AN INTEGRATED SYSTEM FOR SEMI-AUTOMATED SEGMENTATION OF REMOTELY SENSED IMAGERY

E. H. KOK[a] and M. TURKER[b]

a Geodetic and Geographic Information Technologies, METU, 06531 Ankara, Turkey - (emrehkok@gmail.com)
b Hacettepe University, Faculty of Engineering, Department of Geodesy and Photogrammetry, 06800-Beytepe, Ankara, Turkey - (mturker@hacettepe.edu.tr)

Commission I, WG I/V and I/VI

KEY WORDS: Field-Based, Image Segmentation, Perceptual Grouping, Classification, Agriculture

ABSTRACT:

The recent trend in image classification of agricultural areas is towards the use of per-field approaches that work in a manner that a crop label is assigned to each field. The existing field boundaries are commonly used as ancillary information for performing field-based image classification. However, the presence of multiple crops within the fields may cause problems during per-field image classification. To obtain more reliable results, within field image segmentation is needed to be performed for detecting sub-field boundaries. In this study, an integrated system was developed for performing semi-automated segmentation of remotely sensed imagery. The system works based on a field-based logic to perform image segmentation within agricultural fields in an integrated environment of Geographic Information System (GIS) and Remote Sensing (RS). The agricultural fields are selected through a database query and within field segments are detected. First, an edge detection process is performed on the image and the detected edges are vectorized to extract the straight line segments. Then, these line segments are correlated with the existing field boundaries through perceptual grouping and the closed sub-fields are formed. The sub-fields represent within field segments, each of which contains a distinct crop type. The proposed approach was implemented using a 10-m resolution SPOT5 HRV image and a 20-m resolution SPOT4 HRV image covering a part of Karacabey Plain, Turkey. The results illustrate a high degree of agreement with the reference data.

1. INTRODUCTION

Image classification is a frequently used technique for detecting agricultural crop types from remotely sensed data. Most of the current automatic classification techniques operate on per-pixel basis in isolation from other pertinent information. However, various errors are involved in the classification performed on per-pixel basis. The mixed pixels, the simplicity of the basic assumptions made for the classification algorithms, the sensor and the atmospheric effects, and the spectral overlaps between land cover types lead to wrong classification in image analysis. To overcome these drawbacks, field-based classification methods have been developed that take the segments of pixels as the main element and try to label the fields individually. These methods provide significantly accurate results when compared with the pixel-based approaches (De Wit *et al.*, 2004; Tso *et al.*, 1999, Turker and Arikan, 2005).

The field-based classification approaches generally use the spatial vector data (e.g. cadastral maps) in order to access the boundaries of the fields in which the classification will be performed. Each field described by the spatial vector data is utilized during classification and a single label is assigned to the pixels contained within the field. However, multiple crops may exist within the stable boundaries of agricultural fields.

The existence of multiple crops within the fields causes an accuracy problem for the field-based classification approaches. To overcome this, the homogeneity of the fields to be processed through a classification procedure can be satisfied using an image segmentation technique. Janssen and Molenaar (1995) discussed the terrain object hierarchies and the need for segmentation in the fields whose boundaries were defined from cadastral maps. The field, parcel, and farm district hierarchies were described with the demonstrations of crop variations within the fields. The discrepancies and disadvantages of the standard image segmentation techniques were described and the need for a knowledge base for the segmentation was clarified. The general segmentation techniques for crop mapping that operate without a priori knowledge deal with some common errors such as the existence of non-crop features (roads, rivers, ditches, etc.) in the image, the fact of inappropriate resolution for the size of terrain objects, the lack of proper theoretical framework to decide when to use which technique, etc.

However, there are some segmentation techniques that use the available semantic knowledge to refine the segmentation and to obtain more accurate results. The integration of GIS and RS and the use of vector field boundary data within the image segmentation process constitute a crucial point for detecting the boundaries of the homogenous crop fields.

This study proposes a field-based image segmentation technique to segment crop fields in an agricultural area using existing field boundary data as prior information. The segmentation process is carried out within the fields whose boundary data are available from cadastral maps. An edge based methodology was used to detect the dynamic boundaries of the homogeneous regions within the fields representing the different land cover types. The missing field boundaries are extracted from the output of the edge detected image. The detected edges are vectorized and the constructed line segments are modified in order to form the closed regions through a rule based perceptual grouping procedure. The main steps for the

proposed field-based image segmentation process are described in Figure 1.

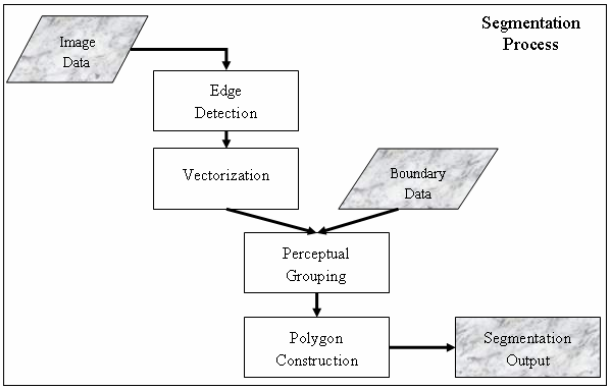


Figure 1. Segmentation Process

To implement the concept, Field-Based Image Segmentation Software (FBISS) was developed. FBISS provides the full capability of using the raster and vector data together and performing the proposed segmentation procedure within the fields and displaying the inputs and the outputs. In addition, several analysis functionalities for the evaluation of the results are provided by the software.

The proposed field-based image segmentation methodology was implemented using the 10-meter resolution SPOT5 image and the 20-meter resolution SPOT4 image covering a part of the Karacabey Plain in Turkey. The cadastral maps of Karacabey Plain were used as the existing vector boundary data of the fields. Different band combinations of the SPOT5 and SPOT4 images were used and the results were evaluated through accuracy assessment.

2. METHODOLOGY

Field-based image segmentation is the process of performing the segmentation procedure within each field separately to determine the homogenous regions by means of detecting the sub-field boundaries that exist within the field. As a first step of the procedure, the agricultural fields to be analyzed are selected one by one from the existing boundary information. The geographic locations of the vertices of the field geometries are available as a formatted text file. The input raster image is geographically referenced. Therefore, the geographic locations of the upper-left and lower-right corners of the image and the spatial resolution are known. The vertex coordinates of the field boundaries are registered to the pixel locations in the image and, for each field, the image area is extracted and processed individually.

When extracting the image patches, the small and thin fields are excluded from further processing. The shape factor (1) and the area of those fields falling below predefined threshold values are not included in the segmentation process. The shape factor is computed as;

$$SH = \frac{\sqrt{4\pi \times Area}}{Perimeter} \quad (1)$$

Figure 2 represents a couple of images extracted from the input image with the existing field boundaries overlaid.

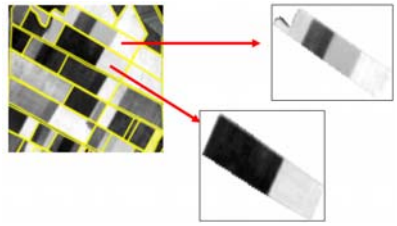


Figure 2. The Extracted Field Images.

Upon capturing the image patches, for each field, the within field segmentation procedure is carried out.

2.1 Edge Detection

Edge detection is a common image processing operation for detecting the rapid gray level variations in an image. Due to the good performance of the Canny Edge Detector with respect to other gradient operators, this technique was chosen to be used for detecting the edges. The effect of the Canny operator is determined by three parameters; (i) the width of the Gaussian mask used in the smoothing phase, (ii) the upper threshold, and (iii) the lower threshold used by the tracker.

In the present case, the lower threshold was selected to be very low and the upper threshold was selected to be rather high. The smooth transitions between different crop types would not be detected if a narrower threshold range was chosen. However, the noise is recoverable and can be cleared through further processings of the segmentation. But the undetected edges corresponding to sub-field boundaries can never be realized and recovered through further processing. As a general approach, over-segmentation is preferred to under-segmentation in the proposed segmentation procedure. The threshold values were adaptively used for the fields according to their sizes. After performing the edge detection operation, a binary image was obtained, which consists of white pixels forming the edge lines and the black pixels representing the others.

2.2 Boundary Masking

Many pixels close to existing field boundaries will also be detected as the edge pixels through the edge detection process. This is because there is a difference in the brightness between the boundary pixels and the pixels falling outside the boundaries. The pixels outside the field are set to white pixels when the distinct image patches are constructed (see Figure 2). Here, the main concern is to extract the lines which may be the candidates to form the sub-field boundaries within the fields. Therefore, the white pixels in the edge detected image, that are close to existing boundaries, were masked out. The field boundaries that already exist as the vector data and those line segments are included in the detected line segments further.

2.3 Vectorization

The vectorization is the process of detecting the coordinates of the end points of the line segments that may form the missing sub-field boundaries from the boundary masked binary image. It is basically a conversion process from raster form to vector form. There are several known methods and algorithms to perform such a conversion (Zenzo *et al.*, 1999).

In the present case, the Suzuki algorithm is used for the vectorization process (Suzuki, 1988). First, the thinning of the binary image is performed. Then, a chain graph is constructed

from the white pixels, which have another white pixel in its 8 neighbourhood. The neighbouring pixels are associated with each other and a chain graph is constructed from the raster data. All the lines that exist in the raster data can be extracted from these graphs.

The detected edges are converted to the line segments using the vectorization process and two pixel coordinates for the ending points of each line segment are calculated.

2.4 Line Simplification

The output of the vectorization process is a set of contours containing the connected line segments, each of which has a constant slope. In order to determine the closed regions within a field, it is necessary to make several associations between the line segments and the existing field boundaries, or between the line segments themselves. However, the set of line segments obtained through contour detection becomes very large and may be cumbersome to be used without any simplification. Thus, a line simplification procedure is needed to be performed for reducing the number of line segments and to obtain longer straight lines.

For this process, the Douglas–Peucker algorithm (Hershberger *et al.*, 1992), which is known to be one of the most popular methods for line simplification, is used. The line segments are simplified and grouped according to the connectivity and intersection relations between each other.

Figure 3 illustrates the process for an input image and the obtained intermediate outputs for the procedures described above.

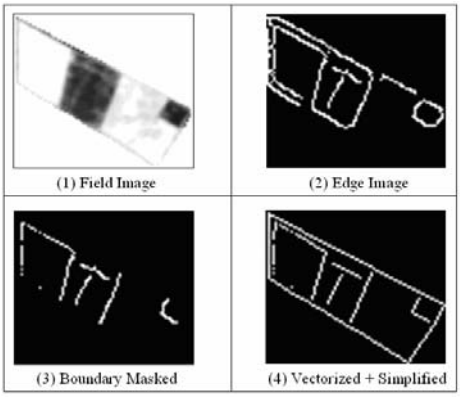


Figure 3. Intermediate Outputs for an Input Field

2.5 Perceptual Grouping of the Line Segments

The vectorized and the simplified line segments do not represent the closed segments. As can be seen in Figure 3, the noisy line segments and the unconnected sub-field boundary line segments still exist. In order to generate the closed polygons, the vertices of the line segments must be associated with the existing field boundaries or with the other line segments. To do that, a rule based perceptual grouping mechanism designed for this study is used. Basically, the process consists of two main steps;

- removing the noisy line segments, and
- modifying the vertices of the remaining line segments.

The input structure for the perceptual grouping can be described as follows:

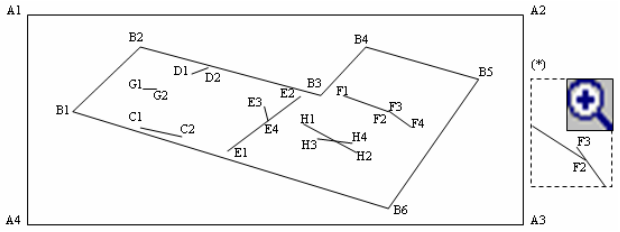
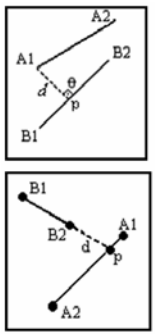


Figure 4. A Sample Field with the Detected Line Segments

The main input set consists of contours. A contour line contains a group of connected line segments. The input set for the field illustrated in Figure 4 is given below.

MS = { Contour-B, Contour-C, Contour-D, Contour-E,
Contour-F, Contour-G, Contour-H },
e.g: Contour-E = ([E1-E2], [E3-E4]) and
Contour-B is the existing field boundary (2)

The line segment pairs within a contour set and between the contour sets are analyzed and the end points of the lines are modified in order to resolve the noisy features and generate closed polygons. The distance between the end points to the lines, the slopes, the existing and possible intersections are analyzed using the pre-defined threshold values. The analysis parameters are given in Figure 5.



Analysis Parameters Between Two Line Segments:

$\|L_2\|$	Length of second line segment
$\mu(L_2)$	Slope of second line segment
$\Delta\mu$	Slope difference between line segments
$d(V_1L_1 - L_2)$	Distance from V_1L_1 to line segment L2
$\theta(V_1L_1 - L_2)$	Angle of line from V_1L_1 to line segment L2
$p(V_1L_1 - L_2)$	Closest point to V_1L_1 on line segment L2
$Int(L_1 - L_2)$	Intersection status of two line segments
$IntP(L_1 - L_2)$	Intersection point of two line segments
$IntD(L_1 - L_2)$	Minimum distance to intersection point
$IntVI(L_1 - L_2)$	Extended vertex index to satisfy intersection

Figure 5. The Analysis Parameters.

2.5.1 Rules

A sequential rule based process is carried out using the analysis parameters. The rules can be summarized as follows:

Rule 1 - Remove the overlapping and the intersecting line segments in each contour set
Rule 2 - Remove the close line segments that are in different contour sets
Rule 3 - Extend the line segments that may intersect with the existing field boundary
Rule 4 - Extend the line segments that may intersect with each other
Rule 5 - Remove the line segments that are not extended and shorter than the pre-defined threshold
Rule 6 - Modify the vertices of the line segments that have open ends by moving the vertex to intersect with the closest line segment
Rule 7 - Remove dangling arcs
Rule 8 - Remove the overlapping line segments and resolve the deviations.

Each rule is defined as an algorithmic expression. The algorithm for the first rule, which is used to remove the

overlapping and the intersecting line segments in each contour set is given below using the input structure described in equation 2.

$$
\begin{aligned}
&\forall CS_i \in MS \text{ where } BI(CS_i) = FALSE \\
&\quad \forall LS_m, LS_n \in CS_i \\
&\quad analyze(LS_m, LS_n) \\
&\quad if\ (\Delta\mu \le ConnectedLineSlopeThreshold) \\
&\qquad if\ (d(V_1L_1 - L_2) \le ConnLineDistThreshold\ and \\
&\qquad\quad d(V_2L_1 - L_2) \le ConnLineDistThreshold) \\
&\qquad\quad remove(LS_m) \\
&\qquad else\ if\ (d(V_1L_2 - L_1) \le ConnLineDistThreshold\ \text{and} \\
&\qquad\quad d(V_2L_2 - L_1) \le ConnLineDistThreshold) \\
&\qquad\quad remove(LS_n)
\end{aligned}
\tag{3}
$$

In this rule, each contour except existing field boundary (boundary indicator - $BI(CS_i)$ equals to false) in main set is processed and each line segment is analyzed against all the other line segments in a contour set. Those that are very close to a longer line segment and have similar slope with that line are removed. In these comparisons, the predefined thresholds are used. The detailed definitions and algorithmic expressions for the other rules can be found in Kök (2005). Several sample fields, in which the perceptual grouping rules are applied, can be seen in Figure 6.

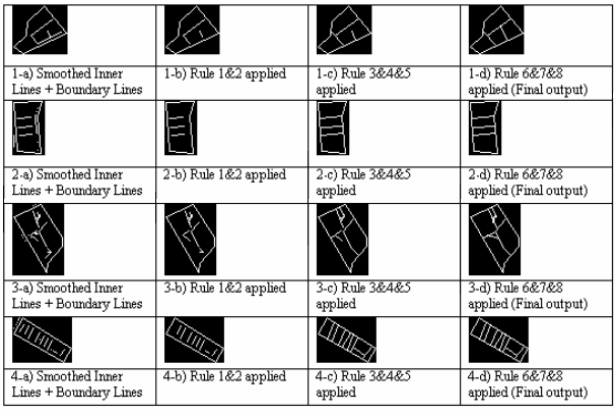


Figure 6. The Outputs of Perceptual Grouping

2.6 Polygonization

The coordinates of the vertices of each line segment and the connectivity relations between the line segments are determined. However, this information is not sufficient for determining the closed polygons within the fields. The connected line segments must be grouped together such that each group defines a disjoint polygon.

Constructing the polygons is applied by using a chain tree of the line segments through the connectivity relations and by finding the cyclic paths from a point back to itself in this tree. A cyclic path from a point to itself represents a closed region. In this tree structure, each node is a vertex of a line segment and this node has child nodes which can be directly reached from that vertex. Finding all the possible cyclic paths for a point means that all the possible polygons having that point as a vertex are constructed.

2.7 Merging

Small polygons may be generated as a result of over-segmentation. This is mainly due to the noisy lines formed through the edge detection operation. These polygons are considered to be noise in the output and may not be the distinct segments that contain different land cover types. Therefore, the small polygons are merged to the adjacent polygons. The two parameters used as hints for detecting whether the polygon is a regular segment or not are; (i) the shape and (ii) the area of the polygon.

If a polygon does not satisfy the thresholds specified for the area and the shape, it is merged with the adjacent polygons. The detected segments represent homogenous areas that have distinct crop types. Therefore, these fields must not be too small and must not have extra-ordinary shapes. Hence, small and triangular shaped fields are merged with the adjacent polygons.

Merging is the last step in the segmentation process and the final output is obtained after this procedure. Figure 7 illustrates several sample final outputs obtained after the merging step.

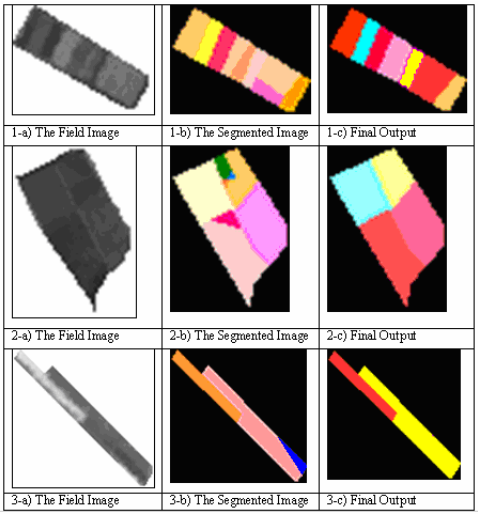


Figure 7. The Outputs of Several Fields After Merging

2.8 Software

The proposed field-based image segmentation process was automated by Field-Based Image Segmentation Software (FBISS). The software is developed using Visual C++ 6.0 and Open Computer Vision (OpenCV, Version 4 Beta) Library which is a powerful C++ library for the basic image processing operations, such as reading images from files and writing them back, edge detection, contour detection, etc.

The software implements the whole segmentation process and includes several analysis functions. The following operations can be performed using the developed software:

- Open/Save/Save As/Print Images (several formats)
- Zoom In/Out, Fit to Window, Full Screen Display
- Load Vector File (Formatted Text File)
- Determine Application and Segmentation Parameters
- Perform Segmentation
- Display the Results and Intermediate Outputs
- Comparison Between Truth Segments and Results
- Generate Reports of Results (Formatted Text File)
- Merging Segments or Parcels

3. IMPLEMENTATION

3.1 Study Area and Data

The Karacabey Plain, which consists of the agricultural fields of various sizes, was selected as the study area to implement the segmentation process. The Karacabey Plain is situated in Marmara Region of Turkey, near the city of Bursa. An area with 4600×7200 m size was selected from the Karacabey Plain. The proposed segmentation procedure is based on the integrated analysis of the raster image and existing field boundary information. Therefore, in the implementation, two input data sets were used. These are;

- raster images (SPOT4 and SPOT5), and
- existing vector boundaries (formatted text file) of 514 fields

Those fields that have the area and shape factor values falling below the predefined thresholds were not included in the processings (222 of 514 fields). Since the small fields generally contain only one crop type (210 of 222 fields), excluding them from further processing was also reasonable for better analyzing the accuracy of the segmentation process.

Since the segmentation process can be applied on single bands only, the four spectral bands (Green/Red/NIR/SWIR) of the SPOT4 and SPOT5 images were combined using two different methods and two separate single band images were derived for both image data sets.

1. First Component of Principle Component Analysis (SPOT4_PCA, SPOT5_PCA)
2. Intensity Image - (Green + Red + NIR) / 3 (SPOT4_Intensity, SPOT5_Intensity)

The within field missing boundaries between different crop types were manually delineated in a previous study conducted in the department (Özdarıcı, 2005). Therefore, the updated existing field boundary data set was used as the reference data to assess the accuracy of the applied segmentation procedure. Figure 8 displays the SPOT4 PCA image with the existing field boundaries overlaid (a) and manual segmentation outputs (b).

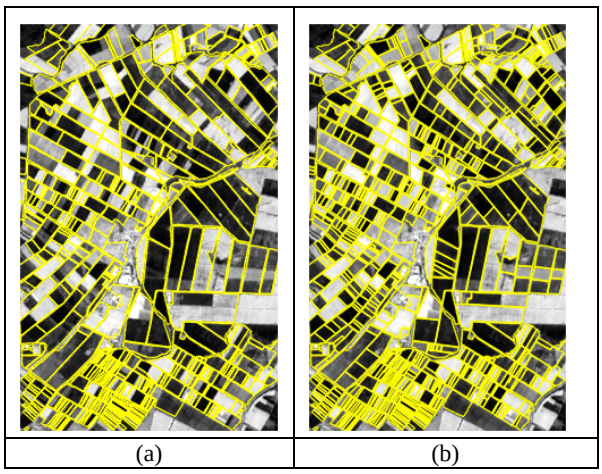


Figure 8. SPOT4 PCA Image Overlaid with (a) Existing Field Boundaries (b) Manually Delineated Boundaries

3.2 Accuracy Assessment

The accuracy assessment is based on overlaying the field geometries derived through automated segmentation process (the result segments) with the geometries of the manually segmented field geometries (the truth segments). The match between the two objects M_{ij} can be expressed as a geometrical mean of the two conditional probabilities of M_i and M_j (Janssen *et al.*, 2001).

$$\begin{aligned} M_{ij} &= \sqrt{(M_i \bullet M_j)} \\ M_i &= Area(i \cap j)/Area(i) \\ M_j &= Area(i \cap j)/Area(j) \end{aligned} \quad (4)$$

M_{ij} gets a value between 0 and 1, where 0 means no matching at all and 1 indicates a complete match. For each parcel, a mean percentage (MP) was calculated by selecting the overlapping pairs between its sub-fields (truth segments) and the obtained result segments. For each field, the mean of the obtained MP values was accepted as the assessed **overall accuracy** (Verification Parameter - VP1).

There are several other parameters to be considered for analyzing the results of the segmentation. First, a success criterion is determined by defining a threshold value (75%) for the matching percentage (Janssen *et al.*, 2001). The truth segments that have a matching percentage with the result segments higher than the predefined threshold are accepted as the successfully detected segments. The outputs for the other truth segments are considered as unsuccessful. The ratio of successfully detected truth segments to all truth segments is calculated and used as another verification parameter (VP2).

In addition, the matching percentage averages are calculated just for the successfully detected segments (VP3) and for the unsuccessfully detected segments (VP4). Finally, a quantitative analysis is applied between the results and the truth segments in terms of over-segmentation and under-segmentation.

3.3 Results and Analyses

The process was applied on each of the four single band images. Figure 9 illustrates the (a) output segments and their integration with (b) SPOT4_PCA image.

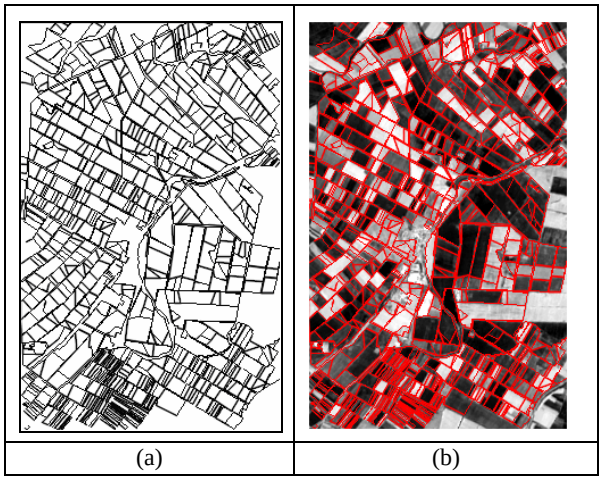


Figure 9. (a) The Detected Boundaries and (b) The Detected Boundaries Overlaid on SPOT4_PCA image

The results and the verification parameters are summarized in Table 1 and Table 2, respectively. Table 1 includes the number of fields that are over-segmented (OS), under-segmented (US) and equally segmented (ES - number of truth segment = number

of result segment for a field) over the processed 292 fields. The geometric errors (GE) for the equally segmented fields are also given in percentages in Table 1.

	US	OD	ES	GE (%)
SPOT5_PCA	52	60	180	3.5
SPOT5_Intensity	62	53	177	2.8
SPOT4_PCA	81	46	165	2.6
SPOT4_Intensity	95	43	154	2.0

Table 1. The Quantitative Results

The quantitative results indicate that neither a significant under-segmentation nor a significant over-segmentation occurs in the outputs. In the segmentation of the SPOT4 images, the under-segmented fields are found to be slightly more than those obtained for the segmentation of the SPOT5 images.

	VP1 (%)	VP2 (%)	VP3 (%)	VP4 (%)
SPOT5_PCA	**83.8**	70.6	94.6	54.8
SPOT5_Intensity	**82.6**	67.5	94.6	54.1
SPOT4_PCA	**78.8**	61.5	94.2	52.1
SPOT4_Intensity	**76.2**	57.6	93.9	49.3

Table 2. The Results Based On Geometry

As can be seen in Table 2, the overall accuracy (VP1) is the highest for SPOT5_PCA image (83.8 %). The values for VP2 which is another accuracy metric seem lower than the overall accuracy. However this parameter must be considered together with VP3 and VP4. The matching percentage averages for the successfully detected truth segments (VP3) are generally very high. This means that the successfully segmented fields have the geometric accuracy of about 95%. Also the unsuccessful segments have the geometric accuracy of about 50% (VP4), which means that those fields are not completely unsuccessful.

It is evident that a better performance was achieved for the segmentation of the SPOT 5 data despite its higher resolution when compared with the SPOT 4 data. In addition, the outputs of the PCA images that contain the spectral variability information of all bands resulted in better accuracies. The PCA images contained higher contrasts and sharper transitions between the crop fields. Therefore, it is believed that this might have caused to achieve better results for the segmentation of the PCA images.

4. CONCLUSION

The performance of the proposed segmentation technique is strongly correlated with the performance of the edge detection. Better results were obtained where the Canny edge detector provided appropriate lines after the edge detection procedure. On the other hand, unrealistic results were obtained where the output of the Canny edge detector either contained a large amount of noisy edges or did not contain the proper edges, which might form the missing boundaries.

As a general performance evaluation, the accuracy of the segmentation process was computed to be $80 \pm 5\%$ for each of the SPOT4 and SPOT 5 images. The results seem to be quite promising. The erroneously detected edges within the fields and the erroneously modified line segments through perceptual grouping can be said to be the main reasons for the over-segmentation. The missing lines that were not able to be detected by the Canny edge detector, the erroneously deleted line segments through the perceptual grouping rules, and the erroneously merged sub-fields were the main reasons for the under-segmentation. In addition, the conversion of multi-spectral satellite images to a single band image means loss of some information. It is believed that the loss might have had significant affects on the accuracy of the segmentation.

The proposed segmentation procedure can be further enhanced in order to increase the accuracy. The Canny edge detector can be applied on single band images only. So, the spectral bands are needed to be reduced to be a single band. However, the use of a multi-spectral edge detection method might be better to detect the edges effectively. In addition, the improvements may be necessary for the perceptual grouping rules.

It can be stated that including the topographical maps in the segmentation process and performing the segmentation in a field-based manner seems to be the most promising way for detecting the subfield boundaries. The authors believe that the proposed segmentation strategy is a starting point for the development of high performance field based image analysis operation that includes both the segmentation and classification procedures.

REFERENCES

De Wit, A. J. W. and Clevers, J. G. P. W. 2004. “Efficiency and Accuracy of Per-Field Classification for Operational Crop Mapping”, International Journal of Remote Sensing, 25(20).

Hershberger, J. and Snoeyink, J. 1992. “Speeding Up the Douglas-Peucker Line-Simplification Algorithm”, Technical Report, In Proc. 5th Intl. Symp. Spatial Data Handling.

Janssen, L. L. F. and Molenaar M. 1995. “Terrain Objects, Their Dynamics, and Their Monitoring by The Integration of GIS and Remote Sensing”, IEEE Transactions On Geoscience and Remote Sensing, 33(3): 749- 758.

Kök, E. H. 2005. “Developing An Integrated System For Semi-Automated Segmentation of Remotely Sensed Imagery”, M.Sc. Thesis, GGIT, METU.

Özdarıcı, A. 2005. “Accuracy Assessment of the Field-Based Classification Using the Different Spatial Resolution Images”, M.Sc. Thesis, GGIT, METU.

Suzuki, S. 1988. “Graph-based Vectorization Method for Line Patterns”, Computer Vision and Pattern Recognition, Computer Society Conference, 616-621.

Tso, B., and Mather, P. M. 1999. “Crop Discrimination Using Multi-temporal SAR Imagery”, International Journal of Remote Sensing, 20(12):2443-2460.

Turker, M. and Arikan M. 2005. “Sequential masking classification of multi-temporal Landsat7 ETM+ images for field-based crop mapping in Karacabey, Turkey” International Journal of Remote Sensing, 26 (17):3813-3830.

Zenzo, S. D., Cinque, L. and Levialdi S. 1996. “Run-Based Algorithms for Binary Image Analysis and Processing”. IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(1):83-89.