DOI 10.34185/1562-9945-2-157-2025-01
UDC 004.8
S.V. Mashtalir, D.P. Lendel

DRONE VIDEO PROCESSING BY FRAGMENT ANALYSIS

Annotation. This study focused on analyzing drone video to address the inherent complexity of processing videos captured with a moving camera. Each frame is divided into smaller fragments through a grid-based segmentation technique, enabling a localized and detailed motion analysis. Singular Value Decomposition (SVD) is applied to compute the Ki Fan norm for each fragment, enabling the detection of dynamic changes between consecutive frames. The fragment-level analysis allows the algorithm to robustly identify regions of interest and differentiate between global motion (camera movement) and local motion (object movement) despite challenges introduced by camera instability. The motion state is classified into four categories: stable camera with no object movement, stable camera with object movement, moving camera with no object movement, and moving camera with object movement. This fragment-based methodology enhances precision in dynamic scene analysis, offering a scalable and efficient solution for applications such as video stabilization, object tracking, and real-time motion detection in complex environments.
Key words Video stream fragmentation; Ky Fan norm; Singular value decomposition; Drones, Object detection; Moving camera; Data Analysis, Video processing

Statement of the problem. Drones, also known as Unmanned Aerial Vehicles (UAVs), have become invaluable tools in various fields, including agriculture, ensuring the safety of construction workers, environmental monitoring, disaster response, surveillance [1,2], and military applications. Their true potential lies in their ability to analyze their surroundings and make intelligent decisions. This is made possible by integrating computer vision algorithms into drones, enabling them to perform various tasks based on real-time camera feeds [3,4]. A lot of drone frameworks presented in the last researches [5,6,7]

Many of these algorithms inherently rely on a moving camera [8]. For instance, in intelligent video surveillance systems, camera movement techniques such as pan-tilt-zoom (PTZ) enhance focus and track targets [9]. Recent advancements in drone technology, particularly the availability of relatively inexpensive drones with advanced imaging capabilities, promise significant commercial applications in the future [10]. These drones can operate with varying levels of movement and autonomy.

Additionally, the progress in smartphone camera technology has fueled interest among consumers in capturing video sequences capable of detecting and tracking moving objects [11]. These cameras often operate with unrestricted movement. Even in outdoor scenes

© Mashtalir S.V., Lendel D.P., 2025

captured by a fixed camera, the environment's dynamic nature may cause the camera to jitter or shake, preventing the camera from being considered entirely stationary [12].

As the use of moving cameras becomes more widespread and interest in detecting moving objects grows, developing robust methods for moving object detection in dynamic environments becomes increasingly essential.

Analysis of the latest research and publications. In the case of a fixed camera, the changes between consecutive frames are typically caused by moving objects. However, not all these changes relate to the objects of interest or the intended application. For indoor scenes, even in controlled environments, unwanted factors like shadows or changes in lighting conditions can interfere with the detection of moving objects [13]. In outdoor scenes, where the environment is usually uncontrollable, issues such as branch movements, cloud motion, and illumination variations can pose significant challenges for moving object detection [14].

Previous research [15,16] has addressed the problem of detecting moving objects in video sequences captured by fixed cameras, with or without such disturbances. The key idea involves building a stable background model and applying a background subtraction technique, where the current frame is subtracted from the background to detect moving objects.

For moving cameras, the detection method has to address the challenges faced by fixed cameras and account for difficulties related to compensating for camera motion. A simple background subtraction with a basic motion compensation model is insufficient for moving cameras. Inaccuracies in motion compensation, likely in free camera movements, lead to failures in building a reliable background model that distinguishes between background and foreground pixels.

One approach to detecting moving objects with a moving camera is to differentiate the motion caused by the camera from that of the moving objects. Solutions generally fall into two categories: background modeling and trajectory classification. Background modeling [17] uses motion compensation to construct a suitable background for each frame. At the same time, trajectory classification [19,18] calculates long-term trajectories for feature points using a tracker and applies clustering techniques to distinguish object-related trajectories from background movement.

Another strategy is to extend background subtraction methods based on low-rank and sparse matrix decomposition, initially developed for static cameras [20,21], to handle moving cameras [22]. The core idea is that coherent information across image frames is captured in a low-rank representation, while outliers—such as moving objects—appear in the sparse representation. However, this assumption relies on a static background across frames. To adapt this approach to moving cameras, a transformation can be incorporated to compensate for background motion, such as affine transformations for pan-tilt-zoom motions or perspective transformations for free camera movements [23].

Object tracking can also serve as a method for detecting moving objects, although its primary goal differs. In tracking, a target object is identified in the initial frame, and its location is updated in subsequent frames. This involves extracting features like color, texture, or

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

statistical information from the target and using a similarity model or classifier to find the best match in the next frame. The target's characteristics are then updated for continued tracking.

An alternative approach is fragment processing. The fragment-based analysis offers a promising approach to address these challenges by dividing video frames into smaller, meaningful regions or fragments for localized and robust object detection. Fragment processing is a broad concept. The approach defines fragments as groups of pixels with similar color or texture properties, simplifying the image into coherent regions for analysis [24]. Or the fragment can be the video frame geometry part. Fragments are defined as video frame geometric parts represented as matrices with arbitrary dimensions. The method significantly reduces computational costs by skipping traditional transformation steps like matrix vectorization. Reducing the load of the drone processor allows not only to optimize data processing but also to reduce the load on the power system. In the research [25], SVD (Singular value decomposition) of the matrix and the Ky Fan norm are proposed for scene change analysis. In the context of motion detection, this approach was expanded [26]. Dividing the frame into 5x5 or 10x10 allowed identify the fragments in which motion occurred. An analysis of the effectiveness of the obtained descriptor across various video data sizes demonstrates that changes in the descriptor for each fragment are independent of both the video resolution and aspect ratio. [27]. The rectangle frame has been transformed into a square matrix by SVD, where each element is a Ky Fan norm value used as an object detection descriptor. The fragments number is changing the analysis scale. In research [28], dividing the frame into 100x100 detected contour moving objects.

The Singular Value Decomposition (SVD) for m×n matrix A is a factorization of the form

$$A = USV^* \tag{1}$$

where U is an m × m complex unitary matrix, S is an m×n diagonal matrix with non-negative real numbers on the diagonal, and V is an n×n complex unitary matrix. If A is real, U and V can be guaranteed to be also real orthogonal matrices.

The SVD is closely associated with several common matrix norms and offers an efficient method for their computation. This relationship is derived from the foundational existence proof for the SVD, which establishes its applicability

$$\|A\|_2 = (\rho\,(\mathrm{A} * \mathrm{A}))^{1/2} = \sigma_1(\mathrm{A}) \tag{2}$$

It follows from our existence the sum first k singular values:

$$\|A\|_k^{KF} = \sigma_1(\mathrm{A}) + .. + \sigma_k(\mathrm{A}) \tag{3}$$

is a matrix norm, called the Ky Fan k-norm. In our approach, the Ki Fan norm is a fragment descriptor. There is no need to do a full SVD transformation to obtain the norm value. It is enough to obtain a matrix of singular values. The numpy library function linalg.svdvals returns the singular values of a matrix (or a stack of matrices) x. When x is a stack of matrices, the function will compute the singular values for each matrix in the stack.

Objective. The purpose of this study is to develop an efficient and robust algorithm for detecting and differentiating between camera motion and object movement within video frames. By leveraging grid-based frame segmentation and Singular Value Decomposition

(SVD) to compute Ki Fan norms, the algorithm identifies dynamic fragments and classifies global (camera) and local (object) motion. This enables accurate motion state assessment, offering insights into the stability of the camera and the activity of objects within the scene.

Presentation of the main material of the research. In this section, we will consider the results produced by the developed application. Our experiment used a drone video camera. Codec is h264, frame size is 1920.x1080, frame rate is 30. The camera position is unstable. To visualize the results of utilizing the Ky Fan norm for video analysis, a Python 3.10.11 application was developed and executed on a system equipped with an Intel Core i5 processor, 16 GB of RAM, and running the Windows operating system. The application relies on two open-source libraries licensed under Apache License: OpenCV version 4.10.0 and NumPy version 2.2.1.

We treat the video as a sequence of frames (Fig. 1). Each frame is converted from RGB to a grayscale model, and a Gaussian blur filter is applied so that the value of each pixel carries only intensity information.

Figure 1 - Video source as a sequence of frames. Source: compiled by the authors

Each frame is divided into G = (n × m) rectangular fragments. For each fragment g_{ij}, its spatial coordinates are determined, and the intensity matrix of its pixels is extracted. The obtained matrix fragment, having an appropriate size for SVD transformation, undergoes singular value decomposition, allowing the calculation of singular values. The Ky Fan norm is then determined for each fragment. The rectangle frame has been transformed into a square matrix 5x5 by SVD where each element is Ky Fan norm value as a descriptor. The difference in norms between the two frames is calculated as:

$$\Delta_{ij} = \left|\sigma_{ij}^{t+1} - \sigma_{ij}^{t}\right| \quad (4)$$

A fragment g_{ij} is marked as dynamic if:

$$\Delta_{ij} > T_{motion} \quad (5)$$

The global mean change across all fragments is calculated as

$$\Delta_{global} = (1/|G|) \sum \Delta_{ij} \quad (6)$$

If more than 80% of fragments have

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

$$\Delta_{ij} < T_{motion}/2 \tag{7}$$

the camera is considered stable:

$$\frac{\left|g_{ij} : \Delta_{ij} < \frac{T_{motion}}{2}\right|}{|G|} > 0.8 \tag{8}$$

Local movement is detected if more than 20% of fragments have

$$\Delta_{ij} > T_{motion} \tag{9}$$

$$\frac{\left|g_{ij} : \Delta_{ij} > T_{motion}\right|}{|G|} > 0.2 \tag{10}$$

Based on the global and local motion analyses, the algorithm determines one of four possible states:

- Stable Camera, No Object Movement
- Stable Camera, Object Moving
- Moving Camera, No Object Movement
- Moving Camera, Object Moving

We selected fragment (2,3) for the result demonstration in details. 2 is Y, 3 is X indexes of the 5x5 matrix. Fragment (2,3) is marked by a white spot (Fig. 2). The camera and object state are on the top side. Fragment with moving object marked by white spot. On frame number 29, the Ky Fan norm value exceeds the threshold, and moving objects have been detected. This zoomed-in block can illustrate the result better.

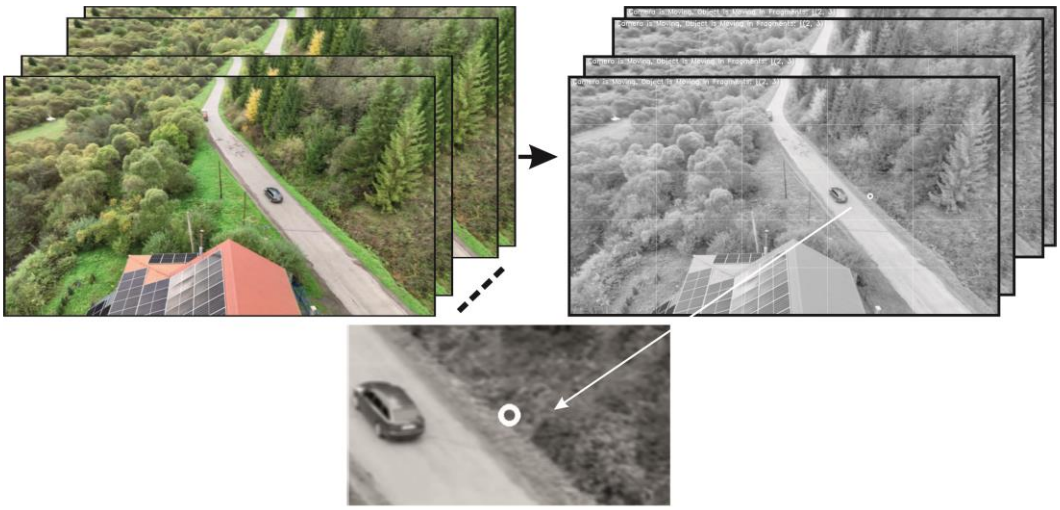
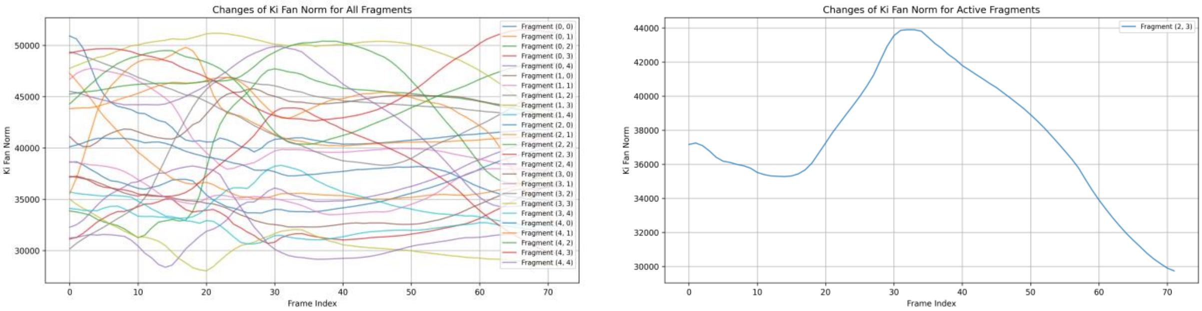


Figure 2 - The result of frame-by-frame processing is a new video source in a grayscale model. The camera and object state are on the top side. Fragment with moving object marked by white spot. Source: compiled by the authors

The drone's flight was smooth, with brief hangs. The camera is considered stable if the deviation is less than the threshold value for more than 80% of the fragments. Ky-Fan norm fluctuations for all fragments are presented in Fig.3 A. Local motion was defined for fragments with significant changes. Fig.3 B presents Ky-Form fluctuation for the fragment (2,3). On frame number 29, the Ky Fan norm value exceeds the threshold, and moving objects have been detected. A comparison of the norm fluctuations in all fragments shows a significant deviation for the fragment (2,3). This indicates the local movement of the object in the fragment. The smooth movement of the drone leads to uniform changes in the normal fluctuations of Ky Fan in all fragments, which can be labeled as a "moving camera." A significant deviation of the norm fluctuations in the segment indicates the local movement of the object.



A B

Figure 3 - Ky-Fan norm fluctuation for all fragments A. Ky-Fan norm fluctuation for fragment (2,3). B presents Ky-Form fluctuation for the fragment (2,3). On frame number 29, the Ky Fan norm value exceeds the threshold, and moving objects have been detected. Y is Ky Fan norm value, X is frame numbers. Source: compiled by the authors

Conclusions. The study introduces an approach for processing drone video footage, analyzing dynamic changes in frames captured by a moving camera. The method achieves localized motion analysis by segmenting video frames into smaller fragments through a grid-based technique. Singular Value Decomposition (SVD) is employed to compute the Ki Fan norm for each fragment, enabling the detection of both global (camera) and local (object) motion. The methodology aims to support applications such as video stabilization, object tracking, and dynamic scene analysis.

LITERATURE / REFERENCES

1. Madhavan, R., Silva, T., Farina, F., Wiebbelling, R., Renner, L., & Prestes, E. (2018). Unmanned aerial vehicles for environmental monitoring, ecological conservation, and disaster management. In *Technologies for Development: From Innovation to Social Impact* (pp. 31-39). Springer International Publishing. DOI: 10.1007/978-3-319-91068-0_3
2. Mohsan, S. A. H., Othman, N. Q. H., Li, Y., Alsharif, M. H., & Khan, M. A. (2023). Unmanned aerial vehicles (UAVs): Practical aspects, applications, open challenges, security issues, and future trends. *Intelligent Service Robotics*, *16*(1), 109-137. DOI: 10.1007/s11370-022-00452-4
3. Akbari, Y., Almaadeed, N., Al-Maadeed, S., & Elharrouss, O. (2021). Applications, databases and open computer vision research from drone videos and images: a survey. *Artificial Intelligence Review*, *54*, 3887-3938. DOI: 10.1007/s10462-020-09943-1

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

4. Arafat, M.Y., Alam, M.M., & Moh, S. (2023). Vision-based navigation techniques for unmanned aerial vehicles: Review and challenges. *Drones*, *7*(2), 89. DOI: 10.3390/drones7020089
5. Heakl, A., Youssef, F., Parque, V., & Gomaa, W. (2024). DroneVis: Versatile Computer Vision Library for Drones. *arXiv preprint arXiv:2406.00447*. DOI: 10.48550/arXiv.2406.00447
6. Rohan, A., Rabah, M., & Kim, S. H. (2019). Convolutional neural network-based real-time object detection and tracking for parrot AR drone 2. *IEEE access*, *7*, 69575-69584. doi: 10.1109/ACCESS.2019.2919332
7. Han, S., Shen, W., & Liu, Z. (2016). Deep drone: Object detection and tracking for smart drones on embedded system. *URL https://web. stanford. edu/class/cs231a/prev_projects_2016/deepdrone-object__2_. pdf*.
8. Yazdi, M., & Bouwmans, T. (2018). New trends on moving object detection in video images captured by a moving camera: A survey. *Computer science review*, *28*, 157-177. DOI: 10.1016/j.cosrev.2018.03.001
9. Yang, J., Xie, X., & Wang, Y. (2017, April). Design of video surveillance and tracking system based on attitude and heading reference system and PTZ camera. In *2017 5th International Conference on Computer-Aided Design, Manufacturing, Modeling and Simulation (CDMMS 2017)* (Vol. 1834, No. 1, p. 040016). DOI: 10.1063/1.4981612
10. Chen, P., Dang, Y., Liang, R., Zhu, W., & He, X. (2017). Real-time object tracking on a drone with multi-inertial sensing data. *IEEE Transactions on Intelligent Transportation Systems*, *19*(1), 131-139. doi: 10.1109/TITS.2017.2750091
11. Dames, P., Tokekar, P., & Kumar, V. (2017). Detecting, localizing, and tracking an unknown number of moving targets using a team of mobile robots. *The International Journal of Robotics Research*, *36*(13-14), 1540-1553. DOI: 10.1177/0278364917709507
13. Leal-Taixé, L., Milan, A., Schindler, K., Cremers, D., Reid, I., & Roth, S. (2017). Tracking the trackers: an analysis of the state of the art in multiple object tracking. *arXiv preprint arXiv:1704.02781*. DOI: 10.48550/arXiv.1704.02781
14. Risse, B., Mangan, M., Del Pero, L., & Webb, B. (2017). Visual tracking of small animals in cluttered natural environments using a freely moving camera. In *Proceedings of the IEEE international conference on computer vision workshops* (pp. 2840-2849). https://openaccess.thecvf.com/content_ICCV_2017_workshops/papers/w41/Risse_Visual_Tracking_of_ICCV_2017_paper.pdf
15. Bouwmans, T. (2014). Traditional and recent approaches in background modeling for foreground detection: An overview. *Computer science review*, *11*, 31-66. DOI: 10.1016/j.cosrev.2014.04.001
16. Yilmaz, A., Javed, O., & Shah, M. (2006). Object tracking: A survey. *Acm computing surveys (CSUR)*, *38*(4), 13-es. DOI: 10.1145/1177352.1177355
17. Wren, C. R., Azarbayejani, A., Darrell, T., & Pentland, A. P. (1997). Pfinder: Real-time tracking of the human body. *IEEE Transactions on pattern analysis and machine intelligence*, *19*(7), 780-785. doi: 10.1109/34.598236

18. Zhai, W., Xiong, X., Mo, G., Xiao, Y., Wu, C., Xu, Z., & Pan, J. (2024). A Bagging-SVM field-road trajectory classification model based on feature enhancement. *Computers and Electronics in Agriculture*, *217*, 108635. DOI: 10.1016/j.compag.2024.108635
19. Zhai, W., Xu, Z., Pan, J., Guo, Z., & Wu, C. (2024). A general image classification model for agricultural machinery trajectory mode recognition. *Computers and Electronics in Agriculture*, *227*, 109629. DOI: 10.1016/j.compag.2024.109629
20. Bouwmans, T., Sobral, A., Javed, S., Jung, S. K., & Zahzah, E. H. (2017). Decomposition into low-rank plus additive matrices for background/foreground separation: A review for a comparative evaluation with a large-scale dataset. *Computer Science Review*, *23*, 1-71. DOI: 10.1016/j.cosrev.2016.11.001
21. T. Bouwmans, T., Aybat, N. S., & Zahzah, E. H. (Eds.). (2016). *Handbook of robust low-rank and sparse matrix decomposition: Applications in image and video processing*. CRC Press. https://dl.acm.org/doi/abs/10.5555/2994445
22. Ebadi, S. E., Ones, V. G., & Izquierdo, E. (2015, September). Efficient background subtraction with low-rank and sparse matrix decomposition. In *2015 IEEE International Conference on Image Processing (ICIP)* (pp. 4863-4867). IEEE. doi: 10.1109/ICIP.2015.7351731.
23. Wu, Y., He, X., & Nguyen, T. Q. (2015). Moving object detection with a freely moving camera via background motion subtraction. *IEEE Transactions on Circuits and Systems for Video Technology*, *27*(2), 236-248. doi: 10.1109/TCSVT.2015.2493499
24. Tomasi, C. (2012). Histograms of oriented gradients. *Computer Vision Sampler*, 1-6.
25 Koliada, M. “Ky fan norm application for video segmentation”. *Herald of Advanced Information Technology*, 2020;*1*(3), 345-351. DOI: 10.15276/hait.01.2020.1
26 Mashtalir, S., & Lendel, D. “Video pre-motion detection by fragment processing”. CEUR Workshop ProceedingsVolume. 2024; 3790, pp. 342 – 351. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85207837522&partnerID=40&md5=26529ee16efb141face273bd943660d5
27. Mashtalir, S. V. & Lendel D. P. “Video fragment processing by Ky Fan norm”, Appl. Asp. Inf. Technol. 2024; 7.1: 59 68. DOI: DOI: 10.15276/aait.07.2024.5
28. Sergii V. Mashtalir, Dmytro P. Lendel (2024). Moving object shape detection by fragment processing. Herald of Advanced Information Technology. 7(4), 414-423. DOI: 10.15276/hait.07.2024.30

Received 26.02.2025.
Accepted 28.02.2025.

Обробка відео з дронів шляхом фрагментного аналізу

В цьому дослідженні ми зосередилися на аналізі відео з дронів для вирішення задачі обробки фрагментів відео, знятих рухомою камерою. Кожен кадр розділений на менші фрагменти за допомогою методу сегментації на основі сітки, що дає змогу локалізувати та детально аналізувати рух. Сингулярне розкладання (SVD) застосовується для обчислення норми Кі Фан для кожного фрагменту, що дозволяє виявити динамічні зміни між послідовними кадрами. Аналіз на рівні фрагментів дозволяє алгоритму надійно визначати області інтересу та розрізняти глобальний рух (рух камери) і локальний рух (рух об’єкта), незважаючи на проблеми, пов’язані з нестабільністю

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

камери. Стан руху класифікується на чотири категорії: стабільна камера без руху об'єкта, стабільна камера з рухом об'єкта, рухома камера без руху об'єкта та рухома камера з рухом об'єкта. Ця методологія на основі фрагментів підвищує точність динамічного аналізу сцени, пропонуючи масштабоване та ефективне рішення для таких підходів, як стабілізація відео, відстеження об'єктів і виявлення руху в реальному часі в складних середовищах.

Машталір Сергій Володимирович – доктор технічних наук, професор, професор кафедри інформатики Харківського національного університету радіоелектроніки.
Лендьел Дмитро Павлович – аспірант ІУСТ Ужгородського національного університету.

Mashtalir Sergii Volodymyrovych– Doctor of Engineering Science, professor, professor of informatics department Kharkiv National University of Radio Electronics, ORCID: https://orcid.org/0000-0002-0917-6622, Scopus Author ID: 36183980100.
Lendel Dmytro Pavlovich – PhD student in Uzhhorod National University, ORCID: https://orcid.org/0000-0003-3971-1945, Scopus Author ID: 59390876900.