

МЕТОДИ ПОПЕРЕДНЬОЇ ОБРОБКИ ІНФОРМАЦІЇ В ПРОЦЕСІ РОЗПІЗНАВАННЯ ОБРАЗІВ

**В. В. Хом'юк, асист. кафедри прикладної математики,
Вінницький державний технічний університет,
тел. 8-0432-440-125**

Розпізнавання образів і сигналів в останні роки стало одним з найбільш популярних і «прикладних» напрямків прикладної математики. Це перетворення являється наслідком дії цілого ряду факторів, зокрема природності зведення множини різноманітних задач обробки та аналізу даних до задачі розпізнавання, інтересу до вивчення складних або погано формалізованих систем, для яких не вдається синтезувати моделі традиційними математичними засобами через недоліки апріорної інформації, виникнення в розпізнаванні цікавої внутрішньої математичної проблематики [5,7]. Звертання до методів і прийомів розпізнавання - характерна риса досліджень, пов'язаних з математичною обробкою інформації в тих випадках, коли точні математичні моделі досліджуваних об'єктів, явищ, процесів неможливо будувати або неможливо реалізувати за допомогою існуючих обчислювальних засобів, а розв'язок шукається на базі правдоподібних евристичних міркувань. Сьогодні розпізнавання досягло такої стадії розвитку, що виникла необхідність, а також і можливість, переходити від опису окремих евристичних процедур до вивчення моделей таких процедур, до підстановки відповідних оптимізаційних задач і т. д.

Терміном «розпізнавання образів» позначається поле діяльності, яке пов'язане як із реальними життєвими потребами (розпізнаванням людьми і тваринами «образів» у їх навколишньому середовищі), так і з вирішенням наукових і технічних задач за участю людини та автоматичних систем. Розпізнавання образів завжди передбачає наявність деякої системи обробки інформації, що має вхід і вихід. На вхід системи дані можуть надходити від безлічі різних джерел: фізичного об'єкта, що знаходиться в деякому середовищі, деякого процесу або як результат експерименту. Дані, що поступають на вхід системи розпізнавання, як правило, дуже складні і включають сигнали, які являються просторовими і/або часовими функціями. Вихідна інформація порівняно проста: вона зводиться до вказівки одного з декількох класів. Наприклад, якщо вхідною інформацією служать результати виміру коефіцієнта відображення деякої поверхні, які отримані за допомогою сканування, то вихідна інформація вказує назви символів, які нанесені на цю поверхню.

Процес попередньої обробки має забезпечувати перетворення та стиск даних (часто в декілька етапів) для того, щоб визначити властивості або ознаки, які найбільш притаманні образам, що розглядаються. Одним з цих процесів являється процес глобальних перетворень[5], при яких вхідні дані використовуються для визначення вихідного значення. Глобальні перетворення застосовуються для отримання ряду характеристик на основі вхідних даних у цілому для того, щоб сформувані ознаки, усунути надмірність, виправити систематичні помилки (наприклад, за допомогою відновлення зображення) і т.д. Отже, деякий набір функцій $y_1(x_1, \dots, x_n), \dots, y_m(x_1, \dots, x_n), m \leq n$, забезпечує отримання значень $y_i, i = 1, 2, \dots, m$, на основі усіх вхідних значень $x_j, j = 1, 2, \dots, n$. У реальних

задачах використовуються головним чином лінійні функції[6] і тому $y_i = \sum_{j=1}^n c_{ij} x_j$ або

$y = Cx$, де x і y - вектори, компонентами яких являються x_j і y_i , а C - деяка матриця. Застосуванню лінійних перетворень може передувати лінійна чи нелінійна

зміна масштабу x , наприклад, логарифмування; подібна операція може згодом бути застосована до y [1]. Якщо в перетворенні $y = Cx$, матриця C - неособлива, $m = n$ і C^{-1} - обернена матриця до матриці C , то $x = C_1^{-1}y_1 + C_2^{-1}y_2 + \dots + C_n^{-1}(n)y_n$, де C_i^{-1} - відповідний стовпець оберненої матриці C^{-1} ; $i = \overline{1, n}$. Цю процедуру можна розглядати як розклад або розбиття вхідних даних x по нормальним функціям або базисним векторам y_i . Використовуються базисні вектори багатьох типів в залежності від математичної і фізичної природи задачі, а також зручності реалізації обчислювальних процедур [1]. У випадку усіченого розкладання (як це робиться при стиску даних) тип і число m ($m < n$) базисних векторів визначають точність розбиття. Оскільки вхідні дані часто мають дуже великий об'єм (особливо при роботі із зображеннями), зручність реалізації обчислювальних процедур іноді стає основною вимогою. Ця обставина стимулювала розробку спеціалізованих швидкодіючих процедур, наприклад швидкого (дискретного) перетворення Фур'є (ШПФ) [2]. Завдяки зручності обчислень можна також здійснювати розбиття матриці C за більше число (але більш простих) кроків і використовувати унітарні, ермітові матриці, а також інші різновиди нормальних матриць [1]. Розглянемо основні розклади матриці на множники. Елементарні унітарні та неунітарні перетворення широко використовуються для отримання самих різних розкладів матриці на множники. Серед цих розкладів найбільш важливими являються розклад на два трикутних множника і на трикутний та унітарний множники. При цьому слід зауважити, що неможливо дати однозначну відповідь на питання "Який з розкладів являється кращим?", оскільки різні задачі ставлять до розкладів різні вимоги. Для порівняння розглянемо деякі розклади квадратних матриць (табл. 1), за умови, що дані для процесу ортогоналізації при перетворенні Фур'є вказані за відсутністю додаткової переортогоналізації [3]:

Таблиця 1

Спосіб отримання співмножників	Режим обчислень	Число операцій	Точність	Додаткова пам'ять
Метод Гауса	f^1	$(2/3)n^3$	αn	0
Компактна схема	f^2	$(2/3)n^3$	β	0
Метод квадратного кореня	f^2	$(1/3)n^3$	1,0	0
Метод обертання	f^1	$2n^3$	2,9 n	0
Метод відображення	f^2	$(4/3)n^3$	2,9 n	2 n
Нормалізований метод відображення	f^2	$(4/3)n^3$	2,9 n	3 n
Метод ортогоналізації	f^2	$2n^3$	1,0	0,5 n^2
Нормалізований метод обертання	f^1	$2n^3$	2,9 n	n

В даній таблиці параметри α, β визначають зростання елементів; символ f^1 означає, що для досягнення відповідної точності можна обмежитись обчисленнями з одинарною точністю; символ f^2 означає, що необхідно використовувати операції накопичення. Загальний час для отримання розкладу визначається числом арифметичних операцій, які необхідно при цьому виконати. Зауважимо, що задачі з матрицями високих порядків вимагають для свого розв'язування значних ресурсів пам'яті ЕОМ, одним з шляхів економії якої являється розміщення інформації про співмножники на місці початкової матриці, в основному на місцях, де отримуються нульові елементи.

Якщо матриця C - нормальна, то вона може бути представлена таким чином: $C = U \Lambda U^{-1}$, де Λ - діагональна матриця характеристичних чисел матриці C , а стовпцями матриці U являються власні вектори матриці C . Деяка матриця U застосовується для представлення різних матриць C , якщо усі вони перестановочні з деякою нормальною матрицею C_0 , характеристичні числа якої різні. Важливий приклад - випадок циклічної перестановки

$$C_0 = \begin{bmatrix} 0 & 1 & . & . & 0 \\ 0 & 0 & 1 & . & 0 \\ . & . & . & . & . \\ 0 & 0 & . & . & 1 \\ 1 & 0 & . & . & 0 \end{bmatrix}, \text{ яка являється апроксимацією зсуву } \begin{bmatrix} 0 & 1 & . & . & 0 \\ 0 & 0 & 1 & . & 0 \\ . & . & . & . & . \\ 0 & 0 & . & . & 1 \\ 0 & 0 & . & . & 0 \end{bmatrix}.$$

Багато операцій обробки сигналів являються інваріантними щодо зсуву[6] (не залежать від часу), тобто $CC_0 = C_0C$, і, отже, їх власні вектори – комплексні показникові функції. Матриці всіх подібних операцій «фільтрації» - циркулянтні; а відповідна матриця U представляє собою перетворення Фур'є. У випадку двовимірних сигналів або зображень часто виявляється можливим представити матрицю C як кронекеровий добуток матриць, що суттєво спрощує розв'язування відповідних задач [4]. В основі багатьох процедур обробки лежить ковариційна матриця K векторів вхідних даних, яка являється ермітовою[4]. У випадку векторів сигналів x з нульовим математичним сподіванням розбиття x за власними векторами ковариційної матриці K з використанням деякої матриці U приводить до деякої дійсної діагональної матриці $\Lambda = U^{-1}KU$, яку можна розглядати також у якості ковариційної матриці $U^{-1}x$; при цьому перетворені дані некорельовані. Це перетворення Карунена-Лоєва, яке цілком залежить від властивостей вхідного сигналу і тому не може вибиратися довільно; часто воно викликає ускладнення при реалізації обчислень [5]. Якщо в кожній циклічній перестановці з'являються усі вхідні сигнали, то ковариційна матриця K являється циркулянтною і таке перетворення збігається із перетворенням Фур'є (вінеровська фільтрація). Прикладами процедур, заснованих на використанні ковариційної матриці K являється узагальнена вінеровська фільтрація [1] і стиск даних, що забезпечує мінімум середньоквадратичної помилки (усічене розкладання по власним векторам ковариційної матриці K з найбільшими характеристичними числами забезпечує мінімізацію середньоквадратичної помилки перетворених даних). У декількох практично використовуваних унітарних матрицях елементами стовпців являються нулі, одиниці, мінус одиниці або невелике число коефіцієнтів, що забезпечує можливість застосування простих алгоритмів (перетворення Адамара, Уолта,

Хаара [1]). Часто вони використовуються для отримання наближено діагональних матриць, які дозволяють спростити обчислення.

Зовсім іншою задачею являється розкладання даних по мінімальному числу стандартних векторів, які вибираються з деякої залежної множини. Вона виникає, наприклад, при описі кривих, що характеризують спектральну щільність і щільність розподілу хромосом, за допомогою деякої суми кривих нормального розподілу, за допомогою яких, як впливає з фізики, потрібно представляти ці дані. Різні методи обробки даних мають з'ясувати ті закономірності, яким підпорядковуються явища або об'єкти, що вивчаються.

Одним з методів, які використовуються у розпізнаванні образів, являється також і кластерний аналіз. Кластерним називають аналіз неklasифікованих об'єктів, який проводиться з метою виділення структур, класів, образів, множини подібних об'єктів і т.п. Дотепер немає точного визначення кластерного аналізу. Однією з причин цього являється та обставина, що кластерний аналіз використовується та удосконалюється в настільки різних областях, як розпізнавання образів та сигналів, біомедицина, фізика, тощо. У кластерному аналізі не існує однозначного кількісного критерію, подібного до помилки класифікації в дискримінантному аналізі. Такий критерій не можна сформулювати, оскільки в різних прикладних задачах різною може бути і мета аналізу. Іноді необхідно виділити групи із високою щільністю розподілу і малою дисперсією, а іноді - знайти зв'язні точкові структури. У літературі описані сотні методів кластеризації. По суті, кластерний аналіз тісно пов'язаний з дослідженнями процесу породження даних. Багато з тих, хто хоче використовувати прийоми цього аналізу, самі створюють нові методи кластеризації. Так і в задачах розпізнавання образів та сегментації зображень широко використовуються методи кластерного аналізу, особливо в тих випадках, коли про дані є досить мало апріорної інформації. Саме тому методи кластеризації відіграють важливу роль в даних процесах. Більшість алгоритмів кластеризації [6] базуються на трьох методах. До першого відносяться методи, які основані на відшуванні моди розподілу. Ідея цих методів полягає у тому, що кластери відповідають максимумам щільності розподілу джерела, що породжує дані. Кожний максимум щільності розподілу відповідає деякому кластеру. Кожний об'єкт повинен бути віднесений до одного з кластерів. В результаті деякі кластери можуть бути об'єднані. Іншими словами відбувається кластеризація за методом мінімальної квадратичної похибки. Алгоритми кластеризації на основі мінімуму квадратичної похибки містять велике число матричних та векторних перетворень [6], для реалізації яких надають перевагу систолическим архітектурам. Якщо число кластерів відомо, можна скористатися другим різновидом аналізу - методом, у якому в якості критерію використовується відношення внутрішньої кластерної дисперсії до міжкластерної. Можуть використовуватися і інше критерії. Майже у всіх посібниках із кластерного аналізу [6] описується ітеративний метод аналізу даних, який самоорганізується. В його основу покладена ідея про те, що об'єкти належать кластеру із найбільш близьким середнім значенням. Розглянемо відповідну процедуру, яка являється збіжною. Середні значення для початкової розбивки вибираються випадковим образом, але так, щоб центри були оточені об'єктами. На другому кроці процедури проводиться віднесення об'єктів до тих кластерів, центри яких знаходяться від них ближче усього. Після цього підраховуються нові середні значення для кластерів і повторюється другий крок процедури. Остання швидко сходиться. Результат кластеризації може залежати від вибору початкових середніх значень для кластерів. Якщо кластери, які отримано в результаті застосування цієї процедури, не являються "істинними", то різне значення вихідних середніх може призвести до різних результатів кластеризації. Якщо число кластерів апріорі невідомо, процедуру варто повторити для різного числа кластерів. І нарешті, відзначимо ієрархічні схеми кластеризації. При їх використанні, як правило, спочатку кожен об'єкт розглядається як окремий кластер. Потім два найближчих (найбільш схожих) кластери об'єднують; цей крок повторюється до отримання

єдиного кластера. Використання цього методу припускає, що деяким чином визначений спосіб виміру відстані між кластерами, наприклад відстанню між двома кластерами можна вважати відстань між середніми значеннями кластерів, між двома найближчими об'єктами або між двома найбільш відпаленими об'єктами. Існує багато інших способів завдання відстані, і кожний з них породжує свій різновид кластерного аналізу. У результаті застосування цього методу будується деякий різновид дерева, який називаний часто «дендрограмою»[7]; остання відображає процес послідовного об'єднання кластерів з урахуванням відповідної міри кластеризації. Остаточний варіант кластеризації вибирається по цій дендрограмі вручну. Задача знаходження оптимального числа кластерів, на основі яких можна розбити вхідний набір даних, являється достатньо складною і займає важливе місце в теорії кластеризації [5]. У статистичному розпізнаванні образів об'єм вхідних даних має велике значення. Якщо об'єм вибірки малий, то властивості класів оцінюються грубо і варто застосовувати лише глобальні методи, що вимагають визначення незначного числа параметрів. Як правило, це зводиться до використання лінійного розподілу і простих методів вибору ознак, причому методи кластерного аналізу дозволяють виділити тільки дуже явно виражені кластери [7]. Однак вибірки великого об'єму важко не тільки отримувати, але й обробляти. Для роботи з ними потрібні великі ЕОМ і багато машинного часу. Тому маленький набір ознак корисний із статистичної точки зору. Передбачається, що невелике число змінних вибирається за допомогою засобів нестатистичного характеру і що цього набору достатньо для вирішення задачі розпізнавання. Тому відпадає необхідність у виборі або виділенні ознак. Якщо додати до наявного набору нові ознаки, то розділяюча сила множини ознак може збільшитися і відповідно покращиться якість розпізнавання. Можливий і протилежний ефект: збільшення числа ознак означає збільшення розмірності і, отже, числа параметрів, що вимагають оцінювання. Унаслідок цього помилки оцінювання будуть зростати, що може призвести до зниження якості розпізнавання. Отже, як висновок - існує оптимальний набір ознак. Як при стиску, так і при розширенні цього набору знижується якість розпізнавання. Це явище називають утворенням гострих максимумів [7]. Одна з головних проблем, пов'язаних з утворенням гострих максимумів, полягає в тому, що навряд чи можливо на підставі заданої початкової множини вибрати оптимальну підмножину ознак. Методи вибору ознак часто не враховують помилок оцінювання параметрів. Отже, необхідно створювати нові методи вибору ознак, що враховують існування гострих максимумів.

Таким чином, задачі, які пов'язані з процесом "розпізнавання образів" дуже різноманітні. При цьому вузьким місцем вдалих реальних систем розпізнавання являється необхідність виконання великого об'єму обчислень, особливо при обробці зображень та швидкодія, яка необхідна для систем, які працюють у реальному часі. Саме тому слід приділяти увагу таким концепціям, як попередня, паралельна, матрична та потокова обробка інформації (особливо стосовно до роботи у реальному часі та обробці великих об'ємів даних за обмеженого часу). І оскільки останнім часом активно ведуться роботи, пов'язані з побудовою моделей обробки інформації у нервовій системі, більшість яких оснований на схемі формального нейрону

У. МакКаллока і У. Піттса, згідно якої нейрон являє собою пороговий елемент, то постає задача створення "особливої методики" для методів попередньої обробки інформації з подальшим їх використанням, наприклад, у нейронних системах, в задачах навчання машин розпізнаванню образів. Отже, питання дослідження перетворення початкових даних при побудові подібних моделей залишається актуальним.

Література

1. Воеводин В. В. Математические модели и методы в параллельных процессах. - М.: Наука, 1986. - 296 с.

2. Кун С. Матричные процессоры на СБИС: Пер. с англ. - М.: Мир, 1991. -672с.
3. Матрицы и вычисления. Воеводин В. В., Кузнецов Ю. А. - М.: Наука. Главная редакция физико-математической литературы, 1984. - 320 с.
4. Сверхбольшие интегральные схемы и современная обработка сигналов: Пер. с англ./ Под ред. С. Гуна, Х. Уайтхауса, Т. Кайлата. - М.: Радио и связь, 1989.-472 с.: ил.
5. Структурные методы обработки эмпирических данных. Браверман Э. М., Мучник И. Б. - М.: Наука. Главная редакция физико-математической литературы, 1983. -464 с.
6. СБИС для распознавания образов и обработки изображений: Пер. с англ./Под ред. К. Фу.- М.: Мир, 1988. - 248с., ил.
7. Распознавание образов: состояние и перспективы: Пер. с англ./ К. Верхаген, Р. Дейн, Ф. Грун и др. - М.: Радио и связь, 1985. - 104 с., ил.

SUMMARY

УДК 681.3

Методы предварительной обработки информации в процессе распознавания образов / В. В. Хом'юк // Вісник Сумського державного університету. – 2002. - № __ . – с.

Статья представляет собой изложение теоретических положений о некоторых методах предварительной обработки информации в процессе распознавания образов. В частности, рассматриваются разложение произвольной матрицы на множители и методы кластеризации исходных данных.

УДК 681.3

Методи попередньої обробки інформації в процесі розпізнавання образів / В. В. Хом'юк // Вісник Сумського державного університету. – 2002. - № __ . – с.

Стаття представляє собою виклад теоретичних положень про деякі методи попередньої обробки інформації в процесі розпізнавання образів. Зокрема, розглядаються розклад матриць на множники і методи кластеризації початкових даних.

UDK 681.3