2nd International Conference on Information and Communication Technology and Its Applications (ICTA 2018)
Federal University of Technology, Minna, Nigeria
September 5 – 6, 2018

Upright FAST-Harris Filter

Abdulmalik Danlami Mohammed[1], Adam Muhammed Saliu[1], Idris Mohammed Kolo[1], Adama Victor Ndako[1], Shafi'i Muhammed Abdulhamid[2], Abdulkadir Baba Hassan[3], and Abubakar Saddiq Mohammed[4]
[1]Department of Computer Science, Federal University of Technology, Minna, Nigeria
[2]Department of Cyber Security Science, Federal University of Technology, Minna, Nigeria
[3]Department of Mechanical Engineering, Federal University of Technology, Minna, Nigeria
[4]Department of Electrical/Electronic Engineering, Federal University of Technology, Minna, Nigeria
abdulmalik.mohammed@futminna.edu.ng, adambinlaf@futminna.edu.ng, idris.kolo@futminna.edu.ng, vnadama@futminna.edu.ng, shafii.abdulhamid@futminna.edu.ng, abhassan@futminna.edu.ng, abu.sadiq@futminna.edu.ng

Abstract—**The traditional approaches to the classification of image regions suffer drawbacks in the face of imaging conditions (occlusion, illumination changes, rotation, viewpoint changes and image blurring) and thus contribute to the poor performance of several vision based applications such as object recognition, object tracking, image retrieval, pose estimation, camera calibration, 3D reconstruction, Structure from motion, stereo images and image stitching. In this work however, feature points extraction method by decomposition of image structure is employed in order to overcome these challenges. The decomposition of an image structure into feature set enhances the performance of many vision-based applications and system. Our feature point extraction method which we refer to as Upright Feature from Accelerated Segment Test with Harris filter (UFAH) in this text, works by combining Feature from Accelerated Segment Test detector with Harris filter. The result obtained in the evaluation process shows that UFAH is robust and also invariant to imaging conditions (i.e rotation, illumination changes and image blurring).**

Keywords--image analysis; feature points; repeatability; Harris filter; FAST; Upright FAST

I. INTRODUCTION

In image analysis, the focus is on the extraction of information about the contents of an image and thus serves as the first step required to simplify additional tasks such as object recognition, object tracking, pose estimation, 3D reconstruction, camera calibration, structure from motion, image stitching, image retrieval and stereo images. Typically, such valuable information corresponds to an image area or region that has unique properties or distinctive structures. These structures include edges, blobs, corners and object contours. The collection of these structures or patterns is referred to as features. In feature extraction method, large dataset is decomposed into smaller ones in order to enhance the performance and speed of processing of many visions based applications.

According to the work presented by [1], feature points, which correspond to a particular structure or pattern in an image, should have some or all of the following properties:

A good feature point should be repeatable between two images captured under different imaging conditions such as illumination change and image rotation.

A good feature point should be surrounded by local image structure that is highly informative and distinctive to enable feature matching.

The location of a good feature point in image should be well defined.

While different feature extractors exist for different image structures, many of the extractors tend to detect feature points with the inclusion of some or all of the properties mentioned above. These extractors are categorized according to the type of image structure that they are designed to detect. For instance, feature extractors that are designed to detect an edge-like feature in an image can be referred to as edge detectors, while blob and corner based feature extractors are referred to as blob and corner detectors respectively. It is however worth mentioning here that the success of implementing a particular extractor in one application may not necessarily yield the desire result in another application because of the kind of image structure each extractor is designed to detect. In addition, the unstable state of some of these image structures under varied imaging conditions such as illumination, rotation and viewpoint can degrade the performance of most of these feature extractors. For example, the disparities in the gradient values in certain directions of an edge caused by an image rotation can led to the extraction of false features and thus degrade the performance of a feature extractor. In this work we extend the work of Feature from Accelerated Segment Test detector known as FAST to include Harris filter in order to overcome the challenges faced with the extraction of feature points from an image. The purpose of this integration is to extract stable features that will enhance object recognition and suppress those features that are considered to be false. We call this improved version of FAST an Upright FAST with Harris filter(UFAH).

The remainder of this text is as follows: Section 2 describes some of the related works with regards to feature extraction and detection. In Section 3, an overview of the

Feature from accelerated Segment Test detector known as FAST is presented. Section 4 introduces our enhance feature extractor-an extension of the original FAST detector. While Section 5 presents performance evaluation of the UFAH and other state of the art detectors, Section 6 discusses further on the outcome of the evaluation process. Finally, Section 7 concludes the discussion of this study.

II. RELATED WORKS

While feature extraction is at the core of many vision based application, developing a robust and efficient feature extractor in the context of vision based application on camera phones is a challenging task due to the low processing power of these devices. However, in recent time, few works have gone a step further to simplify the computation of feature points on a mobile phone with the aim of achieving real time performance and invariance to image transformations (e.g scale change, image rotation, illumination variation, and image blurring). For example, the Oriented FAST and rotated BRIEF proposed by Rublee et al [2], uses the FAST keypoints detector to detect corners, while the orientation of the detected keypoints is computed using the intensity centroid making the work invariance to rotation. However, the computation of the feature orientation using the intensity centroid is computational expensive and thus increases the rate of processing in low memory devices. Another promising work referred to as Binary Robust Invariant Scale Keypoint (BRISK) is presented by Leutenegger et al [3]. BRISK is a scale invariant feature detector in which keypoints are localized in both scale and image plane using the modified version of FAST proposed by [4]. In BRISK, the strongest keypoints are found in octaves by comparing 8 neighboring scores in the same octave and 9 scores in each of the immediate neighbouring layers above and below. To determine the true scale of the keypoints, the maximum score is sub-pixel refined in all the three layers followed by a 1D parabola fitting along the scale-axis. In BRISK, the computational requirement for locating feature points in both scale and image plane is a drawback in particular for devices with low memory capacity. The Fast Retina Keypoints (FREAK) proposed by [5] is an improvement over the sampling pattern and binary comparison test approach between points used in BRISK. The pattern of FREAK is motivated by the retina pattern of the eye. In contrast to BRISK, FREAK employs a cascade approach for comparing pairs of points and uses 128 bits as against the 512 bits obtained in BRISK to enhance the matching process. The Maximal Stable Extremal Regions referred to as MSER is proposed by Matas et al [6]. In this method, blob-like feature points are extracted from a set of thresholded images. The method is invariant to affine transformation but sensitive to illumination changes. Speeded up Robust Feature known as SURF is a robust feature detector and descriptor based on the Hessian matrix and proposed by Bay et al [7] . It has a wide area of applications that include object recognition, camera calibration, image registration, 3D reconstruction and objet tracking. While SURF is partly motivated by SIFT, the computational requirement as characterized in the computation of the local points has made this detector unsuitable for devices with low memory resources. The Scale Invariant Feature Transform referred to as SIFT is a scale and rotation invariant feature detector and descriptor that is proposed by David Lowe [8]. While SIFT has a wide area of applications in object recognition, image stitching, stereo image, image tracking and 3D reconstruction, it is computational requirement is unsuitable for devices with low processing capability. Center Surround Extrema kown as CenSurE is a another feature detector proposed by [9]. Here feature points are computed by finding the extrema of the Laplacian across multiple scales using the full spatial resolution of the original image. The method is invariant to rotation and fast to compute. However, its implementation on a mobile phone is yet to be reported as the time of this text.

III. FEATURES FROM ACCELERATED SEGMENTED TEST

Features from Accelerated Segment Test (FAST) detector proposed by Rosten et al [10] are used to extract a local image structure that corresponds to a corner. FAST works by comparing the intensity values of a pixel with its circular neighborhood of pixels. Given an image pixel P with intensity $\boldsymbol{I_p}$ surrounded by a circle of 16 pixels labeled from 1-16 in a clockwise direction and a threshold value $\boldsymbol{T}$, pixel $\boldsymbol{P}$ is considered a corner, if a set of $\boldsymbol{N}$ consecutive pixels in the circle are above $\boldsymbol{I_p + T}$ or below $\boldsymbol{I_p - T}$. The two conditions that have to be met for a point to be considered a feature in AST can be expressed as follows:

$$\forall\, \boldsymbol{x} \in \boldsymbol{S}, \boldsymbol{I_x} > \boldsymbol{I_p} + \boldsymbol{T} \quad (1)$$

$$\forall\, \boldsymbol{x} \in \boldsymbol{S}, \boldsymbol{I_x} < \boldsymbol{I_p} - \boldsymbol{T} \quad (2)$$

Where $\boldsymbol{S}$ denotes the set of N consecutive pixel and $\boldsymbol{x}$ denotes any pixel within **S**, and $\boldsymbol{I_x}$ intensity value of the $\boldsymbol{x}$ pixel. The initial implementation of FAST set $\boldsymbol{S}$ to 12 because of the high-speed test it offers and thus removing a significant number of non-corners. However, the high- speed FAST detector has weaknesses in terms of the number of rejected candidate points when S < ***12*** and the efficiency is affected by the order in which the 16 pixels are compared. To solve these challenges a machine learning approach is employed.

It is important to mention here that the parameter $\boldsymbol{T}$ used for the test determines the sensitivity of the corner response. For example, a small value of $\boldsymbol{T}$ will result in large numbers of corners and vice versa. The Feature from Accelerated segment test detector is not only simple but also computationally efficient about feature detection. It is thus widely used in real-time object recognition.

In practice however, some of the features returned by the FAST detector are not accurate representation of an image feature given their instability in the face of image deformation. Hence, in our proposed method which is described in the section that follows, we extend the FAST feature detector to include corner filter that will detect features that have high gradient value in all directions thus providing a distinctive and robust features required for object recognition and image retrieval.

IV. UPRIGHT FAST WITH HARRIS FILTER

In the previous section, it is observed that, not all feature points detected by FAST have strong corner strength, since some of them represent edges. In order to overcome this challenge, we extend the FAST feature extractor by integrating the Harris filter to the original implementation of the FAST method. In this way, the corner strength of each of the detected features as returned by the FAST method is measured using the Harris filter [1] and the corner with the strongest strength is extracted as being the strongest feature. We referred to this extended version of the FAST as Upright FAST with Harris filter (UFAH).

Given a set of key points locations, the corner strength of each feature points at that location is computed by comparing the value $\boldsymbol{R}$ (minimum between the eigenvalues λ_1, λ_2 of the second order matrix) with a threshold value $\boldsymbol{T}$ such that if the value of $\boldsymbol{R}$ is greater than a threshold $\boldsymbol{T}$ a corner is found.

$$\text{R} = \min(\lambda_1, \lambda_2) \tag{3}$$

$$CornerStrength = \begin{cases} R > T = 1 \\ R < T = 0 \end{cases} \tag{4}$$

However, we observed that comparing the minimum value of the eigenvalues against a threshold to find corner from a set of detected feature points does not give an accurate result and the desired number of keypoints. Hence, in order to obtain an accurate estimate of corner strength we employ the following corner response function [10]:

$$F(x, y) = \det(M) - k\ trace^2\ (M) \tag{5}$$

Where $\det(M)$ denotes the determinant of the matrix M and $trace\ (M)$ represents the trace of the matrix. $F(x, y)$ is called the corner response function. This function returns a maximum value at isolated corners. k is assigned a constant value of 0.04 which determines the sensitivity of the detector. $\det(M)$ and $trace^2\ (M)$ can be estimated using the following eigenvalue decomposition theory:

$$\det(M) = \lambda_1 * \lambda_2 \tag{6}$$

$$trace(M) = \lambda_1 + \lambda_2 \tag{7}$$

Figure 1b shows images of feature points detected using the Accelerated Segmented Test detector without corner response function, while figure 1c shows the same image containing feature points that are detected using AST combined with the corner response function.

In Figure 1b, a total of 416 keypoints were identified using the AST detector. A large number of features are expected since they not only represent corners but also edges. When the image containing the initial detected points is passed to AST with corner response function (see algorithm 4), a total number of 200 good feature points were returned. These feature points are drawn in red in the image shown in figure 1c. This shows that only the strongest points corresponding to corners are returned.

V. PERFORMANCE EVALUATION OF UPRIGHT FAST WITH HARRIS FILTER

The performance of the Upright FAST with Harris filter(UFAH) as proposed in this text along with the other techniques designed specifically for mobile devices is evaluated on a dataset of images obtained from [12] (provided by the robotics research center of the University of Oxford). These images are observed under different image transformations such as rotation, illumination and image blur. Each dataset is made up of five images for which the first image in the set is considered as the reference image. Figure 2 shows the reference image in each of the dataset.

It is worth mentioning here that, in order to give a fair comparison, the default values of all the detectors as described in their original implementations are use in this experiment. Furthermore, to ensure that an equal number of feature correspondence is obtained for each detector irrespective of the image transformation, the detectors are configured in a way that they can only detects a maximum of 1000 feature points per every image. In this work however, we use a standard metric referred to as repeatability to evaluate the performance of the UFAH detector and compare it with other state of the art detectors. Repeatability is expressed as the number of repeatable feature points between images. Since the images are planar, the feature correspondence (i.e repeatability) is computed in the overlap area where the transformed images are correctly mapped to the reference image (the first images in each dataset). The result of the evaluation using the different image transformation (rotation, illumination and blur) is shown in figure 3.

The repeatability test obtained for all detectors (BRISK, ORB, SIFT, UFAH and USURF) on a dataset of images observed under different angle of rotation is shown In figure 3a. As can be seen from the graph (figure 3a), UFAH has the highest number of feature correspondence and thus outperformed BRISK, SIFT, USURF and ORB. A similar test is performed but on different images with varying illumination changes and the repeatability score obtained for all the detectors is shown in figure 3b. The graph (figure 3b) indicates that UFAH has the best performance compared to the remaining detectors. Figure 3c shows the repeatability test obtained for both UFAH, SIFT, USURF, ORB and BRISK using images with increasing amount of blurring. The graph in figure 3c shows that UFAH has the highest repeatability score and as a result perform better than the remaining detectors.

VI. DISCUSSION

As observed in our performance evaluation section, UFAH shows better performance in terms of repeatability and the number of feature correspondence. For instance, in figure 3a, the graph shows UFAH with the highest repeatability score when images are observed under different angles of rotation-an indication that UFAH is invariant to rotation. The repeatability score and the number of correspondence under illumination changes were obtained for all detectors by decreasing the brightness of the images. In the test, UFAH recorded the highest repeatability score and has the highest number of feature correspondence followed by BRISK, and USURF. The test result in this case indicates also that UFAH is invariant to illumination changes. In the final test, we considered the repeatability score and the number of feature correspondence for all detectors on image blur and the results obtained show that UFAH has the highest repeatability score followed by BRISK and USURF. The result further affirmed that UFAH is invariant to image blur.

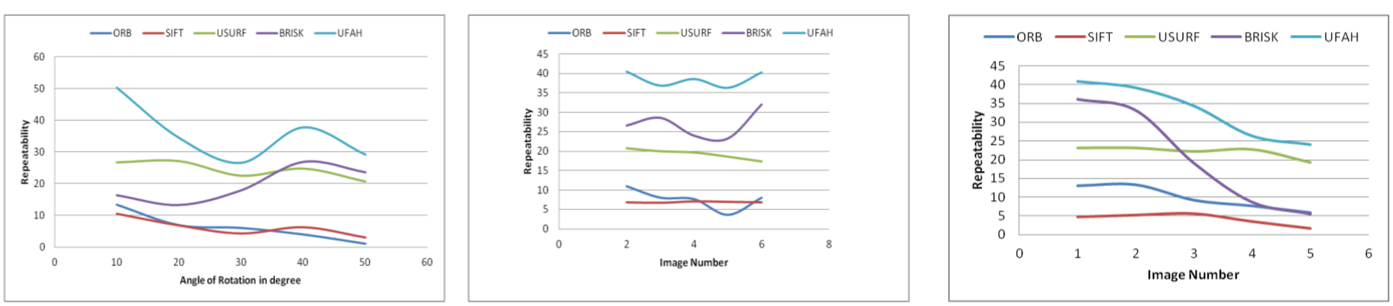
VII. CONCLUSION

In this paper, we are able to demonstrate the simplicity and effectiveness of our proposed feature extraction (UFAH) technique and by extension its suitability on mobile devices. We achieved this by integrating the Features from Accelerated Segment Test detector (FAST) with Harris filter in order to enhance feature extraction and improve feature points matching between a pair of images. This is an important step especially for tasks such as object recognition, image retrieval and 3D reconstruction. The results are promising as demonstrated from the performance evaluation section. However, as the advancement in mobile phone technology continues, future work will include the expansion of the UFAH to allow for the description of the extracted feature with minimal computational requirement.

Figure 1. (a) Original image showing an exit sign along a corridor of the Kilburn building of the University of Manchester. (b) Keypoints detected using FAST detector. The white circles indicate keypoints that also include other interest points such as edges. (c) In this Image, the red circles show the strongest corner returned after applying a corner response function to the detected feature from AST

(a) (b) (c)

Figure 2. (a) Image rotation obtained by rotating the camera around its optical axis. (b) The image with illumination change is obtained through the camera aperture. (c) Image blur is obtained using the camera focus[12]



(a) (b) (c)

Figure 3. The repeatability curve obtained for UFAH, ORB, SIFT, USURF and BRISK detectors on dataset of images observed under (a) Image rotation (b) Illumination changes and (c) Image blurrin

REFERENCES

[1] T. Tuytelaars and K. Mikolajczyk,. Local Invariant Feature: A survey. Foundation and Trends in Computer Graphics and Vision, vol. 3, no. 3, pp177-280. 2007.

[2] E. Rublee, V. Rabaud, K. Konolige, G. Bradski. ORB: Oriented FAST and Rotated BRIEF. IEEE International Conference on Computer Vision (ICCV), 2011.

[3] S. Leutenegger, M. Chli and R. Y. Siegwart. BRISK: Binary robust invariant scalable keypoints. IEEE International Conference on Computer Vision (ICCV), 2011.

[4] E. Mair, G.D.Hager, D.Burschka, M. Suppa, and G.Hirzinger. Adaptive and generic corner detection based on the accelerated segment test. In Proceeding of the European Conference on Computer Vision, vol 2 no. 5, pp. 2010.

[5] A. Alahi, R. Ortiz, P. Vandergheynst. FREAK: Fast retina keypoint. IEEE Conference on Computer Vision and Pattern Recognition, pp510-517. 2012.

[6] M. Donose and H. Bischof. MSER: Efficient maximally stable extremal region tracking. In Proceeding of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), pp. 553-560, 2006.

[7] H.Bay, T. Tuytelaars and L.Van Gool. SURF: Speeded up robust features. In Proceeding of the European Conference on Computer Vision, pp404-417. 2006.

[8] D.G. Lowe. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, vol 2, no. 60, pp91-110. 2004.

[9] M. Agrawal, K. Konolige and M. R. Blas. CenSurE: Centre surround extrema for real time feature detection and matching. In Proceeding of the European Conference on Computer Vision, 2008.

[10] E. Rosten, R.Porter and T. Drummond. Faster and better: A Machine learning approach to corner detection. IEEE Transaction In Pattern Analysis and Machine Intelligence, vol 32, pp105-119. 2010.

[11] C. Harris and M. Stephens. A Combined Corner and Edge Detector. Alvey vision conference, pp147-151. 1988.

[12] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. IEEE Transaction on Pattern Analysis and and Machine Intelligence, vol. 27, no, 10 pp1615-1630. 2005.