Journal of Engineering and Applied Sciences 13 (Special Issue 1): 2413-2419, 2018
ISSN: 1816-949X

Determining the Best Method to Extract Interest Point for Video Hidden Object Recovery

[1]Israa Hadi Ali and [2]Qasim Jaleel
[1]ITS College, [2]IT College,
Babylon University, Hillah, Iraq

Abstract: One of the main problems with video processing is the extraction of feature. Interest points are one of the feature types in the video. There are a number of algorithms that depend on the interest points in the stitching process. The aim of this study is to compare three methods to extract interest points for video hidden target recovery. Hidden target makes the target non-clear for a reason such as security cases or to save the privacy of the target in the video. The proposed method is implemented in two stages. Firstly, determining the hidden region in the target and create a large window around the desired target to track and discover the target parts that appear. Secondly, the stage is the stitching stage of the parts using one of the best three methods namely Harris corner detection, SURF and FREAK. After selecting the best image resulting from those methods and comparing the results. Experiments have demonstrated the efficiency and robustness of the Harris algorithm compared to the SURF and the FREAK in the proposed method.

Key words: Mean shift method, Harris corner algorithm, SURF, FREAK image stitching, interest point, homography, RANSAC and blend

INTRODUCTION

There are many ways to hide the target in the video and one those methods use blur (Vankawala *et al*., 2015). The blur in a video is one of the problems in video processing. In some cases, it is a result of fast motion object, camera shake, noise treatment that makes the object to the smooth (blur). It may be used for hidden face or object (Singh and Sahu, 2013). It is either non-intentional blur or intentional blur. Non-intentional blur such as motion blur or it's a result from noise treatment (Singh and Sahu, 2013). Intentional blur makes the target non-clear for a reason such as security cases or to save the privacy of the target in the video. It is creating by application programs such as editor video, Camtasia program and Wonder share Fillmore. This type of blur cannot be treated by traditional methods. In other cases, the blur either be a dark or a semi-transparent as shown in Fig. 1.

Software that creates the blur sometimes can't cover all intended target but it covers most part of it. So, these parts of the target appear beside the blur (hidden region). And sometimes the whole target is suddenly appearing in one of the frames. This problem explained by the simple example shown in Fig. 2. The proposed method exploits the release of the target parts and collects these parts by tracking methods and recover blur (hidden region) of the target.

Tracking the target of the main tasks in computer vision applications such as monitoring, robotics, etc.

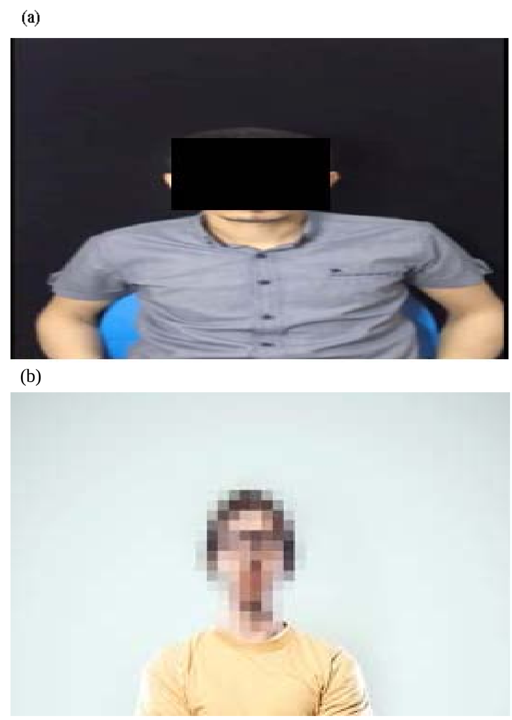

Fig. 1: The types of blur: a) Dark blur and b) Semi-transparent

Tracking of target is to specify the path and location of the target of the frame to frame in the video (Singh and Sahu, 2013; Divya and Ravi, 2016). In the proposed

Corresponding Author: Israa Hadi Ali, ITS College, Babylon University, Hillah, Iraq

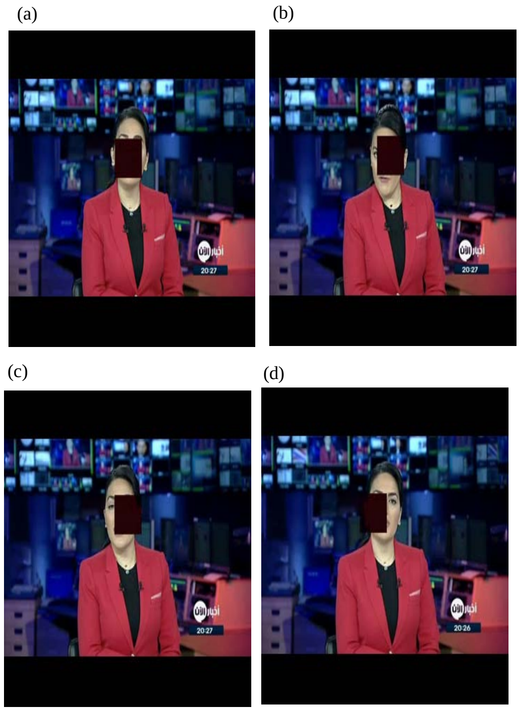

Fig. 2: Position of the hidden region: a) Up part of the target is appeared; b) Bottom part of the target appears; c) Right part of the target appears and d) Left part of target appears

method, use mean shift method to track the target and find the parts that may appear. Mean shift method used to track the non-static target based on the histogram. Tracking is based on the similarity of the target model by comparing its histogram. This method calculates the similarity between the target model and the candidate model through the histogram (Rawat and Raja, 2016; Lakshmeeswari and Karthik, 2016).

After the process of tracking the target and finding the parts that appear as a result of tracking the hidden target, then comes the stitching stage which uses one of the best three methods namely Harris corner detection, SURF (Speeded Up Robust Features) and FREAK (Fast Retina Key Points). These three methods are used to detection the interest points (features) in the image that are used to stitch process those images (parts).

Harris corner algorithm is using to detect the interest point in the image. By computing autocorrelation matrix method. The auto-correlation matrix is computing by Eq. 1. The 2*2 symmetric autocorrelation matrix used for detecting the interest point in the image (Peng *et al*., 2016; Awad and Hassaballah, 2016):

$$M(x,y) = \sum_{u,v} w(u,v) * \begin{pmatrix} I_x^2(x,y) & I_x I_y(x,y) \\ I_x I_y(x,y) & I_y^2(x,y) \end{pmatrix} \quad (1)$$

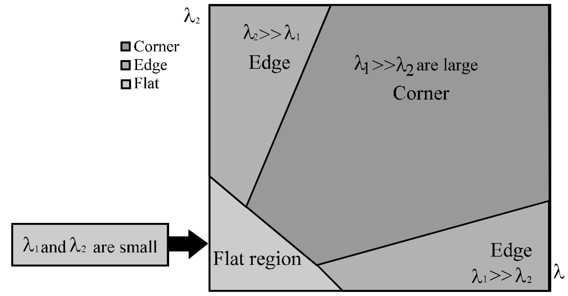

Fig. 3: Classification eigenvalue of the autocorrelation matrix

where, I_x and I_y are local part derivatives in the x and y directions. The eigenvalue of the matrix M is using to find the interest point by computing this matrix for each pixel in the image. If both eigenvalue is large that means is interesting point (corner) the shown in Fig. 3.

The SURF (Speeded Up Robust Features) has been developed by Bay *et al*. (2008) and Xu and Namit (2008). The SURF is depending integral image and Hessian blob detector. The main objective of integral image is to make the box filter easy to calculate. The filter box is sometimes called (filter mean) used to find an interest points by applying three filter types in three directions x, y, xy. The Hessian blob detector is called Hessian matrix that square matrix 2*2 in the second order derivation of density image (Oyallon and Rabin, 2015).

The Fast Retina Key Points Descriptor (FREAK) is developed by Alahi *et al*. (2012) is used to obtain interest points (key points) by describing binary descriptors (Youssef and Atta, 2016). FREAK works on the basis of gradient density as in the SURF algorithm and compare them in the form of binary attributes. Binary attributes are fast in calculations and few in storage (Emadi and Khachikian, 2016).

Homography is a mathematical relationship used as a map of pixel coordinates between the images to be stitched. The homography is mapping between two images in the same scene. Homography is used after detection of the interest point and a find the correlation the interest point between the images (Sturm, 2009).

RANSAC (Random Sample Consensus) is used to estimate robust model in the homography to reduce the calculation time during images stitching process. It is used to eliminate undesirable features in the area between the images to be stitched. Also, RANSAC is used to estimate the parameters of a mathematical model. It is an iterative method and is based on data that contains outlier values when outlier values do not affect the guessing process in the mathematical model (Szpak *et al*., 2015).

Literature review: The following are the selected works for their relationship to the proposed method. Talmale and Janwe (2016), the proposed shift monitoring set of rules is a method no parameters estimation. The proposed method is using mean shift for tracking the target was to identify the region to be tracked and to determine the search window and its location that containing the region to be tracked. Peng *et al.* (2016), the proposed improvement is the Harris algorithm in extracting the feature (detection of interest points). This proposal depends on two steps, the first extracting the target region to significant region detection. The second step is to use the Harris algorithm on the extracted area and this helps to detect the real corner which represents the feature.

Bay *et al.* (2008), the proposal is to create a new algorithm called Speeded Up Robust Features (SURF). The basis of the research of this algorithm represents a rotation-invariant and novel scale to detection the interest points and create descriptor of those points. Alexandre, the proposed is creating algorithm an FREAK (Fast Retina Key point) which takes a small memory and is suitable for complex calculations.

MATERIALS AND METHODS

Figure 4 explains the main steps of the mention method.

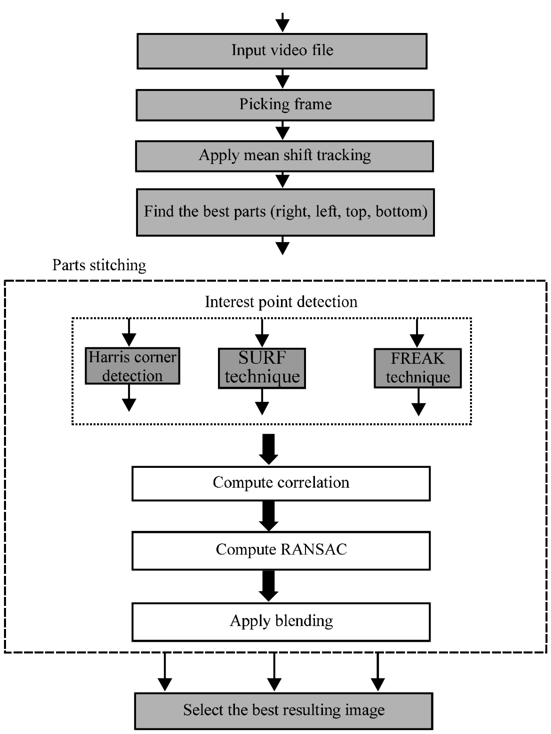

Fig. 4: Block diagram of the proposed method

Input video file: The first step of the method is taking video file that has extension "AVI".

Picking frame: The video file consists a series of frames (images). This proposed method takes all frames (images) in the video file to track and find the best parts of the target that might appear.

Apply mean shift tracking: The mean shift algorithm is used to track a target in a video scene. The algorithm in the proposed method is based on the search window and the target region. The search window is a large window that surrounds the target region and its main function is to restrict the search for the new location only within this window. In general, the search window has two functions. The first function is used in the tracking algorithm by the mean shift. The second function detects the target parts that appear by tracking the target. This tracking phase depends on the determining hidden region (blurring region). Determining hidden region (blurring region) is the first and very important step for tracking of the target. In this step locating the blurring region manually only in the first frame of video. Locating blur region that covers the target shows in Fig. 5. This step determines only blur region that it must be tracking. Next step, draw large

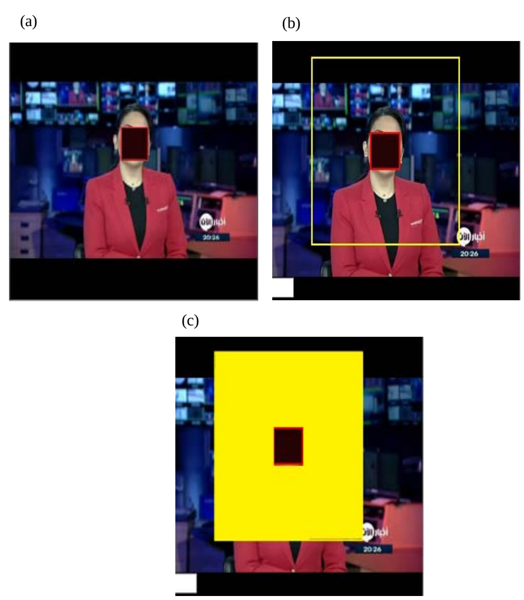

Fig. 5: Locating the target in the first frame: a) Locating the blur region which is the index in red; b) Determining window is greater than the blur region as the index in yellow and c) Represents an example showing the region of the search window with yellow color and this window is used to search for the parts that may appear

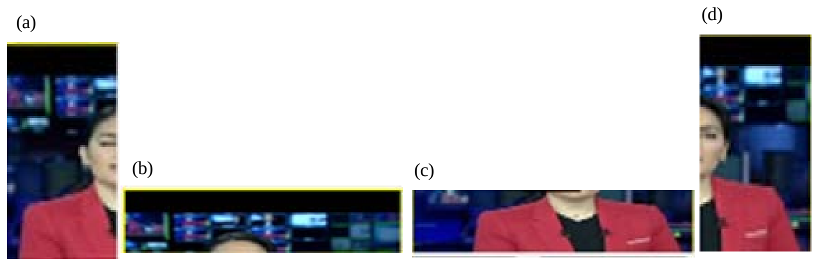

Fig. 6: The best parts have been found: a) Left part; b) Up part; c) Bottom part and d) Right part

window around blur region (search window). The large window is taken because the target is only required that it is covered with this the blurring region. This window is giving it capability for the following the target and assembles parts of this target which appear through video tracking.

Finding best parts (right, left, top, bottom): There are main steps to find the best parts of the target that appear. As explained in the following.

Finding difference frames: In the proposed method, use the background subtraction method. It is extensive use to detect object movement. In the proposed method, the method of subtracting the background is used to detect the target parts that may appear in a series of frames in the video. The background subtraction process is only in the search window. Using the mean shift algorithm, the target is tracking and at the same time, the search window is extracting and stored in a matrix. Each search window in each frame becomes a separate image. In addition, the size of the comparison and the time spent in finding parts decreased.

Splitting the search window into four parts: After collecting the frames that are differing from background frame. Next step, the search windows are divided in to four parts (right, left, top, bottom) each the frames only differed. The purpose of this division is to locate the part of the target that may appear. This division also gives ease of comparison and accuracy in the detection of parts of the target.

Finding the best parts: After that, compared each part in the search window against part of another search window that is differing. The aim from this compare is detection locate each part that appears and its locate. The resulting of this process is only four parts (right, left, top, bottom) (Fig. 6).

Parts stitching: After the process of the detection of the target parts that appear through the process of tracking after that comes the stage of stitching those parts. In the proposed method, the best method is chosen to perform the stitching process. There are three methods in the proposed method for the performance of the parts stitching process: Harris corner detection, SURF and FREAK. These methods are the main function is to detection the interest points in the image according to special criteria. Stitching the parts depends on several steps:

- The first step is the detection of interest points depending on the Harris corner detection, SURF and FREAK
- The second step is the matching of the extracted characteristics (interest points) between the parts to be stitching
- The third step is the use of RANSAC in the homography which is using to find the best match and noise processing
- The final step is a blending that addresses the problem of the edges of the parts when performing a stitching process

RESULTS AND DISCUSSION

The proposed method implementation has three main stage. The first stage determine hidden region (blur region) and it is tracked which is shown in Fig. 5. The second stage is detecting the best part that appeared which shown in Fig. 6. The third is stitching of the parts using one of the best three methods namely Harris corner detection, SURF and FREAK. The image of resulting also is depending in parts that appear which shown in Fig. 7. The resulting image using Harris method depends on the following steps first step is the detection of interest points which shown in Fig. 8. The second step is the matching of the interest points between the parts to be stitching which shown in Fig. 9. The third step is the compute of the RANSAC which shown in Fig. 10. The final step is a blending which shown in Fig. 11. The same steps are repeated for stitching the resulting parts with the rest of the parts where the resulting final image is

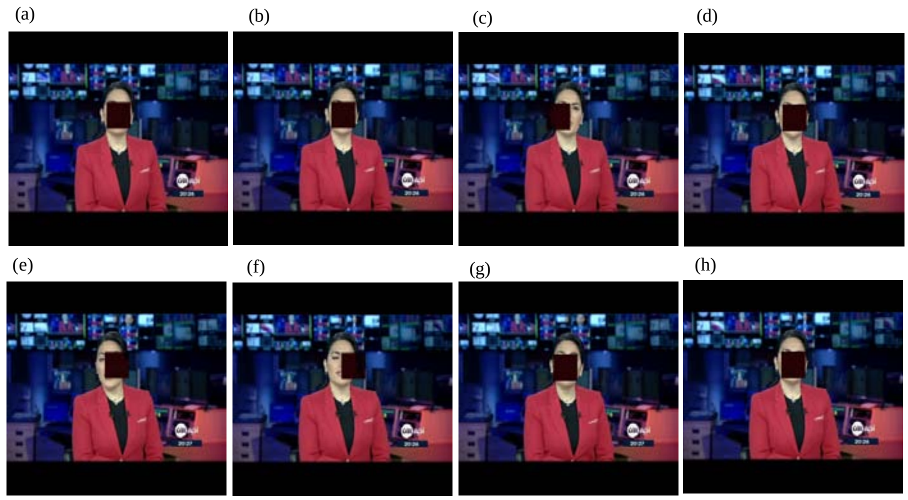

Fig. 7: a-h) Representing the shot that means a sequence of frames in the same scene

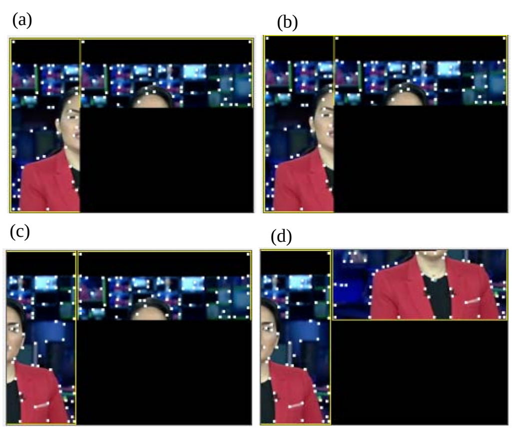

Fig. 8: The detection of interest points between each two parts: a) Left part and up part; b) Left part and bottom part; c) Right part and up part and d) Right part and bottom part

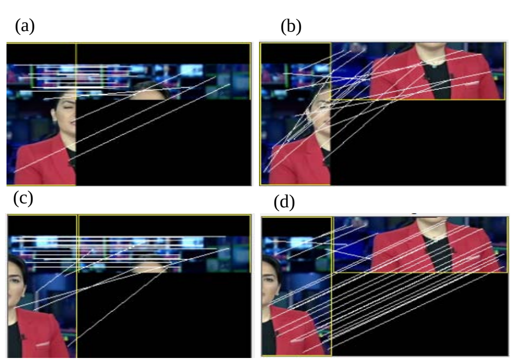

Fig. 9: The correlation between each two parts: a) Left part and up part; b) Left part and bottom part; c) Right part and up part and) Right part and bottom part

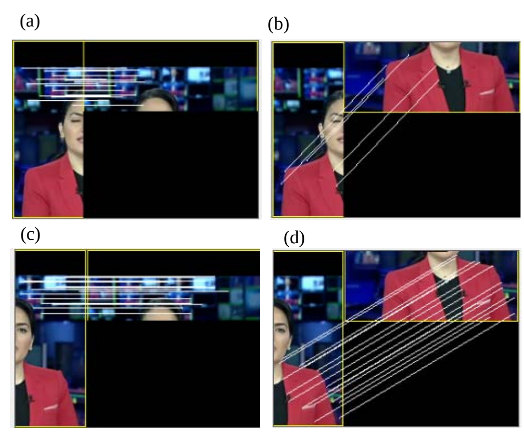

Fig. 10: The RANSAC between each two parts: a) Left part and up part; b) Left part and bottom part; c) Right part and up part and d) Right part and bottom part

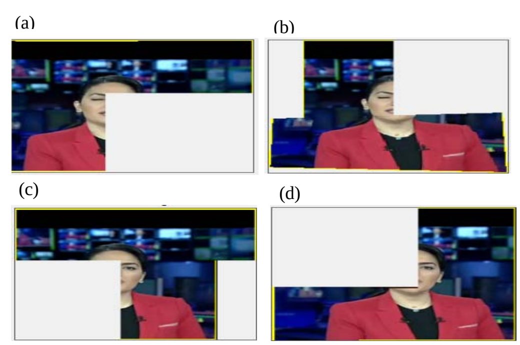

Fig. 11: Blending between each two parts using Harris: a) Left part and up part; b) Left part and bottom part; c) Right part and up part and d) right part and bottom part

Table 1: Results of the comparison of the methods (Harris, SURF and FREAK)

Variables	No. of interest Harris	No. of interest SURF	No. of interest FREAK	No. of correlation Harris	No. of correlation SURF	No. of correlation FREAK	No. of RANSAC Harris	No. of RANSACS URF	No. of RANSAC freak	Result Harris	Result SURF	Result FREAK
Left partand bottom part	104	42	357	17	12	87	5	4	Rang (5-7)	Accurate	Inaccurate	Accurate
Left partand up part	125	61	335	24	20	89	20	7	18	Accurate	Accurate	Accurate
Right partand bottom part	110	44	384	23	12	97	16	4	26	Accurate	Inaccurate	Accurate
Right partand up part	129	63	362	24	19	85	17	5	5	Accurate	Accurate	Inaccurate

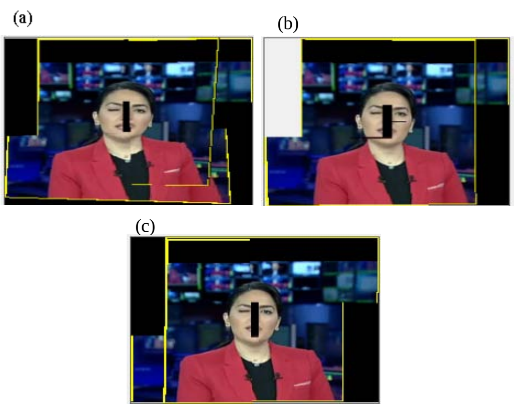

Fig. 12: Blending between each two parts using Harris: a) First part = (Left+bottom)+up +second part = right; b) First part = (Left+up)+right+second part = bottom and c) First part = (Right+bottom)+ up+second part = Left

Fig. 13: The blending using SURF (the left+the up part)+ bottom part with right

shown in Fig. 12. The resulting image using SURF method is based on the same previous steps and the resulting final image is shown in Fig. 13. One image is produced using SURF where the image is made up of parts (the left+the up part)+bottom part with right) and the other parts cannot be stitching. The resulting image using FREAK method is based on the same previous steps and the resulting final image is shown in Fig. 14. Table 1 shows the results of comparing the methods. The

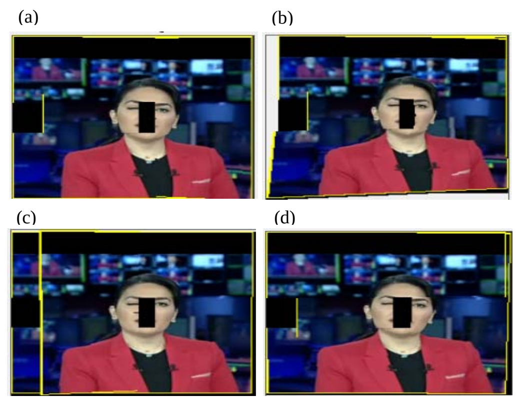

Fig. 14: The blending using FREAK: a) (The left+the up part)+bottom part with right; b) (The left part+the up part) and right with bottom; c) (The right part+the bottom part) and up part with left and d) (The left part+the bottom part) and up part with right

comparison is only made between two parts of the parts (left, bottom, right, up) to illustrate the efficiency of each method. The results of the table show that the number of interest points detected in the Harris is more than FREAK and SURF.

CONCLUSION

The proposed method is stitching of parts that appear. There are some conditions for stitching of parts to get image high accuracy, the first condition, it must be the overlapping between parts from 15-30% and the second condition the file of the video is high quality. The efficiency and accuracy of the resulting image depend on the size of the parts that appear by the target tracking in a series of frames in the video. The three methods (Harris, SURF and FREAK) are influenced by the quality of the video that will extract the frames (images). If the detected parts are flat (free of interest points), the blending process is not done correctly. Interest points are either corner or blob or as a result of a significant change in the color value of adjacent pixels and other situations. The proposed method can be developing to use another method for stitching of parts.

REFERENCES

Alahi, A., R. Ortiz and P. Vandergheynst, 2012. Freak: Fast retina keypoint. Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (CVPR'12), June 16-21, 2012, IEEE, Providence, Rhode Island, USA., ISBN:978-1-4673-1226-4, pp: 510-517.

Awad, A.I. and M. Hassaballah, 2016. Image Feature Detectors and Descriptors: Foundations and Applications. Vol. 630, Springer, Berlin, Germany, ISBN:978-3-319-28852-9, Pages: 436.

Bay, H., A. Ess, T. Tuytelaars and L.V. Gool, 2008. Speeded-Up Robust Features (SURF). Comput. Vision Image Understanding, 110: 346-359.

Divya, M. and A.V. Ravi, 2016. Single object tracking system by using labview. Intl. J. Recent Innovation Trends Comput. Commun., 4: 52-56.

Emadi, M. and S. Khachikian, 2016. Applying fast and freak algorithms in selected object tracking. Intl. J. Adv. Res. Electr. Electron. Instrum. Eng., 5: 5829-5839.

Lakshmeeswari, G. and K. Karthik, 2016. Survey on algorithms for object tracking in video. Intl. J. Comput. Appl., 141: 17-22.

Oyallon, E. and J. Rabin, 2015. An analysis of the SURF method. Image Process. Line, 5: 176-218.

Peng, W., X. Hongling, L. Wenlin and S. Wenlong, 2016. Harris scale invariant corner detection algorithm based on the significant region. Intl. J. Signal Process. Image Process. Pattern Recognit., 9: 413-420.

Rawat, N. and R. Raja, 2016. Moving vehicle detection and tracking using modified mean shift method and kalman filter and research. Intl. J. Technol. Res., 2: 96-100.

Singh, D. and M.R. Sahu, 2013. A survey on various image deblurring techniques. Int. J. Adv. Res. Comput. Commun. Eng., 2: 4736-4739.

Sturm, P., 2009. Geometric computer vision. Inria Grenoble Rhone-Alpes, Montbonnot-Saint-Martin, France. https://team.inria.fr/steep/files/2015/03/poly_3D_eng.pdf.

Szpak, Z.L., W. Chojnacki and A.V.D. Hengel, 2015. Robust multiple homography estimation: An ill-solved problem. Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 2015), June 7-12, 2015, IEEE, New York, USA., ISBN:978-1-4673-6964-0, pp: 2132-2141.

Talmale, S. and N.J. Janwe, 2016. Object tracking in images and videos. Intl. J. Eng. Comput. Sci., 5: 15482-15486.

Vankawala, F., A. Ganatra and A. Patel, 2015. A survey on different image deblurring techniques. Intl. J. Comput. Appl., 116: 15-18.

Xu, A. and G. Namit, 2008. SURF: Speeded-Up robust features. Master Thesis, McGill University, Montreal, Quebec.

Youssef, B.A. and E.H. Atta, 2016. Image forgery detection using freak binary descriptor and level set segmentation. Intl. J. Sci. Eng. Res., 7: 1-8.