

Textural pattern recognition:
*distinguishing between occupation densities in the centre of Dar
Es Salaam, Tanzania.*

DIP.5 final project

November 1998

E. Mac Gillavry

*Student of Human Geography - Cartography
Department of Cartography
Faculty of Geographical Sciences
University of Utrecht*

Supervising staff: ir. B. Gorte PhD

*Image Processing Laboratory
Department of Geoinformatics and Spatial Data Acquisition
Institute for Aerospace Survey and Earth Sciences (ITC)*



1	FINAL PROJECT: INTRODUCTION	4
2	SPOT- IMAGERY	5
3	IMAGE FUSION	6
4	SEGMENTATION OF REMOTELY SENSED IMAGERY	7
5	TEXTURE IN REMOTELY SENSED IMAGERY	7
6	METHODOLOGY	9
6.1	Site selection	9
6.2	Image fusion	9
6.2.1	Fusion in ILWIS 2.1.	9
6.2.2	Fusion in ERDAS Imagine 8.3.1.	10
6.3	Ground truth	11
6.3.1	Land use map modification	11
6.3.2	Sample site selection	12
6.4	Classification	12
6.5	Segmentation	13
6.6	Texture analysis	13
6.7	Integration of texture image into the classification	15
7	DISCUSSION	16
8	CONCLUSION AND RECOMMENDATION	16
9	REFERENCES	17
10	APPENDICES	18



Figure 1. Overview of Tanzania	4
Figure 2. SPOT PAN image	6
Figure 3. Intensity image	6
Figure 4. Adjacency	7
Figure 5. Dar Es Salaam, SPOT XS with study area	9
Figure 6. Fused SPOT-image (XS+P)	10
Figure 7. Stretched image	10
Figure 8. Land use map of study area	11
Figure 9. 3x3 -windows and their respective arrays	12
Figure 10. Classified image using ML	12
Figure 11. Flowchart image segmentation	13
Figure 12. Accuracy per class using different segmentation thresholds	13
Figure 13. Artificial texture image	14
Figure 14. Classified image using texture	15
Table 1. Land use classes	11
Table 2. Classification results	12
Equation 1 Computing the intensity image	19
Equation 2 Substituting the intensity image with the SPOT PAN	19



1 Final project: introduction

As part of the course on Digital Image Processing of Remotely Sensed Data (DIP.5) at the International Institute of Aerospace Survey and Earth Sciences (ITC) in Enschede, the Netherlands, the participant has to do a final project in which the principles and techniques learnt during the course are applied. The final project takes place in the last 7 weeks of the 16-week course.

This project covers part of the project carried out by the “Beleidscommissie Remote Sensing” (BCRS): “NRSP 1990-2000, Urbanization in Dar Es Salaam.” One of the topics of this project is to distinguish between occupation densities using digital image analysis of a data set consisting of SPOT XS and SPOT PAN imagery.

Dar Es Salaam has undergone rapid growth during the last decades. Not all, new settlements were planned by the urban planning authorities, but came into existence rather spontaneously. However, now that these settlements are there, the municipal administration wants to have better insight into the layout of these unplanned settlements. One of the instruments at their use is remotely sensed imagery.



Figure 1. Overview of Tanzania

A conventional supervised classification method Maximum Likelihood (ML) can be performed, but spectral classes are not easily discriminated within an urban setting. Urban subclasses have similar spectral reflectance properties: a mixture of bare soil and buildings. Results are therefore often very poor.

In order to be better able to distinguish between the urban subclasses texture analysis is introduced into the classification process. Special attention is given to the discrimination between planned and unplanned settlements. From visual inspection, one can quite easily notice the textural difference between the urban subclasses. Large plots of lands arranged in an orderly fashion characterize planned settlements. This results in a low texture. The grey level values do not change very rapidly from one pixel to the other. Unplanned settlements on the other hand are characterized by a high building density and the absence of any structure in the layout of the plots. This subclass has a high texture, for grey levels in the image change very fast from one pixel to the other. Therefore, these spatial patterns are supposed to be characteristic for a type of settlement: planned or unplanned. Texture information should then be included into the classification process to improve the accuracy. This works very well for high-resolution imagery, which show higher spectral variability that affects negatively classification results. **During this final project, texture-algorithms are applied on areas of the SPOT PAN data set to extract textural patterns for discrimination different occupation densities with special regard to the distinction between planned and unplanned settlements.**

To extract spatial objects from this image, an image-segmentation is applied on the data. However, as the segmentation algorithm is only able to cope with three input bands and there are more input bands available, image fusion is performed.

Classification is performed on SPOT XS fused with SPOT PAN and SPOT XS with a SPOT PAN-derived texture image. The resulting accuracy matrices are subsequently evaluated to see which one will produce the best results. Furthermore, not only the fusion input-features are a changing parameter in this study, but also the texture analysis algorithm parameters being used.



2 SPOT- imagery

The SPOT payload comprises two identical HRV (High-Resolution Visible) imaging instruments, two tape recorders for image data, and a payload telemetry package for image transmission to ground receiving stations.

The position of each HRV entrance mirror can be commanded by ground control to observe a region of interest not necessarily vertically beneath the satellite. Thus, each HRV offers an oblique viewing capability, the viewing angle being adjustable through $\pm 27^\circ$ relative to the vertical. Two imaging modes are employed, panchromatic (P) and multi-spectral (XS). Both HRVs can operate in either mode, either simultaneously or individually.

The SPOT payload comprises two identical HRV (High-Resolution Visible) imaging instruments, two tape recorders for image data, and a payload telemetry package for image transmission to ground receiving stations.

The position of each HRV entrance mirror can be commanded by ground control to observe a region of interest not necessarily vertically beneath the satellite. Thus, each HRV offers an oblique viewing capability, the viewing angle being adjustable through $\pm 27^\circ$ relative to the vertical. Two imaging modes are employed, panchromatic (P) and multi-spectral (XS). Both HRVs can operate in either mode, either simultaneously or individually.

In the "P" (panchromatic) mode, imaging is performed in a single spectral band, corresponding to the visible part of the spectrum without the blue. The band covers the electromagnetic spectrum from 0.51 to 0.73 μm . This single channel imaging mode supplies only black and white images with a pixel size of 10 meter. Therefore, this band is primarily intended for applications calling for fine geometrical detail.

In the "XS" (multi-spectral) mode, imaging is performed in three spectral bands. The bands used are band XS1 covering 0.50 to 0.59 μm (green), band XS2 covering 0.61 to 0.68 μm (red) and band XS3 covering 0.79 to 0.89 μm (near infrared). By combining the data recorded in these channels, colour composite images can be produced with a pixel size of 20 meters.



3 Image fusion

Image fusion is the process of combining digital images, by modifying the data values, using a certain procedure. Before starting the fusion process, the data should be properly co-registered. Combining data through image fusion techniques results in a new data set. A commonly used technique is the red-green-blue transform into intensity-hue-saturation. The intensity component is replaced with another image and this data set is transformed back into the three primary colours. However, the data set replacing the intensity component should have similar characteristics. This particular transformation is also used by Talukdar. In this study, the intensity component of the data set (SPOT XS) is also replaced with another data set (SPOT PAN).

A very common reason to perform image fusion used to be the drawback of large data sets with regard to computer capability. Image fusion often was a means of data reduction. However in this particular case, the bottleneck is not the computer capability, but the fact that the image segmentation algorithm (see **chapter 4**) is only able to cope with 3 input features. Another reason for the image fusion is the increasing spatial resolution. Finally, the introduction of the SPOT PAN data set into the classification leads to a better and more fair comparability of a straightforward common classification based on spectral characteristics and a classification involving textural information (see **chapter 5**).

Instead of the RGB-IHS transformation as described in the first paragraph, in this study the image fusion is performed using the HSV-transformation (Hue, Saturation, Value). The intensity in this case is not computed using complex transformation matrices, but is very clearly defined as being the total reflected energy per pixel (see **Appendix A, Equation 1**).

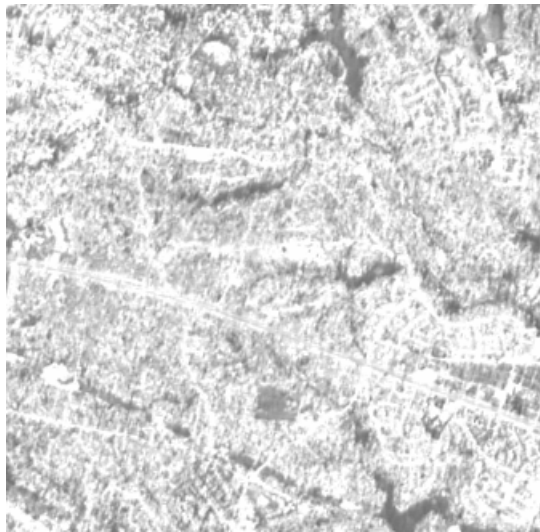


Figure 2. SPOT PAN image

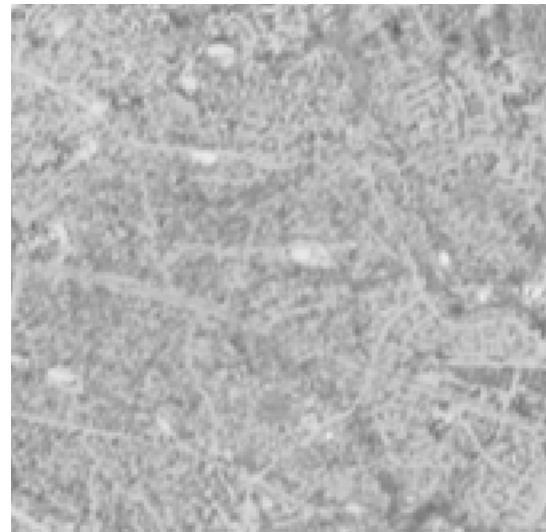


Figure 3. Intensity image

As the SPOT PAN-sensor detects all reflected energy in the broad band covering almost the entire width of the visual range, it can be stated that this data set has similar spectral characteristics. Therefore, it is justifiable to fuse this data set with the SPOT XS data set. This replacement is performed using the calculation as described in **Appendix A, Equation 2**.

Now the three SPOT XS one SPOT PAN bands are fused into three features which can be introduced into the Linux-based C-program.



4 Segmentation of remotely sensed imagery

The purpose of multi-spectral image segmentation is to subdivide an image into regions that are homogeneous according to certain criteria, such that these regions correspond to objects in the terrain. Each region has a unique identifier along with a list that links each identifier to the spectral properties of the corresponding region. Especially with the increasing resolution of remotely sensed imagery where an object is represented by a significant number of pixels, the need to combine these pixels into relevant objects is becoming more and more important.

The segmentation algorithm makes a bottom up quadtree reversal. It starts at single pixels (or larger quadtree leaves in which the pixel values are constant) and recursively merges adjacent quadtree leaves, forming irregular segments at all stages. The adjacency is defined in a 4-neighbourhood, as opposed to a 8-neighbourhood (see **Figure 4.**)

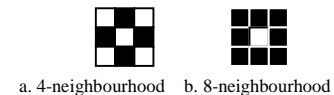


Figure 4. Adjacency

However, due to the recursive Z-scan order, the algorithm has a slight tendency to create segments of regular shapes. This can be seen when a very high threshold is selected. The resulting objects are very regular. Most of them are even just squares!

To bypass this (order dependency) problem several iterations have to be performed. For each iteration the homogeneity criterion is relaxed until a user-defined degree of segmentation is reached. Irregular shapes are formed in the first iterations. These irregular shapes are subsequently used as input for the next iterations. The homogeneity criterion is a user-selected spectral distance threshold θ . For two adjacent quadtree leaves to merge two requirements have to be met. First, the Euclidian distance between the feature vectors should be smaller than 2θ . Second, every entry of the variance-covariance matrix of the combined quadtree leaves should not exceed θ^2 . So, to relax the merging criterion, the value of the threshold θ has to be increased. Unfortunately the user has to apply the “trial and error-method” to find suitable threshold values and to visually inspect the segmentation results.

5 Texture in remotely sensed imagery

High-resolution imagery shows high spectral variability. This reduces statistical separability of land cover classes in spectral data space and tends to diminish classification accuracies. For example in a Landsat TM-scene, an urban agglomeration consists of mixels, containing various objects. These objects form in their spatial constellation the land cover class “urban area”. However, with increasing spatial resolution of the imagery, every pixel is more likely to record just a subset of the constellation of objects characterizing the class “urban area”.

Almost 90% of all computer algorithms for image analysis developed heretofore deal exclusively with spectral variations, as in the spectral signature concept. In most cases, this concept has failed to meet the classification accuracy achieved by experienced photo-interpreters. This indicates, that spectral evidence alone is not enough to adequately classify remotely sensed data (Austrom, 1990, p.3). Inclusion of spatial information (texture) into the classification algorithm should improve the results, for now the constellation of the separate pixels each recording just a subset of all the objects forming the class “urban area” is taken into account.

Texture is based on spatial variations of tone. Tone refers to the brightness or darkness of an area. Texture refers to an impression of roughness or smoothness created by variation in tone or repetition of visual patterns across a surface. From these statements, it can be derived that texture is a neighbourhood property of an image pixel. For a texture to be characterized, it is therefore necessary to characterize the tonal primitive properties as well as to characterize the spatial inter-relationship between them. Texture measures are therefore amongst others dependent on the size of the neighbourhood taken into account by the formula.



Texture analysis can be approached in two different ways: structural and statistical. In the statistical approach, texture is considered as defined by a set of statistics extracted from a large set of local image properties. Statistical pattern analysis is performed by representing the way pixels are interrelated in an image by means of an intermediate matrix, from which a statistical feature is computed and used to build a texture image. Three analysis methods with the best results are:

- grey level dependence (co-occurrence) method
- grey level difference method
- grey level run length method

In this particular study, special attention is given to the **grey level difference method**. Texture analysis algorithms have different parameters (Serrano, 1990, Ch.6):

- displacement vector (4 directions, 3 distances)
- window size (contribute 90% to classification accuracy!)
- statistical features (contribution of 7% to final results)
- spatial resolution (parameter setting is more crucial)
- averaging effect
- radiometric resolution (3% contribution to final results)

Based on the influence each parameter has on the final outcome, the texture analysis performed here focuses on the window size and the statistical features.

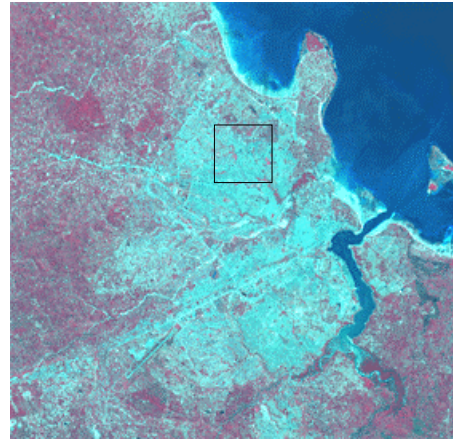


6 Methodology

6.1 Site selection

The entire image is a data set recorded by a SPOT-satellite in the month of May 1998 in both multi-spectral (XS) and panchromatic (PAN) mode. It covers a large area (even larger than the one shown in **Figure 5**) of the city of Dar Es Salaam and its surroundings.

For this project, a sub-scene of the entire image is selected near the town centre of Dar Es Salaam. This sub-scene not only contains both unplanned and planned settlements, but also swamp areas, infrastructure and bare soil.



ULX: 524830.0	ULY:9250580.0
LRX:528190.0	LY:9247300.0

Figure 5. Dar Es Salaam, SPOT XS with study area

It was not just for the reason of image information that this site was selected. Another factor was the availability of aerial photographs of the same scene. These could be used as ancillary data.

The data sets were available in ERDAS 7.x LAN-format, both geometrically and radiometrically corrected. They were first imported into ERDAS Imagine 8.3.1. and subsequently converted to the native IMG-format. Comparing the available aerial photograph with the SPOT-PAN image, the subset is selected. To define the subset, a box-drawing tool from the AOI-tools (Area Of Interest) in the viewer-menu indicates the subset. This Area Of Interest is input for the Inquire Box from the Utility-menu. This Inquire Box provides the co-ordinates for the Subset-dialog box. The same “mask” is applied on the imported SPOT XS-stack. To be able to provide these data sets to the ILWIS 2.1. package, the IMG-files have to be exported to the LAN-format.

The subset procedure could have been executed within the ILWIS 2.1. package. However, ERDAS Imagine 8.3.1. provides the user with an interactive tool to define the co-ordinates for the “masking”. Within ILWIS, the user first has to find out the co-ordinates of the upper-left and lower-right corners and use these co-ordinates to compute the number of rows and columns that cover the subset.

6.2 Image fusion

6.2.1 Fusion in ILWIS 2.1.

In the first instance, the ILWIS 2.1. package is used for the image fusion stage in the project. The LAN-files containing the sub-scene first have to be imported into ILWIS and converted to the native ILWIS MP# file format. The image-stack containing the XS-bands is split into 3 separate files.

To be able to fuse the SPOT PAN band (10 metre resolution) with the three XS-bands (20 metre resolution) all four have to have the same geometry. Based on the georeference-file of the SPOT-PAN the 3 XS-bands are resampled to 10 metre resolution.

Then the HSV-transformation algorithm is performed. First, the intensity of the colour composite has to be computed. The algorithm described by **Appendix A, Equation 1** is implemented into ILWIS using the Map Calc dialog box in the operation list. This intensity image has to be defined as an “image domain” file. This intensity image is input for the HSV-transformation. The intensity image is subsequently replaced by the SPOT PAN data, resulting in three new features. For the algorithm described by **Appendix A, Equation 2**, the Map Calc dialog box is also used. Now, the resulting files have to be defined as “value domain” files. These files are then stretched using a linear stretch, excluding the lower and upper one percent and saved as “image domain” files.



To be able to provide these files to the segmentation programme running under LYNEX, the ILWIS 2.1. MP# -files are exported to the ILWIS 1.4. MPI and MPD file format.

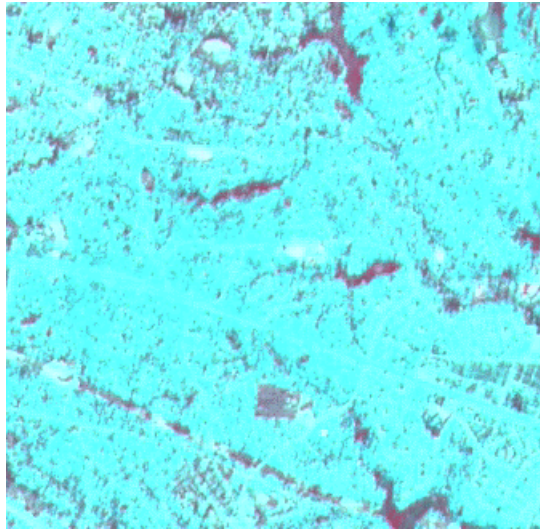


Figure 6. Fused SPOT-image (XS+P)

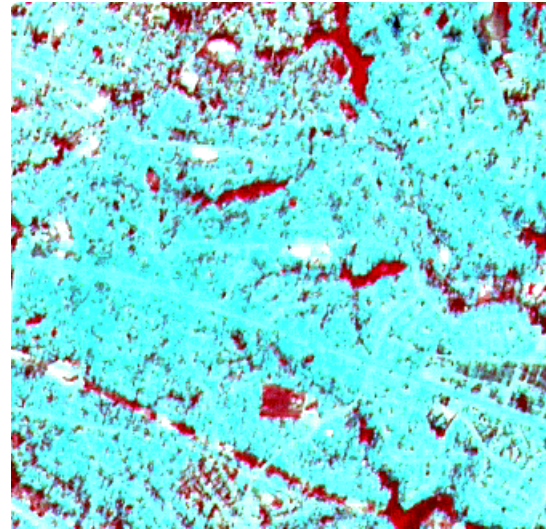


Figure 7. Stretched image

6.2.2 Fusion in ERDAS Imagine 8.3.1.

On second thought however, the ILWIS 2.1. procedure provided a very “push button”-approach. Especially the subsequent necessary stretching techniques are much more sophisticated within the ERDAS Imagine 8.3.1. package. Therefore, the same image fusion procedure covered in the previous section was also performed using this package.

In order to execute the resampling step, the two files containing respectively the SPOT XS and SPOT PAN data set were put into one layer stack. The resampling to the smallest spatial resolution (that of the SPOT PAN) is performed automatically by this stack operation.

Although ERDAS Imagine 8.3.1. provides a tool for IHS-transformation, the image fusion was performed using the HSV-transformation to deviate as little as possible from the ILWIS-procedure. It should be borne in mind, that there might be differences between the outcome due to rounding-off. Differences can not be attributed to transformation parameters for the same calculations were performed.

By choosing the HSV-transformation, the problem was encountered that ERDAS Imagine neither provides this tool as a “push-button” operation (as is the case for the IHS-transformation), nor could any of the calculations as described in **Appendix A (Equation 1/Equation 2)** be performed using the Utility-tools. This tool only provides tools to define calculations with just 2 input files. Therefore, the ERDAS Modeller-tool was used.

After calculation, the fused images were stretched. This was performed by choosing the contrast-tool from the Viewer-menu. By selecting the piece-wise linear stretch, full control was gained over the stretching. The breakpoints were interactively set by visual inspection. The result was exported to the .LAN-format to provide it for ILWIS 2.1. with inclusion of the stretch-look up table (LUT). Within ILWIS 2.1. the .LAN-format was imported to the native .MP#-format and subsequently exported to the ILWIS 1.4. formats .MPI and .MPD. Now the three stretched features could be handled by the UNIX software.



6.3 Ground truth

6.3.1 Land use map modification

From aerial photographs taken in 1992, land uses for the whole of Dar Es Salaam were established and digitized using Arc/Info. This land use polygon-file was clipped to extract the specific location of the study area. The subset was rasterized using the ILWIS 2.1. package to be able to compare the land use with the satellite imagery.

However, neither were all land use classes encountered in the study area, nor was this level of detail necessary for the purpose of this study (see **Figure 8**). These reasons necessitated a reclassification of the land use classes. For the non-residential land uses reclassification was performed by aggregation. Some classes that did not occur within the study area were just straightforwardly removed. The level of detail of classification for the residential classes was not altered. Probably optical data will not be able to distinguish between these levels of occupation density, but texture data hopefully will. By leaving some degrees of freedom within this classification, the strength and applicability of texture can be tested. The resulting land use classes encountered within the study area can be derived from **Table 1**.

Class number	Class description	Class number	Class description
11	High density planned residential	4	Institutional
12	Medium density planned residential	6	Recreational
13	Low density planned residential	71	Vacant Land
2	Industrial	8	Unplanned residential
3	Commercial		

Table 1. Land use classes

It should be noted from **Table 1** that within the class “Unplanned residential” there is no distinction in occupation density. Hopefully with the use of a texture data, this distinction can be introduced. Most of the area classified as vacant land coincides with the swamps often encountered in this area. Also, the classes covering the infrastructure were left out for the reason that on this coarse imagery these linear features are not easily recognized. For the classification the classes “recreational”, “commercial” and “institutional” were not taken into account for they do not relate to a land cover, but land use. For example, the class “recreational” might include both a basketball field and a soccer field. However, the land covers related to these land uses differ from one another. The basketball field falls in the imaginary land cover class “concrete”, whereas the soccer field is classified as “grass.”

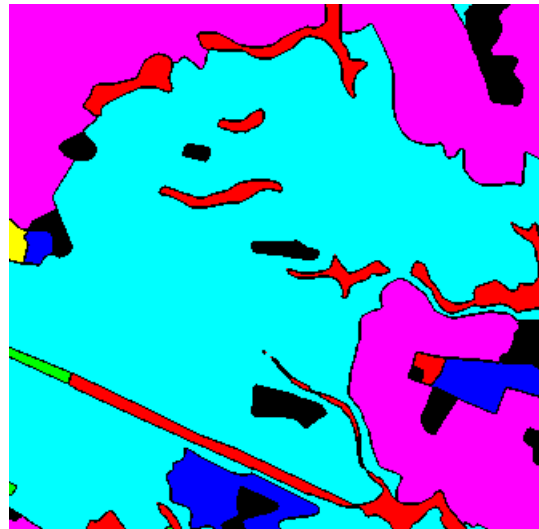


Figure 8. Land use map of study area

To facilitate a fair comparison of the classified image and the land use map, the boundaries between differently classified areas were removed. The land use map was exported from ILWIS 2.1. to the ILWIS 1.4. .MPD and .MPI format to provide them for the UNIX software.



First, the numbering of the classes was modified to a consecutive order using a map calculation algorithm. Then the borders were removed. Also for this procedure a map calculation was performed using a rank-order filter. It is assumed that within a homogeneous area of a 3 by 3 window, the 3rd and seventh entry of a rank order array having nine entries are equal, in a heterogeneous area they are not. Therefore, pixels with different outputs for both rank-order filters, meaning they have a heterogeneous neighbourhood, are classed as “unclassified” (having DN = 0). This removes the boundaries, so for comparison mixels are not taken into account.

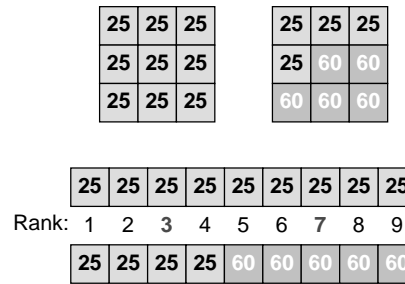


Figure 9. 3x3 -windows and their respective arrays

6.3.2 Sample site selection

The classification of the remotely sensed imagery is not performed using common image processing packages like ERDAS Imagine, ILWIS or IDRISI, but a Linux-based C-program. This programme does not define a training sample set as a spatial database itself, but derives the attributes from two different sources: one file covering the spatial and attribute (thematic) characteristics of the training sample set, another file containing the spectral characteristics. These are combined to create that spatial data set, containing both spatial, attribute and spectral data.

Since the spectral characteristics are derived from the remotely sensed imagery, the spatial and attribute characteristics have to be defined. Sites are selected by means of comparison of the land use map, the SPOT PAN-image and the aerial photograph covering the study area. Training samples for all classes are distributed randomly across the image to avoid bias and subjectivity. Class signatures were statistically derived from the three fused features. To be able to evaluate the classification result a test sample set is needed. In this case the whole land use map of the study area is defined as being the test sample set.

6.4 Classification

Now all data is available on the UNIX-platform. First, the geometric (sample site) and attribute (fused features) aspects of the sample set are combined. Then, the statistics are computed for each training sample, resulting in the feature signatures. As these files are created, the classification is performed using the Maximum Likelihood method. This resulted in a very poor classification (see Figure 10, Table 2 and Appendix C), albeit for the precautions taken regarding the land use map as described in the previous section (compare Figure 10 with Figure 8).

It is just for the class “vegetation” that both accuracy and reliability are passable. So, it can be concluded that the classification was only able to distinguish between built-up and non-built-up area. Within the built-up area no distinction was possible. However, exactly this distinction was the objective of the

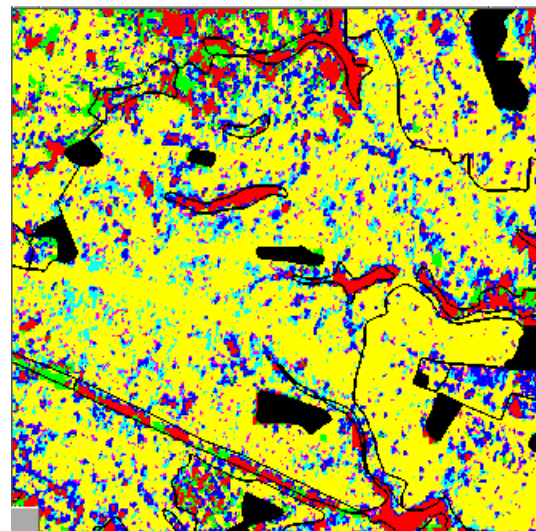


Figure 10. Classified image using ML

average accuracy	34.79 %
average reliability	25.66 %
overall accuracy	13.07 %
overall reliability	13.07 %

Table 2. Classification results

Alternative approaches are needed. First, classification will be improved in the spectral domain (segmentation), later in the spatial domain (texture).



6.5 Segmentation

The region-merging algorithm used for the segmentation requires three image bands - which are first combined into a feature vector quadtree - as input and gives one segmentation as output. The segmentation results in a set of objects with multi-spectral properties (mean vector and variance-covariance matrix).

The three feature files created within ILWIS are first converted into quadtrees to be able to provide them to the subsequent algorithms. Then these quadtree-files are combined using a quadtree calculation algorithm. Now, this combined quadtree is input for the region-merging algorithm. Several thresholds are supplied as well. From the object-table new objects with corresponding multi-spectral properties are derived using the same quadtree calculation algorithm. Finally, these quadtrees are converted back into raster-images to provide them for classification.

The classification is performed using the same algorithms as for the common classification however with different class signatures. These were not derived from the spectral information of the fused features, but derived from the segmented features.

However, also this approach exclusively uses the spectral domain for the classification. This explains the fact that with increasing segmentation thresholds the overall reliability and overall accuracy decrease (see Appendix D). As the defined classes do not differ spectrally, pixels with the same spectral characteristics, but with differing labels, are merged into bigger segments with only one label. Only for the class covering the non-built-up area the accuracy increases (see Figure 12). Segmentation will only facilitate a better classification of the two “hyper”-classes: “built-up” and “vegetation”. Therefore, the methodology used in the spectral domain can not be applied to answer the objective of this study. Another methodology is necessary in this case: the spatial domain.

6.6 Texture analysis

To maximize the separability of the two subclasses, the texture analysis algorithm should result in two highly different texture measures for the two subclasses. First, a texture analysis algorithm has to be selected which is most able to produce such a result. Especially texture coarseness seems to be the most appropriate dimension of texture to be investigated, for it relates to the rate of alternation of the grey levels. This points to the Grey Level Difference Method (Serrano, 1992, p.14). Talukdar also applies this texture measure algorithm in his urban research (Talukdar, 1997, p.15).

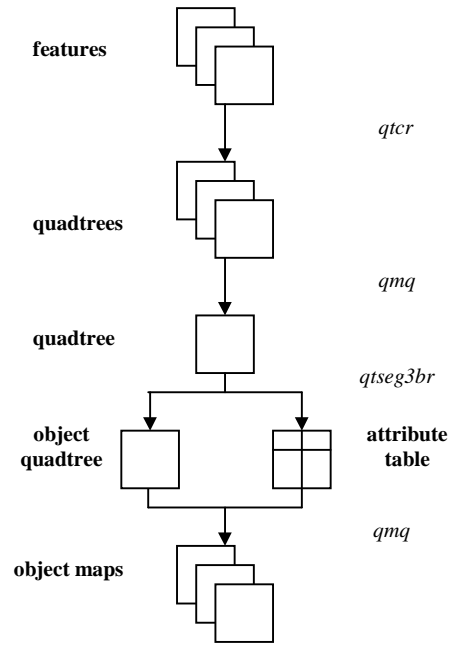


Figure 11. Flowchart image segmentation

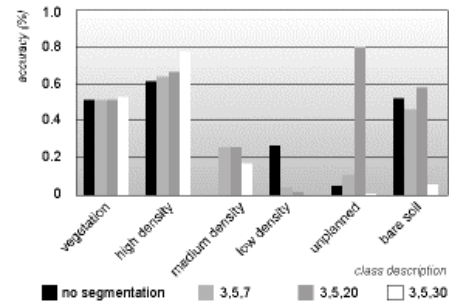


Figure 12. Accuracy per class using different segmentation thresholds



This algorithm is based on absolute differences between pairs of grey levels separated by δ . The absolute difference values are used to compute an array of dimension n , where n is the possible number of grey levels. The entries are frequencies of the absolute difference between two specified pixels in the image. In order to make a proper decision an artificial image is created with different textures on which all parameters can be variably applied under controlled conditions.

From all four statistical methods usually computed from the array, the ASM-method produced the most easily interpretable texture images. With increasing irregularity of texture, the pixels forming the elements of a texture got increasingly higher DN-values. However, the contrast within the texture sub-image had to be sufficiently high.

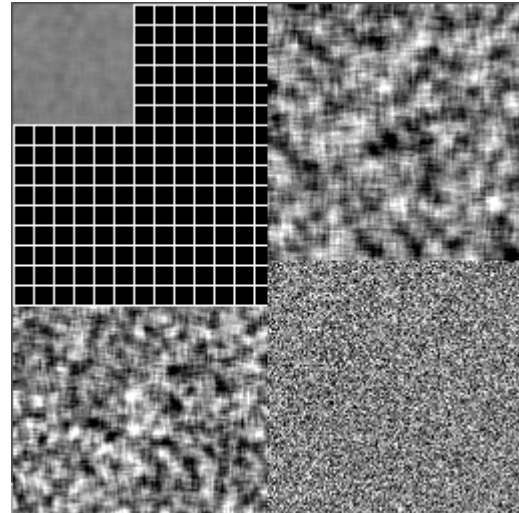


Figure 13. Artificial texture image

Having for every possible difference the same frequency (as is the case in a fully random image), the algorithm resulted in the lowest possible texture measure. When only a few differences occurred (as in the case of a very low texture), this resulted in the highest possible texture measures (See Appendix). The nearer the window size got to the repetition rate of the textural elements, the better the change from one texture to another was distinguished.

For this texture image the window size was set to “11”. By choosing this windows size, the texture sub-images were easily distinguished and also the borders between the texture sub-images were retained. The displacement vector had a length of $\Delta x=1$ and $\Delta y=1$, resulting in a direction of 45° . By choosing these parameter values, the spatial correlation of the pixels was maximal and the direction was not biased towards any of the texture sub-images. Choosing these parameters for the algorithm, the ASM-method was pretty good in discriminating between the different textures, even if their patterns seemed to be very much the same.

Applying the ASM-algorithm with the same parameters - be it for the window size, which was set to “17” - on a sub-image of the Dar Es Salaam scene, the distinction between areas with different texture was very poor. Also, the ENT and MEAN-method resulted in poor texture images after application on the Dar Es Salaam sub-image. The ENT-method produces the largest features for equal matrix entries and they are small when they are very unequal. So in an image with a regular pattern compared to the window-size, this will result in low features, for there are just a few outliers, whereas most of the input values are similar. If it is a fully random image, the entries are all the same, resulting in a large feature. However, as the image shows all really random value-patterns the outcoming features are all pretty much the same. The same goes for the MEAN-method, as it just computes, as its name indicates, the mean of all the entries of the array, using a weighted mean, depending on the grey level difference. As grey levels do not really change over the image within a window, the outcoming features as computed by the algorithm will differ significantly.

The drawback of the ASM-method, i.e. that the contrast within the texture had to be sufficiently high, is tackled by using the CON-method, whereas on the artificial image this did not proved to be very adequate. The larger the differences within the texture are, the larger the texture measure will be produced. Even if the frequency of the smaller differences is higher, this will be the outcome, for the difference itself contributes to the texture measure in a geometric progression whereas the COM-method is just a linear function of the frequency (See Appendix). After applying this image onto the Dar Es Salaam sub-image a moderate distinction was established between settlements with differing occupation densities, especially between planned settlements and the unplanned settlements, for the contrast change differs a lot between these urban sub-classes.



6.7 Integration of texture image into the classification

The texture image is derived from the SPOT PAN image using a Linux based C-programme. To be able to fuse the image with the multi-spectral data set and apply a stretch, the texture image was imported into ERDAS Imagine via ILWIS 2.1. After the fusion and stretching as performed with the SPOT PAN data integration, the three fused features were again imported into the Linux-based C-programme via ILWIS 2.1. Now a classification was performed on this data set consisting of the SPOT XS data fused with the SPOT PAN-derived texture image. The results (see **Appendix G**) were better than using just the spectral information content of the SPOT data set. This indicates that texture analysis might contribute substantially to classification of urban settlements.

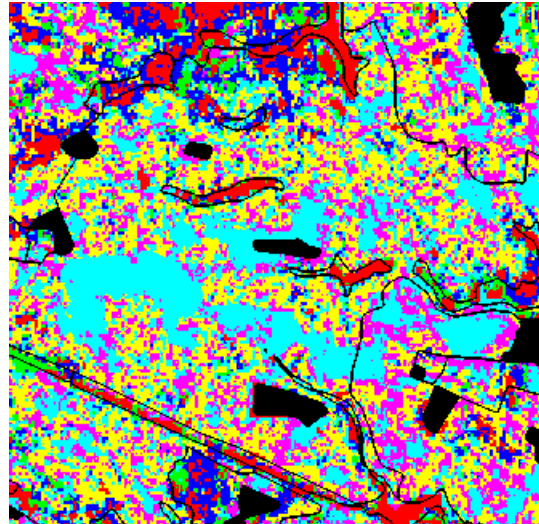


Figure 14. Classified image using texture



7 Discussion

On first sight, a classification improvement of almost 20 percent seems to be a very good result of this study. However, this result should be seen in perspective. To be able to evaluate the real meaning of this, a few caveats have to be borne in mind.

First, one has to be aware of the drawbacks of the ground truth data, a constant parameter in the comparison of the confusion matrices, although precautions were taken in advance. Already some inconveniences are reviewed in **section 6.3.1**. The fact that ground truth (**land cover**) is derived from a **land use** map is certainly a factor to be frowned upon. Furthermore, the level of detail encountered in the satellite imagery exceeds the generalization level of the land use map by far. So, pixels with different land cover may be grouped into one land use class. This might explain the low reliability measures. A last remark on this subject should focus on the temporal friction. The land use data was established from aerial photography taken in 1992. The remote sensing imagery however was recorded in the month of May of 1998. Occupation densities of the different settlements within the study area might have increased due to the overall scarcity of new, low-value construction sites and consolidation activities. Also, especially those zones classified as being “vacant land” might be intruded by construction activities, negatively influencing the classification results.

Secondly, to take the point on generalization level mentioned in the previous paragraph even further and to hint at the last issue, care should be taken regarding the resampling procedure. As mentioned in **section 6.2** image fusion can be performed to attain various objectives. In this study, the reason is data reduction, for the segmentation algorithm is only able to cope with three input features, whereas four inputs are available. To be able to fuse the data, the four inputs have to have the same geometry. This implies that the multi-spectral data set has to be resampled (see **chapter 2**). This leads to a higher level of spatial detail in the data set. This however introduces several inconveniences with regard to classification of urban classes as discussed in **chapter 5**.

Another point stressing this very issue is concerned with texture. As described in **chapter 5** and further elaborated upon in **section 6.6**, texture is a spatial information component of a data set. The texture measure is not an attribute of a pixel itself, but an attribute depending on its neighbourhood, taking into account numerous pixels. Based on this hypothesis, the classification of a texture image on the pixel level (10-meter resolution) is a misleading methodology.

Finally, care must be taken while considering the texture image. From **section 6.6** it can be derived that the methodology described, from a scientific point of view, is not a valid method to estimate algorithm parameters for texture analysis by applying them first on an artificial image. An explanation might be that in a real image the textures and grey level differences are much more similar than in an artificial image. Therefore, it should be concluded that another methodology has to be followed to come to good parameter estimation for a result. Whereas the theory behind texture analysis seems to be rather clear on this subject, the practical experience proves otherwise.

8 Conclusion and recommendation

With regard to planning issues, insight into the discrimination of settlements based upon occupation density is a valuable asset to administration. Remote sensing might provide a useful tool. However, straightforward image processing and analysis techniques do not suffice this need. Texture analysis of remotely sensed data might be more apt to extract the information required. But, based upon this study, no substantial evidence was generated to support this hypothesis.

The results presented here are not persuasive by itself, but taking into account the caveats mentioned, a more fundamental research should be carried out to evaluate the real contribution of texture information to classification improvement in an urban setting.



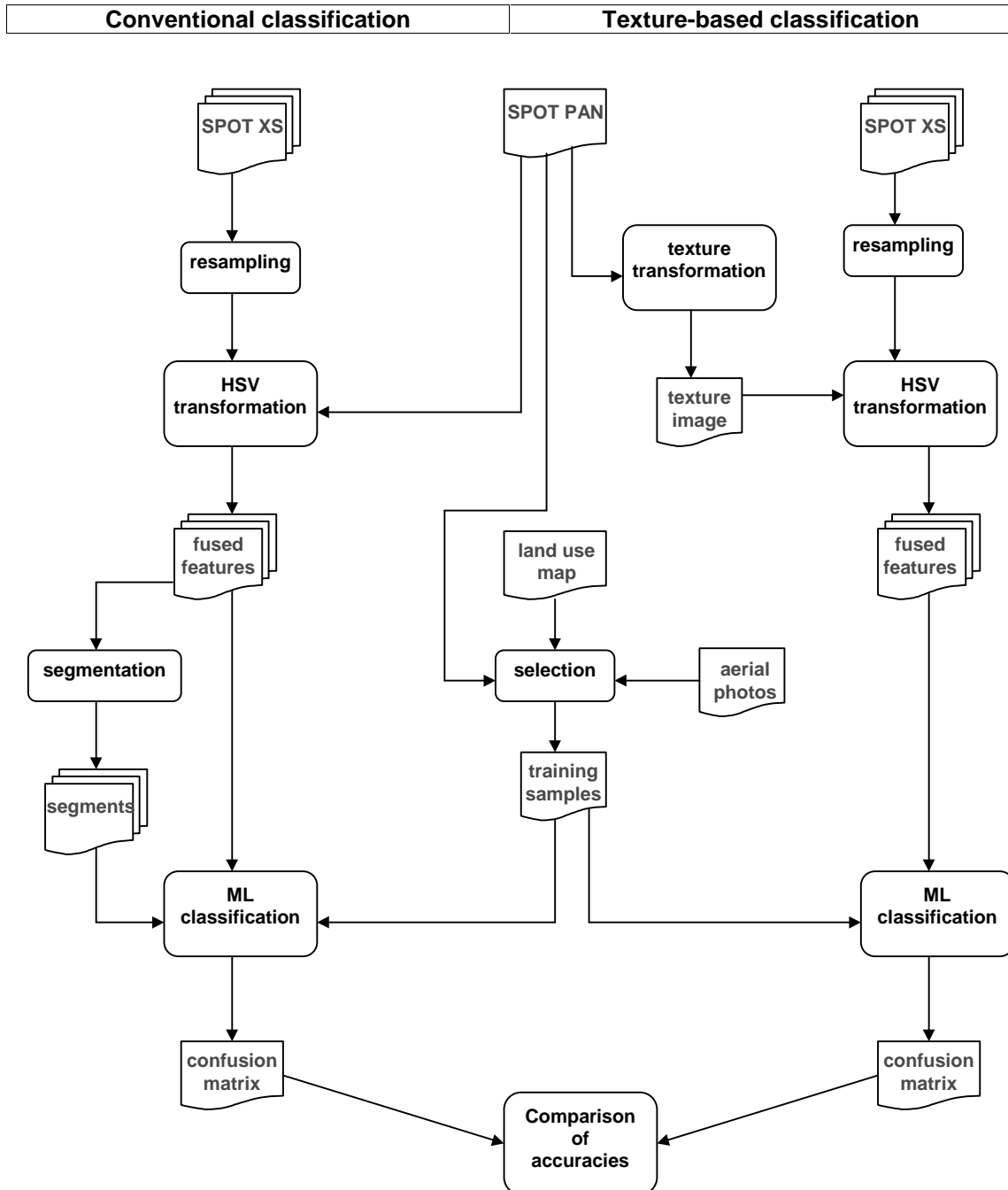
9 References

- Atikah Mohamed Hashim, S. (1998), Research proposal: *Monitoring the urban growth of Dar Es Salaam using Remote Sensing*.
- Gorte, B. (1998), Probabilistic Segmentation of Remotely Sensed. Ph.D. -thesis ITC-Enschede.
- Project proposal NRSP 1990-2000: Urbanization in Dar Es Salaam. (1992) BCRS-Programme-Bureau
- Serrano, C.M.P.(1990), Performance evaluation of Texture Analysis Procedure for Remote Sensing Images. M.Sc.-thesis. ITC-Enschede.
- Talukdar, K.K. (1997), Recognition and Extraction of Spatial Objects from Satellite Data using GIS and Image Processing Techniques for Urban Monitoring: *case from Delhi Mega-City, India*. M.Sc.-thesis. ITC-Enschede.
- Westen, C. van & J. Farifteh (1997), ILWIS 2.1. for Windows User's Guide. ILWIS Department, ITC-Enschede.



10 Appendices

Appendix A: Project flow chart





Appendix B: Image fusion by HSV-transformation

$$DN_{intensity} = \frac{DN_{XS1} + DN_{XS2} + DN_{XS3}}{3}$$

Equation 1 Computing the intensity image

$$DN_{feature_N} = \frac{DN_{XSN}}{DN_{intensity}} * DN_{PAN}$$

Equation 2 Substituting the intensity image with the SPOT PAN

Appendix C: Land use classes

Class number	Class description	Class number	Class description
21	High Density	46	Institutional
11	Medium Density	51	Roadways
12	Low Density	52	Railways
13	Small Scale Industrial	53	Airport
22	Utilities	54	Transshipment
23	Other Industrial	61	Built-up Center
24	Central Business District	62	Organised open space
31	Other Commercial	63	Other recreation
32	Commercial	71	Vacant land
33	Institutional	72	Gardens
40	Training	73	Water Bodies
41	Administration	74	Agriculture
42	Community Facilities	80	Unplanned residential
43	Other Institutional	90	Forest
44	Large Scale Industrial		Island

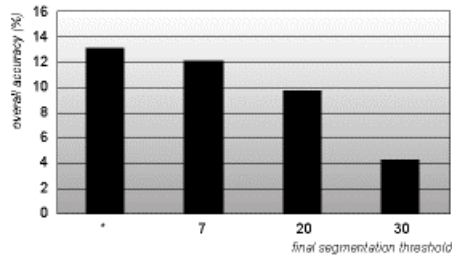
Appendix D: Confusion matrix of ML-classification

	1	2	3	4	5	6	Accuracy
1	2834	797	906	158	273	561	0.51
2	18	146	48	10	4	12	0.61
3	382	1387	993	245	263	461	0.27
4	1695	19474	3127	1366	1751	1727	0.05
5	2681	39907	6237	3060	7559	1344	0.12
6	44	36	23	9	6	130	0.52
Reliability	0.37	0.00	0.09	0.28	0.77	0.03	
Average accuracy:		34.79 %		overall accuracy:		13.07 %	
Average reliability:		25.66 %		overall reliability:		13.07 %	

Class number	Class description
1	Vegetation
2	High density
3	Medium density
4	Low density
5	Unplanned
6	Bare Soil



Appendix E: Overall accuracy for different segmentation thresholds



Appendix F: Grey level difference method

Angular Second Moment (ASM)

$$ASM = \sum_{i=0}^n P\delta(i)^2$$

Contrast (CON)

$$CON = \sum_{i=0}^n i^2 P\delta(i)$$

Entropy (ENT)

$$ENT = \sum_{i=0}^n P\delta(i) * \log P\delta(i)$$

Mean (MEAN)

$$MEAN = \frac{\sum_{i=0}^n i * P\delta(i)}{n}$$

Appendix G: Confusion matrix of classification using texture

	1	2	3	4	5	6	Accuracy
1	2750	429	830	374	617	529	0.50
2	44	94	35	26	37	2	0.39
3	375	631	1012	577	750	386	0.27
4	2141	8189	3399	7448	6771	1192	0.26
5	2502	20163	3742	11319	21265	1797	0.35
6	52	13	44	12	3	124	0.50
Reliability	0.35	0.00	0.11	0.38	0.72	0.03	
Average accuracy:		37.82 %			overall accuracy:		32.80 %
Average reliability:		26.58 %			overall reliability:		32.80 %