A Genial Method of Color Features Extraction in Remote Sensing Images using Multi Kernel Principal Component Analysis with Multi-Level 2-D Wavelet Construction

D. Napoleon[1] K. Ragul[2] U. Lakshmi Priya[3]
[1]Assistant Professor [2,3]Research Scholar
[1,2,3]Department of Computer Science & Engineering
[1,2,3]School of Computer Science and Engineering, Bharathiar University, Coimbatore, India

Abstract— One of the most significant requirements in image retrieval, classification, clustering and etc. is extracting efficient features from an image. The color feature is one of the visual features. In this proposed work, various features in a remote sensing image can be distinguished based on their color. The features are extracted as object and distinguished with color. Initially, Gaussian noise is added to image and multi-level 2-D wavelet construction is applied to get denoised image. Next, proposed Multi Kernel Principal Component Analysis preserves local and global structure of data sets and also handles heterogeneous characteristics in an image. Finally, Fuzzy C-Means clustering algorithm partitions the datasets into clusters so that data in each cluster shares some common characteristic which is integrated with color conversion method to extract feature based on color present in the image. The performance of this proposed work is measured through various performance metrics to analysis best result for feature extraction of remote sensing images.
Key words: Image segmentation, Multi kernel principal component Analysis, Fuzzy C-Means clustering algorithm, Remote sensing Image

I. INTRODUCTION

The Computer Industry has seen a large growth in technology such as access, storage and processing fields. This combined with the fact that there are a lot of data to be processed has paved the way for analyzing data to derive potentially useful information. One of the issues is the effective identification of features in the images and the other one is extracting them. The features can be term in different perspective such as color features plays a vital role in the Remote sensing Image Classification where Shape, Texture and pattern are some of the other feature which can be extracted from an image. This proposed work discusses the methods used for extracting features effectively and to evaluate the efficiency of these features. Initially noise is added to remote sensing image to remove unwanted signal present in it. Image Denoising is an important task in image processing, use of multi-level 2D wavelet to improve the quality of an image and to reduce noise level. The specific characteristics feature in an image is enhanced through image enhancement techniques for visual interpretation.

Principal Component Analysis is to standardize the data in image. Real-world data sets usually exhibit relationships among their variables and also allow linear dimensionality reduction [11]. Proposed Multiple Kernel Learning methods aim to construct a kernel model where the kernel is a linear combination of fixed base kernels. Its shows that using multiple kernels is more useful when combining kernels in a nonlinear or data-dependent way seems more efficient than linear combination in fusing information provided by simple linear kernels. Fuzzy C-Means clustering algorithm finds the similar points without actually knowing the labels and attributes may be similar to others as well as those which make it dissimilar from other can be distinguished through color. Finally, various parameter metrics such as Peak signal to noise ratio, Structural Similarity Measure and Kappa Analysis is used to analysis the performance of proposed work and its shows efficiency of feature extraction based on color.

II. METHODOLOGY

The following system architecture is represented as our proposed work is given below in the Figure 1.

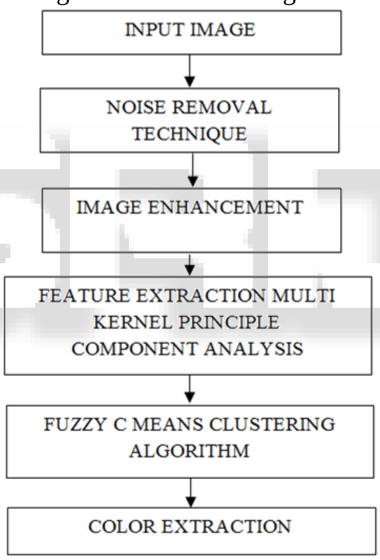


Fig 1: System architecture

III. NOISE REMOVAL

The main aim of image restoration is to remove noise from the original image. Gaussian noise is added in which each pixel in the image will be changed from its original value by a small amount. The multi-level 2D wavelet decomposition will select pixels form from image and decomposes into blocks, and multilevel 2D wavelet decomposition features such as horizontal, vertical, diagonal bands of wavelet transform details [4].

IV. IMAGE ENHANCEMENT

Image enhancement is to improve the quality of the image. It is quite easy, for example, to make an image lighter or darker, or to increase or decrease contrast [12].

V. MULTI KERNEL PRINCIPAL COMPONENT ANALYSIS

Multiple Kernel Learning methods aim to construct a kernel model where the kernel is a linear combination of fixed base kernels. Learning the weighting coefficients for each base kernel, rather than optimizing the kernel parameters of a single kernel.

A. *Constructing the Kernel Matrix:*

Nonlinear transformation Ø(x) from the original D feature space to an M feature space, are usually denoted as M > D. Then each data point x_n is projected to a point $Ø(x_n)$. Traditional PCA can perform in the new feature space, but extremely costly. Thus kernel methods are used to simplify the computation.

By assuming that the projected new features have zero mean:

$$\sum_n Ø(x_n). = 0, \quad (1)$$

The covariance matrix of the projected features is M *M, calculated by

$$C = 1/N\sum^N_{n=1} Ø(x_n). Ø(x_n)T \quad (2)$$

And its eigenvalues and eigenvectors are

$$Cv_i = \lambda_i v_i; \quad (3)$$

Where i = 1.2…M. From (2) and (3), we have

$$1/N\sum^N_{n=1} Ø(x_n)\{ Ø(x_n)^T v_i\} = \lambda_i v_i \quad (4)$$

Which can be written as

$$v_i = \sum^N{}_n{=}1\ a_{in}\text{-}\ Ø(x_n). \quad (5)$$

Now by substituting $v_i$in (4) with (5), we have

$$(x_n)Ø(x_n)^T 1/N\sum^N_{m=1} a_{im} Ø(x_m) = \lambda_i 1/N\sum^N_{n=1} a_{in} Ø(x_n) \quad (6)$$

By defining the kernel function

$$k(x_n,x_m) = Ø\ (x_n)^T\ Ø\ (x_m); \quad (7)$$

And multiplying both sides of Equation (6) by $Ø(x_l)^T$, we have

$$1/N\sum^N_{n=1} k(x_l, x_n) \sum^N_{m=1} a_{im}\ k(x_l, x_n) = \lambda_i 1/N\sum^N_{n=1} k(x_l, x_n) \quad (8)$$

Or the matrix notation

$$K^2 a_i = \lambda_i\ NKa_i \quad (9)$$

Where

$$K_{n,m} = k(x_n, x_m) \quad (10)$$

and a_i is the N-dimensional column vector of ani. a_i can be solved by

$$K\ a_i = \lambda_i\ Na_i; \quad (11)$$

And the resultant kernel principal components can be calculated using

$$y_i(x) = Ø\ (x_n)^T\ v_i = \sum^N_{n=1} a_m\ k(x,x_n), \quad (12)$$

The power of kernel methods is that we do not have to compute Ø (x_n) explicitly. Kernel matrix can be constructed directly from the training data set (x_n) .

VI. FUZZY C-MEANS CLUSTERING ALGORITHM

A cluster is a collection of data objects that are similar to one another within the same cluster and are dissimilar to the objects in other clusters. By clustering, one can identify dense and sparse regions and therefore, discover overall distribution patterns and interesting correlations among data attributes [10]. Fuzzy C-Means (FCM) is a method of clustering which allows one portion of data to two or more clusters. It is depend on minimization of the following objective function

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^m\ ||x_i - c_j||^2\ ,1 \le m < \infty$$

In this *m* is any real number < than 1, *uij* is the degree of membership of *xi* in the cluster *j*, *xi* is the *i*th of d-dimensional measured data, *cj* is the d-dimension center of the cluster, and ||*|| is any norm expressing the similarity between any measured data and the center.

- Select a number of clusters.
- To each point coefficients assign randomly for being in the clusters.
- Repeat until the algorithm has meet
- Assign the centroid for each cluster, using the formula above.
- For each point, compute its coefficients of being in the clusters.

The algorithm minimizes intra-cluster variance and local minimum, and the results depend on the initial choice of weights [7].

VII. FEATURE EXTRACTION

A. *Color:*

Color is an important visual factor. Based on the different applications different color spaces are available such as RGB, LAB, LUV, HSV (HSL), YCrCb and the hue-min-max difference (HMMD).

B. *Converting From RGB to HSV:*

Color vision can be processed using RGB color space or HSV color space. RGB color space describes colors in amount of red, green, and blue present. HSV color space explains colors of Hue, Saturation, and Value. Color display of remote sensing data is of significance for effective visual interpretation. So color conversion method is integrated with the result of Fuzzy C-Means clustering algorithm to extract feature based on color present in the image. Now the result shows that this proposed work is an effective method for extracting features with respect to color even in high resolute remote sensing images.

VIII. PERFORMANCE METRICS

Image Quality Measurement (IQM) is imperative in the progress of image processing algorithms such as enhancement; deblurring, Denoising etc. It is used to evaluate the performances in terms of quality of processed image.

A. *Peak Signal to Noise Ratio (PSNR):*

PSNR is the evaluation standard of the reconstructed image quality, and is important feature. The PSNR is defined as:

$$\text{PSNR} = 10\log\left(\frac{255}{\sqrt{\text{MSE}}}\right)$$

B. *Structural Similarity Measure:*

The structure similarity index is a method for measuring the similarity between two images. The SIM index is a full reference metric, in other words, measuring of image quality based on an initial uncompressed or distortion-free image as reference. The measure between two images x and y of common size N*N is:

$$\text{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

C. *Execution Time:*

The execution time or CPU time of a given task is defined as the time spent by the system executing that task.

D. *Kappa Analysis:*

Kappa can be used as a measure of agreement between model predictions and reality or to determine if the values contained in an error matrix represent a result significantly better than random.

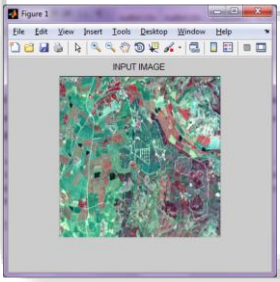
$$\kappa = \frac{N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r}(x_{i+} \times x_{+i})}{N^2 - \sum_{i=1}^{r}(x_{i+} \times x_{+i})}$$

Where N is the total number of sites in the matrix, r is the number of rows in the matrix, x_{+ii} is the number in row i and column i, x_{+i} is the total for row i, and x_{i+} is the total for column [13].

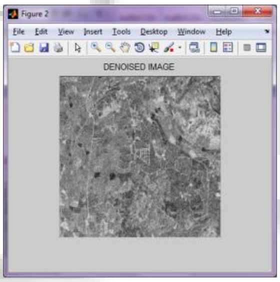
IX. RESULT

In this proposed work the testing is done for segmentation and color feature extraction of Landsat ETM Dataset, Quick Bird Dataset, MODIS Dataset, Aster dataset and SPOT 5 dataset with proposed Multi Kernel PCA. The images are tested using algorithms to calculate the PSNR value, Execution time, Structural Similarity Measures, and Kappa analysis. The experimental result of MODIS and QuickBird datasets are shown below

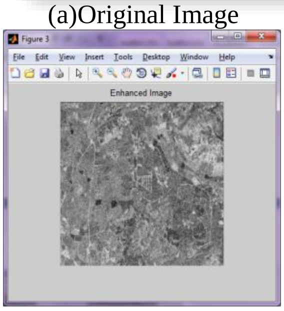
A. *MODIS Dataset:*



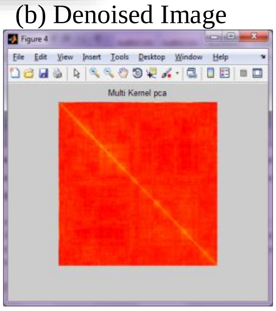
(a)Original Image



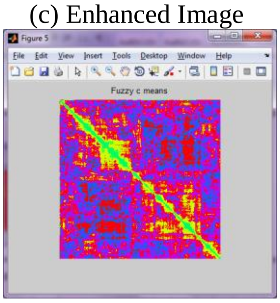
(b) Denoised Image



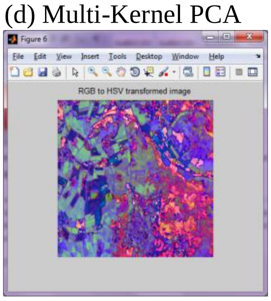
(c) Enhanced Image



(d) Multi-Kernel PCA



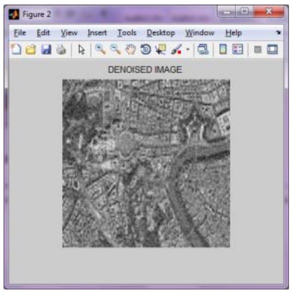
(e) Fuzzy C-Means



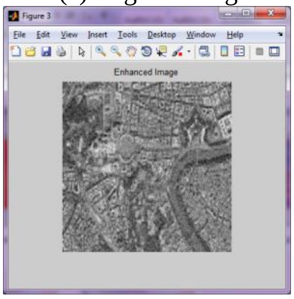
(f) Color Extraction

B. *Quickbird Dataset:*

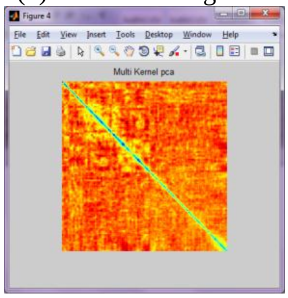
(a)Original Image



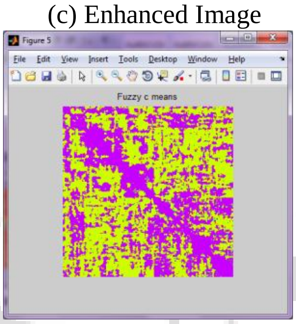
(b) Denoised Image



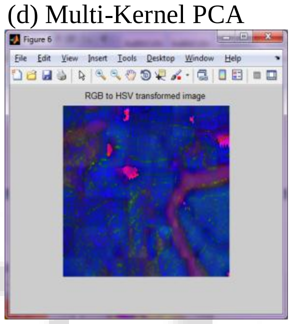
(c) Enhanced Image



(d) Multi-Kernel PCA



(e) Fuzzy C-Means



(f) Color Extraction

PERFORMANCE METRICS	MODIS	QUICK BIRD
PSNR	55.547	62.84
Execution Time	9.317	11.524
Structural Similarity Measure	6.014	1.402
Kappa Analysis	124.031	2.106

Table 1: Performance metrics for MODIS and QUICKBIRDS Datasets

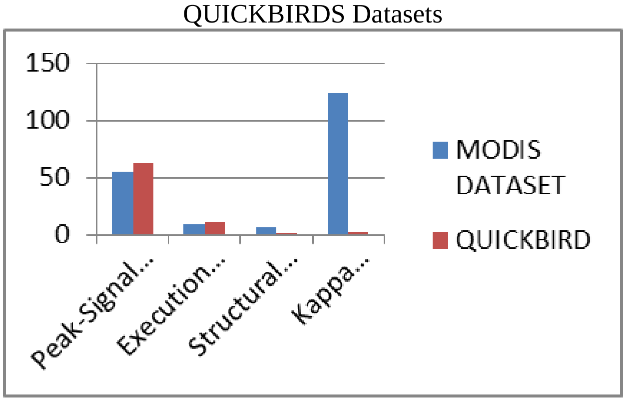


Fig. 2: Graphical representation of MODIS and QUICKBIRD datasets values

X. CONCLUSION

Extracting the Features from the Remote Sensing Image been a challenging task for the researcher from few decades. In which the color is one of the considered feature to extract from an image basically. On such perspective, in this research work, proposed Multi Kernel principal Component Analysis has been developed for the corresponding remote sensing image feature extraction and Fuzzy C-Means clustering algorithm for image segmentation scheme has been presented. By using these two algorithms the PSNR, SSIM, Kappa and Execution time are deliberate for different images color features.

REFERENCES

[1] Licciardi "Linear Versus Nonlinear PCA for the Classification of Hyper spectral Data Based on the Extended Morphological Profiles" IEEE Geosciences and Remote Sensing Letters, Volume: 9, Issue: 3, pp: 447 - 451, May 2012.
[2] Isa, "Sleep Apnea Detection from ECG Signal: Analysis on Optimal Features, Principal Components, and Nonlinearity" IEEE Conference Publications, pp: 1 - 4, May 2011.
[3] Wang, "Machine condition monitoring by nonlinear feature fusion based on kernel principal component analysis with genetic algorithm", IEEE Third International Conference, Volume: 2, pp: 665 – 670, 2007.
[4] Burhan Ergen , "Signal and Image Denoising Using Wavelet Transform".
[5] Wen Ge "Multi-kernel PCA based high-dimensional images feature reduction".
[6] Nock, R. and Nielsen, F. (2006) "On Weighting Clustering", IEEE Trans. on Pattern Analysis and Machine Intelligence, 28 (8), 1–13
[7] Ahmed, Mohamed N.; Yamany, Sameh M.; Mohamed, Nevin; Farag, Aly A.; Moriarty, Thomas (2002). "A Modified Fuzzy C-Means Algorithm for Bias Field Estimation and Segmentation of MRI Data". IEEE Transactions on Medical Imaging 21 (3): 193–199. doi:10.1109/42.996338. PMID 11989844.
[8] Ramanpreet Kaur Sukhpreet Kaur, "Enhancement of Colored Images in Digital Image Processing" International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 7, July 2014 ISSN: 2277 128X
[9] Chu XiaoLi, ,Zhu Ying ,Shi JunTao , Song JiQing, "Method of image segmentation based on Fuzzy C-Means Clustering Algorithm and Artificial Fish Swarm Algorithm".
[10] Amiya Halder, Soumajit Pramanik, Arindam Kar, "Dynamic Image Segmentation using Fuzzy C-Means based Genetic Algorithm", International Journal of Computer Applications (0975 – 8887) Volume 28– No.6, August 2011
[11] Ada, Rajneet Kaur, "Feature Extraction and Principal Component Analysis for Lung Cancer Detection in CT scan Images" International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 3, March 2013 ISSN: 2277 128X.
[12] Muna Abdul Hussain Radhi, Ban Sabah, Aurass Musa Omran Al-Hsniue , "Enhancement Of The Captured Images Under Different Lighting Conditions Using Histogram Equalization Method".
[13] Adubofour Frimpong, "Application of Remote Sensing and GIS for Forest Cover Change Detection".
[14] Yogesh Jain, Amit kumar Nandanwar "Comparative Analysis of K-Means and Fuzzy C-Means Algorithms for team assignment" International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 6, June 2014 ISSN: 2277 128X.