

Geological Feature Extraction Using Independent Component Analysis

Sanjay Kumar Singh

`sanjaykumar.singh@vit.ac.in`

Vellore Institute of Technology University

Research Article

Keywords: Seismic Image Processing, Edge Detection, Independent Component Analysis, Sparse Coding, FastICA Algorithm

Posted Date: August 3rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9926550/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Geological Feature Extraction Using Independent Component Analysis

Sanjay Kumar Singh

Associate Professor, Electronics & Communication Engineering Department,
Vellore Institute of Technology, Katpadi, Vellore, Tamil Nadu - 632006, India.
Email: sanjaykumar.singh@vit.ac.in, Contact: +919897888970

ABSTRACT

This paper presents a resilient and adaptive method for edge detection in seismic images via independent component analysis (ICA). ICA is used on seismic image patches, where the fixed-point FastICA technique is used to learn the basis functions. The majority of these sparse basis functions are used as patterns or template to extract features. Like Gabor wavelet bases, these basis functions are band-limited and confined. The proposed algorithm first transforms a grayscale seismic image into different classes representing edges and background. The edges are extracted by thresholding using sparse components which in turn generates an image into a pattern or feature map. Further, this paper investigates the capabilities of ICA compared with classical methods of edge detection based on seismic images through simulation results.

Keywords: Seismic Image Processing, Edge Detection, Independent Component Analysis, Sparse Coding, FastICA Algorithm.

1. INTRODUCTION

Image processing offers vast opportunities in different areas to process and extract useful information from natural images, for years. Natural images are recognised when both dimensions are spatial. In contrast, a seismic image has one spatial dimension and one temporal dimension. The two-way transit time on the temporal dimension along the offset is represented by a point on a seismic image of a reflecting geological interface at a depth [1]. In order to deduce a description of the earth's subsurface geology, seismic images are now processed similarly to natural images with gray level replacing signal amplitudes due to interest in interactive feature extraction [2-8].

Edge detection is a crucial image processing technique that offers vital visual cues for comprehending and interpreting image elements. The areas where color or intensity drastically changes in one way over a limited number of pixels are called edges. An image size is reduced by the edge detection procedure that show only outlines of the objects which can be used in pattern or object recognition for subsequent image analysis. Edge detection preserves the most important structural features (high frequencies) and filters out less relevant information or redundancies available in the form of image background (low frequencies) which is replaced when it is reconstructed. As a result, fewer pixels or bits are needed to store an image, which lowers storage, bandwidth, complexity, and cost. At the receiving end, the original image is fully recovered for further examination and interpretation [9-12].

Even though numerous edge detectors have been documented in the literature, they frequently have issues including high computing costs and excessive noise sensitivity. Sobel, Prewitt and Kirsch are the most popular classical methods use first directional derivative for edge detection. These first order derivative edge detectors are obviously simple, but highly sensitive to noise and are usually inaccurate. Canny detector is among the widely popular, optimal, multistage and high-performance edge detector in the literature is considered comparatively superior especially in noisy conditions because it uses Gaussian and hysteresis which finds maxima of the first order gradient resulting into thin edges however, slower and computationally intensive. The Laplacian of Gaussian operator, sometimes referred to as zero-crossing edge detectors, uses Laplacian operator, a second order derivative that smoothes the image by convolving with a Gaussian function to lessen the impact of noise before identifying edges in the form of zero crossings. Although the Gaussian edge detectors work better and are comparatively quick, they are computationally complex and produce double or thick edges. [13-16]. These edge detection techniques have all been successful in identifying weak and strong edges to various degrees however, the main drawbacks are lack of computational regularity and imprecise results due to noise. Although numerous widely used and successful edge detection techniques have been documented over the past forty years, still it is highly desirable to realize an optimal edge detection strategy to extract edges in an efficient manner [17-18].

Geological feature extraction is usually a tedious process to locate the sub-surface faults by simple geological investigations or studies of air photographs. Additionally, as the seismic images are huge in size (several terabytes) that limits the storage and transmission frequently. Edge detection alleviates such problems by not only extracting features but also, reducing the

image sizes only to few kilobytes. Thus, this paper's primary contribution is to suggest a revolutionary and efficient edge detection method using ICA, a powerful technique in machine learning which can extract edges with reduction in noise comparatively better than the classical methods.

2. INDEPENDENT COMPONENT ANALYSIS ON IMAGES

ICA is a statistical method that disintegrate multidimensional data into statistically independent parts. A perceptual system is typically exposed to a sequence of small image patches that are sampled from one or more bigger images in image processing applications. The vector $X = [x_1, x_2, \dots, x_N]^T$, which is created by linearly combining N mutually statistically independent basis functions, can be used to describe each image patch. This vector is known as the "observation vector". These basis functions make up a fixed $N \times N$ complete rank matrix's columns known as the "mixing matrix", $A = [a_1, a_2, \dots, a_N]$.

As shown in Fig. (1), a vector $S = [s_1, s_2, \dots, s_N]^T$, also known as the "source vector", determines the weights of these linear combinations for each image. Every element of the source vector has a basis function that corresponds to a root "cause" of the image. The model for linear image synthesis can be provided by

$$X = AS \quad (1)$$

which represents the set of equations in matrix form

$$x_i = \sum_{j=1}^N a_{ij} s_j \quad (2)$$

Where x_i is a pixel in an image that has contributions from a sequence of N image "sources", s_j , which are weighted linearly by a coefficient, a_{ij} , for $1 \leq i, j \leq N$. In such a simple architecture, the perceptual system's primary goal is to modify the images, X , linearly using a matrix of filters $W = [w_1, w_2, \dots, w_N]$ to obtain the resulting vector as,

$$\hat{S} = WX \cong S \quad (3)$$

It is rescaled, statistically independent, and may recover the fundamental causes, S , in a distinct sequence. The “separating matrix” is the straightforward relation $W = A^{-1}$ between the basis functions represented by columns of A and the filters that recover the causes represented by rows of W .

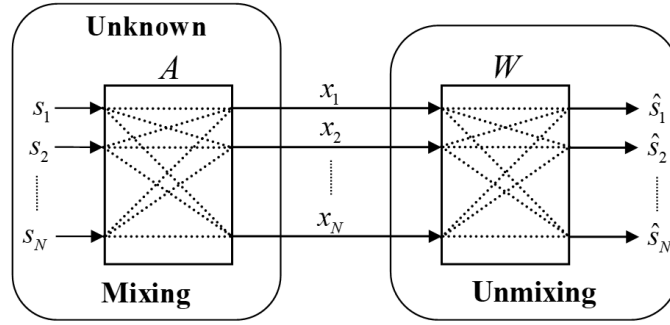


Fig. (1): Model of ICA

From the standpoint of image processing, ICA is considered a fundamental statistical model for natural image statistics. It models images as linear superpositions of basis images with non-Gaussian and independent weight coefficients [19–23]. As illustrated in Fig. (2), several image patches are randomly selected from an image, each of which is described by a linear combination of “basis” images a_{ij} .

$$\text{Image Patch} = s_1 \times \text{Basis Image}_1 + s_2 \times \text{Basis Image}_2 + \dots + s_N \times \text{Basis Image}_N$$

The equation represents the linear image synthesis model. It shows a target image patch on the left, followed by an equals sign, and then a sum of products. Each product consists of a weight coefficient s_i multiplied by a corresponding basis image patch. The basis image patches are shown as small grayscale images with distinct features. The weights s_i are represented by small black squares.

Fig. (2): The Linear Image Synthesis Model

Each component of the sparse code, a type of neural code, is encoded by the intense activity of a comparatively small set of neurons. Decorrelation is the process used to estimate the independent components s_j . It can be said that s_j is independent of the other components if it offers no information about them. The required independent components, s_j , have the fewest Gaussian components under the non-Gaussian distribution assumption. When estimating the independent components or the unmixing matrix, a measure of non-Gaussian behavior is crucial. Kurtosis, a 4th-order statistic with a zero value for the Gaussian distribution, is the traditional indicator of non-Gaussian features represented as,

$$kurt(y) = E\{y^4\} - 3(E\{y^2\})^2 \quad (4)$$

“Sub-Gaussian” refers to random variables with negative kurtosis, whereas “super Gaussian” refers to ones with positive kurtosis. Kurtosis, however, is susceptible to anomalies. Negative entropy is employed as a measure of non-Gaussian since a Gaussian variable has the highest entropy among all the random variables with equal variance. The negative entropy can be defined as

$$J(y) = H(p_{y_{\text{gaussian}}}) - H(p_y) \quad (5)$$

where $H(p_{y_{\text{gaussian}}})$ is the entropy of a Gaussian random variable y having identical covariance and correlation matrix. The differential entropy $H(p_y)$ is defined as follows:

$$H(p_y) = - \int p_y(\eta) \log p_y(\eta) d\eta \quad (6)$$

where $p_y(\eta)$ denotes the probability density function of y . An approximation is utilized instead because it is difficult to quantify the negative entropy.

$$J(y) \approx \{E[G(y)] - E[G(y_{\text{gaussian}})]\}^2 \quad (7)$$

where $G(y)$ is denoted by a function which is non-quadratic and y_{gaussian} as zero-mean, unit-variance Gaussian variable. The estimation would be reliable, if $G(y)$ is selected to grow slowly. The two proven useful forms of $G(y)$ are,

$$G(y) = \frac{1}{a_1} \log \cosh(a_1 y) ; 1 \leq a_1 \leq 2 \quad (8)$$

$$G(y) = - \exp\left(\frac{y^2}{2}\right) \quad (9)$$

As a result, contrast functions in ICA are employed to attain approximations of negative entropy that provide an excellent balance between the two traditional non-gaussian measurements provided by negative entropy and kurtosis. They have attractive statistical qualities, particularly resilience, and are conceptually straightforward and quick to compute. The constraint of decorrelation results from the optimization with a single unit contrast function, which is defined as

$$\text{Maximize, } \sum_{i=1}^N J(y_i) = \sum_{i=1}^N J(w_i^T x), \text{ under constraint } E\{(w_k^T x)(w_j^T x)\} = \delta_{jk} \quad (10)$$

where, w_i , for $i = 1, 2, \dots, N$ represents the rows of matrix, W and, δ is the Kronecker delta function.

One by one, the independent components are estimated to optimize the single unit deflation algorithm represented by equation (10). The algorithm needs to be run multiple times in order to estimate multiple independent components. The fixed-point iteration technique is the foundation of the FastICA algorithm that prevents convergence to the same maxima by decorrelating the vectors after each iteration. To make the ICA method simpler, first the images are pre-processed by whitening and centering. The observed images are turned into zero-mean variables by centering, and the observed and centered images are linearly changed by whitening so that the transformed images have unit variance, zero mean, and are decorrelated.

2.1 FASTICA ALGORITHM

A recursive fixed-point algorithm was developed by Hyvarinen [28-30] for maximising negative entropy, called FastICA, provides a convergence rate ten to one hundred times quicker than that of conventional gradient descent methods, is incredibly computationally efficient. This algorithm, which estimates ICA, is dependable, stable, robust, and easy to use. FastICA's conversion is nonlinear and cubic, or at least quadratic, since no parameters are selected, in contrast to gradient descent techniques. The FastICA algorithm's formal derivation considers the pre-whitening of the data through spheroidizing and centering. Additionally, it takes into account the negative entropy approximation with zero mean and unit variance. The learning rule for FastICA as illustrated in Fig. (3), in search of non-gaussian can be described as follows:

$$w_{i+1} = E\{Xg(w_i^T X)\} - E\{\bar{g}(w_i^T X)\}w_i \quad (11)$$

The one-unit process and the decorrelation process are two processes that can be used to summarize the FastICA algorithm. The following is how the one-unit procedure estimates a single independent component:

1. Set randomly the initial weight vector w_i , with $\|w_i\| = 1$
2. Determine the observed vector X 's covariance matrix $C = E\{XX^T\}$
3. Update the weight vector w_i by

$$w_{i+1} = C^{-1}E\{Xg(w_i^T X)\} - E\{\bar{g}(w_i^T X)\}w_i \quad (12)$$

where, $g(\cdot) = \tanh(\cdot)$.

4. Normalizing the wight w_i using,

$$w_{i+1} = \frac{w_{i+1}}{\sqrt{(w_{i+1})^T C w_{i+1}}} \quad (13)$$

5. Go to step (2) till convergence,

$$\text{i.e., } \Delta w_i = (w_i - w_{i+1})^2 < \epsilon \text{ (a small number)}. \quad (14)$$

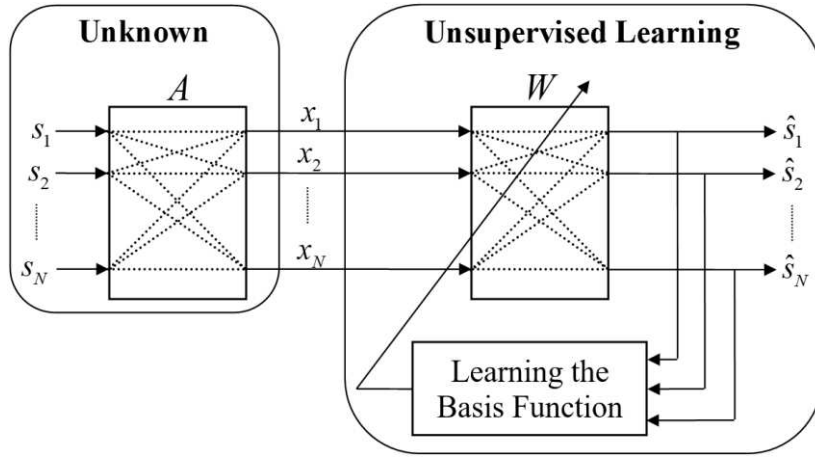


Fig. (3): Learning the Basis Function using FastICA

3. EXPERIMENT AND SIMULATION RESULT

A set of 10,000 random samples with an 8×8 patch size on a set of 15 seismic images were used in the experiment. Basis or features $a_1, a_2, \dots, a_{64}$ that were produced by FastICA method on 8×8 pixel patches of the seismic images are depicted in Fig. (4) which shows that most of the features, that is, the columns of 'A', resemble "edges". This makes it very evident that the rows of $W^{-1} = A$ must likewise match to the same 'edges'.

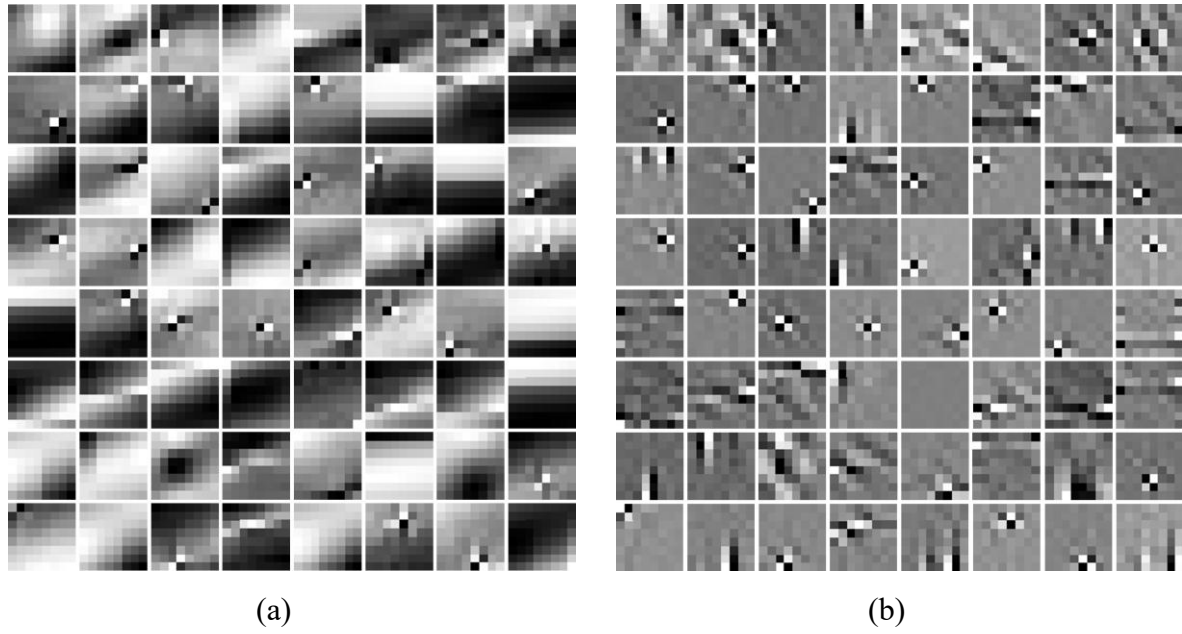


Fig. (4): (a) ICA bases and (b) ICA Filters (Reconstruction) of the seismic image.

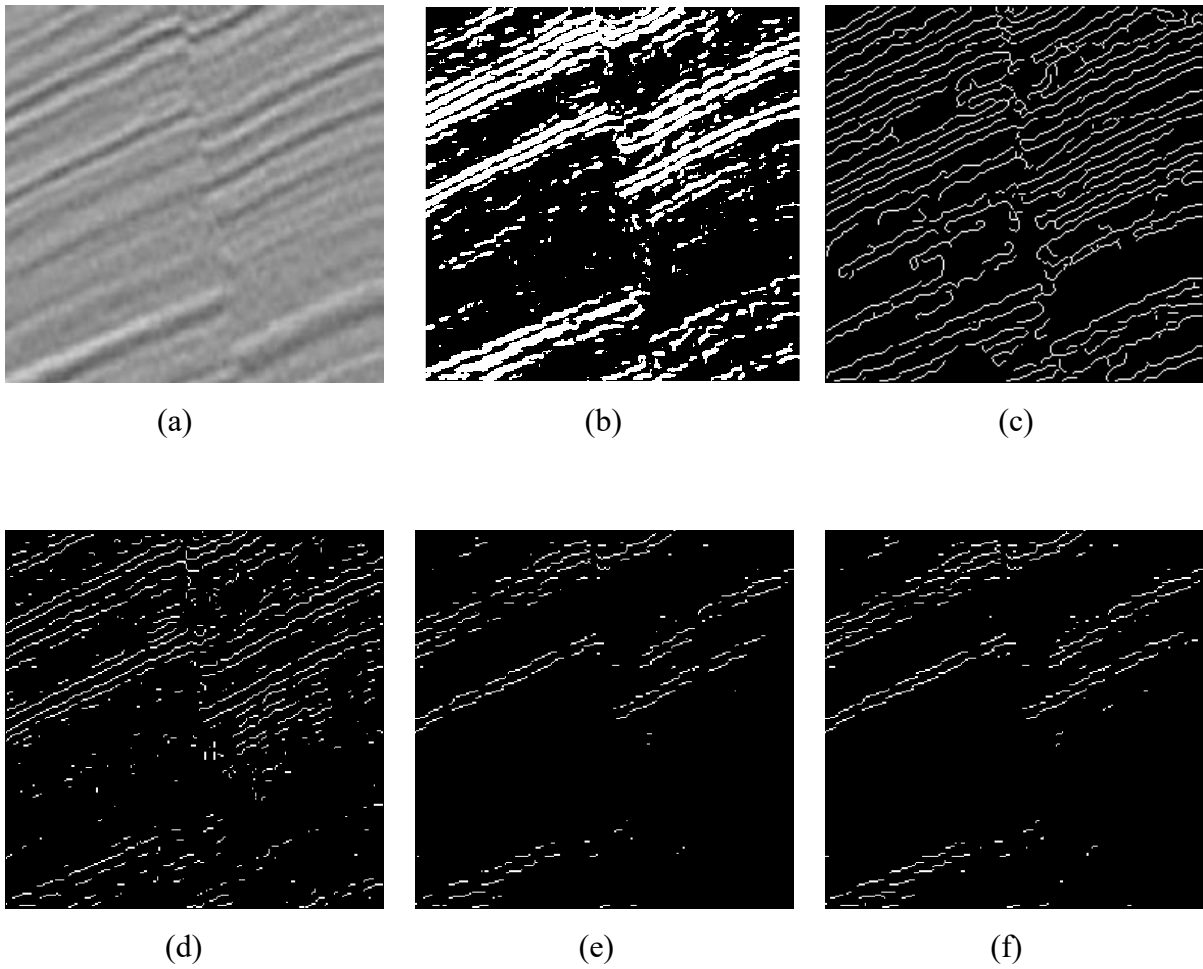


Fig. (5): The edge maps of the seismic image.

(a) Seismic Image (b) Proposed ICA Method (c) Canny (d) LoG (e) Sobel (f) Prewitt.

Moreover, majority of the estimated source signals $s_1, s_2, \dots, s_{64}$ are strongly leptokurtic (mean kurtosis =103.2 which is sufficient to demonstrate their distribution and sparseness). The resulting feature map has the benefit of being insensitive to illumination and can be utilized for pattern extraction. The edges and background pixels in the feature map are categorized into various classes that correspond to the corresponding values between 1 and 64. Only the sparse components take part in the edge extraction of objects or the foreground, and every element is assigned binary 1 whereas, all non-sparse components to binary 0 representing the background hence, a binary image pattern is obtained. Edges representing misalignment of horizons revealing seismic faults has been subsequently recognized using the proposed method, Canny, LoG, Sobel, and Prewitt giving a baseline comparison as shown in Fig. (5).

4. CONCLUSION

This paper suggested a novel feature extraction method for edge detection using ICA to recognize seismic image features such as faults by tracking horizons. ICA has been effectively used to extract sparse basis functions that represent edges. Because statistical processes are inherently data-adaptive, they produce superior outcomes than other transform techniques like wavelet. An estimate of the probability distribution function of random vectors of seismic images using different parameters is known as a statistical model of random variable. Parameters are the foundations or characteristics of seismic image patches. Because they may be decomposed as a linear combination of the eigenvectors linked to the correlation matrix's smallest eigenvalues, the ICA filters represented by rows of a full rank matrix are high pass. This characteristic results in a sparse distribution of the independent components, or high frequency components, whose insignificant elements correspond to patches with lines, contours, or edges. The simulation results of ICA compared to classical methods prove robustness of the proposed algorithm.

The future of edge detection in seismic interpretation is extremely promising due to the development of artificial intelligence, high-performance computation, and contemporary imaging techniques. These developments will play a crucial role in oil and gas exploration, earthquake monitoring, environmental studies, and subsurface imaging technologies.

Declarations:

1. Funding Declaration

No funds, grants, or other support were received during the preparation of this manuscript.

2. Author Contribution Declaration

- Conceptualization of the idea of how image processing could be used to identify geological features.
- Design of methodology based on Independent Component Analysis to meet the requirement.
- Manage to extract features from the use of the seismic image data and perform the analysis.
- Original draft preparation, creation and presentation or finalization.

3. Consent to Publish declaration: Not Applicable

4. Ethics and Consent to Participate declarations: Not Applicable

REFERENCES

- [1]. Bose, N.K., Biemond, J., and Simaan, M., 1990. Editorial on Multidimensional Systems and Signal Processing, Kluwer Academic Publishers, Boston.
- [2]. Keskes, N., Boulanmar, A. and Faugeras, O., 1982. "Application of Image Analysis Techniques to Seismic Data, Proceeding of IEEE Int. Conf. ASSP, Paris, pp. 855-858.
- [3]. Love, P.L., and Simaan, M., 1984. Seismic Signal Character Recognition using Texture Analysis Techniques, Proc. IEEE Int. Conf., ASSP, San Diego, 48.12.1- 48.12.4.
- [4]. Pitas, I. and Kotropoulos, C., 1989. Texture Analysis and Segmentation of Seismic Images, Proceeding, IEEE International Conference on ASSP, Glasgow, pp. 1437-1440.
- [5]. Simon, M. and Zhang, Z., 1990. Automatic Segmentation of Seismic Data Via Knowledge-Based Image Processing Techniques, IEEE.

- [6]. Pitas, I., and Venetsanopoulos, A.N., 1993. Knowledge - Based Image Analysis for Geophysical Interpretation, *Journal of Intelligent and Robotic Systems* 7: 115-137.
- [7]. Wiseman, S., June 7-11, 1999. Seismic Image Segmentation, European Association of Geoscientists and Engineers Conference, Helsinki, Finland.
- [8]. Bakker, P., Verbeek, P.W., and Van Vliet, L.J., Jun. 1999. Edge Preserving Orientation Adaptive Filtering, Vol. 2, pp. 535 - 540, CVPR, Colorado, U.S.A.
- [9]. Pratt, W.K., 1991. Digital Image Processing, John Wiley and Sons, New York, 2nd Ed.
- [10]. Gonzalez, R.C., and Woods, R.E., 1998. Digital Image Processing, Addison - Wesley.
- [11]. Jain, A.K., 2000. Fundamentals of Digital Image Processing, Prentice - Hall of India.
- [12]. Nixon, M.S. and Aguado, A.S., 2008. Feature Extraction and Image Processing, 2nd Ed., Elsevier Ed.
- [13]. Canny J., A Computational Approach to Edge Detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, NO. 6, pp. 679-698, Nov. 1986.
- [14]. Davis, L.S., 1975. A Survey of Edge Detection Techniques, *Computer Graphics and Image Processing*, Vol. 4, pp. 248-270.
- [15]. Ramesh, V., and Haralick, R.M., 1992. Performance Characterization of Edge Detectors, *Applications of Artificial Intelligence X: Machine Vision and Robotics*, Kevin W. Bowyer; Ed, *Proceedings SPIE*, Vol. 1708, pp. 252-266.
- [16]. Heath, M.D., Sarkar, S., Sanocki, T., and Bowyer, K.W., Dec., 1997. A Robust Visual Method for Assessing the Relative Performance of Edge - Detection Algorithms, *IEEE Transactions on PAMI*, Vol. 19, pp. 1338-1359.
- [17]. Maini, R., and Aggarwal, H., 2009. Study and Comparison of Various Image Edge Detection Techniques, *International Journal of Image Processing*, Vol. 3(1), pp.1-12.
- [18]. Panda, C.S., and Patnaik, S., 2009. Filtering Corrupted Image and Edge Detection in Restored Grayscale Image using Derivative Filters, Vol. 3, Issue No. 3, pp. 105-119.
- [23]. Comon, P., 1994. Independent Component Analysis - A New Concept? *Signal Processing*, 36, pp. 287-314.
- [24]. Hyvarinen, A., Karhunen, J., and Oja, E., 2001. Independent Component Analysis, Wiley: London.
- [25]. Hyvarinen, A., 1999. Survey of Independent Component Analysis, *Neural Computing Surveys*, 2:94-128.
- [26]. Hyvarinen, A., Hurri, J., and Hoyer, P.O., 2009. Natural Image Statistics, Berlin : Springer - Verlag.
- [27]. Lee, T.W., 1998. Independent Component Analysis - Theory and Applications, Kluwer

Publishers.

- [28]. Hyvärinen, A., 1999. Fast and Robust Fixed - Point Algorithms for Independent Component Analysis, IEEE Transactions on Neural Networks, 10(3), 626-634.
- [29]. Hyvärinen, A., and Oja, E., 2000. Independent Component Analysis : Algorithms and Applications, Neural Networks, 13(4-5):411-430.
- [30]. Hyvärinen, A., 1997. One-Unit Contrast Functions for Independent Component Analysis : A Statistical Analysis, In Neural Networks for Signal Processing VII (Proc. IEEE NNSP Workshop'97, Amelia Island, Florida), pp. 388-397.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SWsPrograms.zip](#)