Intercontinental Geoinformation Days

http://igd.mersin.edu.tr/2020/

A Comparative Analysis of Speeded Up Robust Features (SURF) and Harris Algorithms in Point Cloud Generation

Ramazan Alper Kucak[1], İrem Yakar[*1], Serkan Bilgi[1], Serdar Erol[1], Adalet Dervisoglu[1]

[1]dÔÔÁÎÂÕÌ 4ÅÃÈÎÉÃÁÌ 5ÎÉÖÅÒÓÉÔÙȟ &ÁÃÕÌÔÙ ÏÆ #ÉÖÉÌ %ÎÇÉÎÅÅÒÉÎÇȟ $ÅÐÁÒÔÍÅÎÔ ÏÆ 'ÅÏÍÁÔÉÃÓ %ÎÇÉÎÅÅÒÉÎÇȟ)ÓÔÁÎÂÕÌȟ 4ÕÒËÅÙ

Keywords
Structure from Motion
Photogrammetry
SURF Algorithm
BRISK Algorithm
Harris Algorithm

ABSTRACT

The precise documentation and preservation of cultural heritage elements are of fundamental importance since they represent a social, historical and cultural improvement of societies throughout history. There have been different documentation techniques from past to present such as classical architectural drawing methods, classical surveying methods, terrestrial photogrammetry and aerial photogrammetry, etc. Photogrammetry emerges as a popular method with the developing technology. On the other hand, different evaluation techniques have been evolved to evaluate the collected photogrammetric data especially with the improvements in computer vision technologies. In this aspect, Structure from Motion (SfM) became a widely used approach for producing a 3D view from photogrammetric images. This paper presents the comparison of point clouds produced by Speeded Up Robust Features (SURF) and Harris Algorithms. The point cloud production was carried out in MATLAB environment. 20 images at total captured by Nikon Coolpix A10 digital camera were used for the point cloud generation. Therefore, the resulted points clouds were investigated comparatively.

1. INTRODUCTION

Cultural heritage according to UNESCO can be described as the product and the process that provides societies with rich resources that are transmitted from the past, created in the present, and endowed for the benefit of future generations. It consists of three parts: tangible, natural, and intangible heritage. Tangible cultural heritage refers to monuments such as sculptures, paintings, structures of an archeological nature, historical buildings, archeological sites, etc.

The management of these tangible cultural heritage elements is crucial due to cultural and social elements that they contain besides their historical importance. In this context, modeling the historical monuments in a detailed way is a key point in terms of protecting these features and ensuring intergenerational transmission. Many different techniques have been used throughout history ranging from classical surveying methods to photogrammetry etc. for acquiring and evaluating data. In this context; photogrammetric techniques have become highly preferred for documenting the tangible cultural heritage compared with the classical methods.

Photogrammetry is a method that relies on the determination of the 3-dimensional (3D) geometry of objects by measuring and analyzing the 2-dimensional (2D) images. Usually, it is divided into two categories as close-range photogrammetry and aerial photogrammetry. In close-range photogrammetry, images are acquired from the locations near or on the surface of the earth. The close-range photogrammetric data provide detailed dimensional information of objects. In contrast, aerial photogrammetric data are collected via overhead shots from an aircraft and provides land use details (Jiang et al. 2008).

On the other hand, there are various approaches and innovations for the evaluation of photogrammetric data. Especially with the developing computer vision technology, different approaches have been developed for the production of 3D models from 2D photogrammetric images. These approaches usually are based on key point detection and descriptor algorithms. The most used of these algorithms can be sorted as SURF (Bay et al. 2008), Harris (Harris, C. G., & Stephens, M.,1988), Shift (Lowe 2004), BRISK (Leutenegger et al. in 2011), AGAST (Mair et al., 2010), etc.

* **Corresponding Author**

(kucak15@itu.edu.tr) ORCID ID 0000 - 0002 - 1128 - 1552
* (yakari@itu.edu.tr) ORCID ID 0000 – 0002 – 7823 – 9674
(bilgi@itu.edu.tr) ORCID ID 0000 – 0002 – 8396 – 3202
(erol@itu.edu.tr) ORCID ID 0000 – 0002 – 7100 - 8267
(adervisoglu@itu.edu.tr) ORCID ID 0000 – 0001 – 7455 – 4282

Cite this study

Kucak A., Yakar İ., Bilgi S., Erol S & Dervisoglu A. (2020). A Comparative Analysis of Speeded Up Robust Features (SURF) and Harris Algorithms in Point Cloud Generation. Intercontinental Geoinformation Days (IGD), 5-8, Mersin, Turkey

It is possible to say the following determinations in the literature research about the algorithms. SURF algorithm is fast in the image matching process, but it is not enough in the feature point extraction (Cheng et al. 2017). However, the Harris Algorithm is an efficient corner detection algorithm, but it cannot enough the issue of scale (Cheng et al. 2017). Binary Robust Invariant Scalable Keypoints (BRISK) is presented by Leutenegger et al. in 2011. The corners are detected by the use of the AGAST algorithm. Filtering is realized with the FAST Corner score while determining the maxima in the scale-space pyramid. The BRISK algorithm is constructed as a binary string to provide illumination invariance. BRISK features are scale and rotation invariant while it has limited affine invariance (Tareen and Saleem 2018).

In addition, Structure from Motion (SfM) is one of the widely used methods for 3D modeling in computer vision software packages (Aicardi et al. 2018). Despite that SfM generally uses computer sciences, it relies on photogrammetry (Kuçak et al. 2017; Jebara et al. 1999). In this study; using SURF and Harris Algorithms in MATLAB environment, point clouds were produced with the SfM technique. The results were examined comparatively.

2. METHOD

2.1 Structure from Motion (SfM)

In the traditional photogrammetric approach; the 3D location and pose of the camera(s) or 3D coordinates of various control points are necessary to determine scene triangulation and to reconstruct the related object. Nevertheless, the Structure from Motion (SfM) method determines the camera pose and scene geometry simultaneously by the use of bundle adjustment which relies on matching features in adjacent overlapping images (Westoby et al. 2012). The initial estimation of camera positions and object coordinates can be done by tracking common features on adjacent images. Then, these positions and coordinates are refined iteratively by the use of non-linear least-squares minimization (Snavely 2008). The SfM approach basically stands for the images taken from a moving sensor. Thus, the method is most proper where the image pairs have a high degree of overlap. In such cases, the scene would be viewed from a wide range of positions that provide a fully three-dimensional structure (Westoby et al. 2012).

SfM Method is based on the epipolar geometry since it relies basically on the photogrammetric approach.

The intersection of epipolar line as well as COP1, COP2, and P points are illustrated in "Fig.1". Suppose that $P_1(u_1, v_1, 1)$ and P_2 $(u_2, v_2, 1)$ are the projection of an individual point in two different images. The point P_2 which is the projection of point P_1 in a single image is limited with the corresponding epipolar plane.

The fundamentals of epipolar geometry are as follows:

$$P_1^T F P_2 = 0 \qquad \textbf{(1)}$$

Equation 1 represents a least-squares minimization where; F is the fundamental matrix with 9 parameters that are restrained to have rank 2 ($\|F\| = 0$). Also, the matrix specifies an epipolar line from an image for a point in an image. These perspective projections constitute the basics of SfM algorithms (Kuçak et al. 2017; Jebara et al. 1999).

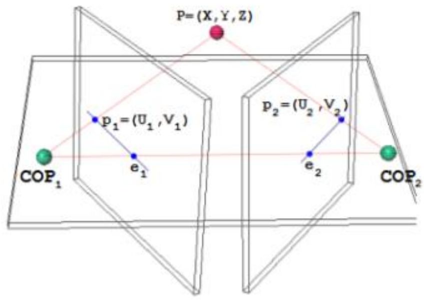

Figure 1. Epipolar Geometry (Jebara et al. 1999)

2.2 Speeded Up Robust Features (SURF)

The Speeded Up Robust Features (SURF) algorithm was presented by Bay et al. 2008. The method is basically based on Gaussian scale-space analysis. SURF relies on the determinant of the Hessian Matrix. The detector improves the feature-detection speed by utilizing integral images. Each feature in a specific neighborhood is described with a distribution of Haar wavelet in the 64 bin descriptor of SURF. The SURF features are rotation and scale-invariant. The descriptor of 128 bin values can be used to deal with the large changes in viewpoints. Hessian Matrix was given in Equation 2:

$$H(x,\sigma) = \left[L_{xx}(x,\sigma)\ L_{xy}(x,\sigma)\ L_{xy}(x,\sigma)\ L_{yy}(x,\sigma)\right] \quad (2)$$

Where; "Lxx(x,σ): the convolution of Gaussian second-order derivative with the image "I" in point "x", as well as "Lxy(x, σ)" and "Lyy (x, σ)" (Tareen and Saleem 2018).

2.3. Harris Algorithm

Harris and Stephens have presented the Harris detector in 1988. It basically specifies the corner points in an image. The gradient of an image is estimated for each pixel. Then, the correlation matrix is obtained. The robustness of the detection is improved by smoothing the elements in the matrix with Gaussian function. Equation 3 is used for calculating the response of each pixel. The pixel is considered a corner, only when both eigenvalues λ_1 and λ_2 are large (Kok and Rajendran 2018).

$$H = \lambda_1 \lambda_2 - k * (\lambda_1 + \lambda_2)^2 \quad (3)$$

3. RESULTS

Image registration can be defined as matching, aligning and overlaying multiple integral images captured from different locations. It is widely used in various vision-based applications. The registration phase can be divided into the following stages: feature detection and description, feature matching, outlier rejection, and image reconstruction. The selection of feature detector-descriptor plays an important role in registration since the timing and accuracy are affected by the computational capability and robustness of the selected detector-descriptor. Hence, it is crucial to choose the most proper detector-descriptor in feature-matching applications (Tareen and Saleem 2018).

In this study, SURF, Harris and BRISK Algorithms were performed on MATLAB environment and resulted in point clouds were examined comparatively.

SURF detects the points based on integral images and the Hessian matrix. The matrix approximates the Gaussian second-order derivative with box filters. The second phase of SURF is the orientation assignment. A circular region is constructed around the detected point, to determine orientation. Then, interest point descriptors are constructed by computing Haar wavelet responses and by extracting square windows around the points. On the other hand, the calculation of each pixel gradient is the basis of the Harris corner detector algorithm. The pixel is determined as a corner if the absolute gradient values in 2 directions are both great (Dawood et al. 2012).

The images used in the study were captured by using Nikon Coolpix A10 digital camera featured a 16.1 megapixel 1/2.3" CCD sensor. 20 images that belong to one facade of the monument were used during the study.

The selected monument for 3D reconstruction is the Fountain of Ahmed III (Figure 2). This fountain was built in 1729 by the Sultan Ahmet III, upon the proposal of the grand vizier Nevşehirli Damat İbrahim Pasha. It is located in front of the Bab-ı Hümayun gate of Topkapı Palace in the Fatih district of Istanbul.

Figure 2. The selected monument for 3D reconstruction

The camera calibration was performed in MATLAB environment by using 12 images at total of MATLAB camera calibration checkerboard captured with Nikon Coolpix A10 digital camera as shown in figure 3.

14.635 key points have been used while aligning the photos in the SURF detector. 1143 key points have been used in the same phase in the Harris detector (Figure 4 & 5). On the other hand, the BRISK Algorithm with the same properties was not successful in aligning the photos since it could not find any key points.

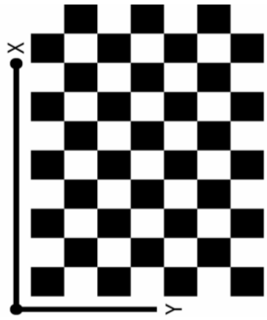

Figure 3. MATLAB camera calibration checkerboard

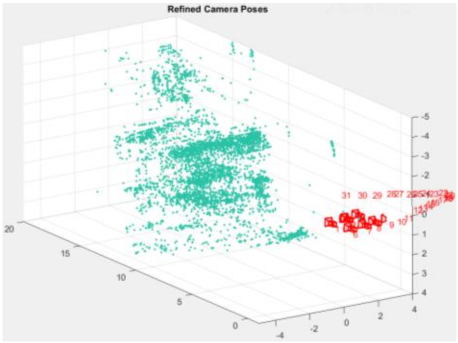

Figure 4. The refined camera poses in the SURF Algorithm

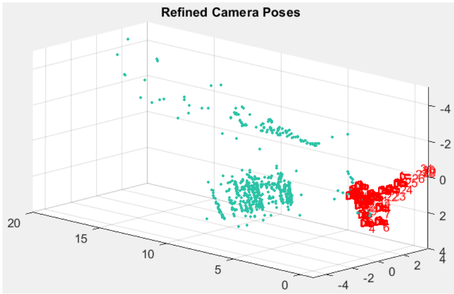

Figure 5. The refined camera poses in Harris Algorithm

In feature matching, ORB (Oriented FAST and Rotated BRIEF) (Rublee, E., et al., 2011) and BRIEF descriptors were used. The Hamming distance was used for the comparison of binary features. In this aspect, an operation in collaboration with a bit count on the result was performed to compute the binary data. The distance between the pairs of binary features could be computed effectively. On the other hand, performing a linear search for matching can be useful only for small datasets. The solution in such cases would be changing the linear search with an approximate matching algorithm (Muja and Lowe 2012).

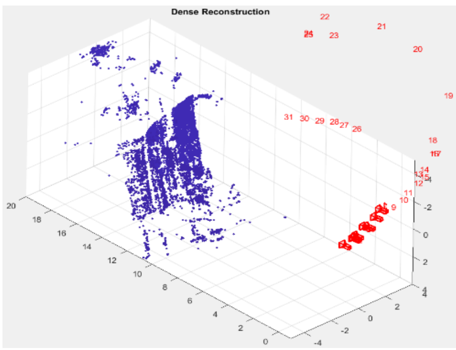

Figure 6. Dense reconstruction in SURF algorithm.

The resulted dense cloud in figure 6 in the SURF Algorithm belongs to one facade of the monument was exported ".ply" file format and imported to Cloud compare software, as shown in figure 7.

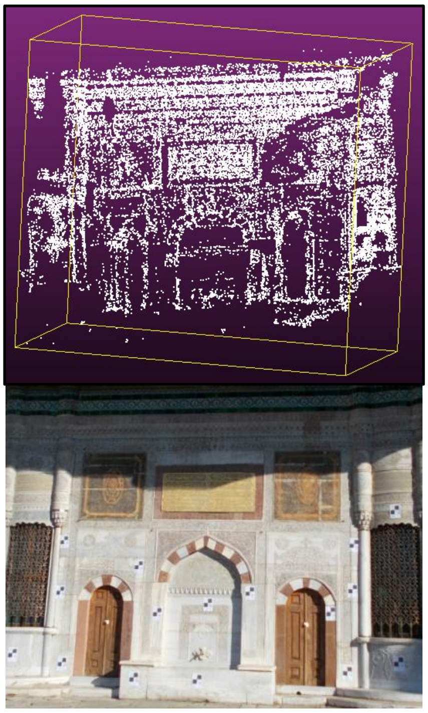

Figure 7. Resulted point cloud in Cloud Compare Software

4. DISCUSSION & CONCLUSION

In this study; using SURF and Harris in a MATLAB environment, point clouds were produced using the SfM Technique. The results were examined comparatively. It was found that that the number of incorrect points in Harris Algorithm was larger in comparison with the SURF Algorithm while the number of dense points was less. The integrity and representation of the point cloud obtained by the Harris Algorithm were also weak. The dense point cloud can be obtained by using different key point detection algorithms in the MATLAB environment. Also, more precise work could be obtained by changing the limit values in SURF and Harris Algorithms.

Different algorithms can be tried for point cloud studies in future studies and more precise 3D models can be produced. Besides, an SfM Algorithm with different detection and descriptor with the experiences gained here; It will be tested in the production of point clouds in C++ or Python environment in future research.

REFERENCES

Aicardi I, Chiabrando F, Lingua, A M & Noardo F (2018). Recent trends in cultural heritage 3D survey: The photogrammetric computer vision approach. Journal of Cultural Heritage, 32, 257-266.

Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). Speeded-up robust features (SURF). Computer vision and image understanding, 110(3), 346-359.

Cheng C, Wang X & Li X (2017). UAV image matching based on surf feature and Harris Corner Algorithm. 4th International Conference on Smart and Sustainable City (ICSSC 2017).

Dawood M, Cappelle C, El Najjar M E, Khalil M & Pomorski D (2012, October). Harris, SIFT and SURF features comparison for vehicle localization based on virtual 3D model and camera. In 2012 3rd International Conference on Image Processing Theory, Tools and Applications (IPTA), pp. 307-312. IEEE.

Harris, C. G., & Stephens, M. (1988, August). A combined corner and edge detector. In *Alvey vision conference* (Vol. 15, No. 50, pp. 10-5244).

Jebara T, Azarbayejani A &Pentland A (1999). 3D structure from 2D motion. IEEE Signal Processing Magazine, 3D and Stereoscopic Visual Communication, Vol. 16., 66-84.

Jiang R, Jáuregui D V & White K R (2008). Close-range photogrammetry applications in bridge measurement: Literature review. Measurement, 41(8), 823-834.

Kok K Y& Rajendran P (2018). Validation of Harris Detector and Eigen Features Detector. In IOP Conference Series: Materials Science and Engineering, Vol. 370, No.1, IOP Publishing.

Leutenegger, S., Chli, M., & Siegwart, R. Y. (2011, November). BRISK: Binary robust invariant scalable keypoints. In 2011 International conference on computer vision (pp. 2548-2555). IEEE.

Lowe D G (2004). Distinctive image features from scale-invariant keypoints. International journal of computer vision, 60(2), 91-110.

Mair, E., Hager, G. D., Burschka, D., Suppa, M. and Hirzinger, G., 2010. Adaptive and generic corner detection based on the accelerated segment test. In: Computer Vision ECCV2010, Springer, pp. 183–196.

Muja M & Lowe D G (2012). Fast matching of binary features. In 2012 9th conference on computer and robot vision (pp. 404-410). IEEE.

Kuçak R A, Özdemir E & Erol S (2017). The segmentation of point clouds with K-means and ANN (Artifical Neural Network). The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 42, 595.

Snavely, N., 2008. Scene reconstruction and visualization from Internet photo collections, unpublished PhD thesis, University of Washington, USA.

Tareen, S.A.K., & Saleem, Z. (2018, March). A comparative analysis of SIFT, SURF, kaze, akaze, ORB, and BRISK. In 2018 International conference on computing, mathematics and engineering technologies (iCoMET) (pp. 1-10). IEEE.

Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., & Reynolds, J. M. (2012). 'Structure-from-Motion' Photogrammetry: A low-cost, effective tool for geoscience applications. Geomorphology, Vol. 179, 300-314.