IJETCR
Open Access

ISSN: 2348 - 2117

International Journal of Engineering Technology and Computer Research (IJETCR)
Available Online at www.ijetcr.org
Volume 2 Issue 5 Page No. 21-26

# Improvement in Stitched Image using Image Fusion

**Ajay Kumar Vishwakarma, Ajay Pal Singh, Jagpreet Singh**

Department of Electronics and Communication Engineering

SLIET Longowal, Sangrur (Punjab) 148106

akguru007@yahoo.com, ajaypal_singh@yahoo.com, singh.jagpreet86@gmail.com

**ABSTRACT**

It is often said that one picture is worth more than ten thousand words. Since the last few decades, image stitching in real time application has been a challenging field for image processing. In this digital era, images are an integral part of human life and play a big role. Image stitching is a process of merging two images in order to create a panoramic image. To perform the whole process, some steps should be followed such as image acquisition, feature extraction, image registration and blending. This paper leads to the study of the techniques of image stitching and how to improve the stitched image quality using image fusion techniques. The quality of image can be measure some parameters like PSNR, NAE, MI, MSE and SD etc. Image stitching is a traditional image process to create seamless panorama which shows new perspective of the world around us. Some applications of this technique are satellite imaging metrological and environment monitoring, 3D rebuilding objects, digital objects video stitching, and 3D image stitching [6].

*Keywords: feature extraction, image registration, warping, blending and image fusion.*

## INTRODUCTION:

Stitching of multiple images can create beautiful high resolution panoramas. This is so popular to increase the field of view of camera by allowing several views of a scene to be combined in to a single view. In general, stitching includes feature extraction, registration matching and blending. The motion relationship between two images or several images can be found through matching and that relates the success rate and the speed of image stitching process [1].

There are many algorithms for matching of images and blending them. For matching purpose, there are two ways in which one is direct method and other is the indirect or feature detection method. Direct method is sometimes inconvenient because it always need a high quality image. Harris corner detector [5], SIFT and SURF algorithm are mostly used. Harris corner detector is used for low level feature detection and while SIFT & SURF is used for high level feature based detection.

## SIFT ALGORITHM

It stands for scale invariant feature transform algorithm developed by D.G. Lowe in 2004, which is a scale invariant and rotation based method. It is also widely used in object recognition, registration, stitching *etc.* in images. The feature detected by the algorithm is accurate, stable, and precise as compared to other methods [15].

The steps involve in image stitching using SIFT algorithm is shown in Fig 1.

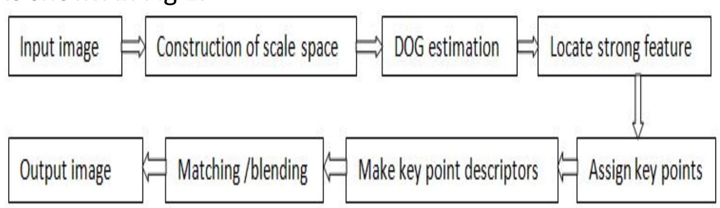


**Figure 1: Flow chart of the SIFT algorithm**

The basic steps which are involved in this algorithm are:

### Input image

First we acquire the two sequential image of a scene through a digital camera for which we are going to apply stitching process. It is necessary that they must have common overlapping area.

### Scale space key point detection

In the first stage, the key point is to identify location and scales that can be repeatedly assigned under different views of same object available in the image. The detected locations are non-variant to scale change of the image may be achieved by searching for fixed features across all possible scales, that is known as scale space. That under variety of reasonable assumption only possible in the Gaussian functions. Scale spaced image is a function of x and y, that can be represented by the Gaussian function. Output $L(x,y,\sigma)$, that is developed from the convolution of a variable-scale Gaussian function G(x,y,$\sigma$), with an input image, I(x,y):

$$L(x,y,\sigma) = G(x,y,\sigma)*I(x,y) \quad ...(1)$$

Where * is the convolution in x and y, and

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad ...(2)$$

To efficiency detect stable key point location in scale space, Daviod G. low proposed using scale-space extrema in the difference-of-Gaussian function convolved with the image:

$$D(x,y,\sigma) = (G(x,y,k\sigma) - G(x,y,\sigma))*I(x,y)$$
$$= L(x,y,k\sigma) - L(x,y,\sigma) \quad ...(3)$$

In order to detect local maxima and minima of D(x,y,$\sigma$), each sample point is compared to its eight neighbors in the current image and nine neighbor in the Gaussian scale space that are shown by the Fig. 2.

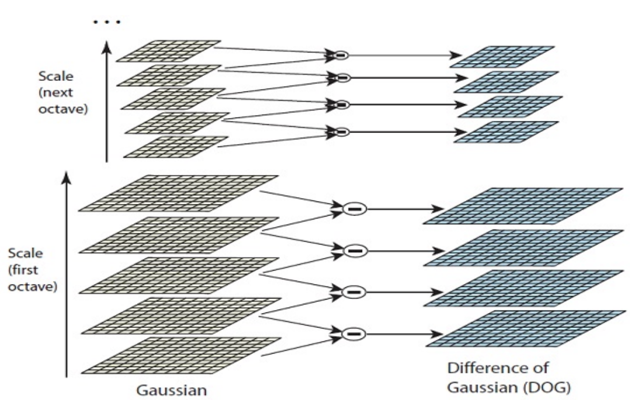


**Figure 2: difference of two adjacent intervals in the Gaussian scale-space**

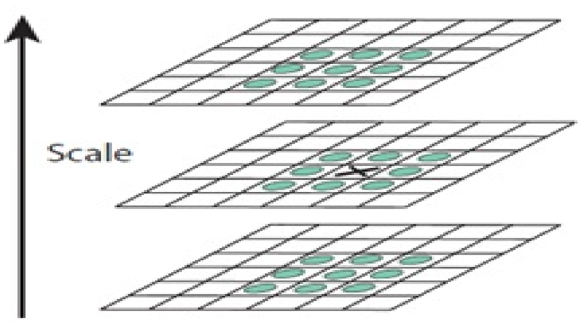


**Figure 3: Max and min detection in the difference-of- Gaussian image**

To increase the stability of matching, remove the points that have lower contrast. The Taylor expansion of the scale-space, function, shown by eq. 4

$$D(X) = D + \frac{\partial D^T}{\partial X}X + X^T \frac{\partial^2 D}{\partial X^2} X \quad ... (4)$$

D and its derivatives are evaluated at the sample points and X=$(x, y, \delta)^T$ is the offset from this point.

**The descriptor representation**

After assigning a consistent orientation to each key-point based on local image properties, the key-point descriptors are relative to orientation of image. The gradient magnitude, m(x,y), and orientation, $\theta(x,y)$, can be computed using pixels differences:

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + L(x,y+1) - L(x,y-1))^2} \quad ...(5)$$

$$\theta(x,y) = tan^{-1}(L(x,y+1) - L(x,y-1))/(L(x+1,y) - L(x-1,y)) \quad ...(6)$$

Where L is the scale of the key-point.

**Feature matching**

The correspondence between feature points in two images will be evaluated. The best match for each feature points is found by identifying its nearest neighbor in the data set of feature vectors from input image. The nearest neighbor is defined as the feature point with minimum Euclidean distance [3].

**Image blending**

Once we found the best transformation we can blend the input image together with reference image. To overcome the visible artifacts that are to hide the edges of the component images we use weighted average method of pixels [15].

The weighting function is

$$f(x,y) = \begin{cases} f_1(x,y) & (x,y) \in f_1 \\ d_1 f_1(x,y) + d_2 f_2(x,y) & (x,y) \in (f_1 \cap f_2) \\ f_2(x,y) & (x,y) \in f_2 \end{cases}$$

Where $f_1$, $f_2$ is the reference image and input image, f is the stitched image. $d_1$ $d_2$ is the weighted value and $d_1 + d_2 = 1$, $0 \le d_1, d_2 \le 1$.

**SURF ALGORITHM**

SURF stands for Speeded up Robust Features is the famous feature-detection algorithms. SURF detects the robust local feature detector technique; it was presented by Herbert Bay et al. that can be used in image processing, computer vision tasks like object recognition in image [15]. It is partly inspired by the SIFT descriptor. The original SURF algorithm is many times faster than SIFT algorithm and it is faster against different image transformations than SIFT. SURF algorithm is basically based on sums of two dimensional Haar wavelet responses and efficiently uses the integral images. An integer approximation method is used here to determine the determinant of Hessian blob detector that can be computed very fast with an integral image. For strong features, here the sum of the Haar wavelet response is used around the strong feature point of interest. Flow chart of surf algorithm is shown in Fig. 4. Below

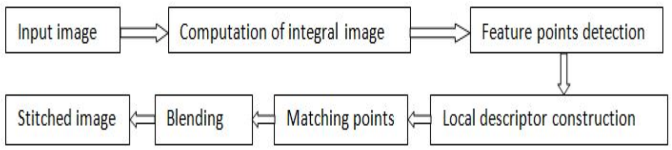


**Figure 4: Flow chart of surf algorithm**

### SURF Feature extraction

In the feature extraction process, features are detected in both images. It is based on scale space theory. The SURF detector is based on the determinant of the Hessian matrix. The given point X= (x,y) in an image I the matrix in X at scale $\sigma$ is defined below:

$$H(X, \sigma) = \begin{bmatrix} L_{xx}(X,\sigma) & L_{xy}(X,\sigma) \\ L_{xy}(X,\sigma) & L_{yy}(X,\sigma) \end{bmatrix} \quad \text{...(7)}$$

Where $L_{xx}(X, \sigma)$ is the convolution of the Gaussian second order derivative $\frac{\partial^2}{\partial x^2} g(\sigma)$ with the image I in the point X and similarly for $L_{xy}(X,\sigma)$ and $L_{yy}(X,\sigma)$ the scale space box filter shown in Fig. 5.

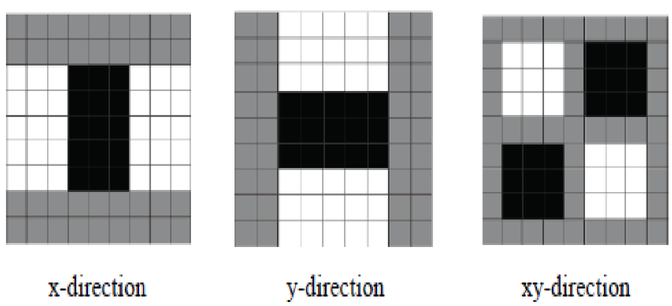


**Figure 5: Box filters**

These box filters can be evaluated very fast using integral image, independent of size. Therefore, the scale space is analyzed by up-scaling the filter size rather than iteratively reducing the image size, which greatly reduces the computing time. To localize interest point in the image and over scales, a non-maximal suppression in a 3x3x3 neighborhood is applied. In order to do this, the determinant of the Hessian function, $H(x,y,\sigma)$ is expressed by Taylor expansion up to quadratic terms centered at detected location;

$$H(X) = H + \frac{\partial H^T}{\partial x} + \frac{1}{2} X^T \frac{\partial^2 H}{\partial x^2} X \quad \text{... (8)}$$

The interpolated location of the extreme is found by taking derivative of this function and setting it to zero such that:

$$\hat{x} = -\frac{\partial^2 H^{-1}}{\partial x^2} \frac{\partial H}{\partial x} \quad \text{... (9)}$$

### Surf descriptor

The extraction of descriptor by the interpolation operation in scale space. Gaussian weighting coefficients are merged with HARR wavelet response to extract the interest points. Let the HARR wavelet responses in horizontal and vertical direction are dx,dy respectively. After summing them they give absolute value of the response that is

$$Vsub = (\Sigma dx, \Sigma dy, \Sigma|dx|, \Sigma|dy|) \quad \text{... (10)}$$

Therefore each sub-region has four values to the descriptor vector that are leading to an overall vector length.

### SURF feature matching

Lowe proposed the ratio between the Euclidean distance to the nearest and the second nearest neighbors pixels [16]. However, large number of matching points creates the complexity and decreases the speed of operation so find true and genuine matching points such that computational efficiency is increased.

### image fusion

The image fusion is the process of creating new high resolution image which is taking from the different image sensors. Or more images into single image having important features from each. In the image fusion process two images are blended together, holding all the complementary as well as common information from each of original images. Result of Fusion process creates high quality resolution image that are free from blurring effects. Fusion is an important technique within many fields such as robotics and medical application. The resultant image fusion is single image which is more suitable for machine perception and human or image processing field [19]. The algorithm of image fusion is based on wavelet transform that proves the geometric resolution of the geometric resolution of images, in which two images to be processed, first of all breakdown into sub image part and then the information is performed using these images under the definite criteria and finally these sub part of images are rebuilt into resultant image with high quality information image.

### Wavelet transform

Wavelet transform depicts a frame-work in which a signal is break into different level, with each level corresponding to coarser resolution or (LF) lower-frequency band and (HF) high frequency bands there are two main types of transforms discrete and continuous, we especially interested in DWT discrete wavelet transform which applies a two channel filter bank (with down sampling) iteratively to a low pass-band the wavelet representation then consist of the low pass-band at lowest resolution and the high pass-bands obtained at each step. This transform is invertible and non redundant.

The DWT is a spatial frequency decomposition that gives a flexible and multi-resolution analysis of an image.

In one dimension (1D) basic idea of DWT is to shows the signal as a superposition of wavelet. Suppose that a discrete signal is represented by function f(t); the wavelet decomposition is defined as below in eq.11

$$f(t) = \sum_{m,n} C_{m,n} \quad \psi_{m,n}(t) \; ; \qquad ...(11)$$

where $\psi_{m,n}(t)=2^{-(m/2)}\psi[2^{-m}\, t-n]$ and m and n are integers. There are some choices are available for $\psi$ such that $\psi_{m,n}(t)$ constitutes an ortho-normal basis, so that the wavelet transform coefficient can be achieved by an inner calculation:

$$C_{m,n} = < f, \psi_{m,n} > = \int \psi_{m,n}(t) f(t)dt \qquad ...(12)$$

In order to develop a multi resolution analysis a scaling function $\phi$ is required, together with the dilated and translated version of it;

$$\phi_{mn}(t)= 2^{-(m/2)} \psi[2^{-m}\, t-n]$$

According to the characteristics of the scale space spanned by $\phi$ and $\psi$, the signal f(t) can be break in its coarse part and details of various size by projecting it into the corresponding spaces.

Therefore to find such decomposition explicity, additional coefficients, $a_{m,n}$ are required at each and every scale. At each scale $a_{m,n}$ and $a_{m-1,n}$ describe the approximation of the function of at resolution $2^m$ and at the coarse resolution $2^{m-n}$ respectively, while coefficient $c_{m,n}$ and $a_{m,n}$ at each scale and position a scaling function is required that is similarly defined to (2) equation the approximation coefficient can be obtained:

$$a_{m,n}=\sum_k h_{2n-k}\, a_{m-1},k \; ; \qquad ...(13)$$

$$c_{m,n}=\sum_k g_{2n-k}\, a_{m-1},k \; ;$$

... (14)

Where $h_n$ is a low-pass filter FIR filter and $g_n$ is related high-pass FIR filter. To rebuild the original signal the analysis filter can be selected from a bi-orthogonal set which have a related set of synthesis filters $\tilde{h}$ and $\tilde{g}$ can be used to perfectly rebuild the signal using the reconstruction formula:

$$a_{m-1,l}(f) = \sum_n [\, \tilde{h}_{2n-l}\, a_{m,n}(f) + \tilde{g}_{2n-l}\, c_{m,n}(f)] \qquad ... (15)$$

Eq. (13) and Eq. (14) are implemented by filtering and down sampling.

Conversely Eq. (15) is a implemented by an initial up sampling and subsequent filtering. The discrete wavelet trans form for 2-D, and discrete wavelet transform for 1-D is first performed operation on the rows and then column of the input data one by one, filtering and down sampling this result is one set of approximation $I_a$ and three sets of coefficient as shown in Fig. 7. where $I_b$, $I_c$ and $I_d$ represent the horizontal vertical and dialog direction of the image I respectively in the language of filter theory these four sub image correspond to the output of low-low(LL), low-high (LH), high-low(HL) and high-high(HH) bands. By recursively applying the same scheme to (LL) sub band multi resolution decomposition with the desire level can then be achieved. Therefore, with K decomposition levels in DWT will have M=3*k+1 such frequency band. Figure 1(a) show the 2-D structure of the wavelet transform with two decomposition level. Therefore for a transform with K levels of decompositions, there is always only one low-frequency band ($LL^k$ in Fig 8); the rest of the bands are high frequency bands in given decomposition levels.

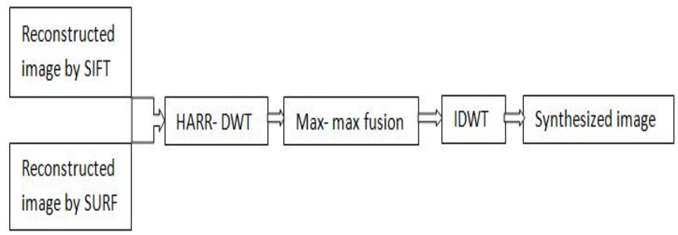


**Figure 6: Image fusion algorithm**

**Fusion with wavelet transform**

In this subsection, to better understand the wavelet-based technique, a schematic diagram is first given in Fig. 8, in general the basic concept of image fusion based on wavelet transform is to perform multi resolution decomposition on each source image, the coefficient of both the high-frequency bands and low-frequency band are then performed with a definite fusion rule as shown in the middle block of Fig. 8. The most widely used fusion rule is maximum selection method. This scheme firstly selects the largest coefficient of absolute value of wavelet at each location from the input image as the coefficient at the location in the fused image. After all the fused image is obtained by performing the IDWT operation for the corresponding combined wavelet coefficients therefore, as shown in Fig. 7, the steps involved fusion based on wavelet transform can be explained below [13].

**Step1.** The image to be fused must be registered to assure that the corresponding pixels are aligned (same line).

**Step2.**These images are decomposed into wavelet transformed images, based on wavelet transformation.

**Step3.** The transform coefficients of different position or bands are operated with a certain fusion rule.

**Step4.** The fused image is constructed by performing an (IWT) based on the combined transform coefficient from step3.

The unique feature of Harr wavelet transform are best processing speed, memory management simplicity, as well as reversibility. Mother Wavelet function can be presented by:

$$\psi(t) = \begin{cases} 1 & 0 \le t \le 0.5 \\ -1 & 0.5 \le t \le 1 \\ 0 & otherwise \end{cases}$$

The scaling function

$$\phi(t) = \begin{cases} 1 & 0 \le t \le 0.5 \\ 0 & otherwise \end{cases}$$

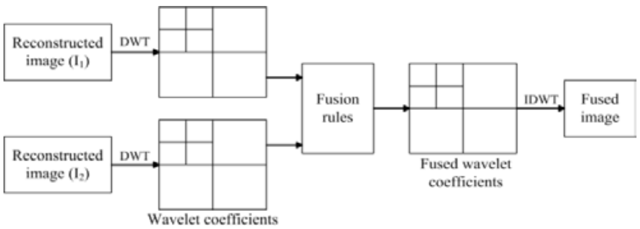


Figure 7: Image fusion scheme using wavelet transform

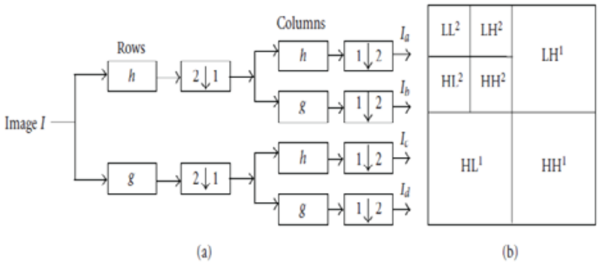


**Figure 8: Discrete wavelet decomposition with filter bank**

**RESULTS AND DISCUSSIONS:**

In this section, Performance analysis of sift surf and proposed method

| Sr. no. | algorithm /Parameters | SURF Constructed image | SIFT Constructed image | Proposed method Image |
|---|---|---|---|---|
| 1 | PSNR (db) | 69.7535 | 71.0869 | 74.8957 |
| 2 | MSE | 289.6812 | 269.4514 | 230.5967 |
| 3 | MI | 2.1078 | 2.0922 | 2.3576 |
| 4 | NAE | 0.3448 | 0.3417 | 0.2860 |
| 5 | SD | 58.6261 | 58.7952 | 58.9997 |

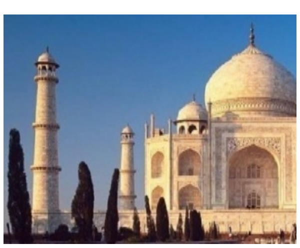

**Figure 9a: left portion of the scene**

**Figure 9b: right portion of the scene**

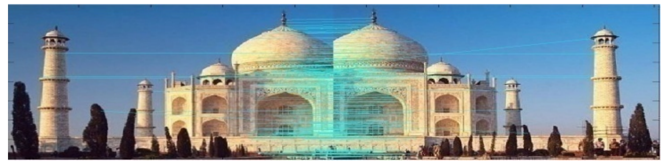

**Figure 10: 276 Points are genuine match out of 1365 points.**

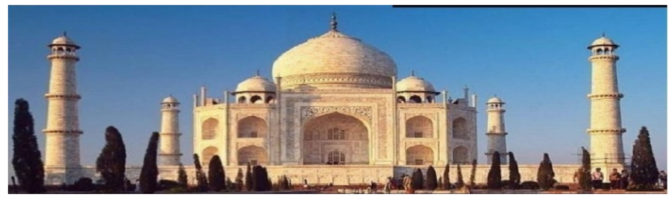

**Figure 11: Resulting panoramic image using sift algorithm**

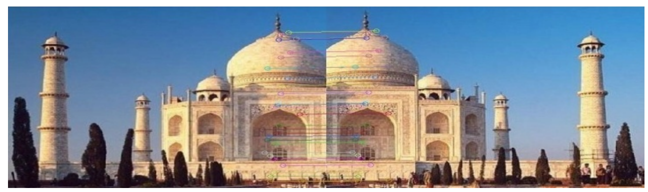

**Figure 12: 30 Correspondence matched feature between Two input images**

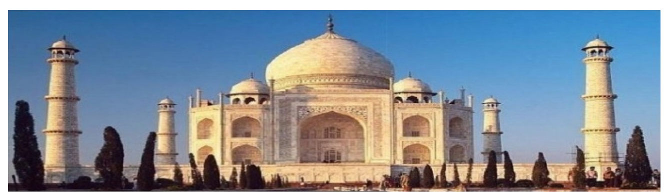

**Figure 13: Resulting panoramic image using surf algorithm**

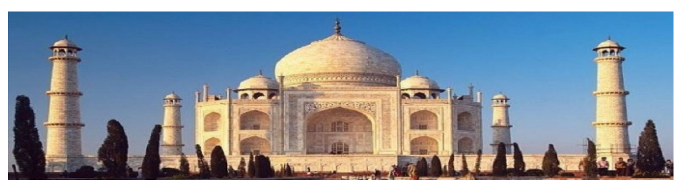

**Figure 14: Proposed resultant image using image fusion technique.**

**CONCLUSION:**

Stitching operation applied on two input images using the popular stitching algorithms i.e. SIFT and SURF were performed. To extract the best features from the images, the blending process was done. The quality of the above stitching techniques is tested by means of three-dimensional rotational images. The panoramic image generated by the fusion process compensates the complementary features and boosts up the common features of separate stitching images. SURF algorithm has the distinctive property of illumination invariance along with good scale and rotation invariance property, whereas, SIFT is the most effective algorithm for scale and rotate image stitching. But, it cannot cope up with illumination variation. Therefore, the resultant image proves superior as compared to the SIFT as well as SURF algorithms in terms of PSNR, MSE, MI, NAE, and SD.

**REFERENCES:**


1. Jyun Hong Chen and Cheng Ming Huang. "Image Stitching on the Unmanned Air Vehicle in the Indoor Environment" SICE Annual Conference August 20-23 Akita University Akita Japan SICE pp. 402-406, 2012.
2. Joon Hyuk Cha Young Sun Jeon."Seamless and Fast Panoramic Image Stitching" IEEE pp. 29-30, 2012.
3. Athreya S. Bhat and Amith V. Shivaprakash."Template Matching Technique for Panoramic Image Stitching" IEEE, 2013.
4. Junguk Cho Joon Hyuk Cha."Stereo Panoramic Image Stitching with a Single Camera" International Conference on Consumer Electronics (ICCCE) IEEE pp. 256-258, 2013.
5. D. K. Jain, G. Saxena, V.K. Singh. "Image Mosaicing Using Corner Technique", Proceeding of International Conference on Communication System and Network Technologies, IEEE pp. 79-84, 2012.
6. Ying zhang, Zhujun and Lei yang. "Research on Video Image Stitching Technology Based on SURF" fifth international symposium on computational intelligence and design IEEE pp. 335-338, 2012.
7. Yang zhang-long, Guo bao-long. "Image Mosaic Based on SIFT" International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IEEE pp. 1422-1425, 2008.
8. Luo Juan, Oubong Gwun. "SURF Applied in Panorama Image Stitching" Conference Image Processing Theory, Tools and Applications pp. 2010.
9. Jingxin Hong, Wu Lin, Hao Zhang, Lin Li. "Image Mosaic Based on SURF Feature Matching" IEEE the 1st International Conference on Information Science and Engineering (ICISE) pp. 1287-1290, 2009.
10. Prajkta Sangle and Krishnan Kutty. "A Method for Generation of Panoramic View based on Images Acquired by a Moving Camera" IJCA Special Issue on "Artificial Intelligence Techniques - Novel Approaches & Practical Applications" AIT, pp. 11-14, 2011.
11. Gao Guandong and Jia Kebin. "A New Image Mosaics Algorithm Based on Feature Points Matching" IEEE conference pp.1-7, 2007.
12. Mirajkar Pradnya and Ruikar Sachin. "Wavelet Based Image Fusion Techniques" International Conference on Intelligent Systems and Signal Processing (ISSP) pp. 77-81, 2013.
13. Soma Sekhar and Dr. M. N. Giri Prasad. "A Novel Approach of Image Fusion on MR and CT Images Using Wavelet Transforms" IEEE conference pp. 172-176, 2011.
14. D.G Lowe."Object Recognition from Local Scale-Invariant Features", International Conference on Computer Vision, vol. 2, pp.1150– 1157, 1999.
15. Herbert Bay. "Speeded up Robust Features", Proceedings of the 9th European Conference on Computer Vision, Cambridge, U.K., pp. 404-417, 2006.
16. Z. Zhang and R. S. Blun. "A Categorization of Multi Scale Decomposition-Based Image Scheme with a Performance Study for a Digital Camera Application," proceedings of the IEEE vol.87, no.8, pp.1315-1326, 1999.
17. B. Zitova and J. Flusser. "Image registration methods: A survey", Image and Vision Computing, vol.21, pp.977-1000, 2003.
18. D. G. Lowe. "Distinctive Image Features from Scale-Invariant Keypoints", International Journal of Computer Vision, vol. 60, pp. 91- 110, 2004.
19. G. Pajares and J. Cruz. "A Wavelet-Based Image Fusion Tutorial", Elsevier Pattern Recognition. vol. 37, pp. 1855-1872, 2004.