AN EVALUATION OF VARIOUS VIDEO PROCESSING TECHNIQUES FOR THE STABILIZATION OF DRONE VIDEOS

Stephanie Cahail

The University of Texas at Austin

ABSTRACT

The use of Unmanned Aerial Vehicles (UAVs), more commonly known as drones, has become increasingly popular for capturing high-quality aerial videos for both professional and recreational purposes. However, many factors can have a negative impact on drone video quality, such as outdoor conditions and human error. This study explores and evaluates a variety of digital video processing methods to increase video stabilization for drone videos. The results of this study indicate that these methods are also useful for stabilizing drone videos.

***Index Terms*—** Video Stabilization, Video Processing, Unmanned Aerial Vehicles, Optical Flow, Computer Vision

1. INTRODUCTION

The ability for drones to capture images and videos has been a recent development over the past ten years, leading to increased popularity of drone use for both professional and recreational purposes. Because of the recent interest in drones, they have evolved to be more accessible and affordable to a variety of consumers. Now, drones are commonly used to capture aerial images and videos for a variety of uses, from real estate to cinematography. Although drones have come a long way in terms of accessibility, there are still many challenges that can prevent pilots from capturing stable, high-quality footage. Some of these challenges are environmental conditions, such as wind during flight, and human error from navigating the drone. Many studies have evaluated image and video processing techniques that are most optimal for stabilization. This experiment explores how some of these techniques perform when applied to drone videos. The following sections include details about related works that have influenced this study, the video processing techniques implemented, a description of the data used for video processing, and the results and takeaways from the conducted experiment.

2. RELATED WORKS

Many existing works influenced the motivation and methodology behind this study. Many prior works have proposed successful video stabilization techniques, whereas other studies provide context for how improving stabilization of drone videos can further increase the accessibility of drones in the future. Many prior works introduce and/or analyze various algorithms for motion estimation, motion smoothing, and motion compensation that stabilize videos well. Many existing works, including [1] and [2], use the Lucas-Kanade method for motion estimation [3]. In fact, Lim et at. [4] refer to this motion estimation algorithm as the "golden standard" for video stabilization. Therefore, it is an ideal algorithm to use to calculate motion estimation on drone videos. Some smoothing algorithms that are commonly used in existing works in this domain are a wide variety of Gaussian filters [5], [6], a sliding-window average filter [4], and the Kalman filter [1], [2], [7]. Most of these prior works mentioned also use an affine motion model to assist in image and video stabilization. This study aims to use multiple variations of these algorithms to stabilize drone videos and compare their performances.

Various studies have analyzed human-drone interactions to identify potential opportunities to make drones more intuitive and accessible. Garcia et al. [8] developed Handifly, a software that adjusts drone pilot controls to adapt to the needs of drone pilots that live with sensory, cognitive, and motor impairments. Video stabilization is one of the implementations included in this software to assist these pilots, which indicates that evaluating various drone video stabilization algorithms could make drones even more accessible in the future.

3. METHODOLOGY

3.1. Data Description

The data set used in this study was generated by the Naval Postgraduate School for a team of researchers from Purdue University to utilize in a UAV-related study [1]. The data set includes 50 video sequences of 70,250 frames with 30 fps frame rate. They are recorded by a GoPro 3 camera mounted on a custom delta-wing airframe. Since the videos are captured in an outdoor environment, they contain numerous characteristics that are useful for the methods in this study to minimize, such as jitter. Additionally, the author of this paper provided a video shot on her drone to be used in the experiment. The existing data set can be accessed using this link: https://engineering.purdue.edu/~bouman/UAV_Dataset/.

3.2. Algorithm Overview

The intention of this study is to implement commonly used video stabilization functions on drone videos and evaluate their performance. Although different techniques are used, the overall process for video stabilization remains constant. There are three key steps to stabilizing a video. The first step is motion estimation, which is the process of estimating global motion trajectory in each frame of the video. The next step is motion smoothing, which involves applying a filter to the trajectory to eliminate high frequencies that cause unwanted motion. Finally, motion compensation is executed. This step encapsulates the process of warping the video by applying the filtered motion trajectory.

The different techniques used for motion estimation in this study are the Lucas-Kanade algorithm complemented by either the rigid transformation or the affine transformation. The algorithms executed for motion smoothing are the Laplacian of Gaussian filter and sliding window average filter. Efforts were made to implement the Kalman filter, but since this filter more suited for object tracking, which is outside the scope of this study, it was difficult to choose optimal parameters for the experiment data. The Laplacian of Gaussian filter was chosen over other Gaussian filters because it excels in edge detection [9]. For the final step of video stabilization, the affine transformation is implemented when Lucas-Kanade is used for motion estimation. Since the Laplacian of Gaussian filter warps the video in addition to smoothing it, using an affine transform is not necessary.

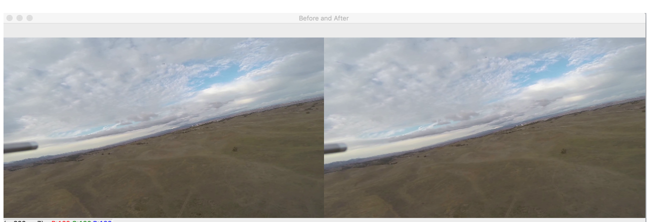

Fig. 1. The output window that contains a side-by-side comparison of the original video and the transformed video

4. EXPERIMENTAL SETUP

Each algorithm instance was written in Python within the Jupyter notebook environment. Some libraries that were utilized in the software development are OpenCV and Scipy's ndimage module. For each algorithm executed, a side-by-side video comparison between the original video and the transformed video is displayed, as seen in Figure 1. Additionally, as seen in [10], peak-noise-to-signal-ratio (PNSR) was computed for each video input and output as a means to compare the computational video quality against inputs and outputs from other methods.

A few input videos were used. One input video used was the drone video provided by the author of this paper. The remaining two input videos were arbitrarily chosen from the data set described earlier. The various instance of algorithms implemented are as follows: 1) Lucas-Kanade with rigid transform, sliding window average filter, and affine transform; 2) Lucas-Kanade with affine transform sliding window average filter, and affine transform; 3) Lucas-Kanade with rigid transform and Laplacian of Gaussian filter; and 4) Lucas-Kanade with affine transform and Laplacian of Gaussian filter.

Input Videos:	Original PSNR	Sliding Window Average Filter PSNR	Laplacian of Gaussian Filter PSNR
Drone Video	37.6069	37.7136	38.0456
Data Set Video #35	32.6686	33.9196	33.1572
Data Set Video #12	32.8113	34.9414	33.3820

Fig. 2. The peak-signal-to-noise ratios for each video input in the experiment for different instances of filtering

5. RESULTS

The results of this study varied slightly depending on the type of motion present in the input videos. The video provided by the author of this paper doesn't have much jitter, but it does have very sharp and abrupt movements (for example, at one point the camera is tilted down and immediately back up). The Go Pro videos from the existing data set we borrowed have smoother movements, but much more jitter present due to the wind in these videos. For all instances, the output video quality computed is slightly greater than the input video quality. Additionally, the type of transform used during the motion estimation process yielded results with negligible differences in quality.

When implementing the instances with the sliding window average filter, it can be observed that the difference in quality between the data set videos and their transformations is greater than the difference in quality between the drone video and its transformation. Also, the sliding window average filter performs better than the Laplacian of Gaussian filter when applied to the data set videos, but performs worse than the Laplacian of Gaussian filter when applied to the drone video. As for the case of implementing the Laplacian of Gaussian filter, the difference in quality between input and output videos is roughly the same for both types of video inputs. However, this filter performed better than the sliding window average filter when applied to the drone video and performed worse than the sliding window average filter when applied to the data set videos. Figure 2 displays the numerical results from this study.

The observations from this experiment indicate that commonly used video stabilization algorithms that exist today also perform well on drone videos. The sliding window average filter performs slightly better than the Laplacian of Gaussian filter on videos with smooth movements and rotations, and the latter filter performs better on videos with abrupt and sudden changes in movement and rotation than the former.

6. DISCUSSION

In this section, the various lessons learned through conducting these experiments are highlighted. Additionally, future directions of this study are discussed.

6.1. Lessons Learned

Through conducting this study, many lessons were learned. First, a better understanding of the main concepts implemented for video stabilization was grasped. Next, this study provided the opportunity to learn what video stabilization methods are popular in the research community and which of these methods work best given the type of video input. Additionally, the capabilities of common image and video processing libraries utilized in Python, such as OpenCV, were explored extensively. Finally, exploring existing works provided a taste of how the study of drone video stabilization algorithms can lead to an improved user experience with drones in the future.

6.2. Future Work

There are many opportunities to expand on this study. First, it would be very insightful to apply additional video stabilization techniques that were not implemented in this paper, such as dense optical flow and Kalman filtering. Applying deep learning methods to drone videos and observing their effect on drone video stabilization is a very exciting future direction, especially considering the rising trend of deep learning techniques. This study had a specific focus on improving video stabilization, but there are other characteristics that play a role in video quality that could be included in a study of this nature moving forward. For example, lighting conditions can impact the quality of drone footage. It would be interesting to evaluate what common video processing techniques would be useful for correcting exposure in drone videos.

7. CONCLUSION

Over the past ten years, drones have become an increasingly popular tool for capturing high-quality aerial videos. The various applications in which drone footage is useful has led to drones becoming more accessible to consumers. However, outdoor conditions during a drone flight and human error introduced by the drone pilot can minimize the quality of drone footage. This study explores the potential of various existing video stabilization algorithms performing well on drone videos. The observations from the experiment indicate that common video stabilization algorithms are also useful for drone video stabilization.

8. REFERENCES

[1] Dong Hye Ye, Jing Li, Qiulin Chen, Juan Wachs, and Charles Bouman, "Deep learning for moving object detection and tracking from a single camera in unmanned aerial vehicles (uavs)," *Electronic Imaging*, vol. 2018, no. 10, pp. 466–1–466–6, Jan 2018.

[2] Zilong Deng, Dongxiao Yang, Xiaohu Zhang, Yuguang Dong, Chengbo Liu, and Qiang Shen, "Real-time image stabilization method based on optical flow and binary point feature matching," *Electronics*, vol. 9, no. 1, 2020.

[3] Bruce D Lucas, Takeo Kanade, et al., "An iterative image registration technique with an application to stereo vision," 1981.

[4] Anli Lim, Bharath Ramesh, Yue Yang, Cheng Xiang, Zhi Gao, and Feng Lin, "Real-time optical flow-based video stabilization for unmanned aerial vehicles," *Journal of Real-Time Image Processing*, vol. 16, no. 6, pp. 1975–1985, Jun 2017.

[5] Amit Goldstein and Raanan Fattal, "Video stabilization using epipolar geometry," *ACM Transactions on Graphics*, vol. 31, no. 5, pp. 1–10, Aug 2012.

[6] Feng Liu, Michael Gleicher, Jue Wang, Hailin Jin, and Aseem Agarwala, "Subspace video stabilization," *ACM Trans. Graph.*, vol. 30, no. 1, Feb. 2011.

[7] Dong Wang, Bin Wang, Sicheng Zhao, Xiaoshuai Sun, Hongxun Yao, and Hong Liu, "Dual-mode video stabilization based on adaptive motion clustering," 2015.

[8] Jérémie Garcia, Luc Chevrier, Yannick Jestin, and Anke M. Brock, "Handifly: Towards interactions to support drone pilots with disabilities," p. 1–6, 2019.

[9] D. Marr and E. Hildreth, "Theory of edge detection," *Proceedings of the Royal Society of London. Series B. Biological Sciences*, vol. 207, no. 1167, pp. 187–217, Feb 1980.

[10] Semi Jeon, Inhye Yoon, Jinbeum Jang, Seungji Yang, Jisung Kim, and Joonki Paik, "Robust video stabilization using particle keypoint update and l1-optimized camera path," *Sensors*, vol. 17, no. 2, pp. 337, Feb 2017.