

Estimating the Orientation of Planar Surfaces: Algorithms and Bounds

Haim Permuter and Joseph M. Francos, *Senior Member, IEEE*

Abstract—This paper presents a computationally and statistically efficient parametric solution to the problem of estimating the orientation in space of a planar textured surface from a single, noisy, observed image of it. The coordinate transformation from surface to image coordinates, due to the perspective projection, transforms each homogeneous sinusoidal component of the surface texture into a sinusoid whose frequency is a function of location. The functional dependence of the sinusoid phase in location is uniquely determined by the tilt and slant angles of the surface. From the physical model of the perspective projection, we derive the Cramér–Rao lower bound on the error variance of estimating the tilt and slant of the observed surface in the presence of observation noise. It is shown in this paper that the phase of each of the sinusoids can be expressed as a linear function of some variables that are related to the surface tilt and slant angles. Using the Phase Differencing Algorithm, we fit a polynomial phase model to a sinusoidal component of the observed texture. Substituting in the derived linear relation, the unknown phase with the one estimated using the Phase Differencing Algorithm, we obtain a closed-form, analytic, and computationally efficient solution to the problem of estimating the tilt and slant angles. The algorithm performance is shown to be close to the Cramér–Rao bound, even for low signal-to-noise ratios, at computational complexity which is considerably lower than that of any existing algorithm.

Index Terms—Cramér–Rao bound, inhomogeneous two-dimensional signals, parametric texture modeling, perspective estimation, two-dimensional polynomial phase models.

I. INTRODUCTION

PERSPECTIVE projection has a dominant and fundamental role in any imaging process, whether by the human visual system or some type of a camera. Hence, perspective is one of the prominent clues in image interpretation and understanding. This makes perspective estimation a key problem in many image modeling and analysis applications. In this paper, we address a special case of this general problem: we consider the problem of estimating the orientation in space of a planar textured surface, from a single observed image of it.

A solution to this problem is an essential component in many image processing and multimedia data processing applications.

Manuscript received August 31, 1999; revised May 3, 2000. This work was supported in part by the Israel Ministry of Science, the Eshkol Fellowship Program in Applied Mathematics under Grant 0616196, in part by the Israel Ministry of Science and the French Ministry of Research and Technology under Grant 8814297, and in part by the Israel Ministry of Science under Grant 8635297.

The authors are with the Department of Electrical and Computer Engineering, Ben-Gurion University, Beer-Sheva 84105, Israel (e-mail: haimp@ee.bgu.ac.il; francos@ee.bgu.ac.il).

Communicated by J. A. O'Sullivan, Guest Editor.
 Publisher Item Identifier S 0018-9448(00)06071-5.

For example, the segmentation of two- (2-D) or three-dimensional (3-D) images and video for content-based coding and representation is considerably simplified if the effects of the perspective projection are eliminated first, thus reducing the inhomogeneity of the image. By estimating and then canceling the effect of the perspective projection on a *given image*, we avoid the difficulty of segmenting and coding an image where each of its patches is inhomogeneous. Thus the original problem is replaced by the simpler problem of segmenting and coding an image where large patches are homogeneous. Furthermore, multifunctional coding of visual information is a desirable feature to provide not only an efficient representation of the information itself, but also to enable additional multimedia functionalities, such as content-based indexing for retrieval from multimedia databases. Clearly, to enable such indexing, the effect of the specific perspective projection in each image has to be nulled in order to “normalize” all images with respect to some “common basis.” In particular, in indexing and retrieval systems of multimedia data that employ the textural information in the imagery components of the data, e.g., [12], the identification of similar textured surfaces as being such is impossible unless the effects of the different perspective projections involved in the process of creating each of the images are estimated and then removed. We, therefore, conclude that all the foregoing applications require an accurate estimate of the perspective transformation to become available at a moderate computational complexity, so that perspective estimation could be conveniently integrated into the higher level applications.

Existing solutions to problems where perspective estimation is involved attempt to extract the projection parameters based on the observed variations in the image generated by the perspective projection. Structure-based approaches attempt to recognize the structure of the surface texture (the “true” texture) from the observed projected image of that surface. In other words, in order to estimate the projection parameters, these methods must first (or jointly as proposed in [5]) obtain the characterizing properties of the surface texture such as regularity, periodicity, symmetry, collinearity, etc. (see, e.g., [6] and the references therein). This task is very difficult in general, and more so in the presence of noise, since due to the projection distortion and the noise contribution the observed surface does not exhibit the expected “regularities,” and texture elements (“texels”) are difficult to identify [5]. An alternative approach, that does not require the initial recognition of the structure of the surface texture is based on statistical assumptions regarding the distribution of textural properties of the surface texture. Typical assumptions are isotropy, [1], [2], [8], [13] and homogeneity, [3], [6], [14]. Thus an observed preferred orientation of an isotropic surface

texture, or an observed inhomogeneity of a homogeneous surface texture, is employed to estimate the surface orientation. We note that many of the foregoing methods (see, e.g., [6], [11], [3], [4]) are derived for binary images, or are based on an initial local analysis of the image, using its edge information. The algorithm proposed in [14] evaluates the dominant frequency at each image point using the wavelet transform, and then employs the spatial dependence of this frequency component to estimate the surface orientation.

Generally speaking, all these algorithms consider the estimation of the *instantaneous frequency* at every image point (or a related quantity) as a first step in a procedure for estimating the tilt and slant of the observed surface from the variations of the instantaneous frequency. This approach is further pursued with the introduction of novel space–frequency methods, see, e.g., [14], [7]. However, given an inhomogeneous signal $d(x_i, y_i)$, the question of the unique determination of its instantaneous phase, frequency, and amplitude is not a straightforward one, as even in the one-dimensional case, incoherent definitions of instantaneous phase and amplitude are common. We refer the interested reader to [19] for a detailed discussion on the one-dimensional problem. Following [19], it is clear that starting from a given signal $d(x_i, y_i)$, it is possible to introduce an infinite number of pairs $[a(x_i, y_i), \Phi(x_i, y_i)]$ such that

$$d(x_i, y_i) = a(x_i, y_i) \cos(\Phi(x_i, y_i)). \quad (1)$$

Nevertheless, in order to be able to interpret $a(x_i, y_i)$ as the instantaneous amplitude of the signal and $\Phi(x_i, y_i)$ as its instantaneous phase, the instantaneous phase and amplitude should be defined in such a way that only a *single*, well-defined, pair $[a(x_i, y_i), \Phi(x_i, y_i)]$ will correspond to any given signal $d(x_i, y_i)$, so that the representation (1) is unique. We shall further elaborate on this point in Section II.

A maximum-likelihood (ML) estimator for the tilt and slant parameters is proposed in [9]. In this framework, the homogeneous surface texture is modeled by a Gauss–Markov random field. A probability distribution function for the observed textured image, assuming a *linear* projection model (instead of the nonlinear perspective projection transformation) is derived. The joint problem of estimating the surface orientation parameters and the texture model is then solved by a ML estimator.

In this paper, we elaborate on the problem of estimating the orientation in space of a planar textured surface, such that, in its own coordinate system the surface texture is *homogeneous*. In particular, we derive universal performance bounds on the accuracy of estimating the tilt and slant parameters in the presence of observation noise, and we propose computationally and statistically efficient estimators for these parameters in the presence of noise. To the best of our knowledge, the derivation of universal performance bounds for this problem and the derivation of estimation algorithms in the presence of noise have never been considered. Furthermore, since the model of the homogeneous surface texture is based on the 2-D Wold decomposition of homogeneous random fields, the proposed algorithms provide a unifying framework for both the structural and statistical methods. In addition, the extremely difficult task of identifying the texture elements from the perspective projected noisy image, as required by the structure-based approaches, is avoided.

The 2-D Wold decomposition implies that the deterministic component of any homogeneous texture field can be approximated by a sum of 2-D sinusoids, [20]. Thus an approximate model of the surface texture deterministic component is given by

$$t(x_s, y_s) = \sum_{l=1}^L A_l \cos(x_s u_l + y_s v_l + \varphi_l) \quad (2)$$

where (x_s, y_s) denote the surface coordinates. The coordinate transformation from surface to image coordinates, due to the perspective projection, transforms each homogeneous sinusoidal component to a sinusoid whose frequency is a function of location. In the case of a planar surface, the functional dependence of the sinusoid phase in location is uniquely determined by the tilt and slant angles of the surface [15]. Hence, the surface tilt and slant angles can be recovered from the phase of the sinusoidal component measured in the image plane.

Denote the image coordinates by (x_i, y_i) . We use the slant–tilt system for representing the orientation of the planar surface (see Fig. 1). The slant σ is the angle between the surface normal and the optical axis z_w . The tilt τ is the angle between the x_i -axis and the projection of the surface normal onto the image plane. Substituting into the texture model (2), the inverse coordinate transformation expression which expresses (x_s, y_s) in terms of the image coordinates (x_i, y_i) , the tilt, the slant, and some known parameters of the camera, we obtain a model of the texture deterministic component, projected onto the image plane. (See Appendix A for the derivation.) In Appendix A it is also shown that under the perspective transformation, the phase of a sinusoidal component whose phase function is given in surface coordinates by $\Phi_s(x_s, y_s) = u x_s + v y_s + \varphi$ becomes

$$\Phi(x_i, y_i) = \frac{\frac{x_i}{f} (\tilde{u} \cos \tau - \tilde{v} \cos \sigma \sin \tau) + \frac{y_i}{f} (\tilde{u} \sin \tau + \tilde{v} \cos \sigma \cos \tau)}{\tan \sigma \left(\frac{x_i}{f} \cos \tau + \frac{y_i}{f} \sin \tau \right) + 1} + \varphi \quad (3)$$

in the coordinate system of the observed image, where $\tilde{u} = u z_0$, $\tilde{v} = v z_0$, z_0 is the z_w -coordinate where the surface crosses the optical axis, and f is the focal length.

From the physical model of the perspective projection, we derive the Cramér–Rao lower bound on the error variance in estimating the tilt and slant of the observed surface. Two computationally efficient algorithms for estimating the tilt and slant angles from the estimated phase of a sinusoidal component of the surface texture are derived.

The perspective projection results in a continuous coordinate transformation from the surface coordinate system to the image coordinate system. Hence, the phase function of each sinusoidal component of the surface texture is transformed by the perspective projection into a nonlinear, continuous function of the image coordinates. Since continuous functions can be approximated by polynomials, a natural choice for modeling the continuous phase function of each sinusoidal component is by a polynomial function of the image coordinates. In [15] we derive an algorithm for estimating the slant and the tilt of the planar surface directly from the estimated parameters of a polynomial model of the phase. The estimated polynomial phase model is

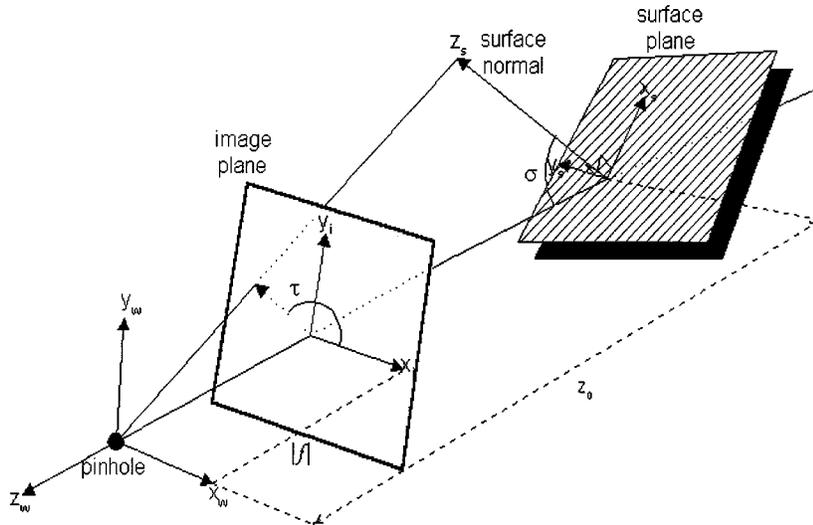


Fig. 1. The perspective projection.

obtained using the Phase Differencing (PD) algorithm, [16], [17]. In order to obtain the estimates of the tilt and slant angles, it is assumed that the estimated polynomial phase coefficients are, in fact, the coefficients of a Taylor series expansion of the phase. However, the algorithm, though computationally efficient, has a relatively high bias and error variance. To improve accuracy, an iterative refinement procedure is applied in a small neighborhood of the previously obtained tilt and slant estimates. The combined two-stage algorithm produces estimates of lower bias and standard deviation, at the cost of higher computational complexity. The algorithms proposed in this paper employ the PD algorithm to estimate the *phase* of a sinusoidal component of the observed surface texture. It is shown that the phase of each of the sinusoids can be expressed as a *linear* function of some variables that are related, in a rather simple form, to the surface tilt and slant angles. Hence, the tilt and slant can be estimated by solving a linear least squares problem. The performance of the two algorithms proposed in this paper is shown to be close to the Cramér–Rao bound, at a computational complexity which is considerably lower than that of any existing algorithm.

The paper is organized as follows. In Section II, the 2-D polynomial phase model is introduced and the considerations involved in applying the PD algorithm for estimating the model are described. In Section III, it is shown that the phase of each of the sinusoids observed in the image plane can be expressed as a linear function of some variables that are related in a rather simple form to the surface tilt and slant angles. This functional dependence is the basis of two algorithms presented in Section III for estimating the surface tilt and slant. In Section IV, we derive the Cramér–Rao lower bound (CRB) on the error variance in estimating the tilt and slant of the observed surface. The bound is derived directly from the physical model of the perspective projection. The performance of the proposed algorithms is illustrated using synthetic and photographed images in Section V. In particular, we investigate the performance of the algorithms in the presence of noise and analyze their performance through Monte Carlo simulations and by comparing the Monte Carlo results with the CRB.

II. THE PARAMETRIC PHASE MODEL AND ITS ESTIMATION

In Section I it is shown that the phase function of any sinusoidal component of the homogeneous surface texture is transformed by the perspective projection into a nonlinear function of the image coordinates. For a given focal length, the transformation is a unique function of the surface tilt and slant angles [15]. Hence, in principle, the surface tilt and slant can be recovered from the phase of the projected sinusoidal component. However, due to its 2π periodicity the phase wraps around, and only its principle value is observable. Therefore, any use of the phase information is limited by the need to unwrap the phase of the observed signal first.

In this paper, we propose to use a parametric model as an alternative to the need to employ phase unwrapping methods. Since continuous functions can be approximated by polynomials, a natural choice for modeling any *continuous* 2-D phase function is by a 2-D polynomial of the coordinates. Since the assumption of phase smoothness is implicit to this model, no *explicit* phase unwrapping is required in estimating the observed phase. In this section we briefly study the model of a single-component constant-amplitude exponential of a polynomial function of the field coordinates. The model, described in this section, as well as the properties of the parametric phase estimation algorithm are studied in detail in [16] and [17]. The proposed phase estimation algorithm is suboptimal (relative to the maximum-likelihood (ML) estimator) but computationally efficient (since no multidimensional search in the parameter space is required). The algorithm is based on the properties of a 2-D phase difference operator.

Let us first define the type of signal for which this operator was designed. Let $\{t(x_i, y_i)\}$ be a discrete 2-D constant amplitude polynomial phase signal, i.e.,

$$t(x_i, y_i) = A \exp\{j\phi_{Q+1}(x_i, y_i)\}, \\ x_i = 0, 1, \dots, N-1, \quad y_i = 0, 1, \dots, M-1 \quad (4)$$

where

$$\phi_{Q+1}(x_i, y_i) = \sum_{\{0 \leq k, \ell: 0 \leq k+\ell \leq Q+1\}} c(k, \ell) x_i^k y_i^\ell. \quad (5)$$

We call $\phi_{Q+1}(x_i, y_i)$ 2-D polynomial of *total degree* $Q + 1$. The amplitude A is a real-valued positive constant. To simplify the presentation, we assume there is no observation noise and $A \equiv 1$. Hence $t(x_i, y_i) = \exp\{j\phi_{Q+1}(x_i, y_i)\}$.

Next we define the basic phase differencing operators.

Definition 1: Let τ_y and τ_x be some positive constants. Define

$$\begin{aligned} \text{PD}_{y^{(q)}}[t(x_i, y_i)] &= t(x_i, y_i), \\ x_i &= 0, 1, \dots, N-1, \quad y_i = 0, 1, \dots, M-1 \end{aligned} \quad (6)$$

and, in general,

$$\begin{aligned} \text{PD}_{y^{(q)}}[t(x_i, y_i)] &= \text{PD}_{y^{(q-1)}}[t(x_i, y_i)] \\ &\quad \times (\text{PD}_{y^{(q-1)}}[t(x_i, y_i + \tau_y)])^* \end{aligned} \quad (7)$$

where the resulting 2-D signal $\text{PD}_{y^{(q)}}[t(x_i, y_i)]$ exists for $x_i = 0, 1, \dots, N-1$, $y_i = 0, 1, \dots, M-1 - q\tau_y$. The phase differencing operator along the x_i -axis, $\text{PD}_{x^{(p)}}[t(x_i, y_i)]$ is defined in a similar way.

Assume we have sequentially applied P times the phase difference operator $\text{PD}_{x^{(1)}}$, and $Q - P$ times the phase difference operator $\text{PD}_{y^{(1)}}$, to some complex-valued 2-D signal $t(x_i, y_i)$. We will denote the resulting signal by $\text{PD}_{x^{(P)}, y^{(Q-P)}}[t(x_i, y_i)]$.

Theorem 1: Let $t(x_i, y_i)$ be given by (4) and (5). Then, the signal $\text{PD}_{x^{(P)}, y^{(Q-P)}}[t(x_i, y_i)]$ is a 2-D exponential given by

$$\begin{aligned} \text{PD}_{x^{(P)}, y^{(Q-P)}}[t(x_i, y_i)] &= \exp\{j[\omega_Q x_i + \nu_Q y_i + \gamma_Q(\tau_x, \tau_y)]\}, \\ x_i &= 0, 1, \dots, N-1 - P\tau_x, \\ y_i &= 0, 1, \dots, M-1 - (Q-P)\tau_y \end{aligned} \quad (8)$$

where

$$\omega_Q = (-1)^Q c(P+1, Q-P)(P+1)!(Q-P)! \tau_x^P \tau_y^{Q-P} \quad (9)$$

$$\nu_Q = (-1)^Q c(P, Q+1-P)P!(Q+1-P)! \tau_x^P \tau_y^{Q-P} \quad (10)$$

and $\gamma_Q(\tau_x, \tau_y)$ is not a function of x_i nor y_i .

Theorem 1 implies that applying in some arbitrary sequence P times the operator $\text{PD}_{x^{(1)}}$ and $Q - P$ times the operator $\text{PD}_{y^{(1)}}$ to the observed signal (4), the resulting signal is the 2-D exponential

$$\text{PD}_{x^{(P)}, y^{(Q-P)}}[t(x_i, y_i)] = \exp\{j[\omega_Q x_i + \nu_Q y_i + \gamma_Q(\tau_x, \tau_y)]\}$$

where ω_Q and ν_Q are given by (9) and (10), respectively. Hence, estimating (ω_Q, ν_Q) using any standard frequency estimation technique, results in estimates of $c(P+1, Q-P)$ and $c(P, Q+1-P)$. In this paper, we estimate the frequency of the exponential using a search for the maximum of the absolute value of the signal's 2-D discrete Fourier transform (2-D DFT). Repeating the procedure which was described above assuming

some arbitrary P , for all P such that $0 \leq P \leq Q$, we obtain estimates of all the parameters of the highest order layer, $Q + 1$, of the phase model. Multiplying $t(x_i, y_i)$ by

$$\exp\left\{-j \sum_{k=0}^{Q+1} \hat{c}(k, Q+1-k) y_i^{Q+1-k} x_i^k\right\}$$

results in a new polynomial phase signal whose total degree is Q . By applying to the resulting signal, a procedure similar to the one used to estimate the parameters $c(k, \ell)$ for $k + \ell = Q + 1$, we obtain an estimate of the $Q + 1$ parameters in the Q "layer."

In general, let $t^{(q+1)}(x_i, y_i)$ denote the 2-D signal, where $q + 1$ denotes the *current* total degree of its phase polynomial. By repeating for all $q = Q, \dots, 0$, the two basic steps of estimating the $c(k, \ell)$ parameters of "layer" $q + 1$ through finding the maxima of

$$|\text{DFT}(\text{PD}_{y^{(q-P)}}[\text{PD}_{x^{(P)}}[t^{(q+1)}(x_i, y_i)])|$$

for all $0 \leq P \leq q$, followed by multiplying the already reduced order 2-D polynomial phase signal by

$$\exp\left\{-j \sum_{k=0}^{q+1} \hat{c}(k, q+1-k) y_i^{q+1-k} x_i^k\right\}$$

in the next step, we obtain estimates for all the phase parameters except $c(0, 0)$. The resulting signal after this processing, $t^{(0)}(x_i, y_i)$, is a constant phase 2-D signal. Taking now the average of the imaginary part of the logarithm of this signal, we obtain an estimate for $c(0, 0)$. We have thus completed the estimation of all the coefficients of the 2-D phase polynomial of total degree $Q + 1$. In the following, we refer to the algorithm as the *Phase Differencing Algorithm* (PD Algorithm).

So far we have described the parameter estimation algorithm for the case in which no observation noise exists. However, in many practical situations the signal is observed in the presence of additive noise. Thus a straightforward but computationally prohibitive alternative to the PD Algorithm is to develop a maximum-likelihood estimator (MLE) for the polynomial phase parameters. It turns out [16], that although the PD Algorithm is suboptimal (relative to the ML algorithm), its performance in the presence of additive white noise, is close to the Cramér-Rao lower bound on the error variance in estimating the parameters of the polynomial phase model, for moderate to high signal-to-noise ratios. Optimal selection rules of the PD algorithm parameters τ_x and τ_y are derived in [22]. These rules are employed in this paper to achieve optimal performance of the phase estimation.

However, the PD Algorithm is designed to work with complex-valued constant-amplitude polynomial phase monocomponent signals. In our application, the 2-D signal is real, and in general it has more than a single component. In [15] we derive an algorithm that isolates a single component from the observed signal and converts it into a complex form through the Hilbert transform, (see, e.g., [18] and the references therein). This procedure produces a 2-D analytic signal of the form

$$z(x_i, y_i) = a(x_i, y_i) \exp\{j\Phi(x_i, y_i)\}$$

such that the instantaneous amplitude and phase of the real-valued component $d(x_i, y_i)$ in (1) are unambiguously defined.

The component selection procedure is based on the results of Section IV on the Cramér–Rao lower bound on the error variance in estimating the tilt and slant of the observed surface. More specifically, it is shown in Section IV that the bounds are nearly linear functions of $\frac{1}{\text{SNR}}$, where SNR denotes the signal-to-noise ratio of the sinusoidal component to which the estimation procedure is applied. It is further shown that the bounds on both the tilt and slant parameters are high when the frequencies of the observed inhomogeneous components are low. The bounds rapidly decrease as the spatial frequencies become higher. Hence, the selection rule selects the highest energy component among the components for which at least four periods are observed in the image. In other words, it may very well be that a higher frequency component will be chosen even if its amplitude is lower than that of a lower frequency component. The selection result is verified using the CRB: if there is more than a single candidate component, the estimation algorithm described in the next sections is applied to each of these components. The estimates obtained based on the alternative choices are substituted into the CRB equations and the estimate that provides the lowest CRB is chosen.

III. LINEAR LEAST SQUARES ESTIMATION OF THE TILT AND SLANT ANGLES

A. Tilt and Slant Estimation Using the Estimated Phase

As was indicated earlier, the perspective transformation involved in the imaging process causes a texture which is homogeneous on the planar surface to appear inhomogeneous in the image plane. Let us consider a single sinusoidal component of the observed surface texture, and assume that the functional dependence of its phase in the image coordinates is available to us. In this section we show that the phase at each image coordinate can be expressed as a *linear* function of some variables that are related in a rather simple form to the required unknown tilt and slant angles. This linear relation enables us to derive a *closed-form, analytic, and computationally efficient* solution to the problem of estimating the tilt and slant angles. Previously derived algorithms, e.g., [5], [9], [14], [15], require a computationally intensive iterative search for the minimum of an objective function.

Define

$$\beta_1 = \frac{(\tilde{u} \cos \tau - \tilde{v} \cos \sigma \sin \tau)}{\cos \sigma} \quad (11)$$

$$\delta_1 = \frac{(\tilde{u} \sin \tau + \tilde{v} \cos \sigma \cos \tau)}{\cos \sigma} \quad (12)$$

$$l_1 = \tan \sigma \cos \tau \quad (13)$$

$$l_2 = \tan \sigma \sin \tau \quad (14)$$

and

$$\beta_2 = \beta_1 + \varphi l_1 \quad (15)$$

$$\delta_2 = \delta_1 + \varphi l_2. \quad (16)$$

Using these notations (3) can be written in the *linear* form

$$\Phi(x_i, y_i) = \frac{x_i}{f} \beta_2 + \frac{y_i}{f} \delta_2 - \frac{x_i}{f} \Phi(x_i, y_i) l_1 - \frac{y_i}{f} \Phi(x_i, y_i) l_2 + \varphi. \quad (17)$$

In (17) the terms $\frac{x_i}{f}$ and $\frac{y_i}{f}$ are *a priori* known, while $\Phi(x_i, y_i)$ is assumed known for all (x_i, y_i) in the observed image. All other terms of the linear equation (17), i.e., $\beta_2, \delta_2, l_1, l_2, \varphi$, are unknown variables. These variables, except φ , are functions of the spatial frequency (u_s, v_s) of the selected sinusoidal component measured in the surface coordinate system, the distance z_0 , and the tilt and slant angles. On a planar surface, the tilt and slant angles are constant. Since the observed surface texture is homogeneous, the variables $\tilde{u}, \tilde{v}, \varphi$ are independent of (x_i, y_i) . Hence, we conclude that all the unknown variables in (17) are independent of (x_i, y_i) .

Note that (17) holds for every (x_i, y_i) in the image of the observed surface. However, in practice the phase function $\Phi(x_i, y_i)$ is unknown, and hence in solving (17) for the tilt and slant angles $\Phi(x_i, y_i)$ must be substituted with its estimate. In this paper we employ the PD Algorithm to estimate the phase as a function of the image coordinates. Since the phase estimation procedure is subject to errors due to observation noise and model mismatch, (17) holds only approximately. Thus let $e(x_i, y_i)$ denote the modeling error of (17). Rewriting (17) for every (x_i, y_i) in the image of the observed surface we obtain the following matrix form:

$$\mathbf{A}\boldsymbol{\lambda} - \mathbf{b} = \mathbf{e} \quad (18)$$

where (see (19)–(21) at the top of the following page), and the unknown parameter vector is

$$\boldsymbol{\lambda} = [\beta_2 \quad \delta_2 \quad l_1 \quad l_2 \quad \varphi]^T. \quad (22)$$

The least squares solution of (18) is given by

$$\boldsymbol{\lambda} = [\mathbf{A}^T \mathbf{A}]^{-1} \mathbf{A}^T \mathbf{b}. \quad (23)$$

Having estimated $\boldsymbol{\lambda}$, the tilt and slant angles can be computed using the estimated values of l_1 and l_2 . Let

$$\tau_1 = \arctan\left(\frac{l_2}{l_1}\right) \quad (24)$$

$$\sigma_1 = \arctan\left(\frac{l_1}{\cos \tau_1}\right). \quad (25)$$

However, four possible solutions for τ and σ satisfy (24), (25). These are given by

$$\tau, \sigma = \left\{ \begin{array}{l} \tau_1, \sigma_1 \\ \tau_1, \sigma_1 + \pi \\ \tau_1 + \pi, -\sigma_1 \\ \tau_1 + \pi, -\sigma_1 + \pi \end{array} \right\}. \quad (26)$$

Yet, only one of the four possible solutions satisfies the constraint that $0 \leq \sigma < (\pi/2)$. This is the desired solution for the tilt and slant angles.

In Appendix B it is further shown that very often the solution for the tilt angle does not require *a priori* knowledge of the origin of the image coordinate system nor of the focal length f .

In the foregoing discussion it was assumed that the phase function $\Phi(x_i, y_i)$ can be estimated from the observed data. Next, we propose two possible algorithms for obtaining the phase estimate. The first method is a direct one. Having estimated the 2-D polynomial model of the observed signal phase using the PD Algorithm, and evaluating the estimated phase by substituting the estimated coefficients into (5), these

$$\mathbf{b} = [\Phi(0,0), \dots, \Phi(0, M-1), \Phi(1,0), \dots, \Phi(1, M-1), \dots, \Phi(N-1,0), \dots, \Phi(N-1, M-1)]^T \quad (19)$$

$$\mathbf{e} = [e(0,0), \dots, e(0, M-1), e(1,0), \dots, e(1, M-1), \dots, e(N-1,0), \dots, e(N-1, M-1)]^T \quad (20)$$

$$\mathbf{A} = \frac{1}{f} \begin{bmatrix} 0 & 0 & -0 \cdot \Phi(0,0) & -0 \cdot \Phi(0,0) & f \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & M-1 & -0 \cdot \Phi(0, M-1) & -(M-1) \cdot \Phi(0, M-1) & f \\ 1 & 0 & -1 \cdot \Phi(1,0) & -0 \cdot \Phi(1,0) & f \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & M-1 & -1 \cdot \Phi(1, M-1) & -(M-1) \cdot \Phi(1, M-1) & f \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ N-1 & 0 & -(N-1) \cdot \Phi(N-1,0) & -0 \cdot \Phi(N-1,0) & f \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ N-1 & M-1 & -(N-1) \cdot \Phi(N-1, M-1) & -(M-1) \cdot \Phi(N-1, M-1) & f \end{bmatrix} \quad (21)$$

estimated values are substituted into (19), (21) instead of the unknown phase values. Next, (23)–(26) are solved using these estimated phase values. The second method is an extension of this method. It is described next.

B. Tilt and Slant Estimation Using the Unwrapped Phase

Since the observed phase of any complex-valued signal is always in the interval $[-\pi, \pi]$, a meaningful interpretation of the phase information is possible only if successful unwrapping of the phase function can be performed to remove the inherent 2π ambiguities of the observed phase. In [21] we present a model-based, 2-D phase unwrapping algorithm for complex-valued 2-D signals with continuous phase functions. The basic building block of this phase unwrapping algorithm is the PD Algorithm, [16], [17]. Since 2-D continuous functions can be approximated by 2-D polynomials, the first step of the phase unwrapping algorithm is to fit a 2-D polynomial model to the observed phase. The estimated phase is then used as a reference information that directs the actual phase unwrapping process: the phase of each sample of the observed field is unwrapped by increasing (decreasing) it by the multiple of 2π which is the nearest to the difference between the principle value of the phase and the estimated phase value at this coordinate.

More specifically, let $\phi(x_i, y_i)$, $\phi_{\text{PV}}(x_i, y_i)$, $\psi(x_i, y_i)$ denote the phase function of the noiseless signal, the principle value of the observed phase, and the unwrapped phase obtained by the proposed procedure, respectively. Also let $\hat{\phi}(x_i, y_i)$ denote the estimated phase obtained using the PD Algorithm. In the absence of noise we have that

$$\phi(x_i, y_i) - \phi_{\text{PV}}(x_i, y_i) = 2\pi k \quad (27)$$

where k is some integer. However, in practice, $\phi(x_i, y_i)$ is unknown to us, and we only have $\hat{\phi}(x_i, y_i)$, which is estimated from the observed noisy measurements. Hence, replacing $\phi(x_i, y_i)$ by $\hat{\phi}(x_i, y_i)$ we obtain the basic unwrapping formula for the observed signal phase

$$\psi(x_i, y_i) = 2\pi \cdot \text{ROUND} \left(\frac{\hat{\phi}(x_i, y_i) - \phi_{\text{PV}}(x_i, y_i)}{2\pi} \right) + \phi_{\text{PV}}(x_i, y_i). \quad (28)$$

Next, the unwrapped phase values are substituted into (19), (21) instead of the unknown phase, and (23)–(26) are solved.

IV. THE CRB ON THE ERROR VARIANCE IN ESTIMATING THE TILT AND SLANT ANGLES

In this section we derive the Cramér–Rao bound (CRB) on the error variance in estimating the tilt and slant angles when the planar surface is observed in white additive Gaussian noise. The CRB provides a well-known lower bound on the achievable variance of any unbiased estimator of these parameters.

Assuming the energy of the texture purely indeterministic component is much smaller than that of the observation noise, we have that an approximate model of the homogeneous surface texture is given by (2). Hence, the observed field model is given by

$$t(x_i, y_i) = \sum_{l=1}^L A_l \cos(\Phi_l(x_i, y_i)) + n(x_i, y_i) \quad (29)$$

where $n(x_i, y_i)$ denotes the zero-mean white Gaussian observation noise, whose variance is ρ^2 . The dimensions of the observed field are $N \times M$. Rewriting (3) we have

$$\Phi_l(x_i, y_i) = [u_l \quad v_l] \tilde{z} \begin{bmatrix} \cos \tau & \sin \tau \\ -\cos \sigma \sin \tau & \cos \sigma \cos \tau \end{bmatrix} \begin{bmatrix} \frac{x_i}{f} \\ \frac{y_i}{f} \end{bmatrix} + \varphi_l \quad (30)$$

where $\tilde{z} = (z_0 / (\sin \sigma ((x_i/f) \cos \tau + (y_i/f) \sin \tau) + \cos \sigma))$. Define

$$\mathbf{t} = [t(0,0), \dots, t(0, M-1), t(1,0), \dots, t(1, M-1), \dots, t(N-1,0), \dots, t(N-1, M-1)]^T. \quad (31)$$

The vectors $\Phi_l, l = 1, \dots, L$ and \mathbf{n} are similarly defined. Let also $\mathbf{D}_l = \cos \Phi_l$ where $\cos \Phi_l$ is an MN -dimensional column vector such that each of its entries is the cosine of corresponding entry of Φ_l . Let $\mathbf{a} = [A_1, A_2, \dots, A_L]^T$ and let $\mathbf{D} = [\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_L]$. Using these definitions, we can rewrite the observations equation (29) in the following matrix representation: $\mathbf{t} = \mathbf{D}\mathbf{a} + \mathbf{n}$. Note that \mathbf{t} is a linear function of

the amplitude parameter vector \mathbf{a} , while the tilt and slant parameters enter nonlinearly through \mathbf{D} . Let $\tilde{u}_l = u_l z_0, \tilde{v}_l = v_l z_0$

$$\tilde{\boldsymbol{\theta}} = [\sigma, \tau, \tilde{u}_1, \dots, \tilde{u}_L, \tilde{v}_1, \dots, \tilde{v}_L, \varphi_1, \dots, \varphi_L]^T \quad (32)$$

and let

$$\boldsymbol{\theta} = [\tilde{\boldsymbol{\theta}}^T, \mathbf{a}^T]^T \quad (33)$$

denote the vector of the unknown problem parameters. Also, let $\mathbf{V} = \mathbf{D}\mathbf{a}$. Since the observation noise is assumed to be a real-valued Gaussian white noise field with zero mean and variance ρ^2 , the probability density function of the observations is given by

$$p(\mathbf{t}; \boldsymbol{\theta}) = \frac{1}{(2\pi\rho^2)^{\frac{MN}{2}}} \exp\left\{-\frac{1}{2\rho^2}\|\mathbf{t} - \mathbf{V}\|^2\right\}. \quad (34)$$

To derive the CRB we use the well-known formula which states that the elements of the Fisher Information Matrix (FIM) are given by

$$\mathbf{F}_{ij} = -E\left\{\frac{\partial^2 \Lambda}{\partial \theta_i \partial \theta_j}\right\} \quad (35)$$

where Λ denotes the log-likelihood function. The CRB is simply the inverse of the FIM [23]. Thus to evaluate the FIM we need to compute the derivatives of the log-likelihood function with respect to the various parameters of interest, and take their expected value. Thus

$$-E\left\{\frac{\partial^2 \Lambda}{\partial \theta_i \partial \theta_j}\right\} = \frac{1}{\rho^2} \left(\frac{\partial \mathbf{V}^T}{\partial \theta_i} \frac{\partial \mathbf{V}}{\partial \theta_j}\right). \quad (36)$$

Rewriting (36) using matrix notations we conclude that the FIM has the form

$$\mathbf{F} = \mathbf{f}^T \mathbf{f} \quad (37)$$

where

$$\mathbf{f} = \frac{1}{\rho} \left[\frac{\partial \mathbf{V}}{\partial \sigma}, \frac{\partial \mathbf{V}}{\partial \tau}, \frac{\partial \mathbf{V}}{\partial \tilde{u}_1}, \dots, \frac{\partial \mathbf{V}}{\partial \tilde{u}_L}, \frac{\partial \mathbf{V}}{\partial \varphi_1}, \dots, \frac{\partial \mathbf{V}}{\partial \varphi_L}, \frac{\partial \mathbf{V}}{\partial A_1}, \dots, \frac{\partial \mathbf{V}}{\partial A_L} \right]. \quad (38)$$

Note that care should be taken in evaluating the performance of an algorithm for estimating angular parameters and in comparing it with the CRB. More specifically, in evaluating the distance between an angular parameter and its estimate, the distance should be evaluated as the minimal Euclidean distance between the parameter, say θ and $\hat{\theta} + kT$ where $\hat{\theta}$ is the estimated value of the parameter, T is the period length, and k can assume any integer value. This is due to the fact that angular parameters are defined on an interval. Hence, if a parameter assumes a value close to the interval boundary, even a small estimation error could be interpreted as a large deviation (whose magnitude is close to the period length T) because of the periodic folding of the parameter space.

The results of numerically evaluating the CRB on estimating the tilt and slant angles as a function of the focal length for some specific examples of surface textures indicate that as the absolute value of the focal length is getting smaller, so does the

CRB on the error variance in estimating the slant and the tilt. This conclusion is intuitively explained by the fact that as the absolute value of the focal length is smaller, we see ‘‘more’’ of the planar surface in the same image. In the limiting case, where $|f| \rightarrow \infty$, the perspective projection becomes an orthogonal projection. In that case, it is impossible to estimate the surface orientation from a single observation of its texture.

A similar numerical analysis of the dependence of the CRB on the orientation of the planar surface, expressed through τ and σ , indicates that the bounds on estimating the tilt and slant have relatively low values (less than 1° standard deviation for an SNR of 10 dB). The only exception is the bound on the error variance in estimating the tilt, which becomes very high as the slant tends to zero. Indeed, as $\sigma \rightarrow 0$ the tilt is losing its physical meaning. Furthermore, the results indicate that the bounds on both the slant and the tilt are strongly influenced by the value of the slant angle but are nearly constant functions of the tilt. Evaluation of the relation between the CRB and the distance of the planar surface from the pinhole z_0 indicates that as the camera is getting further away from the surface, the CRB becomes smaller. This is due to the fact that as the camera is further away from the surface, there are more periods of the surface texture sinusoidal components in the same size of an image.

Note from (30) that the distance of the planar surface z_0 always appears in the phase equation as a multiplication factor of the surface texture spatial frequencies (u_l, v_l) . Thus the functional dependence of the CRB on the texture’s frequencies (u_l, v_l) is similar to its dependence on z_0 . We therefore conclude that employing harmonic components with higher frequencies in an estimation procedure of the slant and tilt should result in better estimates than when low-frequency components with the same amplitude are used. More specifically, our experimental results show that even for medium to high SNRs, if we observe less than four cycles of the texture sinusoidal component, the values of the CRB are very high and the estimation of the tilt and slant is difficult. On the other hand, when the number of observed periods is higher than four, the CRB is considerably lower. Hence, the component to which the phase estimation is applied should be selected as the one with the highest energy among those components for which the number of observed periods is no less than four or five. (Clearly, the selection maybe a matter of a tradeoff between the number of cycles and the amplitudes of the competing components.)

V. NUMERICAL EXAMPLES

In this section we present some results of applying the proposed parametric model and algorithms to synthetic as well as to real images. We also perform some Monte Carlo simulations to analyze the statistical properties of the different algorithms.

A. Statistical Performance Analysis of the Estimation Algorithms

In this subsection we illustrate the performance of the proposed parameter estimation algorithms using Monte Carlo simulations. We compare the estimation error variances of the suggested algorithms with the CRB derived in Section IV. The sur-

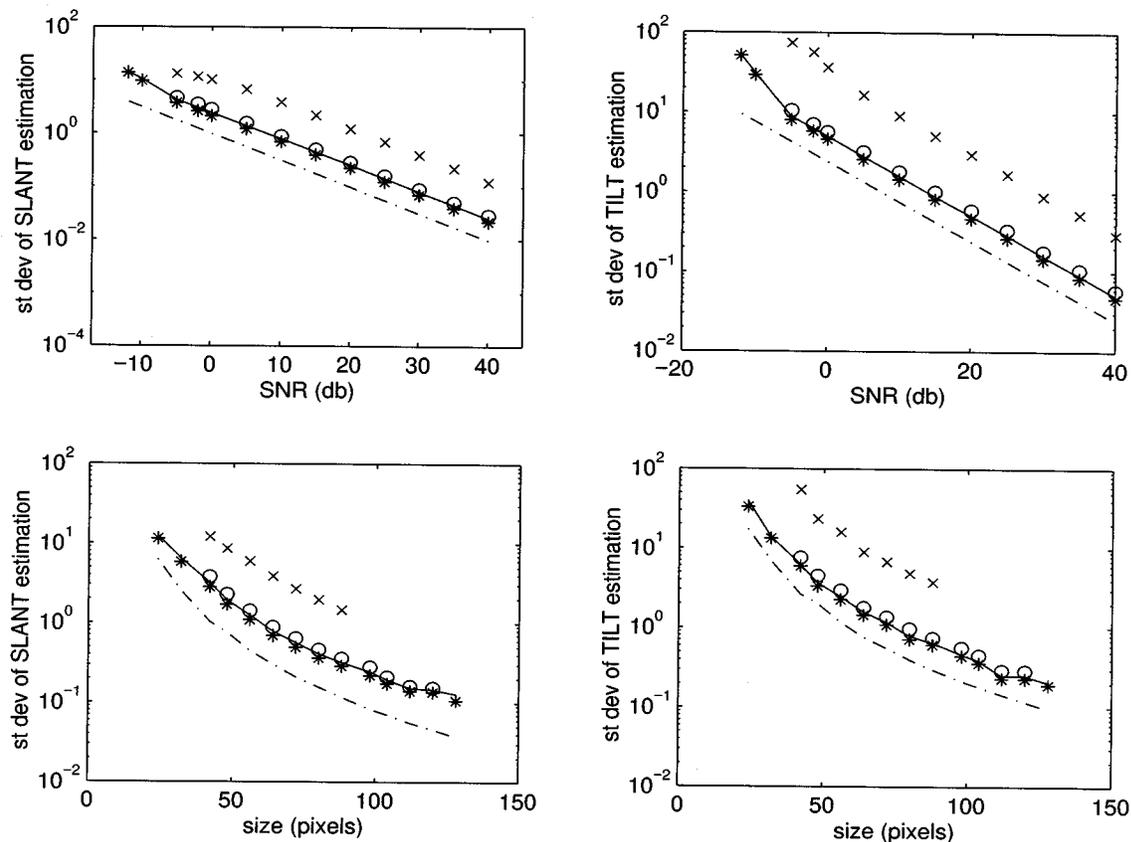


Fig. 2. Performance of the proposed algorithms as a function of SNR_D and data dimensions, in comparison with the corresponding CRB. Dashed-dotted line denotes the CRB, \times denotes the performance curve of the algorithm that employs the Taylor series expansion [15], \circ denotes the performance curve of the search algorithm initialized by the algorithm based on the Taylor expansion [15], solid line denotes the performance curve of the linear least squares estimator that employs the estimated phase, $*$ denotes the performance curve of the linear least squares estimator that employs the unwrapped phase.

face texture being considered in these experiments is composed of three sinusoids, given by

$$t(x_s, y_s) = \sin(x_s u_0 + y_s v_0) + \frac{1}{3} \sin(3x_s u_0 + 3y_s v_0) + \frac{1}{5} \sin(5x_s u_0 + 5y_s v_0) \quad (39)$$

where $(u_0, v_0) = (0.25, 0)$ cycles/cm. To generate the image of the planar surface, the intensity of each pixel in the image plane was evaluated by projecting the intensity levels of the surface texture using (45). More specifically, the intensity of each image pixel (x_i, y_i) is that of the surface coordinate which is mapped by the perspective projection to (x_s, y_s) . The surface orientation parameters are $\sigma = 30^\circ$ and $\tau = 90^\circ$. The focal length of the camera is $f = -60$ mm, $z_0 = 6$ m, and the image plane dimensions are $30 \text{ mm} \times 30 \text{ mm}$, with the origin being located at the center of the image plane. The observation noise is a zero-mean additive white Gaussian noise. Define $\text{SNR}_D = A_D^2 / \rho^2$ where A_D is the amplitude of the selected sinusoidal component (the dominant one in this example) and ρ^2 is the variance of the observation noise. We investigate the performance of the algorithms as a function of the selected component signal to noise ratio SNR_D , and the dimensions of the observed square image.

The experimental standard deviation results depicted in Fig. 2 are based on 500 independent realizations of the image for each SNR_D and data dimensions. For comparison, we have included

in this figure the results obtained by the two estimation algorithms derived in [15]. Recall that the CRB is a lower bound on the error variance of any unbiased estimator of the problem parameters. Hence, the Monte Carlo results in Fig. 2 are depicted only for SNR values (data dimensions) above the threshold SNR (data dimensions). In other words, for each of the algorithms, the leftmost point of its performance curve is the threshold point. Lower SNRs (data dimensions) result in biased estimates. The results indicate that while the computational complexity of the two algorithms derived in this paper is equivalent to the complexity of the algorithm that estimates the tilt and slant using the Taylor series approach, their bias and standard deviation are considerably lower than the bias and standard deviation of the algorithm that employs the Taylor series expansion. Moreover, the performance of these two algorithms is slightly better than that of the computationally demanding iterative search algorithm. We note that the computational complexity of the search algorithm in [15] is still much lower than that of the exhaustive search suggested in [14] due to its initialization by the Taylor series based estimation algorithm.

The Monte Carlo simulations summarized in Fig. 2 demonstrate that the proposed algorithms are essentially unbiased and their error variance very close to the CRB even for low SNR values, and relatively small dimensions of the observed image (such as 24×24 pixels). Note that the algorithm that employs the unwrapped phase produces slant and tilt estimates

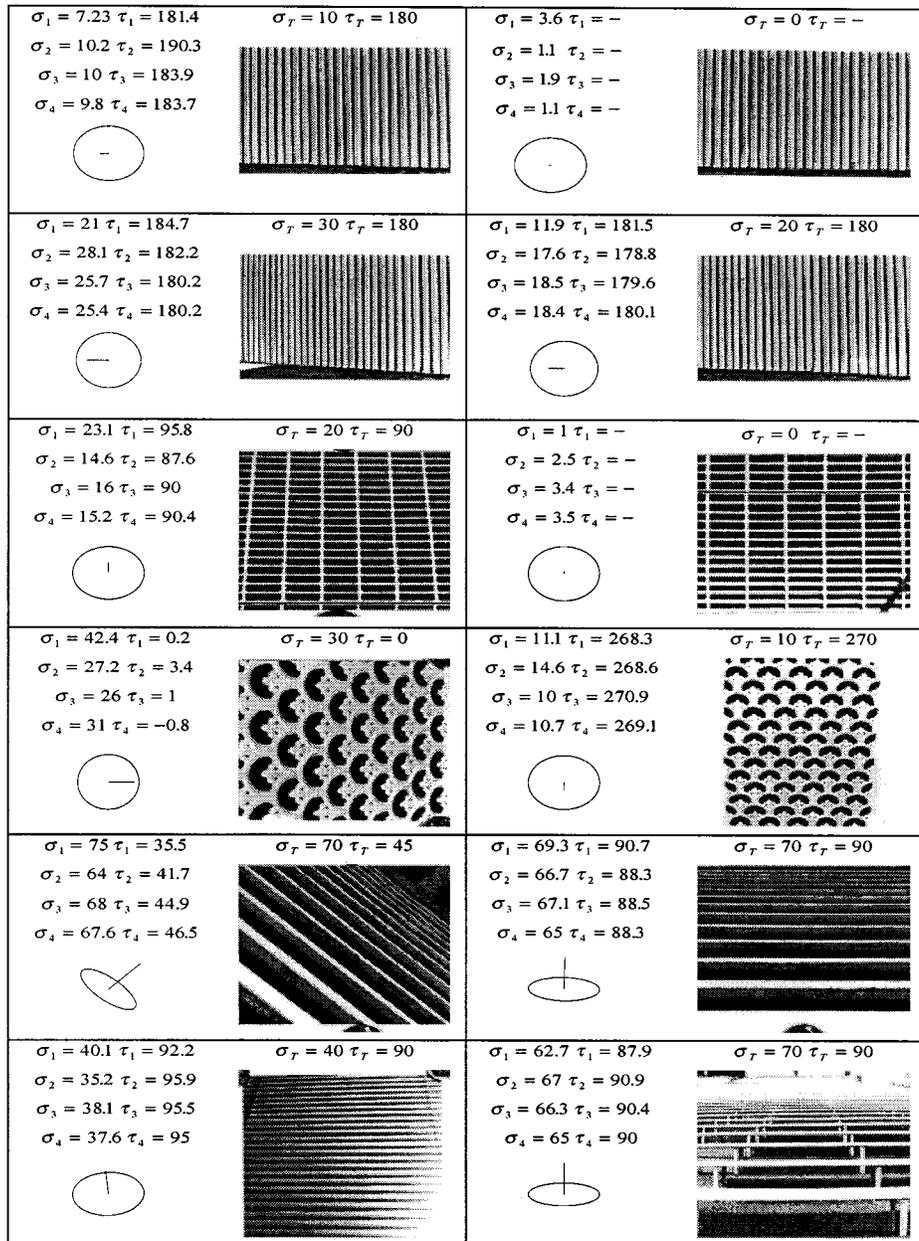


Fig. 3. The images and the corresponding estimated tilt and slant parameters using the four algorithms. Here, $(\tau_1, \sigma_1) - (\tau_4, \sigma_4)$ denote the estimated tilt and slant produced by the four algorithms. The measured tilt, τ_T , and slant, σ_T , are given in the first row. Based on the estimated τ_4 and σ_4 we depict the orientation of the surface normal as seen in the image plane. The ellipse illustrates how a circle drawn on the planar surface would appear in the image plane whose tilt and slant are τ_4 and σ_4 , respectively.

with smaller bias and error variance than any other algorithm. We therefore conclude that the proposed algorithms produce nearly optimal estimates at computational complexity which is considerably lower than that of any existing algorithm.

B. Experimental Results with Real-World Textured Surfaces

In this subsection, we evaluate the performance of the algorithms by applying them to photographed textured surfaces. The images are those used in [14]. The focal length of the camera and the image coordinate system in common units, are known. The algorithms were applied to a 64×64 segment of each original 128×128 image. Fig. 3 shows the images, the estimated ori-

entation produced by each of the proposed algorithms, and the measured ("true") orientation. Based on the tilt and slant estimated using the algorithm that employs the unwrapped phase, we depict the orientation of the surface normal as seen in the image plane. The ellipse illustrates how a circle drawn on the planar surface would appear in the image plane. For comparison, we have also included in this figure the results obtained by the two estimation algorithms derived in [15]. Thus we use in this example the following notations: σ_1, τ_1 are the estimated slant and tilt produced by the algorithm that employs the Taylor series expansion, [15]; σ_2, τ_2 are the estimated slant and tilt produced after an iterative minimization stage is applied to improve the re-

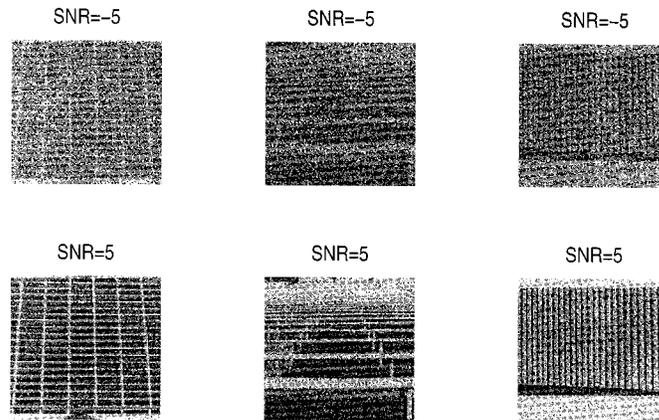


Fig. 4. Observed noisy images at SNR values of 5 dB and -5 dB.

TABLE I
THE RESULTS OF ESTIMATING THE ORIENTATION OF PLANAR SURFACES FROM REAL NOISY IMAGES USING THE ALGORITHM THAT EMPLOYES THE UNWRAPPED PHASE IMAGES ARE NUMBERED IN CORRESPONDENCE WITH FIG. 4

Image	(a) $\sigma_T = 20, \tau_T = 90$				(b) $\sigma_T = 70, \tau_T = 90$				(c) $\sigma_T = 20, \tau_T = 180$			
	SNR	10	5	0	-5	10	5	0	-5	10	5	0
mean σ	15.4	15.5	15.4	17.0	65.0	65.1	64.2	62.8	18.3	18.0	18.0	17.8
st. dev σ	0.7	1.6	2.5	4.2	0.3	0.6	7.6	9.9	0.4	0.8	1.4	2.1
mean τ	90.6	90.0	90.7	89.6	89.9	90.0	91.3	91.7	180.0	180.1	179.8	180.6
st.dev τ	2.5	5.8	9.6	15.2	0.7	1.2	9.4	10.7	1.0	2.2	3.4	6.2

sults of the Taylor series expansion based algorithm, [15]; σ_3, τ_3 are the estimated slant and tilt produced by the least squares solution based on the estimated phase, Section III-A; while σ_4, τ_4 provide the estimation results produced by the least squares solution based on the unwrapped phase, Section III-B. Note that the measured tilt τ_T and slant σ_T are subject to a measurement error of 1° – 3° .

C. Experimental Results with Real-World Textured Surfaces in the Presence of Noise

In this example, we illustrate the performance of the proposed algorithm in the presence of noise for real-world textured surfaces observed at various SNRs. Since the models of these real-world images are unknown, the SNR here is evaluated as the ratio between the experimental variance of the noiseless image (shown in Fig. 3) and the noise variance. The performance of the algorithm that employs the unwrapped phase (Algorithm 4) is evaluated using Monte Carlo simulations for SNR values of 10, 5, 0, and -5 dB. For illustration purposes, Fig. 4 shows a single realization of each test image at SNR values of 5 dB and -5 dB. Note that the algorithm is applied to a 64×64 segment of the 128×128 image shown in Fig. 4. The results of the Monte Carlo simulations, summarized in Table I, indicate that the proposed algorithm is effective even at relatively low SNR's.

D. Recovery of the Homogeneous Surface Texture from the Perspective Image

Once the tilt and slant angles of the observed surface have been estimated, it becomes possible to recover the homoge-

neous surface texture from the perspective viewed image of that surface through nonuniform resampling of the observed image. As explained in Section I, applying this procedure to the entire image, of which the textured surface is a part, considerably simplifies further processing such as coding, and content-based indexing and retrieval of images. In the following, we summarize the main steps of the algorithm for “normalizing” the observed image so that the effect of the perspective projection is eliminated: using the inverse coordinate transformation (45) the coordinates of the image boundaries, expressed in surface coordinates are found. The surface coordinate system is then uniformly sampled, and the image coordinate \mathbf{x}_i that corresponds to each \mathbf{x}_s on the surface sampling grid is evaluated using (44). Finally, the gray level of each sample in the surface coordinate system is set to the gray level of the corresponding observed image sample \mathbf{x}_i (using interpolation since in general the resulting x_i and y_i are not integers). The results of applying this procedure to three real-world images are shown in Fig. 5. It is easily seen that the recovered textures are indeed nearly homogeneous.

VI. CONCLUSIONS

We have presented a parametric solution to the problem of estimating the orientation in space of a planar textured surface, from a single, noisy, observed image of it. Based on the nonlinear physical model of the perspective projection, we derive the Cramér–Rao lower bound on the error variance of estimating the tilt and slant of the observed surface. The algorithms derived in this paper are based on the observation that the coordinate transformation from surface to image coordinates, due to

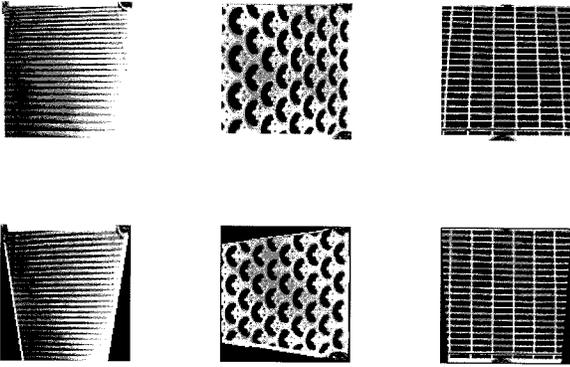


Fig. 5. Recovered homogeneous surface texture from the observed perspective projected image.

the perspective projection, uniquely transforms each homogeneous sinusoidal component of the surface texture into a sinusoid whose frequency is a function of location in the image coordinate system. It is shown that the phase of each of the sinusoids can be expressed as a linear function of some variables that are related to the surface tilt and slant angles. The two-dimensional Hilbert transform is employed to guarantee that the instantaneous phase of each observed inhomogeneous sinusoidal component is uniquely defined. Using the Phase Differencing Algorithm, the unwrapped phase function of a sinusoidal component of the observed texture is evaluated. Substituting in the derived linear relation the unknown phase with its unwrapped estimate, and solving the resulting system by a linear least squares solution, the tilt and slant are estimated. The proposed algorithm is shown to produce unbiased estimates with error variance which is close to the CRB for a wide range of SNRs, at computational complexity which is considerably lower than that of any existing algorithm. To the best of our knowledge, the derivation of universal performance bounds for this problem, and the estimation of the parameters of the perspective projection model in the presence of noise, have never been considered in the literature. Finally, once the tilt and slant angles of the observed surface have been estimated it becomes possible to recover the homogeneous surface texture from the perspective viewed image of that surface, through nonuniform resampling of the observed image. Applying this procedure to the entire image, of which the textured surface is a part, eliminates its inhomogeneities due to the perspective projection, thus considerably simplifying further processing such as coding, segmentation, and content-based indexing and retrieval of images.

APPENDIX A THE PERSPECTIVE TRANSFORMATION

This appendix defines the viewing geometry we use. In the following, we adopt the notations used by Super and Bovik in [14] and assume a pinhole perspective projection model since it provides a good approximation to a lens-type imaging system.

Assign a world coordinate system $\mathbf{x}_w = [x_w \ y_w \ z_w]^T$ to the imaging system such that its origin is at the focal point and the $-z_w$ -axis is the optical axis (see Fig. 1). The image plane is

located at $z_w = f < 0$ where $|f|$ is the focal length. Define the image plane coordinate system $\mathbf{x}_i = [x_i \ y_i]^T$ such that $x_i = x_w$ and $y_i = y_w$.

We use the slant-tilt system for representing the orientation of the plane. The slant σ is the angle between the surface normal and the optical axis z_w . The tilt τ is the angle between the x_i -axis and the projection of the surface normal onto the image plane. To describe a texture on the surface, we must define a coordinate system $\tilde{\mathbf{x}}_s = [x_s \ y_s \ z_s]^T$ on the surface. This coordinate system is formed by

- 1) setting the z_s -axis to be the surface normal;
- 2) setting the x_s -axis to be the back-projection onto the surface of the image tilt vector $(\cos \tau, \sin \tau)$;
- 3) setting the y_s -axis so as to form right-handed orthogonal coordinate system; and
- 4) setting the origin at the intersection of the surface with z_w -axis.

Thus the coordinate transformation from the surface coordinate system to the world coordinate system is given by

$$\mathbf{x}_w = \begin{bmatrix} \cos \sigma \cos \tau & -\sin \tau & \sin \sigma \cos \tau \\ \cos \sigma \sin \tau & \cos \tau & \sin \sigma \sin \tau \\ -\sin \sigma & 0 & \cos \sigma \end{bmatrix} \tilde{\mathbf{x}}_s + \begin{bmatrix} 0 \\ 0 \\ z_0 \end{bmatrix} \quad (40)$$

where z_0 is the z_w -coordinate of the surface where it crosses the optical axis.

The coordinate transformation of a point in the world coordinate system to image coordinates due to the perspective projection is given by

$$\mathbf{x}_i = \frac{f}{z_w} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \mathbf{x}_w. \quad (41)$$

Since for any surface point we have by definition that $z_s = 0$, let us define $\mathbf{x}_s = [x_s \ y_s]^T$ to be the coordinate vector of a surface point. Therefore, the surface to world coordinate transformation of a point *on* the surface is given using (40) by

$$\mathbf{x}_w = \begin{bmatrix} \cos \sigma \cos \tau & -\sin \tau \\ \cos \sigma \sin \tau & \cos \tau \\ -\sin \sigma & 0 \end{bmatrix} \mathbf{x}_s + \begin{bmatrix} 0 \\ 0 \\ z_0 \end{bmatrix}. \quad (42)$$

For any point of the surface we have that its z_w coordinate is given by

$$z_w = z_0 - x_s \sin \sigma. \quad (43)$$

Substituting (42) and (43) into (41) we obtain the surface-to-image coordinate transformation of a point *on* the surface to a point on the image plane due to the perspective projection:

$$\frac{\mathbf{x}_i}{f} = \frac{1}{z_0 - x_s \sin \sigma} \begin{bmatrix} \cos \tau & -\sin \tau \\ \sin \tau & \cos \tau \end{bmatrix} \begin{bmatrix} \cos \sigma & 0 \\ 0 & 1 \end{bmatrix} \mathbf{x}_s. \quad (44)$$

The matrix

$$\begin{bmatrix} \cos \tau & -\sin \tau \\ \sin \tau & \cos \tau \end{bmatrix}$$

is a rotation matrix, and the matrix

$$\begin{bmatrix} \cos \sigma & 0 \\ 0 & 1 \end{bmatrix}$$

provides the projection of \mathbf{x}_s to \mathbf{x}_i/f for a zero tilt. The term $\frac{1}{z_0 - x_s \sin \sigma}$ is a varying scaling factor due to the different distance of each surface point from the pinhole.

The inverse of (44) is given by

$$\mathbf{x}_s = z_w \begin{bmatrix} \sec \sigma & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \tau & \sin \tau \\ -\sin \tau & \cos \tau \end{bmatrix} \frac{\mathbf{x}_i}{f} \quad (45)$$

where substitution of (45) into (43) yields

$$z_w = \frac{z_0}{\tan \sigma \left(\frac{x_i}{f} \cos \tau + \frac{y_i}{f} \sin \tau \right) + 1}. \quad (46)$$

Substituting the inverse coordinate transformation expression (45) into the model (2), we obtain the model of the texture sinusoidal components, projected onto the image plane, i.e.,

$$t_i(\mathbf{x}_i) = t_s[\mathbf{x}_s(\mathbf{x}_i)] = \sum_{l=1}^L A_l \cos(\Phi_l(\mathbf{x}_i)) \quad (47)$$

where the phase of a sinusoidal component whose phase function is given in surface coordinates by $\Phi_l(x_s, y_s) = x_s u_l + y_s v_l + \varphi_l$, becomes

$$\Phi_l(x_i, y_i) = \frac{\frac{x_i}{f} (\tilde{u}_l \cos \tau - \tilde{v}_l \cos \sigma \sin \tau) + \frac{y_i}{f} (\tilde{u}_l \sin \tau + \tilde{v}_l \cos \sigma \cos \tau)}{\tan \sigma \left(\frac{x_i}{f} \cos \tau + \frac{y_i}{f} \sin \tau \right) + 1} + \varphi_l \quad (48)$$

in the coordinate system of the observed image, and we define $\tilde{u}_l = u_l z_0$, $\tilde{v}_l = v_l z_0$.

Since the origin of the observed surface is projected onto the origin of the image, we conclude that for each sinusoidal component of the surface texture, its projection on the image has the same initial phase φ_l as on the surface. This is because the initial value of each cosine function, i.e., its value at $(0, 0)$, remains unchanged under a projection that keeps the origin.

APPENDIX B

A PROPERTY OF THE PERSPECTIVE PROJECTION

In this appendix we elaborate on a special property of the perspective projection, that may be very useful in some special estimation problems.

Theorem 2: Assume that the signs of z_0 and the focal length are known. Then, tilt estimation is independent of the magnitude of the focal length and is invariant to translations of the image coordinate system.

Proof: Using (11)–(14) we can rewrite (3) in the form

$$\Phi(x_i, y_i) = \frac{\frac{x_i}{f} \beta_1 + \frac{y_i}{f} \delta_1}{\frac{x_i}{f} l_1 + \frac{y_i}{f} l_2 + 1} + \varphi. \quad (49)$$

Since in any given problem setting β_1 and δ_1 are fixed, the numerator of (49) is a linear function of the coordinates. Therefore, the nonlinearity of $\Phi(x_i, y_i)$ due to the perspective projection (and hence the inhomogeneity of the signal) is expressed only by the denominator of (49). In other words, by estimating l_1 and l_2 we can uniquely find the slant and the tilt from (13) and (14), as suggested in Section III. The solution for the tilt is given by

$$\tau = \left\{ \begin{array}{l} \arctan \frac{l_2}{l_1} \\ \arctan \frac{l_2}{l_1} + \pi \end{array} \right\}. \quad (50)$$

Using the constraint that the slant is between zero and $\pi/2$, only one solution for tilt is possible. (The one for which substitution of τ into (25) yields slant angle between zero and $\pi/2$).

Assume now that the magnitude of the focal length is unknown. This results in a scaling problem where the true coordinates $\frac{x_i}{f}$ and $\frac{y_i}{f}$ are replaced by $\alpha \frac{x'_i}{f}$ and $\alpha \frac{y'_i}{f}$, with $\alpha > 0$ being unknown. Similarly, if the pinhole location is unknown, we end up with an unknown location of the origin of the image coordinate system, and hence with an unknown shift of the coordinates x_i and y_i by b_x and b_y , respectively, where (b_x, b_y) is the coordinate of the assumed origin in the image coordinate system. (The origin of the image coordinate system is defined as the point where the optical axis crosses the image plane). The joint effect in the case where both the focal length and the origin of the image coordinate system are unknown is expressed by substituting $\frac{x_i}{f}$ and $\frac{y_i}{f}$ by $\alpha \frac{x'_i}{f} + \frac{b_x}{f}$ and $\alpha \frac{y'_i}{f} + \frac{b_y}{f}$. We thus have the relation

$$\frac{x_i}{f} = \alpha \frac{x'_i}{f} + \frac{b_x}{f} \quad (51)$$

$$\frac{y_i}{f} = \alpha \frac{y'_i}{f} + \frac{b_y}{f}. \quad (52)$$

The effect of the unknown parameters on the denominator of the phase function (49) can be evaluated by substituting (51) and (52) into (49). Following this substitution the denominator gets the form

$$\left(\alpha \frac{x'_i}{f} + \frac{b_x}{f} \right) l_1 + \left(\alpha \frac{y'_i}{f} + \frac{b_y}{f} \right) l_2 + 1. \quad (53)$$

The shift and the scale constants modify the numerator of (49) as well. However, it remains a linear function of the coordinates and hence it has no influence on the nonlinear characteristics of the phase in which we are interested.

We note that (53) can be written in the form

$$\left(\frac{x'_i}{f} \tilde{l}_1 + \frac{y'_i}{f} \tilde{l}_2 + 1 \right) \tilde{c} \quad (54)$$

where

$$\tilde{c} = l_1 \frac{b_x}{f} + l_2 \frac{b_y}{f} + 1, \quad (55)$$

and

$$\tilde{l}_1 = \alpha \frac{l_1}{\tilde{c}} \quad (56)$$

$$\tilde{l}_2 = \alpha \frac{l_2}{\tilde{c}}. \quad (57)$$

Estimating \tilde{l}_1 and \tilde{l}_2 the tilt can be evaluated up to a π ambiguity by

$$\tau = \left\{ \begin{array}{l} \arctan \frac{\tilde{l}_2}{\tilde{l}_1} \\ \arctan \frac{\tilde{l}_2}{\tilde{l}_1} + \pi \end{array} \right\}. \quad (58)$$

We therefore conclude that knowledge of the signs of f and \tilde{c} is sufficient to enable us to unambiguously estimate τ by using the constraint that the slant is between zero and $\pi/2$.

We finally show that if $z_0 < 0$, then $\tilde{c} > 0$. Indeed, since (b_x, b_y) is a point on the image plane, there exists a surface co-

ordinate (\bar{x}_s, \bar{y}_s) which is projected by the perspective projection onto (b_x, b_y) , i.e., from (44) we have

$$\begin{bmatrix} \frac{b_x}{f} \\ \frac{b_y}{f} \end{bmatrix} = \frac{1}{z_0 - \bar{x}_s \sin \sigma} \begin{bmatrix} \cos \sigma \cos \tau & -\sin \tau \\ \cos \sigma \sin \tau & \cos \tau \end{bmatrix} \begin{bmatrix} \bar{x}_s \\ \bar{y}_s \end{bmatrix}. \quad (59)$$

Using (43), (13), (14), (55), and (59) we find that

$$\begin{aligned} \tilde{c} &= l_1 \frac{b_x}{f} + l_2 \frac{b_y}{f} + 1 \\ &= \frac{\tan \sigma}{z_0 - \bar{x}_s \sin \sigma} (\bar{x}_s \cos \sigma \cos^2 \tau - \bar{y}_s \cos \tau \sin \tau \\ &\quad + \bar{x}_s \cos \sigma \sin^2 \tau + \bar{y}_s \cos \tau \sin \tau) + 1 \\ &= \frac{\tan \sigma}{z_0 - \bar{x}_s \sin \sigma} (\bar{x}_s \cos \sigma) + 1 = \frac{\bar{x}_s \sin \sigma}{z_0 - \bar{x}_s \sin \sigma} + 1 \\ &= \frac{z_0}{z_0 - \bar{x}_s \sin \sigma} = \frac{z_0}{\bar{z}_w} \end{aligned} \quad (60)$$

where from (43), \bar{z}_w is the value of the z_w coordinate of the planar surface point (\bar{x}_s, \bar{y}_s) . Since the surface is in front of the camera, \bar{z}_w must be negative as, otherwise, (\bar{x}_s, \bar{y}_s) will not be projected onto the image plane. Hence, from the assumption that $z_0 < 0$, we have that $\tilde{c} > 0$. \square

We finally note that in a standard camera $f < 0$. Also, in nearly every case the intersection of the optical axis and the photographed surface occurs in front of the lens, and hence $z_0 < 0$. We therefore conclude that often, the tilt of the planar surface can be estimated from the phase measured on the image plane, even when the focal length and the origin of the image coordinate system are unknown.

ACKNOWLEDGMENT

The authors wish to thank Prof. B. Super for providing them with the images of Fig. 3.

REFERENCES

- [1] A. P. Witkin, "Recovering surface shape and orientation from texture," *Artificial Intell.*, vol. 17, pp. 17–45, 1981.
- [2] K. Kanatani, "Detection of surface orientation and motion from texture by a stereological technique," *Artificial Intell.*, vol. 23, pp. 213–237, 1984.
- [3] J. Aloimonos and M. Swain, "Shape from patterns: Regularization," *Int. J. Computer Vision*, vol. 2, pp. 171–187, 1988.
- [4] J. V. Stone and S. D. Isard, "Adaptive scale filtering: A general method for obtaining shape from texture," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 713–719, 1995.
- [5] D. Blostein and N. Ahuja, "Shape from texture: Integrating texture-element extraction and surface estimation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 1233–1251, 1989.
- [6] K. Kanatani and T. C. Chou, "Shape from texture: General principle," *Artificial Intell.*, vol. 38, pp. 1–48, 1989.
- [7] L. Y. Jau and R. T. Chin, "Shape from texture using the winger distribution," *Computer Vision, Graphics and Image Process*, vol. 52, pp. 163–248, 1990.
- [8] L. G. Brown and H. Shvaytser, "Surface orientation from projective foreshortening of isotropic texture autocorrelation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, pp. 584–588, 1990.
- [9] M. A. S. Patel and F. Cohen, "Local surface shape estimation of 3-D textured surfaces using Gaussian Markov random fields and stereo windows," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, pp. 1091–1098, 1993.
- [10] H. K. Hong, Y. C. Myung, and J. S. Choi, "3-D analysis of projective textures using structural approaches," *Pattern Recogn.*, vol. 32, pp. 357–364, 1999.
- [11] D. H. Ballard and C. M. Brown, *Computer Vision*. Englewood Cliffs, NJ: Prentice Hall, 1982.
- [12] R. Stoica, J. Zerubia, and J. M. Francos, "The two-dimensional wold decomposition for segmentation and indexing in image libraries," presented at the Int. Conf. Acoust., Speech, Signal Processing, Seattle, WA, 1998.
- [13] J. Garding, "Shape from texture for smooth curved surfaces," in *Proc. Europ. Conf. Computer Vision*, Italy, 1992, pp. 630–638.
- [14] B. Super and A. C. Bovik, "Planar surface orientation from texture spatial frequencies," *Pattern Recogn.*, vol. 28, pp. 729–743, 1995.
- [15] J. M. Francos and H. Permuter, "Parametric estimation of the orientation of textured planar surfaces," *IEEE Trans. Image Processing*, accepted for publication.
- [16] J. M. Francos and B. Friedlander, "Two-dimensional polynomial phase signals: Parameter estimation and bounds," *Multidim. Syst. and Signal Processing*, vol. 9, pp. 173–205, 1998.
- [17] B. Friedlander and J. M. Francos, "An estimation algorithm for 2-D polynomial phase signals," *IEEE Trans. Image Processing*, vol. 5, pp. 1084–1087, 1996.
- [18] Y. M. Zhu, F. Peyrin, and R. Goutte, "The use of a two-dimensional Hilbert transform for Wigner analysis of two-dimensional real signals," *Signal Processing*, vol. 19, pp. 205–220, 1990.
- [19] B. Picinbono, "On instantaneous amplitude and phase signal," *IEEE Trans. Signal Processing*, vol. 45, pp. 552–560, 1997.
- [20] J. M. Francos, A. Z. Meiri, and B. Porat, "A unified texture model based on a 2-D wold like decomposition," *IEEE Trans. Signal Processing*, vol. 41, pp. 2665–2678, 1993.
- [21] B. Friedlander and J. M. Francos, "Model based 2-D phase unwrapping," *IEEE Trans. Signal Processing*, vol. 45, pp. 2999–3007, 1997.
- [22] J. M. Francos and B. Friedlander, "Optimal parameter selection in the phase differencing algorithm for 2-D phase estimation and unwrapping," *IEEE Trans. Signal Processing*, vol. 47, pp. 273–279, 1999.
- [23] B. Porat, *Digital Processing of Random Signals*. Englewood Cliffs, NJ: Prentice-Hall, 1994.