

Potential application of remote sensing in monitoring informal settlements in South Africa where complimentary data does not exist

Karishma Busgeeth^{a*}, Frans van den Bergh^b, Jarrell Whisken^a, André Brits^a

^a Built Environment, Council for Scientific and Industrial Research (CSIR), P.O. Box 395, Pretoria 0001, South Africa;

^b Remote Sensing Research Unit, Meraka Institute, CSIR

ABSTRACT

Remotely sensed images are used for many purposes in today's world. In this paper, we explore the potential application of high resolution satellite images in extracting features and classifying urban settlements. The test area is Soweto, an urban area in the Greater Johannesburg Metropolitan area, in Gauteng, South Africa. We propose a new settlement typology for efficient classification of formal and informal settlements via QuickBird satellite images. Following on, an automated classification procedure based on the local binary pattern texture features is introduced. Using a convenience sample of 25 images, we show the feasibility of the new typology by applying it to both a manual classification procedure and an automated one. The manual classification procedure was conducted by a group of five experts who interpreted the images and classified them according to formal and informal settlements. Analysis of the results revealed an overall mean classification accuracy of 99.2% with a standard deviation of 1.79%. The automated method involved extracting tiles at random positions within the 25-sample dataset. The features extracted from these tiles were classified using a support vector machine. Classification accuracy on new samples was 56.27%, but cross-validation on the training data reached classification accuracies of 98%.

Keywords: informal settlements, planning, remote sensing, classification, QuickBird, built-up environment, South Africa

1. INTRODUCTION

Historically, the management of South African human settlements has been driven by the ideology of 'separate development' rather than by a concern to create a healthy, viable urban environment. This legacy has produced a complex set of spatial and physical problems created by the previous apartheid planning system resulting in inefficient city functioning in a context of high rates of unemployment, increasing rates of urbanisation, backlogs in the provision of basic services and a large range of social problems. Differential rights to the city were further polarized into separate forms of urban accommodation, a) those with formal leases to public housing, b) those residing legally in single-sex hostels, and c) those residing illegally, whether in back yard shacks of formal housing areas, in overcrowded hostels, or in informal settlements on invaded land. Informal settlements are known by a variety of names such as shacks, slums, squatter areas, shanty towns, and irregular settlements. Regardless of the name, the common phenomena that distinguish informal settlements from formal settlements are as follows: *do not adhere to local building codes, have either low levels of infrastructure or no infrastructure altogether, are either poorly serviced or not serviced at all, have no security of tenure and are characterized by a rather non-functional pattern.* According to Statistics South Africa (2004), informal settlements countrywide have increased from 1.049 million dwellings (in 1994) to 1.376 million (in 2004) and 'slum' housing is projected to continue increasing to some 2.4 million by 2008 (Hemson and O'Donovan ^[1], as cited by Richards *et al.* ^[2]).

The South African government recognises that the existence of informal settlements is a serious problem as they accommodate a large proportion of the urban population who live in sub-standard living conditions. In addition it is also realised that an increase in migrants to urban areas inevitably leads to a shortage of basic engineering services such as water, sewerage, solid waste removal and places essential services such as health and education under pressure. One of the fundamental difficulties that authorities face when planning a response to the formation and growth of informal settlements is the paucity of spatial and temporal data to assist in recognising and quantifying the understanding of

* kbusgeeth@csir.co.za; phone +27 12 841 2259; fax +27 12 841 4036; www.csir.co.za

settlement morphology, services/infrastructure, growth, population distribution and emerging settlement patterns. Reasons for the inability to obtain data include:

1. *Informal settlements are generally characterized by a dysfunctional settlement structure.* The distribution of plots follows no planned structure or conventional planning principles and streets and technical infrastructure are not catered for. Plot boundaries are unknown or non-existent, and plot sizes varies greatly;
2. *Settlements are highly condensed and difficult to access for surveys.* In certain situations, it is also considered to be too dangerous for official enumerators to do house-to-house data collection;
3. *Informal settlements are dynamic.* Population fluctuations and the erection and removal of structures over short time periods, mean that traditional survey methods cannot effectively capture temporal reality. Traditional surveys may take several months to process, rendering the results outdated at the time of release;

Other reasons are changes in municipal boundaries and overlapping administrative responsibilities. Deficits in manpower, lack of finance and technical equipment are additional challenges. As a result of these difficulties informal settlements are often not spatially documented. There are no maps indicating the position, patterns, size, complexity and influence of the settlements. Maps are regarded as unbiased and neutral sources of information about the world. Such is the power of the map that if authoritative, official maps and atlases fail to map places, the impression is given that those places do not exist ^[3]

The challenge therefore lies in having appropriate methods to identify and monitor the spatial behaviour of informal settlements reliably. One of the possible solutions is to use remote sensing imagery as the primary data source. Recent advances in computing power and the increasing availability of remote sensing imagery have revived renewed interest in remote sensing as potential data ware for monitoring informal settlement behaviour. This category of satellite commonly includes IKONOS (1999), EROS (2000), QuickBird (2001), SPOT – 5 (2002), ALOS (2006). Urban geographers have recognised the potential of this information for various applications including updating of maps, extraction of urban features such as road networks and other engineering and social infrastructure, generation of urban models, land cover mapping and a wide range of other possible applications ^[4, 5]. Despite its potential application Herold ^[6] noted that remote sensing imagery remained an under-utilised data source in urban studies. In addition, the majority of studies that have been undertaken concentrate on developed world examples. Research in the area of observing informal settlement behaviour in South Africa, using high resolution remotely-sensed data is limited as only the study conducted by Hofmann was available ^[7]. The study reported how informal settlements can be detected from other land-use-forms by describing typical characteristics of colour, texture, shape and context using remote sensed data from IKONOS in Cape Town. In cases where the settlements were not appropriately classified, visual inspection was carried out or a final correction by hand was performed with eCognition. He concluded that while high resolution IKONOS data is well suited to detect informal settlements areas using pure IKONOS image data was not sufficiently feasible to detect a single shack.

The objective of this paper is to explore the potential application of high resolution remotely-sensed data such as QuickBird for monitoring informal settlements. The paper uses the greater Soweto area in South Africa as a case study area to explore what kind of morphological human settlement attributes can be observed from QuickBird. These attributes are used to (a) assess their applicability to support the existing urban settlement typology, applied by the South African National Land Cover Classification Legend; (b) propose an urban settlement typology for informal settlements based on morphological attributes with the aim of incorporating this in (c) an automated data extraction and classification procedure which could then be combined with socio-economic data and statistical methods to render planning support and monitor informal settlements. The paper concludes with lessons learned and remaining challenges.

2. CASE STUDY AREA: SOWETO, SOUTH AFRICA

2.1 Introduction

Soweto is an urban area in the Greater Johannesburg Metropolitan area, in Gauteng, South Africa. Recent demographics are not readily available at the time of writing this paper but South Africa's 2001 census put its population at approximately 900, 000 which translates to about one-third of the city's total population.

To analyse the complexity of the informal settlements, satellite images of QuickBird dated from 2005 with 0.60m spatial resolution respectively were sourced and polygons were created to delineate homogeneous areas (See Fig. 1 below).



Fig. 1. QuickBird images of different types of settlements in Soweto: Informal housing (left), Formal suburb with backyard shacks (centre), and Formal suburb (right)

2.2 Observing Urban Settlement Attributes

A list of physical settlement attributes were compiled and used as a basis to assess the extent to which these attributes can be observed from QuickBird. Table 1 specifies the list of urban settlement attributes with the corresponding results.

Table 1. Observed human settlement attributes, using QuickBird.

SPATIAL ATTRIBUTES	CHARACTERISTICS	DESCRIPTION	MANUAL EXTRACTION FROM QUICKBIRD (Y/N)
Structure Size	Uniform size vs. different size	Building size	Y
Settlement Layout	Planned vs. irregular layout of settlements	Road layout; Erf/lot demarcation	Y
	Open space	Amount of open space available; Presence of trees	Y
Housing structure	Building materials	Corrugated iron sheets, wood, plastic, cardboard, bricks etc	Y
	Colour of roof	Homogeneous vs heterogeneous	Y
	Density	Low vs. high/ Adequate space between houses	Y
Engineering services	Roads	Tarred vs. gravel / paths	Y
	Telecommunication	Telkom/Vodacom/MTN/Cell C poles	N
	Electricity	Electric poles	N
		Substation	Y
	Water & sanitation	Toilet facilities & reservoir	Y
	Storm water drainage	Presence of manholes	N
Waste management	Presence of dumping grounds/landfills	Y	
Infrastructure	Education	Presence of schools, colleges	Y
	Business opportunities	Presence of shops, food outlets	Y
	Social facilities	Presence of sports ground/stadium, clinics, community centres, police stations	Y
	Transport facilities	Presence of bus/taxi ranks, stations/railways	Y

The results of the manual extraction of the attributes show that size, layout, housing structure and infrastructure could be identified from the QuickBird images whilst establishing whether the settlements had access to different types of engineering services were not conclusive.

2.3 Existing Urban Settlement Typologies

Using the above attributes in Table 1, the applicability of using the existing urban settlement typology, applied by the South African National Land Cover Classification Legend (See Fig. 2), was assessed. The typology for the urban areas of the National Land Cover Classification is very limited and this hinders its application to effectively classify urban settlements.

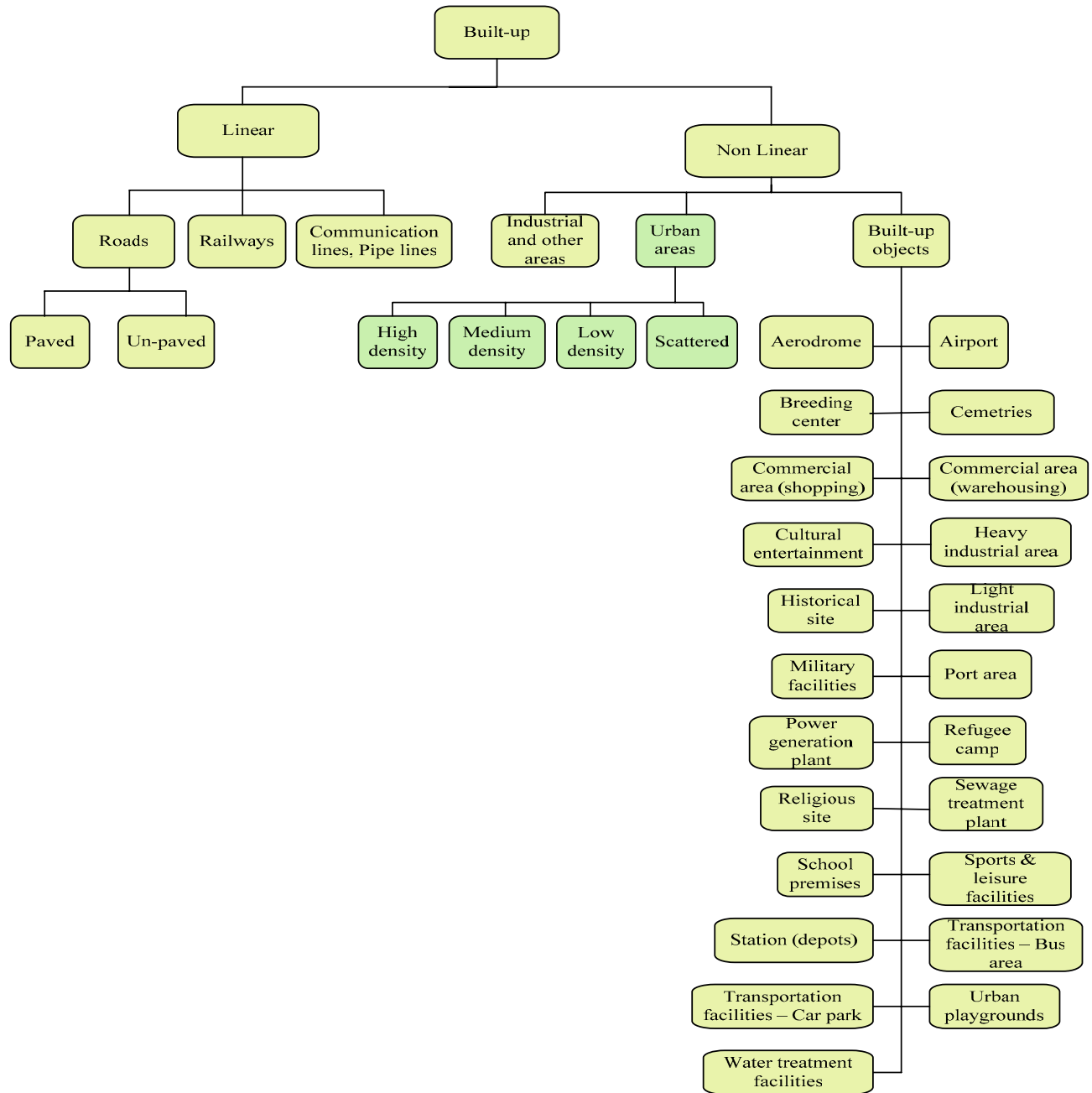


Fig. 2. The South African National Land Cover Classification Legend

3. PROPOSED SETTLEMENT TYPOLOGY

The need to propose a refined settlement typology to harmonize the diversity of urban settlements therefore arises. The purpose of the proposed settlement typology is to allow the classification of settlements primarily using satellite images. The proposed settlement typology integrates three elements which are of particular importance. The *first* refers to the physical features describing the building size, the layout of the roads and settlements; the *second* are the intrinsic features which describe the building materials, the colour of roof and the building density; the *third* describes the contextual features relating to the geographical relationship between the different attributes i.e. presence of engineering services and presence of infrastructure. The proposed settlement typology follows a hierarchical approach which means that each class has one or more sub-classes which enables inheritance of class features thus making the typology more versatile. Fig. 3 shows the proposed settlement typology¹.

The typology incorporates four levels of classification: the first level of classification differentiates *built-up environment* from *non built-up environment*; the second level of classification distinguishes *informal settlements* from *formal settlements* whilst the third level of classification refers to sub-classes. *Informal squatter* and *informal township* are sub-classes of *informal settlements*, whilst *RDP subsidised housing*, *high-rise buildings/flats*, *mining hostels*, *traditional settlement*, *formal township/ suburb plus backyard shack* and *formal township/suburb* are all sub-classes of *formal settlements*. The typology also consists of a collective set of rules called *rules of segregation* which allow the interpretation of the remotely sensed images to easily classify whether a particular settlement falls into the informal or formal settlement category. These rules have been constructed based on the knowledge about characteristics of the selected class. The complex nature and pattern of informal settlements in South Africa reveals that the typology should be flexible enough to accommodate the different types of informal dwellings, hence the fourth level of classification. However, defining the third rules of segregation involved more than the three elements (physical, intrinsic and contextual) mentioned above. A fourth element, socio-economic variables, was factored in. This includes household income, age of household, location and land ownership.

4. AUTOMATED CLASSIFICATION PROCEDURE

4.1 Overview of texture feature extraction and classification of urban settlements

Automated classification of urban settlements can be implemented using a two-stage process: The first stage extracts relevant features (e.g., texture attributes) from the satellite imagery, and the second stage employs a classifier (e.g., a neural network) to label the extracted features according to the classes defined in the settlement typology. Urban mapping performed at moderate to high resolution (>10m pixel size) often relies on spectral features, such as sub-pixel spectral unmixing^[8].

Very high resolution satellite data, like that obtained from the QuickBird satellite, lends itself well to the extraction of spatial features in urban environments. An intuitive approach would be to perform building delineation, followed by a method that describes the relative positions and sizes of the buildings. This approach may be appropriate for larger buildings, but Hofmann found that individual shacks could not be identified on IKONOS imagery^[7]. QuickBird has a higher spatial resolution than IKONOS, at 0.6m vs 1m in the panchromatic band; this represents a significant increase in spatial resolution, but the pixel size is still too small to reliably detect the space between shacks in informal settlements, making the identification of individual shacks infeasible.

One solution to this problem is to consider spatial features that do not rely on explicit object delineation. Such image features are collectively referred to as “texture features” in the machine vision community. The best known set of texture features are those proposed by Haralick^[9], which are derived from a gray-level co-occurrence matrix (GLCM) built over a region of interest in the image. Pesaresi investigated the use of these GLCM texture features to perform automated classification of 16 urban land cover types, with types such as “built-up, regular pattern, high density”^[10]. Using these classes, the contrast feature derived from the GLCM achieved a classification accuracy of around 98%. Shackelford presented an integrated textural and spectral approach to urban land cover classification, again making use of GLCM-derived features^[11].

¹ The Proposed Settlement Typology is an ongoing piece of work and is continuously being refined.

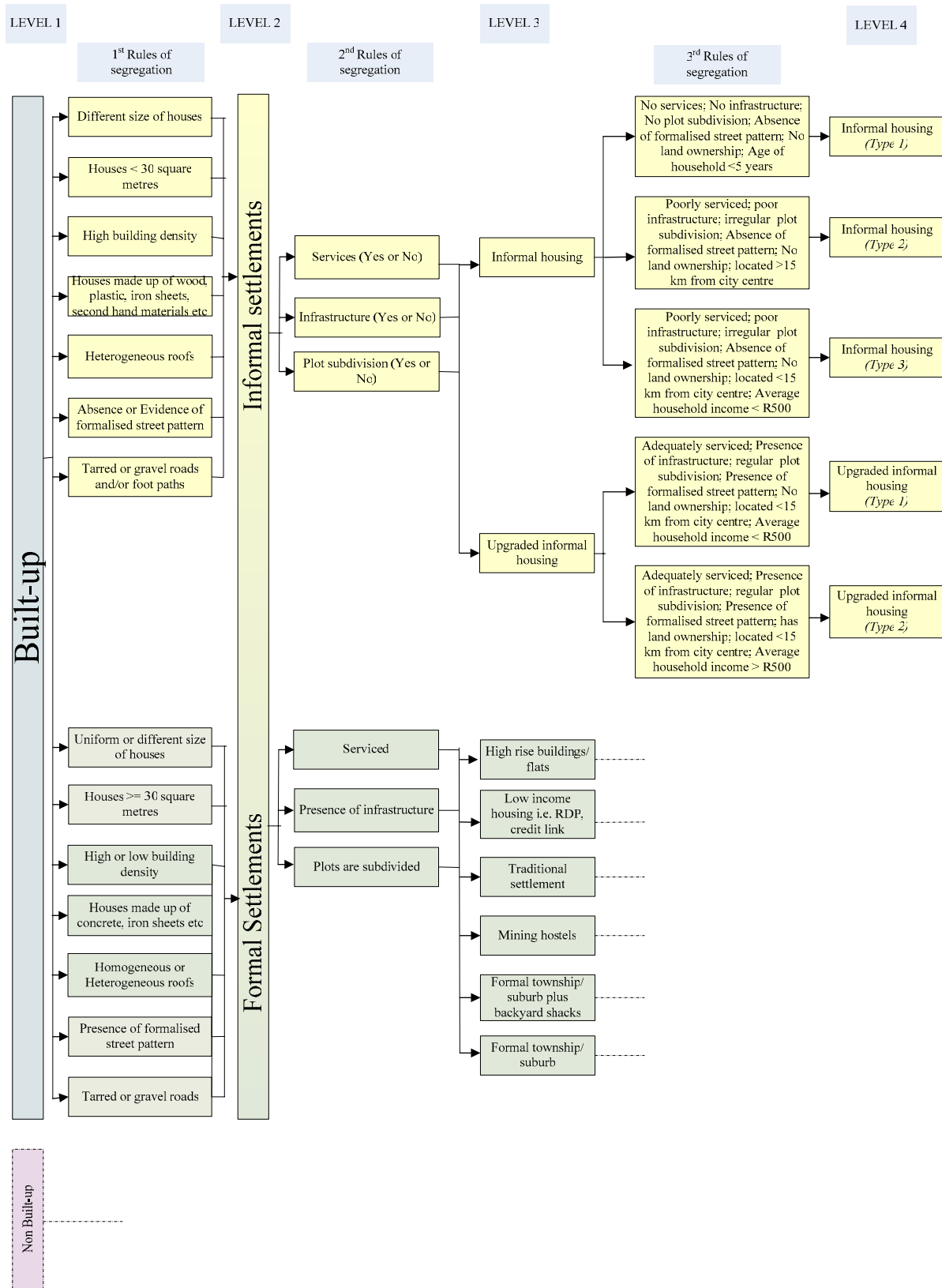


Fig. 3. The Proposed Settlement Typology

Although the GLCM-derived features appear to work well in remote sensing applications, it has been surpassed by the Local Binary Pattern (LBP) method in a variety of machine-vision texture classification problems [12]. The application of the LBP texture features to remote sensing was investigated by Lucieir in a multi-spectral automated segmentation application [13], and found to yield an accuracy of approximately 77%.

The procedure proposed here will employ LBP texture features in an automated settlement classification system.

4.2 The local binary pattern method

The local binary pattern features are extracted from individual image tiles, as illustrated in Fig. 4.

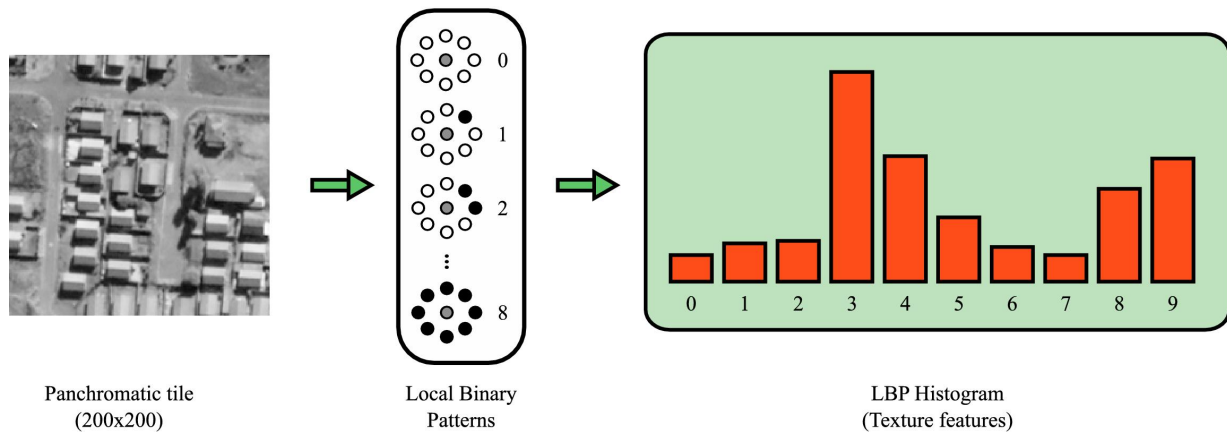


Fig. 4. An illustration of the Local Binary Pattern feature extraction process. Binary patterns, such as those illustrated in the centre of the diagram, are detected at each pixel position in the image tile. A histogram is constructed by accumulating the codes associated with the detected patterns.

These features are extracted by calculating a code value at each pixel position in a tile, accumulating them over the entire tile to construct a histogram, which is used as the final texture feature. The codes are computed as follows:

1. Define the code value as a string of n bits. Set them to “0”.
2. Define a list of n sampling positions, arranged in a circle of radius r around the central pixel p .
3. Compare the pixel intensity at each sampling position to the intensity of the central pixel p . If the value at the sampling position is greater than the value at p , then append a “1”-bit to the code; otherwise append a “0”-bit.
4. Normalise the bit-string code by applying the bit-wise rotate operation to find the orientation with the smallest binary value. This is effectively rotates the pattern around the central pixel until it reaches the reference orientation \square an illustration of the reference orientation can be seen in Fig. 4.
5. The codes now fall into two classes: those with a single “run” of “0”-bits followed by a “run” of “1”-bits, the so-called *uniform* patterns, and those with more than a single “run” of “1”-bits. Map all these non-uniform codes to the value $n+1$.

This method yields the so-called *ru12* uniform local binary patterns [12], which are designed to be invariant to rotation of the image, as well as invariant to contrast changes and offsets in the intensity of the pixels.

A multi-resolution texture feature vector is constructed by concatenating the histograms obtained by applying the LBP method with different combinations of the parameters r (circle radius) and n (number of samples).

4.3 Urban settlement classification using local binary patterns

Lucieir *et al.* demonstrated the application of LBP texture features to automated segmentation [13], however, these features can also be used to perform automated classification. The experiment presented here is designed to measure the ability of the LBP texture features, derived from panchromatic QuickBird imagery, to correctly separate the various settlement types. A subset of the classes defined in Fig. 3 was selected for this experiment: Informal Housing (IH),

Upgraded Informal Housing (UIH), Formal Suburb (FS), Formal Suburb with Backyard shacks (FSB), and Non-Urban (NU). Examples of the IH, FSB and FS classes can be seen in Fig. 1.

Texture features can only be computed over a region, since they often involve statistics that can not be collected at only a single point. The simplest way of defining these regions is to specify a grid of overlapping tiles that cover the image. The texture features are then calculated separately for each tile. A tile size of 200x200 QuickBird pixels, corresponding to a 120x120m region on the ground, was selected based on the findings of Pesaresi^[10].

To measure the classification accuracy of the LBP texture features a set of tiles were extracted from polygons labelled by class type. The tiles were extracted from random positions within these polygons, provided that the entire tile fitted inside the polygon, so that each tile contained only a single settlement type.

Three resolution levels of LBP texture features were extracted; at each resolution level 8 sampling points (i.e., $n = 8$) were selected, yielding a histogram with 10 bins. The corresponding radii of the three resolution levels were 1.0, 3.0, and 8.0 pixels, respectively. The final LBP texture feature vector of dimension 30 was formed by concatenating the histograms obtained at the three resolution levels. A support vector machine^[14] was trained using the libSVM library².

5. EVALUATION OF THE PROPOSED SETTLEMENT TYPOLOGY

5.1 Manual data extraction and classification procedure

To prevent bias, a group of five experts were approached to apply the proposed settlement typology of Fig. 3 to a representative sample of 25 satellite images of the study area. The experts visually interpreted the images and used the first rules of segregation as classification criteria for differentiating between formal and informal settlements. Analysis of the results revealed an overall mean classification accuracy of 99.2% with a standard deviation of 1.79%.

Based on this exercise, it is noted that classifying the settlements down to level 4 is challenging. This is because visual interpretation alone cannot identify features like household income, age of household, location and land ownership by merely looking at the satellite imagery. In the case of informal settlements clearly assessing whether they have access to engineering services is not so straightforward from the images. Further, in the case of formal settlements, the morphology of flats and mining hostels is very similar. Hence, one should have either prior knowledge of the area or be aware of the context for effective differentiation and classification.

5.2 Automated data extraction and classification procedure

Section 4.3 described the method used to generate the training and test data used to evaluate the performance of the proposed automated classification. Cross-validation (10-fold) training was performed to obtain an estimate of the expected generalisation classification accuracy. An overall mean classification accuracy of 98.37% with a standard deviation of 0.156% was achieved over 10 repetitions of the cross-validation. Table 2 provides the confusion matrix obtained on the training data.

Table 2. Classification confusion matrix obtained with 10-fold cross-validation on tiles extracted from the training set. The associated Kappa value is 0.98.

Algorithm classification →	FSB	FS	IH	UIH	NU
Ground truth classification ↓					
FSB	421	9	0	0	2
FS	13	492	0	0	3
IH	0	0	201	0	0
UIH	0	0	0	227	0
NU	6	1	0	4	509

² <http://www.csie.ntu.edu.tw/~cjlin/libsvm>

In a second test, tiles were extracted at random positions within the 25-sample test dataset used in Section 5.1. This dataset was generated from a different region of Soweto. It is therefore possible that the shift in location could have resulted in subtle changes in the observed patterns in some of the classes. The features extracted from these tiles were classified using the support vector machine trained on the training data set. Table 3 presents the corresponding confusion matrix. Overall classification accuracy was only 56.27%, which is much lower than what was obtained on the training set with cross-validation. If a cross-validation experiment is performed on only the 25-sample test dataset, the overall classification accuracy increases to 99.55%. This large difference in classification performance can possibly be attributed to actual differences between the training and the test datasets. It is possible that the training set did not include representative samples from all the possible patterns found in all the classes. As a last check, the training and test sets were combined to form a dataset with 2950 instances. This combined data set was used to train the support vector machine using cross-validation, yielding an overall classification accuracy of 97.29%. This appears to support the conclusion that the local binary pattern method is able to accurately describe the patterns observed in the various settlement classes, but that the quality and completeness of the training set is critical to obtaining good results on regions not seen during training.

Table 3. Classification confusion matrix obtained by applying the trained support vector machine to tiles extracted from the 25-sample test dataset. The associated Kappa value is 0.43.

Algorithm classification →	FSB	FS	IH	UIH	NU
Ground truth classification ↓					
FSB	184	18	0	0	2
FS	495	223	0	41	1
IH	0	0	50	0	0
UIH	0	0	19	30	0
NU	1	1	0	2	257

6. CONCLUSION

This study focused on the potential application of QuickBird images for monitoring informal settlements in South Africa. Using the greater Soweto area, South Africa, as a case area, the authors manually extracted a list of human settlement attributes that can be observed from QuickBird images. Based on the findings of this study, the authors noted that applying the urban typology of the South African National Land Cover Classification Legend was too limited to effectively classify urban settlements. The research was motivated by the need to develop an urban settlement typology to facilitate the effective classification of informal settlements in South Africa. The proposed settlement typology integrates four elements of settlements which are of particular importance, namely: the physical features, the intrinsic features, the contextual features and the socio-economic features. It provides a tool for the visual, systematic and representative analysis of South African settlements from remote sensing images.

An automated procedure for settlement classification was proposed: Local binary pattern texture features are computed from tiles extracted from panchromatic QuickBird imagery, which are classified using a support vector machine. The proposed settlement typology was evaluated both manually and automatically using a convenience sample of 25 images. The manual classification procedure yielded an overall mean classification accuracy of 99.2% with a standard deviation of 1.79%. Evaluating the automated procedure on a separate training set using 10-fold cross-validation yielded an overall mean classification accuracy of 98.37% with a standard deviation of 0.156%. The same automated system yielded a classification accuracy of only 56.7% when applied to the test set of 25 sample images used in the evaluation of the manual classification procedure.

6.1 Lessons learned

Whether manual or automated, extraction and classification can be a very involved process. Humans are good at scanning large areas and recognising objects, and can easily identify the settlement classes once they have seen a few examples. In contrast, the local binary pattern texture features used in the automated classification had sufficient

descriptive power to separate the classes during the training phase, but they did not generalise well to new samples. This implies that a very complete training set would be required to ensure that the automated system performs well when applied to new areas. Examples of every subtle variation in the pattern of each settlement type must be present in the training set to produce human-competitive results.

6.2 Remaining Challenges

The samples in the test set that were misclassified by the automated system must be manually compared to the existing samples in the training set to verify that they are actual variations of the patterns associated with the relevant settlement class.

Additional experiments must be performed with the automated system using different training and test sets to measure the typical degradation in classification accuracy caused by the natural variations in the structure of the settlements.

ACKNOWLEDGEMENTS

The authors would like to thank the five experts namely Dr. S. Biermann, T. Kruger, Dr. K. Landman, M. Mmonwa and Dr. L. Waldeck for their valuable time, efforts and comments.

REFERENCES

- ¹ Hemson, D. and O'Donovan, M., "Putting numbers to the scorecard: Presidential targets and the state of delivery", in *State of the Nation: South Africa*, 11–45 (2005).
- ² Richards, R., O'Leary, B., and Mutsonziwa, K., "Measuring quality of life in informal settlements in South Africa," *Social Indicators Research* **81**(2), 375–388 (2007).
- ³ Stickler, P. J., "Invisible towns: A case study in the cartography of South Africa," *GeoJournal* **22**(3), 329–333 (1990).
- ⁴ Volpe, F. and Rossi, L., "QuickBird high resolution satellite data for urban applications," in 2nd *GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*, 1–3 (2003).
- ⁵ Paulsson, B., *Urban Applications of Satellite Remote Sensing and GIS Analysis*, World Bank, Washington, DC (1992).
- ⁶ Herold, M., Hemphill, J., Liu, X., and Clarke, K. C., "Urban patterns and processes: a remote sensing perspective," in *1st EARSeL Workshop of the SIG Urban Remote Sensing*, (Mar. 2006).
- ⁷ Hofmann, P., "Detecting informal settlements from IKONOS image data using methods of object oriented image analysis — an example from Cape Town (South Africa)," in *Remote Sensing of Urban Areas*, Jürgens, C., ed., Regensburg: University Regensburg (2001).
- ⁸ Powell, R. L., Roberts, D. A., Dennison, P. E., and Hess, L. L., "Sub-pixel mapping of urban land cover using multiple endmember spectral mixture analysis: Manaus, Brazil," *Remote Sensing of Environment* **106**, 253–267 (Jan. 2007).
- ⁹ Haralick, R. M., Dinstein, I., and Shanmugam, K., "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics* **3**, 610–621 (1973).
- ¹⁰ Pesaresi, M., "The remotely sensed city," tech. rep., Space Applications Institute, Joint Research Centre (2000).
- ¹¹ Shackelford, A. and Davis, C., "A hierarchical fuzzy classification approach for high-resolution multispectral data over urban areas," *Geoscience and Remote Sensing, IEEE Transactions on* **41**(9), 1920–1932 (2003).
- ¹² Ojala, T., Pietikäinen, M., and Mäenpää, T., "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(7), 971–987 (2002).
- ¹³ Lucieer, A., Stein, A., and Fisher, P., "Multivariate texture-based segmentation of remotely sensed imagery for extraction of objects and their uncertainty," *International Journal of Remote Sensing* **26**(14), 2917–2936 (2005).
- ¹⁴ Cortes, C. and Vapnik, V., "Support-vector networks," *Machine Learning* **20**(3), 273–297 (1995).