IMAGE DIFFERENCE HISTOGRAM: A NEW TOOL FOR IMAGE ANALYSIS APPLIED TO CLASSIFICATION OF URBAN SETTLEMENTS

G.Noel [1,2], B.Van Wyk [1,2], A. Van Wyk [1,3]

[1] French South African Institute of Technology, South Africa
[2] Tshwane University of Technology, South Africa
[3] University of the Witwatersrand, South Africa

1. INTRODUCTION

Automatic classification of urban settlement from satellite images has various applications such as urban development planning and the gathering of environmental statistics. Abeigne Ella *et al.* [1] showed that Local Binary Patterns (LBP) features extracted from QuickBird images of the Soweto region in South Africa is superior to other feature extraction methods described in [2],[3], [4], [5], [6]. Following the introduction by Van Wyk *et al.* of the Difference Histogram for feature extraction in time-series [7], the present paper introduces an extension and generalization of this technique to images. The Image Difference Histogram is introduced and applied to the classification of urban settlement images. The results are compared with those presented in [1].

2. IMAGE DIFFERENCE HISTOGRAM

The concept of Difference Histogram for time series is to extract the length of segment of increase in the signal [7]. Unlike the case of time series, the definition of increase (or decrease) in an image requires the definition of a view direction $\vec{d}$. The image I is decomposed in signals $\phi_j(I,d)$ using the view direction $\vec{d}$. If we call Y the operator decomposing the image, then

$$Y\colon I \xrightarrow{d} [\phi_1(I,d),\dots,\phi_N(I,d)] \tag{1}$$

with

$$\vec{d} = \alpha\vec{\imath} + \beta\vec{\jmath} \tag{2}$$

where $\vec{\imath}$ and $\vec{\jmath}$ are the *x* and *y* axis of the image.

The idea of Image Difference Histogram is summarized by the following definitions:

Definition 1. An **Image Difference Histogram**, $\Omega(I,d)$, is defined as a scaled representation of the number of occurrences of the area of **Surfaces of Increase**.

Definition 2. A **Surface of Increase** is the area of neighboring **Segment of Increase**. Two segments are considered neighbors if at least one pixel from one segment has a neighbor in the other segment. The neighborhood depends on the choice of the pixel connectivity.

Definition 3. A **Segment of Increase** is a group of consecutive pixels of $\phi_j(I,d)$ such that

$$\phi_j^{n+1} - \phi_j^n > \phi_j^n - \phi_j^{n-1} - \varepsilon \tag{3}$$

where ε is a **Tolerance Parameter.**

Definiton 4. The **Tolerance Parameter,** ε, is defined as a positive real number chosen to maximize some distance measurement between $\Omega_i(I), i = 1, \dots, C$ are the Image Difference Histograms obtained from images belonging to C different classes.

Each $\phi_j(I,d)$ is computed as follows:

Consider the 3D space $\mathfrak{J}$ with an orthonormal basis $(\vec{\imath}, \vec{\jmath}, \vec{k})$, O being the origin. The following notations are used:

- The usual scalar product is defined on $\mathfrak{J}$ and denoted$\langle\,.\,\rangle$.
- $\mathscr{P}$ is the hyperplane of dimension 2 generated by $(\vec{\imath}, \vec{\jmath})$.
- The image I defines a 3-D surface composes by the points P, where $\overrightarrow{OP} = x\vec{\imath} + y\vec{\jmath} + I(x,y)\vec{k}$. The projection of I on the plane $\mathscr{P}$ is called I_{xy}.
- The contour $\mathfrak{C}$ of I_{xy} is closed. For each point of the contour $\mathfrak{C}$, a vector $\overrightarrow{v_P}$ is defined. $\overrightarrow{v_P}$ is perpendicular to the tangent of $\mathfrak{C}$ at P and points in the opposite direction of the closed contour.

The initial points of $\phi_j(I,d)$ denoted $\phi_j^1(I,d)$are composed by the points of the contours with the following property:

$$\langle \overrightarrow{v_P}.\, \vec{d} \rangle < 0 \tag{4}$$

If $\mathscr{N}$ is the neighborhood defined as the union of the neighboring pixels of $\phi_j^1(I,d)$ belonging to I_{xy}, there are pixels which are in $\mathscr{N}$ and not classified in $\phi_j(I,d)$. For each point P belonging to $\mathscr{N}$ and not to $\phi_j(I,d)$, P belongs to $\phi_K(I,d)$ if the absolute value of the angle between $\overrightarrow{P\phi_K^1(I,d)}$ and $\vec{d}$ is minimal for $K \epsilon [1, \dots, N]$.

The points P are then filed in $\phi_j(I,d)$ and denoted $\phi_j^2(I,d)$. The same process is iterated until all the points all the image have be put in the $\phi_j(I,d)$.

The following algorithm is applied to compute $\Omega(I,d)$:

1: Choose $\vec{d}$

2: $Y: I \overset{d}{\rightarrow} [\phi_1(I,d), \dots, \phi_N(I,d)]$

3: Compute the segments of increase for all the $\phi_j(I,d)$

4: Determine the surface of increase by looking at the area defined by neighboring segments of increase.

5: Compute the histogram $\Omega(I,d)$, the centers of the bins representing the area of the surface of increase.

The information contained in an image will be better extracted if the image is looked at using various points of view or directions, the Image Difference Histogram $\Omega(I,d)$ will the then defined by for example the operators *Hmean*:

$$Hmean: [\Omega(I,d_1),\ldots,\Omega(I,d_M)] \xrightarrow{mean} \Omega(I) \tag{5}$$

which computes the mean of each bins among the $\Omega(I,d_j)$.

The value of the tolerance parameters ε has a great impact on the efficiency of the feature extraction. In our case, the optimal value of ε is determined on a training set, the chosen value being the one maximizing the separation between the histograms. The same reasoning is applied to determine the most suitable number of directions.

3. SIMULATIONS

The simulations are run on QuickBird images of the Soweto are in South Africa. The urban settlement are classified into eight categories, namely Formal Township type 1 (FT1), Formal Township type 2 (FT2), Informal Settlement (IS), Formal Township Informal Settlement type 1 (FTIS1), Formal Township Informal Settlement type 2 (FTIS2), Formal Township Informal Settlement type 3 (FTIS3), Formal Suburbs (FS) and Informal Township (IT).

The experiments showed that the computations of four histograms using four orthogonal directions (typically 0°, 90°, 180°, and 270°) gives proper results while keeping the computational time reasonable.

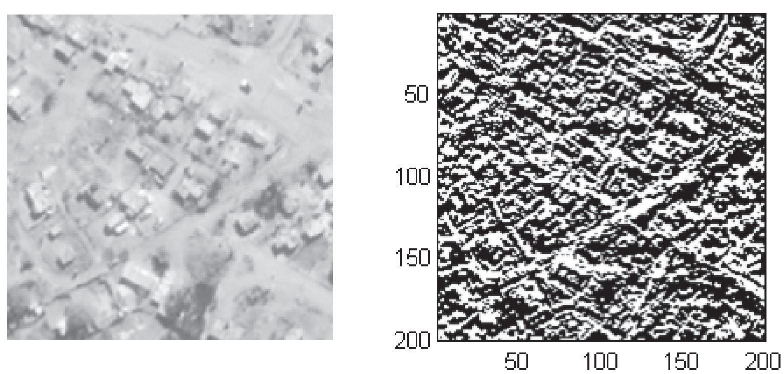

Figure 1: a) Original Image b) Surface of Increase ($\vec{d}=\vec{\imath}$, $\varepsilon = 12$)

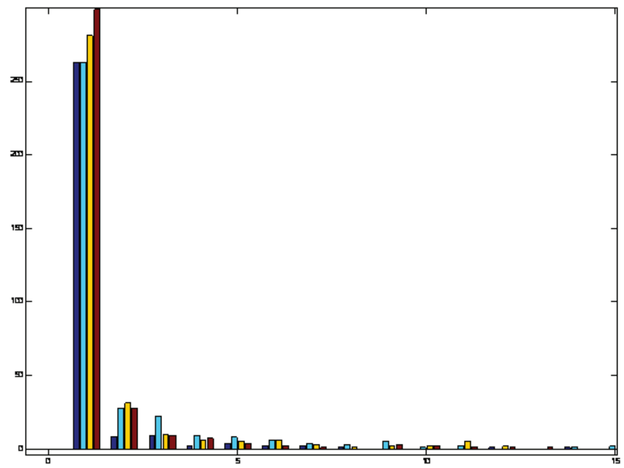

Figure 2: Image Difference Histogram using four view directions $[\Omega(I, d_1), \ldots, \Omega(I, d_4)]$, each color representing one direction.

Fig2. shows the variability of the histogram in respect of the view direction. Using an unsupervised classifier, the classification success for two classes is above 90%.

4. CONCLUSION

The Image Difference Histogram gives a powerful tool to extract features in texture images. Future works includes classifying the features using a properly trained neural network and to apply the method to other texture database. The real-time capability of the method would be highlighted.

5. REFERENCE

[1] L. P. Abeigne Ella, B. J. van Wyk, M.A. van Wyk and F. van den Bergh, “A Comparison of Texture Feature Algorithm Algorithms for Urban Settelement Classification”, *Proceedings of the 2008 IEEE International Geoscience & Remote Sensing Symposium* July 6-11, 2008 | Boston, Massachusetts, U.S.A.

[2] J.A. Benediktsson, M. Pesaresi, and K. Amason, “Classification and feature extraction for remote sensing images from urban areas based on morphological transformations,” *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 41, no.9, pp. 1940–1949, 2003.

[3] M. Pesaresi, “Texture Analysis for Urban Pattern Recognition Using Fine-resolution Panchromatic Satellite Imagery,”*Geographical and Environmental Modelling*, vol. 4, no. 1, pp. 43–63, 2000.

[4] J. Spiker and T. Warner, *Geo-Spatial Technologies in Urban Environments*, pp. 197–213, Springer, 2007.

[5] R. M. Haralick, I. Dinstein, and K. Shanmugam, “Textural features for image classification,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, pp. 610–621, 1973.

[6] Y. Chen and E. R. Dougherty, “Gray-scale morphological granulometric texture classification,” *Optical Engineering*, vol.33, pp. 2713–2722, 1994.

[7] B.J. Van Wyk, M.A. Van Wyk, and G.Qi, “Difference histograms: a new tool for time series analysis applied to bearing fault diagnosis,” *Pattern recognition letters*, 30 (2009) 595-599, 2009.