

УДК 004.5

ТЕХНОЛОГІЯ МАСШТАБОНЕЗАЛЕЖНОГО ПЕРЕТВОРЕННЯ ОЗНАК ЗОБРАЖЕНЬ

Р. Г. Зацерковний

Українська академія друкарства,
вул. Під Голоском, 19, Львів, 79020, Україна

Роз'яснено суть задачі відстеження об'єктів у межах галузі комп'ютерного зору. Обґрунтовано необхідність розробки спеціалізованих алгоритмів, адаптованих для відстеження об'єктів на противагу виявлення об'єктів. Здійснено формалізований опис алгоритму SIFT для масштабонезалежного перетворення ознак. Описано важливість інваріантності зображень щодо переміщення і масштабу та сформовано процес створення інваріантних зображень через функцію лапласіану гаусіана. Здійснено теоретичний опис кожного кроку програмного алгоритму. Описано можливі галузі використання алгоритму.

Ключові слова: машинне навчання, відстеження об'єктів, computer vision, SIFT.

Постановка проблеми. Одна з основних задач комп'ютерного зору — відстеження об'єктів, що переміщуються в межах декількох фотографій або відеокадрів. Ця задача використовується в галузях поліграфії, відеоспостереження, робототехніки, віртуальної реальності, розробки безпілотних автомобілів тощо. Хоча деякі експерти вважають, що для відстеження об'єктів (object tracking) можна застосовувати наявні методи виявлення об'єктів (object detection) з деякими незначними змінами, під час обробки відеоряду з'являються додаткові проблеми, які потрібно врахувати у процесі розробки алгоритму, наприклад:

- об'єкти можуть зникати з відеоряду і згодом знову з'являться у ньому;
- об'єкти можуть змінювати розмір, нахил або освітлення в межах одного відео;
- об'єкти можуть несподівано перетинатись один між одним;
- тощо.

Хоча людина може легко порівняти декілька зображень і візуально виділити на них об'єкти або відстежити рух об'єкта в межах одного відеозапису, комп'ютерний алгоритм не має «інтуїтивного» знання про природу зображення.

Найпопулярнішим способом визначення і зіставлення об'єктів є виявлення так званих ключових точок — особливих точок, які стосуються певного об'єкта і можуть бути використані для його ідентифікації на інших зображеннях. Для визначення та опису таких точок використовуються алгоритми комп'ютерного зору, одним з яких є масштабонезалежне перетворення ознак (scale-invariant feature transform, SIFT).

Мета статті — детально описати алгоритм масштабнонезалежного перетворення ознак зображень та обґрунтувати необхідність його використання для вирішення завдань комп'ютерного зору.

Виклад основного матеріалу дослідження. Загалом в алгоритмі SIFT є чотири основні кроки:

- створення масштабо-інваріантних зображень;
- локалізація ключових точок;
- орієнтація ключових точок;
- створення опису ключових точок.

Детальніше розглянемо кожен з цих кроків.

Перший крок роботи алгоритму полягає в перетворенні звичайного зображення в зображення, ознаки якого є інваріантними до переміщення та масштабу. Отже, SIFT позбувається небажаної характеристики деяких інших алгоритмів: характеристики зображення (кути та ін.) можуть відрізнитись при збільшенні масштабу; наприклад, кутова характеристика може перетворитись на декілька ліній, а інші характеристики — з'являться або зникати.

З метою створення інваріантного зображення обчислюється функція лапласіан гаусіана (LoG) для вхідних даних, що отримується в результаті операції згортки (convolution) зображення $I(x, y)$ з ядром гаусіана:

$$L(x, y, \sigma) = G(x, y, \sigma) \cdot I(x, y);$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma} e^{-\frac{(x^2+y^2)}{2\sigma^2}}.$$

Ця операція перетворює зображення на розмите, при цьому використовує параметр σ для регулювання ступеня розмиття: що більше значення параметра, то менша чіткість кінцевого зображення. У подальшому знаходження країв та кутів різних фігур здійснюється за допомогою обчислення суми похідних другого порядку для вхідного зображення, так званих лапласіанів. Зазвичай цей крок дещо спрощується через високу обчислювальну складність. Замість похідних обчислюються різниці гаусіанів — відмінності між гаусіановими розмиттями з двома різними значеннями σ . За допомогою порівняння значень декількох різниць апроксимуються значення необхідних похідних другого рівня.

На другому кроці роботи алгоритму ключові точки, а саме локальні екстремуми, локалізуються. Цей процес фактично полягає у фільтрації так званих недостатньо локалізованих ключових точок, які лежать вздовж краю лінії або недостатньо контрастні, тому не мають цінності для роботи алгоритму.

Алгоритм SIFT використовує значення ряду Тейлора другого порядку у ключових точках з метою фільтрації. Якщо значення не перевищує деякий заданий поріг (зазвичай це 0.03), то ключова точка відкидається:

$$D(x) = D + \frac{dD^T}{dx}x + \frac{1}{2}x^T \frac{d^2D^T}{dx^2}x;$$

$$|D(\hat{x})| < 0.03.$$



Рис. 1. Гаусіани зображення з $\sigma = 3$ і $\sigma = 5$ та різниця між ними

Крім того, відкидаються крайові точки, багато з яких знаходяться під час апроксимації з використанням DoG. Для погано визначених крайових точок знаходять слід та визначник матриці Хессіана другого порядку. Недостатньо локалізованими вважаються крайові точки, для яких відношення $\frac{Tr^2(H)}{Det(H)}$ не перевищує інше порогове значення; зазвичай $r = 10$.

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix};$$

$$Tr(H) = D_{xx} + D_{yy};$$

$$Det(H) = D_{xx}D_{yy} - D_{xy}^2.$$

Третій крок полягає у присвоєнні кожній ключовій точці орієнтації з метою досягнення інваріантності щодо повертання зображення. Інакше кажучи, характеристики не повинні змінюватись у зображеннях, де предмет розташований під іншим кутом. Для цього зчитуються деякі сусідні точки з ключовою (їх кількість залежить від масштабу зображення), і в цьому регіоні обчислюється градієнт — різниця між центральною та сусідніми точками. Зазвичай градієнт обчислюється для усіх варіацій поворотів сусідніх точок, для збільшення точності зіставлення характеристики зберігається найвищий з усіх отриманих.

На четвертому і останньому кроці створюється дескриптор характеристики: береться регіон розміром 16 на 16 пікселів і розділяється на підблоки розміром 4 на 4. Для кожного підблока створюється гістограма, яка зберігається у вигляді вектора і в подальшому використовується для зіставлення характеристик з іншими зображеннями. На цьому етапі робота алгоритму завершується.

Зіставлення характеристик між двома зображеннями виконується за допомогою алгоритму найближчих сусідів — кожен градієнт з одного зображення по чергово порівнюється з градієнтами іншого, і у випадку знаходження близького сусіда два градієнти зіставляються. У деяких випадках другий найближчий сусід схожий на перший (внаслідок шуму або з інших причин), але перебуває значно даліше від першого. У такому разі градієнти не зіставляються взагалі.



Рис. 2. Приклад роботи алгоритму SIFT для зіставлення двох зображень

Висновки. Внаслідок здатності SIFT знаходити характеристики зображення, інваріантні відносно переміщення, масштабу та повороту зображення алгоритм широко використовується у галузі розпізнавання об'єктів. Серед інших галузей, де використовується SIFT, є панорамна фотографія, робототехніка, 3D-моделювання та інші задачі комп'ютерного зору.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Ives Rey-Otero, Delbracio M. Anatomy of the SIFT Method. Image Processing On Line. 2014. Vol. 4. Pp. 370–396.
2. Lowe D. G. Object recognition from local scale-invariant features. Kerkyra, Greece : IEEE, 1999.
3. Lowe D. G. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision. 2004. Vol. 2. Pp. 91–110.
4. OpenCV. Open Source Computer Vision Library. 2015.
5. Lindeberg T. Image Matching Using Generalized Scale-Space Interest Points. Journal of Mathematical Imaging and Vision. 2015. Vol. 52. № 1. Pp. 3–36.

REFERENCES

1. Ives, Rey-Otero, & Delbracio, M. (2014). Anatomy of the SIFT Method: Image Processing On Line, 4, 370–396 (in English).
2. Lowe, D. G. (1999). Object recognition from local scale-invariant features. Kerkyra, Greece : IEEE (in English).

3. Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints: International Journal of Computer Vision, 2, 91–110 (in English).
4. OpenCV. Open Source Computer Vision Library. 2015 (in English).
5. Lindeberg, T. (2015). Image Matching Using Generalized Scale-Space Interest Points: Journal of Mathematical Imaging and Vision, 52, 1, 3–36 (in English).

doi: 10.32403/2411-3611-2020-1-37-74-78

SCALE-INVARIANT IMAGE FEATURE TRANSFORMATION TECHNOLOGY

R. H. Zatserkovnyi

*Ukrainian Academy of Printing,
19, Pid Holoskom St., Lviv, 79020, Ukraine
zatserkovnyi.rostyslav@gmail.com*

Object tracking in computer vision is the task of tracking of an entity such as a person, a car or another moving object through a video feed or series of images. Object tracking algorithms are used in the fields of video surveillance, robotics, virtual reality, self-driving cars and other high-tech industries, making it one of computer vision's most frequently used problems. Although sometimes object detection algorithms can be used for object tracking problems with relatively minor modifications, many additional issues are known to arise in object tracking, the main one among which is the need to track an individual object without regard for its scale, position, or rotation within an image. Commonly, this is done through identifying so-called "keypoints" within an image that best describe an object, and then using them to identify the same object in other images. However, many computer vision methods rely on the images that are being compared having a similar scale or rotation, and even minor changes in illumination can significantly impair the effectiveness of an algorithm.

In this article, the SIFT (scale-invariant feature transformation) algorithm often used for keypoint detection is described. Theoretical steps of the algorithm – the creation of scale-invariant images, the localization of keypoints that achieves scale invariance, the orientation of keypoints that achieves rotation invariance, and the turning of keypoints into a gradient descriptor – are denoted. Matching of key points between two images through the nearest-neighbor method is described, and practical examples of the algorithm's step-by-step operation are displayed. Finally, general comparisons are made between SIFT as well as other keypoint detection algorithms that do not satisfy the requirement for scale-invariance.

Keywords: machine learning, object tracking, computer vision, SIFT.

Стаття надійшла до редакції 22.04.2020.

Received 22.04.2020.