A MULTI-RESOLUTION FEATURE REDUCTION TECHNIQUE FOR IMAGE SEGMENTATION WITH MULTIPLE COMPONENTS

Michael Unser and Murray Eden

Biomedical Engineering and Instrumentation Branch, Bldg 13, Room 3W13,
National Institutes of Health, Bethesda MD 20892, USA.

Abstract - This paper presents a linear feature reduction technique for multi-component or textured image segmentation. The transformation matrix is computed by simultaneously diagonalizing scatter matrices evaluated at two different spatial resolutions. Under reasonable conditions, this transform closely approximates the generalized Fisher's linear discriminants which are optimal for region separability. Experimental examples suggest that this technique is superior to the KLT for texture segmentation.

I. INTRODUCTION

The goal of image segmentation is to divide an image into regions that are uniform or homogeneous with respect to certain characteristics. Most techniques use a single scalar property, such as the gray level [1]. There is considerable interest in methods that can handle additional information as, for example, is to be found in multiple component images. In such a representation, each pel is characterized by a feature vector. Typical examples are color images with three color variables, multiband images such as LANDSAT or Thematic Mapper, or images of distributions of chemical elements in electron micrographic sections [2].

Local texture properties may be represented in a similar fashion, deriving feature vectors by means of local linear transforms [3]. In this approach, the input image is first processed by a bank of filters associated with some linear transform; for example, the Hadamard or Sine transform. A non-linear operator (absolute value or square) is then applied to each filtered channel and is followed by a lowpass filter. This procedure yields a multi-component characterization of texture properties in terms of local image statistics associated to some window centered on the current spatial index. Other approaches to texture analysis and segmentation are summarized in [4].

In any of these applications, the use of multiple features results in a much greater complexity in decision making. It is therefore advisable to work in a lower order projection space. The most commonly used feature reduction technique is the Karhunen-Loève transform (KLT) [5]. Although this transform is optimal for minimum error data representation, it is usually not the most nearly adequate for discriminating between image regions. In the particular case where the mean vectors and covariance matrices associated to the different regions are known, a much more satisfactory feature reduction is obtained by using multiple or generalized Fisher's linear discriminant functions (GFLD) [6]. The GFLD is optimal for a large variety of separability criteria [7]. Unfortunately, this technique is not applicable in most practical situations, due to the lack of a priori knowledge.

Here, we present an alternative to the KLT that usually provides a much better approximation of the optimal linear discriminant functions without requiring any knowledge of the region statistics. This approach uses multi-component image representations at various levels of spatial resolution. The method is fully described in section 2 and its major properties are given in section 3. Finally, we present some examples of texture segmentation and compare the performance of our approach with the KLT.

II. MULTI-RESOLUTION FEATURE REDUCTION

The technique described here applies to the processing of an N component image $\{u_{k,l}\}$, where $u_{k,l}$ is an N-dimensional vector defined for every spatial index (k,l), and n_0 is the number of pels in the image. It is assumed that the spatial support R_0 on which the image is defined can be partitioned into r mutually exclusive homogeneous regions : $R_0 = R_1 \cup R_2 \cdots \cup R_r$, with corresponding number of pels $n_0 = n_1 + n_2 \cdots + n_r$, and mean vectors $E\{u_{k,l} | (k,l) \in R_n\} = \mu_n$ $(n=1,..,r)$. It is also required that adjacent regions differ in their mean values $\mu_n \neq \mu_m$ $(m \neq n)$, and that all covariance matrices are non-zero.

A. Multi-resolution representation

The first step is to compute a series of images at different resolutions, $\{u_{k,l}^{<i>}\}$, by processing the original multi-component image with a sequence of lowpass filters :

$$u_{k,l}^{<i>} = u_{k,l} * g_{k,l}^{<i>} \qquad (1)$$

The convolution kernels $\{g_{k,l}^{<i>}\}$, which are normalized to unity, are applied to all components simultaneously. The index i denotes the level of spatial resolution, and the initial level is $\{u_{k,l}^{<0>}\} = \{u_{k,l}\}$, with the convention that $g_{k,l}^{<0>} = \delta_{k,l}$. Typically, the smoothed sequences are computed iteratively from a cascade of operators. Therefore, for $j>i$, we have that

$$u_{k,l}^{<j>} = u_{k,l}^{<i>} * g_{k,l}^{<ij>} \qquad (2)$$

The convolution kernels may be viewed as estimation windows and are particularly useful for the evaluation of texture features that are not defined at the single pel level but always associated to some elementary region. In practice, we may choose a succession of Gaussian operators with a octave or half-octave scale progression. These may be implemented efficiently using the method described by Burt (or some slight modification for a half octave progression) that is based on the cascaded convolution with a separable gaussian-like kernel that is progressively expanded and filled with zeros to provide the desired scale progression [8].

B. Linear feature reduction

Assuming that the multi-component signal has a global zero mean, the scatter matrix at resolution level i is evaluated as

$$S_u^{<i>} = \frac{1}{n_0} \sum_{(k,l) \in R_0} u_{k,l}^{<i>} \, u_{k,l}^{<i>T} \qquad (3)$$

Our feature reduction techniques requires the estimation of two such matrices $S_u^{<i>}$ and $S_u^{<j>}$ at distinct levels of resolution i and j (for example, i=0 and j=1) and is determined by solving the generalized eigenvector problem :

$$\lambda^{ij} S_u^{<i>} t^{ij} = S_u^{<j>} t^{ij} \qquad (4)$$

An NxN transformation matrix $U = [t_1 \ldots t_N]^T$ is constructed from the N eigenvectors, which are ordered according to their decreasing eigenvalues. This transform, which is generally non-orthogonal, diagonalizes both scatter matrices $S_u^{<i>}$ and $S_u^{<j>}$. The eigenvalues $\{\lambda_m, m=1,..,N\}$, represent the ratio between the energies of the multi-component image representations at resolutions i and j projected on the axes specified by their corresponding eigenvectors. Since smoothing decreases signal variability and $i<j$, the eigenvalues will be smaller than one. Furthermore, the projected features are generally ordered according to their decreasing discrimination power. The number of significant features (M) is determined by retaining only those eigenvalues greater than a threshold α, which is the expected energy reduction obtained with noise only (e.g., $\alpha=1/2$ for a half-octave scale progression); this number is usually inferior or equal to $(r-1)$, as justified by properties C and E.

III. PROPERTIES

A. Invariance on linear transformations

Unlike the KLT, the transformation of the data defined by Eq. (4), which we refer to as the multi-resolution KLT (MKLT), is invariant to any non-singular linear transformation.

Proof : Let us consider the linearly transformed multi-component sequence $w_{k,l} = T u_{k,l}$, where T is a NxN full rank transformation matrix. We first note that the linear transformation can be applied at any level of the multi-resolution representation without affecting the end result, that is :

$$\begin{aligned} \boldsymbol{w}_{k,l}^{<i>} &= \boldsymbol{w}_{k,l} * g_{k,l}^{<i>} = (\boldsymbol{T}\, \boldsymbol{u}_{k,l}) * g_{k,l}^{<i>} \\ &= \boldsymbol{T}\, (\boldsymbol{u}_{k,l} * g_{k,l}^{<i>}) = \boldsymbol{T}\, \boldsymbol{u}_{k,l}^{<i>} \end{aligned} \quad (5)$$

Consequently, the scatter matrix of $\boldsymbol{w}_{k,l}^{<i>}$ is given by $\boldsymbol{S}_w^{<i>} = \boldsymbol{T}\, \boldsymbol{S}_u^{<i>}\, \boldsymbol{T}^{\mathrm{T}}$ and can be substituted in Eq.(4). It is then not hard to show that the eigenvalues of this modified equation and the diagonalized scatter matrices of the transformed data remain invariant.

B. Minimal energy reduction

Let us consider a single projected component $v_{k,l}^{<i>} = \boldsymbol{t}^{\mathrm{T}}\, \boldsymbol{u}_{k,l}^{<i>}$. The energy ratio of the sequences $\{v_{k,l}^{<j>}\}$ and $\{v_{k,l}^{<i>}\}$ is measured by $\lambda^{ij} = (\boldsymbol{t}^{\mathrm{T}}\boldsymbol{S}_u^{<j>}\boldsymbol{t})/(\boldsymbol{t}^{\mathrm{T}}\boldsymbol{S}_u^{<i>}\boldsymbol{t})$ and can be shown to be minimized when $\boldsymbol{t}$ is the first eigensolution of Eq.(4). Since $v_{k,l}^{<j>} = v_{k,l}^{<i>} * g_{k,l}^{<ij>}$, it follows that our feature reduction method will extract the components for which spatial smoothing produces the least energy reduction.

C. Approximation of the generalized linear Fisher-discriminant

In multiple discriminant analysis, the scatter matrix $\boldsymbol{S}$ is decomposed as the sum of a between-region scatter matrix $\boldsymbol{B}$ and within-region scatter matrix $\boldsymbol{W}$. The GFLD, which maximizes the between to within-region variance ratio, is obtained from the eigen-solutions of [6,7] :

$$\boldsymbol{B}\boldsymbol{t} = \beta \boldsymbol{W}\boldsymbol{t}. \quad (6)$$

The eigenvalues represent the between to within variance ratios along the axes of this transform. In the case unsupervised segmentation, $\boldsymbol{B}$ and $\boldsymbol{W}$ are unknown and this technique cannot be used. We will show, however, that the eigenvectors of Eq.(6) and Eq.(4) are equivalent, provided that :

$$\boldsymbol{B}_u^{<i>}\boldsymbol{t} = \beta^{<i>}\boldsymbol{W}_u^{<i>}\boldsymbol{t} \;=\; \boldsymbol{B}_u^{<j>}\boldsymbol{t} = \beta^{<j>}\boldsymbol{W}_u^{<j>}\boldsymbol{t} \quad (7)$$

This is equivalent to requiring that : (i) the between-region scatter matrices at both levels of resolution are equal, which is quite reasonable since smoothing only modifies the average in the border regions which generally account for a small portion of the total image area, and (ii) that the eigenvectors of (6) are identical (or at least not changed significantly) from one level to another, which is generally well satisfied.

Proof : Eq. (7) implies that :

$$[\boldsymbol{B}_u^{<i>} + \boldsymbol{W}_u^{<i>}]\, \boldsymbol{t} = (\beta^{<i>} + 1)\, \boldsymbol{W}_u^{<i>}\boldsymbol{t} \quad \text{(i)}$$

$$[\boldsymbol{B}_u^{<j>} + \boldsymbol{W}_u^{<j>}]\, \mathbf{t} = (\beta^{<j>} + 1)\, \boldsymbol{W}_u^{<j>}\boldsymbol{t} \quad \text{(ii)}$$

$$\boldsymbol{W}_u^{<i>}\boldsymbol{t} = (\beta^{<j>}/\beta^{<i>})\, \boldsymbol{W}_u^{<j>}\boldsymbol{t} \quad \text{(iii)}$$

Eq.(iii) can be substituted in (i). The next step is to isolate $\boldsymbol{W}_u^{<j>}\boldsymbol{t}$ on the right hand side of Eqs.(i) and (ii) which yields

$$\boldsymbol{W}_u^{<j>}\boldsymbol{t} = \frac{(\beta^{<i>}/\beta^{<j>})}{(\beta^{<i>} + 1)} [\boldsymbol{B}_u^{<i>} + \boldsymbol{W}_u^{<i>}]\boldsymbol{t} \quad \text{(iv)}$$

$$\boldsymbol{W}_u^{<j>}\boldsymbol{t} = \frac{1}{(\beta^{<j>} + 1)} [\boldsymbol{B}_u^{<j>} + \boldsymbol{W}_u^{<j>}]\, \boldsymbol{t} \quad \text{(v)}$$

We finally complete our proof by equating those two expressions and rewriting the resulting equation as

$$[\boldsymbol{B}_u^{<i>} + \boldsymbol{W}_u^{<i>}]\, \boldsymbol{t} = \lambda' \,[\boldsymbol{B}_u^{<j>} + \boldsymbol{W}_u^{<j>}]\, \boldsymbol{t} \quad \text{(vi)}$$

where λ' is given by

$$\lambda' = \frac{1 + 1/\beta^{<i>}}{1 + 1/\beta^{<j>}} \quad (8)$$

D. Relationship with the Karhunen-Loeve transform

The solution of Eq.(4) is equivalent to the KLT in the particular case where $\boldsymbol{S}_u^{<i>}$ is proportional to the identity matrix. This property leads to a particularly efficient implementation. In a first step, the KLT is computed at resolution level i and the rotated eigen-components are standardized by their eigenvalues, which results in an identity scatter matrix. The transformed components are then convolved with the kernel $\{g_{k,l}^{<ij>}\}$ to reach resolution level j. The MKLT is now given by the KLT associated to the scatter matrix of the transformed sequence. This procedure may be iterated for successive levels of resolution.

E. Ordering of the components of the MKLT

An advantage of the MKLT over the KLT is that the eigensolutions of Eq.(4) are almost always ordered according to their discriminability. If the conditions of property C are satisfied, the relationship between the eigenvalues of Eq.(4) and (6) is the following :

$$\lambda_n^{ij} = \frac{1 + 1/\beta_n^{<j>}}{1 + 1/\beta_n^{<i>}} = \frac{\beta_n^{<i>} + \alpha_n^{ij}}{\beta_n^{<i>} + 1}, \quad (n=1,..,N) \quad (9)$$

where $\alpha_n^{ij}=\beta_n^{<i>}/\beta_n^{<j>}$. If we now assume that α_n^{ij}, which is approximately equal to the ratio of the within region variances at resolution levels j and i, is equal to some constant, λ_n^{ij} is an increasing function of $\beta_n^{<i>}$ which implies that the eigen-solutions of Eq.(4) are ordered according to their decreasing within to between variance ratios. When the within-region variation is due to noise alone, the value of $\alpha_n^{ij}=\alpha$ can be derived analytically and used as a threshold for the determination of the number of significant components (M).

IV. EXPERIMENTAL EXAMPLE

To illustrate this technique, we have applied it to the segmentation of a 128x128 test image composed of two predefined texture regions (c.f. Fig. 1a-b). Both texture regions had their histograms equalized with 32 equiprobable gray level values to guarantee that segmentation is based on texture characteristics exclusively. After subtraction of the global mean, the image was filtered by the local 2x2 discrete Hadamard transform and thereafter rectified, which is the simplest application of the technique described in [3]. The corresponding multi-component image representation is shown in fig. 1-c1 to 1-c4. The first operator ($\begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix}$) is a lowpass filter while the three others are vertical ($\begin{vmatrix} 1 & -1 \\ 1 & -1 \end{vmatrix}$), horizontal ($\begin{vmatrix} 1 & 1 \\ -1 & -1 \end{vmatrix}$) and diagonal ($\begin{vmatrix} 1 & -1 \\ -1 & 1 \end{vmatrix}$) edge detectors, respectively. Two global indicators of individual channel performance have been included: η_i is the current relative energy contribution in channel i, and β_i is the corresponding between to within-region variance ratio computed using the pre-defined regions displayed in Fig. 1-b. This latter quantity is an objective measure of region separation. The channel histograms are also represented.

A multi-resolution image representation was obtained using an iterative Gaussian smoothing similar to the technique described by Burt [8]. The original scheme was slightly modified to provide a half octave scale progression. For the initial sequence $\{u_{k,l}^{<0>}\}$, the discrimination power is very small, although it can be seen that the texture regions differ slightly in their mean values. The effect of smoothing is to decrease the within-region feature variance as illustrated in Fig. 1-d1 to 1-d4 which displays the local texture features $\{u_{k,l}^{<3>}\}$, after 3 iterations of the Gaussian filter. The equivalent diameter of the spatial window is approximately 7. Channel No. 2 is the most discriminative with β_2=1.08. The first reduced components obtained with the the KLT and the MKLT are shown in Fig.1-e and 1-f, respectively, and must be compared to the optimal solution obtained with the GFLD (Fig.1-g) which was computed using the pre-defined texture region given in Fig.1-b. The performance of the KLT is fair with β_1=0.87 but is not better than channel 2 taken on its own. The MKLT was determined by simultaneously diagonalizing the scatter matrices at resolution level 2 and 3 (e.g. $S_u^{<2>}$ and $S_u^{<3>}$). With β_1= 1.67, it provides an excellent approximation of the GFLD (β^*=1.71). Unlike the KLT, the corresponding histogram is bimodal and makes accurate image segmentation feasible by simple thresholding.

In this example as well as in all other cases that we have considered using different textures, the discrimination power of the first component of the MKLT was always found to be superior to any individual channel or to any component of the KLT and was almost as good as that of the GFLD. The texture regions usually appeared to be more nearly uniform within the different image regions and the modes in histogram were always more pronounced. We have also found the MKLT to be quite robust in the sense that segmentation performance was less sensitive to changes in structural parameters such as the size and the texture properties of the image regions than for the KLT. It should be pointed out, however, that the method breaks down when the size of the averaging window approaches the size of the smallest regions in the image. This result is not surprising and is due to the fact that smoothing with too large a kernel tends to modify the region averages so that they can no longer be assumed to remain constant, as required for property C.

The MKLT does also perform well when there are more than two regions in an image. In such cases, the number of significant components (M) of the MKLT (or GFLD) can be greater than one ($M \leq \max\{r\text{-}1, N\}$), which usually requires the use of more sophisticated classification techniques than simple

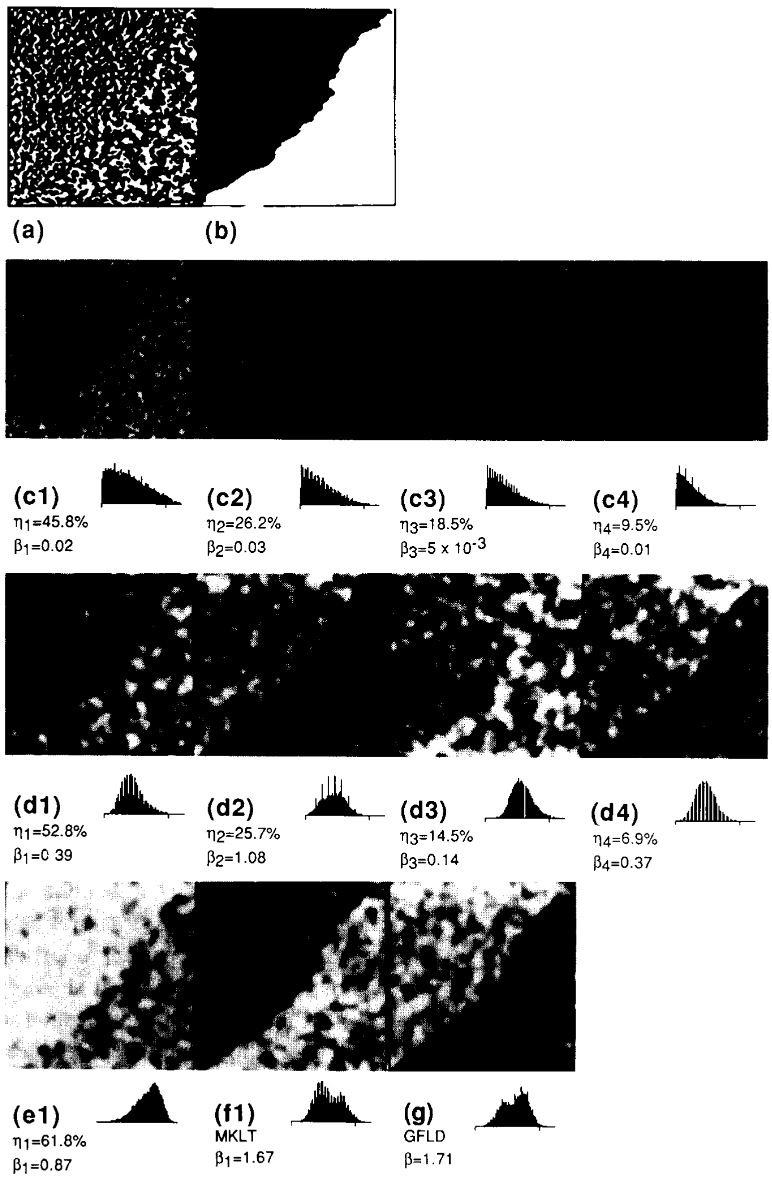

Fig. 1 : Example of feature reduction. (a) : 128x128 test image created using D57 and D9 Brodatz textures [10], (b) definition of image regions, (c1-4) : filtered channels using masks obtained from the 2x2 DHT , (d1-4) : feature planes at iteration 3 of the smoothing algorithm, (e) : first component of the KLT, (f) first component of the MKLT, (g) Fisher linear discriminant.

histogramming. The MKLT may also be used with other texture features than those discussed in the example. These features, however, should be evaluated from the succession of an operator (usually non-linear) working at the pel level and a smoothing window which provides a local estimate of the expected value of some region property. This particular decomposition defines a broad class of measurements that includes most of the features commonly used in texture analysis [9].

V. CONCLUSION

In this paper, we have discussed the use of feature reduction techniques to simplify segmentation of images with multiple components. We have described a new multi-resolution linear transform (MKLT) that appears to be more powerful than the conventional KLT. This transform has a number of very attractive properties such as its good approximation of the optimal generalized Fisher's linear discriminants and its ordering of the reduced components according to their separability. It is specially suited for segmentation by texture but is also applicable to all kinds of multi-component signals such color or multiband images.

We have presented an illustrative example for which efficient texture segmentation could be achieved by simple thresholding of a reduced component. This procedure could be potentially improved by using the MKLT in conjunction with more sophisticated pel classification schemes such as a coarse-fine strategy to refine the location of edges or relaxation labeling.

References

[1] R.M. Haralick and L.G. Shapiro, "Image segmentation techniques", *Computer Vision, Graphics and Image Processing*, vol. 29, 1985, pp. 100-132.

[2] R.D. Leapman, C.E. Fiori, K.E. Gorlen, C.C. Gibson, and C.R. Swyt, "Combined elemental and STEM imaging under computer control", *Ultramicroscopy* 12, 1981, pp. 281-292.

[3] M. Unser, "Local linear transforms for texture measurements", *Signal Processing*, Vol. 11, No. 1, July 1986, pp. 61-79.

[4] L.Van Gool, P. Dewaele, and A. Oosterlinck, "Texture analysis Anno. 1983", *Computer Vision, Graphics, Image Processing*, vol. 29, 1985, pp. 336-357,.

[5] S. Watanabe, "Karhunen-Loève Expansion and Factor Analysis", Trans. Fourth Prague Conf. on Information Theory, etc., 1965, Czechoslovak Academy of Sciences, Prague, 1967.

[6] R. O. Duda and P.E. Hart, *Pattern Classification and Scene Analysis*, John Wiley & Sons, New York, 1973.

[7] D.W. Peterson and R.L. Mattson, "A method of finding linear discrimant functions for a class of performance criteria", *IEEE Trans. Info. Theory*, IT-12, July 1966, pp. 380-387.

[8] P. J. Burt, "Fast algorithms for estimating local image properties", *Computer Graphics and Image Processing* Vol. 21, 1983, pp. 368-382,.

[9] M. Unser, "Thesis alert : Statistical characterization of texture with an application to automated inspection", *Signal Processing*, Vol. 8, No. 1, February 1985, p. 137.

[10] P. Brodatz, "Textures - A photographic album for artists and designers", Dover, New-York, 1966.