Hybrid Human-machine Non-linear Filter Design Using Envelopes

MARCEL BRUN[1,3], ROBERTO HIRATA JR.[1,2], JUNIOR BARRERA[1], EDWARD R. DOUGHERTY[3]

[1]Departamento de Ciência da Computação, Instituto de Matemática e Estatística - USP
Rua do Matão, 1010 - 05508-900 - São Paulo, BRAZIL
{mbrun,hirata,jb}@ime.usp.br

[2]SENAC College of Computer Science and Technology
Rua Galvão Bueno, 430 - 01506-000 - São Paulo, BRAZIL

[3]Department of Electrical Engineering, Texas A&M University
College Station - 77843-3128 - TX - United States
edward@ee.tamu.edu

Abstract. Machine design of a signal or image operator involves estimating the optimal filter from sample data. In principle, relative to the error measure used, the optimal filter is best; however, owing to design error, the designed filter might not perform well. In general it is suboptimal. The envelope constraint involves using two humanly designed filters that form a lower and upper bound for the designed operator. The method has been employed for binary operators. This paper considers envelope design for gray-scale filters, in particular, aperture filters. Some basic theoretical properties are stated, including optimality of the design method relative to the constraint imposed by the envelope. Examples are given for noise reduction and de-blurring.

Keywords: Morphological operators, statistical design, constraint, optimal filters

1 Introduction

Strictly data-driven automatic design of operators works well only for relatively small windows. To incorporate more variables, and therefore more information, it is necessary to impose constraints on the class of filters from which the designed operator is to be taken. Hence, rather than try to estimate the optimal operator, one tries to estimate the optimal operator in the constraint class. There is an advantage to constraint if the expected error of the constrained designed operator is less than the expected error of the unconstrained designed operator. Typically, constraint involves the application of some heuristics. For envelope design, the heuristics take the form of lower and upper bounds on the designed operator. These are humanly designed and imposed on the automatic design procedure. If they are chosen in such a way that the optimal filter lies fully between them, then there is always an advantage to using the envelope. Unfortunately, if the lower and upper bound are too far apart, this advantage is negligible (or in the extreme null because we could always choose zero and infinity for the lower and upper bounds). To gain real advantage requires imposing meaningful constraints. If the sample size is sufficiently large, then the constraint will not be beneficial, but for large windows such a prospect can be unreasonable and therefore a well-chosen constraint can be beneficial. Previously, envelope constraint has been studied for binary filters [1]. Here we extend the concept to gray-scale filters, state some basic properties, and provide applications to both de-noising and de-blurring. In particular, we apply envelope design to aperture filters [2]. Theorems are stated without proof. The proofs, further theoretical results, and more applications are given in [3].

2 W-Operators and Apertures

Digital images can be formally defined and represented by functions from a non-empty set E that is an Abelian group with respect to a binary operation $+$ to an ordered chain L. Usually, E is a subset of $Z \times Z$ (where Z is the set of integers) and L is a positive interval of Z, i.e., $L = [0, l-1]$, where $l \in Z$. A binary image is an element of $\mathcal{P}(E)$ (the class of subsets of E). It can also be represented as a function of $E \to [0,1]$ via the indicator function [2]. The set of all possible functions from E to L will be denoted by L^E and a mapping Ψ from L^E to L'^E (where L' is a nonnegative interval of Z, not necessarily equal to L) will be called an image operator.

A finite subset W of E, customarily containing the origin (of E), will be called a *window* and the number of points of the window W will be denoted by $|W|$. A *configuration* is a function from $W \to L$ and the space of all possible configurations from $W \to L$ will be denoted by L^W. Configurations are usually the result of translating a window W by t, $t \in E$, and observing the values of an image h, $h \in L^W$. Formally, if $W = \{w_1, w_2, \ldots, w_n\}$, $n = |W|$, and we associate the points of W to a vector $(w_1, w_2, \ldots, w_n)$, then a configuration $h(W_t)$, seen by W translated by t, denoted W_t, can be written as,

$$h(W_t) = (h(t + w_1), h(t + w_2), \ldots, h(t + w_n)),$$

Since digital images can be modeled by digital random functions, $h(W_t)$ is a realization of a random vector $\mathbf{X} = (X_1, X_2, \ldots, X_n)$, i.e., $h(W_t) = \mathbf{x} = (x_1, x_2, \ldots, x_n)$, where $\mathbf{x}$ denotes a realization of $\mathbf{X}$ and $x_1, x_2, \ldots, x_n$ are the values observed in h under W_t. An important subclass of operators from L^E to L'^E is the class of *W-operators.*

These are translation invariant (t.i.) and locally defined (l.d.) by a window W. If an image operator Ψ is a W-operator, then it can be characterized by a function $\psi : L^W \to L'$, called a *characteristic* function, by,

$$\Psi(h)(t) = \psi(h(t+w_1), h(t+w_2), \ldots, h(t+w_n)) = \psi(\mathbf{x})$$

An aperture configuration is a function from $W \to K$ ($K = [-k,k], k \in Z^+$), and the set of all possible aperture configurations on W is denoted by K^W. These configurations are usually the result of a spatial translation of a window W by t, $t \in E$, range translating W by z, $z \in Z$, and truncating the observed values to values inside K. In this case, a configuration can be written as

$$h^*_{-z}(W_t) = (h^*_{-z}(t+w_1), h^*_{-z}(t+w_2), \ldots, h^*_{-z}(t+w_n))$$

where $h_{-z}(t) = h(t) - z$, $z = z(h(W_t))$ is a function of $h(W_t)$, and h^*_{-z} is defined by

$$h^*_{-z} = \begin{cases} h_{-z} & : \ -k \le h_{-z} \le k \\ k & : \ h_{-z} > k \\ -k & : \ h_{-z} < -k \end{cases} \qquad (1)$$

The class of aperture operators is a subclass of the class of W-operators that are t.i. and l.d. by a window W, and are also locally defined by a range window $K = [-k, k], k \in Z^+$. Let $\mathbf{X}^*$ be a truncated random variable $\mathbf{X}$ (following the truncation rule of Eq. 5), i.e., $\mathbf{X}^* = (X_1^*, X_2^*, \ldots, X_n^*)$, and let $\mathbf{x}^*$ be a realization of $\mathbf{X}^*$. An aperture operator Ψ defined via the aperture $\mathcal{A} = W \times K$ is characterized by an *aperture characteristic function* $\psi_{\mathcal{A}} : \mathcal{A} \to L'$ according to the representation

$$\Psi_{\mathcal{A}}(h)(t) = \psi_{\mathcal{A}}(h^*_{-z}(t+w_1), \ldots, h^*_{-z}(t+w_n)) = \psi_{\mathcal{A}}(\mathbf{x}^*)$$

Given two gray-level images on E, h to be observed and g to be estimated, the basic filtering problem is to find a filter Ψ that minimizes an error measure between $\Psi(h)(t)$ and $g(t)$, where $t \in E$. Assuming h and g are jointly stationary [4], the *mean-square error* (MSE) of Ψ is given by $E[|g(t) - \Psi(h)(t)|^2]$. If Ψ is a W-operator, these are equivalent to $E[|Y - \psi(\mathbf{X})|]$ and $E[(Y - \psi(\mathbf{X}))^2]$, respectively. For an operator Ψ defined by ψ (defined on L^D, $D = L^W$), the MSE is

$$MSE(\psi) = \sum_{\mathbf{x} \in D} \sum_{y=0}^{l-1} (y - \psi(\mathbf{x}))^2 P(y, \mathbf{x}) \qquad (2)$$

where $P(y, \mathbf{x})$ is the joint probability of $(y, \mathbf{x})$.

The optimal MSE filter is given by the median and the mean of the distribution of $Y|\mathbf{x}$, respectively.

When using a suboptimal filter instead of the optimal filter, there is an increase in error. The total increase in MSE error from using ψ in place of the optimal filter ψ_{opt} is

$$\Delta(\psi, \psi_{opt}) = \sum_{\mathbf{x} \in D} (\psi(\mathbf{x}) - \psi_{opt}(\mathbf{x}))^2 P(\mathbf{x}). \qquad (3)$$

If $\psi_{opt,N}$ is an estimate of ψ_{opt} based on N sample pairs $(\mathbf{X}^1, Y^1), \ldots, (\mathbf{X}^N, Y^N)$, then there is a design (estimation) cost $\Delta(\psi_{opt,N}, \psi_{opt})$. The MSE of the designed filter is given by

$$MSE(\psi_{opt,N}) = MSE(\psi_{opt}) + \Delta(\psi_{opt,N}, \psi_{opt}). \qquad (4)$$

Since the estimated filter $\psi_{opt,N}$ depends on the training sample, it is random. Estimation error (*precision*) is defined by the expected cost $E[\Delta(\psi_{opt,N}, \psi_{opt})]$ and depends on the estimation procedure. The expected MSE of $\psi_{opt,N}$ is found by taking the expected value in Eq 3, in which $MSE(\psi_{opt})$ is constant.

3 Design of W-operators under envelope constraint

In this section we introduce the notion of *envelope* and quantify its effects on the quality of the operator designed.

3.1 Envelope constraint

The envelope constraint has been originally defined on binary W-operators[1], that is, operators that transform subsets of $P(W)$ (the power set of W). There we defined two subsets, $A, B \subset P(W)$ such that we have confidence that the W-operator is zero if is applied to a subset $\mathbf{x}$ of A and one if it is applied to a subset $\mathbf{x}$ of B^c. In reality, the sets A and B define two operators α and β, respectively. Therefore, the envelope of a function $\psi : P(W) \to 0, 1$ was defined as $\psi_{con} = (\psi \vee \alpha) \wedge \beta)$. It is always true that $\alpha \leq \psi_{con} \leq \beta$.

We now generalize this approach to gray level operators. Let $D = L^W$ be the configuration space, and let $\alpha, \beta : D \to L$ be two gray level characteristic functions with $\alpha \leq \beta$. Our prior knowledge is some confidence in that $\alpha \leq \psi_{opt} \leq \beta$. In such case, let ψ be a machine designed characteristic function, for each configuration $\mathbf{x} \in D$, if $\psi(\mathbf{x}) < \alpha(\mathbf{x})$ then we will prefer to use $\alpha(\mathbf{x})$ in place of $\psi(\mathbf{x})$. The same situation arises if $\psi(\mathbf{x}) > \beta(\mathbf{x})$. Formally, for each operator $\psi : D \to L$, we define its *constrained* operator ψ_{con} in the following way:

$$\psi_{con}(\mathbf{x}) = \begin{cases} \alpha(\mathbf{x}) & : \psi(\mathbf{x}) < \alpha(\mathbf{x}) \\ \psi(\mathbf{x}) & : \alpha(\mathbf{x}) \leq \psi(\mathbf{x}) \leq \beta(\mathbf{x}) \\ \beta(\mathbf{x}) & : \psi(\mathbf{x}) > \beta(\mathbf{x}) \end{cases} . \quad (5)$$

As in the binary case, $\psi_{con} = (\psi \vee \alpha) \wedge \beta$. We will call the pair (α, β) an *envelope.* An envelope (α, β) defines an *envelope constraint* "$\mathbf{Q}_{\alpha,\beta} \subset L^D$" as the sub collection of the operators in L^D that lies between α and β, i.e. $\mathbf{Q}_{\alpha,\beta} = \{\psi \in L^D : \alpha \leq \psi \leq \beta\}$. For any operator $\psi \in L^D$, its constrained operator ψ_{con} is an element of the restriction $\mathbf{Q}_{\alpha,\beta}$.

3.2 Error Analysis for envelope constraint

The way to combine human and machine design is by machine designing of operators under envelope constraint. In place of the best operator in L^D, we will seek the best operator in $\mathbf{Q}_{\alpha,\beta}$. A key problem in restriction design is the existence of a easy way to train constrained operators. The solution to the problem of proving this existence has been shown for resolution constraint [5, 6], and also for increasing binary filters [7]. In the case of gray-level envelopes, the next theorem shows that the filter $\psi_{opt-c} = (\psi_{opt} \vee \alpha) \wedge \beta$ obtained by envelope constraining the optimal filter ψ_{opt} is the optimal filter in the class $\mathbf{Q}_{\alpha,\beta}$:

Theorem: Let $\psi_{opt-c} = (\psi_{opt} \vee \alpha) \wedge \beta$, then for any $\psi \in \mathbf{Q}_{\alpha,\beta}$, $\Delta(\psi, \psi_{opt}) - \Delta(\psi_{opt-c}, \psi_{opt}) \geq 0$

The importance of this theorem is that it allows to compute the optimal constrained filter by constraining the estimationn of the optimal non-constrained one.

Given the envelope (α, β) and the estimate ψ of ψ_{opt}, we need to determine the advantage of using the envelope constrained filter $\psi_{con} = (\psi \vee \alpha) \wedge \beta$ instead of ψ. The advantage "$A_\psi(\alpha, \beta)$" is given by the difference between the MSE increase for ψ over ψ_{opt} and the MSE increase for ψ_{con} over ψ_{opt}:

$$A_\psi(\alpha, \beta) \quad = \quad \Delta(\psi, \psi_{opt}) - \Delta(\psi_{con}, \psi_{opt}). \quad (6)$$

The next theorem shows that the envelope constraint must be beneficial if well designed.

Theorem: If the optimal filter ψ_{opt} lies in the envelope, meaning $\alpha \leq \psi_{opt} \leq \beta$, then

$$A_\psi(\alpha, \beta) \geq 0. \quad (7)$$

In general the optimal filter does not lie within the envelope and the advantage can be positive or negative. The potential for disadvantage is illustrated by letting $\psi = \psi_{opt}$, meaning that before constraint, the estimated filter is actually optimal. The next theorem shows that the envelope constraint will be detrimental if not well designed.

Theorem: If $\psi = \psi_{opt}$, then

$$A_\psi(\alpha, \beta) \leq 0. \quad (8)$$

We can see that using the constrained operator ψ_{con} in place of an operator ψ, the error will decrease (or not change) if the optimal operator lies in the envelope, but it can increase too if the optimal operator is not completely bounded by the envelope (α, β).

3.3 Precision of estimation under envelope constraint

When designing the optimal operator from pairs of samples, the designed operator is only an approximation of the optimal operator. For different sets of samples, different estimators of the optimal operator can be created. To analyze the goodness of an envelope, one must compute the average value of the MSE increase $A_\psi(\alpha, \beta)$ for all the possible operators that can be created from a sample of a given size.

As with other types of constraints, the envelope constraint is detrimental when we have a large number of samples pairs. In such cases, is better to train the optimal operator of the whole class L^D than the optimal constrained operator in $\mathbf{Q}_{\alpha,\beta}$. However, for small samples and a good envelope (α, β), it can be better to design envelope constrained operators than non constrained ones.

To examine estimation precision, we suppose that ψ results from a statistical estimation of ψ_{opt} using N pairs of samples, so that $\psi = \psi_N$. Assuming that ψ_N is a strongly consistent estimator of ψ_{opt}, $\psi_N \rightarrow \psi_{opt}$ almost surely as $N \rightarrow \infty$. We are using a strongly consistent estimators. The next theorem bounds the expectation $E[A_{\psi_N}(\alpha, \beta)]$ of the MSE increase "$A_{\psi_N}(\alpha, \beta)$" when N goes to infinity

Theorem: If ψ_N is a strongly consistent estimator of $\psi_o pt$

$$\mathrm{Lim}_{N \rightarrow \infty} E[A_{\psi_N}(\alpha, \beta)] \leq 0. \quad (9)$$

In the limit, the envelope constraint is not beneficial and its disadvantage is increased (negatively) by making α larger and β smaller. The influence of a larger α is in two senses, first, by increasing the number of configurations $\mathbf{x} \in D$ where $\psi_{opt}(\mathbf{x}) < \alpha(\mathbf{x})$, and second, by increasing the value $(\alpha(\mathbf{x}) - \psi_{opt}(\mathbf{x}))^2$ for such configurations.

Equations 7 and 9 can seem as opposites. In the hypothesis of equation 7 the advantage is always non negative, hence at the limit when N goes to infinity, it can be only zero. In this case, lower values for N will give better advantage of the envelope.

4 Experimental results

In this section we demonstrate the performance of the hybrid approach to design filters for noise reduction and deblurring on signals and images. Blurring is accomplished by an 11 point flat convolution kernel for signals and a 3×3 non flat convolution kernel for images (Fig. 1). The noise is independent Gaussian noise added at 5% of the points.

1	2	1
2	3	2
1	2	1

Figure 1: Kernel for images blurring

The signals used for the experiments reported in this paper are "saw" signals (Fig. 2) which are generated by the model described in [8].

The images used for the experiment follow a random Boolean function model [9] whose primary function is pyramidal with at most 16 gray levels range.

We have generated 20 images, each of size 256×256 points, and their respective blurrings. Figures 3 and 4 show part of an original and blurred image, respectively.

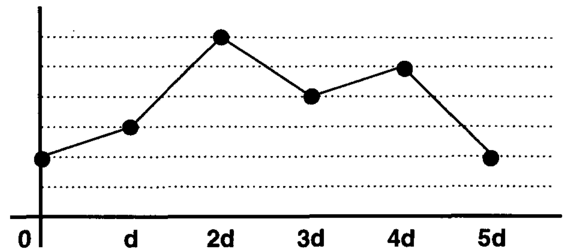

Figure 2: Saw function passing through equally spaced random points

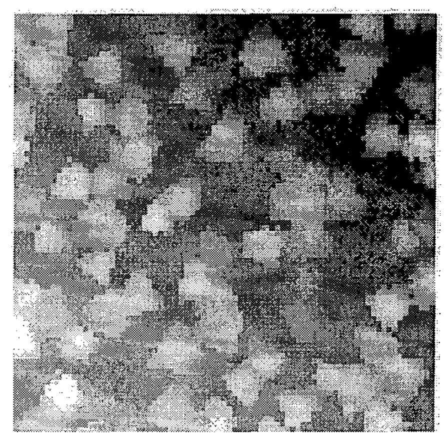

Figure 3: Part of a random Boolean function

4.1 Noise Filtering in Signals

This experiment shows the performance of the envelope approach in comparison with aperture filters for noise filtering. Thirty saw signals have been generated (d = 15, size = 1024 points) and corrupted by noise with mean 0 and variance 3 at 5% of their points. From this set, one to ten pairs of signals (noise, original) are given to the system to design the aperture filters and twenty pairs to test the performance. Two hybrid filters have been tested, *env₁* and *env₂*. For *env₁*, α and β are given by,

$$\alpha(f) = \varepsilon_B(f) \quad (10)$$

$$\beta(f) = \delta_B(f) \quad (11)$$

where B is a 1×7 structuring element [10], ε_B is an erosion [10] by B, δ_B is a dilation [10] by B, and ψ is an aperture filter with $W = 1 \times 7$ and $k = k' = 15$.

For env_2, α and β are given by,

$$\alpha(f) = \wedge\{\varphi_B \phi_B(f), \phi_B \varphi_B(f)\} \quad (12)$$

$$\beta(f) = \vee\{\varphi_B \phi_B(f), \phi_B \varphi_B(f)\} \quad (13)$$

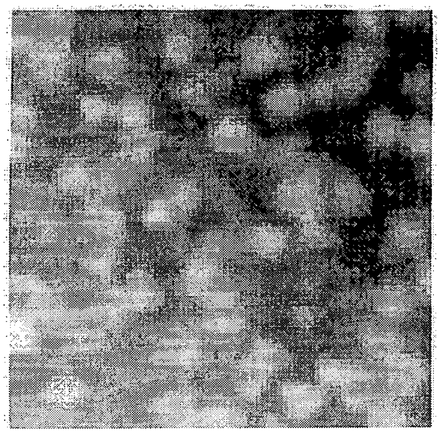

Figure 4: Blurring of the image shown in Fig. 3

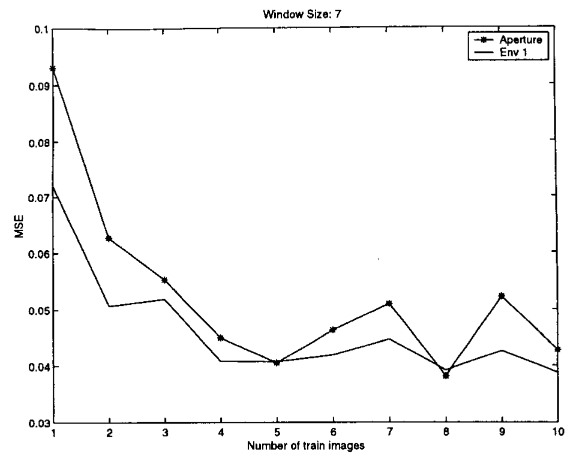

Figure 5: MSE errors for env_1

where φ_B is an opening [10] by B, ϕ_B is a closing [10] by B, $\wedge$ is the infimum [10], $\vee$ is the supremum [10], B, ψ, W, k and k' are the same as defined before. The composition of closings and openings are sometimes called *alternating sequential filters* [2]

The aperture filters and the hybrid filters have been applied to 20 test signals and compared to the respective ideal images. The MSE errors as a function of the number of training examples have been computed and averaged. The average MSE for the noisy images is 0.27. Figure 5 shows the MSE errors for the aperture filters and the env_1 filters. The results of ε_B and δ_B do not depend on the number of examples and are constant equal to 0.62 and 0.44, respectively. It is interesting to note that the hybrid filter gives a better result than the aperture filter. Moreover, it is less affected by noise in the training set than the correspondent aperture filter. For instance, for 6 and 7 images, the error of the aperture filter increases (due to noise in the training set). The error of the hybrid filter also increases but not in the same proportion.

Figure 6 shows the MSE errors for the designed aperture filters, $\varphi_B\phi_B$, $\phi_B\varphi_B$, and the env_2 filters. The results of the alternating sequential filters do not depend on the number of examples and are constant equal to 0.075 and 0.056, respectively. Now, the hybrid filter is not better than the aperture filter for more than 4 training images, just more stable. However, for less than 4 images, the hybrid filter is very good, showing that this approach can be a good alternative when the number of training examples is limited and we have a good α and β.

Figure 7 shows the results of both env_1 and env_2 filters, together with the results of the aperture filters.

4.2 Deblurring Boolean Model Images

This experiment shows the same technique applied to 2D random Boolean model images. We compared aperture filters to linear and hybrid filters. From the set of 20 images, we used one to ten pairs (blurred, original) to design the aperture filters and ten pairs to test the performance. The hybrid filter tested, *env*, has α and β given by,

$$\alpha(f) = \Psi_{optlin}(f) - 1 \quad (14)$$

$$\beta(f) = \Psi_{optlin}(f) + 1 \quad (15)$$

where Ψ_{optlin} is the optimal restricted linear filter with window 9×9. ψ is an aperture filter with $W = 3 \times 3$ and $k = k' = 5$.

Figure 8 shows the initial MSE error and the MSE errors for the aperture filters, for the optimal restricted linear filter and the *env* filter. The results confirm the advantage of the envelope.

5 Conclusions

Envelope design provides a systematic means by which humanly designed operators can be used to assist in machine design. If the lower and upper bounds are well chosen, then envelope design can be advantageous in the common situation in which there is insufficient data to adequately design filters requiring more than a small number of variables.

6 Acknowledgments

Marcel Brun is supported by FAPESP under grant 98/15586-9. Roberto Hirata Jr. is supported by CNPq under grant 145157/1998-9.

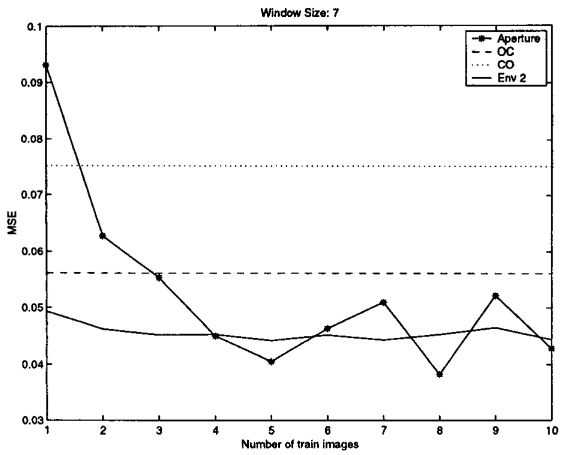

Figure 6: MSE errors for env_2

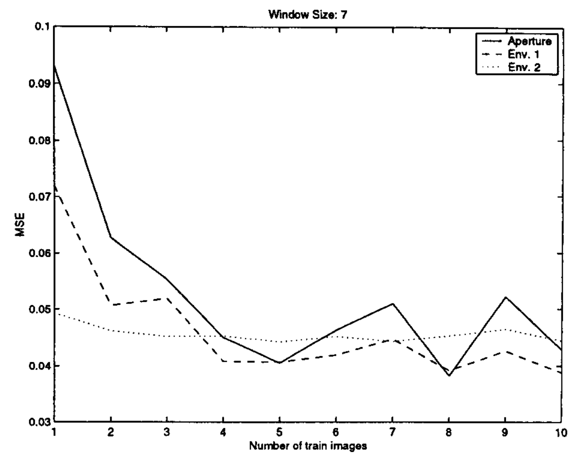

Figure 7: MSE errors for filtering noise

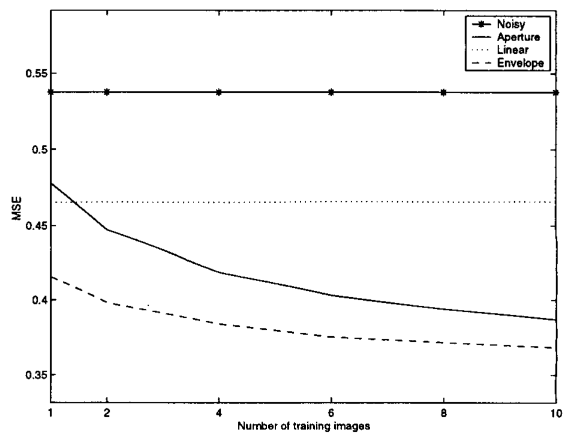

Figure 8: MSE errors for *env*

References

[1] J. Barrera, E. R. Dougherty, and M. Brun. Hybrid human-machine binary morphological operator design. An independent constraint approach. *Signal Processing*, 80(8):1469–1487, August 2000.

[2] H. J. A. M. Heijmans. *Morphological Image Operators.* Academic Press, Boston, 1994.

[3] M. Brun, R. Hirata Jr., J. Barrera, and E. R. Dougherty. Non-linear Filter Design Using Envelope. *In preparation*, 2001.

[4] E. R. Dougherty. *Random Processes for Image and Signal Processing.* SPIE and IEEE Presses, Bellingham, 1999.

[5] E. R. Dougherty, J. Barrera, G. Mozelle, S. Kim, and M. Brun. Multiresolution Analysis for Optimal Binary Filters. *Mathematical Imaging and Vision*, (14):53–72, 2001.

[6] Roberto Hirata Jr., Marcel Brun, Junior Barrera, and Edward Russel Dougherty. Image restoration by multiresolution aperture filters. In *Nonlinear Image Processing and Pattern Analysis XII*, volume 4304 of *Proceedings of SPIE*, January 2001.

[7] N. S. T. Hirata, E. R. Dougherty, and J. Barrera. A Switching Algorithm for Design of Optimal Increasing Binary Filters Over Large Windows. *Pattern Recognition*, 33(6):1059–1081, June 2000.

[8] R. Hirata Jr., E. R. Dougherty, and J. Barrera. Aperture Filters. *Signal Processing*, 80(4):697–721, April 2000.

[9] M. Schmitt and F. Preteux. *Image Analysis and Mathematical Morphology*, volume 2, chapter Boolean texture analysis and synthesis. Ed. J. Serra, Academic Press, New York, 1988.

[10] J. Serra. *Image Analysis and Mathematical Morphology.* Academic Press, 1982.