TuEngr Group

International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies

http://TuEngr.com

PAPER ID: 11A06S

A ROBUST VIDEO STABILIZATION ALGORITHMS FOR GLOBAL MOTION ESTIMATION USING BLOCK MATCHING

Syed Shuja Hussain [1], **Muhammad Nazir** [1], **Sunil Kumar Sharma** [2], **M. A. Rahim Khan** [2*]

[1] *Department of Computer Science, Majmaah University, Majmaah 11952, SAUDI ARABIA.*
[2] *College of Computer and Information Sciences, Majmaah University, Majmaah 11952, SAUDI ARABIA.*

ARTICLE INFO

Article history:
Received 02 September 2019
Received in revised form 03 January 2019
Accepted 14 January 2020
Available online 25 January 2020

Keywords:
Global motion vectors; Video stabilization; Block matching, Camera motion; Smooth motion vector; Low pass filter; Block Matching Algorithm.

ABSTRACT

In a camera dictionary, shakiness jitters are normally used. Jitteriness problem is encountered due to frequent use of a phone camera. Shooting video from moving vehicles or by using a handheld phone camera causes Shakiness. Block matching algorithms can be used for implementing the video stabilization process. The proposed algorithm in this paper avoids shakiness from a raw video of YUV and results in stabilizing in YUV video. The proposed method works on extracting motion among consecutive frames of any raw video. With respect to every scene of the video, premeditated camera motions like panning and translation motion are smooth with slow variations in the time realm from frame to frame. Fast motion variation with respect to frame to frame is increased while sponging the camera motion. A low pass filter can be used to find the recovery of the calculated motion parameters. The corrective shifting of noisy components of the motion vector stabilized the video images. The accurate result can be achieved by a successful run on shaky video without hampering its quality. The parallel processing techniques have been applied for the estimation of motion vectors, the results improved by applying to modify the proposed algorithm for dropping the operation time of the program.

Disciplinary: Computer Science and Engineering.

1 INTRODUCTION

Video quality is reduced while taking the video from moving vehicles or by the handheld camera due to shaking. Video stabilization by using digital image processing methods is useful to avoid the camera motion (Zhang, et al, 2015). In this paper, video stabilization & various real word stabilizing video applications along with YUV videos i.e. Raw format videos are discussed. This is the preprocessing of data captured by the camera before saving it in standard formats as .avi or .mpeg for the stabilization process. There are some handshaking effects in the videos captured by these devices (Rawat, et al, 2013). These effects from videos can be removed by using digital stabilization methods.

*Corresponding author (M. A. Rahim Khan) Email: m.khan@mu.edu.sa

Video stabilization plays an important role in dynamic image coding and processing to remove hand-shaking effects in a handheld camera or PC (Yang et al., 2015). The proposed application can be installed in these handheld cameras or PC to get converted stabilized video and its further processing from a raw video (Okayed et al., 2015).

2 BACKGROUND OF PROBLEM

The vibrations and undesired motion occur in the video during the video acquisition process make the video unstable so that the video stabilization techniques are required to fix the issues related to detecting and correcting the malfunctions. Several video stabilization techniques have been developed till today for a wide range of performance, the fast strategies are the most significant to get real-time video stabilization on embedded devices (Giovanni et al. 2011). The adaptive Gaussian filters are a powerful technique for video stabilization (Souza, 2018).

The challenges in the video stabilization in the block –matching algorithm to reduce the computing requirement and complexity to obtain high accuracy with various constraints to achieve the optimum video stabilization and tracking of the path of the original video (Li et al., 2014).

Many schemes have been proposed for video stabilization that uses various transformations and estimation approaches. Hung-Chang et al. (2004) proposed to use the Affine Transformation in the video stabilization process. The motion estimation between sequential frames has been calculated by the 6-parameter affine transformation $f(x, y, t)$ and $f(x, y, t-1)$. The proposed affine model can capture various attributes of the motion like scale, rotation & translation, etc. This model can be extended to incorporate polynomials with higher orders that can handle motion in a better manner. A quadratic error function is minimized to estimate parameters of an affine transformation. The proposed model can manage and recognize rotation, translation & scaling motions among the various frames however it is computationally extensive as large no of parameters is to be examined.

The k-means clustering algorithm helps in stabilized the video by eliminating the local motion vector due to the shaking videos (Wu et al., 2017). This method's advantage is to provide robust security effectively.

Liu et al. (2009) discussed a procedure that changes a video from a camcorder, so it shows up as though it were taken with a coordinated camera movement. The proposed technique changes the video to show up as though it were taken from adjacent perspectives, permitting 3D camera developments to be reenacted. The structure form – motion is a very powerful technique for the smoothing motion of a three-dimensional model (Liu et al., 2011) in this technology the quality of stabilization is high but it required the high computing requirement.

This strategy first recuperates the first 3D camera movement and a meager arrangement of 3D, static scene focuses on utilizing an off-the-rack structure-from movement framework. At that point, an ideal camera way is figured either consequently (e.g., by fitting a straight or quadratic way) or intelligently. At long last, its strategy plays out a slightest squares advancement that registers a spatially-changing twist from each info video outline into a yield outline.

Battiato et al., (2009) proposed a framework in which the SHIFT feature tracked to handle inter-frame motion in consecutive frames. The proposed approach doesn’t lean on a point detector taken before the criteria of the SHIFT descriptor. So the performance is measured depending on various point detection approaches by selecting proper detectors for increasing the quality of stabilization (Xie et al., 2015).

First, the Feature points are identified & then by using fuzzy logic modeling and geometric error measures, its stability is calculated. For evaluating the matching quality, two types of error measures are used (Abdullah et al., 2012):

i. The approach Angle between the two local motion vectors is acquainted with rotational components.

ii. The other approach Euclidean remove among expected and genuine points: well suited for coordinating that does not agree with got translational parts is rejected. It may not create the right outcome for fringe focuses amid the event of turn. The fluffy model can be utilized to remunerate this which takes two blunder measures as info and gives a solitary quality record as yield. The yield is a genuine incentive in the range [0, 1] that speaks to the level of coordinating among the different sets of focuses.

Another approach for video stabilization is content-based video retrieval [CBVR] is widely used for the data likewise visual, audio or metadata. CBVR techniques have the three main features: Key-frame detection, Shot boundary detection and video retrieval. The techniques are effective to extract the texture, local and global color and preserve the motion features of the videos (Mohamad et al., 2016).

3 FORMULATION OF PROPOSED RESEARCH WORK

The proposed research problem is formulated sequentially as per shown with the following steps in Figure 1.

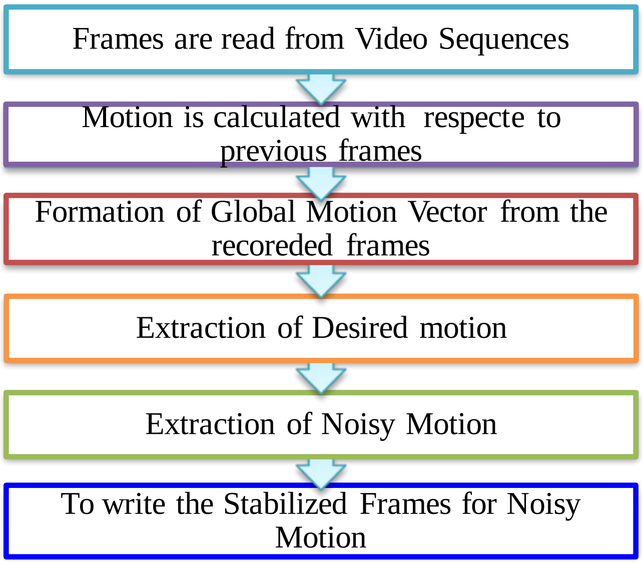

Figure 1: YUV Format Process flowchart

There are three parts in YUV arrange that are utilized to encode a video or shading picture i.e. one Luma segment Y' (White and Black Component) and the important is a feature that the color chrominance components U and V. Y'UV was presented by researchers amid the structure of shading TV in the highly contrasting framework. A flag transmission technique was required that was good with high contrast TV and ready to include hues [Kalvein Rantelobo, et al,2016]. The Luma part is now there as a dark and white flag. The UV flag is added as an answer to the said necessity. The structure of the YUV420 outline appears in Figure 2.

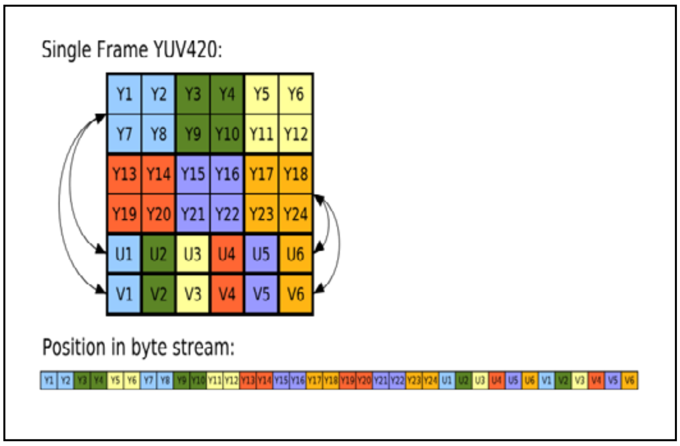

Figure 2: UV420 signal Frame.

For every pixel, Y esteems are extraordinary. As appeared in Figure 2, 2*2 squares of pixels are utilized to share U and V esteems. Y esteems are put away first, at that point U esteems lastly V esteems at adjoining memory areas to frame a byte stream. Since the proposed research work needs to appraise the movement between persistent edges, so just Y esteems are utilized and remunerated U and V outline esteems are adjusted by assessed movement (Yang et al., 2015).

3.1 DIGITAL VIDEO STABILIZATION ALGORITHMS

For Digital Video Stabilization, several different approaches are available. Few of them are Differential based, matching based, Energy-based, and Phase based approaches (Suneja et al, 2010). The video stabilization approaches are - NR Iterative Approach for Video Stabilization, Affine Transformation Based Video Stabilization and Block Matching Algorithm and Dual-Tree complex wavelet transform (DT-CWT) (Pang, 2010).

The proposed research focuses on Block Matching Algorithm. In Block Matching Algorithm, motions between consecutive frames are estimated. It is done by matching the current frame-blocks with previous frames.

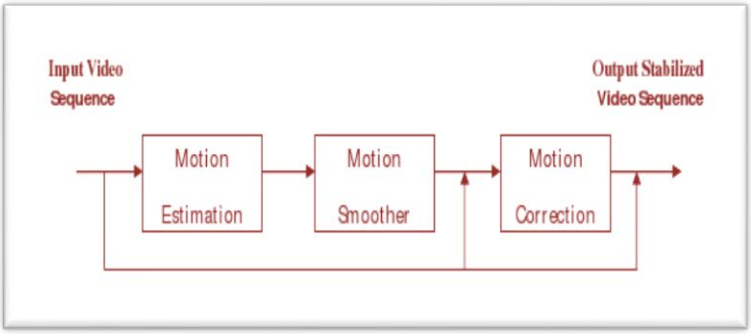

Figure 3: Motion Stabilization Architecture Block Diagram.

The Motion Estimator reads video frame sequences first and then compute the differences between consecutive frames. Further, the possible motion is expected, and a motion vector is generated. This generated motion vector is produced to Motion Smoother that eliminates unwanted motions. Finally, based on smoothed motion vector sequences, the Motion Corrector does require adjustments to produce output as stabilized video sequences or frames (as shown in Figure 3). In the proposed research, the basic assumption is taken that the motion is translational only.

3.2 VIDEO STABILIZATION PROCESS: MOTION ESTIMATION

In motion estimation there are thee three steps, In the first step, the image is divided into a small block of the size16 * 16 pixels. The frame comparison iteratively takes place. It has been compared that the next frames matched a block to the previous frame with the proximity of 8 pixels in each direction (Mostafa et al., 2016). The Sum of Squared Differences (SSD) is calculated to decide the local and global motion vector of each block.

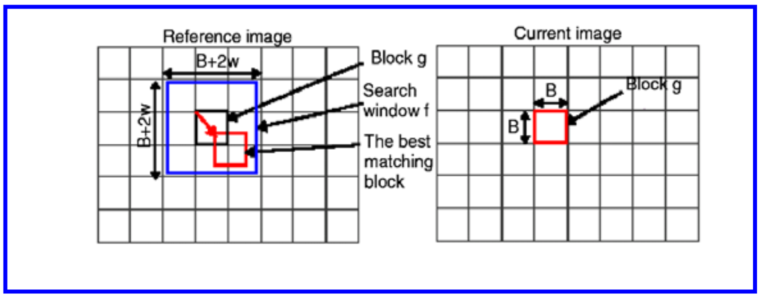

Figure 4: Key Block for the reference and Current image for Best Match

The video captured have the sequential frames, the key block is important in the whole image for the block matching. The local motion vectors have been obtained by the fast block and used to generate the Global motion vector between the iterative frames through the affine model using the method of least squares (Hong et al., 2010).

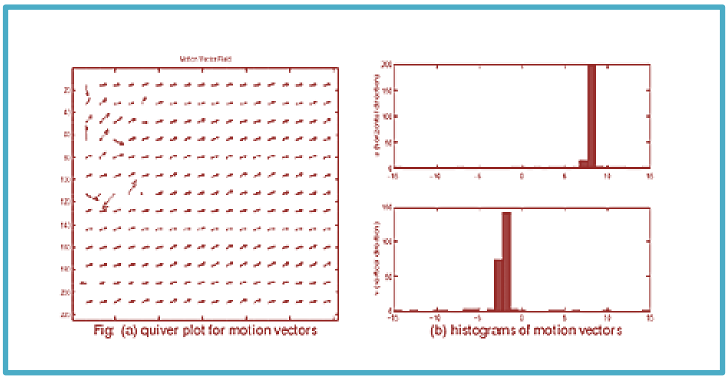

Figure 5: (a) Quiver plot of the motion vector, (b) Histogram of the motion vector.

Figure 5, the quiver plot in the form arrows displays the local motion vectors for the image. If the components $(u(i), v(i))$ transform in $(x(i), v(i))$. The dimension of the matrices $A = [u, v]$ and $B = [x, y]$ should be the same size and hold the correspondence between position and velocity components. The Global Motion of video sequences is represented by these Local Vectors. Some random values in the graph represent the variation from the desired motion. Local motions of the objects in the frame are represented by these local vectors. The local motion in a video has been considered for the objects and global motion as the motion of the background.

Figure 5(b) depicts the histogram of the Local Motion Vectors for the objects in the horizontal & vertical direction of the frames. The horizontal component Global Motion $GHM_i = |\sum_i^n HC_i|$ where HC_i is the of Local Motion Vectors in the horizontal direction. Similarly, the other components in The horizontal component Global Motion $GVM_i = |\sum_i^n VC_i|$ where VC_i is the of Local Motion Vectors in the vertical direction.

The sequence of the frame makes the series of vectors that represent the motion of the sequential blocks of the key-frame. The histogram approach has been used for the robustness of the model. The bivariate variance of the blocks of the image is

$$\sigma^2 = \frac{1}{mn}\sum_{j=1}^{m}\sum_{i=1}^{n}\left(B_k(i,j) - \overline{B_k}\right)^2 \quad (1),$$

where, B_k denotes the block need to register. Higher values of the standard deviation result that the image has a high level of residual.

3.2.1 LOCAL MOTION VECTOR (LMV)

The components of the LMV in the horizontal and vertical are denoted by the symbols $X \text{ and } Y$ respectively. The local motion vector has the instances the same as the total number of the frames in the video's sequences reduced by one. The size of the vectors is calculated by the number of blocks discretize in a frame. Each frame of the video's sequences has a motion of the blocks of the frame and results in mapping with a correspondence block.

3.2.2 GLOBAL MOTION VECTOR (GMV)

The GMV has the horizontal and vertical components and represents by X and Y. The size of GMV is obtained by the count of the video frames and reduced by one. i.e. n-1 relative motions are stored if there exist n frames in a video. This global motion vector is used to find the complete motion of the new frame with the corresponding old frame.

3.2.3 MOTION SMOOTHER

The motion smoothing the videos the global motions plays a key role. The Global for the consecutive frames of the video sequences the following method is applied due to the block matching of the consecutive frames.

$$NGV(i) = NGV(i-1) + OGV(i) \quad (2).$$

For removing the shakes and jitters, this New GMV is required to be smoothed. In this work, the window size $n = 5$ is considered and MA (Moving average) is taken as a smoothing filter. The GMV gives the displacement between any two frames that can be found. This vector has the limitation; the motion cannot be compensating due to the involvement of all the desired motion in smoothing the motion. All unwanted motion must be compensated. There is no need to remove the desired motion. Smooth Motion Vector can be found by smoothing New Global Motion Vector which contains only the required motion as all unwanted random motion has been removed by its smoothing.

Figure 6 shows that by using the smoothening process, all unwanted random shakes are removed that resulting only in the synthesized frame for the corresponding vector, in the smoothening process known as New GMV. The camera motion should be used in such a way that no black borders appear in the stabilized video frame. This new GMV is used to smoothing the motion vector. However, the size and structure of the new GMV do not affect in the smoothing process.

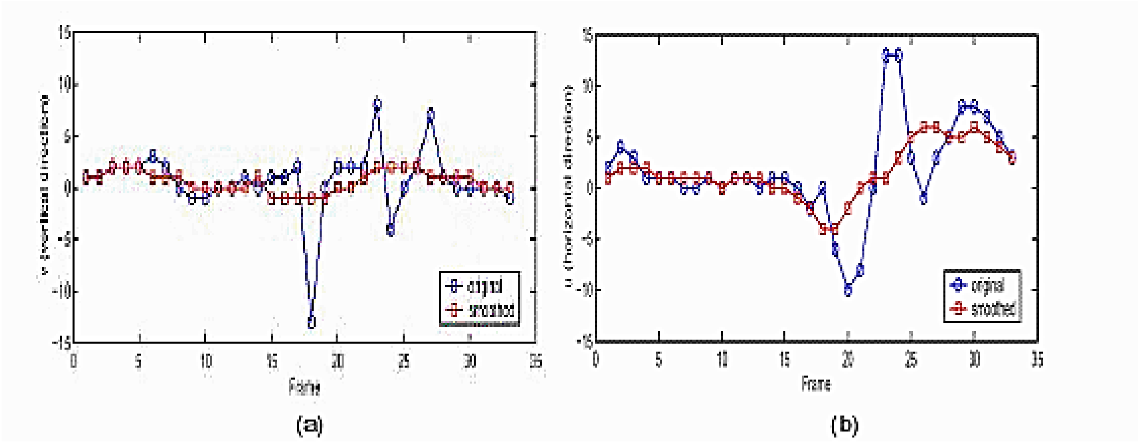

Figure 6: Smooth motion & original new global motion vector (a) vertical and (b) horizontal directions.

3.2.4 MOTION CORRECTION

In the frames of the video, the time elapsed for the two frames is very less, this small-time acquisition results in the poor quality of the images. In many applications like brain imaging lot of physiological change takes place results in poor images. The series of images must produce some functional information. The information gathered with the help of the statistical analysis of the data obtained by the series of the frames. To fit the data for the analysis of the motion it should be well corrected. Motion compensation is performed through Motion corrector. It is also used to construct the smoothed frames. The error in motion must be found after calculating all the required motion is called as the smooth motion vector. The stored difference vector calculated as

$$dv\ [i] \ = \ NGV[i] - \ SV\ [i] \qquad (3).$$

When unwanted motions are found, this motion is compensated while finding the new stabilized video sequences.

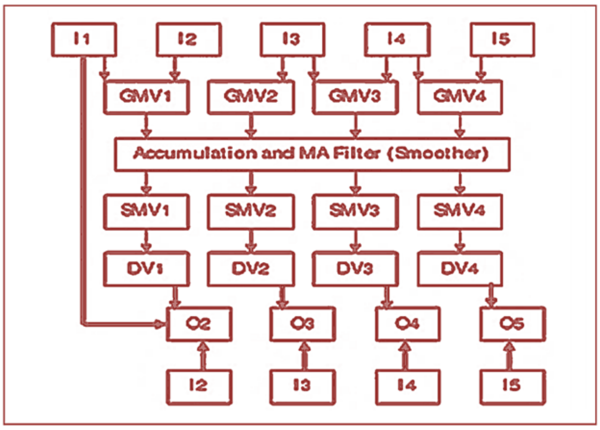

Figure 7: Process of entire video stabilization.

In Figure 7, Input frames are represented by I1, I2, I3, components of Global Motion Vector are represented by GMV1, GMV2, components of smooth motion vectors are represented by SMV1, SMV2, components of difference Vectors are represented by DV1, DV2 and final smoothed stabilized frames are represented by O1, O2. DV1 is applied by Motion corrector on frame 1 to generate smoothed frame 2 and DV2 is applied on input frame 2 (I2) to generate smoothed frame 3 and continued sequentially. The thick lines represent the motion correction in the entire video stabilization process.

4 DESIGN

Following modules are used to modularize the proposed application:

i. GrabFrame: Frames from input video sequences are garbed through this module and total numbers of frames are calculated. Some data structure is used to store Y, U, V pixel values separately.

ii. Get motion: this module compares two consecutive frames. Further local motion vector and global motion vectors are calculated through it.

iii. Calmotion: A block matching algorithm is implemented through this module. Also, the current and previous frames are compared. The frame is divided into a block of 16 * 16 pixels. Each current frame block and its corresponding previous frame-blocks are compared in the proximity of 16 pixels.

iv. Mode: Mode function is implemented through this module. The global vector depends on the mode value of the Local motion vector's mode.

v. MovingAvg: This module is important and has a similarity with the low pass filter. An accumulated version of GV and new GV are the inputs of this module. Due to movement of the camera & desired motion, shaky motions are contained in this vector and it is the component of a high pass in new GV which has been removed because of the low pass filter. So smooth vector is the output of this module. The element of vector that gives the noisy component information can be achieved by Subtracting the smooth vector from the new global vector that is called diffvector.

vi. Write: Through this module, stabilized frames are written back to the corresponding output file. The first input frame is written as same. Due to camera vibration, diffvector's ith element is like the forward shift operator of the frames. Therefore, the backward shift of the frame associated with the diffvector [i]. The black color is used to fill the edges (Suneja, et al, 2010).

5 TESTING AND ANALYSIS

Virtual videos are used to examine all the code first. Even small jitter ware removed with high precision so we get results of fine quality. Then using a digital camera, some videos were shot with heavy shakes & vibrations. By some debugging and modifications, we get similar results when tested the code with these videos.

Video stabilization results for YUV video was shown in the given demo. The video was captured by Cannon digital camera in mp4 format that is converted to YUV format by using yuv_tools with resolution 176 * 144 pixels.

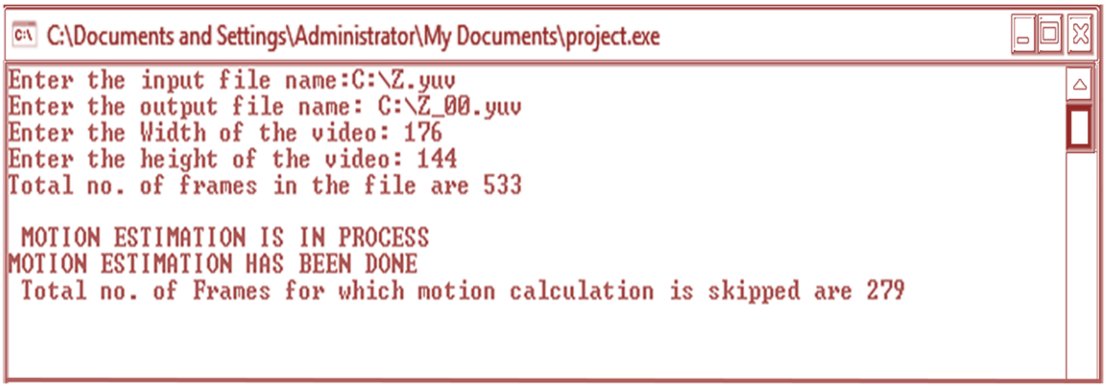

Figure. 8: Support Window after running code on information video Z.yuv

Figure 8 shows the video file name (input), width & height of the video output video name, which are the inputs to the program.

Figure 9: Sequential frames of Input Video.

Figures 9 and 10 show the mapping between the sequential frames and stabilized frames. The figure clearly shows that after applying the block matching video stabilization techniques the image is less shaky.

Figure 10: Stabilized frames of corresponding output video.

The shaky motion, all frames have been compensated. Since some frames are shifted corresponding to the defector, so some pixels are required to be filled to keep the same frame size. That's why there is some black color is there at the frame's edges.

6 CONCLUSION

The proposed worldwide movement assessment calculation depends on 3 stage assessment demonstrate that can be utilized to expel hand-shaking impacts with great execution for translational video movements. C++ is utilized to actualize the proposed calculation. Recordings can be balanced out correctly by evacuating instability with the created Video Stabilizer with no quality debasements. The outcomes delivered by the proposed calculation are worthy of genuine and virtual or mimicked recordings. The proposed calculation can be utilized to expel the hand-shaking impacts by video adjustment that is difficult to dispense with close by held gadgets like cameras or cell phones. The square coordinating calculation is utilized to execute a video adjustment process that requires tremendous calculation for SSD calculation. Parallel preparing (like Compute Unified Device Architecture (CUDA®)) for estimation of movement vectors can be utilized to change the proposed calculation for dropping the activity time of the program.

7 DATA AND INFORMATION AVAILABILITY

Relevant information can be made available by contacting the corresponding author.

8 ACKNOWLEDGMENT

The authors thank Deanship of Scientific Research at Majmaah University for supporting this work under Project Number No. 1441-88.

9 REFERENCES

Mohamad, A., and H. Farsi. (2016). Content Based Video Retrieval Based HDWT and Sparse Representation, Image Anal Stereo. 35, 67-80.

Suneja A., and G. Kumar. (2010). An Experimental Study of Edge Detection Methods in Digital Image, Global Journal of Computer Science. 10(2), 17-19.

Battiato S., Gallo G., Puglisi G. (2009). Fuzzy based Motion Estimation for Video stabilization Using SIFT interest points. Conference on Digital Photography, CA, USA.

Pang, D., Huizhong Chen and Sherif Halawa. (2010). Efficient Video Stabilization with Dual-Tree Complex Wavelet Transform. EE368 Project Report.

Liu, F., M Gleicher, J Wang, H Jin, A Agarwala. (2011). Subspace video stabilization. ACM Trans. Graph. 30(1).

Liu, F., M. Gleicher, H. Jin, and A. Agarwala. (2009). Content Preserving Warps for 3D Video Stabilization, ACM Transactions on Graphics. 28(3).

Puglisi, G., S. Battiato. (2011). A Robust Image Alignment Algorithm for Video Stabilization Purposes, IEEE Transactions on Circuits and Systems for Video Technology. 21(10), 1390–1400.

Hung-Chang Chang, Shang-Hong Lai, Kuang-Rong Lu, (2004). A robust and efficient Video Stabilization Algorithm, Multimedia and Expo, IEEE International Conference on Multimedia and Expo (ICME), Taiwan.

Kalvein Rantelobo, Hendro Lami, Wirawan, (2016).Video Transmission using Combined Scalability Video Coding over MIMO-OFDM Systems, Indonesian Journal of Electrical Engineering and Computer Science, 4(2), 390 -396.

L. M. Abdullah, N. M. Tahir & M. Samad. (2012). Video Stabilization based on Point Feature Matching Technique, IEEE Control & System Graduate Research Colloquium, (ICSGRC 2012). Shah Alam, Selangor, Malaysia.

Lengyi Li, Xiaohong Ma, and Zheng Zhao. (2014). Real-Time Video Stabilization Based On Fast Block Matching And Improved Kalman Filter, Fifth International Conference on Intelligent Control and Information Processing, Dalian, Liaoning, China.

M. Okade, G. Patel, P.K.Biswas. (2015). Robust Learning Based Camera Motion Characterization Scheme with applications to video stabilization, IEEE Transactions on Circuits and Systems for Video Technology. 26(3), 453–466.

Mostafa, S.A. K., Abdelrahman Ali, Multiresolution. (2016).Video Watermarking Algorithm Exploiting the Block-Based Motion Estimation, Journal of Information Security, 7, 260-268

Rawat, P., J. Singhai. (2013). Efficient Video Stabilization Technique for Hand-Held Mobile Video, International Journal of Signal Processing, Image Processing and Pattern Recognition. 6(3), 17-32

Souza, M. & Pedrini, H. (2018). Digital video stabilization based on adaptive camera trajectory smoothing" J Image Video Proc. 37.

Hong, W., Wei D., and Batur, A. U. (2010). Video Stabilization and Rolling Shutter Distortion Reduction. 2010 IEEE 17th International Conference on Image Processing. 3501-3504.

Wu, M., Li, X., Liu, C., Liu, M., Zhao, N., Wang, J., & Zhu, L. (2017). Robust global motion estimation for video security based on improved k-means clustering. Journal of Ambient Intelligence and Humanized Computing, 1-10.

Zheng, X., C.Shaohui, W.Gang, L. Jinlun. (2015). Video Stabilization System Based on Speeded-up Robust Features, International Industrial Informatics and Computer Engineering Conference, Published by Atlantis Press, 1995-1998.

Zhang, Y., Y Leng, Xu He. (2015). A fast video stabilization algorithm with unexpected prediction strategy, IEEE International Conference on Advanced Intelligent Mechatronics (AIM), 571- 576

Zhang, Y., Y.Leng, X. He. (2015). A Fast Video Stabilization Algorithm with Unexpected Motion Prediction Strategy, IEEE International Conference on Advanced Intelligent Mechatronics, Busan, Korea.

Syed Shuja Hussain is a Lecturer in Computer Science at Majmaah University, Kingdom of Saudi Arabia, and information security specialist in the Department of Information Security in the Deanship of Information Technology at Majmaah University, Saudi Arabia. He received a BS in Computer Engineering from Sir Syed University of Engineering & Technology, and MS (Telecom Engineering) from University of Engineering and Technology (UET), Peshawar. His research interests include Network Security, Information Security, Cybersecurity and IoT Security.

Muhammad Nazir is a Lecturer in Computer Science at Majmaah University Kingdom of Saudi Arabia. He received a Bachelor of Science Degree from the University of Punjab, Pakistan, and a Master's Degree from International Islamic University, Islamabad Pakistan. His research interest includes Data Mining, Distributed Data Base and Image Processing.

Dr. Sunil Kumar Sharma is an Assistant Professor of Mathematics at the College of Computer and Information Sciences. Majmaah University, Saudi Arabia. He received his Ph.D. in Mathematics from Gautam Buddh Technical University, Lucknow India. His areas of research interest are to deal with the Mathematical Modeling, Numerical Computation of Biomechanics of Diarthrodial Joints, Multicriteria Decision Model for Production Systems, Robot in the field of education. The current research involves Mathematical Modeling of Physical and Biological Problems in general and Numerical Analysis.

Mohd Abdul Rahim Khan is a Lecturer at Majmaa'h University. He got a Master's degree in Technology in Information Technology from GGSIP University Delhi. He is pursuing a Ph.D. at Lingaya's University Faridabad. His research areas are Information Security, Android security, Computer vision, Artificial Intelligence, etc.