Sequential Estimation of Two-dimensional Sinusoidal Models

Anurag Prasad[†] & Debasis Kundu[†] & Amit Mitra[†]

Abstract

Estimating the unknown parameters of a two-dimensional (2-D) sinusoidal signal is an important and a difficult problem. In this paper, we propose a simple sequential procedure for estimating the unknown frequencies and amplitudes of the 2-D sinusoidal components when the signal is affected by noise. When there are p components in the signal, the k-th step of the procedure provides strongly consistent estimators of the k dominant sinusoids when $k \leq p$. When $k > p$, the supernumerary amplitude estimators converge to zero almost surely. The asymptotic distribution of the proposed estimators coincides with the asymptotic distribution of the least-squares estimators. Numerical simulations are performed for various sample sizes and for various model parameters. Some real texture data and some synthesized texture data are analyzed.

Keywords: Frequencies; Amplitudes; Least-squares estimators; Strongly consistent estimators; Asymptotic distributions; Bayesian Information Criterion.

[†] Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Kanpur, Pin 208016, INDIA.

Corresponding Author: Debasis Kundu, e-mail:kundu@iitk.ac.in

1 Introduction

In this paper we consider the problem of estimating the parameters of the following 2-D sinusoidal signal;

$$y(m,n) = \sum_{k=1}^{p} \left\{ A_k^0 \cos(m\lambda_k^0 + n\mu_k^0) + B_k^0 \sin(m\lambda_k^0 + n\mu_k^0) \right\} + X(m,n). \tag{1}$$

Here A_k^0 and B_k^0 are unknown real numbers, usually known as amplitudes, $\lambda_k^0, \mu_k^0 \in (0,\pi)$ are unknown frequencies. The additive error $\{X(m,n)\}$ is from a stationary random field and satisfies Assumption 1, which will be described later. The number of components 'p' may be known or unknown. The problem is to estimate the unknown parameters given a sample $\{y(m,n); m = 1,\ldots,M, \quad n = 1,\ldots,N\}$.

The first term on the right hand side of (1) is the signal component and the second term as the noise component. Detection of the signal component in presence of noise is an important problem in Statistical Signal Processing. Moreover, the 2-D sinusoidal model (1) has received considerable attention in the signal processing literature because of its widespread applicability in the texture analysis. It is observed by Francos *et al.* [3] that (1) can be used very effectively to model 2-D texture images and they estimated the unknown frequencies by selecting the sharpest peaks of the periodogram function $I(\lambda,\mu)$, of the observed signal $y(m,n)$. The two dimensional periodogram function $I(\lambda,\mu)$ is defined as follows;

$$I(\lambda,\mu) = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \left| y(m,n) e^{-i(m\lambda + n\mu)} \right|^2. \tag{2}$$

For some of the theoretical developments of the 2-D sinusoidal model, the readers are referred to Rao *et al.* [12], Kundu and Mitra [7], Bansal *et al.* [1], Zhang and Mandrekar [17], Mitra and Stoica [9] and Kundu and Nandi [8].

The corresponding one-dimensional problem has received considerable attention in the statistical signal processing and time series literature. Fisher [2] and Hartley [5] first studied

testing for hidden periodicities obscured by Gaussian white noise. Later on several articles have appeared dealing with the theoretical developments of the different estimators and developing several efficient algorithms for one-dimensional model. See for example the list of references by Stoica [14], the review article by Kundu [6] or the recent article by Prasad *et al.* [10] and the references cited there.

The 2-D frequency estimation problem is well known to be numerically difficult. Problems can arise when $p \geq 2$. If $p = 2$ and if the distance between (λ_1, μ_1) and (λ_2, μ_2) is small, then most methods will have difficulty in distinguish between the two pairs of frequencies. The most efficient estimators, are the least-squares estimators. If the number of components p is known, the rate of convergence of the least-squares estimators of the frequencies are $O_p(M^{-\frac{3}{2}}N^{-\frac{1}{2}})$ and $O_p(N^{-\frac{3}{2}}M^{-\frac{1}{2}})$ respectively, see for example Rao *et al.* [12], Kundu and Mitra [7] and Kundu and Nandi [8]. But finding the least-squares estimators may not be easy, since there may be several local minima on the surface of the criterion function. Making the initial guesses is a difficult task if the number of components is high. One way of starting is to take the maxima of the periodogram function (2) as the initial guesses. Asymptotically, the periodogram has local maxima at the true frequencies. But, if two pairs of frequencies are very close, then this method may not work.

Let us consider the following synthesized texture data;

$$\begin{aligned} y(m,n) &= 4.0\cos(1.8m + 1.1n) + 4.0\sin(1.8m + 1.1n) + \\ &\quad 1.0\cos(1.7m + 1.0n) + 1.0\sin(1.7m + 1.0n) + X(m,n), \end{aligned} \tag{3}$$

where

$$X(m,n) = e(m,n) + 0.25e(m-1,n) + 0.25e(m,n-1).$$

Here $e(m,n)$s are independent and identically distributed ($i.i.d.$) normal random variables with mean zero and variance 2.0. The periodogram function (2) of the data generated from

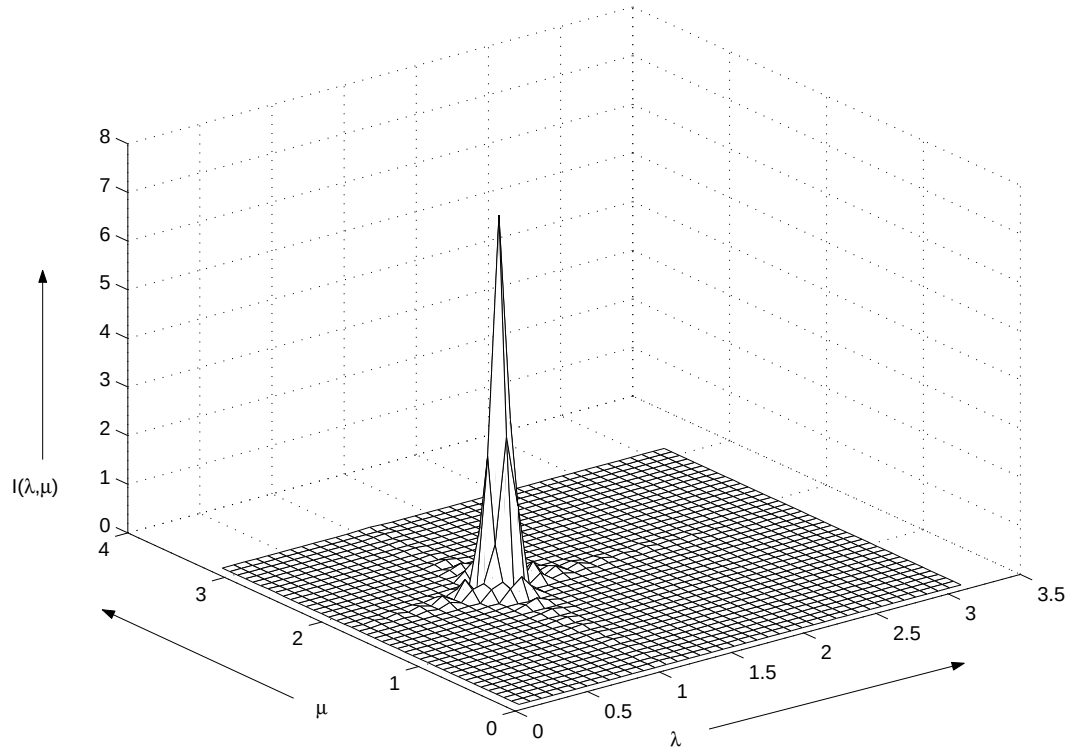

Figure 1: Periodogram plot of the synthesized signal.

(3) is provided in Figure 1. In this case, the two pairs of frequencies are not resolvable. Therefore, it is difficult to know where to start an iterative process for finding the least-squares estimates.

Another practical problem occurs while finding the least-squares estimators when the number of components 'p' is very large. For some texture data, p can be as large as 20. In such situations, the initial guess can be crucial. We have two major aims in this paper. The first aim is to estimate the unknown parameters, when the number of components, p, is known. If p is known, then we can propose a sequential procedure to estimate the unknown parameters. The $2p$-dimensional optimization problem can be reduced to p, two-dimensional optimization problems. The estimators obtained by the proposed method have the same rate of convergence as the least-squares estimators.

The second aim of this paper is to study the properties of the estimators if p is not known. If p is not known and a lower order model is fitted, the proposed estimators are consistent estimators of the dominant components with the same rate of convergence as the least-squares estimators. If we fit a model of higher order than that of the underlying process, then the estimators up to the p-th step are consistent. To illustrate the proposed method, we analyze some real texture data and some synthesized data. We also perform some computer simulations.

The rest of the paper is organized as follows. In section 2, we provide the necessary assumptions on the model parameters and on the errors. Consistency results of the proposed estimators are provided in section 3. Asymptotic distributions and the convergence rates of the proposed estimators are provided in section 4. Simulation results are provided in section 5. In section 6, data analysis results have been provided and finally the conclusions appear in section 7.

2 Model Assumptions and Methodology

2.1 Assumptions

In this section, first we provide the necessary assumptions on the model parameters and particularly on the errors. It is assumed that the observed data $\{y(m,n); m = 1, \ldots, M, n = 1, \ldots, N\}$ is of the form (1). The additive error $\{X(m,n)\}$ is from a stationary random field and it satisfies the following Assumption 1;

Assumption 1: Let us denote the set of positive integers by $\mathcal{Z}$. It is assumed that $\{X(m,n); m,n \in \mathcal{Z}\}$ can be represented as follows;

$$X(m,n) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} a(j,k) e(m-j, n-k), \tag{4}$$

where $a(j,k)$s are real constants such that

$$\sum_{j=-\infty}^{\infty}\sum_{k=-\infty}^{\infty}|a(j,k)|<\infty, \tag{5}$$

and $\{e(m,n); m,n \in \mathcal{Z}\}$ is a double array sequence of *i.i.d.* random variables with mean zero and finite variance σ^2.

Assumption 2: The frequency sets $\{\lambda_i^0, \mu_i^0\}$ are distinct *i.e.* for $i \neq j$, $(\lambda_i^0, \mu_i^0) \neq (\lambda_j^0, \mu_j^0)$ and $(\lambda_i^0, \mu_i^0) \in (0,\pi)\times(0,\pi)$ for all $i = 1,\ldots,p$.

Assumption 3: The amplitudes satisfy the following restriction;

$$0 < A_p^{0^2} + B_p^{0^2} < \ldots < A_1^{0^2} + B_1^{0^2} < 2S^2 < \infty. \tag{6}$$

2.2 Methodology

The method that is proposed for estimating the unknown parameters is a stepwise regression procedure that begins by minimising

$$Q_1(A,B,\lambda,\mu) = \sum_{m=1}^{M}\sum_{n=1}^{N}\left[y(m,n) - A\cos(m\lambda+n\mu) - B\sin(m\lambda+n\mu)\right]^2, \tag{7}$$

with respect to (*w.r.t.*) A, B, λ and μ. This may be achieved by concentrating the function with respect to A and B. That is to say, these parameters are replaced by the estimating equations $\widehat{A}(\lambda,\mu)$ and $\widehat{B}(\lambda,\mu)$ that are specified in Appendix A. Then, the resulting function

$$\begin{aligned} R_1(\lambda,\mu) &= Q_1(\widehat{A}(\lambda,\mu),\widehat{B}(\lambda,\mu),\lambda,\mu) \\ &= \sum_{m=1}^{M}\sum_{n=1}^{N}\left[y(m,n) - \widehat{A}(\lambda,\mu)\cos(m\lambda+n\mu) - \widehat{B}(\lambda,\mu)\sin(m\lambda+n\mu)\right]^2 \end{aligned} \tag{8}$$

is minimised with respect to λ and μ.

The estimates from the first stage are $\widehat{\lambda}_1$, $\widehat{\mu}_1$, $\widehat{A}_1 = \widehat{A}(\widehat{\lambda}_1,\widehat{\mu}_1)$ and $\widehat{B}_1 = \widehat{B}(\widehat{\lambda}_1,\widehat{\mu}_1)$, and the corresponding residual sequence is

$$y^{(1)}(m,n) = y(m,n) - \widehat{A}_1\cos(m\widehat{\lambda}_1 + n\widehat{\mu}_1) - \widehat{B}_1\sin(m\widehat{\lambda}_1 + n\widehat{\mu}_1). \tag{9}$$

The second-stage estimates, $\widehat{\lambda}_2$, $\widehat{\mu}_2$, $\widehat{A}_2$, and $\widehat{B}_2$ may be obtained by minimising the function $R_2(\lambda,\mu) = Q_2\{\widehat{A}(\lambda,\mu), \widehat{B}(\lambda,\mu), \lambda, \mu\}$, which comes from replacing the data sequence $y(m,n)$ in Q_1 by the residual sequence $y^{(1)}(m,n)$. Further stages of the procedure follow in like manner.

Iterative regression procedures that employ a concentrated criterion function in the manner of the present procedure have been investigated by Richards [13].

3 Consistency of Proposed Estimators

In this section we provide the consistency results of the proposed estimators when the number of components is known. We consider two cases (a) when the number of components of the fitted model, q, is less than the actual number of components, p and (b) when $q > p$. We need the following lemma;

Lemma 1: Let $\{X(m,n); m \in \mathcal{Z}, n \in \mathcal{Z}\}$ be a sequence of stationary random variables satisfying Assumption 1, then as $M \to \infty$, $N \to \infty$,

$$\sup_{\alpha,\beta} \left| \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} X(m,n) e^{i(m\alpha+n\beta)} \right| \to 0 \quad a.s. \tag{10}$$

Proof: This can be obtained from Kundu and Nandi [8].

Lemma 2: Consider the set $S_c = \{\theta; \theta \in \Theta, \text{ and } |\theta - \theta_1^0| \geq c\}$, where $\theta = (A, B, \lambda, \mu)$ and $\theta_1^0 = (A_1^0, B_1^0, \lambda_1^0, \mu_1^0)$, $\Theta = [-S, S] \times [-S, S] \times [0, \pi] \times [0, \pi]$. If for any $c > 0$,

$$\liminf_{M,N\to\infty} \inf_{\theta \in S_c} \frac{1}{MN} \left\{ Q_1(\theta) - Q_1(\theta_1^0) \right\} > 0 \quad a.s., \tag{11}$$

then $\widehat{\theta}_1$ which minimizes $Q_1(\theta)$, is a strongly consistent estimator of θ_1^0.

Proof: The result was proved by Wu [15] when θ is of dimension one. A similar proof is available for cases of multiple parameters.

THEOREM 1: If the Assumptions 1-3 are satisfied, then $\widehat{\theta}_1$ is a strongly consistent estimator of θ_1^0.

PROOF: Consider the following expression

$$\frac{1}{MN}\left[Q_1(\theta) - Q_1(\theta_1^0)\right] = f(\theta) + g(\theta), \tag{12}$$

where

$$\begin{aligned} f(\theta) &= \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}(A_1^0\cos(m\lambda_1^0 + n\mu_1^0) + B_1^0\sin(m\lambda_1^0 + n\mu_1^0) \\ &- A\cos(m\lambda + n\mu) - B\sin(m\lambda + n\mu))^2 \\ &+ \frac{2}{MN}\left\{A_1^0\cos(m\lambda_1^0 + n\mu_1^0) + B_1^0\sin(m\lambda_1^0 + n\mu_1^0) - A\cos(m\lambda + n\mu) - B\sin(m\lambda + n\mu)\right\} \\ &\times \left[\sum_{k=2}^{p}\left\{A_k^0\cos(m\lambda_k^0 + n\mu_k^0) + B_k^0\sin(m\lambda_k^0 + n\mu_k^0)\right\}\right], \end{aligned}$$

and

$$\begin{aligned} g(\theta) &= \frac{2}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}X(m,n)[A_1^0\cos(m\lambda_1^0 + n\mu_1^0) + B_1^0\sin(m\lambda_1^0 + n\mu_1^0) \\ &- A\cos(m\lambda + n\mu) - B\sin(m\lambda + n\mu)]. \end{aligned}$$

Now using Lemma 1, it easily follows that

$$\sup_{\theta\in S_c}|g(\theta)| \to 0 \quad a.s..$$

Using lengthy but straight forward calculations and splitting the set S_c similar to Kundu and Nandi [8], it can be shown that

$$\liminf_{M,N\to\infty}\inf_{\theta\in S_C} f(\theta) > 0.$$

This proves the result. ∎.

Now we will show that, at the second step, the proposed method produces consistent estimates. We need the following results;

LEMMA 3: If the Assumption 1-3 are satisfied, then

$$M(\widehat{\lambda}_1 - \lambda_1^0) \to 0 \quad a.s. \quad \text{and} \quad N(\widehat{\mu}_1 - \mu_1^0) \to 0 \quad a.s.$$

PROOF: Consider the 4×4 diagonal matrix $D_1 = \text{diag}\left\{1, \quad 1, \quad M^{-1}, \quad N^{-1}\right\}$. Using the multivariate Taylor series expansion we can write

$$Q_1'(\widehat{\theta}_1) - Q_1'(\theta_1^0) = \left(\widehat{\theta}_1 - \theta_1^0\right) Q_1''(\bar{\theta}). \tag{13}$$

Here

$$Q_1'(\theta) = \left[\frac{\partial Q_1(\theta)}{\partial A}, \frac{\partial Q_1(\theta)}{\partial B}, \frac{\partial Q_1(\theta)}{\partial \lambda}, \frac{\partial Q_1(\theta)}{\partial \mu}\right]$$

and

$$Q_1''(\theta) = \begin{bmatrix} \frac{\partial^2 Q_1(\theta)}{\partial A^2} & \frac{\partial^2 Q_1(\theta)}{\partial A \partial B} & \frac{\partial^2 Q_1(\theta)}{\partial A \partial \lambda} & \frac{\partial^2 Q_1(\theta)}{\partial A \partial \mu} \\ \frac{\partial^2 Q_1(\theta)}{\partial B \partial A} & \frac{\partial^2 Q_1(\theta)}{\partial B^2} & \frac{\partial^2 Q_1(\theta)}{\partial B \partial \lambda} & \frac{\partial^2 Q_1(\theta)}{\partial B \partial \mu} \\ \frac{\partial^2 Q_1(\theta)}{\partial \lambda \partial A} & \frac{\partial^2 Q_1(\theta)}{\partial \lambda \partial B} & \frac{\partial^2 Q_1(\theta)}{\partial \lambda^2} & \frac{\partial^2 Q_1(\theta)}{\partial \lambda \partial \mu} \\ \frac{\partial^2 Q_1(\theta)}{\partial \mu \partial A} & \frac{\partial^2 Q_1(\theta)}{\partial \mu \partial B} & \frac{\partial^2 Q_1(\theta)}{\partial \mu \partial \lambda} & \frac{\partial^2 Q_1(\theta)}{\partial \mu^2} \end{bmatrix}.$$

Moreover, $\bar{\theta}$ is a point between $\widehat{\theta}_1$ and θ_1^0. Since $Q_1(\widehat{\theta}_1) = 0$,

$$\left(\widehat{\theta}_1 - \theta_1^0\right) D_1^{-1} \left[\frac{1}{MN} D_1 Q_1''(\bar{\theta}) D_1\right] = -\left[\frac{1}{MN} Q_1'(\theta_1^0) D_1\right]. \tag{14}$$

Using Theorem 1 and trigonometric identity, it follows that;

$$\lim_{M,N\to\infty} \left[\frac{1}{MN} D_1 Q_1''(\bar{\theta}) D_1\right] = \lim_{M,N\to\infty} \left[\frac{1}{MN} D_1 Q_1''(\theta_1^0) D_1\right] = 2\Sigma_1 > 0 \quad a.s.,$$

where

$$\Sigma_1 = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{4}B_1^0 & \frac{1}{4}B_1^0 \\ 0 & \frac{1}{2} & -\frac{1}{4}A_1^0 & -\frac{1}{4}A_1^0 \\ \frac{1}{4}B_1^0 & -\frac{1}{4}A_1^0 & \frac{1}{6}\left(A_1^{0^2} + B_1^{0^2}\right) & \frac{1}{8}\left(A_1^{0^2} + B_1^{0^2}\right) \\ \frac{1}{4}B_1^0 & -\frac{1}{4}A_1^0 & \frac{1}{8}\left(A_1^{0^2} + B_1^{0^2}\right) & \frac{1}{6}\left(A_1^{0^2} + B_1^{0^2}\right) \end{bmatrix}.$$

Using Lemma 1, it follows that

$$\frac{1}{MN} Q_1'(\theta_1^0) D_1 \to 0 \quad a.s..$$

Therefore, the result follows. ∎

THEOREM 2: If the Assumption 1-3 are satisfied, and $p \geq 2$, then $\widehat{\theta}_2$ obtained by minimizing $Q_2(A, B, \lambda, \mu)$ as defined in Section 2, is a strongly consistent estimator of θ_2^0.

PROOF: Using Theorem 1 and Lemma 3 we obtain;

$$\widehat{A}_1 = A_1^0 + o(1) \;\; a.s., \quad \widehat{B}_1 = B_1^0 + o(1) \;\; a.s., \quad \widehat{\lambda}_1 = \lambda_1^0 + o(N) \;\; a.s., \quad \widehat{\mu}_1 = \mu_1^0 + o(N) \;\; a.s.$$

Here the random variable $U = o(1)$ means, $U \to 0 \;\; a.s.$, and $U = o(N)$ means $UN \to 0 \;\; a.s.$. Therefore, the result follows using the same method as in Theorem 1, by using

$$\widehat{A}_1 \cos(\widehat{\lambda}_1 m + \widehat{\mu}_1 n) + \widehat{B}_1 \sin(\widehat{\lambda}_1 m + \widehat{\mu}_1 n) = A_1^0 \cos(\lambda_1 m + \mu_1 n) + B_1^0 \sin(\lambda_1^0 m + \mu_1^0 n) + o(1) \;\; a.s.$$

4 Asymptotic Distribution of the Estimators

In this section, we obtain the asymptotic distributions of the proposed estimators at each step. Consider the following 4×4 diagonal matrix D as follows;

$$D = \begin{bmatrix} \frac{1}{\sqrt{MN}} & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{MN}} & 0 & 0 \\ 0 & 0 & \frac{1}{M^{3/2}N^{1/2}} & 0 \\ 0 & 0 & 0 & \frac{1}{M^{1/2}N^{3/2}} \end{bmatrix}.$$

Since $Q_1'(\widehat{\theta}_1) = 0$, from (13), we obtain

$$(\widehat{\theta}_1 - \theta_1^0) D^{-1} \left[D Q_1''(\bar{\theta}) D \right] = - \left[Q_1'(\theta_1^0) D \right].$$

Now observe that as $\widehat{\theta}_1 \to \theta_1^0 \; a.s.$, and using Lemma 3,

$$\lim_{M,N \to \infty} D Q_1''(\bar{\theta}) D = \lim_{M,N \to \infty} D Q_1''(\theta_1^0) D \quad a.s.$$

We also have the following results;

$$Q_1'(\theta_1^0)D \xrightarrow{d} N_4(0, 4\sigma^2 c_1\ \Sigma_1) \tag{15}$$

$$\lim_{M,N\to\infty} DQ_1''(\theta_1^0)D = 2\Sigma_1, \tag{16}$$

where

$$c_1 = \left| \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} a(j,k) e^{-(j\lambda_1^0 + k\mu_1^0)} \right|^2 .$$

For the proof of (15) and (16), see the Appendix B. Therefore, we can state the following result;

Theorem 3: If the Assumption 1-3 are satisfied, then

$$\left[M^{\frac{1}{2}}N^{\frac{1}{2}}(\widehat{A}_1 - A_1^0), M^{\frac{1}{2}}N^{\frac{1}{2}}(\widehat{B}_1 - B_1^0), M^{\frac{3}{2}}N^{\frac{1}{2}}(\widehat{\lambda}_1 - \lambda_1^0), M^{\frac{1}{2}}N^{\frac{3}{2}}(\widehat{\mu}_1 - \mu_1^0)\right] \xrightarrow{d} N_4(0, \sigma^2 c_1 \Sigma_1^{-1}).$$

By proceeding in the same manner, and using Theorem 2, it can be shown that the result holds for any $k \leq p$. Thus;

Theorem 4: If the Assumption 1-3 are satisfied, then

$$\left[M^{\frac{1}{2}}N^{\frac{1}{2}}(\widehat{A}_k - A_k^0), M^{\frac{1}{2}}N^{\frac{1}{2}}(\widehat{B}_k - B_k^0), M^{\frac{3}{2}}N^{\frac{1}{2}}(\widehat{\lambda}_k - \lambda_k^0), M^{\frac{1}{2}}N^{\frac{3}{2}}(\widehat{\mu}_k - \mu_k^0)\right] \xrightarrow{d} N_4(0, \sigma^2 c_k \Sigma_k^{-1}),$$

where c_k, Σ_k can be obtained from c_1 and Σ_1 by replacing A_1^0, B_1^0, λ_1^0 and μ_1^0 with A_k^0, B_k^0, λ_k^0 and μ_k^0.

5 Numerical Results

In this section, we present some numerical results to demonstrate the performance of the estimation procedure in relation to the asymptotic results. The model of equation (3) has been adopted and various values have been specified for the sample dimensions M and N. The random number generator RAN2 of Press *et al.* [11] has been used in generating the

additive error. The programs, which are written in FORTRAN-77, can be obtained from the corresponding author on request.

At each step, the minimisation *w.r.t.* λ and μ has been achieved using the downhill simplex method described in Press *et al.* [11]. The frequency values of the model (3) are close. Therefore, it is difficult to find appropriate starting values. In fact, the periodogram of Figure 1 shows that, at the sample size in question, they cannot be separated.

Table 1 shows the results from pursuing the first step ($k = 1$) for the estimation procedure when the data are generated by the model of equation (3). The procedure has been applied to samples of sizes $M = N = 20$, 30, 40, 50. In each case, there have been 1000 trials. The table reports the average values of the estimates (AEs) their mean square errors (MSEs) and their asymptotic variances (ASYVs). Table 2 shows the analogous results from pursuing the second step ($k = 2$) of the procedure, and Table 3 relates to the supernumerary estimates that arise is when $k = 3$, which is in excess of the model order.

Some points are evident in the table. As the sample size increases the biases and the MSEs decrease. This verifies the consistency of the estimates. The biases of the linear parameters A and B are greater than those of the non-linear parameters λ and μ as might be expected. The MSEs match quite well with the asymptotic variances for large sample sizes. From the table values, it is clear that the proposed sequential procedure is working quite well, when the number of component is exactly estimated, over-estimated or under-estimated.

6 Data Analysis

In this section we present two illustrative data analyses. One concerns some real texture data and the other concerns some synthesized texture data.

Table 1: Model 1 is considered with $k = 1^*$.

		$A_1 = 4.0$	$B_1 = 4.0$	$\lambda_1 = 1.8$	$\mu_1 = 1.1$
	AE	2.4084	4.6440	1.8071	1.1072
M= 20	MSE	(0.257E+01)	(0.428E+00)	(0.570E-04)	(0.581E-04)
N= 20	ASYV	(0.267E-01)	(0.267E-01)	(0.625E-05)	(0.625E-05)
	AE	1.9834	4.7000	1.8124	1.1123
M= 30	MSE	(0.408E+01)	(0.495E+00)	(0.154E-03)	(0.153E-03)
N= 30	ASYV	(0.119E-01)	(0.119E-01)	(0.124E-05)	(0.124E-05)
	AE	3.4178	4.3148	1.8037	1.1037
M= 40	MSE	(0.347E+00)	(0.105E+00)	(0.143E-04)	(0.145E-04)
N= 40	ASYV	(0.667E-02)	(0.667E-02)	(0.391E-06)	(0.391E-06)
	AE	4.3241	3.6970	1.7988	1.0988
M= 50	MSE	(0.109E+00)	(0.972E-01)	(0.168E-05)	(0.170E-05)
N= 50	ASYV	(0.427E-02)	(0.427E-02)	(0.160E-06)	(0.160E-06)

* The average estimates and the MSEs are reported for each parameter. The first row represents the true parameter values. In each box corresponding to each sample size, the first row represents the average estimates, the corresponding MSEs and the asymptotic variances (ASYV) are reported below within brackets.

Table 2: Model 1 is considered with $k = 2^*$.

		$A_2 = 1.0$	$B_2 = 1.0$	$\lambda_2 = 1.7$	$\mu_2 = 1.0$
	AE	0.6518	-0.5174	1.6409	0.9416
M= 20	MSE	(0.128E+00)	(0.232E+01)	(0.382E-02)	(0.373E-02)
N= 20	ASYV	(0.285E-01)	(0.285E-01)	(0.107E-03)	(0.107E-03)
	AE	0.6855	-0.5170	1.6545	0.9544
M= 30	MSE	(0.102E+00)	(0.231E+01)	(0.210E-02)	(0.211E-02)
N= 30	ASYV	(0.127E-01)	(0.127E-01)	(0.211E-04)	(0.211E-04)
	AE	1.1287	0.6318	1.6912	0.9912
M= 40	MSE	(0.211E-01)	(0.145E+00)	(0.841E-04)	(0.844E-04)
N= 40	ASYV	(0.713E-02)	(0.713E-02)	(0.669E-05)	(0.669E-05)
	AE	1.0901	0.8734	1.6983	0.9983
M= 50	MSE	(0.112E-01)	(0.200E-01)	(0.531E-05)	(0.566E-05)
N= 50	ASYV	(0.456E-02)	(0.456E-02)	(0.274E-05)	(0.274E-05)

* The average estimates and the MSEs are reported for each parameter. The first row represents the true parameter values. In each box corresponding to each sample size, the first row represents the average estimates, the corresponding MSEs and the asymptotic variances (ASYV) are reported below within brackets.

Table 3: Model 1 is considered with $k = 3^*$.

		$A_3 = 0.0$	$B_3 = 0.0$
M= 20	AE	0.1229E-01	0.115E-01
N= 20	VAR	(0.216E-01)	(0.350E-01)
M= 30	AE	0.2291E-02	-0.312E-02
N= 30	VAR	(0.122E-01)	(0.112E-01)
M= 40	AE	0.1151E-02	0.110E-02
M= 40	VAR	(0.191E-02)	(0.344E-02)
N= 50	AE	0.110E-02)	0.100E-02
N= 50	VAR	(0.117E-02)	(0.103E-02)

* The average estimates and the variances are reported for each parameter. The first row represents the parameter values. In each box corresponding to each sample size, the first row represents the average estimates, and the corresponding variances (VAR) are reported below within brackets.

TEXTURE DATA: The real texture data of Figure 2 is modelled by equation (1); but, in this case, the number of the components, p, is unknown. We have plotted the 2-D periodogram in Figure 3, to give an indication of number of components, but the matter remains unclear. Therefore, we have estimated the model sequentially for $k = 1, \ldots, 50$ and we have used the Bayesian Information Criterion (BIC) to estimate p.

In this case the BIC takes the following form;

$$BIC(k) = (MN) \ln \widehat{\sigma_k^2} + \frac{1}{2}(4k+1) \ln(MN)$$

where $\widehat{\sigma^2}_k$ is the innovative variance, when the number of components is k. In this case the number of parameters to be estimated is $4k+1$. We plot the $BIC(k)$ as a function of k in Figure 4. It is observed that for $k = 20$, $BIC(k)$ gives the minimum value, therefore in this case the estimate of p, say $\widehat{p} = 20$. We have fitted the model (1) with $p = 20$ to the texture data. We estimate the parameters sequentially as proposed in Section 2. The estimated texture are plotted in Figure 5. It matches reasonably well. We have plotted the residuals in Figure 6. It does not show any pattern. It looks like random patterns.

To test the randomness of the residual noise pattern, we have used Hopkins' test (see for example Zeng and Dubes [16]). The Hopkins test statistic is

$$T = \frac{\sum_{k=1}^{M} U^d(k)}{\sum_{k=1}^{M} U^d(k) + \sum_{k=1}^{M} W^d(k)},$$

where $U(k); k = 1, \ldots, M$ denote the distances from the M sampling origins to the nearest patterns, $W(k); k = 1, \ldots, M$ denote the distances from M patterns selected at random inside the sampling frame to their nearest patterns and d denotes the dimension, see Zeng and Dubes [16] for details.

Under the null hypothesis of randomness, T has a beta distribution with parameters (M, M). The Hopkins statistic is relatively small under regularity (when the patterns are pictured as falling into a mosaic in which any two patterns cannot be too close) and relatively

large under aggregation (when the patterns are generated in separate balls). Testing for random pattern, Hopkins has suggested to perform the following two one sided tests:

Test 1: H_0: Pattern is random vs. H_1 : The patterns are generated under regularity.

Test 2: H_0: Pattern is random vs. H_2 : The patterns are aggregated.

Hopkins suggested to perform both the tests (one sided) and if H_0 cannot be rejected in both cases, then the patterns can be described as random. A one sided test of H_0 vs. H_1 has form 'reject H_0 if $T < t_a$' and one of H_0 vs. H_2 has form 'reject H_0 if $T > t_b$'.

For the noise pattern under consideration, the value of test statistic is $T = 0.4999$. Here we have taken $M = 20$ sampling origins. The cutoff values are $t_a = 0.37$ and $t_b = 0.63$ for beta(20, 20) distribution. Hence both the tests fail to reject the null hypothesis of randomness of the noise pattern in Figure 6. Therefore, in this case the error structure satisfies the model assumptions.

Note that it was possible to fit such a large order model, because it has been done sequentially, otherwise it would have been a difficult task to estimate all the parameters simultaneously.

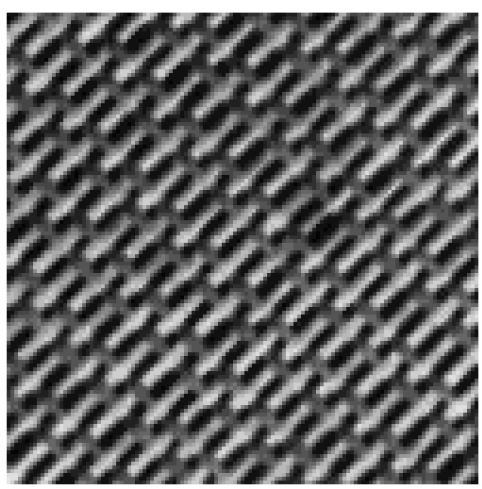

Figure 2: Original Texture.

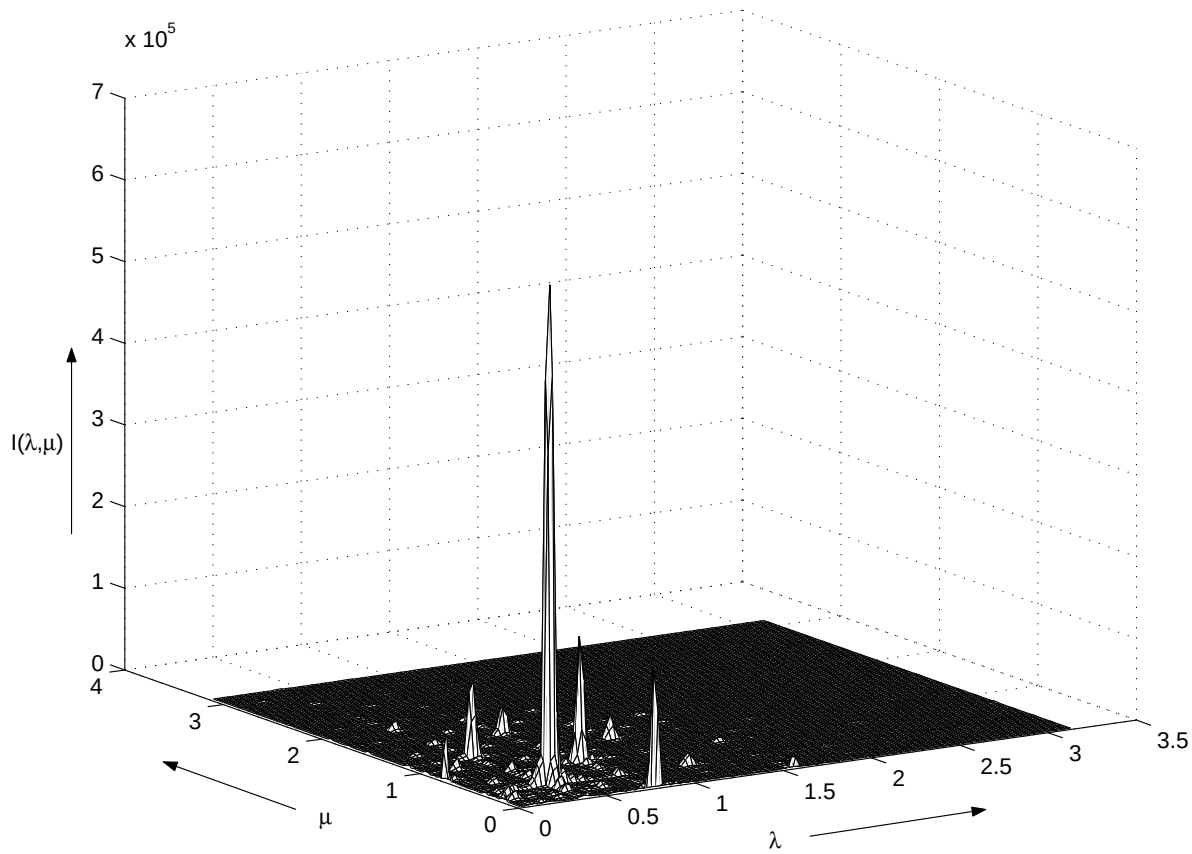

Figure 3: 2-D Periodogram of the Original Texture.

Synthesized Data: Now we analyze a synthesized data obtained from the model (3). The noisy texture is plotted in Figure 7, when the original texture (without the noisy component $X(m,n)$), is plotted in Figure 8. Our problem is to extract the original texture Figure 8 from the noisy texture Figure 7. We have used our sequential procedure to estimate the unknown parameters and obtained the following estimates

$$\widehat{A}_1 = 3.4155, \quad \widehat{B}_1 = 4.3671, \quad \widehat{\lambda}_1 = 1.8039, \quad \widehat{\mu}_1 = 1.1038$$

$$\widehat{A}_2 = 1.1536, \quad \widehat{B}_2 = 0.6645, \quad \widehat{\lambda}_2 = 1.6880, \quad \widehat{\mu}_2 = 0.9955.$$

Based on the above estimates, the estimated texture is plotted in Figure 9. From Figure 8 and Figure 9 it is quite clear that they match quite well.

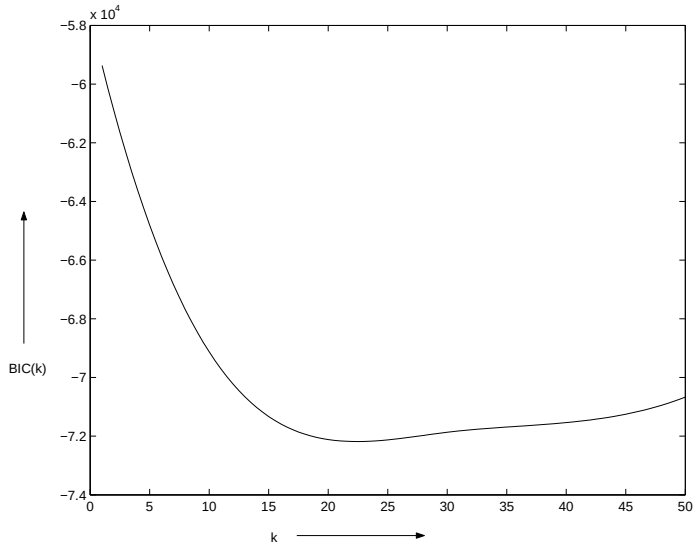

Figure 4: BIC(k) values for different k.

7 Conclusions

In this paper we have provided a sequential procedure for estimating the unknown parameters of the 2-D sinusoidal model. This is an extension of the one dimensional (1-D) sequential estimation procedure recently proposed by the authors, see Prasad *et al.* [10]. It is well known that the 1-D or 2-D sum of sinusoidal models do not satisfy the standard regularity conditions for the consistency and asymptotic normality properties of the least-squares estimators to hold for the general non-linear regression models, see for example Kundu and Nandi [8]. Moreover, due to the complexity of the higher dimensional model, the extension of 1-D results to higher dimensional model is not trivial and hence separate special attention is required for each case. Because of this, extensive work has been done separately for 1-D and 2-D sum of sinusoidal models, to provide different estimation procedures and to establish their properties, see for example the review article by Kundu [6].

Although, the least-squares estimators are the most efficient estimators, but finding the least-squares estimators is a challenging problem. Numerically, it is well known to be a

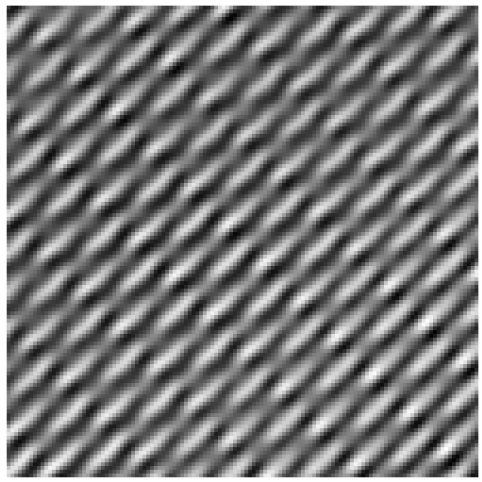

Figure 5: Estimated Texture.

difficult problem, particularly if the two sets of frequencies are very close or if the number of components is very high. For example, if $p = 20$, as we have observed in the texture data, then to find the least-squares estimators, we need to use a forty-dimensional optimization procedure, which is quite difficult to implement. On the other hand in our proposed sequential procedure, we need to solve twenty, 2-D optimization procedures. It is observed that our proposed sequential procedure at each stage produces efficient estimators which are asymptotically equivalent to the least-squares estimators. At each stage we require only a two dimensional optimization procedure, therefore our method is easy to implement and its performance is satisfactory.

Acknowledgements

Part of the work of the second and third authors has been supported by a grant from the Department of Science and Technology, Government of India.

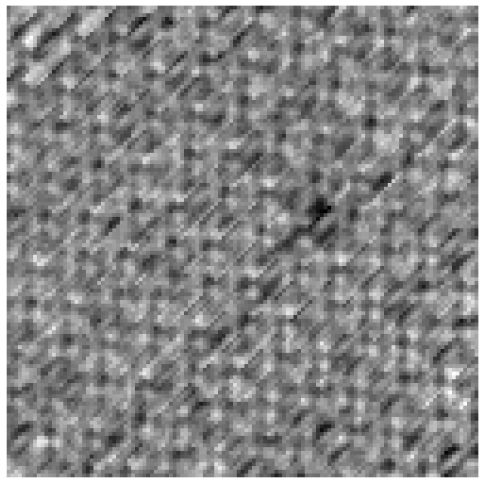

Figure 6: Noise.

Appendix A

Explicitly $\widehat{A}(\lambda,\mu)$ and $\widehat{B}(\lambda,\mu)$ can be written as

$$\begin{bmatrix} \widehat{A}(\lambda,\mu) \\ \\ \widehat{B}(\lambda,\mu) \end{bmatrix} = \left(U^T U\right)^{-1} U^T Y,$$

where

$$U^T U = \begin{bmatrix} \sum_{m=1}^{M}\sum_{n=1}^{N}\cos^2(m\lambda+n\mu) & \sum_{m=1}^{M}\sum_{n=1}^{N}\cos(m\lambda+n\mu)\sin(m\lambda+n\mu) \\ \sum_{m=1}^{M}\sum_{n=1}^{N}\cos(m\lambda+n\mu)\sin(m\lambda+n\mu) & \sum_{m=1}^{M}\sum_{n=1}^{N}\sin^2(m\lambda+n\mu) \end{bmatrix}$$

and

$$U^T Y = \begin{bmatrix} \sum_{m=1}^{M}\sum_{n=1}^{N} y(m,n)\cos(m\lambda+n\mu) \\ \\ \sum_{m=1}^{M}\sum_{n=1}^{N} y(m,n)\sin(m\lambda+n\mu) \end{bmatrix}.$$

Appendix B

First we show here that

$$Q_1'(\theta_1^0)D \xrightarrow{d} N_4(0, 4\sigma^2 c_1 \Sigma_1), \tag{17}$$

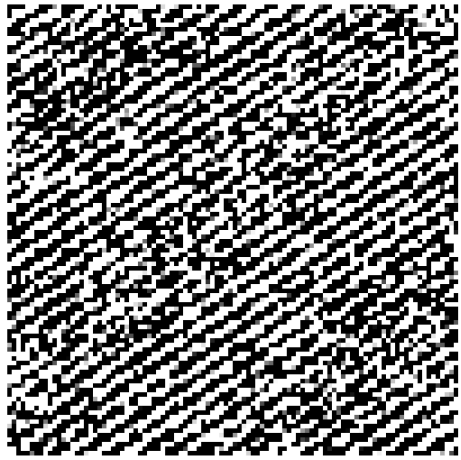

Figure 7: Noisy Synthesized Texture.

To prove (17), we need different elements of $Q_1'(\theta_1^0)$. Note that

$$\begin{aligned}
\frac{\partial Q_1(\theta_1^0)}{\partial A} &= -2\sum_{m=1}^{M}\sum_{n=1}^{N}\cos(m\lambda_1^0+n\mu_1^0)\left[\sum_{j=2}^{p}\left[A_j^0\cos(m\lambda_j^0+n\mu_j^0)+B_j^0\sin(m\lambda_j^0+n\mu_j^0)\right]\right. \\
&\quad \left. + X(m,n)\right] \\
\frac{\partial Q_1(\theta_1^0)}{\partial B} &= -2\sum_{m=1}^{M}\sum_{n=1}^{N}\sin(m\lambda_1^0+n\mu_1^0)\left[\sum_{j=2}^{p}\left[A_j^0\cos(m\lambda_j^0+n\mu_j^0)+B_j^0\sin(m\lambda_j^0+n\mu_j^0)\right]\right. \\
&\quad \left. + X(m,n)\right] \\
\frac{\partial Q_1(\theta_1^0)}{\partial \lambda} &= -2\sum_{m=1}^{M}\sum_{n=1}^{N} m\{A_1^0\sin(m\lambda_1^0+n\mu_1^0)+B_1^0\sin(m\lambda_1^0+n\mu_1^0)\} \\
&\quad \times\left[\sum_{j=2}^{p}\left[A_j^0\cos(m\lambda_j^0+n\mu_j^0)+B_j^0\sin(m\lambda_j^0+n\mu_j^0)\right]+X(m,n)\right]
\end{aligned}$$

Since for $(\alpha_1,\beta_1)\neq(\alpha_2,\beta_2); (\alpha_i,\beta_i)\in((0,\pi)\times(0,\pi)); i=1,2,$

$$\lim_{M,N\longrightarrow\infty}\frac{1}{\sqrt{M}\sqrt{N}}\sum_{m=1}^{M}\sum_{n=1}^{N}\cos(m\alpha_1+n\beta_1)\cos(m\alpha_2+n\beta_2) = 0, \tag{18}$$

$$\lim_{M,N\longrightarrow\infty}\frac{1}{\sqrt{M}\sqrt{N}}\sum_{m=1}^{M}\sum_{n=1}^{N}\sin(m\alpha_1+n\beta_1)\sin(m\alpha_2+n\beta_2) = 0, \tag{19}$$

$$\lim_{M,N\longrightarrow\infty}\frac{1}{M^{\frac{3}{2}}N^{\frac{1}{2}}}\sum_{m=1}^{M}\sum_{n=1}^{N} m\sin(m\alpha_1+n\beta_1)\sin(m\alpha_2+n\beta_2) = 0, \tag{20}$$

$$\lim_{M,N\longrightarrow\infty}\frac{1}{M^{\frac{3}{2}}N^{\frac{1}{2}}}\sum_{m=1}^{M}\sum_{n=1}^{N} m\cos(m\alpha_1+n\beta_1)\cos(m\alpha_2+n\beta_2) = 0, \tag{21}$$

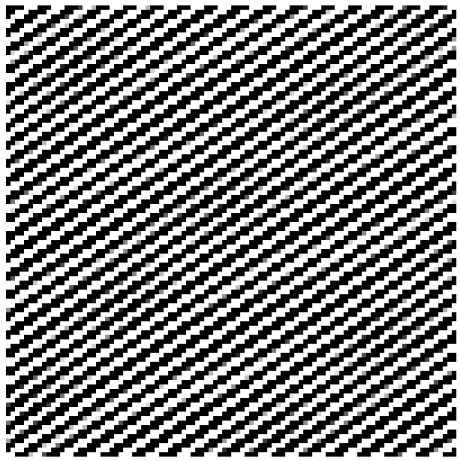

Figure 8: Original Synthesized Texture.

therefore,

$$Q_1'(\theta_1^0)\, D \overset{a.eq.}{=} -2 \begin{bmatrix} M^{-\frac{1}{2}} N^{-\frac{1}{2}} \sum_{m=1}^{M} \sum_{n=1}^{N} \cos(m\lambda_1^0 + n\mu_1^0) X(m,n) \\ M^{-\frac{1}{2}} N^{-\frac{1}{2}} \sum_{m=1}^{M} \sum_{n=1}^{N} \sin(m\lambda_1^0 + n\mu_1^0) X(m,n) \\ M^{-\frac{3}{2}} N^{-\frac{1}{2}} \sum_{m=1}^{M} \sum_{n=1}^{N} m X(m,n) \{A_1^0 \sin(m\lambda_1^0 + n\mu_1^0) + B_1^0 \sin(m\lambda_1^0 + n\mu_1^0)\} \\ M^{-\frac{1}{2}} N^{-\frac{3}{2}} \sum_{m=1}^{M} \sum_{n=1}^{N} n X(m,n) \{A_1^0 \sin(m\lambda_1^0 + n\mu_1^0) + B_1^0 \sin(m\lambda_1^0 + n\mu_1^0)\} \end{bmatrix} \tag{22}$$

Here $\overset{a.eq.}{=}$ means asymptotically equivalent. Now using the Central Limit Theorem (CLT) of the stochastic processes (see Fuller [4]), the right hand side of (22) tends to a 4-variate normal distribution with mean vector 0 and dispersion matrix $4\sigma^2 c_1 \Sigma_1$. Therefore, the result follows.

To prove

$$\lim_{M,N \longrightarrow \infty} D Q_1''(\theta_1^0) D \longrightarrow 2\Sigma_1, \tag{23}$$

we use the following results in addition to (18), (19), (20) and (21),

for $(\alpha_1, \beta_1) \neq (\alpha_2, \beta_2); (\alpha_i, \beta_i) \in ((0,\pi) \times (0,\pi));\ i = 1, 2$ and for $0 < \alpha \neq \beta < \pi$,

$$\lim_{M,N \longrightarrow \infty} \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} m \sin^2(m\alpha + n\beta) = \frac{1}{2}, \tag{24}$$

$$\lim_{M,N \longrightarrow \infty} \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \sin(m\alpha_1 + n\beta_1) \sin(m\alpha_2 + n\beta_2) = 0, \tag{25}$$

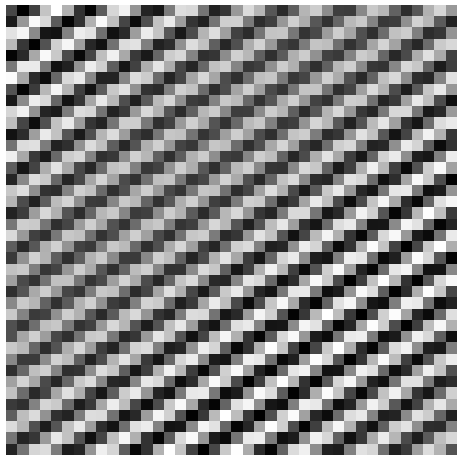

Figure 9: Estimated Synthesized Texture.

$$\lim_{M,N\longrightarrow\infty} \frac{1}{M^2N} \sum_{m=1}^{M} \sum_{n=1}^{N} m \sin^2(m\alpha + n\beta) = \frac{1}{4}, \tag{26}$$

$$\lim_{M,N\longrightarrow\infty} \frac{1}{M^3N} \sum_{m=1}^{M} \sum_{n=1}^{N} m^2 \sin^2(m\alpha + n\beta) = \frac{1}{6}, \tag{27}$$

and similar results for cos function also. Now the results can be obtained by routine calculations mainly considering each element of $Q_1''(\theta_1^0)$ matrix and using the above equalities.

References

[1] Bansal, N.K., Hamedani, G.G. and Zhang, H. (1999), "Non-linear regression with multidimensional indices", *Statistics and Probability Letters*, vol. 45, 175 - 186.

[2] Fisher, R.A. (1929), "Tests of significance in Harmonic analysis", *Proc. Royal Soc. London, Ser A*, 125, 54-59.

[3] Francos, J.M., Meiri, A.Z. and Porat, B. (1993), "A united texture model based on 2-D Wald like decomposition", *IEEE Transactions on Signal Processing*, vol. 41, 2665 - 2678.

[4] Fuller, W. A. (1976), *Introduction to Statistical Time Series*, John Wiley and Sons, New York.

[5] Hartley, H.O. (1949), "Tests of significance in harmonic analysis", *Biometrika*, vol. 36, 194 - 201.

[6] Kundu, D. (2002), "Estimating parameters of sinusoidal frequency; some recent developments", *National Academy of Sciences Letters*, vol. 25, 53-73, 2002.

[7] Kundu, D. and Mitra, A. (1996), "Asymptotic Properties of the Least Squares Estimates of 2-D Exponential Signals", *Multidimensional Systems and Signal Processing*, vol. 7, 135 - 150.

[8] Kundu, D. and Nandi, S. (2003), "Determination of discrete spectrum in a random field", *Statistica Neerlandica*, vol. 57, 258 - 283.

[9] Mitra, A. and Stoica, P. (2002), "Asymptotic Cramer-Rao bound for 2-D superimposed exponential signals", *Multidimensional Systems and Signal Processing*, vol. 13, 317 - 331.

[10] Prasad, A., Kundu, D. and Mitra, A. (2008), "Sequential estimation of the sum of sinusoidal model parameters", *Journal of Statistical Planning and Inference*, vol. 138, 1297 - 1313.

[11] Press, W.H., Teukolsky, S.A., Vellerling, W.T., Flannery, B.P. (1992), *Numerical recipes in FORTRAN, the art of scientific computing*, 2nd. Edition, Cambridge University Press, Cambridge.

[12] Rao, C.R., Zhao, L.C. and Zhou, B. (1994), "Maximum likelihood estimation of 2-D superimposed exponential signals", *IEEE Transactions on Signal Processing*, vol. 42, 1795-1802.

[13] Richards, F. S. G. (1961), "A method of maximum likelihood estimation", *Journal of the Royal Statistical Society*, B, 469-475.

[14] Stoica, P. (1993), "List of references on spectral analysis", *Signal Processing*, vol. 31, 329-340.

[15] Wu, C.F.J. (1981), "Asymptotic theory of non-linear least squares estimation", *The Annals of Statistics*, vol. 9, 501 - 513.

[16] Zeng G. and Dubes R. (1985), "A comparison of tests for randomness", *Pattern recognition*, vol. 18, 191-198.

[17] Zhang, H. and Mandrekar, V. (2001), "Estimation of hidden frequencies for 2D stationary processes", *Journal of Time Series Analysis*, vol. 22, 613 - 629.