплярів для кожної літери S_i . В третій - середня ширина кожної літери, котра визначається діленням сумарної ширини екземплярів літери на кількість цих екземплярів:

$$d_i = S_i / N_i.$$

Виділена таким чином по середній ширині літера знову подається на розпізнавання (контроль відділення). Якщо літера при цьому розпізнається невірно, то оператор, змінюючи розмір вікна, визначає істинну ширину літери. Середня ширина даної літери відповідно корегується, і дана літера записується у файл для подальшого використання її при навчанні мережі. Потім у вікно записується наступна літера і виконуються аналогічні дії. При досягненні наперед заданої кількості помилок відбувається процес навчання мережі. У процесі навчання мережі використовуються усі літери, котрі розпізнавалися неправильно. Для розв'язання задачі розпізнавання літер застосовуються алгоритми скелетизації, потовщення літер, виділення ознак, зсувного кодування, навчання, що використовувалися при розв'язанні задачі розпізнавання окремих літер (розділ 3), виділених в інтерактивному режимі.

Система автоматичного розпізнавання креслярських специфікацій на основі нейронної мережі персеptronного типу розміром 8192 нейрони з використанням зсувного кодування була реалізована у вигляді комплексу програм. Проведені експерименти підтвердили ефективність запропонованих методів та алгоритмів.

Довжина навчальної вибірки у проведених експериментах збільшувалась до 1200 (використовувались різні варіанти написання символів). При цьому проводились експерименти по розпізнаванню символів, написаних різними почерками. Розмір нейронного ансамблю збільшувався, починаючи з 250 нейронів, при цьому зростає відсоток правильно розпізнаних символів. І при досягненні розміру нейронного ансамблю 650 нейронів ймовірність розпізнавання становила 98% при об'ємі контрольної вибірки 1000 символів. При подальшому збільшенні розміру нейронного ансамблю настає переповнення мережі та ймовірність розпізнавання падає. Результати розпізнавання рукописних стилізованих текстів підтверджують правильність вибраного підходу моделювання окремих функцій мозку та ефективність запропонованих методів обробки інформації у технічних системах розпізнавання текстів.

РЕЗУЛЬТАТИ РОБОТИ ТА ВИСНОВКИ

1. При моделюванні нейронних мереж людини та тварини доцільно використовувати нейронні мережі персеptronного типу та застосовувати розподілений спосіб кодування інформації в них.

2. За допомогою запропонованого методу зсувного кодування інформації можна будувати математичні моделі розподілених нейронних кодів, внутрішня структура яких не залежить від положення об'єкта. При такому методі кодування близькі ознаки та об'єкти кодуються корельованими векторами.

3. Розроблено ефективні алгоритми попередньої обробки зображень стилізованих рукописних символів.

4. Для розв'язання задачі розділення злитих символів доцільно застосовувати алгоритм навчання системи, що полягає в послідовному уточненні середньої ширини символу за рахунок операторної корекції ширини вікна для виділення символу.

5. Для розпізнавання стилізованих рукописних символів достатнім є набір ознак, що містить відрізки прямих ліній різної орієнтації. Алгоритм виділення таких ознак моделює роботу складних клітин первинної зорової кори головного мозку людини та тварини.

6. На базі розроблених алгоритмів створено інтерактивну систему розпізнавання стилізованих рукописних символів, робота якої підтверджує ефективність розроблених алгоритмів.

7. Побудовано та реалізовано у вигляді програми систему автоматичного розпізнавання креслярських специфікацій на базі математичної моделі нейронної мережі людини з використанням розподіленого кодування інформації.

Основні результати роботи опубліковано в працях:

1. Лавренюк А.Н. О выборе информативных признаков для нейрокомпьютерного распознавания рукописных текстов// Нейросетевые технологии и нейрокомпьютеры. - Киев: Ин-т кибернетики им.В.М.Глушкова АН Украины, 1994. - С. 36-41.

2. Лавренюк А.Н. Применение нейронных сетей для распознавания надписей на чертежах// Нейрокомпьютинг. Вопросы теории и практики. - Киев: Ин-т кибернетики им.В.М.Глушкова НАН Украины, 1995. - С.24-31.

3. Лавренюк А.Н. Нейросетевая система для автоматического чтения чертежных спецификаций// Нейрокомпьютинг. Вопросы теории и практики. - Киев: Ин-т кибернетики им.В.М.Глушкова НАН Украины, 1996. - С. 28-35.

4. Куссиль М.Э., Лавренюк А.Н. Программно-аппаратный ассоциативно-проективный нейрокомпьютер// Тезисы докладов Международной конференции по проблемам моделирования в бионике "Биомод-92". - С.-Петербург, 1992. - С. 258-259.

5. Куссиль Н. М., Куссиль М. Е., Лавренюк А.М., Чумак В. В. Ассоциативно-проективный нейрокомпьютер та деякі питання створення пристроїв для задач розпізнавання образів// Праці I Всеукраїнської конф. "Обробка сигналів і зображень та розпізнавання образів". - Київ, 1992. - С. 209-210.

6. Лавренюк А.Н. Автоматическая система чтения чертежных спецификаций// Труды III Украинской конф. по автоматическому управлению "АВТОМАТИКА-96". - Севастополь, 1996. - т. 2. - С. 73.

АННОТАЦИЯ

Лавренюк А. Н.

Исследование нейросетевой системы распознавания чертежных спецификаций.

Диссертация (рукопись) на соискание ученой степени кандидата технических наук по специальности 05.13.02 - математическое моделирование в научных исследованиях, Институт кибернетики им. В.М.Глушкова НАН Украины, Киев, 1997.

Защищается 6 научных работ, в которых содержатся теоретические и практические результаты по моделированию нейронных сетей и кодированию информации в них. Предложен метод сдвигового кодирования, разработаны алгоритмы предварительной обработки изображений стилизованных рукописных символов, обучения и распознавания. Разработана и реализована в виде программного обеспечения автоматическая система чтения чертежных спецификаций.

Lavrenyuk A.N.

Investigation of neural network system for drawing specifications recognition.

Thesis (manuscript) for scientific degree of Candidate of Technical Sciences, speciality code 05.13.02 - Mathematical Modelling in Scientific Investigations, V.M.Glushkov Institute of Cybernetics, National Academy of Sciences of Ukraine, Kiev, 1997.

Six publications are defended containing theoretical and practical results on neural network modelling and coding of information in the networks. Shift coding method is proposed, algorithms for image pre-processing of drawing symbols, learning and recognition are developed. Automatic system for drawing specifications reading is developed and realised in software.

Ключові слова:

Нейронні мережі, розпізнавання символів, виділення ознак, кодування інформації, попередня обробка зображень, алгоритм навчання.

Підписано до друку 03.03.97. Формат 60x90 1x16.

Папір друкарський № 2. Друк офсетний.

Умов. друк. л. 1, 16. Фіз. друк. л. 1, 25.

Замовлення 279. Тираж 100.

Вул. Суворова, 4/6. КОС.

435437

AB 37.163