International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 22 (2017) pp. 12460-12466

Image Mosaic with Rotated Camera and Book Searching Applications Using SURF Method

Jaejoon Kim

School of Computer and Communication, Daegu University,
201 Daegudae-ro, Gyeongsan, Gyeongbuk, 38453, South Korea.
Orcid: 0000-0002-0444-4106

Abstract

In image processing or computer vision, feature extraction algorithms are used for various applications and the performance of these extracted characteristics depend greatly on the field they are applied to. Ever since the most notable feature extraction algorithm which is SIFT was introduced, studies for comparing and analyzing the performance of feature extraction algorithms have been continuously carried out to compare robustness and computation complexity. In this paper, we show the feasibility of exploiting SURF feature extraction algorithm to be applied for panorama image and book searching applications. Book searching applications in particular are limited, but research is underway to help the visual handicapped find the book that they want in the future.

Keywords: Feature extraction; Rotated Camera; SIFT; SURF; Visual Application.

INTRODUCTION

The expansion of the smartphone and the development of the network environment are leading to a rapid increase of the image data and various application fields such as the conversion of the image information through the mobile device, for example, the panorama image, the face recognition, or the object matching are increasing. Recently, the high technological mobile terminal can utilize the built-in advanced camera function to improve the image quality, use various augmented reality applications, or panoramic image production without moving the camera separately.

The panoramic image shows a phenomenon in which a wide angle of view shows the disadvantage of visually recognizing only objects within a limited area existing in one photograph. In addition, as shown in the paper [1], the accident scene of the vehicle can be viewed from a wide angle, so that it can be used as evidence for resolving a dispute between the perpetrator and the victim. Another technical issue can be rotated camera function which can take a photo by rotating camera angle by fixed amount. This approach can be compared with moving picture for generating panoramic view based on storage amount.

A panoramic image or an object matching application is a process of extracting feature points existing in two or more images and finding a desired result. The feature points may be corners or blobs as vectors having their own scales and directions, respectively. The scale and direction that are characteristics of this vector need to be robust to changes in relative scale and direction. SIFT (Scale Invariant Feature Transform) and SURF (Speeded Up Robust Feature) [2] are among the most notable algorithms used for extracting features and these are used in applications such as object detection [3], location recognition [4-5] and facial recognition [6-7]. In general, feature point extraction algorithms are more applied to recognize objects. In this paper, we describe how to apply SURF method, one of the representative feature extraction algorithms, to panorama image generation and book searching application.

In this paper, we will explore image stitching and book searching applications using SIFT and SURF and show possible and useful applications. Section 2 contains the summary of our research and interest into the background of our experimental approach related SURF method. Section 3 will discuss design of application system. Section 4 will show our results and Section 5 will conclude our experiment and discuss future works.

RELATED WORK

The feature points of an image can be like corners, letters, or shapes to represent own object. It is considered that this feature point requires a complicated process to become a feature point that is invariant to size and rotation. The characteristics of SIFT and SURF, which are representative feature point extraction algorithms, are examined.

A. Image Stitching

In Gaussian scale space, SIFT [8-9] broadly, which extracts features that are invariant to changes in size or rotation, is a method to find the local pole for the response of the Laplacian filter and reduce the amount of computation of LoG (Laplacian of Gaussian) through DoG (Difference of Gaussian) as shown in Figure 1. SIFT uses scale σ terminology which is defined in scale space function $L(x, y, \sigma)$. Depending on usage of scale value, the image to be processed can be blurred and smoother.

When we use variable scale value on considering images of the scale lower and higher, we can take account on local extrema of 26 around a given point. In order to identify local extrema, Taylor series approximation can be used. On corresponding to the scale, the size of the blob can be detected

and the edge responses are eliminated by using the Hessian matrix. After this process, the image is subsampled into a different octave and the similar scale space detection is carried out. This enables to identify the stable key points from the detected set. To achieve rotation invariance a weighted direction histogram is created for 4x4 neighborhoods of the interest point into 36 bins. This serves as the local descriptor for an image. Figure 2 described keypoint descriptor.

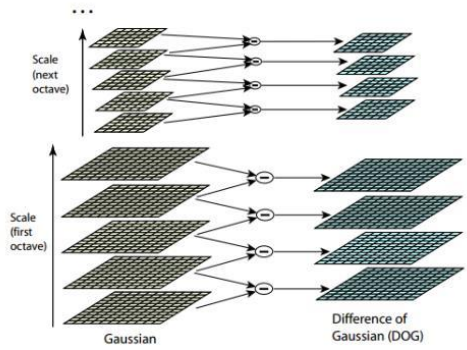


Figure. 1: The configuration of DoG generation process [8].

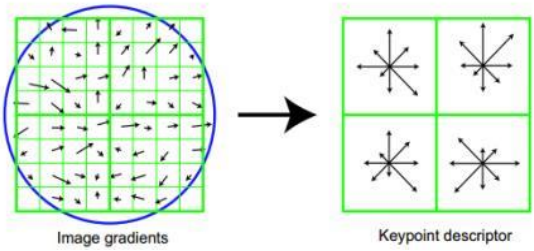


Figure. 2: The concept of keypoint descriptor: The gradient magnitude and orientation is first calculated at each image sample point around the location of the keypoint which is used to generate a keypoint descriptor [8].

The fundamental and key stages for SIFT algorithm shows the following characteristics.

- Scale Space Extrema Detection
- Determine scale and location of feature points
- Gaussian Pyramid
- Difference of Gaussian
- Apply different sigma value on each octave to be robust for a scale
- Key Point Localization
- Remove unsatisfied points on feature points
- Orientation Assignment
- Determine a directional component
- Key Point Descriptor
- Find orientation of gradient values around keypoint
- Key Point Matching
- Match by identifying nearest neighbors

The advantage of SIFT detector is strong against image scale, illumination, translation, rotation, and occlusion, but it is also hindered since the amount of computations required is large.

B. Review of Speeded-Up Robust Feature

SURF is one of the algorithms to find feature points that are invariant to environmental changes such as environmental changes such as scale, illumination, and viewpoint from multiple images. It is an algorithm that greatly improves speed with comparable performance to SIFT method. Although the processing speed of SURF is several times faster than SIFT, its accuracy is somewhat lower. However, SURF has shown better performance during real-time processing. SURF also has the advantage of not using features given in the color space as it only uses information in gray space [10-11].

Notable SURF features can be described as following.

- Exploit the integral image
- SURF detector is an approximation of the Hessian
- The calculations needed for detection are reused in descriptor computation
- Maintains robustness to scale, rotation, and illumination change
- Approximately 2 times faster than DoG and 10 times faster Hessian-Laplace detector [11]

Integral image

In SURF algorithm, unlike SIFT, we use integral images to find points of interest and areas. To calculate the point of interest, we use Fast Hessian Detector, which uses an integral image. The integral image means literally integrating the image. At any point (i, j) in the original image as shown in Figure. 3, we sum up all the pixels to the left and up from that point (i, j) [13].

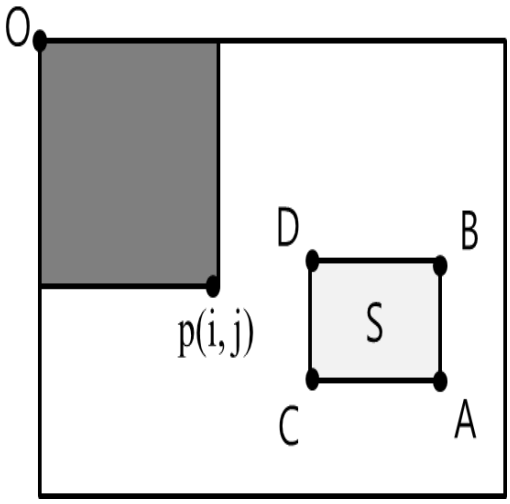


Figure 3: The concept of integral image [12].

$$I(x) = \sum\sum I(i, j) \quad (1)$$

Looking an integral image denoted as I(x), we first calculate the area for the image, then calculate the area **S** using 4 known points: A, B, C, and D. If an integral image of an image is generated and a partial sum of the image is known, the area of the area can be directly calculated and the calculation time can be shortened.

Reducing dimensional size using reduced detector and descriptor

To calculate the point of interest, SURF uses a Hessian matrix. Given a point p=(x, y) in an image I, the Hessian matrix H(p, σ) at point p and scale σ can be described in equation (2).

$$H(p,\sigma) = \begin{vmatrix} L_{xx}(p,\sigma) & L_{xy}(p,\sigma) \\ L_{xy}(p,\sigma) & L_{yy}(p,\sigma) \end{vmatrix} \quad (2)$$

where $L_{xx}(p,\sigma)$, $L_{xy}(p,\sigma)$ and $L_{yy}(p,\sigma)$ are the convolution of the second-order derivative of gaussian with the image I(x, y) at the point x.

In the paper [11], it is useful to find the blob structure of the Hessians matrices based method with good accuracy and maximum determinant. In this process, if the determinant is a negative number and the eigenvalue is a different sign, then the coordinates are not a point of interest and the determinant is a positive number. If both eigenvalues are negative or positive, the coordinate is determined as a point of interest. SURF is considered as an interest point when calculating the Hessian matrix for the x axis, y axis, and xy axis during the process of analyzing the scale-space using Gaussian for the second derivative. We also use approximated D_{xx} and D_{yy} box filters in x and y direction to further simplify the computations when computing the Gaussian distribution Hessian matrix.

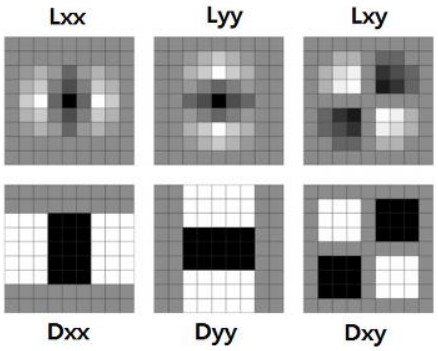


Figure. 4: The explanation for Gaussian second order partial derivative and approximation for the second order Gaussian partial derivative. Upper part describes Gaussian second order partial derivative in x, y, xy-direction denoted in L_{xx}, L_{yy}, L_{xy}. Lower part describes the approximation for the Gaussian partial derivative in x, y, xy-direction denoted in D_{xx}, D_{yy}, and D_{xy} [3].

Using the approximated box filter as shown in the Figure 4 above, the area width is obtained quickly by integral image calculation, and the Hessian matrix is calculated with the obtained area to obtain the interest points. SURF also extracts points of interest from multiple scales similar to SIFT in order to be robust. As shown in Figure 5 below, the SIFT method fixes the filter size and reduces the image size to detect the point of interest. On the other hand, the SURF method has the feature of extracting the point of interest while fixing the image scale and increasing the filter size as opposed to the SIFT method.

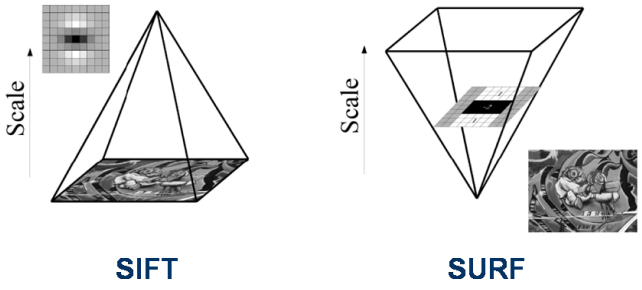


Figure. 5: The configuration of scale space on SIFT and SURF algorithm. Both SIFT and SURF divides the scale space (x,y,s), a mapping of x,y coordinates and scale, into octaves and levels. While SIFT downscales the image, SURF uses increasingly larger files. This allows SURF to calculate responses with arbitrarily large kernels [3].

The difference from SIFT is due to the efficiency of the calculation. Because it uses an integral image, the user can adjust the filter size simply by changing the position. In addition, because down sampling is not performed, there is an advantage since there is no aliasing. However there is also a disadvantage that scale invariance can be limited. The

maximum value for the size space (detected point of interest) is selected through non-maximum suppression of the point selected in the scale-space. Also, by interpolating the maximum value, interpolated points of interest are stored as feature vectors.

Since, in the x-y direction, the Haar-wavelet response is calculated with respect to the neighbors around the final interpolated point of interest, it is possible to apply fast calculation which is an advantage of the integral image. In the descriptor construction process, the area around the feature points is divided into 4 × 4 blocks, with each block being sampled at 5 × 5. First obtain the Haar-wavelet response for each sample in the block and obtain the sum of dx, dy, $|dx|$, $|dy|$. Then normalize the feature descriptor vector to make it robust to lighting changes.

Fast matching using contrast

After the descriptor stage is processed and the descriptors are calculated, we need to find the corresponding features between the two descriptor vectors by measuring distance component. In the work, we exploit the Laplacian code that computes the Hessian matrix or find the matching by searching for the closest neighbours by Euclidean distance. In SURF algorithm, the computation time can save by calculation of the Laplacian and depending on its positive or negative sign it is decided whether they can be matched or not. This one dimensional-like and simple calculation processing could be helpful for reducing overall processing time. At this stage, the sign can decide whether the selected point is a light blob inside a dark background or a dark blob in a light background.

DESIGN OF APPLICATION SYSTEM

This section discusses the applied implementation technique. The applied applications are consisting of two parts: first, building panoramic image using SURF and blending method and second, matching for interested book search. In many feature extraction related applications, we can take account into various characteristics such as ratio changing, removal or pan and zoom in the image. However, the implemented application in this work is more focus on image mosaic with SURF and possible book searching application for blind people.

A. ***Panorama image and multi-band blending***

Panoramic images can be photographed using multiple cameras with high quality of wide viewing angle (180 degrees). It can give a feeling of being in the field by filling in the field of view from indoor multi-monitor as well as large outdoor billboards and exhibition halls. The SURF algorithm, which gives similar properties, can produce panorama images in software. Especially, the panorama image can be used for various services. For example, there may be large-scale device services such as dialogue digital gallery services, digital walking and live panorama services, monitoring for services such as military and security, or display and performance.

This paper adopts various research results that the SURF has faster processing speed than the SIFT [13]. The unidirectional CCTV or blackbox used in the general parking system has a disadvantage that it cannot see scenes that occur over a certain angle. In other words, it is possible to guess the situation at the scene by using CCTV installed at different angles. However by synthesizing the images captured from a camera that rotates at a certain angle we can produce results that are more accurate and lively. Figure 6 described the block application diagram for panoramic image generation with multi-band blending.

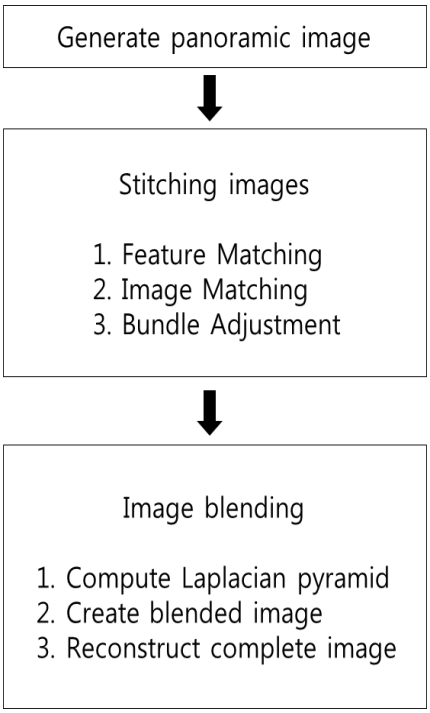


Figure. 6: The block application diagram for panoramic image generation with multi-band blending.

Figure 7 shows a dataset taken using a rotated camera rotated by 10 degrees. The fundamental characteristic of the rotated camera is that it can reduce the size of the data stored in the image of the head without shooting the moving image where the data is stored. When the SURF algorithm and the cropping process are performed, the difference in background intensities due to the time difference occurs. This difference in background intensities can be used to dilute the intensity difference between adjacent images using a multi-band blending technique [14]. Multiband blending allows blending the overall color of the images and mitigate the differences in the scale and the exposure existing between the juxtaposed images to remove sharp differences in the exposure between pictures and get a panorama as homogeneous as possible.

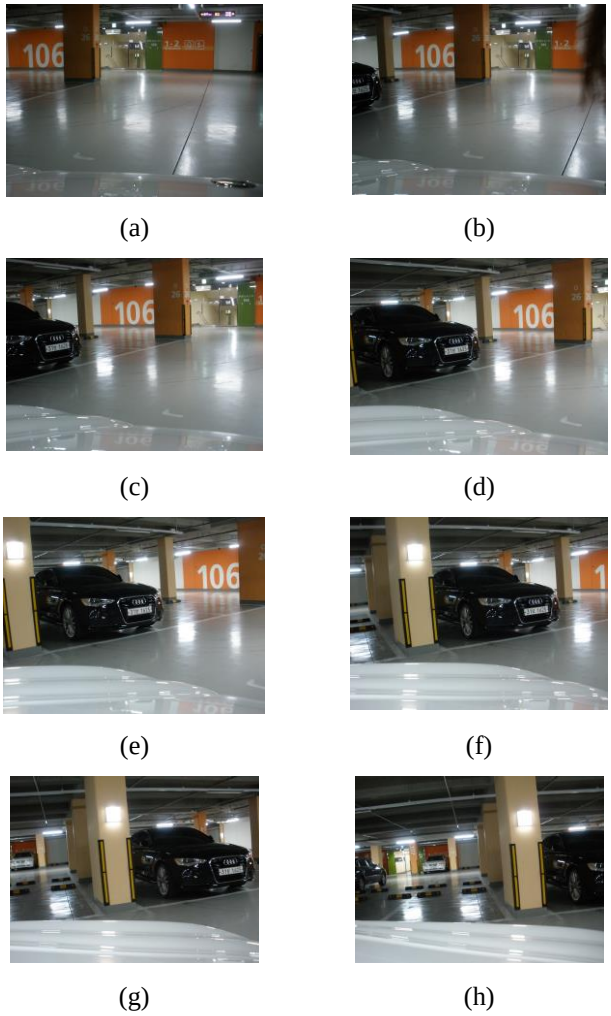


Figure. 7: Test data for generation panoramic image application. The test images are generated with rotated camera by 10 degree rotation from (a) to (h).

B. ***Book searching (Library Application)***

Matching of feature points is a process of finding corresponding points between two images. It can be considered that object pairs having the most corresponding points are recognized as the same object by finding corresponding points between a model database of an object previously created and an input image in recognition of a specific object. For each key point of two images, the closest point is searched, which can be obtained by calculating the distance between the key points. Figure 8 and 9 show the book searching application procedure and applied bookshelf image, respectively.

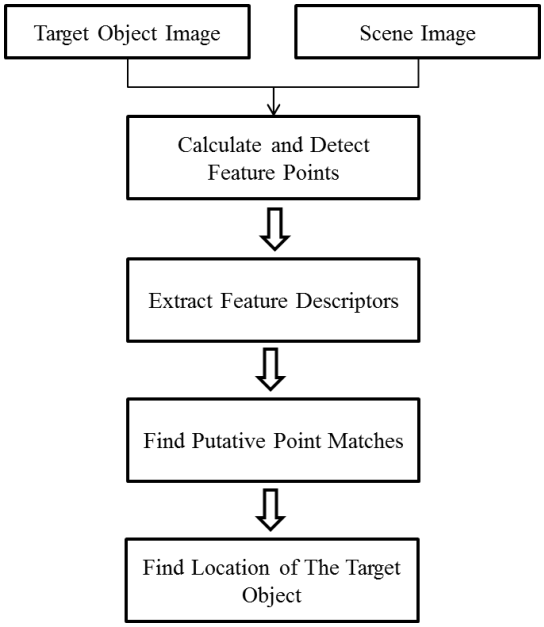


Figure. 8: The block application diagram for book searching and matching application.

(a)

(b)

Figure. 9: Test data for book searching and matching application. (a) target book image (b) scene image for a bookshelf.

EXPERIMENTAL RESULTS

To test the applicability of the SURF algorithm, we

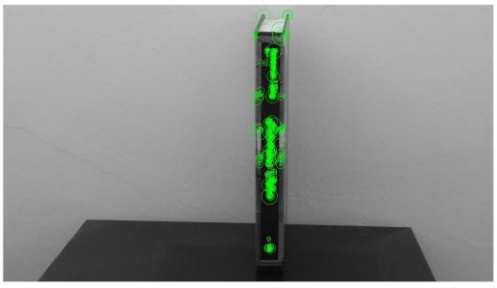
constructed a panoramic mosaic image of a still image rotated by 10 degrees using a rotated camera and experimented with matching in book searching applications. Such experiments can be applied to the book searching application of the visually impaired in the future. The experiment was implemented using MATLAB 2016A on a PC with an Intel i5-2400 3.1 GHz processor, 4 GB of memory, and a 64-bit Windows 7 operating system.

Figure 10 shows the results of the image data of the test taken from the rotated camera and the result of the clipping and multiband blending. Figure 10-(a) described generated panoramic image with 8 test images by 10 degree with a rotated camera. As we can see, this result does not apply multiband blending. However, Figure 10-(b) shows smoothed ghost between each image boundary. Figure 11 shows the process-specific results for the book searching application. Figure 11-(a) and (b) show 100 and 300 strong feature point for target image and 300 strong feature point for scene book shelf image. Figure 11-(c) and (d) show putative matched points with outlier and putative matched points with inlier only, respectively. Figure 11-(e) shows final result for interested target book searching and matching.

(a)

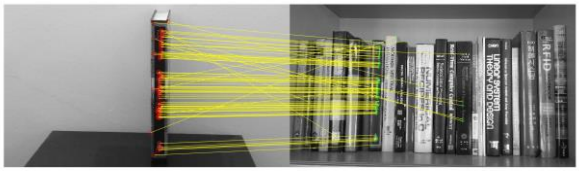
(b)

Figure 10: Generation of panoramic image. (a) result without blending method and (b) result with cropping and muti-band blending method.



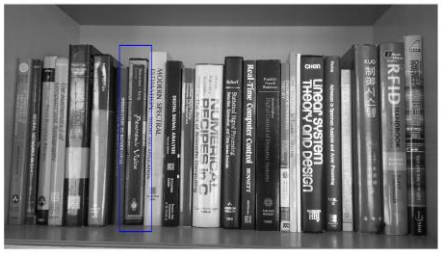
(a)

(b)



(c)

(d)



(e)

Figure 11: Book searching application. (a) 100 strong feature point for target image, (b) 300 strong feature point for scene book shelf image, (c) putative matched points (outlier) (d) putative matched points (inlier) (d) detection for interested book searching.

CONCLUSION

In this paper, we have explored SURF approach for mosaic image and book matching application. A panorama movie can be photographed using a mobile terminal or an advanced camera to help a dispute that may occur at a traffic accident site. However, being able to synthesize images photographed while rotating a certain angle, which can reduce the amount of heavy data, can be taken advantage of. Book searching application using SURF algorithm shows excellent object matching results and shows possibility to be useful for visually impaired people. However, since the linear search among the steps that are important in the matching result is inefficient in

terms of the amount of calculation, it is necessary to examine a method for searching at a higher speed. Also, as a future application field of this study, it is necessary to study mapping between virtual objects and real world objects in augmented reality.

REFERENCES

[1] Jaejoon Kim, "Design and Application for Stitching Images from Blackbox with Multiple Lenses," International Journal of Applied Engineering Research, Vol. 10, No. 23, pp. 43741-43746, 2015.

[2] Richard Szeliski, "Image Alignment and Stitching: a tutorial," Foundations and Trends in Computer Graphics and Vision, Vol. 2, Issue 1, pp. 1-104, 2006.

[3] H. Wersing and E. Korner, "Learning Optimized Features for Hierarchical Models of Invariant Object Recognition," Neural Computation, Vol. 15, No, 7, pp. 1559 -1588, 2003.

[4] B. Sirmacek and C. Unsalan, "Urban-area and Building Detection Using SIFT Keypoints and Graph Theory," IEEE Transactions on Geoscience and Remote Sensing, Vol. 47, No. 4, pp. 1156-1167, 2009.

[5] C. Tao, Y. Tan, H. Cai, and J. Tian, "Airport Detection From Large IKONOS Images Using Clustered SIFT Keypoints and Region Information," IEEE Geoscience and Remote Sensing Letters, Vol. 8, No. 8, pp. 128-132, 2010.

[6] C. Tan, H. Wang, and P. Deli, "SWF-SIFT approach for infrared face recognition," Tsinghua Science and Technology, Vol. 15, No. 3, pp. 357-362, 2010.

[7] R. Gonzalez and R. Woods, Digital Image Processing, Addison Wesley, 1992.

[8] David G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, Vol. 60, Issue 2, pp. 91-110, 2004.

[9] M. Brown and D. Lowe, "Recognising panoramas," in *Proc. of the 9th International Conference on Computer Vision (ICCV03)*, 2003, Vol. 2, pp. 1218-1225.

[10] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding*, Vol. 110, No. 3, pp. 346-359, Jun. 2008.

[11] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool, "SURF: Speeded Up Robust Features," in *proc. on European Conference Computer Vision*, Vol. 1, pp. 404-417, 2006.

[12] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features", in *proc. of IEEE CS Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp. 511-518, 2001.

[13] Shambhavi Jain, B. L. Sunil Kumar, and Ramesha Shettigar, "Comparative Study on SIFT and SURF face feature descriptors," in *proc. of International Conference on Inventive Communication and Computational Technologies (ICICCT),* pp. 200-205, 2017.

[14] P. J. Burt and E. H. Adelson, "A multiresolution spline with application to image mosaics", *ACM Transactions on Graphics (TOG)*, Vol. 2, No. 4, pp. 217-236, 1983.