# Experimental Performance Characterization of Adaptive Filters

Guido Appenzeller
IPR, University of Karlsruhe, Kaiserstr. 12,
76128 Karlsruhe, Germany
EMail: G.Appenzeller@IEEE.org

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

## Abstract

*Adaptive filters and image enhancement techniques have repeatedly been suggested to make feature extraction more robust. Few comparative analysis exists between competing techniques and even the existing ones evaluate 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. Our method is applied to adaptive filters. By analyzing the system architecture we chose an evaluation level and derive the evaluation criteria and parameter control strategies. The results show that none of the tested techniques performs better than linear filters or the Canny edge detector.*

## 1. Introduction

In this diverse field of adaptive filters (or image enhancement techniques) 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 publications only include example pictures to characterize the performance of the technique. Mastin [9] uses 30 reviewers to judge the images. Migeon [10] gives grades for certain characteristics (e.g. edge enhancement, corner preservation) but presents no experimental data. Wu et al. [16] define a number of statistical measures to characterize adaptive techniques and give quantitative results. However, none of these results make it possible to estimate the effect of these filters on the performance of a vision system.

The work described here concerns the development of a method to experimentally evaluate how the choice of a low level vision component affects the quality of the final output of a vision system. It has already been recognized [13] that this can not be done by analyzing the component alone but requires system engineering methodology as the interactions of the component we want to test with other system components must be investigated and parameter control strategies must be found.

## 2. Evaluation Strategy

The performance of adaptive filters has typically been characterized in terms of their effect on images. Image enhancement was evaluated by measuring by how close the restored image was to the original one. This was either judged by a human observer or measured by, for example, a mean square error criterion. However, in computer vision applications one is not interested in the images itself but in the information that is contained in the image which is relevant to performing a vision task.

Communications theory defines noise as everything that obscures the information in the signal. For a segment extraction process this definition would include not only camera and acquisition noise but also textures or different lightning conditions. The goal of adaptive filtering here is not to get 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.

The system oriented approach to the evaluation of a vision system component we suggest is the following:

1. **Evaluation Level**. Select the stage of the vision system (e.g. image acquisition, edge detection, segment extraction) after which the results will be evaluated in order to judge the component. We do this in section 3.

2. **Evaluation Criteria**. Define a set of quantitative criteria that characterizes the quality of the result and design an algorithm to measure these criteria. We present these in section 4.

3. **System Control**. Analyze what components or parameters of the system interact with the component we are testing (e.g. stronger smoothing requires different thresholds) and find ways to adapt the parameters of the system accordingly. This is done in section

5.

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

## 3. Choosing the Evaluation Level

The vision system we are interested in is depicted in figure 1. After image acquisition, a filtering stage is used to reduce the noise in the image. As filters we will try both, adaptive image enhancement techniques and linear filters. A gradient operator 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 (e.g. Canny's operator) can be used. The gradient image is thinned using gradient maxima thinning and contour chains are extracted using hysteresis thresholding [3]. Using a recursive splitting algorithm the chains are transformed into segments. The segments are matched and the 3D scene description is calculated.

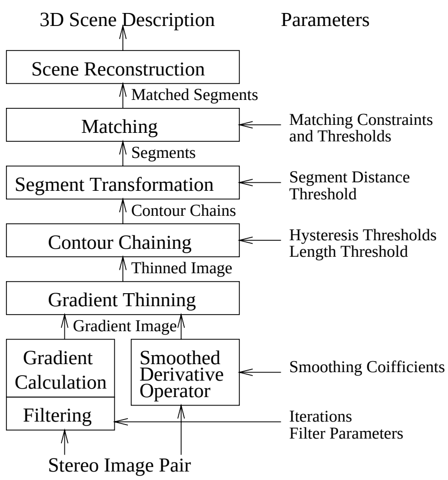


**Figure 1. The Vision System**

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.

- The output of higher levels of vision systems (e.g. 3D Scenes, Feature Positions) tend to be more task specific and less general than the output of lower levels (e.g. gradient images). Thus we face a tradeoff of a broadly applicable evaluation result against a result that is of high significance with respect to a certain task.

- Evaluation of a vision system at a higher level takes more calculation time as all steps below have to be executed and, more important, their parameters have to be controlled.

- For many evaluation methods, ground truth about the input data is required. This is most conveniently done with artificial images, however, such an approach requires that disturbances and image contents can be modeled sufficiently well. While this is easy for lower levels (e.g. contour level) it is increasingly difficult at higher levels (e.g. for the 3D scene level one would have to generate random scenes and render them using ray-tracers).

From the above considerations we have decided to evaluate the system by looking at the result of the segment extraction step.

## 4. The Evaluation Criteria

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 well the vision system is able to identify segments. We verify for 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. For reasonable parameter settings it is always below 1 (as setting the detection thresholds to infinity will not detect any pixels). We define the robustness of the vision system as the inverse 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 adaptive filters. For example, if we use a large Gaussian filter matrix (e.g. 100x100) for smoothing we will obtain an extremely high robustness when detecting a single, infinitely long edge. The strong smoothing however will destroy smaller details which for many applications are necessary. We therefore need a second measure to describe the ability of the filter to preserve small details.

If we had only linear filters we could describe the effect of the filter on the frequency spectrum of the image. For adaptive filters the situation is more complicated. While edge preserving filters will not change the frequency spectrum of a single edge they do change small structures made of edges (e.g. two parallel edges with a distance of one pixel). What we need is a common measure that can be used for both, linear and adaptive filters.

The theoretical framework for this measure is provided by a scale space analysis [15]. In Gaussian scale space, edges (zero crossings of the gradient) disappear at a certain value of the scale axis. We will now take a feature and vary its size until it is no longer preserved by the adaptive filters. We then look at which scale this feature would disappear and relate this scale space location to the adaptive filter. It is clear that the image that is filtered with the adaptive filter is not equal to the image at its scale space position. However coarseness of the the features extracted from the two images will be similar.

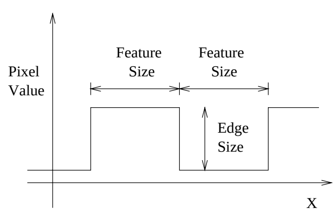


**Figure 2. The feature used for scale space measurement.**

As we are interested in edge segments, the feature we use will consist of edges. The feature depicted in figure 2. We reduce the size of the feature with sub-pixel accuracy until it disappears (the number of gradient zero crossings in the quantized image becomes less than 3). It should be noted that the disappearance of the feature in scale space involves quantization effects which had to be considered in the experiment.

## 5. Controlling the System

### 5.1. Component Interactions

When we exchange the filter of our segment extraction system we will usually measure a change in the systems performance. This change can be due to two reasons. First because the performance of the filter is truly different. Second because the new filter causes another system component to behave differently (e.g. a thresholding stage now discards all data because the new filter would require a different threshold). If no precautions are taken the second effect will usually be dominant.

As one of the properties of the filters we want to test is their smoothing ability we have to make sure that there are no further components that smoothes the data. The only other component that is able to do this is the edge detector. We will therefore use the least smoothing symmetric edge detector, the simple 2x2 derivative operator (-1 1).

All steps in the vision system have free parameters. When one component of the system is exchanged the optimal settings of these parameters will usually change. If we want to create an even field for all tested filters we have to adjust the parameters for each filter individually. Bad parameter control or a static parameter setting can have an effect that is significantly stronger than the differences in the tested filters [1].

One solution would be to let the parameters for every filter be adjusted by a human operator. To do this for each filter and noise level is not only very time consuming but the results obtained could hardly be called scientific as they are not reproducible. What we therefore need are automatic parameter control strategies.

### 5.2. Automatic Parameter Control

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. For our experiments we have shown [1] that only three parameters, the two hysteresis thresholds and the chain length threshold have to be controlled dynamically. A fourth parameter, the segment length threshold is controlled by a simple feed forward loop.

Two of the adaptive filters have a free parameter, the anisotropic diffusion technique and the Sigma Nearest Neighbor (SIG) filter. As for the parameter of anisotropic diffusion no effective control strategy was known it was tested with several parameter settings. For the SIG filter control of the parameter was possible when the signal to noise ration was above 3.0. Below that value or in the presence of blur any parameter setting either smoothed the edges or did no longer reduce the noise. The SIG filter was therefore excluded from the tests.

To control the number of iterations of a filter two different strategies are used. For the linear techniques the filters have been tested for each iteration count separately. For the adaptive filters we determined after how many iterations the images converged to a stable (and usually optimal) state and used this number for the experiments.

We have tested several methods to control the two hysteresis thresholds and the length threshold. All use the error function defined in the last section.

- **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).
- **Model Based Tuning.** Following an idea of Ramesh [14] 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 [2]. While the error is small it is in the order of magnitude of the difference between different filters.
- **Gradient Descent.** Gradient descent in the parameter space is difficult as the error of the statistical measurements creates local minima and the evaluation function is constant for too small or too large thresholds. 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 [1].

For the measurements gradient descent parameter control is used.

## 6. Experiments and Discussion

The adaptive filters that were implemented are the Symmetric Nearest Neighbor (SNN) filter [6], the K-Nearest Neighbor (KNN) filter [4], the Sigma Nearest Neighbor (SIG) filter [8], the Nagao filter [11], Lee's filter [7] and Anisotropic Diffusion [12]. The SNN, KNN and Nagao techniques have been combined with Median, Mean and Lee filter as in [16]. 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.

To allow comparison with linear filters Mean filtering (size 3x3, 5x5 and 7x7 between 1 and 10 iterations) and Binomials (same configurations as mean) have been tested. As reference detectors Sobel's operator and Deriche's detector [5] are included.

The name of each adaptive or linear filter is followed by the window size and the number of iterations (e.g. SNN-Mean3-15 is a 3x3 filter with 15 iterations). The experiments where made on artificial edges of a single height that where distorted by zero mean white Gaussian noise.

Figure 3 shows the inverse robustness (the signal to noise ratio for an error of 0.1) on the X-Axis and the size of the smallest detail for that was still detected for each filter on the Y-Axis. The ideal filter would be at (0,0). It would detect infinitely small details at any signal to noise ratio. 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.

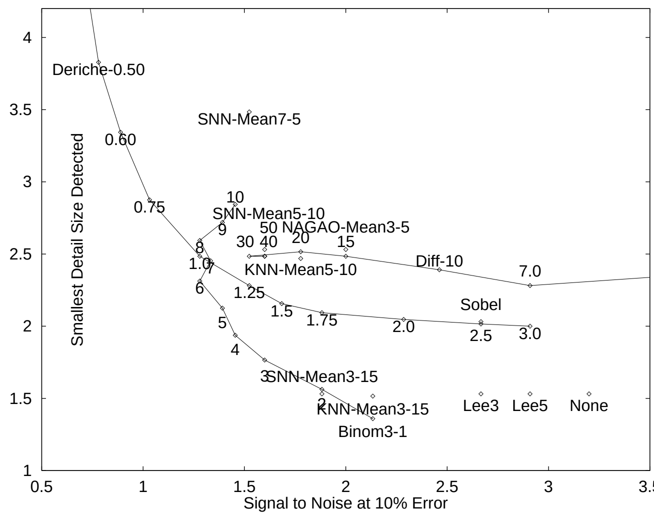


**Figure 3. Robustness and Scale Space Position of the tested Filters**

The result for the adaptive techniques are somewhat surprising. None of them performs better than either the Canny detector or the use of Binomials for smoothing. The Canny 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 Canny detector is symmetric to the pixel center while the edge detector that uses Binomials is symmetric to the pixel boundary. As a result the Canny detector produces two maximal responses to an edge which is on a pixel boundary while the Binomial based detector produces only one.

The SNN filter provides the best robustness of the adaptive techniques with KNN and Nagao following. 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.

Lee's filter increases the robustness of the segment extraction only slightly. This is because its smoothing strength depends on the variance of the data in its window which is always high on edges.

The edges for the conducted tests had a height that caused them to be smoothed by anisotropic diffusion if its parameter was 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. It is no longer edge preserving. The scale space location of the diffusion filtering changes little as the value of the parameter is varied. This again illustrates that diffusion scale space as proposed in [12] is fundamentally different to the Gaussian scale space as proposed by [15].

These integer approximations of Gaussian provide the best results for low noise. The robustness is maximal for about 8 iterations. For stronger smoothing the robustness decreases because of multiple responses generated by the filters and quantization effects. Multiple responses equally caused problems for filters of a size of more than 3x3 and the single iteration of the 3x3 filter.

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 is due to the measurement technique). The Sobel detector performs very similar to a Canny detector with a smoothing coefficient of 2.5.

Deriche's recursive implementation of Canny's detector offers the best overall robustness of the field as well as an excellent overall tradeoff. It additionally has the shortest calculation time except for the single iteration Binomial, mean and Sobels filters.

From additionally measured data we find that adaptive and linear techniques (the latter include the Canny-Deriche detector) try to enhance the signal to noise ratio in very different ways. 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. The linear techniques reduce the value of the gradient dramatically but the relative variance of it is lowered to an even larger extent. As the smoothing increases, the response to a single edge is becoming broader and the number of pixels with an almost equal gradient value increase leading to multiple responses.

An argument in favor of the adaptive techniques is that the white zero mean Gaussian noise 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).

## 7. Conclusion

We have presented an experimental method for the performance characterization of low-level vision algorithms in vision systems. The goal of evaluating the effect of the algorithm on the overall output of the system poses a number of new problems. To guarantee that the measured effect is really due to the tested component is a complex problem. System engineering methodology and strategies for the automatic control of parameters are necessary to solve it. We have implemented several methods for the automatic control of parameters. The experimental results show that for the detection of segments in the presence of white Gaussian noise no adaptive filter gives better results than Gaussian filtering or the Canny-Deriche edge detector.

## References


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

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

[3] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.

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

[5] R. Deriche. Fast algorithms for low-level vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(1):78–87, Jan. 1990.

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

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

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

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

[10] B. Migeon. Pretraitement d'image par filtrage epsf. Technical report, Direction de Recherches Etudes et Techniques INRIA, Jan. 1994.

[11] M. Nagao and T. Matsuyama. Edge preserving smoothing. *Computer Graphics and Image Processing*, 9:394–407, 1979.

[12] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, July 1990.

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

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

[15] 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.

[16] W.-Y. Wu, M.-J. J. Wang, and C.-M. Liu. Performance evaluation of some noise reduction methods. *CVGIP Graphical Models and Image Processing*, 54(2):134–146, Mar. 1992.