

Технические науки

УДК 550.34.06.013.3

Лукиянчук Анатолій Анатолійович

аспірант (ад'юнкт) Військового інституту

Київського національного університету імені Тараса Шевченка

Лукиянчук Анатолий Анатольевич

аспирант (адъюнкт) Военного института

Киевского национального университета имени Тараса Шевченко

Lukiyanchuk Anatoliy

PhD Student of the

Military Institute of Taras Shevchenko University of Kyiv

МЕТОДИ РОЗПІЗНАВАННЯ ОБРАЗІВ, ЯК ІНСТРУМЕНТ

ІДЕНТИФІКАЦІЇ ПРОТИВНИКА

МЕТОДЫ РАСПОЗНАВАНИЯ ОБРАЗОВ, КАК ИНСТРУМЕНТ

ИДЕНТИФИКАЦИИ ПРОТИВНИКА

METHODS OF IMAGE RECOGNITION AS AN INSTRUMENT OF

ENEMY IDENTIFICATION

***Анотація.** В статті розглядається алгоритм прийняття рішення за яким розв'язується проблема розпізнавання образу. Надано визначення «образу в інформаційному просторі». Наведені умовно графічне представлення розпізнавання образів та можливості нейронно-мережевого та синтаксичного або структурного методів обробки даних. Визначені основні недоліки та переваги наведених методів розпізнавання образів.*

***Ключові слова:** образ, розпізнавання образів, сейсмоакустика, нейронні мережі, периметрові системи, ідентифікація, дискримінантні методи, синтаксичні або структурні методи.*

Аннотация. В статье рассматривается алгоритм принятия решения по которому решается проблема распознавания образа. Дано определение «образа в информационном пространстве». Дано условно графическое представление распознавания образов и возможности нейронно-сетевого и синтаксического или структурного методов обработки данных. Определены основные недостатки и преимущества приведённых методов распознавания образов.

Ключевые слова: образ, распознавание образов, сейсмоакустика, нейронные сети, периметровые системы, идентификация, дискриминантные методы, синтаксические или структурные методы.

Summary. The article deals with the investigation of the algorithm of making decisions for solving the problem of image recognition. The definition of "image in the information space" is given. The graphic representation of image recognition and possibilities of neural network and syntactic or structural data processing methods are given. The main disadvantages and advantages of the above methods of image recognition are determined.

Key words: image, image recognition, seismoacoustics, neural networks, perimetric systems, identification, discriminant methods, syntactic or structural methods.

Вступ. Периметрові системи виявлення, що призначені для підвищення ефективності охорони промислових, військових та цивільних об'єктів, займають особливе місце в галузі спеціальної техніки. Сейсмоакустичні системи спостереження мають в порівнянні з рештою систем виявлення можливість контролювати обстановку не лише в вузькій смузі уздовж контрольованого периметру, але й в широкій зоні на підступах до нього. Робота сейсмоакустичної системи заснована на реєстрації сейсмічних хвиль, що збуджуються в ґрунті різними фізичними діями. Сейсмоакустична система виконує функції по отриманню

інформації із зовнішнього середовища, її обробці і ухваленню рішення про вторгнення порушника в зону, що охороняється. Для підвищення ефективності роботи сейсмоакустичних систем при вирішенні подібних задач використовують сучасні технології з елементами штучного інтелекту.

Метою статті є аналіз систем розпізнавання образів та виявлення їх недоліків та переваг.

Основна частина. Одним із завдань, що вирішуються сейсмоакустичною системою спостереження є ідентифікація об'єкту порушення. Ідентифікувати об'єкт означає поставити у відповідність сигналу на вході системи будь-який об'єкт (людина, група людей, автомобіль або інша техніка). При цьому вирішується завдання класифікації за сейсмограмою, пов'язаною з визначенням приналежності цього об'єкту до одного із заздалегідь виділених класів об'єктів, тобто з визначенням приналежності сигналу на вході системи до виділеного класу сигналів. Її рішення передбачає розробку певного методу (вирішального правила), який з певною вірогідністю міг би віднести записану подію до одного з класів. Більшість математичних методів розпізнавання образів поділяють на дві групи [2]: дискримінантні та синтаксичні або структурні. Особливе місце в розпізнаванні належить методам, які ґрунтуються на використанні штучних нейронних мереж [3; 4; 5; 6]. Такі методи в літературі відносять до дискрименантних методів або виділяють в окрему групу.

Дамо тепер визначення образу. Образ в інформаційному просторі – це деякий сигнал, навколо якого сформовано суцільну та однозв'язну множину, до того ж усі сигнали множини асоціюються з ним. Образ формується шляхом узагальнення усіх сигналів, близьких до нього. При такому визначенню образу проблема розпізнавання вирішується просто.

Якщо сигнал, що досліджується потрапляє в образ (в його множині) то говоримо, що він виявлений, якщо ні – це сторонній образ.

Нехай маємо деякий сигнал S . Визначимо, чи є в ньому образ F та яка його амплітуда. Будемо віднімати сигнал F від S з деякою амплітудою A та знайдемо енергію залишку W . При повному відніманні енергія, що залишилася мінімальна.

$$\sum_{i=0}^N (S_i - AF_i)^2 = W \rightarrow \min \quad (1)$$

Знайдемо амплітуду образу. Це варіаційна задача. В точці мінімуму енергії приватна похідна за амплітудою дорівнює нулю

$$\frac{\delta W}{dA} = 0 \quad (2)$$

Виконаємо диференціювання і після простих перетворень отримаємо

$$A = \frac{1}{W_F} \sum_{i=0}^N S_i F_i \quad (3)$$

$$W_F = \sum_{i=0}^N F_i^2 \quad (4)$$

W_F - енергія образу. Амплітуда сигналу F , який присутній в S , обчислюється як згортка цих сигналів. Амплітуда є безрозмірною величиною та являє собою масштабний фактор наявності образу F в сигналі S (операція виявлення деякого сигналу за допомогою згортки добре відома в техніці обробки сигналів під назвою оптимальний кореляційний фільтр).

Тепер необхідно встановити кордон безлічі сигналів, що належать образу та визначити прийняття рішення "свій-чужий". Спочатку знайдемо енергію вхідного сигналу та його амплітуду.

$$W_S = \sum_{i=0}^N S_i^2 \quad (5)$$

Визначається відносна амплітуда вхідного сигналу, до того ж за одиницю прийнята амплітуда образу F . Це необхідно для подальшого порівняння A_S та A .

$$A_S = \sqrt{\frac{W_S}{W_F}} \quad (6)$$

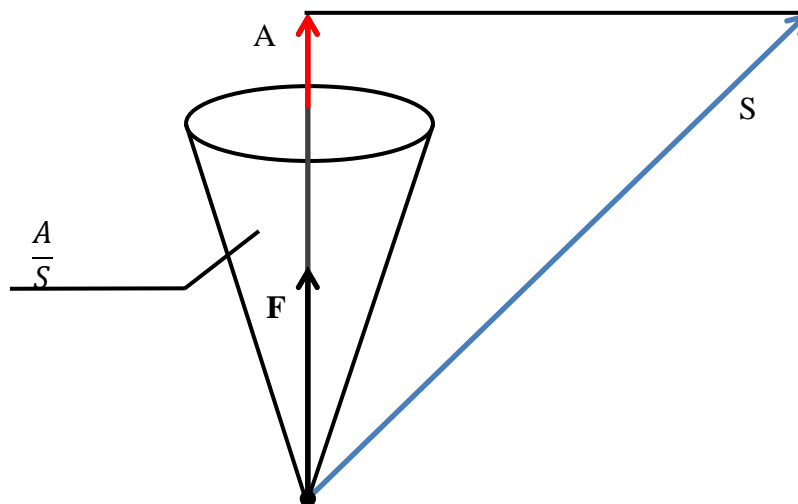


Рис. 1. Умовне графічне представлення операції розпізнавання

Умовне графічне представлення операції розпізнавання показано на рисунку 1. Під час розпізнавання необхідно визначити амплітуду образу, якщо ж необхідного образу немає, то в якості негативного результату приймається амплітуда вхідного сигналу. Алгоритм прийняття рішення виглядає наступним чином:

$$\frac{A}{A_S} \geq (1 - K) \begin{cases} \text{Вірно } YES = A; NO = 0 \\ \text{Невірно } YES = 0; NO = A_S \end{cases} \quad (7)$$

Множина сигналів, що розташовані всередині конуса узагальнюється в образ, а сам сигнал F є віссю симетрії конусу. Якщо вхідний сигнал потрапляє у внутрішній простір конусу, умова (7) виконується (*Вірно*), на вихід YES передається амплітуда образу. В протилежному випадку (*Невірно*) на виході NO встановлюється амплітуда вхідного сигналу.

K – коефіцієнт узагальнення. Він приймає значення в інтервалі від 0 до 1. Якщо $K=0$, узагальнення відсутні. Найменші відхилення вхідного сигналу від образу F призводять до видачі негативного результату $NO=A_S$. Якщо $K=1$ – максимальне узагальнення. Даний підхід відображає простий

факт, що чим більша амплітуда образу, тим більші можуть бути допустимі значення.

Нехай сигнал від об'єкта описаний m -розмірним вектором ознак $X = \{x_1, \dots, x_m\}$. Безліч реалізацій сигналів від об'єкту, описаних за допомогою такого вектора, утворюють навчальну множину. По ній можна визначити закономірності і зв'язки між значеннями ознак. Для кожного об'єкта, маючи навчальну вибірку, можна знайти еталонний вектор (або кілька векторів) і поставити у відповідність номер класу z . Тоді, скориставшись деяким правилом і задавшись мірою подібності, можна вхідний (контрольний) сигнал, описаний за допомогою того ж вектора ознак, віднести до одного з класів z .

Оптимальна за критерієм мінімуму середнього ризику класифікація виконується на основі умовних ймовірностей $P(z/x)$. При наявності апріорної невизначеності синтез класифікатора здійснюється з використанням навчальної вибірки D , яка складається з пар значень $x(n), z(n), n = \overline{1, N}$. Завдання класифікації при наявності декількох альтернатив може бути зведена до сукупності задач бінарної класифікації. В цьому випадку $z \in \{0, 1\}$. Рішення \hat{z} виноситься на основі апостеріорної вірогідності $P(z = 1/x)$, що обчислюється на основі теореми Байєса (8):

$$P(z = 1|x) = \frac{p(x|z=1) P(z=1)}{p(x|z=1) P(z=1) + p(x|z=0) P(z=0)} \quad (8)$$

Незважаючи на те, що для вирішення даного завдання в даний час успішно застосовуються методи статистичного аналізу, триває пошук більш ефективних алгоритмів, які б дозволили проводити класифікацію точніше і з меншими витратами. В якості таких методів пропонується використовувати апарат нейронних мереж. Застосування нейронних мереж забезпечує такі властивості систем розпізнавання образів: нелінійність, відображення вхідної інформації у вихідну, адаптивність, контекстність

інформації, паралельність обчислень, аналогію з нейробіологією, універсальність в предметних областях.

Оскільки вихідні дані представлені в вигляді сейсмограм - тимчасових відображень коливань земної поверхні, то в такому вигляді аналізувати інформацію, оцінювати різні фізичні характеристики зафіксованого події достатній точно важко. Існують методи, спеціально призначені для обробки сигналів, які дозволяють виділяти певні ознаки і, в подальшому, по ним проводити аналіз записаного події.

У сейсмограмі аналізованої події можна виділити досить багато особистих характеристик, але далеко не всі з них можуть дійсно нести інформацію, істотну для надійної ідентифікації об'єкта. Кілька ретельно відібраних ознак можуть забезпечити ймовірність помилкової класифікації істотно меншу, ніж при використанні повного набору.

При необхідності використовувати нейронно-мережеві методи для вирішення конкретних задач перше з чим доводиться стикатися - це підготовка даних. Успіх навчання нейронної мережі (НМ) може вирішальним чином залежати від того, в якому вигляді представлена інформація для її навчання. Загальний принцип попередньої обробки даних для навчання полягає в максимізації ентропії входів і виходів. Стиснення вхідних даних, зменшення ступеня їх надмірності, що використовує існуючі в них закономірності, може істотно полегшити подальшу роботу, виділяючи дійсно незалежні ознаки.

Для того щоб зрозуміти які з вхідних змінних несуть максимум інформації, а якими можна знехтувати необхідно або порівняти всі ознаки між собою і визначити ступінь інформативності кожного з них, або намагатися знайти визначені комбінації ознак, які найбільш повно відображають основні характеристики вхідних даних.

Основний принцип попередньої обробки даних - це зниження існуючої надмірності усіма можливими способами. Це підвищує інформативність прикладів і, тим самим, якість роботи НМ.

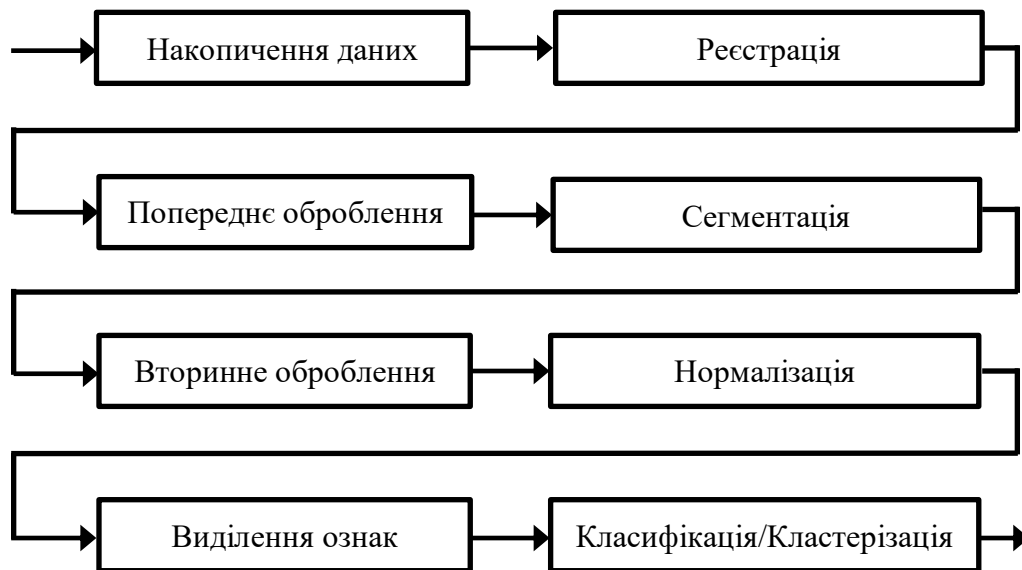
Найбільш інформативними ознаками сигналу, що використовуються для його класифікації є:

- енергія вибірки;
- скважність сигналу (якщо він має імпульсний характер)
- коефіцієнт заповнення вибірки;
- відношення енергії високочастотної частини спектра до низькочастотної.

З вихідних даних формуються дві вибірки - навчальна і тестова. Навчальна вибірка потрібна для алгоритму налаштування вагових коефіцієнтів, а наявність перевіркової, тестової вибірки потрібно для оцінки ефективності навченої нейронної мережі.

Умовний розподіл $P(z/x, w, s)$ ймовірності приналежності вектора ознак x до класу z може бути задано за допомогою НМ з радіальними базисними функціями, що характеризується вектором параметрів w та структурою s . Синтез класифікатора включає процедури визначення, як параметрів так і структури НМ. На основі байєсівської методології розроблено сімейство алгоритмів синтезу нейромережових класифікаторів, які були використані для вирішення завдання ідентифікації об'єктів в сейсмоакустичній системі. При цьому всі алгоритми забезпечували синтез класифікаторів, що характеризуються досить близькими значеннями коефіцієнтів помилок.

Зазвичай, процес розпізнавання здійснюється у кілька етапів, основні з них наведено нижче [1]:



Застосування нейронних мереж [7] забезпечує такі властивості систем розпізнавання образів: нелінійність, відображення вхідної інформації у вихідну, адаптивність, контекстність інформації, паралельність обчислень, аналогію з нейробиологією, універсальність в предметних областях.

Для навчання мереж використовують різноманітні варіанти алгоритмів випадкового пошуку. Навчання проводиться на обмеженій кількості сигналів. Мережа надасть гарні результати, якщо їй надати сигнали, що близькі до навчальних вибірок. Якщо ж надати сильно відмінний сигнал, то реакція мережі буде непередбачуваною. Якість навчання оцінюється лише за допомогою великої кількості тестів. Розмір образу, який необхідно виявити, як правило набагато разів менший, ніж розмір масиву експериментальних даних. Отже, процес розпізнавання повинен бути реалізований у вигляді віконної операції, в якій алгоритм аналізує дані в обмеженій вибірці з великого масиву (у вікні) та відбувається переміщення вікна по масиву (сканування). В результаті розпізнавання повинна бути отримана інформація про положення образу в масиві даних та його амплітуда. Ми аналізуємо, наскільки близький сигнал, що досліджується до очікуваного та знаходимо відмінності між

ними. Окрім того, ми влаштуємо інтуїтивний кордон відмінностей. Якщо різниця між образом та сигналом, що досліджується дуже велика, вважаємо, що образ не розрізняється. А якщо різниці допустимі, асоціюємо сигнал, що досліджується з образом. Фактично прирівнюємо їх один до одного, відкидаючи різницю.

Якщо використовувати підхід, що базується на теорії рішень [8], то кожний об'єкт, що розпізнається (образ) є у вигляді вектора ознак $x=(x_1, x_2, \dots, x_k)T$, де k – спільне число ознак, що визначають точку у k -мірному просторі ознак, а класифікація полягає у розбитті цього простору на області, що не перетинаються. При використанні синтаксичного підходу кожний образ описується пропозицією (або іншим подібним об'єктом деревоподібна структури) a_1, a_2, \dots, a_l , де a_i ($1 \leq i \leq l$) – синтаксичний символ (або примітив), а класифікація зводиться до аналізу синтаксису пропозиції. Такий аналіз можна проводити будь-яким шляхом співвідношення образу з еталоном або за допомогою граматичного розбору. У першому випадку потребується завдання етальонних образів, а в другому – граматичних правил. Синтаксис пропозиції дає структурну інформацію про вихідний образ, що дуже важливо в задачах розпізнавання сейсмічних сигналів.

Сейсмічні записи отримують у вигляді одномірних масивів даних (сейсмічних трас). Хоча для представлення одномірних даних пропонуються різні способи синтаксичного опису, найбільш природно представляти їх у вигляді строк примітивів. Для того щоб спростити аналіз, розділимо процедуру синтаксичного опису на три етапи: розбиття (сегментація) образів, вибір ознак та розпізнання примітивів.

Сегментація образів. Цифровий запис отримують після дискредитації неперервного сигналу, який являє собою відклик на вплив імпульсу джерела, ускладнений адитивним зовнішнім шумом. У ряді додатків, де важливу роль відіграють окремі піки і западини на записі, сегментація

сигналів проводиться у відповідності з їх формою. В інших випадках, такі особливості форми, як окремі піки і западини, не дуже інформативні, особливо при низькому співвідношенні сигнал/завада. У таких задачах зазвичай використовують сегментацію по довжині. Кожний сегмент (відрізок запису) описується набором ознак, за якими і відбувається класифікація. Відрізки можуть бути фіксованої або довільної довжини. Хоча сегментація з довільною довжиною більш ефективна та забезпечує більш високу точність представлення даних, однак пошук оптимального варіанту сегментації зазвичай являє собою важку та часомістку задачу. Якщо довжина відрізка обрана вдало, отримують адекватне представлення вихідних записів. В кожному окремому випадку при виборі довжини відрізків повинен забезпечуватись компроміс між точністю представлення і ефективністю аналізу. Чим коротше відрізок, тим простіше примітив. Однак довжина строки при цьому зростає, що призводить до зниження ефективності аналізу та росту часу обчислень (об'єм обчислень пропорційний довжині строки). Проте чим коротші відрізки, тим більше вони чутливі до шуму.

Вибір ознак. Будь-які лінійні або нелінійні перетворення вихідних даних за умови, що вони відображають відмінності між сигналами, можна розглядати як ознаки. Достатньо двох інформативних ознак: кількість нуль-перетинань та логарифм енергії. Кількість переходів через нуль наближено характеризує домінуючу частотну компоненту сигналу, а логарифм енергії – його амплітуду. Таким чином ці дві ознаки характеризують відрізок запису. Переваги синтаксичного підходу проявляється в тому, що процедура відбору ознак спрощується, так як вони визначаються за більш коротким фрагментам, чим при статистичному підході. Крім того, самі вимоги до відбору ознак тут менш жорстокі. Таким чином при синтаксичному підході в першу чергу використовується структурна інформація, а не результати складних вимірювань.

Розпізнання примітивів. Після сегментації та вибору ознак можна проводити класифікацію. Розбиття множини примітивів на класи виконується або експертом (кваліфікованим інтерпретатором), або автоматично на основі аналізу еталонних (навчальних) сегментів. У зв'язку із тим, що наявність кваліфікованого експерта можливо забезпечити не завжди та й такий спосіб може в деяких випадках виявитися ненадійним, використовують процедуру кластеризації. Ця процедура полягає в розбитті множини даних на групи схожих між собою сегментів. Ступінь схожості зазвичай визначається відстанню між сегментами, що представлені у вигляді векторів $x=(x_1, x_2, \dots, x_p)T$.

Нехай нам необхідно розбити n сегментів x_1, x_2, \dots, x_p на k окремих підмножин (кластерів) C_1, C_2, \dots, C_k . Слід відзначити, що для класифікації примітивів ми все-таки використовуємо підхід, заснований на теорії рішень. Типовий спосіб состоит у визначенні критеріальної функції, яка для кожного варіанту розбиття оцінює якість кластеризації. Тоді задача зводиться до максимізації або мінімізації критеріальної функції.

Граматичний аналіз з корекцією помилок. Граматика образів, що необхідна для проведення аналізу, може бути або сконструйована виходячи з евристичних мислень, або виведена з навчаємої (еталонної) вибірки. Помилки при сегментації та розпізнаванні примітивів, що викликані шумом та іншими факторами, на практиці виникають дуже часто. В таких випадках використовують граматичний аналіз з корекцією помилок. Для цього на початку граматики образів перетворюється в покриваючу граматику, яка генерує як правильні так і будь-які помилкові пропозиції, а для випадку детермінованої граматики аналіз з кореляцією помилок може виконуватись на основі використання критерію мінімальної відстані.

Висновок. Таким чином, зазвичай, процес розпізнавання здійснюється у кілька етапів: накопичення даних, реєстрація, попереднє

оброблення, сегментація, вторинне оброблення, нормалізація, виділення ознак, класифікація/кластерізація.

Застосування нейронних мереж є більш ефективним методом розпізнавання образів в порівнянні з іншими. Однією з основних переваг нейронних мереж є те, що всі елементи можуть функціонувати паралельно, тим самим істотно підвищуючи ефективність розв'язання поставленої задачі. Застосування нейронних мереж забезпечує такі властивості систем розпізнавання образів: нелінійність, відображення вхідної інформації у вихідну, адаптивність, контекстність інформації, паралельність обчислень, аналогію з нейробиологією, універсальність в предметних областях. Зниження існуючої надмірності усіма можливими способами є основним принципом попередньої обробки даних, це підвищує інформативність прикладів і, тим самим, якість роботи нейронних мереж.

Література

1. Neural network systems, techniques and applications in pattern recognition: Research reports B1 / Laboratory of computational engineering Helsinki university of technology; J. Lampinen. – Helsinki, 1997. – 61 p.
2. Мазуров Вл. Д. Математические методы распознавания образов: учебное пособие / Вл.Д. Мазуров. – СПб.: Питер, 2010 г. – 101 с.
3. Нейронные сети. Statistica Neural Networks. Методология и технологии современного анализа данных / Под редакцией В.П. Боровикова. - 2-е издание, переработанное и дополненное – М.: Горячая линия – Телеком, 2008. – 392 с.
4. Нейронные сети. Полный курс. Второе издание / Саймон Хайкин – Перевод с английского Н.Н. Куссуль, А.Ю. Шелестова – М.: Издательский дом "Вильямс", 2008. – 1104 с.

5. Искусственный интеллект: современный подход / Стюарт Рассел, Питер Норвиг – Перевод с английского К.А. Птицын – М.: Издательский дом “Вильямс”, 2006. – 1408 с.
6. <http://www.deeplearningbook.org/>
7. Когнітивні методи кібернетики: Навчальний посібник / О.С. Бичков. – К.: Видавничо-поліграфічний центр центр “Київський університет”, 2006. – 127с.
8. Анализ и выделение сейсмических сигналов / Под редакцией Ч. Чжання – Перевод с английского А.Л. Малкина, Г.Н. Гогоненкова. – М.: Мир, 1986. – 240 с.