On a Two-Dimensional
Maximum Entropy Spectral
Estimation Method for the
Texture-Image Analysis

by
C. H. Chen
Gia-Kinh Young
Department of Electrical Engineering
Southeastern Massachusetts University
North Dartmouth, Massachusetts 02747

The support of the Statistics and Probability Program of
the Office of Naval Research on this work is gratefully
acknowledged. The authors also would like to thank Mr.
N. A. Malik for his kind assistance and Mr. D. AbouJind for valuable discussions.
1. Introduction

Although it is generally recognized that texture images contain statistical, spectral and structural domain information, the use of spectral information alone can be quite effective in the texture-image analysis studies such as texture discrimination and segmentation. Bajcsy and Liberman [1] expressed the power spectrum in polar coordinates, then integrate over $r$ and $\theta$ to obtain the two one-dimensional functions. The location of peaks in these functions indicates prominent texture coarseness and directionality. Weszka et. al. [2] integrated the power spectrum within 16 spatial frequency zones which were combinations of four 1-octave frequency ranges and four $45^\circ$ orientation sectors. They also computed eight "contrast" measures based on the cooccurrence matrix, and obtained better discrimination than with the power spectrum measures. Laws [3] computed a number of energy measures by filtering the texture with sets of small linear operators, then squaring and summing the output of each filter. He reported better discrimination with the energy than with the cooccurrence measures.

A fundamental problem with the power spectrum analysis is the computational accuracy and computational complexity. For texture study, accurate power spectrum must be computed from the small image segments. In this case, the two-dimensional Fourier analysis cannot provide sufficient accuracy as the Fourier analysis is more accurate with a large number of pixels. The two-dimensional
maximum entropy spectral analysis, however, is very suitable for a small number of pixels. The computational complexity has been a drawback in using the two-dimensional maximum entropy spectral estimation procedures. Recently, Lim and Malik [4, 5, 6] have proposed an efficient iterative algorithm for the two-dimensional maximum entropy power spectrum estimation suitable for the minicomputer implementation. Their method is adapted and generalized for use in our PDP 11/45 minicomputer for the texture-image analysis. A three-dimensional graphics software is developed for the spectral display at the different viewing positions. Extensive computer results on the spectral analysis of texture images are also reported.

II. Two-Dimensional Power Spectrum Estimation

To obtain the power spectrum of a two-dimensional signal, the direct method is to calculate the two-dimensional Fourier transform of the autocorrelation function, i.e.:

\[
P_x(w_1, w_2) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} R_x(n_1, n_2) e^{-j(n_1\omega_1 + n_2\omega_2)}
\]  

(1)

The following notations will be used in the report:

- \(x(n_1, n_2)\): A 2-D random signal whose power spectrum we wish to estimate.
- \(R_x(n_1, n_2)\): Autocorrelation function of \(x(n_1, n_2)\)
- \(\hat{R}_x(n_1, n_2)\): An estimate of \(R_x(n_1, n_2)\)
- \(P_x(w_1, w_2)\): Power spectrum of \(x(n_1, n_2)\)
- \(\Lambda(n_1, n_2)\): Autocorrelation function whose power spectrum is \(1/P_x(w_1, w_2)\)
- \(A\): A set of points \((n_1, n_2)\) for which \(R_x(n_1, n_2)\) is known.
- \(F\): Discrete time Fourier Transform
- \(F^{-1}\): Inverse discrete time Fourier Transform
From (1), it is important to note that the determination of the power spectral density entails complete knowledge of the generally infinite extent autocorrelation function. For the finite signal, this method is proved to have poor resolution due to the truncated and sampled autocorrelation function set.

There are various techniques to estimate the power spectrum for the one-dimensional signal [7]. One technique that has been recognized as the best due to its high resolution is the Maximum Entropy Method (MEM). The basic idea of this approach is to extrapolate the autocorrelation function of a random process by maximizing the entropy $H$ of the corresponding probability density function:

$$ H = \int_{-\pi}^{\pi} \log P_x(w) \, dw $$

where $P_x(w)$ is the power spectrum density. The characteristics of the maximum entropy method are equivalent to the autoregressive signal modeling [8] which requires solving a set of linear equations for the filter coefficients. This can be expressed as:

$$ \hat{P}(w) = \frac{1}{1 + \sum_{k=1}^{M} a_k e^{-j\omega}} $$

and the filter coefficients $a_k$ are obtained by solving the normal equations:

$$ \hat{R}(i) = -\sum_{k=1}^{M} a_k \hat{R}(i-k) $$

There are different algorithms proposed to find the solutions of the normal equations. The most efficient ones are the Levinson recursive algorithm and the Burg recursive algorithm [7]. But for the two-dimensional case, the problem is different since the normal equations become a highly nonlinear problem [4]. For the general form of the two-dimensional case,
\[ \sum_{(i,j) \in B} a_{ij} R_x(r-i, s-j) = R_x(r, s) \quad \text{for } (r, s) \in B \]  

Here the set $B$ consists of all points where the filter mask has non-zero values, and the power spectrum obtained from $a_{ij}$ is given by

\[ \hat{P}_x(w_1, w_2) = \left| \sum_{(r,k) \in B} a_{rk} \exp(-j\omega_1 r - j\omega_2 k) \right|^2 \]  

In this case, we can see from (5) that the size of independent values of $R_x(n_1, n_2)$ required to solve the above set of equations is greater than the size of the filter mask. For example, Fig. 1(a) shows the autoregressive filter mask size as $3 \times 3$, and Fig. 1(b) shows a larger size of independent values of $R_x(n_1, n_2)$ required to solve for $a_{ij}$ in Fig. 1(a) by equation (5).

![Diagram](image)

Fig. 1

Clearly, the number of correlation points needed is greater than the number of filter coefficients. Since the estimated power spectrum given by (6) is completely determined by the coefficients alone, it does not possess enough degrees of freedom to solve for the spectrum. Therefore, the normal equations are not linear as in the one-dimensional case. Many methods have been proposed to
extend the Levinson's and the Burg's algorithms in two dimensions. However, those algorithms are not computationally attractive, and there is no guarantee that a solution or an approximate solution can be obtained. For instance, Burg [13] has proposed an iterative solution which requires the inversion of a matrix in each iteration where the dimension of the matrix is of the order of the number of the given autocorrelation points. No experimental results using this technique have yet been reported. Wernecke and D'Addario [14] have proposed a scheme in which an attempt is made to numerically maximize the entropy. The maximization is done by continuously adjusting the power spectrum estimate and evaluating the expressions for the entropy and its gradient. The procedure is computationally expensive and is not guaranteed to have a solution. Woods [11] expresses the Maximum Entropy Method as a power series in the frequency domain and attempts to approximate the ME PS estimate by truncating the power series expansion. Even though such an approach has some computational advantages relative to others, the method is restricted to the class of signals for which the power series expansion is possible.

Based on the reason that the closed form solution of the two-dimensional ME method is hard to obtain, Lim and Malik developed a new iterative algorithm, using adaptive filtering concept. This algorithm can correctly estimate the true spectrum and is computationally simple due to the utilization of Fast Fourier Transform (FFT). The basic idea of this algorithm is on the notion that the given correlation points in region A is consistent and the correc-
pending coefficient should be zero outside region $A$, and proceed this iteration repeatedly until the optimal solution is obtained. That is, given $R_x(n_1, n_2)$ for $(n_1, n_2) \in A$, determine $\hat{P}_x(w_1, w_2)$ such that $\hat{P}_x(w_1, w_2)$ has the form

$$\hat{P}_x(w_1, w_2) = \frac{1}{\sum_{(n_1, n_2) \in A} \lambda(n_1, n_2) e^{-j \omega n_1} e^{-j \omega n_2}}$$

and

$$R_x(n_1, n_2) = F^{-1} \left[ P_x(w_1, w_2) \right] \text{ for } (n_1, n_2) \in A$$

A simple flowchart is shown in Fig. 2. We begin with some initial estimate of $\lambda(n_1, n_2)$, obtain the corresponding correlation function, correct the resulting correlation function for $(n_1, n_2) \in A$ with the known $R_x(n_1, n_2)$, obtain the corresponding $\lambda(n_1, n_2)$ from the correct correlation function, and then replace the resulting $\lambda(n_1, n_2)$ with $0$ for $(n_1, n_2) \notin A$. This completes one iteration and the corrected $\lambda(n_1, n_2)$ is a new estimate of $\lambda(n_1, n_2)$.

**Fig. 2**
However, to prevent the zero crossing problem when taking the inverse of \( R_y(n_1, n_2) \) and \( R_y(n_1, n_2) \) due to the correction of \( R_y(n_1, n_2) \) and \( \Lambda(n_1, n_2) \), and also to keep the correlation function to be positive definite, they modify the procedure by linearly interpolating some parameters to prevent divergence of the error and to increase the rate of convergence [4]. Thus they have added some constraints to make the practical algorithm as shown in Fig. 3.

III. Examples

The following are some examples of using Lim and Malik's algorithm which is implemented in our PDP 11/45 minicomputer. The input signals are two-dimensional autocorrelation function originated from sinusoids buried in white noise, so that the correlation function given has the form of

\[
R_x(n_1, n_2) = \sigma^2 \delta(n_1, n_2) + \sum_{i=1}^{M} a_i^2 \cos(w_{i1} n_1 + w_{i2} n_2) \quad (8)
\]

where \( \sigma^2 \) is the white noise power, \( M \) is the number of sinusoids, \( a_i^2 \) is the power of the \( i \)th sinusoid, and \( w_{i1} \) and \( w_{i2} \) represent the frequency of the \( i \)th sinusoid. Fig. 4 - Fig. 6 are the cases for 1, 2, 3 sinusoids respectively, which correspond to the input data shown in Tables 1, 2, 3. From these results, we can see this iterative procedure can easily predict the true spectrum although the matrix of autocorrelation function is not very large.

However, in Lim and Malik's examples and the examples shown in Figs. 4-6, the expression used in the autocorrelation-function calculation is Eq. (8) which is an analytical form. This expression is based on the consideration of continuous and infinite sinusoids. So its values are completely symmetric. For the practical two-
GIVEN $R(n_1,n_2) = R(n_1,0) \delta(n_2,0)$, DPT LENGTH $N$, $k = 0.5$

$\alpha_0 = 0$, $\beta_0 = 0$, $\epsilon = 10^{-8}$

INITIAL ESTIMATE:
$R(n_1,n_2) = R(n_1,0) \delta(n_2,0)$

$\lambda(n_1,n_2) = \frac{1}{N} \sum_{n_1,n_2} R(n_1,n_2)$

Fig. 3 A detailed flowchart of the Lim-Malik iterative algorithm for 2-D ME PSE implemented in the report.

(IEEE ASSP Trans, June 1981)
Table 1

<table>
<thead>
<tr>
<th>A</th>
<th>M</th>
<th>$\sigma^2$</th>
<th>$a_i^2$</th>
<th>$(w_{11}/2\pi, w_{12}/2\pi)$</th>
<th>$\epsilon$</th>
<th>NDIT</th>
<th>NITR</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>5.0</td>
<td>1.0</td>
<td>0.375, 0.125</td>
<td>$10^{-4}$</td>
<td>32</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 2

<table>
<thead>
<tr>
<th>A</th>
<th>M</th>
<th>$\sigma^2$</th>
<th>$a_i^2$</th>
<th>$(w_{11}/2\pi, w_{12}/2\pi)$</th>
<th>$\epsilon$</th>
<th>NDIT</th>
<th>NITR</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.7</td>
<td>2</td>
<td>6.0</td>
<td>1.0</td>
<td>0.125, 0.125</td>
<td>$10^{-3}$</td>
<td>32</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>0.9</td>
<td>1.0</td>
<td>0.375, 0.125</td>
<td>$10^{-3}$</td>
<td>32</td>
<td>12</td>
</tr>
</tbody>
</table>

Table 3

<table>
<thead>
<tr>
<th>A</th>
<th>M</th>
<th>$\sigma^2$</th>
<th>$a_i^2$</th>
<th>$(w_{11}/2\pi, w_{12}/2\pi)$</th>
<th>$\epsilon$</th>
<th>NDIT</th>
<th>NITR</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>3</td>
<td>6.0</td>
<td>1.0</td>
<td>0.1, 0.1</td>
<td>$10^{-4}$</td>
<td>32</td>
<td>41</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>6.0</td>
<td>1.0</td>
<td>0.3, 0.1</td>
<td>$10^{-4}$</td>
<td>32</td>
<td>21</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>6.0</td>
<td>1.0</td>
<td>0.2, 0.2</td>
<td>$10^{-4}$</td>
<td>32</td>
<td>21</td>
</tr>
</tbody>
</table>
dimensional signal, this is no longer the case. To generalize this
algorithm so that it can be fitted for any two-dimensional signal,
we use the well known unbiased formula for the calculation of
text image data.

\[ R_{n_1, n_2} = \frac{1}{N \times N} \sum_{k=1}^{N-n_1} \sum_{l=1}^{N-n_2} x(k, l) x(k+n_1, l+n_2) \]  

To make sure that this expression can fit the real autocorrelation
function, we have used it to generate a simulated autocorrelation
function of the two-dimensional sinusoids, and the spectrum estimated
is almost the same as that calculated from Eq. (6).

In this section, we will present the results of the two-dimensional
maximum entropy method Power Spectrum Estimation of the real texture
images. These data are taken from the U. S. C. data base. The size
of each image is 64 x 64. The original textures are shown in Fig. 7.
These pictures will reappear but 5 times larger in Fig. 8a - Fig. 16a
with its corresponding estimated spectrum shown in Fig. 8b - Fig. 16b.
Fig. 8c - Fig. 16c and Fig. 8d - Fig. 16d.

Fig. 6b shows that the picture has one main frequency component
near (0.01, 0.22), one small dc component and one frequency at
(0.5, 0.5) and other very small ripples. Because of the three-
dimensional display, Fig. 8c and Fig. 8d cannot accurately indicate
the (0.01, 0.22) point, but we can see the entire power spectrum
function distribution in Fig. 8c and Fig. 8d. Fig. 9b shows that
the main frequency component is approximately equal to (0.02, 0.126)
and other small ripples. Comparing Figs. 8 and 9, we can see clearly
the difference in power spectrum for different textures.

Fig. 10 shows that the texture contains a main frequency
component around (0.0, 0.0) and other small amplitude components.
From the corresponding picture (upper right segment in Fig. 1), we can see that it contains mostly high grey level pixels, and we note that it has a high dc component. By taking a longer DFT length, we can predict the other small amplitude frequency terms. This effect can be shown in Fig. 13 which is a reconstructed picture with parts of Fig. 11 and Fig. 12. From Fig. 13a, we can see the spectrum estimated is the composition of Fig. 11d and Fig. 12a. The frequency term near (0.06, 0.14) is affected by the contribution of Fig. 12. When we take a larger autocorrelation function matrix and longer DFT length, this term can be recovered. Fig. 17 clearly demonstrates this phenomenon. To prove the accuracy of the two-dimensional 2D FSA algorithm, three sets of "nearly periodic" texture data were tested. From Figs. 14-16, the results obtained are satisfactory and appear to indicate truly the real power spectrum.

IV. Discussion

For a real two-dimensional signal, this algorithm can accurately predict the main frequency components with a moderate size autocorrelation function (ACF) matrix and short DFT length. However, for the other frequency terms with smaller amplitude, it must take a larger ACF matrix and longer DFT length to discriminate, since the larger ACF matrix will add more information about the signal. But this will take too much computational time. Also the spectral peaks that are very close cannot be resolved by this algorithm with the moderate size ACF matrix and DFT length. To understand and to improve the spectral resolution capability of the algorithm, we will analyze the true spectrum. We are now continuing the research by generating an
artificial two-dimensional signal which is predicted by the coefficients of the AR model suggested by Chassow and et al. [10]. It can be expressed as

\[ x(n_1, n_2) = - \sum_{k=0}^{N} \sum_{m=0}^{N} a_{km} x(n_1-k, n_2-m) \]

where the real power spectrum can be calculated by using

\[ P_x(w_1, w_2) = \frac{1}{\left| \sum_{k=0}^{N} \sum_{m=0}^{N} a_{km} e^{-j(kw_1 + mw_2)} \right|^2} \]

We expect that by comparing the estimated spectrum with the true one calculated from Eq. (11), we can find some special properties which can be used as the reference for improvement.

For the texture images studied in this report, the two-dimensional maximum entropy spectrum method has provided reasonably well spectral discrimination for different types of textures. The spectrum is indeed far superior to the conventional FFT algorithm for the small image size considered. The spectral features suggested for texture discrimination may be the powers computed for the various frequency bands. With its computational efficiency and accuracy, this algorithm is presently the only one available for minicomputer implementation. Further improvement in spectral resolution will make this algorithm even more powerful for the texture-image analysis.
References


Fig. 4(a) The contour map of the estimated power spectrum, \( \{k_{12}\} \). Note: \( \Delta \) is the dB value difference between each contour.

Fig. 4(b) Three-dimensional display of the estimated power spectrum. The viewer's eye is at the upper side of \( (-c_{11}=0, \ c_{12}=0) \) point.

Fig. 4(c) Three-dimensional display of the estimated power spectrum with the viewer's eye being at the upper side of \( (-c_{11}=0.5, \ c_{12}=0.5) \) point.
Fig. 5(a) The contour map of the estimated power spectrum with \( g_0 = 6 \).

Fig. 5(b) Three-dimensional display of the estimated spectrum at \((0.0, 0.0, 0.0)\) point.

Fig. 5(c) Three-dimensional display of the estimated power spectrum with point of view at \((0.5, 0.5)\).
Fig. 6(a) The contour map of the estimated power spectrum with 20=0.

Fig. 6(b) Three-dimensional display of the estimated spectrum at (0.05,0.05) point of view at (0.5,0.5).

Fig. 6(c) Three-dimensional display of the estimated power spectrum with point of view at (0.5, 0.5).
Fig. 7  The original test texture data (taken from USC data base). Each data format is 64x64. The right one in second row is reconstructed from the left and center pictures of the second row.
Fig. 8(a) The original test data.

Fig. 8(b) The contour map with $4dE=3$ with main frequency around $(0.01, 0.22)$.

Fig. 8(c) Three-dimensional display at $(0.0,0.0)$ point of view.

Fig. 8(d) Three-dimensional display at $(0.5,0.5)$ point of view.
Fig. 9(a) The original test data. Fig. 9(b) The contour map with $\Delta dB=3$ with main frequency around $0.02, 0.126$.

Fig. 9(c) Three-dimensional display at $(0.0,0.0)$ point of view. Fig. 9(d) Three-dimensional display at $(0.5,0.5)$ point of view.
Fig. 10(a) The original test data.

Fig. 10(b) The contour map with $\Delta dB=3$ with main frequency around $(0,0,0)$.

Fig. 10(c) Three-dimensional display at $(0,0,0)$ point of view.

Fig. 10(d) Three-dimensional display at $(0.5,0.5)$ point of view.
Fig. 11(a) The original test data.

Fig. 11(b) The contour map with \(\Delta B=1.0\) with main frequencies around \((0.0,0.0)\) and \((0.06,0.125)\).

Fig. 11(c) Three-dimensional display at \((0.0,0.0)\) point of view.

Fig. 11(d) Three-dimensional display at \((0.5,0.5)\) point of view.
Fig. 12(a) The original test data.

Fig. 12(b) The contour map with $\Delta f=2$, with main frequency around $(0.03, 0.07)$.

Fig. 12(c) Three-dimensional display at $(0.0, 0.0)$ point of view.

Fig. 12(d) Three-dimensional display at $(0.5, 0.5)$ point of view.
Fig. 13(a) The test data, as constructed from Fig. 11a and Fig. 12a).

Fig. 13(b) The contour map with $\Delta dB=1$, with main frequency around (0.0, 0.0).

Fig. 13(c) Three-dimensional display at (0.0, 0.0) point of view.

Fig. 13(d) Three-dimensional display at (0.5, 0.5) point of view.
Fig. 14(a) The original test data.

Fig. 14(b) The contour map with $\Delta B = 3$, with main frequency around $(0.08, 0.09)$.

Fig. 14(c) Three-dimensional display at $(0.0, 0.0, 0.0)$ point of view.

Fig. 14(d) Three-dimensional display at $(0.5, 0.5)$ point of view.
Fig. 15(a) The original test data.

Fig. 15(b) The contour map with $|dB|=3$, with main frequency around $(0.065,0.20)$.

Fig. 15(c) Three-dimensional display at $(0.0,0.0)$ point of view.

Fig. 15(d) Three-dimensional display at $(0.5,0.5)$ point of view.
Fig. 16(a) The original test data.

Fig. 16(b) The contour map with $\phi_B = \lambda$, with main frequency around $(0.3125, 0.3125)$.

Fig. 16(c) Three-dimensional display at $(0.0,0.0)$ point of view.

Fig. 16(d) Three-dimensional display at $(0.5,0.5)$ point of view.
Fig. 17(a) The contour map of the 3-dB-pole-crossing area at (0.0,0.0) and (0.5,0.5).

Fig. 17(b) Three-dimensional display at (0.0,0.0) point of view.

Fig. 17(c) Three-dimensional display at (0.5,0.5) point of view.
The two-dimensional maximum entropy spectral estimation method proposed by Lim and Malik is studied and modified for spectral analysis of texture images. The computational efficiency of the algorithm makes it feasible for minicomputer implementation to extract effective spectral domain texture features for texture discrimination. Extensive computer results are presented.