SIFT, SURF, Gabor and Fused Feature Classification Using SVM for Multispectral Satellite Image Retrieval

Chandani Joshi[1], G N Purohit[2] and Saurabh Mukherjee[3]

AIM&ACT, Department of Computer Science

Banasthali University, Rajasthan, India

Abstract: *Content based image retrieval system is becoming a challenging task in every field, as the volume of the data is growing day by day. The earth observatory system is producing immense high resolution images daily, these satellite missions demands for the new approaches to manage and retrieve the satellite images efficiently. The experiment is conducted on the multispectral satellite images, of Landsat 8 sensor. The experiment is based on high level feature extraction techniques i.e. Scale Invariant Feature Transform (SIFT), Speeded Up Robust Features (SURF) and Gabor descriptors. These techniques are fused together to bridge the semantic gap of the retrieval. The feature vector created by the technique used is compared by the feature vector of the query image using the Mahalanobis similarity measure technique. The Precision and Recall is computed for the data set. The results have shown the improved retrieval rate. The retrieval efficiency is further increased by using the SVM classifier by classifying the satellite images based on Urban area, Water body and Vegetation. The experimental results shows that the fusion technique gives better result and more accuracy can be obtained by classifying the dataset using SVM*

Keywords: *SIFT, SURF, CBIR, Landsat 8, Precision, Recall, Mahalanobis*, *and SVM.*

1. Introduction

The retrieval and matching of features taken from different sensors at different time and viewpoints in remote sensing had become a challenging task [9].The need arises for such a system, which can handle the task efficiently and accurately. Content based image retrieval based on the low level feature extraction techniques does not bridge the semantic gap [15]. It is found that for a satellite image, the texture feature extractor plays a vital role. The recent research also focuses on the texture feature extraction techniques. In this paper the Gabor filter is implemented for the feature extraction. In a human visual system, the visual information processing is done by the multi-channel filtering theory, the Gabor technique is enthused by the same concept. In this theory, the image is decomposed into a number of filtered images of a specified amplitude, frequency, and orientation. Gabor filters have been used extensively in image analysis due to their nature of orientation selectivity, spatial locality and frequency characteristic [2]. Along with the Gabor the SIFT and SURF descriptor is used for the feature extraction. SIFT descriptor is invariant to orientation, uniform scaling and illumination changes while the SURF is relevant for its fastest speed of retrieval. Up to some extent invariant to affine distortion as well. These properties make these descriptors significant for the satellite images. In the experiment the results of both the extractors are compared with the proposed technique.

A satellite image consists of multiple classes, for retrieving the matched results accurately the images should be classified. Satellite image classification is a prevailing technique to extract information from massive number of satellite images and is also a process of grouping pixels into meaningful classes [10]. For the classification SVM is used in the paper, as it is designed for the searching of the optimal solution of a problem as compared to other classification techniques. The researchers have found that SVM produces more accurate results then the other techniques such as decision tree and neural network [11]. The obtained results are matched using the Mahalanobis distance, according to the ranking of the images.

https://doi.org/10.17758/ERPUB.ER1217401

The experiment is performed on the Landsat 8 sensors data of 30 meter resolution, of nearby Banasthali region, district Tonk, Rajasthan, India. In the experiment the images are classified based on three classes i.e. the Urban area, Water Body and Vegetation.

2. Proposed Methodology

Fig.1, shows the methodology chart of the proposed system. Initially, the features of the images are extracted using the Gabor, SURF and SIFT descriptors. Then the proposed extraction technique is employed on the data set. The steps for the processing are given below:

- Perform the pre-processing of the satellite image in Erdas Imagine, 2014.
- Convert the RGB images into graylevel images.
- Store the images in the database.
- Input the query image.
- Extract the features of the image using the Gabor, SURF and SIFT technique to form feature vector.
- Apply the Fused model to form the feature vector of the corresponding image.
- Calculate the similarity measure using the Mahalanobis distance measure.
- Retrieve the relevant images based on the similarity measures.
- Calculate the Precision and Recall,

  Where, Precision is calculated by:

  $$\frac{\textit{Number of relevant images retrieved}}{\textit{Total number of images retrieved}}$$

  And Recall by:

  $$\frac{\textit{Number of relevant images retrieved}}{\textit{Total number of relevant images in the database}}$$

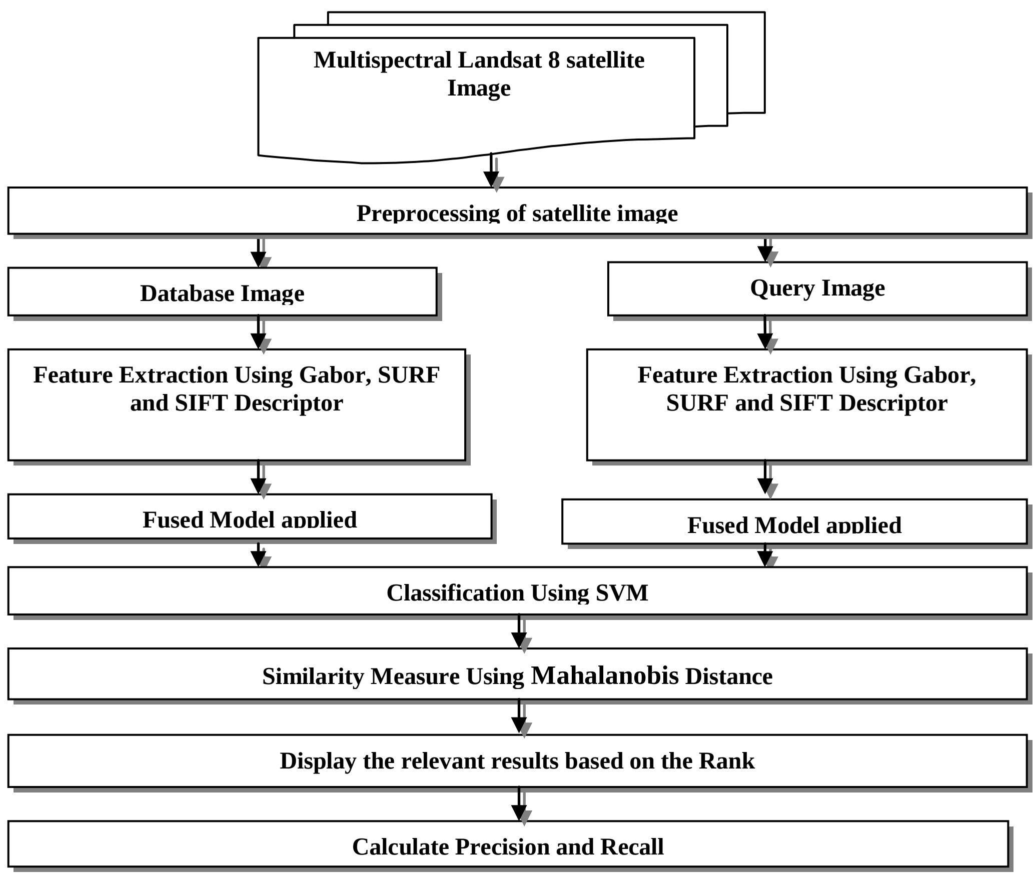


Fig. 1:Methodology chart

2.1. Image descriptors

2.1.1. SIFT Descriptors

In our work local SIFT features are used. The SIFT proposed by Lowe is" invariant to rotation, scaling, affine deformation, slightly to noise, change in illumination and also viewpoint change" [16]. The working of SIFT is similar to that of human vision system. Therefore, it is robust to occlusion and clutter. The four major computational stages for generating the image features are:

- **Scale-space extrema** – In the first step, with all possible scales and locations the image is searched. This is done by efficient usage of the "difference-of-Gaussian" function to identify the scale and orientation invariant interest points. This step contributes for identifying the potential locations that can be considered for finding features.
- **Keypoint localisation** – in this stage at each interest point detected, the location and scale are determined by using a detailed model. Based on the stability, the keypoints are selected. Step 2 eliminates the outlier by considering the extrema and thus aids in accurately locating the feature points.
- **Orientation assignment** – based on the directions of the local image gradients all the keypoint locations are assigned one or more orientations. This step deals with the rotation invariance. The central derivate, gradient magnitude and the direction of smooth image at the keypoint are calculated. Based on these parameters a weighted direction histogram is constructed in the neighborhood of the keypoint. The keypoint direction stands by the direction of the peak. The operations in future on these images are performed on the assigned scale, location and orientation, thus making the images invariant to transformations.
- **Keypoint descriptor** – for each keypoint, the local region around the keypoint is considered for measuring the local gradients that are further transformed to a representation, which can withstand significant levels of local shape distortion and also illumination changes. The gradient oriented histograms are computed for the 16 × 16 neighborhood at the keypoints. Then eight bin weighted histogram is computed for each 4 × 4 regions. The resultant 16 histograms are concatenated to form a 128 dimensional vector [13].

2.1.2 Gabor Texture Features

The use of Gabor feature extractor has proved very effective in analyzing remotely sensed imagery. By applying the orientation and scale selective Gabor filters to an image, the Gabor texture features are extracted. The filter bank consists of S scales and R orientation results, which gives the total of RS filter image as given below [1]:

$$f'_{11}(x, y), \ldots\ldots f'_{RS}(x, y) \tag{1}$$

A 2 dimensional RS feature vector, i.e. global Gabor texture feature, $Gabor_{Global}$ is formed by calculating the standard deviation and mean of the set local Gabor texture feature filtered images.

$$Gabor_{Global} = [\mu_{11}, \sigma_{11}, \mu_{12}, \sigma_{12} \ldots \mu_{RS}, \sigma_{RS}] \tag{2}$$

Where, μ_{RS} and σ_{RS} are the mean and standard deviation of $f'_{RS}(x, y)$. At the end, to normalize the difference in ranges, each of the 2-RS components is scaled to have a mean of zero and a standard deviation of one across a dataset [1].

2.1.3 SURF Descriptor

In order to detect interest points, the Hessian matrix is approximated using a set of box-type filters shown in Fig. 2. These 9 x 9 box filters approximate second order Gaussian derivatives in y- and xy-direction with s = 1.2 and represent the lowest scale for computing blob response maps. These derivatives are referred as D_{yy} and D_x respectively the singular

points calculation of the Hessian matrix of SURF is based on the computation of the determinant of the Hessian matrix.

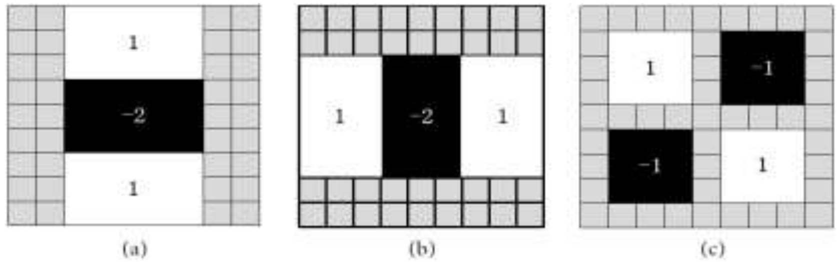


Fig. 2(a), (b) and (c): Filters used to find the Hessian Matrix in SURF.

The white areas correspond to the value +1, the black ones –2 (in the third filter – 1), gray – zero. So the Hessian calculation of SURF can be calculated as:

$$det(H_{approx}) = D_{xx}D_{yy-} \tag{3}$$

where, D_{xx}, D_{yy}, D_{xy} are the convolution products by the filters shown in the above figure.

For orientation assignment, wavelet responses are used with adequate Gaussian weights. The dominant orientation is estimated by calculating the sum of all responses within a sliding orientation window of angle 60 degrees.

Also the sign of Laplacian (trace of Hessian Matrix) is used for underlying interest point, which is already computed during detection. It distinguishes bright blobs on dark backgrounds from the reverse situation. In the matching stage, we only compare features if they have the same type of contrast, as shown in Fig. 3.This minimal information allows for faster matching, without reducing the descriptor's performance.

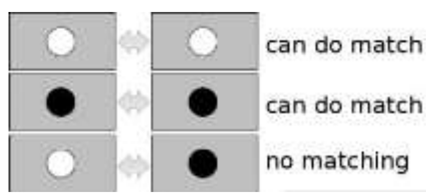


Fig 3:.Matching of blobs on the basis of Contrast in SURF

2.2 Classification

2.2.1 SVM

SVM is a supervised classification method based on statistical learning. "Structural Risk Minimization" (SRM) principle is in use in SVM. In SVM the low dimension feature space is transformed to high dimension feature space which maximizes the margin [7].

The study by the researcher shows that SVM outperforms as compared to the ANN because of the some problems encountered in ANN such as the over-fitting, local minima and sensitivity to the dimensionality of the data , while the SVM has given more accurate results even with a small number of training samples [3].

SVM is exemplified by an efficient hyperplane searching technique, the technique consumes the less processing time by using the minimal training area. The method is able to evade over fitting problem and requires no assumption on data type [4].

Formerly, SVM was used as the binary classifiers that use to correctly divide the data points into two classes, by identifying the optimal hyperplane. Among the infinite hyperplane, the highest margin hyperplane will be selected by the SVM. The distance between the training points (support vector) and the classifier are indicated by the margins. Fig. 4 explains the fundamental concept of support vector machine [4].Many techniques can be implemented to develop the classifier from binary to multiclass i.e. one against all and one against one [9]. In two categories the data is classified in SVM i.e. linearly or nonlinearly. For nonlinear data "Kernel function" is used. In case of linear data, it tries to reduce the training inaccuracy by locating along all hyperplanes. The researchers have found that SVM produces higher accuracy rate as compared to other classifiers [3], [5], [6].

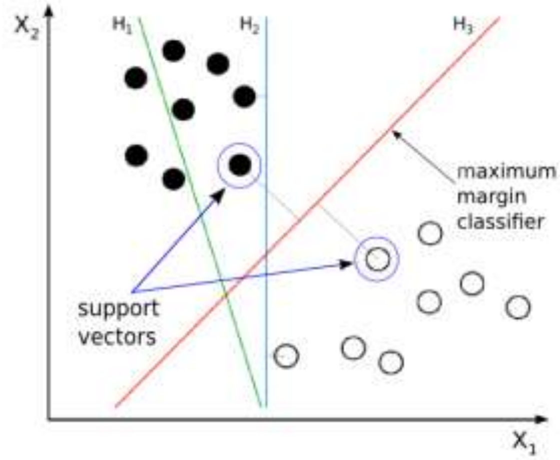


Fig.4. SVM Concept Representation.

The SVM is aimed to separate input pixels in a plan using a hyper plane shown in Fig. 4. The hyper plane is a "plane in a multidimensional space and is also called a decision surface or an optimal separating hyper plane or an optimal margin hyper plane" [8]. The hyper plane is described by the equation as:

$$w.x+b = 0 \tag{4}$$

where, b is a constant (bias or threshold), w is the normal to the hyper-plane (weights).

2.3 Similarity Measure

2.3.1 Mahalanobis Distance

The Mahalanobis distance is defined as the distance between a point P and a distribution D. It has been introduced by P. Mahalanobis in year 1936. It is a multidimensional generalization technique to measure, how many standard deviations away the P is from the mean of D. If the P is at mean D, the distance is Zero, and the value grows as the P moves away from the mean. Along each principal component axis, it measures the number of standard deviations from P to the mean of D. If each of these axes is rescaled to have unit variance, then Mahalanobis distance corresponds to standard Euclidean distance in the transformed space. Mahalanobis distance is thus unit less and scale-invariant, and takes into account the correlations of the data set. The Mahalanobis distance of an observation $X= (X_1, X_2, X_3 \ldots. X_n)^T$ from a group of observation with mean $\mu = (\mu_1, \mu_2, \mu_3 \ \ldots.. \mu_n)^T$ and the Covariance matrix is defined as:

$$D_M(x) = \sqrt{(X-\mu)^T R^{-1} (X-\mu)} \tag{5}$$

3. Results and Discussion

Confusion matrix is detailed in Fig. 5. It consisted of three classes i.e. Urban area, Water body and Vegetation. These parameters are well recognized due to their particular textural appearance. Classified rate is more than 87.00% for Urban area and Vegetation. For the corresponding input image, the existence of water body is not denser.The Precision and Recall computed for the SIFT, SURF, Gabor and the proposed technique is shown in Table I. Among the SIFT, SURF and Gabor descriptor, the Precision and Recall value of Gabor descriptor is high, because satellite images consists more phenomenal textural feature The proposed method has obtained the higher retrieval rate. It shows that the fusion of all three techniques yield better results. Table II, shows the obtained results with respect to the SVM classifier used. The outcomes are better then, the outcomes shown in Table I. Fig. 7 depict the graphical representation of the Precision and Recall values corresponding to the techniques used with and without using the classifier. Top 20 ranked images of the proposed technique with and without classifier are shown in the Fig. 6.

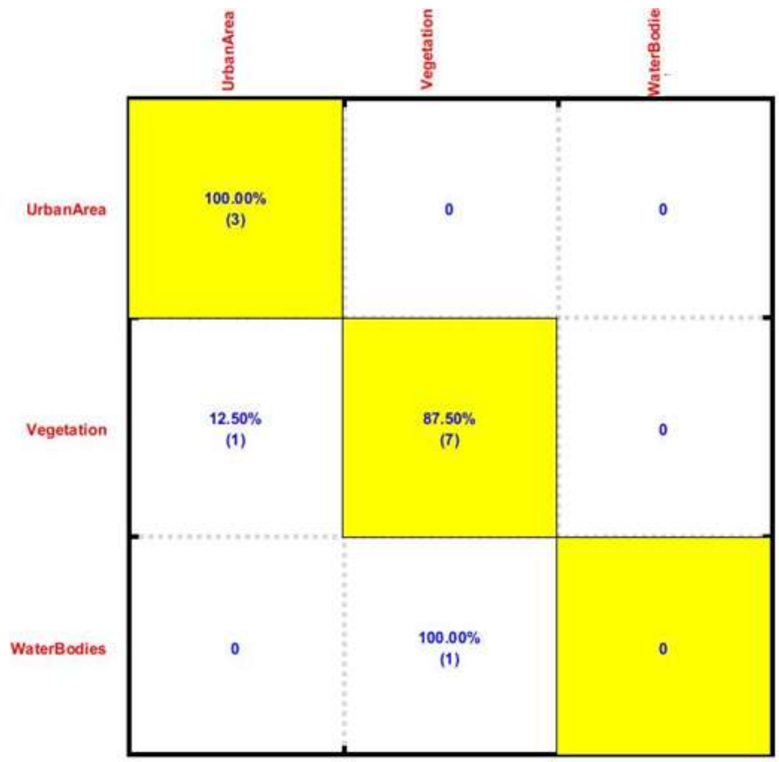


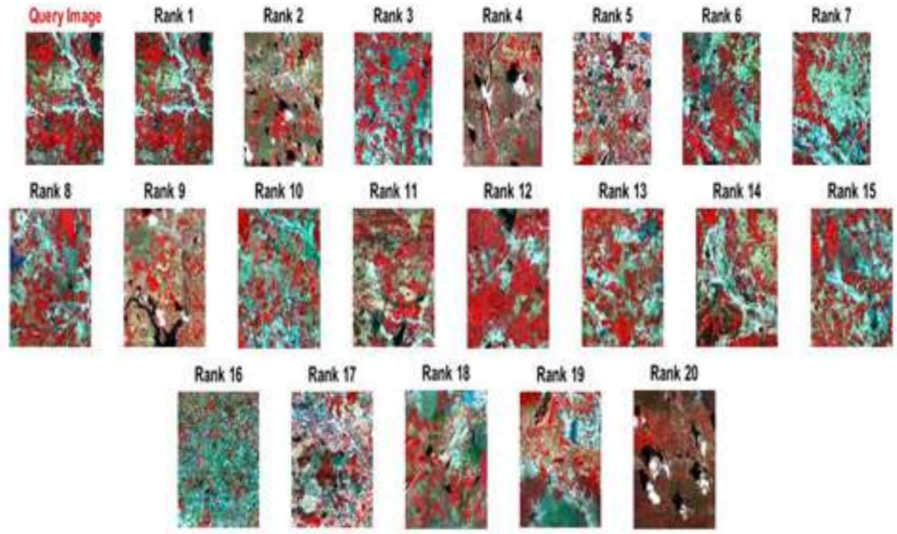
Fig.5 Confusion Matrix for three classes.

TABLE I. Precision and Recall without using SVM classifier

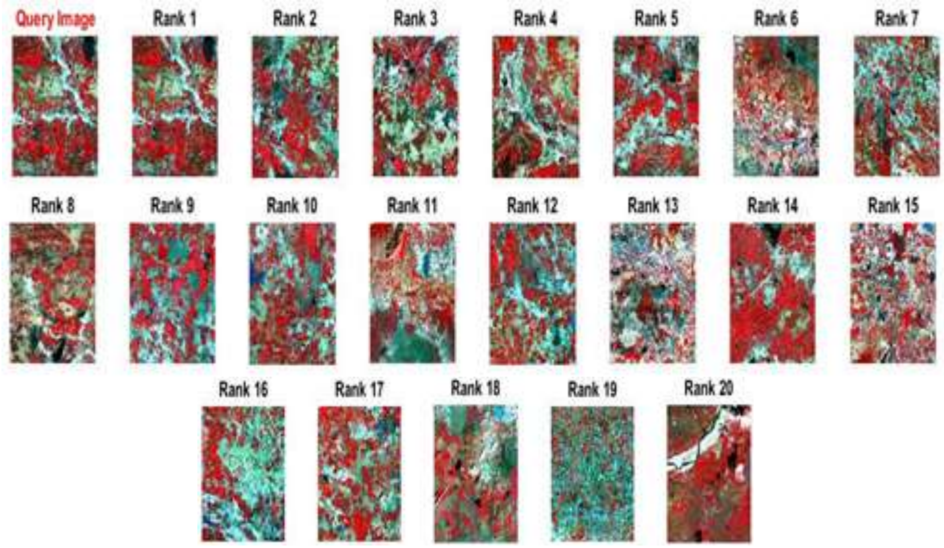
Feature Descriptors without classification	**Categories**					
	Urban		***Vegetation***		***Water Body***	
	Precision	***Recall***	***Precision***	***Recall***	***Precision***	***Recall***
SIFT	0.71	0.59	0.79	0.56	0.70	0.62
SURF	0.73	0.60	0.88	0.60	0.64	0.61
Gabor	0.80	0.67	0.85	0.65	0.68	0.59
Fused	0.90	0.70	0.91	0.71	0.76`	0.69

TABLE II. Precision and Recall with using SVM classifier

Feature Descriptors with classification	**Categories**					
	Urban		***Vegetation***		***Water Body***	
	Precision	***Recall***	***Precision***	***Recall***	***Precision***	***Recall***
SIFT	0.74	0.64	0.79	0.61	0.70	0.63
SURF	0.79	0.70	0.85	0.65	0.65	0.62
Gabor	0.82	0.76	0.90	0.70	0.72	0.60
Fused	0.93	0.81	0.95	0.75	0.80	0.71

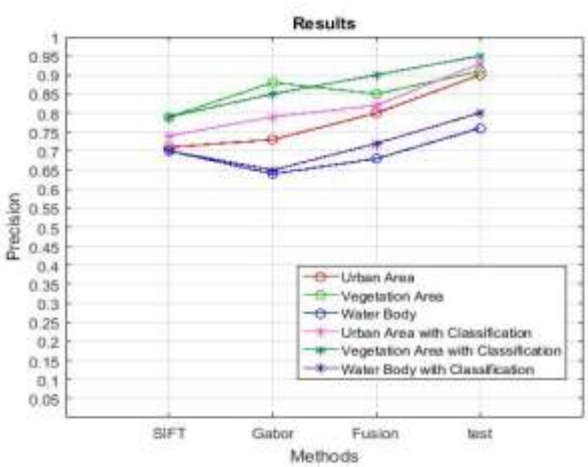


(a)

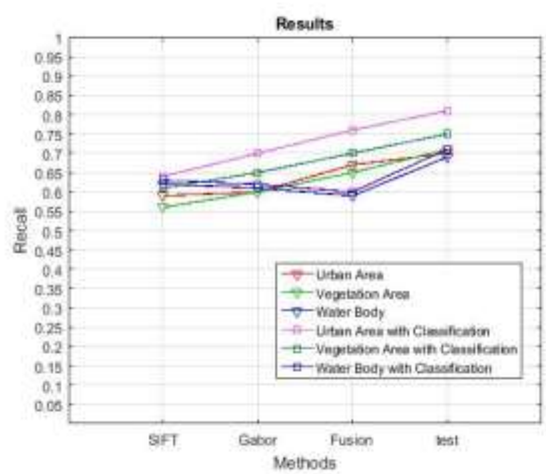


(b)

Fig 6 (a) and (b) Proposed technique based extracted ranked results with and without using SVM respectively.



(a)



(b)

Fig 7(a) and (b) Comparison of Precision and Recall values of the proposed method with other existing method with and without using classifier respectively.

4. Conclusion and future work

Retrieval of the accurate images in the image retrieval system with less computational time is an exigent task. In the experiment we have tried to propose a novel algorithm using the SIFT, SURF and Gabor descriptors for the feature extraction. From the experimental results, it can be concluded that the proposed algorithm have shown better results as compared to the existing techniques being used individually. Since all the three descriptors are well suited for the image type i.e. satellite images, so they have been used for the present experiment. This experiment aimed to find the reliable feature extraction and classification technique.SVM classifier is used with the existing techniques as well as with the proposed techniques. Table 2, shows the increased accuracy with SVM in all the techniques.

In future the same techniques can be applied with the other soft computing techniques. Other feature extraction technique can be used together to yield much better results. The techniques can be tested over the hyperspectral data.

5. References

[1] Yi. Yang,, and S. Newsam. "Comparing SIFT descriptors and Gabor texture features for classification of remote sensed imagery." Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on. IEEE, 2008.
https://doi.org/10.1109/ICIP.2008.4712139

[2] Alexandra, Noramiza, Chikannan. "Content Based Image reterival and classification using Speeded-Up Robust feature (SURF) and grouped Bag-of-Visual-Words(GBoVW)."International Conferecne on Technology , Informatics, Management, engineering &Environment June 23-26 ,2013.

[3] B. Dixon, and C. Nivedita . "Multispectral landuse classification using neural networks and support vector machines: one or the other, or both?." International Journal of Remote Sensing 29.4 :1185-1206, 2008.
https://doi.org/10.1080/01431160701294661

[4] N.I.S. Bahari, , A. Ahmad, and B. M. Aboobaider. "Application of support vector machine for classification of multispectral data." IOP Conference Series: Earth and Environmental Science. Vol. 20. No. 1. IOP Publishing, 2014.
https://doi.org/10.1088/1755-1315/20/1/012038

[5] G.M. Foody, and A. Mathur. "A relative evaluation of multiclass image classification by support vector machines." IEEE Transactions on geoscience and remote sensing 42.6 : 1335-1343, 2014.
https://doi.org/10.1109/TGRS.2004.827257

[6] D. Shi and X. Yang. "Support Vector Machine for Landscap e Mapping from Remote Sensor Imagery." AutoCarto 2012 .

[7] S.M. Jog, M. Dixit, A. Rajgopalan, and S.D. Ranade "KNN and SVM based Satellite Image Classification.", JIREEICE, ISSN (Online) 2321 – 2004.

[8] A. Bekaddour, A. Bessaid, and F. T. Bendimerad. "Multi Spectral Satellite Image Ensembles Classification Combining k-means, LVQ and SVM Classification Techniques." Journal of the Indian Society of Remote Sensing 43.4 (2015): 671-686.
https://doi.org/10.1007/s12524-014-0445-x

[9] M. Tahoun, A.R. Shabayek, H.Nassar, M.M. Giovenco, R. Reulke, E. Emary, and A. E. Hassanien "Satellite Image Matching and Registration: A Comparative Study Using Invariant Local Features." Image Feature Detectors and Descriptors. Springer International Publishing, 2016. 135-171, 2016.
https://doi.org/10.1007/978-3-319-28854-3_6

[10] U. Gandhimathi, and AN. Valliyapan, International Journal for trends in Engineering and technology Voulme 13 Issue 1 – MAY 2016 - ISSN: 2349 – 9303.

[11] A. Bekkari, S. Idbraim, A. Elhassouny, D. Mammass, M.E. Yassa and D. Ducort. "SVM classification of high resolution urban satellite images using Haralick feature “.International Journal of Scientific & Engineering Research, Volume 3, Issue 6, June-2012.

[12] A. Hasnat, S. Halder, A. Hoque, D. Bhattacharjee, and M. Nasipuri. “A FAST FPGA architexcture for measuring the distance between two colorv images using Manahattan distance metric.” International Journal of Electronics and Communication Engineering & Technology (IJECET), ISSN 0976 – 6472, June, 2013.

[13] H. Kavitha, and M. V. Sudhamani. "Experimental analysis of SIFT and SURF features for multi-object image retrieval." International Journal of Computational Vision and Robotics 7.3 (2017): 344-356.
https://doi.org/10.1504/IJCVR.2017.10004039

[14] Shodhganga.inflibnet.ac.in/bitstream/10603/33597/12/12_chapter4.pdf

[15] C. Joshi, G.N Purohit, S. Mukherjee., "Imapct of CBIR journey in Satellite Imaging", Communication and Computing Systems by Taylor and Fransic Group,London, ISBN 978-1-138-02952-1, 2017.

[16] Lowe, G. David "Distinctive image features from scale-invariant keypoints." International journal of computer vision 60.2 (2004): 91-110, 2004.