Modeling and Estimation of Symmetric Color Textures

Anurag Prasad[†] & Debasis Kundu[†]

Abstract

In this paper we have considered the sum of three-dimensional (3-D), sinusoidal model. It is observed that this model can be used quite effectively to model symmetric color textures. The estimation problem is particularly difficult when the adjacent frequency-sets are not well separated or when the number of components is very large. Using the fact that the regressor vectors are orthogonal, we propose a simple sequential procedure to estimate the unknown frequencies and amplitudes of the 3-D sinusoidal signals. It is observed that if there are p components in the signal then our procedure at the k-th ($k \le p$) stage produces strongly consistent estimators of the k dominant sinusoids. For $k > p$, the amplitude estimators converge to zero almost surely. Asymptotic distribution of the proposed sequential estimators is established and it is observed that it coincides with the asymptotic distribution of the least squares estimators. Numerical simulations are performed to observe the performance of the proposed estimators for different sample sizes and models. One color texture data and one synthesized data are analyzed for illustration.

Keywords: Sinusoidal signals; least squares estimators; asymptotic distribution; over and under determined models, strongly consistent estimators; 3-D sinusoidal signals.

[†] 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

We consider the following problem;

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

where $m = 1, \ldots M; n = 1, \ldots, N$ and $s = 1, \ldots, S$. Here the amplitudes A_k^0 and B_k^0 are such that $A_k^{0^2} + B_k^{0^2} \neq 0$, the frequencies λ_k^0, μ_k^0, $\nu_k^0 \in (0, \pi)$ for $k = 1 \ldots, p$, and they are distinct. The error random variables $X = \{X(m,n,s); 1 \leq m \leq M, 1 \leq n \leq N, 1 \leq s \leq S\}$ are from a stationary random field with mean zero and a finite variance. The explicit assumptions on the random field X will be given in the next section. The problem is to estimate the unknown parameters $A^0 = (A_k^0)_{1\leq k\leq p}$, $B^0 = (B_k^0)_{1\leq k\leq p}$, $\lambda^0 = (\lambda_k^0)_{1\leq k\leq p}$, $\mu^0 = (\mu_k^0)_{1\leq k\leq p}$ and $\nu^0 = (\nu_k^0)_{1\leq k\leq p}$ given an observed sample $\{y(m,n,s); \quad m = 1, \ldots, M, \quad n = 1, \ldots, N, \quad s = 1, \ldots, S\}$.

Several two-dimensional (2-D) models under various assumptions on error have been used to describe gray-scale textures. Zhang and Mandrekar [17], Kundu and Gupta [8] considered the following model;

$$y(m,n) = \sum_{k=1}^{p} [A_k^0 \cos(m\lambda_k^0 + n\mu_k^0)] + X(m,n)$$

where $\{X(m,n); 1 \leq m \leq M, 1 \leq n \leq N\}$ are assumed to be independent and identically distributed ($i.i.d$) random variables. Kundu and Nandi [9] used a simple generalization of the above model as;

$$y(m,n) = \sum_{k=1}^{p} [A_k^0 \cos(m\lambda_k^0 + n\mu_k^0) + B_k^0 \sin(m\lambda_k^0 + n\mu_k^0)] + X(m,n), \quad (2)$$

where the additive error $X(m,n)$ is assumed to be from a stationary random field. For some related work the readers are referred to Bansal *et al.* [1], Francos *et al.* [2], Mitra and Stoica [10], Rao *et al.* [13] and the references cited therein.

In each of the above two models, $y(m, n)$ represents a specific intensity in the gray-scale of the pixel at the (m, n)-th position in the picture. For instance, if 0 represents black and 1 represents white, then any real number $\in [0, 1]$ corresponds to a particular intensity of gray. It is clear that the size of picture is $M \times N$ pixels.

Here is a brief description of how the color pictures are stored digitally in RGB (Red-Green-Blue) format. In the digital representation of RGB type color pictures, the color at each pixel is determined by the red, green and blue intensity at that pixel. It means we need to store three values to determine the actual color at each pixel. We can store a color picture of size $M \times N$ in a $M \times N \times 3$ array, where the first layer (composed of the elements $y(1, 1, 1), y(1, 2, 1), \ldots, y(M, N, 1)$) of size $M \times N$, represents intensities of red, second layer, that of green and the third layer represents blue intensity.

Figure 1 represents an original symmetric color texture obtained from the following cite http://local.wasp.uwa.edu.au/∼pbourke/texture_colour.

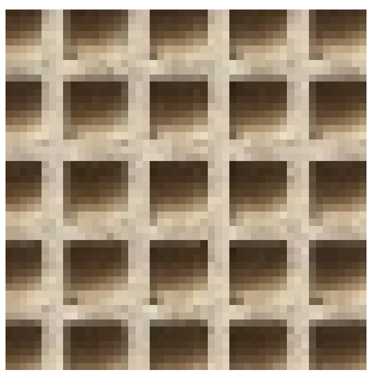


Figure 1: An original Color Texture

Its size is 50×50 pixels. Consider the pixel at position $(1, 1)$. Its color is determined by three elements $y(1, 1, 1) = 0.2471$, $y(1, 1, 2) = 0.1725$ and $y(1, 1, 3) = 0.1059$, which represents intensities of red, green and blue respectively. These three numbers, collectively, are called an RGB-triplet. We have an RGB-triplet for each of the MN pixels, which collectively determine the color of the full picture of size $M \times N$.

We try to model the above symmetric color texture using (1). Here a large sample of observations will correspond to a bigger picture, which in turn implies that M and N tend to infinity, but the third dimension S always remains fixed at 3.

The problem of estimating the parameters even for the one-dimension (1-D) or two-dimension (2-D) sinusoidal models, is well known to be numerically difficult. Even for 1-D sinusoidal model, when $p \geq 2$ and the separation of the two frequencies is not much, it is very difficult to estimate the unknown parameters, see Kay [5]. Several methods are available in the literature for estimating the parameters of the sinusoidal signals, see for example the review article by Kundu [7]. The most efficient estimators, of course, are the least squares estimators. The rates of convergence of the least squares estimators are $O_p(N^{-\frac{3}{2}})$ and $O_p(N^{-\frac{1}{2}})$ respectively for the frequencies and amplitudes, see Hannan [4], Walker [15] or Kundu [6], where $U = O_p(N^{-\delta})$ means UN^{δ} is bounded in probability. But it is well known that the least squares estimators are difficult to compute numerically, since there are several local minima on the least squares surface. It is observed that if two frequency-sets are very close to each other or if the number of components is very large, then finding the initial guess itself is very difficult and therefore starting any iterative process for finding the least squares estimators is not easy. One of the standard methods to find the initial guesses of the frequencies is to find the maxima at the Fourier frequencies of the periodogram function. It is known that asymptotically the periodogram function has local maxima at the true frequencies, see Walker [15] and Kundu and Nandi [9].

Another practical problem occurs while using the least squares estimators when p is very large. It is observed in this paper that for some of the color textures the value of p can be as large as 66. Therefore, in a high dimensional optimization problem, the choice of initial guess can be very crucial. Further, since several local minima are present on the objective function surface, often the iterative process will converge to a local rather that the global optimum point.

The aim of this paper is two-fold. First, if the number of components, p, is known, using the fact that the regressor vectors are orthogonal, we propose a step-by-step sequential procedure to estimate the unknown parameters. It is observed that the $3p$-dimensional optimization problem can be reduced to p 3-D optimization problems, which are computationally tractable. Therefore, if p is large then the proposed method can be very useful. Moreover, it is observed that the estimators obtained by the proposed method have the same asymptotic rate of convergence as the least squares estimators.

The second aim of this paper is to study the properties of the estimators if the number of components p is unknown. If p is not known and we fit a lower order model, it is observed that the proposed estimators are consistent for the dominant components with the same convergence rate as the least squares estimators. If we fit a higher order model, then the estimators obtained up to p-th step are consistent for the unknown parameters with the same convergence rate as the least squares estimators and both the amplitude estimates after the p-th step converge to zero almost surely. We perform some numerical simulations to study the behavior of the proposed estimators. One synthesized data and one color texture data have been analyzed for illustrative purpose.

The rest of the paper is organized as follows. In section 2, we provide the necessary assumptions and also the methodology. Consistency of the proposed estimators are obtained in section 3. Asymptotic distributions or the convergence rates are provided in section 4. Numerical examples are provided in section 5. Data analysis results are presented in section 6 and finally we conclude the paper in section 7.

2 Model Assumptions and Methodology

2.1 Assumptions

It is assumed that the observations are obtained from model (1). We make the following assumptions on the parameters and the error. The additive error $\{X(m,n,s)\}$ is from a stationary linear process with mean zero and finite variance. It satisfies the following Assumption 1. From now onward, we will denote the set of positive integers as $\mathcal{Z}$.

Assumption 1: $\{X(m,m,s); m,n \in \mathcal{Z}, s = 1,\ldots,S\}$ can be represented as;

$$X(m,n,s) = \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} \sum_{w=-\infty}^{\infty} a(u,v,w) e(m-u, n-v, s-w), \tag{3}$$

where $a(u,v,w)$s are real constants that satisfy;

$$\sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} \sum_{w=-\infty}^{\infty} |a(u,v,w)| < \infty, \tag{4}$$

and $\{e(m,n,s)\}$ is a sequence of $i.i.d.$ random variables with mean zero and variance $\sigma^2 < \infty$.

Assumption 2: The frequencies λ_k^0, μ_k^0, $\nu_k^0 \in (0,\pi)$ for $k = 1\ldots,p$, and $\lambda_k^0 \neq \lambda_q^0$, $\mu_k^0 \neq \mu_q^0$, $\nu_k^0 \neq \nu_q^0$ for $1 \leq k \neq q \leq p$.

Assumption 3: The amplitudes satisfy the following restrictions;

$$0 < A_p^{0^2} + B_p^{0^2} < \ldots < A_1^{0^2} + B_1^{0^2} < K^2 < \infty \tag{5}$$

2.2 Methodology

We propose the following sequential procedure to estimate the unknown parameters. The method can be applied even when p is unknown. At the first step we minimize the following quantity;

$$Q_1(A,B,\lambda,\mu,\nu)$$

$$= \sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{s=1}^{S}\left[y(m,n,s) - A\cos(m\lambda + n\mu + s\nu) - B\sin(m\lambda + n\mu + s\nu)\right]^2 \qquad (6)$$

with respect to A, B, λ, μ and ν. Using the separable regression technique of Richards [14], it can be easily shown that for fixed λ, μ and ν, $\widehat{A}(\lambda, \mu, \nu)$ and $\widehat{B}(\lambda, \mu, \nu)$, where the explicit expressions of $\widehat{A}(\lambda, \mu, \nu)$ and $\widehat{B}(\lambda, \mu, \nu)$ are provided in the Appendix 1, minimize (6). Replacing A and B by $\widehat{A}(\lambda, \mu, \nu)$ and $\widehat{B}(\lambda, \mu, \nu)$, respectively, we obtain;

$$\begin{aligned} R_1(\lambda, \mu, \nu) &= Q_1(\widehat{A}(\lambda, \mu, \nu), \widehat{B}(\lambda, \mu, \nu), \lambda, \mu, \nu) \\ &= \sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{s=1}^{S}\Big[y(m,n,s) - \widehat{A}(\lambda, \mu, \nu)\cos(m\lambda + n\mu + s\nu) \\ &\quad - \widehat{B}(\lambda, \mu, \nu)\sin(m\lambda + n\mu + s\nu)\Big]^2. \end{aligned} \qquad (7)$$

Therefore, if $(\widehat{\lambda}, \widehat{\mu}, \widehat{\nu})$ minimize $R_1(\lambda, \mu, \nu)$, then $(\widehat{A}(\widehat{\lambda}, \widehat{\mu}, \widehat{\nu}), \widehat{B}(\widehat{\lambda}, \widehat{\mu}, \widehat{\nu}), \widehat{\lambda}, \widehat{\mu}, \widehat{\nu})$ minimize (6). We will denote these estimators as $(\widehat{A}_1, \widehat{B}_1, \widehat{\lambda}_1, \widehat{\mu}_1, \widehat{\nu}_1)$.

Now we consider the following sequence for $m = 1, \ldots, M$, $n = 1, \ldots, N$ and $s = 1, \ldots, S$;

$$y^{(1)}(m,n,s) = y(m,n,s) - \hat{A}_1\cos(m\widehat{\lambda}_1 + n\widehat{\mu}_1 + s\widehat{\nu}_1) - \hat{B}_1\sin(m\widehat{\lambda}_1 + n\widehat{\mu}_1 + s\widehat{\nu}_1), \qquad (8)$$

and minimize $Q_2(A, B, \lambda, \mu, \nu)$ which is obtained in (6) by replacing $y(m,n,s)$ with $y^{(1)}(m,n,s)$. The estimators obtained by minimizing $Q_2(A, B, \lambda, \mu, \nu)$ are denoted by $(\widehat{A}_2, \widehat{B}_2, \widehat{\lambda}_2, \widehat{\mu}_2, \widehat{\nu}_2)$.

If p is the number of terms of the model (1) and it is known, continue the process up to p steps. If p is not known then we fit sequentially a k order model, where k may not be equal to p. In the next section we provide the properties of the proposed estimators in both cases when p is known/ unknown.

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, k, is less than or equal to the actual number of components, p and (b) when $k > p$. We need the following lemma;

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

$$\sup_{\alpha,\beta,\gamma} \left| \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{s=1}^{S} X(m,n,s) e^{i(m\alpha+n\beta+s\gamma)} \right| \longrightarrow 0 \quad a.s. \tag{9}$$

PROOF See the Appendix 2. ∎

LEMMA 2: Consider the set $S_c = \{\theta; \theta \in \Theta, \text{ and } |\theta - \theta_1^0| \geq c\}$, where $\theta = (A, B, \lambda, \mu, \nu)$ and $\theta_1^0 = (A_1^0, B_1^0, \lambda_1^0, \mu_1^0, \nu_1^0)$, $\Theta = [-K, K] \times [-K, K] \times [0, \pi] \times [0, \pi] \times [0, \pi]$. If for any $c > 0$,

$$\liminf_{M,N\to\infty} \inf_{\theta \in S_c} \frac{1}{MN} \left[Q_1(\theta) - Q_1(\theta_1^0) \right] > 0 \;\; a.s., \tag{10}$$

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

PROOF: It follows from Lemma 1 of Wu [16]. ∎

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

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

where

$$\begin{aligned} f(\theta) = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{s=1}^{S} & (A_1^0 \cos(m\lambda_1^0 + n\mu_1^0 + s\nu_1^0) + B_1^0 \sin(m\lambda_1^0 + n\mu_1^0 + s\nu_1^0) \\ & - A \cos(m\lambda + n\mu + s\nu) - B \sin(m\lambda + n\mu + s\nu))^2 \\ & + \frac{2}{MN} \left\{ A_1^0 \cos(m\lambda_1^0 + n\mu_1^0 + s\nu_1^0) + B_1^0 \sin(m\lambda_1^0 + n\mu_1^0 + s\nu_1^0) \right. \\ & \qquad \left. - A \cos(m\lambda + n\mu + s\nu) - B \sin(m\lambda + n\mu + s\nu) \right\} \times \end{aligned}$$

$$\times \left[\sum_{k=2}^{p} \left\{ A_k^0 \cos(m\lambda_k^0 + n\mu_k^0 + s\nu_k^0) + B_k^0 \sin(m\lambda_k^0 + n\mu_k^0 + s\nu_k^0) \right\} \right],$$

and

$$g(\theta) = \frac{2}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{s=1}^{S} X(m,n,s)[A_1^0 \cos(m\lambda_1^0 + n\mu_1^0 + s\nu_1^0) + B_1^0 \sin(m\lambda_1^0 + n\mu_1^0 + s\nu_1^0)$$
$$- A\cos(m\lambda + n\mu + s\nu) - B\sin(m\lambda + n\mu + s\nu)].$$

Now using Lemma 1, it easily follows that

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

Note that (see Appendix 3)

$$\liminf_{M,N\to\infty} \inf_{\theta \in S_C} f(\theta) > 0, \tag{11}$$

that proves the result. ∎.

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

LEMMA 3: If the Assumption 1-3 are satisfied, then as $M \to \infty$ and $N \to \infty$,

$$M(\widehat{\lambda}_1 - \lambda_1^0) \to 0 \quad a.s., \quad N(\widehat{\mu}_1 - \mu_1^0) \to 0 \quad a.s. \tag{12}$$

PROOF: Consider the 5×5 diagonal matrix $D_1 = \text{diag}\left\{1,\ 1,\ M^{-1},\ N^{-1},\ \tau^{-\frac{1}{2}} S^{\frac{1}{2}}\right\}$. Where $\tau = \sum_{s=1}^{S} s^2$. Here τ has been chosen so as to make the $(5,5)$-th element of the matrix Σ_1 (defined in (15)) independent of S.

Using the multivariate Taylor series expansion we can write

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

Here

$$Q_1'(\theta) = \left[\frac{\partial Q_1(\theta)}{\partial A}, \frac{\partial Q_1(\theta)}{\partial B}, \frac{\partial Q_1(\theta)}{\partial \lambda}, \frac{\partial Q_1(\theta)}{\partial \mu}, \frac{\partial Q_1(\theta)}{\partial \nu}\right], \tag{13}$$

and

$$Q_1''(\theta) = \begin{bmatrix} \frac{\partial^2 Q_1(\theta)}{\partial A^2} & \frac{\partial^2 Q_1(\theta)}{\partial A \partial B} & \frac{\partial^2 Q_1(\theta)}{\partial A \partial \lambda} & \frac{\partial^2 Q_1(\theta)}{\partial A \partial \mu} & \frac{\partial^2 Q_1(\theta)}{\partial A \partial \nu} \\ \frac{\partial^2 Q_1(\theta)}{\partial B \partial A} & \frac{\partial^2 Q_1(\theta)}{\partial B^2} & \frac{\partial^2 Q_1(\theta)}{\partial B \partial \lambda} & \frac{\partial^2 Q_1(\theta)}{\partial B \partial \mu} & \frac{\partial^2 Q_1(\theta)}{\partial B \partial \nu} \\ \frac{\partial^2 Q_1(\theta)}{\partial \lambda \partial A} & \frac{\partial^2 Q_1(\theta)}{\partial \lambda \partial B} & \frac{\partial^2 Q_1(\theta)}{\partial \lambda^2} & \frac{\partial^2 Q_1(\theta)}{\partial \lambda \partial \mu} & \frac{\partial^2 Q_1(\theta)}{\partial \lambda \partial \nu} \\ \frac{\partial^2 Q_1(\theta)}{\partial \mu \partial A} & \frac{\partial^2 Q_1(\theta)}{\partial \mu \partial B} & \frac{\partial^2 Q_1(\theta)}{\partial \mu \partial \lambda} & \frac{\partial^2 Q_1(\theta)}{\partial \mu^2} & \frac{\partial^2 Q_1(\theta)}{\partial \mu \partial \nu} \\ \frac{\partial^2 Q_1(\theta)}{\partial \nu \partial A} & \frac{\partial^2 Q_1(\theta)}{\partial \nu \partial B} & \frac{\partial^2 Q_1(\theta)}{\partial \nu \partial \lambda} & \frac{\partial^2 Q_1(\theta)}{\partial \nu \partial \mu} & \frac{\partial^2 Q_1(\theta)}{\partial \nu^2} \end{bmatrix}. \tag{14}$$

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

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

Using Theorem 1 and trigonometric identities, see for example Appendix 5.D (Page no. 160) of Prasad [11] for details, it follows that;

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

which is a positive definite matrix, given as $\Sigma_1 =$

$$\begin{bmatrix} \frac{1}{2} & 0 & \frac{B_1^0}{4} & \frac{B_1^0}{4} & \frac{B_1^0}{4}\tau^{-\frac{1}{2}} S^{\frac{1}{2}}(S+1) \\ 0 & \frac{1}{2} & -\frac{A_1^0}{4} & -\frac{A_1^0}{4} & -\frac{A_1^0}{4}\tau^{-\frac{1}{2}} S^{\frac{1}{2}}(S+1) \\ \frac{B_1^0}{4} & -\frac{A_1^0}{4} & \frac{(A_1^{0\,2}+B_1^{0\,2})}{6} & \frac{(A_1^{0\,2}+B_1^{0\,2})}{8} & \frac{(A_1^{0\,2}+B_1^{0\,2})}{8}\tau^{-\frac{1}{2}} S^{\frac{1}{2}}(S+1) \\ \frac{B_1^0}{4} & -\frac{A_1^0}{4} & \frac{(A_1^{0\,2}+B_1^{0\,2})}{8} & \frac{(A_1^{0\,2}+B_1^{0\,2})}{6} & \frac{(A_1^{0\,2}+B_1^{0\,2})}{8}\tau^{-\frac{1}{2}} S^{\frac{1}{2}}(S+1) \\ \frac{B_1^0}{4}\tau^{-\frac{1}{2}} S^{\frac{1}{2}}(S+1) & -\frac{A_1^0}{4}\tau^{-\frac{1}{2}} S^{\frac{1}{2}}(S+1) & \frac{(A_1^{0\,2}+B_1^{0\,2})}{8}\tau^{-\frac{1}{2}} S^{\frac{1}{2}}(S+1) & \frac{(A_1^{0\,2}+B_1^{0\,2})}{8}\tau^{-\frac{1}{2}} S^{\frac{1}{2}}(S+1) & \frac{(A_1^{0\,2}+B_1^{0\,2})}{2} \end{bmatrix}. \tag{15}$$

Using Lemma 1, it follows that

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

Therefore, $(\widehat{\theta}_1 - \theta_1^0) D_1^{-1} \to 0 \quad a.s.$, hence the lemma. ∎

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

Proof: Using Theorem 1 and Lemma 3 we obtain;

$$\widehat{A}_1 = A_1^0 + o(1) \quad a.s., \quad \widehat{B}_1 = B_1^0 + o(1) \quad a.s.$$

$$\widehat{\lambda}_1 = \lambda_1^0 + o(M) \quad a.s., \quad \widehat{\mu}_1 = \mu_1^0 + o(N) \quad a.s.$$

$$\widehat{\nu}_1 = \nu_1^0 + o(1) \quad a.s.$$

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

$$\widehat{A}_1 \cos(\widehat{\lambda}_1 m + \widehat{\mu}_1 n + \widehat{\nu}_1 s) + \widehat{B}_1 \sin(\widehat{\lambda}_1 m + \widehat{\mu}_1 n + \widehat{\nu}_1 s)$$
$$= A_1^0 \cos(\lambda_1 m + \mu_1 n + \nu_1 s) + B_1^0 \sin(\lambda_1^0 m + \mu_1^0 n + \nu_1 s) + o(1) \ \ a.s.,$$

and using the fact that the regressor vectors are orthogonal. ∎

The result can be extended up to the k-th step for $1 \leq k \leq p$. We can formally state the result as follows.

Theorem 3: If the Assumptions 1-3 are satisfied for $k \leq p$, then $\widehat{\theta}_k$, the estimator obtained by minimizing $Q_k(A, B, \lambda, \mu, \nu)$, where $Q_k(A, B, \lambda, \mu, \nu)$ is defined analogously to $Q_2(A, B, \omega, \lambda, \mu, \nu)$ for the k-th step, is a consistent estimator of θ_k^0.

It may be mentioned that the estimators obtained at different steps produce consistent estimates because here the regressor vectors are asymptotically orthogonal. Now we will investigate the properties of the estimators if the sequential process is continued even after p-th step. We need the following lemma for this;

Lemma 4: If $\{X(m, n, s)\}$ is same as defined in Assumption 1, and $\widehat{A}$, $\widehat{B}$, $\widehat{\lambda}$, $\widehat{\mu}$ and $\widehat{\nu}$ are obtained by minimizing

$$\sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{s=1}^{S} \left(X(m, n, s) - A\cos(m\lambda + n\mu + s\nu) - B\sin(m\lambda + n\mu + s\nu)\right)^2,$$

then

$$\widehat{A} \longrightarrow 0 \quad a.s. \quad \text{and} \quad \widehat{B} \longrightarrow 0 \quad a.s.$$

Proof: Using the similar steps as in Walker [15], it easily follows that

$$\widehat{A} = \frac{2}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{s=1}^{S} X(m, n, s) \cos(m\widehat{\lambda} + n\widehat{\mu} + s\widehat{\nu}) + o(1) \qquad a.s.$$

$$\widehat{B} = \frac{2}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{s=1}^{S} X(m,n,s) \sin(m\widehat{\lambda} + n\widehat{\mu} + s\widehat{\nu}) + o(1) \quad a.s.$$

Now using Lemma 1, the result follows. Therefore, we have the following result.

THEOREM 4 If the Assumptions 1-3 are satisfied, then for $k > p$, if $\widehat{\theta}_k = (\widehat{A}_k, \widehat{B}_k, \widehat{\lambda}_k, \widehat{\mu}_k, \widehat{\nu}_k)$ minimizes $Q_k(A, B, \lambda, \mu, \nu)$, then

$$\widehat{A}_k \longrightarrow 0 \quad a.s. \quad \text{and} \quad \widehat{B}_k \longrightarrow 0 \quad a.s..$$

4 Asymptotic Distribution of the Estimators

We obtain the asymptotic distribution of the proposed estimators at the k-th step, $(k \leq p)$ in this section. We will denote $Q_1(A, B, \lambda, \mu, \nu)$ by $Q_1(\theta)$, *i.e.*,

$$Q_1(\theta) = \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{s=1}^{S} \left(y(m,n,s) - A\cos(m\lambda + n\mu + s\nu) - B\sin(m\lambda + n\mu + s\nu)\right)^2. \quad (16)$$

Now from multivariate Taylor series expansion, we obtain;

$$Q_1'(\widehat{\theta}_1) - Q_1'(\theta_1^0) = (\widehat{\theta}_1 - \theta_1^0) \; Q_1''(\bar{\theta}), \quad (17)$$

where $Q_1'(\theta_1)$ and $Q_1''(\theta_1)$ are given by (13) and (14), respectively and $\bar{\theta} = (\bar{A}, \bar{B}, \bar{\lambda}, \bar{\mu}, \bar{\nu})$ is a point on the line joining $\widehat{\theta}_1$ and θ_1^0. Note that $Q_1'(\widehat{\theta}_1) = 0$. Consider the following 5×5 diagonal matrix D as follows;

$$D = \begin{bmatrix} M^{-\frac{1}{2}}N^{-\frac{1}{2}}S^{-\frac{1}{2}} & 0 & 0 & 0 & 0 \\ 0 & M^{-\frac{1}{2}}N^{-\frac{1}{2}}S^{-\frac{1}{2}} & 0 & 0 & 0 \\ 0 & 0 & M^{-\frac{3}{2}}N^{-\frac{1}{2}}S^{-\frac{1}{2}} & 0 & 0 \\ 0 & 0 & 0 & M^{-\frac{1}{2}}N^{-\frac{3}{2}}S^{-\frac{1}{2}} & 0 \\ 0 & 0 & 0 & 0 & M^{-\frac{1}{2}}N^{-\frac{1}{2}}\tau^{-\frac{1}{2}} \end{bmatrix}. \quad (18)$$

Therefore, (17) can be written as

$$(\widehat{\theta}_1 - \theta_1^0)D^{-1} = -Q_1'(\theta_1^0) \; D \left[D \; Q_1''(\bar{\theta}) \; D\right]^{-1}. \quad (19)$$

Now observe that $\bar{\theta} \longrightarrow \theta_1^0 \ \ a.s.$, therefore,

$$\lim_{M,N \longrightarrow \infty} DQ_1''(\bar{\theta})D = \lim_{M,N \longrightarrow \infty} DQ_1''(\theta_1^0)D.$$

It can be shown by straightforward but tedious calculations that

$$\lim_{M,N \longrightarrow \infty} DQ_1''(\theta_1^0)D \longrightarrow 2\Sigma_1, \tag{20}$$

where Σ_1 is given by (15), see for example Appendix 5.D (Page no. 160) of Prasad [11] for details. Further using the Central Limit Theorem of the linear process (see Fuller [3], Page 251-256) it follows that

$$Q_1'(\theta_1^0)\ D \stackrel{d}{\longrightarrow} N_5(0, 4\sigma^2 c_1 \Sigma_1), \tag{21}$$

where

$$\begin{aligned} c_1 = & \left| \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} a(i,j,k) \cos(i\lambda_1^0 + j\mu_1^0 + k\nu_1^0) \right|^2 \\ & + \left| \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} a(i,j,k) \sin(i\lambda_1^0 + j\mu_1^0 + k\nu_1^0) \right|^2, \end{aligned} \tag{22}$$

and '$\stackrel{d}{\longrightarrow}$' means convergence in distribution. Therefore, we have the following result;

Theorem 5: If the Assumption 1-3 are satisfied, then $\Big[M^{\frac{1}{2}}N^{\frac{1}{2}}S^{\frac{1}{2}}(\widehat{A}_1 - A_1^0), M^{\frac{1}{2}}N^{\frac{1}{2}}S^{\frac{1}{2}}(\widehat{B}_1 - B_1^0), M^{\frac{3}{2}}N^{\frac{1}{2}}S^{\frac{1}{2}}(\widehat{\lambda}_1 - \lambda_1^0), M^{\frac{1}{2}}N^{\frac{3}{2}}S^{\frac{1}{2}}(\widehat{\mu}_1 - \mu_1^0), M^{\frac{1}{2}}N^{\frac{1}{2}}S^{\frac{3}{2}}(\widehat{\nu}_1 - \nu_1^0)\Big] \stackrel{d}{\longrightarrow} N_5\left(0, \sigma^2 c_1 \Sigma_1^{-1}\right)$, where $M, N \to \infty$, $S < \infty$, and Σ_1 and c_1 are given in (15) and (22) respectively.

Proceeding exactly in the same manner, and using Theorem 2, it can be shown that similar result holds for any $k \leq p$ and it can be stated as follows.

Theorem 6: If the Assumption 1-3 are satisfied, then, for $k \leq p$, $\Big[M^{\frac{1}{2}}N^{\frac{1}{2}}S^{\frac{1}{2}}(\widehat{A}_k - A_k^0), M^{\frac{1}{2}}N^{\frac{1}{2}}S^{\frac{1}{2}}(\widehat{B}_k - B_k^0), M^{\frac{3}{2}}N^{\frac{1}{2}}S^{\frac{1}{2}}(\widehat{\lambda}_k - \lambda_k^0), M^{\frac{1}{2}}N^{\frac{3}{2}}S^{\frac{1}{2}}(\widehat{\mu}_k - \mu_k^0), M^{\frac{1}{2}}N^{\frac{1}{2}}S^{\frac{3}{2}}(\widehat{\nu}_k - \nu_k^0)\Big] \stackrel{d}{\longrightarrow} N_5\left(0, \sigma^2 c_k \Sigma_k^{-1}\right)$, where $M, N \to \infty$, $S < \infty$, and Σ_k and c_k are obtained from Σ_1 and c_1 by replacing $A_1^0, B_1^0, \lambda_1^0, \mu_1^0, \nu_1^0$ by $A_k^0, B_k^0, \lambda_k^0, \mu_k^0, \nu_k^0$.

5 Numerical Results

Here we present some numerical results to see the asymptotic behavior of the estimates for different sample sizes and different models. All the computations are performed at the Indian Institute of Technology Kanpur, using the random deviate generator RAN2 of Press *et al.* [12]. All the programs are written in FORTRAN-77. We have considered the following model;

$$y(m,n,s) = \sum_{k=1}^{2} \{A_k \cos(m\lambda_k + n\mu_k + s\nu_k) + B_k \sin(m\lambda_k + n\mu_k + s\nu_k)\} + X(m,n,s).$$

Here $A_1 = 10.0$, $B_1 = 10.0$, $\lambda_1 = 2.0$, $\mu_1 = 2.0$, $\nu_1 = 2.0$, $A_2 = 5.0$ $B_2 = 5.0$, $\lambda_2 = 1.5$, $\mu_2 = 1.5$, $\nu_2 = 1.5$ and

$$X(m,n,s) = e(m,n,s) + 0.33e(m-1,n,s) + 0.33e(m,n-1,s) + 0.33e(m,n,s-1), \quad (23)$$

where $e(m,n,s)$ are *i.i.d.* normal random variables with mean zero and variance one. For each model, we considered different k values and different sample sizes. Mainly we have reported the following cases: $k = 1$, $k = 2$ and $k = 3$, *i.e.* lower order model, exact order model and higher order model respectively. In each case we have generated the sample using the model parameters and the error structure as provided in (23). Then for each fixed k, we estimate the model parameters using the sequential procedure provided in Section 2. At each step, 3-D optimization has been performed using the downhill simplex method as described in Press *et al.* [12]. In each step we repeat the procedure 1000 times and report the average estimates and mean squared errors for all the unknown parameters. For comparison purposes, we report the asymptotic variances also for all the estimates. The results are reported in Tables 1, 2 and 3.

Some of the points are quite clear from the table values. As the sample size increases the biases and the mean squared errors (MSEs) decrease. That verifies the consistency properties

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

		$A_1 = 10.0$	$B_1 = 10.0$	$\lambda_1 = 2.0$	$\mu_1 = 2.0$	$\nu_1 = 2.0$
S= 3	AE	13.8202	4.7191	1.9614	1.9613	1.9794
M= 10	MSE	(0.147E+02)	(0.283E+02)	(0.151E-02)	(0.151E-02)	(0.530E-03)
N= 10	ASYV	(0.261E-01)	(0.261E-01)	(0.224E-05)	(0.224E-05)	(0.280E-04)
S= 3	AE	10.5653	9.5766	1.9988	1.9988	1.9953
M= 20	MSE	(0.360E+00)	(0.224E+00)	(0.226E-05)	(0.221E-05)	(0.545E-04)
N= 20	ASYV	(0.653E-02)	(0.653E-02)	(0.140E-06)	(0.140E-06)	(0.700E-05)
S= 3	AE	10.0248	9.8697	1.9998	1.9998	2.0026
M= 30	MSE	(0.201E-01)	(0.346E-01)	(0.273E-06)	(0.236E-06)	(0.202E-04)
N= 30	ASYV	(0.290E-02)	(0.290E-02)	(0.277E-07)	(0.277E-07)	(0.311E-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 2: Model 1 is considered with $k = 2^*$.

		$A_2 = 5.0$	$B_2 = 5.0$	$\lambda_2 = 1.5$	$\mu_2 = 1.5$	$\nu_2 = 1.5$
S= 3	AE	6.3076	1.7128	1.4571	1.4570	1.4707
M= 10	MSE	(0.177E+01)	(0.110E+02)	(0.191E-02)	(0.191E-02)	(0.159E-02)
N= 10	ASYV	(0.751E-01)	(0.751E-01)	(0.258E-04)	(0.258E-04)	(0.322E-03)
S= 3	AE	5.0444	4.8795	1.4997	1.4998	1.4939
M= 20	MSE	(0.100E+00)	(0.488E+00)	(0.699E-05)	(0.724E-05)	(0.178E-02)
N= 20	ASYV	(0.188E-01)	(0.188E-01)	(0.161E-05)	(0.161E-05)	(0.805E-04)
S= 3	AE	5.0013	4.9908	1.5000	1.5000	1.5002
M= 30	MSE	(0.320E-01)	(0.327E-01)	(0.116E-05)	(0.110E-05)	(0.931E-04)
N= 30	ASYV	(0.835E-02)	(0.835E-02)	(0.318E-06)	(0.318E-06)	(0.358E-04)

* 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 $k = 3^*$.

		$A_3 = 0.0$	$B_3 = 0.0$
S= 3	AE	-0.834E+00	-0.892E+00
M,N= 10	VAR	(0.260E-01)	(0.358E-01)
S= 3	AE	-0.109E-01	0.703E-02
M,N= 20	VAR	(0.111E+00)	(0.749E-01)
S= 3	AE	-0.246E-02	0.118E-01
M,N= 30	VAR	(0.376E-01)	(0.360E-01)

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

of the estimates. The biases of the linear parameters are much more than the non-linear parameters as 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 data analysis for illustrative purpose. One is the original color texture data which has been discussed in the introduction and it is of size 50×50 pixels, and the other one is the synthesized texture data.

Color Texture Data:

We model the color texture in Figure 5(a), using the model in (1). The number of components p is unknown here. We calculate the values of the periodogram function, $I(\lambda, \mu, \nu)$, where

$$I(\lambda, \mu, \nu) = \left| \frac{1}{MNS} \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{s=1}^{S} y(m, n, s) e^{-i(m\lambda + n\mu + s\nu)} \right|^2, \tag{24}$$

at the Fourier frequencies, *i.e.*, at $(\lambda, \mu, \nu) = (\frac{m}{M}\pi, \frac{n}{N}\pi, \frac{s}{S}\pi)$ for $m = 1, \ldots, M-1$, $n = 1, \ldots, N-1$ and $s = 1, \ldots, S-1$. Note that the true frequencies are strictly between 0 and π. Since $S = 3$, we have to compute only an $(M-1) \times (N-1) \times 2$ array of periodogram values. We look for presence of sharp peaks among the calculated periodogram values to get an approximate number of sinusoidal components in the signal. Although a three dimensional matrix cannot be plotted directly, this problem is simplified by noting that we have only two layers along the third dimension. We can plot the two layers, each of size $(M-1) \times (N-1)$ and look at them simultaneously for presence of sharp peaks. We have plotted the two layers periodogram of the original texture for $\nu = \frac{\pi}{3}$ and $\nu = \frac{2\pi}{3}$ in Figure 2 and Figure 3, respectively. Here the second figure does not show many peaks, that is, there are not many

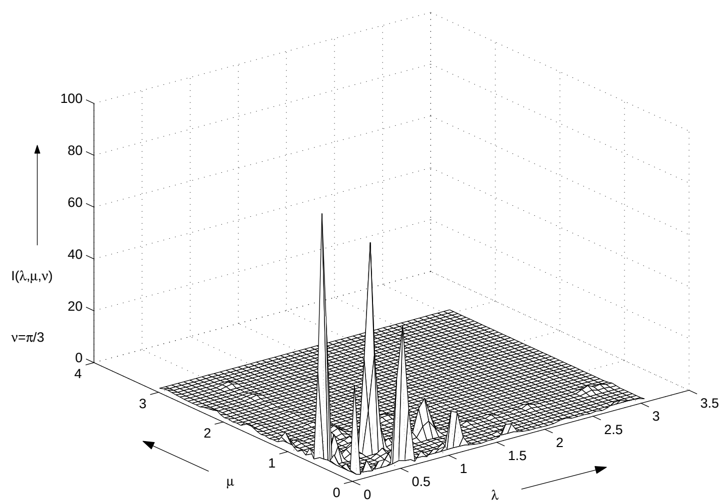


Figure 2: Periodogram of Original Texture (Layer 1)

components for which $\nu = \frac{2\pi}{3}$. In the first layer of the periodogram, there are not only several dominant peaks but also several other smaller peaks. Therefore, it is very difficult to decide the number of components from the peaks of the periodogram values. It would be better to look at some other criteria for choosing the number of components. We have fitted the model sequentially for $k = 1, \ldots, 200$ and use the BIC to estimate p. In this case under the assumptions of *i.i.d.* errors (due to the computational difficulties involved) the BIC takes the following form;

$$BIC(k) = (MNS) \ln \widehat{\sigma_k^2} + \frac{1}{2}(5k+1)\ln(MNS), \tag{25}$$

where $\widehat{\sigma^2}_k$ is the innovations variance, when the number of components is k. In this case the number of parameters to be estimated is $5k+1$. We plot the $BIC(k)$ as a function of k in Figure 4. It is observed that for $k = 66$, $BIC(k)$ gives the minimum value, therefore in this case the estimate of p, say $\widehat{p} = 66$. We have fitted the model (1) with $p = 66$ to the texture data. We estimate the parameters sequentially as proposed in Section 2. The initial estimates are the periodogram estimates obtained by maximizing the periodogram function in (24). The original texture, initial estimated texture obtained by maximizing periodogram and the final estimated textures obtained by the method proposed in Section 2, are plotted in Figure 5. The final estimated texture is observed to match the original texture reasonably well.

It may be noticed that it was possible to fit such a large order model, because the components are estimated sequentially, otherwise it would have been a much difficult task to estimate all the parameters simultaneously.

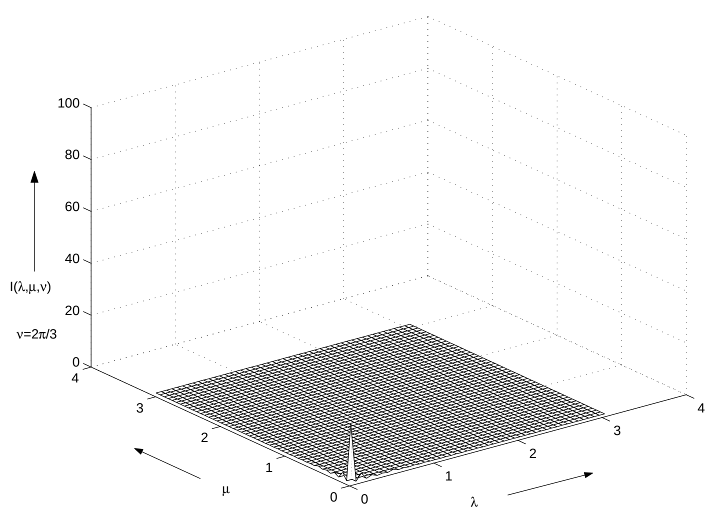


Figure 3: Periodogram of Original Texture (Layer 2)

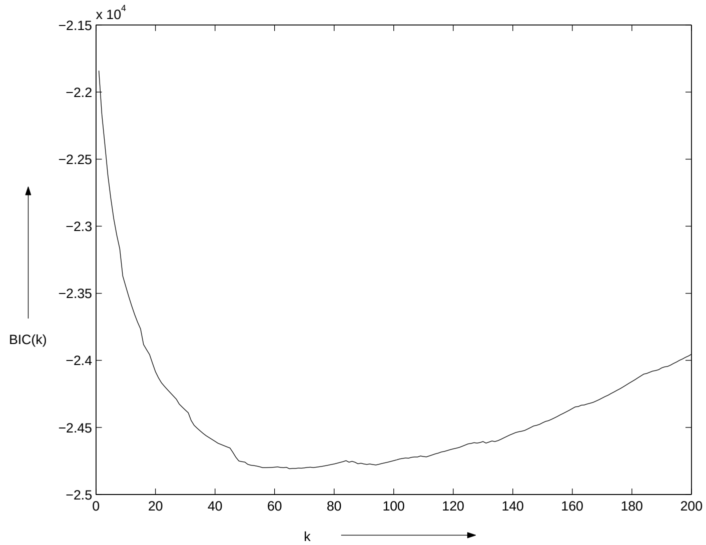
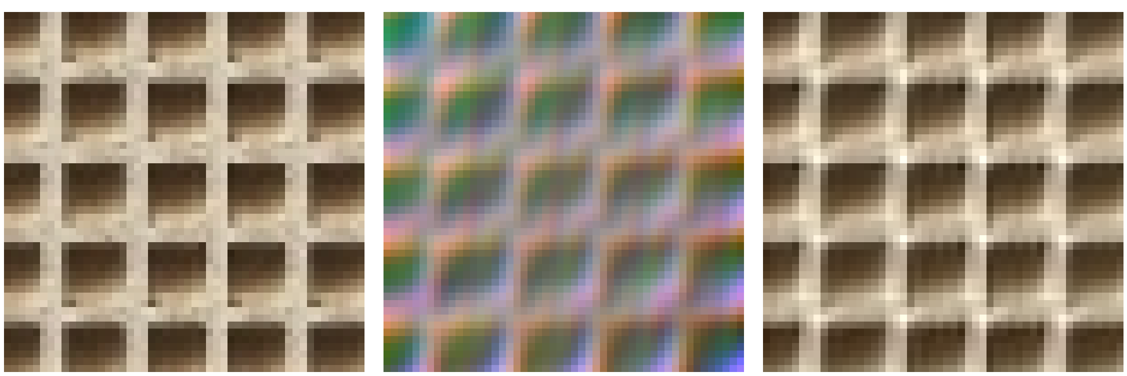


Figure 4: BIC plot for the Original Texture

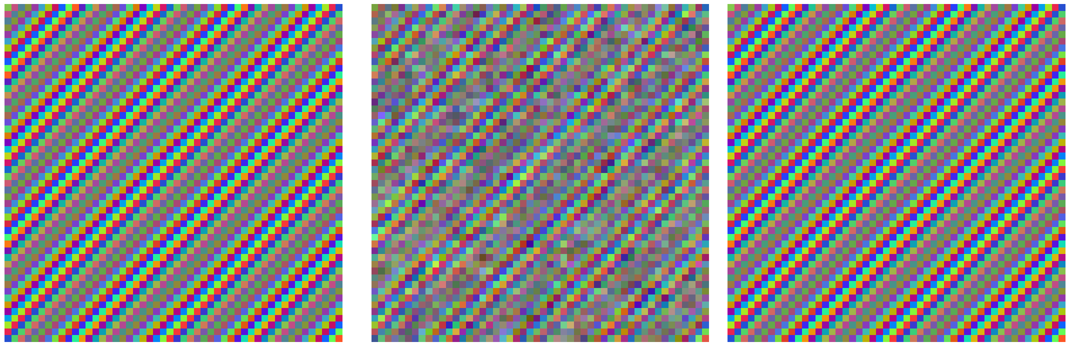


(a) Original Texture

(b) Initial Estimated Texture

(c) Final Estimated Texture

Figure 5: Original, Initial Estimated and Final Estimated Textures



(a) Actual Texture (b) Noisy Texture (c) Estimated Texture

Figure 6: Synthesized Textures

Synthesized Data: Now we analyze a synthesized data obtained from the model in (1). We have take the parameter values as;

$$\begin{array}{lllll} A_1^0 = 10.0, & B_1^0 = 10.0, & \lambda_1^0 = 2.0, & \mu_1^0 = 2.0, & \nu_1^0 = 2.0 \\ A_2^0 = 5.0, & B_2^0 = 5.0, & \lambda_2^0 = 1.5, & \mu_2^0 = 1.5, & \nu_2^0 = 1.5, \end{array}$$

and $M = N = 50$. The error structure is the following;

$$X(m,n,s) = e(m,n,s) + 0.33e(m-1,n,s) + 0.33e(m,n-1,s) + 0.33e(m,n,s-1)$$

Here $e(m,n,s)$s are *i.i.d.* normal random variables with mean 0 and variance 10. The noisy color texture is plotted in Figure 6(b), when the original texture (without the noisy component $X(m,n)$), is plotted in Figure 6(a). Our problem is to extract the original texture Figure 6(a) from the noisy texture Figure 6(b).

We have used the sequential procedure to estimate the unknown parameters and obtained the following estimates

$$\begin{array}{lllll} \widehat{A}_1 = 10.0307, & \widehat{B}_1 = 10.1660, & \widehat{\lambda}_1 = 2.0004, & \widehat{\mu}_1 = 2.0002, & \widehat{\nu}_1 = 1.9933 \\ \widehat{A}_2 = 5.1458, & \widehat{B}_2 = 4.7853, & \widehat{\lambda}_2 = 1.5001, & \widehat{\mu}_2 = 1.4976, & \widehat{\nu}_2 = 1.5125. \end{array}$$

Based on the above estimates, the estimated texture is plotted in Figure 6(c). Figures 6(a) and 6(c) match quite well.

7 Conclusions

In this paper we have provided a sequential estimation procedure to estimate the unknown parameters of the 3-D sinusoidal model. 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 difficult problem, particularly when the number of components is very high. For example, if $p = 66$, as we have observed in the color texture data, then to find the least squares estimators, we need to use a 66×3 dimensional optimization procedure, which is not very easy to implement. On the other hand in our proposed sequential procedure, we need to solve 66, 3-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. Since at each stage we need to use only a three-dimensional optimization procedure, therefore our method is very easy to implement and it is observed that its performance is also quite satisfactory.

Acknowledgements

The authors would like to thank two referees for their helpful suggestions which had improved the earlier version of the paper. Part of this work of the second author has been supported by a grant from the Department of Science and Technology, Government of India.

Appendix 1:

Expressions of $\widehat{A}(\lambda,\mu,\nu)$, $\widehat{B}(\lambda,\mu,\nu)$

Consider the following $MNS \times 2$ matrix

$$Z = \begin{bmatrix} \cos(\lambda+\mu+\nu) & \sin(\lambda+\mu+\nu) \\ \cos(\lambda+\mu+2\nu) & \sin(\lambda+\mu+2\nu) \\ \vdots & \vdots \\ \cos(\lambda+\mu+S\nu) & \sin(\lambda+\mu+S\nu) \\ \cos(\lambda+2\mu+\nu) & \sin(\lambda+2\mu+\nu) \\ \cos(\lambda+2\mu+2\nu) & \sin(\lambda+2\mu+2\nu) \\ \vdots & \vdots \\ \cos(\lambda+2\mu+S\nu) & \sin(\lambda+2\mu+S\nu) \\ \vdots & \vdots \\ \cos(\lambda+N\mu+\nu) & \sin(\lambda+N\mu+\nu) \\ \cos(\lambda+N\mu+2\nu) & \sin(\lambda+N\mu+2\nu) \\ \vdots & \vdots \\ \cos(\lambda+N\mu+S\nu) & \sin(\lambda+N\mu+S\nu) \\ \vdots & \vdots \\ \cos(M\lambda+N\mu+S\nu) & \sin(M\lambda+N\mu+S\nu) \end{bmatrix}$$

and the $MNS \times 1$ vector of observations

$Y = [y(1,1,1), y(1,1,2), \ldots, y(1,1,S), y(1,2,1), y(1,2,2), \ldots, y(1,2,S), y(1,N,1), y(1,N,2),$
$\ldots, y(1,N,S), \ldots, y(M,N,S)]'$.

Then the estimates are given by,

$$\begin{bmatrix} \widehat{A}(\lambda,\mu,\nu) \\ \widehat{B}(\lambda,\mu,\nu) \end{bmatrix} = (Z'Z)^{-1}Z'Y.$$

Appendix 2:

Proof of Lemma 1

$$\sup_{\alpha,\beta,\gamma} \left| \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{s=1}^{S} X(m,n,s) e^{i(m\alpha+n\beta+s\gamma)} \right|$$

$$= \sup_{\alpha,\beta,\gamma} \left| \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} X(m,n,1) e^{i(m\alpha+n\beta)} \cdot e^{i\gamma} + X(m,n,2) e^{i(m\alpha+n\beta)} \cdot e^{i2\gamma} + \ldots \right.$$
$$\left. \ldots + X(m,n,S) e^{i(m\alpha+n\beta)} \cdot e^{iS\gamma} \right|$$
$$\leq \sup_{\alpha,\beta} \left| \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} X(m,n,1) e^{i(m\alpha+n\beta)} \right| + \sup_{\alpha,\beta} \left| \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} X(m,n,2) e^{i(m\alpha+n\beta)} \right|$$
$$\ldots \sup_{\alpha,\beta} \left| \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} X(m,n,S) e^{i(m\alpha+n\beta)} \right|.$$

Since each term goes to zero $a.s.$, see Kundu and Nandi [9], the result follows. ∎

Appendix 3:

Proof of (11)

Let us denote the dummy variable θ by θ_1. It enough to prove (11) for $p = 2$.

$$f(\theta_1) = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{s=1}^{S} \{A_1^0 \cos(m\lambda_1^0 + n\mu_1^0 + s\nu_1^0) + B_1^0 \sin(m\lambda_1^0 + n\mu_1^0 + s\nu_1^0)$$
$$- A_1 \cos(m\lambda_1 + n\mu_1 + s\nu_1) - B_1 \sin(m\lambda_1 + n\mu_1 + s\nu_1)\}\}^2$$
$$+ \frac{2}{MNS} \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{s=1}^{S} \{A_1^0 \cos(m\lambda_1^0 + n\mu_1^0 + s\nu_1^0) + B_1^0 \sin(m\lambda_1^0 + n\mu_1^0 + s\nu_1^0)$$
$$- A_1 \cos(m\lambda_1 + n\mu_1 + s\nu_1) - B_1 \sin(m\lambda_1 + n\mu_1 + s\nu_1)\}$$
$$\times \{A_2^0 \cos(m\lambda_2^0 + n\mu_2^0 + s\nu_2^0) + B_2^0 \sin(m\lambda_2^0 + n\mu_2^0 + s\nu_2^0)\}$$

We want to prove that

$$\liminf \inf_{\theta_1 \in S_c} f(\theta_1) > 0 \text{ a.s.} \tag{26}$$

Proving the above is equivalent to proving it separately with S_c replaced by A_c, B_c, Λ_c, M_c and N_c where

$$A_c = \{\theta : \theta = (A_1, B_1, \lambda_1, \mu_1, \nu_1); |A_1 - A_1^0| \geq c\}$$

$$B_c = \{\theta : \theta = (A_1, B_1, \lambda_1, \mu_1, \nu_1); |B_1 - B_1^0| \geq c\}$$

$$\Lambda_c = \{\theta : \theta = (A_1, B_1, \lambda_1, \mu_1, \nu_1); |\lambda_1 - \lambda_1^0| \geq c\}$$

$$M_c = \{\theta : \theta = (A_1, B_1, \lambda_1, \mu_1, \nu_1); |\mu_1 - \mu_1^0| \geq c\}$$

$$N_c = \{\theta : \theta = (A_1, B_1, \lambda_1, \mu_1, \nu_1); |\nu_1 - \nu_1^0| \geq c\}$$

The set A_c can be subdivided into the following mutually exclusive and exhaustive sets:

$$\Omega_A^1 = \{\theta = (A_1, B_1, \lambda_1, \mu_1, \nu_1); \lambda_1 = \lambda_1^0, \mu_1 = \mu_1^0, \nu_1 = \nu_1^0, |A_1 - A_1^0| \geq c\}$$

$$\Omega_A^2 = \{\theta = (A_1, B_1, \lambda_1, \mu_1, \nu_1); \lambda_1 = \lambda_1^0, \mu_1 = \mu_1^0, \nu_1 \neq \nu_1^0, |A_1 - A_1^0| \geq c\}$$

$$\Omega_A^3 = \{\theta = (A_1, B_1, \lambda_1, \mu_1, \nu_1); \lambda_1 = \lambda_1^0, \mu_1 \neq \mu_1^0, |A_1 - A_1^0| \geq c\}$$

$$\Omega_A^4 = \{\theta = (A_1, B_1, \lambda_1, \mu_1, \nu_1); \lambda_1 = \lambda_2^0, \mu_1 = \mu_2^0, \nu_1 = \nu_2^0, A_1 = A_2^0, B_1 = B_2^0, |A_1 - A_1^0| \geq c\}$$

$$\Omega_A^5 = \{\theta = (A_1, B_1, \lambda_1, \mu_1, \nu_1); \lambda_1 = \lambda_2^0, \mu_1 = \mu_2^0, \nu_1 = \nu_2^0, A_1 = A_2^0, B_1 \neq B_2^0, |A_1 - A_1^0| \geq c\}$$

$$\Omega_A^6 = \{\theta = (A_1, B_1, \lambda_1, \mu_1, \nu_1); \lambda_1 = \lambda_2^0, \mu_1 = \mu_2^0, \nu_1 = \nu_2^0, A_1 \neq A_2^0, B_1 = B_2^0, |A_1 - A_1^0| \geq c\}$$

$$\Omega_A^7 = \{\theta = (A_1, B_1, \lambda_1, \mu_1, \nu_1); \lambda_1 = \lambda_2^0, \mu_1 = \mu_2^0, \nu_1 = \nu_2^0, A_1 \neq A_2^0, B_1 \neq B_2^0, |A_1 - A_1^0| \geq c\}$$

$$\Omega_A^8 = \{\theta = (A_1, B_1, \lambda_1, \mu_1, \nu_1); \lambda_1 = \lambda_2^0, \mu_1 = \mu_2^0, \nu_1 \neq \nu_2^0, |A_1 - A_1^0| \geq c\}$$

$$\Omega_A^9 = \{\theta = (A_1, B_1, \lambda_1, \mu_1, \nu_1); \lambda_1 = \lambda_2^0, \mu_1 \neq \mu_2^0, |A_1 - A_1^0| \geq c\}$$

$$\Omega_A^{10} = \{\theta = (A_1, B_1, \lambda_1, \mu_1, \nu_1); \lambda_1 \neq \lambda_1^0, \lambda_1 \neq \lambda_2^0, |A_1 - A_1^0| \geq c\}$$

For $\theta \in \Omega_A^1$, it can be shown that, $f(\theta_1) \longrightarrow \frac{1}{2}[(A_1^0 - A_1)^2 + (B_1^0 - B_1)^2] > 0$.

For $\theta_1 \in \Omega_A^2, \Omega_A^3, \Omega_A^8, \Omega_A^9$ or Ω_A^{10}, $f(\theta_1) \longrightarrow \frac{1}{2}[({A_1^0}^2 + {B_1^0}^2) + (A_1^2 + B_1^2)] > 0$.

For $\theta \in \Omega_A^4$, $f(\theta_1) \longrightarrow \frac{1}{2}[({A_1^0}^2 + {B_1^0}^2) - ({A_2^0}^2 + {B_2^0}^2)] > 0$.

For $\theta \in \Omega_A^5$, $f(\theta_1) \longrightarrow \frac{1}{2}[(B_2^0 - B_1)^2] > 0$.

For $\theta \in \Omega_A^6$, $f(\theta_1) \longrightarrow \frac{1}{2}[(A_2^0 - A_1)^2] > 0$.

For $\theta \in \Omega_A^7$, $f(\theta_1) \longrightarrow \frac{1}{2}[(B_2^0 - B_1)^2 + (A_2^0 - A_1)^2] > 0$.

See Appendix 3.C of Prasad [11] for details. Note that the main idea about dividing the set A_c is to consider all possibilities of the rest of the parameters when $|A_1 - A_1^0| \geq c$. Similarly for the other sets also.

The set B_c, Λ_c, M_c and N_c can be similarly broken down into smaller subsets as A_c above and it can be shown separately for each of the above subsets that $f(\theta_1) > 0$ as $M, N \to \infty$.

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] 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.

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

[4] Hannan, E. J. (1971), "Non-linear time series regression", *Jour. Appl. Prob.*, Vol. 8, 767-780.

[5] Kay, S.M. (1988), *Modern Spectral Estimation: Theory and Applications*, Prentice Hall,

[6] Kundu, D. (1997), "Asymptotic properties of the least squares estimators of sinusoidal signals", *Statistics*, 30, 221-238.

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

[8] Kundu D. and Gupta R. D. (1998) , "Asymptotic properties of the least squares estimators of a two dimensional model", *Metrika*, vol. 48, 8397.

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

[10] 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.

[11] Prasad, A. (2008), *Some non-linear regression models and their applications in statistical signal processing*, Ph. D. thesis, Indian Institute of Technology Kanpur, *http* : //172.28.64.70 : 8080/*jspui*/*handle*/123456789/10707

[12] 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.

[13] 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.

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

[15] Walker, A.M. (1971), "On the estimation of a Harmonic components in a time series with stationary residuals", *Biometrika*, 58, 21-36.

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

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