Design of Optimal Filters for Image Texture Feature Separation

Trygve Randen and John Håkon Husøy

Høgskolen i Stavanger
P.O. Box 2557 Ullandhaug, 4004 Stavanger, Norway
t.randen@ieee.org, http://www.hsr.no/˜tranden

ABSTRACT

The use of filters is one of the major techniques for texture feature extraction. In this paper, we present approaches to filter design for optimal texture feature extraction. A generic feature extraction scheme of a finite impulse response filter and a local energy function is modelled, and the filter is optimized with respect to the Fisher criterion. We develop the necessary expressions for a gradient search optimization. Based on some assumptions, we also find a closed form solution to the problem. Then we extend the closed form solution to a more generalized unifying framework. Finally, the practical utilities of the filter design techniques are shown through experiments.

1. INTRODUCTION

The design of *optimal* filters for texture discrimination is gaining increased interest, and several recent articles propose methods for the design of such filters. Dunn et al. [1] and Teuner et al. [2] present approaches to the design of optimal Gabor filters and Husøy et al. [3] present an approach to the determination of the optimality of a bank of critically sampled perfect reconstruction IIR filters. A common problem with all these approaches is that the number of free parameters is low, and there is no guarantee that a true global optimum will be reached.

Opening for a much less restricted set of possible solutions, Jain and Karu [4] propose an approach to optimal filter design using a multi layer perceptron neural network for the filtering and local energy operations, and the back propagation training rule for the determination of the optimal parameters for multiple textures. Mahalanobis and Singh [5] uses yet another approach to the design of the filters, determining the optimal filter by stochastic texture analysis and maximizing the ratio of the energies of the filter responses of two textures. A related approach was presented by Unser [6], for design of an optimal *transform* for the two-texture classification problem.

In this paper we propose a new technique for the design of filters for two-class feature extraction which are optimal with respect to the Fisher criterion [7]. By assuming that the filter outputs are Gaussian random variables, we are able to solve the optimization problem.

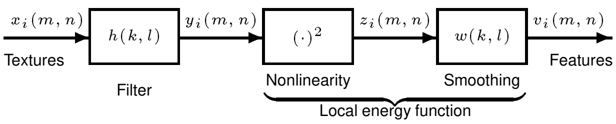


Figure 1. The experimental setup.

If we make the assumption that the filter outputs are separable AR(1) processes (a rather restrictive assumption), we are able to find a closed form solution to the optimization problem. Otherwise, the optimization problem is solved by a gradient search.

We finally develop a generalized formulation and a unifying framework for the closed form solution of several optimization criteria. We also show how a minimum classification error probability approach may fit into the unifying framework.

2. THE FISHER CRITERION FUNCTION

The generic experimental setup for texture feature extraction with filters [8], Figure 1, is used as our model. We assume that we have an input image consisting of two textures, $x_1(m,n)$ and $x_2(m,n)$. The output for texture $x_i(m,n)$ is the feature image $v_i(m,n)$. In the sequel, we denote the feature mean for texture i by μ_{v_i} and the feature variance by $\sigma_{v_i}^2$. Expressions for the mean and variance will be given later.

The filtered image, $y_i(m,n)$, is given by the two-dimensional convolution of $h(m,n)$ and $x_i(m,n)$. We define $\mathbf{h}$ as the vector formed by lexicographically ordering the rows of $h(m,n)$, and $\mathbf{x}_i(m,n)$ as $[x_i(m,n)\cdots x_i(m,n-N+1)\cdots x_i(m-M+1,n-N+1)]^T$. Then, the output at pixel location (m,n) for texture i may be written

$$y_i(m,n) = \mathbf{h}^T\mathbf{x}_i(m,n).$$

Similarly, the smoothing operation may be defined as

$$v_i(m,n) = \mathbf{w}^T \mathbf{z}_i(m,n).$$

We assume that the smoothing filter is an $N_w \times N_w$ averaging filter.

The obvious aim of an optimal filter for two-texture discrimination is to make the local energies of the filter outputs maximally separable. If the feature separation is good we will achieve good segmentation – and the segmentation may be performed by a simple two-class linear discriminant classifier [7]. In order to achieve good separation, Mahalanobis and Singh [5] proposed the object function

$$J_1(\mathbf{h}) = \frac{\mu_{v1}}{\mu_{v2}} \tag{1}$$

and Unser [6] proposed to use

$$J_2(\mathbf{h}) = \frac{(\mu_{v1} - \mu_{v2})^2}{\mu_{v1}\mu_{v2}} \tag{2}$$

for optimal filter and transform design, respectively. (The Unser criterion [6] can easily be adapted to filter design.) The major drawback with both these criteria, is that they do not incorporate the feature variances. Hence we may achieve large separation of the feature *means*, but so large feature *variances* that the feature distributions overlap considerably. One object function which incorporates the variance is

$$J(\mathbf{h}) = \frac{(\mu_{v1} - \mu_{v2})^2}{{\sigma_v}_1^2 + {\sigma_v}_2^2}, \tag{3}$$

known as the Fisher criterion [7, p. 135]. It is previously not used in optimal texture feature extraction. In this section we will develop the necessary theory for optimal filter design with respect to the Fisher criterion.

In order to find the optimal filter, we have to find the partial derivative of $J(\mathbf{h})$ with respect to $\mathbf{h}$,

$$\begin{aligned} \frac{\partial J(\mathbf{h})}{\partial \mathbf{h}} &= \frac{2(\mu_{v1} - \mu_{v2})\left(\frac{\partial \mu_{v1}}{\partial \mathbf{h}} - \frac{\partial \mu_{v2}}{\partial \mathbf{h}}\right)}{({\sigma_v}_1^2 + {\sigma_v}_2^2)} \\ &- \frac{(\mu_{v1} - \mu_{v2})^2 \left(\frac{\partial {\sigma_v}_1^2}{\partial \mathbf{h}} + \frac{\partial {\sigma_v}_2^2}{\partial \mathbf{h}}\right)}{({\sigma_v}_1^2 + {\sigma_v}_2^2)^2}. \end{aligned} \tag{4}$$

Before we can calculate this expression, we have to develop expressions for the feature mean and variance.

2.1. Feature mean

Assuming wide sense stationarity, the expected energy at the output of the filter is defined by

$$\begin{aligned} \mu_{vi} &= E\{v_i(m,n)\} = E\{w(k,l) * z_i(m,n)\} \\ &= E\{z_i(m,n)\} = E\left\{y_i^2(m,n)\right\} \\ &= E\left\{(\mathbf{h}^T \mathbf{x}_i(m,n))^2\right\} = \mathbf{h}^T \mathbf{R}_{xxi} \mathbf{h}, \end{aligned}$$

where

$$\mathbf{R}_{xxi} = E\left\{\mathbf{x}_i(m,n)\mathbf{x}_i^T(m,n)\right\}.$$

The partial derivative of μ_{vi} with respect to the filter vector $\mathbf{h}$ is needed in the subsequent sections

As shown in [9], if a matrix $\mathbf{A}$ is symmetric, then

$$\frac{\partial \mathbf{h}^T \mathbf{A} \mathbf{h}}{\partial \mathbf{h}} = 2\mathbf{A}\mathbf{h}, \tag{5}$$

hence, since $\mathbf{R}_{xxi}$ is symmetric we get

$$\frac{\partial \mu_{vi}}{\partial \mathbf{h}} = 2\mathbf{R}_{xxi}\mathbf{h}. \tag{6}$$

2.2. Feature variance

The variance of the feature image, $v_i(m,n)$, is

$$\begin{aligned} {\sigma_v}_i^2 &= E\left\{(v_i(m,n) - \mu_{vi})^2\right\} \\ &= \mathbf{w}^T E\left\{\mathbf{z}_i(m,n)\mathbf{z}_i^T(m,n)\right\} \mathbf{w} \\ &- 2\mu_{vi}\mu_{vi} + {\mu_{v}}_i^2 \\ &= \frac{1}{N_w^2} \sum_{k=0}^{N_w-1} \sum_{l=0}^{N_w-1} r_{zzi}(k,l) - {\mu_v}_i^2, \end{aligned} \tag{7}$$

where

$$\begin{aligned} r_{zzi}(k,l) &= E\{z_i(m,n)z_i(m+k,n+l)\} \\ &= E\{y_i(m,n)y_i(m+k,n+l) \\ &\qquad y_i(m,n)y_i(m+k,n+l)\}. \end{aligned} \tag{8}$$

For a one-dimensional Gaussian signal $y(k)$, we have $E\{y(k)y(l)y(m)y(n)\} = r_{yy}(k-m)r_{yy}(l-n) + r_{yy}(l-m)r_{yy}(k-n) + r_{yy}(k-l)r_{yy}(m-n)$ [10]. If we assume that the filter output, $y(m,n)$, has a Gaussian distribution the fourth order moment in Equation (8) is easily found, and we get

$$r_{zzi}(k,l) = 2r_{yyi}^2(k,l) + r_{yyi}^2(0,0).$$

Then, from Equation (7),

$$\begin{aligned} {\sigma_v}_i^2 &= \frac{1}{N_w^2} \sum_{k=0}^{N_w-1} \sum_{l=0}^{N_w-1} \left(2r_{yyi}^2(k,l) + r_{yyi}^2(0,0)\right) \\ &- {\mu_v}_i^2 \\ &= \frac{2}{N_w^2} \sum_{k=0}^{N_w-1} \sum_{l=0}^{N_w-1} r_{yyi}^2(k,l) + {\mu_v}_i^2 - {\mu_v}_i^2 \\ &= \frac{2}{N_w^2} \sum_{k=0}^{N_w-1} \sum_{l=0}^{N_w-1} r_{yyi}^2(k,l). \end{aligned} \tag{9}$$

Since

$$\begin{aligned} r_{yyi}(k,l) &= E\left\{y_i(m,n)y_i(m+k,n+l)\right\} \\ &= \mathbf{h}^T\mathbf{R}_{xxi}(k,l)\mathbf{h}, \end{aligned}$$

where

$$\mathbf{R}_{xxi}(k,l) = E\left\{\mathbf{x}_i(m,n)\mathbf{x}_i^T(m+k,n+l)\right\},$$

we have from Equation (9) that

$$\sigma_{v_i}^{\ 2} = \frac{2}{N_w^2}\sum_{k=0}^{N_w-1}\sum_{l=0}^{N_w-1}\left(\mathbf{h}^T\mathbf{R}_{xxi}(k,l)\mathbf{h}\right)^2.$$

We need the derivative of $\sigma_{v_i}^{\ 2}$, but since $(\mathbf{h}^T\mathbf{R}_{xxi}(k,l)\mathbf{h})$ is non-symmetric for $k,l \neq 0$, we cannot use Equation (5). However, for an arbitrary $L \times L$ matrix $\mathbf{A}$ we have

$$\begin{aligned} \frac{\partial}{\partial h_i}\left(\mathbf{h}^T\mathbf{A}\mathbf{h}\right) &= \frac{\partial}{\partial h_i}\left(\sum_{k=1}^{L}\sum_{l=1}^{L}a_{kl}h_kh_l\right) \\ &= \sum_{k=1,\ k\neq i}^{L}a_{ki}h_k + \sum_{l=1,\ l\neq i}^{L}a_{il}h_l \\ &+ \ 2a_{ii}h_i, \end{aligned}$$

and therefore

$$\frac{\partial \mathbf{h}^T\mathbf{A}\mathbf{h}}{\partial \mathbf{h}} = \mathbf{A}\mathbf{h} + \mathbf{A}^T\mathbf{h}.$$

Then

$$\begin{aligned} \frac{\partial \sigma_{v_i}^{\ 2}}{\partial \mathbf{h}} &= \frac{4}{N_w^2}\sum_{k=0}^{N_w-1}\sum_{l=0}^{N_w-1}\left(\mathbf{h}^T\mathbf{R}_{xxi}(k,l)\mathbf{h}\right) \\ &\quad \left(\mathbf{R}_{xxi}(k,l)\mathbf{h} + \mathbf{R}_{xxi}(k,l)^T\mathbf{h}\right). \end{aligned}$$

2.3. Gradient search optimization

For any given filter, $\mathbf{h}$, we may now compute the object function $J(\mathbf{h})$ and its partial derivative. Hence, we may perform gradient search with some gradient search algorithm to find the filter yielding the optimum filter $\mathbf{h}$.

2.4. Closed form optimization

With a gradient search optimization of the filter, we have no guarantee of convergence to a global optimum. In addition, depending on the surface of the search space, the convergence may be slow (remember that $\mathbf{h}$ typically has large dimensionality). Hence a closed form solution to the optimization problem would be desirable. We were, unfortunately, not able to find any closed form solution, i.e. any solution to the equation $\partial J(\mathbf{h})/\partial \mathbf{h} = 0$ from the results above.

If we assume that $y_i(m,n)$ is a separable AR(1) random process, we have $r_{yyi}(k,l) = r_{yyi1}(k)r_{yyi2}(l)$ where $r_{yyij}(k) = \rho_{ij}^k r_{yyij}(0)$ for $k > 0$, and hence

$$r_{yyi}(k,l) = \rho_{i1}^k\rho_{i2}^l r_{yyi}(0,0).$$

Thus, from Equation (9) we have

$$\begin{aligned} \sigma_{v_i}^{\ 2} &= \frac{2}{N_w^2}\sum_{k=0}^{N_w-1}\sum_{l=0}^{N_w-1}\rho_{i1}^{2k}\rho_{i2}^{2l}r_{yyi}^2(0,0) \\ &= \frac{2}{N_w^2}\frac{1-\rho_{i1}^{2N_w}}{1-\rho_{i1}^2}\frac{1-\rho_{i2}^{2N_w}}{1-\rho_{i2}^2}r_{yyi}^2(0,0) \\ &= \frac{2}{N_w^2}\frac{1-\rho_{i1}^{2N_w}}{1-\rho_{i1}^2}\frac{1-\rho_{i2}^{2N_w}}{1-\rho_{i2}^2}\left(\mathbf{h}^T\mathbf{R}_{xxi}\mathbf{h}\right)^2, \end{aligned}$$

which, since $\mathbf{R}_{xxi}$ is symmetric, has the derivative (Equation (5))

$$\begin{aligned} \frac{\partial \sigma_{v_i}^{\ 2}}{\partial \mathbf{h}} &= \frac{8}{N_w^2}\frac{1-\rho_{i1}^{2N_w}}{1-\rho_{i1}^2}\frac{1-\rho_{i2}^{2N_w}}{1-\rho_{i2}^2} \\ &\quad (\mathbf{h}^T\mathbf{R}_{xxi}\mathbf{h})\mathbf{R}_{xxi}\mathbf{h} \\ &= \frac{4\sigma_{v_i}^{\ 2}}{\mu_{v_i}}\mathbf{R}_{xxi}\mathbf{h}. \end{aligned} \tag{10}$$

Inserting the expressions from Equations (6) and (10) into the expression for $\partial J(\mathbf{h})/\partial \mathbf{h}$, Equation (4), and equating to zero yields

$$(\mathbf{R}_{xx2}^{\ -1}\mathbf{R}_{xx1})\mathbf{h} = \frac{1+\frac{(\mu_{v1}-\mu_{v2})\sigma_{v2}^2}{(\sigma_{v1}^2+\sigma_{v2}^2)\mu_{v2}}}{1-\frac{(\mu_{v1}-\mu_{v2})\sigma_{v1}^2}{(\sigma_{v1}^2+\sigma_{v2}^2)\mu_{v1}}}\mathbf{h},$$

which we may rewrite as

$$(\mathbf{R}_{xx2}^{\ -1}\mathbf{R}_{xx1})\mathbf{h} = \lambda\mathbf{h} \tag{11}$$

with

$$\lambda = \frac{1+\frac{(\mu_{v1}-\mu_{v2})\sigma_{v2}^2}{(\sigma_{v1}^2+\sigma_{v2}^2)\mu_{v2}}}{1-\frac{(\mu_{v1}-\mu_{v2})\sigma_{v1}^2}{(\sigma_{v1}^2+\sigma_{v2}^2)\mu_{v1}}}.$$

We identify Equation (11) as an eigenvalue equation, where the filter, $\mathbf{h}$, is the eigenvector and λ is the eigenvalue. The optimal filter is therefore the eigenvector, $\mathbf{h}$, yielding maximum $J(\mathbf{h})$.

Note that the we do not expect the real world case to meet the assumptions above exactly, and we suspect that the assumption that $y_i(m,n)$ is a separable AR(1) process is the least realistic. However, we do expect that useful results may be obtained by making the above assumptions. Our experimental results indicate that such is the case.

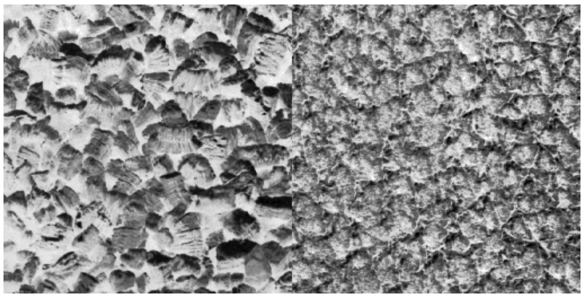


Figure 2. The texture pair D5-D92

2.5. Experimental results

We do not propose any way of finding optimal filter sizes, so for demonstration purposes we optimize an 8×8 finite impulse response (FIR) filter with a 40×40 smoothing operator. As the initial value for the gradient search approach, we use the closed form solution.

Examples of the feature surfaces with the two techniques described here for the texture pair D5-D92 in Figure 2 from the Brodatz [11] album are given in Figure 3. We see clearly that it is feasible to use a simple thresholding classification scheme and get a satisfying segmentation in both cases.

We evaluated the proposed techniques on a large set of Brodatz [11] textures, and were not able to find any situation where either of the techniques based on the Fisher criterion failed. However, as expected the gradient optimized solution was somewhat better (it is based on fewer restrictions) and using estimated means and variances from the true feature images, we found it to yield averagely around 50% higher $J(\mathbf{h})$, Equation (3).

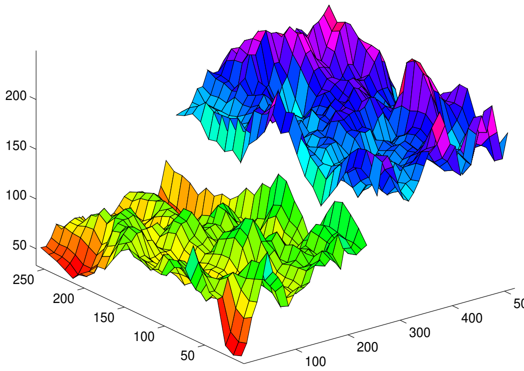
We expect that filters optimalized with respect to object functions which do not include the feature variances may yield more feature overlap. For the D5-D92 texture pair, this is also the case, and the features from the filter optimalized with respect to the object function (1) had a considerable feature overlap due to higher variance, as we see in Figure 4. We hasten to note that although the object function of Equation (1) was unsuccessful for this texture pair, is yielded good filters for several other texture pairs.

3. OTHER CRITERION FUNCTIONS

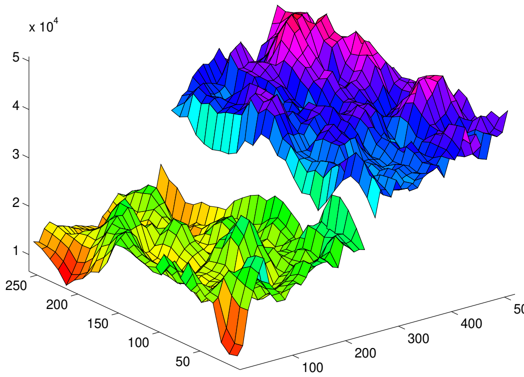
The closed form solution above was based on a specific object function, the Fisher criterion. In this section, we'll examine a broader range of criterion functions.

3.1. A generalized closed form solution

We'll start by showing that any object function on the feature means and variances may yield the same eigenproblem as we find in the closed form solution for the Fisher criterion.



(a)



(b)

Figure 3. The feature surface from the closed form optimization (a), and from the gradient optimization (b). The discontinuity in the surface indicates the true texture boundary.

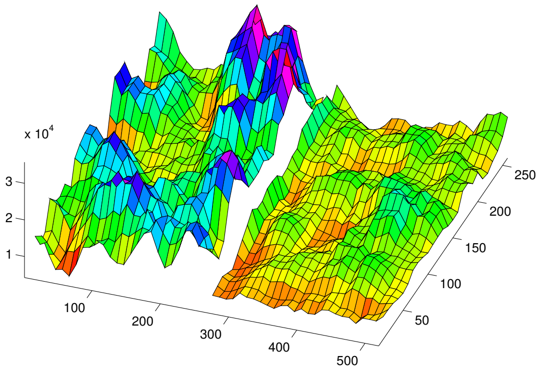


Figure 4. The feature surface from the filter optimalized with respect to the energy ratio object function. The discontinuity in the surface indicates the true texture boundary.

In Section 2.1 we found that the feature mean, μ_{v_i}, is a function of $(\mathbf{h}^T\mathbf{R}_{xxi}\mathbf{h})$. Furthermore, in Section 2.4, we found that the feature variance, $\sigma_{v_i}^2$, is a function of $(\mathbf{h}^T\mathbf{R}_{xxi}\mathbf{h})$, if we assume $y_i(m,n)$ to be a separable AR(1) process. Hence, *any* object function on the feature means and variances may be expressed as $f\left((\mathbf{h}^T\mathbf{R}_{xx1}\mathbf{h}), (\mathbf{h}^T\mathbf{R}_{xx2}\mathbf{h})\right)$, given the appropriate assumptions. The optimal solution then entails the solution of

$$\frac{\partial f\left((\mathbf{h}^T\mathbf{R}_{xx1}\mathbf{h}), (\mathbf{h}^T\mathbf{R}_{xx2}\mathbf{h})\right)}{\partial \mathbf{h}} = 0.$$

which can be expanded to

$$\begin{aligned} &\frac{\partial f\left((\mathbf{h}^T\mathbf{R}_{xx1}\mathbf{h}), (\mathbf{h}^T\mathbf{R}_{xx2}\mathbf{h})\right)}{\partial(\mathbf{h}^T\mathbf{R}_{xx1}\mathbf{h})}\frac{\partial(\mathbf{h}^T\mathbf{R}_{xx1}\mathbf{h})}{\partial \mathbf{h}} \\ &= -\frac{\partial f\left((\mathbf{h}^T\mathbf{R}_{xx1}\mathbf{h}), (\mathbf{h}^T\mathbf{R}_{xx2}\mathbf{h})\right)}{\partial(\mathbf{h}^T\mathbf{R}_{xx2}\mathbf{h})} \\ &\quad \frac{\partial(\mathbf{h}^T\mathbf{R}_{xx2}\mathbf{h})}{\partial \mathbf{h}}. \end{aligned} \tag{12}$$

Since $\mathbf{R}_{xxi}$ is symmetric, Equations (5) and (12) yield

$$\begin{aligned} &{\mathbf{R}_{xx2}}^{-1}\mathbf{R}_{xx1}\mathbf{h} \\ &= -\frac{\partial f\left((\mathbf{h}^T\mathbf{R}_{xx1}\mathbf{h}), (\mathbf{h}^T\mathbf{R}_{xx2}\mathbf{h})\right)}{\partial(\mathbf{h}^T\mathbf{R}_{xx2}\mathbf{h})} \\ &\quad \left(\frac{\partial f\left((\mathbf{h}^T\mathbf{R}_{xx1}\mathbf{h}), (\mathbf{h}^T\mathbf{R}_{xx2}\mathbf{h})\right)}{\partial(\mathbf{h}^T\mathbf{R}_{xx2}\mathbf{h})}\right)^{-1}\mathbf{h}, \end{aligned}$$

and since

$$\frac{\partial f\left((\mathbf{h}^T\mathbf{R}_{xx1}\mathbf{h}), (\mathbf{h}^T\mathbf{R}_{xx2}\mathbf{h})\right)}{\partial(\mathbf{h}^T\mathbf{R}_{xx1}\mathbf{h})}$$

is a scalar, we may rewrite this as

$$({\mathbf{R}_{xx2}}^{-1}\mathbf{R}_{xx1})\mathbf{h} = \lambda\mathbf{h}, \tag{13}$$

with

$$\begin{aligned} \lambda &= -\frac{\partial f\left((\mathbf{h}^T\mathbf{R}_{xx1}\mathbf{h}), (\mathbf{h}^T\mathbf{R}_{xx2}\mathbf{h})\right)}{\partial(\mathbf{h}^T\mathbf{R}_{xx2}\mathbf{h})} \\ &\quad \left(\frac{\partial f\left((\mathbf{h}^T\mathbf{R}_{xx1}\mathbf{h}), (\mathbf{h}^T\mathbf{R}_{xx2}\mathbf{h})\right)}{\partial(\mathbf{h}^T\mathbf{R}_{xx2}\mathbf{h})}\right)^{-1}. \end{aligned}$$

We identify Equation (13) as the same eigenvalue equation as we had in the previous closed form solution, but with a different expression for λ.

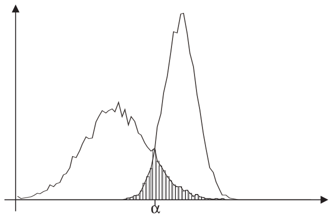


Figure 5. Example feature distributions for two textures. The error probability is illustrated by the shaded area.

3.2. Minimum classification error probability

A special form of the generalized solution above may arise if we use the classification error probability as our criterion function. If the feature probability density functions (PDF) are as given in Figure 5, the classification error probability, P_ϵ, is as illustrated by the area of the shaded region.

If we assume that the $y_i(m,n)$'s are Gaussian, the PDF of $y_i(m,n)$ is determined by its mean, μ_{y_i}, and variance, $\sigma_{y_i}^2$. If we assume that the $x_i(m,n)$'s are zero mean, we have $\mu_{y_i} = E\{y_i(m,n)\} = E\left\{\mathbf{h}^T\mathbf{x}_i(m,n)\right\} = 0$, and since $v_i(m,n)$ is a window-averaged version of $z_i(m,n) = y_i^2(m,n)$ with a fixed averaging operator, the PDF of $v_i(m,n)$ is determined by only $\sigma_{y_i}^2$. We denote the PDF of $v_i(m,n)$ by $p_{v_i}(\xi) = p_v(\xi, \sigma_{y_i}^2)$.

Assume that $p_v(\xi, \sigma_{y_1}^2)$ and $p_v(\xi, \sigma_{y_2}^2)$ have one point of intersection, $\alpha = \alpha(\sigma_{y_1}^2, \sigma_{y_2}^2)$, as illustrated in Figure 5. Assuming known a priori probabilities P_1 and P_2, the error probability, illustrated by the shaded area in Figure 5, is given by

$$\begin{aligned} P_\epsilon &= P_2\int_{-\infty}^{\alpha} p_v(\xi, \sigma_{y_2}^2)d\xi \\ &+ P_1\int_{\alpha}^{\infty} p_v(\xi, \sigma_{y_1}^2)d\xi. \end{aligned}$$

Finding the filter, $\mathbf{h}$, yielding minimum P_ϵ entails the solution of

$$\frac{\partial P_\epsilon}{\partial \mathbf{h}} = 0.$$

Since all the input parameters to $\frac{\partial}{\partial \mathbf{h}}P_\epsilon$ are functions of $\sigma_{y_i}^2$, and since

$$\begin{aligned} \sigma_{y_i}^2 &= E\left\{(y_i(m,n) - \mu_{y_i})^2\right\} \\ &= E\left\{y_i^2(m,n)\right\} \\ &= (\mathbf{h}^T\mathbf{R}_{xxi}\mathbf{h}), \end{aligned}$$

P_ϵ is a function of $(\mathbf{h}^T\mathbf{R}_{xx1}\mathbf{h})$ and $(\mathbf{h}^T\mathbf{R}_{xx2}\mathbf{h})$, thus $P_\epsilon = f\left((\mathbf{h}^T\mathbf{R}_{xx1}\mathbf{h}), (\mathbf{h}^T\mathbf{R}_{xx2}\mathbf{h})\right)$. Hence the

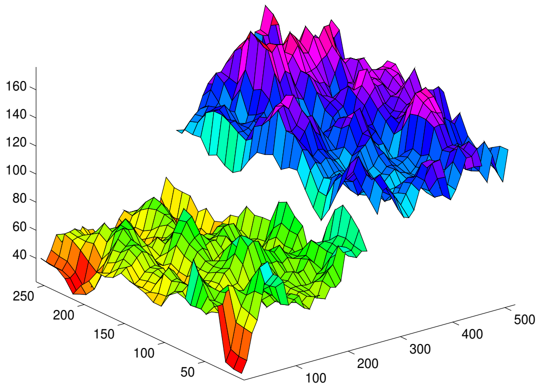


Figure 6. The feature surface from the closed form optimization of the filter yielding minimum classification error probability. The discontinuity in the surface indicates the true texture boundary.

optimal filter is found as one of the solutions to the eigenproblem $(\mathbf{R}_{xx2}^{-1}\mathbf{R}_{xx1})\mathbf{h} = \lambda\mathbf{h}$, as shown in Section 3.1.

The discussion above says nothing about how to pick the filter yielding minimum classification error from the set of eigenvectors. One practical solution to this is to test all candidates, and select the filter yielding minimum feature histogram overlap. The experimental results from this approach for the texture pair D5-D92 is shown in Figure 6. Tests on a large dataset indicated that this approach is also very robust.

3.3. A unified framework

As we have seen, all closed form solutions entail the solution of the eigenproblem $(\mathbf{R}_{xx2}^{-1}\mathbf{R}_{xx1})\mathbf{h} = \lambda\mathbf{h}$. This eigenproblem also occurs as the solution to the optimization with respect to the object function (1) [5], and if we apply the results of Section 3.1, we see that the optimization of the object function (2) also entails this eigenproblem.

Hence, we see that the closed form solution of various optimization problems all entail the same eigenproblem. Note, however, that different eigenvectors may be selected as the optimal filters in the various cases.

4. CONCLUSION

We have developed several new approaches to optimal feature extraction for two textures with one filter, based on modelling the textures with their autocorrelation functions. First we developed one gradient search and one closed form technique for the design of filters optimal with respect to the Fisher criterion. Then we extended the closed form solution to a more generalized formulation, including basically the same mathematical technique as for the Fisher criterion. Finally we proved that also the minimum classification error criterion may fit into this generalized formulation.

REFERENCES

[1] D. F. Dunn, W. E. Higgins, and J. Wakeley, "Texture segmentation using 2-D Gabor elementary functions," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, pp. 130–149, Feb. 1994.

[2] A. Teuner, O. Pichler, and B. J. Hosticka, "Unsupervised texture segmentation of images using tuned matched Gabor filters," *IEEE Trans. Image Processing*, vol. 43, July 1995.

[3] J. H. Husøy, T. Randen, and T. O. Gulsrud, "Image texture classification with digital filter banks and transforms," in *Proc. SPIE International symposium on optical instrumentation and applied science: Applications of Digital Image Processing XVI*, (San Diego), pp. 260 – 271, July 1993.

[4] A. K. Jain and K. Karu, "Automatic filter design for texture discrimination," in *Proc. Int. Conf. on Patt. Rec.*, pp. 454–458, Int. Assoc. for Patt. Rec., Oct. 1994.

[5] A. Mahalanobis and H. Singh, "Application of correlation filters for texture recognition," *Applied Optics*, vol. 33, no. 11, pp. 2173–2179, 1994.

[6] M. Unser, "Local linear transforms for texture measurements," *Signal Processing*, vol. 11, no. 1, pp. 61 – 79, 1986.

[7] K. Fukunaga, *Statistical Pattern Recognition*. San Diego: Academic Press, 2 ed., 1990.

[8] T. Randen and J. H. Husøy, "Multichannel filtering for image texture segmentation," *Optical Engineering*, vol. 33, pp. 2617–2625, Aug. 1994.

[9] S. Haykin, *Adaptive filter theory*. Englewood Cliffs: Prentice Hall, 2 ed., 1991.

[10] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*. Englewood Cliffs: Prentice-Hall, 1975.

[11] P. Brodatz, *Textures: A photographic album for artists and designers*. NY: Dover, 1966.