Wavelet based Marker-Controlled Watershed Segmentation Technique for High Resolution Satellite Images

Imdad Rizvi
Satellite Image Processing Research Laboratory
CSRE, Indian Institute of Technology Bombay
Mumbai 400076, India

B K Mohan
Satellite Image Processing Research Laboratory
CSRE, Indian Institute of Technology Bombay
Mumbai 400076, India

ABSTRACT
Image analysis requires a segmentation step to distinguish the significant components of the image, i.e., the foreground from the background. As a step prior to image classification the quality of the segmentation is of significant importance. High-resolution satellite image classification using standard per-pixel approaches is difficult because of the high volume of data, as well as high spatial variability within the objects. One approach to deal with this problem is to reduce the image complexity by dividing it into homogenous segments prior to classification. This has the added advantage that segments can not only be classified on basis of spectral information but on a host of other features such as neighborhood, size, texture and so forth. Segmentation of the images is carried out using the region based algorithms such as marker-based watershed transform by taking the advantage of multi-resolution and multi-scale gradient algorithms. This paper presents an efficient method for image segmentation based on a multi-resolution application of a wavelet transform and marker-based watershed segmentation algorithm. It also addresses the issue of excessive fragmentation into regions of watershed segmentation, which is avoided by the multi-resolution analysis fact. The most significant components perceived in the highest resolution image will remain identifiable also at lower resolution. Hence the seeds for watershed segmentation on the lower resolution levels are identified and then used them to identify the significant seeds in the highest resolution image. Experimental result of proposed technique gives promising result on QuickBird images. It can be applied to the segmentation of noisy or degraded images as well as reduce over-segmentation.

General Terms
Algorithms, Measurement and Design.

Keywords
Multi-resolution Analysis, Image Segmentation, Watershed Transform, High resolution satellite image.

1. INTRODUCTION
Image segmentation is the partitioning of an image into meaningful regions based on homogeneity or heterogeneity criteria, respectively [1]. It is the process by which an image is decomposed into several regions such that each region is homogeneous with respect to its tone, texture or color and adjacent regions are bounded by an edge that is indicative of strong difference in the property characterizing each region. The goal of the segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. With the advances in geo-spatial technology and spaceborne imaging, extracting the target information from the high resolution remote sensing image has become a challenging problem. Features that can be extracted from imagery include facilities (e.g., buildings, roads, waterbodies etc.), transportation features, land use/land cover, vegetation types (agriculture, forest types), disturbed features due to exploration activities, water bodies, oil spills in water, quarrying sites, soil types, and more. However, the performance of classical segmentation techniques are limited by various factors like shadow problems, within-object variability and noise leading to over-segmentation at some places, under-segmentation in others, resulting in poor overall accuracy. To overcome these problems, analyzing the images at multiple resolution levels or multi-resolution is being considered (excluding shadow problems). Primarily due to the progresses in spatial resolution of satellite imagery, the methods of segment-based image analysis for generating and updating geographical information are becoming more and more important [2]. Until now, various techniques and algorithms have been proposed for image segmentation. They are classified into three main categories: clustering, edge detection and region extraction [3][4]. Most segmentation techniques are based on this third category [5][6]. Other conventional segmentation approaches range from the split-and-merge method to morphological segmentation. Among them, morphological segmentation techniques are of particular interest because they rely on morphological tools, which are very useful to deal with object-oriented criteria such as size and contrast [7]. But a major problem with the watershed segmentation algorithm is that it produces severe over-segmentation due to the great number of minima and various noise within an image or its gradient. There are two main drawbacks when applying the watershed algorithm to image segmentation: sensitivity to strong noise and high computational requirements to merge the over-segmented regions. To overcome these problems, a strategy was proposed by Meyer and Beucher [8]. The strategy is called marker-controlled segmentation and is based on the idea that machine vision systems often roughly ‘‘know’’ from other sources the location of the objects to be segmented. These problems can be overcome when this segmentation algorithm is integrated within a multi-resolution approach [9]. A second difficulty is constructing object regions from their boundaries. Watersheds

have also been used in multi-resolution methods to produce resolution hierarchies of image ridges and valleys. Although these methods were successful in segmentation certain classes of images, they require significant interactive user guidance or accurate prior knowledge on the image structure [7].

A method to reduce over-segmentation in watershed partitioned images based on the use of a multi-resolution representation of the input image is also proposed by Frucci *et al.* (2005). The underlying idea is that the most significant components perceived in the highest resolution image will remain identifiable also at lower resolution. Thus, starting from the image at the highest resolution, a multi-resolution representation by building a resolution pyramid was obtained. Then, seeds were identified for watershed segmentation on the lower resolution pyramid levels and suitably use them to identify the significant seeds in the highest resolution image [12].

A robust image segmentation method is proposed which is based on multi-resolution analysis and marker-controlled watershed segmentation algorithm. In the presented approach, over-segmentation and noise related problems are minimized as the watershed operation is carried out on the low-pass filtered low-resolution images from the wavelet transform. In turn, the computational complexity is simplified and reduced dramatically by operating on a low-resolution image. After image segmentation, the segmented low-resolution image with label is projected at full-resolution (original) image by an inverse wavelet transform. Experimental results show that the proposed method is effective in the segmentation of images. The description of methodology adopted in the formulation and data used in the present work are discussed in Section II. The results obtained by implementing the methodology proposed are presented in Section III. Summary and conclusions of the present work are presented in Section VI. Last section contains the conclusions drawn from this work and scope for further work in this area.

2. METHODOLOGY

Image segmentation based on watershed transformation can potentially provide more accurate segmentation with low computational cost. It got its name from the manner in which the algorithm segments regions into catchment basins. There are two basic approaches to watershed image segmentation.

1. The first one starts with finding a downstream path from each pixel of the image to a local minimum of the image surface altitude. A catchment basin is then defined as the set of pixels for which their respective downstream paths all end up in the same altitude minima.

2. The second approach is dual of the first one; a local minimum is identified for each region and then the topographic surface is immersed in water; water starts filling all the catchment basin. If two catchment basins would merge as a result of further immersion, a dam is built to prevent it.

2.1 Marker-Controlled Watershed Transform

Watershed applied directly to the gradient image results in over-segmentation due to irreverent minima or noise patches or other image irregularities. The concept of markers can be used to solve this over-segmentation problem whose goal is to detect the presence of the homogenous regions from the image by a set of morphological simplifications. They spatially locate object and background ensuring to keep up the interior of the object as a whole. The Markers are connected components belonging to an image [8]. Image objects resulting from segmentation represent image object primitives, serving as information carriers and building blocks for further classification or other segmentation processes. Each image object has a large number of characteristic properties; the so-called object features or attributes. In this sense, the best segmentation result is that which provides optimal information for further processing. This has the added advantage that segments can not only be classified on basis of spectral information but on a host of other features such as neighborhood, size, texture and so forth. Segmentation of the images is carried out using morphological marker-controlled watershed transform by employing the advantages of multi-resolution analysis and multi-scale gradient algorithms. The segmentation of the color images is obtained using watershed transform to get its homogenous regions. The proposed algorithm is given below

- Apply multi-resolution framework to input image. The Daubechies wavelet is considered because it's an orthogonal transform, compact support and requires small computational complexity. Daubechies wavelet uses overlapping windows, so the high frequency coefficient spectrum reflects all high frequency changes. Therefore Daubechies wavelet are useful in compression and noise removal of audio signal processing. Daubechies 4-tap wavelet has been chosen for this implementation.
- Image is observed at different resolutions, only the most significant regions will be perceived at all resolutions, even if in a more coarse way at lower resolution. Regions that, at the highest resolution image, can be interpreted as noise or constitute fine details are generally not preserved when resolution decreases. Thus, if the seeds for watershed segmentation of the highest resolution image are identified in a lower resolution level, the resulting partition is expected to be characterized by a reduced number of regions, corresponding to the most significant image parts [12].
- Seeds at one of the lower resolution levels are identified. These seeds are suitably projected onto the highest resolution level of the pyramid and are used to select among the seeds originally detected at that resolution, only those corresponding to the most significant regions. All other seeds originally found in the highest resolution image undergo a suitable removal process, aimed at merging the corresponding partition regions. The watershed segmentation of the highest resolution image is finally accomplished, by using only the seeds that survived the removal process.
- Use multi-scale gradient algorithms to calculate color gradient.
- Consider morphological gradient of each band of the image calculated using equation (1)

$$G(f) = (f \oplus B) - (f \Theta B) \tag{1}$$

where $G(f)$ = Morphological color gradient,

f = Given image

B = Structuring element.

The multi-scale morphological color gradient is then calculated using the formula

$$MG(f) = \frac{1}{n}\sum_{i=1}^{n}[G(f)\Theta B_{i-1}] \quad (2)$$

where

B_{i-1} = Structuring element of size (2i+1)x(2i+1)

The multi-scale morphological color gradient is dilated with a square structuring element of size 2x2. The constant h is then added to the dilated image. A final gradient image, $FG(f)$ is obtained by reconstructing the multi-scale gradient image, $MG(f)$ with its dilated image as a reference image.

$$FG(f) = \Phi^{rec}\left(\left(MG(f) \oplus B\right) + h, MG(f)\right) \quad (3)$$

The markers are the interior of the objects of the interest and the backgrounds are to be extracted from the image to get the segmented image. Too many markers results in over-segmentation and too few markers results in under-segmentation. The markers can be extracted from white top-hat or black top-hat transform. But extracted markers from either white or black top-hat will miss some of the objects. So, to utilize the advantage of both top-hat, markers are extracted using morphological laplacian, which can be defined as:

$$L(f) = g^{+}(f) - g^{-}(f) \quad (4)$$

where $g+(f)$ = White top hat transform and

$g-(f)$ = Black top hat transform

For utilizing the spectral property of the image, markers are extracted from morphological color laplacian of the image; and is calculated using equation (5)

$$L(f) = \sqrt{\left(L_r(f)^2 + L_g(f)^2 + L_b(f)^2\right)} \quad (5)$$

where $L(f)$ = Morphological color gradient,

$L_r(f)$ = Gradient of the red band,

$L_g(f)$ = Gradient of the green band and

$L_r(f)$ = Gradient of the blue band.

- The markers extracted from image using morphological laplacian, are labeled using connected component labeling. Each connected marker is assigned with a unique label.
- Morphological Watershed segmentation algorithm applied to the image $C(f)$ which is obtained from the marker image $M(f)$ and final multi-scale morphological color gradient image $FG(f)$. For any pixel p at position (i,j), C is obtained by
- As marker image $M(f)$, provides rough partition of the objects and the final gradient image $FG(f)$ avoids over merging. Thus, average of the marker and final gradient multi-scale morphological color images preserves the contours of the object. The algorithm used for the implementation of the watershed transform is ordered queue as it based on maker-controlled watershed transform and easy to implement.
- Marker-controlled watershed transform algorithm is used for region segmentation [10].
- Region merging is done to avoid over-segmentation.
- Each segmented object or region is assigned the average grayscale of each band to generate the mosaic color image.
- To get the final segmentation at high resolution image; low frequency coefficient of the wavelet is replaced with mosaic image; while detailed coefficients of the wavelet are modified so as to avoid noise introduced back into the finer image. Inverse wavelet transform is then applied on these modified images to get the high resolution segmented image.
- The output of the watershed transform may result in over-segmentation. To merge the adjacent region or the homogenous regions; region merging using criterion is implemented.

2.1 Study Area

The study areas used in present study are subsets of Mumbai and Rome City images from QuickBird satellite image which are illustrated as Study Area 1 and Study Area 2 respectively. Its spatial resolution is of 0.61m for the panchromatic layer and 2.44m for the multi-spectral ones with an area of 2000x2000 and 228x434 pixels as shown in Figure 1 and Figure 2 respectively.

Figure 1. High resolution satellite image of Study Area 1

Figure 2. High resolution satellite image of Study Area 2

3. RESULTS

Gray scale image of both the study areas are shown in Figure 3 and Figure 4.

Figure 3. Gray scale image of Study Area 1

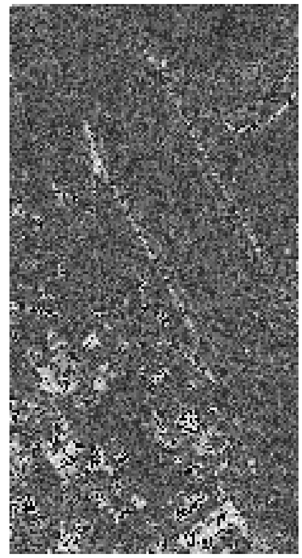


Figure 4. Gray scale image of Study Area 2

Noise removal and generation of low resolution image are done by Daubechies wavelet transform. The multi-resolution image is generated by the two-scale Daubechies 4-tap wavelet transform and the markers for the watershed segmentation algorithm were extracted from a low-resolution image. Three level decomposition is chosen. Markers extraction from morphological color laplacian is done by the threshold from -30 to 40. Region merging is done with size less than 48 pixels per region and the maximum mean difference allowed for region merging is 8. The number of regions reduced to 1654 from 2568. To get the segmentation results at original high resolution image, inverse wavelet transform is then applied on the each mosaic image. Here watershed segmentation is affected by excessive fragmentation into regions. This, besides requiring a suitable complex process to reduce the number of markers from which the partition originates, may bias the successive assignment of the partition regions to the foreground and the background.

The regional minima found in ∇_k are generally used as the seeds starting from which watershed transformation generates a partition of ∇_k (and, hence, of G_k) into regions characterized by some gray-level homogeneity. We note that the image at level 1 is affected by excessive fragmentation, caused by the very large number of regional minima. Some (heavy) process is generally accomplished to select among the seeds found in ∇_1 only those that are significant to correctly partition G_1 as shown in Figure 5. Since G_1 is well represented even at the lowest resolution level of the pyramid and, in turn, the seeds found in ∇_3 are considerably less than those found in ∇_1, we will use the seeds found in ∇_3 to select among the seeds detected in ∇_1 the most significant ones and obtain, in this way, a less fragmented partition of G_1 as shown in Figure 6.

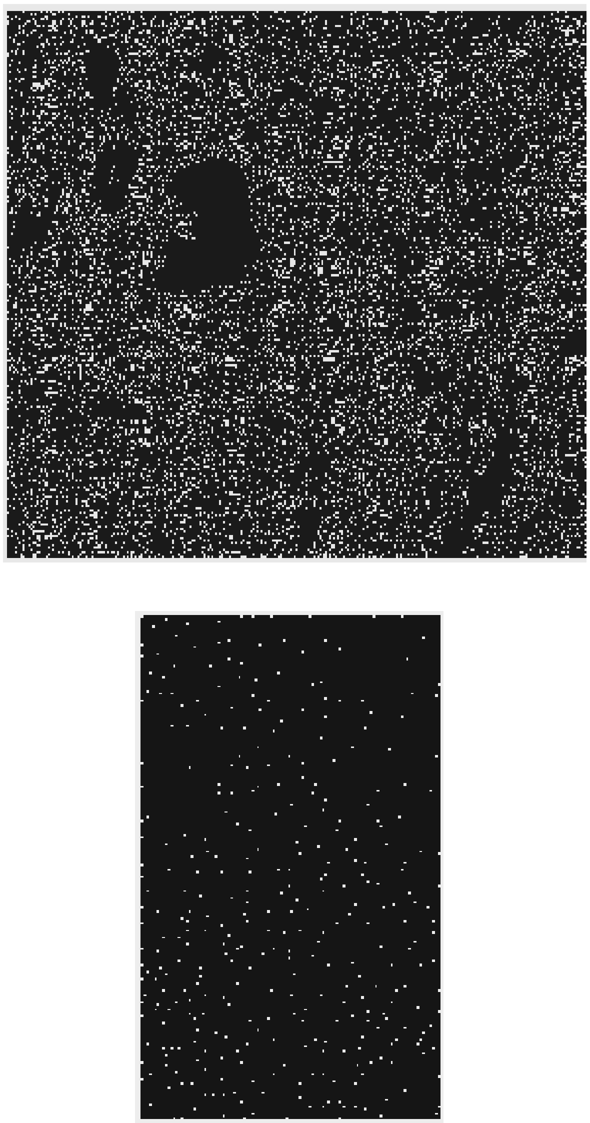


Figure 5. Seeds of L1. Study Area 1(Top) and Study Area 2 (Bottom)

To this aim, we project the seeds from level 3 to level 1. This is possible due to the fact that our pyramid construction method preserves the links parent-children. Thus, for each pixel at level 3 we can easily identify its descendants at level 1. Obviously, since any parent pixel at level 3 has four children at level 2 and each of these children has in turn four children at level 1, for each seed pixel found in ∇_3, which is obvious from Figure 7.

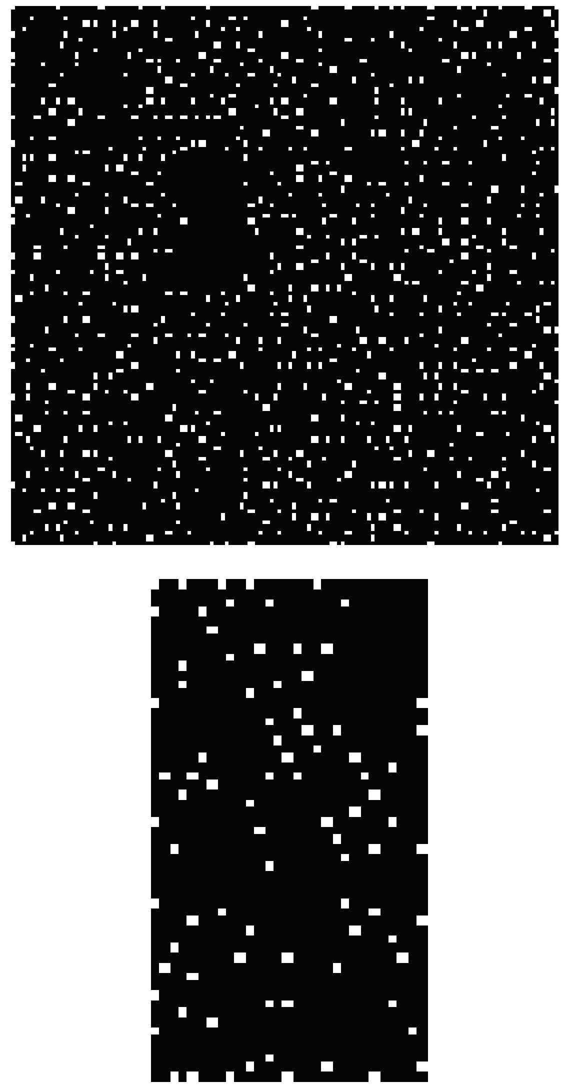
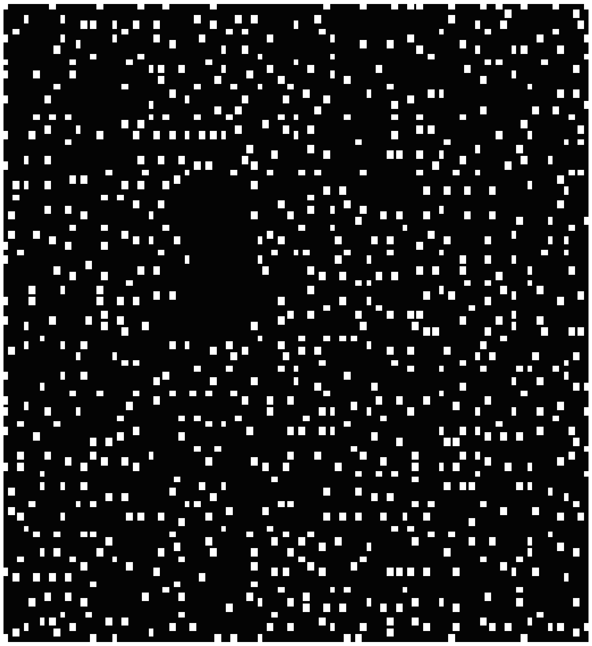


Figure 6. Seeds of L3. Study Area 1(Top) and Study Area 2 (Bottom)

Our idea is to regard as significant a seed originally detected in ∇_1 as shown in Figure 5 only provided that its associated partition region includes at least one descendant of the seeds found at level 3 as shown in Figure 6. All other seeds detected in ∇_1 are regarded as non significant and, by means of a flooding process, the corresponding partition regions are merged. Due to the large size of the sets of descendants originated from the seeds found at level 3, still too many seeds would be preserved at level 1. To reduce their number, we do the following process. Let M be the number of connected components of descendants, CCDi, found at level 1. In the gradient image ∇_1, we inspect the M sets Ci of pixels with homologous positions with respect to the pixels of the sets CCDi. In each set Ci, we identify and preserve as seeds only the pixels, whose gray-level is minimal with respect to the gray-levels of the other pixels of Ci.



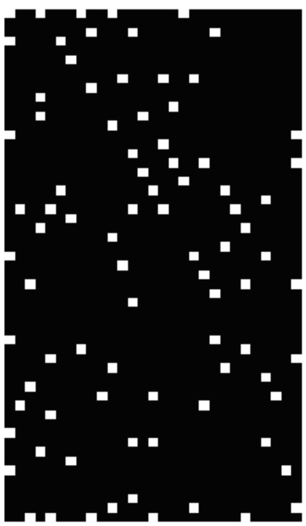


Figure 7. Seeds of L3 projected on L1. Study Area 1(Top) and Study Area 2 (Bottom)

All other descendants are removed (Figure 7). Flooding is then applied at level 1 to merge all partition regions that do not include at least one descendant that survived the removal process. The watershed lines of the partition of G_1, obtained by using the seeds found at level 3 to identify the significant seeds at level 1, are superimposed in white on G_1 in Figure 7 for Study Area 1 and Figure 8 for Study Area 2. By selecting a different lower resolution level, we can use the seeds found there to identify the significant seeds among those detected at level 1. The segmentation results obtained using proposed algorithm are better and visually appealing. Thus, it can be concluded that using the multi-resolution framework for classification and segmentation of the image is a desirable approach for high resolution images. The proposed algorithm is implemented in C++ language on Windows XP platform, and has been successfully tested with various multi-spectral images.

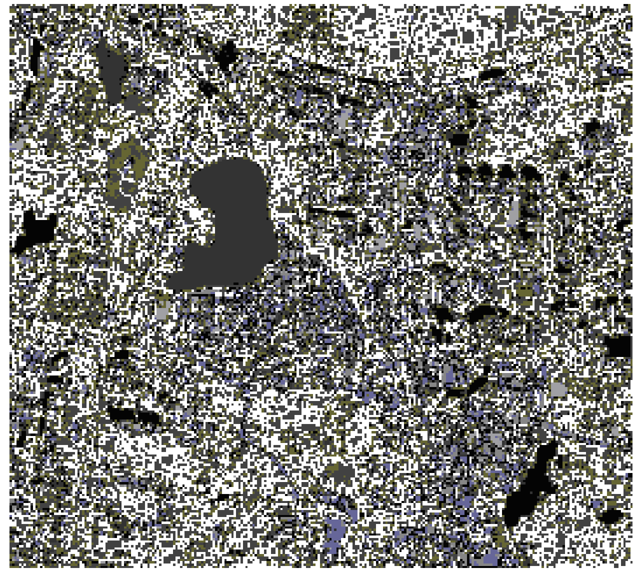


Figure 7. Segmented image of Study Area 1

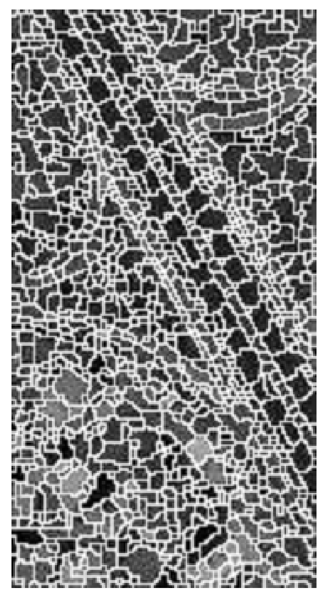


Figure 8. Segmented image of Study Area 2

Evaluation of segmentation results is done on number of segmented regions, PSNR, Goodness function *F*, which is defined by Liu and Yang [11]

$$F(I) = \sqrt{M} \times \sum_{i=1}^{M} \frac{e_i^{\,2}}{\sqrt{A_i}} \tag{6}$$

where *I* is the image to be segmented, *M* is the number of regions in the segmented image, A_i is the area or i^{th} region number of pixels and e_i is the sum of the Euclidean distance of the color vectors between the original image and the segmented image of each pixel in the region. Equation (6) is composed of two terms: the first term, $\sqrt{M}$, penalizes segmentation that forms too many regions; the second penalizes segmentations containing non-homogeneous regions. A larger value of e_i indicates that the feature of the region was not well segmented during the image segmentation process.

3.1 QUALITY ASSESSMENT

General criterions, like the delineation of varying land cover types, segmentation of linear objects, occurrence of faulty segmentations and a description of the overall segmentation quality were in the focused by visual survey. A detailed comparison based on visual delineated and clearly definable reference areas was carried out. Therefore 20 different areas (varying in location, form, area, texture, contrast, land cover type etc.) were selected and each was visually and geometrically compared with the segmented pendants. The geometrical comparison is a combination of formal factors (area, perimeter, and shape index) and the number of segments. Additionally the quality of segmentation was visually rated (0 poor, 1 medium, 2 good). A good segmentation quality is reached, when the overall differences of all criteria between the segmentation results and the associated reference objects are as low as possible. Furthermore the objects of interest should not be over-segmented too much. To quantify the fit of each of the reference objects with the largest segment overlapping them, the Area-Fit-Index (AFI), was introduced by [13] and is defined as follows:

$$AFI = \frac{A_{reference\ object} - A_{l\arg est\ object}}{A_{reference\ object}} \quad (7)$$

where A is the Area of the reference and the largest segment of the result respectively. If the AFI equals zero a perfect segmentation is indicated. The AFI range can be from [-∞, 1] and the value of AFI= -0.16 for 20 reference objects.

4. SUMMARY AND CONCLUSIONS

In this paper we have evaluated wavelet based marker-controlled watershed segmentation technique for high resolution satellite images. The proposed technique is based on the use of a multi-resolution representation of the input image and on the selection of markers for segmenting the highest resolution image, guided by the markers found at lower resolution. Then, the seeds for watershed segmentation are identified on one of the lower resolution pyramid levels and suitably use them to identify the significant seeds in the highest resolution image. The procedure toward complete segmentation consists of various steps like creating multi-resolution images using Daubechies wavelet transform, image segmentation using a marker-controlled watershed segmentation algorithm, merging of the segmented regions. Then, markers are identified for watershed segmentation on the lower resolution levels and suitably used to identify the significant markers in the highest resolution image. The experimental results indicate that the over-segmentation problem, which is typical of the watersheds technique, can be significantly attenuated by use of wavelet transform. False contours due to low-contrast edges within the regions of interest are also effectively reduced with proposed technique. It is robust when applied to noisy and/or blurred images, performing better than other segmentation techniques proposed in the literature. The post-processing stage eliminates effectively the remaining over-segmented regions. This image is finally partitioned by watershed segmentation, providing a satisfactory result. Since different lower resolution can be used to identify the markers at the highest resolution, an alternative segmentation is obtained of the highest resolution image, among which the user can select the best suited one for the specific task.

5. SCOPE FOR FUTURE WORK

As a future watch we will concentrate on choosing an adaptive threshold. We must also note that beyond purely spectral information, image objects contain a lot of additional attributes which can be used for classification and this method is more suitable and will be the trend for the high resolution remotely sensed data. Object-based approach has the advantage to produce compact objects which correspond to human eye perception of the environment and it reduces the variance problem of very high resolution satellite data. Therefore, dividing the image into regions and then opting for classification is better than per pixel classification. The segmentation results obtained are not able to preserve the boundaries of the regions. It also provides possibilities to bring in additional knowledge on the image objects of interest, on object inter-relations and relations to external map or GIS information. The further work can be done in a same direction to preserve the boundaries of the regions. We also intend to further investigate the application of proposed technique hyper-spectral images, which comprises of more then 200 bands.

6. ACKNOWLEDGEMENTS

Authors are grateful to Digital Globe Inc., USA for granting permission to use the browse images on their website for this research work.

7. REFERENCES

[1] Haralick, R. M.; Shapiro, L. G. (1992): *Computer and robot vision*. Vol. I, Addison-Weasley, Reading, 672 p.

[2] Meinel, G.; Neubert, M. (2004): A Comparision of segmentation programs for high resolution remote sensing data. *Int. Arch. of Photogrammetry, Remote Sensing and Spatial Information Sciences* XXXV-B4, pp. 1097-1102.

[3] Gonzalez, R.C., Woods, R.E., 2002. *Digital Image Processing*, 2nd ed. Prentice-Hall, Reading, NJ, USA.

[4] Pal, N.R., Pal, S.K., 1993. A review on image segmentation techniques. *Patt. Recognit*. 26 (9), 1277–1294.

[5] Bhandarkar, S.M., Hui, Z., 1999. Image segmentation using evolutionary computation. *IEEE Trans. Evolut. Comput*. 3 (1), 1–21.

[6] Kim, H.J., Kim, E.Y., Kim, J.W., Park, S.H., 1998. MRF model based image segmentation using hierarchical distributed genetic algorithm. *IEE Electron. Lett*. 34 (25), 1394–1395.

[7] Kim, J.B., Kim, H.J., 2003. Multi-resolution -based watersheds for efficient image segmentation. *Patt. Recogni Lett.* 24, 473-488.

[8] Beucher, S., Meyer, F., 1993. The morphological approach to segmentation: the watershed transformation. In: Dougherty, E. (Ed.), *Mathematical Morphology in Image Processing*. Marcel Dekker, New York.

[9] Wang, J.Z., Li, J., Gray, R.M., Wiederhold, G., 2001. Unsupervised multi-resolution segmentation for images with low depth of field. *IEEE Trans. Patt. Anal. Mach. Intell*. 23(1), 85-90.

[10] Meyer, F,. and Beucher, S., 1990, "Morphological Segmentation," *Journal of Visual Communication and Image Representation*, v. 11, p. 21–46.

[11] Liu, J., Yang, Y. H., Multi-resolution color image segmentation, *IEEE Trans. Patt. Anal. Intell*. 16 (7), 689-700, 1994.

[12] Frucci, M., Ramella, G. and Sanniti di Baja, G., 2005, Over segmentation reduction via multiresolution image representation, *Pattern Recognition, Image Analysis and Applications*, LNCS 3773, pp. 989–996.

[13] Lucieer, A., 2004. Uncertainties in Segmentation and Their Visualisation. PhD Thesis, *Utrecht University, ITC* Dissertation 113, Enschede, 174 p. http://www.itc.nl/library/Papers_2004/phd/lucieer.pdf (Accessed on 18th November 2010)