

КОМП'ЮТЕРНІ НАУКИ

УДК 519.6

<http://dx.doi.org/10.30970/vam.2024.32.12418>ПРО ЗАСТОСУВАННЯ МЕТОДІВ ВІЗУАЛЬНОЇ
ОДОМЕТРІЇ ДЛЯ АВТОНОМНОЇ НАВІГАЦІЇ
ЛІТАЛЬНИХ АПАРАТІВ

С. Іванов, А. Музичук

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

Розглянуто алгоритми, за допомогою яких розширено застосування методів візуальної одометрії для автономної навігації безпілотного літального апарата під час виконання поворотів і зміни висоти, використовуючи roll- та pitch-нахили. Розглянуто побудову матриць проєктивного перетворення, які використовують відповідні кути нахилів з польотного контролера літального апарата. Відповідно до типу сцени та її освітленості запропоновано алгоритм динамічного вибору параметра порогу Threshold в методі SURF, який дає змогу утримувати кількість точок інтересу у діапазоні, характерному типові місцевості, над якою пролягає маршрут апарата. Запропоновано також алгоритм вибору ключових відеофреймів з відеопотоку, який враховує характеристики ділянок маршруту й оптимізує використання обчислювальних ресурсів бортового комп'ютера. Запропоновані алгоритми протестовано на симуляторі Gazebo. Отримані результати свідчать про ефективність підходу візуальної одометрії за умов відсутності каналу зв'язку з оператором літального апарата, а також без використання сигналу GPS.

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

1. ВСТУП

Візуальна одометрія (VO) [4] є поширеним підходом до виконання автономної навігації безпілотного літального апарата (UAV – Unmanned Aerial Vehicle), який ґрунтується лише на аналізі зображень з відеопотоку його бортової камери. Ключовим компонентом його реалізації є детектор так званих точок інтересу (IP – interest points) на зображеннях місцевості (сцени), над якою пролітає UAV. Одним з ефективних детекторів виявився метод SURF [2], в якому пошук IP виконується шляхом аналізу гаусіана функції, що описує яскравість зображення на відеофреймах. Від часу винаходу цей метод зазнав кілька модифікацій, виконаних з метою його застосування у вигляді відповідної програмної реалізації у навігаційних системах UAV.

Проведені у [1] дослідження стосовно застосування UAV були виконані у припущенні, що апарат рухається паралельно до поверхні землі, тобто на одній висоті, і, крім того, об'єктиві відеокамери спрямований перпендикулярно до земної поверхні. Ці обставини давали змогу підібрати деяке стале значення кроку для вибору фреймів з відеопотоку, за якого фрейми перекривалися і робили успішною процедуру вибору точок інтересу, потрібних для обчислення параметрів афінного перетворення фрейма, який попередньо вибрали з відеопотоку, у поточний. Як свідчать результати

застосування такого підходу на практиці, він надавав також змогу відстежувати повороти UAV.

Зазначимо, що під час польоту на великі відстані сумарна довжина ділянок маршруту із зазначеним характером руху може бути значною, тому запропоновані в [1] рішення забезпечували високу надійність навігації і, крім того, завдяки вдосконаленому методу пошуку IP на зображенні RANSAC [3] вдалося суттєво зменшити час пошуку потрібних ключових точок (ILP – inliers). У підсумку це дало змогу виконати програмну реалізацію описаного підходу, яка виконує задачі навігації UAV у реальному часі.

Однак на певних ділянках маршруту UAV виникає потреба у виконанні маневрів, які виходять за рамки згаданих вище припущень. Зокрема, це pitch- та roll-повороти [7], які супроводжуються зміною висоти польоту та кута нахилу об'єктива відеокамери, а також сповільнення чи зависання UAV над деякою локацією зі зміною висоти. Такі випадки зумовлюють потребу у відповідних алгоритмах, яким і присвячена стаття.

2. ВІЗУАЛЬНА ОДОМЕТРІЯ ЗІ ЗМІННИМ КРОКОМ ВИБОРУ ФРЕЙМІВ

У задачах навігації після детекції IP на новому відеофреймі наступним етапом методу SURF є зіставлення знайдених точок з відповідними точками на попередньому відеофреймі. Завершальним етапом на кожному кроці є визначення зміщення, що відповідає поточній локації відеокамери. Вибір фреймів цифрового зображення з послідовного відеопотоку бортової камери виконують з деяким кроком s . За звичай цей крок є сталим, його величину підбирають експериментально на основі таких характеристик сцени, як її структура й освітлення. Намагаючись досягнути максимальної швидкодії методу, крок прагнуть вибирати якомога більшим s .

Метод SURF добре працює при рівномірному горизонтальному русі UAV над сценами різних видів (див. результати експериментів у [1]). Однак у разі зависання апарата, коли в об'єктив відеокамери потрапляє та сама сцена, її зображення на послідовних фреймах не буде змінюватися. Очевидно, що у цьому випадку вибір фреймів з постійним кроком недоцільний, передусім з огляду на даремну затрату обчислювальних ресурсів бортового комп'ютера. Крім того, як виявилось на практиці, відбувається нагромадження похибки локації, якої можна уникнути, якщо не брати до уваги непотрібні фрейми.

Для вирішення зазначеної проблеми VO алгоритм вибору фреймів з відеопотоку можна модифікувати, використовуючи введення поняття ключових фреймів (key-frames), які визначаються на основі отриманих раніше параметрів руху UAV. На відміну від класичного алгоритму, тепер зміщення обчислюється не стосовно сусіднього (з точністю до фіксованого значення кроку s) фрейму, а відносно відповідного ключового фрейма. В такий спосіб можна пропустити довільну кількість подібних фреймів з відеопотоку й уникнути нагромадження похибки для визначення матриці відповідного афінного перетворення.

3. ДИНАМІЧНЕ ВИЗНАЧЕННЯ ПОРОГУ ГЕССІАНА В МЕТОДІ SURF

Як випливає з практики використання методу SURF, не лише в задачах навігації (див., наприклад, [6]) запорукою ефективної побудови матриці взаємного проєктив-

ного перетворення відеофреймів є знаходження оптимальної кількості ІЛР, для якої надалі будемо використовувати позначення N . Передусім їх має бути достатньо для успішної побудови шуканої матриці перетворення. Але надмірна кількість таких точок суттєво сповільнюватиме роботу SURF в цілому, оскільки під час зіставлення точок здебільшого використовується комбінаторний перебір.

SURF-детектор можна побудувати різними способами. Широкого використання набув детектор, який функціонує шляхом пошуку локальних максимумів гессіана для функції, що є результатом згортки функції яскравості зображення на відеофреймі з гауссіаном (див., наприклад, [1])

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

Функція, яка описує яскравість зображення, дискретна і стала в межах кожного пікселя, тому її згортка з гауссіаном існує і визначає нескінченно диференційовну функцію. За допомогою параметра σ у такій згортці керують процесом згладжування зображення, його також зручно використовувати як параметр масштабу в матриці проєктивного перетворення [8].

Зазначимо, що згладжування зображення приводить до зменшення кількості точок інтересу, і, у підсумку, так можна досягнути зменшення значення параметра N . Заразом у методі SURF для цього передбачено спеціальний параметр Threshold, яким задають порогове значення розглянутого вище гессіана. Точки, в яких гессіан набуває меншого за визначений поріг значення, не беруться до розгляду. Результати застосування на практиці методу SURF (див. [1]) дають змогу виробити рекомендації щодо вибору прийнятного значення порога для різного виду місцевості, над якою виконується політ, а також це значення залежить від освітленості цієї сцени та від роздільної здатності камери. Враховуючи такі характеристики, значення порога вибирається сталим на маршруті. Цей поріг повинен гарантувати достатню кількість ІЛР, тому для польотів над сценою зі змінними характеристиками значення Threshold значно занижене, щоб не відсікти потрібні точки.

Описаний підхід до керування значенням порога зумовлює надмірне використання обчислювальних ресурсів. Зокрема, це буде породжувати довший час обробки відеофрейма, зменшуючи кількість опрацьованих за секунду кадрів, що може зробити алгоритм непридатним на використання на реальному апараті. Тому для забезпечення адаптивності алгоритму до різної місцевості та висоти запропоновано алгоритм динамічного вибору значення порога. Він ґрунтується на вимозі, щоб значення N на досліджуваному відеофреймі перебувало в характерному для сцени діапазоні $[range_{min}, range_{max}]$, інакше для порога рекурентно підбирається нове значення h_{n+1} через попереднє h_n за формулою

$$h_{n+1} = \begin{cases} h_n/k, & N > range_{max}, \\ h_n, & range_{min} \leq N \leq range_{max}, \\ \max(h_n \cdot k, h_*), & N < range_{min}, \end{cases} \quad (2)$$

де величина h_* означає мінімальне значення порога, $k \in (0, 1)$ – характерна для сцени стала, ці величини підбирають експериментально. На початку рекурсії значення порога те саме, що й на попередньому фреймі.

Продемонструємо динамічний вибір значення порога Threshold, використавши симулятор Gazebo [5]. На рис. 1 подано зображення проєкції заданого в симуляторі

маршруту (синій колір) над сценою, характерною для міської забудови. У цьому випадку моделювалася зміна висоти польоту, про що видно з характеру зміни масштабу зображення на відеофреймах (значення масштабу обернено пропорційне висоті польоту).

Отримані графіки демонструють динамічну зміну значення порога, яка забезпечує знаходження значення параметра N у заданому діапазоні, хоч воно і коливається зі зміною фреймів. Зазначимо також, що застосування запропонованого алгоритму динамічного вибору порога Threshold у методі VO дає змогу ефективно обчислювати реальну траєкторію польоту (жовтий колір на рис. 1) зі скороченням часу розрахунків у кілька разів (приблизно у 4 рази для розглянутого прикладу).

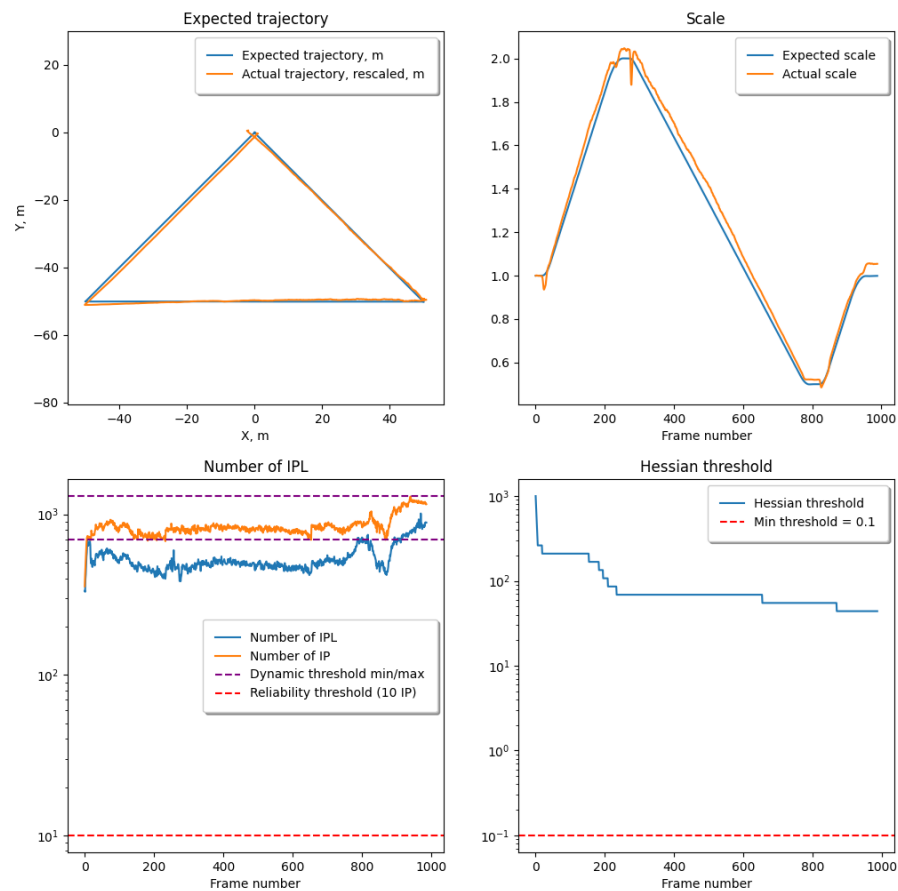


Рис. 1. Тестування алгоритму візуальної одометрії з динамічним визначенням порога Threshold для значень параметрів $range_{min} = 700$, $range_{max} = 1300$, $h_* = 0.1$, $k = 0.8$

Testing the visual odometry algorithm with dynamic Threshold determination for parameter values $range_{min} = 700$, $range_{max} = 1300$, $h_* = 0.1$, $k = 0.8$

4. ПЕРСПЕКТИВНА КОРЕКЦІЯ

Для керування, розгону чи набору висоти літальні апарати використовують нахили, які можна описати так званими pitch- та roll-кутами (див., наприклад, [7]). Внаслідок цього зображення, яке фіксуватиме камера, відрізнятиметься від зображення, яке формувала би камера з об'єктивом, спрямованим вертикально вниз. Щоб компенсувати зазначену зміну перспективи, описаний у [1] підхід VO потребує певного препроцесингу. З цією метою будемо розглядати віртуальну відеокамеру зі спрямованим вниз об'єктивом і перетворювати зображення, отримане реальною камерою, до відповідного віртуального, яке називатимемо відкорегованим, а сам процес такого перетворення – перспективною корекцією.

Для виконання перспективної корекції будемо використовувати кути нахилу, які можна отримати безпосередньо з польотного контролера UAV. Введемо позначення α і β для нахилів pitch та roll, відповідно. Вважаємо, що відеофрейм, на якому фіксується зображення, має розмір $w \times h$, і також позначимо f_x, f_y фокусні відстані камери в пікселях у напрямку осей x та y , відповідно. Тоді згідно з моделлю відеокамери (див., наприклад, [6, Chapt.6]) матриця перспективного перетворення, яка проектує реальне зображення об'єктів у їхнє зображення у віртуальній відеокамері, набуде вигляду

$$H = K_1 R^{-1} K_2^{-1}, \quad (3)$$

де $K_1, K_2 \in \mathbb{R}^{3 \times 3}$ – калібрувальна матриця віртуальної камери та її модифікована матриця, яка враховує зміщення оптичного центра між камерами у пінхоул моделі при застосуванні гомографії, відповідно

$$K_1 = \begin{pmatrix} f_x & 0 & w/2 \\ 0 & f_y & h/2 \\ 0 & 0 & 1 \end{pmatrix}, \quad K_2 = \begin{pmatrix} f_x & 0 & \theta_x \\ 0 & f_y & \theta_y \\ 0 & 0 & 1 \end{pmatrix}. \quad (4)$$

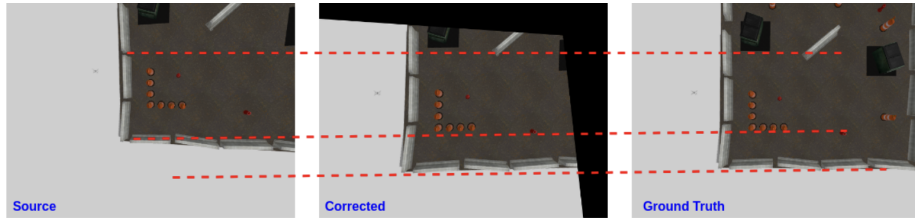


Рис. 2. Реальне зображення (source) з відеофрейма UAV, кореговане зображення (corrected) та зображення з віртуальної камери (ground truth) у тій самій точці локації, але з вертикально спрямованим об'єктивом. Червоним пунктиром позначені характерні риси, які збігаються у віртуальному та відкорегованому зображеннях

Real image (source) from UAV video frame, corrected image (corrected) and image from virtual camera (ground truth) at the same location point, but with the lens pointed vertically. The red dotted line indicates the characteristic features that coincide in the virtual and corrected images

Тут $\theta_x = w/2 - f_x \sin(\alpha)$, $\theta_y = h/2 - f_y \sin(\beta)$ – зміщення відтвореного зображення з врахуванням нахилу камери. Матрицю повороту R отримуємо як добуток $R = R_x R_y$ матриць повороту за кожною віссю, які визначені через кути нахилу

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix}, \quad R_y = \begin{pmatrix} \cos(\beta) & 0 & -\sin(\beta) \\ 0 & 1 & 0 \\ \sin(\beta) & 0 & \cos(\beta) \end{pmatrix}. \quad (5)$$

Описану послідовність перетворень демонструють зображення відеокадрів на рис. 2. Отримане після виконання перспективної корекції зображення добре узгоджене з зображенням, яке формувала б у тій самій точці спостереження віртуальна камера з вертикально спрямованим об'єктивом. Маючи такі відкореговані зображення, можемо застосовувати до них описаний алгоритм VO.

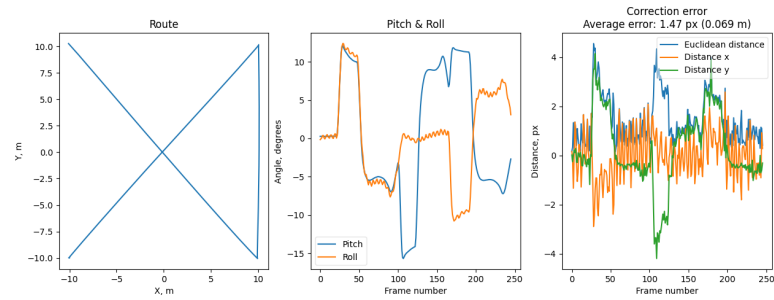


Рис. 3. Схема маршруту та кути нахилу під час польоту UAV (перші два графіки) і похибки обчислення координат локалізації з використанням алгоритму перспективної корекції.

Початкова висота польоту $H = 15$ метрів

Route diagram and inclination angles during UAV flight (first two graphs) and errors in calculating localization coordinates using the perspective correction algorithm.

Initial flight altitude $H = 15$ meters



Рис. 4. Частина карти з Google Maps місцевості, над якою заплановано політ UAV. Червоним кольором позначено межі кількох ділянок сцени, які попадали на відеофрейми під час польоту

A portion of a Google Maps of the area over which the UAV is scheduled to fly. The boundaries of several areas of the scene that were captured in the video frames during the flight are marked in red

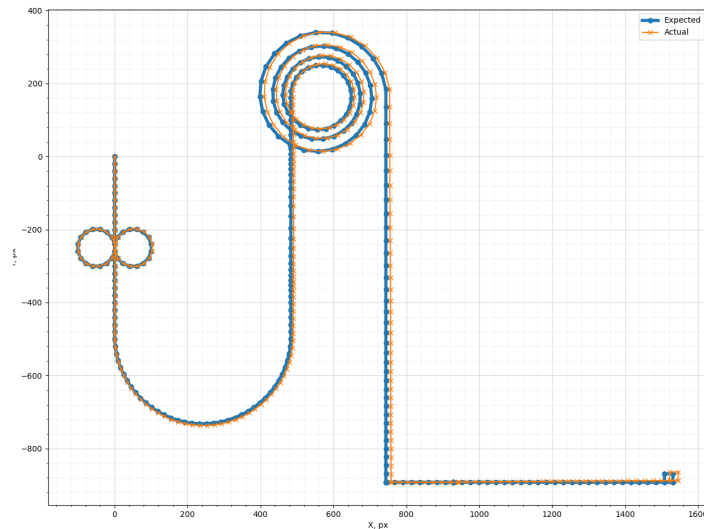


Рис. 5. Схема заданого маршруту (синій колір, Expected) польоту UAV за картою на рис. 4 і графік обчисленої траєкторії (жовтий колір, Actual).

Спіралеподібна ділянка відповідає наборові висоти з використанням
The scheme of the given route (blue color, Expected) of the UAV flight according
to the map in Fig. 4 and the graph of the calculated trajectory (yellow color,
Actual). The spiral section corresponds to the climb using

Для детальнішого аналізу похибки, яку вносить запропонований алгоритм перспективної корекції, було проведено моделювання на симуляторі Gazebo [5] польоту UAV, схему заданого маршруту з відповідними нахилами під час польоту подано на рис. 3. Як бачимо з графіків, відхилення на кожному відеофреймі обчислених координат центра камери (correction error) від очікуваних значень (що тактувалися як ground truth у попередньому прикладі) не перевищує кількох пікселів, а середнє відхилення становить 1.47 пікселя, що з врахуванням внутрішніх параметрів відеокамери та початкової висоти польоту відповідає 7 сантиметрам, що цілком допустимо для таких задач. Після перетворення відкореговане зображення можна використовувати у методах VO.

Розглянемо ще один приклад застосування методів VO для моделювання автономної навігації UAV за картою реальної місцевості, взятої з Google Maps. Її частина зображена на рис. 4 і є характерною для сцени з міською забудовою. Схема заданого в симуляторі Gazebo маршруту зображена на рис. 5. Крім прямолінійного руху, цей маршрут передбачає виконання різних поворотів, а також набір висоти на подібній до спіралі ділянці. На рис. 4 червоними прямокутниками позначено частини сцени, які попадали на кілька послідовних відеофреймів, зміна їхніх розмірів відповідає зміні масштабу зображення з набором висоти.

5. ВИСНОВКИ

Розроблені алгоритми дають змогу поширити застосування методів візуальної одометрії до задач автономної навігації літальних апаратів під час виконання ма-

неврів, які виконують за допомогою зміни кутів roll- та pitch-нахилів, що надаються польотним контролером. Запропонований підхід до побудови матриць проєктивного перетворення дає змогу уникати складного перетворення координат під час виконання складних маневрів і зводити проблему навігації до застосування класичних методів візуальної одометрії.

Відповідно до типу сцени та її освітленості запропоновано використовувати алгоритм динамічного вибору параметра порога Threshold в методі SURF, який дає змогу утримувати кількість точок інтересу у діапазоні, характерному типові місцевості. На ділянках маршруту, де літальні апарати сповільнюють свій горизонтальний рух, або зависають, вибір зображень з відеопотоку доцільно виконувати з ключових відеофреймів.

Усі розроблені алгоритми оптимізовані щодо використання обчислювальних ресурсів бортового комп'ютера. Отримані результати апробації цих алгоритмів свідчать про ефективність їхнього використання у візуальній одометрії за умов відсутності каналу зв'язку з оператором літального апарата, а також без використання сигналу GPS.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Іванов С. Про використання кластерної вибірки в алгоритмах візуальної одометрії / С. Іванов, А. Муzychuk // Вісник ЛНУ. Серія прикл. матем. та інформ. Вип. 31. – 2023. – С. 45–55. – <http://dx.doi.org/10.30970/vam.2021.29.11337>
2. 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.
3. 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. – <https://doi.org/10.1145/358669.358692>.
4. 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. – <https://doi.org/10.1109/MRA.2012.2182810>.
5. Gazebo: Features and Benefits. – <https://gazebo-sim.org/features>.
6. Hartley R. Multiple View Geometry in Computer Vision / R. Hartley, A. Zisserman Cambridge University Press. – 2004.
7. How Things Fly. Pitch, Roll and Yaw. – <https://howthingsfly.si.edu/media/roll-pitch-yaw>.
8. Lindeberg T. Scale Selection Properties of Generalized Scale-Space Interest Point Detectors // J. Math. Imaging Vis. – 2013. – Vol. 46. – P. 177–210. – <https://doi.org/10.1007/s10851-012-0378-3>.

Стаття: надійшла до редколегії 17.10.2024

доопрацьована 25.10.2024

прийнята до друку 14.11.2024

ON THE APPLICATION OF VISUAL ODOMETRY METHODS FOR AUTONOMOUS NAVIGATION OF AIRCRAFTS

S. Ivanov, A. Muzychuk

*Ivan Franko National University of Lviv,
1, Universytetska str., Lviv, 79000, Ukraine,
e-mail: serhii.ivanov@lnu.edu.ua, anatoliy.muzychuk@lnu.edu.ua*

The paper considers algorithms that extend the use of visual odometry methods for autonomous navigation of an unmanned aerial vehicle when performing turns and altitude changes due to roll and pitch angles. The construction of projective transformation matrices using the corresponding tilt angles from the flight controller of the aircraft is presented. In accordance with the type of scene and its illumination, an algorithm for dynamic selection of the Threshold parameter in the SURF method is proposed, which allows keeping the number of points of interest in the range characteristic of the type of area over which the vehicle route runs. We also propose an algorithm for selecting key video frames from a video stream that takes into account the characteristics of the route sections and optimises the use of on-board computer computing resources. The proposed algorithms are tested on the Gazebo simulator. The obtained results demonstrate the effectiveness of the visual odometry approach in the absence of a communication channel with the aircraft operator, as well as without using the GPS signal.

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