

Теорія розпізнавання образів

Література

1. Форсайт Д., Понс Ж. — Компьютерное зрение. Современный подход. — М.:, 2004.
2. Шапиро Л., Стокман Дж. — Компьютерное зрение. — М.:, 2006.
3. Ту Дж., Гонсалес Р. Принципы распознавания образов. — М.: Мир, 1978.
4. Фисенко В.Т., Фисенко Т.Ю. — Компьютерная обработка и распознавание изображений. — СПб.:, 2008.
5. Васильев В.И. — Распознающие системы. — К.: Наукова думка, 1983.
6. Анисимов Б.В., Курганов В.Д., Злобин В.К. — Распознавание и цифровая обработка изображений, 1983.
7. Аркадьев А.Г., Браверман Э.М. Обучение машины распознаванию образов. М.: Наука, 1964.

Основні поняття теорії розпізнавання образів

Теорія розпізнавання образів — розділ кібернетики, що розвиває теоретичні основи й методи класифікації і ідентифікації предметів, явищ, процесів, сигналів, ситуацій і т.п. об'єктів, які характеризуються скінченним набором деяких властивостей і ознак. Такі задачі вирішуються досить часто, наприклад, при переході або проїзді вулиці за сигналами світлофора. Розпізнавання кольору лампи світлофора, що засвітилася, і знання правил дорожнього руху дозволяє прийняти правильне рішення про те, можна, чи не можна переходити вулицю в цей момент.

У більшості випадків, сприймаючи явища навколишнього світу, людина здійснює їх класифікацію, тобто розбиває ці явища (предмети, ситуації) на групи схожих явищ (саме схожих, а не тотожних). По тим чи іншим причинам необхідно віднести в одну групу у чомусь "подібні" явища чи предмети, які можуть при цьому значно відрізнятися один від одного. Наприклад, усі фігури, зображені на наступному малюнку, ми називаємо "літерою А", незважаючи на їх значну відмінність у їх написанні.



Рис. 1.

Суттєвим є той факт, що виділивши такі групи (множини) об'єктів, ми отримуємо здатність "упізнавати", тобто встановлювати належність до вже відомої множини, нові об'єкти, які ще не зустрічалися нам раніше, наприклад впізнавати букви, написані новим для нас почерком.

Отримавши уявлення про те, що являє собою буква "А" на основі деякої, зазвичай невеликої кількості екземплярів цієї букви, ми спроможні упізнати як завгодно велику кількість інших її екземплярів.

Проте, далеко не всі множини об'єктів дають змогу на основі невеликої частини множини упізнати як завгодно багато інших невідомих нам її представників. Наприклад, фотографії студентів деякого вузу утворюють множину. Проте неможливо після ознайомлення зі скажімо десятьма фотографіями студентів визначити по новій фотографії є людина студентом цього вузу чи ні.

Таким чином існують множини деякого особливого типу. Ці множини мають *характерну властивість*, яка виявляється у тому, що *після ознайомлення із скінченною частиною об'єктів цих множин, можна упізнавати* як завгодно велику кількість інших їх представників. Множини такого типу будемо називати **образами**.

Прикладами образів можуть бути такі множини: чоловіки, дитячі портрети, ссавці, картини Пікассо, цифри 5, зображення винищувачів МІГ. Застосовувати до них термін "образ" ми можемо тому, що ознайомлення з образом не пов'язано із запам'ятовуванням окремих об'єктів, а упізнавання нового об'єкта відбувається без безпосереднього порівняння із кожним раніше відомим.

Характерна властивість образів об'єктивна у тому сенсі, що різні люди (живі істоти), які навчалися на різних групах об'єктів образу, у переважній більшості однаково і незалежно один від одного класифікують одні і ті самі нові об'єкти. Саме об'єктивність цієї властивості образів дає змогу людям, які вчилися у різних школах, успішно розпізнавати раніше невідомий їм почерк.

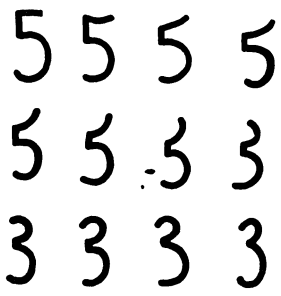


Рис. 2.

Проте, самі образи у той самий час є у певній мірі "розпливчастими" і питання про належність об'єкта до даного образу не завжди може бути вирішене однозначно. На рис. 2 показано, як за допомогою невеликих послідовних змін можна перетворити цифру 5 у цифру 3.

Рис. 2 ілюструє важливу особливість об'єктів кожного образу. Можна в деяких межах змінювати об'єкт, але він усе одно залишиться об'єктом того самого образу.

Створення штучних систем *розпізнавання образів* залишається складною теоретичною й технічною проблемою. Необхідність у такому розпізнаванні виникає в самих різних областях — від військової справи й систем безпеки до оцифрування різних аналогових сигналів.

Сприйняття явищ у формі образів відіграє надзвичайно важливу роль у процесах пізнання зовнішнього світу. У процесі біологічної еволюції багато тварин за допомогою зорового й слухового апарата вирішили задачу *розпізнавання образів* досить добре. Як впливає із самого означення образу "розпізнаванню" нових для нас об'єктів передуює процес навчання. Під час навчання істоти ознайомлюються із деякою кількістю об'єктів і, крім цього, із якогось джерела (наприклад від батьків, старших і т.п.) отримують інформацію про те, до якого образу відноситься кожний із цих об'єктів. Цей процес отримав назву "**навчання з учителем**".

Більш загальний характер має "**навчання без учителя**", у процесі якого система вчиться спонтанно виконувати поставлене завдання без втручання з боку "вчителя". Навчання машин без вчителя формулюється як задача **кластерного аналізу**. Вибірка об'єктів розбивається на кластери (множини, що мають порожній перетин), таким чином, що кожний кластер складається із "схожих" об'єктів, а різні кластери "суттєво" відрізняються один від одного. Кластеризація часто використовується в якості допоміжного засобу розв'язування задач класифікації та регресійного аналізу. Деякі алгоритми розв'язування задач класифікації комбінують навчання з учителем та навчання без учителя (наприклад навчання мереж векторного квантування).

Традиційно задачі розпізнавання образів включають у коло задач **штучного інтелекту**. Можна виділити два основних напрямки:

- Вивчення здібностей до розпізнавання, якими володіють живі істоти, їхнє пояснення й моделювання;
- Розвиток теорії й методів побудови пристроїв, призначених для розв'язання окремих задач у прикладних цілях.

Формальна постановка задачі. Розпізнавання образів — це віднесення вихідних даних до певного класу за допомогою виділення істотних ознак, що характеризують ці дані, із загальної маси несуттєвих даних. При постановці задач розпізнавання намагаються користуватися математичною мовою.

Підходи до розпізнавання образів. Для оптичного розпізнавання образів можна застосувати метод перебору вигляду об'єкта під різними кутами, масштабами, зсувами й т. д. Для букв потрібно перебирати шрифт, властивості шрифту й т. д.

Другий підхід — знайти контур об'єкта й досліджувати його властивості (зв'язність, наявність кутів і т. д.)

Ще один підхід — використовувати штучні нейронні мережі (багатошарові перцептрони, мережі квантування, мапи Кохонена, рекурентні мережі). Цей метод вимагає або великої кількості прикладів задачі розпізнавання (із правильними відповідями), або спеціальної структури нейронної мережі, що враховує специфіку даної задачі.

Перцептрон як метод розпізнавання образів

Ф. Розенблатт уводячи поняття про *модель мозку*, завдання якої полягає в тому, щоб показати, як у деякій фізичній системі, структура й функціональні властивості якої відомі, можуть виникати психологічні явища та описав найпростіші *експерименти з розрізнення*. Дані експерименти цілком стосуються до методів розпізнавання образів, але відрізняються тим, що алгоритм розв'язання не детермінований.

Найпростіший експеримент, на основі якого можна одержати психологічно значиму інформацію про деяку систему, зводиться до того, що моделі пред'являються два різних стимули й потрібно, щоб вона реагувала на них різним чином. Метою такого експерименту може бути дослідження можливості спонтанного розрізнення стимулів системою при відсутності втручання з боку експериментатора, або,

навпаки, вивчення примусового розрізнення, при якому експериментатор прагне навчити систему здійснювати необхідну класифікацію.

У досвіді з навчанням перцептронів зазвичай пред'являється деяка послідовність образів, у яку входять представники кожного із класів, що підлягають розрізненню. Відповідно до деякого правила *модифікації пам'яті* правильний вибір реакції *підкріплюється*. Потім перцептронів пред'являється контрольний стимул і визначається ймовірність одержання правильної реакції для стимулів даного класу. Залежно від того, збігається чи не збігається обраний контрольний стимул з одним з образів, які використовувалися в навчальній послідовності, отримують різні результати:

1. Якщо контрольний стимул не збігається з жодним із навчальних стимулів, то експеримент пов'язаний не тільки з *чистим розрізненням*, але містить у собі й елементи *узагальнення*.
2. Якщо контрольний стимул збуджує деякий набір сенсорних елементів, цілком відмінних від тих елементів, які активізувалися при впливі раніше пред'явлених стимулів того ж класу, то експеримент є дослідженням *чистого узагальнення*.

Перцептрони не мають здатності до чистого узагальнення, але вони цілком задовільно функціонують в експериментах із розрізнення, особливо якщо контрольний стимул досить близько збігається з одним з образів, щодо яких перцептрон уже нагромадив певний досвід.

Приклади задач розпізнавання образів

- Розпізнавання літер.
- Розпізнавання штрих-кодів.
- Розпізнавання автомобільних номерів.
- Розпізнавання осіб.
- Розпізнавання мови.
- Розпізнавання зображень.
- Розпізнавання локальних ділянок земної кори, у яких знаходяться родовища корисних копалин.

Розпізнавання образів застосовується в наступних областях:

- Біоінформатика: пошук шаблонів в ДНК.
- Бази даних: пошук і класифікація .
- Обробка текстів: тематична класифікація .
- Аналіз зображень: розпізнавання символів, робота з картами, розпізнавання осіб, поділ об'єктів.
- Виробництво: контроль якості (візуальна перевірка коректності мікросхем) .
- Пошук по мультимедіа: визначення жанрів.
- Біометрія: ідентифікація людини за відбитками пальців, по райдужній оболонці ока .
- Прогнозування: погода, сейсмологія, геологія.
- Обробка мови: переклад аудіо-сигналів в текст.

Повернемося до означення. *Розпізнавання образів* — процес віднесення об'єкта з фіксованою групою його властивостей до одного об'єкту з множини образів за задалегідь обумовленим правилом. Наприклад, рибу у тенетах треба поділити на окунів і лососів. Припустимо, що це робиться по довжині риби. Тобто у нас є об'єкт "риба", і за значенням властивості "довжина" ми відносимо рибу або до образу "лосось", або до образу "окунь".

Процедура розпізнавання

Виділимо найбільш важливі кроки у процедурі розпізнавання:

1. Сприйняття образу.
На цьому етапі проводять отримання значень характеристичних властивостей об'єкта (вимірювання лінійних вимірів, фотографування, оцифровка звуку).
2. Попередня обробка.
Видалення шумів, представлення зображення в чорно-білому варіанті, обрізання непотрібних частин зображення.
3. Виділення характеристик (індексація).
На цьому етапі вимірюються характеристичні властивості об'єкта (вимірюємо довжину риби та її колір).
4. Класифікація (прийняття рішення).

Розробка системи розпізнавання

- Отримати навчальну вибірку (тренувальну колекцію) (навчальна вибірка — множина об'єктів, для яких відомі їхні образи.
Наприклад колекція аудіо записів для кожного звуку, або колекція зображень кожної букви латинського алфавіту).
- Вибрати модель представлення об'єктів.
- Вибрати значущі характеристики.
Це один з найважливіших етапів розробки системи розпізнавання. Наприклад, якщо у випадку ідентифікації риби окунь / лосось в якості характеристики вибрати тільки довжину риби, то ніяке правило класифікації не зможе точно визначити тип риби, оскільки можна зустріти лосося і окуня однакової довжини.
- Розробити правило класифікації.
Правило класифікації — правило, яке за значеннями характеристичних властивостей об'єкта відносить його до одного з образів.
- Навчання розпізнавальної системи.
На цьому етапі *алгоритм навчання* "збирає досвід" на основі розпізнавання навчальної вибірки. Для того, щоб правильно виставити коефіцієнти (параметри) розпізнавальної системи, алгоритм навчання застосовують на навчальній вибірці, контролюючи результат роботи алгоритму.
- Перевірка якості навчання. Повернення назад до кроків 2 (3, 4) ...
Якщо частота помилок алгоритму не влаштовує, то необхідно повернутися до п. 2, 3, 4. Інтуїтивно зрозуміло, що збільшення кількості характеристичних властивостей, збільшення тренувальної колекції покращують якість роботи алгоритму.
- Оптимізація розпізнавальної системи.
Після того, як якість роботи системи розпізнавання підходить під умову розглянутої задачі, іноді доводиться провести його оптимізацію. Початковий алгоритм розпізнавання може бути занадто довгим або ресурсоємним. Прискорити процедуру розпізнавання можна зменшивши кількість характеристичних властивостей об'єкта, вибравши інші характеристичні властивості, використовуючи інше правило класифікації.

Методи розпізнавання

Виділяють 4 групи методів розпізнавання:

- Порівняння із зразком.
Застосовуємо геометричну нормалізацію і вважаємо відстань до прототипу. Найбільш наочно застосування цього методу в розпізнаванні тексту.
Завдання. У нас є зображення відсканованого символу і колекція зображень зразків (усіх букв абетки), ми хочемо визначити, який букві алфавіту відповідає відскановане зображення. Розв'язок. Масштабуємо зображення символу до розмірів зразків і виберемо той, відстань до якого мінімальна.
- Нейронні мережі.
Вибираємо вид мережі і налаштовуємо коефіцієнти. На вхід нейронної мережі подається об'єкт для розпізнавання. Група рецепторів мережі відповідає за прийом своєї характеристичної властивості.
- Статистичні методи.
- Структурні та синтаксичні методи.
Розбираємо об'єкт на елементи. Будуємо правило, в залежності від входження окремих елементів та їх послідовностей.
Приклад. Припустимо, нам необхідно визначити вид риби, поданої на вхід. Спростимо задачу до випадку 2 видів (окунь і лосось). Перший розв'язок, який приходить в голову — зробити припущення про вигляді риби по її довжині.

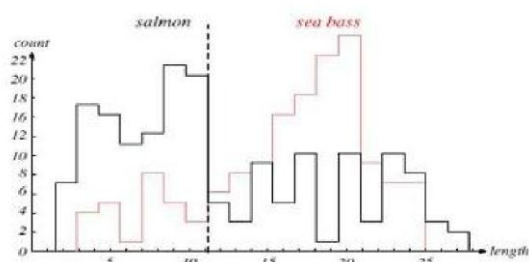


Рис. 3.

Інша характеристична властивість — забарвлення. Керуючись лише однією характеристикою не можливо точно визначити вид риби.

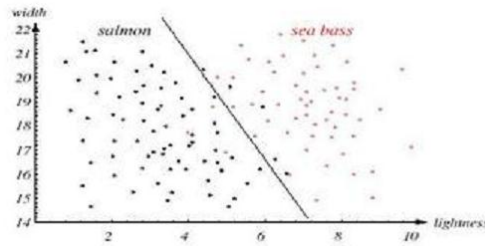


Рис. 4.

Але керуючись обома ознаками, можна визначити вид риби більш точно з використанням гiперплощини.

Перетворення зорових образів у цифровий код

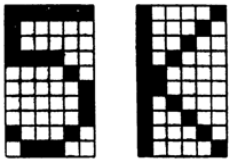


Рис. 5

Зображення розташовується у деякому прямокутному полі, розбитому на однакові квадрати (елементи поля або рецептори). Квадрати, на які потрапило зображення, цілком зафарбовуються у чорний колір і позначаються одиницею. Інші елементи позначаються нулем. Якщо пронумерувати елементи зліва направо і зверху униз, то цифри 5 та літери "К" на рис. 5 будуть відповідати коди

«5» — 11111110000010000011110000001000000100000100000
1100010011100

«К» — 1000011000101001001010001100001100001010001001
00100010100001.

У загальному випадку код фігури записується у вигляді x_1, x_2, \dots, x_n , де $x_i \in \{0,1\}$, $i = \overline{1, n}$, n — кількість елементів поля рецепторів.

Точка у просторі рецепторів називається *граничною точкою* множини, якщо у її коді є хоча б одна цифра, зміна якої переводить точку у іншу множину. У протилежному випадку точка називається *внутрішньою*.

Якщо для множини точок виконуються умови

- число граничних точок невелике порівняно із їх загальним числом,
- довільні дві внутрішні точки множини можуть бути з'єднані достатньо плавною лінією, яка проходить лише через внутрішні точки,
- майже кожна внутрішня точка має достатньо великий окіл, який цілком міститься у множині,

то така множина точок називається **компактною множиною**.

Описаний метод кодування не враховує колір і тому придатний лише для чорно-білих зображень. У загальному випадку кожний розряд коду може приймати значення з деякого відрізка $[a, b]$.

Гіпотеза компактності образів: Простому зоровому образу відповідає компактна множина точок у просторі рецепторів. Поняття компактності проілюстровано на рис. 6.

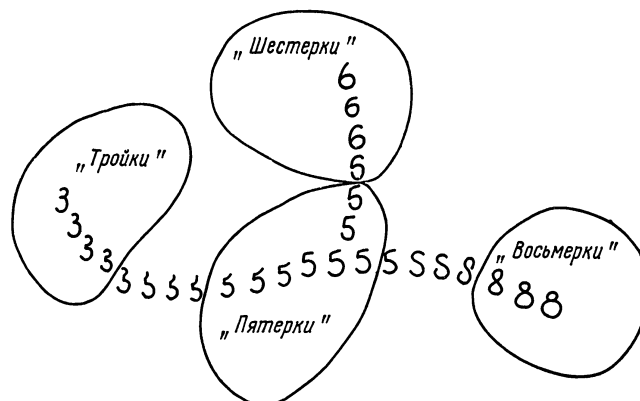


Рис. 6.

Гіпотеза компактності дає змогу звести задачу навчання розпізнавання образів до задачі побудови відокремлюючих поверхонь по відомим точкам кожного образу. У результаті простір рецепторів розбивається на кілька областей, про кожну з яких відомо, що вона переважно містить точки, які відповідають певному образу. Якщо тепер потрібно розпізнати нову невідому фігуру, то потрібно перетворити цю фігуру у точку простору рецепторів і визначити, у якій частині простору ця точка розташована.

Метод січних площин

Алгоритм навчання, заснований на методі січних площин, полягає у відокремленні образів за допомогою частин гіперплощин. Алгоритм складається з наступних етапів.

1. Навчання (формування гіперплощин для відокремлення множин):

- a. проведення січних площин;
- b. вилучення зайвих гіперплощин;
- c. вилучення зайвих частин площини.

2. Розпізнавання нових об'єктів.

Припустимо, що потрібно навчити комп'ютер розпізнавати 3 образи a , b , та c .

Проведення січних площин. У комп'ютер вводять коди двох точок, які належать різним образам та проводять довільну пряму, яка їх відокремлює.

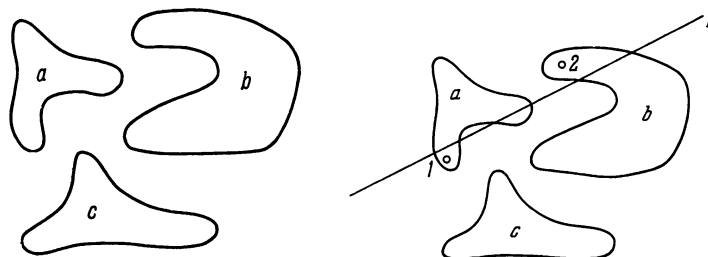


Рис. 7.

Далі береться третій об'єкт і перевіряється правильність класифікації відносно проведеної прямої.

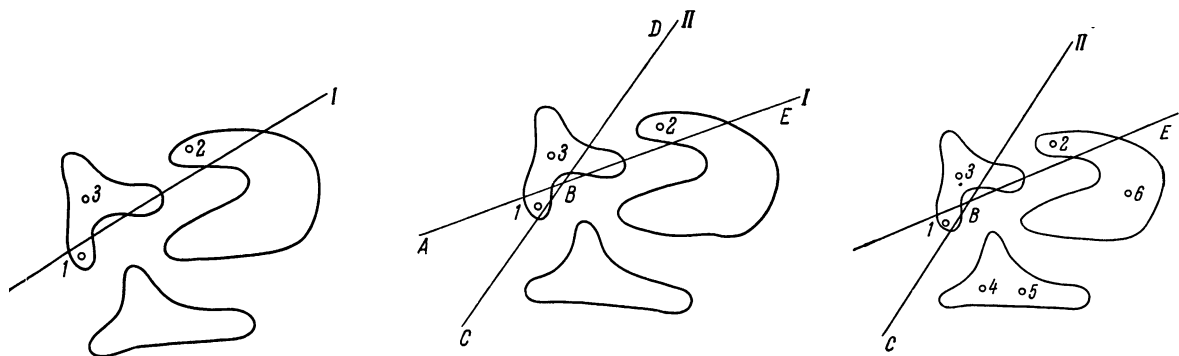


Рис. 8.

Оскільки третя точка належить тій самій півплощині, що й друга, то приходимо до висновку, що поточний класифікатор працює невірно. Для відокремлення точок 2 та 3 проводиться пряма II. Після цього площина розбивається на 4 частини. Ті частини, у якій лежать об'єкти 1 та 3 (області ABC та ABD) відносимо до класу a , область DBE — до класу b .

При перевірці нової точки можливі три випадки:

- 1) виникає протиріччя (як це було для точки 3);
- 2) протиріччя не виникає, точка потрапляє у свою частину простору;
- 3) протиріччя не виникає, точка потрапляє у вільну (непозначену) частину простору (точки 4 і 5);

У третьому випадку (точки 4 та 5 на рис. 8) машина відносить вільну частину площини до відповідного образу.

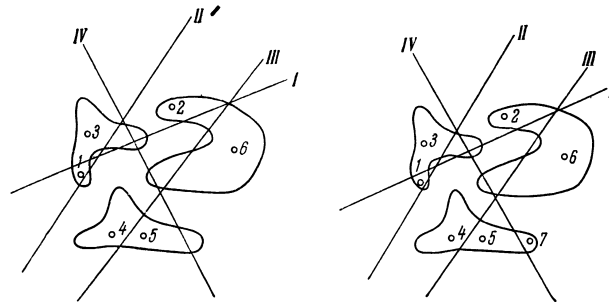


Рис. 9.

Для відокремлення точки 6 (образ *b*) від точок 4 та 5 (образ *c*) потрібно провести дві нові прямі (прямі III та IV на рис. 9).

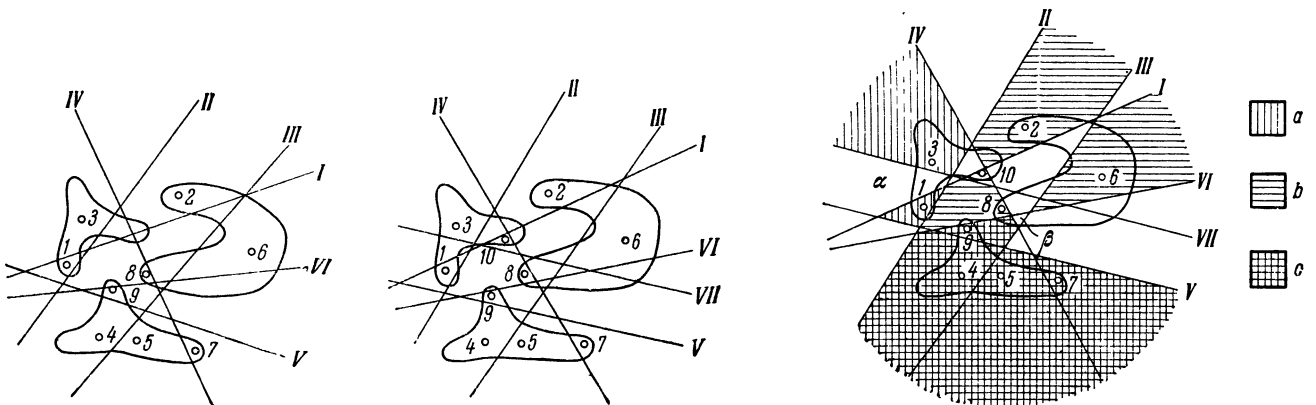
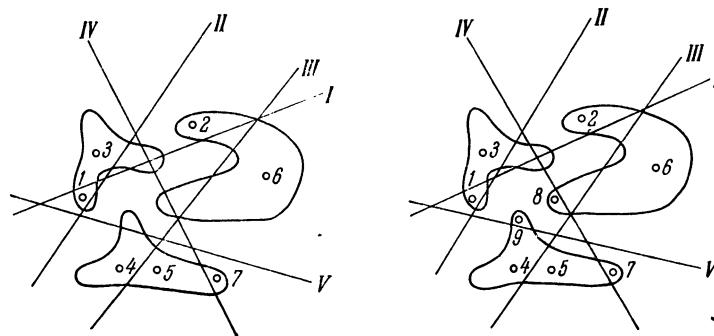


Рис. 10.

Подальші кроки вищенаведеного процесу навчання наведені на рис. 10., де вказано проміжкові та кінцевий результат стадії а). Незафарбовані ділянки не відносяться розпізнавальною системою до жодного із трьох наведених образів

Видалення зайвих площин. З рис. 10 видно, що гіперплощини (прямі) I, III та V можуть бути вилученні цілком, оскільки вони не містять частин, відкидання яких призводить до появи протиріч. Наприклад, пряма I використовується лише для відокремлення об'єктів 1 та 2 — представників різних образів *a* та *b*. Але ці об'єкти відокремлює також пряма IV (або пряма VI). Розбиття після відкидання зайвих площин наведено на рис. 11. Слід зазначити, що у результаті частка вільних частин площини зменшилися.

Проте не можна вважати, що процес навчання завершений, оскільки ще залишилися "порожні" частини площини. Для їх вилучення використовується гіпотеза компактності. Природнім є віднести "порожні" області до того образу, що і яка-небудь суміжна зафарбована частина. Цей процес називають процесом вилучення зайвих частин. Для деяких вільних частин площини цей процес однозначний (область α на рис. 10), для інших — ні (усі області на рис. 11). Рекомендується [7] проводити це вилучення поступово проглядаючи усі гіперплощини (прямі на рис. 11) та "порожні шматки", які до них прилягають.

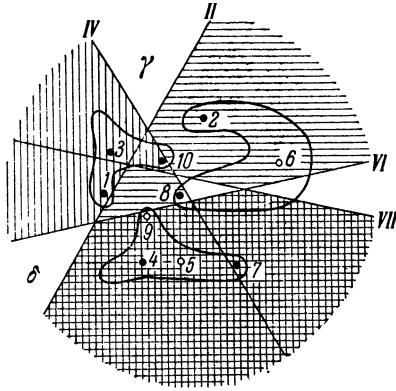


Рис. 11.

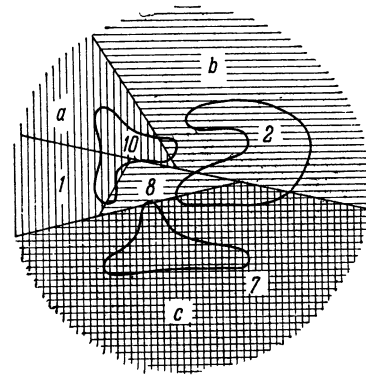


Рис. 12.

Тому спочатку виконуємо перевірку усіх частин площини II, потім усіх частин площини IV і т.д. Кінцевий результат наведений на рис 12.

Опис алгоритму

Випадковим чином вибираємо n чисел $\lambda_1, \dots, \lambda_n$. Далі для координат відповідних точок обчислюємо значення сум

$$\sigma^{(1)} = \sum \lambda_i x_i^{(1)}$$

$$\sigma^{(2)} = \sum \lambda_i x_i^{(2)},$$

Після цього вибирається число $\lambda_{n+1} = (\sigma^{(1)} + \sigma^{(2)})/2$ (вважаємо, що значення сум різні). Тоді рівняння

$$\lambda_1 x_1 + \dots + \lambda_n x_n - \lambda_{n+1} = 0.$$

Будемо позначати положення точки відносно площини знаком відхилення $\lambda(\mathbf{x}) = \lambda_1 x_1 + \dots + \lambda_n x_n - \lambda_{n+1}$. Далі процес описується за допомогою таблиць знаків:

Таблиця 1.

Таблиця знаків			
Номер точки	Образ	Номер площини	
		I	II
		Знак точки	
1	a	0	
2	b	1	

Таблиця 2.

Таблиця знаків			
Номер точки	Образ	Номер площини	
		I	II
		Знак точки	
1	a	0	
2	b	1	
3	a	1	

Таблиця 3.

Таблиця знаків				
Номер точки	Образ	Номер площини		
		I	II	III
		Знак точки		
1	a	0	1	
2	b	1	0	
3	a	1	1	

Таблиця 4.

Таблиця знаків			
Номер точки	Образ	Номер площини	
		I	II
		Знак точки	
1	a	0	1
2	b	1	0
3	a	1	1
4	c	0	0
5	c	0	0
6	b	0	0

Таблиця 5.

Таблиця знаків					
Номер точки	Образ	Номер площини			
		I	II	III	IV
		Знак точки			
1	a	0	1	1	1
2	b	1	0	1	0
3	a	1	1	1	1
4	c	0	0	1	1
5	c	0	0	0	1
6	b	0	0	0	0

Кінцевий результат кроку а. наведений у таблиці 6.

Таблиця 6.

Таблиця знаків								
Номер точки	Образ	Номер площини						
		I	II	III	IV	V	VI	VII
		Знак точки						
1	a	0	1	1	1	0	1	1
2	b	1	0	1	0	0	1	0
3	a	1	1	1	1	0	1	0
4	c	0	0	1	1	1	0	1
5	c	0	0	0	1	1	0	1
6	b	0	0	0	0	0	1	0
7	c	0	0	0	0	1	0	1
8	b	0	0	1	1	0	1	1
9	c	0	0	1	1	0	0	1
10	a	0	0	1	1	0	1	0

Таблиця 7.

Таблиця знаків					
Номер точки	Образ	Номер площини			
		II	IV	VI	VII
		Знак точки			
1	a	1	1	1	1
2	b	0	0	1	0
3	a	1	1	1	0
4	c	0	1	0	1
5	c	0	1	0	1
6	b	0	0	1	0
7	c	0	0	0	1
8	b	0	1	1	1
9	c	0	1	0	1
10	a	0	1	1	0

Далі відбувається перевірка можливості видалення стовпців. Виконується перевірка того, чи залишаються унікальними рядки після відкидання стовпчика. Якщо так, то відповідну площину можна відкинути. Результат зображений у таблиці 7. У цій таблиці є однакові рядки знаків. Зайві точки також можна виключити. Результат наведений у останній таблиці.

Розпізнавання нових об'єктів. Обчислюються знаки об'єкту відносно усі гіперплощин і отриманий код порівнюється з усіма рядками таблиці знаків. При співпаданні рядків об'єкт відноситься до відповідного образу.

Таблиця 8

Таблиця знаків					
Номер точки	Образ	Номер площини			
		II	IV	VI	VII
		Знак точки			
1	a	1	1	1	1
2	b	0	0	1	0
3	a	1	1	1	0
4	c	0	1	0	1
7	c	0	0	0	1
8	b	0	1	1	1
10	a	0	1	1	0

Метод потенціалів для розпізнавання образів

Точковий електричний заряд у однорідному середовищі створює електричне поле, зображене на рис. 13. Радіальні лінії — це силові лінії поля, концентричні кола — лінії однакового потенціалу.

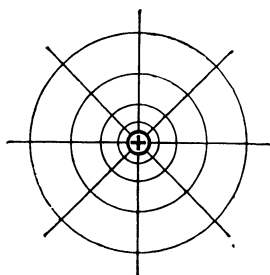


Рис. 13.

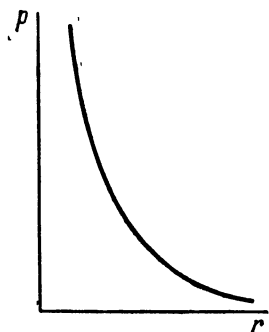


Рис. 14.

Потенціал p у кожній точці простору визначається співвідношенням

$$p = a \frac{q}{r^2},$$

де a — деяка стала, q — величина заряду, r — відстань від заданої точки до заряду.

Крива зміни потенціалу як функції відстані наведена на рис. 14. Потенціал зменшується по мірі віддалення від його джерела. Якщо поле утворене кількома зарядами, то потенціал в кожній точці рівний сумі потенціалів, які створюються кожним із зарядів.

Припустимо, що у просторі розташовані дві компактні групи зарядів. У одній групі — позитивні, у другій — негативні. На рис. 15 показаний розподіл потенціалів у околі цих зарядів, на рис. 16 ці потенціали алгебраїчно просумовані.

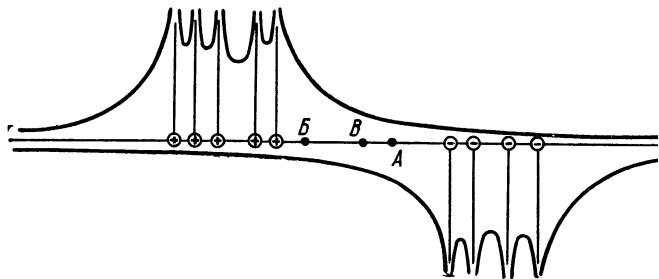


Рис. 15.

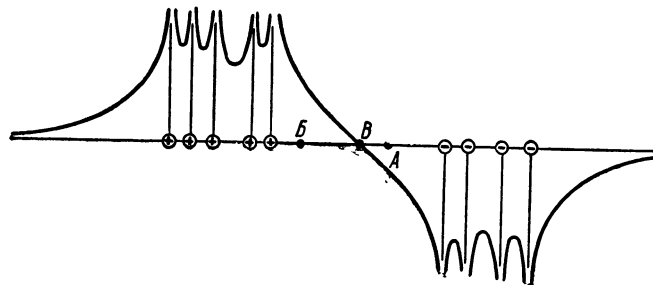


Рис. 16.

Точку можна віднести до тієї чи іншої множини точок у залежності від того, який знак має сумарний потенціал поля у цій точці.

Вищенаведені міркування по аналогії можна перенести на точки простору рецепторів. Кожній точці, яка з'являється у процесі навчання, поставимо у відповідність деяку функцію, аналогічну по формі до електричного потенціалу. Такою функцією може бути, наприклад, функція

$$\varphi(D) = \frac{1}{1 + \alpha D^2},$$

де α — деякий коефіцієнт, D — відстань (у деякій метриці) між точкою-джерелом потенціалу та точкою, у якій обчислюється потенціал. Наприклад, у якості D можна використовувати евклідову відстань, віддаль Хеммінга, манхеттенську метрику, тощо.

Нехай джерелами потенціалів є група точок, які відповідають деякому образу a . Тоді можна вважати, що середній потенціал, які створюють у деякій точці простору джерела цього образу, характеризує віддаль від цієї точки до усього образу у цілому.

Спираючись на гіпотезу компактності, можна запропонувати наступне правило розпізнавання: точку відносимо до того образу, середній потенціал якого у цій точці є максимальним.

Алгоритм розпізнавання на основі методу потенціалів

1. У процесі навчання запам'ятовуються координати усіх точок та належність їх до відповідних образів a_1, \dots, a_m .
2. Розпізнавання.
 - а. Для точки \mathbf{x} , яка підлягає розпізнаванню, обчислюється потенціали кожного образу, тобто суми

$$\Phi_i(\mathbf{x}) = \frac{1}{n_i} \sum_{\mathbf{a} \in a_i} \varphi_{\mathbf{a}}(\mathbf{x}), \quad i = 1, 2, \dots, m,$$

де m — кількість різних образів, n_i — кількість точок відповідного образу, використаних у процесі навчання, $\varphi_{\mathbf{a}}(\mathbf{x}) = \frac{1}{1 + \alpha D^2(\mathbf{a}, \mathbf{x})}$ — потенціал, який створює точка \mathbf{a} у точці \mathbf{x} .

- б. Порівнюються величини $\Phi_1(\mathbf{x}), \Phi_2(\mathbf{x}), \dots, \Phi_m(\mathbf{x})$ і точка \mathbf{x} відноситься до того образу, потенціал якого є найбільшим.

У випадку двох образів a_1 та a_2 рішення можна приймати на основі значення знаку функції $\Delta\Phi(\mathbf{x}) = \Phi_{a_1}(\mathbf{x}) - \Phi_{a_2}(\mathbf{x})$.

Застосування вищевказаного алгоритму для розпізнавання цифр, зображених за допомогою рецепторного поля 6×10 у середньому дало змогу досягнути правильного розпізнавання у 85% випадків за умови, що $n_1 = n_2 = \dots = n_{10} = 12$, причому подальше збільшення обсягу навчальної вибірки практично не впливає на якість розпізнавання.

Недоліки найпростішого алгоритму виявляються у тому випадку, коли точки навчальної вибірки розподілені нерівномірно. Відповідний приклад наведений на рис. 17-19.

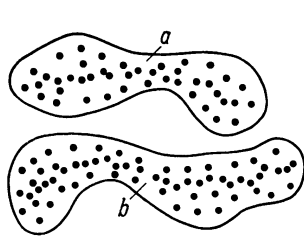


Рис. 17.

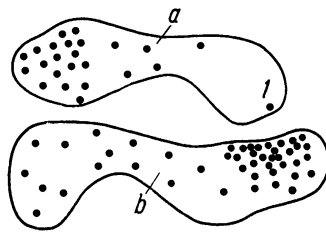


Рис. 18.

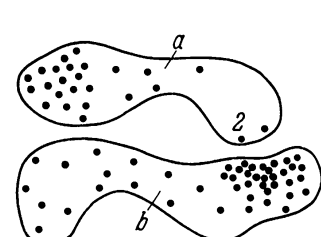


Рис. 19.

Навчальна вибірка для образів a та b , зображених на рис. 17, наведена на рис. 18. У цьому випадку у зв'язку із «невдалою» навчальною вибіркою, нова точка 2 (рис. 19) буде віднесена до образу b . Більш того, машина може не впізнати навіть відомий їй об'єкт, який зустрічався у процесі навчання, наприклад, точку 1, яка належить до образу a .

Покращена модифікація алгоритму навчання. Алгоритм розпізнавання може бути покращений, шляхом введення поняття «ваги точки» та доповнення кроку навчання наступною операцією. Після запам'ятовування усіх об'єктів їм присвоюють початкову вагу 1. Далі до елементів навчальної вибірки застосовують крок 2 алгоритму навчання, причому потенціали середні потенціали обчислюють за формулою

$$\Phi_i(\mathbf{x}) = \frac{1}{n_i} \sum_{\mathbf{a} \in a_i} \varphi_{\mathbf{a}}(\mathbf{x}) \cdot w(\mathbf{a}), \quad i = 1, 2, \dots, m,$$

де $w(\mathbf{a})$ — вага точки \mathbf{a} .

Якщо при розпізнаванні об'єкту \mathbf{a} відбувається помилка, то вагу $w(\mathbf{a})$ відповідного об'єкту збільшують на деяку величину, наприклад на одиницю. Потім застосовується ще один такий самий цикл перевірки та корекції ваг і т.д. Цикли повторюються до тих пір, поки усі відомі фігури не будуть розпізнані правильно.

Застосування покращеного алгоритму до розпізнавання дало змогу підвищити відсоток правильного розпізнавання нових зображень цифр до 89%.

Удосконалення кодування графічних об'єктів. У випадку бінарних зображень традиційні функції відстані Хеммінга або Евкліда не завжди точно відображають відмінність між зображеннями. Відповідний приклад наведений на рис. 20. Шляхом зсуву вертикальної лінії із п'ятірки (рис. 20, а), можна отримати «трохи гіршу» п'ятірку (рис. 20, б) та трійку (рис. 20, в), кожна з яких відрізняється від початкової фігури у шести розрядах. Тобто, елемент «свого» образу та елемент «чужого» образу знаходяться на однаковій відстані від початкової фігури. Схожий приклад наведений на рис. 21 для випадку зображень одиниць та двійок, які відрізняються у десяти розрядах.

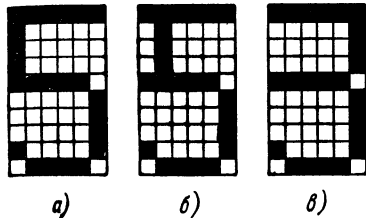


Рис. 20.

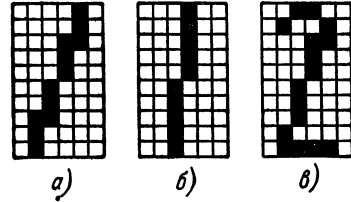


Рис. 21.

Ці приклади показують, що значні зміни елементів зображень, які значно спотворюють зображення, приводять тих самих змін у відстані, що і незначні. Тому застосований метод кодування недостатньо добре передає відмінність фігур.

Надійність розпізнавання може бути підвищена шляхом удосконалення методу кодування. Поставимо у відповідність кожному рецептору функцію, яка рівна одиниці на цьому рецепторі, спадаючу по усім напрямкам від нього, тобто функцію, аналогічну до потенціалу (рис. 22, а). Ця функція може бути апроксимована ступінчастими функціями (рис. 22, б-в).

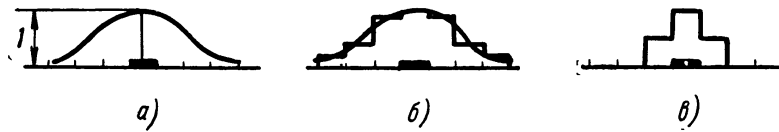
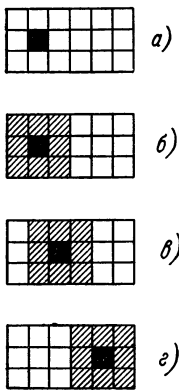


Рис. 22.



У випадку «грубої апроксимації» функція рівна одиниці на відповідному елементів рецепторного поля, $\frac{1}{2}$ — на сусідніх елементах, нулю — на всіх інших. Для прикладу розглянемо одну кліткову фігуру, зображену на рис. 23, а. На рис. 23, б зображено значення ступінчастої функції відстані. Код такої фігури має вигляд

$$\frac{1}{2} \frac{1}{2} \frac{1}{2} 000 \frac{1}{2} 1 \frac{1}{2} 000 \frac{1}{2} \frac{1}{2} \frac{1}{2} 000.$$

Посунемо фігуру на один елемент управо (рис. 23, в). Отримаємо фігуру із кодом

$$0 \frac{1}{2} \frac{1}{2} \frac{1}{2} 000 \frac{1}{2} 1 \frac{1}{2} 000 \frac{1}{2} \frac{1}{2} \frac{1}{2} 00.$$

Рис. 23. Евклідова відстань між цими фігурами рівна $\sqrt{2}$. Тепер посунемо фігуру іще на дві клітинки вправо. Отримаємо фігуру (рис. 23, в) із кодом

$$000 \frac{1}{2} \frac{1}{2} \frac{1}{2} 000 \frac{1}{2} 1 \frac{1}{2} 000 \frac{1}{2} \frac{1}{2} \frac{1}{2} \cdot$$

Відстань від цієї фігури до початкової рівна $\sqrt{6}$. Як видно із прикладів, при такому методі кодування «більш віддаленим» фігурам відповідає більша відстань у просторі рецепторів.

Описаний вище прийом можна застосувати для більш складних фігур з використанням сумарних потенціалів рецепторів. На рис. 24 зображено кодування цифр із прикладу, наведеного на рис. 20.

2	2,5	2	2	2	1,5
2,5	2,5	1,5	1,5	1,5	1
2	1,5	0	0	0	0
2,5	2,5	1,5	1,5	1	0,5
2	2,5	2	2	2	1
1	1,5	1,5	1,5	2	2
0	0	0	0	1,5	2
0,5	0,5	0	0	1,5	2
1,5	1,5	1,5	1,5	2	2
1	2	2	2	2	1

2	2,5	2,5	2	2	1,5
2	3	2,5	1,5	1,5	1
1,5	2	1,5	0	0	0
2	3	2,5	1,5	1	0,5
2	2,5	2,5	2	2	1
1	1,5	1,5	1,5	2	2
0	0	0	0	1,5	2
0,5	0,5	0	0	1,5	2
1,5	1,5	1,5	1,5	2	2
1	2	2	2	2	1

1,5	2	2	2	2,5	2
1	1,5	1,5	1,5	2,5	2,5
0	0	0	0	1,5	2
1	1,5	1,5	1,5	2	2
1,5	2	2	2	2	1,5
1	1,5	1,5	1,5	2	2
0	0	0	0	1,5	2
0,5	0,5	0	0	1,5	2
1,5	1,5	1,5	1,5	2	2
1	2	2	2	2	1

Рис. 24.

Тоді відстань між «п'ятірками» рівна 2,25, а відстань між першою п'ятіркою та трійкою — 5,22.

Застосування удосконаленого методу кодування дозволило збільшити середню надійність розпізнавання букв до 94%.

Метод нормалізації плоских зображень

Зображення об'єктів часто отримуються у вигляді проєкцій. Із множини усіх проєкцій можна виділити підмножини R_1, R_2, \dots , для яких виконується наступна закономірність: усі елементи цих множин можна отримати у результаті деяких квазіізоморфних перетворень однієї проєкції. Ця проєкція називається *стандартною*, а загадані вище перетворення — допустимими.

Одним із підходів, який дозволяє ототожнювати проєкції підмножин R_1, R_2, \dots полягає у *нормалізації* — зведені довільної проєкції $Z \in R_i$ до стандартної проєкції $Z^0 \in R_i$, $i = 1, 2, \dots$, тобто

$$H_\Gamma(X) = Z^0,$$

де H — оператор нормалізації, Γ — множина груп допустимих перетворень. Останні можуть бути поділені на групи Γ_1 — трансляцій, Γ_2 — обертань, Γ_3 — гомотетій.

Довільне бінарне зображення на площині однозначно описується функцією яскравості $B(\mathbf{z})$, де $\mathbf{z} = (x, y)^T$. Тоді нормалізація, яка проводиться по групам перетворень Γ_1 - Γ_3 , описується наступним чином:

$$\mathbf{z}_0 = M_k M_\psi (\mathbf{z} - \mathbf{z}_c),$$

де

$$\mathbf{z}_c = (x_c, y_c), \quad M_\psi = \begin{pmatrix} \cos \psi & -\sin \psi \\ \sin \psi & \cos \psi \end{pmatrix}, \quad M_k = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix},$$

де \mathbf{z}_c — вектор перетворень на етапі 1, M_ψ , M_k — матриці перетворень на етапах 2-3 нормалізації.

Вектор \mathbf{z}_c характеризує паралельний зсув проєкції відносно стандартної проєкції (центрування). Вважають, що цент мас стандартної системи співпадає із початком координат. Тоді

$$x_c = \frac{m_{10}}{m_{00}}, \quad y_c = \frac{m_{01}}{m_{00}}$$

де m_{01} , m_{10} — моменти зображення першого порядку.

Моменти порядку $\alpha + \beta$ обчислюються за формулою

$$m_{\alpha\beta} = \sum_i \sum_j B(x_i, y_j) x_i^\alpha y_j^\beta.$$

З останньої формули випливає, що m_{00} — кількість забарвлених точок зображення.

У результаті центрування проєкція буде описуватися зміщеною функцією яскравості $B_1(x, y) = B(x - x_c, y - y_c)$. Довільний центральний момент порядку $\alpha + \beta$ позначимо $\mu_{\alpha\beta}$.

Для однозначності орієнтування доцільно фіксувати на центрованому зображенні деякий вектор ψ з початком у точці O . Будемо вимагати, щоб для стандартного зображення цей вектор співпадав із віссю OX . Алгоритм пошуку вектора ψ має бути інваріантним відносно елементів груп перетворень Γ_2 та Γ_3 . Для орієнтації використовуються центральні моменти другого порядку. При повороті центрованого зображення навколо початку координат на кут ψ проти годинникової стрілки центральні моменти другого порядку наступним чином виражаються через початкові моменти $\mu'_{20}, \mu'_{02}, \mu'_{11}$:

$$\begin{aligned}\mu_{20}(\psi) &= \mu'_{20} \cos^2 \psi + \mu'_{02} \sin^2 \psi - \mu'_{11} \sin 2\psi \\ \mu_{02}(\psi) &= \mu'_{20} \sin^2 \psi + \mu'_{02} \cos^2 \psi + \mu'_{11} \sin 2\psi \\ \mu_{11}(\psi) &= 0,5 (\mu'_{20} - \mu'_{02}) \sin 2\psi + \mu'_{11} \cos 2\psi\end{aligned}$$

Для останньої залежності можна вказати такий кут ψ_0 , для якого $\mu_{11} = 0$. Очевидно, що

$$\psi_0 = 0,5 \operatorname{arctg} \frac{2\mu'_{11}}{\mu'_{02} - \mu'_{20}} + \nu \frac{\pi}{2} \quad (\nu = 0, 1, 2, \dots).$$

Але виконання умови $\mu_{11} = 0$ ще не забезпечує однозначності орієнтування, оскільки при повороті зображення на кут 2π є чотири положення, для яких $\mu_{11} = 0$. Можна показати, що якщо $\mu_{11} = 0$, то один моментів $\mu_{20}(\psi)$, $\mu_{02}(\psi)$ набуває своє найбільше значення, а інший — найменше. Тому, якщо доповнити умову $\mu_{11} = 0$ умовою $\mu_{20}(\psi) > \mu_{02}(\psi)$, то множина усіх можливих орієнтацій може бути зведена до двох, які відрізняються на кут π .

Для однозначного орієнтування необхідна така характеристика, яка різко змінюється при повороті на кут π . Таким є момент $\mu_{30}(\psi)$, який обчислюється наступним чином:

$$\mu_{30}(\psi) = \mu'_{30} \cos^3 \psi - 2\mu'_{21} \cos^2 \psi \sin \psi + 2\mu'_{12} \cos \psi \sin^2 \psi - \mu'_{03} \sin^3 \psi,$$

де початкові центральні моменти позначені штрихом. Із останнього співвідношення випливає, що $\mu_{30}(\pi + \psi) = -\mu_{30}(\psi)$. Тому, вважаючи, що стандартному зображенню відповідає значення $\mu_{30} > 0$, отримаємо наступні умови для однозначної орієнтації проєкції:

$$\mu_{11} = 0, \quad \mu_{20}(\psi) > \mu_{02}(\psi), \quad \mu_{30} > 0. \quad (1)$$

Після знаходження координат x_c та y_c для обчислення моментів можна скористатися формулою:

$$\mu_{\alpha\beta} = \sum_{i=1}^{M_c} \left\{ \left[\sum_{\nu=0}^{\alpha} (-1)^\nu C_{\alpha}^{\nu} x_i^{\alpha-\nu} x_c^{\nu} \right] \left[\sum_{\xi=0}^{\beta} (-1)^\xi C_{\beta}^{\xi} y_i^{\beta-\xi} y_c^{\xi} \right] \right\},$$

де x_i та y_i — координати i -го елемента зображення до центрування.

Для зображень, які мають центр симетрії вище, ніж другого порядку, умова (1) не є застосовною, оскільки $\mu_{11}(\psi) \equiv 0$, $\mu_{20}(\psi) = \mu_{02}(\psi)$. Для таких зображень застосовують умову

$$\mu_{30} + \mu_{12} = 0.$$

Для виконання цієї умови виконують поворот на кут

$$\psi_1 = \operatorname{arctg} \frac{\mu'_{30} + \mu'_{12}}{\mu'_{21} + \mu'_{03}} + \nu \frac{\pi}{2} \quad (\nu = 0, 1, 2, \dots).$$

Для однозначного орієнтування тоді використовується наприклад умова $\mu_{21}(\psi) + \mu_{03}(\psi) > 0$, яку можна досягти додатковим обертом на кут π .

Якщо $\mu'_{21} + \mu'_{03} = \mu'_{12} + \mu'_{30} = 0$, то кут ψ_1 не можна обчислити. У цьому випадку накладають умову $\mu_{40} = \mu_{04}$, яка виконується за умови повороту на кут

$$\psi_2 = \frac{1}{2} \operatorname{arctg} \frac{\mu'_{40} - \mu'_{04}}{2(\mu'_{31} + \mu'_{13})} + \nu \frac{\pi}{2} \quad (\nu = 0, 1, \dots),$$

Для однозначної орієнтації тоді накладають умови

$$\mu_{31} + \mu_{13} > 0, \quad \mu_{50} > 0.$$

Для визначення коефіцієнту k використовують масштабування по площі. Тоді якщо S_c^* та S_c — площі відповідно стандартного зображення та зображення, яке потребує нормалізацію, то

$$k = \sqrt{\frac{S_c^*}{S_c}}.$$

Класифікація образів із використанням функції відстані

У якості критерію класифікації використовується підхід заснований на класифікації образів за критерієм мінімуму відстані.

Класифікації образів за критерієм мінімуму відстані

1. Випадок одного еталону. Метод найближчого сусіда.

В деяких задачах об'єкти деяких класів (образів) мають тенденцію до групування навколо деякого об'єкту, який є типовим або репрезентативним для відповідного образу. Типовим прикладом є задача зчитування банківських чеків, для якої вектори, які відповідають образам кожного класу, будуть майже ідентичними.

Розглянемо M класів, які допускають зображення за допомогою еталонних представників $\mathbf{z}_1, \dots, \mathbf{z}_M$. Евклідова відстань між довільним вектором \mathbf{x} та цими образами обчислюється за формулою

$$D_i = \sqrt{\sum_{j=1}^n (x_j - z_{ij})^2},$$

або

$$D_i = \|\mathbf{x} - \mathbf{z}_i\| = \sqrt{(\mathbf{x} - \mathbf{z}_i)'(\mathbf{x} - \mathbf{z}_i)}.$$

Вектор \mathbf{x} відноситься до класу ω_i , якщо умова $D_i < D_j$ виконується для усіх $j \neq i$. Перетворимо формулу обчислення відстані

$$\begin{aligned} D_i^2 &= \|\mathbf{x} - \mathbf{z}_i\|^2 = (\mathbf{x} - \mathbf{z}_i)'(\mathbf{x} - \mathbf{z}_i) = \\ &= \mathbf{x}'\mathbf{x} - 2\mathbf{x}'\mathbf{z}_i + \mathbf{z}_i'\mathbf{z}_i = \mathbf{x}'\mathbf{x} - 2\left(\mathbf{x}'\mathbf{z}_i - \frac{1}{2}\mathbf{z}_i'\mathbf{z}_i\right). \end{aligned}$$

Остання формула показує, що вибір мінімальної відстані до класу еквівалентний максимізації величини $\mathbf{x}'\mathbf{x} - 2\left(\mathbf{x}'\mathbf{z}_i - \frac{1}{2}\mathbf{z}_i'\mathbf{z}_i\right)$. Тому функції рішень можна визначити як

$$d_i(\mathbf{x}) = \mathbf{x}'\mathbf{z}_i - (1/2)\mathbf{z}_i'\mathbf{z}_i, \quad i = 1, 2, \dots, M,$$

де $d_i(\mathbf{x})$ — лінійні функції рішень. Якщо покласти

$$\begin{aligned} w_{ij} &= z_{ij}, \quad j = 1, 2, \dots, n, \\ w_{i, n+1} &= - (1/2)\mathbf{z}_i'\mathbf{z}_i \end{aligned}$$

і $\mathbf{x} = (x_1, \dots, x_n, 1)$, то функції рішень можна записати у вигляді

$$d_i(\mathbf{x}) = \mathbf{w}'_i \mathbf{x}, \quad i = 1, 2, \dots, M,$$

де $\mathbf{w}_i = (w_{i1}, \dots, w_{in}, w_{i,n+1})$. На рис. 25 зображена відокремлююча межа для випадку двох класів.

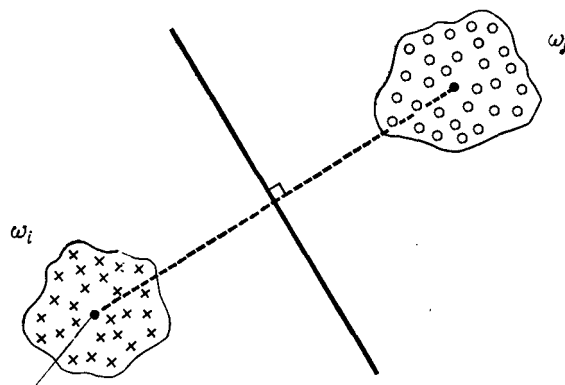


Рис. 25.

Границя між класами є гіперплощиною, яка рівновіддалена від еталонних представників класів. Описаний підхід ще називають кореляцією або співставленням із кластером.

2. Випадок кількох еталонів.

У цьому випадку кожний клас ω_i можна охарактеризувати еталонами $\mathbf{z}_i^1, \dots, \mathbf{z}_i^{N_i}$, N_i — кількість еталонів i -го класу. Одним із найпростіших підходів є метод "найближчого сусіда" (НС-правило). Відстань між i -им класом та об'єктом \mathbf{x} можна обчислити за формулою

$$D_i = \min_l \|\mathbf{x} - \mathbf{z}_i^l\|, \quad l = 1, 2, \dots, N_i;$$

тобто з використанням відстані до найближчого до \mathbf{x} еталона класу. Функції рішень можна записати у вигляді

$$d_i(\mathbf{x}) = \max_l \{(\mathbf{x}' \mathbf{z}_i^l) - (1/2)(\mathbf{z}_i^l)' \mathbf{z}_i^l\}, \quad l = 1, 2, \dots, N_i,$$

і процедура віднесення нового об'єкта до класу така сама, як і раніше. На рис. 26 показані границі класів у випадку двох класів, які містять по два еталони кожний.

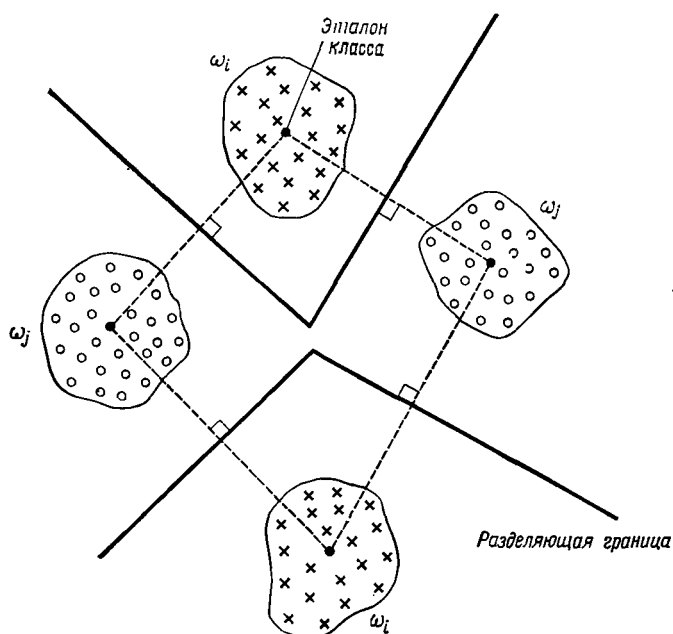


Рис. 26.

Отриманий класифікатор є частинним випадком *кусково-лінійного класифікатора*.

3. Узагальнення принципів класифікації по мінімуму відстані. Метод k -найближчих сусідів.

НС-правило може бути узагальнено на k -НС правило, згідно якого обчислюються k -найближчих сусідів і об'єкт відноситься до того класу, до якого відноситься найбільша кількість з k -найближчих еталонів.

Кластеризація

Кластерний аналіз полягає у поділі заданого набору об'єктів на кластери — групи об'єктів, схожі між собою за тим чи іншим критерієм подібності. Близькими по змісту до терміну "кластер" є поняття "клас" та "таксон". Тому кластеризацію часто називають таксономією або класифікацією без учителя, оскільки, на відміну від задачі "звичайної" класифікації, класи апріорно не задані.

Слід зазначити, що у кластеризації присутні як об'єктивні, так і суб'єктивні моменти. Це добре ілюструється прикладом Богнарда. Фігури, зображені на рис. 27 можна розбити на різні класи у залежності від критерію кластеризації: а) по кольору — на два кластери, по кількості кутів — на три, по розміру — на два або три. Тому результат кластеризації визначається багато у чому завданнями та засобами, які використовують у її процесі

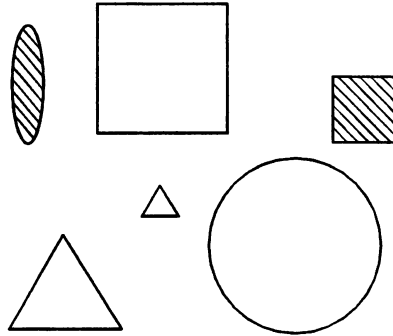


Рис. 27.

Міри близькості. Для виділення кластерів потрібно у першу чергу ввести функцію подібності об'єктів. Розглянемо міри близькості у випадку кластеризації об'єктів, заданих у векторній формі. Найбільш часто використовується евклідова відстань. Для булевих векторів використовують відстань Хеммінга. Також часто використовують відстань Махаланобіса

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(\mathbf{x} - \mathbf{y})^T S^{-1} (\mathbf{x} - \mathbf{y})}$$

де S — коваріаційна матриця, \mathbf{x}, \mathbf{y} — об'єкти-змінні.

Можна також використовувати неметричні функції близькості

$$s(\mathbf{x}, \mathbf{z}) = \frac{\mathbf{x}'\mathbf{z}}{\|\mathbf{x}\| \|\mathbf{z}\|}, \quad (2)$$

яка рівна косинусу кута між векторами \mathbf{x} та \mathbf{z} (у останній формулі через \mathbf{x}' позначається вектор-рядок транспонований, до вектора-стовпчика \mathbf{x}).

У випадку двійкових векторів величина $\mathbf{x}'\mathbf{z}$ рівна кількості співпадаючих ознак для об'єктів \mathbf{x} та \mathbf{z} . Іншим варіантом формули (2) є міра Танімото

$$s(\mathbf{x}, \mathbf{z}) = \frac{\mathbf{x}'\mathbf{z}}{\|\mathbf{x}\|^2 + \|\mathbf{z}\|^2 - \mathbf{x}'\mathbf{z}}.$$

Слід зазначити, що у зв'язку із неоднорідністю шкал, у яких вимірюються ознаки об'єктів, при розрахунках відстаней може виявитися, що деякі ознаки набагато сильніше впливають на близькість об'єктів, ніж інші. Тому бажано попередньо провести нормалізацію змінних.

Основні підходи до побудови методів кластеризації. Методи кластеризації поділяють на ієрархічні та неієрархічні. У ієрархічних методах будується дерево вкладених кластерів (дендограма). Приклад зображення дендограм наведено на рис. 28.

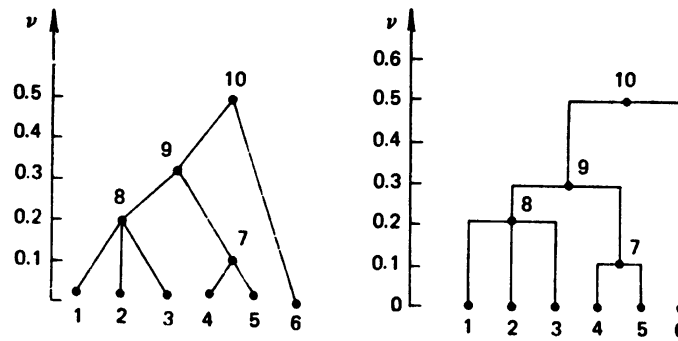


Рис. 28. Різні способи зображення дендограм

Ієрархічні методи поділяються на агломеративні методи та методи поділу.

Агломеративні методи полягають у послідовному об'єднанні початкових кластерів (побудова кластерів знизу вверху). Методи поділу — у розщепленні існуючих кластерів (побудова кластерів зверху вниз).

Агломеративні алгоритми кластеризації

Нехай X — скінченна множина, яку потрібно кластеризувати. На першому кроці кожному елементу множини відповідає одноелементний кластер. На кожному наступному кроці вибираються два найближчі між собою кластери C_i та C_j , які об'єднуються у новий кластер $C_i \cup C_j$. Після кожного такого кроку кількість кластерів зменшується. У кінці процедури отримується кластер, який складається із одного об'єкта.

Для оцінки відстані $d(C_i, C_j)$ між двома множинами використовують наступні функції:

а) Відстань між найближчими об'єктами множин

$$d(C_i, C_j) = \min \{d(\mathbf{x}, \mathbf{y}) \mid \mathbf{x} \in C_i, \mathbf{y} \in C_j\}.$$

б) Відстань між найбільш віддаленими об'єктами.

$$d(C_i, C_j) = \max \{d(\mathbf{x}, \mathbf{y}) \mid \mathbf{x} \in C_i, \mathbf{y} \in C_j\}.$$

в) Середня відстань між множинами

$$d(C_i, C_j) = \frac{1}{|C_i| |C_j|} \sum_{\mathbf{x} \in C_i} \sum_{\mathbf{y} \in C_j} d(\mathbf{x}, \mathbf{y}).$$

г) Відстань між центрами мас множин

$$d(C_i, C_j) = d(\mathbf{z}_i, \mathbf{z}_j).$$

д) Статистична відстань

$$d(C_i, C_j) = \sqrt{\frac{|C_i| \cdot |C_j|}{|C_i| + |C_j|}} \|\mathbf{z}_i - \mathbf{z}_j\|,$$

де $\mathbf{z}_i, \mathbf{z}_j$ — центри мас відповідних множин, які обчислюються за формулою $\mathbf{z}(C) = \frac{1}{|C|} \sum_{\mathbf{x} \in C} \mathbf{x}$.

Відстань $d(\mathbf{x}, C)$ між та об'єктом та множиною покладають рівною $d(\{\mathbf{x}\}, C)$.

Агломеративні алгоритми кластеризації із функцією г) запропонував **Сокал**, алгоритм із функцією д) — **Уорд**.

Алгоритм Боннера

Випадковим чином вибирається одноелементний кластер, який об'єднується з тими одноелементними кластерами, відстань до яких не перевищує заданий поріг T . Далі випадково виби-

рається наступний одноелементний кластер і так само проходить злиття. Процес продовжується до тих пір, поки ще є одноелементні кластери.

Алгоритми поділу

Алгоритм Макнаутона-Сміта

У алгоритмі багато разів застосовується наступна процедура поділу множини C на два кластери C_1 та C_2 :

На початку усі об'єкти множини заносяться у кластер C_1 , а кластер C_2 — порожній.

На кожному кроці знаходиться об'єкт $\mathbf{x} \in C_1$, для якого величина $d_{\mathbf{x}} = d(\mathbf{x}, C_1 \setminus \{\mathbf{x}\}) - d(C_2, \mathbf{x})$ приймає найбільше значення, де $d(\mathbf{x}, C_1 \setminus \{\mathbf{x}\})$, $d(\mathbf{x}, C_2)$ — відстані між об'єктом \mathbf{x} та відповідними множинами.

Якщо $d_{\mathbf{x}} > 0$, то об'єкт \mathbf{x} виключається із кластера C_1 і заноситься у кластер C_2 і відбувається перехід на наступний крок процедури поділу, інакше робота процедури завершується.

У алгоритмі Макнаутона-Сміта спочатку усі об'єкти заносяться у один кластер C . Потім цей кластер за допомогою вищенаведеної процедури розбивається на два кластери C_1 та C_2 . Далі процес розбиття застосовується до одного з кластерів C_1 та C_2 . У результаті отримують три кластери і т. д. Рекомендується на кожному кроці вибирати для розбиття кластер найбільшого діаметру.

Неієрархічні алгоритми кластеризації

Для неієрархічних методів притаманним є використання поняття центра кластера. Об'єкти відносяться до того кластера, відстань до центра якого є найменшою.

1. Пороговий алгоритм кластеризації

Нехай $S = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_r\}$ — множина об'єктів, яку треба розбити на кластери і нехай також задано деякий поріг T і визначено вектор \mathbf{z}_1 — центр першого кластера (у найпростішій модифікації $\mathbf{z}_1 = \mathbf{x}_1$).

Припустимо, що об'єкти $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_i\}$, ($i < r$) вже віднесені до одного із m кластерів C_1, \dots, C_m із центрами у $\mathbf{z}_1, \dots, \mathbf{z}_m$, ($m \leq i$). Для вектора \mathbf{x}^{i+1} потрібно обчислити величини $d(\mathbf{x}_{i+1}, \mathbf{z}_1), \dots, d(\mathbf{x}_{i+1}, \mathbf{z}_m)$ — відстані від \mathbf{x}_{i+1} до центрів кластерів C_1, \dots, C_m . Якщо усі ці відстані більші за поріг T , то створюється новий кластер C_{m+1} із центром $\mathbf{z}_{m+1} = \mathbf{x}_{i+1}$. У протилежному випадку об'єкт \mathbf{x}_{i+1} включається у кластер із найближчим до нього центром.

Результат роботи порогового алгоритму залежить від вибору першого центру та порогу T . Ця залежність проілюстрована на рис. 29.

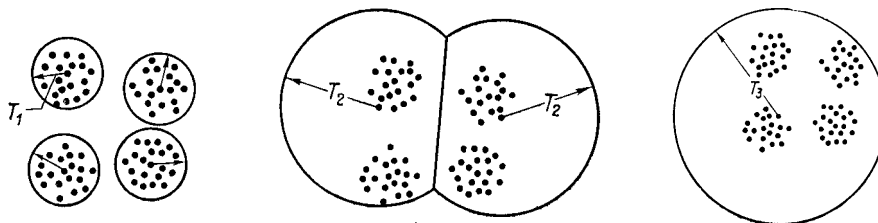


Рис. 29.

Існує велика кількість модифікацій порогового алгоритму. Наприклад, можна задавати бажану кількість кластерів та корегувати величину порогу T . Інша модифікація полягає у тому, що звичайний пороговий алгоритм застосовують r разів, змінюючи центр першого кластера (на i -ій ітерації покладають $\mathbf{z}_1 = \mathbf{x}_i$, ($i = 1, \dots, r$)). Для порівняння якості кластеризації (для фіксованого порогу T) обчислюють величини

$$F_i = \sum_{j=1}^{m_i} \sum_{\mathbf{x} \in C_j^{(i)}} d(\mathbf{x}, \mathbf{z}_j), \quad i=1, \dots, r,$$

де $C_j^{(i)}$ — j -ий кластер, m_i — кількість кластерів, які було отримано у результаті i -ої кластеризації. Очевидно, що величини F_i зменшуються із зменшенням порогу T .

Вибирають той варіант кластеризації $\{C_1^{(i)}, \dots, C_{m_i}^{(i)}\}$, для якого величина F_i є найменшою.

2. Алгоритм максимінної віддалі

Вхід. Множина об'єктів $S = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_r\}$, $\mathbf{x}_i \in \mathbb{R}^n$ ($i = \overline{1, r}$, $r > 2$).

Крок 1. $\mathbf{z}_1 \leftarrow \mathbf{x}_1$, $S' \leftarrow S \setminus \{\mathbf{x}_1\}$.

Крок 2. У множині S' знаходиться об'єкт \mathbf{x}_2 , який знаходиться на найбільшій відстані d_1 від центра першого кластера \mathbf{z}_1 .

$\mathbf{z}_2 \leftarrow \mathbf{x}_2$, $S' \leftarrow S' \setminus \{\mathbf{x}_2\}$, $m \leftarrow 2$.

Крок 3. Якщо множина S' не порожня, то

1) Для усіх елементів $\mathbf{x} \in S'$ обчислити відстані $d(\mathbf{x}, \mathbf{z}_i)$, $i = \overline{1, m}$ до центрів кластерів $\{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m\}$

2) Знайти $d(\mathbf{x}) = \min_{1 \leq i \leq m} \{d(\mathbf{x}, \mathbf{z}_i)\}$.

3) Знайти $d' = \max_{\mathbf{x}} d(\mathbf{x}) = d(\mathbf{x}')$ — максимум із знайдених мінімальних відстаней.

4) Вибрати коефіцієнт k_i ($0 < k_i < 1$). Якщо

$$d' \leq k_i \cdot (d_1 + \dots + d_{m-1}) / (m-1),$$

то перейти на крок 4.

інакше

$$d_m \leftarrow d', \quad m \leftarrow m+1, \quad \mathbf{z}_m \leftarrow \mathbf{x}', \quad S' \leftarrow S' \setminus \{\mathbf{x}'\}$$

(кількість кластерів збільшується на одиницю, відповідний об'єкт \mathbf{x}' вилучається із навчальної вибірки та стає центром нового кластера).

5) перейти на початок кроку 3.

інакше перейти на крок 4.

Крок 4. На основі правила найближчого сусіда провести розбиття навчальної вибірки S на кластери із використанням заданої множини центрів $\{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m\}$.

Робота алгоритму проілюстрована на рис. 30-31.

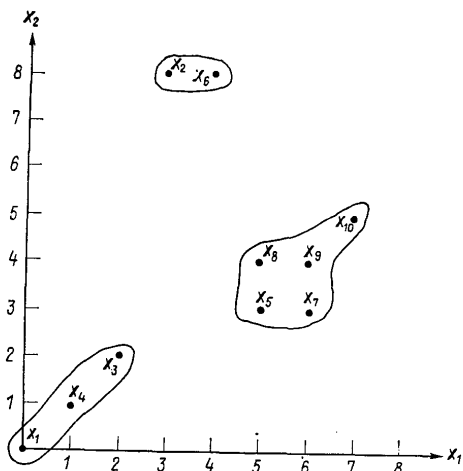


Рис. 30.

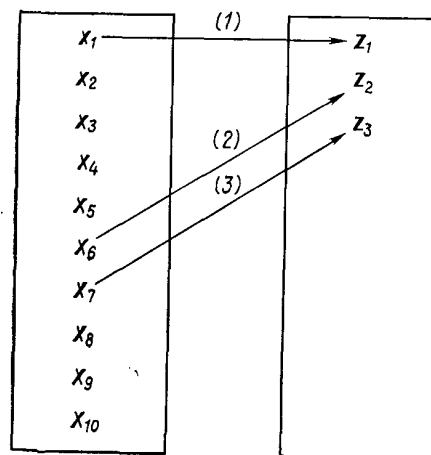


Рис. 31.

Зауваження. При практичному застосуванні алгоритму максимінної віддалі в якості k_i можна використовувати константу $1/2$.

3. Алгоритм k -means (групових середніх)

У цьому алгоритмі кількість кластерів скінченної множини $X \subset \mathbb{R}^n$ задається наперед та мінімізується показник якості кластеризації, який вважається рівним сумі квадратів відстаней усіх елементів кластера до його центра.

Крок 1. $r \leftarrow 0$ та вибрати вектори $\mathbf{z}_1^{(1)}, \dots, \mathbf{z}_k^{(1)}$ — початкові центри кластерів.

Крок 2. $r \leftarrow r + 1$ та розподілити усі об'єкти \mathbf{x} множини X по кластерам за наступним правилом:

$$\mathbf{x} \in C_j^{(r)} \Leftrightarrow \|\mathbf{x} - \mathbf{z}_j^{(r)}\| < \|\mathbf{x} - \mathbf{z}_i^{(r)}\|, \quad i = 1, \dots, k, \quad i \neq j.$$

Крок 3. Обчислити нові центри кластерів $\mathbf{z}_1^{(r+1)}, \dots, \mathbf{z}_k^{(r+1)}$ за формулою (3).

$$\mathbf{z}_i^{(r+1)} = \frac{1}{|C_i^{(r)}|} \sum_{\mathbf{x} \in C_i^{(r)}} \mathbf{x}, \quad i = 1, \dots, k, \quad (3)$$

де $|C_i^{(r)}|$ — кількість елементів кластера $C_i^{(r)}$.

Крок 4. Якщо для усіх центрів виконується умова $\mathbf{z}_i^{(r+1)} = \mathbf{z}_i^{(r)}$, $i = 1, \dots, k$, то робота алгоритму завершена, інакше перейти на крок 2.

Вибір центрів кластерів на кроці 1 може проводитися довільним чином. Можна показати, що центри кластерів, координати яких обчислюються за формулою (3), є оптимальними згідно до показника якості. Кавер та Шлезінгер довели, що алгоритм групових середніх збігається за скінченну кількість кроків.

Приклад. Проілюструємо алгоритм k -means на прикладі розбиття 20-елементної множини X , елементи якої наведені на рис. 32, на два кластери.

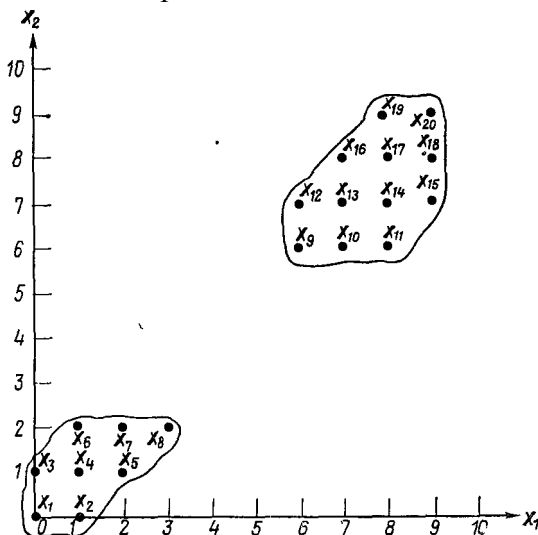


Рис. 32. Ілюстрація роботи алгоритму групових середніх

Крок 1. Покладаємо $\mathbf{z}_1^{(1)} = \mathbf{x}_1 = (0, 0)$, $\mathbf{z}_2^{(1)} = \mathbf{x}_2 = (1, 0)$.

Крок 2. $C_1^{(1)} = \{\mathbf{x}_1, \mathbf{x}_3\}$, $C_2^{(1)} = \{\mathbf{x}_2, \mathbf{x}_4, \mathbf{x}_5, \dots, \mathbf{x}_{20}\}$.

Крок 3. Перерахунок центрів кластерів

$$\mathbf{z}_1^{(2)} = \frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_3) = (0; 0,5), \quad \mathbf{z}_2^{(2)} = \frac{1}{18}(\mathbf{x}_2 + \mathbf{x}_4 + \dots + \mathbf{x}_{20}) = (5,67; 5,33).$$

Крок 4. Так як $\mathbf{z}_i^{(2)} \neq \mathbf{z}_i^{(1)}$, $i = 1, 2$, то виконується повернення на крок 2.

Крок 2. Використовуючи нові центри, отримуємо $C_1^{(2)} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_8\}$, $C_2^{(2)} = \{\mathbf{x}_9, \mathbf{x}_{10}, \dots, \mathbf{x}_{20}\}$.

Крок 3. Перерахунок центрів кластерів

$$\mathbf{z}_1^{(3)} = \frac{1}{8}(\mathbf{x}_1 + \mathbf{x}_2 + \dots + \mathbf{x}_8) = (1, 25; 1, 13), \mathbf{z}_2^{(3)} = \frac{1}{12}(\mathbf{x}_9 + \mathbf{x}_{10} + \dots + \mathbf{x}_{20}) = (7, 67; 7, 33).$$

Крок 4. Так як $\mathbf{z}_i^{(3)} \neq \mathbf{z}_i^{(2)}$, $i = 1, 2$, то виконується повернення на крок 2.

Крок 2. Використовуючи нові центри, отримуємо $C_1^{(3)} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_8\}$, $C_2^{(3)} = \{\mathbf{x}_9, \mathbf{x}_{10}, \dots, \mathbf{x}_{20}\}$.

Крок 3. Перерахунок центрів кластерів

$$\mathbf{z}_1^{(4)} = \frac{1}{8}(\mathbf{x}_1 + \mathbf{x}_2 + \dots + \mathbf{x}_8) = (1, 25; 1, 13), \mathbf{z}_2^{(4)} = \frac{1}{12}(\mathbf{x}_9 + \mathbf{x}_{10} + \dots + \mathbf{x}_{20}) = (7, 67; 7, 33).$$

Крок 4. Так як $\mathbf{z}_i^{(4)} = \mathbf{z}_i^{(3)}$, $i = 1, 2$, то робота алгоритму завершена.

Результатом алгоритму роботи буде розбиття $\{C_1, C_2\}$ множини X на кластери $C_1 = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_8\}$ та $C_2 = \{\mathbf{x}_9, \mathbf{x}_{10}, \dots, \mathbf{x}_{20}\}$ із центрами $\mathbf{z}_1 = (1, 25; 1, 13)$, $\mathbf{z}_2 = (7, 67; 7, 33)$

Алгоритм k -середніх має багато модифікацій.

Одна з найпростіших модифікацій — алгоритм k -медіан (partitioning around medoids). На кроці 3 цього алгоритмі у якості центру i -го кластера використовують вектор $\mathbf{z}_i^{(r+1)}$, кожна координата якого є медіаною відповідних координат векторів кластеру $C_i^{(r)}$, ($i = 1, \dots, k$).

4. Алгоритм ISODATA (алгоритм Болла-Холла)

Алгоритм ІЗОДАТА є модифікацією алгоритму групових середніх. Спочатку випадковим чином із множини X , яка підлягає кластеризації, вибираються центри k кластерів. Інші об'єкти множини X відносяться до кластера із найближчим центром. Далі знаходяться центри кластерів так само, як у методі групових середніх. Два кластери C_i та C_j зливаються, якщо статистична відстань $d(C_i, C_j)$ менша за задане порогове значення T_1 . Якщо середньо-групова дисперсія деякого кластера C_i по деякій координаті r ($1 \leq r \leq n$) більша за порогове значення T_2 , то відповідний кластер розбивається. Далі процедура застосовується до отриманих кластерів.