Experimental Performance Characterization of Low Level Vision Components in Vision Systems - Theory and Application

Guido Appenzeller
IPR, University of Karlsruhe, Kaiserstr. 12,
76128 Karlsruhe, Germany
EMail: appenz@imag.fr

James L. Crowley
IMAG-PRIMA, 46 Ave Félix Viallet
38031 Grenoble, France
EMail: jlc@imag.fr

Abstract

Adaptive filters or image enhancement techniques have repeatedly been suggested to make feature extraction more robust. Few comparative analysis exists and even the existing ones concern only the effect of the filter on the image but not its effect on a feature extraction process. In this paper we present an experimental approach for the evaluation of low-level vision system components in a system framework and use it on adaptive filters. As we are interested in the overall performance of the system the interactions between the algorithms and other parts of the system have to be understood. We design a general evaluation strategy and derive from it an algorithm to measure the results. In order to have objective and reproducible results the automatic control of the parameters of the vision system is necessary for which we present several techniques. The results show that none of the tested techniques performs better than linear filters or the Canny edge detector.

1 Introduction

The derivative based detection of segments becomes unstable in the presence of noise. This can be explained on a theoretical level by the fact that the detections of extrema or zero crossings in the derivatives is an ill posed problem. Several approaches exist to make the segment detection process robust. They are most notably smoothed edge detection operators [1], tracking the segments at several scales [2], statistical methods as local hough transforms or line support regions and the use of image enhancement techniques or adaptive filtering before the segment extraction takes place. Adaptive filters seem particularly interesting because they claim to offer a tool that is independent of the type of feature that is used [3].

A great variety of image enhancement techniques have been suggested in the past years. In this diverse field comparative analysis is important to identify techniques that are useful for computer vision applications. Yet very few comparative evaluations or reports about their successful usage exist. Most reports only include example pictures to characterize the performance of the technique. Mastin [4] uses 30 reviewers to judge the images. Migeon [5] gives grades for certain characteristics (e.g. edge enhancement, corner preservation) but presents no experimental data. While both introduce some objectiveness but the results are not reproducible.

Wu et al. [6] define a number of statistical measures to characterize adaptive techniques and give quantitative results. The results however tell us nothing about the effect of the adaptive techniques on the quality of segment extraction.

The work described here concerns the development of a method to experimentally evaluate the performance of low level vision components in general and adaptive filters in particular. It has already been recognized [7] that this can not be done by analyzing the the component alone but requires system engineering methodology as the interactions of the component we want to test with other system components has to be investigated and parameter control strategies have to be found. The approach suggested by Ramesh et al. [8] to find a model of the steps of the vision system and derive from it a mapping between disturbances in the input data to disturbances in the output data did not seem applicable as models for the highly nonlinear filters are not known and seem difficult to develop. We have however implemented a related technique for parameter control.

In section 2 we will first develop the overall evaluation strategy. Section 4 we will present a measure for the quality of the segment extraction and an algorithm to measure it. The problem of the control of the free parameters of the vision system is treated in section 5.2 which is followed by the experimental results and the conclusion.

2 Evaluation Strategy

2.1 General Requirements

One main goal of performance characterization is to describe algorithms in a way that allows a vision engineer to select the appropriate components for his application. In experimental evaluation the results are obtained by measuring properties of the output data of the system. We believe a good experimental performance characterization should meet the following goals:

- **Valid Criteria Definition.** For each criterion we use to characterize the performance of the algorithm we have to define why we consider it important for the performance of the algorithm and how exactly we measure it.

- **Comparable.** The results of the experiments should be directly comparable to new experiments. Only relative comparisons of algorithms against each other would make it necessary to rerun the old experiments to for new algorithms. The only way to guarantee comparability is quantitative evaluation.

- **Reproducible.** All information that is necessary to reproduce the results should be given. For simple algorithms this means specifying the parameter settings, for complex systems this can mean that the system has to be made parameter free by using automatic parameter control strategies.

- **Error Estimation.** Any measurement should include at least a rough estimation of the precision of the measurement. Measurements can have large errors due to statistical or methodical errors (e.g. random noise, unprecise parameter adjustment). These errors can have an effect that is more significant than the choice of the algorithm. In order to obtain valid results we should at least show that these errors are smaller than the measured effect.

So far the performance of adaptive filters has only been characterized in terms of their effect on images. Image enhancement was measured by how close the restored image was to original one. This was either judged by a human observer or measured by, for example, a mean square error criterion. In computer vision applications one is however not interested in the images itself but the symbolic representation that is derived from them and if it contains the information that is necessary to perform the desired task.

Communications theory (e.g. [9]) defines noise as everything that obscures the information in the signal. For a segment extraction process this definition would include camera and acquisition noise but also textures or different lightning conditions. The goal of adaptive filtering in this case would not mean getting as close to the original image as possible but to remove everything that hinders the further feature extraction steps. From this point of view a mean square error criterion between the original and the restored image does not make much sense.

One main goal of performance characterization is to provide a scientist or engineer with information by which he is able to select algorithms that are suitable for his application or system. In order to provide this information we have to characterize adaptive filters by their effect on the overall performance of the vision system. The most straightforward approach would now be to evaluate directly the performance of the system for a given task. For a complex system such an evaluation will be extremely time consuming and the results will be difficult to transfer to other applications. It is often a better idea to evaluate the system at an intermediate level. Therefore the basic evaluation strategy will be:

1. Choose an intermediate (or the final) result of the vision system as the result we will use to judge the performance of the system component we want to test.

2. Analyze what components or parameters interact with the components we are testing and find ways to adapt them.

3. Define a set of quantitative criteria that characterizes the quality of the intermediate result.

4. Design an algorithm to measure these criteria.

Finally we should make a rough estimation how precise our measurements are and compare this to the differences between the tested components.

3 Evaluation in a System Framework

A vision systems classically consists of a number of steps. Each step operates as a filter and passes only that part of the data up to higher levels that it considers important. Which each step the amount of data decreases while the level of abstraction of the data increases.

For each step different algorithms can be used. If we want to obtain optimal results with the system we have to choose the right combination of algorithms. Additionally each step has a number of free parameters (e.g. thresholds, smoothing coefficients) that have to be adjusted. Their optimal setting will depend on the algorithms as well as on the parameters of the other steps.

The vision system we are interested in is a segment based image understanding system, depicted in figure 1. After image acquisition a filtering stage is used to reduce the noise in the image. As filters we will try adaptive techniques as well as linear filters. A gradient operator

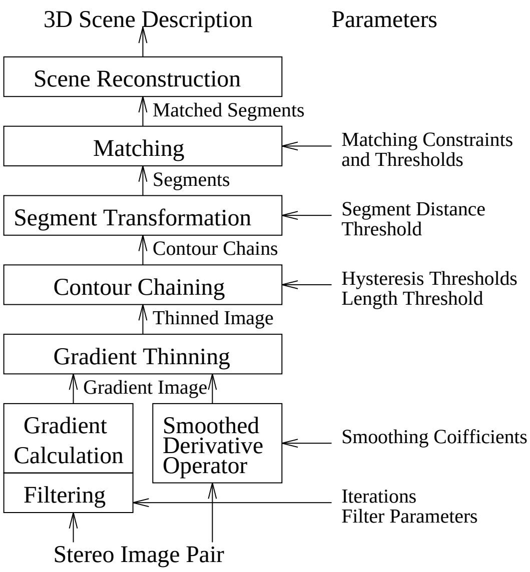

Figure 1: The Vision System

is used to calculate the absolute value and direction of the gradient. Instead of the separate filtering and gradient calculation steps a smoothing derivative operator as Canny's operator can be used. The gradient image is thinned using gradient maxima thinning and contour chains are extracted using hysteresis thresholding [1]. Using a recursive splitting algorithm the chains are transformed into segments. The segments are matched and the 3D scene description is calculated.

Before we can start with the evaluation of the vision system we have to decide after which step we want to evaluate, define a quality measure for its results and understand the interaction of its components and parameters.

3.1 Choosing the Evaluation Level

For evaluation of the vision system the results of any of its steps can be used. The decision for the level at which we want to evaluate the performance of a component faces us with a number of tradeoff problems.

- Evaluation of vision systems is a task dependent problem. The final evaluation of any vision system is how well it is able to perform the task it was designed for. We are however looking for a more general measure that is usable for a broader range of applications. Generally the output of higher level steps of a vision system (e.g. object recognition, scene reconstruction) are more task dependent the components and parameters of the steps below have been adjusted to perform optimally for the given task while the intermediate results of lower level steps (e.g. features, groups) are more generally applicable. When choosing the best level at which we want to evaluate the vision system we are faced with a tradeoff between generality and a less valid evaluation criterion.

- The higher the level is at which we evaluate the longer the calculation for each measurement will take. For each measurement all the steps below the evaluation level have to calculated and their parameters have to be optimized. While the increase in calculation time due to the computation of more steps will be manageable the increase of the parameter space can lead to a dramatic increase of calculation time for some parameter control strategies (e.g. exponential for exhaustive search).

- For many evaluation methods ground truth about the used data is required. If the disturbances and image contents can be modeled sufficiently good this is most conveniently done with artificial images. While this is easy for low levels (e.g. contour level) it is increasingly difficult at higher levels (e.g. for the 3D scene level one would have to generating random scenes and render them using ray-tracers).

We have decided to evaluate our vision system at the segment level. The results we obtain are relatively general as segment extraction systems today are often similar while higher up steps (especially matching) still use very different techniques. The generation of example images is relatively easy and calculation time is still manageable.

The evaluation of the segments now requires a definition of what we consider good segments. We will define our criteria in section 4.

3.2 Interaction of System Components

As the only component we will exchange are the adaptive filters only their interaction with other components has to be considered. The main effect of the adaptive filters is to reduce the noise in the image data. The only other component that is able to smooth is the edge detector. Most convolution based edge detectors can be seen as the derivative of a smoothing function (e.g. a Gaussian for the Canny Operator or a mean filter of size two for the Sobel Operator). As we only want to study the effect of the smoothing of the adaptive filters we will use the least smoothing edge detector, the simple 2x2 derivative operator (-1 1).

When we exchange the adaptive filters the optimal parameter settings for the other steps will be different. To take this into account we will optimize the parameters for each filter and each signal to noise level. This is discussed in section 5.

4 The Evaluation Criteria

We need two different evaluation criteria for different purposes.

First we need a set of criteria to characterize the performance of the filters in the experiments. These criteria should include as much information that might be significant for later processing steps as possible and should give the vision engineer clear guidelines which algorithms to use.

Second we need a criterion that is used for parameter control. This criterion must be a scalar function that is minimized to find the best parameter setting. As this criterion has to be calculated for each iteration of the parameter control cycle its calculation has to be fast.

We will characterize every filter by two characteristics, the robustness of the segment extraction and its scale space position.

4.1 The Robustness

The robustness measures how good the vision system is able to identify segments. It does not measure how the quality of the segments is (e.g. how precise their location is).

The most simple implementation of this measure would be to judge for every detected segment if it corresponds to a real segment in the image. Such a definition however would require thresholds (maximal angular difference, maximal length difference, breaks etc.).

A simpler definition is to verify every detected pixel of a every detected segment if it belongs to a real segment or not. We consider all pixels that are within a two pixel radius of real segments as being correctly identified. To avoid that edge detectors that produce multiple responses to an edge receive good ratings in a second phase multiple responses are detected along a line perpendicular to the segment and a penalty for each multiple response is introduced.

All pixels that are not within a two pixel radius are considered to be due to noise (false alarms). If we denote the real number of segment pixels by N_{Truth}, the number of correctly detected segment pixels $N_{Segment}$ and the number of false alarms to N_{False} the overall error function is:

$$Q = \frac{N_{Truth} - N_{Segment} + N_{False}}{N_{Truth}} \tag{1}$$

This error function is zero if all image pixels are correctly identified. It can always kept below 1 by setting the detection thresholds to values that will not detect any pixels. For parameter control we try to minimize this function.

We define the robustness of the vision system as the signal to noise ratio where the error rate is ten percent.

4.2 Scale Space Position

Robustness alone however is not a sufficient criterion to characterize the performance of the adaptive filters. Figure 2 shows the error rate for the Deriche Filter for different smoothing parameters. As we decrease the parameter α, we increase the smoothing of the filter and with it the robustness. The strong smoothing however will destroy small details and round off corners which for many applications is harmful.

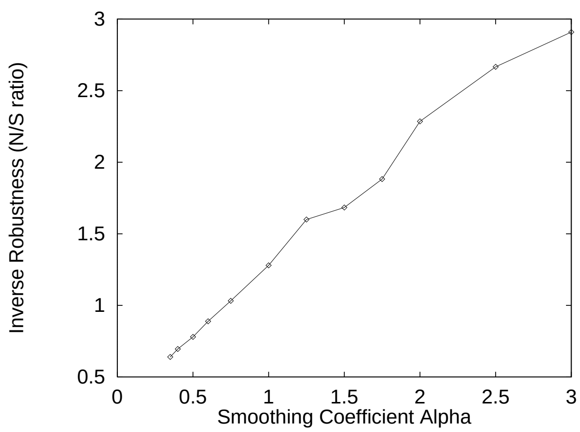

Figure 2: The Inverse Robustness for the Deriche Filter against the Smoothing Parameter

Locally adaptive filters show a similar behavior [10]. For linear filters this behavior is easy to understand. Gauss and Deriche filters both serve as low pass filters in the frequency spectrum of the image. The effect of adaptive filters on the frequency spectrum can not be easily characterized. We will therefore use a different approach and measure the ability of a filter to detect small details. The theoretical framework for this measure is provided by the scale space theory by Witkin [2].

We will characterize each adaptive filter by its approximate position on the z-axis in a reference scale space. It is clear that the image that is filtered with the adaptive filter is not equal to the image at its scale space position. We can however devise a test to compare the detail preservation ability of one filter against a different filter that does have scale space. This test will use a scalable feature. As we are interested in segments this feature will consist of edges.

If we use an energy preserving FIR filter (e.g. a Gaussian) for our scale space all features that consist of one or two edges will never disappear [10]. With three edges we can however construct a feature that disappears somewhere in the scale space. It is depicted in figure 3.

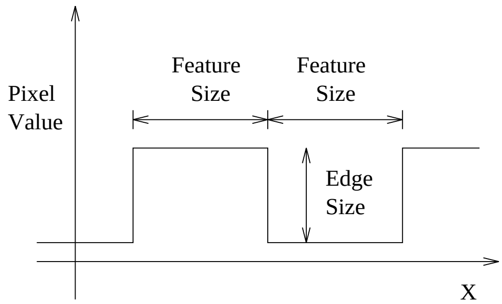

Figure 3: The feature used for scale space measurement.

We reduce the size of the feature with sub-pixel accuracy until it disappears (the number of gradient zero crossings becomes less than 3). The scale space location of the filter is the position at which the feature disappears in the reference scale space.

5 Parameter Control

The automatic control of parameters is necessary for performance characterization. If the parameters are not controlled or controlled manually the effect of the bad parameter tuning can be stronger than the effect of the different tested components.

It could be thought that while static parameter settings do not give optimal results they provide an even field for testing the filters. Figure 4 shows the average response of the edge detector (i.e. the gradient value) with and without Median filtering for a range of signal to noise ratios. The plot shows that while the average gradient value increases with noise if no filtering is applied it decreases if median filtering is used. To extract the contour chains from the image only those pixels of the gradient image are take into account that are above a certain threshold. To receive optimal results this threshold obviously had to be different for no filtering and the Median filter and it has to change with the varying noise.

In computer vision today parameters are often controlled by a human operator. To do this for each filter and noise level is not only very time consuming but the results obtained can hardly be called scientific as they are not reproducible.

The vision system we use has a large number of parameters. However some of them are only weakly correlated to other parameters and can therefore be treated separately while for others optimal settings can be derived for the theory of the system. Our first step therefor is to analyze the parameters of the system.

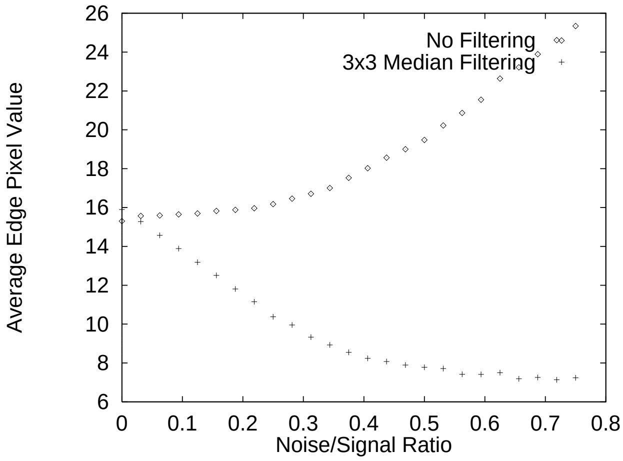

Figure 4: Minimum Values of Canny Response along a Edge Part

5.1 Parameters of our System

The parameter of the vision system are shown in figure 1. As we are evaluating at segment level the parameters of the stages above (matching and reconstruction) do not have to be controlled.

There is one parameter for the contour to segment transformation that specifies how big the distance of a contour chain of a segment from the location of the segment may be before the segment is split into two segments. We control this parameter by using ground truth to measure the deviation of the contour chain from the real position of the segment and to set the parameter accordingly.

The two hysteresis thresholds and the length threshold have a significant impact on the performance of the system. They have to be adjusted dynamically to the filter used and the noise in the image. We have tested several strategies to control them as described below.

The smoothing coefficient of smoothed gradient operators, in our case the Canny/Deriche detector, specifies a tradeoff between the signal to noise ratio of the response, the localization of the edge and the detail resolution. We have measured the Deriche detectors performance for a number of discrete settings from which other values can be interpolated.

Two of the adaptive filters have a free parameter, the anisotropic diffusion technique and the Sigma Nearest Neighbor (SIG) filter. Both parameters are relative to the grey values in the image and specify up to what edge heights edges should be preserved and when smoothing should occur. We have made experiments on the setting of this parameter for both filters. For anisotropic diffusion the effect of this parameter changes little if the images are distorted by blur or noise. The edge is restored perfectly. For large noise the edge restoration process becomes unstable. For the SIG filter the parameter works well if the signal to noise ratio is at least 3.0. If the noise is increased the parameter has either be set to a value that will no longer smooth the noise or the filter will smooth the edge itself. The presence of blur makes the filter even more unstable. Worst of all, local defects in edges will propagate as the filter is iterated eventually smoothing the complete edge. Due to this problems the SIG filter was excluded from the experiments.

All of the filters can be used in an iterative way. The behavior of the linear and the adaptive techniques however is different. For adaptive techniques the image converges against a final image. After a certain number of iterations the image becomes stable or oscillates between similar states. Linear filters change the image with every further iteration until quantization effects halt the process. For this reason two different strategies are used. For the linear techniques the filters have been tested for each iteration count separately. For the adaptive filters we first determined after how many iterations the performance of the filters did not change anymore. This number of iterations was then used for the experiments.

5.2 Methods for Parameter Control

We have tested several methods to control the two hysteresis thresholds and the length threshold.

- **Exhaustive Search.** Searching the complete parameter space for the setting that minimizes the error function necessarily finds the optimal parameter setting. The problem is that this method is computationally very expensive (about one day to find the optimal parameter settings). It however provides us with some important insight on the geometry of the error function in the parameter space.

- **Model Based Tuning.** Following an idea of Ramesh [11] we have constructed a model of the vision process. Unlike the author we do not model the edges, noise and edge detector but only the chaining process itself and use statistics collected from the edge image as initial parameters. Results have been published in [12]. While the error is small it is in the order of magnitude of the difference between different filters.

- **Gradient Descent.** Gradient descend in the parameter space is difficult for several reasons. First the error of the statistical measurements creates many local minima. Second the evaluation function has many constant regions for too small or too large thresholds. Finally the calculation time is very slow. However, if initial parameters are carefully selected it is applicable. The error of the quality function against exhaustive search is very small (< 0.05). A detailed evaluation can be found in [10].

For the measurements gradient descent parameter control is used.

6 Experiments and Discussion

6.1 The Experimental Setup

The adaptive filters that were implemented are the Symmetric Nearest Neighbor (SNN) filter [13], the K-Nearest Neighbor (KNN) filter [14], the Sigma Nearest Neighbor (SIG) filter [15], the Nagao filter [16], Lee's filter [17] and Anisotropic Diffusion [18]. The SIG and diffusion techniques have a single parameter while the other techniques are parameter free. The SNN, KNN and Nagao techniques have been combined with Median, Mean and Lee filter as in [6]. SNN, KNN and Lee's techniques have been tested in three sizes (3x3, 5x5 and 7x7) for each combination. Diffusion has been tested for a number of possible settings for its parameter k.

As separate linear filters mean filtering (size 3x3, 5x5 and 7x7 between 1 and 10 iterations) and Binomials (integer approximations of Gaussians, same configurations as mean) have been tested. As a reference Sobel's operator and Deriche [19] implementation of the Canny detector is included, the latter with several different smoothing coefficients.

The name of the adaptive and linear filters is followed by the window size and the number of iterations (e.g. SNN-Mean3-15 is a 3x3 filter with 15 iterations).

Overall about 90 different filter configurations have been tested. The experiments where made on artificial edges of a single height that where distorted by zero mean white Gaussian noise. For each filter for 64 signal to noise level and two levels of blur the parameters were tuned and a number of properties were measured (see figure 6). Calculation of the experiments took several weeks on a cluster of workstations.

6.2 The Experimental results

Figure 5 shows a combined plot of the robustness and the detail resolution of the most successful filters of the test. The X-Axis shows the robustness of the filter and the Y-Axis the position in scale space characterized by the smallest feature size where the feature was still being detected. Due to pixel quantization the detection of features of a size less than one is not possible. Each filter is characterized by a point in the diagram. Filters for which a parameter is varied are connected by lines.

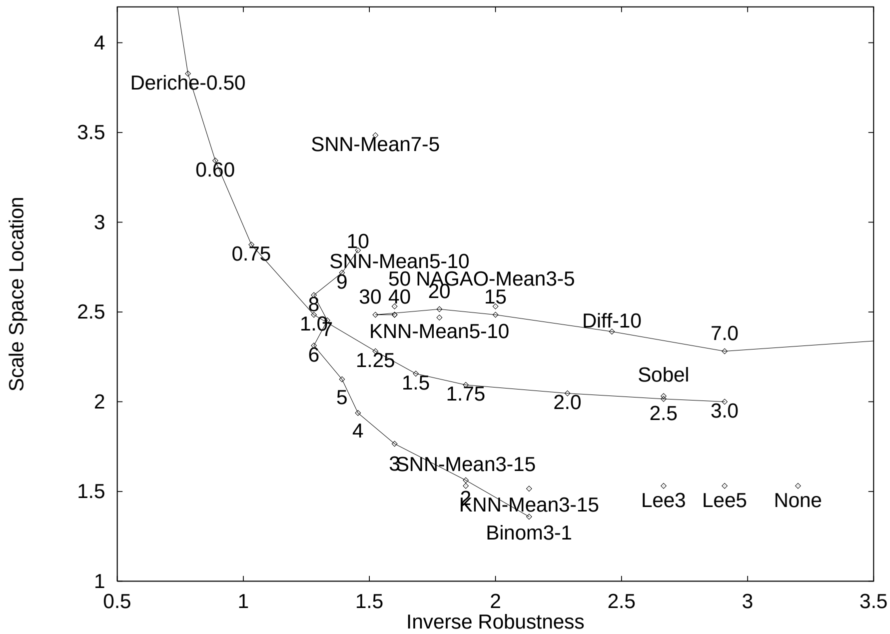

Figure 5: The Robustness - Scale Space Results

For the Deriche detector the tradeoff between robustness and detail resolution is clearly visible. The result for the adaptive techniques are somewhat surprising. None of them performs better than either the Deriche detector or the use of Binomials for smoothing. The Deriche detector offers the best robustness while the Binomials offer a better detail resolution for big signal to noise ratios. This is due to the fact that our implementation of the Deriche detector is symmetric to the pixel center while the edge detector that uses Binomials is symmetric to the pixel boundary. As a result the Deriche detector produces two maximal responses to an edge which is on a pixel boundary while the Binomial based detector produces only one.

For the SNN, KNN and Nagao techniques the combination with the Lee filter always provided the weakest smoothing with the Median and the Mean combinations following. The tradeoff for the Mean filter was usually best so only its result is given.

We now discuss each of the the different techniques briefly.

Filter Name	Robu	Scal	EVl	EdVr	Loc	Zero	Low	Tear	Fork
None	3.200	1.531	16.7	0.232	0.184	0.008	0.005	0.020	0.054
SNN-Mean3-15	1.882	1.531	13.2	0.077	0.363	0.005	0.062	0.027	0.075
KNN-Mean3-15	2.133	1.516	13.4	0.099	0.293	0.003	0.041	0.023	0.071
Nagao-Mean3-5	2.000	2.531	13.8	0.095	0.401	0.004	0.033	0.017	0.130
Mean3-5	1.333	2.438	3.5	0.012	0.616	0.023	0.003	0.015	0.053
Binom3-8	1.280	2.594	3.3	0.010	0.619	0.024	0.004	0.023	0.053
Deriche-0.50	0.780	3.828	4.1	0.010	—	0.018	0.006	0.006	0.018

Robu: The Robustness of the filter.
Scale: The smallest feature size the filter can still detect.
EVl: The average value of the gradient maximum of the edge.
EdVr: The standard deviation of the gradient maximum of the edge.
Loc: The standard deviation of the position of the edge pixel.
Zero: The probability of no edge pixel being found.
Low: The probability of edge pixels being discarded because they are below the lower hysteresis threshold.
Tear: The probability of a segment being torn perpendicular to its direction.
Fork: The probability of forks in a contour.

Figure 6: Selected Experimental Results for some Filters

6.2.1 SNN, KNN and Nagao Filter

The SNN filter provides the best robustness of the adaptive techniques with KNN and Nagao following. The measurements show that the 3x3 SNN-Mean filter is even slightly better than the Binomial filters the difference however is smaller than the error interval for the measurements. All filters are able to restore degraded edges perfectly while the noise is low. As the noise rises his ability to preserve edges becomes unstable and the filters introduce edges in areas where none exist.

6.2.2 Lee's Filter

Lee's filter only slightly increases the robustness of the segment extraction but as the scale space location equally only changes very little it might be interesting for low-noise applications. The reason why it is not effective is that its smoothing effect depends on the variance of the data in its window which is always high on edges. The edge reason therefore is only hardly changed.

6.2.3 Anisotropic Diffusion

Anisotropic diffusion is conceptually similar to the SIG filter. The edges for the conducted tests had a hight that caused them to be smoothed away for a parameter above 16. It can be seen in the plot that the best results are obtained for parameter settings above 30. For this setting anisotropic diffusion works almost like a mean filter.

A second interesting observation is that the scale space location of the diffusion filtering is relatively low for its high number of iterations (20) and that it changes little as the value of the parameter is varied. The diffusion scale space as proposed in [18] is fundamentally different to the scale space as proposed by [2].

6.2.4 Mean and Binomial Filters

These integer approximations of Gaussian provide the best results for low noise. As the number of iterations increases the robustness first increases, reaches a peak at about 8 iterations and then decreases. While part of this effect can be explained with the quantization noise than is generated from the integer implementation the additional data that has been collected suggests that the main reason are multiple responses and subsequent thinning errors due to multiple responses of the detector. Binomials with bigger sizes had very similar behaviors as their smaller counterparts of which they can be constructed, excluding quantization effects, by convolving the smaller ones with themselves. Mean filters of a size of more than 3x3 as well as a single iteration of the 3x3 filter had severe problems with multiple responses making them unusable.

6.2.5 Simple Detectors

The simple derivative operator has, as it can be expected the worst robustness and the best detail resolution of the field (the slightly better detail resolution of the Binomial filter is only due to the measurement technique used). The Sobel detector performs very similar to a Deriche detector with a smoothing coefficient of 2.5.

6.2.6 The Deriche Detector

Deriches recursive implementation of Canny's detector offers the best overall robustness of the field. Its smoothing parameter makes it adjustable to the amount of noise in the image. The recursive implementation additionally gives it the fastest calculation time except for the single iteration Binomial, mean and Sobels filters.

6.3 Discussion

From the filter characteristics that were measured additionally can give us a certain insight about why the adaptive techniques fail.

Adaptive and linear techniques (the latter include the Deriche detector) show a very different behavior. The adaptive techniques weaken the strength of the gradient only slightly, however are not able to control its variance effectively. The techniques typically fail because the edge itself becomes too disturbed and can no longer be discriminated from the noise as the high number of low thresholds indicate.

The linear techniques reduce the value of the gradient dramatically but the relative variance of it is lowered to an even larger extent. Errors due to the gradient values of noise and edge becoming similar are not a problem. As the smoothing increases the response to a single edge is becoming broader ad the number of pixels with an almost equal gradient value increases. It can be seen in the data by the number of zero pixels (which are mainly thinning errors due to an unstable gradient direction) increases and the localization becomes worse.

As it could have been expected, the smallest feature the filters are still able to detect is half their window size. The adaptive behavior is not changing this. However is the scale space position independent of the number of iterations. This is not the case for the Median and Lee based SNN, KNN and Nagao filters. Here the size of the smallest feature can be significantly smaller than the size of the filter window. The anisotropic diffusion technique has a region of influence of 5x5 as the the calculation of the gradient for each of the pixels in the 3x3 window enlarges its size by one.

The measured deviation of the detected edge from its real location was always relatively small. This indicates that bad localization is already a problem for the contour extraction process itself and does not have to be considered separately.

We believe one of the reasons why the adaptive techniques fail is that they do not take spatial information into account. All pixels, independently from their distance to the central pixel are considered equal. By introducing weights that depend on the distance this could be changed. Such techniques have been suggested by [20] and Kimia et al. [21] and first results look promising.

Another argument in favor of the adaptive techniques is that the white zero mean Gaussian noise is is a very rough model for the type of noise that is contained in images. For other noise types nonlinear filters are known to work better than linear techniques (e.g. Median filters for impulsive noise). Additionally the technique that performed best here, the Deriche detector, was designed using Canny's criteria which include the assumption that the noise is white additive Gaussian.

7 Conclusion

We have presented a method for the performance characterization of low-level vision algorithms in vision systems. As theoretical models for the algorithms we wanted to test did not exist we chose experimental evaluation.

The goal of evaluating the effect of the algorithm on the overall output of the system poses a number of new problems. Due to the complexity of the system task based evaluation is often not a good choice. Intermediate results are sometimes better for evaluation. To guarantee that the measured effect is really due to the tested component is a difficult problem. System engineering methodology and strategies for the automatic control of parameters are necessary. We have implemented several methods for the automatic control of parameters and have shown that valid results can be obtained using gradient descend in the parameter space.

We have tested a group of image enhancement techniques as a preprocessing stage for a segment extraction system and compared them against linear filters and the Canny/Deriche detector. Parameter control was performed using gradient descend in the parameter space. Evaluation criteria was the robustness of the extraction and the detail resolution of the technique.

None of the tested techniques performs better than the linear techniques or the integrated edge detector for white zero mean Gaussian noise. A possible reason for this are the neglect of spatial information.

References

[1] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.

[2] A. Witkin, "Scale space filtering," *Proc. Int. Joint Conf. Artificial Intell., held at Karsruhe, West Germany, 1983, published by Morgan-Kaufmann, Palo Alto, California*, 1983.

[3] P. S. Marc, J.-S. Chen, and G. Medioni, "Adaptive smoothing: A general tool for early vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, pp. 514–529, June 1991.

[4] G. A. Mastin, "Adaptive filters for digital image noise smoothing: An evaluation," *Computer Vision, Graphics, and Image Processing*, vol. 31, pp. 103–121, 1985.

[5] B. Migeon, "Pretraitement d'image par filtrage epsf," tech. rep., Direction de Recherches Etudes et Techniques INRIA, Jan. 1994.

[6] W.-Y. Wu, M.-J. J. Wang, and C.-M. Liu, "Performance evaluation of some noise reduction models," *CVGIP Graphical Models and Image Processing*, vol. 54, pp. 134–146, Mar. 1992.

[7] V. Ramesh, *Performance Characterization of Image Understanding Algorithms.* PhD thesis, University of Washington, 1995.

[8] V. Ramesh and R. M. Haralick, "Random perturbation models and performance charcterization in computer vision," in *Computer Vision and Pattern Recognition*, pp. 521–527, 1992.

[9] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering.* John Wiley and Sons, 1965.

[10] G. Appenzeller, "Image filtering for 3d reconstruction," Master's thesis, University of Karlsruhe, Sept. 1995.

[11] V. Ramesh and R. M. Haralick, "A methodology for automatic selection of iu algorithm tuning parameters," in *ARPA Image Understanding Workshop*, 1994.

[12] G. Appenzeller and J. L. Crowley, "Experimental evaluation of vision systems and automatic parameter control," in *Asian Conference of Computer Vision*, Dec. 1995.

[13] D. Harwood, M. Subbararao, H. Hakalahti, and L. Davis, "A new class of edge preserving smoothing filters," *Pattern Recongition Letters*, vol. 5, pp. 155–162, 1987.

[14] L. S. Davis and A. Rosenfeld, "Noise cleaning by iterated local averaging," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 8, pp. 705–710, 1978.

[15] J.-S. Lee, "Digital image smoothing and the sigma filter," *Computer Vision, Graphics, and Image Processing*, vol. 24, pp. 165–168, 1983.

[16] M. Nagao and T. Matsuyama, "Edge preserving smoothing," *Computer Graphics and Image Processing*, vol. 9, pp. 394–407, 1979.

[17] J.-S. Lee, "Digital image enhancement and noise filtering by use of local statistics," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 2, pp. 165–168, Mar. 1980.

[18] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 629–639, July 1990.

[19] R. Deriche, "Fast algorithms for low-level vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 78–87, Jan. 1990.

[20] Y. fai Wong, "A clustering filter for scale-space filtering and image restauration," in *Conference on Computer Vision and Pattern Recognition*, pp. 668–669, 1993.

[21] B. B. Kimia and K. Siddiqi, "Geometric heat equation and nonlinear diffusion of shapes and images," in *Conference on Computer Vision and Pattern Recognition*, pp. 113–120, 1994.