Evaluation of a segmentation algorithm designed for an FPGA implementation

Kurt Schwenk[a], Maria Schönermark[a] and Felix Huber[a]

[a]German Aerospace Center (DLR), German Space Operations Center (GSOC), Münchner Str. 20, 82234 Wessling, Germany

ABSTRACT

The present work has to be seen in the context of real-time on-board image evaluation of optical satellite data. With on board image evaluation more useful data can be acquired, the time to get requested information can be decreased and new real-time applications are possible. Because of the relative high processing power in comparison to the low power consumption, Field Programmable Gate Array (FPGA) technology has been chosen as an adequate hardware platform for image processing tasks. One fundamental part for image evaluation is image segmentation. It is a basic tool to extract spatial image information which is very important for many applications such as object detection. Therefore a special segmentation algorithm using the advantages of FPGA technology has been developed. The aim of this work is the evaluation of this algorithm. Segmentation evaluation is a difficult task. The most common way for evaluating the performance of a segmentation method is still subjective evaluation, in which human experts determine the quality of a segmentation. This way is not in compliance with our needs. The evaluation process has to provide a reasonable quality assessment, should be objective, easy to interpret and simple to execute. To reach these requirements a so called Segmentation Accuracy Equality norm (SA EQ) was created, which compares the difference of two segmentation results. It can be shown that this norm is capable as a first quality measurement. Due to its objectivity and simplicity the algorithm has been tested on a specially chosen synthetic test model. In this work the most important results of the quality assessment will be presented.

Keywords: Image segmentation, Segmentation evaluation, FPGA, on-board image processing

1. INTRODUCTION

On fundamental task in image processing is image segmentation. An image segmentation is the partition of an image in different non-overlapping regions, based on heterogeneity and homogeneity criteria.[1] It is a basic tool to extract spatial image information which is very important for many applications such as object detection. For an on-board processing framework of optical satellite data, an image segmentation method was developed.[2] The on-board processing framework is based on Field Programmable Gate Array (FPGA) technology. This technology was chosen because image processing tasks, like image filtering are often quite computational intensive. FPGA can deliver enough computational power at a relative small power usage. But therefore the algorithm has to be well suited to the FPGA environment. Processing steps should be static and the algorithm should have high parallelism potential. Basic window filters like the medium, median or sobel filter are quite good applicable.[3] But it is difficult and time intensive to implement more complex filters like for example a mean shift filter[4] or component labelling.[5] Therefore a lot of work has to be done to find an appropriate segmentation method and adapt this in a benefiting way. The chosen segmentation method is designed for gray image segmentation based on contour extraction. The segmentation approach is shortly described in the second section. The main outcome of in this paper is the quality performance evaluation of this segmentation algorithm. Quality and quantitative analysis is very important, as for the construction as for the user to estimate the suitability for his purposes. The main difficulty is that there is no theory for image segmentation and no standard method for quality evaluation.[6] One of the most common ways for evaluating the performance of a segmentation method is still a subjective evaluation, in which human experts determine the quality of the segmentation result.[6] But this is not in compliance with our needs. A method for a quantitative, objective performance evaluation, including the theory, the method and results is presented in the third section.

Send correspondence to Kurt Schwenk
E-mail: Kurt.Schwenk@dlr.de, Telephone: (+49) 0 8153 28 3493

Image and Signal Processing for Remote Sensing XIX, edited by Lorenzo Bruzzone,
Proc. of SPIE Vol. 8892, 889208 · © 2013 SPIE · CCC code: 0277-786X/13/$18
doi: 10.1117/12.2028660

2. SEGMENTATION METHOD

First a short description of the employed segmentation method is given. In Jahn/Halle[7–9] a segmentation method for gray value segmentation in the context of on-board processing is presented. The main part is an edge preserving smoothing filter. It is a kind of dynamic bilateral filter,[10] choosing it's parameters in dependency of the image contents. This method provides a basis for the developed method. It uses a lot of floating point operations and square root operations which are not ideal for an FPGA implementation. Therefore a special form mainly based on integer values was created. Its main part is the so called gap smoothing filter. It is an edge preserving smoothing 3x3 window filter. It performs a cluster analysis on the filter window to decide the histogram has an uni-modal or a bimodal histogram distribution. If there is a bimodal distribution, the average of the grey value's of the cluster containing the center pixel is computed. If there is no bimodal distribution it acts like a medium filter averaging over all window pixel gray values. The whole segmentation module can be seen in Figure 1. The filter quality can be clearly improved if a pre-filter (Level 1 filter), in this case a medium

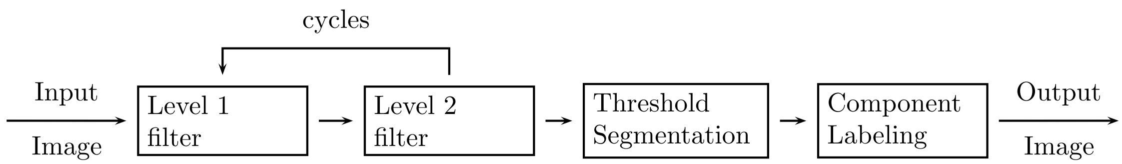
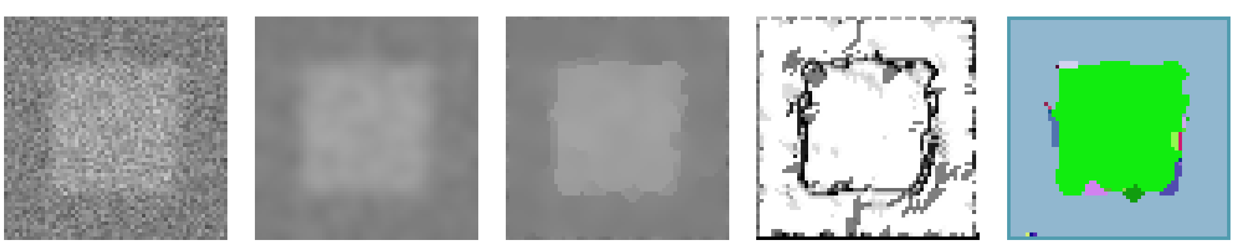


Figure 1. segmentation module

filter, is used and if the filter steps are executed several times. A simple threshold segmentation is used to determine the neighbourhood relation of the pixels. Adjacent pixels are connected if their grey colour difference is less than a predefined threshold. The neighbourhood relation is coded in the local connection map. In the end a component labelling[1] is carried out. The result is a false colour image, the component map, in which every segment is labelled with an unique colour. An example is given in figure 2. Here an image with Gaussian noise is smoothed by the filter. For the Level 1 filter a 3x3 medium filter (2 iterations) and for the Level 2 filter the gap smoothing filter (15 iterations) is used. The segmentation threshold is set to 1. The result should be a perfect square in the middle of the image. In figure 2(e) it can be seen that this result is not perfect. In the following chapter it will be shown how this failure is measured.

For a demonstration the whole process chain was implemented on a Xilinx Spartan 3E-1200 FPGA[11] running at 12.5 MHz. In the first design for each pixel and smoothing filter iteration (Level 1 and Level 2) three clocks are required. Therefore the the smoothing process needs $3 * 1024 * 1024 * iterations * clock$ cycles for a $1024 * 1024$ image. Note that in the above example for the whole smoothing process 17 iterations are needed. In other words in this form four smoothing filter iterations per seconds can be executed. Therefore approximately four seconds for the whole smoothing process of a 1024*1024 image are required. On more advanced hardware for example the Xilinx Spartan 6 with a faster and wider memory system, only one clock is required for each pixel and filter iteration and frequencies higher than 50MHz can be reached. Then about three 1024*1024 images per second can be processed. If more frames per second are needed or very large images have to be processed, two or more smoothing filter modules can be implemented in parallel on the FPGA, linearly increasing the performance.



(a) (b) (c) (d) (e)

Figure 2. Segmentation process: (a) noisy test image, (b) after Level 1 filter, (c) after Level 2 filter, (d) local connection map, (e) component map

3. IMAGE SEGMENTATION EVALUATION

3.1 Goals and concepts

An image segmentation evaluation is important for a later user and for the developer. For a developer it is inevitable for optimising and further developing a segmentation method. For a user it is helpful to choose the right segmentation method, to perform a reasonable implementation and to be able to estimate the behaviour of a method. Here, the segmentation evaluation was performed firstly to determine the (optimal) parameters of the segmentation method, secondly to make a performance measurement. In which cases it is applicable? What are the limits? And what is the general behaviour of the method? In the end it should enable a comparison with other approaches. To address these points, the following requirements were determined. The first and most important point is that with the segmentation evaluation a reasonable quality assessment is possible. Furthermore it should be objective in terms, that the evaluation criteria and the quantitative determination of these are precise and definite. This is very important for the understanding and the interpretation of the evaluation results. Furthermore the evaluation results should be easy to interpret and understand. In the literature, often a lot of different performance measurements are presented in a table or a graph. The meaning of the performance measurements is poorly presented. Therefore it is very difficult to understand the results. In the end the measurement method should be both simple in the construction and in the execution. The main problem is, that even nowadays there is no general segmentation theory and no standard procedure for a performance evaluation.[6,12] A composition of different evaluation methods is presented in.[6] They are categorised in *subjective* and *objective* methods. Objective methods are further classified in *system-level* and *direct* evaluation methods. A direct evaluation can be further divided into *analytical* and *empirical* methods. And in the end empirical methods are catalogued in *supervised* and *unsupervised* evaluation methods. The method presented here is a supervised method. To an image, a reference segmentation has to be created. With a discrepancy measure the segmentation result and the reference segmentation are compared. The difficulty is to find a useful discrepancy measure and to create a reference image. The creation of a reference segmentation can be simplified if a synthetic test image is chosen. The elementariness of the test image is also a key point in the analysis and interpretation of the measurement results. In the following chapter a mathematical description of the terms image segmentation, segmentation criteria and unsupervised evaluation is presented.

3.2 A small theory of image segmentation and evaluation

In the following, general and basic principles of an image segmentation and evaluation are introduced. It is very important to clarify what should be understand under an image segmentation quality. First some notations. Let $\mu : \Omega \to C$ an image. Ω is the position (coordinate) space of an image, in general a subset of $\mathbb{R}^2$ or $\mathbb{N}^2$. Let C be the color space of the image. For example an 8 bit grey value image is of the form $\mu : [0, \text{width}]_{\mathbb{N}} \times [0, \text{height}]_{\mathbb{N}} \to [0, 255]_{\mathbb{N}}$. The set of all images over the position space Ω and the colour space C is called $Img(\Omega, C)$ or if the colour space is determined or not important $Img(\Omega)$.

A segmentation of an image $\mu : \Omega \to C$ is a partition of the position space. A definition of the partition of a set can be found in every mathematical introduction course, for example.[13] The set of all segmentations/partitions of Ω is called $Seg(\Omega)$. It is important that the external structure of a segmentation of an image is only dependent on the position space Ω. For the description of how a segmentation should look like, a segmentation rule is needed. Here it is called segmentation criteria and represented through its quality function

$$c : Img(\Omega) \times Seg(\Omega) \to \mathbb{R}$$

For an image and an segmentation the criteria quality function $c(\mu, s)$ describes the goodness of the segmentation s in context to the criteria, if it is supposed to be a segmentation of μ. Here the agreement is that higher values of c are better. If there are more criteria, they must be merged to have the form of c. Only in this way a grading and comparison is possible. One of the most problems is to determine the criteria quality function c. One of the first image segmentation criteria can be found in[14] - *"Regions of an image segmentation should be uniform and homogeneous with respect to some characteristic such as gray tone or texture. Region interiors should be simple and without many small holes. Adjacent regions of a segmentation should have significantly different values with respect to the characteristic on which they are uniform. Boundaries of each segment should be simple, not ragged, and must be spatially accurate"*. These criteria are rough descriptions. The question is, what is the right way to render these criteria more precisely and formulate them in the proposed way. This is necessary because only in that way a exact and objective description can be expressed in general. An exact criteria, modelling some of the above criteria is the the Mumford and Shah model presented in.[15] It can be expressed as a combination of three partial criteria.

$$c_1(\mu, s, \mu_0) = \int_\Omega (\mu_0 - \mu)^2 dxdy, \quad c_2(\mu, s, \mu_0) = \int_{\Omega - \delta_s} | \bigtriangledown \mu_0^2 | dxdy, \quad c_3(s) = |\delta_s|$$

The Mumford and Shah model is then given through

$$c(\mu, s) = -\inf\{c_1(\mu, S, \mu_0) + \lambda * c_2(\mu, S, \mu_0) + \nu * c_3(s) \mid \mu_0 \in Img(\Omega)\}$$

where $\lambda, \mu \in \mathbb{R}$ are free selectable parameters and μ_0 is a linking parameter. In this work this criteria is not used, but it should give a first impression of the difficulty to create useful segmentation criteria. The unsupervised segmentation measures presented in[6] can in the most cases be interpreted as segmentation criteria too.
For a criteria c the best segmentation s_{opt} of an image μ has the highest goodness

$$c(\mu, s_{opt}) = \max\{c(\mu, s) | s \in Seg(\Omega)\}$$

One important point is that often there is not the one "best solution", there can be more "best solutions". Another important point is that every segmentation method M can be interpreted as a criteria in the following way, $c_M(\mu, s) = 1$ if $s = M(\mu)$ else $c_M(\mu, s) = 0$. This is not only an academic statement, it should make clear that for every segmentation method there is a criteria for which the segmentation method is the optimal one.

To measure the performance of a segmentation algorithm we need a criteria and a (weight) test set $T_\lambda = \{\mu_i \in Img(\Omega), \lambda_i \in \mathbb{R}\}_{i \in I}$ where i is a finite index set and λ_i a weight parameter. The performance or goodness of a segmentation method A is than given through

$$G_{T_\lambda, c}(A) = \sum_{i \in I} \lambda_i * c(\mu_i, A(\mu_i))$$

It is important to notice that a performance analysis is intrinsically tied to the (weight) test set and the criteria.

Now the supervised way of constructing a criteria is presented. Let $T = \{\mu_i \in Img(\Omega)\}_{i \in I}$ be a test set. Let $R = \{s_i \in Seg(\Omega)\}_{i \in I}$ be a set of reference segmentations. If we have a difference function (discrepancy measure)

$$d : Seg(\Omega) \times Seg(\Omega) \to \mathbb{R}$$

on $Seg(\Omega)$ and a set of reference segmentation R, a criteria on T can be constructed in the following way: $c_{(T,d)}(\mu_i, s) = d(\mu_i, s_i)$. If necessary an additional minus can be used to guarantee that a higher criteria value can interpreted as a higher criteria goodness. Note that this criteria can only be applied on the test set T. This is the main constraint to unsupervised criteria as for example the Mumford and Shah Model, but it enables us to build a criteria without understanding its intrinsics and the need to formulate them.

At the end it should be seen that there a two problems in the field of image segmentation. The first challenge is the formulation of an appropriate image segmentation criteria. The second challenge is the search of a segmentation algorithm, performing optimal under the chosen criteria. The problem of an evaluation of a segmentation measurement is equal to the problem of finding a segmentation criteria. Here, because the supervised way is chosen, the problem of finding a criteria is shifted to finding a test set with reference segmentations and a difference function. This is addressed in the next section.

3.3 The measurement method

In this section the measurement method is presented. As stated in the last section's a supervised method is used. Therefore a discrepancy measure $d : Seg(\Omega) \times Seg(\Omega) \to \mathbb{R}$ is needed. The method on which the measurement function used in the work is based was utilized in,[16] probably going back to.[17] The segment accuracy function has the form

$$SA(a|S) = \sum_{m \in S} \frac{|a \cap m|}{|m|} * \frac{|a \cap m|}{|a|}$$

Note, that for a set M, $|M|$ is the number of elements of M. Let $S \in Seg(\Omega)$ be the reference segmentation and let $a \subset \Omega$ be an arbitrary segment. $SA(a|S)$ gives a rating on how well this segment fits in the segmentation S. This function can be extended to a discrepancy measure of two segmentations in the following way:

$$SA(S_2|S_1) = \frac{\sum_{a \in S_2} SA(a|S_1) \ * \ |a|}{\sum_{a \in S_2} |a|}$$

where $S_2, S_1 \in Seg(\Omega)$ are two segmentations. The following conditions have been proven:

1. $SA(S_2|S_1) \in (0, 1]$
2. $SA(S_2|S_1) = 1 \Leftrightarrow S_1 \leq S_2$

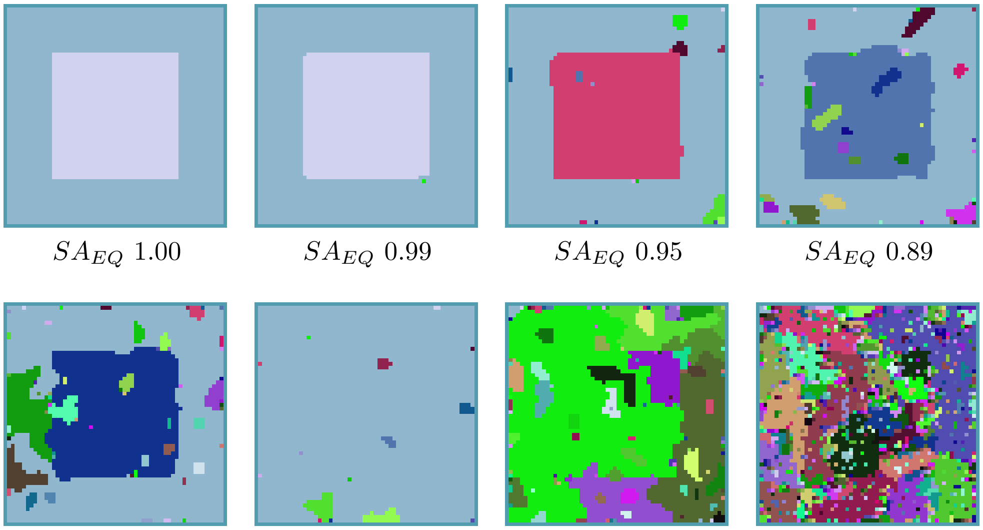
$S_1 \leq S_2$ means that S_1 is a refinement of S_2 in mathematical terms $\forall m \in S_1 \ \exists n \in S_2 \ m \subseteq n$. In the language of segmentation S_1 is a (pure) over-segmentation compared to S_2. In words explained $SA(S_2|S_1)$ measures the degree of over-segmentation of S_2 compared to S_1. Because of the statement above, it is is not possible to measure the difference of S_1, S_2 if S_2 is an over-segmentation compared to S_1. Therefore in[16] a second measurement function, the over-segmentation index $OSD(S_2|S_1) = \frac{|S_2|}{|S_1|}$ was used based on the quotient of the numbers of segments in each segmentation. The two measurement function were not merged and therefore for each measurement two graphs were presented. The presented measurement results are therefore not easy to interpret. Hence it was decided to take another approach. In the following the Segmentation Accuracy Equality measurement function (SA_{EQ}) is presented:

$$SA_{EQ}(S_2, S_1) = \sqrt{SA(S_1|S_2) * SA(S_2|S_1)}$$

The following properties have been proven:

1. $SA_{EQ}(S_1, S_2) \in (0, 1]$
2. $SA_{EQ}(S_1, S_2) = SA_{EQ}(S_2, S_1)$
3. $SA_{EQ}(S_1, S_2) = 1 \Leftrightarrow S_1 = S_2$

In figure 3 it can be seen that with this function a reasonable quality measurement is possible. The segmented images presented here are compared with a reference segmentation. The reference segmentation is the image in the top left, therefore as stated in the third property it has SA_{EQ} value 1.00. Now going from left to right the quality gets worse. In SA_{EQ} 0.99 image without the edges of the inner square, everything is perfect. Little disturbances appear in the SA_{EQ} 0.95 image whereas in the SA_{EQ} 0.89, 0.86 cases the discrepancy gets larger. In the last three cases great failures appear. The inner square could not be separated from the background in the SA_{EQ} 0.72 case. In the last two cases the inner square is barely visible. Overall the here presented examples are in agreement with the subjective impression.



SA_{EQ} 1.00 | SA_{EQ} 0.99 | SA_{EQ} 0.95 | SA_{EQ} 0.89

SA_{EQ} 0.86 | SA_{EQ} 0.72 | SA_{EQ} 0.48 | SA_{EQ} 0.24

Figure 3. SA_{EQ} example images

3.4 The test set

In this chapter the chosen test set is presented. It was decided to use synthetic test images. A perturbation model based on random Gaussian noise and image blurring is used to simulate real world situations. The reason for using synthetic images is that a reference image is available and that a perturbation model can be exactly described and modified. Another important point is that the elementariness of the test model, allows the interpretation and understanding of the measurement results. Real image scenes were analysed. In a lot of cases the segmentation behaviour could be explained with help of the measurement results, produced with the synthetic images. On the test image only one square is placed in the middle. The square has the size 34x34 pixel and the whole image 60x60 pixel. The grey value difference between the square and the background is varied between 0-64. Three perturbation models were chosen:

- **Model A:** Gaussian noise
- **Model B:** Gaussian noise, after a blurring of the image two times with medium 3x3 filter
- **Model C:** blurring the image two times with medium 3x3 filter and then adding Gaussian noise

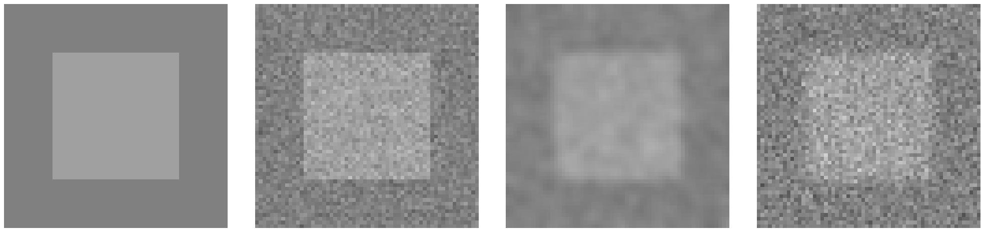
In the Model B and Model C image blurring is used to smear out edges. In that way neighbourhood effects at edges should be simulated. To specify the level of the Gaussian noise the standard variance is used. Note that the standard derivation is the square root of the standard variance ($\delta = \sqrt{var}$). Example test images are presented in figure 4.

4. EVALUATION RESULTS

In this chapter some measurement results are presented. It should be demonstrated that the quality analysis provides great benefit in developing and understanding the segmentation method. For clarity in this work only Model A test images are utilised.

4.1 Parameter estimation

The first reason the quality assessment was carried out, has been the determination of optimal algorithm parameters. The gap smoothing algorithm has the form X_1 * mean filter, X_2 * gap(g_1,g_2,g_3), threshold segmentation t. X_1 * Mean Filter is the first filter step. X_2 * Gap(g_1,g_2,g_3) the second filter step, and threshold segmentation



(a) (b) (c) (d)

Figure 4. Perturbation models: (a) reference image with 32 gray value difference, (b) Model A, (c) Model B, (d) Model C, all perturbation models with noise level 256

the segmentation step. X_1, X_2 are the number of filter runs. g_1, g_2, g_3 are gap smoothing filters parameters, set constant to $g_1 = 1.5$ $g_2 = 1.5$ $g_3 = 4$. In the following the determination of the number of runs of the second filter step X_2 is presented. The test image set consists of Model A perturbation images, each with an grey value difference 32 and Gaussian random noise with a variance ranging form 0 ... 1024. For each noise level 129 different sample images were examined to take into account the random character of the noise. In figure 5 it can be easily seen that the number of runs has a great impact. Higher noise levels required more runs. With this

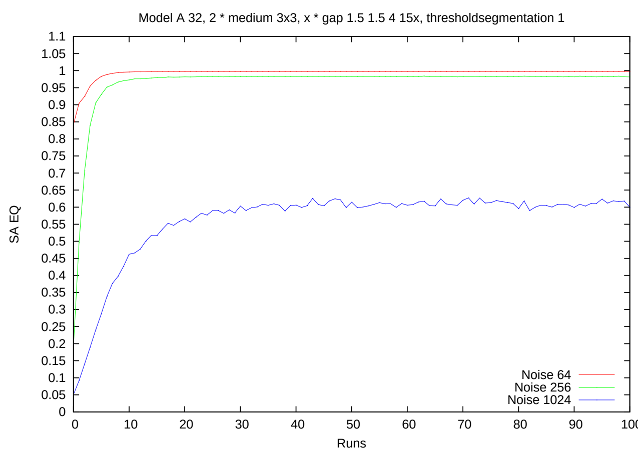


Figure 5. level 2 filter runs, Model A, grey value difference 32

filter configuration noise level 1024 image cannot be segmented satisfactorily. A high quality, over 0.95 SA_{EQ}, could be reached on low level noise images. Another point which should be mentioned is that the filter is stable in that sense that at a certain run number optimal quality is reached and remains at this level. This is a special feature of that filter. For a pure medium filter this is not the case. If the processing time is a critical point here 5 runs should be enough to get a good segmentation performance. In this work 15 runs are chosen because model 2, 3 require more runs to achieve an optimal quality and the processing speed was not the key factor.

4.2 Segmentation characterisation graph

In this chapter the segmentation characterisation graph is presented. It is the main result of this work. This graph is a first description of the overall performance of a segmentation model. The SA EQ quality is measured in dependency of the noise level and grey value difference of the test image. It can be determined which grey level difference is necessary at a certain noise level to achieve good segmentation results. An example is given in figure 6. Two parameter configurations are shown. On the right images the filter used in the left image is repeated four times, with a skip of the very first medium 2x filter step. It can be seen that on the left image at high noise levels (576, 1024) no satisfying quality can be reached. On the right image even high noise level could be processed but at the cost of some sensitivity. It can be seen that with higher noise level more grey value difference is needed

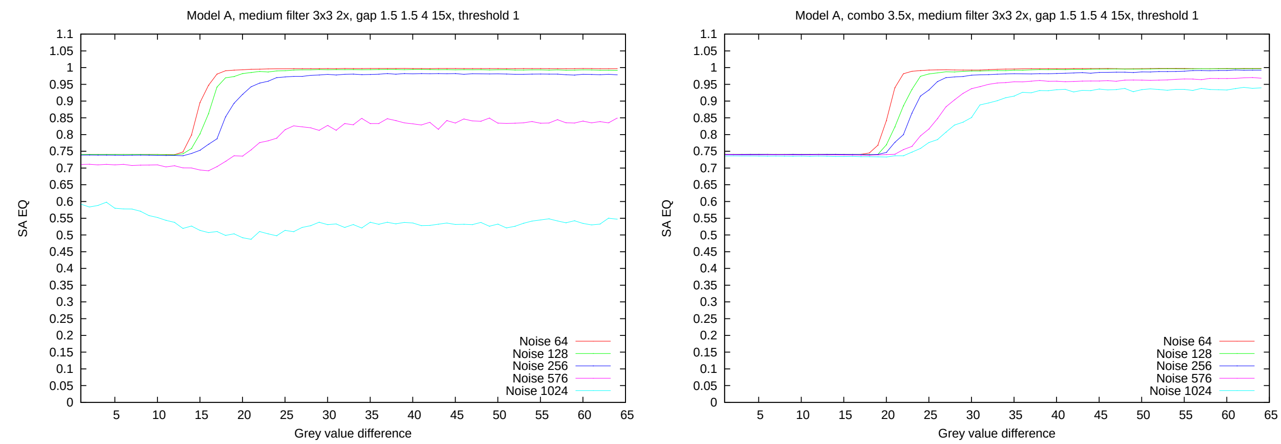


Figure 6. filter sensitivity, model a) two parameter configurations

and that the maximal achievable quality drops. A lot of segmentation characterisation graphs were analysed. The structure is very similar to the presented ones. Generally three regions could be located. The "insensitive region". In this region the grey value is too small to divide the square from the background (in that case about 0.75 SA_{EQ} is achieved). In our example on the left image it's from 0 to 15 for the noise level 64, 128. 256. After that we have a transition region. Here the dividing is heavily dependent on the noise distribution. At the same noise level in one case no dividing of the inner square is possible in another a good dividing is achieved. The quality in this region is very dependent on the grey value difference. The third region is the maximal region. In this region maximal quality is reached.

4.3 Comparison of different segmentation methods

A reasonable way to compare different segmentation methods is to compare their segmentation characterisation graphs. In this section the presented method is compared with two alternative methods.

1. 2x Medium 3x3, 15 x gap 1.5 1.5 4, segmentation threshold 1
2. 2x Medium 3x3, 15 x threshold smoothing 6, segmentation threshold 1
3. Combo 30x, bilateral-filter 6 3, median filter, threshold segmentation 1

The first alternative approach is using a simple threshold filter instead of the gap smoothing filter. In a 3x3 filter window only the gray values with an absolute difference to the center pixel smaller then a given threshold are averaged. The second alternative is a bilateral median combination (Combo). First a bilateral filter is executed and then a median filter. This process is repeated 30 times. More details about the bilateral filter can be found in.[10] The test sets are Model A images with a noise level 256. The filter parameters are adjusted to have the same sensibility, starting at around 15 grey value difference.
It can be seen in figure 7 that the gap filter and threshold filter have nearly the same sensitivity, but at an higher gray value difference the threshold filter quality drops. As seen in figure 8 the threshold filter cannot dissolve the border region of the inner square at higher grey value differences as good as the gap smoothing filter. This is due to the static threshold that is not adjusted as it is done by the gap smoothing filter. The bilateral median filter has the same sensibility but for high quality results a lager gray value difference is necessary. The problem of comparing different segmentation methods is the interpretation. To come back to the theory chapter 3.2 the overall performance of a method can be computed as the average SA_{EQ} value of all test images ranging from grey value difference of 0 to gray value difference of 64. The goodness values are $G_{T,SA_{EQ}}(gap) = 0,91265$, $G_{T,SA_{EQ}}(thres) = 0.88969$ and $G_{T,SA_{EQ}}(bil) = 0.89922$.
Overall the test result provides great assistance for getting a first insight and detecting interest spots. If necessary the kind of failure can be analysed further by visual inspection, as done in figure 8.

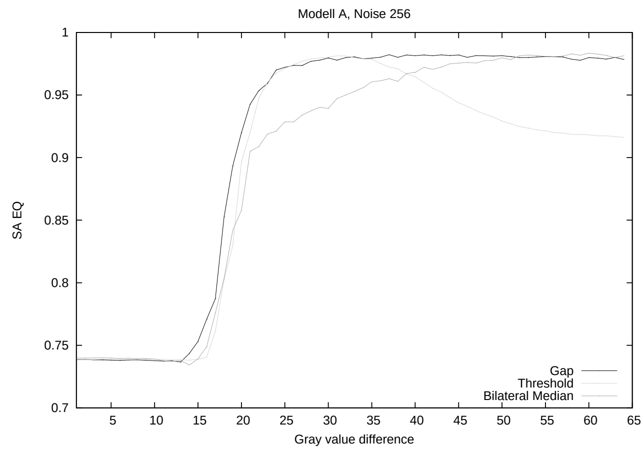


Figure 7. Filter comparison

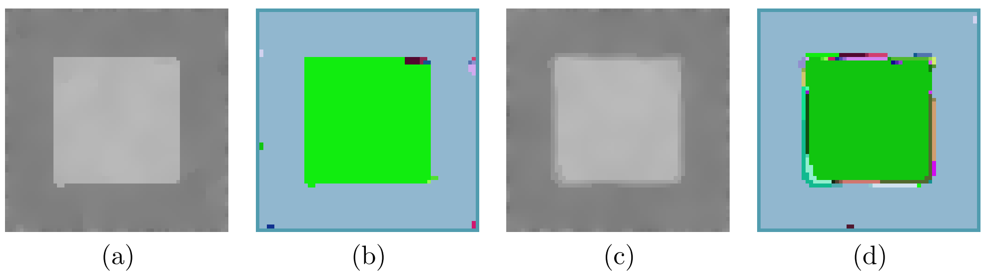


Figure 8. difference gap-threshold smoothing and gap-threshold segmentation: (a) gap smoothing result, (b) gap segmentation result, (c) threshold smoothing result, (d) threshold segmentation result

4.4 Discussion

In the last sections a method for the evaluation of a segmentation method was presented. The requirements are reasonable quality assessment, objectivity, good interpretability and a practical applicable measurement process. With the chosen method, these points are fulfilled in most parts. With the help of the evaluation results filter optimization were carried out, which could be hardly achieved only with visual inspection. One point is the possibility to evaluate a lot of test images. For example 64 * 129 images were analysed to get one filter characteristic curve. Another point is that with SA_{EQ} the quality of a segmentation could be expressed in one number, which helps to interpret the data. Besides it is important to mention the usage of an elementary synthetic test image. This was primarily done to get a reference image. But it turned out that this was very helpful for the interpretation of the measurement data. Without the use of a simple test model, consisting of only two segments, background and foreground, the filter characteristics graph would simply not have that simple structure. The filter characteristics graph allows the determination of the sensibility of a segmentation method. The correlation of sensibility, noise and segmentation accuracy and the influence of edge sharpness (not presented here) could be analysed. Furthermore the filter characteristic graph provides an option to compare different segmentation methods in an objective way, pointing out their weaknesses and strength. With the help of the sensitivity graph it could be understand why in a real satellite image in one case a segmentation of a real world object occurred, in another not. This work is a rough first evaluation of a segmentation method. There is space for further investigations. For example the spatial effect could be examined more in detail. What is the smallest object that can be identified at a given grey value difference? What is the influence of texture?

4.5 Summary new

In this paper a new method for the evaluation of an image segmentation method, designed for an advantageous implementation on an FPGA chip is presented. In the first part the segmentation method is shortly explained.

In the second part the whole evaluation process is demonstrated. Starting with the description, in mathematical terms, what should be understood under image segmentation and image segmentation quality in general, a overview of that topic and a comprehensive theoretical background is given. It was pointed out that an image segmentation evaluation is strictly connected to a (goodness) criteria and a test set. The construction of a (goodness) criteria with the help of a discrepancy measure is presented. With that construction also the connectivity and difference of supervised and unsupervised evaluation methods is explained. After the theoretical introduction a concrete way to evaluate a segmentation method is presented. A new discrepancy measure called SA_{EQ} was introduced in detail and it is shown with some objective mathematical facts as well with subjective examples that it is feasible for quality measurement. The test sets and the perturbation models are presented and the advantage of using elementary synthetic test images is illustrated. In the end the benefit of the proposed segmentation evaluation method is demonstrated. Three exemplary measurement results are shown. In the first example a graph is presented, which is showing the influence of one algorithm runtime parameter towards the segmentation quality. With that the behaviour of the algorithm in context to that parameter could be analysed and a good value for that parameter could be chosen. In the second example, the most important result of the quality assessment is presented: a filter characterisation graph, describing the segmentation performance in dependency of the gray value difference between the fore and background object of the test set at various noise levels. With that measurement, the impact of contrast and noise level on the segmentation process is objectively and quantitatively shown. In the end a comparison of the segmentation method with two alternative approaches is presented and the possibility to detect differences is shown. Overall this paper gives an input in the understanding of image segmentation and serves as an example for segmentation evaluation. This helps booth, developer and later users in the understanding and optimisation of segmentation methods.

REFERENCES

[1] R. Haralick and L. Shapiro, *Computer and robot visión*, no. Bd. 1 in Computer and Robot Vision, Addison-Wesley Pub. Co., 1992.

[2] K. Schwenk, K. Goetz, M. von Schoenermark, and F. Huber, "Real-time evaluation of remote sensing data on board of satellites," in *2011 21st International Conference on Field Programmable Logic and Applications*, pp. 399–400, IEEE, 2011.

[3] D. V. Rao, S. Patil, N. A. Babu, and V. Muthukumar, "Implementation and evaluation of image processing algorithms on reconfigurable architecture using c-based hardware descriptive languages," *International Journal of Theoretical and Applied Computer Sciences* **1**(1), pp. 9–34, 2006.

[4] D. B. K. Trieu and T. Maruyama, "An implementation of the mean shift filter on fpga," in *Field Programmable Logic and Applications (FPL), 2011 International Conference on*, pp. 219–224, 2011.

[5] M. Jablonski and M. Gorgon, "Handel-c implementation of classical component labelling algorithm," in *Digital System Design, 2004. DSD 2004. Euromicro Symposium on*, pp. 387–393, IEEE, 2004.

[6] H. Zhang, J. Fritts, and S. Goldman, "Image segmentation evaluation: A survey of unsupervised methods," *Computer Vision and Image Understanding* **110**(2), pp. 260–280, 2008.

[7] H. Jahn, "Graph structure for image segmentation," *Proc. SPIE* **3026**, pp. 198–208, 1997.

[8] H. Jahn, "Segmentation of remote sensing images with a layered graph network," *International Archives of Photogrammetry and Remote Sensing* **31**, pp. 360–364, 1996.

[9] W. Halle, "Ausgewählte algorithmen der segmentierung und klassifikation zur thematischen bildverarbeitung in der fernerkundung," tech. rep., Deutsches Zentrum für Luft- und Raumfahrt e. V. Institut für Weltraumsensorik und Planetenerkundung, 1999. ISSN 1434-8454.

[10] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Computer Vision, 1998. Sixth International Conference on*, pp. 839–846, IEEE, 1998.

[11] X. Corporation, "Spartan-3e fpga family: complete data sheet," 2013.

[12] M. Neubert, H. Herold, and G. Meinel, "Evaluation of remote sensing image segmentation quality–further results and concepts," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **36**(4/C42), 2006.

[13] J. Lucas, *Introduction to Abstract Mathematics*, G - Reference, Information and Interdisciplinary Subjects Series, Ardsley House, 1990.

[14] R. Haralick and L. Shapiro, "Image segmentation techniques," *Computer vision, graphics, and image processing* **29**(1), pp. 100–132, 1985.

[15] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Communications on Pure and Applied Mathematics* **42**(5), pp. 577–685, 1989.

[16] M. Neuenhahn, H. Blume, and T. Noll, "Pareto optimal design of an fpga-based real-time watershed image segmentation," in *Proceedings of the Conference on Program for Research on Integrated Systems and Circuits (ProRISC'04)*, 2004.

[17] M. Everingham, H. Muller, and B. Thomas, "Evaluating image segmentation algorithms using the pareto front," in *Computer Vision ECCV 2002*, A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, eds., *Lecture Notes in Computer Science* **2353**, pp. 34–48, Springer Berlin Heidelberg, 2002.