

Знаходження точок сходу на зображеннях міських кварталів

Віктор Ю. Сдобніков, Богдан Д. Савчинський

Міжнародний науково-навчальний центр
інформаційних технологій та систем, Київ
viktor.fti@gmail.com, bogdan.savchynsky@gmail.com

Анотація

Робота присвячена обчисленню положення точок перетину проєкцій паралельних прямих (точок сходу) на площині зображення, отриманого за допомогою фотокамери. Знання точок сходу дозволяє оцінити поворот та внутрішні параметри камери, позбутись перспективних спотворень тощо, тому їх знаходження є важливим у комп'ютерному баченні.

В роботі запропоновано постановку задачі обчислення точок сходу як задачі самонавчання та наведено ефективний алгоритм її розв'язання.

Експериментальна перевірка запропонованого методу здійснювалась на зображеннях житлових кварталів, що використовуються в інтерактивній картографічній інтернет-системі www.Mappy.com.

Abstract

This paper is dedicated to the estimation of the parallel lines projections intersections (vanishing points) on the image's plane, obtained with the projective camera. Vanishing points allow the estimation of camera parameters, rectifying images and others features, so their finding is important problem in computer vision.

The task of vanishing points estimation is formalized as an unsupervised learning problem and an effective algorithm of it's solving is propounded in current paper.

Experimental testing of a propounded method was implemented over the images of town quarters which are used in interactive cartographic system www.Mappy.com.

1. Вступ

Прямі, паралельні у тривимірному просторі, взагалі кажучи, перестають бути паралельними при центральному їх проєктуванні на площину зображення (як це відбувається у традиційних фотокамерах), і перетинаються в одній точці. Її прийнято називати *точкою сходу* (vanishing point). Приклади точок сходу наведено на Рис. 2, 3.

Точки сходу, що відповідають відомим напрямкам в тривимірному просторі, дозволяють оцінити поворот та внутрішні параметри камери [1], ректифікувати зображення, позбуваючись проєктивних спотворень [1], та вирішувати інші прикладні задачі машинного бачення.

Через свою важливість задача знаходження точок сходу вже давно досліджується в машинному баченні. Існує значна кількість підходів до її розв'язання [2, 3, 4, 5, 6, 7, 8].

В роботі запропоновано нову формальну постановку і розв'язок задачі знаходження точок сходу як задачі навчання без учителя. Вхідними даними для алгоритму самонавчання [9, гл. 6] є вибірка прямих, виділених на зображенні за допомогою перетворення Хафа [10]. Задача полягає у розділенні представленої вибірки на класи, кожен з яких відповідає підмножині паралельних у тривимірному просторі прямих, та оцінці координат точок перетину (точок сходу) прямих, що входять до кожного з класів.

Другий розділ присвячено постановці задачі та її розв'язанню, в третьому представлено експериментальні результати.

2. Постановка задачі знаходження точок сходу як задачі самонавчання

Перший етап розв'язання задачі полягає у виділенні певної множини прямих на вхідному зображенні. Для цього було використано перетворення Хафа [10]. Перетворення Хафа дозволяє знаходити параметрично задані криві на зображенні, зокрема прямі.

Пряма

$$x \cdot \cos \theta + y \cdot \sin \theta - \rho = 0 \quad (1)$$

на площині визначається кутом нахилу відносно осі ординат θ та відстанню від початку координат до прямої ρ .

Результатом перетворення Хафа є множина точок у просторі параметрів (θ, ρ) . Кожній точці приписано значення її інтенсивності, яке пропор-

ційне кількості точок на зображенні, що належать відповідній прямій. Для оцінки положення точок сходу використовуються лише певна наперед задана кількість точок з найбільшими значеннями інтенсивності. Інтенсивність приймає цілі значення і для спрощення запису формул ми замінимо кожну точку інтенсивністю I на I точок одиначної інтенсивності.

Вважатимемо, що кожна пряма на зображенні належить до одного з m класів. Множину класів позначимо $K = k_1, k_2, \dots, k_m$. Кожен з класів $k_i, i = 1, \dots, m-1$, містить прямі, що відповідають одній, певній точці сходу. До класу k_m належать всі інші прямі.

Якби розбиття прямих на класи було відоме, то знаходження точок сходу, що відповідають кожному з класів, могло би бути здійснене за допомогою методу найменших квадратів [5]. Формальну постановку задачі розбиття прямих на класи і оцінки точок сходу за ними наведено у наступному підрозділі.

2.1. Постановка задачі розбиття прямих на класи та оцінки точок сходу

Нехай Ω — множина всіх можливих прямих на вхідному зображенні (точок в просторі перетворення Хафа), а K — множина класів прямих. На декартовому добутку $\Omega \times K$ вважатимемо визначеним з точністю до певних параметрів розподіл ймовірностей $p_{\Omega K}$. Величина

$p_{\Omega K}(\omega, k; a_k) = p_K(k) \cdot p_{\Omega|k}(\omega; a_k), \omega \in \Omega, k \in K$ визначає спільну ймовірність прямої ω і її належності до класу k в залежності від значення параметра a_k . Величина $p_K(k)$ визначає апіорну ймовірність класу k , а $p_{\Omega|k}(\omega; a_k)$ — ймовірність прямої ω за умови її належності до класу k . Нехай $\omega_1, \omega_2, \dots, \omega_n \in \Omega$ — множина виділених на зображенні прямих. Ймовірність цієї сукупності дорівнює:

$$l(p_K(k), a_k | k \in K) = \prod_{i=1}^n \sum_{k \in K} p_K(k) \cdot p_{\Omega|k}(\omega_i; a_k). \quad (2)$$

Сукупність величин $M = (p_K(k), a_k | k \in K)$ називатимемо *моделью*.

Як буде показано далі, задача розбиття сукупності $\omega_1, \omega_2, \dots, \omega_n$ виділених на зображенні прямих на класи та оцінка відповідних точок сходу полягає в пошуку такої моделі M^* , що максимізує ймовірність цієї сукупності прямих (2).

$$M^* = \arg \max_M \left(\prod_{i=1}^n \sum_{k \in K} p_K(k) \cdot p_{\Omega|k}(\omega_i; a_k) \right) \quad (3)$$

Вкажемо тепер, який вигляд має розподіл $p_{\Omega|k}(\omega; a_k)$.

Розглянемо довільну сукупність паралельних прямих в тривимірному просторі. У випадку, коли площина проектування камери не паралельна їм, проекції прямих на цю площину перетинаються в одній точці — точці сходу цих прямих. Внаслідок впливу відомих факторів (неточності виділення прямих на зображенні, викривлень, зумовлених оптичною системою фотокамери тощо) ці прямі не перетинаються в одній точці. Ми шукаємо точку, сума квадратів відстаней від якої до всіх прямих сукупності мінімальна. Відхилення (відстань, взята зі знаком) точки (x_0, y_0) від прямої, що задається параметрами θ та ρ дорівнює $x_0 \cdot \cos \theta + y_0 \cdot \sin \theta - \rho$. Вважатимемо розподіл відхилення прямих сукупності від точки сходу гаусівським. Таким чином, для класу k , що відповідає точці сходу (x_k, y_k) , розподіл ймовірностей $p_{\Omega|k}(\omega; a_k)$ набуває вигляду:

$$p_{\Omega|k}(\omega; a_k) = \frac{1}{\sqrt{2\pi}\sigma_k} \cdot e^{-\frac{(x_k \cos \theta + y_k \sin \theta - \rho)^2}{2\sigma_k^2}} \quad (4)$$

де σ_k — дисперсія гаусівського розподілу, а $a_k = (x_k, y_k, \sigma_k)$.

У випадку, коли площина проектування камери паралельна сукупності паралельних у просторі прямих, їх проекції на неї також паралельні, а, значить, задаються лише одним параметром — кутом нахилу θ . Міркування, аналогічні щойно наведеному, приводять до гаусівського розподілу з математичним сподіванням θ_k кута нахилу прямих k -того класу:

$$p_{\Omega|k}(\omega; a_k) = \frac{1}{\sqrt{2\pi}\sigma_k} \cdot e^{-\frac{(\theta - \theta_k)^2}{2\sigma_k^2}} \quad (5)$$

де σ_k — дисперсія гаусівського розподілу, а $a_k = (\theta_k, \sigma_k)$.

У цьому випадку ми казатимемо, що відповідна точка сходу знаходиться на нескінченності.

Окрім двох розглянутих класів паралельних в просторі прямих ми розглядатимемо також окремий клас всіх інших прямих, взагалі кажучи не паралельних в тривимірному просторі. До нього будемо відносити всі ті прямі, які не попали в перші $m-1$ класів. Розподіл цих прямих вважатимемо рівномірним на прямокутнику

$$\{(\theta, \rho) \mid \theta_{min} < \theta < \theta_{max}, \rho_{min} < \rho < \rho_{max}\}$$

$$p_{\Omega|k}(\omega) = \frac{1}{(\theta_{max} - \theta_{min}) \cdot (\rho_{max} - \rho_{min})} \quad (6)$$

Цей розподіл заданий повністю, і не залежить від параметрів.

При постановці задачі пошуку моделі, що максимізує ймовірність (2) сукупності $\omega_1, \omega_2, \dots, \omega_n$ виділених на зображенні прямих вважатимемо відомими кількість класів, що відповідають скінченим (розподіл вигляду (4)) та нескінченим (розподіл (5)) точкам сходу. Окремий клас прямих непаралельних у тривимірному просторі (розподіл (6)) завжди один.

Знаходження моделі $(p_K(k), a_k | k \in K)$ означає в тому числі і знаходження точок сходу, які містяться в параметрах $a_k | k \in K$.

2.2. Розв'язок задачі розбиття прямих на класи та оцінки точок сходу

Для розв'язання задачі (3) було використано алгоритм самонавчання [9, гл. 6], відомий також як EM алгоритм [11].

Цей алгоритм є ітеративним. Він починає роботу з певної моделі $M^0 = (p_K^0(k), a_k^0 | k \in K)$ і на кожному кроці t будує нову модель M^t та для кожної прямої ω_i з навчальної сукупності оцінює апостеріорні ймовірності $\alpha^t(i, k)$ належності її до кожного з класів $k \in K$ за формулами:

$$\alpha^t(i, k) = \frac{p_K^t(k) \cdot p_{\Omega|k}(\omega_i; a_k^t)}{\sum_{k' \in K} p_K^t(k') \cdot p_{\Omega|k'}(\omega_i; a_{k'}^t)} \quad (7)$$

$$p_K^{t+1}(k) = \frac{\sum_{i=1}^n \alpha^t(i, k)}{\sum_{k' \in K} \sum_{i=1}^n \alpha^t(i, k')} = \frac{\sum_{i=1}^n \alpha^t(i, k)}{n} \quad (8)$$

$$a_k^{t+1} = \arg \max_a \sum_{i=1}^n \alpha^t(i, k) \cdot \log p_{\Omega|k}(\omega_i; a), \quad k \in K. \quad (9)$$

Величини $\alpha^t(i, k)$ та $p_K^{t+1}(k)$ обчислюються безпосередньо за формулами (7) та (8). Обчислення формули (9) здійснюється за методом найменших квадратів.

Про вказаний алгоритм відомо (див. [9, гл. 6]), що: (i) він монотонно зменшує (не збільшує) значення функції правдоподібності $l(p_K^t(k), a_k^t | k \in K)$ (див. (2)) вхідної сукупності прямих; (ii) збігається в локальний мінімум, який залежить від початкової моделі M^0 . Експериментальні результати використання алгоритму та задання початкової моделі M^0 обговорюються в наступному розділі.

3. Експериментальні результати

Запропонований метод було перевірено на тестовій вибірці з 1005 зображень міських кварталів. Приклади таких зображень та знайдені точки сходу наведено на Рис. 1, 2, 3.

Зображення з тестової вибірки можна поділити на два типи. На зображеннях першого типу міститься лише один фасад (одна площина). Йому відповідають дві точки сходу: „вертикальна“, в якій перетинаються лінії, перпендикулярні до поверхні землі, та „горизонтальна“, в якій перетинаються лінії, паралельні земній поверхні та площині фасаду. Зображення другого типу містить кілька фасадів, розташованих в різних площинах. Кожній з площин відповідає своя „горизонтальна“ точка сходу, вертикальна точка сходу для всіх площин одна і та ж. Таким чином кількість точок сходу більша за кількість фасадів на зображенні на одиницю.

Всі тестові зображення пов'язані з базою даних, де вказано тип зображення. Окрім цього, для зображень першого типу в базі даних вказано розташування точки зйомки відносно площини фасаду, а саме вказано зліва, зправа чи по центру відносно фасаду знаходилася камера. Для лівого розташування камери горизонтальна точка сходу є скінченою і знаходиться справа (відповідний розподіл описується формулою (4)), як на Рис. 2, для правого — зліва, при центральному розташуванні (Рис. 1) камери точка сходу є нескінченою (розподіл описується формулою (5)). Зображення другого типу, як правило, відповідають кутовим будинкам, яким відповідають три точки сходу: зліва, справа та вгорі (див. Рис. 3). Кількість, тип та початкове положення точок сходу обиралися згідно з цією апріорною інформацією.

На зображеннях першого типу всі точки сходу знаходилися правильно приблизно на 90% тестових зображень, на зображеннях другого типу — близько 78%. Основні помилки були зумовлені недостатньою кількістю виділених на фасаді ліній, неправильним визначенням типу точки сходу та вибором початкових умов.



Рис. 1: Результат роботи алгоритму при центральному розташуванні камери (знайдені обидві точки сходу, але показана лише точка сходу на нескінченності).

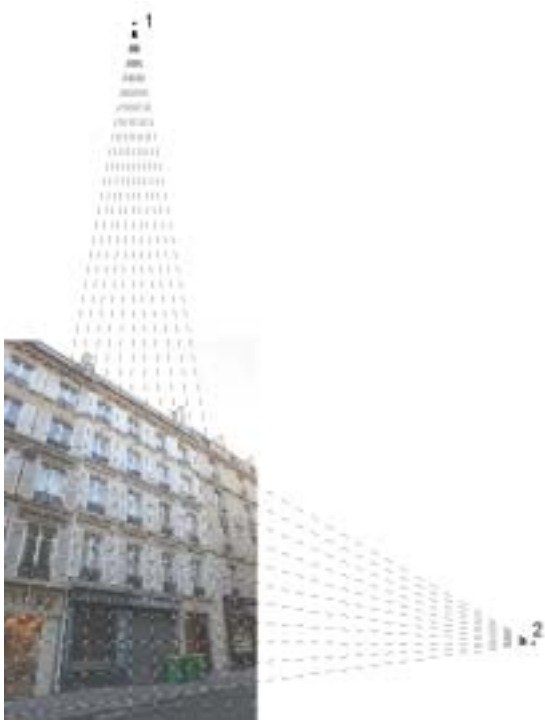


Рис. 2: Результат роботи алгоритму для зображення першого типу, 1 — „вертикальна“, 2 — „горизонтальна“ точки сходу.

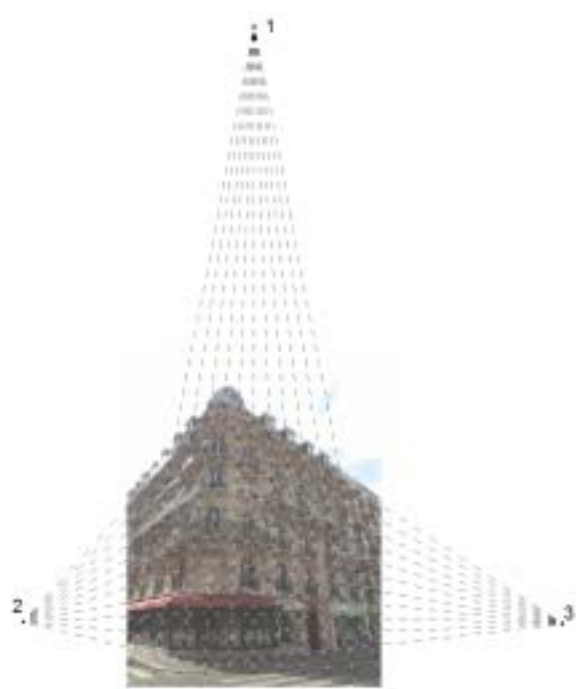


Рис. 3: Результат роботи алгоритму для зображення другого типу, 1 — „вертикальна“, 2 та 3 — „горизонтальні“ точки сходу.

4. Література

- [1] R. Hartley, A. Zisserman Multiple View Geometry in Computer Vision (Second Edition).—United Kingdom, Cambridge University Press, 2003.—656 p.
- [2] Thorsten Suttorp and Thomas Bucher. Robust vanishing point estimation for driver assistance. // In Proceedings of the 9th International IEEE Conference on Intelligent Transportation Systems (ITSC 2006), 2006. accepted.
- [3] O. Barinova, A. Kuzmishkina, A. Vezhnevets, V. Vezhnevets. Learning class specific edges for vanishing point estimation // Proc. of Graphicon'2007, pp. 162-165, Moscow, Russia, June 2007
- [4] A. Almansa, A. Desolneux, S. Vamech. Vanishing point detection without any a priori information // Pattern Analysis and Machine Intelligence, April 2003 (Vol. 25, No. 4) pp. 502-507
- [5] D.G. Aguilera, J. Gemez Lahoz, J. Finat Codes. A new method for vanishing points detection in 3D reconstruction from a single view // ISPRS Commission V, WG V/2. Mestre (Venice), 2005
- [6] Carsten Rother. A new approach for vanishing point detection in architectural environments. // In Proceedings of the British Machine Vision Conference, volume 20, pages 647Ц656, 2002.
- [7] Cantoni, V.; Lombardi, L.; Porta, M.; Sicard, N. Vanishing point detection: representation analysis and new approaches. // Image Analysis and Processing, 2001. Proceedings. 11th International Conference on Volume, Issue, 26-28 Sep 2001 Page(s) : 90 - 94
- [8] GF McLean, D Kotturi. Vanishing point detection by line clustering. IEEE Trans. Pattern Anal. and Machine Intell., Vol. 17, No. 11, pp. 1090-1095, 1995.
- [9] М. И. Шлезингер, В. Главач Десять лекцій по статистическому и структурному распознаванию.—Киев: Наук. думка, 2004. — 545 с.
- [10] R. O. Duda, P. E. Hart. Use of the Hough Transformation to Detect Lines and Curves in Pictures. // Comm. ACM, Vol. 15, pp. 11-15, January, 1972.
- [11] Demster A., Laird N., Rubin D. Maximum likelihood from incomplete data via the EM algorithm // Journal of the Royal Statistic Society.- 1977.- Vol. B39.- P. 1-38.