

І.С. Дмитрієва, Ю.М. Мусна

ДОСЛІДЖЕННЯ ІСНУЮЧИХ ПІДХОДІВ ДО ВИРІШЕННЯ ЗАДАЧІ РОЗПІЗНАВАННЯ ГРАФІЧНИХ ОБРАЗІВ

Анотація: Робота присвячена дослідженню існуючих підходів до вирішення задачі розпізнавання графічних образів, їх класифікації та узагальненню, огляду можливостей їх застосування.

Ключові слова: графічні образи, задача розпізнавання.

Розпізнавання образів можна визначити як віднесення вихідних даних до певного класу за допомогою виділення суттєвих ознак або властивостей, що характеризують ці дані, із загальної маси несуттєвих деталей.

Постановка задачі розпізнавання образів [1]: Дано безліч об'єктів. Щодо них необхідно провести класифікацію. Безліч представлено підмножинами, які називаються класами. Задано: інформація про класи, опис усієї множини і опис інформації про об'єкт, належність якого до певного класу невідома. Потрібно за наявною інформацією про класи і описі об'єкта встановити - до якого класу належить цей об'єкт.

Найбільш часто в задачах розпізнавання образів розглядаються монохромні зображення, що дає можливість розглядати зображення як функцію на площині. Вводячи поняття подібності між образами можна поставити задачу розпізнавання. Конкретний вид такої постановки сильно залежить від послідовних етапів при розпізнаванні відповідно до тих чи інших підходів.

Мета роботи: дослідити підходи до вирішення задачі розпізнавання текстових символів, як графічних образів.

Різні автори дають різну типологію методів розпізнавання графічних образів. Одні автори розрізняють параметричні, непараметричні й евристичні методи, інші - виділяють групи методів, виходячи з шкіл, що історично виникли, і напрямків у даній області. Наприклад, у роботі [2], описано два основні способи представлення знань: інтенціональне представлення - у вигляді схеми зв'язків між

атрибути (ознаками) та екстенціональне представлення - за допомогою конкретних фактів (об'єкти, приклади). Інша класифікація базується на реалізації трьох основних типів методології: евристичний, математичний й лінгвістичний (синтаксичний). Доволі часто системи розпізнавання графічних образів створюються на основі комбінації цих методів.

Евристичні методи. За основу евристичного підходу взяті інтуїція й досвід людини. У ньому використовуються принципи перерахування членів класу й спільності властивостей. Звичайно, системи, побудовані такими методами, включають набір специфічних процедур, розроблених виключно до конкретних задач розпізнавання. Хоча евристичний підхід відіграє більшу роль у побудові систем розпізнавання образів, неможна узагальнити принципи синтезу, тому що розв'язок кожної конкретної задачі вимагає використання специфічних прийомів розробки системи. Це означає, що структура і якість евристичної системи в значній мірі визначаються обдарованістю й досвідом розробників.

Структурні (лінгвістичні) методи. Ці методи базуються на використанні спеціальних граматики мов що породжують, за допомогою яких можна описувати сукупність властивостей об'єктів, що розпізнаються. Кожен об'єкт представляється сукупністю непохідних елементів «сполучених» між собою тими або іншими способами або, іншими словами «реченням» деякої «мови». При цьому важливою є інформація, яка описує структуру кожного об'єкта, а від процедури розпізнавання вимагається, щоб вона давала можливість не лише віднести об'єкт до певного класу, але й описати ті властивості об'єкта, які унеможливають віднесення його до іншого класу. Наприклад, зображення на рисунку 1 може бути описано за допомогою ієрархічних структур, показаних на рисунку 2.

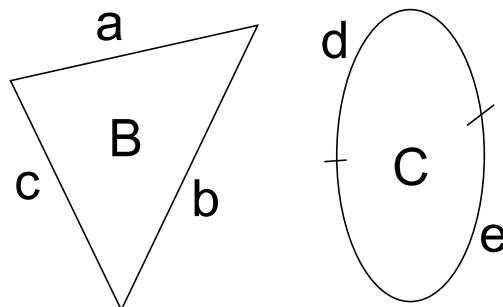


Рисунок 1 - Зображення А

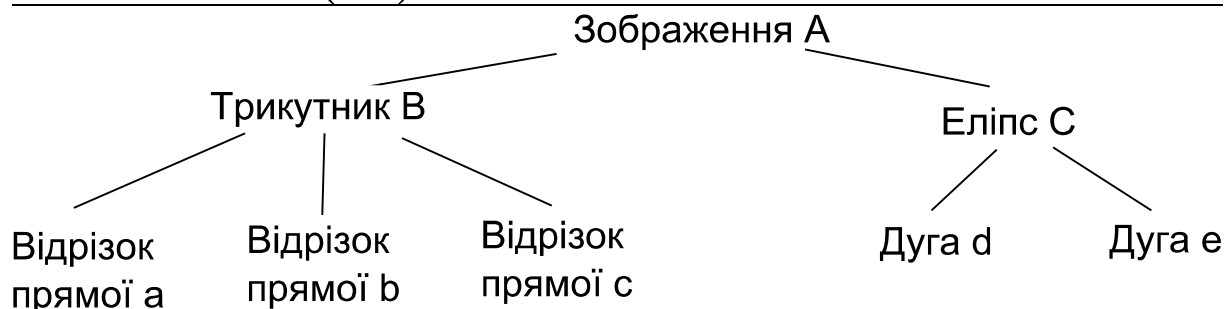


Рисунок 2 - Ієрархічний структурний опис зображення А

Процес розпізнавання здійснюється після ідентифікації в об'єкті непохідних елементів і складанні опису об'єкту. Розпізнавання полягає в синтаксичному аналізі, або граматичному розборі, «речення», що описує даний об'єкт. Ця процедура встановлює, чи є ця речення синтаксично правильним по відношенню до заданої граматики. Синтаксичний аналіз дає деякий структурний опис речення (зазвичай у вигляді дерева).

Для того щоб отримати граматику, що описує структурну інформацію класу образів, необхідний пристрій виводу граматики, що дозволяє відновити її по заданій множині об'єктів у вигляді, подібному мовному. Функції цього пристрою аналогічні процесу «навчання» у дискримінантних системах розпізнавання образів. У результаті навчання на прикладах об'єктів розглянутого класу формується структурний опис цього класу. Потім отриманий у вигляді граматики опис використовується для опису образів і синтаксичного аналізу.

Математичні методи. В основу математичного підходу покладені правила класифікації, які формулюються й виводяться в рамках певного математичного формалізму за допомогою принципів спільності властивостей і кластеризації. Математичні методи побудови систем розпізнавання можна розділити на два класи: детерміністські й статистичні.

Детерміністський підхід базується на математичному апараті, що не використовує в явному вигляді статистичні властивості досліджуваних класів образів. Статистичний підхід ґрунтується на математичних правилах класифікації, які формулюються й виводяться в термінах математичної статистики. Побудова статистичного класифікатора в загальному випадку припускає використання байєсівського класифікаційного правила і його різновидів. Це правило забезпечує одержання оптимального класифікатора в тих випад-

ках, коли відомі щільності розподілу для всіх сукупностей образів і ймовірності появи образів для кожного класу. Істинні ймовірності характеристик класів зазвичай невідому. Тому практично використовуються оцінки, отримані по відомим прикладам класів. Цю, звичайну в статистиці процедуру оцінки по вибірці імовірнісних характеристик генеральної сукупності інтерпретують як навчання. Вирішуючи правила фактично не залежать від специфіки об'єктів, що розпізнаються, оскільки вони мають справу не з самими об'єктами, а з деякими числовими характеристиками. Внаслідок цього ці правила мають відому загальність. Критерієм оптимальності є вимога, щоб загальна ймовірність неправильного розпізнавання була мінімальною.

Серед можливих оптимальних вирішальних функцій найбільше докладно розглянуті її лінійні наближення, тобто лінійні вирішальні функції. Лінійність тут означає, що пари узагальнених образів дихотомічно розділяється в просторі образів однієї й тільки однієї гіперплощиною виду $\sum_{i=1}^n \lambda_i x_i - \lambda_{n+1} = 0$, де x_i - координати простору образів.

За основу процесу розпізнавання образів прийнятий перерозподіл вихідних ймовірностей $\{p_i\} = p_1, p_2, \dots, p_w$ взаємовиключних, і таких що й представляють повну групу класів $\{S_i\} = S_1, S_2, \dots, S_w$, до яких може бути віднесений розпізнаваний образ, тобто $\sum_{i=1}^w p_i = 1, 0 \leq p_i \leq 1$, і, крім того $\{p(s_x \in S_i)\} = p(s_x \in S_1), p(s_x \in S_2), \dots, p(s_x \in S_w)$, для яких виконується умова нормування $\sum_{i=1}^w p(s_x \in S_i) = 1$.

Приналежність s_x до одного з класів встановлюється при умові, що множина ймовірностей приналежності s_x до класів $S_1, S_2, \dots, S_w$ вдовольняє вимогам, що спричинені необхідною достовірністю відповідей і полягають, наприклад, в тому, що для деякого значення i виконується всі нерівності

$$\frac{p(s_x \in S_i)}{p(s_x \in S_j)} \geq A_{ij}, \quad j = \overline{1, w},$$

де A_{ij} - деяке додатне число.

Інакше кажучи, мається на увазі, що за основу процесу розпізнавання беруть процес пере розподілення вихідних ймовірностей за допомогою аналізу сукупності отриманих реалізацій випадкового процесу, що представляє собою образ, що розпізнається, і отримання за рахунок цього необхідної інформації.

Необхідно зазначити, що всі опубліковані статистичні моделі розпізнавання образів, що використовують фактично різні статистики, дають дуже гарні результати розпізнавання в різних задачах - медичної діагностики, розпізнаванні рукописних символів, класифікації даних сейсморозвідки і т.д.

Формалізм побудови узагальнених графічних моделей поєднує в собі багато методів статистичного моделювання, таких як факторний аналіз, аналіз розподілів, моделі сумішей розподілів, приховані марковські моделі, фільтри Калмана, моделі Айзинга та деякі інших. Всі зазначені моделі можна розглянути в рамках графічних моделей байєсівського типу як окремі приклади загального формалізму. Перевагою такого підходу є те, що методи дослідження процесів та обробки даних, розроблені в одній області, можуть бути успішно перенесені в інші.

Одним з найбільш популярних типів графічних моделей є мережа Байєса [3]. Головна її особливість в тому, що ребра в мережі є направленими та відображають умовну залежність між змінними.

Вираз

$$p(H_k | E) = \frac{p(E | H_k) \cdot p(H_k)}{\sum_{i=1}^n p(E | H_i) \cdot p(H_i)}, \quad (1)$$

який представляє собою формулу Байєса, є основою, на якій будуються мережі Байєса. В (1) H_k означає будь-яку гіпотезу з n можливих. Ймовірності $p(E | H_k)$ задаються експертами апіорно. Ці ймовірності є дуже корисними, тому що, як правило, легше знайти ймовірність послідовності подій типу причина-наслідок, ніж навпаки. Значення $p(H_k)$ називають апіорними ймовірностями, вони визначають початкові ймовірності для всіх гіпотез. Сила байєсового методу полягає в тому, що апіорні ймовірності можна уточнювати (оновлювати) у відповідності до реалій протікання процесу, що досліджується. Це дозволяє уточнювати ймовірності подій при

надходженні додаткової інформації. Знаменник виразу (1) можна розглядати як нормуючий член.

Типи байєсівських мереж (БМ).

Дискретні БМ – мережі, у яких змінні вузлів представлені дискретними величинами. Вони мають такі властивості: кожна вершина представляє собою подію, яка описується випадковою величиною, яка може мати кілька станів; всі вершини, пов'язані з «батьківськими», визначаються таблицею умовних ймовірностей або функцією умовних ймовірностей; для вершин без «батьків» ймовірності її станів є безумовними. Інакше кажучи, у байєсівських мережах довіри вершини представляють собою випадкові змінні, а дуги – імовірнісні залежності, які визначаються через таблиці умовних ймовірностей. Таблиці умовних ймовірностей кожної вершини містить ймовірності станів цієї вершини за умови станів її «батьків».

Динамічні БМ – мережі, у яких значення вузлів змінюються з часом.

Неперервні БМ – мережі в яких змінні вузлів – це неперервні величини. У багатьох випадках події можуть приймати будь-які стани з деякого діапазону. Тобто змінна X буде неперервною випадковою величиною, простором можливих станів якої буде весь діапазон її припустимих значень $X = \{x \mid a \leq x \leq b\}$, яке містить нескінченну множину точок. В цьому випадку некоректно говорити про ймовірності окремого стану, тому що при їх нескінченно великій кількості вага кожного буде дорівнювати нулю. Неперервні БМ використовуються для моделювання стохастичних процесів у просторі станів з неперервним часом.

Гібридні БМ – мережі, які містять вузли з дискретними і неперервними змінними.

При використанні БМ, що містять неперервні і дискретні змінні, існує ряд обмежень:

1. дискретні змінні не можуть мати неперервних батьків;
2. неперервні змінні повинні мати нормальний закон розподілу, умовний на значеннях батьків;
3. розподіл неперервної змінної X з дискретними батьками Y та неперервними батьками Z є нормальним розподілом:

$$P(X \mid Y = y, Z = z) = N(\mu_x(\mu_y, \mu_z), \sqrt{\sigma_x(\sqrt{\sigma_y})}),$$

де μ_x, μ_y, μ_z – математичні очікування, σ_x, σ_y – дисперсії, $\sqrt{\sigma_x}, \sqrt{\sigma_y}$ – середньоквадратичне відхилення; μ_x лінійно залежить від неперервних батьків, а σ_x взагалі не залежить від неперервних батьків. Однак, μ_x та σ_x залежать від дискретних батьків. Це обмеження гарантує можливість формування точного висновку.

Висновки. В роботі було проведено дослідження існуючих підходів до вирішення задачі розпізнавання графічних образів, їх класифікації. Пропонується для задачі розпізнавання текстових символів, як графічних образів, застосувати підхід з використанням різноманітних мереж Байеса.

ЛІТЕРАТУРА

1. Гонсалес Р., Вудс Р. Цифровая обработка изображений. - М.: Техносфера, 2005. —1072 с.
2. Лебедев А.А. Введение в анализ и синтез систем - М.: Изд-во МАИ, 2001. — 352 с.
3. Згуровський М. З., Бідюк П. І., Терентьев О. М., Просянкіна-Жарова Т. І. Байєсівські мережі в системах підтримки прийняття рішень — К: ТОВ«Видавниче Підприємство «Едельвейс», 2015. — 300 с.