Finding Thresholds for Image Segmentation

Theo E. Schouten, Maurice S. klein Gebbinck,
Ronald P.H.M. Schoenmakers†, Graeme G. Wilkinson‡

Katholieke Universiteit Nijmegen,
Informatica,
Toernooiveld 1, 6525 ED Nijmegen,
The Netherlands
Email ths@cs.kun.nl
tel +31-80-653175, fax +31-80-553450
†Telecom Italia, Divisione Spazio, via Tiburtina 965, 00156 Roma, Italy
Currently attached to:
Joint Research Centre,
Institute for Remote Sensing Applications,
21020 Ispra, (VA), Italy
Email: ronald.schoenmakers@cen.jrc.it
tel +39-332-785260, fax +39-332-789074
‡Joint Research Centre,
Institute for Remote Sensing Applications,
21020 Ispra, (VA), Italy
Email: graeme.wilkinson@cen.jrc.it
tel +39-332-785103, fax +39-332-789469

ABSTRACT

Segmentation methods for images often have cost functions which evaluate the (dis)similarity between pixels or segments. Thresholds on cost values are then used to decide whether or not to grow, join or split segments. The results for a given image critically depend on the selection of the threshold values. In remote sensing, a too low threshold will split up regions of constant ground cover and a too high threshold will join adjacent regions of different ground cover.

Optimal thresholds can be determined using different classes of methods:

- generating cost value distributions from the original image.
- obtaining statistical distributions from segmented images.
- comparing a "true" segmentation with the results of segmentation using a range of thresholds.

A so-called "true" segmentation can be derived from human expert segmentations or from maps obtained by ground surveys or segmentation of higher resolution images.

Several methods for threshold determination are described for a hybrid segmentation method developed by us. Measures are described for comparison of two segmentations. Results are evaluated using several real and artificially generated images.

1 Introduction

Segmentation methods for images often use cost functions which evaluate the (dis)similarity between pixels or segments. A cost function can be based on statistical measures of the intensity value distributions of the regions, on the size of the regions, on their shapes and boundaries, etc. Thresholds on cost values provided by the functions are then used to decide whether or not to grow, join or split segments. The segmentation quality for a given image critically depends on the selection of the threshold values. In remote sensing, a too low threshold (requiring that regions must be very similar in order to be merged) will split up regions of constant ground cover. A too high threshold will join adjacent regions of different ground cover. Intermediate threshold values influence the handling of mixed pixels: pixels that receive reflective radiance from different types of ground cover. This can be from two different fields or fields and small objects that separate fields like roads, waterways, hedges or bare ground. In this paper various methods for determining an optimal threshold setting for a certain segmentation method are considered.

In section 2 the segmentation method adopted is described. Also some simple quality measures to compare two segmentations of a given image are described. In section 3 the real images and the artificially generated image used for testing the threshold determination are described. The way of generating the artificial image is given as well as the way of obtaining a perfect or true segmentation of it. In the following three sections methods are described and evaluated on the images to obtain threshold values from cost value distributions, from distributions of the number of regions produced by segmentation and from the quality measures.

In the last section (7) the threshold determination methods are evaluated and areas of further results are indicated.

2 Segmentation Method

The segmentation method adopted is based on the optimized best merge approach as originally proposed by Beaulieu [1] and Tilton [10] and further developed by Schoenmakers [6, 7, 8]. Initially each pixel is considered as a separate region. As a first step neighboring regions having the same intensity value in all the channels are merged. Then all pairs of neighboring regions are considered and have a merging cost calculated. The two neighboring regions having the lowest merging cost are then merged into a new region. When there is more than one group of regions with minimum cost value, they are merged in parallel. This merging of regions with the lowest merging cost is repeated until the lowest cost value is above a set threshold value. The cost function can be based on statistical measures of the intensity value distributions of the regions, on the size of the regions, on their shapes and boundaries, etc. In this paper the function proposed by Schoenmakers for the cost between regions i and j is used:

$$\sqrt{\frac{\sum_{c=1}^{n}(v_c^i - v_c^j)^2}{n}}$$

with:

- n the number of channels
- v_c^i the intensity value of channel c of region i (the average value of all the pixels in the region)

This cost function only depends on the average intensity per region, not on the size of the regions.

The best merge method becomes too slow for large images because of the exponentially growing number of neighboring region pairs to consider. A hybrid segmentation method was therefore developed combining edge detection and best merge region growing. Edges are detected, thinned and followed (if needed with a jump over a single pixel) to divide a large image into a number of closed polygons. The borders of the polygons consist of pixels with a large edge value, which should not be crossed by a region growing process. The best merge method can therefore be applied separately on the inner pixels of each closed polygon. In an additional step [9] the polygon pixels can then be merged with neighboring regions or with other polygon pixels. For this merging the same cost function as for the best merge step can be used.

This hybrid segmentation method also eliminates the disadvantages of segmentation based on edge detection or region growing alone:

- segments are not allowed to over-grow clear edges.
- incomplete boundaries as detected by edge detection methods are completed by region growing.

In remote sensing, fields are often separated from each other by small objects like roads, waterways, hedges or bare ground. The width of those objects are usually smaller than or comparable to the width of a pixel, which depends on the resolution of the images (10 m for SPOT-P, 20 m for SPOT-XS, 30 m for LANDSAT-TM). Region growing tends to merge fields separated by such objects if the fields have comparable intensities, because there are many boundary pixels with intensities which differ only a little from the field intensities. Edge detection methods often do not fully find such boundaries because of the many mixed pixels in them. The edge linking methods fill some of those holes and make it more difficult for the best merge to join the fields. Currently edge pixels not being part of a polygon border take part in region growing but receive no additional processing. This could be changed in a next version of the region growing program, to further reduce over-growing of field separators.

Although quality measures to evaluate segmentation results are sometimes used [2, 3, 4, 5], there is no generally accepted standard. Liu and Yang [4] use the following evaluation function to measure the badness of a segmentation:

$$F = \sqrt{R} \times \sum_{i=1}^{R} \frac{e_i^2}{\sqrt{A_i}}$$

with:

- R the number of regions in the segmented image
- e_i the intensity error of region i: the sum of the Euclidean distance of the intensity vectors between the original image and the segmented image of each pixel in the region.
- A_i the number of pixels in region i.

F is normalized by the size of the image and by some factor. The lower the value of F, the better is the segmentation result.

Because the true segmentation of the artificial image can be easily generated, some simple quality measures were developed and used to compare two segmentations (called the true and test segmentations) of a single image. For each field i of the true segmentation the number of inner pixels nin_i is determined, thus leaving out the pixels having a border pixel as 4-connected neighbor. Then the largest number $ncov_i$ of those inner pixels which belong to a single segment in the test segmentation is determined. The split quality is calculated as:

$$qsplit = \frac{\sum_{i=1}^{n} ncov_i}{\sum_{i=1}^{n} nin_i}$$

The split quality is 1 if true fields are not split into segments in the test segmentation and goes to 0 if they are heavily split. Then in the above the roles of the true and test segmentations are reversed to obtain the *qmerge* quality value. This merge quality is 1 if true fields are not merged together in the test segmentation and goes to 0 if all of them are merged with at least one other true field. The average of the two quality values can then be taken as the overall quality of the segmentation. Note that mixed pixels are not taken into account in this simple quality measures, one of the reasons being that not much information about them is available.

The current implementation of this hybrid segmentation method is heavily optimized for speed and is able to segment full size (6000 by 6000 pixels, 6 channels) LANDSAT images on the current generation of workstations in a reasonable amount of time (estimated 10 hours on a SUN10/30). To be able to handle these large images, only a small part of images plus intermediate and region information is kept in memory at any one time. A parallel implementation on a set of workstations is under development which will distribute the polygons over the workstations for the best merge and will also parallelize as much as possible the other parts of the processing chain.

3 Images

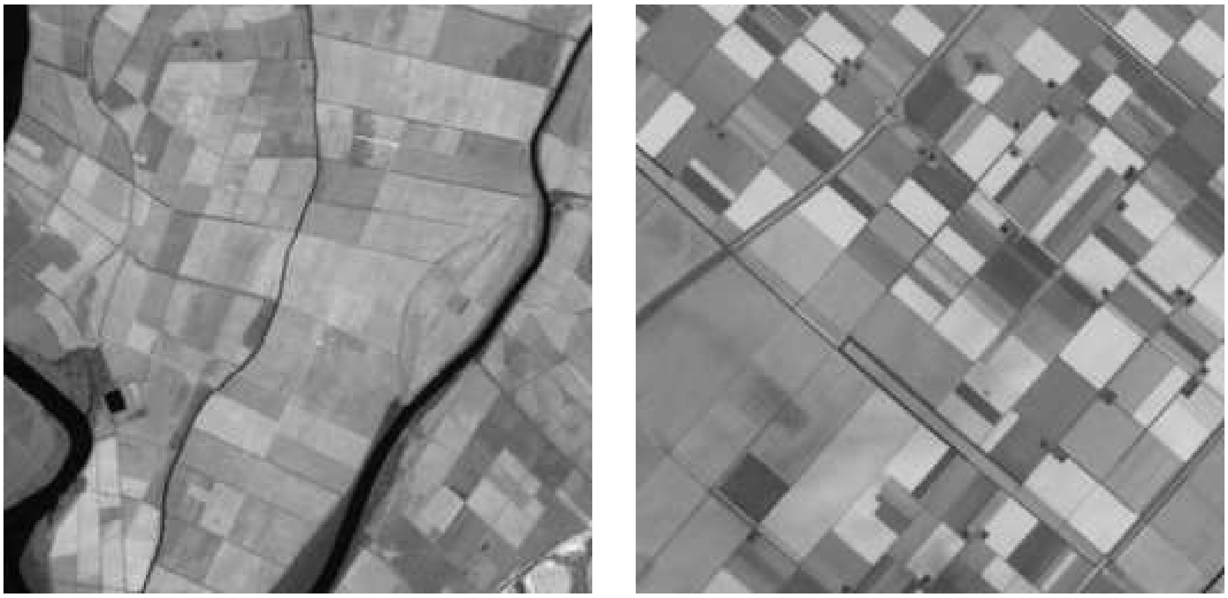
For this paper several 256 by 256 pixel images have been used for testing the developed threshold determination methods:

port part of a 6 channel LANDSAT-TM image of Portugal (see figure 1(a)).

neth part of a 3 channel SPOT-XS image of an agricultural area in the Netherlands. (see figure 1(b)).

gree part of a 1 channel SPOT-P image of Greece (see figure 2(a)).

veld an artificial generated 1 channel image with 68 fields (see figure 2(b)).



(a) Channel 4 of the *port* image (©Eurimage) (b) Channel 2 of the *neth* image (©Spotimage)

Figure 1: Images *port* and *neth*

The artificial image was generated [11] in the following way. Using a drawing program (Xfig was used), field boundaries were drawn and a X11-bitmap image was generated. For each field the average intensity value and the standard deviation (σ) for each channel were stored in a readable file. A program than generated a satellite image (ERDAS format) on a smaller scale as the X11-bitmap image, thus generating mixed pixels which received radiance from different fields. Also an information file was generated, describing each pixel in the large image. Thus for each pixel of the generated satellite image it was known which fields contributed to it. This file together with the file describing the segmentation results on the generated image can then be used by another program to generate performance measures of the segmentation program and its parameters. These performance measures are still under development and not used in this paper, only the quality measures described in the previous section were used.

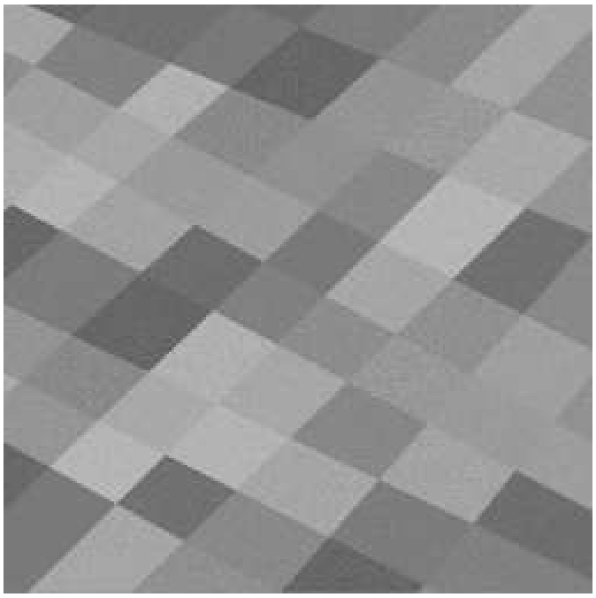
The image *veld* was generated using 68 fields and 1 channel. The average intensity values and their $\sigma's$ were set according to values obtained from fields in channel 4 of the *port* image. The $\sigma's$ were set to 3 for 45 and to 4 for 23 fields. No small field separators, like roads, waterways, etc., were drawn.

A so called "true" segmentation of the *veld* image was generated in the following way. An image was generated using the same X11-bitmap image as for *veld* but another field characteristics file with the $\sigma's$ set to 0 and the intensity values of neighboring fields different by at least 5. This image was then region growned using a low threshold value of 2.0, all the pixels fully covered by a certain field were thus merged into one segment. The mixed pixels were either merged with a large "true" region or form small regions consisting only of mixed pixels. As the used quality measures (see section 2) don't consider mixed pixels, the segmentation of the mixed pixels is not important.

4 Threshold from cost values

The used cost function depends only on the intensity values of pixels and of regions, the intensity value of a region is just the average intensity of all the pixels belonging to that region. Therefore it is possible to estimate from an image the cost value distribution inside fields of constant ground cover and between such fields, without knowing a true segmentation.

(a) *gree* image (©Spotimage)



(b) *veld* image

Figure 2: Images *gree* and *veld*

A pixel has a large chance to belong to the same ground cover as its 8 neighbors (spatial redundancy property). Some constraints can be imposed on the intensity distribution of the 9 pixels to decrease the probability that pixels covering different ground covers are included. These constraints should not be too severe in order not to bias the to be obtained cost value distributions. The chosen constraints are that in each channel the distribution must not be too skewed:

$$(median - min) < (max - median) * 4$$
$$(max - median) < (median - min) * 4$$

where *median* is the median, *min* the minimum and *max* the maximum value of the 9 intensity values. It was checked that the final cost value distributions are rather independent of the selected factor in this skewness test. For each pixel which fulfills the above constraints, the costs between it and its 8 neighbors are calculated. If one of the pixels does not (fully) belong to the same field as the central pixel, the corresponding cost value will be usually higher. The peak of the distribution of the 8 cost values is thus expected to be closer to the minimum than to the maximum value. This leads to the additional requirement that the median value (the average of the 4th and 5th value) should be closer to the minimum than to the maximum value. If this is the case, the cost values between the minimum and median + (median - minimum) values are taken as cost values inside a field.

Then randomly a 3 by 3 area in the image is selected, of which the intensity value distribution must also be not too skewed. The cost value between the two area's using the median intensity values, is taken as a cost value between fields. This random selection of an area is repeated several times in order to increase the statistics.

The inside and outside distributions are then scaled to contain the same number of entries and shown in figures 3 and 4 for the 4 images. Note that due to having only one channel and the way the cost values were calculated, the distributions of the *veld* and *gree* images have only integer cost values.

The inside field cost distributions show a peak located at a value of 1.5, 1.0, 2.0 and 1.5 for the *port*, *neth*, *gree* and *veld* images. The distributions are higher and show a tail on the high cost side, caused by mixed pixels and pixels of other fields. Taking the reasonable assumption that the real inside field cost distributions are symmetric, they end at twice the peak value. That cost value can then be regarded as the value where fields are not split due to internal variations in intensity values.

The between field cost distributions are slowly rising and then decreasing (except for the *veld image*, which however has a low number of fields) as could be expected and not much information can be extracted from them.

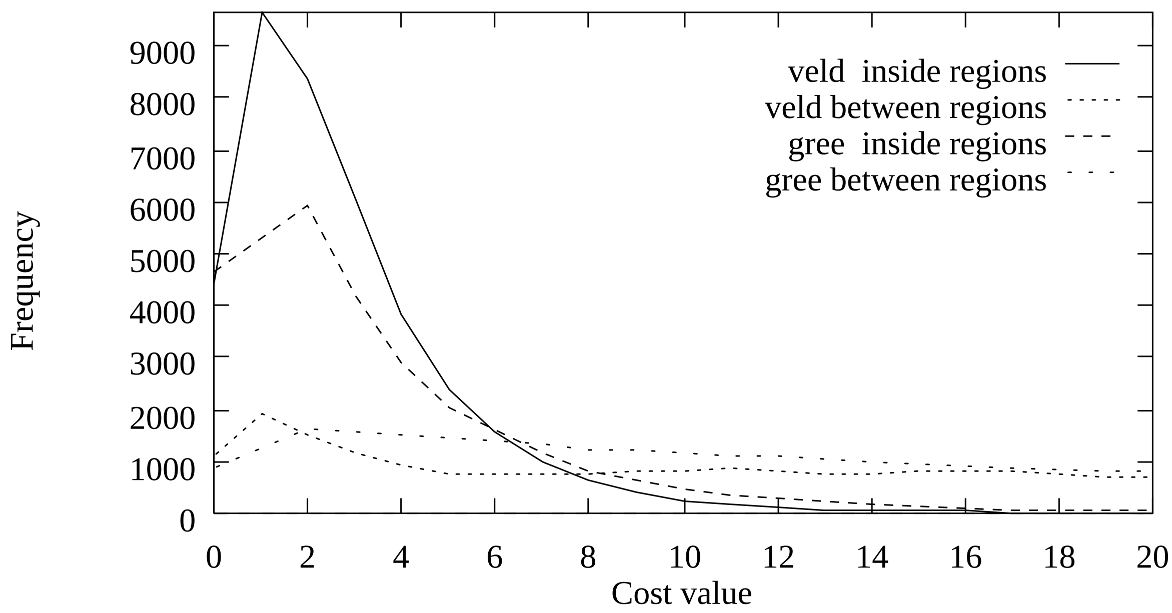


Figure 3: Cost distribution of *veld* and *gree* images

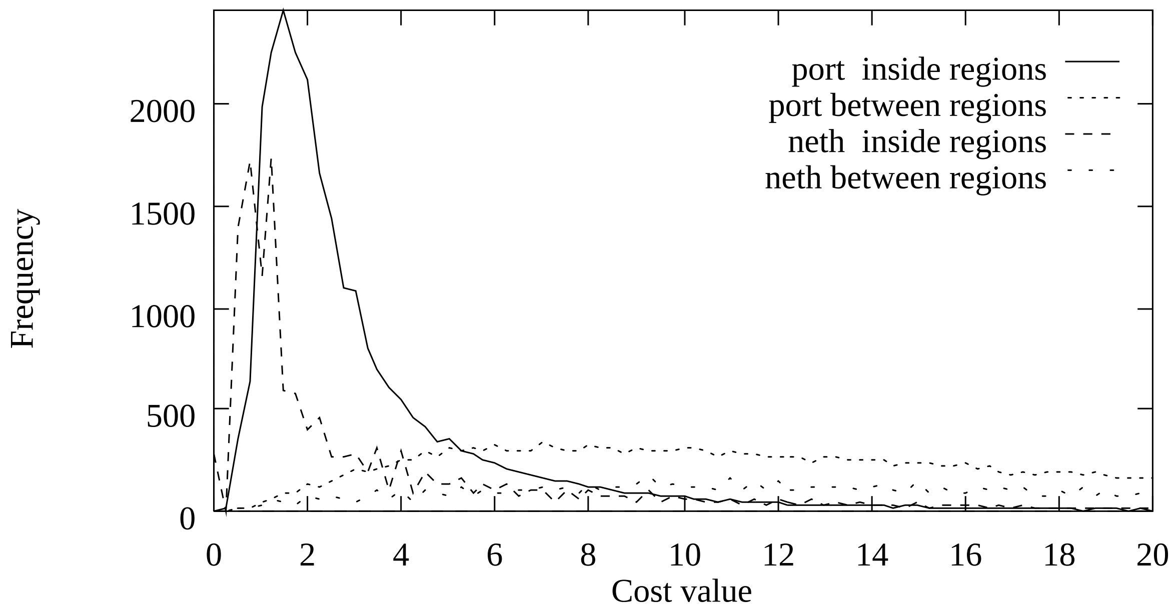


Figure 4: Cost distribution of *port* and *neth* images

At the crossing point of the inside and between field distributions there is an equal amount of neighboring mixed or different field pixels and further away pixels with the same cost function. Above this point the merging of different fields probably can start. These cross-overs occur at cost values of 5.0, 6.5, 6.5 and 7.5 for the *port*, *neth*, *gree* and *veld* images.

5 Threshold from the number of regions

After segmentation the number of regions found can be shown as a function of the used threshold. To study the formation of regions one can leave out regions consisting of only one pixel. Taking only regions having inner pixels reduces the influence of mixed pixels and small boundary objects. To be able to study the over-growing of regions in this and the following chapter the best merge region growing was applied without using polygons produced by edge-detection. Using polygons limits the over-growing to fields which happen to be in the same polygon as regions in different polygons can not be merged and thus makes determining the threshold where over-growing starts more difficult. In figures 5 and 6 the distributions of the number of regions are given.

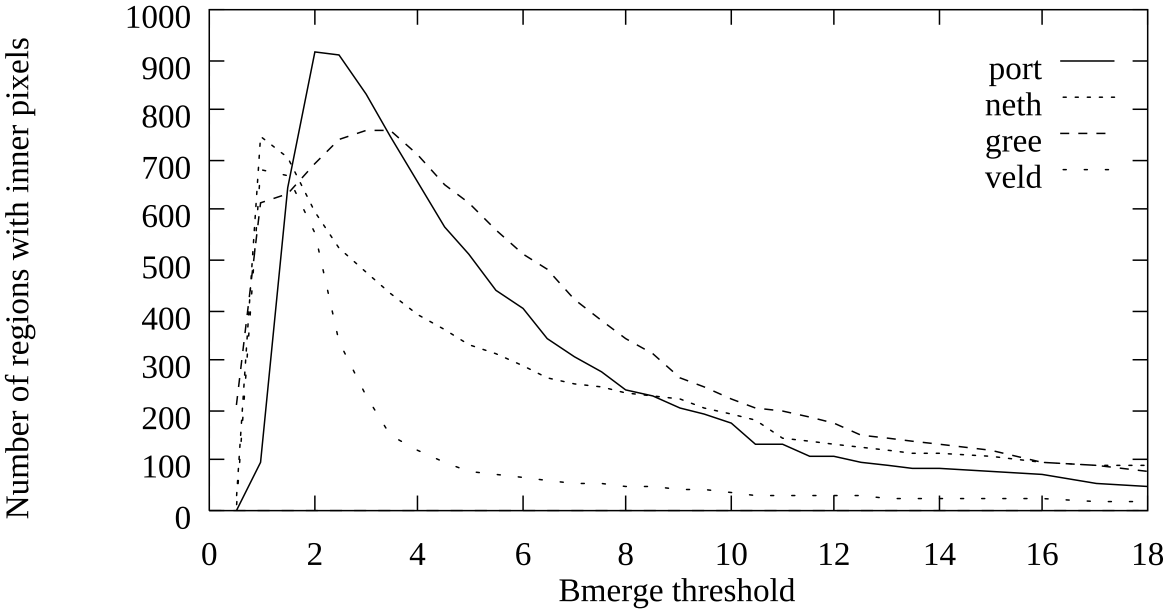


Figure 5: Number of regions with inner pixels

The distributions of the number of regions with inner pixels peak at values of 2.25, 1.75, 3.5 and 1.5 for the *port*, *neth*, *gree* and *veld* images. These are a bit below twice the peak values found in the inside field cost distributions. Pixels of the true fields are there merged into one or a few regions having inner pixels and a lot of smaller regions. For the *veld* image the peak is at the peak of the inside cost distribution, this is due to the fact that all fields are large and there are some 10 larger regions per field. Thus cost values a factor 1 to 2 (depending on the size of the fields) higher than the peak value can be regarded as the threshold value where pixels of a field are merged into one region.

The distributions of the number of regions with more than 1 pixel often do not show a real peak at low threshold values. There most pixels are already merged with one or a few neighboring pixels at the lowest threshold values.

After the peak, the distributions decrease, this decrease becomes slower with increasing threshold. With increasing threshold first regions in the same field are added, then mixed pixels and boundary pixels are included and finally different regions are merged. The decrease becomes slower because the costs inside a field are smaller than the costs between fields and their mixed and boundary pixels, which in turn are smaller than the costs

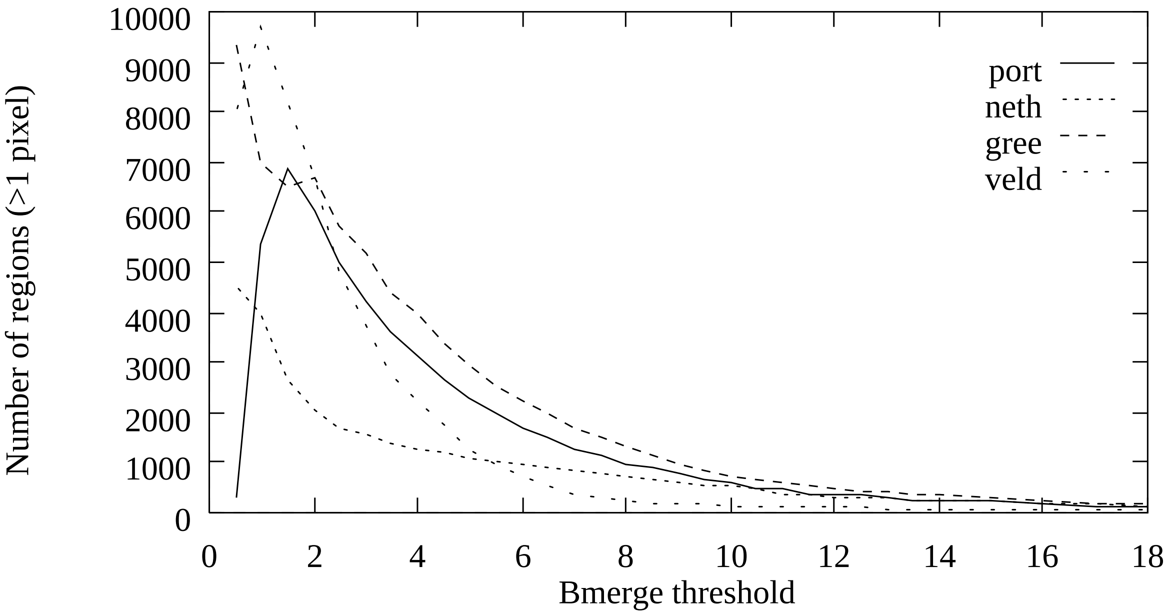


Figure 6: Number of regions with more than 1 pixel

between different fields. The inside and the between field cost distributions (figures 3 and 4) decrease more or less nearly exponential, as can be seen when plotting them on a logarithmic scale. The decrease in the number of regions is therefore expected to be also exponential in the threshold ranges where regions inside a field are merged and where different fields are merged. For the threshold range in between, where mixed and boundary pixels are included, a gradual change in exponent should occur.

On the number of regions distributions with a logarithmic scale (figures 7 and 8), one remarks that the decreases in the number of regions show more or less the expected behavior.

Drawing by eye the two exponential decreases (linear on the logarithmic scale) one can estimate the threshold value where the two slopes intersect and the value where the decrease starts deviating from the first slope. The deviation points indicate the points where mixed pixels start being included but they are rather difficult to determine from the plots. The intersections are more easily determined and denote the points where merging of different fields becomes a problem. For the four images the so determined threshold values are given in table 1.

value	*port*	*neth*	*gree*	*veld*
inner cross	10	8	10	4.5
inner deviate	8	6	9.5	4
>1 cross	10	8	10	7.5
>1 deviate	7	4.5	7	5

Table 1: Threshold values from number of regions plots

Note that around a cost value of 5 the number of regions with inner pixels of the *veld* image comes near the 68 fields in the image and the best segmentation threshold for that image is thus in the neighborhood of 5. The determined values for the *veld* image from the number of regions with more than 1 pixels are thus too high.

For the *veld* image the intersection point from the number of regions with inner pixels is at the expected position (see also section 6). The corresponding values for the real images are higher than the thresholds determined for them in section 4. This can be explained by the fact that the best merge must grow over the mixed pixels and

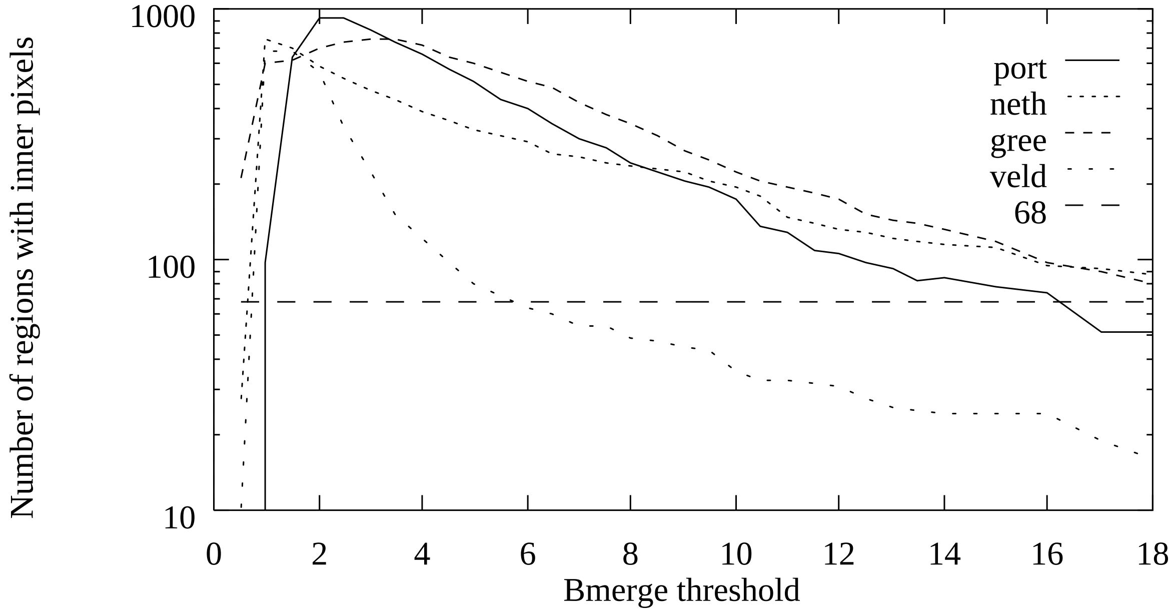


Figure 7: Number of regions with inner pixels, log scale

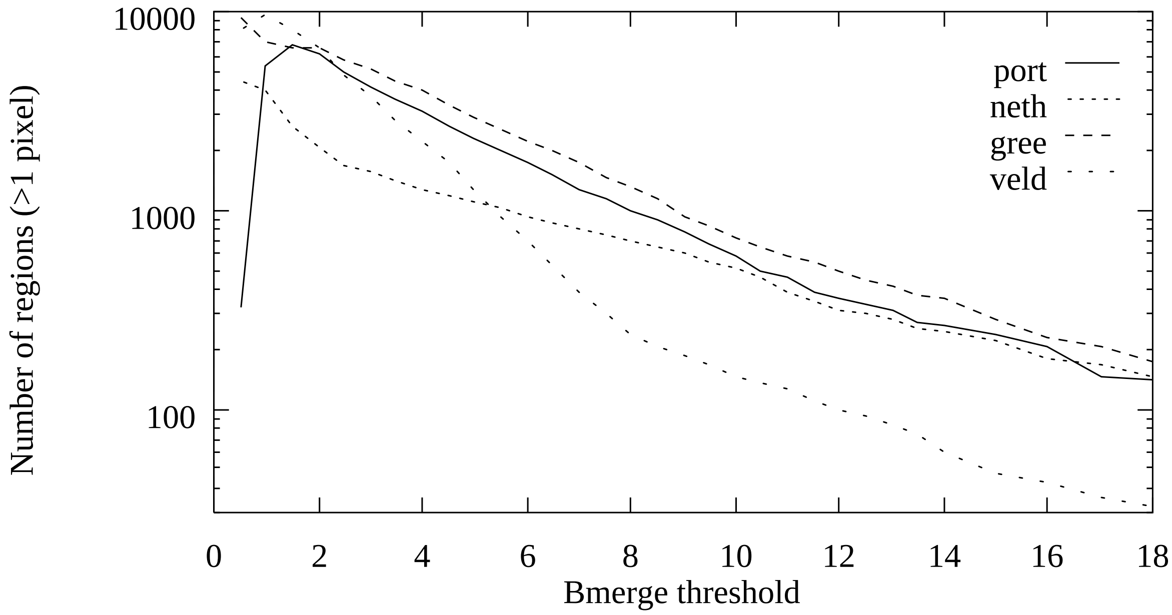


Figure 8: Number of regions with more than 1 pixel, log scale

small boundary objects to merge different fields. In section 4 only the cost between the fields was taken into account. The intersection point of the two slopes of the distribution of the number of regions with inner pixels can thus be taken as the point where merging of different fields has become a problem.

The deviation points from the number of regions with inner pixels look too high. They probably indicate the threshold where mixed pixels caused by the small boundary objects start to be included.

6 Threshold from quality measures

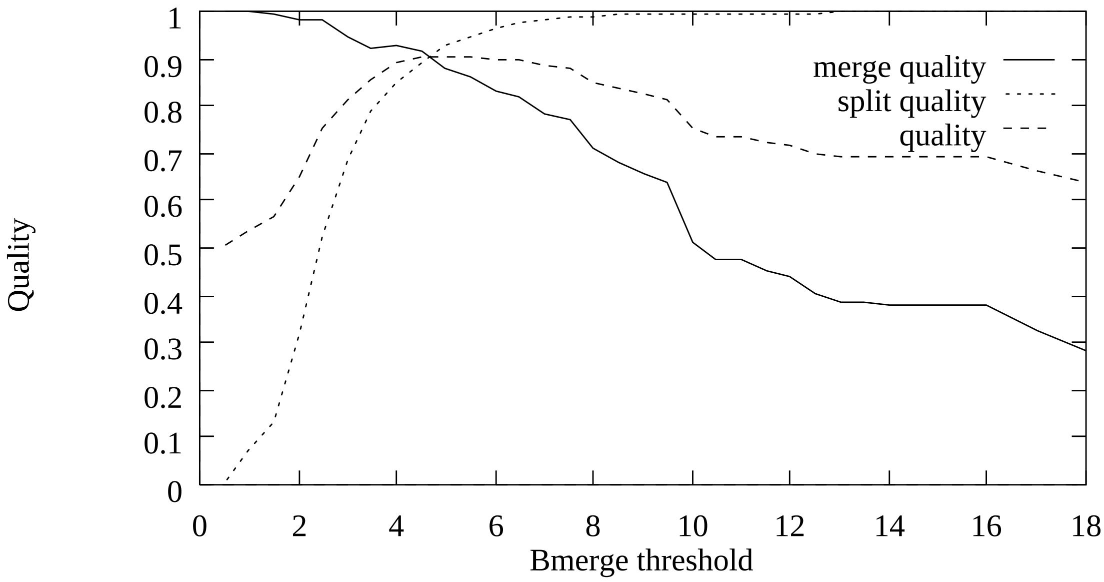


Figure 9: Segmentation quality for the *veld* image

For the artificial image a true segmentation was generated (see section 3) where all pixels fully covered by a field are correctly segmented. The quality measures (see section 2) were then calculated and shown in figure 9.

As expected the split and merge quality increases resp. decreases, they intersect at a cost value of 4.6. The total quality (the average of the split and merge quality) shows a peak with a flat top between cost values 4.5 and 5.5. Thus in this range the optimal threshold for segmentation of the artificial *veld* image can be found. Note that the mixed pixels are not considered in these quality measures.

The results of the evaluation function F of Liu and Yang (see section 2) are shown in figure 10. No minimum is observed at the place where is should be, e.g. around 5 for the *veld* image. Also several variations on F were tried (using the number of regions with more than 1 pixel, dividing by A_i), they show the same negative results.

For the real images a true segmentation is not available. Therefore in figure 11 the split quality is shown taking the segmentation at a threshold of 16 as the reference or true image. After a first rise a region with slower rise (with exception of the *port* image) and then a faster increase towards the end can be observed. The middle region is in the range 2-10.5, 1-7 and 2.5-9.5 for the *neth*, *gree* and *veld* images. For the *port* image the final rise occurs at a value of 10. The slope difference between the middle and last region is rather small for the *veld* image, caused by the missing of boundary pixels and the small number of mixed pixels. When the final rise occurs there are already many fields merged and the threshold at that point is too high.

7 Discussion

Methods were developed to determine thresholds for a segmentation program by considering cost value distributions and distributions of the number of regions as function of the threshold. They were applied to LANDSAT,

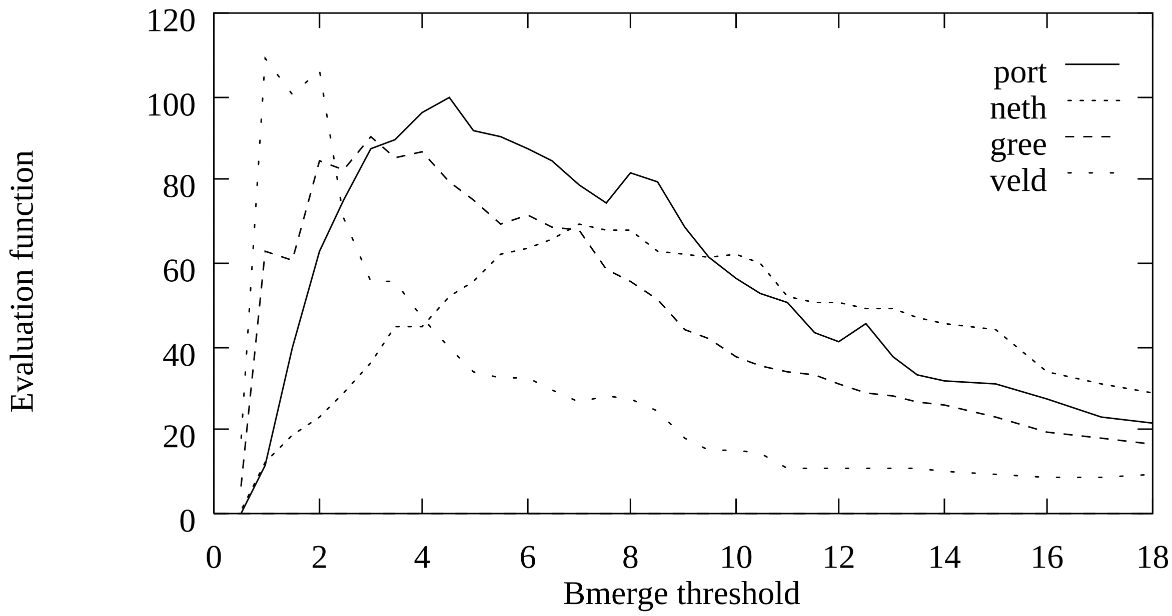


Figure 10: Evaluation function F from Liu and Yang

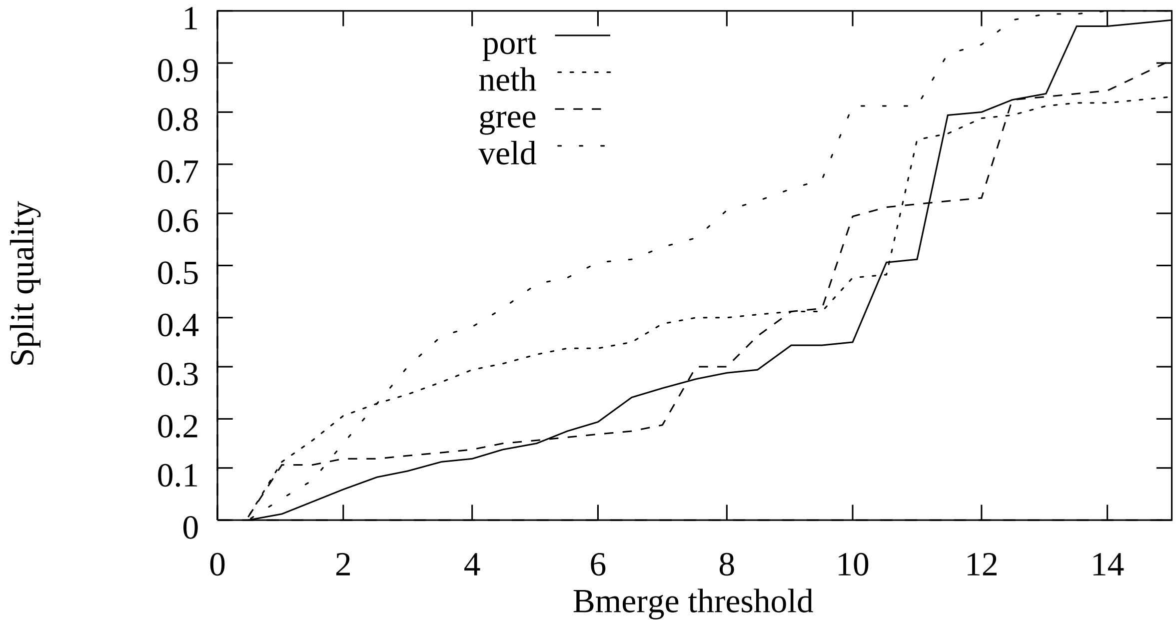


Figure 11: Split quality taking threshold 16 as a reference

SPOT and artificially generated images.

The cost value distributions give indications of thresholds where fields are not split due to internal variations in intensity values and of thresholds where merging of different fields can occur. This method can be applied when the used cost function depends only on intensity values of pixels and regions.

The number of region distributions provide threshold values where the merging of different regions (over-growing) starts. Also thresholds, where fields are not split due to internal variations in intensity values, are indicated up to a factor 1 to 2 which depends on the average size of the fields.

Quantitative measures were developed to measure the quality of a segmentation with respect to a true segmentation. Mixed pixels (receiving intensity distributions from different ground cover) are not taken into account in these measures. Distributions of those measures as function of the segmentation threshold provide an optimal threshold range, as shown for the artificially generated image. Optimal regards here the unwanted splitting or merging of fields, not how the mixed pixels or pixels from small boundary objects (roads, waterways, tree's, etc.) are handled. Using one of the quality measures taking segmentations with a very high threshold as the "true" segmentation, gives indications of thresholds where over-growing is already a problem thus where the thresholds are already too high.

The evaluation function F as used by Liu and Yang[4] was shown not to produce an useful indication of a threshold value.

Further research will be directed towards incorporating mixed pixels into the quality measures. For real images human expert segmentation will be obtained assigning pixels either to a field or to a mixed pixel category. More realistic artificial images with small field separating objects will be generated where the true segmentation is known to sub-pixel level[11]. Also segmentation quality measures using only a single segmented image[4, 3] will be furthur examined.

References

[1] J-M. Beaulieu and M. Goldberg. Hierarchy in Picture Segmentation: A Stepwise Optimization Approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 11(2):150-163, 1989

[2] Sang Uk Lee, Seok Yoon Chung and Rae Hong Park. A Comparative Performance Study of Several Global Thresholding Techniques for Segmentation. *CVGIP 52:171-190, 1990*

[3] M.D. Levine and A.M. Nazif. Dynamic measurement of computer generated image segmentations. *IEEE PAMI 7:155-164, 1985*

[4] J. Liu and Y-H Yang. Multiresolution Color Image Segmentation. *IEEE PAMI 16:689-700, 1994*

[5] W.A. Yasnoff, J.K. Mui and J.W. Bacus. Error measures for scene segmentation. *Pattern Recognition 9:217-231,1977*

[6] R.P.H.M Schoenmakers. Segmentation of Remotely Sensed Imagery, draft version 2.0 *PhD thesis* Katholieke Universiteit Nijmegen, 1994

[7] R.P.H.M Schoenmakers, G.G. Wilkinson and Th. E. Schouten. Segmentation of Remotely-Sensed Imagery: A Re-Definition for Operational Applications. *IGARSS91* pages 1087-1090, Espoo, Finland, 1991

[8] R.P.H.M Schoenmakers, G.G. Wilkinson and Th. E. Schouten. Results of a Hybrid Segmentation Method *This conference* Rome 1994

[9] Th. E. Schouten. Adding Edges to Regions *Internal Note* Katholieke Universiteit Nijmegen, 1994

[10] J.C. Tilton. Image Segmentation by Iterative Parallel Region Growing and Splitting. *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS89)*, pages 2235-2238, Vancouver, 1989

[11] H.F.J.M. Voogd. Error measurement for segmentation techniques *Master thesis* Katholieke Universiteit Nijmegen, 1994