

III

Теоретичні основи обробки та розпізнавання зображень

Theoretical Basis of Image Processing and Recognition

Метод підвищення розмірності сигнатур при обробці зображень в оптико-цифровій системі

Роман Бачевський, Леонід Муравський,
Юрій Камінський, Аркадій Стефанський

Фізико-механічний інститут АН України

Україна, 290601, Львів, МСП
бул. Наукова, 5
тел.: (032) 263-33-88

При розв'язанні різноманітних задач обробки зображень і розпізнавання образів широке застосування знаходить гібридні оптико-цифрові системи (ОЦС), в яких висока швидкодія та паралелізм оптичних методів обробки поєднуються з точністю та програмно-алгоритмічною гнучкістю [1]. В багатьох гібридних ОЦС, в яких проводиться аналіз і просторово-частотна дискретизація оптичних фур'є-спектрів досліджуваних зображень чи об'єктів, використовуються фотоелектронні мозаїки або скануючі діафрагми з комбінованою секторно-кільцевою топологією. Наявність в системі такого пристрою дозволяє проводити ізочастотно-азимутальний аналіз спектрів, який вважається оптимальним при відборі ознак, інваріантних до масштабно-кутових перетворень. Разом з тим, при дослідженні складноструктурованих зображень, зокрема матеріалів космоаерофотозйомки, розмірність сигнатур, тобто кількість базових ознак, що формуються безпосередньо в результаті просторово-частотної дискретизації спектру, може виявитись недостатньою для реалізації в цифровій частині ОЦС алгоритмів структурометричного аналізу, морфоструктурної класифікації, розпізнавання, тематичного дешифрування тощо. З метою значного підвищення ефективності і достовірності виконання вказаних алгоритмів було запропоновано підвищити розмірність сигнатур за рахунок додаткового розбиття кільцевих зон фотоелектронної мозаїки на секторні ділянки [2].

Однак технологія виготовлення такого роду мозаїки набагато складніша і достовірність адекватного відтворення вихідними сигналами відповідних дискретизованих ділянок енергетичного спектру значно нижча. Крім того, значне збільшення фотоприймальних елементів у мозаїці приводить по суттєвого ускладнення електронних блоків управління і комутації, а також до підвищення шумів комутації.

В данній роботі запропоновано новий метод формування сигнатури розмірністю $M \times N$ на базі сигнатури розмірністю $M + N$, що утворюється в результаті дискретизації енергетичного спектру вхідного зображення секторно-кільцевою фотоелектронною мозаїкою, котра складається з M секторних і N кільцевих елементів. Вказаний метод дозволяє обчислювати в цифровій частині ОЦС нову геометрію дискретизації, яка утворюється внаслідок взаємного перекриття m -го сектора дискретизації з n -м кільцем ($m = 1, \dots, M$; $n = 1, \dots, N$) і адекватна дискретизації кожної з півплощин спектру на $M + N$ секторно-кільцевих зон. Кожен елемент такої дискретизації можна описати потужністю спектру в

області, обмеженій інтервалом частот ($\rho_m, \rho_m + \Delta \rho$) і сектором радіальних напрямів (азимутів) [$\Theta_n, \Theta_n + \Delta \Theta$]:

$$K_{m,n} = \int_{\rho_m}^{\rho_m + \Delta \rho} \int_{\Theta_n}^{\Theta_n + \Delta \Theta} I(\rho, \Theta) \rho d\rho d\Theta, \quad (1)$$

де $I(\rho, \Theta)$ — інтенсивність спектру в деякій довільній точці. Індекс m, n вказує на те, що дана область знаходиться в місці перетину n -го кільця та m -го сектору.

Сигнал, отриманий від m -го сектору, можна представити як суму

$$S_m = \sum_{n=1}^N K_{m,n}, \quad (2)$$

а сигнал, отриманий від n -го кільця

$$r_n = \sum_{m=1}^M K_{m,n}. \quad (3)$$

Система, утворена з рівнянь (2) і (3), має $M+N$ рівнянь і $M \times N$ невідомих. В загальному випадку $M \times N > M+N$ при $M > 2$ і $N > 2$ і система є невизначеню. Пропонований метод ґрунтуються на можливості розбити досліджувану область на меншу кількість дискретних частин і перейти до системи 4-х рівнянь з 4-ма невідомими для визначення потужності будь-якого (m, n) -го елементу дискретизованого спектру.

Процедура знаходження n' -го ($1 \leq n' \leq N$) елемента I-го сектора вимагає виконання двох обчислювальних кроків. На першому кроці формується система

$$\left\{ \begin{array}{l} \sum_{n=1}^{n'} K_{1n} + \sum_{n=n'+1}^N K_{1n} = S_1 \\ \sum_{m=2}^M \sum_{n=1}^{n'} K_{mn} + \sum_{m=2}^M \sum_{n=n'+1}^N K_{mn} = \sum_{m=2}^M S_m \\ \sum_{n=1}^{n'} K_{1n} + \sum_{m=2}^M \sum_{n=1}^{n'} K_{mn} = \sum_{n=1}^{n'} r_n \\ \sum_{n=n'+1}^N K_{mn} + \sum_{m=2}^M \sum_{n=1}^{n'} K_{mn} = \sum_{n=n'+1}^N r_n. \end{array} \right. \quad (4)$$

Оскільки дана система лінійних рівнянь містить 4 невідомих $\sum_{n=1}^{n'} K_{mn}$,

$$\sum_{n=n'+1}^N K_{mn}, \quad \sum_{m=2}^M \sum_{n=1}^{n'} K_{mn}, \quad \sum_{m=2}^M \sum_{n=n'+1}^N K_{mn}, \text{ то вона має розв'язок.}$$

На другому кроці формуються дві системи лінійних рівнянь

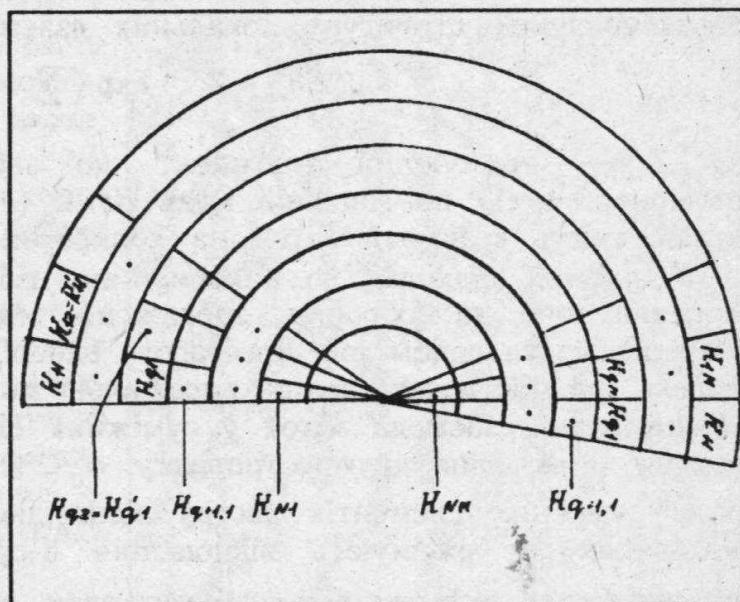
$$\left\{ \begin{array}{l} \sum_{n=1}^{n-1} K_{1n} + K_{1n'} = \sum_{n=1}^{n'} K_{1n} \\ \sum_{m=2}^M K_{1n} \sum_{n=1}^{n'-1} K_{mn} + \sum_{m=2}^M K_{mn'} = \sum_{m=2}^M \sum_{n=1}^{n'} K_{mn} \\ \sum_{n=1}^{n-1} K_{1n} + \sum_{m=2}^M \sum_{n=1}^{n'-1} K_{mn} = \sum_{n=1}^{n'-1} r_n \\ K_{1n'} + \sum_{m=2}^M K_{mn'} = r_{n'} \end{array} \right. , \quad (5)$$

$$\left\{ \begin{array}{l} K_{m,n'+1} \sum_{m=2}^N K_{mn} = \sum_{n=n'+1}^N K_{mn} \\ \sum_{m=2}^M \sum_{n=n'+2}^N K_{mn} + \sum_{m=2}^M K_{m,n'+1} = \sum_{m=2}^M \sum_{n=n'+1}^N K_{mn} \\ \sum_{n=n'+1}^N K_{mn} + \sum_{m=2}^M \sum_{n=n'+1}^N K_{mn} = \sum_{n=n'+1}^N r_n \\ K_{1,n'+1} + \sum_{m=2}^M K_{m,n'+1} = r_{n'+1} \end{array} \right. , \quad (6)$$

з яких поряд з іншими невідомими сумами знаходимо значення $K_{n'+1}$ і $K_{n'+1,m}$. Провівши в системах (1), (2) і (3) заміну n' на $(n'+2)$ або n' на $(n'-2)$, здійснююмо обчислення $(n'+2)$ -го і $(n'+3)$ -го або $(n'-2)$ -го і $(n'-1)$ -го елементів і т. д., тобто знаходимо значення всіх елементів першого сектору.

Враховуючи центральну симетричність спектру в площині Фур'є-перетворення, нумерацію можна починати з будь-якого сектора. Ця обставина дозволяє першому сектору присвоїти останній номер, а другому відповідно перший і аналогічним чином за допомогою систем рівнянь (4) – (6) знайти всі його елементи $K_{n',2} = K_{n',1}'$. Повторивши описану процедуру для решти секторів, знаходимо весь масив $M \times N$ елементів. Нова геометрія дискретизації наведена на рисунку.

Застосування цього алгоритму дозволяє значно збільшити ступінь дискретизації дифракційних картин без додаткових витрат на виготовлення складнішої фотоприймальної матриці. Час запису аналізованого зображення і необхідна для цього машинна пам'ять при цьому не зростають. Крім цього, значно збільшується розмірність сигнатур, що дозволяє суттєво підвищити достовірність алгоритмів структурного аналізу, класифікації, кластеризації, дешифрування зображень і розпізнавання образів.



Синтезована геометрія дискретизації спектру вхідного зображення $(q - n')$

Література

1. Применение методов Фурье-оптики / Под ред. Г. Старка. — М. : Радио и связь, 1988. — 536 с.
2. Янущ Д.А., Скуброва Н.В. Структуро-зональный анализ изображений при геологическом дешифрировании аэро- и космических снимков // Структурный анализ снимков в аэрокосмических исследованиях Земли. — М., 1985. — С. 77–97.



Гіббсівські випадкові поля як породжуючі моделі текстурних карт сегментів

Георгій Гімельфарб, Олексій Залесний

Інститут кібернетики АН України

Україна, 252207, Київ
просп. Академіка Глушкова, 40
тел.: (044) 266-25-69

Пропонуються нові ймовірносні моделі зображення, що породжують цифрові карти сегментів, на які поділено зображення, як вибіркові реалізації певного марківського випадкового поля (МВП) з гіббсівським розподілом імовірностей (ГРІ) сигналів. Кarta сегментів визначається як цифрове зображення $\bar{h}: \Omega \rightarrow K$, задане на растрі або скінченних числових решітках $\Omega = \{ i : i = 0, \dots, N - 1 \}$. Сигнали h_i в елементах растра приймають значення із скінченої множини $K = \{ 1, \dots, q \}$ і інтерпретуються як мітки або коди назв сегментів, у які входять ці елементи.

ГРІ $P(\bar{h})$ на множині H всіх можливих карт \bar{h} сегментів факторизується по мітках h_i для певної підмножини $\alpha \subset \Omega$ елементів растра. Підмножини α покривають весь растр і розглядаються як повні підграфи, або кліки, що характеризують структуру локальних взаємодій міток на растрі:

$$P(\bar{h}) = Z^{-1} \exp \left(\sum_{\alpha \subset \Omega} V_{\alpha}(h_i : i \in \alpha) \right), \quad (1)$$

де Z — нормуючий множник, що забезпечує одиничне значення суми імовірностей (1) по множині карт H ; $V_{\alpha}(\cdot)$ — задані функції, що визначають інтенсивність взаємодії міток на конкретних кліках $\alpha \subset \Omega$.

У відомих підходах до використання гіббсівських породжуючих моделей для моделювання та обробки зображень множина карт сегментів описується здебільшого за допомогою просторово однорідної багаторівневої логістичної (БРЛ) моделі, що має досить великі виражальні можливості [1–3]. Ця модель враховує тільки парні взаємодії міток у суміжних елементах растра, для чого задаються чотири сім'ї клік другого порядку $\alpha_a \subset \Omega$, $a = \{ h, v, d_1, d_2 \}$, що об'єднують пари суміжних елементів растра відповідно по горизонталі, вертикалі й двох діагоналях та формують відношення 8-сусідства на растрі. Функція $V_{\alpha}(\cdot)$ однаакова для всіх клік і набуває двох значень: λ при однакових i , 0 при неоднакових мітках h_i та h_j , $i, j \in \alpha$, на кліці. Значення λ визначає інтенсивність взаємодії міток. ГРІ для БРЛ-моделі може бути записаний як

$$P(\bar{h}) = Z^{-1} \exp \left(\lambda \sum_a \sum_{\alpha_a \subset \Omega} \delta(h_i, h_j : i, j \in \alpha_a) \right), \quad (2)$$

де $\delta(\cdot)$ — функція Кронекера. У цій моделі при $\lambda > 0$ більш імовірні карти сегментів, що мають меншу кількість граничних елементів растра (елементів, з якими межують елементи з відмінними мітками).

Карти сегментів, описані за допомогою ГРІ типу (1) та (2), можуть породжуватися шляхом поелементної стохастичної релаксації [2, 3].

Можна запропонувати подальші узагальнення моделі (2), що поширяють її виражальні можливості на текстурні карти: БРЛ-модель з близькими та далекими взаємодіями міток (БДВ-модель) і модель виділених структурних взаємодій міток (СВ-модель). У цих моделях більш імовірні карти, на яких сегменти мають відносно велику кількість граничних елементів растра, але відрізняються певною регулярністю розміщення та складністю форм зв'язних фрагментів кожного сегмента.

У БДВ-моделі попередні чотири сім'ї клік близької взаємодії доповнюються аналогічними чотирма сім'ями двоелементних клік далекої взаємодії, елементи яких розташовані на відстані a один від одного. Відстань визначається як максимальна із різниць позицій по горизонталі та вертикальні для пари елементів.

Для кожної групи сімей клік — близької і далекої взаємодії — вводиться свій параметр λ (відповідно λ_{sr} і λ_b) у ГРІ типу (2). У більш загальному випадку кліки далекої взаємодії можуть бути повернуті відносно клік близької взаємодії шляхом такого зсування одного з елементів на b позицій по вертикальні або горизонтальні, яке зберігає незмінною відстань a між елементами однієї кліки. Коли $b = 0$ і відсутні близькі взаємодії ($\lambda_{sr} = 0$), растр поділяється на a^2 незалежних підрастрів, на кожному з яких формується своя карта, що відповідає БРЛ-моделі (2). При $b > 0$ кількість незалежних підрастрів скорочується (у певних випадках — до єдиного взаємозалежного растра Ω).

Подальше узагальнення БДВ-моделі полягає в тому, що для деяких сегментів вводиться анізотропність по напрямках, тобто залежність функції $V(\cdot)$ від номера a сім'ї клік для окремих значень міток. Найпростіший приклад такої анізотропії — відсутність взаємодії у певних напрямках для деяких значень міток.

У СВ-моделі задається одна сім'я клік дев'ятого порядку, що утворюють на растрі квадратні «вікна» розміром 3×3 елементів. Функція $V_\alpha(\cdot)$, що характеризує інтенсивність взаємодій міток на цих кліках, однаакова для клік кожного конкретного сегменту і набуває двох можливих значень: λ для кількох вилучених конфігурацій своїх і чужих міток на кліці та 0 для всіх інших конфігурацій. Своєю вважається мітка, що збігається з міткою центрального елемента вікна, чужою — будь-яка інша у межах кліки. Усього існує 256 таких конфігурацій міток.

На картах, які породжує ця модель при $\lambda > 0$, більш імовірні сегменти, на яких домінують вилучені конфігурації міток.

Експерименти підтверджують можливість опису і породження різноманітних текстурних зображень за допомогою запропонованих гіббсівських моделей карт сегментів.

Література

- Derin H., Elliot H. Modelling and segmentation of noisy and textured images using Gibbs random fields // IEEE Trans. on Pattern Analysis and Machine Intellig. — 1987. — 9, № 1. — P. 39–55.

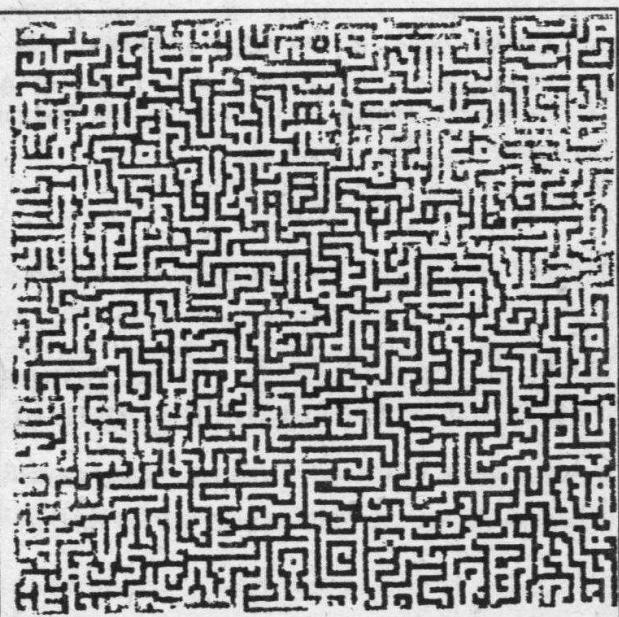


Рис. 1. СВ-модель

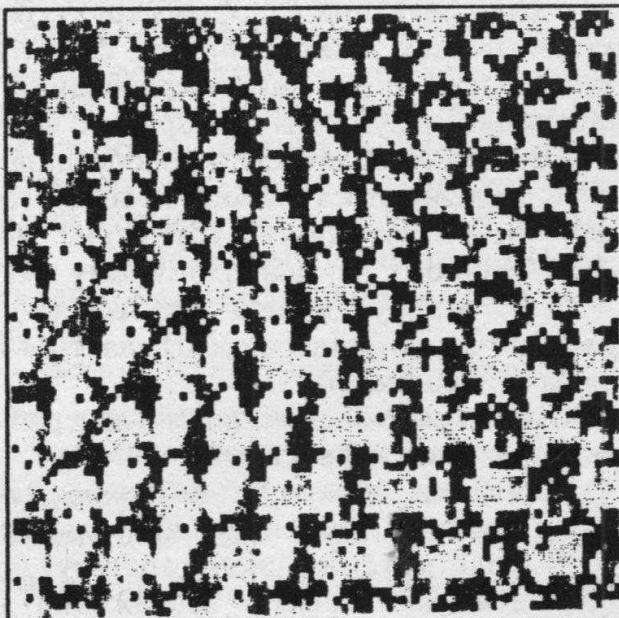


Рис. 2. БДВ-модель

2. Marroquin J., Mitter S., Poggio T. Probabilistic solution of ill-posed problems in computational vision // J. of the Amer. Statistical Ass. — 1987. — 82, № 397. — P. 76–89.
3. Gimelfarb G.L., Zalesny A.V. Low-level Bayesian segmentation of piecewise-homogeneous noisy and textured images // Intern. J. of Imaging Systems and Technology. — 1991. — 3, № 3. — P. 227–243.



Модель сприйняття та інтерпретації образної візуальної інформації

Володимир Грицик, Роман Бачевський, Елеонора Захарко

Фізико-механічний інститут АН України

Україна, 290601, Львів, МСП
бул. Наукова, 5
тел.: (032) 263-31-44

У роботі розвивається семантико-асоціативний підхід до проблем інтерпретації та дешифрування складноструктурзованих зображень, акцентується увага на взаємозв'язку синтаксичних, семантичних і прагматичних аспектів сприйняття та подається варіант моделювання механізмів, принципів й закономірностей переробки візуальної інформації людиною.

Основні особливості пропонованої моделі зводяться до таких:

1. Моделюються деякі процеси сприйняття візуальної інформації людиною, але не аналітичні структури мозку. Модель не включає пояснень на фізіологічному й нейрологічному рівнях.

2. В основу моделі покладено діяльність висококваліфікованого експерта. Експерт-інтерпретатор розглядається як величезний і досконалій перетворювач інформації.

3. Центральний механізм переробки візуальної інформації приймається паралельним багаторівневим.

4. Модель враховує не лише синтаксис, але і семантику та прагматику, тобто зміст і тематичну актуальність вхідної інформації.

У ролі математичного апарату представлення синтаксичних, семантичних та прагматичних аспектів зображення використовується теорія образів.

Модель сприйняття візуальної інформації людиною розглядається як у спонтанному (природному) режимі, так і в наяв'язаному, тобто із використанням технічних систем, що здійснюють обробку та представлення зображень. Це дозволяє поставити конкретні вимоги як до умов здійснення інтерактивного режиму (ідеється про визначення набору та логіки задання послідовності алгоритмів перетворення та представлення зображень, що відповідає логіці сприйняття та інтерпретації візуальної інформації людиною), так і до фізичних принципів реалізації відповідної технічної системи.

Ефективність використання моделі та семантико-асоціативного підходу в цілому проілюстровані на прикладах обробки, аналізу та дешифрування складноструктурзованих зображень в задачах геології, екології, медицини та інших.



Структурний алгоритм розпізнавання об'єктів на графічних зображеннях

Володимир Кийко

Інститут кібернетики АН України

Україна, 252207, Київ 207
просп. Академіка Глушкова, 40
Тел.: (044) 266-25-69

Припускається, що зображення на вході алгоритму подано у вигляді графа, дужкам якого відповідають відрізки «скелету» зображення, а вершинам — точки перетину цих відрізків [1]. Задача розпізнавання полягає у виявленні на цьому графі всіх підграфів, визначених за допомогою граматики $G = (V_T, V_N, F, E)$, де V_T — множина термінальних символів, V_N — множина нетермінальних символів, F — множина так званих первинних фрагментів зображення і, нарешті, E — сукупність додаткових метричних і структурних обмежень на шуканий об'єкт зображення.

Множина V_T складається з відрізків горизонтальних, вертикальних і похилих (під кутами кратними 45°) ліній, узагальнених опуклих і увігнутих дуг, а також нульового (початкового) і пустого («байдужого») термінальних елементів. Узагальнена опукла (увігнута) дуга є послідовністю відрізків ліній на зображені, кожен із яких повернутий за (проти) годинниковою стрілкою по відношенню до попереднього відрізка. Нульовий термінальний елемент є деякою точкою на скелеті зображення. Пустий термінальний елемент є «білим» відрізком лінії, тобто є частиною фону і використовується для задання групи незв'язаних між собою частин зображення. Кожний термінальний елемент має дві кінцеві точки: початкову та кінцеву, а також задовільняє у загальному випадку деяким обмеженням, таким наприклад, як наявність або відсутність певних ліній, що проходять через кінцеву точку цього елемента.

Множина $V_N = S_E \cup S_c$ нетермінальних символів (структурних елементів) є множиною міток, що відповідають кінцевим точкам термінальних елементів граматики. Початковій точці термінального елемента при цьому ставиться у відповідність один структурний елемент $s \in S_c$, а кінцевій точці — 0 або 1 структурний елемент $s \in S_E$, а також довільна, загалом кажучи, кількість структурних елементів $s \in S_C$. Пара $C = \{ \text{кінцева точка}, \text{структурний елемент } s \in S_C \}$ називається стиком.

Множина первинних фрагментів F є головною складовою частиною граматики G . Деякий первинний фрагмент $f(v, s_1, s_2) \in F$ вважається заданим, якщо:

- задано термінальний елемент $v \in V_T$;
- кінцевим точкам цього елемента поставлені у відповідність деякі підмножини $S_1 \subset S_N$ і $S_2 \subset S_N$ структурних елементів.

Справедливе просте правило допустимої композиції фрагментів: два фрагменти можна стикувати тільки вздовж їх спільного стику C . Наслідком цієї операції є породження деякого фрагменту F_n , який не належить до множини первинних фрагментів F і не має стику C . Процес породження зображення граматикою визнається закінченим, якщо на деякому кроці отримано фрагмент зображення, що не має жодного стику і задовільняє множині E додаткових метричних і структурних обмежень на цей фрагмент.

Початковий етап процесу розпізнавання полягає в перегляді вершин графа, що представляє скелет зображення, і пошуку такої вершини, яка задовільняє ознакам нульового первинного фрагмента. Первінний фрагмент граматики,

відповідний нульовому термінальному елементу, має один чи два стики і представляє на даному етапі так званий проміжний фрагмент F_n об'єкту зображення, який необхідно виділити. Наступна робота алгоритму полягає в змінюванні фрагмента F_n шляхом пошуку (прослідження) на зображенні та приєднання до його стиків інших первинних фрагментів із застосуванням відомих процедур граматичного розбору [2] під час пошуку кожної допустимої послідовності фрагментів. Цей процес визнається успішно закінченим, якщо вдається сформувати фрагмент F_n , який не має вільних стиків і задовільняє всім обмеженням $r \in E$. В протилежному разі процес виділення об'єкта зображення припиняється і виконується перехід до пошуку наступного нульового термінального елемента.

Розглянутий алгоритм був реалізований у програмному комплексі введення, обробки і розпізнавання принципіальних електрических схем [3] і забезпечує можливість гнучкого настроювання системи на необхідний алфавіт складових елементів схеми, а також оперативного вводу непередбачених раніше об'єктів. Проведені численні експерименти показали високу швидкодію і достатньо велику надійність розпізнавання графічних зображень.

Література

1. Кийко В.М. Швидкодіючий алгоритм скелетизації графічних зображень // Теоретичні та прикладні питання розпізнавання зображень. — Київ : ІК АН України, 1991. — С. 5–15.
2. Фу К. Структурні методи в розпізнаванні образів. — М. : Мир, 1977. — 320 с.
3. Система введення, обробки та розпізнавання графічних зображень на персональному комп'ютері / С.Л. Александрова, В.М. Кийко, В.В. Мацелло, М.І. Шлезінгер // У цьому ж збірнику.



Перетворення типу кореляційних при виділенні геометро-топологічних ознак бінарних зображень

Микола Кириченко, Андрій Куценко

Чернівецький університет

Україна, 274012, Чернівці
вул. Університетська, 2
тел.: (037) 229-84-48

Було проведено дослідження перетворення типу автокореляційної згортки функції освітленості за напрямками. В результаті інтерпретації першої похідної згортки, запропоновано декілька нових конструктивних алгоритмів виділення геометро-топологічних ознак площинних зображень.

Більшість ознак, одержаних після застосування розроблених алгоритмів, має властивість інваріантності відносно широких класів перетворення площини. Інтегральний характер перетворень, а отже однорідність операцій при обчислennях, дозволяють використовувати при моделюванні цих перетворень на ЕОМ режим паралельної обробки інформації.

В деяких випадках передбачається можливість ефективної реалізації запропонованих алгоритмів на базі сучасних засобів когерентної оптики. Множина запропонованих алгоритмів виділення ознак систематизована, і у вигляді деревоподібної структури представлена методика обробки бінарних зображень.

Частина алгоритмів запрограмована, і було проведено обчислювальний експеримент по знаходженню характеристик зображень.

Література

1. Кириченко Н.Ф., Куценко А.А. Автокорреляционные методы выделения геометро-топологических признаков бинарных изображений — Київ, 1991. — 21 с. — Деп. в УкрНИИНТИ 08.05.91 № 645-Ук 91.

Морфологічні пристрої картиної обробки та аналізу зображень

Володимир Красиленко, Наталя Дем'янчук

Вінницький політехнічний інститут

Україна, 286021, Вінниця
Хмельницьке шосе, 95
Тел.: (043) 222-57-18

При обробці зображень у багатьох випадках застосування класичних підходів, заснованих на принципах та теорії лінійних систем та на перетвореннях Фур'є та йому подібних, обмежене тим, що вони не дозволяють чисельно описати форму або геометричну структуру сигналів. Теоретико-множинний підхід по аналізу зображень, що розвивається в останній час, відомий під назвою математичної морфології, дозволяє здійснювати строгий кількісний опис особливостей геометричної структури сигналів в зручному для людини та систем технічного зору вигляді [1].

Незважаючи на велику кількість успішних застосувань математичної морфології, основу методів якої складають теоретико-множинні принципи: алгебра зображень, нелінійна суперпозиція та клас нелінійних морфологічних систем (НМС), застосування її гальмується відсутністю машинних архітектур для ефективної реалізації морфологічних перетворень.

В роботі розглядається архітектура оптоелектронного процесора обробки картичних операндів [2] та принципи реалізації двомірних обробляючих обчислювальних середовищ та структур з оптичними входами-виходами, які ефективно реалізують основні операції математичної морфології. Такі структури названо морфологічними операційними пристроями картичного типу (МОП КТ).

Двійкові сигнали та бінарні зображення можуть бути представлені за допомогою множин, наприклад булевих матриць A_i , а півтонові зображення B_i — за допомогою ансамблю його порогових, лінійно упорядкованих множин, кількість яких для зображення з квантовими 2^n рівнями інтенсивності відповідає числу рівнів

$$B_i = \max \{ p : (n, m) \in \Pi_p (B_i) \},$$

де $\Pi_p (B_i) = \{ n, m : B_i(n, m) \geq p \}$, причому p перекриває діапазон $B_i(n, m)$, а

$$\Pi_p (n, m) = \begin{cases} 1, & \text{якщо } B_i(n, m) \geq p, \\ 0, & \text{якщо } B_i(n, m) < p. \end{cases}$$

Таким чином, для введення та перетворення напівтонових зображень необхідно мати пороговий пристрій картичного типу або АЦП КТ [3], які незалежно від числа елементів, пікселів в зображені перетворюють вихідні двомірні сигнали в набір множин (булевих матриць Π_p), а для їх зберігання необхідна сторінкова пам'ять, яка може бути реалізована за допомогою оптоелектронного реєстра зсуву, описаного в [4].

Ряд морфологічних операторів, включаючи такі, як розмикання, замикання, дають змогу здійснювати нелінійну фільтрацію, багатомасштабне зглажування тощо. Вони можуть бути визначені через фундаментальні оператори приросту, ерозії, які традиційно визначаються через об'єднання та перетин множин вихідного та зсунутого, перенесеного відповідно з структуруючим елементом, зображень.

В зв'язку з цим для найпростіших структуруючих множин (більш складні можна представляти суперпозицією простих) необхідно, щоб обробляюче середовище могло здійснювати функції зсуву мінімум у чотирьох напрямках та двомірні функції «І», «АБО». Таким середовищем може бути оптоелектронний реєстр зсуву з паралельним оптичним входом-виходом [2].

При обробці цифрових багаторівневих зображень можна здійснювати морфологічні операції як за допомогою набору МОП КТ, що працюють з бінарними зображеннями, представленними пороговими множинами P_p ($p \in 0 + 2^n - 1$) за допомогою суперпозицій, так і з МОП КТ, що працюють в багатозначній, гібридній або неперервній (з неперервнологічними функціями) логіках. В останньому випадку операції приросту та ерозії зводяться до поелементних операцій мінімуму, максимуму, інверсії над вихідними напівтоновими матрицями та їх зсунутими копіями, причому параметри зсуву задаються структуруючою множиною.

Крім розглянутих фундаментальних операторів для розширення виконуваних перетворень, зокрема для виділення контурів та ліній, знайдення піків, долин, горбів, скелетизації тощо, обробляюче середовище повинне здійснювати операції обчислення не тільки різниці множин ($A_i \setminus A_j$) для визначення границь двійкового зображення, але й алгебраїчну поелементну різницю числових матриць при знаходженні градієнтів еrozії та приросту при обробці напівтонових зображень. Враховуючи сказане вище, було визначено набір базових операцій над дворівневими і багаторівневими двомірними сигналами та були сформульовані системні вимоги, які дозволили розробити універсальне середовище, на базі якого можуть бути побудовані практично всі МОП КТ. В основу функціонування такого середовища покладений принцип, який дозволяє обробляти як аналогові, так і цифрові (дво-та багаторівневі з різними засобами кодування) сигнали і здійснювати одночасно перетворюючі, порогові, нелінійні операції та операції різних логік. Структурна схема такого оптоелектронного середовища складається з однієї або декількох матриць, оптично керованих фазоімпульсних модуляторів з оптичними картичними входами запуску та виходами, керованих оптоелектронних затворів (ОЕЗ) та картичних елементів пам'яті, з'єднаних оптично, в тому числі за допомогою обернених оптичних картичних зв'язків. Вихідне зображення вводиться в картичні елементи пам'яті (КЕП) та фіксується там. Бінарні (багаторівневі) картичні з КЕП, впливаючи на ФІМ КТ, перетворюють рівні оптичні сигнали. Вибором послідовності керуючих сигналів, які подаються до ОЕЗ, можна легко перебудовувати середовище на виконання тої чи іншої функції [5].

В доповіді розглядаються конкретні варіанти реалізації МОП КТ, приводяться схеми базових елементів (ФІМ) матриць, результати експериментальних досліджень, які підтверджують можливість досягнення продуктивності 10^9 біт/с і частоти обробки кадрів 0,1 мГц.

Викладаються також нові математичні моделі основних операцій математичної морфології, зображені через операції лінійної алгебри та деякі допоміжні операції:

1) операція приросту для бінарних зображень:

$$X \Pi B = \bigcup_{b \in B} \left(h^{-by''} X h^{bx} + h^{by} X \cdot h^{bx} + h^{-by''} + X h^{-bx''} + h^{by} X \cdot h^{-bx''} \right) = \\ = \lambda_p \left\{ \sum h^* \cdot X + \sum X \cdot h^{**} \right\}^2,$$

де U — об'єднання, λ_p — порогова функція від числової матриці;

2) операція еrozії для бінарних зображень:

$$X \epsilon B = \lambda_p = \sum \langle 1 \rangle \text{ в } B \left\{ \sum (h^* X + X h^{**}) \right\};$$

3) для напівтонових приrostів:

$$\tilde{f} \Pi B = \max_{y \in B} \{ \tilde{f} \cdot h^y \},$$

де \tilde{f} — чисрова матриця, $\tilde{f} \cdot h^y$ — зсувні матриці.

Запропоновані оператори здійснення нелінійної крос-кореляції зображень є зручними для реалізації на запропонованих МОП КТ. Обчислення такої нелінійної двомірної функції над зображенням розмірністю $N \times N$ та еталоном розмірністю $M \times M$ ($M < N$) здійснюється за M^2 тактів, які включають порогову обробку вихідного зображення, зсув та накопичення, причому відліки вихідної функції розмірністю $(N - M + 1)$ формуються паралельно. Такий морфологічний корелятор півтонових зображень працює в реальному масштабі часу, а його реалізація не потребує великих витрат.

Література

1. Марагос П. Морфологические системы для многомерной обработки сигналов // ТИИЕР. — 1990. — № 4. — С. 109–131.
2. Дубчак В.Н., Бойко Р.В., Красиленко В.Г. Разработка и применение оптоэлектронных двумерных бистабильных структур // Вестн Академии наук БССР, Сер. физ.-мат. наук. — 1990. — № 6. — С. 69–71.
3. А. с. 1702129 (СССР). Способ записи и хранения изображений и устройство для его осуществления / В.Г. Красиленко и др. // БИ. — 1991. — № 48.
4. А. с. 1476503 (СССР). Оптоэлектронное устройство для логической обработки информации / В.Г. Красиленко и др. // Там же. — 1989. — № 16.
5. Дубчак В.Н., Колесницкая О.К., Красиленко В.Г. Принципы создания и схемотехнические вопросы построения многоканальных устройств и систем для параллельной обработки и анализа изображений // Распознавание образов и анализ изображений : Тез. докл. Всесоюз. конф. — Минск, 1991. — Ч. 3. — С. 83–87.



Ефективний засіб розпізнавання зображень

Людмила Ліщинська

Вінницький політехнічний інститут

Україна, 286021 Вінниця 21
Хмельницьке шосе, 95
Tel.: (043) 223-93-80

Конструювання найбільш інформативних ознак — одна з основних задач при розв'язанні питань теорії та практики розпізнавання образів [1, 2]. Пов'язано це з тим, що отримана в результаті конструювання система оптимальних ознак буде визначати не тільки алгоритм розпізнавання, але і структуру побудови високопродуктивних розпізнаючих пристрій (спеціалізованих процесорів).

Ефективною системою математичного опису при розпізнаванні образів об'єктів є система моментних ознак зображення. Інтегральний опис за допомогою моментів $m_{\alpha\beta}$ ($\alpha, \beta = 0, 1, 2, \dots$) та можливість його викристання обумовлені в [1].

У доповіді розглядається підхід, який дає можливість визначити в реальному масштабі часу моментні ознаки зображення як молодших показників степеня, так і старших. Цей підхід зводиться до операції перемноження матриць: вхідної та специфічних (керуючих), причому можливо підібрати такий набір керуючих матриць, що з'являється можливість розв'язання широкого класу задач.

У загальному випадку для моментних ознак:

$$m_{\alpha\beta} = || \Phi^\beta B \Phi^\alpha ||,$$

де $B(x, y)$ — двомірна функція, що описує вхідне зображення, Φ — матриця перетворення, α, β — показники степеня моментної ознаки, $||A||$ — евклідова норма матриці A .

Сьогодні, коли більш перспективним є створення оптоелектронних цифрових процесорів обробки зображень, що передбачають картичне (паралельне) введення інформації, доцільно будувати спецобчислювач на основі матричного перемно-

ження, що виконує операції над вхідними та керуючими операндами (матрицями), котрі задають вид та функцію перетворення, матричними операндами, зміна яких у відповідності з алгоритмом розв'язання задачі дозволить досягнути більшої універсалізації.

Література

1. Анисимов Б.В., Курганов В.Д., Злобин В.К. Распознавание и цифровая обработка изображений. — М. : Выш. шк., 1983. — 295 с.
2. Горелик А.А., Скрипкин В.А. Методы распознавания. — М. : Выш. шк., 1989. — 232 с.



Двовимірні графові граматики для розпізнавання графічних зображень

В'ячеслав Мацелло

Інститут кібернетики АН України

Україна, 252207 Київ 207,
просп. Акад. Глушкова, 40
Тел.: (044) 266-25-69

Ця робота є прикладом застосування ідеї двовимірних граматик [1] для розпізнавання одного класу графічних зображень.

За допомогою різноманітних методів попередньої обробки графічне зображення можна задати у вигляді сукупності відрізків прямих ліній. Цілком зрозуміло, що задане таким чином зображення фактично є плоским графом. Причому кінці відрізків, з яких сформовано зображення, є вершинами графа, а самі відрізки складають множину дуг графа. Такий граф задає симетричне бінарне відношення сусідства на множині вершин $R \subset T \times T$. Графічне зображення T_R є множиною точок на площині з заданим на цій множині відношенням сусідства.

Характерною ознакою графа T_R є те, що його вершини мають фіксовані координати на площині, а дуги — фіксовані нахили (напрямки). На напрямки дуг накладають обмеження у вигляді множини I допустимих напрямків. Якщо визначена множина з N допустимих напрямків, тоді кожній вершині $t \in T$ можна поставити у відповідність N -роздільний двійковий код. Причому j -й розряд цього коду дорівнює 1, якщо з цієї вершини виходить дуга у напрямку, який має номер j , і дорівнює 0 у протилежному випадку. Множину допустимих кодів w назначимо W .

Зображення T_R можна перетворити в деякий інший граф за такими правилами. Вершинами u нового графа є вершини графа T_R , які помічені відповідними кодами w , тобто $u = (t, w)$. Кожну дугу неорієнтованого графа T_R міняємо на пару орієнтованих дуг протилежних напрямків. Кожну орієнтовану дугу помітимо трійкою, яка має вигляд (u', I_K, u'') . Така відмітка означає, що у графі T_R є пара вершин $(t', t'') \in R$, які мають коди відповідно w' та w'' , а з'єднуюча їх дуга має напрямок I_K . Наявність у дуги відмітки у вигляді (u', I_K, u'') можна розглядати, як ознаку виконання певного відношення структурного сусідства P , яке задане на множинах U та I і має вигляд $P \subset U \times I \times U$. Перетвореному таким чином графу дамо назву структурного графа і позначимо T_P . Структурний граф T_P задає інформацію про структурно-топологічні властивості зображення T_R у вигляді, зручному для подальшого синтаксичного аналізу.

Для структурних графів, які відповідають зображенням одного класу, задамо множину S — алфавіт структурних елементів. З елементів множини S будується

опис графа T_P . Описом структурного графа T_P називається відображення $D : U \rightarrow S$ множини структурних вершин графа T_P на алфавіт структурних елементів.

Для структурних елементів, які формують описи графів T_P заданого класу, задаються обмеження у вигляді двох множин.

1. Множина $Z_W \subset W \times S$ пар, які мають вигляд (w, s) . Ця множина для кожного типу структурної вершини $w^* \in W$ задає набір структурних елементів $s \in S$, допустимих для опису структурної вершини цього типу.

2. Множина $Z_S \subset S \times I \times S$ трійок, які мають вигляд (s, I_k, s') . Ця множина задає відношення структурного сусідства на елементах алфавіту S . При цьому трійка (s', I_k, s'') може увійти до складу множини Z_S тільки у тому разі, якщо на структурному графі цього класу може знаходитись пара вершин u' та u'' , які мають коди відповідно w' та w'' , для яких виконуються умови

$$(u', I_k, u'') \in P, (w', s') \in Z_W, (w'', s'') \in Z_W.$$

Двовимірною графовою граматикою звуться п'ятірка введених вище множин

$$G = \langle I, W, S, Z_W, Z_S \rangle.$$

Зображення T_R називається допустимим у граматиці G , якщо

- 1) Будь-яка дуга зображення T_R має напрямок, допустимий у множині I ;
- 2) Існує допустимий опис відповідного графа T_P . Опис D є допустимим, якщо:

a) кожній вершині графа T_P , яка має тип w^* , поставлений у відповідність єдиний структурний елемент $s^* \in S$, такий, що $(w^*, s^*) \in Z_W$;

b) для будь-якої пари структурних елементів $s' \in S$ та $s'' \in S$, які відповідають парі сусідніх структурних вершин u' та u'' , таких, що $(u', I_k, u'') \in P$, виконується $(s', I_k, s'') \in Z_S$.

Синтаксичний аналіз зображення T_R за допомогою двовимірних графових граматик полягає у пошуку допустимого опису D для відповідного структурного графа T_P . Розроблено швидкодіючий алгоритм синтаксичного аналізу, який дозволяє відповісти на питання, чи існує для зображення T_R допустимий опис відповідного структурного графа T_P з точки зору граматики G , і в разі позитивної відповіді отримати цей опис.

Алгоритм складається з двох етапів. На першому етапі для кожної вершини u структурного графа T_P , яка має тип w^* , створюється список структурних елементів s^* , таких, що $(w^*, s^*) \in Z_W$. На другому етапі для сусідніх вершин багаторазово виконується процедура усунення зі списків тих структурних елементів, для яких не виконується відношення структурного сусідства Z_S . Якщо на черговій ітерації який-небудь список стає пустим, це показує, що зображення T_R не є допустимим з точки зору граматики G . Якщо після виконання чергової ітерації всі списки залишилися незмінними, а також у кожному списку залишилося рівно по одному структурному елементові, це є ознакою того, що допустимий опис існує і структурні елементи, що лишилися у списках, складають цей опис.

Доведено, що обчислювальна складність алгоритму пропорційна квадрату кількості вершин структурного графа T_P , що набагато ефективніше, ніж у відомих алгоритмів синтаксичного аналізу для традиційних формальних граматик.

Існують дві можливості організації технології розпізнавання за допомогою двовимірних графових граматик. Першу з них доцільно використовувати в тому разі, коли необхідно виконати класифікацію окремих, не дуже складних об'єктів. В такому разі для кожного з K класів зображень будеться своя двовимірна

графова граматика G_i , $i = 1 + K$. Зображення перевіряється на допустимість у кожній з граматик і відноситься до i -го класу, якщо воно визнане допустимим у граматиці G_i .

Інший спосіб використовується, якщо потрібно проаналізувати складне зображення. В цьому випадку будеся одна двовимірна графова граматика, яка описує всі можливі зображення, які підлягають аналізу. Головна ціль синтаксичного аналізу — пошук допустимого опису D . Далі на базі отриманого опису можна за допомогою простих логічних функцій виконати аналіз структури зображення, а також змістовну інтерпретацію його складових частин і зображення в цілому.

Література

1. Шлезингер М.И. Математические средства обработки изображений. — Киев : Наук. думка, 1989. — 198 с.

