ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 4, Issue 6 , June 2017

Spatial and Transform domain based Content Based Image Retrieval for classification of Very High Resolution images

Rashmi R V, Sheela Sridhar

P.G. Student, Department of Computer Science, B.N.M.Institute of Technology, Bangalore, Karnataka, India
Associate Professor, Department of Computer Science, B.N.M.Institute of Technology, Bangalore, Karnataka, India

ABSTRACT: In this paper, explore the combination of both spatial and transform domain methods for the satellite image which are very high resolution by applying content base image retrieval concept. The methodology is first applied to supervised classification framework. It is applied to texture database before to classification in order to find the top n best results in the context of the application. Once the best results are identified then supervised classification method is performed by extracting texture features from both the domains and learning database. Then the classification is applied to the selected classifier. This strategy is illustrated with satellite image with different types of images on it. Totally there are six different types of images that are used as the experimental analysis. The advantage of doing is its high adaptability and low parameters.

KEYWORDS: Classification, wavelet, very high resolution (VHR), CBIR

I. INTRODUCTION

Data mining is the process of discovering interesting patterns from massive amounts of data. It is also known as Knowledge Discovery from Data (KDD). As a knowledge discovery process, it typically involves data cleaning, data integration, data selection, data transformation, pattern discovery, pattern evaluation, and knowledge presentation. The major dimensions of data mining are data, knowledge, technologies, and applications.

Image processing is a method to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. It is a type of signal dispensation in which input is image, like video frame or photograph and output may be image or characteristics associated with that image. Usually Image Processing system includes treating images as two dimensional signals while applying already set signal processing methods to them.

Supervised(human guided) classification is based on the idea that a user can select sample pixels in an image that are representative of specific classes and then direct the image processing software to use these training sites as references for the classification of all other pixels in the image. Training sites (also known as testing sets or input classes) are selected based on the knowledge of the user. The user also sets the bounds for how similar other pixels must be to group them together. These bounds are often set based on the spectral characteristics of the training area, plus or minus a certain increment (often based on "brightness" or strength of reflection in specific spectral bands). The user also designates the number of classes that the image is classified into. Many analysts use a combination of supervised and unsupervised classification processes to develop final output analysis and classified maps.

Statistical modeling is much easier if some pre-processing is carried out on the input images. Typical preprocessing is done via transformation of image pixel values into a suitable space where simple models with a small number of parameters can describe the data. Wavelets have recently emerged as an effective tool to analyze texture information as they provide a natural partition of the image spectrum into multiscale and oriented sub bands via efficient transforms. Furthermore, since wavelets are used in major future image compression standards and are also shown to be prominent in searching for images based on texture, a wavelet-based texture retrieval system can be used effectively in conjunction with a compression system and retrieval systems using other image features.

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 4, Issue 6 , June 2017

The satellites have been designed with urgent tasking in mind; images can be requested from the Pleiades satellites less than six hours before they are acquired. This functionality will prove invaluable in situations where the expedited collection of new image data is crucial, such as crisis monitoring. This improved flexibility also leads itself to corridor and persistent surveillance acquisition modes.

II. RELATED WORK

Feature extraction can be done in two different phases: Testing phase and Training phase. In training phase there are some set of satellite with different classes of images. the feature extraction of these images can be done by one of the most popular texture descriptors is GLCM(Grey Level Co-occurrence Matrix) which creates a matrix with distances and directions among the pixels, and then removes significant figures from the matrix as texture features. In this paper, there are four features including energy, contrast, correlation, homogeneity. Apart from these features covariance, sum, standard deviation, mean and many more are there. Totally there are 11 features are extracted for each images in the dataset. The output of the feature extraction will give some values from both the domains. Then similarity measurements are performed based on the signatures generated from feature extraction.

B. Beguet et al proposed in [1] presents a new feature selection method which aims to effectively maps remote sensing data. Feature selection is pre-processing step for classification which provides the most relevant attributes as input features. It will reduce the classification complexity and increase classification accuracy. High dimensional data is generally used as a pre-processing step for classification. It could be hyper spectral image classification or texture analysis. There are mainly three types of feature selection methods are exists they are: filters, wrappers and embedded methods. Filters are independent of the classifier, exploring only the intrinsic discriminative power of features. These methods are fast but suboptimal. Wrappers use the classification algorithm as an evaluation mean. They are powerful but induce very high computational costs. Embedded methods include the feature selection in the learning process.

N-E Lasmar[2] et.al. proposed the framework of texture retrieval, the challenge is to provide baseline algorithms making a system able to retrieve, from a textured image databases, the relevant candidates similar to a given query according to the texture cue. A typical retrieval scheme consists of two major tasks. The first one is devoted to feature extraction, where signatures are estimated from each image in the database and from the query. The second task evaluates a similarity measure, based on previous features, to decide which images of the database are close to the query. Texture browsing or searching systems is to provide a tractable mathematical description of natural textures. Standard random field modeling consists in providing a parametric Probability Density Functions (PDF) which enables us to fit the empirical histograms. Stochastic model-based approaches are theoretically justifiable since information divergences such as the Kullback-Leibler Divergence (KLD) are asymptotic limits of likelihood functions that can be used to measure the similarity between data drawn from different distribution families; the stochastic framework has proven to be asymptotically optimal in terms of the retrieval rate when the KLD between PDF models.

I. Champion[3] et al. proposed for the environmental and forest management applications data on the forest variables such as biomass, height, trunk are required to prove that texture can be used instead of density age or stand age to retrieve forest plantation. The main goal of this paper is to explore the relationship between stand ages which are related to the forest variables. Synthetic Aperture Radar (SAR) systems have demonstrated their potentials for discriminating biomass volumes in young to mature stands, especially at low frequencies. Significant relationship has been established between radar mean intensity and biophysical variables. The importance of this paper is to investigate the relationships between texture descriptors and forest stand age values of different values. Texture descriptors like Grey Level Concurrence Matrix (GLCM) are experimented on a SAR images of pine forest, at cross and parallel polarizations. Various pixel pair configurations were considered for calculation of the GLCM, i.e., the relative orientation (α) and distance (d) separating the two pixels. The results showed that energy and entropy were highly correlated with stand age. For both indicators, texture/stand age regressions were not markedly impacted by the method used to calculate the GLCM and regressions did not reveal any marked optimum or minimum for distance d or direction α that would enhance or reduce regression quality.

L.Gueguen proposed in paper [4] tells about the new compact representation for the fast query or classification of compound structures from very high resolution optical remote sensing imagery. Pixel based and object based classification is becoming widely used for extracting information from imagery. This bag-of-features representation relies on the multiscale segmentation of the input image and the quantization of image structures pooled into visual word distributions for the characterization of compound structures. A compressed form of the visual word distributions is described, allowing adaptive and fast queries/classification of image patterns. The proposed representation and the query methodology are evaluated for the classification of the UC Merced 21-class data set, for the detection of informal settlements and for the discrimination of challenging agricultural classes. The high fidelity of image data provided by

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 4, Issue 6 , June 2017

the new and advanced space borne sensor constellations provides a unique access to the information encoded in the observed image patterns over broad areas. Large amount of data and the large surface covered by very high resolution (VHR) images, it became obvious over the past few years that automatic or semiautomatic detection/classification technique are mandatory to get timely access to the image information content.

III. METHODOLOGY

A. CBIR

The main components of CBIR are the features which includes the Geometric shape, colours and the texture of the image. Features can be of two types like local features and global features. Object recognition can be done easily using the local features. The next component is the associated text in which the images can also be retrieved using the text associated with the image. The other component is the relevant feedback where it helps to be more precise in searching the relevant images by taking up the feedbacks of the user. Biomedicine, Military, Education, Web image classification and searching are some of the areas where the CBIR technique finds its prime importance. Some of the examples for the current CBIR are Viper which is Visual Information Processing for Enhanced Retrieval, QBIC which is Query by Image Content and Visual seek which is a web tool for searching images and videos. CBIR mainly decreases the heavy workload and overcomes the problem of heavy subjectivity.

Texture is one of the important characteristics used in identifying the object in an image. Image analysis involves investigation of the image data for a specific application. Normally, the raw data of a set of images is analysed to gain insight into what is happening with the images and how they can be used to extract desired information. In image processing and pattern recognition, feature extraction is an important step, which is a special form of dimensionality reduction. When the input data is too large to be processed and suspected to be redundant then the data is transformed into a reduced set of feature representations. The process of transforming the input data into a set of features is called feature extraction. Features often contain information relative to colour, shape, texture or context.

CBIR had been a very important and effective research area in many fields. Searching images from large database and giving appropriate results, increased bandwidth availability will help to increase the use of internet by user in future. Therefore a significant difficulty that needs to be look after is fast retrieval of images from large databases. Image retrieval system searches for images from large database and tries to find exact or nearly same images. CBIR can greatly enhance the precision of the data being returned and is an effective to conventional textual-based image searching. Color, texture and histogram features are used to differentiate an image from other images. Research and development issues in CBIR cover a range of topics, many shared with mainstream image processing and information retrieval. Typical pre-processing is done via transformation of image pixel values into a suitable space where simple models with a small number of parameters can describe the data. In early days because of very large image collections the manual annotation approach was more difficult. In order to overcome these difficulties Content Based Image Retrieval (CBIR) was introduced. Content-based image retrieval (CBIR) is the application of computer vision to the image retrieval problem. In this approach instead of being manually annotated by textual keywords, images would be indexed using their own visual contents .The visual contents may be colour, texture and shape. This approach is said to be a general framework of image retrieval .There are three fundamental bases for Content Based Image Retrieval which are visual feature extraction, multidimensional indexing and retrieval system design. The colour aspect can be achieved by the techniques like averaging and histograms. The texture aspect can be achieved by using transforms or vector quantization .The shape aspect can be achieved by using Wavelets have recently emerged as an effective tool to analyze texture information as they provide a natural partition of the image spectrum into multiscale and oriented subbands via efficient transforms wavelet based texture retrieval system can be used effectively in conjunction with a compression system and retrieval systems using other image features which is shown in below fig 1.

Steps involved in CBIR:
1) Features are extracted from each image in the database
2) Same features are extracted from the query image
3) Similarity between query features and database features are compared
4) Images with maximum similarity are retrieved

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 4, Issue 6 , June 2017

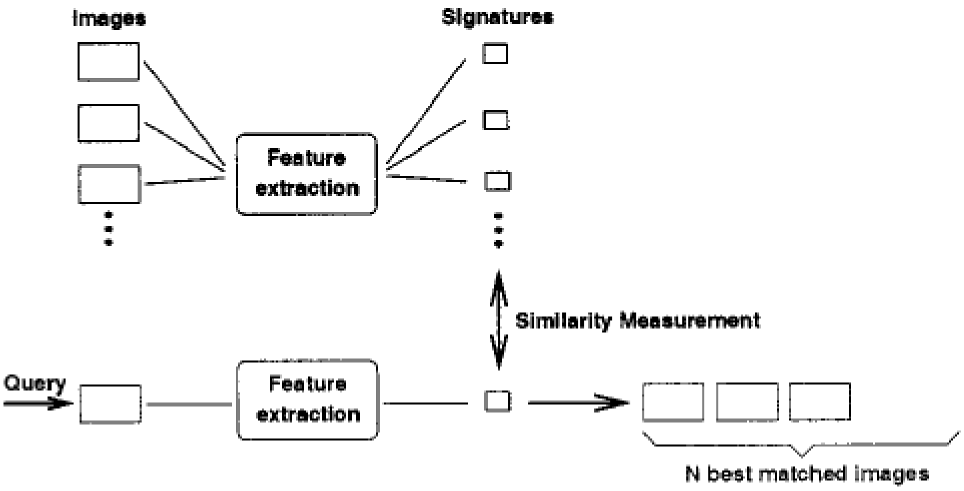


Fig. 1. Image retrieval system architecture.

B. Feature Extraction

Image analysis involves investigation of the image data for a specific application. Normally, the raw data of a set of images is analyzed to gain insight into what is happening with the images and how they can be used to extract desired information. In image processing and pattern recognition, feature extraction is an important step, which is a special form of dimensionality reduction. When the input data is too large to be processed and suspected to be redundant then the data is transformed into a reduced set of feature representations. The process of transforming the input data into a set of features is called feature extraction. Features often contain information relative to colour, shape, texture or context.

Guiying Li (2012) defined texture is a repeated pattern of information or arrangement of the structure with regular intervals. In a general sense, texture refers to surface characteristics and appearance of an object given by the size, shape, density, arrangement, proportion of its elementary parts. A basic stage to collect such features through texture analysis process is called as texture feature extraction. Due to the signification of texture information, texture feature extraction is a key function in various image processing applications like remote sensing, medical imaging and content based image retrieval. There are four major application domains related to texture analysis namely texture classification, segmentation, synthesis and shape from texture. Texture classification produces a classified output of the input image where each texture region is identified with the texture class it belongs. Texture segmentation makes a partition of an image into a set of disjoint regions based on texture properties, so that each region is homogeneous with respect to certain texture characteristics. Texture synthesis is a common technique to create large textures from usually small texture samples, for the use of texture mapping in surface or scene rendering applications.

GLCM which is introduced firstly by Haralick is one of the oldest and prominent statistical textual feature extraction method applied in many fields for texture analysis. GLCM is the matrix that holds the distribution of co-occurring intensity patterns at a given offset over a given image. Second-order statistical (Haralick) features are extracted to analyze the texture of the image which can subsequently be used for classification task. GLCM incorporates the spatial relationships of intensity values with each other as well as their occurrence quantities. Let f is an image whose intensity values vary in the interval [0, L-Ij. Each element of GLCM indicates the number of times that the pixel pair (Zi, Zj) occurred in f with orientation Q. The orientation represented with Q eventually represents a displacement vector d=(dx,dy Idx=dy=dgJ where dg is the number of gaps between the pixels of interest. For the situation of adjacency dg=O. Orientation can also be represented with two parameters as the distance d that the intensities Zi, Zj apart from each other with angle a [19-20]. d can take values between 0 and L-2 theoretically. Orientation of the pixel pattern can be at four different directions as 0°, 45°, 90° and 135° that a can take. That is, each image can have four different GLCMs for each angle (0°, 45°, 90° and 135°) for a specific d. The size of a GLCM depends on the discrete intensity values in the image. If the intensity values of the image vary in the interval [0, L-I} then the size of the GLCM becomes (L-1) x(L-I). An example of configuration of the four GLCM matrixes (GLCM 0°, GLCM 45°, GLCM 90°, GLCM 135°) of an image for d=O is demonstrated in Fig. 2.

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 4, Issue 6 , June 2017

0	2	3	0
1	0	2	3
0	1	3	2
2	0	1	3

(a)

#	0	1	2	3
0	0	2	3	0
1	1	0	0	2
2	1	0	0	2
3	1	0	1	0

(b)

#	0	1	2	3
0	1	0	0	1
1	0	0	3	0
2	1	1	0	0
3	0	0	0	1

(c)

#	0	1	2	3
0	0	2	1	0
1	2	0	0	1
2	1	0	0	2
3	1	0	2	0

(d)

#	0	1	2	3
0	2	0	0	0
1	0	2	0	0
2	0	0	2	0
3	1	0	0	2

(e)

Fig. 2. GLCM construction based on a (a) test image along four possible directions (b) 0° (c) 45° (d) 90° and (e) 135° with a distance d =0.

The flowchart of the analysis is shown in Fig. 3. As shown in the corresponding flowchart, DWT of the images is performed for feature extraction which is subsequently followed by Euclidean Distance-based classification.

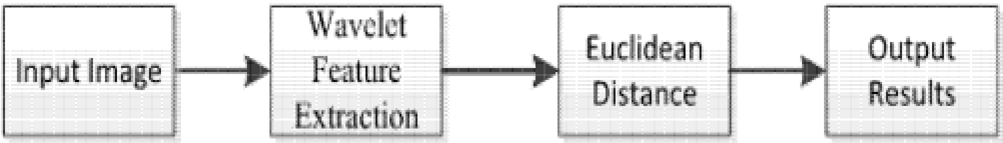


Fig. 3. Flowchart of the first analysis method.

C. Similarity Measurement

KNN is an *non parametric lazy learning* algorithm. That is a pretty concise statement. When you say a technique is non parametric , it means that it does not make any assumptions on the underlying data distribution. This is pretty useful , as in the real world , most of the practical data does not obey the typical theoretical assumptions made (eg gaussian mixtures, linearly separable etc) . Non parametric algorithms like KNN come to the rescue here. It is also a lazy algorithm. What this means is that it does not use the training data points to do any *generalization*. In other words, there is *no explicit training phase* or it is very minimal. This means the training phase is pretty fast . Lack of generalization means that KNN keeps all the training data. More exactly, all the training data is needed during the testing phase. (Well this is an exaggeration, but not far from truth). This is in contrast to other techniques like SVM where you can discard all non support vectors without any problem. Most of the lazy algorithms – especially KNN – makes decision based on the entire training data set (in the best case a subset of them).

KNN assumes that the data is in a *feature space*. More exactly, the data points are in a metric space. The data can be scalars or possibly even multidimensional vectors. Since the points are in feature space, they have a notion of distance. This need not necessarily be Euclidean distance although it is the one commonly used. Each of the training data consists of a set of vectors and class label associated with each vector. In the simplest case , it will be either + or – (for positive or negative classes). But KNN , can work equally well with arbitrary number of classes. A single number "k" . This number decides how many neighbors (where neighbors is defined based on the distance metric) influence the classification. This is usually a odd number if the number of classes is 2. If k=1 , then the algorithm is simply called the nearest neighbor algorithm.

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 4, Issue 6 , June 2017

Case 1 : k = 1 or Nearest Neighbor Rule

This is the simplest scenario. Let x be the point to be labeled . Find the point closest to x . Let it be y. Now nearest neighbor rule asks to assign the label of y to x. This seems too simplistic and some times even counter intuitive. If you feel that this procedure will result a huge error , you are right – but there is a catch. This reasoning holds only when the number of data points is not very large. If the number of data points is very large, then there is a very high chance that label of x and y are same. An example might help – Lets say you have a (potentially) biased coin. You toss it for 1 million time and you have got head 900,000 times. Then most likely your next call will be head. We can use a similar argument here. Let me try an informal argument here - Assume all points are in a D dimensional plane . The number of points is reasonably large. This means that the density of the plane at any point is fairly high. In other words , within any subspace there is adequate number of points. Consider a point x in the subspace which also has a lot of neighbors. Now let y be the nearest neighbor. If x and y are sufficiently close, then we can assume that probability that x and y belong to same class is fairly same – Then by decision theory, x and y have the same class.

The book "Pattern Classification" by Duda and Hart has an excellent discussion about this Nearest Neighbor rule. One of their striking results is to obtain a fairly tight error bound to the Nearest Neighbor rule. The bound is

$$P^* \leq P \leq P^*(2 - \frac{c}{c-1}P^*)$$

Where P^* is the Bayes error rate, c is the number of classes and P is the error rate of Nearest Neighbor. The result is indeed very striking (atleast to me) because it says that if the number of points is fairly large then the error rate of Nearest Neighbor is less that twice the Bayes error rate.

Case 2 : k = K or k-Nearest Neighbor Rule

This is a straightforward extension of 1NN. Basically what we do is that we try to find the k nearest neighbor and do a majority voting. Typically k is odd when the number of classes is 2. Lets say k = 5 and there are 3 instances of C1 and 2 instances of C2. In this case , KNN says that new point has to labeled as C1 as it forms the majority. We follow a similar argument when there are multiple classes. One of the straight forward extension is not to give 1 vote to all the neighbors. A very common thing to do is *weighted kNN* where each point has a weight which is typically calculated using its distance. For eg under inverse distance weighting, each point has a weight equal to the inverse of its distance to the point to be classified. This means that neighboring points have a higher vote than the further points.

IV. EXPERIMENTAL RESULTS

In this there are totally 6 different types of satellite images are considered. Among which aero plane, buildings, forest, freeway, harbor, storage tanks are different categories of satellite images. There are two phase are present in the CBIR: training phase and testing phase. The feature extracted for training feature are already calculated as it consumes more time to calculate all the features in this phase so, results are already calculated for those images. Based on those features we are going to calculate the first and second order parameter for the testing features.

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 4, Issue 6 , June 2017

The overall picture of these paper will look like as shown in the below fig 3.

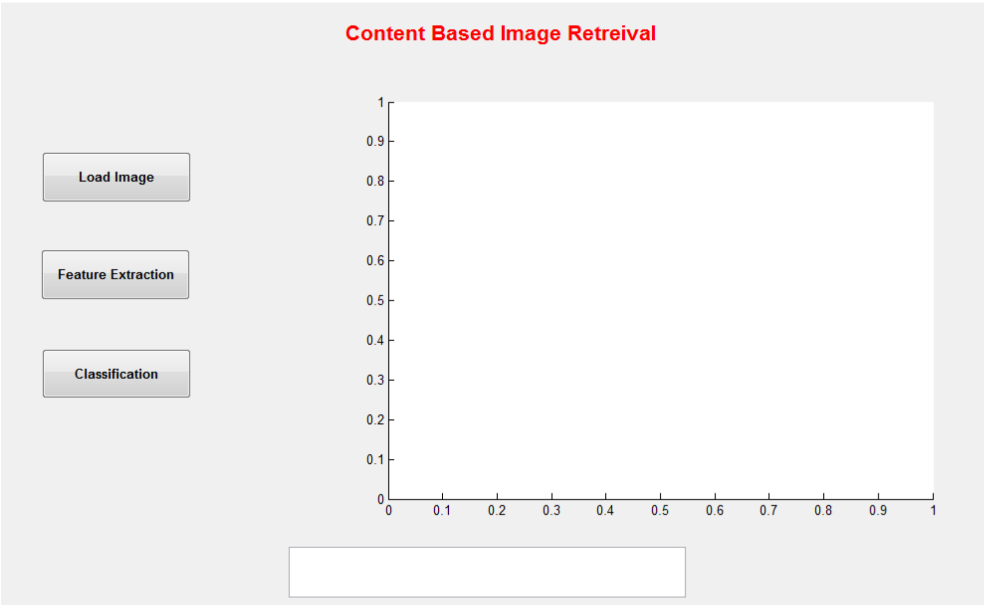
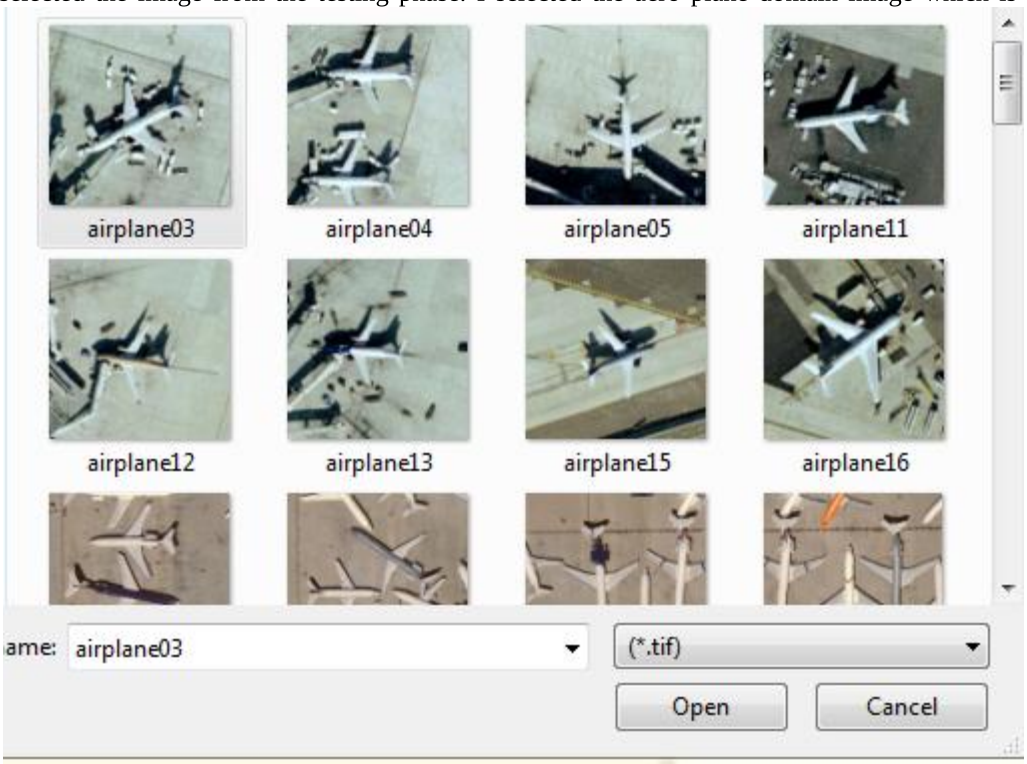


Fig 3. CBIR architecture

First we need to load the images, so click on the load image button it will generate option to choose the image which needs to be extracted. The user wants to select either from testing or training features, here in the below example I have selected the image from the testing phase. I selected the aero plane domain image which is shown in below Fig 4.



ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 4, Issue 6 , June 2017

Since it is supervised classification, already there will be some traning dataset based on the result the query image will be identified and classified to the particular domain. Now its time to reveal the result of the work i.e user clicks classification button then it will generate the result shown in the below Fig 5. The distance between other images present in the database is shown in Fig 6.

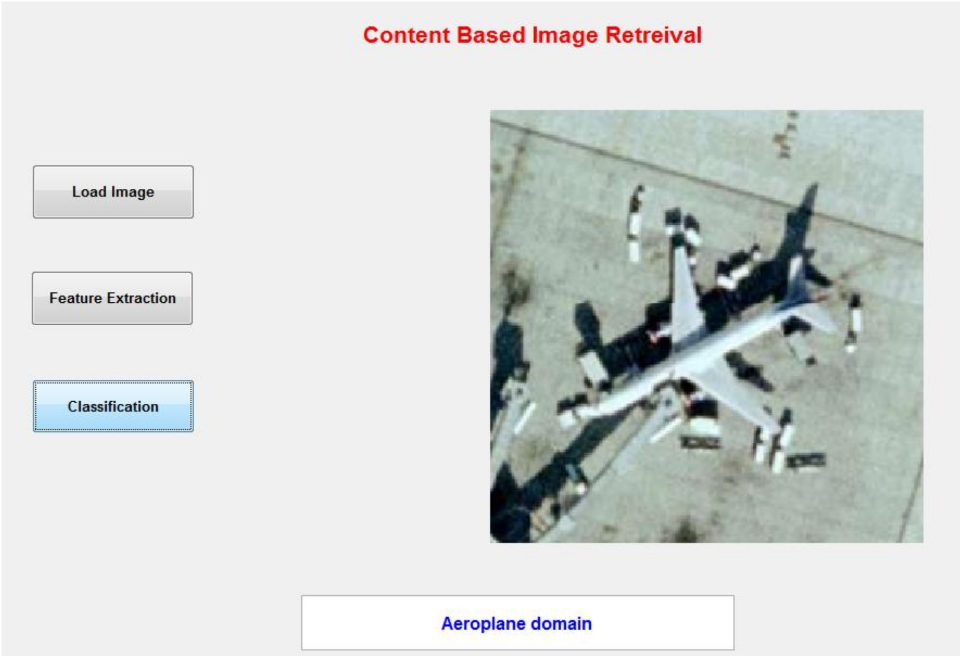


Fig 5. Classification of the image

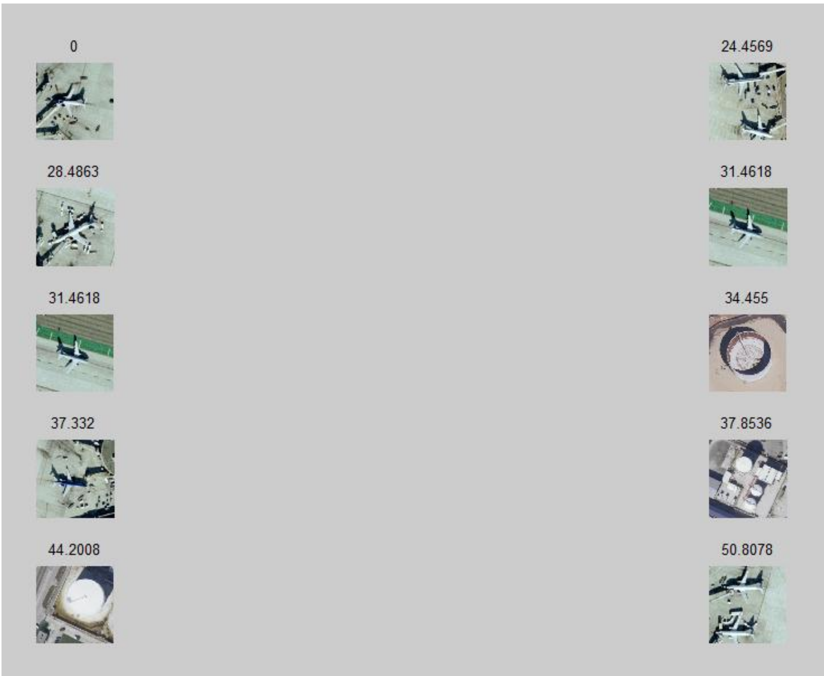


Fig 6 Distance between other images

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 4, Issue 6 , June 2017

V. CONCLUSION

Although wavelet-based probabilistic models gained in popularity in the image processing and computer vision communities in the recent years, they have not yet been extensively explored for applications in VHR optical remote sensing data. Here, we gathered evidence that this approach is also perfectly suitable for the supervised classification of land covers in such VHR optical data. Moreover, we proposed a complete strategy to apply wavelet-based multivariate models in a supervised classification procedure of textures in VHR panchromatic data. Texture features are extracted from the learning database and from the regions of the prepartition by using multivariate models to represent the distribution of observed local spatial dependencies in wavelet subbands in a multiscale and multiorientation framework. A CBIR analysis carried on the learning database is first conducted to identify the most efficient models to retrieve textures in the context of application. A classifier based on a similarity measure or a likelihood criterion is next used to produce classification results with the most relevant models. The applicability of this strategy was tested in two distinct contexts. In both applications, the use of the proposed strategy has enabled to achieve satisfactory classification results with at least one of the tested multivariate models displaying higher classification accuracies than the standard texture analysis using GLCM descriptors. These results confirm the pertinence of using multivariate modeling in the wavelet domain to capture potentially complex texture patterns and improve the classification of VHR optical remote sensing data. Moreover, the classification is almost straightforward with only a few parameters to be set.

REFERENCES

[1]. B. Béguet, S. Boukir, D. Guyon, and N. Chehata, “Modelling-based feature selection for classification of forest structure using very high resolution multispectral imagery,” in Proc. IEEE Int. Conf. SMC, 2013, pp. 4294–4299.

[2]. N. E. Lasmar and Y. Berthoumieu, “Gaussian copula multivariate modeling for texture image retrieval using wavelet transforms,” IEEE Trans. Image Process., vol. 23, no. 5, pp. 2246–2261, May 2014.

[3]. I. Champion, C. Germain, J. P. Da Costa, A. Alborini, and P. Dubois-Fernandez, “Retrieval of forest stand age from SAR image texture for varying distance and orientation values of the gray level co-occurrence matrix,” IEEE Geosci. Remote Sens. Lett., vol. 11, no. 1, pp. 5–9, Jan. 2014.

[4]. L. Gueguen, “Classifying compound structures in satellite images: A compressed representation for fast queries,” IEEE Trans. Geosci. Remote Sens., vol. 53, no. 4, pp. 1803–1818, Apr. 2015.

[5]. L. Fang, S. Li, X. Kang, and J. A. Benediktsson, “Spectral–spatial classification of hyperspectral images with a superpixel-based discriminative sparse model,” IEEE Trans. Geosci. Remote Sens., vol. 53, no. 8, pp. 4186–4201, Aug. 2015.