

EFFECT OF WAVELET COMPRESSION ON THE AUTOMATIC CLASSIFICATION OF URBAN ENVIRONMENTS USING HIGH RESOLUTION MULTISPECTRAL IMAGERY AND LASER SCANNING DATA

J. B. K. KIEMA

Institute for Photogrammetry and Remote Sensing (IPF)
University of Karlsruhe, Germany
kiema@ipf.uni-karlsruhe.

KEY WORDS: Data Compression, Wavelets, Data fusion, Classification.

ABSTRACT

This paper examines the influence of data fusion and wavelet compression on the automatic classification of urban environments. The principal data used is airborne Daedalus scanner imagery. Laser scanning data is introduced as an additional channel alongside the spectral channels thus effectively fusing the local height and multispectral information. The feature base is expanded to include both spectral (e.g., spectral signature and texture) and spatial features (e.g., shape, size, topology etc.). This enables the incorporation of context information into the feature extraction. A maximum likelihood classification is then applied. It is demonstrated that the classification of urban scenes is significantly improved by integrating multispectral and geometric datasets. The fused imagery is then systematically compressed (channel by channel) at compression rates ranging from 5 to 100 using a wavelet-based compression algorithm. The compressed imagery is then classified. Analysis of the results obtained indicates that a compression rate of up to 20 can conveniently be employed without adversely affecting the segmentation results.

1 INTRODUCTION

Besides cartographic generalisation, map updating defines a basic problem in cartography that is yet to be solved in a satisfactory manner (Bähr and Vögtle, 1999). The automatic classification of object features from remotely sensed imagery may be viewed as a contribution towards the solution of this problem. This is of particular interest in urban environments, especially given the high concentration of man-made features in such areas. That most of the research effort in the automatic segmentation of geospatial imagery is currently focused on man-made features in general, and buildings and roads in particular, testifies to this fact. However, in view of their size, structure, spectral characteristics, as well as simple diversity, the segmentation of man-made features defines a unique problem.

Airborne laser scanning (ALS) is well suited for the production of Digital Surface Models (DSMs). The geometric information contained in this data can be used to support the discrimination between objects that are projected higher than the terrain (e.g., buildings, trees etc.) from those that are basically at terrain level (e.g., streets, gardens, parks etc.). Although ALS data provides a rich source for geometric information, the use of this data alone is nonetheless, of limited applicability in the extraction of urban objects. In order to enhance its value and thereby exploit its full potential, it is often necessary to integrate this with other data sources e.g., multispectral data.

Data compression is routinely employed in remote sensing within the context of transferring recorded spatial data from satellite sensors to ground stations, as well as in the ultimate storage of this data. In addition to this, the examination of the influence of data compression techniques on the quality and further processing of satellite imagery represents an important and contemporary research and development topic (Shiewe, 1998). In this regard, (Fritz, 1997) observes that compression rates above a factor of 10 will be required in order to handle future high-resolution commercial sensors.

The field of wavelets has opened up new opportunities for the compression of remotely sensed data. In contrast to the standard JPEG lossy compression technique which works best on "continuous tone" (homogeneous) data, wavelet compression is well suited for data which is characterised by sharp spectral discontinuities. It can be argued that by virtue of being comprised of a mixture of different spatial objects (e.g., buildings, streets, gardens, water-fountains etc.), urban environments often result in heterogeneous data. This type of data is therefore best compressed using wavelet-based compression schemes.

Against the above background, the main objectives of the study presented here are twofold. Firstly, to consider within the perspective of data fusion the automatic classification of urban environments from high-resolution remotely sensed imagery and ALS data. And secondly, to investigate the influence of wavelet compression on the classification accuracy of urban objects.

2 CLASSIFICATION OF URBAN ENVIRONMENTS

2.1 Study area and data used

The selected test area for this study covers an area of about 400m x 400m in the city of Karlsruhe which lies in the south-western part of Germany. This includes the famous Karlsruhe castle and part of the main campus of the University of Karlsruhe. The principal data used is airborne Daedalus scanner imagery. Marketed under the acronym Daedalus ATM (Airborne Thematic Mapper), this is basically an opto-mechanical line scanner with 11 different spectral channels and 8 bit data quantisation ideally developed to match the successful Landsat TM. The system specifications for this sensor are given in (Kramer, 1996). In addition, ALS data is also employed. In general, two ALS sensor types can be distinguished: pulse and continuous wave laser scanning systems. The operating principle and basic system overview for ALS is discussed in (Wehr and Lohr, 1999).

2.2 Fusing the different datasets

Data fusion is defined as a formal framework in which are expressed means and tools for the alliance of data originating from different sources (Wald, 1999). This aims at obtaining information of greater quality. Hence, data fusion ought not to be viewed as a mere collection of tools and means but rather, a comprehensive framework through which the integration of spatial data is realised. Although the concept of data fusion is easy to understand, its exact meaning and use often varies from one scientist to another. Different approaches to this may theoretically be employed (e.g., RGB colour composites, IHS transformation, Principal Component Analysis, wavelets etc.) (Pohl, 1999). The particular method adopted depends on several factors including: the structure of the data to be fused; the specific image characteristics that need to be enhanced or preserved etc.

Two different approaches to the fusion of multi-sensor data may be adopted for applications like the one described in this study. It is possible to use either the hierarchical classification approach or the additional channel concept. The hierarchical or layered classification approach is basically a structured technique through which the different datasets to be fused are applied in such a way as to successively divide the working area into more detailed object classes (Savian and Landgrebe, 1991). In principle, this begins with basic object classes before progressively zeroing in on more detailed ones. On the other hand, in the additional channel method the different datasets to be fused are introduced as separate channels in an integrated fashion within an expanded dataset. This of course necessitates the co-registration of the different datasets to a uniform georeference system.

Comparing the above two methods, the main disadvantage of the hierarchical classification approach is the propagation of the classification errors in the subsequent classification steps. On the other hand, the main advantage of the additional channel procedure is the enhanced flexibility in the data processing. For this reason, the additional channel concept is adopted for the data fusion in this study. Hence, the normalised ALS data is introduced as an additional channel alongside the multispectral channels of the Daedalus ATM imagery.

2.3 Classification Approach

spectral features	spectral signature texture
spatial features	structure size shape/contour topology

Table 1: Variables for integrated image segmentation

Spectral information is conventionally employed in the classification of multispectral imagery. The inadequacy of this in the extraction of urban objects has been acknowledged in many studies ((Gong and Howarth, 1990); (Fung and Chan, 1994); (Haala and Brenner, 1999)) etc. In applications where high accuracy is required, it is imperative to enhance the object feature base. This needs to be expanded to include both spectral and spatial feature characteristics ((Schilling and Vögtle, 1996); (Bähr and Vögtle, 1998)). Table 2 shows some of the variables that need to be considered. These may be included either, explicitly through the integration of

different multi-sensor data source(s) or, implicitly through the use of appropriate segmentation methods. For example, it

is possible to implicitly incorporate topological information through the use of triangulation segmentation procedures e.g., Delaunay triangulation.

Different classifiers may be employed in the supervised classification of remotely sensed data, namely: maximum likelihood, minimum distance and parallelepiped classifiers (Lillesand and Kiefer, 1994). For the study presented here, a maximum likelihood classification approach is adopted. The selection of the training data is basically done using manual digitising. Five basic urban object classes are identified: *Buildings*, *Pavement*, *Trees*, *Grass-covered areas* and *Special*. The class *Special* is introduced in order to take care of the many miscellaneous urban objects of limited dimension that are imaged (e.g., vehicles, tents, water fountains, sculptures etc.). Because of diverse spectral diffusion in the multispectral data depending on different sunlight conditions, the digitisation of the training areas is often a very time consuming exercise. Consequently, it would be preferable to procure the training data automatically using existing GIS databases as proposed in (Walter, 1998).

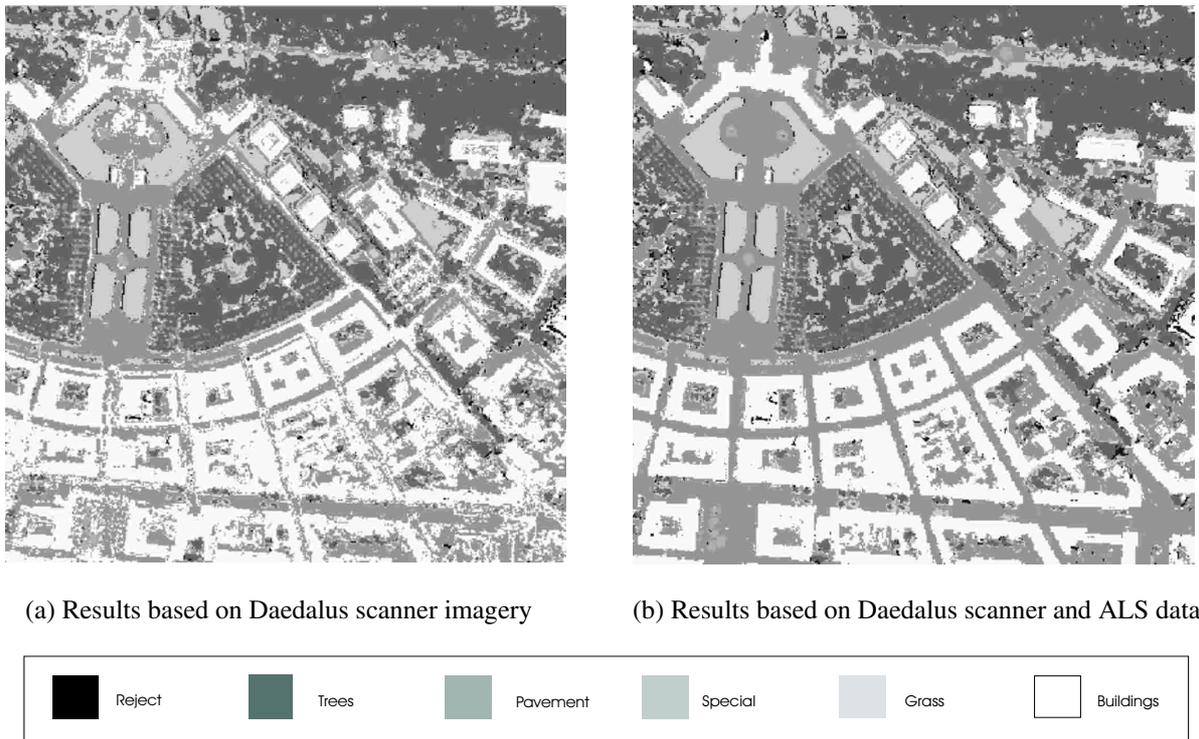


Figure 1: Classification results

For the purpose of this study, the elimination of shadow effects in the classification is carried out using an approach similar to that described in (Haala and Walter, 1999). Firstly, the shadow areas on the imagery are automatically identified. A separate class *Shadow* is then introduced for each of the a priori defined object classes. This means that each object class is divided into one separate subclass for shadow areas and another for non-shadow areas. Separate training data is digitised for each subclass. The classification is then carried out after which the shadow and non-shadow subclasses for each object class are combined. This results in one unique class for each a priori defined object class. Figs. 1a and b compare the final classification results in the absence and presence of the ALS data respectively. A comparison of these figures clearly demonstrates the improvement realised in the classification upon fusing the multispectral and geometric datasets. In particular, the improved ability to discriminate between the low lying urban object classes (e.g., pavements, grass-covered areas etc.) from features that are significantly above the terrain (e.g., buildings, trees etc.) is noted.

Triangulation-based segmentation methods are usually employed in mid-level image processing procedures in order to combine structured image regions into semantically homogeneous clusters. In general, these methods use the delaunay triangle as the basic image segmentation primitive. Through this, the topological relationship between the image segments is implicitly exploited. This is in contrast to traditional segmentation methods which use the spectral information of the pixels. Details on spatial tessellations in general, and delaunay triangulation in particular, are discussed in (Okabe *et. al*, 1992). In order to smoothen the contours of the identified segments as well as minimise the effect of noise, use is also made of the connected components technique and morphological operations. A more comprehensive treatment of mathematical morphology and connected components is given in (Serra, 1986) and (Haralick and Shapiro, 1992) respectively. Segments are extracted for the different urban object classes. A delaunay

segmentation is then applied. Fusing the valid object segments for the class *Buildings* results in Fig. 2. In general, the delaunay triangulation results compare very well with similar results obtained using the first segmentation approach and can therefore be used in the automatic verification of GIS databases.

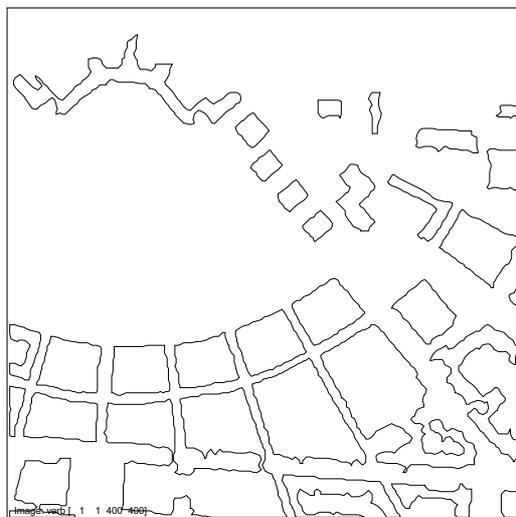


Figure 2: Selection and fusion of valid building segments

3 WAVELET COMPRESSION

Applications of wavelet analysis in remote sensing are rapidly increasing. The main areas of this include: the fusion of satellite sensor imagery using multi-resolution analysis ((Garguetduport *et. al*, 1996); (Ranchin and Wald, 1997); (Wald *et. al*, 1998)), image compression (Werness *et. al*, 1994), speckle noise reduction in SAR data (Horgan, 1998) etc. The basic difference between wavelet compression and the *Joint Photographic Experts Group* (JPEG) compression standard is in the transformation of the image from a spatial to a frequency domain. Whereas this is realised in JPEG through the use of the Discrete Cosine Transform (DCT), the same is achieved in wavelet compression through the Discrete Wavelet Transform (DWT). In general, wavelet compression results in a better image quality than the JPEG method, especially in applications where high compression rates are required (Shiewe, 1998). Moreover, the undesirable blocky artifacts that often result from the use of the 8x8 pixel blocks in the JPEG method are avoided. The wavelet method is also significantly faster than fractal image compression at similar compression rates.

The mathematical framework behind wavelets is beyond the scope of the work presented here. A more comprehensive discussion of this is given in ((Daubechies, 1992); (Chui, 1996); (Louis *et.al*, 1997)) etc. In general, wavelets are defined as mathematical functions that partition data into different frequency components and then facilitate the analysis of each component with a resolution matched to its scale. This takes into account the fact that the level of detail in an image may not be constant. In contrast to Fourier methods, wavelets allow the partition of the input signal not only with respect to its spectral, but also spatial properties. Therefore, whereas Fourier methods are ideal for „continuous tone” (homogeneous) data, wavelet methods on the other hand, are suited for input signal that is characterised by discontinuities and sharp spikes. Hence, remotely sensed data that depict a high concentration of heterogeneous objects (e.g., urban environments) are best compressed using wavelet compression schemes.

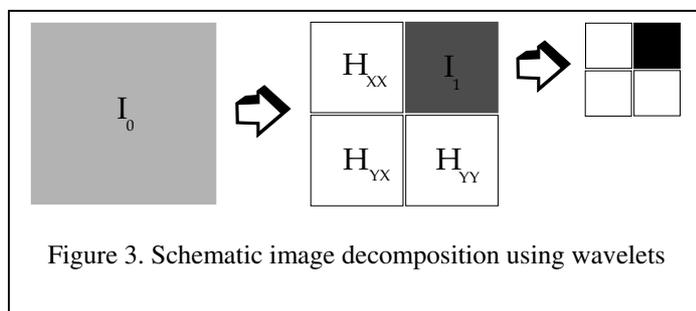


Figure 3. Schematic image decomposition using wavelets

In principle, the DWT method operates by splitting the image into four new image quadrants within the wavelet domain. Three of these new images contain high-frequency information in the horizontal, vertical and diagonal orientations respectively, while the fourth contains low-frequency information. The three high-frequency information quadrants are subsequently eliminated. The transformation is then repeated once again on the low-frequency image quadrant. Ideally, this is iterated until the last low-pass image contains only a

single pixel. Fig. 3 shows a schematic description of this procedure up to the second level of wavelet decomposition.

Through the use of an appropriate threshold function, the floating wavelet coefficients are approximated into corresponding integer values. Incidentally, it is this quantisation process that constitutes loss of information in lossy compression schemes. The number of bits required to represent each quantised coefficient are then reduced through binarisation.

4 SEGMENTATION RESULTS

For the study presented here the wavelet compression software *Lurawave* is used. This is a software package developed precisely for the compression of remotely sensed imagery (Luratech, 1999). The fused imagery is systematically compressed (channel by channel) at compression rates (K) ranging between 5 and 100. The compressed fused imagery is then classified using the approach described in section 2.3. Results obtained are then analysed and their geometric and semantic quality evaluated.

4.1 Geometric analysis

Sixty (60) identical image points are manually measured in the classification results for both the original and compressed imagery. The same object points (e.g., building corners) are identified and positional errors estimated. In general, relatively large error values are obtained. This is basically because of the concomitant pointing problems encountered, especially with classification results at higher compression rates. Fig. 4 clearly illustrates this limitation.

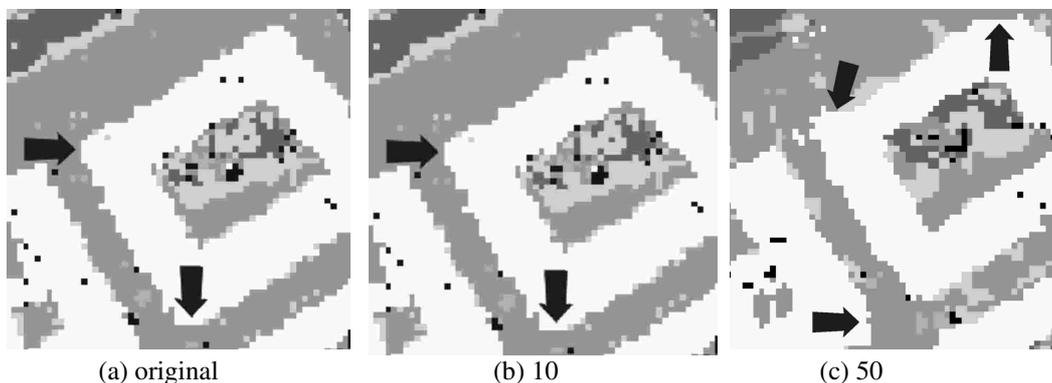


Figure 4: Pointing problems for various classification results

A Fisher test is then carried out in order to establish whether the positional errors obtained for the classification results at the different compression rates (s) are statistically different from similar results for the uncompressed data (s_0). Tests are performed at 5% level of significance. Table 2 shows the general test scheme applied.

Null hypothesis:	$H_0 : \sigma = \sigma_0$ (\Rightarrow no difference)
Alternative hypothesis:	$H_A : \sigma > \sigma_0$ (\Rightarrow difference)
Quantile (FISHER):	$F_{\nu_1, \nu_2, 0.05} = 1.53$ ($\nu_1 = \nu_2 = 59$)
Test value:	$Z_F = s^2 / s_0^2$
Test decision:	$H_0 : Z_F < F \quad / \quad H_A : Z_F \geq F$

Table 2: Fisher test scheme for evaluating geometric quality

For the original fused dataset, the positional accuracy is estimated at $s_0=0.44$ Pixel ($\pm 0.66m$). A comparison of this with similar estimates for different compression rates is illustrated in Fig. 5. From this it is apparent that the compression rate of 20 represents the critical compression rate. Beyond this value, the classification results for the compressed data are statistically different from those of the uncompressed data.

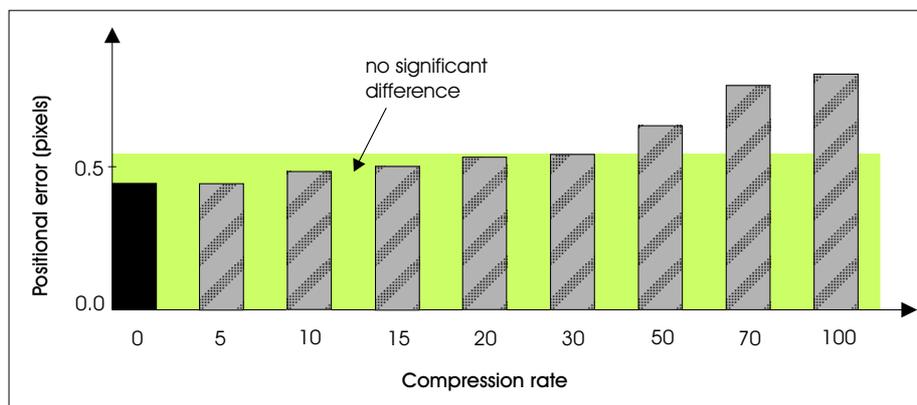


Figure. 5: Geometrical errors for different classification results

4.2 Semantic quality

The efficiency of data compression algorithms can be evaluated in several different ways. For example, it is possible to estimate the relative complexity of the algorithm, the memory required to implement the algorithm, how fast the algorithm performs on a given machine, the amount of compression achieved, or even how closely the reconstruction of the compressed data resembles the original data (Sayood, 1996). For this study, the assessment of the semantic quality of the classification results is basically done using the last of the evaluation options given above. Further, both visual and automatic methods are compared.

4.2.1 Visual interpretation

There is hardly any significant visual difference between the classification results from the original fused imagery and the compressed imagery for compression rates of up to about 30. However, for higher compression rates ($K \geq 50$) the smoothing effect characteristic of wavelet compression progressively increases and the quality of the classification results quickly deteriorates.

4.2.2 Automatic interpretation

Analysis of the grey value histograms results in the popular quantitative parameter, the *Peak-Signal-to-Noise-Ratio* (*PSNR*). This describes the relationship between the maximum grey value within the original image and the noise resulting from the compression. Considering an image of dimension ($n \times m$), the *PSNR*, expressed in decibels (*dB*), is given by:

$$PSNR(dB) = 10 \log \left(\frac{x_{peak}^2}{noise} \right);$$

$$\text{with } noise = \frac{1}{(n.m)} \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} (g_{ij} - g'_{ij})^2$$

where x_{peak} is the maximum grey value within the original image

g_{ij} - grey value for the (i, j)th element in the original image

g'_{ij} - grey value for the (i, j)th element in the compressed image

The relationship between the *PSNR* and the compression rate for the different classification results is shown in Fig. 6. The level of noise increases with increasing compression rate up to a compression rate of about 15. Consequently, the *PSNR* decreases accordingly. Between the compression rates of 15 and 20, no change in the *PSNR* is noted. At compression rates higher than 20, the smoothing effect of wavelet compression becomes apparent. As a direct result of this, the level of noise reduces and the *PSNR* increases conformably. However, after a compression rate of about 30 the level of noise increases once again. Subsequently, this results in lower *PSNR* values most probably due to the instability of the classification results.

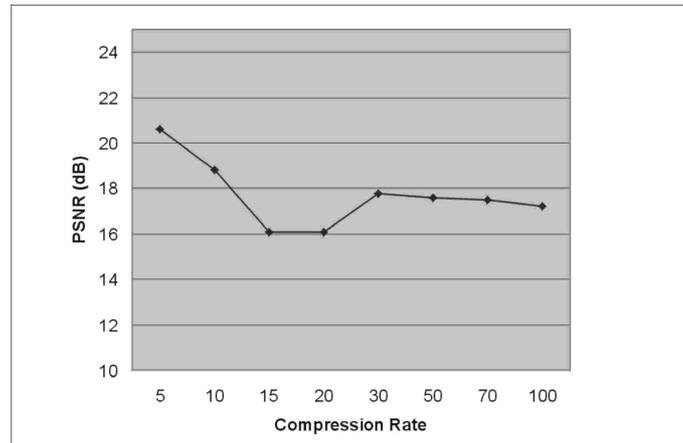


Figure 6: Relationship between *PSNR* and compression rate

5 SUMMARY AND CONCLUSIONS

This study underlines the importance of multi-sensor data fusion in the classification of urban environments. In particular, the need to integrate multispectral and geometric datasets is underscored. The need to incorporate context information in image segmentation is also highlighted. This is achieved by expanding the object feature base in order to exploit both spectral and spatial feature characteristics. Through this, higher classification accuracy and better semantic differentiation between the various object features can be achieved.

Data compression is conventionally employed in remote sensing within the context of data transmission and storage. However, the examination of the influence of this on the quality and further processing of satellite sensor imagery represents an important research and development topic. This is basically because of the existing dilemma between the huge amount of data often captured by remote sensing sensors and the technical restrictions in using this. Compression rates greater than 10 have been proposed for the next generation of commercial sensors in view of their higher spatial resolution and larger swath widths (Fritz, 1997). In this regard, lossy compression schemes provide the only viable solution. The superiority of wavelet-based methods over the standard JPEG technique in the compression of remotely sensed data has been demonstrated (Shiewe, 1998). Nonetheless, the smoothing effect of wavelet compression, especially at higher compression rates ($K \geq 50$) is undesirable.

The compression rate beyond which the smoothing effect becomes evident represents the critical compression rate that defines the range within which the compressed imagery can be used in further processing e.g., classification. It is clear from the results obtained in this study that a compression rate of up to 20 can be adopted for the classification of urban environments using Daedalus ATM imagery fused with airborne laser scanning data without adversely affecting the classification results. Further studies on this are still required to determine: (1) whether the cut-off value is influenced by the spatial resolution of the imagery, and (2) whether this depends on the scale at which the actual variation exists or on that which the user is interested in.

ACKNOWLEDGMENTS

The author wishes to acknowledge the support of Prof. Dr.-Ing H.-P. Bähr and the Deutsche Akademischer Austauschdienst (DAAD) in the preparation of this paper.

REFERENCES

- Bähr, H.-P., and Vögtle, T., 1998. *Digitale Bildverarbeitung: Anwendung in Photogrammetrie, Kartographie und Fernerkundung*. Wichmann Verlag, Heidelberg., pp. 220-228.
- Bähr, H.-P., and Vögtle, T., 1999. *GIS for Environmental Monitoring*. E. Schweizerbart'sche Verlagsbuchhandlung, Stuttgart.
- Chui, C. K., 1996. *Introduction to Wavelets*, Academic Press, San Diego.

- Daubechies, I., 1992. Ten Lectures on Wavelets, Vol. 61, CBMS-NSF Regional conference series in Applied Mathematics, Philadelphia, USA.
- Fritz, L.W., 1997. Status of new commercial Earth observation satellite systems. *Photogrammetrie. Fernerkundung. Geoinformation*, pp. 369-382.
- Fung, T. and Chan, K.C., 1994: Spatial composition of spectral classes: A structural approach for image analysis of heterogeneous land-use and land-cover types, *Photogrammetric Engineering and Remote Sensing*, Vol. 60(2), pp. 173-180.
- Garguetduport, B., Girel, J., Chassery, J. and Pautou, G., 1996. The use of multiresolution analysis and wavelet transform for merging Spot panchromatic and multispectral image data. *Photogrammetrie. Fernerkundung. Geoinformation*, Vol. 62, pp. 1057-1066.
- Gong, P. and Howarth, P. J., 1990: The use of structural information for improving land-cover classification accuracies at the rural-urban fringe, *Photogrammetric Engineering and Remote Sensing*, Vol. 56(1), pp. 67-73.
- Haala, N. and Brenner, C., 1999: Extraction of buildings and trees in urban environments, *Photogrammetric Engineering and Remote Sensing*, Vol. 54(2-3), pp. 130-137.
- Haala, N. and Walter, V., 1999: Automatic classification of urban environments for database revision using Lidar and colour aerial imagery, In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 32, Part 3/1, pp. 76-82.
- Haralick, R.M. and Shapiro, L.G., 1992. *Computer and Robot Vision*. Vol. 1. Addison-Wesley, Reading Massachusetts., pp.157-255.
- Horgan, G., 1998. Wavelets for SAR image smoothing. *Photogrammetrie. Fernerkundung. Geoinformation*, pp. 1171-1177.
- Kramer, H.J., 1996. Observation of the Earth and its Environment. Survey of Missions and Sensors. Third Ed., Springer-Verlag Berlin Heidelberg.
- Lillesand, T. and Kiefer, R., 1994: Remote sensing and image interpretation, John Wiley and Sons, New York.
- Louis, A. K., Maass, P. and Rieder, A. , 1997. Wavelets: Theory and Applications, John Wiley and Sons, Chichester
- Luratech., 1999. <http://luratech.com> (28th Apr. 1999)
- Okabe, A., Boots, B. and Sugihara, K., 1992. Spatial tessellations: concepts and applications of voronoi diagrams. John Wiley Press, Chichester, UK., pp.175-194.
- Pohl, C., 1999: Tools and methods for fusion of images of different spatial resolution, In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 32, Part 7-4-3, pp. 76-82.
- Ranchin, T. and Wald, L., 1997. Data fusion of remotely sensed images using the wavelet transform: The ASSIS Solution. In: *SPIE 1997 International Symposium on Optics, Imaging and Instrumentation., Mathematical Imaging: Wavelet Application in Signal and Image Processing V*, San Diego, California, USA, pp. 272-280.
- Savian, F. R., and Landgrebe, D., 1991. A survey of decision tree classifier methodology. *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 21, No. 3, pp. 660-674.
- Sayood, K., 1996. Introduction to data compression. Morgan Kaufmann Inc., USA., pp.217-230.
- Schiewe, J., 1998. Effect of lossy data compression techniques on geometry and information content of satellite imagery. In *International Archives of Photogrammetry and Remote Sensing*, Vol. 32, Part 4, Stuttgart, pp. 540-544.
- Schilling, K.-J., and Vögtle, T., 1996. Satellite image analysis using integrated knowledge processing. In *International Archives of Photogrammetry and Remote Sensing, XVIII Congress of the ISPRS*, Part B3, pp. 752-757.
- Serra, J., 1986. An Introduction to Mathematical Morphology. In *Computer Vision, Graphics and Image Processing*, Vol. 35, pp. 283-305.
- Wald, L., 1999. Definition and terms of reference in data fusion. In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 32, Part 7-4-3, W6, Valladolid, Spain, pp. 2-6.
- Wald, L., Ranchin, T. and Mangolini, M., 1998. Fusion of satellite images of different spatial resolutions: Assessing the Quality of resulting images, *Photogrammetric Engineering and Remote Sensing* Vol. 63(6), pp. 691-699.
- Walter, V., 1998. Automatic classification of remote sensing data for GIS database revision. In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 32, Part 4, pp. 641-648.
- Werness, S., Wie, S. and Carpinella, R., 1994. Experiments with wavelets for compression of SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 32, pp. 197-201.
- Wehr, A. and Lohr, U., 1999: Airborne Laser Scanning: An Introduction and Overview. *Photogrammetric Engineering and Remote Sensing*, Vol. 54, pp. 68-82.