Primary Image Segmentation

Ullrich Köthe
Fraunhofer-Institut für Graphische Datenverarbeitung
D-18059 Rostock, Joachim-Jungius-Str. 9, Email: koethe@egd.igd.fhg.de

Abstract: This paper introduces the notion of *primary image segmentation* which serves as a well defined link between low- and high-level image analysis. A general algorithmic framework based on priority queues is proposed that allows for the integration of a variety of different segmentation algorithms. A seeded region growing approach, along with a number of improved seed selection methods and foveation of critical areas, is chosen to realize this framework. Experimental evaluation shows very good performance of these algorithms on a relatively large number of outdoor photographs without the need to adjust parameters.

1. Introduction

The segmentation of images has always been a key problem in computer vision. Up to the early 80ies bottom-up techniques like edge detection and split-and-merge algorithms were the primary focus of research. However, by that time people realized that „perfect“ segmentation would not be possible without incorporation of higher level knowledge. Thus the focus shifted towards model based techniques like snakes [1] and methods based on geometric models [2]. These techniques generally give very satisfying segmentation results. However, on a number of reasons which have been underemphasized in the past, the bottom-up techniques are still necessary as a natural complement of the top-down approaches:

- If the model base consists of a great number of different models, say 10 000, we need effective strategies to choose, among all models in the database, those models which have a great chance of being a correct description of the image content. If we lack a-priori scene knowledge this can only be done by bottom-up analysis of the image itself.
- Many model based techniques are based on iterative improvement of an initial model configuration. Appropriate low-level information is needed to find an initial estimate of the model that ensures convergence.
- Often one and the same image may be analyzed using several different models depending on the task to be fulfilled. This is typically the case with content based retrieval in image databases. A primary segmentation could be stored as a common starting point for all those tasks.

These examples show that the full power of top-down methods can only be exploited if they are built upon a well defined bottom-up analysis. Therefore we introduce the notion of *primary image segmentation* which is meant to serve as a link between bottom-up and top-down approaches of segmentation. The goal of primary segmentation is therefore similar to the goal of Marr’s *primal scetch* [3]. However, our approach will be based on regions, since regions, as 2 dimensional entities, provide a much richer image description than edges.

The work described here is based on the work of Lanser [4] and Adams and Bischof [5]. Lanser [4] describes a technique that combines edge oriented and region oriented algorithms to improve segmentation results. Adams and Bischof [5] introduce a new region growing algorithm that overcomes severe limitations of older approaches. This article will extend those ideas by putting them in a general framework which allows for the combination of a variety of segmentation algorithms, resulting in great potential to optimize all aspects of the segmentation process.

2. Definition of primary image segmentation

Primary image segmentation is defined as an optimal segmentation obtained in a pure bottom-up fashion that provides the information necessary to initialize and constrain high-level segmentation methods. Although the details of primary segmentation methods will depend on the application domain, we require that they do not depend on a priori knowledge about the objects present in a particular scene or image specific parameter adjustments. These claims become realistic because we do not seek for a perfect segmentation result but rather for the best possible support for more intelligent methods to be applied afterwards.

Unfortunately up to now there is no theory which defines the quality of a segmentation. Therefore we have to rely on some heuristic constraints which the primary segmentation should meet:

- The segmentation should provide regions that are homogenous with respect to one or more properties, i.e. the variation of measurements within the regions should be considerably less than the variation at borders.
- The position of the borders should coincide with local maxima, ridges and saddle points of the local gradient of the measurements.
- Areas that perceptually form only one region should not be splitted into several parts. In particular this applies to smooth shading and texture.
- Small details, if clearly distinguished by their shape or contrast, should not be merged with their neighboring regions.

Fig. 1:
Wrongly connected regions resulting from inversion of an edge image obtained using Chen and Castans algorithm [8].
The regions shown form only 4 connected components which is obviously semantically incorrect. However, detection of the missed boundaries during postprocessing is very difficult.
(For the original see upper left of fig. 5)

3. Algorithmic Framework for Primary Segmentation

As was already mentioned primary segmentation should be a region based segmentation. The simplest idea to obtain this would be the complement of a standard edge detection (see [4]). Unfortunately edge detection algorithms tend to miss important edges if their contrast is poor which results in an undersegmentation (see fig. 1). Even an extended algorithm that tries to close edge fragments starting from terminating points as proposed in [4] will not reliably find all region boundaries.

In our opinion variants of region growing have the greatest potential to meet all requirements outlined in section 2. Here the segmentation process is split into two stages: at first an initial incomplete segmentation forming seed regions is performed. A subsequent region growing stage assigns the yet unlabeled pixels to one of the regions until the segmentation is complete. This leads to a four-step framework for primary segmentation (fig. 2): preprocessing, seed selection, region growing and optional postprocessing to improve the regions. To emphasize the role of primary segmentation as a link between low-level and high-level processing the figure also includes the high-level part which builds upon primary segmentation.

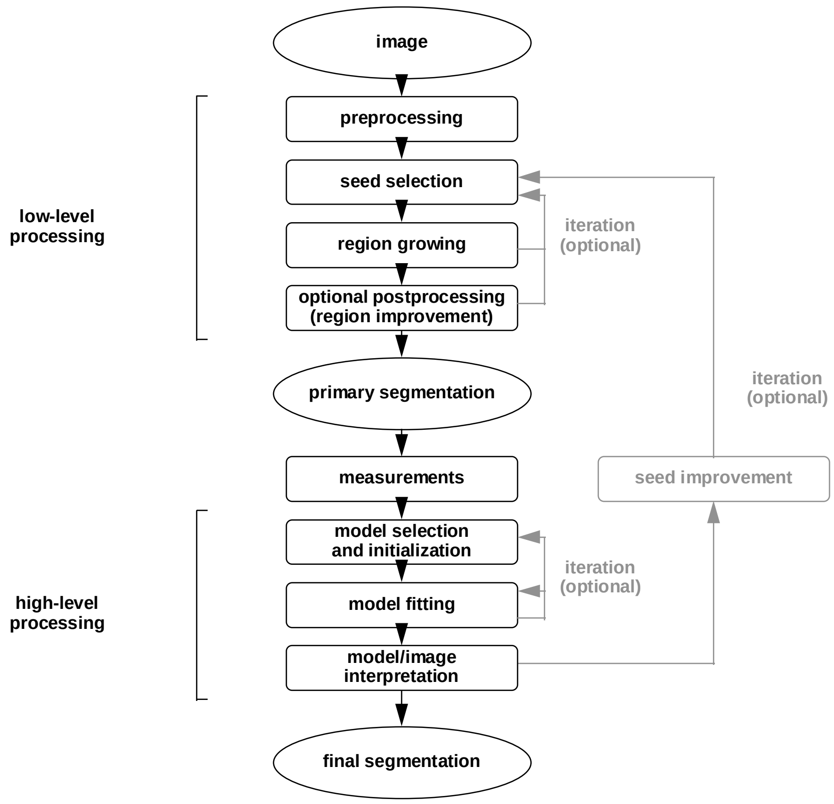

Fig: 2:
Proposed vision pipeline where primary segmentation forms the center of the algorithmic framework

In its simplest form this scheme may be regarded as a *vision pipeline* which starts at an image and ends at a final interpreted segmentation. However, for most images *iteration* would be necessary to obtain satisfying results.

4. Specifying the Algorithms

Seeded Region Growing

Region growing algorithms start from an initial, incomplete segmentation and try to aggregate the yet unlabelled pixels to one of the given regions. The initial regions are usually called *seed regions* or *seeds*. The decision whether a pixel should join a region or not is based on some fitness function which reflects the similarity between the region and the candidate pixel. As proposed in [5], the order in which the pixel are processed is determined by a global priority queue which sorts all candidate pixels by their fitness values. This approach elegantly mixes local (fitness) and global (pixel order) information. The algorithm proceeds as follows:

1. Find initial set of candidate pixels, calculate their fitness and put them into the priority queue.
2. While(Queue is not empty.)

   2a) Get candidate pixel with best fitness from queue.

   2b) If (Candidate has more then one neighboring regions.) then

   Mark pixel as border region.

   else

   Mark pixel with label of its neighboring region.

   Identify new candidates among the neighbors of the pixel just processed,

   calculate their fitness and put them into the queue.

Candidate pixels are those pixels which are neighbors (4- or 8-connectivity) of exactly one labeled region. When regions meet in the curse of growing, a one pixel wide border is left (step 2b).
Possible fitness functions include:

- The fitness is equal to the local gradient. Then regions always meet at local gradient maxima. This is exactly the behavior of the well known *watershed algorithm* [6], which thus can be viewed as a special case of seeded region growing.
- In [5], Adams and Bischof use the magnitude of the difference between the graylevel of the pixel and the mean graylevel of the region. Slightly worse edge localization is paid off by better resistance of the segmentation against noise.
- Generalizations to color images include the maximum of the gradient among all color bands or any of distance measures between the pixel's color and the mean color of the region (e.g. Euclidean distance in RGB- or CIE-Lab-space, angular color distance).
- Complex fitness functions can be defined using linear combinations of simpler functions.

For very thin regions of about 3 pixel width we encounter aliasing problems which are solved by *foveation* [9]: the resolution of those critical areas is increased by a factor of 2 in either direction. Fig. 7 shows a significant improvement of segmentation accuracy obtained this way.
The most critical part of the algorithm turns out to be the selection of the seeds. As the growing does not change the number of regions a region lost during seed selection can not be recovered later. The same holds for regions that have mistakenly been splitted into more than one seed. Although some postprocessing may be performed to repair these errors, special care must be taken when the seeds are selected.

Seed selection

Seed selection may be viewed as an incomplete segmentation procedure where pixels are only assigned to regions if this decision can be made with high confidence. However, even so we do not expect one single seed selection method to be good enough on all occasions. (If we had such a multi-purpose seed finder we could probably use it to segment the image directly.) Therefore we propose a procedure which *combines* raw seeds from a number of different sources. The decision which source to choose in a particular area of the image will be based on a comparison of the confidence measures and on the constraint that different seeds may not be connected (and thus merged into a big region) by a seed from another source. Again we choose a priority queue to decide which region should be processed next. The algorithm may be summarized as follows:

1. For each seed finding method:
   Find candidate regions, calculate confidence and put them into the priority queue.
2. While (Queue is not empty.)
   2a) Get next region from queue.
   2b) If (Candidate region will not connect distinct regions that are already in the final seed set.) then
   Put region into the final seed set.

We use three classes of basic seed selection methods: region based, edge based and topological:

1. Since we want the regions to be homogenous we can specify an absolute threshold on the image gradients. We estimate it using a result from Pratt [7] who points out that in many images the histogram of the gradient magnitude approximately shows an exponential (Laplacian) distribution many images except for a peak at low gradients - see fig. 3. This peak corresponds to homogenous regions in the image and we can choose the gradient at the transition point between the peek and the Laplacian line as a threshold for homogeneity.
   Seeds resulting from thresholding of gradient magnitudes are usually reliable if they are relatively large. A typical example for seeds chosen this way is shown on the left in figure 4.
2. Although edge detectors tend to miss important borders (as was discussed above, fig. 1), we may still use them to identify seed regions where we are confident that no border was missed. Experiments show that seeds obtained this way are typically very good if they are relatively small and have high contrast at their borders. Thus those seeds ideally complement the seeds resulting from 1. Typical results are shown on the right in figure 4.
3. According to the original watershed algorithm [6] we can also choose local minima (valleys) of the gradient magnitude as seeds. Although these valleys in general are noisy and result in oversegmentation we may use them in areas were no other seeds have been found.

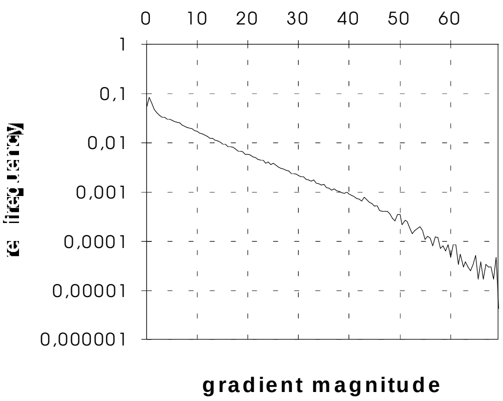

Fig. 3:
Logarithmic plot of a typical histogram of gradient magnitudes showing a Laplacian distribution except for a peak at low gradients due to homogeneous regions. (The noise at high gradients results from sampling effects for low histogram counts.)
The transition between the peak and the Laplacian line usually lies between 3 and 4 (for images with 256 graylevels).

Preprocessing and postprocessing

The issues of pre- and postprocessing shall be discussed only briefly. As we require small detail to be preserved we can not use a simple smoothing algorithm during preprocessing since all detail would be smoothed away. Instead we have chosen an adaptive smoothing algorithm according to [10] which keeps high-contrast detail while removing noise and low-contrast structures.

A conventional split-and-merge procedure can be used for postprocessing to identify regions, that have mistakenly been splitted or merged during seed selection, and to correct this errors.

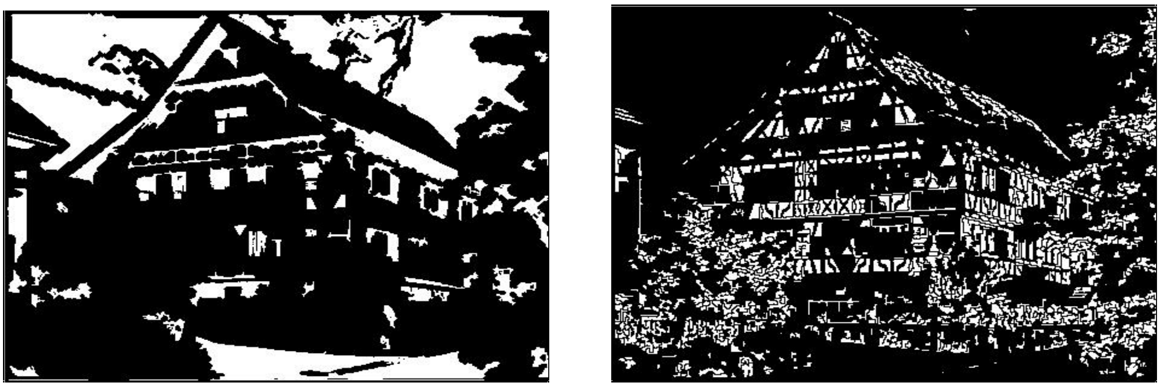

Fig. 4: Reliable seeds obtained from different sources (left: homogenous regions, right: small regions in inverted edge image, original: see upper left of fig. 5)

5. Experimental evaluation

The experiments to verify the proposed framework concentrate on the following questions:

1. Does it perform well on different images without the need to adjust parameters ?
2. Are regions wrongly connected or split ?
3. Is significant detail preserved ?
4. How accurate are the edges located ?

The algorithms have been evaluated on about 40 different outdoor photographs with resolutions between 300 and 800 pixels in each direction. Given identical parameter settings during the tests, the segmentation quality was comparable on all images, as is illustrated on three examples in fig. 5.

Small detail is preserved very well. On the examples this can be seen especially in the half timbered structures and in the branches of the trees in the lower image. This results from careful seed selection and from our use of foveation. Fig. 7 illustrates that the original resolution is clearly not sufficient to correctly locate the borders between regions.

A few areas in the images suffer from oversegmentation, e.g. the roofs of the buildings and the mast of the ship. This is due to two different effects. Firstly, large, medium textured regions may be not quite homogeneous enough to result in one single seed. The same holds for large variations in shading. Secondly, long narrow structures may not be recognized as single regions by the seed finding algorithms. Hence, improvement of those algorithms is still needed.

In the image of the ship one recognizes that the chimney is merged with a part of the sea, since the contrast between these regions is very low. This could be improved using color as an additional cue because the chimney is originally black, whereas the sea is dark blue. However, such undersegmentation occurs much less frequently then the oversegmentation as discussed above.

Fig. 6 illustrates the accuracy of edge localization. The estimated region borders are overlaid over the original image. Visual inspection shows a very good match between these borders and the image. For a more objective analysis we compared these borders with the edges obtained by Chen and Castans edge detector [8] which has been proven to optimally localize step edges. Most borders coincide with the optimal edges or lie within 1 pixel distance. The medium distance between the borders and the optimal edges is 0.6, indicating that about half of the border pixels match exactly with an optimal edge, whereas most of the rest lie at a distance of just 1 pixel.

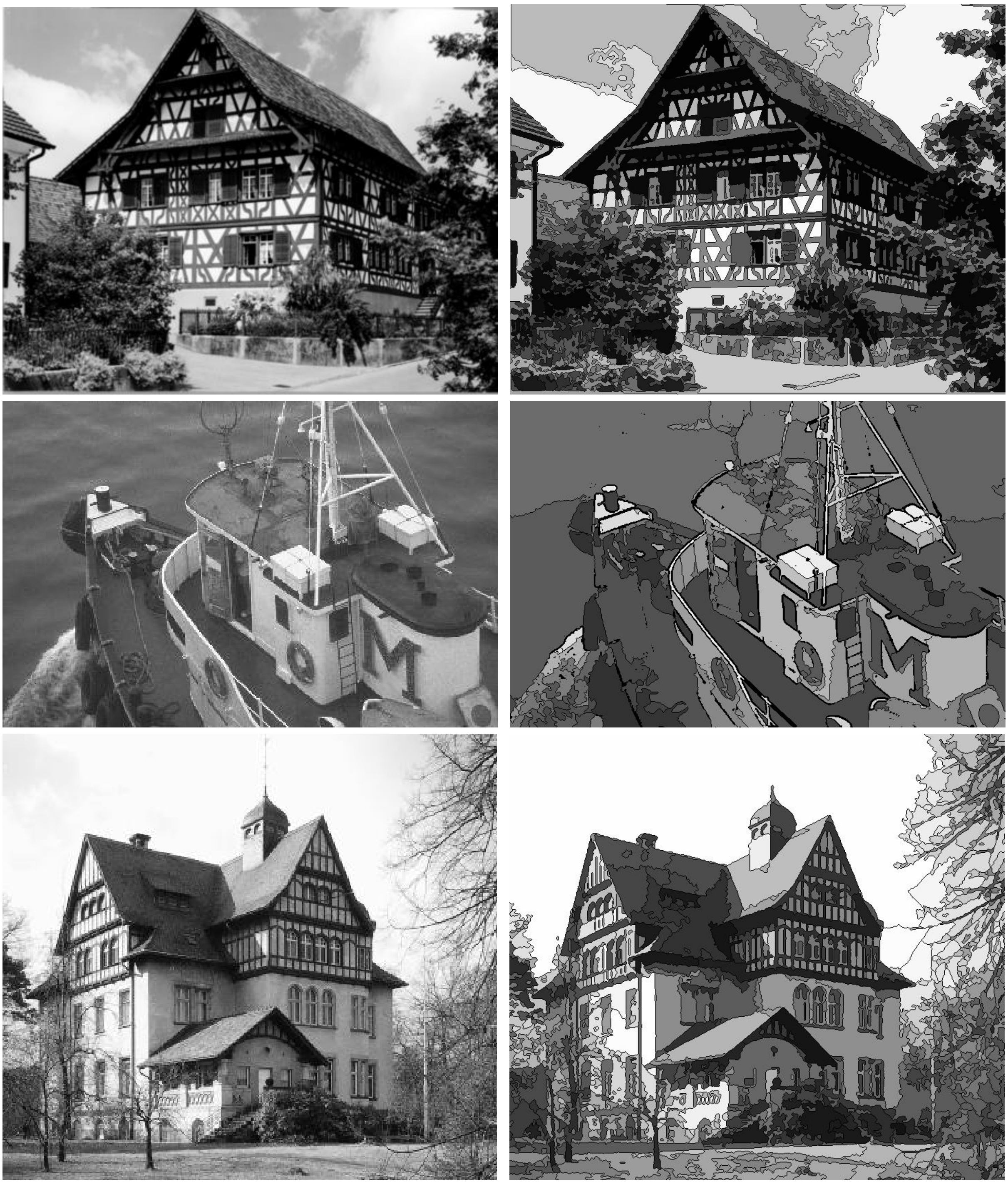

Fig. 5: Examples for primary segmentation using the proposed framework. (Left: Original image, Right: resulting segmentation) All images were processed with identical parameter settings. (Number of resulting regions from top to bottom: 2804, 385, 3874 respectively)
(middle image: from [4], available as http://wwwradig.informatik.tu-muenchen.de/horus/horus1.html)

Fig. 6:
Accuracy of edge localization: Region borders overlaid over the original image
(Detail from fig. 5 top)

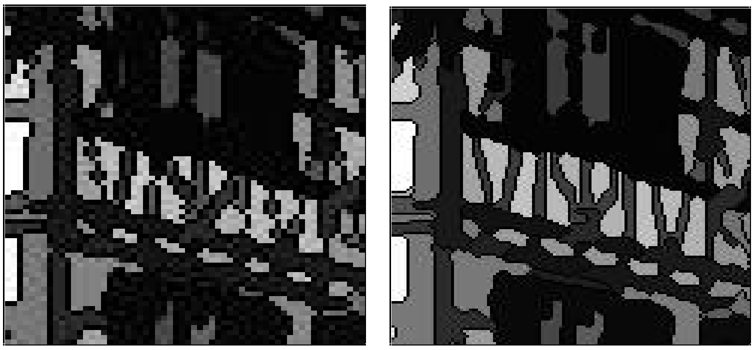

Fig. 7:
Comparison of segmentation results using original resolution (left) and foveation (right)
(Detail from fig. 5 top right)

6. Conclusions

In this paper a generalized framework for primary segmentation has been proposed and implemented. The strength of this framework lies in the possibility to smoothly integrate a variety of segmentation algorithms. Experiments have shown very promising results with respect to preservation of detail and robustness on various different images. Seed selection and foveation have been identified as key issues for further improvement of the algorithms.
Another direction of research lies in the implementation of the high-level, model based part of our framework. Above all we have to identify attributes that should be measured on a primarily segmented image to guide the high-level algorithms. These may include border attributes like straightness, position and direction, photometric and geometric attributes of the regions, invariants and texture. These attributes have to be organized in a way that they can index a large model database and initialize appropriate models.
The possibilities of feedback and iteration between the different stages of the framework constitute a third very interesting field of further investigation. We believe that higher level feedback could especially be useful during seed selection.

7. Literature

[1] M. Kass, A. Witkin, D. Terzopoulos: „Snakes: Active Contour Models", Proc. 1st Intl. Conf. on Computer Vision 1987, pp. 259-269
[2] A. R. Pope: „Model-Based Object Recognition: A Survey of Recent Research", Univ. of British Columbia, Dept. of Computer Science, Techn. Report CS-TR 94-04
[3] D. Marr: „Vision", Freeman, 1982
[4] S. Lanser: „Kantenorientierte Farbsegmentation im CIE-Lab Raum", in: Mustererkennung '93, 15. DAGM Symposium, Springer 1993
[5] R. Adams, L. Bischof: „Seeded Region Growing", PAMI 16(6), pp. 641-647, 1994
[6] S. Beucher: „Segmentation Tools in Mathematical Morphology", in: C.H. Chen, L.F. Pau, P.S.P. Wang: „Handbook of Pattern Recognition and Computer Vision", World Scientific Pub., 1994
[7] W.K. Pratt: „Digital Image Processing", John Wiley & Sons, 1991
[8] J. Shen, S. Castan: „An Optimal Linear Operator for Step Edge Detection", CVGIP: Graphical Models and Image Processing, vol. 54, no. 2, pp. 112-133, 1992
[9] P. Burt: „Smart Sensing in Machine Vision", in: H. Freeman: Machine Vision, Academic Press 1988
[10] P. Perona, J. Malik: „Scale Space and Edge Detection using Anisotropic Diffusion", PAMI, vol. 12, p. 629, 1990