TEXTURE-BASED REMOTE-SENSING IMAGE SEGMENTATION

Dihua Guo, Vijayalakshmi Atluri and Nabil Adam
{devaguo,atluri,adam}@cimic.rutgers.edu

ABSTRACT

Typically, high-resolution remote sensing (HRRS) images contain a high level noise as well as possess different texture scales. As a result, existing image segmentation approaches are not suitable to HRRS imagery. In this paper, we have presented an unsupervised texture-based segmentation algorithm suitable for HRRS images, by extending the local binary pattern texture features and the lossless wavelet transform. Our experimental results using USGS 1ft orthoimagery show a significant improvement over the previously proposed LBP approach.

1. INTRODUCTION

With the advances in remote sensing technology high-resolution remote-sensing (HRRS) imagery has been drawing more attention from a wider audience, beyond the traditional scientific user community. The state-of-the-art remote sensing image retrieval supports image search based on metadata (i.e. sensor type, sensor name, image type, image date, and location) or on perceptual features (e.g. color, texture, and shape). However, for most non-scientific users, articulating a query using these metadata or low-level features can be non-intuitive and difficult. We believe that adding semantic annotations to HRRS images would greatly benefit in effectively formulating queries to retrieve the data that users are interested.

The first step to performing automatic annotation is to segment HRRS images *properly* and *accurately*, which is quite a challenging task. To segment an image accurately means to avoid any false segmentation, whereas to segment properly implies that the segmentation algorithm should terminate at a point to eliminate over-segmenting or under-segmenting. While under-segmenting may cause loss of information, over-segmenting may bring out too many ignorable details.

We believe that to accurately and properly segment a HRRS image, no single segmentation technique is suitable. This is because, in a HRRS image, different objects appear in different land use categories. For example, in a residential area, houses and roads are the main elements. To partition objects in this land use category, lines, rectangles, and directions would be helpful for further analysis. On the other hand, in a vegetation covered area, it is hard to extract any lines. Therefore, it would be more efficient to apply a multi-step segmentation procedure, where a HRRS image is segmented at the land use level first, and then different segmentation algorithms are applied based on the the category of the land use. In essence, in this paper, we argue that, to be able to segmenting HRRS images correctly and properly at land use level is an important preprocessing step to identify objects. Even though our ultimate purpose is to recognize objects in a more detailed level, in this paper, we limit our focus to the problem of *efficiently* and *effectively* segmenting remote sensing images at land-use level, which include heavy residential areas, industrial area, pastures and water bodies.

In general, existing remote sensing image segmentation algorithms can be categorized into color-based and texture-based algorithms. Color is a distinctive global feature, which is successful when the resolution of images is not very high. HRRS images are different in the sense that each pixel is a mixed measure of various ground objects. Therefore, using illumination alone is not enough for many land covers. In addition, based on our observation, most of the land cover has texture patterns. The repeating occurrence of homogeneous regions of images is texture. Texture image segmentation identifies image regions that have are homogeneous with respect to a selected texture measure. Recent approaches to texture based segmentation are based on linear transforms and multiresolution feature extraction [1], Markov random filed models [2, 3], Wavelets [4–6] and fractal dimension [7]. Although unsupervised texture-based image segmentation is not a novel approach, these have limited adoption due to their high computational complexity. Some of the existing methods perform well on small set of fine-grained texture mosaics, but were not examined on HRRS images. Some approaches used remote-sensing images, but they require *apriori* knowledge of the image to achieve satisfactory results.

We have found that the Local Binary Pattern (LBP) approach [8–10] is most suitable to be used as a base idea to build our segmentation approach, since it has been proved to be a computationally efficient, rotation invariant and robust in handling multi-resolution images. The texture feature, LBP, is constructed by the joint distribution of values of a circularly symmetric neighbor set of pixels in a local neighborhood. However, our empirical studies show that, even though the LBP approach has been proved to be a robust algorithm in segmenting regular scenery images, it is not suitable for HRRS images since texture description in LBP is a highly discriminating texture features and LBP treats the reflectance distortion as a different texture feature. In contrast, HRRS images taken under natural circumstances, landscapes are usually distorted by natural phenomena. As a result, if we use LBP, images tend to be over-segmented.

The enhanced LBP unsupervised texture segmentation algorithm [9] that uses the repeating patterns as another description of texture, is also not suitable to segmenting HRRS images. This is because, HRRS images comprise of repeating patterns at different scales, as illustrated in figure 1. Both these images are cropped from the same 1ft resolution aerial images of Spring Field, MA, where figures 1a and figure 1b show woods and suburban residential neighborhood, respectively. As we can see, the repeating patterns of woods exist in a smaller neighborhood compared to that of the residential area.

0-7803-9332-5/05/$20.00 ©2005 IEEE

(a) Woods (b) Residential Area

Fig. 1. Textures with Different Scales

Since the different scale textures of the repeating patterns cannot be detected by the enhanced LBP, we have extended this algorithm to suit to these different scales of repeating patterns.

Even though the algorithm proposed by [9] is efficient and overcomes the problem of requiring apriori knowledge, our preliminary experimental results show that it does not perform well on HRRS images that contain non-uniform textures. Unlike the segmentation of mosaic texture images used in the experimentation of LBP, segmenting HRRS images is different in the following aspects: (i) Presence of various sources of noise in the acquisition process. (ii) Presence of different scales of textons (the repeating unit that forms texture).

In this paper, we propose an unsupervised texture-based segmentation algorithm for HRRS images suitable for segmenting at the land use level. It employs the splitting-and-merging approach [11] to texture-based segmentation. In the preprocessing stage, each HRRS image goes through Haar wavelet transformation. The segmenting approach consists of three major steps: (i) hierarchical splitting that recursively splits the original image into children blocks by comparing texture features of blocks, (ii) optimizing, which adjusts the splitting result, if the results of the reduced resolution images have dramatically reduced segments, (iii) merging, in which the adjacent regions with similar texture are merged until a stopping criterion is met.

Our approach enhances the LBP approach [9] in the following aspects: (1) We preprocess images using wavelet transformation to eliminate the noise in the acquisition process. (2) Instead of segmenting HRRS images under original resolution alone, we also apply the algorithms on the wavelet-transformed images for comparison. If the segmenting results change dramatically, we would adopt the part with fewer components. (3) In addition, we adjust the termination thresholds according to the size of blocks, which is decided based on the homogeneity of sibling blocks.

2. OUR TEXTURE-BASED SEGMENTATION APPROACH

In this section, we present our unsupervised texture-based segmentation approach, which is depicted in Figure 2, and elaborate the wavelet transform, quad-tree-based splitting, splitting results optimization, and agglomerative merging steps. A longer version of this paper is available in [12].

2.1. The Texture Features

In the following, we introduce the texture features used in our approach. Recent studies by [9, 10] show that excellent texture discrimination can be obtained by local texture operators

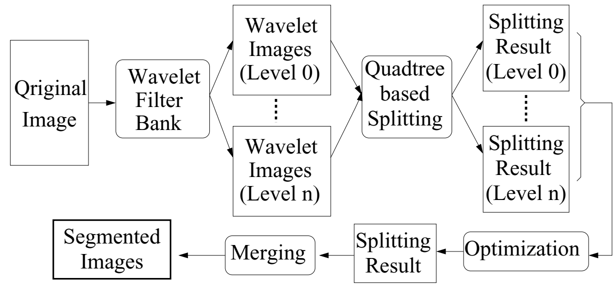

Fig. 2. Unsupervised Texture-based Segmentation Algorithm

and nonparametric statistical discrimination of sample distribution. Because it is a straightforward approach, the computation cost of texture segmentation using distributions of local binary patterns (LBP) is better than other methods [9, 10]. The texture contents of an image region are characterized by the joint distribution of the LBP and contrast (C). By definition, LBP is invariant to any monotonic gray-scale transform. This is an important feature, because every HRRS image is taken under different illumination and atmospheric condition. A simple contrast measure C is taken into consideration to address the contrast of the texture, which is not described by LBP. Here C is the difference between the average gray-level of those pixels. To compare two images, we actually compare the distributions of LBP/C. The LBP/C distribution is approximated by a discrete two-dimensional histogram of size 256xb, where b is the number of bins for C. According to [9, 10], b is chosen as a trade-off between the discriminative power and the stability of the texture transform. A very large b would make the 2-D histogram very sparse and incomparable. If b is too small, the histogram will lack resolution and contrast feature C will add very little discriminative information to the process. In [9], 32 bins were chosen, based on the consideration that their average image size is 32x32, whereby an average number of 32 entries per bin can be obtained. In our experiment, we use the HRRS images, whose average size is 1024x1024. Based on our experiment, 16-bin generates reasonable results.

A log-likelihood-ratio, the G statistic [9] is used as a pseudo-metric for comparing LBP/C distribution.

$$G = 2 \times ([\sum_{s,m}\sum_{i=1}^{n} f_i \log f_i] - [\sum_{s,m}(\sum_{i=1}^{n} f_i)\log(\sum_{i=1}^{n} f_i)] - [\sum_{i=1}^{n}(\sum_{s,m} f_i)\log(\sum_{s,m} f_i)] + [(\sum_{s,m}\sum_{i=1}^{n} f_i)\log(\sum_{s,m}\sum_{i=1}^{n} f_i)])$$

where s, m are the two sample histograms, n is the number of bins and f_i is the frequency at bin i. The value of the G statistic indicates the probability that the two sample distributions come from the same population: the higher the value, the lower the probability that the two samples are from the same population. We measured the similarity of two histograms with a two-way test of interaction or heterogeneity. The more alike the histograms s and m are, the smaller is the value of G.

2.2. Wavelet Transformation

Compared to the ordinary scenery digital images, HRRS images contain noise that deteriorates the result of unsupervised segmentation approach. We have chosen wavelet transformation to traditional low-pass filter because of its multi-resolution property. Moreover, the advantage of wavelet transformation over Fourier analysis is that it allows better resolution in space and frequency. We have studied two basic wavelet transforms – Haar Wavelet and db5 Wavelet and found that the Haar decomposition is faster than the daubechies. Since the performance of the two wavelet transforms do not have critical difference, we chose the faster Haar wavelet transform.

After decomposing an image by wavelet transformation for n level, the size of approximation of level n is $1/(2^n)$ of the original size, and the resolution of approximation of level n is 2^n, the original ground resolution. Therefore, we needed to perform inverse discrete wavelet transform to reconstruct approximation that has original size as well as reduced resolution. The compressed images are stored in the database and split along with the original images. We refer to them as wavelet images.

2.3. Quad-tree Based Splitting

Split-and-merge is a top-down approach, which is an attempt to improving computational time when compared to pixel-based approach. Another advantage of split-and-merge approach is that, one can compare block-based features instead of pixel-based features as the latter one would have difficulty when image pixels border several textured regions

Given an original (1024x1024) HRRS image, we have divided it into four square subblocks recursively until each individual subblock is uniform in texture. Whether we need to divide any block into 4 subblocks is based on a uniformity test. As we pointed out in the previous subsection, the extracted block feature, the LBP/C distribution, is represented by a 2-D histogram and a pseudo-metric G is used to measure the distance between two blocks. To compare among four blocks, we compute the six pairwise G metrics between the LBP/C histograms of the four subblocks. If we denote the largest of the six G values by G_{max} and the smallest by G_{min}, the block is found to be non-homogeneous, and therefore is split further into four subblocks, if the measure of relative dissimilarity R within region is greater than a threshold X [9], where $R = \frac{G_{max}}{G_{min}} > X$. This procedure is repeated recursively on each subblock until a predetermined minimum block size S_{min} is reached. The minimum block size cannot be determined arbitrarily, as the block has to contain a sufficient number of pixels for the LBP/C histogram to be reliable. According to [9] 8x8 blocks is stable enough. We define S_{min} as 16x16 to reduce the computational burden.

To determine the proper threshold X is a challenge. [9] argues that one should choose a small value for X, and suggested 1.2 for their experimental dataset. However, for HRRS images, X should be adjusted according to the block size. Obviously, the more pixels each block contains, the larger chance for the two blocks to be far away. The situation would be exaggerated when the texton is small. Since smaller texture blocks will contain fewer data points from which to derive statistical features, there will be a larger deviation in features extracted from these blocks of a similar texture. In words, possibly the G metric of two blocks, each with 128x128 pixels is much larger than that of the 2 blocks with 32x32 pixels, even though the first pair are perceptually similar grasslands.

To develop the distance threshold to be used for homogeneity test, we have examined the training texture cuts taken from our image database. We observed a comparatively strong (0.63) correlation between the G metric and image size. Performing a linear regression analysis on this data, the threshold function is obtained. To simplify our computation, equation 1 is used in our experiment.

$$X = 1.2 \quad if\sqrt{\frac{S_{original}}{S_{block}}} > 4$$
$$X = 5.3 - \sqrt[4]{\frac{S_{original}}{S_{block}}} \; if\sqrt{\frac{S_{original}}{S_{block}}} <= 4 \qquad (1)$$

Here, X is the threshold to be used in the homogeneity test, $S_{original}$ the size of the original image, and s_{block} the number of pixels in the image block.

2.4. Optimization

Since we subject all the wavelet images as well as the original ones to this recursive splitting procedure, we would get L+1 splitting results, given that our wavelet transform is L level. If we compare the results of the split, some regions would have dramatic differences in the number of divided blocks. Dramatic change is defined as follows:

$$if \quad \frac{N_{original}}{N_{wavelet}} \times \frac{S_{original}}{S_{block}} >= 100 \quad thenDr = 1$$
$$if \quad \frac{N_{original}}{N_{wavelet}} \times \frac{S_{original}}{S_{block}} < 100 \quad thenDr = 0$$

Here $N_{original}$ and $N_{wavelet}$ are the number of split blocks for original image and wavelet compression image at wavelet transform l, $Dr = 1$ implies that we consider the change is dramatic, otherwise the change is not remarkable. Usually, the original images would have the most divided regions, because they carry more noise. In this algorithm, we take the blocks at the level where noticeable decrease occurs to replace the corresponding areas in the original splitting result. For example, assume that the upper right block is split into hundreds of small areas, while the splitting result from level one wavelet compression image consist of only 4 blocks, we would take the particular block's result at level one as the final splitting result. Note with the substitution of the splitting results, the corresponding texture feature, LBP/C distributions should also be updated accordingly.

2.5. Merging

Once the image is split into blocks of roughly uniform texture, we apply a homogeneous blocks merging procedure, which merges similar adjacent regions until a stopping criterion is satisfied. The merging procedure is a bottom up procedure. At a particular stage of the merging, we merge that pair of adjacent segments, which satisfy the following criteria:

(i) The two blocks have the same parent: this situation is comparatively simple. If the distance of the feature vectors of the two blocks is smaller than a certain threshold Y. The two blocks are merged and the two respective LBP/C histograms are summed to be the histogram of the new region.

(ii) The two blocks have different parents: a block is merged with the neighbor blocks if the current block and the neighbor blocks are homogeneous, and the two neighbor blocks have distance smaller than threshold X after adjusted by the level weight. We can express the second condition by $G_{s,m} < G_{homogeneity} \times \omega_L$ Here $G_{s,m}$ is the distance metric between the two involved blocks. The weight factor ω_L takes into account the spatial proximity influence when merging blocks with smaller size to upper level. Consider the recursively split image as a pyramid, ranging from 0 (original image) to top level L, ω_L is equal to $\omega_l = 1 - (\frac{L-1}{L}) \times 0.1$, thus, this weight factor eases the merging process at lower levels by decreasing the product.

3. EXPERIMENTAL RESULTS AND OBSERVATIONS

Data Set: Our image database consists of (i) 1ft High-Resolution Orthoimagery, which cover several different areas of the US (downloaded from US Geological Survey), and (ii) 1ft resolution aerial photo of New Jersey Meadowlands. The selected areas include the following land use categories: heavy residential areas, suburban residential areas, water body, grassland, woods. Out of this database, we have manually selected 1024x1024 TIFF images, which are cropped from the original images, and contain two or more land use categories.

Experimental Results: The proposed unsupervised segmentation algorithm is *efficient* because we chose the texture feature [8, 9] that has been proven to be computationally efficient. The results of the original LBP and our proposed algorithm can be compared in Figure 3, we presented the 2 sets of images. The upper left one is the original image, the upper right one is the splitting result of the level one wavelet transform approximation, the lower left image is the final segmentation result of the proposed algorithm, and the lower right image is the segmented results when LBP [9]'s LBP, while the one on the right are the segmented results using the proposed approach. As we can observe from Figures 3 and 4, the segmented region layout is basically correct.

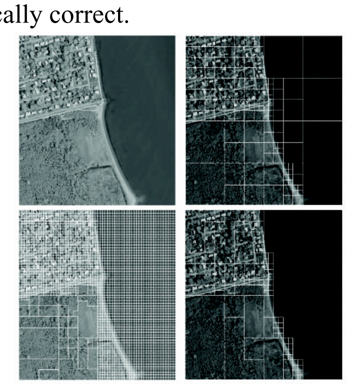

Fig. 3. Comparison of Segmentation Results. upper-left: Original USGS orthoimage; upper-right: result after quad-tree splitting and applying our approach; lower-left: splitting result applying the LBP algorithm; lower-right: segmentation results after merging

Some Observations: Although the proposed method only roughly discriminates the boundary regions, our results are fairly accurate. As shown in Figure 3 the original segmentation algorithm tends to over segment the remote sensing images. With preprocessing and optimization, we have dramatically improved the segmentation results for some areas, such as water and woods, that suffer from high level of noise. In residential area, the number of divided regions do not improve dramatically with the reduction in resolution. To solve this problem, instead of using 8 neighbors of a given pixel, we have increased the number of regions to 25, through which we capture more neighborhood information than before.

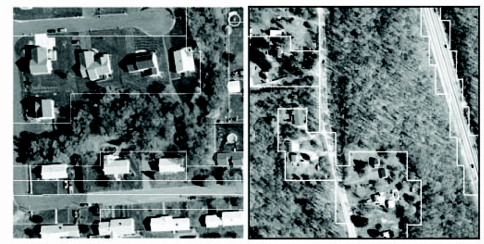

Fig. 4. Additional Segmentation Results

Next Steps: The study presented in this paper is part of the on going research on semantic-based HRRS image retrieval. Since the ultimate goal is to annotate HRRS images automatically to show interesting objects, we are currently working on identifying the land use of each region, and conducting experiments to find the distinctive features of each land use. Even though we have noticed that some land use categories are distinguishable using LBP/C distribution, others cannot be recognized by simply using one feature.

4. REFERENCES

[1] M. Unser, "Texture classification and segmentation using wavelet frames," *IEEE transaction on Image Processing*, vol. 4, no. 11, pp. 1549–1560, 1995.

[2] B. Wang and L. Zhang, "Supervised texture segmentation using wavelet transform," in *Proceedings of the 2003 International Conference on Neural Networks and Signal Processing*, 2003, vol. 2, pp. 1078–1082.

[3] Chang-Tsun Li and Roland Wilson, "Unsupervised texture segmentation using multiresolution hybrid genetic algorithm," in *Proc. IEEE International Conference on Image Processing ICIP03*, 2003, pp. 1033–1036.

[4] T.R. Reed and J. M. H. Du Buf, "A review of recent texture segmentation, feature extraction techniques," in *CVGIP Image Understanding*, 1993, pp. 359–372.

[5] A. K. Jain and K. Karu, "Learning texture discrimination masks," *IEEE transactions of Pattern Analysis and Machine Intelligence*, vol. 18, no. 2, pp. 195–205, 1996.

[6] Éldman de Oliveira Nunes and Aura Conci, "Texture segmentation considering multi band, multi resolution and affine invariant roughness.," in *SIBGRAPI*, 2003, pp. 254–261.

[7] L.F. Eiterer, J. Facon, and D. Menoti, "Postal envelope address block location by fractal-based approach," in *17th Brazilian Symposium on Computer Graphics and Image Processing*, D. Coppersmith, Ed., 2004, pp. 90–97.

[8] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions," *Pattern Recognition*, vol. 29, no. 1, 1996.

[9] T. Ojala and M. Pietikäinen, "Unsuperrvised texture segmentation using feature distribution," *Pattern Recognition*, vol. 32, no. 3, pp. 477–486, 1999.

[10] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresoltion gray-scale and rotation invariant texture classification with local binary patterns," *IEEE transactions on pattern analysis and machine intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

[11] R. C. Gonzalez and R. E. Woods, *Digital Image Processing, 2nd Edition*, Prentice Hall, 2001.

[12] D. Guo, V. Atluri, and N. R. Adam, "Texture-based remote-sensing image segmentation," Tech. Rep. CIMIC0501, Rutgers University, http://cimic.rutgers.edu/ devaguo/techReport0501.pdf, 2005.