Implementing a procedure to extract urban areas based on Multispectral Classification and Mathematical Morphology

Alberta Bianchin (*), Stefano Foramiti (*)

(*) DP-IUAV, S.Croce 1957, 30135 Venezia, tel 041.2572163, fax: 041.5240403,
e-mail: alberta@iuav.it e-mail: foramiti@iuav.it

Abstract
We present an exploratory project aiming at progressing in studies concerning the following issue: when assessing the thematic classification accuracy, which should be the appropriate scale of the ground truth in relation to the resolution of satellite images used in the classification procedure?
The work presented is the implementation stage of the procedures necessary to carry on this kind of studies with regard to the urban theme.
The scheme developed includes: 1) the implementation of a classification procedure to extract urban areas based on multispectral classification and mathematical morphology, 2) a series of morphological transformations on the original vector map used as training and test areas, 3) the calculation of kappa index.
The whole procedure has been satisfactory tested through a case-study. As to the problem stated above, the outcomes of the case study need to be further investigated.

Research Questions
This work is part of a larger study on the quality assessment of the urban object recognition in remote sensing image processing. General issue is : as urban objects are scale-dependent, which is the appropriate scale of the ground truth for which resolution of image processed in order to validate the result of image classification.
To carry out the research, a procedure to extract built areas performed in our previous work (Bianchin, Pesaresi, 1993) has been implemented using GRASS and completed by a validation procedure of the result through Kappa index.
The ground truth is provided by a vector map at scale 1:5 000, produced by the Regional Administration called CTRN (Carta Tecnica Regionale Numerica), which includes different layers (building, roads, etc). The building layer of the CTRN has been processed by using some morphological operators in order to obtain different resolution of the urban theme.
The urban area of Treviso and its surroundings, in the Venetian region, has been selected as study-case. This is an area representative of the urban phenomena called "dispersive city".

The procedure adopted
The procedure adopted to extract urban areas from satellite image (Bianchin, Pesaresi, 1993) is explained by the schema below. It is based on two different processing chains for the two images: Landasat-TM and SPOT-Pan. Their results are merged to obtain an intersection image.
The first chain is a two-theme (urban – no-urban) supervised classification of multispectral Landsat-TM image.
The second chain is based on mathematical morphological operators applied to the SPOT-Pan image to extract a "structural' information related to the same theme. It starts from the monospectral SPOT-Pan image to generate the morphological gradient image PAN1. Building areas show a high value of the morphological gradient. PAN1 is re-sampled to be at the same resolution of TM so it

may be intersected with TM. PAN2 is binarized by a threshold determined interactively on the basis of ground truth, to obtain a two-theme classification. PAN3 image and TM classified generate the intersection image IOUT.
The basic idea of the procedure is to improve the accuracy of the TM classification trough the "structural" information extracted from the morphological processing of SPOT-Pan.

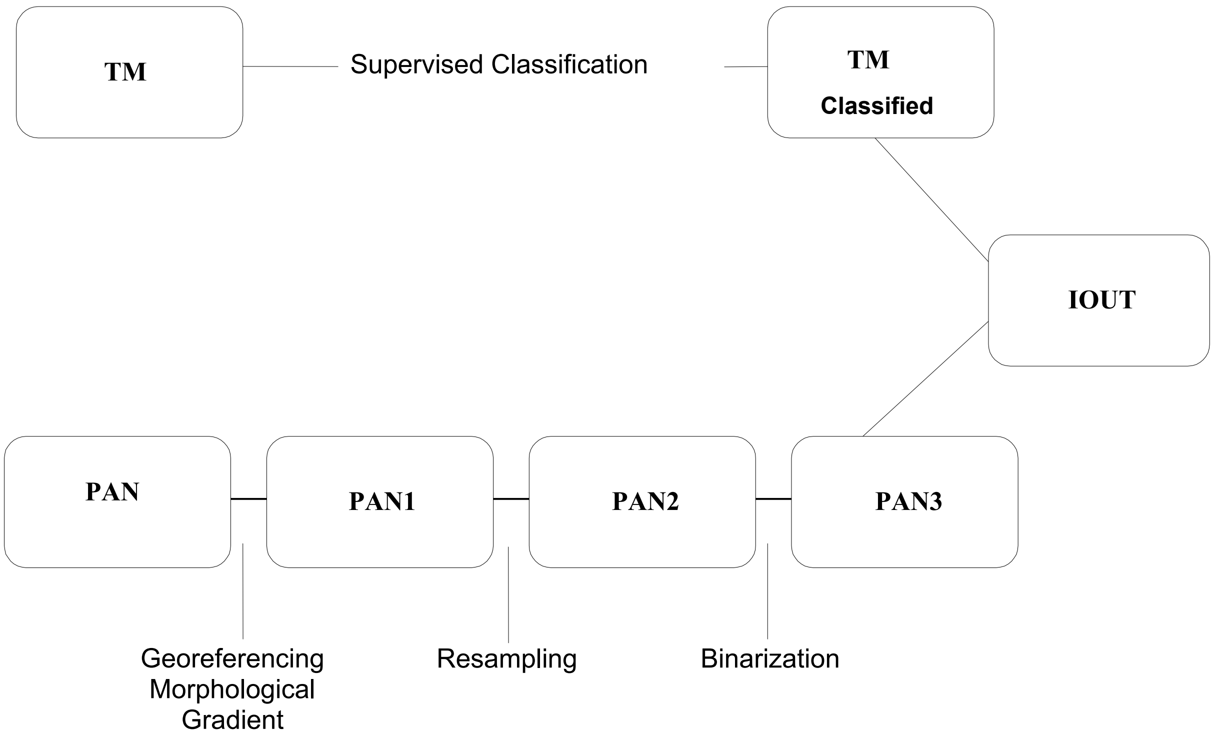

Figure 1 – Scheme of the procedure

The empirical case-study: Treviso
In the area selected - Treviso and its surroundings - to experiment our research questions, image data available are:
Landsat5 TM, 14^{th} July 1994 (TM940714),
SPOT-Pan , 17^{th} August 1988 (063258),
CTRN, 1:5 000, BIM-Piave, 1995, (sheets: 105 072, 105 083, 105 111, 105 112, 105 123, 105 124).

The operating phases of the present job have been the following: ·

- pre-processing of the input data;
- extraction of urban objects by a radiometric supervised classification of multispectral image Landsat-TM and analysis of its accuracy;
- implementing the procedure of mathematical morphology to be applied to the SPOT-Pan image;
- checking accuracy on the intersection image obtained.

The research has been developed using the software MapInfo, Idrisi and Grass. We took advantage of the peculiar characteristics of each program in the different phases of job in order to improve efficiency and operating flexibility.
The process for the extraction of urban areas has been developed entirely with Grass on Linux platform. The whole procedure has been automated by scripts containing iterative and conditional instructions.

The stage of data pre-processing included the following operations:

1. Selection of the study-area, the relating map sheets and information layers of interest from the CTRN.
2. Conversion into raster format of the extracted layers of the CTRN.
3. Preparation of the training and test maps for the supervised classification of Landsat-TM and following accuracy assessment.
4. Geo-referencing and rectifying the layer of the CTRN and the Landsat and SPOT images.

The processing chain applied to Landsat-TM image consisted of:
1. Geometric correction in order to have a common reference grid for Landsat-TM and CTRN, which allows the geometric correspondence of urban objects of two images (TM and CTRN).
2. Supervised classification of the Landsat-TM multispectral images for urban areas extraction, using the CTRN building layer as training set and test set later in the accuracy assessment.

The procedure based on mathematical morphology for the urban areas extraction, applied to the SPOT-Pan image, consisted of the following operations:
1. Geo-referencing the SPOT image using the CTRN as reference system.
2. Application of the morphologic gradient to panchromatic image SPOT-Pan.
3. Re-sampling the morphologic gradient image by a pass-low filter " erosion " (inferior function) with kernel 3x3.
4. Binarization of the image through an opportune threshold value.

Ultimate stage:
1. Intersection (AND operation) between the classified image Landsat (urban class) and binary image processed from SPOT-Pan.
2. Calculation of confusion matrix and Kappa index.

Implementation in GRASS

Urban areas extraction from Landsat TM
In Grass have been created three databases (locations):
- Landsat TM-location containing the Landsat TM image (7 bands),
- SPOT-PAN location containing the SPOT-PAN image,
- Treviso-Location containing the CTRN (vector and raster geo-referenced maps – DATUM WGS 84 - cartographic projection UTM 33).
-

The multispectral image Landsat5 TM, 14th July 1994 (TM940714), is used to extract urban areas of Treviso zone. The study-area has been cut from the original scene and introduced in the GRASS location-Landsat TM.

The raster binary image of Treviso zone, processed from CTRN, have been imported in GRASS in the Treviso -Location.

Rectification of the CTRN on the Landsat TM reference system.
The images of the CTRN whose buildings layer will be used as training and test areas, have been rectified on the Landsat-TM reference system so that the two grids could be superposed one upon another.
This was made in order to avoid geo-referencing the Landsat-TM as it would have alter the radiance and the co-occurrence of adjacent pixel, effects that are normally determined from the re-sampling operation necessary in the geo-referencing phase.

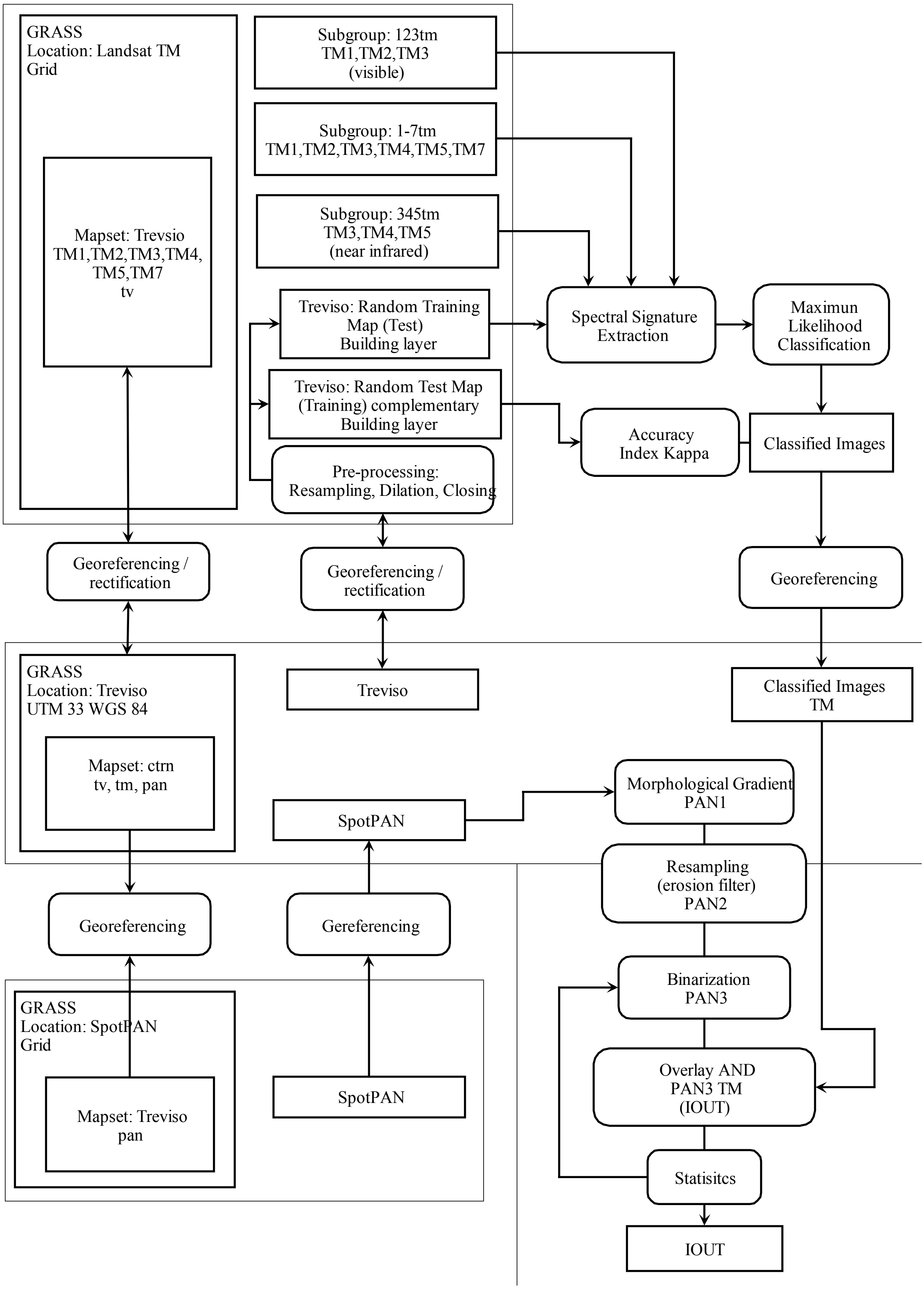

Figure 2 – Scheme of the whole procedure implemented in GRASS.

Figure 3 - Two classes (buildings, no-buildings), geo-referenced image TV.X, resolution 5m x 5m (Location Treviso), turned out after rasterization of layer 1 (buildings) of the CTRN.

To facilitate the detection of points suitable for the image rectification (i.points), concerning the CTRN, they have been selected into the layers : roads (s), railroad (fe), hydrography (f), in the CTRN- Location.
Concerning the Landsat-TM-Location, points have been selected in the second principal component image resulted by a Principal Component Analysis (i.pca) of the set of TM-images.

The method adopted for the rectification phase is based on the 2nd order polynomial transformation. The sampling operation uses the nearest neighbors (i.rectify2) function.
The resulting Overall RMS Error : **1.31** (pixels).

The rectification error is normally acceptable for standard deviation (RMS) values inferior or equal to 0,5 pixel (the Landsat cell dimension is of 1 x 1). The characteristics and the type of data used do not allow to come under the error value obtained without renouncing to points essential to the referencing phase.

Supervised classification of urban areas
The image of buildings layer obtained from the CTRN (tv.x), once rectified and inserted in the Landsat-TM-location has been used as a two-classes (buildings, no-buildings) training area in the phase of multispectral radiometric classification of the Landsat-TM scene of Treviso.
Three sets of bands have been used for the classification (group):

- the set of bands of the visible region TM1; TM2; TM3 (indicated with .123);
- the set of bands of the near-infrared region (TM3; TM4; TM5 - indicated with .345);
- the whole set of bands TM, excepted the TM6 of the thermal infrared region (TM1; TM2; TM3; TM4; TM5; TM7 - indicated with .1_7).

In detail, the procedure of extraction of urban theme, through the radiometric classification, includes the following steps:

1. the preparation of the training image at two-classes (buildings, no-buildings) obtained from the CTRN through a random sampling of cells (maps random - r.random),
2. the application of the learning algorithm (i.gensig), using the random map as training image and each of the three sets of bands, specified above TM (123, 345, 123457);
3. the classification, in the strict meaning, using the maximum likelihood (i.maxlik) function and each of the three sets of images TM (123, 345, 123457);
4. the calculation of the Kappa index (r.kappa) for the accuracy assessment using the image complementary to the training one as test image.

The different dimension the cells of the raster grid between the map of buildings extracted from the CTRN (pixel of 0.25 x 0.25) and Landsat-TM images (pixel of 1 x 1 equal to 30m x 30m), has forced to modify the training map and to experience a certain number of solutions in order to make the information derived from the CTRN consistent to the Landsat-TM resolution.
Therefore the following transformations have been applied on the training map (Location LandsatTM, grid 0.25 x 0.25):

- the grid transformation in dimension, from cells of 0.25 unit to cells of 1 unit (Landsat TM dimension), with re-sampling (RES);
- numerical closing (CLOSE) operations with kernel of 3x3, 5x5, 7x7, 9x9 applied on the grid of 0.25 unit resolution;
- numerical dilation (DILATE) operations with kernel of 3x3, 5x5, 7x7, 9x9 applied on the grid of 0.25 unit resolution.

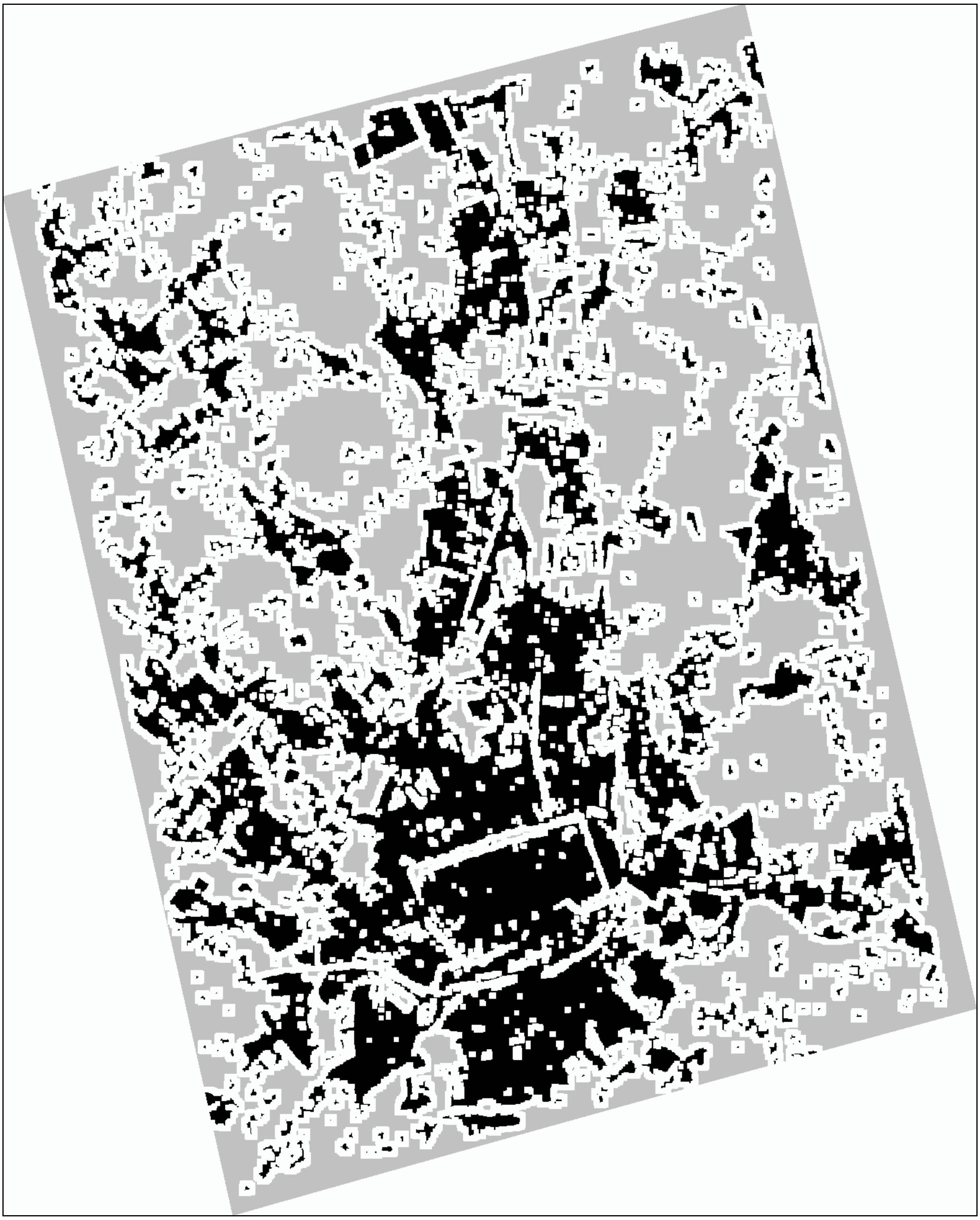

Figure 4 - TV.CLOSE.7.PRE image, resulted from the numerical closing transformation of TV.X ***by 7 x 7 kernel (CLOSE.7) and the erosion of no-urban areas (PRE).***

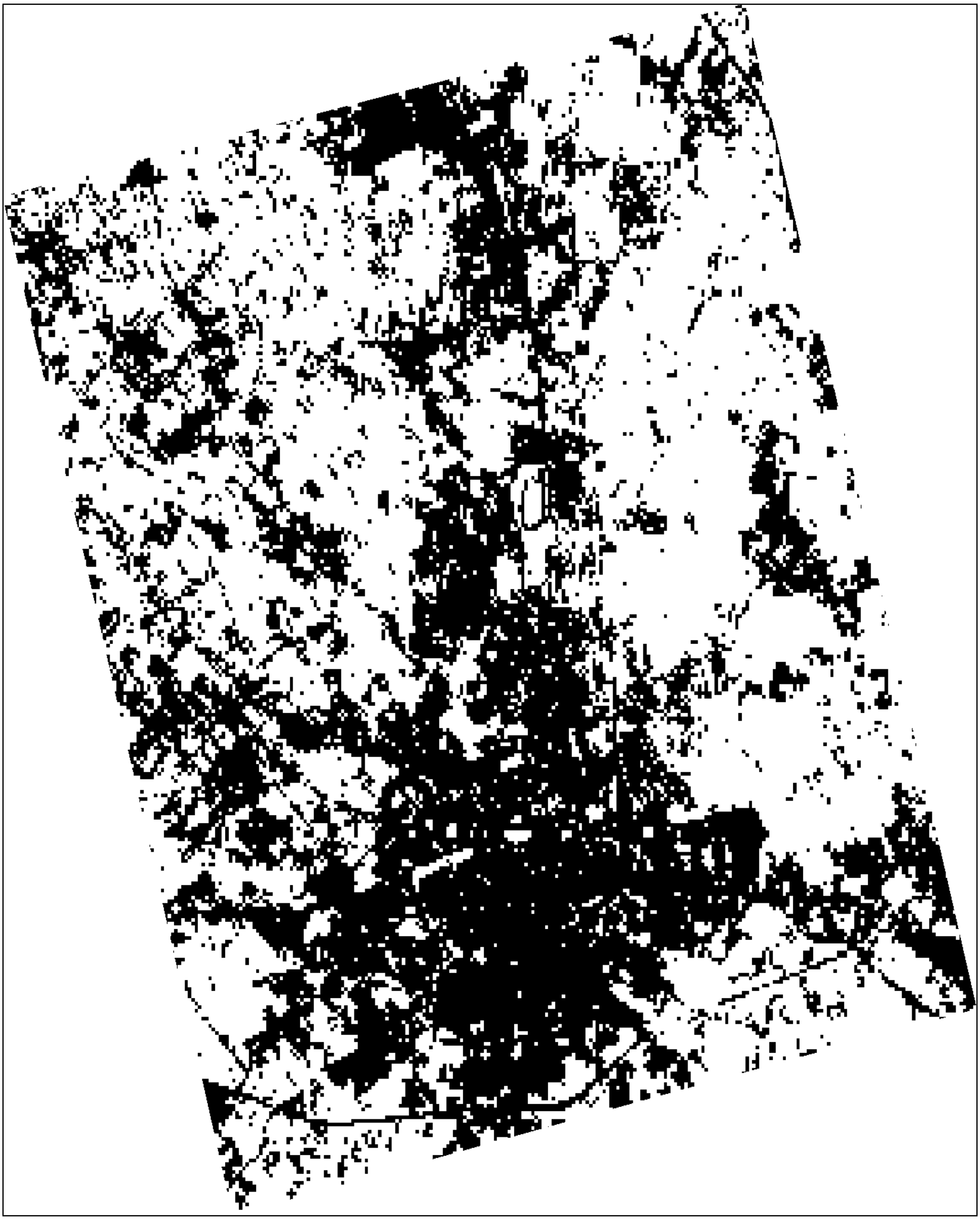

Figure 5 - ML.CLOSE.7.PRE.1_7. - Maximum likelihood supervised Classification of the ***Landsat-TM image – six bands 1_7*** *– training set from TV.CLOSE.7.PRE.*

Processing chain of SPOT-Pan and final intersection

The SPOT-Pan image (pixel of 10mx10m; 0.51-0.73 µm) dated 17th /08/1988, is processed through a procedure for the urban areas extraction based on mathematical morphology operators (Bianchin - Pesaresi, 1993).

The Spot PAN image has been geo-referenced on the CTRN cartographic reference system.
The geo-referencing accuracy : Overall RMS Error : 5,39.

The maximum standard deviation (RMS) value for the referencing operation is supposed to be equal to 5.00 pixel (the SPOT Pan cell dimension is of 10m x 10m). It has not been possible to go under such threshold. This obviously introduces an uncertainty in the procedure

Now, the morphologic gradient operator has been applied on the SPOT-Pan geo-referenced image. The gradient PAN1 image shows all the radiometric boundaries.
The PAN1 image is re-sampled by a low-pass filter " erosion " (inferior function) with kernel 3x3, in order to keep the urban areas characterized by high values of gradient for regions larger than one pixel and to eliminate the gradient regions of less than one pixel
Having the training urban areas from the CTRN, a threshold value of the gradient values, which distinguishes urban areas from non urban areas can be found by comparison with the ground truth. A two classes image is obtained by applying that threshold value.

The search of the threshold value can be optimized by using a recursive calculation of the Kappa index, in function of the threshold value which is incremented at each step. The best threshold value corresponds to the minimum Kappa index.
Then the two results of two processing chains on the TM and SPOT-PAN are intersected in order to have a more accurate classification.

Just two remarks about the procedure:

- the optimizing calculation of threshold value is made on the intersected image (IOUT),
- the TM classified image must be geo-referenced in the CTRN reference system before the intersection with SPOT-Pan chain.

The classified Landsat TM image, resulted from the maximum likelihood supervised classification by using the six bands: TM1; TM2; TM3; TM4; TM5; TM7, training set from TV.CLOSE.7.PRE. Better overall Kappa index : 0,692841.

PAN3 image resulted from the application of the threshold value on the gradient image PAN2, obtained by a low-pass filter re-sampling of the PAN1. The threshold value cuts the gradient low values in order to bring out the urban areas, which are characterized by a texture of high values of gradient and by a marked contrast with surrounding areas.

Figure 6 - SPOT PAN image of the city area of Treviso.

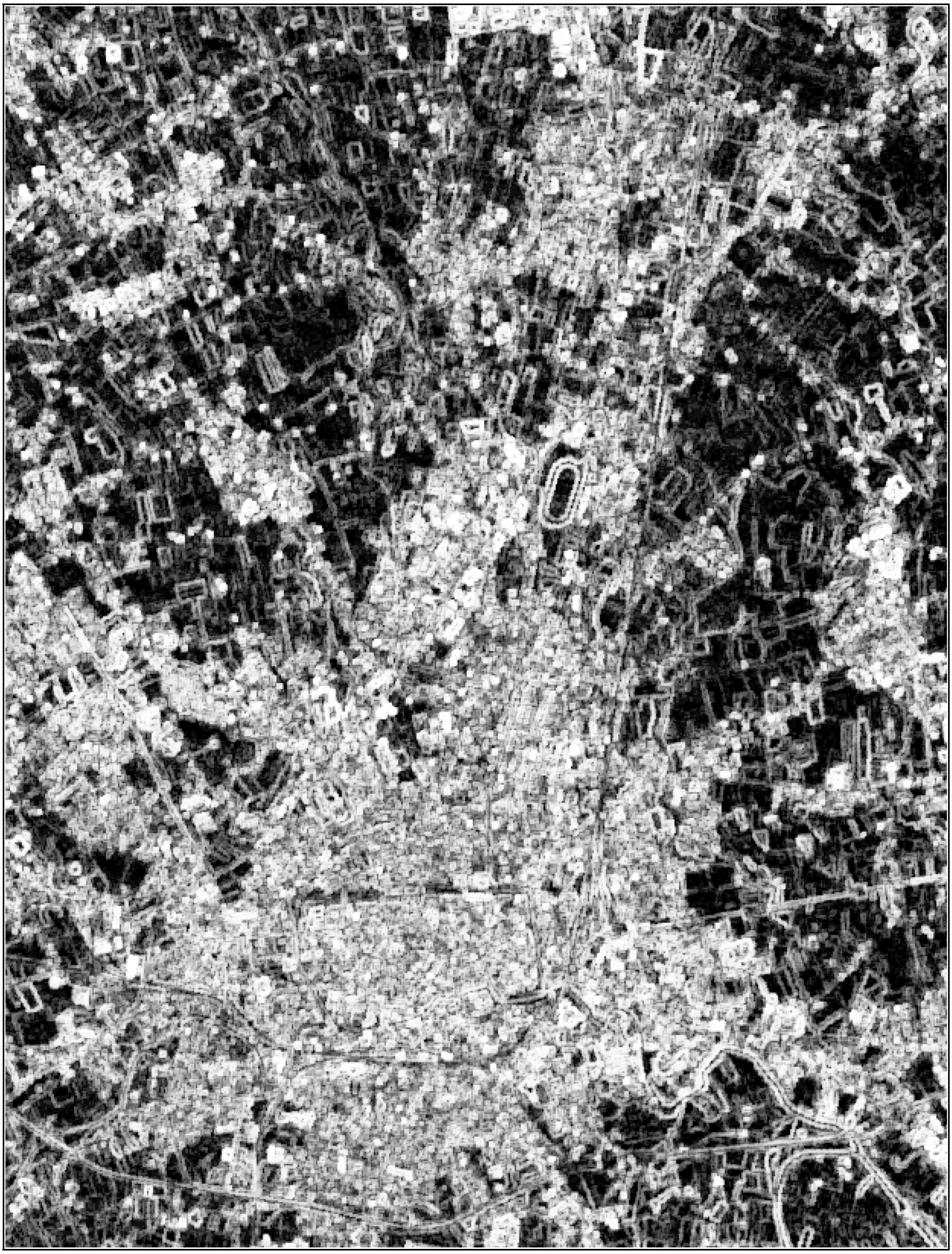

Figure 7 - PAN1 – Morphological Gradient image.

Analysis of the accuracy

The accuracy assessment has been calculated through the error matrix and the different indexes defined in the literature : Estimated Kappa for each class, Overall Kappa index, % of observed correct. Calculation has been done for all the classifications resulted from the different sets of bands of TM and different elaborations of CTRN as training areas, for the total amount of 42 classifications (Bianchin, Foramiti, 1999).

The best result has been obtained for the configuration using six bands of TM (TM.1_7) and TV.Close 7 as training map.

The map of urban areas, produced from the procedure described above, is compared for accuracy assessment with the map TV.Close.7 test. Test areas and training areas are obtained from the same map TV.Close.

The best value of the Kappa index for IOUT is done by the following report.

IOUT (PAN3 - Threshold 19) and TV.Close.7

1 no-urban areas

2 urban areas

Kappa:	0.523905
Kappa Variance:	0.000008
% Observed Correct:	79.758261

In detail:

```
ACCURACY ASSESSMENT

MAPS: MAP1 = close.7 in CTRN (tv in CTRN)
      MAP2 = iout in CTRN)

Error Matrix
Panel #1 of 1
                 MAP1
     cat# 1          2          Row Sum
 M      1  302879    70295     373174
 A      2  32614     102612    135226
Col Sum    335493    172907    508400


Cats % Commission   % Ommission    Estimated Kappa
1    81.162943      90.278784      0.446132
2    75.881857      59.345197      0.634518

Kappa          Kappa Variance
0.523905  0.000008

Obs Correct         Total Obs      % Observed Correct
405491              508400         79.758261
```

It must pointed out that the better threshold value (19), applied to PAN2 to obtain PAN3, that maximizes the Kappa index of IOUT is different from the threshold value (20-24) which maximizes the percentage of total correct observations.

IOUT

Maximum Kappa Index value (0,523905): Threshold 15-19
Maximum % Observed Correct (80,137097 %): Threshold 20 – 24

The graphic below shows the variation of three accuracy parameters, with increment of 5 digital numbers of the threshold value used in the phase of binarization of PAN2.
TV.CLOSE.7 is the test area used as reference.

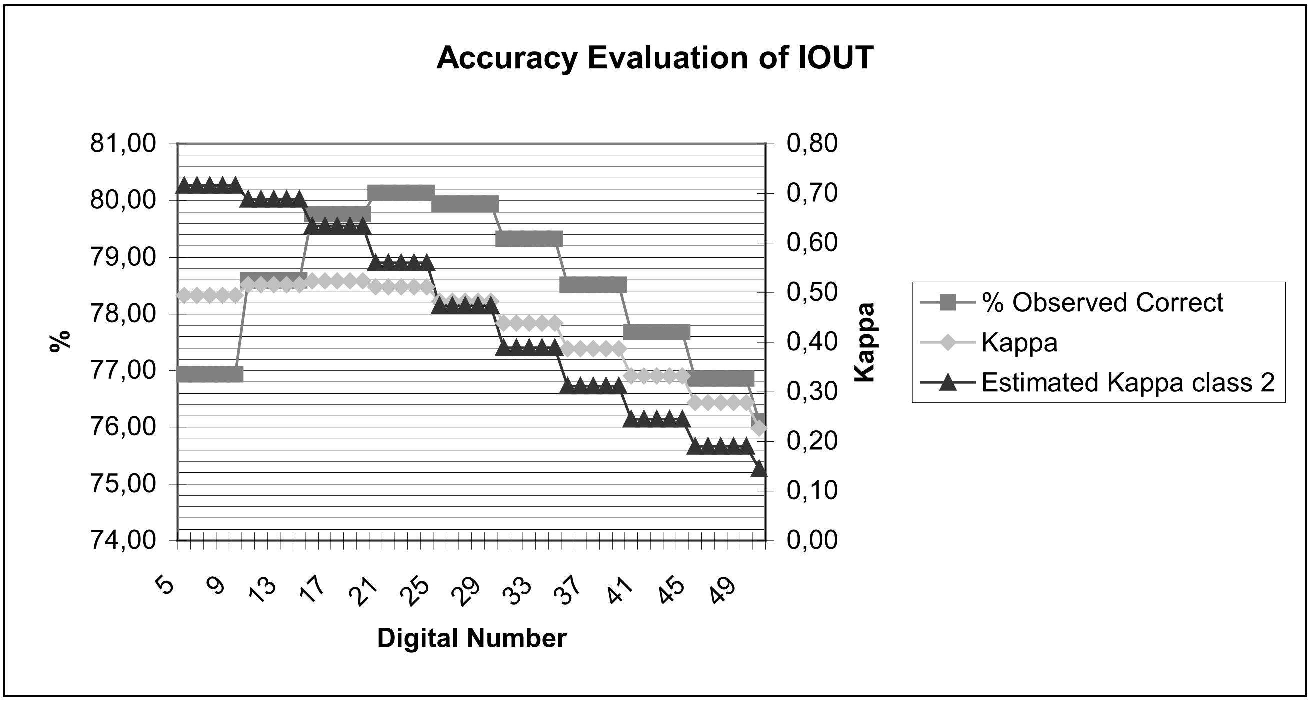

Figure 8 – Accuracy indexes of IOUT with TV.Close.7 as reference, in function of the threshold value, applied to PAN2 to obtain PAN3.

The urban areas present in IOUT which is the overlay of PAN3 (the threshold value is 19) and TM, through the operator AND, is the result of one "reductive " operation. The overall Kappa value of IOUT (0.523905) is less than the one of TM (0.692841) and better than the one of PAN3 (0.325620). The percentage of total correct observations of IOUT (79.8 %) is less than the one of TM (84.9 %) and better than PAN3 (64.1 %).

Conclusions

It can be stated that Grass is an adequate environment for this kind of studies. It provides different modules satisfying the tasks involved in the approach proposed : geo-referenciation and rectification, supervised multispectral classification, map algebra operations, kappa index calculation.

About the results of the case study carried out, it is worth noting that the images available have not been taken in the same period. The SPOT-PAN is 5 years older than the Landsat-TM. This fact can explain some bad results of IOUT accuracy compared to the Landsat-TM accuracy.

As regard the approach chosen to address the multi-scale and multi-resolution issues, many considerations can be done.

First of all, the precision of the superimposition of different images -Landsat-TM, SPOT-PAN and CTRN - is the crucial point in order to obtain good results by this approach. The idea to use the CTRN as ground truth is a practical proposal once the CTRN should be an available and updated document.

Secondly, the discrepancy between the variation of Kappa Index and of the Percentage of total correct observations in function of threshold value which appears from the figure 8 calls for more understanding about the deep significance of such indexes.

Moreover, further investigations need to be made in how the morphological operations applied to the CTRN intervenes in defining the urban structures.

References

AA.VV.,(1994) - *Qualità e aggiornamento nei GIS - Workshop*, Istituto Universitario di Architettura di Venezia - Dipartimento di Urbanistica, Venezia, Ca'Tron, 25 november 1994.

A.A.VV., (1995) *Statistica e Telerilevamento* - Commissione Statistica e Telerilevamento, Quaderno n. 3, CNUCE.

Ahokas, E.; Jaakkola, J.; Stokas, P. (1990) - *Interpretability of Spot data for general mapping*, OEEPE - European organization for experimental photogrammetric research, n.24.

Atkinson, P.M., Curran P.J. (1997) - *Chosing an appropriate spatial resolution for remote sensing investigations*, in Photogrammetric Engineering & Remote Sensing, vol. 63, n.12, pp. 1345-1351.

Bianchin, A.; Pesaresi, M; (1993) - *Procedure to extract built space* - Proceedings of International Symposium Operationalization of Remote Sensing, ITC Enschede (NL) 19-23 April 1993, Vol 2, p.214-224.

Bianchin, A.; Pesaresi, M; (1993) - *Procedure to extract built space* - Proceedings of International Symposium Operationalization of Remote Sensing, ITC Enschede (NL) 19-23 April 1993, Vol 2, p.214-224.

Bianchin,A., Foramiti, S. (1999) - *Problemi di multirisoluzione nella validazione dei risultati dell'elaborazione di immagini* – Proc. ASITA 2000, p.203-208.

Congalton, R.G., (1991) – A review of assessing the accuracy of classifications of remotely sensed data, - Remote Sensing and Environment, 37, p. 35-46.

Congalton, R.G., Green, K., (1999) – Assessing the accuracy of remotely sensed data: principles and practices, Lewis Publishers.
Congalton, R.G., Plourde L., (2000) Sampling Methodology, Sample Placement, and other Important Factors in Assessing the Accuracy of Remotely Sensed Data – Accuracy 2000, p. 117-124.
Giordano, A., Vergin,H. (1994) - *Il controllo di qualità nei sistemi informativi territoriali*, Ed. Il Carro, Venezia.
Goodchild,M., Gopal, S. (1989) - *Accuracy of spatial database*, Taylor&Francis.
Goodchild, M.F., Guoquin S., Shiren Y. (1992) - *Development and test of an error model for categorical data* - IJGIS, vol. 6, n. 2, 87-104.
Heuvelink, G.B.M.(1998) - *Error propagation in environmental modelling with GIS* - Taylor&Francis.
Pesaresi, M., Kanellopulos, I. (1999) – Una nuova tecnica di segmentazione morfologica e una rete neurale per la rilevazione delle strutture urbane, *Rivista Italiana di Telerilevamento,* n°16 – Maggio 99, pp.51-58.
Pesaresi, M., Bianchin, A. (2001) - Recognizing Settlement structure using Mathematical Morphology and Image Texture -, J.P.Donnay,J.P.,.Barnsley, M.J.,Longley,P.A. eds - Remote Sensing and Urban Analysis - Taylor and Francis, London and New York. pp.55-67.
Ripley, B.D. (1981) - *Spatial statistics*, Wiley&Son, UK.
Nesset, E. (1996) - *Use of the weighted Kappa coefficient in classification error assessment of thematic maps* - IJGIS - International Journal of Geographic Information Systems, vol. 10, n.5.
H. Veregin, H. (1995) - *Developing and testing of an error propagation model for GIS overlay operations* - IJGIS, vol. 9, n. 6, 595-619.