

УДК 004.93

ЗІСТАВЛЕННЯ ЗОБРАЖЕНЬ НА ОСНОВІ ПОШУКУ НАЙБЛИЖЧИХ СУСІДІВ В ПРОСТОРИ ПАРАМЕТРІВ

Дашкевич А.О., к.т.н.,

Шоман О.В., д.т.н.

Національний технічний університет

«Харківський політехнічний інститут» (Україна)

В задачах комп'ютерного зору та обробки зображень часто виникає необхідність проведення процесу зіставлення точок двох зображень, наприклад, для пошуку об'єктів на зображеннях або для відновлення тривимірних геометричних моделей сцени. В той же час, не існує загального методу визначення чітких точкових відповідностей на зображеннях.

В роботі запропоновано підхід до визначення пар ключових точок на зображеннях на основі розбиття простору параметрів на регулярну сітку і представлення такої сітки в вигляді просторової хеш-таблиці для прискорення пошуку. Підхід дозволяє визначати стійкі пари ключових точок на двох зображеннях, що дозволяє його використання в задачах стереозору для відновлення тривимірних моделей поверхонь і для пошуку об'єктів на зображеннях.

Представлений підхід складається з наступних кроків: визначення ключових точок з використанням дескрипторів; зіставлення ключових точок на основі побудови двовимірного простору параметрів, який формується з розміру і кута орієнтації дескриптора; розбиття простору параметрів на регулярну сітку; побудова хеш-таблиці на основі сітки, в якості значень в комірках хеш-таблиці містяться кількість пар ключових точок з близькими значеннями розміру і орієнтації дескриптора, що відповідають даній комірці; пошук в хеш-таблиці комірки, що міститиме найбільшу кількість точок, така комірка відповідатиме найбільш стійким відповідностям ключових точок на зображеннях. Представлений алгоритм може бути розширений для роботи з довільною кількістю параметрів. Запропонований алгоритм дозволяє проводити зіставлення ключових точок на зображеннях за час $O(n)$.

Ключові слова: ключові точки, зіставлення зображень, простір параметрів, дескриптор, розбиття на регулярну сітку, просторові хеш-таблиці, стереозір.

Постановка проблеми. В задачах комп'ютерного зору та обробки зображень часто виникає необхідність проведення процесу зіставлення двох зображень, наприклад, для пошуку об'єктів на зображеннях або для відновлення тривимірних геометричних моделей сцени. Для ефективного розв'язання таких задач необхідно визначити стійкі пари ключових точок на вхідних зображеннях, такий процес може бути представлений як послідовність наступних кроків:

- пошук ключових точок на двох вхідних зображеннях;
- зіставлення знайдених ключових точок та формування пар точок, що належать однаковим об'єктам.

Задача пошуку ключових точок на даний час може бути доволі ефективно розв'язана з використанням таких алгоритмів, як scale-invariant feature transform (SIFT) [1], speeded up robust features (SURF) [2], oriented FAST and rotated BRIEF (ORB) [3], методів на основі штучних нейронних мереж [4] та ін. В той же час, не існує загального методу визначення чітких точкових відповідностей на зображеннях. В зв'язку з цим існує необхідність розробки методів визначення стійких відповідностей ключових точок на зображеннях.

Аналіз останніх досліджень і публікацій. Для пошуку відповідностей ключових точок на зображеннях найчастіше використовують спеціальні просторові структури даних, наприклад kd-дерева [5, 6], BSP-дерева [7, 8], R-дерева [9] або алгоритми просторового хешування [10]. Реалізація вказаних методів міститься в бібліотеці fast library for approximate nearest neighbor (FLANN) [11], яку можливо застосовувати для зіставлення ключових точок.

Основними підходами до зіставлення є:

- прямий підхід (метод грубої сили) – побудова матриці відстаней від усіх ключових точок одного зображення до усіх ключових точок другого зображення, але часова складність такого підходу складає $O(n^2)$, що не дає змогу використовувати його у випадках, коли кількість ключових точок завелика;
- оптимізований підхід з використанням просторових структур даних, в якому ключові точки обох множин організуються в єдину структуру і пошук відбувається за час $O(n \log n)$ з використанням дерев і за $O(n)$ з використанням просторового хешування, але в деяких випадках обидва підходи дають складність в найгіршому випадку $O(n^2)$.

В роботі [10] запропоновано підхід до визначення стійких пар ключових точок на сусідніх кадрах відео послідовності за умови відносного руху сцени між кадрами. Недоліком зазначеного підходу є те, що ефективно зіставлення можливе лише за умови руху сцени відносно камери, тобто придатне тільки для роботи з відео-матеріалами і не може бути застосовано для довільних зображень

або зображень тієї ж сцени, але виконаних в різний час, або зі значним відносним рухом між кадрами. Інший недолік – зіставлення ключових точок відбувається в два етапи, що зменшує часову ефективність алгоритму, тому, необхідною є побудова алгоритму без попереднього зіставлення точок.

Формулювання цілей статті. Розробка методу зіставлення ключових точок на зображеннях на основі розбиття простору параметрів дескрипторів зображень на регулярну сітку і пошуку найближчих сусідів на такій сітці.

Основна частина. В представленій роботі запропоновано розвиток підходу з [10]. Наведемо основні кроки алгоритму [10]:

- 1) визначення ключових точок з використанням дескрипторів ORB (або SIFT);
- 2) початкове зіставлення ключових точок з використанням бібліотеки FLANN;
- 3) побудова двовимірного простору параметрів, який формується з напрямку вектора руху ключової точки між кадрами та модуля такого вектора;
- 4) розбиття простору параметрів на регулярну сітку – гіперкуб, з розмірністю $t \times t$, де t – кількість клітин сітки вздовж кожної вісі в просторі параметрів;
- 5) побудова хеш-таблиці на основі сітки, в якості значень в комірках хеш-таблиці місяться кількість пар ключових точок з близькими значеннями напрямку вектору руху і його модуля, що відповідають даній комірці;
- 6) пошук в хеш-таблиці комірки, що міститиме найбільшу кількість точок, така комірка відповідатиме переважному напрямку руху ключових точок між кадрами.

Пропонується проводити зіставлення на етапі отримання дескрипторів (1), що дозволить позбавитись кроку (2). Кожний дескриптор ключової точки має такі основні параметри:

$p(x, y)$ – координати положення точки на зображенні;

s – діаметр (розмір) дескриптору;

α – кут орієнтації дескриптору;

r – ваговий коефіцієнт дескриптору;

o – номер октави даного дескриптору.

Визначальними параметрами є розташування, діаметр і орієнтація ключової точки, але для досягнення інваріантності до паралельного перенесення пропонується використання тільки s та α для формування простору параметрів. Надалі проводиться дискретизація (рис. 1) і робота алгоритму відбувається згідно з кроками 3)–6).

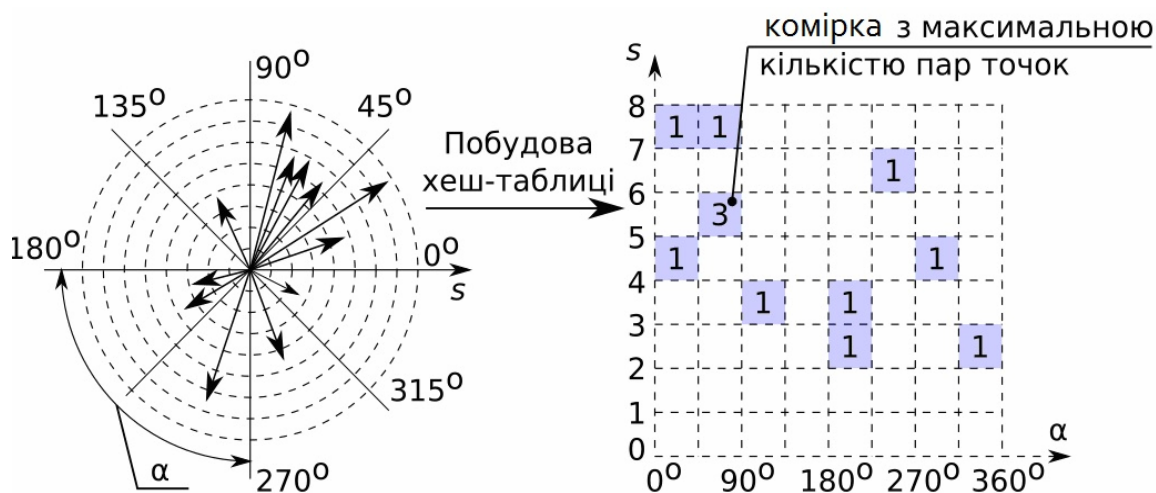


Рис. 1. Схема дискретизації простору параметрів

На рис. 2 представлено результат зіставлення ключових точок для двох зображень з використанням бібліотеки FLANN та за запропонованим підходом. Можна побачити, що результатом роботи алгоритму є менша кількість пар точок, але результат є більш коректним на відміну від зіставлення з FLANN.

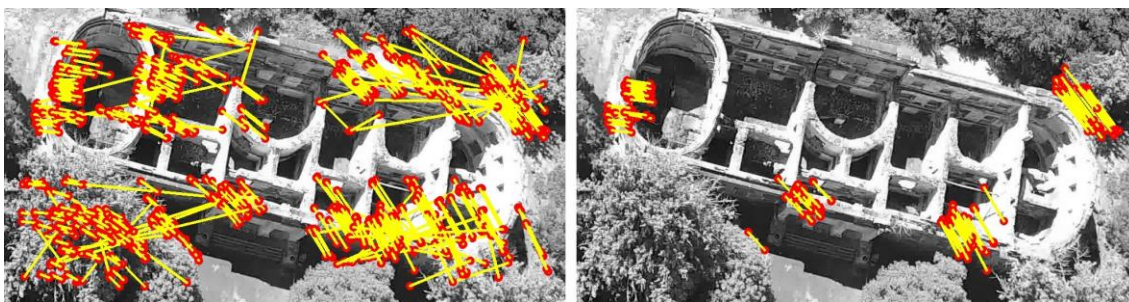


Рис. 2. Візуалізація роботи алгоритму

Висновки. Запропонований алгоритм дозволяє проводити зіставлення ключових точок на зображеннях за час $O(n)$ в середньому. Алгоритм може бути використаний для пошуку об'єктів на зображеннях і для відновлення тривимірної моделі сцени методами стереозору і структури із руху. До недоліків можна віднести невелику кількість використаних параметрів в процесі дослідження, але алгоритм можливо розширювати для роботи з довільною кількістю параметрів.

Література

1. Lowe D.G. Object recognition from local scale-invariant features. Proceedings of the Seventh IEEE International Conference on Computer Vision. IEEE, Kerkyra, Greece, 1999. Vol. 2. P. 1150–1157.
2. Bay H., Tuytelaars T., Van Gool L SURF: Speeded Up Robust Features. Computer Vision – ECCV 2006. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006. P. 404–417.
3. Rublee E., Rabaud V., Konolige K., Bradski G. ORB: An efficient alternative to SIFT or SURF. *International Conference on Computer Vision*. IEEE, Barcelona, Spain, 2011. P. 2564–2571.
4. Riegler G., Ulusoy A. O., Bischof H., Geiger A. OctNetFusion: Learning Depth Fusion from Data. International Conference on 3D Vision (3DV). IEEE, Qingdao, 2017. P. 57–66.
5. Bentley J.L. Multidimensional Divide and Conquer. *Communications of the ACM*, 1980. Vol. 23. Is. 4. P. 214–229.
6. Friedman J.H., Bentley J. L., Finkel R. A. An algorithm for finding best matches in logarithmic expected time. *ACM Transactions on Mathematical Software*, 1977. Vol. 3. Is. 3. P. 209–226.
7. de Berg M. Computational Geometry: Algorithms and Applications Springer Science & Business Media, 2008. P. 259.
8. Castelli V, Lawrence D. . Image Databases: Search and Retrieval of Digital Imagery. John Wiley & Sons, 2004. P. 422.
9. Guttman A. R-Trees: A Dynamic Index Structure for Spatial Searching. Proc. ACM SIGMOD International Conference on Management of Data, 1984. P. 47–57.
10. Дашкевич А.А., Воронцова Д. В. , Скоробогатько Н. В. Алгоритм поиска устойчивых соответствий пар ключевых точек на изображениях и картах глубины. *Вестник НТУ «ХПИ», Серія: Нові рішення в сучасних технологіях*. Харків: НТУ «ХПИ», 2019. № 5 (1330). С. 86–90.
11. Muja M., Lowe D. G. Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration. *International Conference on Computer Vision Theory and Application (VISSAPP'09)*. INSTICC Press, 2009. P. 331–340.

СОПОСТАВЛЕНИЕ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ПОИСКА БЛИЖАЙШИХ СОСЕДЕЙ В ПРОСТРАНСТВЕ ПАРАМЕТРОВ

Дашкевич А.А., Шоман О.В.

В задачах компьютерного зрения и обработки изображений часто возникает необходимость проведения процесса сопоставления точек двух изображений, например, для поиска объектов на изображениях или для восстановления трёхмерных геометрических моделей сцены. При этом не существует общего метода определения устойчивых точечных соответствий на изображениях.

В работе предложен подход к определению пар ключевых точек на изображениях на основе разбиения пространства параметров на регулярную сетку и представления такой сетки в виде пространственной хеш-таблицы для ускорения поиска. Подход позволяет определять устойчивые пары ключевых точек на двух изображениях, что позволяет использовать его в задачах стереозрения для восстановления трёхмерных моделей поверхностей и для поиска объектов на изображениях.

Представленный подход состоит из последовательности шагов: определение ключевых точек с использованием дескрипторов; сопоставление ключевых точек на основе построения двумерного пространства параметров, которое формируется из размера и угла ориентации дескриптора; разбиение пространства параметров на регулярную сетку; построение хеш-таблицы на основе сетки, в качестве значений в ячейках хеш-таблицы содержатся количества пар ключевых точек с близкими значениями размера и ориентации дескрипторов, которые соответствуют данной ячейке; поиск в хеш-таблице ячейки с наибольшим количеством точек, такая ячейка будет соответствовать наиболее устойчивым сочетаниям ключевых точек на изображениях. Представленный алгоритм может быть расширен для работы с произвольным числом параметров и позволяет проводить сопоставление ключевых точек на изображениях за время $O(n)$.

Ключевые слова: ключевые точки, сопоставление изображений, пространство параметров, дескриптор, разбиение на регулярную сетку, пространственные хеш-таблицы, стереозрение.

IMAGE MATCHING BASED ON NEAREST NEIGHBOR SEARCH IN PARAMETER SPACE

Dashkevich A., Shoman O.

In tasks of computer vision and image processing, it is often necessary to carry out the process of comparing points of two images, for example, to search for objects in images or to restore three-dimensional geometric models of the scene. At the same time, there is no common method for determining stable point correspondences in images.

An approach to determine of key point pairs in the images based on a subdivision of parameter space into regular grid and building a spatial hash table over the grid is proposed. Approach, that is developed, provides to determine strong key point pairs in two images, which makes possible to use it in the stereovision problems to surface reconstruction and for object search in the images.

The presented approach consists of a sequence of steps: defining key points using descriptors; matching key points based on the construction of a two-dimensional parameter space, which is formed from the size and orientation angle of the descriptor; partitioning the parameter space into a regular grid; constructing a hash table based on the grid; the values in the cells of the hash table contain the number of pairs of key points with similar values of the size and orientation of the descriptors that correspond to this cell; searching in the hash table of the cell with the greatest number of points, such a cell will correspond to the most stable combinations of key points in the images.

The presented algorithm can be extended to work with an arbitrary number of parameters and allows for the mapping of key points on images with a time complexity of $O(n)$.

Key words: key points, image matching, parameter space, descriptor, subdivision into regular grid, spatial hash-tables, stereovision.