SATellite ImagE RetrIEval: AN APPROACH BASED ON TEXTURE FEATURES

José Luiz de Souza Pio, Camillo Jorge Santos Oliveira and Mario Fernando Montenegro Campos

Federal University of Minas Gerais (UFMG) VerLab – Computer Vision and Robotics Laboratory - Av. Antonio Carlos, 6627 – Pampulha – Belo Horizonte – MG, Brazil.

1. INTRODUCTION

The popularization of the Internet and the growing of large image databases has motivated the research activity in the area of content-based image retrieval. The problem of content-based image retrieval requires semantic interpretation and an approach for this is the computation of visual features, like colors, contours and textures that can be used as quantitative parameters for the identification of similar images. Thus, the problem of retrieving images with some content is transformed in a problem of retrieving images visually close to a target (Johansson, 2000).

Texture is the visual impression of coarseness or smoothness caused by the variability or uniformity of image tone or color. It also can be defined as a function of the spatial variation in pixel intensities. In image interpretation, a pattern is defined as the overall spatial form of related features, and the repetition of certain forms is a characteristic pattern found in many cultural objects and some natural features. One immediate application of image texture is the recognition of homogeneous regions using texture properties (Amir, 2001). In remotely sensed data, texture is defined to be the local scene heterogeneity and this property is used for classification of land use categories such as water, agricultural areas, forest areas, urban areas, etc. This fact also can be used to perform segmentation, boundary detection and content-based image retrieval (Rignot and Kwok, 1990).

Content-based retrieval with texture descriptors will be very useful in Remote Sensing and GIS database context, when answering typical queries such as “find images of Amazon region that are not cloud covered”, “find images of Amazon that are 10% cloud covered” or “find all images in database that have regions that looks like this vegetation pattern” (Zhu and Yang, 1998).

In this work we focus on the retrieval of complex images based on their textural content. The approach we consider is oriented to retrieving textured Landsat images (see Figure 1) from a thematic database, where the semantic content of the images is limited to a specific domain.

2. RELATED WORKS

There are several remarkable achievements within the field of image retrieval, where a wide range of research purposes and applications are found. A survey on content-based search in image and video database is shown in Johansson (2001). The main focus in this work is on what kind of image features are used but also the user interface and the users possibility to interact with a database image system.
Manjunath and Ma (1996), focus on image processing aspects and in particular using texture information for browsing and retrieval of large image data. They use Gabor wavelet features for texture analysis and made comparisons with other multiresolution texture using Brodatz texture database and illustrated an application to browsing large air photos.

Di Sicascio et al. (1998), focus on the retrieval of complex images based on their textural content. They use GMRF for texture discrimination and a region-growing algorithm for texture segmentation. Relevance feedback is introduced to improve retrieval accuracy. If the results of the query are not satisfying, the user can select, within the retrieved subset, the images considered relevant. The system modifies the query by increasing the query feature vector with the contribution of the selected images. By selecting an image as relevant, the most similar image to the query, relevance feedback was applied, stressing distinguished feature in the relevant image. The selected image ranks highest, the one less similar has been withdrawn and the image of the query area was selected from now appears in the retrieved set.

Texture analysis has been extensively used to classify remotely sensed images. Land use classification where homogeneous regions with different types of terrains (such as wheat, Bodies of water, urban regions, etc.) need to be identified is an important application. Haralick have analyzed SAR images using texture features computed from gray level co-occurrence matrices. Wikantika (2001) focuses on radiometric correction of topographically induced effect and combined use of spectral and textural features from Landsat-TM image to discriminate land cover in mountainous area. Texture statistics have been used as the feature dimensions in improving classification accuracy.

Kim (2001) uses the kernel principal component analysis (PCA) for texture classification. The Kernel PCA has recently been proposed as a nonlinear extension of PCA. The basic idea is to first map the input space into a feature space via a nonlinear map and then computes the principal components in that feature space. By adopting a polynomial kernel, the principal components were computed within the product space of the input pixels making up the texture patterns.

The use of wavelets to perform texture classification is addressed by Leguizamon (1996). The wavelet transform has emerged as a very important tool for the analysis of non-stationary one-dimensional and multidimensional signals. In many cases it is a convenient alternative to Fourier analysis due to its property of preserving the information of space and scale. Wavelet transform is a signal decomposition onto a set of basis functions. Basis functions are called wavelets and are obtained from dilation, contraction as well as shifting of a prototype wavelet. The characterization of texture in digital images is based in the multiresolution analysis of the surface under study. This approach stems in the extraction of roughness and coarseness information, which is supposed to be present in the lower resolution images, and the detail images obtained at different scales of wavelet decomposition.

3. TEXTURE DESCRIPTORS

The texture descriptors are computed based on co-occurrence matrices. These matrices estimates properties related to second-order statistics. The \( g \times g \) gray level co-occurrence matrix \( P \) for a displacement \( d = (d_x, d_y) \) in a direction \( \theta \) is defined as

\[
P_{\theta}(i,j) = |{(l(r, s), (t, v)) : ((r, s), (t, v)) = i, l(t, v) = j}|,
\]
where \( I \) is a gray level image, the entry \((i,j)\) of \( P_\delta \) is the number of occurrences of the pair of gray levels \( i \) and \( j \) which are in a direction \( \theta \) and a distance \( d \) apart. For a set of co-occurrence matrices \( P_{0,45,90,135} \) we find the

\[
p(s,t) \text{ matrix given by } p(s,t) = \sum_{j=0,45,90,135} P_j(s,t). \]

The normalized matrix is given by \( p_{\text{norm}} = \left[ p'(s,t) \right] = \frac{1}{M} \left[ p(s,t) \right] \), where

\[
M = \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} p(s,t). \]

Consider that \( p_x(s) = \left[ \sum_{i=0}^{m-1} P_i(s,t) \right] \) and \( p_y(t) = \left[ \sum_{t=0}^{m-1} P_s(s,t) \right] \) are the marginal probability matrices, the six most important texture descriptors can be given by the normalized matrix \( p(s,t) \). These descriptors are shown in Table 1:

<table>
<thead>
<tr>
<th>Principal textures descriptors</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bullet ) Energy (or second angular moment) ( f_1 = \sum_{i=0}^{m-1} \sum_{t=0}^{m-1} (p'(s,t))^2 )</td>
</tr>
<tr>
<td>( \bullet ) Contrast (Inertia) ( f_2 = \sum_{i=0}^{m-1} \sum_{t=0}^{m-1} (s-t)^2 (p'(s,t))^2 )</td>
</tr>
<tr>
<td>( \bullet ) Local homogeneity ( f_3 = \sum_{i=0}^{m-1} \sum_{t=0}^{m-1} (1 + (s-t)^2)^{-1} p'(s,t) )</td>
</tr>
<tr>
<td>( \bullet ) Entropy ( f_4 = \sum_{i=0}^{m-1} \sum_{t=0}^{m-1} p'(s,t) \log(p'(s,t)) )</td>
</tr>
<tr>
<td>( \bullet ) Correlation I ( f_5 = (f_4 - HXY1) / \max {HX, HY} )</td>
</tr>
<tr>
<td>where, ( HX = \sum_{s=0}^{m-1} p_x(s) \log(p_x(s)) ), ( HY = \sum_{t=0}^{m-1} p_y(t) \log(p_y(t)) ), and ( HXY1 = -\sum_{s=0}^{m-1} \sum_{t=0}^{m-1} p'(s,t) \log(p'(s,t)p'(t)) )</td>
</tr>
<tr>
<td>( \bullet ) Correlation II ( f_6 = \sqrt{1 - e^{-2(HXY2-f_5)}} )</td>
</tr>
<tr>
<td>where, ( HXY2 = -\sum_{s=0}^{m-1} \sum_{t=0}^{m-1} p'(s,t) \log(p'(s,t)p'(t)) )</td>
</tr>
</tbody>
</table>

The meaning of these descriptors can be seen as: Energy is a measure of image homogeneity. The higher the energy, the more homogenous it is. The homogeneity measurement is larger for patterns with a smaller contrast. The contrast and the homogeneity measure dual properties. Entropy can be interpreted as a measurement of
uncertainty, variability, and complexity. Correlation is a measurement of the degree of linear relationship between two random variables.

4. AN APPROACH FOR LANDSAT IMAGE RETRIEVAL

Landsat Thematic Mapper images are multispectral electromagnetic energy measured in seven spectral bands ranging from the visible to the thermal infrared. Each pixel represents an area 30 m by 30 m for six of the seven bands whereas pixels in the thermal infrared band represent an area 120 m by 120 m. Some typical spectral band images with different texture patterns are shown in Figure 1.

In our image retrieval model, each image in database is subdivided in a $k \times k$ regular grid. The grid cells in main diagonal are selected to generate the vector $f$, the feature vector based on texture descriptors that is shown in Table 1. These vectors are grouped in a set $F$ that will be used to find an image in the database. The Figure 2 shows the steps to find the feature vectors.

Once image associated features values have been computed and stored in $F$, queries may be processed (see Figure 3). Queries can be submitted either as a sample texture image or selecting a rectangular area $R$ within a submitted image. The co-occurrence based features are computed in the reference vector $f$, and then are used to find similarities.

Some image retrieval methods are based on a relationship between objects. The relationship is described variously as “similarity”, “dissimilarity” and “association”. The measurement of similarity is designed to quantify the likeness between objects so that if one assumes it is possible to group objects in such a way that an object in a group is more like the other members of the group than it is like any object outside the group. The similarity measurement $S$, is given by the angle between nonzero vectors $f$ and $f_i$ in $\mathbb{R}^n$ by the Equation (7) following

![Image](image_url)
$S_i = \frac{\langle f_r, f_i \rangle}{\|f_r\| \|f_i\|}$

(7)

Where $\langle f, f_i \rangle$ is the inner product of $f$ and $f_i$, given by $\sum_{i=1}^{n} f_r f_i$, $\|f_r\|$, $\|f_i\|$ are their respectively norms.

The retrieval is performed by measuring the similarity in the $n = 6$ dimensional space defined by the index terms between the reference vector $f$ of the query and the term vectors $f_i$ in $F$ (step 4 in QUERY procedure – Figure 3). Resulting retrieved image are indexed according to a growing distance score. Retrieval results are considered correct if the original image is retrieved in the first three higher ranking in sorted set $S'$. The procedure to indices generation and to query process is shown in Figure 2 and Figure 3.

**Procedure** INDEXING (image I)

Input: an image $I_{m,n}$ pixels
Output: the set of all feature vectors $F$
1. Divide $I$ in $k \times k$ regular sub-images
2. For each sub-image $I_{ij}$ in diagonal do
3. Find the feature vector $f$ given by Equations (1), (2), (3), (4), (5), and (6).
4. Update the set $F' = f$
5. Return $F'$;

![Fig. 2 – Indices generation procedure.](image)

**Procedure** QUERY (image R)

Input: a sample image $R_{a,b}$ pixels
Output: a sorted set $S'$
1. Find the co-occurrence matrix $C_r(R)$;
2. Find the reference feature vector $f_r$;
3. For each $f_i$ in $F$ do
4. $S_i = \frac{\langle f_r, f_i \rangle}{\|f_r\| \|f_i\|}$
5. $S' =\text{sort}(S_i)$;
6. Return $(S')$;

![Fig. 3 – Query process.](image)
5. PARTIAL RESULTS

Now, we describe our current work on extracting texture features and indices generation to accomplish rapid content-based image retrieval across similar satellite images. We are currently investigating the use of the four principal descriptors as indices in a simple query. For a simulated database with ten Landsat 128 x 128 textured images randomly selected in a pool of 200 similar images, the indices set $F$ was by INDEXING Procedure (Figure 2). The energy, entropy, contrast, and homogeneity descriptors are used to construct the feature vectors $f$. These values are shown in the Table 2 rows.

<table>
<thead>
<tr>
<th>Table 2 A simulated database with indices set $F$.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>-2.9030</td>
</tr>
<tr>
<td>0.0210</td>
</tr>
<tr>
<td>0.0177</td>
</tr>
<tr>
<td>0.1573</td>
</tr>
</tbody>
</table>

The descriptors value variation in database is shown in Figure 4. These values indicate different textures kinds in images. With these texture patterns we can explore different similarity measures for each of different feature vectors.

![Figure 4: Descriptors values variation](image)

A simple query process using image 1 as the reference image ($f = [-2.9030, 0.0210, 0.0177, 0.1573]$) we obtained as result the set $S = [1.9582, 9.1539, 6.1050, 4.4855, 3.9002, 7.8119, 1.8851, 2.9612, 7.6205]$, when sorted given us $S' = [1.8851, 1.9582, 2.9612, 3.9002, 4.4855, 6.1050, 7.6205, 7.8119, 9.1539]$, that indicate best similarity with image 7, image 2 and image 8. Visually this appears a good result. Further we can adopt a norm to measure the result accuracy.

Computing the feature vector for an image (with four descriptors) takes 13.34 seconds of CPU time (in Matlab on a Pentium II – 200 MHz) or 2.1 minutes on average to generate all $F$ set for ten images. The similarity between two feature vectors takes 0.009 seconds. A query takes 0.01 second on average. The experimental results indicate that this approach are quite robust and maintaining a reasonable level of retrieval performance. Our current
work is on incorporating the diagonal segmentation in images to obtain best performance in co-occurrences and indices generation.

6. FINAL REMARKS

Currently available large image repositories require new and efficient methodologies for query and retrieval. Content-based access appears to be a promising direction to increase the efficiency and accuracy of unstructured data retrieval. A prototype is implemented with Matlab and its preliminary results had been shown a good performance for textured Landsat images. Further work is in progress in two main directions: improvement of an automatic merging of similar texture areas and integration of the system with other feature extraction approaches. The limitations include the fact that the system requiring user intervention is strongly dependent on the original image quality, and time consuming in index generation.

7. REFERENCES:


