A Review on Feature Descriptors for Remote Sensing Image Retrieval

Authors

D.R.Dhotre¹, Dr. G. R. Bamnote², Manisha R. Amalkar³

¹Asst. Prof., CSE Department, SSGMCE, Shegaon, India
²Professor, CSE Department, PRMIITR, Badnera, India
³ME Student, CSE Department, SSGMCE, Shegaon, India

Email: grbamnote@mitra.ac.in, d_dhotre@yahoo.co.in, manishaamalkar31@gmail.com

ABSTRACT:
This paper presents the review on results of application of feature descriptors for remote sensing image retrieval. The circular covariance histogram and the rotation-invariant point triplets potential are explored here as a multi-scale texture descriptors.

In remote sensing images, the availability of images containing a significantly higher amount of spatial and spectral details has Cover (a piece of ground) with flat stones or bricks; the way for new applications (e.g., hyper spectral target detection, compound object recognition, etc.) and new commercial products and it has enabled the use of a wider range of image analysis methods. On the other hand, the rapid accumulation of gigabytes worth of remote sensing data on a daily basis has provided robust and automated tools, which was designed for their management, search, and retrieval, as essential for their effective exploitation.

Keywords: Content-based image retrieval (CBIR), mathematical morphology (MM), remote sensing, texture description.

INTRODUCTION

In remote sensing scene, spatial and spectral resolution capacities of image acquisition devices have increased significantly. Here spectral wise, every pixel is described with its spectral response with the advent of hyper spectral sensors, across not a few but hundreds of narrow contiguous spectral bands, and as far as spatial resolutions are concerned, they have improved here from 30 m down to a few tens of centimeters per pixel. The rate of image acquisition has also reached at higher levels, thus leading to the formation of very large remotely sensed data warehouses in private or public establishments.[2]

Mathematical morphology (MM) is another relatively popular tool with remote sensing data classification [4], its application to the CBIR of remote sensing images has been, on the other hand, only application specific [5], which is a shame since, as far as texture description is concerned, morphological approaches are among the top performing solutions [6], [7], as they are both multi-scale and possess invariance against rotation and illumination variations.

In detail, motivated by their texture description performance, we have to apply a couple of recently introduced texture descriptors, namely, the circular covariance histogram (CCH) and the rotation-invariant point triplets (RITs) [7], on content-based remote sensing image retrieval. In this regard, the Fourier power spectrum (FPS) of an image can be of great assistance since it is inherently suitable for describing periodic functions and detecting dominant orientations. Aptoula [5] has recently conducted a comparative study on parameterization measures for morphological image series using the Fourier transform.
LITERATURE REVIEW:
1. Krystian Mikolajczyk and Cordelia Schmid OCTOBER 2005, compare the performance of descriptors computed for local interest regions, as, for example, extracted by the Harris-Affine detector. Many different descriptors have been proposed in the literature it is unclear which descriptors are more appropriate and how their performance depends on the interest region detector. Here, according to author, the number of comparison is carried out for different descriptors, different interest regions, and for different matching approaches.
2. Mauro Dalla Mura, OCTOBER 2010, Propose to characterize the spatial information of VHR data by using a multilevel, multi-attribute approach based on morphological attribute filters. These are proposed to be an extension of the morphological profiles and of their derivative concepts, which are conventionally defined for openings and closings by reconstruction.
3. Katia Stankov and Dong-Chen He Oct. 2010, APs provide a multilevel characterization of an image created by the sequential application of morphological attribute filters that can be used to model different kinds of the structural information. Sophisticated methods that integrate spatial and spectral information are needed to detect and extract buildings from VHSR images. Sophisticated methods that integrate spatial and spectral information are needed to detect and extract buildings from VHSR images.
4. Katia Stankov and Dong-Chen He Jan. 2013, the proposed supervised method included the following steps: Several classes were defined according to the roof colors of buildings in the image. For each class, a gray scale image was generated based on the spectral similarity (color) between roofs to improve the low local contrast of VHSR images, which can cause buildings to be undetected. In these images, potential building locations appeared with bright tones. Next, the grayscale HMT was applied to the images generated in the previous step in order to assign pixels to buildings.
5. E. Aptoula Nov. 2012, Author investigates the visual appearance of real-world surfaces and the dependence of appearance on the geometry of imaging conditions. He discuss a new texture representation called the BTF (bidirectional texture function) which captures the variation in texture with illumination and viewing direction. Author presents a BTF database with image textures from over 60 different samples, each observed with over 200 different combinations of viewing and illumination directions. He describe the methods involved in collecting the database as well as the importance and uniqueness of this database for computer graphics.

CONTENT BASED REMOTE SENSING IMAGE RETRIEVAL
Content based image retrieval (CBIR) which provides a very effective and advanced means to manage and utilize image database and computer vision. The early work of Bretschneider et al. [8] relies on the statistical properties of image classes that determined through fuzzy k-means classification. Scott et al. [9], for instance, are one of the few that employ the shape-related properties of image content, specifically here by means of multiscale bitmap representations.
In particular, Hongyu et al. [10] and Newsam et al. [11] have employed Gabor filters, and Tobin et al. [12] have explored local binary (LBP) and local edge patterns (LEP), while Xute al. [13] have combined spectral features with the power of grayscale co-occurrence matrices. More recently, Samal et al. [14] investigated auto-correlation measures and fractal attributes computed from a hierarchical image representation, and Gleason et al. [15] combined multispectral histograms with LBP and LEP. Liu et al. [16] studied a region-based approach where combinations of color and texture features are computed on image regions after segmentation.
MORPHOLOGICAL TEXTURE DESCRIPTORS

At first this section will recall the definition and principles of CCH and RIT [7] and then proceed to present two new texture descriptors based on the FPS of an image’s QFZ representation.

**A. CCH (circular covariance histogram)**

CCH and RIT aim at detecting periodic patterns in their input. CCH relies on using isotropic structuring elements (SEs) in the form of circles, conversely to morphological covariance which employs point couples at arbitrary predefined directions and distances as SEs. Its working principle is simple; after processing the gray scale input image \( f \) with the circular SEs \( B_i \) of various radii \( i \in [1, n] \) using a morphological operator \( T \), where \( T \) could be erosion (\( \epsilon \)), dilation (\( \delta \)), opening (\( \gamma \)), closing (\( \phi \)), or another morphological operator or filter, we obtain a series \( \Pi^T_n \) of intermediate images. In other words, this step consists of construction of a morphological scale space based on the operator \( T \), similarly to the morphological profiles, where the values that each pixel possesses across the various filtering stages constitute its spectral–spatial signature [1].

**Fig. 1.** (a) Horizontally oriented point pair SE at distance of 8 pixels, as used by morphological covariance, and (b) a circle-shaped SE with a 4-pixel radius as used by CCH (here the SE center is denoted with \( \times \)).

**B. RITs (rotation-invariant point triplets)**

Although similar to CCH, RIT, on the other hand, employs not only a circular SE but a set of point triplets at various orientations of which the outputs are combined by means of various methods. To explain, given a grayscale image \( f \) and circular SEs \( B_i \) of radius \( i \in [1, n] \), we decompose each circle into anti-diametrical point triplets. Since the circle which have a perimeter of \( 8i \) pixels, thus we obtain \( 4i \) point triplets \( B_{ij}, j \in [1, 4 \times i] \) for each circle, at a distinct orientations, each have an angle of \( 2 \times \pi \times k/(8 \times i) \) where \( k \in [0, 4 \times i−1] \). An example for \( i = 2 \) is shown in Fig. 2.

**Fig. 2.** Eight point triplets \( \{B_{2j}\}_{1 \leq j \leq 4 \times 2} \) corresponding to radius \( = 2 \).

**C. FPSs from QFZs**

The FPS of an image can be of great assistance since it is inherently suitable for describing periodic functions and detecting dominant orientations. As such, FPSs have been employed in the context of texture description for CBIR purposes very early [10]. Aptoula [5] has recently conducted a comparative study on parameterization measures for morphological image series using the Fourier transform. Author investigates the visual appearance of real-world surfaces and the conditions. He discusses a new texture representation called the BTF (bidirectional texture function) which captures the variation in texture with illumination and viewing direction.
Fig. 3. Two samples from the UC Merced Land Use data set (golf-course-01 and parking-lot-35, both of size 256 × 256 pixels) with distinct levels of coarseness and the resulting FPS_1 series obtained with seven disks. (a) Golf course. (b) FPS of (a). (c) Parking lot. (d) FPS of (c). (e) FPS_1 of (a) and (c).

<table>
<thead>
<tr>
<th>Similarity metrics</th>
<th>FEATURE LENGTH</th>
<th>DE</th>
<th>DM</th>
<th>DI</th>
<th>DB</th>
<th>DX²</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCH</td>
<td>20</td>
<td>0.726</td>
<td>0.6877</td>
<td>0.677</td>
<td>0.648</td>
<td>0.649</td>
</tr>
<tr>
<td>RIT</td>
<td>20</td>
<td>0.769</td>
<td>0.751</td>
<td>0.751</td>
<td>0.757</td>
<td>0.757</td>
</tr>
<tr>
<td>FPS1</td>
<td>14</td>
<td>0.731</td>
<td>0.74</td>
<td>0.798</td>
<td>0.719</td>
<td>0.726</td>
</tr>
<tr>
<td>FPS2</td>
<td>8</td>
<td>0.805</td>
<td>0.79</td>
<td>0.853</td>
<td>0.8</td>
<td>0.783</td>
</tr>
<tr>
<td>GABOR</td>
<td>60</td>
<td>0.762</td>
<td>0.719+</td>
<td>0.719</td>
<td>0.719</td>
<td>0.733</td>
</tr>
<tr>
<td>LBP</td>
<td>54</td>
<td>0.769</td>
<td>0.762</td>
<td>0.762</td>
<td>0.735</td>
<td>0.745</td>
</tr>
<tr>
<td>CCH+RIT</td>
<td>40</td>
<td>0.7</td>
<td>0.639</td>
<td>0.639</td>
<td>0.625</td>
<td>0.656</td>
</tr>
<tr>
<td>CCH+RIT FPS1</td>
<td>54</td>
<td>.65</td>
<td>.635</td>
<td>.614</td>
<td>.589</td>
<td>.604</td>
</tr>
<tr>
<td>CCH+RIT FPS1 FPS2</td>
<td>62</td>
<td>0.64</td>
<td>0.589</td>
<td>0.609</td>
<td>0.586</td>
<td>0.575</td>
</tr>
<tr>
<td>LOCAL</td>
<td>150</td>
<td></td>
<td></td>
<td></td>
<td>0.601</td>
<td></td>
</tr>
<tr>
<td>LOCAL INVARIENTS[18]</td>
<td>1500</td>
<td></td>
<td></td>
<td></td>
<td>0.561</td>
<td></td>
</tr>
</tbody>
</table>

TABLE 1: Retrieval performances of various descriptors in terms of ANMRR for various similarity metrics. “+” denotes concatenation[1].

D. Gabor Filters: Gabor filter have been implemented with five scales and six orientations, thus leading to a feature vector of length 60, consisting of the mean and standard deviation of the 30 filter outputs. Gabor filters are band pass filters which are used in image processing for feature extraction, texture analysis, and stereo disparity estimation. The impulse response of these filters is created by multiplying an Gaussian...
envelope function with a complex oscillation. Gabor showed that these elementary functions minimize the space (time)-uncertainty product. By extending these functions to two dimensions it is possible to create filters which are selective for orientation. Under certain conditions the phase of the response of Gabor filters is approximately linear.

Figure 4. Class-specific retrieval performance of the various tested descriptors in terms of ANMRR using optimal similarity metrics.[1]

D. LBP (Local Binary Pattern):
LBP on the other hand, we have used the multiscale rotation-invariant definition from [15], equipped with three radii (LBPriu28,1+16,2+24,3), so its feature vectors are 54-dimensional. Local Binary Pattern was introduced by Timo ojala [17]. The standard version of the LBP of a pixel is formed by thresholding the 3X3 neighborhood of each pixel value with the center pixel’s value. Let gc be the center pixel gray level and gi (i=0,1,..7) be the gray level of each surrounding pixel. If gi is smaller than gc, the binary result of the pixel is set to 0 otherwise set to 1. All the results are combined to get 8 bit value. The decimal value of the binary is the LBP feature. The uniform pattern constraint reduces the number of LBP pattern from 256 to 58 and it is very useful for face detection [17].

CONCLUSION
The remote sensing image collection has constantly increasing size attracted a lot of attention to developing solutions for their content-based retrieval. In this context, Content descriptors are of crucial importance as they are expected to capture the higher level properties of their input. Although the local invariants have been successfully used in this regard, it has chosen instead to focus in this paper on global and, specifically, on morphological descriptors. Motivated by their success with gray scale texture description, we have studied CCH and RITs that describe their input’s periodicity and size distribution. Moreover, two new texture descriptors have been developed consisting of simple rotation invariant measures on the FPS of the QFZ-based scale space of their input. These two have been designed specifically to complement CCH and RIT, with additional information on directionality and coarseness. All four have been tested with the UC Merced LULC data set, the largest of its kind, against LBP and Gabor filters, where the morphological approach has achieved the best known results, outperforming SIFT too, despite a much shorter feature vector length.
In conclusion, as content description by means of local invariants is an established paradigm, this paper by no means aims to convince the reader otherwise but rather aims to underline the potential of global descriptors that have been systematically ignored since the advent of local invariants.
REFERENCES


