We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,300

Open access books available

172,000

International authors and editors

190M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

SFE2D: A Hybrid Tool for Spatial and Spectral Feature Extraction

Bahman Abbassi and Li Zhen Cheng

Abstract

A crucial task for integrated geoscientific image (geo-image) interpretation is the relevant geological representation of multiple geo-images, which demands high-dimensional techniques for extracting latent geological features from high-dimensional geo-images. A standalone mathematical tool called SFE2D (spatiospectral feature extraction in two-dimension) is developed based on independent component analysis (ICA), continuous wavelet transform (CWT), k -means clustering segmentation, and RGB color processing that iteratively separates, extracts, clusters, and visualizes the highly correlated and overlapped geological features from multiple sources of geo-images. The SFE2D offers spatial feature extraction and wavelet-based spectral feature extraction for further extraction of frequency-dependent features. We show that the SFE2D is a robust tool for automated pattern recognition, fast pseudo-geological mapping, and detection of regions of interest with a wide range of applications in different scales, from regional geophysical surveys to the interpretation of microscopic images.

Keywords: feature extraction, independent component analysis, continuous wavelet transform, k -means clustering segmentation, pseudo-geological mapping

1. Introduction

Global trends show that the number of major mineral discoveries is drastically declined during the last 25 years because the most outcropping mineralization systems are mapped and already excavated [1]. Therefore, a significant challenge for mineral exploration is the deep targeting of hidden resources. The “deep” targets are usually considered as resources as deep as hundreds of meters to thousand meters of depth. However, occurrences of “deep” targets are not limited to the third dimension. They can be seen as complex obscured mineralization systems barely detectable within a large amount of hyper-dimensional geoscientific data. Consequently, advanced data mining methods are needed to extract useful information from large amounts of data sets. Responding to these challenges of integrating multidisciplinary information is the motivation of present research.

Automated interpretation of multiple images is an important topic in large data interpretation and multivariate pattern recognition. No matter how accurate the geoscientific images (geo-images) are, each geo-image is only sensitive to a limited number of geological features. Therefore, different parts of the subsurface geology can be reconstructed from various geo-images. The critical question is how one can put together the interpreted jigsaw puzzles inside a high-dimensional space to

rebuild a relevant geological model. This has led us to the topic of feature extraction in the treatment of large data sets.

The fundamental problem in integrated geoscientific interpretation is the proper geological understanding of multiple geo-images, which demands high dimensional techniques for the extraction of geological information from hyperdimensional data sets. Several methods are available for feature extraction in the high dimensional space, which can be seen as a dimensionality increase problem [2]. The problem is that the increased dimensionality due to feature extraction leads to an overload of information that creates difficulties for human visual interpretation as well as machine learning optimization [2]. That is to say; latent patterns are lost in high space as one moves to higher dimensions. Several recent studies, mainly in seismic attributes analysis, have investigated the problem of high-dimensional pattern recognition to extract relevant geological features from broadband seismic data [3–7]. In mineral exploration, the treatment of multiple non-seismic data also put forth a similar pattern recognition problem, which seeks to extract relevant geological features from potential field and electromagnetic data in the form of multiple images [8].

This study aimed to provide a tool for spatial and spectral feature extraction from multiple geo-images to facilitate the identification of geoscientific features. The study employed a quantitative approach combining different source separation methods for feature extraction and representation. The output is the SFE2D program, which is a standalone 2D spatio-spectral feature extraction tool based on unsupervised source separation methods like principal component analysis (PCA), independent component analysis (ICA), continuous wavelet transform (CWT), RGB color processing, and k -means clustering segmentation. The program is made in a MATLAB environment for feature extraction and dimensionality reduction of hyperdimensional geoscientific data sets through variance/kurtosis/negentropy maximization, k -mean clustering segmentation, color pick algorithm, and RGB visualizations (SFE2D ver. 2.0).

2. Theoretical background of source separation

2.1 Feature extraction

The theoretical framework underpinning this study is based on blind source separation (BSS) methods. Generally, BSS aims to recover latent source signals (or images) from a set of highly mixed signals (or images). Consider a mixing model, where every geo-image ($\mathbf{g}_i$) is a mixture of several latent features ($\mathbf{f}_j$) with different contributions in constructing the observed image. A mathematical description for this model can be formulated as a linear mixing system that links the latent features to the observed geo-image by a set of mixing weights (a_{ij}) that determine the contribution of each feature in the geo-image construction:

$$\mathbf{g}_i = \sum_j a_{ji} \mathbf{f}_j \quad (1)$$

Where $j = 1, 2, \dots, n$ denotes the number of latent features, and $i = 1, 2, \dots, m$ indicates the number of geo-images. The vector form of Eq. (1) can be rewritten as:

$$\mathbf{g} = A\mathbf{f} \quad (2)$$

Where the unknown mixing weights are defined as the matrix $[A]_{(j, i)} = a_{ji}$. The observed geo-images $\mathbf{g} = [\mathbf{g}_1(x, y), \mathbf{g}_2(x, y), \dots, \mathbf{g}_m(x, y)]^T$ are linear mixtures of the

Figure 1.
 Schematic of feature extraction model.

latent features $\mathbf{f} = [\mathbf{f}_1(x,y), \mathbf{f}_2(x,y), \dots \mathbf{f}_n(x,y)]^T$. The x and y denote the coordinates on a two-dimensional space in the spatial feature extraction scheme. The problem is to find a separation matrix (W) that tends to unmix the geo-images ($\mathbf{g}$) to recover the hidden features ($\mathbf{f}$):

$$\mathbf{f} = W\mathbf{g} = A^{-1}\mathbf{g} \tag{3}$$

Where separation matrix (W) is the inversion of the mixing matrix and cannot be directly determined. Nevertheless, it can be estimated adaptively as an optimization problem and setting up a cost function to be minimized iteratively. This is the general formulation of a feature extraction process that iteratively eliminates or at least reduces the effect of feature overlaps (shadow effect) on the observed geo-images (**Figure 1**).

2.2 Spatial feature extraction through PCA and ICA

Estimating a relevant separation matrix is possible by making assumptions about the statistical measures such as correlation. In some instances, it is common for two geo-images to be statistically correlated. For example, suppose there is a linear relationship between two gamma-ray concentration images (e.g., K vs. eTh). In that case, the amount of information that the first image provides is as same as the second one. Therefore, one can transform the bivariate data to a univariate form without losing any valuable information. This transformation is called dimensionality reduction, which is the basis of PCA algorithms. PCA algorithms utilize the maximization of second-order statistical measure (variance) for image separation and produce linearly uncorrelated images (**Figure 2a**). However, when there is a nonlinear form of correlation (dependency) between images, PCA will not work. In this case, ICA is useful to separate the geo-images into nonlinearly uncorrelated images by maximizing non-Gaussianity (**Figure 2b**).

PCA is, in fact, the first step in the ICA process. PCA looks for a weight matrix D so that a maximal variance of the principal components (y) of the centered geo-images ($\mathbf{g}_C$) are confirmed:

$$y = D\mathbf{g}_C \tag{4}$$

During this preprocessing step, the mean of the input geo-images ($\mathbf{g}_i$) is removed (also known as Centering). Then whitening is performed by eigenvalue decomposition, which results in unit variance and identity covariance matrix. The matrix D that is calculated for linear transformation of centered geo-images into whitened images is:

$$D = \lambda^{-1/2}\mathbf{E}^T \tag{5}$$

Figure 2.

Schematic description of feature extraction process: (a) PCA through maximization of variance. The correlated bivariate data sets (g axes) are transformed into univariate forms ($\hat{f}$ axes) without losing valuable information. (b) ICA through maximization of non-Gaussianity. The correlated and interdependent (overlapped) bivariate data sets (g axes) are transformed into two separate univariate independent components ($\hat{f}$ axes).

Where λ is eigenvalue and E is the eigenvector matrix of the covariance matrix of the geo-images [9].

Whitening solves half of the ICA problem. The second step in the ICA is increasing the non-Gaussianity of the whitened images or principal components (y). The problem is to find a rotation matrix R that during the multiplication with principal components produces the least Gaussian outputs ($\hat{f}$) that are approximations of the latent features (**Figure 3**):

$$\mathbf{f} \approx \hat{\mathbf{f}} = R^T \mathbf{y} \quad (6)$$

2.3 Fast-ICA algorithms

This study has employed two alternative methods to obtain the rotation matrix R based on Fast-ICA through kurtosis maximization and negentropy maximization. Both kurtosis and negentropy maximization methods performed through the Hyvärinen fixed-point algorithm employ higher-order statistics to recover the independent sources [9]. The fixed-point algorithm has a couple of properties that make it superior to the gradient-based methods: (a) It has a cubic speedy convergence criterion. (b) In contrast to gradient-based algorithms, it does not need any learning rate adjustment or parameter tuning, making it easy to implement [9].

Mixing $\rightarrow$ Whitening $\rightarrow$ Rotating

Figure 3.

Generalized steps of independent component analysis for feature extraction.

2.3.1 Fast-ICA through kurtosis maximization

According to the central limit theorem, the distribution of a sum of independent random variables tends toward a Gaussian (normal) distribution [9]. One can use this principle to maximize the non-Gaussianity through kurtosis maximization (kICA). Kurtosis that is the fourth-order cumulant of the whitened images can be expressed as the normalized version of the fourth moment:

$$kurt(y) = E\{y^4\} - 3 \quad (7)$$

Where $E\{\cdot\}$ denotes the expectation over the unknown density of input geo-images. Kurtosis provides a measure of how Gaussian ($kurt = 0$), super-Gaussian ($kurt > 0$) or sub-Gaussian ($kurt < 0$) the probability density functions of the geo-images are. Maximization of the kurtosis of the principal components estimates the latent independent components ($\hat{\mathbf{f}}$). The algorithm for kurtosis maximization in the present SFE2D program incorporates deflationary orthogonalization that estimates the independent components one by one:

- Step 1. Center and whiten the geo-image data $\mathbf{g}$ to give $\mathbf{y} = D\mathbf{g}_C$, where $D = \lambda^{-1/2}\mathbf{E}^T$.
 Step 2. Set N as the number of independent components to be estimated and a counter $I \leftarrow 1$.
 Step 3. Set randomly an initial vector $\hat{W}_I$ of unit norm (rows of the separation matrix).
 Step 4. Let $\hat{W}_I \leftarrow E\left\{y\left[\hat{W}_I^T y\right]^3\right\} - 3\hat{W}_I$.
 Step 5. Rotate by deflationary orthogonalization: $\hat{W}_I \leftarrow \hat{W}_I - \sum_{i=1}^{I-1} \left(\hat{W}_I^T \hat{W}_i\right) \hat{W}_i$.
 Step 6. Normalize: $\hat{W}_I \leftarrow \hat{W}_I / \|\hat{W}_I\|$
 Step 7. Go to step 4 if $\hat{W}_I$ is not converged: $\left|\left\langle \hat{W}_I^{k+1}, \hat{W}_I^k \right\rangle\right| \neq 1$. Otherwise, $\hat{W}_I$ is row I of the estimated separation matrix.
 Step 8. If $I \leq N$ go to step 3 and set $I = I + 1$ to estimate the next row of the separation matrix.
 Step 9. Obtain the latent features by $\hat{\mathbf{f}} = \hat{\mathbf{W}}\mathbf{g}$ where $\hat{\mathbf{W}} = [\hat{W}_1, \hat{W}_2, \dots, \hat{W}_I]^T$.

2.3.2 Fast-ICA through fixed-point negentropy maximization

For large-scale studies, kurtosis maximization is time-consuming because of its computational complications. Also, kurtosis is not a robust measure of non-Gaussianity in the presence of unknown noises and image artifacts. Kurtosis is very sensitive to outliers, that is, a single erroneous outlier value in the tails of the distribution makes kurtosis extremely large. Therefore, using kurtosis is well justified only if the independent components are sub-Gaussian and there are no outliers.

Another principle in information theory enables us to obtain the rotation matrix R more efficiently in the presence of outliers. It states that the spatial complexities of the input images are equal to or greater than that of the least complex latent features [9]. This general principle enables us to maximize the non-Gaussianity of multiple geo-images through negentropy maximization (nICA).

Negentropy is based on the information-theoretic quantity of entropy. The entropy (H) of a geo-image (g) with probability density of $p(g)$ is defined as [9]:

$$H(g) = - \int p(g) \log p(g) dg \quad (8)$$

The more random (unpredictable and unstructured) the variable is, the larger its entropy. Between all random variables, the Gaussian variable has the largest entropy. Negentropy of a geo-image $\mathbf{g}$ is normalized differential entropy of $\mathbf{g}$, which is the difference between the entropy of that geo-image $\mathbf{g}$ and the entropy of a Gaussian random vector of the same covariance matrix as $\mathbf{g}$ ($\mathbf{g}_{\text{gauss}}$):

$$\text{neg}(\mathbf{g}) = H(\mathbf{g}_{\text{gauss}}) - H(\mathbf{g}) \quad (9)$$

Negentropy is always non-negative and is zero when the signal has a Gaussian distribution. In other words, the more random (unpredictable and unstructured) the variable is, the larger its entropy [9]. Negentropy approximated in this study is based on using classical higher order cumulants given as:

$$\text{neg}(\mathbf{y}) \approx \frac{1}{12}E\{\mathbf{y}^3\}^2 + \frac{1}{48}\text{Kurt}(\mathbf{y})^2 \quad (10)$$

Since entropy estimation is not directly performed on geo-images, and instead, centered/whitened images (principal components) are used for entropy calculations, therefore, $\mathbf{g}$ is replaced by $\mathbf{y}$. Replacing polynomial functions $\mathbf{y}^3$ and $\mathbf{y}^4$ by another non-quadratic function G^i that does not grow too fast results in robust negentropy estimation (i is an index denoting odd and even functions):

$$\text{neg}(\mathbf{y}) \approx k_1[E\{G^1(\mathbf{y})\}]^2 + k_2[E\{G^2(\mathbf{y})\} - E\{G^2(\nu)\}]^2 \quad (11)$$

Where ν is a standardized Gaussian variable with unit variance and zero mean, and k_1 and k_2 are positive constants. In the case where only one non-quadratic function G is used, the approximation becomes simpler:

$$\text{neg}(\mathbf{y}) \approx [E\{G(\mathbf{y})\} - E\{G(\nu)\}]^2 \quad (12)$$

In this moment-based approximation, we used a slow-growing exponential function for G :

$$G(\mathbf{y}) = -e^{(-\mathbf{y}^2/2)} \quad (13)$$

The criterion for negentropy maximization is based on bringing an objective function to an approximated maximum value. The algorithm for negentropy maximization in the present SFE2D program is as followed:

- Step 1. Center and whiten the geo-image data $\mathbf{g}$ to give $\mathbf{y} = D\mathbf{g}_C$, where $D = \lambda^{-1/2}\mathbf{E}^T$.
- Step 2. Set N as the number of independent components to be estimated and a counter $I \leftarrow 1$.
- Step 3. Set randomly an initial vector $\hat{W}_I$ of unit norm (rows of the separation matrix).
- Step 4. Let $\hat{W}_I \leftarrow E\{\mathbf{y}\mathbf{g}[\hat{W}_I^T \mathbf{y}]\} - E\{\mathbf{g}'[\hat{W}_I^T \mathbf{y}]\}\hat{W}_I$, where $\mathbf{g}(\mathbf{y}) = \mathbf{y}e^{(-\mathbf{y}^2/2)}$.
- Step 5. Rotate by deflationary orthogonalization: $\hat{W}_I \leftarrow \hat{W}_I - \sum_{i=1}^{I-1}(\hat{W}_I^T \hat{W}_i)\hat{W}_i$
- Step 6. Normalize: $\hat{W}_I \leftarrow \hat{W}_I / \|\hat{W}_I\|$.
- Step 7. Go to step 4 if $\hat{W}_I$ is not converged: $|\langle \hat{W}_I^{k+1}, \hat{W}_I^k \rangle| \neq 1$. Otherwise, $\hat{W}_I$ is the row I of the estimated separation matrix.
- Step 8. If $I \leq N$ go to step 3 and set $I = I + 1$ to estimate the next row of the separation matrix.
- Step 9. Obtain the latent features by $\hat{\mathbf{f}} = \hat{\mathbf{W}}\mathbf{g}$ where $\hat{\mathbf{W}} = [\hat{W}_1, \hat{W}_2, \dots, \hat{W}_I]^T$.

2.4 Spectral feature extraction through wavelet-based PCA and ICA

Spectral decomposition of geo-images provides a unique way for feature extraction in the frequency domain. Although the Fourier transform is a powerful tool for image decomposition, it does not represent abrupt changes efficiently due to the infinite oscillation of the periodic function in any given direction. In contrast, wavelets are localized in space and have finite durations. Therefore, the output of wavelet decomposition effectively reflects the sharp changes in images, and that makes it an ideal tool for feature extraction [10–12].

Mathematically, the 2D CWT of an image $I(x, y)$ is defined as a decomposition of that image (I) on a translated and dilated version of a mother wavelet $\psi(x, y)$. Thus, the 2D CWT coefficients are given by:

$$C_s = (b_1, b_2, a) = \frac{1}{\sqrt{|a|}} \iint I(x, y) \psi^* \left(\frac{x - b_1}{a}, \frac{y - b_2}{a} \right) dx dy \tag{14}$$

Where b_1 and b_2 are controlling the spatial translation, $a > 1$ is the scale, and ψ^* is the complex conjugate of the mother wavelet $\psi(x, y)$.

Analysis of isolated sources in potential field data with CWT was introduced by Moreau [13]. Moreau showed that the maxima lines of the CWT indicate the location of the potential field sources [10]. He also showed that the maxima lines bear the highest signal-to-noise ratios, allowing the treatment of the noisy data sets [14]. An example of a CWT on a geophysical image in which the mother wavelet is shifting and scaling in one direction is shown in **Figure 4**. The scale is inversely

Figure 4. (a) CWT on a geophysical image. CWT increases the dimensionality depending on the choice of scales (S) and directions (D). (b) CWT Mexican Hat mother wavelet in five scales (S₁–S₅). (c) CWT with Cauchy mother wavelet in six directions (D₁–D₆ at S₄).

proportional to frequency. Large-scale factors are corresponding to largely expanded mother wavelets (low frequencies).

As can be seen in different frequencies, different features are detectable (**Figure 4b**). Scaling and shifting the mother wavelet in other directions also reveal other sets of features (**Figure 4c**). If the mother wavelet is isotropic, there is no dependence on the angle in the CWT. The Mexican Hat mother wavelet used in **Figure 4b** is an example of isotropic wavelets.

On the other hand, an anisotropic mother wavelet is dependent on the angle in the analysis; therefore, the CWT acts as a local filter for an image in scale, position, and angle. The Cauchy wavelet is an example of an anisotropic wavelet (**Figure 4c**). To better see the effect of directional transformation, we can use different anisotropic mother wavelets and compare the results.

3. The SFE2D architecture

To separate the spectral features, we created a 2D feature extraction scheme that combines 2D CWT with variance, kurtosis, and negentropy maximization algorithms (PCA, kICA, and nICA) to extract spectral features from wavelet spectra. The algorithm consists of three main stages (**Figure 5**):

1. A preprocessing step where 2D interpolation and filtering of raw datasets prepare the input images for feature extraction.
2. Spatial feature extraction employs PCA/ICA for spatial source separation and dimensionality reduction.
3. Spectral feature extraction that consists of two substages: (a) continuous wavelet transform (CWT) and (b) spectral PCA/ICA (SPCA/SkICA/SnICA in **Figure 5**).

An effective spectral feature extraction depends on computer hardware specifications, specifically for larger numbers of spectral features produced by increasing the number of scales and mother wavelet's directions. In practice, users can create a larger number of features by changing the CWT parameters in the SFE2D program.

Figure 5.
Schematic view of the SFE2D procedure.

For large numbers of CWT features (100–1000 CWT features), PCA/ICA methods provide a way to decorrelate and separate spectral features from CWT images.

If needed, users can also reduce the dimensionality to summarize the features. By running the proposed workflow in the SFE2D program, the redundant frequency volumes could be reduced to a more manageable number of components. Taking advantage of the ICA statistical properties, we can keep the most geologically pertinent information within the spectral decomposed data.

4. Feature representation

The SFE2D program provides interpretation tools for feature representation in two dimensions:

4.1 RGB image compilation and analysis

SFE2D users can select the RGB image compilation method to assemble a colored image from three manually chosen extracted features by PCA and ICA methods. The RGB combination makes it possible to bring out hidden characteristics. To control which extracted features are used for image compilation, the user can graphically set the label number of desired features. For example, in nICA feature extraction, the 1, 2, 3 labels mean the features 1, 2, and 3 are used for image compilation. The user can also change the polarity of the features since PCA/ICA methods distort the polarities randomly. For example, one can set $-2, 1, 3$ labels for RGB compilation.

A color pick algorithm is also developed to select the region of interest (ROI) based on color intensities. To find objects of a specific color, the SFE2D program assigns the low and high thresholds for each RGB color band with several clicks on the regions of interest. The more users click on the specified zones, the more accurate the thresholding works. Then, the program automatically picks up the minimums and maximums from red/green/blue bands. To reduce the effect of the outliers, only a range of 5 to 95 percentiles are kept, which results in smoother color picking tasks. **Figure 6** represents an example of testing the ROI detection method in the SFE2D program to detect the red-colored objects in a standard image.

4.2 The color image segmentation algorithm

SFE2D uses a k -means clustering algorithm for the segmentation of color images. The segmentation helps to reduce the color space size to a manageable number of colors. The process produces a pseudo-geological map based on the spatially extracted RGB features. This eventually helps geologists to detect the hidden geological contacts and structures in geo-images. In addition, segmentation significantly reduces memory usage and speeds up image analysis by focusing on relevant information.

The output of segmentation is a set of k non-overlapping segments $\{S_1, S_2, \dots, S_k\}$ that comprises the whole segmented representation of a dataset X in the form of [15]:

$$X = \bigcup_{i=1}^k S_i \quad (15)$$

For the non-overlapping condition, we should have $S_i \cap S_j = \emptyset$ for $i \neq j$, where $1 \leq i, j \leq k$ which guarantees that each cluster of data belongs uniquely to a segmented centroid.

Figure 6. Color pick method for extracting the red color objects. (a) Original RGB image. (b) ROI detection.

SFE2D incorporates a k -means segmentation for grayscale or RGB images. The program aims to separate extracted features from an image in different clusters iteratively. The algorithm computes a hyperdimensional centroid for each cluster. Each segment S_i is uniquely defined by a center c_i :

$$c_i = \frac{1}{|S_i|} \sum_{s \in S_i} s \quad (16)$$

Where $|S_i|$ is the number of elements in the i th cluster.

The centroid gets modified interactively through minimizing a cost function that is the distance between data pattern p_j and centroid c_i :

$$F(c_1, c_1, \dots, c_k) = \sum_{i=1}^k \sum_{p_j \in C_i} d^2(p_j, c_i) \quad (17)$$

If the image is RGB, the program calculates a 3D Euclidean distance for the RGB color space. The SFE2D program reads data points p and returns k cluster centroids as $C = \{c_1, c_2, \dots, c_k\}$. The k -means algorithm for RGB image segmentation in the SFE2D program is as followed:

- Step 1: Generate k random initial centroids: $C \leftarrow \text{Rand}(p, k)$.
- Step 2: Repeat the following until S_i do not change:
- $S_i \leftarrow \left\{ p : \|p - c_i\|^2 \leq \|p - c_j\|^2 \forall j, 1 \leq i, j \leq k \right\}$ such that $S_i \cap S_j = \emptyset$.
 - $c_i \leftarrow \frac{1}{|S_i|} \sum_{x_j \in S_i} x_j$
- Step 3: Return all clustered segments as $C = \{c_1, c_2, \dots, c_k\}$.

5. Evaluation

Before applying SFE2D algorithms to geoscientific applications, we tested the workflows, including PCA, kICA, nICA, ROI color pick algorithm, and k -mean segmentation in two sections dedicated to grayscale and RGB ICA-based image segmentation.

5.1 Spatial feature extraction performance

In grayscale, we tested the performances of kICA versus nICA algorithms and the effect of feature overlaps (shadow effects) on image segmentation. An example of PCA and ICA procedures is presented in **Figure 7**. We simulated and evaluated the feature extraction processes on an overlapped photo that is mixed with three images: a human face, fruits, and vegetable scenes (features f_1 , f_2 , and f_3 in **Figure 7**). The problem can be formulated as the extraction of the human face (feature f_2) from the mixed images (g_1 , g_2 , and g_3) without any prior information

Figure 7. An example of independent component analysis with kurtosis and negentropy maximization on linear mixtures of three independent images (f_1 , f_2 , and f_3 with 180×180 pixels). ICA goes a step further and rotates the principal components to maximize non-Gaussianity. The results are recovered original images. The column on the right side represents the cross plots of the variables.

from the original face. This simulation helps us to evaluate the performance of the SFE2D algorithm before the implementation on more complex applications.

While PCA whitening fails to recover all three features (y_1, y_2, y_3), the Fast-ICA algorithm could separate the latent features inside the mixed images ($\hat{f}_1, \hat{f}_2, \hat{f}_3$). Further analysis of the non-Gaussianity maximization methods revealed that the Fast-ICA through negentropy maximization (nICA algorithm) compared to kurtosis maximization (kICA algorithm) is more effective and recovers more details of each original photo. The kICA converges at 50 iterations, while the nICA converges at iteration 5 (**Figure 8**). The kICA algorithm is prone to outliers and thus fails to converge to an effective solution. However, restarting the kICA each time might improve the results. Therefore, for large images, the efficiency of the nICA method is superior to the kICA, with about 10 times faster convergence and more accurate results.

Another influencing factor with the same level of added unpredictable noise is the sampling interval. This is crucially important in the geophysical sense since the sampling interval directly influences the geophysical survey budget. Large regional-scale data sets are sampled in smaller intervals and give more details about geological structures. As is shown in **Figure 9**, increasing the number of samples reduces

Figure 8.

Performance of kICA algorithm (a) compared to nICA algorithm (b). The kICA results converge at 10 iterations while the nICA results are converged at iteration 5. The nICA method is superior to the kICA, with about two times faster convergence and more accurate results (180×180 pixels).

Figure 9.

Performance of kICA algorithm (a) compared to nICA algorithm (b) on large images with 720×720 pixels. For the kICA algorithm, a 16 times increase of pixels results in eight times slower convergence with numerous local minima. Therefore, for larger images with smaller sampling intervals, negentropy maximization is superior to kurtosis maximization.

the performance of kICA compared to nICA and indicates that for large datasets with smaller sampling intervals, negentropy maximization is superior to kurtosis maximization.

However, there are two inherent ambiguities in the ICA framework: (1) Ambiguity in permutations of the recovered sources. This problem states that the order of the estimated independent components is unspecified. (2) Ambiguity in the recovered amplitudes of the sources. One can partially solve this problem by assuming unit variance for all sources. However, there is still a 50% chance that the polarities of the sources are not determined correctly. In cases where we have enough prior information, we can multiply sources by -1 to achieve the best results.

The effect of mixing on image segmentation is also explored in **Figure 10**. ICA significantly reduces the impact of feature overlaps and enhances the performance of image segmentation. The proposed approach integrates information from multiple geo-images and makes sure that the combined images are maximally independent and unique. In other words, the feature overlaps are minimal. This has an important implication in image segmentation since a slight presence of image overlaps and artifacts distort the segmentation output. As shown in **Figure 10**, the segmentation after negentropy maximization improves the human face detection.

5.2 Spectral feature extraction

Further feature extraction can be performed on the images to extract spectral features like edges and lineaments inside the photos. **Figure 11a** shows the results of CWT with Cauchy mother wavelet and 20 scales on the second independent component (recovered human face). As can be seen, different spectral features appear in different CWT scales/directions. Low-level features like edges appear in higher frequencies, and high-level features appear in lower frequencies. However, the transition from each scale/direction is relatively smooth, and that allows some features from each scale/direction leak into the following scale/direction, creating unwanted spectral features' overlap. To reduce these effects, we perform source separation methods like PCA and ICA algorithms again. The results are shown in **Figure 11b**. As is shown, spectral source separation with nICA has improved the low/high-level features with crisper edges and more visible boundaries.

Figure 10.
Effect of ICA as a preprocessing step before grayscale image segmentation. (a) Segmentation on the mixture image (g_2 in **Figure 6**). (b) Segmentation after negentropy maximization nICA ($\hat{f}_2$ in **Figure 6**).

Figure 11.

(a) Space-frequency representation with Cauchy mother wavelet in four scales on the second independent component (recovered human face). Different spectral features appear in different CWT scales/directions. (b) Separation of spectral features with nICA. The results are crisper and have more distinct features.

6. Applications

6.1 Integrated interpretation of geophysical data

6.1.1 Fast pseudo-geological mapping

Geophysical surveys are essential for fast geological mapping in mineral exploration programs. Nevertheless, there is not necessarily a one-to-one correlation between geophysical responses and contrasts in lithotypes because the rock-forming minerals do not entirely control geophysical properties. For example, in magnetic surveys, most magnetic anomalies are only related to the distribution of rock-forming magnetite inside the rocks. Therefore, a “geological” map created using geophysical data sets should be referred to as a “pseudo-geological” map [16]. However, the integrated interpretation of multiple data types, for example, magnetic, radiometric, conductivity, and so on, helps reduce uncertainty and acquire a more reliable geological model. SFE2D provides such an integration tool for explorations. A geological map of the exemplified region indicates a very complex geological setting along with complex structural faults and fractures (Figure 12).

The geophysical survey of the area is part of a regional program that aims to complete the geological mapping of the southwestern part of James Bay at a scale of 1: 50,000. The region’s rocks are mainly of Archean age and belong to the Superior Province, which forms the heart of the Canadian Shield, one of the largest Precambrian cratons exposed from the terrestrial globe. There are also a number of dykes of Neoproterozoic to Paleoproterozoic diabase [18]. The region is subdivided into five sub-provinces, and the boundaries between the different sub-provinces are based on lithological contrasts, metamorphic, structural, and geophysical information.

However, the geological compilation is subject to varying interpretations due to intrinsic uncertainties in geoscientific data sets. This section provides an example of feature extraction for integrated geophysical interpretation through a PCA and ICA-based color image segmentation by k -means clustering on three sources of geophysical data sets.

Figure 12.
Geology of the region. (a) Geological sub-provinces and structural features. (b) General geological map of the region. Geological data are compiled from SIGEOM [17].

GGMplus Bouguer gravity data sets [19], high-resolution airborne magnetic data sets, and digital elevation model exerted from SIGEOM [17] are used as inputs for integrated feature extraction and segmentation (**Figure 13**).

The program can perform an image segmentation based on the k -means algorithm over the RGB-merged independent components. Unlike traditional segmentation methods, the integrated k -means segmentation algorithm helps calculate segments based on principal or independent components of the original data sets. The results are pseudo-geological maps that integrate extracted features from the three sources of data sets (**Figure 14a and b**).

6.1.2 Spectral feature extraction

CWT is performed on the magnetic data sets (**Figure 13c**). Decomposing the magnetic data sets with CWT forms raw spectral features that eventually increase dimensionality. 2D wavelet coefficients are calculated here for eight scales with Mexican Hat mother wavelet, containing several frequency-dependent raw features (**Figure 15**). However, the algorithm allows choosing any number of scales and directions (in the case of anisotropic mother wavelets), as long as the computer hardware specifications allow.

As shown in **Figure 15**, spectral decomposition reveals many latent features that are not properly visible in the spatial domain. The major difficulty arises in extracting and selecting the spectral features due to the statistical interdependence of features in various spectra. Spectral decomposition of images also overloads the computational cost of interpretation and demands a methodology to pick the best spectral representation of images. The SFE2D program tackles this problem with the

Figure 13. Three input images. (a) Digital elevation model. (b) GGMplus gravity anomalies. (c) Aeromagnetic data sets.

Figure 14. Color image segmentation after variance and negentropy maximization and RGB image compilation through k -means clustering with eight segments. (a) PCA segmentation with variance maximization. (b) ICA segmentation with negentropy maximization.

Figure 15.
CWT with Mexican Hat mother wavelet with eight scales over magnetic data sets.

spectral independent component analysis. We used Fast-ICA through negentropy maximization to separate the eight raw wavelet features to produce the spectral inputs necessary for interpretation. The Fast-ICA algorithm can reduce the dimensionality of the raw features. This is important when we produce a large number of raw features through CWT, and as a result, our computational hardware resources are not enough to handle the inevitable high dimensionality.

We reduced the dimensionality of the raw spectral features and thus produced three independent features (**Figure 16**). As shown in **Figure 16**, the program uncovered several low-frequency features in higher wavelet scales related to deep sources and high-frequency features related to shallow sources.

2D wavelet coefficients are also calculated with Cauchy mother wavelet on eight scales and eight directions, producing several directional/frequency-dependent raw features (**Figure 17**). As can be seen, most of the wavelet spectral content is redundant, and similar features are repeated in 64 subsequent directions/scales. Three independent spectral components are extracted through spectral feature extraction and dimensionality reduction (**Figure 18a–c**). Therefore, the 64-dimensional hyperspace is reduced to a 3D RGB space to facilitate visual interpretations

Figure 16.
Spectral feature extraction and dimensionality reduction by negentropy maximization on CWT results with Mexican Hat mother wavelet. (a–c) Extracted independent spectral features.

Figure 17.
CWT with Cauchy mother wavelet on eight scales and eight directions over magnetic data sets.

(**Figure 18d**). As can be seen, the process helped to uncover several hidden lineaments in the NE–SW direction.

6.2 Microscopic image analysis

Optical microscopy with polarized light microscopes has become one of the most powerful techniques in petrography and mineral exploration. However, due to the complexity of mineral assemblages, acquiring quality images with details of mineral grains and fine textures is a challenging task. Here we present an application of the

Figure 18.
Spectral negentropy maximization ICA. (a–c) Independent spectral features. (d) RGB image compilation. (e) Extracted lineaments (dashed red lines).

Figure 19.
(a) An example of petrographic thin section imagery used in this study for mineralogical feature extraction. (b–d) Independent components of RGB bands combined to form different representations for enhanced visualization. The thin section image is exerted from Carleton NAGTWorkshops [20].

SFE2D program in the micromorphological characterization of minerals under thin section. An example of an optical microscopic image is presented in **Figure 19a**. The SFE2D program can separate mineral zones on a highly mixed texture.

The algorithm starts with balancing the colors on the original image based on the normal distribution. Then, source separation on color bands helps extract hidden features in principal or independent components. The RGB compilation of independent components of the image is also shown in **Figure 19b–d**. As can be seen, different polarities offer different feature representations.

The program also offers a fast detection of the region of interest (ROI) with a color-pick algorithm based on few clicks on specified color zones. As can be seen in **Figure 20**, subtle mineralogical zones are optimally extracted for geological interpretations. This eventually helps calculate the percentage of specified minerals on thin sections that are very important for mineral exploration studies.

Figure 20.
Regions of interest delineated by the color-pick algorithm in the SFE2D program. (a) First extract mineral. (b) Second extract mineral. (c) Third extracted mineral.

7. Conclusions

The result presented here illustrates the theory, design, performance, and applications of a standalone mathematical tool (SFE2D) for 2D spatial and spectral feature extraction, based on PCA, ICA, CWT, k -means clustering segmentation, and image processing algorithms. SFE2D provides an integration tool for the interpretation of multiple geoscientific data sets at once. SFE2D has straightforward applications in various geophysical and geoscientific explorations where one needs to add value to observed geo-images by recovering hidden features in hyperdimensional data sets.

The program can perform an image segmentation based on the k -means algorithm over the RGB-merged independent/principal components. The results are pseudo-geological maps that integrate extracted features from multiple data sets. Unlike traditional segmentation methods, the integrated k -means segmentation algorithm helps calculate segments based on principal or independent components of the original data sets. The proposed approach integrates information from multiple geophysical data sets and makes sure that the combined images are maximally independent and unique. In other words, the feature overlaps are minimal. This has an important implication in image segmentation since a slight presence of image overlaps and artifacts distort the segmentation output.

Spectral decomposition of images also provides a unique way for feature extraction in the frequency domain. Deploying the SFE2D algorithms helps eliminate redundant frequency volumes and reduce them to a more manageable number of components. Taking advantage of the ICA statistical properties, we can keep the most geologically pertinent information within the spectral decomposed data. The feature extraction algorithms in SFE2D can also be used in deep learning applications where feature extraction is a primary step in optimizing neural network design.

The SFE2D program can also be used in the micromorphological characterization of minerals under thin sections. The program offers a fast detection of the region of interest (ROI) with a color-pick algorithm based on few clicks on specified color zones.

The SFE2D application is straightforward and for immediate use, as long as users already installed the MATLAB runtime library on their computer.

Acknowledgements

This study is funded by FRQNT (Fonds de recherche, Nature et technologies du Québec) and MERNQ (ministère de l'Énergie et des Ressources naturelles du Québec).

Appendix A: SFE2D program environment

The SFE2D is provided as a standalone executable program. To use the executable program (SFE2D.exe), users do not need any previously installed Matlab software on the PC. The only prerequisite is installing the latest Matlab 2021a Runtime library, free to download from MathWorks [21]. To do so, users can install the required runtime codes by double-clicking on the offline program installer "Installer_SFE2D (Offline).exe" or the online version of it "Installer_SFE2D (Online).exe" that is a lightweight version in which the program and the runtime codes are going to be installed automatically through downloading from

MathWorks server. Users can also directly download the runtime from the MathWorks website [21].

After unzipping the file and installing the runtime, the program should work by simply double click on the SFE2D.exe. It is recommended to copy/paste the SFE2D.exe to a project folder where data sets are located with read/write permission. Running the program for the first time needs activation. To activate the program, one needs to double click on the “Activate.exe” file to create a “Key” file. A dialog box (Password Required) appears that asks to insert the password provided for the user. As soon as typing it and pressing Enter, a “Key” file is going to be generated. The user needs to keep that file and save it in the same project folder where “SFE2D.exe” is copied. When running the SFE2D, the program interface appears as in Figure A1. On the right side, the user can control the program parameters, and on the left, the progress of calculations and possible errors can be tracked with the windows command shell console for execution (black screen on the left).

Figure A1.
SFE2D program interface.

IntechOpen

IntechOpen

Author details

Bahman Abbassi* and Li Zhen Cheng

The Research Institute on Mines and Environment (RIME) UQAT-Polytechnique,
University of Quebec in Abitibi-Témiscamingue, Rouyn-Noranda, Quebec, Canada

*Address all correspondence to: bahman.abbassi@uqat.ca

IntechOpen

© 2021 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

References

- [1] Barnett C, Williams P. Mineral exploration using modern data mining techniques. *First Break*. 2006;**24**(7): 295-310
- [2] Murtaugh F, Starck JL, Berry MW. Overcoming the curse of dimensionality in clustering by means of the wavelet transform. *Computer Journal*. 2000; **43**(2):107-120
- [3] Brown AR. Seismic attributes and their classification. *The Leading Edge*. 1996;**15**:10
- [4] Chen Q, Sidney S. Seismic attribute technology for reservoir forecasting and monitoring. *The Leading Edge*. 1997;**16**: 445-456
- [5] Chopra S, Pruden D. Multiattribute seismic analysis on AVO-derived parameters. *The Leading Edge*. 2003;**22**: 998-1002
- [6] Lindseth RO. Seismic attributes—Some recollections. *Canadian Society of Exploration Geophysicists Recorder*. 2005;**30**(3):16-17
- [7] Chopra S, Marfurt KJ. Emerging and future trends in seismic attributes. *Geophysics*. 2008;**27**:298-318
- [8] Abbassi B. Integrated imaging through 3D geophysical inversion, multivariate feature extraction, and spectral feature selection [thesis]. Université du Québec à Montréal and Université du Québec en Abitibi-Témiscamingue; 2018
- [9] Hyvärinen A, Oja E. Independent component analysis: Algorithms and applications. *Neural Networks*. 2000;**13**: 411-430
- [10] Moreau F, Gibert D, Holschneider M, Saracco G. Wavelet analysis of potential fields. *Inverse Problems*. 1997;**13**:165-178
- [11] Honório BCZ, Sanchetta AC, Leite EP, Vidal AC. Independent component spectral analysis. *Interpretation*. 2014;**2**(1):SA21-SA29
- [12] Trauth MH. MATLAB® Recipes for Earth Sciences. 4th ed. Berlin: Springer-Verlag; 2015. 436p. DOI: 10.1007/978-3-662-46244-7
- [13] Moreau F. Méthodes de traitement de données géophysiques par transformée en ondelettes [thesis]. Rennes, France: Université de Rennes I; 1995
- [14] Moreau F, Gibert D, Holschneider M, Saracco G. Identification of sources of potential fields with continuous wavelet transform: Basic theory. *Journal of Geophysical Research*. 1999;**104**: 5003-5013
- [15] Olukanmi P, Nelwamondo F, Marwala T. Rethinking k-means clustering in the age of massive datasets: a constant-time approach. *Neural Computing & Applications*. 2020;**32**: 15445-15467
- [16] Dentith M, Mudge ST. *Geophysics for the Mineral Exploration Geoscientist*. Cambridge: Cambridge University Press; 2014. 516p. DOI: 10.1017/CBO9781139024358
- [17] SIGEOM. Système d'information géominière of Québec [Internet]. 2021. Available from: <https://sigeom.mines.gouv.qc.ca> [Accessed: 01 August 2021]
- [18] Card KD, Ciesielski A. DNAG #1 subdivisions of the superior province of the Canadian shield. *Geoscience Canada*. 1986;**13**:5-13
- [19] Hirt C, Claessens S, Fecher T, Kuhn M, Pail R, Rexer M. New ultrahigh-resolution picture of Earth's gravity field. *Geophysical Research Letters*. 2013;**40**:4279-4283

[20] Carleton NAGTWorkshops. Optical mineralogy and petrography [Internet]. 2021. Available from: https://serc.carleton.edu/NAGTWorkshops/mineralogy/optical_mineralogy_petrography.html [Accessed: 01 August 2021]

[21] MathWorks [Internet]. 2021. Available from: [mathworks.com](https://www.mathworks.com) [Accessed: 01 August 2021]