

Знаходження об'єктів на знімках міських кварталів

Ткачук В.Л.

Міжнародний науково-навчальний центр
інформаційних технологій та систем, Київ, Україна

Анотація

Розглянуто дві практичні задачі сегментації: виділення зображення фасадів окремих будинків на фотознімках і виділення зображень автомобілів на фоні будинків. Для цих задач запропоновано їх формальну постановку та розв'язок. Зображення і його сегментація розглядаються як випадкові величини, що мають спільний розподіл. Шуканою є сегментація з найбільшою імовірністю при заданому зображенні. Розглянуто алгоритми пошуку оптимальної сегментації для цих практичних задач.

1. Вступ

В роботі досліджується проблема виділення окремих об'єктів на зображенні. Для розв'язання цієї задачі використовуються алгоритми і методи структурної сегментації зображень: поле зору розбивається на окремі сегменти, кожен з яких відповідає об'єкту реального світу. В кожному сегменті кольори пікселів пов'язані певними статистичними закономірностями. Проте сегмент визначається не тільки кольорами своїх пікселів, а й такими характеристиками, як зв'язність і форма. Це накладає додаткові умови на множину допустимих сегментацій.

Для кожного конкретного зображення є множина допустимих сегментацій, серед якої треба вибрати одну найкращу. В роботі зображення і сегментації розглядаються як реалізації деяких випадкових величин, а критерієм якості сегментації є її умовна імовірність при заданому зображенні.

Розглянуто дві практичні задачі: виділення зображень фасадів окремих будинків на фотознімках і виділення зображень автомобілів на фоні будинків. Представлено зведення цих задач до задачі сегментації і наведені алгоритми пошуку оптимальної сегментації.

2. Означення і постановка задачі

Поле зору T назвемо прямокутну підмножину

двовимірної цілочисельної ґратки — множину координат пікселів зображення:

$$T = \{(i, j) \mid i = \overline{0, W-1}, j = \overline{0, H-1}\}.$$

Величину W називатимемо шириною, а H — висотою поля зору. Елементи $t \in T$ назвемо пікселами. Будемо називати піксели t, t' сусідніми, якщо одна з їх координат однакова для обох пікселів, а друга відрізняється не більше ніж на 1.

Палітрою X назвемо довільну скінченну множину. Елементи палітри X назвемо кольорами. Для практичних задач будемо використовувати палітру L^*a*b , описану в [1]. В моделі L^*a*b колір представлено вектором $c \in X = \{L, a, b\}; L, a, b \in \mathbb{R}$. Ми будемо виконувати операції додавання і віднімання кольорів, а також множення їх на скаляр. Відстань $d(\cdot, \cdot)$ між кольорами визначена як евклідова відстань між векторами в \mathbb{R}^3 :

Функцію $x_T : T \rightarrow X$ називатимемо зображенням. Значення зображення в пікселі $t \in T$ позначимо через x_t .

Сегментом s будемо називати ім'я довільного об'єкту, а множиною сегментів S — множину імен.

Сегментацією поля зору T на сегменти з S будемо називати функцію $f_T : T \rightarrow S$. Причому, $f_T \in F_T$, де F_T — клас допустимих сегментацій. Значення сегментації в пікселі $t \in T$ зображення позначимо через f_t .

Назвемо піксел $t \in T$ граничним в сегментації f_T , якщо хоча б один з сусідніх з ним пікселів належить іншому сегменту. В такому випадку піксел належить границі свого сегменту: $t \in \partial f_t$. Множину всіх граничних пікселів позначимо через ∂S .

2.1. Модель зображення як випадкової величини

Будемо вважати, що при заданому полі зору існує імовірність кожного зображення $P(x_T)$, кожної сегментації $P(f_T)$ і їх спільна імовірність $P(x_T, f_T)$. В практичних ситуаціях ці розподіли

не відомі, а експериментальних даних недостатньо для побудови достовірної статистики (зважаючи на велику кількість можливих зображень і сегментацій), тому в кожному практичному випадку ми будемо робити припущення щодо деяких розподілів. Крім того, ми можемо *розумним* чином визначити умовну імовірність $P(x_T|f_T)$ зображення x_T при фіксованій сегментації f_T . Вважаємо, що колір і сегментація кожного пікселя зображення не залежать від інших пікселів.

2.2. Знаходження оптимальної сегментації як задача знаходження оцінки максимальної правдоподібності

Метою розпізнавання є знаходження сегментації, що має найбільшу імовірність при заданому зображенні:

$$f_T^* = \arg \max_{f_T \in F_T} P(f_T|x_T). \quad (1)$$

Враховавши те, що $P(x_T)$ — константа, за формулою Байєса отримуємо:

$$f_T^* = \arg \max_{f_T \in F_T} P(x_T|f_T)P(f_T). \quad (2)$$

3. Визначення умовної імовірності зображення

Для визначення ймовірності зображення при заданій сегментації $P(x_T|f_T)$ будемо враховувати ймовірності $P^*(x_T|f_T)$ *породження* пікселів певного кольору сегментами, до яких вони належать, а також ймовірність $P(\partial S(f_T))$ границь сегментів.

$$P(x_T|f_T) = P^*(x_T|f_T) \cdot P(\partial S(f_T)) \quad (3)$$

3.1. Ймовірність кольору пікселя зображення

Можливі різні варіанти визначення імовірності породження зображення заданою сегментацією. Враховуючи незалежність сегментації окремих пікселів, достатньо задати імовірність кольору окремого пікселя при фіксованій сегментації — $P^*(x_t|f_T)$ для кожного пікселя $t \in T$.

В якості ймовірності пікселя візьмемо середні значення частот, з якими в сегменті зустрічаються кольори, *близькі* до кольору пікселя:

$$\forall t \in T : P^*(x_t|f_T) = \frac{\tilde{N}_{f_t}^{x_t}}{N_{f_t}}, \quad (4)$$

де

$$\tilde{N}_{f_t}^{x_t} = \frac{\int_{c \in \hat{O}_a(x_t)} N_{f_t}^c}{m}, \quad (5)$$

N_s — кількість пікселів віднесених до сегменту s , N_s^c — кількість пікселів кольору c віднесених до сегмента s , $\hat{O}_a(x_t) = \{c \in X : d(c, x_t) \leq a\}$ — куля радіусом a і з центром в x_t , a — фіксований параметр, m — об'єм кулі.

3.2. Імовірність границі сегментів

Для пікселів, що лежать на границях сегментів, розглянемо імовірність того, що вони є граничними. Виходячи з припущення, що граничні пікселі сусідніх сегментів мають якнайбільше відрізнятися за кольором, ймовірність того, що піксел є граничним, запишемо так:

$$P(t \in \partial S) = \frac{1}{C} |\nabla x_t|, \quad (6)$$

$|\nabla x_t|$ — модуль значення кольорового градієнту в пікселі t , C — константа нормування: $C = \max_{t \in T} |\nabla x_t|$. Точне визначення, а також методи обчислення кольорового градієнта описані в [2]. На практиці користувалися наближеним значенням:

$$\nabla x_t = \sqrt{X_t^2 + Y_t^2}, \quad (7)$$

де

$$X_{(i,j)} = d(x_{(i+1,j)}, x_{(i-1,j)}), \quad (8)$$

$$Y_{(i,j)} = d(x_{(i,j+1)}, x_{(i,j-1)}). \quad (9)$$

Враховавши ймовірність границі, з 3 отримуємо наступну формулу:

$$\begin{aligned} P(x_T|f_T) &= P^*(x_T|f_T) \cdot \prod_{t \in T, t \in \partial f_t} P(t \in \partial S) = \\ &= \prod_{t \in T} P^*(x_t|f_T) \cdot \prod_{t \in T, t \in \partial f_t} P(t \in \partial S), \end{aligned} \quad (10)$$

де $P^*(x_t|f_T)$ ймовірність пікселя зображення визначена вище за 4.

4. Знаходження фасаду будинку на зображенні

Задача полягає в наступному: є знімок на якому фронтально знаходиться фасад будинку, необхідно визначити координати його лівої і правої границь.

Зведемо задачу до формальної моделі. Множина сегментів S складається з трьох сегментів: *лівого* — все, що на знімку зліва від фасаду; *правого* — все, що справа від фасаду; *центрального* — фасад будинку. До класу допустимих сегментацій належать всі функції виду:

$$f_{(i,j)}^{L,R} = \begin{cases} \text{„ліво“}, & \text{якщо } i < L \\ \text{„право“}, & \text{якщо } i \geq R \\ \text{„центр“}, & \text{якщо } L \leq i < R \end{cases}$$

L — ліва границя зображення фасаду, R — права границя.

$$0 \leq L \leq R \leq W.$$



Рис. 1: Приклад сегментованого зображення: зеленим кольором відмічений лівий сегмент, синім – центральний, а червоним – правий.

Пропонується наступний розв'язок задачі:

1. Припустимо що всі апіорні імовірності сегментацій $P(f_T)$ однакові, тоді оптимальна сегментація знаходиться за формулою:

$$f_T^{*,L,R} = \arg \max_{L,R} P(x_T | f_T^{L,R}). \quad (11)$$

2. Для кожної сегментації обчислити її імовірність за формулою 10.
3. В якості оптимальної вибрати сегментацію, для якої обчислена імовірність максимальна.

Повний перебір сегментацій можливий, беручи до уваги невелику кількість допустимих сегментацій.

$$|F_T| = \frac{(W-1) \cdot (W-2)}{2}.$$

5. Знаходження зображення автомобіля на фоні будинку

Розглянемо випадок, коли знімок містить тільки одне зображення автомобіля. Множина сегментів складається з сегменту, що відповідає за зображення фасаду, і сегменту, що відповідає за зображення автомобіля:

$$S = \{,будинок“, „машина“\}.$$

Визначимо клас допустимих сегментацій F_T . Зображення автомобіля є зв'язною областю,

що завжди знаходиться нижче зображення фасаду, і при перегляді зображення зліва-направо висота контуру машини до певного моменту не зменшується, а потім не збільшується, висота контуру змінюється *плавно*.



Рис. 2: Приклад сегментованого зображення: зеленим кольором відмічено будинок, а червоним – машину.

Кількість допустимих сегментацій є занадто великою для застосування алгоритму повного перебору. Ми не можемо перебрати всі сегментації, але якщо нам відомі апіорні імовірності сегментацій $P(f_T)$ і умовні імовірності зображень при заданій сегментації $P(x_T | f_T)$, то можна знайти оптимальну сегментацію f_T^* більш швидким способом, скористувавшись методом динамічного програмування. Враховуючи незалежність сегментації пікселів, нам достатньо знати лише апіорну імовірність сегментації для кожного пікселя $P(f_t), \forall t \in T$, а також умовну імовірність породження кожним сегментом пікселя заданого кольору $P(c|s), \forall c \in X, s \in S$.

Апіорний розподіл сегментацій $P(f_t)$ – невідомий. Припущення, що всі $P(f_t) = const$ не приведе до прийнятного результату, оскільки в такому випадку $\forall c \in X : P(c, „машина“) = P(c, „будинок“)$, і всі сегментації мають однакову імовірність. Отже необхідно вибрати $P(f_t)$ якимось іншим чином. Будемо вважати істинним такий набір апіорних імовірностей сегментацій $P(f_t)$, що максимізує імовірність зображення x_T :

$$\{P(f_t)\}^* = \arg \max_{\{P(f_t)\}} P(x_T), \quad (12)$$

Імовірність зображення $P(x_T)$ і апіорні імовірності сегментацій $P(f_T)$ зв'язані через формулу:

$$P(x_T) = \sum_{f_T \in F_T} P(x_T, f_T) = \sum_{f_T \in F_T} P(x_T | f_T) P(f_T).$$

Для знаходження шуканого розподілу 12 скористаємося відомим алгоритмом самонавчання, що описаний в [3]. Алгоритм потребує початкових умов — початкового наближення для апіорних імовірностей

сегментацій в кожному пікселі $p^0(f_t)$. В роботі досліджено декілька варіантів вибору цих величин, проте не знайдено алгоритмів для їх побудови без допомоги оператора.

Весь алгоритм розв'язання має наступну схему:

1. Вибрати початковий розподіл $p^0(f_t)$.
2. Ітеративно виконувати наступні кроки (алгоритм самонавчання, кількість ітерацій є параметром):
 - (a) за відомими апріорними імовірностями сегментацій $p^i(f_t)$ знаходимо імовірності породження кожним сегментом піксела заданого кольору $p^i(c|s)$, $c \in X$, $s \in S$,
 - (b) знаючи імовірності $p^i(c|s)$ знаходимо новий розподіл $p^{i+1}(f_t)$.
3. Вважаючи, що нам відомий точні імовірності сегментацій $P(f_T)$ і умовні імовірності зображення при заданих сегментаціях $P(x_T|f_T)$, знаходимо оптимальну сегментацію.

6. Результати експериментів

Результати експериментів представлені на рисунках 3 (приклад виділення фасаду окремого будинку) і 4 (приклад виділення зображення автомобіля).



Рис. 3: Сірі області на зображенні виникли в результаті попередньої обробки фотознімка, піксели з цих областей не враховуються при роботі алгоритму розпізнавання. Коефіцієнт згладження гістограми кольорів в $4 a = 3$.



Рис. 4: Приклад вдалого виділення автомобіля. Кількість неправильно відсегментованих пікселів $\approx 2\%$.

7. Висновки

Основні результати роботи:

- побудовано модель зображення і його сегментації. Об'єкти в даній моделі розглядаються як сегменти, що відповідають найімовірнішій сегментації. Приведені способи визначення імовірностей для моделі;
- для практичної задачі виділення фасадів будинків на зображеннях, представлено зведення її до задачі сегментації і запропоновано алгоритм її розв'язання;
- для практичної задачі виділення автомобілів на фоні фасадів будинків, представлено зведення її до задачі сегментації і запропоновано алгоритм її розв'язання;
- проведено експериментальну перевірку розроблених алгоритмів.

8. Література

- [1] Fairchild M. D. Color and Image Appearance Models. John Wiley and Sons, 2005. P. 408.
- [2] Gonzales R., Woods R. Digital Image Processing. Prentice Hall, 2008. P. 954.
- [3] М. Шлезингер, В. Главач. Десять лекцій по статистическому и структурному распознаванию.— Киев: Наукова думка, 2004.— С. 545.