Study of Change Detection based on Edge Detection of Satellite Images

Sanjay Kumar Singh

Associate Professor, Department of Communication Engineering,
School of Electronics Engineering, VIT University, Vellore, Tamilnadu – 632 014, India.

Orcid Id: 0000-0002-5760-1865

Abstract
This paper focuses on a robust and flexible edge detection technique based on Independent Component Analysis (ICA) on satellite images for recognition and subjective analysis of changes caused due to Japan earthquake 2011 followed by a devastating tsunami. ICA has been applied to satellite image patches to learn the basis functions using the fixed-point FastICA algorithm. As most of the basis functions are sparse, they are used as pattern template for feature extraction. These basis functions are usually localized, band-limited and oriented like Human Visual System (HVS) and resemble as Gabor wavelet basis function. In the proposed edge detection technique GeoEye’s IKONOS satellite images corresponding to pre and post events have been first transformed to pattern maps (feature map) in which edges and background pixels have been classified into different classes. The edges have been extracted by using sparse components only, whereas non sparse components have been suppressed and treated as background.

Keywords: Computer Vision; Edge Detection; Remote Sensing; Tsunami; Principal Component Analysis; Independent Component Analysis; Fast ICA; Human Visual System; Sparse Coding; Gabor Wavelet Functions.

INTRODUCTION
Computer vision plays a very active role in satellite remote sensing for change detection, pattern recognition, feature extraction, etc; of an area or a scene [18]. Edge detection is one of the important tasks in computer vision which extracts most important visual clues from the image, which is significant in understanding and interpreting the image features. The locations where the intensity of gray or color pixels significantly vary in some direction across a small number of pixels is identified as the edge of an image. The process of edge detection reduces an image to show only its edges, which appear as the outlines of objects within the image that can be used in subsequent image analysis operations for feature detection and pattern recognition. It reduces significantly the size of the image and filters out the information that may be regarded as less relevant, preserving the important structural features of the image. Most images contain some amount of redundancies that can be removed when edges are detected and replaced when it is reconstructed. Hence, the size of the image is reduced, while retaining the image’s vital information with high frequency. When an image is represented as edges, the number of bits or pixels required to store an image reduces (hence, cost and complexity also reduces) with the original image is restored completely at the receiving end for further processing and interpretation [10][21][23][26][28].

Although many different methods for edge detection are available in the literature, but some common problems of these methods are a large volume of computation and too much sensitivity to noise. The classical methods such as the Sobel, Prewitt and Kirsch detectors use the first directional derivative to determine edges. These detectors are simple to implement, but they are usually inaccurate and highly sensitive to noise. The zero-crossing edge detectors use the second derivative along with the Laplacian operator. They have fixed detection characteristics in all directions, but they are also very sensitive to noise. The Gaussian edge detectors, such as the Canny, is one of the most popular edge detectors, has been widely used in many applications because of its superior performance especially in noisy conditions, which reduces the influence of noise by smoothing the image before detecting edges. Although the Gaussian detectors exhibit relatively better performance, they are computationally much more complex. All these edge detection algorithms have succeeded in various degrees for the detection of weak and strong edges; however one of the most common drawbacks of these algorithms is that most existing algorithms that produce precise results lack computational regularity. Even though, many effective methods for edge detection has been an active area for more than 40 years, still it may be difficult to detect and realize the optimal edge detection strategy [19][20]. Therefore, an edge detector that is capable of extracting edges from images in an efficient manner is highly desirable [13][24][27].
HUMAN VISUAL SYSTEM

Hubel and Wiesel point out that in the HVS a series of cells from the retina to the cerebral cortex describe the receptive field which is the basic structure and functional unit of information processing in the visual system. It is activated by corresponding photoreceptor cells of the retina and can cause or modulation of visual cells to respond. Neurophysiologic studies have shown that the receptive field of cells in the primary visual cortex is direction sensitivity significantly. Individual neurons are only responsive to the stimulus in the receptive fields. In short, a single neuron shows a strong reflection to only certain band information such as the specific direction of the edge, line segments, stripes and other image characteristics.

Figure 1: Flow Chart of Visual Information Processing


That is, a neuron in the retina responds simply to whatever contrast is present at that point in space, whereas a neuron in the cortex would respond only to a specific spatial configuration of light intensities (e.g. an edge of a particular orientation) [14]. As shown in Fig. (1), light is detected by the photoreceptors of the retina and converted into electrical signals that are sent to the neurons in the visual cortex. The electrical responses of the neurons fire referred to as potential actions or spikes. In the past few years, many contributors have dealt with the problem of determining an explicit relation between the spikes and the signals from the photoreceptors.

Barlow’s principle of redundancy reduction, proposes that a useful goal of sensory coding is to transform the input in a manner that reduces the redundancy due to complex statistical dependencies among elements of the input streams. The usefulness of redundancy reduction can be understood by considering the process of image formation, which occurs by light reflecting off of independent entities (i.e. objects) in the world and being focused onto an array of photoreceptors in the retina. The activities of the photoreceptors themselves do not form a particularly useful signal to the organism because the structure present in the world is not made explicit, but rather is embedded in the form of complex statistical dependencies, or redundancies among photoreceptor activities.

A reasonable goal of the visual system is to extract these statistical dependencies, so that images may be explained in terms of a collection of independent events, so that means forming the sparse representation for a given image. The properties of the receptive fields of simple cells in primary visual cortex were compared which was found to be similar to the properties of independent component filters generated by performing ICA on a large set of natural images [16].

SPARSE CODING

The sparse code is a kind of neural code in which every item is encoded strongly by the activation of a relatively smaller set of neurons. Sparse coding of natural images results in wavelet-like oriented filters which resemble the receptive fields of simple cells of the primary visual cortex, could possibly confer several advantages such as superior information storage capacity, redundancy reduction, intrinsic fault-tolerance, information processing capabilities as well as huge payoffs in power/energy minimization and optimal resource management, makes the structure in natural signals explicit, representing the complex data to be easily read out at various subsequent processing levels with optimal implementation perspective [11][12].

The sparse coding of input images $I(x,y)$ can be described as signal representation in the form:

$$I(x,y) = \sum s_i a_i(x,y)$$

where, coefficients $s$ have probability distributions (of their activity) with heavy tails highly peaked around zero. The functions $a_i$ are called basis functions of the model. In practical terms, it is said that the particular neuron will be inactive most of the time and that it will fire rarely. This means that an image encoding is possible by only small number of active units. It is very important to point out that such sparse distributions reduce statistical dependence among output units, thus providing the link between sparse coding and the ICA. The approximate illustration of the distribution of activity of code coefficients for sparse coding is shown in Fig. (2).
Independent component analysis (ICA) can be defined as a statistical method which linearly transforms an observed multidimensional random vector of concealed latent variables into components that are as statistically independent as possible from each other. The perceptual system for image processing applications is exposed to a number of small image patches, randomly sampled from one or more large images. Let us consider that each image patch represented by the vector

\[ \mathbf{x} = [x_1, x_2, \ldots, x_N]^T \]

has been formed by linear combination of \( N \) mutually statistically independent basis functions, called ‘observation vector’. The basis function form the columns of a fixed full rank \( N \times N \) matrix

\[ \mathbf{A} = [a_{11}, a_{12}, \ldots, a_{NN}] \]

called ‘mixing matrix’. The weighting of this linear combinations which is varied with each image is given by a vector

\[ \mathbf{s} = [s_1, s_2, \ldots, s_N]^T \]

called ‘source vector’. Each component of this vector is associated with its own basis function which represents an underlying ‘cause’ of the image. The linear image synthesis model is therefore given by

\[ \mathbf{X} = \mathbf{A} \mathbf{S} \]  

which is the matrix version of the set of equations

\[ x_i = \sum_{j=1}^{N} a_{ij} s_j \]  

where \( x_i \) is a pixel in an image, contributing to a set of \( N \) image ‘sources’, \( s_j \), linearly weighted by a coefficients, \( a_{ij} \) (\( 1 \leq i, j \leq N \)). The main aim of perceptual system is to linearly transform the images, \( \mathbf{X} \), with a matrix of filters,

\[ \mathbf{W} = [w_1, w_2, \ldots, w_N] \]

so that the resulting vector recovers the underlying causes, \( \mathbf{S} \), possibly in different order, rescaled and statistically independent.

\[ \hat{\mathbf{S}} = \mathbf{WX} \approx \mathbf{S} \]  

The basis functions (columns of \( \mathbf{A} \)) and the filters which recover the causes (rows of \( \mathbf{W} \)) have the simple relation \( \mathbf{W} = \mathbf{A}^{-1} \), called ‘separating matrix’. The generative model for ICA is shown in Fig.(3) [1][2][3][15][22][26].

From an image processing perspective, ICA can be considered as one of the most basic statistical model for natural images which performs a linear superposition of basis images having non-Gaussian and independent weighting coefficients [7][17]. A series of image patches sampled randomly in the image where, each image patch can be represented by a linear combination of ‘basis’ images \( a_{ij} \) as shown in Fig.(4).

Sparse coding attempts to find the weight matrix \( \mathbf{W} \) making the coefficients \( s_j \) as sparse as possible, meaning that from a large set of image patches only a few of them are significantly active (i.e., non-zero) resemble receptive fields of neurons in the visual cortex as being spatially localized, oriented and band pass [8][9][14][25]. The procedure of estimating independent components \( s_j \) is actually a process of decorrelation. If the information on \( s_j \) does not give any information on the other components, then \( s_j \) is considered independent of these components. According to the assumption of nongaussian distribution, the desired...
independent components \( s_j \) contain the least Gaussian components. A measure of non-Gaussianity is the key for estimating the unmixing matrix and therefore the independent components. The classical measure of non-Gaussianity is kurtosis, which is the fourth order statistics and has zero value for Gaussian distribution, as

\[
\text{kurt}(y) = E \{ y^4 \} - 3 \{ E \{ y^2 \} \}^2
\]

The random variables with negative kurtosis are called 'subgaussian' and those with positive kurtosis are called 'supergaussian'. However, kurtosis is sensitive to outliers. Because a Gaussian variable has the largest entropy among all random variables of equal variance, negentropy can be used as a measure of non-Gaussianity. The negentropy is defined as

\[
J(y) = H(p_{y_{gaussian}}) - H(p_y)
\]

where, \( H(p_{y_{gaussian}}) \) is the entropy of a Gaussian random variable with the same covariance and correlation matrix as \( y \). \( H(p_y) \) is the differential entropy and is defined as

\[
H(p_y) = -\int p_y(\eta) \log p_y(\eta) \, d\eta
\]

where \( p_y(\eta) \) is the probability density function of \( y \). Since, the negentropy is difficult to compute, an approximation is used instead.

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

where \( y_{gaussian} \) represents a zero-mean unit-variance Gaussian variable and \( G(y) \) a non-quadratic function. If \( G(y) \) is chosen not to grow too fast, the estimation would be robust. It has been proved that the following two forms of \( G(y) \) are very useful.

\[
G(y) = \frac{1}{a_1} \log \cosh (a_1 y)
\]

where, \( 1 \leq a_1 \leq 2 \)

\[
G(y) = -\exp \left( \frac{y^2}{2} \right)
\]

Thus, the main function of contrast functions in ICA is to obtain approximations of negentropy which provides a tradeoff among the properties of the two conventional nongaussianity measures namely, kurtosis and negentropy. They are considered to be simple, computationally fast, bear good statistical properties and robust. The optimization method using single unit contrast function is subjected to the constraint of decorrelation of the data, can be given as

Maximize, \( \sum_{i=1}^{N} J(y_i) = \sum_{i=1}^{N} J(w_i^T x) \)

under constraint, \( E\{ (w_i^T x)(w_j^T x) \} = \delta_{ij} \)  \( 11 \)

where, \( w_i, i = 1, 2,..., N \) are the rows of the matrix \( W \) and \( \delta \) is Kronecker delta function. The optimization method given in Equation (11) is known as a single unit deflation algorithm, in which, the independent components are estimated one by one stepwise. For the estimation of a large number of independent components, this algorithm is executed on several units in multi-step. After iteration the vectors are decorrelated for preventing the convergence to the same maxima. This particular algorithm is known basically as FastICA algorithm which operates on a fixed-point iteration scheme. The images are first pre-processed by centering and whitening for simplifying the ICA algorithm. Centering is the process in which the observed images are converted to zero-mean variables, whereas, whitening transforms the centered images linearly such that they are uncorrelated or independent by normalizing to unit variance which is obtained by identity covariance matrix.

### A. FastICA Algorithm

Hyvarinen [4][5][6] developed a recursive ‘fixed-point algorithm’ for maximising negentropy, called FastICA is highly computationally efficient whose convergence rate is 10-100 times faster than conventional gradient descent methods. FastICA is also considered to be robust, reliable, stable and user-friendly method for performing the estimation of ICA. FastICA convergence is cubic (or at least quadratic) and nonlinear as no parameters have to be chosen contrary to gradient descent methods. The formal derivation of the algorithm considers that data have been prewhitened by centering and spheroïdizing, and also considers an approximation for the negentropy of a zero-mean unit variance. The learning rule as shown in Fig. (5), for FastICA to search for non-Gaussianity can be described as follows:

\[
w_i^* = E\{ Xg(w_i^T X) \} - E\{ g'(w_i^T X) \} w_i
\]

The FastICA algorithm can be summed up of two processes: the one unit process and the decorrelation process. The one unit process is only used to estimate one independent component as follows:

1. Randomly initialize the weight vector \( w_i \), that has \( \|w_i\| = 1 \)
2. Calculate the Covariance matrix \( C \) of the observed vector \( X \) i.e., \( C = E\{XX^T\} \)
3. Update the weight vector $w_i$ by

$$w_i^+ = C^{-1}E\{Xg(w_i^T X)\} - E\{g'(w_i^T X)\}w_i$$

(13)

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

4. Normalize $w_i$,

$$w_i^+ = \frac{w_i^+}{\sqrt{(w_i^+)^T C w_i^+}}$$

(14)

5. Goto (2), until convergence

i.e., $\Delta w_i = \left( w_i - w_i^+ \right)^2 < \epsilon$ (a small number).

**SIMULATION RESULTS**

The experiment has been performed by taking a set of 50000 random samples of patch size $32 \times 32$. Basis or features have been obtained by applying the FastICA algorithm to $32 \times 32$ pixel patches of the images. The basis image shows that most of the features (i.e., the columns of $A$) resemble to ‘edges’. Therefore, the rows of $W^{-1} = A$ must be matched to (the same) ‘edges’. Moreover, most of the estimated source

**Figure 5:** The Linear Image Synthesis Model Learning the Basis Function

**Figure 6:** Pre-earthquake Satellite Image

**Figure 7:** ICA Bases of Fig.(6)

**Figure 8:** ICA Filters of Fig.(6)

**Figure 9:** Edge Detection of Fig. (6) by ICA

**Figure 10:** Post-earthquake Satellite Image
signals are highly kurtotic which proves that they are sparse and distributed. The features map can be used for pattern extraction and has the advantage that it is insensitive to illumination. The Figs. (7 and 11) is a set of features \( a_i \) obtained by ICA and Figs. (8 and 12) represents a set of symmetrically orthogonalized feature vectors \( w_i \) maximizing sparsity learned from images. In the features image, the edges and background pixels have been classified into different classes and represented by the corresponding value between 1 and 256. Only sparse components participate to extract edges and every element value is set to 1 whereas, non-sparse components are set to 0 hence, a binary image pattern is achieved.

The multi-resolution optical satellite images, Fig. (6) of half-meter resolution and Fig. (10) of one-meter resolution has been captured by GeoEye’s IKONOS satellite before (April 4, 2010) and one day after (March 12, 2011) the 9.0-magnitude earthquake struck the Oshika Peninsula on March 11, 2011 from 423 miles in space as it moved from north to south over Japan at a speed of 4 miles/second.

### Table 1: Standard Image Quality Measures

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Formulae</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Square Error</td>
<td>( \text{MSE} = \frac{1}{MN} \sum_{j=1}^{M} \sum_{k=1}^{N} (x_{j,k} - x'_{j,k}) )</td>
<td>5.2e+03</td>
</tr>
<tr>
<td>Peak Signal to Noise Ratio</td>
<td>( \text{PSNR} = 10 \log \left( \frac{2^n - 1}{\text{MSE}} \right) )</td>
<td>10.97</td>
</tr>
<tr>
<td></td>
<td>( = 10 \log \left( \frac{255}{\text{MSE}} \right) )</td>
<td></td>
</tr>
<tr>
<td>Normalized Cross-correlation</td>
<td>( \text{NK} = \frac{\sum_{j=1}^{M} \sum_{k=1}^{N} x_{j,k} x'<em>{j,k}}{\sum</em>{j=1}^{M} \sum_{k=1}^{N} x_{j,k}^2} )</td>
<td>0.59</td>
</tr>
<tr>
<td>Average Difference</td>
<td>( \text{AD} = \frac{\sum_{j=1}^{M} \sum_{k=1}^{N} (x_{j,k} - x'_{j,k})}{MN} )</td>
<td>33.21</td>
</tr>
<tr>
<td>Structural Content</td>
<td>( \text{SC} = \frac{\sum_{j=1}^{M} \sum_{k=1}^{N} x_{j,k}^2}{\sum_{j=1}^{M} \sum_{k=1}^{N} x_{j,k}^2} )</td>
<td>1.72</td>
</tr>
<tr>
<td>Maximum Difference</td>
<td>( \text{MD} = \max \left(</td>
<td>x_{j,k} - x'_{j,k}</td>
</tr>
<tr>
<td>Normalized Absolute Error</td>
<td>( \text{NAE} = \frac{\sum_{j=1}^{M} \sum_{k=1}^{N}</td>
<td>x_{j,k} - x'_{j,k}</td>
</tr>
</tbody>
</table>
The simulation results have been found to be quite encouraging and prove the robustness of the proposed technique in order to be compared the two edges detected imageries corresponding to before and after the damage caused by Japan earthquake 2011. From the simulation results as shown in Fig. (6) it can be clearly observed that the regular patterns of the edges correspond to pre-tsunami images represent structures such as houses, rooftops, road networks, etc. are intact which is reflected in the corresponding edge image as shown in Fig. (9). On the other hand, the post-tsunami image as shown in Fig. (10) represents the areas filled with water bodies, mud, debris, partially or fully destroyed structures which the same can be visualized in an edge image as shown in Fig. (13).

The standard image quality of the two imageries as shown in Fig. (6) and Fig. (10) can be quantitatively measured which has been provided in Table (I) in order to find differences between them qualitatively. Similarly, the modern similarity index measures can be tabulated as presented in Table (II) to measure the edge image quality and establish a comparison between Fig. (9) and Fig. (13).

CONCLUSION

In this paper, a robust feature extraction technique has been proposed for edge detection based on ICA to extract satellite image features. Statistical processes have inherent property of data adaptiveness which results better as compared to other transform methods like wavelet. The given image has been randomly sampled into small patches. These patches are considered as a random vector for the modelling of the data. Statistical model of a random variable is nothing but the calculation of the probability distribution function of the random vectors of the images using different parameters. Parameters in case of image patches are its features or basis. ICA successfully applied to extract basis functions which are sparse and represent ‘edges’. ICA provides a sparse code in terms of a set of ICA bases that represent oriented and localized lines or edges. ICA filters (rows of matrix A) have been decomposed as a linear combination of the eigenvectors associated to the smallest eigenvalues of the data correlation matrix, meaning that these filters are all high-pass. This property leads to a sparse distribution of the independent components, whose non-negligible elements correspond to patches that contain lines or edges (i.e., high frequency content).

Table II: Edge Image Quality Measures

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edge Based Similarity Index (ESSIM)</td>
<td>0.9941</td>
</tr>
<tr>
<td>Feature Similarity Index (FSIM)</td>
<td>0.6095</td>
</tr>
</tbody>
</table>

REFERENCES


