

ПРО ВИКОРИСТАННЯ КЛАСТЕРНОЇ ВИБІРКИ В АЛГОРИТМАХ ВІЗУАЛЬНОЇ ОДОМЕТРІЇ

С. Іванов, А. Музичук

*Львівський національний університет імені Івана Франка,
вул. Університетська 1, Львів, 79000
e-mail: serhii.ivanov@lnu.edu.ua, anatoliy.muzychuk@lnu.edu.ua*

Розглянуто підхід візуальної одометрії для обчислення траєкторії літаючої платформи на підставі даних з її бортової відеокамери. Він полягає в покроковому застосуванні методу SURF, запропонований у праці Herbert Bay, Andreas Ess, Tinne Tuytelaars і Luc Van Gool, для побудови афінного перетворення, з допомогою якого можна відобразити множини точок з одного кадра на інший. Визначення на кадрі точок інтересу, які беруть до уваги під час побудови матриці перетворення, виконується шляхом пошуку точок максимуму гесіана для функції, що описує яскравість пікселів зображення. Оскільки такий детектор інваріантний до афінних перетворень системи координат, то визначені ним точки інтересу зумовлюють ефективну побудову відповідності між кадрами, які вибирають з відео з деяким кроком. У підсумку це дає змогу оцінити зміну координат рухомої платформи у глобальній системі координат.

Наведено результати апробації розробленого програмного забезпечення, які демонструють застосовність методу SURF для обчислення траєкторій рухомих платформ. Зокрема з'ясовано, що так можна будувати траєкторії, які близькі до траєкторій на основі GPS-даних.

Заразом проаналізовано вплив характеру досліджуваної місцевості та її освітлення на кількість точок інтересу, які виявляє детектор. Означено, що у більшості розглянутих випадків їхня кількість виявилася надмірною на етапі побудови відповідності між суміжними відеокадрами. Ця обставина виявлялася у підборі методом RANSAC значної кількості точок, які не потрібні для обчислення матриці перетворення. У підсумку це приводило до надлишкового часу обчислення траєкторії.

З метою оптимізації кількості потрібних точок інтересу автори запропонували модифікацію RANSAC, яку назвали кластерний RANSAC. Його суть полягає у визначенні на кожному k -му кроці SURF деякого кластера на попередньому кадрі c_{k-1} , на якому алгоритм RANSAC робить спробу побудувати потрібне афінне перетворення для відображення на кадр c_k . Приклади використання програмної реалізації методу SURF із застосуванням кластерного RANSAC демонструють значне скорочення часу обчислення траєкторій порівняно з застосуванням програмної реалізації звичайного методу RANSAC. Таке вдосконалення створює підстави для використання модифікованого методу SURF у режимі реального часу.

Ключові слова: візуальна одометрія, детектор SURF, кластерний RANSAC, оцінка руху.

1. ВСТУП

Сучасний рівень розвитку комп'ютерної техніки та оптичних пристроїв створює умови для розробки навігаційних систем для безпілотних літаючих апаратів (UAV – Unmanned Aerial Vehicle), здатних забезпечувати автономну навігацію на підставі інформації, яка надходить лише з їхніх відеокамер. Такий підхід зумовлює низку задач, які необхідно розв'язувати в режимі реального часу з використанням обмежених обчислювальних ресурсів, що перебувають на борту UAV. Звідси впливає актуальність розробки чисельних методів та відповідного програмного забезпечення для ефективного розв'язування цих задач. Детальний огляд підходів до вирішення

пов'язаних з UAV проблем та отриманих у цій сфері результатів див., наприклад, [4, 10]. Зазначимо, що у переліку таких сучасних ефективних підходів є візуальна одометрія (VO).

Суть підходів VO полягає в оцінці зміни позиції рухомого об'єкта у тривимірному просторі на підставі лише даних, що надходять з відеокамер, встановлених на цьому об'єкті. Сучасні методи VO поєднують технології комп'ютерного зору (CV) та машинного навчання (ML). Термін “візуальна одометрія” був введений у праці [9] для дистанціювання від інших підходів, які знайшли широке застосування в різних галузях робототехніки завдяки використанню даних також від інших сенсорів і відомі у відповідній літературі як методи неперервної локалізації та відображення (SLAM – Simultaneous Localization and Mapping). Огляд методів SLAM див., наприклад, у [5, 8]).

VO також можна розглядати як альтернативу підходам, які отримують координати об'єкта від систем глобального позиціонування (GPS). Зауважимо, що попри широке назагал використання, системи GPS не завжди доступні за різних природних чи штучних причин, а також можуть подавати координати зі значною похибкою. Методи CV у поєднанні з ML дають змогу вирішувати окреслені проблеми з великою точністю в режимі реального часу.

Типовою задачею VO є пошук відповідних множин дискретних точок, які належать двом зображенням тієї самої сцени. Їх розглядають на двох відеокадрах, які з певною періодичністю можуть надходити з того самого відео, або ж належати різним відео, зокрема отриманим різними пристроями за різних обставин. Важливою умовою для розв'язання зазначеної задачі є часткове перекриття зображень, яке містить відповідні точки з певними характеристиками.

Побудова відповідності двох різних зображень сучасними методами VO ґрунтується на розгляді так званих точок інтересу і відбуваються в три етапи. Перший – етап детекції – полягає у виявленні точок інтересу, а саме пікселів, в яких деяка функція, пов'язана з яскравістю зображення, досягає максимуму. На другому етапі для кожної точки інтересу обчислюють характеристики, які описують її зв'язок з яскравістю сусідніх пікселів. Обчислені значення зберігають у векторах-дескрипторах. Третій етап полягає у пошуку множин пар точок інтересу, які є зображенням на послідовних відеокадрах відповідних точок місцевості, і побудові відображення таких множин одна на одну. Для цього зазвичай використовують афінне перетворення. Тобто, задача зводиться до побудови відповідної матриці, за допомогою якої можна послідовно відображати пікселі попереднього відеокадру на наступний і після переходу до глобальних координат оцінювати параметри руху UAV.

Відомі різні підходи до побудови дескрипторів. Їхні найпоширеніші математичні моделі описані в [1, 4, 6]. Основні вимоги до дескрипторів – інваріантність щодо певного класу перетворень координат – переміщень, повороту чи зміни масштабу, а також забезпечення високої швидкості обчислення. Особливістю таких моделей є дискретність функції яскравості, яка описує зображення. Тому під час попереднього опрацювання відеоданих передбачається їх згладжування за допомогою операції згортки з деяким гладким ядром. Це дає змогу перенести операції диференціювання на ядро й отримувати детектори точок інтересу на основі різних комбінацій похідних, які найкраще підходять для конкретних задач VO. На практиці для задач навігації UAV ефективним виявився детектор на основі гесіана функції яскравості, зокрема його реалізація у методі SURF (Speeded-Up Robust Features) [1].

Метод SURF ґрунтується на згортці функції яскравості з ядром у вигляді функції густини ймовірності гауссового розподілу. Він характеризується швидким обчисленням гесіана функції яскравості за рахунок апроксимації ядра та ефективних алгоритмів чисельного інтегрування при обчисленні згорток. У підсумку це визначає переваги SURF над іншими методами VO і зміщує акценти оптимізації методу на описаний вище етап побудови відображень точок інтересу. В сучасних реалізаціях з цією метою широко використовують різні варіанти алгоритму RANSAC (RANdom SAMple Consensus) [2].

Наша мета – вдосконалити метод SURF з врахуванням специфіки відеозображень, отриманих при горизонтальних переміщеннях UAV, а також вироблення методик його застосування за різних обставин. Вона складається зі вступу, трьох розділів і висновків. У першому розділі розглянуто формулювання задачі, стисло описано суть методу SURF та наведено приклад обчислення цим методом траєкторії UAV на підставі відеоданих. Другий розділ присвячений аналізу алгоритмів формування множин точок інтересу для сцен різного виду і з різним освітленням. У третьому розділі описано модифікацію методу RANSAC, яка дає змогу реалізувати швидкі алгоритми пошуку точок інтересу, та на прикладі продемонстровано його ефективність.



Рис. 1. Невідформатований кадр розміром 1920x1080 пікселів у внутрішній просторовій системі координат

2. ВИЗНАЧЕННЯ ОСНОВНИХ ПОНЯТЬ І МОДЕЛЕЙ

Розглянемо деякі особливості предметної області, яка пов'язана з навігацією UAV. Найперше, це використання трьох систем координат: глобальної, локальної та внутрішньої. Глобальну (світову) використовують для відстежування руху апарата (відеокамери) безпосередньо у деякому реальному середовищі. Внутрішня (власна) асоційована у стандартно з конкретним відеокадром за допомогою матриці

відповідного піксельного зображення. У цьому випадку початок координат є у лівому верхньому куті кадра, вісь ординат направлена вниз. Локальну систему прив'язують до конкретної точки на сцені шляхом перенесення початку координат та задання відповідних напрямів осей координат.

Зображення, отримане за допомогою відеокамери відповідно до її роздільної здатності, є дискретним зі сталими на пікселях значеннями яскравості. Приклад такого зображення місцевості на одному з попередньо неопрацьованих кадрів відео зображено на рис. 1.

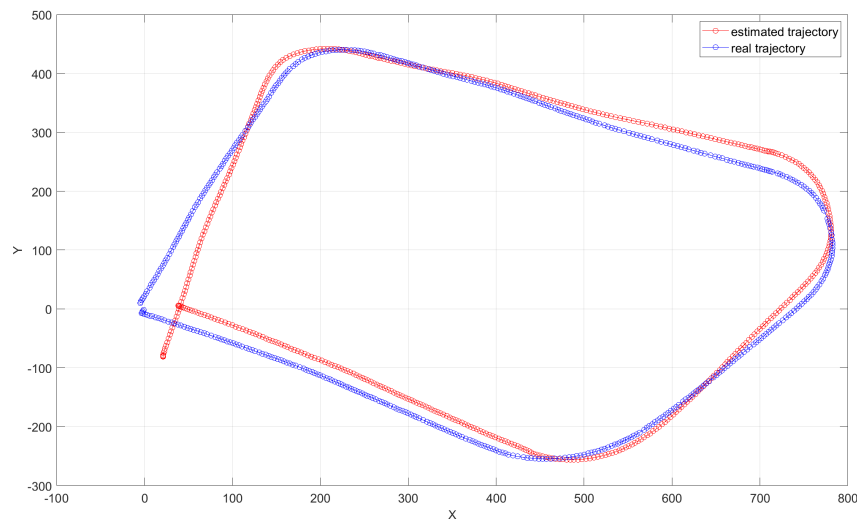


Рис. 2. Обчислена на основі відео траєкторія (червоний колір) і реальна траєкторія (синій колір) на основі GPS координат

Зазначимо, що у внутрішній системі пікселі мають цілочислові координати. Крім того, яскравість пікселів також задають невід'ємним цілим числом. Через ці дві обставини більшість операцій у цій системі координат можна виконувати в арифметиці цілих чисел. З одного боку, це дає змогу будувати високопродуктивні обчислювальні комплекси, з іншого – потребує спеціальних підходів до побудови математичних моделей, зокрема при диференціюванні та інтегруванні.

Нехай на деякій прямокутній області $\Omega \in \mathbb{R}^2$ яскравість зображення у відтінках сірого кольору задана деякою функцією \tilde{f} . Розглянемо її згортку $f(x) := (\tilde{f} * g)(x)$ з гауссіаном

$$g(x; \sigma) := \frac{1}{2\pi\sigma} e^{-|x|/2\sigma}, \quad \sigma > 0, \quad x = (x_1, x_2) \in \Omega. \quad (1)$$

Використовуючи властивість похідних згортки, зокрема перекидаючи при диференціюванні похідні на ядро, для функції $f : \Omega \rightarrow \mathbb{R}$ можемо визначити вектор градієнта

$$\nabla f = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right]^\top \quad (2)$$

і матрицю Гессе

$$H(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} \text{ в } \Omega. \quad (3)$$

Оскільки мішані похідні цієї функції неперервні, то порядок диференціювання в (3) можна змінювати. Тоді матимемо змогу обчислювати гессіан

$$\det(H(f)) = \frac{\partial^2 f}{\partial x_1^2} \frac{\partial^2 f}{\partial x_2^2} - \left(\frac{\partial^2 f}{\partial x_1 \partial x_2} \right)^2 \text{ в } \Omega. \quad (4)$$

Він є інваріантним щодо афінних перетворень координат, що створює підстави для його використання в детекторах точок інтересу [6]. Зокрема детектор методу SURF [1] використовує точки локальних максимумів гессіану в Ω .

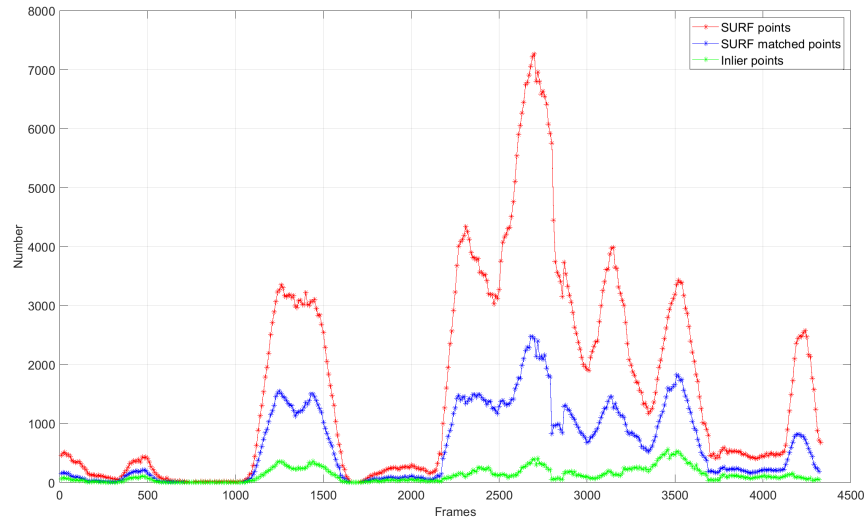


Рис. 3. Кількість точок інтересу, виявлених методом SURF вздовж траєкторії, зображеної на рис. 2

Зазначимо, що за рахунок спеціальних підходів до апроксимації похідних і чисельного інтегрування при обчисленні згорток у формулі (4), а також завдяки швидкому алгоритмові пошуку максимуму цілочислових функцій метод SURF досягає високої швидкодії при побудові множин точок інтересу. Зазначені характеристики зумовлюють широке використання SURF для проведення досліджень у рамках різних фреймворків (див., наприклад, [3]), зокрема MATLAB [7]. На рис. 2 продемонстровано обчислену в MATLAB на основі відео методом SURF траєкторію (червоний колір), яку описує центр відеокамери при русі UAV. Як бачимо, вона добре узгоджена з траєкторією, побудованою з GPS-даних (синій колір), що свідчить про придатність SURF до використання в режимі реального часу в практичних задачах VO.

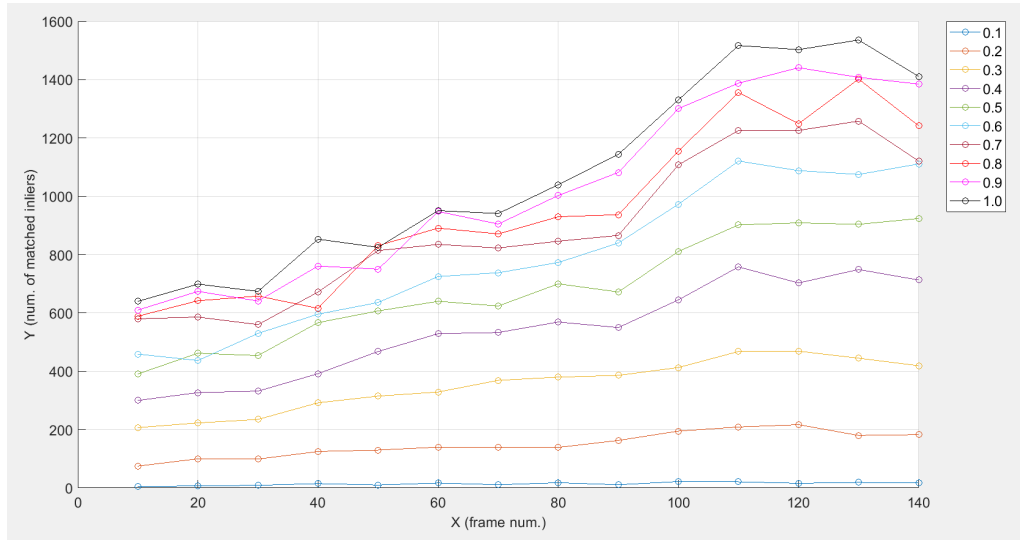


Рис. 4. Кількість точок інтересу, виявлених методом SURF вздовж траєкторії над кварталом міста з освітленістю $f_k := kf$, $k \in (0, 1]$, де f – функція яскравості у відео, знятому у час з найбільшою освітленістю сцени

3. ОСОБЛИВОСТІ ФОРМУВАННЯ МНОЖИН ТОЧОК ІНТЕРЕСУ ДЛЯ ЧЕРГОВОГО ВІДЕОКАДРУ

Детектор методу SURF дає змогу визначати певні області (“blobs” у термінах методу) – характерні частини зображення з точками інтересу, в яких не виникатимуть нові точки максимуму гессіана (4) зі збільшенням масштабу зображення (див. [6]). Причому беруться до уваги лише такі точки інтересу, в яких значення гессіана більше від деякого порогового значення, заданого параметром Threshold.

Позначимо N_i^{SURF} кількість точок інтересу на кадрі з номером i . Як свідчить практика, застосування SURF, значення величини N_i^{SURF} залежить (крім Threshold) також і від вигляду сцени(місцевості), для якої розглядають відеозображення. Параметр N_i^{SURF} може мати широкий діапазон зміни. Наприклад, на рис. 3 графік червоного кольору демонструє динаміку зміни N_i^{SURF} на послідовності кадрів, взятих з відеозображення, за яким обчислювали траєкторію на рис. 2. Для цього кадри вибирали з деяким кроком з відео за деякого фіксованого значення параметра Threshold. Відповідно до такого алгоритму числа по осі абсцис означають номер кадру у відеозображенні. Зміну параметра N_i^{SURF} на порівняно однорідних сценах демонструють графіки чорного кольору на рис. 4 і 5, які стосуються зображень кварталу міста та ділянки поля, відповідно. Як бачимо, у кожному з випадків цей параметр має свій діапазон зміни.

Зауважимо, що точки, які кожного разу беруть для побудови відповідності між послідовними кадрами, становлять лише підмножину з загальних множин точок інтересу на відповідних кадрах. Як видно з графіка синього кольору на рис. 3, їхня кількість може бути значною на окремих ділянках траєкторії. Щоб уникнути сповільнення алгоритму, який визначає відповідність завдяки попарному перебо-

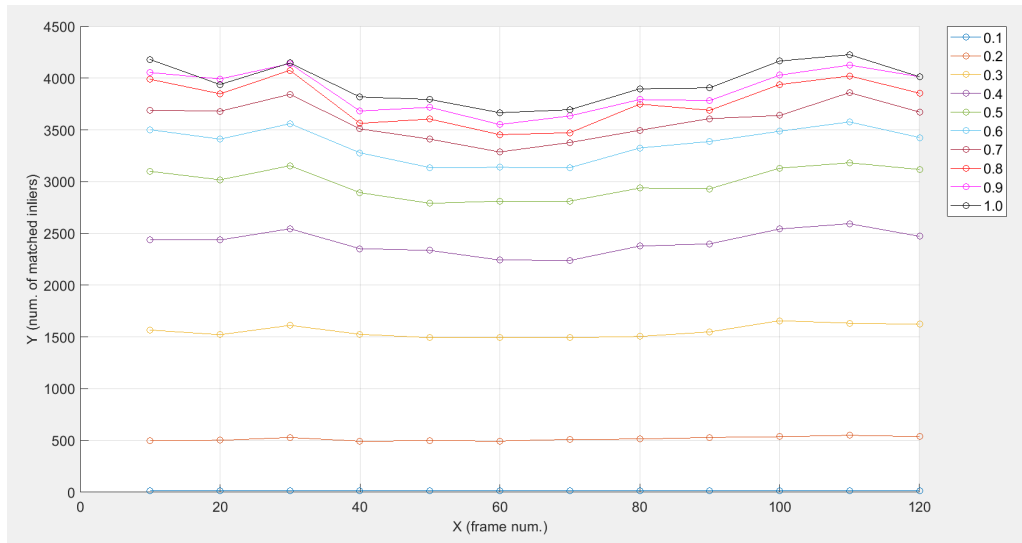


Рис. 5. Кількість точок інтересу, виявлених методом SURF вздовж траєкторії над ділянкою поля з освітленістю $f_k := kf$, $k \in (0, 1]$, де f – функція яскравості у відео, знятому у час з найбільшою освітленістю сцени

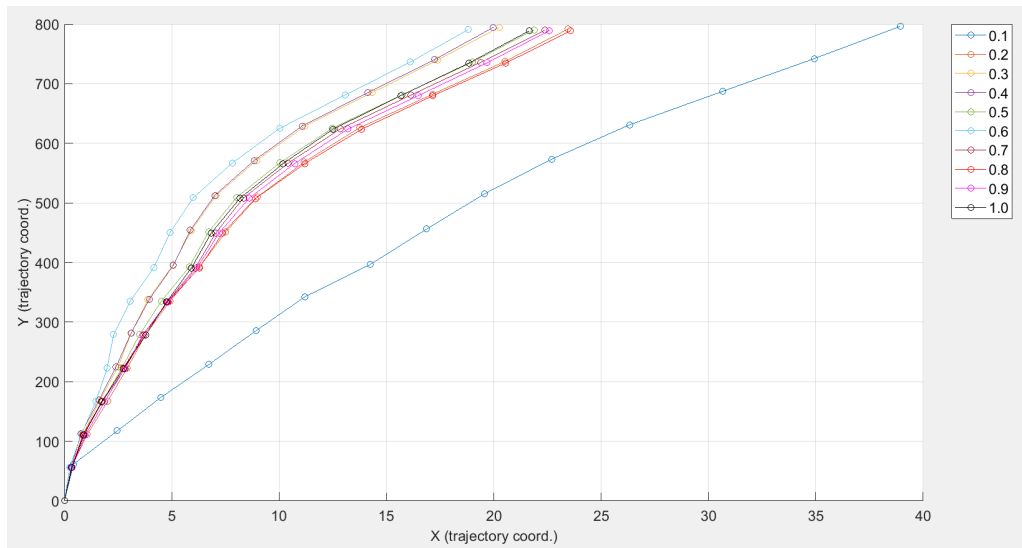


Рис. 6. Фрагменти обчислених траєкторій за різної освітленості $f_k := kf$, $k \in (0, 1]$, де f – функція яскравості у відео, знятому у час з найбільшою освітленістю сцени

ру зайвих точок, в SURF використовують метод RANSAC (RANdom Sample Consensus) [2]. Суть цього методу полягає в ітераційному пошуку потрібних точок шляхом вибору четвірки випадкових точок із зазначених вище множин, які близькі своїми дескрипторами. Результатом роботи методу RANSAC на цьому кроці будуть підмножини, до яких належать точки інтересу суміжних кадрів, що найкраще задовольняють критерій відповідності. Далі з них вибирають точки для побудови матриці афінного перетворення.

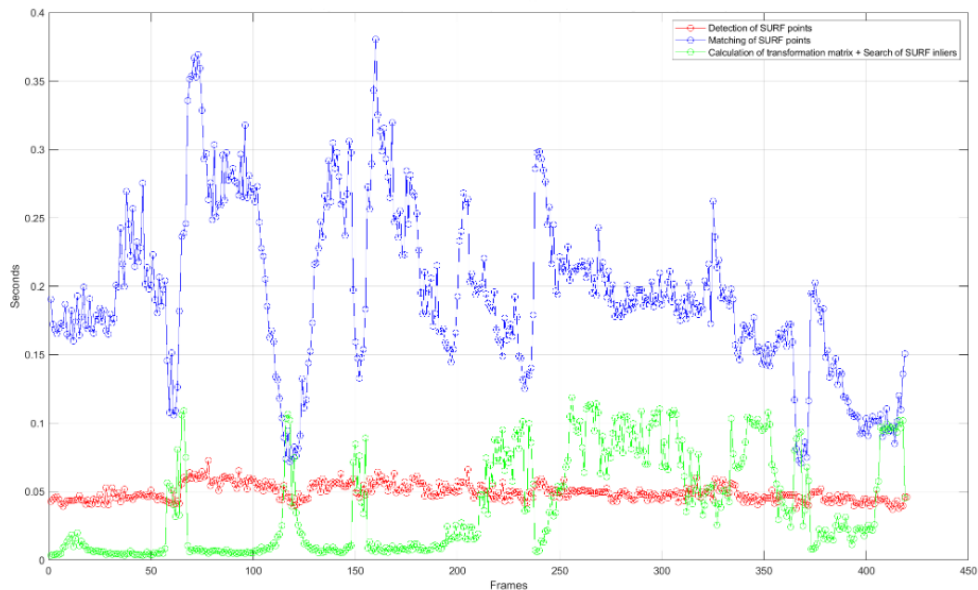


Рис. 7. Час виконання кожного етапу методу SURF з використанням стандартної реалізації методу RANSAC при обчисленні траєкторії

Крім вигляду сцени на значення N_i^{SURF} , також впливає її освітлення. При проведенні експериментів з використання SURF ефект освітлення моделювали шляхом зміни значення яскравості пікселів на кадрах відео за формулою $f_k := kf$, $k \in (0, 1]$. Тут f — функція яскравості у відео, знятому у час з найбільшою освітленістю сцени, тобто вважаємо $f_1 = f$. Графіки на рис. 4 і 5 демонструють, що зі зміною освітленості залежність кількості точок інтересу N_i^{SURF} подібна для різних сцен.

Зазначимо також, що за допомогою методу SURF можна обчислювати траєкторії зі значною зміною освітлення. У цьому випадку можна використовувати ті самі параметри методу, що і у випадку найкращої освітленості. Графіки на рис. 6 демонструють, що обчислені для різної яскравості траєкторії близькі аж до випадку $k = 0.1$, що на практиці моделює фактично повну темряву.

4. ОПТИМІЗАЦІЯ АЛГОРИТМУ ПОШУКУ ТОЧОК ІНТЕРЕСУ

Наведені у попередньому розділі результати застосування методу SURF демонструють перспективність його застосування для реалізації підходу VO щодо обчислення траєкторій UAV з відповідних відеозображень. Заразом очевидно є

надлишковість його операцій, яка зумовлена широким діапазоном значень кількості виявлених точок інтересу, що також є наслідком функціонування детектора SURF зі сталим значенням параметра Threshold при аналізі усіх кадрів відеозображення.

Вплив зазначеного ефекту на втрату продуктивності методу SURF можна зменшити, якщо взяти до уваги характер руху UAV, а саме той факт, що у переважній кількості точок траєкторії рух лінійний. Тоді замість застосування методу RANSAC до усього відеокадру можна перейти до розгляду точок на деякій його частині, яку розбивають на кластери певної форми. Таку модифікацію методу називатимемо кластерний RANSAC.



Рис. 8. Час виконання кожного етапу методу SURF з використанням кластерного методу RANSAC при обчисленні траєкторії

Розглянемо алгоритм методу SURF з використанням кластерного RANSAC. Нехай на n -му кроці SURF маємо два суміжних відеокадри – попередній c_{n-1} і поточний c_n . Вважаємо, що для них сформовані множини точок інтересу і відповідні їм дескриптори. Також відомі параметри афінного перетворення з кадру $n - 2$ на $n - 1$, а саме зміщення (t_u, t_v) початку координат і кут повороту, який на цьому фрагменті траєкторії близький до нуля. Далі кадр c_{n-1} обрізають, відкидаючи від нього прямокутні смуги шириною t_u і t_v , відповідно, на сторонах, протилежних до напрямку руху відеокамери. Після цього отриманий з кадру c_{n-1} прямокутник ділять на 9 однакових прямокутників (називатимемо їх кластерами) і серед них знаходять такий, що містить найменше точок інтересу (за умови, що їх не менше 12). Далі саме на такому кластері і кадрі c_n робимо спробу знайти відповідні точки інтересу методом RANSAC. Якщо спроба виявиться неуспішною, то вибирають інший кластер. Якщо ж усі спроби побудувати матрицю переходу будуть невдалими, то це означає хибність припущення про лінійний рух від c_{n-1} кадру до c_n . У такому випадку застосовуємо звичайний метод RANSAC.

Зауважимо, що використаний тут спосіб розбиття на прямокутні кластери є найпростішим з погляду ресурсозатратності обчислень. Очевидно, що його можна вдосконалювати залежно від припущення стосовно характеру руху UAV, задаючи відповідну геометрію кластерів та їхню кількість.

Звуження множини, з якої метод RANSAC виконує вибір точок інтересу на етапі побудови відповідності, суттєво зменшує час виконання цього етапу на кожному кроці методу SURF. Для порівняння обчислимо траєкторії по тому самому відео з використанням стандартної реалізації методу RANSAC і за допомогою її кластерної модифікації. Графіки часу виконання кожного етапу методу SURF у цих підходах зображено на рис. 7 і 8, відповідно.

Як і очікувалося, найбільшого скорочення (в середньому понад 10 разів) зазнав час на етапі визначення відповідності точок інтересу на сусідніх кадрах відео (на рисунках цьому етапові відповідають графіки синього кольору). Якщо у стандартній реалізації це був найзатратніший етап з найширшим діапазоном зміни часового параметра, то за допомогою використання алгоритму кластеризації він став найшвидшим і з найвужчим діапазоном зміни затрат часу.

5. ВИСНОВКИ

Результати, отримані під час апробації розробленого програмного забезпечення, свідчать про ефективність методу SURF для розв'язування проблем візуальної одометрії, зокрема для керування навігацією UAV над різною місцевістю при різних варіантах освітлення. У цьому випадку запропонована модифікація методу RANSAC – кластерний RANSAC – дає змогу приблизно на порядок швидше будувати відповідність між точками відеокадрів, що дає підстави для виконання необхідних операцій у режимі реального часу.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Bay H. SURF: Speeded Up Robust Features / H. Bay, A. Ess, T. Tuytelaars, L. Van Gool // *Computer Vision and Image Understanding (CVIU)*. – 2008. – Vol. 110, № 3. – P. 346–359.
2. Fischler M. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography / M. Fischler, R. Bolles // *Communications of the ACM*. – 1981. – Vol. 24 (6). – P. 381–395. – doi: 10.1145/358669.358692.
3. Fraundorfer F. Visual Odometry: Part II – Matching, Robustness, and Applications / F. Fraundorfer, D. Scaramuzza // *IEEE Robotics & Automation Magazine*. – 2012. – Vol. 19, № 2. – P. 78–90. – doi: 10.1109/MRA.2012.2182810.
4. Jiang X. A review of multimodal image matching: Methods and applications / Xingyu Jiang, Jiayi Ma, Guobao Xiao, Zhenfeng Shao, Xiaojie Guo // *Information Fusion*. – 2021. – Vol. 73. – P. 22–71. – doi: 10.1016/j.inffus.2021.02.012.
5. Kudriashov A. SLAM Techniques Application for Mobile Robot in Rough Terrain / A. Kudriashov, T. Buratowski, M. Giergiel, P. Malka // *Mechanisms and Machine Science*. – Springer, Cham. – 2020. – Vol. 87. – doi: 10.1007/978-3-030-48981-6.
6. Lindeberg T. Scale Selection Properties of Generalized Scale-Space Interest Point Detectors / T. Lindeberg // *J. Math. Imaging Vis.* – 2013. – Vol. 46. – P. 177–210. – doi: 10.1007/s10851-012-0378-3.
7. MATLAB [Електронний ресурс] – Режим доступу: <https://www.mathworks.com/products/matlab.html>.

8. Mur-Artal R. RB-SLAM: a versatile and accurate monocular SLAM system / R. Mur-Artal, J.-M. Montiel, J.D. Tardos // IEEE Transactions on Robotics. – 2015. – Vol. 31, № 5. – P. 1147–1116.
9. Nister D. Visual odometry / D. Nister, O. Naroditsky, J. Bergen // Proc. Int. Conf. Computer Vision and Pattern Recognition. – 2004. – P. 652–659.
10. Wu X. Deep Learning for Unmanned Aerial Vehicle-Based Object Detection and Tracking: A survey / X. Wu, W. Li, D. Hong, R. Tao, Q. Du // IEEE Geoscience and Remote Sensing Magazine. – 2022. – Vol. 10, № 1. – P. 91–124. – doi: 10.1109/MGRS.2021.3115137.

Стаття: надійшла до редколегії 28.09.2023

доопрацьована 25.10.2023

прийнята до друку 08.11.2023

ON THE USE OF CLUSTER SAMPLING IN VISUAL ODOMETRY ALGORITHMS

S. Ivanov, A. Muzychuk

*Ivan Franko National University of Lviv,
1, Universytetska str., 79000, Lviv, Ukraine*

e-mail: serhii.ivanov@lnu.edu.ua, anatolii.muzychuk@lnu.edu.ua

In this paper, we explore the approach of visual odometry for computing the trajectory of a flying platform based on data from its onboard camera. This approach involves the step-by-step application of the SURF method, as proposed by Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool, to construct an affine transformation that allows mapping a set of points from one frame to another. The selection of interest points on each frame, which are considered when constructing the transformation matrix, is achieved by searching for the maxima of the image pixel brightness using the Hessian matrix. Since such a detector is invariant to affine coordinate transformations, the interest points it defines facilitate an efficient correspondence between frames selected from the video at certain intervals. Ultimately, this enables the estimation of the changing coordinates of the moving platform in a global coordinate system.

The paper presents the results of the software developed for this purpose, demonstrating the applicability of the SURF method for computing trajectories of moving platforms. In particular, it is shown that this approach can be used to construct trajectories that closely resemble those based on GPS data.

Additionally, the influence of the nature of the studied terrain and its lighting conditions on the number of interest points detected by the detector is analyzed. It was noted that in most of the considered cases, their number turned out to be excessive at the stage of building correspondence between adjacent video frames. This manifested itself in the excessive computation time of the trajectory due to the inclusion of a significant number of points that are not needed for estimating the transformation matrix using the RANSAC method.

To optimize the number of required interest points, the authors propose a modification of RANSAC called “clustered RANSAC”. Its essence lies in determining, at every k -th step of SURF, a certain cluster on the previous frame c_{k-1} on which the RANSAC algorithm attempts to construct the necessary affine transformation for mapping to the current frame c_k . Examples of using the software implementation of the SURF method with clustered RANSAC demonstrate a significant reduction in trajectory computation time compared to using the conventional RANSAC method. This improvement paves the way for the application of the modified SURF method in real-time scenarios.

Key words: visual odometry, SURF detector, clustered RANSAC, motion estimation.