COMPARING COLOR AND TEXTURAL INFORMATION IN VERY HIGH RESOLUTION SATELLITE IMAGE CLASSIFICATION

Ewout Vansteenkiste, Abram Schoutteet, Sidharta Gautama, Wilfried Philips

University of Ghent, department of Telecommunication and Information Processing
Sint-Pietersnieuwstraat 41, 9000 Ghent, Belgium
ervsteen@telin.UGent.be

ABSTRACT

With the advent of very high resolution satellite images, such as IKONOS, the question of how we can incorporate textural information in classifying and segmenting different regions has become of great interest. In this paper we compare the power of classifying regions based on using color information alone to using texture and color texture information. We use a 2D and 3D extension of the co-occurrence matrix and the features derived from them. In the latter case the effect of color space reduction is also evaluated. We found that although color features perform best in the easy classification tasks, very high classification rates are obtained using color texture features and the fragmentation degree in the classified areas is smaller.

1. INTRODUCTION

Remote sensing has widespread applications both in the civil as well as the military domain. The emphasis in all these lies in the quick, accurate and reliable extraction of specific information out of the images, if possible in an automatic way [1].

Texture, or the natural, visible patterns present in the images, is a very useful property to segment or classify images in remote sensing [2]. Color texture however, is a relatively new concept [3]. The multi spectral and very high resolution capabilities of sensors like IKONOS and Quickbird make it possible to add more detailed, accurate information about the spatial correlation between the pixels in different textural fields and spectral bands of interest. The aim of this paper is to investigate the power of color texture features on IKONOS color images with a resolution of 4 meters.

In the next sections the methods used to generate color, texture and color texture information are described as well as the color reduction and the template test images. The obtained results are discussed and we conclude with a discussion of the results.

2. IMAGES

In this paper we consider very high resolution (4 to 1 meters) IKONOS images of the areas of Ghent and Brussels, Belgium, taken in September 1999. Since the images cover areas of about 10 squared km. it was not practicle to process the whole images. Therefor we constructed 4 well chosen template images to perform our experiments on, Figure 1. One contains only rural areas such as fields and woods. The others combine urban and rural information. The images are chosen so that we would have ground truth on the classification task. On the other hand different textures are incorporated to investigate the power of the techniques. Our main focus was to see how the techniques could handle difficult urban areas where often different types of building are mixed.

3. METHODS

3.1. Color Features

The easiest way of generating features is by clustering the color information present in the image. For each pixel in a region on interest the three-dimensional feature vector $\mathbf{x} = (x_R, x_G, x_B)$ containing the RGB information is clustered and given as input to the multidimensional Bayesian MAP classifier. well chosen test areas should lead to well separated clusters.

Not only the RGB color space was tested yet we also transformed the images to the Lab, $YCrCb$ and HSV color spaces.

3.2. Texture Features

Texture features were calculated using an adaptation of Haralic‘s co-occurrence matrix [2], which is intuitively a 2-dimensional histogram of the grey values of pixel pairs located at a predefined distance d under a specific angle θ. In our case we made the matrix independent of θ by averaging out over 8 predefined angles $\Theta = \{\frac{2k\pi}{8}|k = 1..8\}$.

0-7803-8554-3/04/$20.00 ©2004 IEEE.

Let I be an $M \times N$ image, $\triangle x = d\cos\theta$ and $\triangle y = d\sin\theta$, then the entry on position (i,j) in the matrix is given by

$$P_d(i,j) = \frac{1}{R}\sum_{m=\triangle x}^{M-\triangle x}\sum_{n=\triangle y}^{N-\triangle y}\sum_{\theta\in\Theta}\delta(f(m,n)=i \wedge f(m+\triangle x, n+\triangle y)=j),$$

with $R = \sum_{m=\triangle x}^{M-\triangle x}\sum_{n=\triangle y}^{N-\triangle y}\sum_{\theta\in\Theta}\delta((m,n)\in I)$ a normalization factor, $\delta(x)$ the Kronecker delta function, and $f(m,n)$ the grey value at pixel position (m,n).
From this matrix multiple first and second order parameters were calculated: Grey Mean Value, Variance, Angular Second Moment, Contrast, Correlation, Entropy, Homogeneity, Inverse Difference Moment, Maximum probability and Kappa [2]. We should note that the co-occurrence matrix is calculated over the intensity image $I = \frac{R+G+B}{3}$.

3.3. Color Texture Features

To extract color texture features we expanded the co-occurrence matrix to 3D [4]. In stead of just looking at the neighboring pixels in the 2D intensity image, we now tend to the look both in the Red, Green and Blue band. Hereby we select specific co-occurrences: define $\Theta = \{\frac{2k\pi}{8}|k = 1..8\}$ and $\Phi = \Theta \setminus \{\frac{k\pi}{2}|k = 1..4\}$, then again the entry on place (i,j) in the matrix, using spatial coordinates $\triangle x = d\cos\theta\cos\phi$, $\triangle y = d\sin\theta\cos\phi$ and $\triangle z = d\sin\phi$, is given by

$$P_d(i,j) = \frac{1}{R}\sum_{m=\triangle x}^{M-\triangle x}\sum_{n=\triangle y}^{N-\triangle y}\sum_{\theta\in\Theta}\sum_{\phi\in\Phi}\delta(f(m,n,0)=i \wedge f(m+\triangle x, n+\triangle y, \triangle z)=j),$$

with $R = \sum_{m=\triangle x}^{M-\triangle x}\sum_{n=\triangle y}^{N-\triangle y}\sum_{\theta\in\Theta}\sum_{\theta\in\Theta}\delta((m,n,0)\in I)$, $\delta(x)$ the Kronecker delta function and $f(m,n,0)$ is the pixel value at pixel position (m,n) in the Green band. Again different parameters are calculated from this matrix, as Angular Second Moment, Contrast, Correlation, Entropy, but also others as Cluster Shade, Inertia and Local Homogeneity [4]. Notice that the angles are chosen so that not the texture information in one band, nor the color information through the bands is taken into account, but solely the color texture information present.

3.4. Color Space Reduction

In order to speed up processing both in the case of the texture analysis as of the color texture as well as to overcome the curse of dimensionality a reduction of the number of grey values, colors was performed.
Grey value reduction: here we apply a simple clustering where we start with an initial cluster containing all grey values of all pixels present in the image, $x_1, x_2, ...x_{MN}$. The centroid c_0 and a distortion D_0 measure are calculated as follows:

$$c_0 = \frac{1}{MN}\sum_{i=1}^{MN} x_i$$

$$D_0 = \sum_{i=1}^{MN} \parallel c_0 - x_i \parallel^2 .$$

In an iteration process the initial cluster is split in 2 clusters, based on the distortion. Then the centroids are recalculated as are the distortions to these centroids. In that way we keep on splitting up till we obtain the a priori defined number of clusters, grey value levels. In our case we reduced to 32 grey values.
Color reduction: here we used the Peer Group Filtering (PGF) algorithm [5], where first Peer Groups around a pixel are calculated to weigh color values, then a non linear filter is applied to remove possible impulse noise (not discussed here) after which the remaining color values are clustered based on local statistics. Let $x_0(i,j)$ be the color vector at position (i,j), take a $w \times w$ window around this pixel and calculate the Euclidian distance $d_l(i,j)$ for all surrounding pixels $l = 1...k = w^2 - 1$. The point of the clustering is now to detect if multiple color clusters are present in the window around $x_0(i,j)$ and to split them if so.
A peer group $P(i,j)$ of $x_0(i,j)$ and size $m(i,j)$ is defined as

$$P(i,j) = \{x_l(i,j)|l = 1...m(i,j)-1\},$$

consisting of $x_0(i,j)$ and the neighboring pixels of the "similar" color. Now we try to define the value of $m(i,j)$ as

$$m(i,j) = \arg\max_l J(l),$$

with

$$J(l) = \frac{|a_1(l) - a_2(l)|^2}{s_1^2(l) - s_2^2(l)}, \qquad l = 1...k$$

$$a_1(l) = \frac{1}{l}\sum_{n=0}^{l-1} d_n(i,j) \wedge a_2(l) = \frac{1}{k+1-l}\sum_{n=l}^{k} d_n(i,j),$$

$$s_1^2(l) = \sum_{n=0}^{l-1}|d_n(i,j) - a_1(l)|^2 \wedge s_2^2(l) = \sum_{n=l}^{k+l-1}|d_n(i,j) - a_2(l)|^2.$$

Once this is done $x_0(i,j)$ is replaced by the weighed average of its Peer Group Neighbors

$$x_{new}(i,j) = \frac{\sum_{l=0}^{m(ij)-1} w_l p_l(i,j)}{\sum_{l=0}^{m(ij)-1} w_l} \quad p_l(i,j) \in P(i,j)$$

with w_l standard Gaussian weights, chosen according to the relative positions of $p_l(i,j)$ with respect to $x_0(i,j)$. Once this Peer Group Filtering is done the clustering or color quantization can begin. A weight $\nu(i,j)$ is accorded to each pixel

$$\nu(i,j) = \exp(-T(i,j))$$

with

$$T(i,j) = d_{m(i,j)-1}(i,j)$$

the maximal distance of each Peer Group, if we would order all $d_l(i,j)$. In this way pixels in a textured area will be less weighed than pixels in a non textured, smooth area, which is natural as we want to reduce colors without removing texture information. The average $T(i,j)$ in a region, T_{avg} is a measure for the smoothness of a region. The bigger T_{avg}, the less smooth an image is, the more clusters are needed. The number of clusters needed can be estimated as $N = \beta T_{avg}$, with β to be defined experimentally, in our case set to 2.
A Generalized Lloyd algorithm is then used with centroid c_0 and distortion D_0 now defined as

$$c_0 = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} x(i,j)\nu(i,j)}{\sum_{i=1}^{M}\sum_{j=1}^{N} \nu(i,j)}$$

$$D_0 = \sum_{i=1}^{M}\sum_{j=1}^{N} \nu(i,j)|c_0 - x_{i,j}|^2.$$

so that centroids are shifted in the direction of points with higher weights end points with smaller weights will be assigned fewer clusters. The clusters are split according to the plane through the centroid perpendicular to the biggest eigenvalue of the cluster. In our experiments we typically had a reduction to 18 colors.

3.5. Classification Algorithms

A Parallelepiped classifier, a maximum distance classifier and a Bayesian MAP classifier for multi-dimensional normal distributions [6] were tested. Well chosen test regions of interest outside of the templates were chosen to train these.

4. RESULTS

On color information alone we found the RGB color space the most performing (up to 97.90 % correctly classified pixels), followed by Lab (96.82%), $YCrCb$ (95.36%), HSV (94.84%). Concerning the classification algorithms used, the MAP classifier (up to 97.90%) outperformed the Parallelepiped (84.77%) and Minimum Distance (up to 90 %) classifier by far. This is expected seen the simple nature of both. There for, all results obtained for texture and color texture are based on the MAP classifier.
When we look at Table 1 we obtain rate above 90% for all template images based on color information. Texture features perform worse, even when a grey value reduction is applied, (not in the table), average classification rates range from 79 to 96% except for template image 3, which is also the less trivial one to classify. Turning to color texture results we obtain results ranging from 71 to 95% without a color reduction. Using a color reduction with PGF leads to increasing classification rates in almost all cases .
Using the co-occurrence matrix different values of d were tested, ranging from 1 to 5. There were no significant difference in classification rates in that. Also different window size in which the co-occurrence matrix was calculated were used ranging from a 6x6, to a 12x12 up to 20x20 and 30x30 window size. As was already mentioned before a grey value reduction was made to 32 grey values, experimentally determined, and a color reduction to 18 colors, based on the parameters extracted from PGF. Figure 2 shows some visual results.

5. DISCUSSION

In this article we set up some experiments on how color texture information could be used in very high resolution image classification tasks. Given the right tuning of the texture extractor parameters, very high classification rates can be obtained. One might argue that simple color features outperform color texture in most cases, this is not always so. The texture feature extractors work on sliding windows which implies that at the boundary of 2 textured areas we get border effects, misclassification up to 10% when the sliding window size is chosen to be very big. This effect can clearly be seen when comparing the boundaries on the upper left 4 images to the others in Figure 2. How we can overcome these effects is currently investigated at our department. The usefulness of a color reduction both in processing speed as in accuracy has been shown using PGF, compare the lower images of Figure 2. Turning to the segmentation of city boundaries we did not succeed as successfully, rates not higher the 70% using texture information. We should remark here that even for the human eye it is not easy to make a correct segmentation in these. The window size and parameters of the co-occurrence matrix here are very crucial. Other techniques based on Gabor Filters are currently investigated [1] on this image.
Next to the classification rate, the fragmentation degree of the classified areas is also of great importance, f.i. if we would want to do object based classification. We clearly notice that color texture classification leads to less fragmented regions, compare template images 1 and 4 trough the different techniques.
As a conclusion we can say that up to now color (or spectral information in general) is a very powerful tool in classifying satellite images. Although in some cases it might not do

	Im. 1	Im. 2	Im. 3	Im. 4
Color	97.9	96.8	90.2	93.4
Tex ws=12	92.2	85.8	68.1	94.9
Tex ws=20	87.7	80.7	67.7	92.4
Tex ws=30	79.7	77.4	64.4	96.7
Col Tex ws=6	85.9	71.0	76.2	75.0
Col Tex ws=12	85.2	79.3	56.7	89.7
Col Tex ws=20	87.0	76.8	66.9	89.4
Col Tex PGF ws=6	96.6	82.3	71.2	67.3
Col Tex PGF ws=12	91.8	76.1	71.2	67.3
Col Tex PGF ws=20	86.0	82.1	63.5	96.9

Table 1. Percentage of pixels correctly classified for the different template images over the different parameters. The window sizes (ws) are also indicated for the texture parameters, 6x6, 12x12, 20x20.

the trick, bad lightning conditions, different textures with same color. Here we have shown that color texture information leads to almost equally high classification rates with less fragmentation.

Fig. 1. The 4 template images, image 1: fields, image 2:city + woods, image 3: city, image 4: fields and city

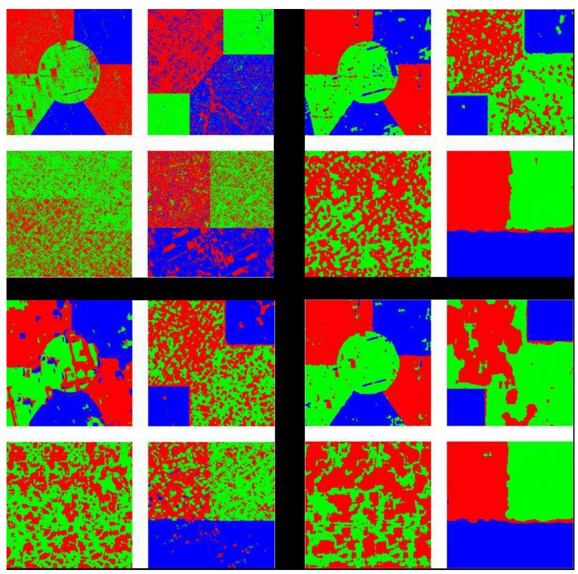

Fig. 2. Upper left : RGB color results, Upper Right: texture results, Lower Left: color texture results, Lower Right: color texture PGF results

6. REFERENCES

[1] Dimai A. Schroder M., "Texture information in remote sensing images: A case study," *Proc. of Workshop on Texture Analysis (WTA '98), University of Freiburg, Germany*, 1998.

[2] K. Shanmugan R.M. Haralick and I. Dinstein, "Textural features for image classification," *IEEE Trans. Systems, Manufact., Cybernet.*, vol. 3, pp. 610–621, 1988.

[3] Parmiggiani F. Baraldi A., "An investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters," *IEEE Trans. on geoscience and Remote Sensing*, vol. 33, pp. 293–403, 1995.

[4] Jaaseainen Lenz R. Hauta-Kasari M., Parkkinen J., "Generalized co-occurrence matrix for multispectral texture analysis," *Proc. of the 13th International Conference on Pattern Recognition (ICPR96), Vienna, Austria)*, 1998.

[5] Moore M.S. Manjunath B.S. Deng Y., Kenney C., "Peer group filtering and perceptual color image quantization," *IEEE International symposium on Circuits and Systems, (VLSI), Orlando, FL*, vol. 4, pp. 21–24, 2001.

[6] Koutrombas K Theodoridis S., "Pattern recognition, elsevier science," vol. 1, 1999.