Aquaculture Feature Extraction From Satellite Image Using Independent Component Analysis

JongGyu Han, KwangHoon Chi, YeonKwang Yeon

Korea Institute of Geosciences & Mineral Resources,
305-350, 30 Gajung-dong Yusong-ku Daejeon, Republic of Korea
{jghan, khchi, ykyeon}@kigam.re.kr

Abstract. In multi-dimensional image, ICA-based feature extraction algorithm, which is proposed in this paper, is for the purpose of detecting target feature about pixel assumed as a linear mixed spectrum sphere, which is consisted of each different type of material object (target feature and background feature) in spectrum sphere of reflectance of each pixel. Landsat ETM+ satellite image is consisted of multi-dimensional data structure and, there is target feature, which is purposed to extract and various background image is mixed. In this paper, in order to eliminate background features (tidal flat, seawater and etc) around target feature (aquaculture) effectively, pixel spectrum sphere of target feature is projected onto the orthogonal spectrum sphere of background feature. The rest amount of spectrum sphere of target feature in the pixel can be presumed to remove spectrum sphere of background feature. In order to make sure the excellence of feature extraction method based on ICA, which is proposed in this paper, aquaculture feature extraction from Landsat ETM+ satellite image is applied. Also, In the side of feature extraction accuracy and the noise level, which is still remaining not to remove after feature extraction, we have conducted a comparing test with traditionally most popular method, maximum-likelihood. As a consequence, the proposed method from this paper can effectively eliminate background features around mixed spectrum sphere to extract target feature. So, we found that it had excellent detection efficiency.

1. Introduction

The image, which is obtained to take a picture for the surface of earth from Landsat ETM+ satellite is consisted of multi-dimensional data structure of multiplex spectrum sphere. A lot of data for the surface of earth are recorded such as an image of multi-dimensional data structure. In order to extract target feature from multiplex spectrum satellite image, there is an image processing method to change the peculiarity of data with the reflection of a certain axis for image data. At this time, the number of axis of reflected are needed to match the number of original data level so that the data level after reflection can be the same with the data before reflection or small to have the effect of compress.

There are methods to decide the axis to get the result as a meaningful outcome value when multi-dimensional data are reflected[1, 2]. ICA (Independent Component Analysis) method, which is expanded method of PCA (Principal Component Analy-

sis) is a statistical technique represented as a multi-dimensional vector of independence component with combination of linear. ICA can eliminate not only mutual relations of data but also higher level of mutual relations. In a consequence, it is the method to transform to independent between dimensions [3]. ICA is primarily applied to analyze data and extract feature. In this regards, we can use BSS (Blind Source Separation) method for original data and tracking down to find original data with mixed data without knowing the data to be mixed processing [4].

In this paper, ICA method is applied to extract wished feature data from multi-dimensional data structure. In this regard, Lee et al [6, 7], Hyvarinen [5] develops ICA-based feature extraction method based on the result of previous research works. An experiment about aquaculture feature extraction from Landsat ETM+ satellite image has been conducted to verify the validity of the result about actual application. In chapter 2, the materials and methods used in this paper are described and, in chapter 3, test result will be reviewed together with them. And, finally, the conclusion will be described in chapter 4.

2. Materials and Methods

This paper is described for algorithm of object detection, which can be classified as object and background from each pixels of multi spectral image consisted more than 2 object peculiarities. Multi spectrum image, which is used in this paper, is obtained from Landsat TM sensor with the following spectrum sphere: 450 ㎚~520 ㎚(band 1), 520 ㎚~ 600 ㎚(band 2), 630 ㎚~690 ㎚(band 3), 760 ㎚~ 900 ㎚(band 4), 1550 ㎚~1750 ㎚(band 5), 1040 ㎚~ 1250 ㎚(band 6), 2080 ㎚~2350 ㎚(band 7).

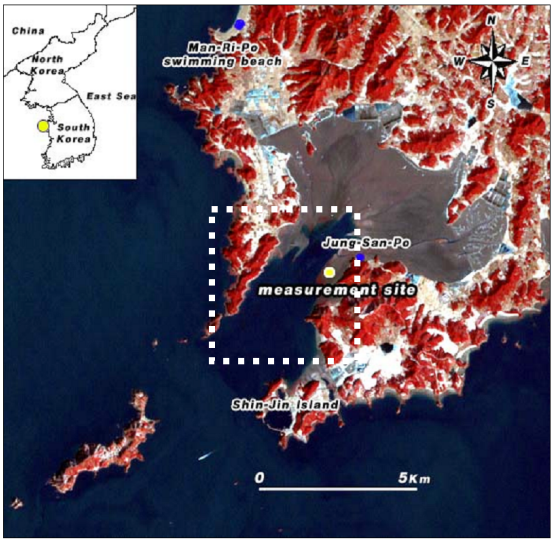

Fig. 1 Experiment data (Landsat ETM+) and research area

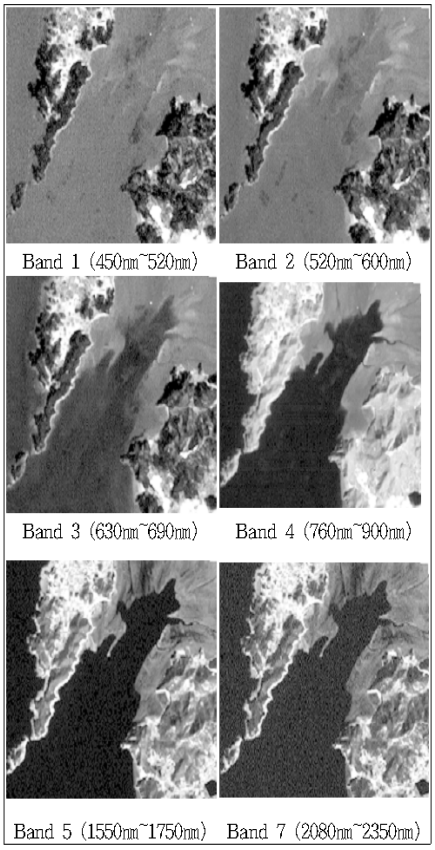

Fig. 2 Band image of different spectrum of LANDSAT ETM+ Satellite Image

In here, band 6, which is spectrum sphere of thermal infrared rays, is excluded in this paper. Fig. 1 shows you the satellite image of Landsat ETM+ used for experiment and, as a place of aquaculture; its location is Jungsanpo Taeahn-gun, Chungnam province. Fig. 2 shows you image data of each different spectrum band photographed from LANDSAT ETM+ Satellite Image around sea area of Jungsanpo Taean-gun, Chungname province when ebb and flow is ebb tide on Feb 16, 2003.

Signal energy is controlled by reflected sunbeams radiant energy rather than discharged sunbeams by observed substance in spectrum sphere. Fig. 3 shows you the curved line of spectral reflectance, which is drawn with Landsat ETM+ spectrum sphere to measure spectral reflectance for aquaculture and around substances.

In algorithm for object detection of partial pixel based on mixed models of linear spectrum, reflected spectrum sphere of each pixel is presumed as linear mixed spectrum sphere from each different kinds of substance in the surface of pixel. If pixel spectrum sphere is controlled by object spectrum sphere, this pixel is the sign as an object pixel. If the contribution extent of object spectrum sphere can neglect in pixel

spectrum sphere, that pixel can be classified as a background. As a consequence, what pixel spectrum sphere is disassembled as an organization spectrum sphere is needed to decide what kinds of substances are occupied extensively as a target or background.

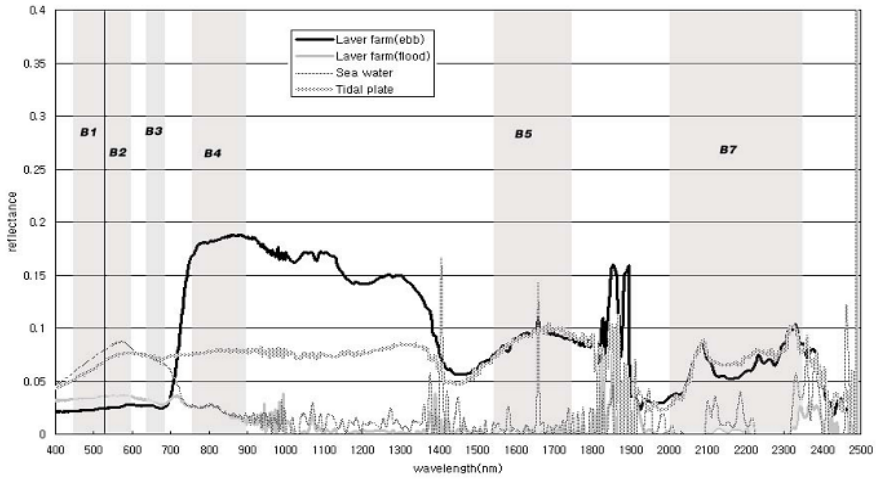

Fig. 3 Curved line of spectral reflection of Landsat ETM+ bands

3. Feature Extraction Procedure Using ICA

ICA is a technique to isolate mutual independent signals statistically from linear mixed signals and, it is applied to not only signal field but also image field vigorously. ICA method seeks for direction to reach independence elements in data.

There are lots of performance algorithms to exist for ICA such as Entropy minimized method, common data minimized method, Gaussian measure most suitable method and, maximum likelihood method. Lee[6, 7] is applied as a mixed model of independent element analyzer by maximum likelihood method. The weakness of this method is not extended easily to find a partial space. That weakness is not applied for fields like image coding. However, in an image segmentation and field of image classification, image is described as much smaller dimension than the numbers of pixels.

Fig. 4 is a sequence chart of data processing to extract target feature based on ICA method, which is indicated in this paper.

At first, in a data preparation stage as a first step, input feature with N unit from Landsat ETM+ Satellite Image data and creation of input data collection with N+1 dimension from output class. And, they go through the process of formalizing for each input features. In a second performing stage of ICA, ICA is performed to new data collections, which have been made in the data preparation step. And, the result is saved as an extra weight matrix W in dimension of (N+1) x (N+1).

In a third step, absolute average of each independent line vector, which has N+1 of extra weight matrix W, is searched. And, among the extra weight line elements, elem-

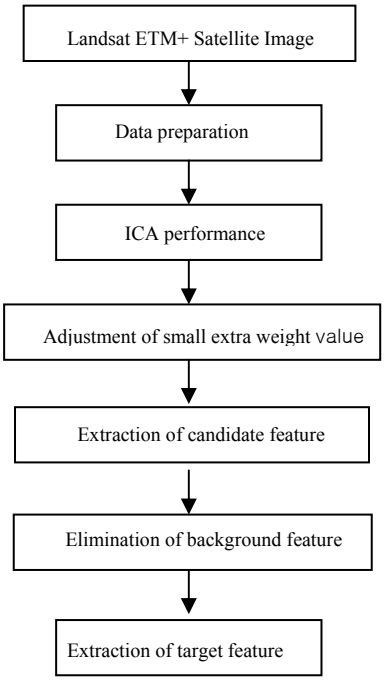

Fig. 4 The sequence chart for data processing of feature extraction

-ents value, which has an extra weight value less then absolute average, is made as zero. Fourth step is the stage of extracting for candidate feature. After all extra weight value line vectors are reflected in an original peculiar space, candidate feature of new N+1 unit is extracted to multiply original input data by new extra weight matrix of (N+1) x N dimension. Extracted candidate features have been mixed with target and background features. Last stage is the step of deleting background feature from extracted candidate features. Candidate feature set F is made in this stage and, if extra weight value, which is coped with each candidate features, is zero, it is considered as background features and, it is removed. As a consequence, only final target features, which are removed background features from candidate features will be extracted.

3. Test Results and Considerations

In this chapter, in order to extract aquaculture feature from Landsat ETM+ multi spectrum image data, image of 200x161 size is applied and, it is formalized to keep gaucian distribution peculiarity of N (m, σ^2). Also, in order to delete noise effectively

from original data, test has been conducted to find how features are extracted with keeping the average as regularly (m=0) and changing (σ^2: 0~1) the dispersion.

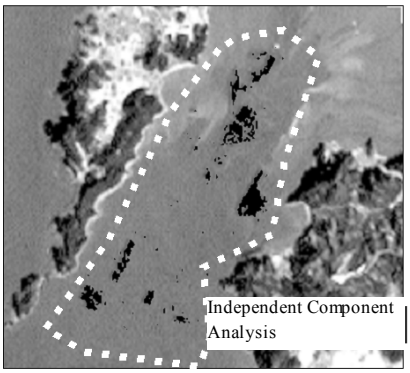

Fig. 5 Aquaculture feature through ICA-based feature extraction method (Black color in white dotted line)

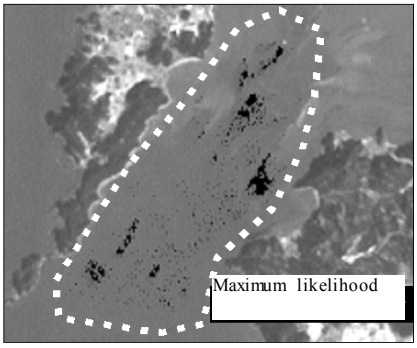

Fig. 6 Aquaculture feature extracted by maximum likelihood distribution method (Black color in white dotted line).

Fig. 5 shows you aquaculture feature extracted to apply feature extraction method based on ICA, which is proposed in this paper. Indicated section with black color in white dotted line is shown for aquaculture feature.

Fig. 6 shows you aquaculture feature extracted through distribution method of maximum likelihood, which is the most popular used. The test has been conducted by the side of noise, which is not eliminated after feature extraction and accuracy of extracted feature.

Noise, which is not eliminated after feature extraction, from ICA-based feature extraction method is shown 10% lower than maximum likelihood method and, also, in the side of accuracy, ICA-based feature extraction method is shown 6% higher than maximum likelihood method. Test result of Fig. 5 and 6 are shown well in image and, it displays the superiority of ICA-based feature extraction method.

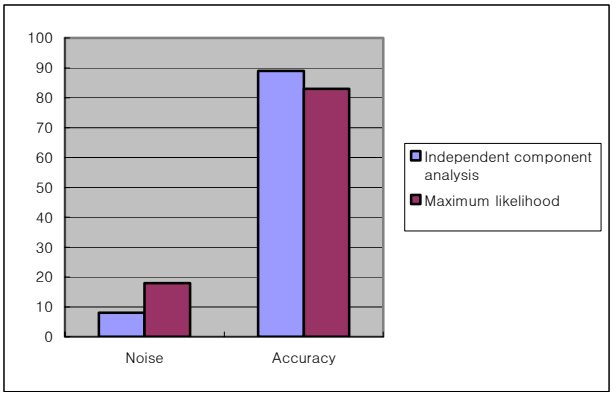

Fig. 7 Accuracy comparison for feature extraction and noise, which is not eliminated from extracted feature after test.

5. Conclusion and Discussion

ICA-based feature extraction algorithm, which is proposed in this paper, is purposed to detect target feature about pixel supposed as a linear mixed spectrum sphere, which is consisted of each different substance types (target feature and background feature) that has different reflection spectrum sphere of each pixels.

On Landsat ETM+ Satellite Image, which is consisted of multi dimensional data structure, ICA-based feature extraction method is indicated to eliminate background features (tidal flat, seawater and etc), which is located around target feature (aquaculture) effectively and, in order to confirm the superiority of proposed method, aquaculture feature can be successfully extracted on Landsat ETM+ satellite image.

In the side of noise level, which is not eliminated after feature extraction and accuracy, comparing test with maximum likelihood method, which is the most popular method traditionally, has been conducted.

As a consequence, the proposed method in this paper shows you superior detection performance in extraction of target feature to extract as background feature is eliminated effectively in mixed spectrum sphere around target feature.

References

1. P. J. Huber, Projection Pursuit, *The Annals of Statistics*, vol. 13, no. 2, pp. 435-475, 1985.

2. A. Hyvarinen, New Approximations of Differential Entropy for Independent Component Analysis and Projection Pursuit, *In Advances in Neural Information Processing Systems 10 (NIPS '97)*, pp. 273-279, 2003/02/18 22:55:40, 1998
3. A. Hyvarinen, J. Karhunen, and E. Oja, Independent ComponentAnalysis, *John Wiley & Sons*, 2001.
4. A.J. Bell and T.J. Sejnowski, An Information-Maximization Approach to Blind Separation and Deconvolution, *Neural Computation*, vol. 7, pp. 1129-1159, 1995.
5. A. Hyv¨arinen. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks, 10(3):626-634, 1999.*
6. T.-W. Lee, M. Girolami, and T.J. Sejnowski. Independent component analysis using an extended infomax algorithm for mixed sub-gaussian and super-gaussian sources. *Neural Computation, 11(2):417-441, 1999.*
7. T.-W. Lee, M.S. Lewicki, and T.J. Sejnowski. Unsupervised classification with non-Gaussian mixture models using ICA. In M.S. Kearns, S.A. Solla, and D.A. Cohn, editors, *Advances in Neural Information Processing Systems 11, Cambridge, MA, 1999. NIPS, MIT Press.*