

ANALYSIS ON VIDEO STABILIZATION TECHNIQUES – A SURVEY

Abstract

Video stabilization is an important component of a visual quality improvement. It restores feature extraction and tracking in unstable videos. Video stabilization is to enhance the quality of consumer-level recordings, such as running, cycling, or driving. Video stabilization contains basic steps viz., Motion Estimation, Outlier Removal, Camera Motion Modeling, Camera Motion Correction, and Video Synthesis. A significant number of machine learning approaches are used to support computer vision techniques for stabilizing videos. Several computer vision difficulties, including super resolution, picture deblurring, style transfer, classification and recognition, among others have been successfully addressed using deep learning techniques. In recent years efficient 3D video stabilization currently demands the use of new deep learning techniques. Provide an analysis of video stabilization techniques in this work.

Keywords: Video Stabilization, Motion Estimation, Outlier Removal, Motion Modeling, Video Synthesis.

Authors

R.Mehala

Ph.D., Research Scholar
Department of Computer Applications
Applications
Alagappa University
Karaikudi.

Dr. K.Mahesh

Professor
Department of Computer
Alagappa University
Karaikudi.

I. INTRODUCTION

Video stabilization is one of the popular topics in the recent years in industry and academia. Many intelligent applications, including transportation, security, and monitoring systems, heavily rely on vision systems. Videos captured using handheld mobile cameras frequently have undesirable and slow motions like track, boom, and pan, which dramatically lower the video's quality. Video stabilization is accomplished by estimating and reducing the unwanted interframe motion between the succeeding frames, synthesizing the new stabilized video sequence and in mobile videos, the interframe action is typically slow and smooth.

II. RELATED WORK

A large number of earlier video stabilization techniques use the same structure and focus on enhancing individual elements. The primary difficulties, practical considerations, and mathematical fundamentals of video stabilization algorithms are presented in this study [1]. By proposing new technique that was motivated by the findings of research on Image Quality Assessment in its broadest sense, particular attention to the Video Stabilization Quality Assessment has been presented. In order to get around the shortcomings of current techniques, several fresh directions for future research have been offered.

Digital video stabilization solutions are relatively cheap and simple to employ because they merely utilize data from previously recorded film and don't require any additional hardware or technical expertise about camera motion [2]. Earth motion, military target tracking, astrophotography, and other fields all require video stabilization. There are numerous methods for stabilizing videos. This essay provides a study of the various digital video stabilization techniques.

Liu, Shuaicheng & Tan et.al [3] presented a model that outputs the stabilizing mesh from the previous frame into a sliding window given a 2D camera path as input, warping the subsequent frame to its stabilized location. Online video can be stabilized by continually using the model for incoming frames. A hybrid loss to preserve the video's stabilized spatial and temporal coherence. This Motion Stab dataset was produced for training.

Umrikar, D.P., & Tade et.al [4] provided a novel method to use CNNs only as an optimizer instead of learning from data. This method also utilizes the overall advantages of CNNs, including conventional gradient-based optimization techniques. This approach outperforms earlier work both computing and visually and it is robust in circumstances where current state-of-the-art approaches have known shortcomings.

Marcos Roberto e Souza et.al [5] present a comprehensive analysis of the video stabilization literature, arranged in accordance with a suggested taxonomy. The problem is introduced with a formal definition and a brief physical interpretation. They also gave a thorough analysis of the key issues and emerging trends in this busy field.

Yiming Wang et.al [6] presents a comprehensive survey of video stabilization. Three stages of video stabilization are briefly reviewed. Besides, conventional methods and the deep learning-based methods are discussed in detail. Furthermore, a special attention to video stabilization quality assessment, benchmark datasets, and state-of-the-art performance has been described. Finally, discussions on current challenges and future directions to overcome

the limitations of the existing methods and better meet the needs of researchers in this active area are provided.

III.A FUNDAMENTAL CONCEPTS OF VIDEO STABILIZATION

Mechanical, optical, and digital stabilization are the various stages that video stabilization has gone through. A gyroscopic stabilizer is a tool used in mechanical stabilization. The camera's gyrostabilizer, an indiscernible tripod, corrects for the videos that are captured and delivers stabilized videos. These gadgets are unsuited for imaging that require minimal power consumption and payload, like hand-held devices, due to their weight and increased energy consumption.

Video stabilization techniques seek to produce new video sequences with minimal interframe motion. The three phases of any digital video stabilization technique are Motion Estimation (ME), Motion Smoother (MS), and Motion Compensation (MC). Motion Estimation estimates the motion between the frames and passes the motion parameters to Motion Smoothing, which eliminates the unwanted camera motions. Motion Compensation determines the overall alteration that will stabilize the current frame.

- 1. Motion Estimation:** Finding motion vectors from nearby frames in a video series is known as motion estimation. The motion vectors may apply to the entire frame or only a certain area, such as rectangular blocks, patches of any shape, or even a single pixel. Motion estimation is an essential component in any technique for video stabilization. It is especially crucial to obtain a precise estimate of the global motion when performing motion estimation since camera motion in real life frequently entails some sort of global transformation.
- 2. Motion Smoothing:** The motion vectors for both frames can be calculated using the motion estimation process' output. The purpose of motion smoothing is to eliminate jitters and unwanted motions from the footage that was captured. Passive filters are an efficient way to eliminate these undesirable motions, which are typically thought of as high frequency motion like vibration. Consequently, motion filtration is another name for motion smoothing.
- 3. Motion Compensation:** In order to create stabilized video, motion compensation involves reducing temporal redundancy between frames. It differs from motion smoothing, which gets rid of spatial repetition in between frames. Motion compensation is done frame by frame using the motion smoother's smoothed motion vectors. By employing the related smoothed motion vector for motion compensation on the original frame₁, for instance, stabilized frame₂ is produced. Stabilized frame₂ and its accompanying smoothed vectors are utilized to obtain stabilized frame₃, and so on. The full stabilized video sequence can be created in this way.

IV. TECHNIQUES OF VIDEO SYNOPSIS

- SIFT(Scale Invariant Feature Transform)
- SURF(Speeded Up Robust Feature)
- FAST(Fast Feature from Accelerated Segment Test)

- BRIEF(Binary Robust Independent Elementary Feature)
- ORB(Oriented FAST and Rotated BRIEF)

- 1. Scale Invariant Feature Transform (SIFT):** SIFT is a computer vision method used to find and describe specific characteristics in images. This feature is frequently applied in image processing. Difference of Gaussians (DoG) Space Generation, Keypoints Detection, and Feature Description are some of the SIFT operations. SIFT key points are taken from a collection of reference images and entered into a database. The SIFT characteristics are invariant to image scaling and rotation because they are local, based on the way the object appears at certain interest points. SIFT algorithm consists of four major stages: scale-space extrema detection, keypoint localization, orientation assignment and keypoint descriptor. The Benefits of extract distinctive features from images that is independent of an image's scale and rotation. The limitations are slow and inefficient at detecting changes.
- 2. Speeded Up Robust Feature (SURF):** It is a local feature detector and descriptor with a patent, is used in computer vision. It can be applied to projects like 3D reconstruction, image registration, classification, and object recognition. Scale-invariant feature transform (SIFT) descriptor is a source of some of its inspiration. SURF performs similarly to SIFT while being faster. The significant speedup results from the use of integral pictures, which, regardless of scale, greatly reduce the number of operations for basic linear convolutions. The Limitations of the rotation and lighting changes make it unstable.
- 3. Feature Accelerated Segment Test (FAST):** A single pixel in an image that has a distinct position and can be reliably recognized is considered to be an interest point. Interest points should preferably be repeated across images and have high local information richness. Application areas for interest point detection include object recognition, image matching, and tracking. The Benefits of the feature point may be extracted and an interest point detector could be created using the edge detection technique FAST for use in real-time frame rate applications. The limitations is High levels of noise cannot be tolerated.
- 4. Binary Robust Independent elementary Feature (BRIEF):** Recent feature descriptors like the BRIEF identifier use simple digital comparisons between pixels in an image patch. In many ways, its performance is comparable to SIFT's, including its resilience to lighting, blur, and perspective distortion. It is, nevertheless, extremely susceptible to in-plane rotation. An image patch created from a series of binary intensity tests is described by a bit string in this sentence. It offers a quick way to locate binary strings without first looking for descriptors. The benefits of constructing and matching are substantially quicker than alternative techniques. Greater recognition rates, provided that they are invariant to significant in-plane rotations, Real-time matching performance with short computational times. The Limitations is the object is not intended to be rotationally invariant.
- 5. Oriented FAST and Rotated BRIEF (ORB):** By adding a quick and precise orientation component and utilizing the rotated BRIEF description, ORB modifies the FAST detector to detect critical spots. Key points are found on each level of a pyramid that has been built

from the image. It is expected that corner intensity is offset from the centre. To calculate orientation, a vector representing this offset is used. The pixel patch is used to smooth out images. The BRIEF descriptor is then directed based on the orientation of each pixel patch to provide rotational invariance. The benefits of almost two orders of magnitude faster than SIFT and SURF, less influenced by image noise, and computationally efficient compared to SIFT. The limitation of the purpose of ORB does not include scale invariance.

V. CONCLUSION

Several feature extraction methods, including scale-invariant feature transform, accelerated robust feature, accelerated segment test, binary robust independent elementary feature, and oriented FAST and robust BRIEF are studied in this paper for the purpose of motion estimation in video stabilization. A thorough explanation of the Techniques used for the global motion estimate stage of video stabilization can be provided for each technique. The results of this evaluation revealed a number of flaws and restrictions in each and every technique.

REFERENCES

- [1] WilkoGuilluy, Laurent Oudre, AzeddineBeghdadi, Video stabilization: Overview, challenges and perspectives, *Signal Processing: Image Communication*, Volume 90, 2021, 116015, ISSN 0923-5965, <https://doi.org/10.1016/j.image.2020.116015>.
- [2] Zhu X, Guo K, Fang H, Chen L, Ren S, Hu B (2021) Cross view capture for stereo image super-resolution. *IEEE Trans Multimed*
- [3] Zhu X, Guo K, Ren S, Hu B, Hu M, Fang H (2021) Lightweight image super-resolution with expectation maximization attention mechanism. *IEEE Trans Circ Syst Video Technol*
- [4] Patel Amisha* ,Ms.HetalVala.“A Survey on video stabilization Techniques”-International Journal of Engineering Sciences & Research Technology-ISSN: 2277-9655. February, 2015
- [5] Liu, Shuaicheng & Tan, Ping & Yuan, Lu & Sun, Jian & Zeng, Bing. (2016). MeshFlow: Minimum Latency Online Video Stabilization. 9910. 800-815. 10.1007/978-3-319-46466-4_48.
- [6] Umrikar, D.P., & Tade, S.L. (2017). Survey paper for Different Video Stabilization Techniques.
- [7] Marcos Roberto e Souza, Helena de Almeida Maia, and Helio Pedrini. 2022. Survey on Digital Video Stabilization: Concepts, Methods, and Challenges. *ACM Comput. Surv.* 55, 3, Article 47 (March 2023), 37 pages. <https://doi.org/10.1145/3494525>.
- [8] Yiming Wang, Qian Huang, Chuanxu Jiang, Jiwen Liu, Mingzhou Shang, Zhuangmiao, Video stabilization: A comprehensive survey, *Neurocomputing*, Volume 516, 2023, Pages 205-230, ISSN 0925-2312, <https://doi.org/10.1016/j.neucom.2022.10.008>.
- [9] David G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints”, *International Journal of Computer Vision*, 2004.
- [10] Herbert Bay , Andreas Ess , Tinne Tuytelaars , Luc Van Gool “Speeded-Up Robust Features (SURF)”, *Computer Vision and Image Understanding* 110 (2008) 346–359.
- [11] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, “ORB: an efficient alternative to SIFT or SURF”, 2011 IEEE International Conference on Computer Vision, pp. 2564- 2571, Nov. 2011.