

**ЗАВДАННЯ ЗІСТАВЛЕННЯ ЗОБРАЖЕНЬ ТА ПОБУДОВА
SIFT ДЕСКРИПТОРІВ**

Маханьов Р.О.

Науковий керівник – к.т.н., доц. Кобилін О.А.

Харківський національний університет радіоелектроніки, каф. ІНФ,
м. Харків, Україна,тел.: (057) 702-14-19, e-mail: ruslan.makhanov@nure.ua

This study investigates the efficacy of detecting key points in images using a method based on Gaussian pyramids and differences. Key points are identified based on local extrema in the Gaussian pyramid levels, and their orientations are determined by analyzing gradient directions in the vicinity. A descriptor is then computed for each key point to provide a compact representation of its local features. These descriptors, computed on Gaussian-smoothed images, enable robustness to rotations. Despite their versatility, these descriptors may encounter limitations in scenarios involving repetitive patterns or reflections. Nevertheless, they find wide applicability in various image analysis tasks.

Головною процедурою при виявленні ключових точок є створення піраміди гаусіанів та різниць між ними [1]. Піраміда гаусіанів – це метод масштабування та згладжування зображень, що використовує послідовність гаусіанських розмиттів на різних рівнях масштабування. Гаусіаном є зображення:

$$L(x, y, \zeta) = G(x, y, \zeta) * I(x, y), \quad (1)$$

де L – значення гаусіана в точці з координатами (x, y) ; ζ – радіус розмиття; G – гаусове ядро; I – значення вихідного зображення.

Простір зображення, що масштабується, представляє собою набір різних версій вихідного зображення, які були згладжені за допомогою певного фільтра. Оскільки досліджено, що гаусово масштабований простір є лінійним та інваріантним відносно зсувів, обертань та масштабу, при цьому не порушуючи локальних екстремумів, а також виконує властивості напівгрупи, то різні рівні розмиття зображення гаусовим фільтром можуть бути розглянуті як вихідне зображення на різних масштабах.

Ключова точка – це точка на зображенні, яка відзначається особливими характеристиками, такими як кути, краї, текстури або яскраві точки, які є унікальними та легко відрізняються від інших областей зображення. Після визначення того, що певна точка є ключовою, необхідно визначити її орієнтацію. Точка може мати кілька можливих напрямків [2].

Орієнтація ключової точки визначається на основі напрямків градієнтів у точках, що оточують особливу точку. Усі розрахунки градієнтів виконуються на зображенні в піраміді гаусіанів, з масштабом, що найближ-

чий до масштабу ключової точки. Величина градієнта у точці (x, y) обчислюються за допомогою формули:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}, \quad (2)$$

де m – величина градієнта.

Напрямок градієнта у точці (x, y) обчислюються за допомогою формули:

$$\theta(x, y) = \tan^{-1}((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))), \quad (3)$$

де θ – напрямок градієнта.

Дескриптор – ідентифікатор ключової точки, що виділяє її з решти маси особливих точок. Він представляє собою компактне числове представлення особливостей околиць цих точок, виокремлюючи їх із загального масиву точок на зображенні. Дескриптор дозволяє описати локальні особливості, та використовується для подальшого порівняння, розпізнавання та відновлення об'єктів на зображеннях.

В цілому, дескриптор може бути будь-яким об'єктом (якщо він виконує свої функції), проте зазвичай ним є інформація про оточення ключової точки.



Рисунок 1 – Приклад порівняння зображень за допомогою SIFT дескрипторів

У методі SIFT дескриптором є вектор, який представляє собою компактне числове представлення особливостей околиць ключової точки на зображенні. Як і напрямок ключової точки, дескриптор обчислюється на гаусіані, що має найближчий масштаб до ключової точки, і базується на градієнтах у певному вікні навколо ключової точки. Перед обчисленням дескриптора це вікно обертається на кут напрямку ключової точки, що забезпечує інваріантність до повороту [3, 4].

У процесі виявлення ключових точок та побудови їх дескрипторів спочатку створюється піраміда гаусіанів та різниць, з використанням гаусіанських фільтрів для створення масштабованого зображення. Кожна ключова точка відображається на окремій області зображення, де кожен піксель утворює квадратне вікно дескриптора. Для кожного пікселя області обчислюються градієнти та враховується вплив гаусівського розмиття. Градієнти кожної області групуються у гістограми, що характеризують напрямки та інтенсивності змін. Ці гістограми об'єднуються, утворюючи фінальний векторний дескриптор, який ідентифікує ключову точку. Дескриптор нормалізується та урізається для зменшення впливу шуму, а потім нормалізується повторно для подальшого використання. Такий алгоритм дозволяє створити компактне та репрезентативне представлення ключових точок зображення, що може бути використане для подальших обчислень та порівнянь [5].

Дескриптори SIFT не є бездоганними, що підтвердилось практичними результатами на рис. 1. Не всі виявлені точки та їхні описи відповідають висунутим вимогам. У деяких ситуаціях може виявитися, що рішення неможливо знайти, навіть якщо воно існує. Наприклад, при спробі визначення афінних перетворень, коли на зображенні додатково присутнє віддзеркалення шуканого об'єкту, може виникнути проблема, оскільки отримуємо повторюванні об'єкти, що робить дескриптори різних ключових точок схожими між собою. Незважаючи на ці обмеження, SIFT дескриптори можливо успішно використовувати у багатьох практичних сценаріях.

Список використаних джерел:

1. Кобилін О. А., Творошенко І. С. Методи цифрової обробки зображень : навч. посіб. Харків : ХНУРЕ, 2021. 124 с. URL: <https://doi.org/10.30837/978-966-659-295-1>. (дата звернення: 10.01.2024)
2. Gorokhovatskyi V., Zamula A. Employment of intelligent technologies in multiparametric control systems. telecommunications and radio engineering. Telecommunications and Radio Engineering. 2016. Vol. 75, No. 19, P. 1775–1785.
3. Yakovleva O., Nikolaieva K. Research of descriptor based image normalization and comparative analysis of SURF, SIFT, BRISK, ORB, KAZE, AKAZE descriptors. Advanced Information Systems. 2020. Vol. 4, No. 4, P. 89–101.
4. Yakovleva O., Kovtunen A., Liubchenko V., Honcharenko V., Kobylin O. Face detection for video surveillance-based security system. CEUR Workshop Proceedings. 2023. Vol. 3, No. 3, P. 69–86.
5. Tvoroshenko I., Pomazan V., Gorokhovatskyi V., Kobylin O. Application of video data classification models using convolutional neural networks. 2023. Vol. 7, No. 11, P. 134–145.