ISSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)
Vol 4, Issue 4, April 2019

Image Feature Matching Using PSSC

[1] Kavitha H, [2] Akshatha S, [3] Bhargavi B, [4] Rachana H J
[1] Assoc. Prof. Department of Information Science and Engineering,
[2] [3] [4] Undergraduates at Department of Information Science and Engineering,
Siddaganga Institute of Technology, Tumakuru.

Abstract: - **Feature matching refers to estimating robust feature correspondences between two images of same scene, which maps the key points from source to target data set. In this work an effective approach is used in the form of Progressive Sparse Spatial Consensus (PSSC) for finding more true matches from a putative set of feature correspondences. The key purpose is performing sparse approximation progressively based on spatial consensus. This significantly reduces the computation complexity as well as covers the more true matches. The spatial transformation between images is characterized by non-parametric thin plate spline kernel which enables our progressive Sparse Spatial Consensus method to handle non-rigid and rigid motions of the image pairs. The Expectation Maximization along with the maximum likelihood model is used to estimate and optimize the degree of true match. The quantitative outcomes obtained on publicly accessible data sets are verified with the results of various algorithms shows that our approach outstands in the rate of precision, recall and f-measure specifically in the case of large-scale outliers.**

Index Terms—Feature matching, Outliers, Progressive, Robust Estimator, Sparse Consensus.

I. INTRODUCTION

A wide range of importance is given for establishing reliable point correspondences between two images including various fields like 3D reconstruction, image retrieval, camera self-calibration, pattern recognition and tracking [1]. This paper mainly focuses on establishing the feature correspondences between the pair of images. Feature can be defined as relevant information for solving the various tasks. The feature based image matching makes use of the local feature based on the descriptors. Features can be represented in various forms such as points, segments and regions.

In this paper, an algorithm called as PSSC is used to establish the point correspondences. Putative correspondences set containing higher outliers in comparison to the inliers will be created by the algorithm. The creation of putative set will take place by matching the points based on various factors like intensity, local descriptors. Most widely used methodology for the removal of outliers in RANSAC [2].

RANSAC maximizes the number of inliers in the putative set by reducing the outliers. RANSAC method works for a parametric model and fails to maximize the result of non-rigid ones. There are various methods for non-parametric models like Identifying Correspondence Function (ICF) [3], Coherent Spatial Relations (CSR) [4], Vector Field Consensus (VFC) [5] and Locality Preserve Matching (LPM) [6].

VFC makes use of vector field interpolation. LPM is a flexible algorithm for feature matching in both rigid and non-rigid transformations. Outliers are removed from the putative set and transformation is performed. CSR estimates the coherent spatial relation between the inliers and treats inlier set as a whole. Though there exists various methodologies it is a tedious task to have a practical algorithm which deals with real time problem. To meet the objective the putative set need to contain higher inliers comparatively, whereas the existing methodologies have higher computation which reduces their efficiencies.

All the above stated problems are overcome in PSSC algorithm which works even when there is maximum number of outliers. Firstly, spatial consensuses between the features are estimated, by using Gaussian distribution. Mismatches can be removed by maximum likelihood problem and then expectation maximization is applied. To increase the performance in case of large number of outliers the algorithm will be applied progressively. Thin plate spline (TPS) kernel [7] is used to handle the rigid and non-rigid transformations. Spatial consensus is computed with sparse application, which reduces the computational complexity and increases the efficiency.

The flow of the paper is as follows: Section II contains the related work, section III describes the proposed methodology, the quantitative result analysis is shown in section IV and work is concluded in section V.

II. RELATED WORK

Various fields like computer vision [8], pattern recognition [9], medical image analysis and remote sensing, makes use of image matching. Outlier removal plays an important role in image matching. Various methods have been employed for the detection and removal of outliers. In [10], authors have introduced the hidden variable to detect the outliers.

Feature based methods extract salient features from the image. A popular approach for feature matching includes a two stage matching technique. In the first stage, a putative set is computed by considering the points matching the similar descriptors. In second step, the outliers are pruned by setting geometric constraint that need to satisfy certain geometric requirement such as homography, epipolar geometry. Feature matching uses another strategy of estimating a correspondence matrix on parametric or non – parametric constraints. Heuristic iterative algorithm makes use of the above feature matching strategy [11, 12]. Chui and Rangarajan adopted a strategy for formulating the correspondences and transformation [13].

Plenty of methods such as statistical regression, resampling and graph matching are employed. Resampling method generates a hypothesis repeatedly on the basis of parametric model by considering the putative correspondences set. Resampling mainly concentrates on obtaining the mismatch free set. Resampling methods such as RANSAC [2], Progressive Sample Consensus (PROSAC) [14] to acquire smallest outlier free subset for estimating the parameter model. Prior estimated probability of the putative set is considered and there application in the field of motion estimation is studied.

The statistical regression method does not work well in case of non – rigid transformation and the efficiency is degraded with the increase in outliers in the putative set. Graph matching [15] uses pair – wise constraint and casted as a quadratic problem.

To overcome these issues various non – parametric methods such as LPM, VFC are recently developed. VFC method interpolates a vector field which estimates the consensus of inlier points. LPM algorithmic method maintains a smoothness transformation of the image by preserving the neighborhood structure.

Many efforts towards generation of a better putative set have taken place. Guo and Cao’s [29] method of pruning the false matches by considering the triangle constraint. Ma et al [30] method of pruning the false match by preserving the local neighborhood. Increasing the true matches by pruning away the false matches will be effectively carried out in this method. The algorithm is applied progressively to obtain maximum number of matches even in case of small set with high outlier ratio. By adopting this method efficiency of the system is increased as the computational complexity stays low.

III. METHOD

The description of the feature matching techniques is in this section. Initially two images are taken, which are usually of the same scene but in different view and they may differ in the amount of noise and outliers. In order to detect the features in the images, the Scale-invariant feature transform (SIFT) [16,17] algorithm is used, followed by the EM solution which is applied for Sparse Spatial Consensus (SSC) and called iteratively to achieve a progressive strategy, Progressive Sparse Spatial Consensus (PSSC).The flow is as follows:

A. Scale Invariant Feature Transform (SIFT)

From the images taken as input, the local features are selected. Local features are nothing but the remarkable patterns of the image which may be a point, patch or an edge which are specific in the surroundings. The local features should be localizable and distinct. The local features selected are converted into a vector representation which is done by descriptor, one among the descriptors used is SIFT. SIFT descriptor works well even if the local features are established in the edge extreme of the image. The algorithmic steps implemented in order to achieve the SIFT matches are as follows:

1) Scale-Space Extrema Detection

The scale space function identifies the unique patterns of the image. Scale space function is interpreted as follows:

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

In the equation (1), $*$ is referred to as a convolution operator. It performs tasks like reduction of noise, edge detection, blurring of image etc. $I(x, y)$ is the input image and $G(x, y, \sigma)$ is the Gaussian scale. Based on the values of Difference of Gaussian (DOG), the key points are taken at maxima/minima at those points. The Difference of Gaussian is calculated as:

$$D(x, y, \sigma) = L(x, y, k\sigma) * L(x, y, \sigma) \quad (2)$$

The difference in the scale space function is calculated in the equation (2) with the difference of a scale k.To identify the maxima/minima from the obtained points, they analyzed with neighbors on the same scale and from the immediate above and below scales. If the obtained value is same at all points then it is considered as a key point.

2)Key Point Localization

The key points which are imperfect or unstable are distinguished and eliminated. The elimination is done by computing value of the Laplacian for each of the point which was calculated during the scale-space extrema detection.

$$z = \frac{\partial^2 D^{-1}}{\partial x^2} \frac{\partial D}{\partial x} \quad (3)$$

Where z is the value of key points which are found to be unstable. If the value of z is less than the considered threshold value which is 1.5, then such points are excluded from the key point set.

3) Orientation Assignment

The stable key points obtained from the key point localization are allocated with a constant orientation which establishes the image properties. The gradient magnitude value m(x,y) is calculated as follows:

$$m(x,y)=\sqrt{\begin{matrix}(L(x,y+1)-L(x,y-1))^2+\\(L(x+1,y)-L(x-1,y))^2\end{matrix}} \qquad (4)$$

The located key points are fixed or unchangeable with respect to the rotation, scale or the location of the image.

4) Key Point Descriptor

The key point descriptors are constructed with the magnitude of each point obtained from the orientation assignment. The Histograms are created from the descriptors and the value for variance. The resulting vectors obtained from the histogram are the required SIFT keys. The SIFT features obtained are with highest accuracies which are independent of the image rotations or the transformations. The Difference of Gaussian (DoG) is calculated during the Scale-space extrema detection, provides high accuracy and contains many interesting points.

B. The Expectation Maximization algorithm

After obtaining the SIFT features from the input images, if the features contain a vast number of outliers, it leads to the result which may be partial. Or there may be certain variables which may not be distinguished directly, but they affect the matching results. Obtaining such inconsistent results won't help the analysis. The EM algorithm when applied for the maximum likelihood estimation helps in the detection of such variables. The EM algorithm works as follows:

E-step: For the data x, a set of unknown variables θ is estimated as:

$$\theta(t)=\{\mu_j(t),\varepsilon_j(t)\} \quad j=1\dots.M \qquad (5)$$

Here, $\mu_j(t)$ represents the mean value and $\varepsilon_j(t)$is the co-variance for a set of M clusters. From the values obtained from equation (5) responsibility, p_n is calculated.

$$P(C_j|x)=\frac{|\sum j(t)^{-1/2}|exp^{nj}\,P_j(t)}{\sum_{k=1}^{M}|\sum j(t)^{-1/2}|exp^{nk}P_k(t)} \qquad (6)$$

$$n_j=-\frac{1}{2}(x-\mu_j(t))\sum j(t)^{-1/2}(x-\mu_j(t)) \qquad (7)$$

M-step: In this step, the values for mean and co-variance from the equation (6) are updated.

$$\mu_j(t+1)=\frac{\sum_{k=1}^{N}P(C_j|x)x_k}{\sum_{k=1}^{N}P(C_j|x)} \qquad (8)$$

$$\varepsilon_j=\frac{\sum_{k=1}^{N}P(C_j|x)(x_k-\mu_j(t))(x_k-\mu_j(t))}{\sum_{k=1}^{N}P(C_j|x)} \qquad (9)$$

From the $\mu_j(t)$ from equation (8) and $\varepsilon_j(t)$ from equation (9), the responsibility $p_n$is updated as:

$$p_j(t+1)=\frac{1}{N}\sum_{k=1}^{N}P(C_j|x) \qquad (10)$$

After the completion of E-step and the M-step, the amount of inliers, I_0 is calculated from the resultant responsibility of equation (10)

$$I_0=\{n|p_n>\tau, n=1\dots.N\} \qquad (11)$$

From equation (11), τ is the threshold value which is in the range of 1.5 for all the data sets.

C. Sparse Spatial Consensus (SSC)

After each iteration, the inlier set is obtained as a result of the Sparse Spatial Consensus algorithm. A putative set $P=\{(x_n,y_n)\}_{n=1}^{N}$ is constructed with x_n and y_n being the feature point location in the image. The putative set, P is acceptable if it has wide amount of true matches and the false matches are discarded. The value of responsibility $p_n$is calculated after each E-step and the transformation **f** is updated after each M-step accordingly to the EM algorithm. The inlier set obtained as an output may contain large number of mismatches and discarding them during the construction of a good putative set may lead to the elimination of some correct matches too [18]. In order to avoid this problem, a progressive procedure is introduced, which is nothing but the Progressive Sparse Spatial Consensus (PSSC). The working of PSSC is as follows:

A putative set $P_0=\{(x_i,y_i)\}_{i=1}^{N}$ is constructed with a high inlier ratio and a small threshold ratio t_0. The SSC algorithm works well and the image matches are detected. Another putative set $P_1=\{(x_i,y_i)\}_{i=1}^{N}$ is built such that inlier ratio as compared to P_0 is typically low and the threshold ratio $t_1$is large. In case of the putative set P_1, the SSC results are undesirable. The cause of SSC not working well as expected is due to the local extrema values of the image. To make the working of SSC efficient in cases such as P_1, the EM algorithm is given a good initialization. The result of P_0 putative set can act as a good initializer for the EM iteration which is then applied SSC for the set P_1. Thereby the local extrema values can be obtained and the matching continues. The value of R_i is calculated from the result I_0. R_i, the responsibility is calculated as follows:

$$R_i=\begin{cases}1 & if\ (x_i,y_i)\ is\ in\ I_0\\ \in & otherwise\end{cases} \qquad (12)$$

The value of I_0 is calculated from the values of responsibility, $p_n$as in equation (11) and ∈ is a small number. The true correspondences can hence be increased by performing this method progressively. This process continues for putative sets with a still higher threshold. In such cases the matching is first done on the set P_0, whose results are applied for SSC and the inlier set is obtained, which later is given as an initializer for P_1, and the process continues. The outline for PSSC algorithm is shown as below:

Input: Putative set P_0 , P_1.

Output: Inlier set I_1

1. SSC is performed onP_0, initial values of I_0 is obtained.
2. From the obtainedI_0, the value of responsibility R_i is initialized.
3. SSC is performed on $P_1$and inlier set I_1 is obtained.

IV .EXPERIMENTAL RESULTS

To estimate the feature matching performance of our proposed PSSC algorithm on several image pairs, the outcomes are compared with other feature matching algorithms like CSM [4], VFC [5], LPM [6]. The putative point correspondences set for all the four matching methods are estimated from SIFT key descriptors using nearest neighbor approach. The procedures are carried out on desktop with 1.62 GHz Intel core CPU, 4GB memory and matlab code. The parameters for all the algorithms are fixed and the validity of correspondence function is estimated. The experiments are carried out on two aspects i) image pairs correlated by homography ii) image pairs correlated by non-rigid transformation.

SIFT features are extracted by using VLFeat [19] open source toolbox. Putative correspondences are constructed based on threshold values of distance ratios. The feature matching performance is specified by precision, recall, f-measure, run time and number of correct matches.

$$Precision = \frac{Number\ of\ reserved\ correct\ matches}{Number\ of\ whole\ reserved\ matches} \quad (13)$$

$$Recall = \frac{Number\ of\ reserved\ correct\ matches}{Number\ of\ correct\ matches} \quad (14)$$

$$F-measure = \frac{2*precision*recall}{precision+recall} \quad (15)$$

SIFT key features are used for object recognition from training dataset and compare each feature to the new image. PSSC is experimented on the dataset which contain image pairs having variations in viewpoint, scaling, illumination, rotation, image compression. The number of correct matches significantly increases when SIFT algorithm is applied along with PSSC as shown in Table 1.

TABLE 1. Feature match values of SIFT and SIFT with PSSC

Dataset	SIFT	SIFT + PSSC
1	30	126
2	320	488
3	42	166
4	480	1150
5	274	458

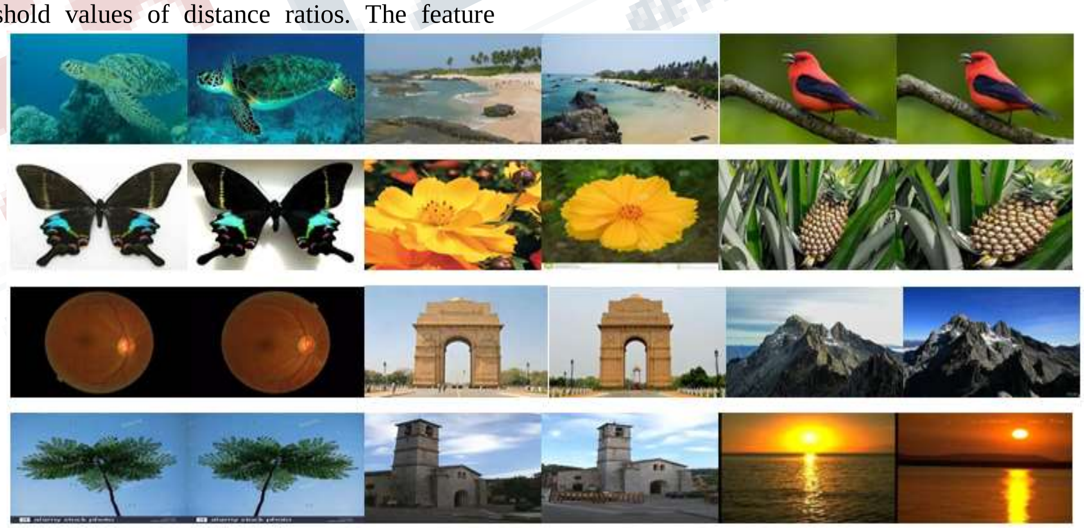

FIGURE 1. Sample dataset of each category (from left to right animals, beaches, birds, butterflies, flowers, fruits, iris, monuments, mountains, trees, buildings, nature).

ISSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)
Vol 3, Issue 4, April 2019

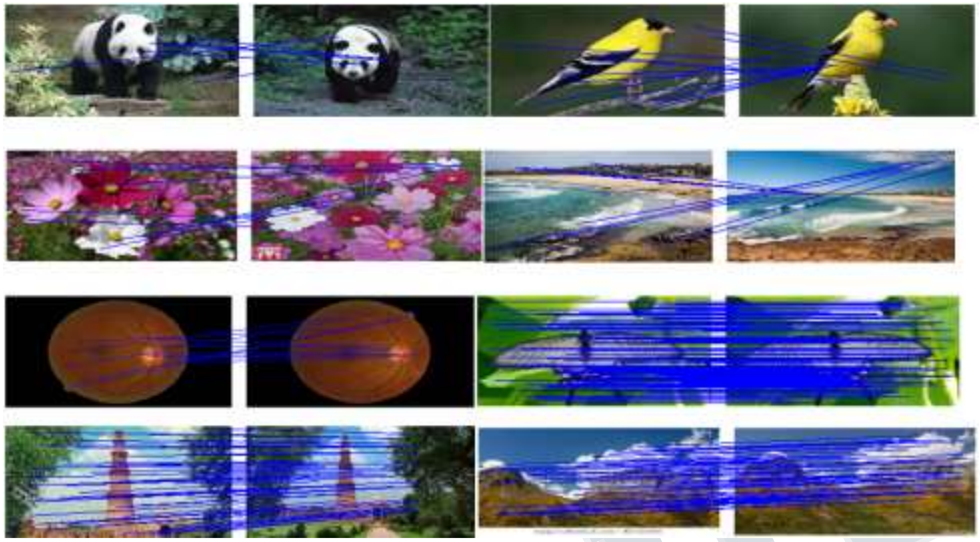

FIGURE 2. Feature matching results of PSSC algorithm on eight image pairs (from left to right and top to bottom: animals, birds, flowers, beaches, iris, butterfly, monuments and mountains) the lines indicates the matches.

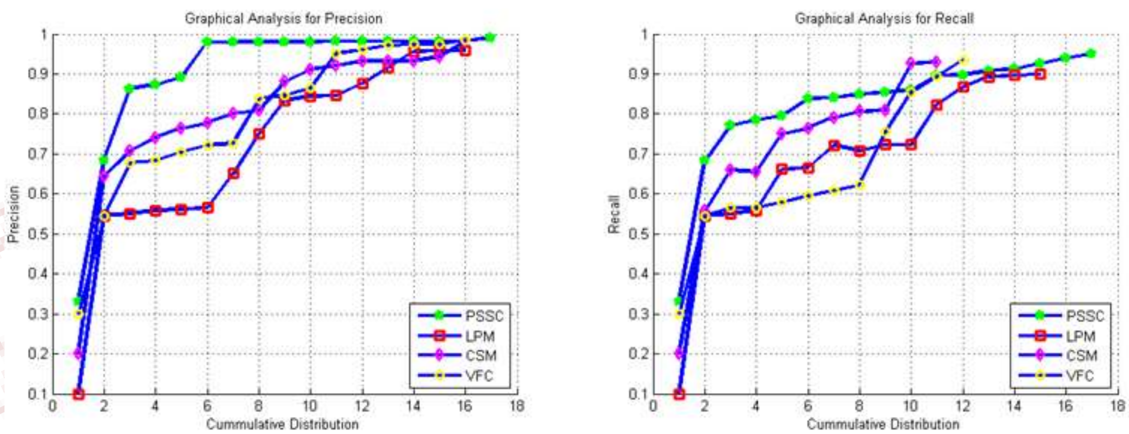

FIGURE 3. Precision (left) and Recall (right) of four feature matching methods applied on different categories of datasets.

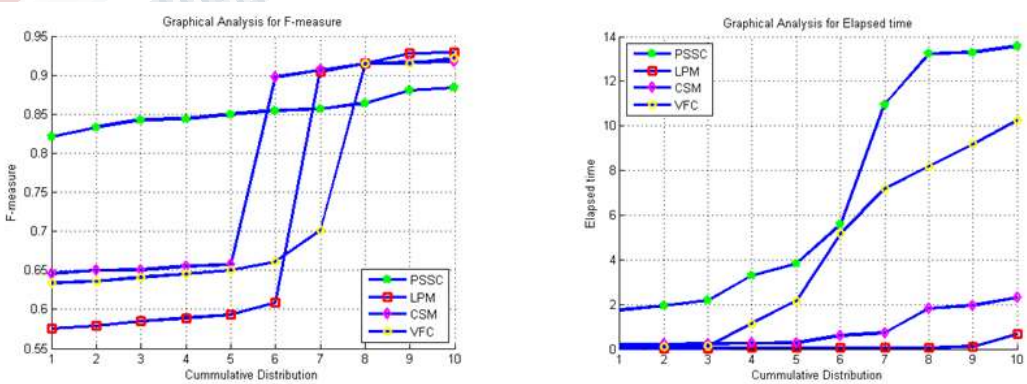

FIGURE 4. F-measure (left) and Elapsed time (right) of four feature matching methods applied on different categories of datasets.

The PSSC is tested on 150 image pairs belonging to 9 different categories some of which are shown in Fig 1. Sparse approximation is employed as a result there is a significant speed up of matching procedure. The quantitative result comparison on the various dataset categories with three feature matching methods such as CSM [4], VFC [5], LPM [6]. The statistical analysis of number of correct matches, precision, recall, f-measure and average run time are calculated.

In terms of number of correct matches and precision as shown in Fig 3 and Fig 4, VFC and CSM resemble each other and they are graded in the middle. LPM uses k-means neighborhood method which cannot work well if the putative set contains many false matches and doesn't produce high precision. In contrast, our PSSC yield better precision and correct matches by adopting progressive strategy for matching. PSSC holds rank in producing significant number of true matches even if putative set contains false matches. The schematic illustration of our PSSC is shown in Fig 2.

The runtime comparison of different methods is analyzed in Fig 4. LPM is most efficient as it employs nearest employs nearest neighbor strategy for each feature key points. VFC, PSSC adopts sparse approximation stands in the middle and they are quite efficient. When the inlier ratio is quite low it's really challenging to solve the matching problem but our PSSC works well even there is more number of false matches.

CONCLUSION

In this work, a progressive strategy called Progressive Sparse Spatial Consensus is used for image feature matching. Mismatch removal is carried out by performing EM algorithm iteratively along with maximum likelihood model. TPS kernel facilitates our PSSC to handle rigid and non-rigid transformation of image pairs. Matching strategy is applied progressively to enable our approach to cover more number of true matches. The quantitative outcomes obtained on publicly accessible data sets are verified with the results of various algorithms shows that our approach outstands in the rate of precision, recall and f-measure specifically in the case of large-scale outliers.

REFERENCES

[1] J. Ma, J. Zhao, J. Jiang, and H. Zhou, "Non-rigid point set registration with robust transformation estimation under manifold regularization," in Proc. AAAI Conf. Artif. Intel, pp. 4218-4224, 2017.

[2] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography," Commun. ACM, vol. 24, no. 6, pp. 381-395, 1981.

[3] X. Li and Z. Hu, "Rejecting mismatches by correspondence function," Int. J. Comput. Vis., vol. 89, no. 1, pp. 1-17, 2010.

[4] J. Chen, J. Ma, C. Yang, and J. Tian, "Mismatch removal via coherent spatial relations," J. Electron. Imaging, vol. 23, no. 4, p. 043012, 2014.

[5] J. Ma, J. Zhao, J. Tian, A. L. Yuille, and Z. Tu, "Robust point matching via vector field consensus," IEEE Trans. Image Process., vol. 23, no. 4, pp. 1706-1721, Apr. 2014.

[6] J. Ma, H. Zhou, J. Zhao, Y. Gao, J. Jiang, and J. Tian, "Robust feature matching for remote sensing image registration via locally linear transforming," IEEE Trans. Geosci. Remote Sens., vol. 53, no. 12, pp. 6469-6481, Dec. 2015.

[7] G.Wahba, Spline Models for Observational Data. Philadelphi, PA, USA: SIAM, 1990.

[8] B.D. Lucas and T.Kanade, "An iterative image registration technique with an application to stereo vision," in Proc. Int. Joint Conf. Artif. Intel. pp. 674-679, 1981.

[9] J. Ashburner, "A fast diffeomorphic image registration algorithm," vol. 38, no. 1, pp. 95-113, 2007.

[10] FujiaJu, Yanfeng Sun, JunbinGao, Siemeng Liu, YongliHu, "Image outlier detection and Feature Extraction via L1-Norm based 2D Probabilistic PCA", IEEE trans. Image Processing, vol.24 ,no. 12, 2015.

[11] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes,"IEEE Trans. Pattern Anal. Mach. Intell., vol. 14, no. 2, pp. 239–256, Feb. 1992.

[12] Y. Liu, "Improving ICP with easy implementation for free-form surface matching," Pattern Recognit., vol. 37, no. 2, pp. 211–226, Feb. 2004.

[13] H. Chui and A. Rangarajan, "A new point matching algorithm for nonrigid registration, "Comput. Vis. Image Understand., vol. 89, no. 2/3, pp. 114–141, Feb. 2003.

[14] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 4, pp. 509–522, Apr. 2002.

[15] H. Liu and S. Yan, "Common visual pattern discovery via spatially coherent correspondence," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 1609-1616, Jun. 2010.

[16] Kavitha H, Dr. M V Sudhamani, "Expirmental analysis of SIFT and SURF features for Multi Object Image Retrieval", in International Journal of Computational Vision and Robotics, vol. 7, pp. 344-356, 2017.

[17] Kavitha H, Dr. M V Sudhamani, "Analysis of Multi Object Image Retrieval based on Combination of HOG & BEMD and SIFT Features", in International Journal

of Applied Engineering Research, ISSN 0973-4562, vol.10, pp. 418-425, Dec. 2015.
[18] O. Pele and M. Werman, “A linear time histogram metric for improved SIFT matching,” in Proc. Eur. Conf. Comput. Vis., pp. 495-508, 2008.
[19] A. Vedaldi and B. Fulkerson, “VLFeat-An open and portable library of computer vision algorithms,” in Proc. ACM Int. Conf. Multimedia, pp. 1469-1472, 2010.