ISSN:0975-9646

Content Based Image Retrieval: Integration of Neural Networks Using Speed-Up Robust Feature and SVM

Sukhmanjeet Kaur [#1], Mr. Prince Verma [*2]

[#]*M.tech CSE Student, CTIEMT*
[2]*Assistant Professor,*
Department of Computer Science, CTIEMT

Abstract--- In this paper, we present an efficient algorithm based on SURF (Speeded up Robust Features), SVM and NN. The method applies the SURF algorithm in the detection and description for image features; first it applies the SURF feature detector in extracting reference images and matching feature points in the image, respectively. In the process of feature points matching; the false matching points are eliminated through this algorithm. Finally, according to the rest of the match point which can estimate the space geometric transformation parameters between two images and thus matching process is completed. In this thesis, SURF algorithm is used to detect and descript the interest points; and match the interest points by using Surf [1, 3]. In this thesis, the same is tried to retrieve with the use of SURF and fed into Support Vector Machine (SVM) and NN (Neural Network) for further classification. The SURF technique is fast and robust interest points detector which is used in many computer vision applications. For the implementation of this proposed work we use the Image Processing Toolbox under MATLAB Software.

Keywords—Image processing, Matching, Surf, Neural Network and SVM.

1. INTRODUCTION

CBIR or **Content Based Image Retrieval** is the retrieval of images based on visual features such as texture, colour and shape. Reasons for its development are that in many large image databases and traditional methods of image indexing have proven to be insufficient, laborious and extremely time consumed. These old methods of image indexing and ranging from storing an image in the database and associating it with a keyword or number to associating it with a categorized description. In CBIR, each image that is stored in the database has its features extracted and compared to the features of the query image. At present, the image matching methods can be roughly divided into two classes; one is the image matching based on image matching and feature matching. Matching method is directly use the image grey value to determine the space geometry transform between the images, this method can make full use of the information of the image, so it is also known as the matching method based on integral image. It has no feature detection steps in the feature matching stage. The size of window is fixed and even whole image matching are adopted in estimation. So the calculation is simple and also easy to be performed. In recent years, very large collections of images and videos have grown rapidly. In parallel with this growth, con-tent based retrieval and querying the indexed collections are required to access the information that are visual. Two of the main components of the visual information are texture and color. The history of the content-based image retrieval can be divided into three phases:

- The retrieval based on artificial notes.
- The retrieval based on vision character of image contents.
- The retrieval image is based on image semantic features.

The image retrieval that is based on artificial notes labels images by using text firstly, in fact it has already changed image retrieval into traditional keywords retrieval. Problem with the approach is that, it brings heavy workload and it still remains subjectivity and uncertainty. Because the image retrieval that is based on artificial notes still remains insufficiency that adapts vision image features has been come up and become the main study. The character of this method is image feature extraction impersonally whether the retrieval is good or not depends on the features extraction accuracy. So the research based on vision features is becoming the focus in the academic community. The feature of vision can be classified by semantic hierarchy into middle level feature and low- level feature. Low-level feature includes texture, color and inflexion. Middle level involves shape description and object feature. Content based Image Retrieval systems try to retrieve images similar to a user-defined specification or pattern. Their goal is to support image retrieval based on content properties. One of the main advantages of the CBIR approach is the possibility of an automatic retrieval process instead of the traditional keyword-based approach which usually requires very laborious and time-consuming previous annotation of database images.

CBIR process involves two steps:

- The **Feature Extraction** is the first step in the process is extracting image features to a distinguishable extent.
- The **Matching** is the second step involves matching these features to yield a result that is visually similar.There are several CBIR systems that are currently exist and are being developed constantly:
- **QBIC** or **Query by Image Content** was developed by IBM and Almaden Research Centre to allow users to graphically pose and refine queries based on multiple visual properties such as colour, texture and shape. It supports queries based on input images, user-constructed sketches, selected colour and texture patterns.

- **VIR Image Engine** by Virage like QBIC enables image retrieval based on primitive attributes such as texture, colour and structure. It examines the pixels in the image and performs an analysis process for deriving image characterization features.
- **Visual SEEK** and **Web SEEK** were developed by the Department of Electrical Engineering of Columbia University where both these systems support colour and spatial location matching as well as texture matching.
- **NeTra** was developed by the Department of Electrical and Computer Engineering, University of California. It supports spatial layout, colour, shape and texture matching as well as image segmentation.
- **MARS** or **Multimedia Analysis and Retrieval System** was developed by the Beckman Institute for Advanced Science and Technology of Illinois University. It supports texture, colour, spatial layout and shape matching.
- **Viper** or **Visual Information Processing for Enhanced Retrieval** was developed at the Computer Vision Group of Geneva University. It supports colour and texture matching.

1.1. SPEEDED UP ROBUST FEATURE

SURF (Speeded up Robust Features) is a robust local feature detector; first presented by Herbert Bay et al in 2006; that can be used in computer vision tasks like object recognition or 3D reconstruction. This is partly inspired by the SIFT descriptor. Therefore standard version of SURF is several times faster than SIFT and claimed by its authors to be more robust against different image transformations than SIFT. And SURF is based on sums of 2D Haar wavelet responses and makes an efficient use of integral images. This uses an integer approximation to the determinant of Hessian blob detector; which can be computed extremely quickly with an integral image. Therefore For features; it uses the sum of the Haar wavelet response around the point of interest. These can be computed with the aid of the integral image. SURF used in this approach to extract relevant features and descriptors from images. This approach is preferred over its predecessor due to its succinct descriptor length for instance 64 floating point values.

In SURF, a descriptor vector of length 64 is constructed using a histogram of gradient orientations in the local neighborhood around each key point. Modified SURF (Speeded up Robust Features) is one of the famous feature-detection algorithms [11, 17]. The panorama image stitching system which combines an image matching algorithm; modified SURF and an image blending algorithm; multi-band blending. This process is divided in the following steps: first; get feature descriptor of the image using modified SURF; secondly; find matching pairs; using correlation matrix; and remove the mismatch couples by RANSAC (Random Sample Consensus); then; adjust the images by bundle adjustment and estimate the accurate homographic matrix; lastly; blend images by Alpha blending. And comparison of SIFT (Scale Invariant Feature Transform) and Harris detector are also shown as a base of selection of image matching algorithm.

And according to the experiments; the present system can make the stitching seam invisible and get a perfect panorama for large image data and it is faster than previous method. SURF approximates or even outperforms previously proposed schemes with respect to repeatability; distinctiveness; and robustness; yet can be computed and compared much faster. And this is achieved by relying on integral images for image convolutions; by building on the strengths of the leading existing detector sand descriptors specially using a Hessian matrix-based measure for the detector; and a distribution-based descriptor and by simplifying these methods to the essential [18,20]. This leads to a combination of novel detection; description; and matching steps. It approximates or even outperforms previously proposed schemes with respect to repeatability; distinctiveness; and robustness; yet can be computed and compared much faster and this is achieved by;

- Relying on integral images for image convolutions
- Building on the strengths of the leading existing detectors and descriptors (using a Hessian matrix-based measure for the detector; and a distribution based descriptor).
- Simplifying these methods to the essential.

1.2. NEURAL NETWORKS

Neural network is set of interconnected neurons. This is used for universal approximation. Artificial neural networks are composed of interconnecting artificial neurons (programming constructs that mimic the properties of biological neurons). And artificial neural networks may either be used to gain an understanding of biological neural networks; or for solving artificial intelligence problems without necessarily creating a model of a real biological system which is highly complex. An Artificial neural network algorithms attempt to abstract this complexity and focus on what may hypothetically matter most from an information processing point of view. Good performance (e.g. as measured by good predictive ability; low generalization error); or performance mimicking animal or human error patterns; can then be used as one source of evidence towards supporting the hypothesis that the abstraction really captured something important from the point of view of information processing in the brain of human. And another incentive for these abstractions is to reduce the amount of computation required to simulate artificial neural networks.

1.2.1. Architecture of artificial neural network

The basic architecture consists of three types of neuron layers: input; hidden; and output. And feed-forward networks; the signal flow is from input to output units; strictly in a feed-forward direction. Therefore data processing can extend over multiple layers of units; but no feedback connections are present. The recurrent networks contain feedback connections. The contrary to feed-forward networks; the dynamical properties of the network are important. In some cases; the activation values of the units undergo a relaxation process such that the network will evolve to a stable state in which these activations do not change anymore [12].

1.2.2. Artificial Neural Networks

Artificial neural networks are composed of interconnecting artificial neurons (programming constructs that mimic the properties of biological neurons). Therefore Artificial neural networks may either be used to gain an understanding of biological neural networks; or for solving artificial intelligence problems without necessarily creating a model of a real biological system that is highly complex: artificial neural network algorithms attempt to abstract this complexity and focus on what may hypothetically matter most from an information processing point of view. Good performance (e.g. as measured by good predictive ability and low generalization error), or performance of mimicking an animal or the human error patterns and can then be used as one source of evidence towards supporting the hypothesis that the abstraction really captured something important from the point of view of information processing in the brain [20]. Another incentive for these abstractions is to reduce the amount of computation required to simulate artificial neural networks; so as to allow one to experiment with larger networks and train them on larger data sets. And application areas of ANNs include system identification and control (vehicle control; process control); game-playing and decision making (backgammon, chess, racing), pattern recognition (radar systems; face identification; object recognition); sequence recognition (gesture, speech, handwritten text recognition); medical diagnosis; financial applications; data mining , visualization and e-mail spam filtering.

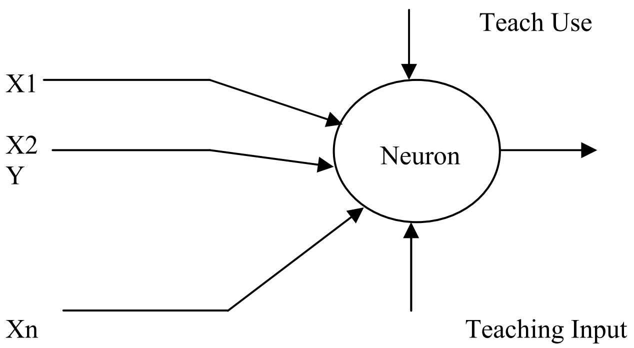

Figure 1: Neural Network

1.2.3. Delta Rule

The delta rule is a gradient descent learning rule for updating the weights of the artificial neurons in a single-layer perceptron. This is a special case of the more general back propagation algorithm. For a neuron j with activation function g(x); the delta rule for j, ith weight is given by

$$\Delta W_{ij} = (t_j - y_j)\, g'(h_j) x_i \quad (1)$$

Therefore delta rule is commonly stated in simplified form for a perceptron with a linear activation function as $\Delta W_{ij} = \alpha\ (t_j - y_j)\ x_i$; where α is known as the learning rate parameter.

1.3. SUPPORT VECTOR MACHINE

The Support Vector Machine (SVM) is a state-of-the-art classification method introduced in 1992 by Boser, Guyon, and Vapnik. The SVM classifier is widely used in bioinformatics (and other disciplines) due to its highly accurate; able to calculate and process the high-dimensional data such as gene expression and exibility in modelling diverse sources of data .SVMs belong to the general category of kernel methods. And a kernel method is an algorithm that depends on the data only through dot-products. This is the case; the dot product can be replaced by a kernel function which computes a dot product in some possibly high dimensional feature space. It has two advantages: First; the ability to generate non-linear decision boundaries using methods designed for linear classifiers. And second; the use of kernel functions allows the user to apply a classifier to data that have no obvious fixed-dimensional vector space representation. Thus prime example of such data in bioinformatics are sequence; either DNA or protein; and protein structure. Using SVMs effectively requires an understanding of how they work. When training an SVM the practitioner needs to make a number of decisions: how to pre-process the data, how to the kernel to use or work; and finally; setting the parameters of the SVM and the kernel [1]. Uninformed choices may result in severely reduced performance. Therefore we aim to provide the user with an intuitive understanding of these choices and provide general usage guidelines [7, 13]. All the examples shown were generated using the PyML machine learning environment, which focuses on kernel methods and SVMs.

1.3.1. Preliminaries: Linear Classifiers

Support vector machines are an example of a linear two-class classifier. This section explains what that means. The data for a two class learning problem consists of objects labelled with one of two labels corresponding to the two classes; for convenience we assume the labels are +1 or -1. In what follows boldface x denotes a vector with components xi. Thus notation xi will denote the ith vector in a dataset, f (xi; yi) gni =1, where yi is the label associated with xi. The boundary between regions classified as positive and negative is called the decision boundary of the classifier. The decision boundary defined by a hyper plane is said to be linear because it is linear in the input examples. A classifier with a linear decision boundary is called a linear classifier. Conversely, when decision of the boundary is a classifier depends upon the data in non-linear the classifier is said to non-linear.

1.3.2. Kernels: From Linear To Non-Linear Classifiers

In many applications a non-linear classifier provides better accuracy. Yet; linear classifiers have advantages; one of them being that they often have simple training algorithms that scale well with the number of examples [9, 10]. This begs the question: Can the machinery of linear classifiers be extended to generate non-linear decision boundaries? Therefore furthermore; can we handle domains such as protein sequences or structures where a representation in a fixed dimensional vector space is not available? The naive way of making a non-linear classifier out of a linear classifier is to map our data from the input space X to a feature space F using a non-linear function.

The approach of explicitly computing non-linear features does not scale well with the number of input features: when applying the mapping from the above example the dimensionality of the feature space F is quadratic in the dimensionality of the original space. The result in a quadratic increase in memory usage for storing the features and a quadratic increase in the time required to compute the discriminant function of the classifier. The complexity of quadratic is feasible for low dimensional data; but when handling gene expression data that can have thousands of dimensions; quadratic complexity in the number of dimensions is not acceptable. And Kernel methods solve this issue by avoiding the step of explicitly mapping the data to a high dimensional feature-space.

Gaussian kernel is defined by:

$$K(\mathbf{x_i},\mathbf{x_j}) = \exp(-\frac{\|\mathbf{x_i}-\mathbf{x_j}\|^2}{2\sigma^2}) \qquad (2)$$

Where k > 0 is a parameter that control the width of Gaussian. It plays a similar role as the degree of the polynomial kernel in controlling the exibility of the resulting classifier. We saw that a linear decision boundary can be kernelized i.e. its dependence on the data is only through dot products. In order for this to be useful, the training algorithms need to be kernelizable as well [6]. It turns out that a large number of machine learning algorithms can be expressed using kernels | including ridge regression, the perceptron algorithm, and SVMs [16].

2. Previous Work

In the previous work, there were multiple methods used to retrieve the images. So that input image will more accurate for understand and better results.eg – one class, two-class, and multiclass SVMs to annotate images for supporting keyword retrieval of image, multiresolution-based signature subspace classifier (MSSC). The MSSC was developed to highly save the computational time. With application to psoriasis images. The essential techniques consist or involved of feature extraction and image segmentation (classification) methods. Other method, confidence-based dynamic ensemble (CDE), which employs a three-level classification scheme At the base-level, the CDE uses one-class Support Vector Machines (SVMs) to characterize a confidence factor for ascertaining the correctness of an annotation (or a class prediction) made by a binary SVM classifier.

3. Proposed Approach

The block diagram of the proposed system is shown in the following Figure.

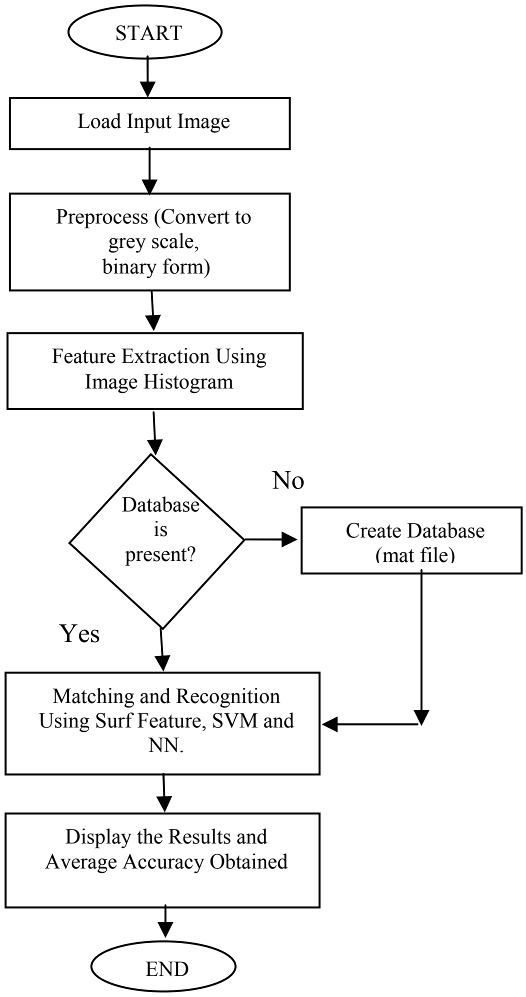

Figure 2: Flow Chart of Proposed Work

4. Results And Discussion

A starting GUI was created to perform all the five operations that is browse input image, process input image, create database, process database and match.

Figure 3: Matched Image

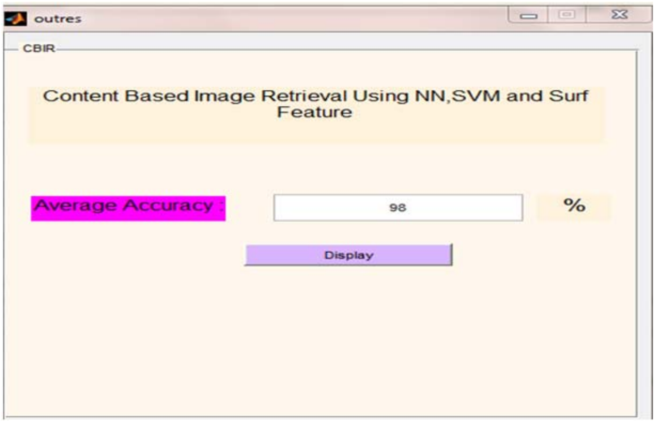

Figure 4: Average Accuracy of CBIR

This figure shows the result of given input image and represents the accuracy of retrieval image.

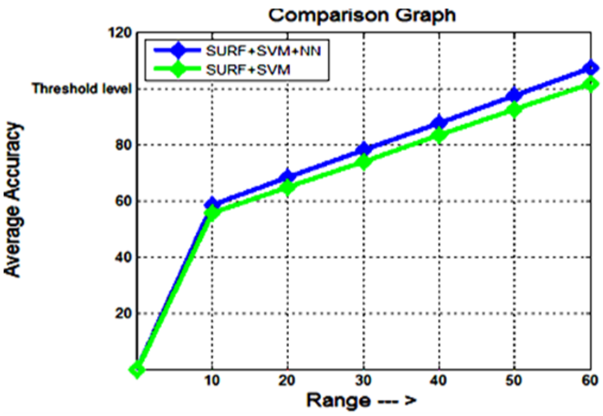

Figure 5: Accuracy between Previous and Proposed work

In this work we are using SURF, SVM and NN for better results. But in previous methods the results are not good. So, in this figure we show that our work is better as compare to previous work.

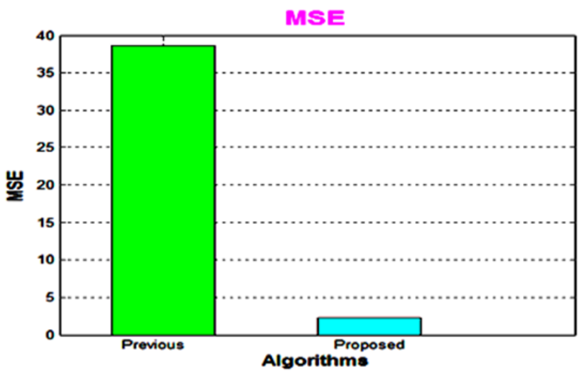

Figure 6: Mean Square Error

Mean square error (MSE) in the previous work was very high and those were reducing the working of accuracy and increase the working of time etc. This figure represents result of our proposed methods.

Table 1: MSE between Previous and Proposed work

	Previous Work	Proposed Work
MSE	38.7000	2.3000

This table show the result of proposed and previous method. In the previous work MSE was very high but in our work MSE result is low as compare to previous method.

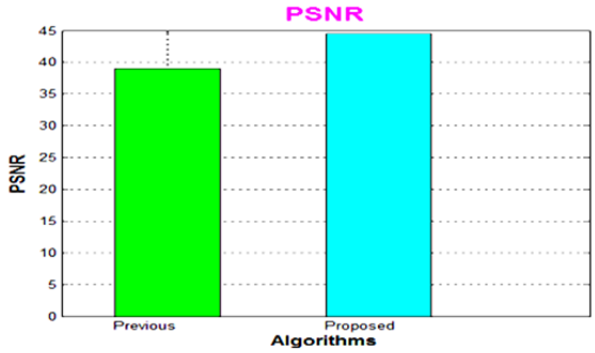

Figure 7: PSNR

Table 2: PSNR between Previous and Proposed work

	Previous Work	Proposed Work
PSNR(db)	39.0585	44.5135

PSNR(Peak signal to noise ratio) increase the working. In the previous work PSNR is low but in proposed method It will increase (shows in table).

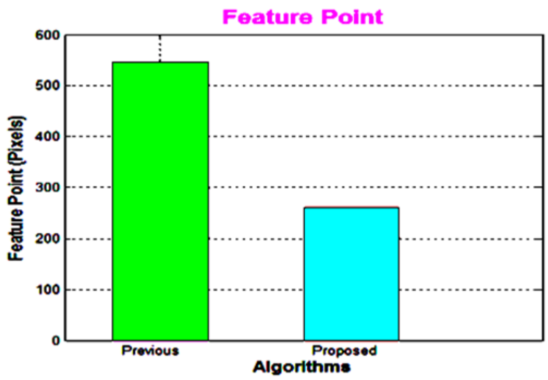

Figure 8: Feature Point

Table 3: Comparison of Feature Point between Previous and Proposed work

	Previous Work	Proposed Work
Feature Point	547	261

Table 4: Matching between Previous and Proposed work

Comparison of Matching time between Previous Work and Proposed Work

	Previous Work	Proposed Work
Time (sec--->)	2.0300	1.8504

This table show the result of proposed and previous method. In the previous work matching time was high but in our work matching time result is low as compare to previous method

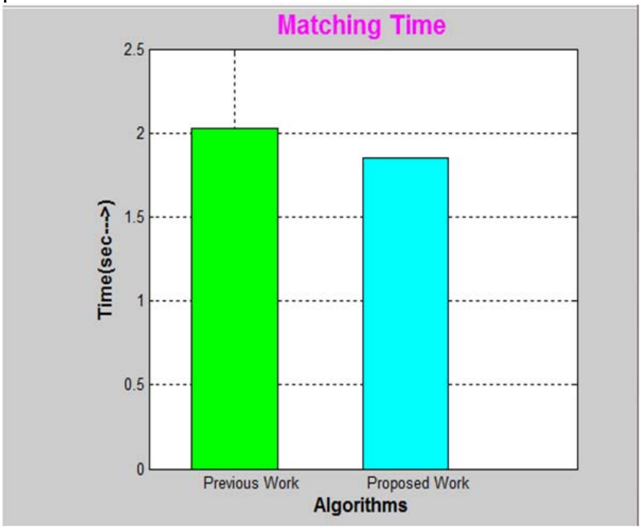

Figure 7: Matching Time

Matching Time in the previous work was very high. This figure represents result of our proposed methods which is better.

5. CONCLUSION

We proposed 'Image Matching Based on Improved SURF Algorithm using SVM Classifier and Neural Network'. Content-Based Image Retrieval (CBIR) is a challenging task which retrieves the similar images from the large database. Therefore most of the CBIR system uses the low-level features such as color; texture and shape to extract the features from the images. Finally it estimates the space geometric transformation parameters between two images and completes the matching according to the rest of the matching point [3, 6]. Many CBIR techniques have been proposed earlier but they were not good enough and can be temporarily tampered with so the task was not fulfilled. CBIR alone with Surf and SVM Method could not provide better results. Therefore we use Content Based Image Retrieval with Surf, SVM and Neural Network providing three level improved results. Even if image is tampered with our Accuracy is not affected and thus our purpose is fulfilled. Better PSNR and CCR results of Content Based Image Retrieval with Surf, SVM and Neural Network.

REFERENCES

[1] B Zitova, J Flusser. Image registration methods: A survey. Image Vis. Comput. 2003; 21(11): 977-1000.

[2] Dai X, Khorram S. Development of a Feature-Based Approach to Automated Image Registration for Multi temporal and Multi sensor Remotely Sensed Imagery. Proceedings of the IEEE Interactional Geosciences and Remoth Sensing Symposium. 1997; 1: 243-245.

[3] Goshtasby A. Mbolieally-assisted approach to digital image registration with application in computer vision. Ph. D. dissertation, Dept. of Computer Science, Miehigan State University, TR83-013. 1983.

[4] Ton J, Jain A. Stering Landsat images by point matehing. IEEE Transactions on Geosciences and Remote Sensing. 1989; 27(5): 642-651.

[5] Li H, Manjunath B, Mitra S. Ntour-Based Approach to Multisensor Image Registration. IEEE Transactions on Image Processing. 1995; 4(3): 320-334.

[6] Tham J, Ranganath S, Ranganath M, Kassim A. Vel Unrestricted Center-Biased Diamond Search Algorithm for Block Motion Estimation. IEEE Transactions on Circuits System and Video Technology. 1998; 8(4): 369-377.

[7] Yuan Z, Wu F, Zhuang Y. I-sensor image registration using multi-resolution shape analysis. Journal of Zhejiang University Science A. 2006; (4): 549-555.

[8] Ziou D, Tabbone S. Dtection techniques-an overview. International Journal of Pattern Recognition and Image Analysis. 1988; (5): 537-559.

[9] Mount D, Netanyahu N, Moigne J. Ent Algorithms for Robust Feature Matching. Pattern Recognition, 1999; 32(1): 17-38.

[10] Bay H, Tuvtellars T, Gool L Van. SURF: speeded up robust features. Proceedings of the European Conference on Computer Vision. 2006: 404 417.

[11] Rong W, Chen H, et al. Icing of Microscope Images based on SURF. 24th International Conference Image and Vision Computing New Zealand (IVCNZ2009). 2009: 272-275.

[12] Hartley R, Zisserman A. View Geometry in Computer Vision, second. ed. Cambridge University Press, Cambridge. 2003.

[13] Tola E, Lepetit V. St local descriptor for dense matching. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society. 2008: 1-8.

[14] Tola E, Lepetit VY: An Efficient Dense Descriptor Applied to Wide-Baseline Stereo. IEEE Transactions on Pattern Analysis & Machine Intelligence. 2010; 32(5): 815-830.

[15] Bracewell R. Fourier Transform and Its Applications. McGraw-Hill, New York. 1965.

[16] Xiaoli Li, Xiaohong Wang and Chunsheng Li: Image Matching Based on Unification, The Fourth International Joint Conference on Computational Science and Optimization (2011), p. 825-828.

[17] Herbert Bay, Andreas Ess, Tinne Tuytelaars and Luc Van Gool: Speeded-Up Robust Features (SURF), Computer Vision and Image Understanding (2008), 110(3), p. 346-359.

[18] David G. Lowe: Distinctive Image Features from Scale-Invariant Keypoints, International Journal of Computer Vision (2004), 60(2), p. 91-110.

[19] Luo Juan and Oubong Gwun: A Comparison of SIFT, PCA-SIFT and SURF, International Journal of Image Processing (2009), 3(4), p. 143-152.

[20] Dorigo M., Maniezzo V. and Colorni A.: Ant System: Optimization by A Colony of Cooperating Agents, IEEE Transaction on Systems, Man, and Cybernetics-Part B (1996), 26(1), p. 29-41.