Sensor Review
An effective hyperspectral image retrieval method using integrated spectral and textural features
Zhenfeng Shao, Weixun Zhou, Qimin Cheng, Chunyuan Diao, Lei Zhang.

Article information:
To cite this document:
Permanent link to this document:
https://doi.org/10.1108/SR-10-2014-0716

Downloaded on: 15 January 2019, At: 01:08 (PT)
References: this document contains references to 25 other documents.
To copy this document: permissions@emeraldinsight.com
The fulltext of this document has been downloaded 261 times since 2015*

Users who downloaded this article also downloaded:

Access to this document was granted through an Emerald subscription provided by emerald-srm:155010 []

For Authors
If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service
information about how to choose which publication to write for and submission guidelines are available for all. Please visit
www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com
Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of
more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online
products and additional customer resources and services.
Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics
(COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

*Related content and download information correct at time of download.
An effective hyperspectral image retrieval method using integrated spectral and textural features

Zhenfeng Shao and Weixun Zhou
State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing,
Wuhan University, Wuhan, China
Qimin Cheng
Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan, China
Chunyuan Diao
Department of Geography, University at Buffalo, The State University of New York, New York, USA, and
Lei Zhang
State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing,
Wuhan University, Wuhan, China

Abstract
Purpose – The purpose of this paper is to improve the retrieval results of hyperspectral image by integrating both spectral and textural features. For this purpose, an improved multiscale opponent representation for hyperspectral texture is proposed to represent the spatial information of the hyperspectral scene.

Design/methodology/approach – In the presented approach, end-member signatures are extracted as spectral features by means of the widely used end-member induction algorithm N-FINDR, and the improved multiscale opponent representation is extracted from the first three principal components of the hyperspectral data based on Gabor filters. Then, the combination similarity between query image and other images in the database is calculated, and the first k more similar images are returned in descending order of the combination similarity.

Findings – Some experiments are calculated using the airborne hyperspectral data of Washington DC Mall. According to the experimental results, the proposed method improves the retrieval results, especially for image categories that have regular textural structures.

Originality/value – The paper presents an effective retrieval method for hyperspectral images.

Keywords Image processing, Spectral analysis

Paper type Research paper

1. Introduction

With the increasing availability of the huge volume of Earth Observation data, remote sensing image retrieval technology has drawn more and more public attention in recent years. Traditionally, metadata, such as sensor type, data acquisition time and geographical position, are commonly used for information retrieval from remote sensing databases (Ferecatu and Boujemaa, 2007). Unfortunately, metadata cannot describe the abundant visual information contained in remote sensing images well. In most applications, people are more concerned with the content of remote sensing images rather than corresponding metadata. In addition, metadata-based retrieval usually requires professional knowledge, which has impeded the development and application of remote sensing image retrieval. In this respect, content-based remote sensing image retrieval (CBRSIR) method is a more desired choice. Ozkan et al. (2014) compared three CBRSIR representation methods and gave a performance analysis with some experiments on the publicly available UC Merced dataset.

For CBRSIR method, images are characterized by low-level visual features, which can be extracted using feature extraction algorithms, and images are associated with corresponding feature vectors. Given one query image, relevant images will then be retrieved and returned in the order of ascending or descending similarity calculated through predefined similarity measures. In image retrieval community, CBRSIR methods are mostly exploited on high-resolution images (Yang and Newsam, 2013; Wang et al., 2014) and synthetic aperture radar (SAR) images (Espinoza-Molina and Datcu, 2013).
Despite the important progress of CBRSIR, few works have been done on hyperspectral images. In recent years, the Computational Intelligence Group of the University of Basque Country has conducted some works on hyperspectral image retrieval (Grana and Veganzones, 2012; Veganzones and Grana, 2012). Grana and Veganzones (2012) extracted end-members as spectral features by end-member induction algorithms and defined an end-member-based image distance to measure the similarity between two hyperspectral images. Both real and synthetic hyperspectral data were used to validate the proposed hyperspectral image retrieval algorithm. It is shown that the proposed algorithm, based on spectral features, is effective in retrieving hyperspectral images. However, the method takes no consideration of the spatial distribution of end-members. To avoid such deficiency and improve retrieval results, Veganzones and Grana (2012) proposed a spectral–spatial retrieval system, which estimated and introduced the fractional abundances of end-members in similarity measure. Plaza et al. (2007) proposed a parallel system for efficient hyperspectral image retrieval.

Though some differences exist among above works, they share similar ideas that end-member signatures extracted from hyperspectral images based on linear mixing model are taken as spectral features in the retrieval system, and the similarity between two images is measured through calculating the distance between corresponding spectral features in a feature space. Nevertheless, due to the phenomena that the same object can have very different spectra, and totally different objects may share similar spectra, some hyperspectral images are still difficult to retrieve, even with high spectral resolution. In hyperspectral image processing techniques, spectral and spatial information are complementary. In some cases, it is even mandatory to integrate them together for some certain applications (Plaza et al., 2009).

To improve the performance of hyperspectral image retrieval, we propose a hyperspectral image retrieval method that integrates spectral and textural features. Specifically, spectral features are extracted by end-member induction algorithm, and textural features are computed from the principal component analysis (PCA)–transformed image with Gabor filters.

### 2. Feature extraction and similarity measure

In this section, the framework of the proposed retrieval method is illustrated first, and details of feature extraction as well as similarity measures used are then introduced.

#### 2.1 Framework of the proposed method

Figure 1 illustrates the framework of the proposed hyperspectral image retrieval method that mainly includes two parts, feature extraction and similarity measure. Both of these are indispensable functional modules of one CBRSIR system. For feature extraction, spectral and textural features are extracted by means of N-FINDR and Gabor filters, respectively, and two feature databases, including a spectral feature database and a textural feature database, are built. Consequently, each image can be represented by the corresponding features stored in the feature databases through indexing mechanism. For similarity measure, the process can be described briefly as follows:

- Calculate the spectrum similarity and the texture similarity between the query image and images in the database through predefined similarity measures.
- Combine the spectrum similarity and the texture similarity, and obtain the combination similarity between these two images.
- Rank the images with descending combination similarity, and return the first k retrieved images.

#### 2.2 Feature extraction procedure

##### 2.2.1 Spectral features

The linear mixing model is widely used for hyperspectral data analysis (Keshava and Mustard, 2002). The core of this model is based on the premise that within a specific hyperspectral scene, the reflectance of each pixel can be described as the linear combination of endmember spectra, and it can be extended to the whole image and denoted by:

$$ I = E\Phi + W $$

(1)

where \( I \) is the hyperspectral image with \( q \) spectral bands, \( E = [e_1, e_2, \ldots, e_n] \) is a \( q \times n \) matrix standing for the end-member signatures of image \( I \) and \( e_i(i = 1, 2, \ldots, n) \) is one of the extracted end-member signatures of image \( I \). \( \Phi = [\psi_1, \psi_2, \ldots, \psi_n] \) is the fractional abundance matrix and \( \psi_i(i = 1, 2, \ldots, n) \) is the fractional abundance vector of end-member \( e_i \). Note that \( \Phi \) is subject to two constraints: each element of \( \Phi \) is greater than 0, and the sum of row elements of \( \Phi \) equals 1. \( W \) is the additive noise.

Figure 2(a and b) includes two hyperspectral image patches with the same size, and both have two land cover classes A and B. Assuming the two image patches happen to have the same end-member signatures, then they will be considered very similar through the pre-defined similarity measure. However, in fact, they have low similarity because of the totally different proportions of A and B in each image patch. To overcome such deficiency, the hyperspectral image \( I \) can be characterized by the tuple \((E, \Phi)\), which takes the spatial
Figure 2 Two hyperspectral image patches with the same land cover classes but totally different proportions.

distribution of end-members into consideration, thus improving the retrieval results (Veganzones and Grana, 2012).

Then, the focus of spectral features extraction turns to the extraction of end-member signatures. Many methods have been developed for end-member determination in practice. Veganzones and Grana (2008) gave a comprehensive enumeration of commonly used end-member determination methods, such as N-FINDR (Winter, 1999), Pixel Purity Index (Boardman et al., 1995), Convex Cone Analysis (Ifarraguerri and Chang, 1999) and Automated Morphological Endmember Extraction (Plaza et al., 2002). In the work by Plaza et al. (2004), the end-member extraction methods mentioned above and some other methods are rigorously compared by a custom-designed quantitative and comparative framework. However, it is unlikely to find an algorithm that is appropriate for all the situations in end-member extraction. In this article, N-FINDR is used for end-member extraction.

Despite N-FINDR being widely used in various domains, it cannot determine the number of needed end-members (Plaza and Chang, 2005). Here, we provide two strategies to determine the number of induced end-members. The first strategy is to manually assign the number of induced end-members, and the second strategy is to estimate the number of end-members using virtual dimensionality (VD) algorithm (Chang and Du, 2004). We use the package provided by Veganzones and Grana (2012) to implement N-FINDR algorithm based on the two strategies.

In view of the situation that the number of identified and used end-members in practice is ranging from three to seven (Keshava and Mustard, 2002) and the hyperspectral image data used in our experiments contains seven ground cover types, we assign the number of end-members ranging from 3 to 7 for the first strategy.

2.2.2 Textural features

There are various texture representations with respect to texture feature analysis, such as Gabor filter (Manjunath and Ma, 1996), Wavelet Transform (Chang and Kuo, 1993), Local Binary Pattern (LBP) (Ojala et al., 2002) and global morphological texture (Aptoula, 2014). While these texture features tend to achieve ideal performance in image retrieval literature, they are generally designed for processing images with a single band, which are not suitable for multispectral or hyperspectral images directly. Though applying texture extraction methods to each band of the hyperspectral image may be one solution, it suffers from high computational complexity and inefficiency. In addition, each band of the hyperspectral data exhibits nearly the same scene structure, and the information will be redundant if all bands are taken into consideration.

Shi and Healey (2003) proposed a multiscale opponent representation for hyperspectral texture recognition, which is based on Gabor filters and combines spatial information across spectral bands. Motivated by this research, we proposed an improved color texture descriptor for aerial image retrieval in our previous work (Shao et al., 2014). In this paper, the improved color texture descriptor was adapted for hyperspectral texture extraction.

Due to the high dimensionality of hyperspectral data, we present a PCA-based textural feature extraction strategy with lower computational complexity and higher efficiency for hyperspectral data. The following steps illustrate the brief procedures of the proposed strategy.

- Apply PCA to reduce the dimensionality of the whole hyperspectral scene to acquire the first three principal components.
- Partition the principal components into image patches of 32 × 32 pixels size and prepare the retrieval dataset (details can be found in Section 3).
- Extract Gabor opponent (GaborOpp, which is named CGOT in our previous work mentioned above) texture of each image patch and represent each image patch using a GaborOpp feature vector.

Considering that GaborOpp textural feature is fused by Gabor kernels and combines spatial information across spectral bands. Motivated by this research, we proposed an improved color texture descriptor for aerial image retrieval in our previous work (Shao et al., 2014). In this paper, the improved color texture descriptor was adapted for hyperspectral texture extraction.

Hyperspectral image retrieval method

Zhenfeng Shao, Weixun Zhou, Qimin Cheng, Chunyuan Diao and Lei Zhang

Sensor Review

Volume 35 · Number 3 · 2015 · 274–281

Downloaded by Wuhan University At 01:08 15 January 2019 (PT)
The Gabor texture of $P$ is characterized by $T = \{\mu_{uv}, \sigma_{uv}\}$, which is one $80 \times 3$ matrix and $i$th column represents the textural feature vector of $i$th principal component.

Let $g_{uv}(x,y)$ and $g_{uv}(x,y)$ be the convolution results of different principal components $P_i$ and $P_j$, $i,j = 1, 2, 3; i \neq j$, respectively, of image patch $P$ using equation (4), and the normalized difference between $g_{uv}(x,y)$ and $g_{uv}(x,y)$ can be defined by:

$$d_{g_{uv}}(x,y) = \frac{g_{uv}(x,y)}{\sqrt{\sum g_{uv}^2(x,y)}} - \frac{g_{uv}(x,y)}{\sqrt{\sum g_{uv}^2(x,y)}}$$

where $u$ and $v$ denote the orientation and scales of Gabor filters used, respectively. Note that $v$ and $v'$ are adjacent scales, which means they meet the condition $|v - v'| \leq 1$.

Thus, 13 scale groups and 8 orientations are available. In terms of scale groups, they are: $(v, v') \in \{(0,0), (1,1), (2,2), (3,3), (4,4), (0,1), (1,0), (1,2), (1,1), (2,3), (3,2), (3,4), (4,3))$. The opponent feature of $P$ can be characterized by $opp = \{\sqrt{\sum d_{g_{uv}}(x,y)}\}$, which is a total of 312 feature vectors. Then GaborOpp texture is defined by:

$$Gabor\text{Opp} = \{\mu_{uv}, \sigma_{uv} \sqrt{\sum d_{g_{uv}}^2}\}$$

As one of the comparative features, the multiscale opponent representation (UniOpp) used for hyperspectral texture recognition by [Shi and Healey, 2003] can be characterized as:

$$Uni\text{Opp} = \{\sum g_{uv}(x,y), \sqrt{\sum d_{g_{uv}}^2(x,y)}\}$$

2.2.3 Similarity measures
Effective image similarity measures are requisite for accurate image retrieval. For spectral features, the spectral–spatial dissimilarity measure is used to calculate the dissimilarity between two image patches (Veganzones and Grana, 2012). Let $E_1 = [e_{11}, e_{12}, \ldots, e_{1n}]$ and $E_2 = [e_{21}, e_{22}, \ldots, e_{2n}]$ be the end-member signatures of image $H_1$ and $H_2$, respectively, where $m$ and $n$ are the number of end-members. Define $D_{ij} = [d_{ij}, i = 1, 2, \ldots, m; j = 1, 2, \ldots, n]$ as the end-member distance matrix, and $d_{ij}$ as the distance between the end-member pair from $E_i$ and $E_j$. In this article, $d_{ij}$ is calculated by means of spectral angle mapper, one common distance metric in hyperspectral image processing as follows:

$$d_{ij} = \arccos\left(\frac{e_i^T e_j}{||e_i||^2}||e_j||^2\right)$$

where $e_i$ and $e_j$ are the end-member vectors of $E_i$ and $E_j$, respectively. The spectrum dissimilarity between $H_1$ and $H_2$ is given by:

$$D^S(H_1, H_2) = \sum_{i,j} \sum_{ij} r_{ij} d_{ij}$$

where $d_{ij}$ is the spectral distance between two end-members from $E_i$ and $E_j$, and $r_{ij}$ is the significance matrix associated with $d_{ij}$. The readers can refer to the work by Veganzones and Grana (2012) for more details about the computation of the significance matrix $r_{ij}$.

For GaborOpp textural feature, the texture dissimilarity between $H_1$ and $H_2$ is given by:

$$D^T(H_1, H_2) = \sum_{i=1}^{p} \left|\frac{f_i - f_j}{\sigma_i}\right|$$

where $f_i(i = 1, 2, \ldots, p)$ and $f_j(i = 1, 2, \ldots, p)$ are corresponding elements of GaborOpp texture vector of image $H_1$ and $H_2$, respectively, and $\sigma_i(i = 1, 2, \ldots, p)$ is the standard deviation of GaborOpp texture over the entire image database.

The combination dissimilarity $D(h_1, h_2)$ between $H_1$ and $H_2$ is defined by:

$$D(h_1, h_2) = \alpha D^S(h_1, h_2) + \beta D^T(h_1, h_2)$$

where $D^S(h_1, h_2)$ and $D^T(h_1, h_2)$ are spectrum dissimilarity and texture dissimilarity, respectively. $\alpha$ and $\beta$ are the weight of spectrum dissimilarity and texture dissimilarity in the combination similarity, respectively, and can be determined by:

$$\left\{\begin{array}{l}
\alpha + \beta = 1 \\
\alpha - \beta = D^S(h_1, h_2) - D^T(h_1, h_2)
\end{array}\right.$$
the strategy used to determine the number of needed end-members is an important factor that affects the spectral retrieval result. To reduce this impact and improve the performance, an improved strategy (IS) is proposed and briefly described as follows. For the categories Grass and Tree, VD is used to estimate the number of induced end-members. For the categories Mixture, Water, Others, Building and Road, the number of induced end-members is assigned 3, 4, 5, 6 and 7, respectively. Estimating the number of needed end-members is still an open question in spectral unmixing literature, but it is not the focus of our work in this paper. Therefore, we call IS “strategy” rather than “method” here, which means IS is introduced to decrease the influence of the number of end-members on the retrieval performance to some extent, and it is not appropriate for other applications as VD algorithm does.

The average precision of each category using above strategies to determine the number of induced end-members is shown in Figure 4. It is apparent that IS improves the retrieval performance of the categories Mixture, Water, Others, Building and Road, the number of induced end-members is assigned 3, 4, 5, 6 and 7, respectively. Estimating the number of needed end-members is still an open question in spectral unmixing literature, but it is not the focus of our work in this paper. Therefore, we call IS “strategy” rather than “method” here, which means IS is introduced to decrease the influence of the number of end-members on the retrieval performance to some extent, and it is not appropriate for other applications as VD algorithm does.

In subsequent experiments, IS is applied to determine the number of induced end-members.

3.2 Experiment II

Despite the good performance of spectral features for some image classes, the overall performance measured by average precision is not satisfying. In addition, those image categories that have regular textural structures, such as Grass, Road and Mixture, have not achieved better performance as expected. From this respect, the image retrieval system should benefit from the incorporation of textural information.

Table II summarizes the performance of GaborOpp texture against three comparative features. For Gabor texture and UniOpp texture, Section 2.2.2 shows the detailed process of feature extraction. For LBP feature, we compute $LBP_{8,1}^{riu}$, which means LBP has been implemented using 8 pixels circular neighbor of radius 1 in our experiments. According to the results in Table II, GaborOpp has better performance on the hyperspectral dataset than the other three texture features, indicating the powerful representation performance of GaborOpp texture.

Table III demonstrates the advantage of hyperspectral image retrieval using integrated spectral and GaborOpp features. For the categories Grass, Road and Mixture, integrated features improve the performance dramatically compared with spectral features. For the categories Others, Tree, Water and Building, which have no obviously regular

Table I Retrieval precision of each category using spectral features with two strategies to determine the number of induced end-members

<table>
<thead>
<tr>
<th>Category</th>
<th>VD</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grass</td>
<td>0.5533</td>
<td>0.3187</td>
<td>0.5278</td>
<td>0.4952</td>
<td>0.4925</td>
<td>0.5071</td>
</tr>
<tr>
<td>Road</td>
<td>0.3827</td>
<td>0.2724</td>
<td>0.3223</td>
<td>0.3393</td>
<td>0.3608</td>
<td>0.4137</td>
</tr>
<tr>
<td>Mixture</td>
<td>0.3444</td>
<td>0.3860</td>
<td>0.3201</td>
<td>0.3819</td>
<td>0.2849</td>
<td>0.3191</td>
</tr>
<tr>
<td>Others</td>
<td>0.2511</td>
<td>0.2393</td>
<td>0.2877</td>
<td>0.2997</td>
<td>0.2862</td>
<td>0.2558</td>
</tr>
<tr>
<td>Tree</td>
<td>0.4426</td>
<td>0.3115</td>
<td>0.4214</td>
<td>0.3732</td>
<td>0.3612</td>
<td>0.3702</td>
</tr>
<tr>
<td>Water</td>
<td>0.7115</td>
<td>0.7462</td>
<td>0.8149</td>
<td>0.7716</td>
<td>0.7387</td>
<td>0.7208</td>
</tr>
<tr>
<td>Building</td>
<td>0.6194</td>
<td>0.5202</td>
<td>0.5514</td>
<td>0.5938</td>
<td>0.6871</td>
<td>0.6368</td>
</tr>
</tbody>
</table>

textural structures, integrated features have similar performance as spectral features indicating that textural features are valuable information for hyperspectral image retrieval, particularly for those categories having regular textural structures.

The precision-recall curve is used to further evaluate the performance of integrated features. Postulating a query that searches for the first k more similar images to the query image, then precision is defined as the fraction of relevant images in k retrieved images and recall is defined as the fraction of relevant images that are retrieved in the entire image database. The precision-recall curve for spectral features, textural features and integrated features is shown in Figure 5. The plots demonstrate that the integrated features outperform spectral features and textural features. In terms of textural features, the improved color texture descriptor GaborOpp has better performance compared to the other three textural features. However, when recall is greater than 0.45, the curve of integrated features decreases quickly. Two reasons can be used to explain this phenomenon. On one hand, retrieving the first k more relevant images becomes more difficult when the query scope k increases, and the complexity of the image scenes compromises the performance. On the other hand, the number of images of each category is uneven. So as the query scope k increases, the proportion of relevant images in the returned images is down rapidly for the categories of small size, and is down slowly for the categories of big size.

The results shown in Figure 5 not only demonstrate the advantages of the integrated features for hyperspectral image retrieval but also provide the insight that the number of images of each category should be as even as possible. In addition, the idea that textural features incorporating discriminative information among color bands have better representation performance may be informative for other applications such as texture recognition and classification.
4. Conclusion and future work

We present an effective retrieval method that integrates spectral and textural features for hyperspectral images. For spectral features, we introduce three strategies to estimate the number of needed end-members when using N-FINDR to extract the end-members. For textural features, we propose an improved color texture descriptor that has been demonstrated to be effective for hyperspectral image retrieval. The experiments show that integrated features outperform spectral features measured by several performance metrics. In addition, the experiments also indicate that textural features can be used as complementary information for hyperspectral image retrieval, particularly for image categories that have regular textural structures.

In future, a more powerful end-member determination strategy allowing for different image categories will be studied, and integration of semantic features to bridge the semantic gap will be considered as well.

References


Hyperspectral image retrieval method
Zhenfeng Shao, Weixun Zhou, Qimin Cheng, Chunyuan Diao and Lei Zhang


Corresponding author
Weixun Zhou can be contacted at: weixunzhou1990@whu.edu.cn
This article has been cited by:


3. Jing Zhang, Wenhao Geng, Xi Liang, Jiafeng Li, Li Zhuo, Qianlan Zhou. 2017. Hyperspectral remote sensing image retrieval system using spectral and texture features. *Applied Optics* **56**:16, 4785. [Crossref]