Texture segmentation using stable texture features

Sylvain Meignen, Valérie Perrier

Abstract. This paper describes a new unsupervised texture segmentation algorithm when the number of textures is a priori unknown. The basic idea of the algorithm is the study of the stability of specific features with texture sample size. which are obtained using Karhunen-Loeve transform of the vector of texture decomposition over specific filters. These filters are wavelet frames associated to orthogonal and biorthogonal filters. The end of the paper is devoted to the description of the segmentation algorithm, and some illustrative experiments on natural textures are presented.

§1. Introduction

Under its simplest form, the problem of texture discrimination is to design a method to decide whether two given samples are similar or significantly different. The first step is to find a qualitative description of textures. Characterizations based on multichannel or multiresolution analysis have received a lot of interest [1][4][6][9]. Orthogonal wavelet or orthogonal wavelet packet decompositions allow multiresolution analyses while preserving the volume of data and have been used for texture classification [3][6]. The main drawbacks of the orthogonal wavelet (packet) decompositions consist in the loss of translation invariance. Once a special kind of decomposition has been chosen (filtering step), a nonlinear operator is applied to the coefficients of the decomposition. Most of the time, it consists in computing local energies at each scale [9]. These quantities are not invariant with the size of the neighborhood used for their computation. This motivates the construction of a new nonlinear transform of the coefficients that will be stable with texture sample size. The paper is organized as follows: section 2 presents texture multiscale decomposition with tight frames, section 3 and 4 is devoted to the definition of new stable features (in a sense to clarify), filter design is explained in section 5 while the unsupervised segmentation algorithm is detailed in section 6. Texture segmentation results conclude the paper.

Curve and Surface Fitting: Saint-Malo 2002
XXX (eds.), pp. 1–3.

ISBN XXX.

§2. Filter bank and wavelet frames

Let I be an image of textures assumed to be a function f of $l_2(\mathbb{Z}^2)$ with support $[1,N]\times[1,N]$. f is filtered by $F=\{F_i\in l_2(\mathbb{Z}^2);\quad i=1,\cdots,P\}$ a frequency selective filter bank. Our approach for the design of filter F is based on dilation and translation of filters h and g that define perfect reconstruction filter banks of $l_2(\mathbb{Z}^2)$ [10], which satisfy standard quadrature mirror filter condition:

$$|\hat{h}(\omega)|^2+|\hat{h}(\omega+\pi)|^2=1$$

Filter h also satisfies condition $\hat{h}(0)=1$. A complementary high-pass filter g is obtained by symmetry with respect to $\frac{\pi}{2}$:

$$\hat{g}(\omega)=\hat{h}(\omega+\pi)$$

In this case, filters h and g are called conjugate mirror filters and h and g can both be even. Even filters are advantageous in alleviating boundary artifacts through simple method of mirror extension [4]. Using filters h and g , it is possible to define a tight frame of $l_2(\mathbb{Z}^2)$. Let $\|.\|$ be the norm in $l_2(\mathbb{Z}^2)$ and $<,>$ the corresponding scalar product. Then, a tight frame c_n of $l_2(\mathbb{Z}^2)$ is defined by:

$$\sum_n <f,c_n>^2=\|f\|^2 \quad \forall f\in l_2(\mathbb{Z}^2)$$

We have the equivalence [8]:

Theorem 1. *The family*

$$F=\{h_J[n-l]h_J[m-k],\{g_j[n-l]h_j[m-k],\\ h_j[n-l]g_j[m-k],g_j[n-l]g_j[m-k]\}_{1\le j\le J}\}_{(l,k)\in Z^2}\}$$

is a tight frame of $l_2(Z^2)$, for all $J\ge 1$ if and only if:

$$|\hat{h}(\omega)|^2+|\hat{g}(\omega)|^2=1 \tag{5}$$

where $h_{j+1}[k]=\sum_n h[n]h_j[k-2^jn]\quad g_{j+1}[k]=\sum_n g[n]h_j[k-2^jn]$ *and* $h_0=h$.

Such tight frame cannot be generated by even filters h and g of finite length that satisfy condition (5). Thus, we consider even filters of infinite length, i.e. Battle-Lemarié filters, which we truncate. An alternative is to use biorthogonal wavelet filters which are of finite length but do not satisfy condition (5). Battle-Lemarié filters are defined by:

$$\hat{h}(\omega)=\exp(\frac{-i\epsilon\omega}{2})\sqrt{\frac{H_{2m+2}(\omega)}{2^{2m+1}H_{2m+2}(2\omega)}}\quad \textit{with}\quad H_m(\omega)=\sum_{-\infty}^{\infty}\frac{1}{(\omega+2k\pi)^m}$$

A criterion to decide which filters are better for energy conservation is to compare $|\hat{h}(\omega)|^2 + |\hat{g}(\omega)|^2$ to 1. For Battle-Lemarié filters, the effect of truncation on energy conservation increases with m. We then compare the frequency selectivity of truncated Battle-Lemarié filters with 7 coefficients to biorthogonal cubic spline filter ($p = 4$, see [5] for a definition of biorthogonal filters). Truncation of Battle-Lemarié filters induces undesirable lobes in their frequency representation when m increases while Biorthogonal filters do not exhibit such lobes.

§3. Stability of local energies

Many segmentation algorithms [1][9] use, as features, local energies associated to images $\tilde{f}_i = f * F_i$, defined by:

$$E_i[n,m] = \frac{1}{\#B([n,m])} \sum_{[k,l]\in B([n,m])} \tilde{f}_i[k,l]^2 \tag{1}$$

where $\#B([n,m])$ denotes the size of $B([n,m])$. When E_i is used for texture segmentation purpose $\#B([n,m])$ should be chosen with respect to the frequency content of the channel defined by $\tilde{f}_i$ [7]. To study the stability of the energy in the different frequency channels, we assume that $\#B([n,m])$ is chosen independent on i. For large frequency component, the estimation of $E_i[n,m]$ is not relevant on small neighbourhoods $B([n,m])$. In most images of textures, the greatest part of the energy is concentrated on the smallest scales of the decomposition (small frequencies), which are also the most stable when $\#B([n,m])$ varies. We are going to introduce a transform on vector $(\tilde{f}_i)_{1\le i\le 3J+1}$ that provides stable features based on the energy content of images E_i.

§4. Features and stability definitions

In this section, we consider an image I of size $N \times N$ made of a single texture. Karhunen-Loeve transform of vectors $\tilde{f}[n,m]$, assumed to have a zero mean, is its projection on the basis of eigenvectors of matrix:

$$S = \frac{1}{N^2} \sum_{[n,m]\in supp(I)} \tilde{f}[n,m]\tilde{f}[n,m]^T \tag{2}$$

where $supp(I)$ denotes the support of image I. If the rank of S is P, which is the case in practice for $N^2 >> P$, the eigenvectors of S, denoted by $\{u_i, \quad i = 1,\cdots,P\}$ correspond to positive eigenvalues $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_P$. We denote S^{I_k} the covariance matrix computed on any subimages I_k of I. For two samples $I_{k'}$ and I_k of I, both the eigenvectors and the eigenvalues of respective matrices S^{I_k} and $S^{I_{k'}}$ should be identical. They are, in practice, strongly dependent on the size of the subimages.

Let assume $N = 2^m$, and split I into non overlapping subimages of equal size $K \times K$, $K = 2^{m'}$ with $m' < m$. Denote A_K the corresponding subset of images, and $\left\{S^k, k \in A_K\right\}$ the set of covariance matrices. We also define the set of eigenvectors u_i^k and the corresponding eigenvalues λ_i^k. Eigenvectors u_i^k should be collinear to eigenvector u_i. To measure this, we introduce the following definition of the stability of eigenvectors:

Definition 1. *Eigenvector u_i is more stable than eigenvector u_j at scale K if*

$$\sum_{k \in A_K} | < u_i^k, u_i > | > \sum_{k \in A_K} | < u_j^k, u_j > |$$

Unfortunately, this definition may not be useful in practice since we often compute eigenvectors locally with no idea of what they are at bigger scales. We now define $D_K = \{(k,p) \in A_K, k \neq p\}$ and:

$$\boldsymbol{S}_1(A_K, i) = \frac{1}{\# D_K} \sum_{(k,p) \in D_K} | < u_i^k, u_i^p > |$$

Using these notations we consider the following definition of the relative stability of eigenvectors at a given scale:

Definition 2. *Eigenvector u_i is said to be relatively more stable than eigenvector u_j at scale K if*

$$\boldsymbol{S}_1(A_K, i) > \boldsymbol{S}_1(A_K, j) \tag{3}$$

Notice that $\boldsymbol{S}_1(A_K, 1) = 1$ means the ith eigenvector is identical on any subimages of A_K. If relation (3) is valid for all K, we say that u_i is unconditionally more stable than u_j. We define stability and relative stability of eigenvalues as follows:

Definition 3. *Eigenvalue λ_i is more stable than eigenvalue λ_j at scale K if*

$$\sum_{k \in A_K} \frac{|\lambda_i^k - \lambda_i|}{\lambda_i^k + \lambda_i} < \sum_{k \in A_K} \frac{|\lambda_j^k - \lambda_j|}{\lambda_j^k + \lambda_j}$$

Let now consider that:

$$\boldsymbol{S}_2(A_K, i) = \frac{1}{\# D_K} \sum_{(k,p) \in D_K} \frac{|\lambda_i^k - \lambda_i^p|}{\lambda_i^k + \lambda_i^p}$$

Definition 4. *Eigenvalue λ_i is said to be relatively more stable than eigenvalue λ_j at scale K if*

$$\boldsymbol{S}_2(A_K, i) < \boldsymbol{S}_2(A_K, j) \tag{4}$$

§5. Stability of features with texture sample size

We experimentally study the stability of the eigenvalues and of the eigenvectors of S using different filter banks F, on various examples of textures extracted from Vitex database [11]. We investigate the stability of the first three eigenvectors and of the corresponding eigenvalues, through the computation of $\boldsymbol{S}_1(A_K, i)$ (3) and $\boldsymbol{S}_2(A_K, i)$ (4) for various K. The filters h and g we use are truncated Battle-Lemarié filters ($m = 1$, $m = 3$ and $m = 5$) and biorthogonal filter ($p = 4$). We notice that u_1 is unconditionally more stable than u_2 and u_3. Hence, if we consider stability, the first eigenvector u_1 seems to be a good candidate for the characterization of textures. This property of stability is is in accordance with the remark of section 2 which says that the most energetic subspaces correspond to small scales. Indeed, the direction of eigenvector u_1 is mainly given by the most energetic subspaces (i.e. those associated to small scales) and therefore are less sensitive to texture sample size.

We then study the influence of J on the stability of eigenvector u_1. The direction of u_1 is given by the most energetic subspace associated to the decomposition on F. When the depth of decomposition is too small, the energy is mainly contained in the approximation subspace (i.e. subspace generated by $h_J[n]h_J[m]$). In this case, the direction of the first eigenvector is very stable, its components being approximately $(1, 0, \cdots, 0)$, whatever texture sample size. An appropriate depth of decomposition is determined saying that the energy contained in the approximation subspace is smaller than that in detail subspaces (i.e. the other subspaces) which translates into the appropriate depth of decomposition is the smallest J such that the first component of u_1 becomes smaller than 0.5. Looking at the stability of eigenvectors, this is interpreted as follows: when the energy is not contained in the approximation subspace, deeper decomposition of the image do not alter $\boldsymbol{S}_1(A_K, 1)$ since new components that appear in u_1 are of weak amplitude.

$\boldsymbol{S}_1(A_K, 1)$ must also increase with K: textures being all the more stable that the sample is large. The study of $\boldsymbol{S}_1(A_K, 1)$ for different F shows that biorthogonal filters as well as Battle-Lemarié filters with $m = 3$ or $m = 5$ are not good candidates because $\boldsymbol{S}_1(A_K, 1)$ may not increase with K.

In applications, we have noticed that choosing a Battle-Lemarié filter with $m = 1$ and 7 coefficients is a good trade-off. With such a filter, we determine the appropriate J for the decomposition of textures over F. With the depth of decomposition thus determined, texture scale is defined as the smallest texture sample size for which the first eigenvector of the covariance matrix S can be considered as stable.

To summarize, properties required for filters h and g and for filter bank F to generate stable u_1's when texture sample size varies are:

- Frequency selectivity for h and g
- Appropriate depth of decomposition J for F
- Relation $|\hat{h}(\omega)|^2 + |\hat{g}(\omega)|^2 \approx 1$ must be satisfied for all ω

§ 6. Unsupervised texture segmentation algorithm

In the previous section, we showed that it is possible, using an appropriate filter bank F, to get a stable texture feature: the eigenvector u_1. However, it may not be enough discriminative for texture problem. We notice that it is profitable to use the corresponding eigenvalue as extra feature.

From this, we derive a segmentation algorithm made of three steps: a hierarchical splitting step, an agglomerative merging step and a final segmentation step. In each step of the algorithm, we need a distance. For both the hierarchical splitting step and the merging step we use the following distance between two subimages I_β and $I_{\beta'}$:

$$d(I_\beta, I_{\beta'}) = \left(1 - |< u_1^{I_\beta}, u_1^{I_{\beta'}} >|\right) \frac{|\lambda_1^{I_\beta} - \lambda_1^{I_{\beta'}}|}{\lambda_1^{I_\beta} + \lambda_1^{I_{\beta'}}} \frac{|\sum_{j>1} \lambda_j^{I_\beta} - \sum_{j>1} \lambda_j^{I_{\beta'}}|}{\sum_{j>1} \lambda_j^{I_\beta} + \sum_{j>1} \lambda_j^{I_{\beta'}}}$$

where $u_i^{I_\beta}$ and $\lambda_i^{I_\beta}$ stand for eigenvector and eigenvalues of matrix S^{I_β}. The choice of distance d is motivated by the fact that u_1 has to be stable within a texture (first term). In such a case, the first eigenvalues must also be close (second term). We, finally, add a term corresponding to the difference between the energy in the subspace orthogonal to u_1 (third term). In the final segmentation step, we will either use distance d or the maximum likelihood distance depending on the textures under consideration.

Let us now detail the algorithm. We consider an image I of size $N \times N$, where N is power of 2, containing several textures T_i whose number is a priori unknown.

Hierarchical splitting step

- Split image I into 4 subimages of equal size
- Compute eigenvectors u_1 over I and over subimages (I_1, I_2, I_3, I_4)
- Compute distance d between I_β and I for $\beta \in \{1, \cdots, 4\}$
- Given a threshold α and a minimum size s_{min} for the sample, the algorithm stops if one of the following conditions is satisfied:

  1. $$\max_{I_\beta} d(I_\beta, I) < \alpha$$

  2. $\#I$ equals s_{min}
- Otherwise, the algorithm splits the subimages recursively

Agglomerative merging step

The problem is to merge the subimages determined at the end of the first step into the exact number of textures. For a given scale K, we determine the set $\tilde{A_K}$ of subimages of A_K that are unsplit at the end of the first step. We then define a relation p on this set of subimages, as follows:

for all $(I_\beta, I_{\beta'}) \in \tilde{A_K}$

if $d(I_\beta, I_{\beta'}) < \delta$, $d(I_\beta \cup I_{\beta'}, I_\beta) < \delta$, $d(I_\beta \cup I_{\beta'}, I_{\beta'}) < \delta$ *and* $(I_\beta \neq I_{\beta'})$

$$p(I_\beta, I_{\beta'}) = d(I_\beta, I_{\beta'})$$

else $p(I_\beta, I_{\beta'}) = 0$

Computing distances between $I_\beta \cup I_{\beta'}$ and I_β and between $I_\beta \cup I_{\beta'}$ and $I_{\beta'}$ insures that the union preserves the features. The merging step then proceeds:

finish = false

$G_K = \tilde{A_K}$

while finish = false

- *for* $I_\beta \in G_K \quad mark[I_\beta] = 0$
- *nmark = true*

  nb = 0

  while nmark = true

  Find $(I_1, I_2) \in {G_K}^2$ *such that:* $(I_1, I_2) = \underset{I_\beta, I_{\beta'} \in G_K}{argmax}\; p(I_\beta, I_{\beta'})$

  with $mark[I_\beta] = 0$ *or* $mark[I_{\beta'}] = 0$

  if $p(I_1, I_2) \neq 0$

  if $mark[I_1] + mark[I_2] = 0$

  $nb = nb + 1,\ mark[I_1] = mark[I_2] = nb$

  else $mark[I_1] = \max(mark[I_1], mark[I_2]),\ mark[I_2] = mark[I_1]$

  else nmark = false
- *Compute features over the nb regions which define a new* G_K
- *Compute the proximity relation over* G_K

  if $\forall I_\beta, I_{\beta'} \in G_K \quad p(I_\beta, I_{\beta'}) = 0$

  finish = true

An advantage of this algorithm is to allow the merging of disconnected subimages. At each scale K, we dispose of a set of textures which are merged by decreasing K. To get a precise definition of the features, we only assess texture content using the largest scales. As we know now what textures T_i are in the image, the segmentation step can be viewed as a supervised segmentation step. The distance to be used in the segmentation step strongly depends on the Gaussian or non Gaussian character of texture features. Distance d performs well when the image is made of non Gaussian textures (i.e. when $\tilde{f}$ is a non Gaussian vector). When one needs to compare Gaussian textures, the maximum likelihood distance performs better. It is defined as follows:

$$\tilde{d}(\tilde{f}[n,m],T_k) = (\tilde{f}[n,m] - \mu^{T_k})^T {S^{T_k}}^{-1}(\tilde{f}[n,m] - \mu^{T_k}) + \log(|det(S^{T_k})|)$$

where μ^{T_k} is the average of $\tilde{f}[n,m]$ over T_k. Each point is assigned a label considering the minimum of the maximum likelihood distance. When the image contains textures for which $\tilde{f}[n,m]$ is either Gaussian or non Gaussian the maximum likelihood distance may be used in the following way:

$$\bar{d}(\tilde{f}[n,m],T_j) = \frac{1}{\#B([n,m])} \sum_{[k,p]\in B([n,m])} \tilde{d}(\tilde{f}[k,p],T_j)$$

where $\#B([n,m])$ should be chosen with respect to the scales of the textures in the image. Segmentation is then performed using "Kmeans" algorithm with initial guess (μ^{T_j}, S^{T_j}).

§ 6. Segmentation results on synthetic images

The examples we consider are images displayed in the first column of Figure 1. Filter h we used in the frame representation (see Theorem 1) is truncated Battle-Lemarié linear filter with 7 coefficients. Only the two biggest values for parameter K determined at the end of the splitting step are taken into consideration in the merging step. All the others parameters are detailed in the caption of Figure 1.

§7. Conclusion

The scope of this paper was to introduce an unsupervised texture segmentation algorithm that is based on the stability with texture sample size of specific features. We investigated the stability of the eigenvectors of the covariance matrix associated to the Karhunen-Loeve transform of the decomposition of images of textures on an tight frame. This study of eigenvectors helped us to determine what filter should be preferred. In that context, we also gave arguments for the choice of the appropriate depth of decomposition for the frame. We put forward that the first eigenvector of the matrix associated to the Karhunen-Loeve transform is the most stable among the most energetic subspaces.

We derived from this a new unsupervised segmentation algorithm made of three steps: a splitting step, an agglomerative merging step and a final segmentation step.

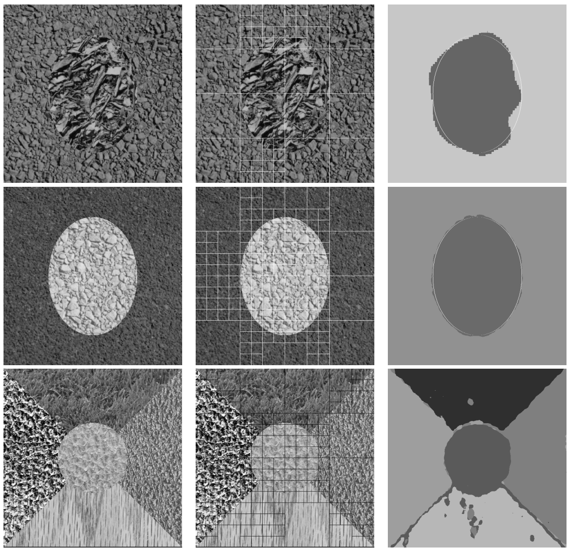

Fig. 1. Example of images of textures (first column). In the second column, are displayed splitting step results using distance d and first row, $\alpha = \delta = 0.3$ second row, $\alpha = \delta = 0.2$ third row, $\alpha = 0.2$ and $\delta = 0.4$, the neighborhood $B([n, m])$ for the computation of the feature is of size 11×11, s_{min} is set to 32×32 pixels each time. Third column displays final segmentation results, first row, using distance d (local computations are made on 64×64 squares), second and third row using maximum likelihood distances ($\tilde{d}$ and $\bar{d}$ respectively).

References

1. Bovik A.C., Clark M. and W.S. Geisler, Multichannel texture analysis using localized spatial filters, IEEE Trans. Patt. Anal. Machine Intell., vol. 12, no. 1, 55-73, 1990
2. Coroyer C. et P. Duvaut, Contribution of third order correlation tools for textures analysis, Traitement du signal, vol. 15,no. 2, 129-138, 1998.
3. Laine A. and J. Fan, Texture classification by wavelet packet signatures, IEEE Trans. Patt. Anal. Machine Intell., vol. 15, no. 11, 1186-1191, 1993.
4. Laine A. and J.Fan, Frame representation for texture segmentation, IEEE Trans. Image Processing, vol. 5, 771-780, 1996.

5. Mallat S., A wavelet tour of signal processing, Academic Press, 1998.
6. Mojsilovic A., Popovic M.V. and D.M. Rackov, On the selection of an optimal wavelet basis for texture characterization, IEEE Trans. Image Processing, vol. 9, no. 12, 2043-2050, 2000.
7. Randen T. and J.Husoy, Filtering for texture classification : a comparative study, IEEE Trans. Patt. Anal. Machine Intell., vol. 21, no. 4, 291-310, 1999.
8. O. Rioul, A discrete-time multiresolution theory, IEEE Trans. Signal Processing, vol. 41, no. 8, 2591-2606, 1993.
9. M.Unser, Texture classification and segmentation using wavelet frames, IEEE Trans. Image Processing, vol. 4, no. 11, 1549-1560, 1995.
10. M. Vetterli, Filter bank allowing perfect reconstruction, Signal Processing, vol. 10, no.3, 219-244, 1986.
11. VisTex, Color image database, http://www-white.media.mit.edu/vismod/imagery/VisionTexture, MIT Media Lab, 1995.

Sylvain Meignen and Valérie Perrier
LMC-IMAG
Univ. Joseph Fourier, BP 53
38041 Grenoble cedex 9, FRANCE
sylvain.meignen@imag.fr