HAL
open science

Shape from texture using continuous wavelet transform

Hartmut Führ, Wen-Liang Hwang, Bruno Torrésani

To cite this version:

Hartmut Führ, Wen-Liang Hwang, Bruno Torrésani. Shape from texture using continuous wavelet transform. Wavelet Applications in Signal and Image Processing VIII, Aug 2000, San Diego, United States. 10.1117/12.408668 . hal-01300326

HAL Id: hal-01300326

https://hal.science/hal-01300326

Submitted on 10 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Shape from texture using continuous wavelet transform

Hartmut Führ[a], Wen-Liang Hwang[b] and Bruno Torrésani[c]

[a] Zentrum Mathematik, Technische Universität München,
D-80290 München, Germany
[b]Institute of Information Science, Academia Sinica
Nankang, Taipei, Taiwan
[c]Laboratoire d'Analyse, Topologie et Probabilités, Université de Provence,
39 rue Joliot-Curie, 13453 Marseille Cedex 13, France

ABSTRACT

We give an algorithm for the estimation of the orientation of a planar surface with a homogeneous texture, viewed under perspective projection. We follow the two step procedure which is usually employed for this type of problem: First, a set of local distortion matrices is estimated – here we use wavelets –, then we determine the surface orientation which best fits the local distortions. In both parts the techniques we use are original.

Keywords: Shape from texture, local distortion, continuous wavelet transforms,local scale.

1. INTRODUCTION

The shape from texture problem can be described as follows: We are given a photographic image of a surface in three-dimensional space. The surface displays a pattern (or texture), which is in some way regular. Then the image gives a distorted view of the surface texture, with the distortion depending on the geometry of the surface and the imaging process. The task consists in reconstructing the surface from the image, using the (suitably defined) regularity assumption.

While the general formulation asks for arbitrary shapes, very often the focus is on the case where the surface is planar. In this case the problem reduces to the computation of the surface normal, usually determined by two angles, called *slant* and *tilt*. (See section 3 for more details, and **Figure 1** for an illustration.) More general shapes are then estimated under the assumption that the surface is locally well approximated by its tangent planes, which allows a local estimation of surface normals. Following this approach, we shall concentrate on the planar case as well.

Among the various types of regular textures, there are, roughly speaking, two main groups under consideration:

- Isotropic textures. In this case the projected texture is no longer isotropic, and introducing and measuring certain degrees of anisotropy allows an estimation of the surface properties. A good example for this approach is foreshortening: Under the projection map, circles become ellipses (at least approximately), and the minor axes of these ellipses point in the tilt direction.[1]

- Homogeneous textures. Homogeneity is understood as translation invariance or periodicity. In this case the usual approach is to consider two neighboring image patches. Under the homogeneity assumption on the original image, one patch is expected to be a distorted version of the other, the distortion being due to the imaging process.

  This naturally leads to a two step procedure for estimating surface orientatio[2]: First estimate distortions from the image, usually in the form of distortion matrices (i.e., the distortion is locally approximated by

Further author information: (Send correspondence to H.F.)
H.F.: E-mail: fuehr@@mathematik.tu-muenchen.de
W.-L.H.: E-mail: whwang@@iis.sinica.edu.tw
B.T.: E-mail: Bruno.Torresani@@sophia.inria.fr

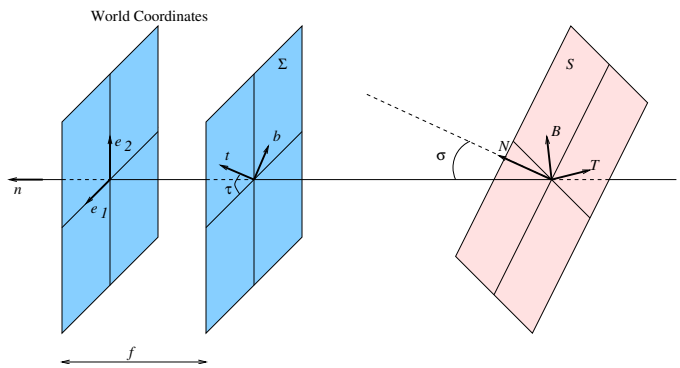

Figure 1. The geometry of the problem: definition of slant (σ) and tilt (τ), and of the three coordinate systems: $(\underline{e}_1, \underline{e}_2, \underline{n})$, $(\underline{t}, \underline{b}, \underline{n})$ and $(\underline{T}, \underline{B}, \underline{N})$.

affine mappings). The second step determines the orientation which best agrees with these distortions. Here it is important to note that a single distortion matrix does not contain sufficient information. It is rather the change in distortion than the distortion alone which is exploited.

Our paper deals with homogeneous textures. Most of the time we concentrate on textures which have one dominant frequency component ("harmonic textures"), or at least a dominant frequency direction. Then the effect of the distortion will be a simple scaling, which considerably simplifies the problem.

This way we obtain a rather simple and somewhat naïve model: Looking at an image of a projected single frequency texture, we perceive the distortion as a change in scale. It seems to be a good first guess to take the direction in which the local scale (suitably defined) decreases most rapidly as tilt direction, and to derive slant from the speed with which the scale changes; or, more formally, to derive tilt and slant from the gradient of the local scale. However there is a catch, as illustrated by **Figure 2**: Whenever the frequency does not point in the direction of tilt, (or perpendicular to tilt), the first tilt estimate will lie somewhere between true tilt and the frequency direction.

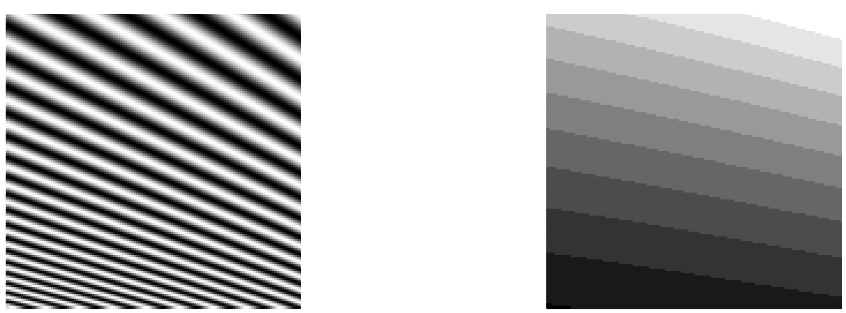

Figure 2. On the left hand side is a projected sinusoid, with frequency pointing away from tilt. On the right hand side is a plot of the corresponding local scales (quantized to better illustrate the effect). In this example, tilt is exactly vertical, but the local scales decrease in a direction between tilt and frequency. The plots were made using the formulas derived in Section 3.

It is however possible to correct the first tilt estimate, using an estimate of the frequency direction as additional information.

For the estimation of the distortions we use the 2D continuous wavelet transform (see Section 5 below for more details). Due to the covariance of CWT with respect to scalar dilations, the distortion results in a displacement of the wavelet coefficients, and we propose an algorithm to estimate this displacement. The continuous wavelet transform is an attractive tool for the shape from texture problem, despite the computational cost it entails

(particularly by comparison to the discrete wavelet transforms), mostly because it is easy to design and allows arbitrary resolution in scale.

Dealing with more general textures requires a more sophisticated approach, mainly because the distortion matrices cannot be assumed to be scalar any more. Adapting the algorithm for the estimation of scalar distortions to the more general case turns out to be nontrivial: The wavelet transform we use is not covariant with respect to arbitrary distortions, in particular not with respect to the distortion matrices one encounters in the shape from texture problem. However, due to the particular structure of the distortion matrices, there might still be a way to adapt the algorithm, by considering the behaviour of the wavelet coefficients along different directions separately. We intend to further investigate this promising generalisation.

An alternative would consist in simply "building in" more distortions into the wavelet transform, which is the approach taken by Clerc and Mallat in Ref. 3, by introducing the so-called "warplet transform". This transform, which maps an image to a function of six real variables, is motivated from a conceptual point of view rather than a practical one, since the actual calculation of a full warplet transform of an image would be very time-consuming. For our algorithm the warplet transform is simply too big.

The paper is structured as follows: Section 2 introduces and discusses the various texture models. Sections 3 and 4 give an overview of the imaging geometry. The behaviour of single frequency textures under perspective projection is described, and we show how the surface orientation can be calculated from the local scales. Section 5 contains a short overview of the CWT in two dimensions. In Section 6 we derive the algorithm for the estimation of local scales and discuss under which circumstances the algorithm can be expected to work properly. Section 7 is devoted to a summary of the complete algorithm for the estimation of slant and tilt. Section 8 finally shows and briefly discusses some examples of the performance of the algorithm on real and synthetic images.

2. THE TEXTURE MODEL

It is difficult to give a precise definition of texture. Roughly speaking, a texture is a feature of a (part of an) image, which exhibits variations at scales much smaller than the image scale. Texture may exhibit a regular pattern (the so-called *texels*), or some statistically regular behavior. The explicit or implicit assumption underlying most shape from texture methods is that a "regular" texture is deformed by the geometry under a perspective projection, i.e. by the shape of a 3D object.

However, the "regularity" of the texture may be understood in different ways. By regular we mean that in the considered textured region, the image has "some" invariance properties with respect to simple geometric transformations on the object's surface: generally translations (stationary textures) and/or rotations (isotropic textures). It is important to notice that in most cases (i.e. except for simple surfaces such as planes or spheroids*), such invariance properties make sense only locally. In this paper, we shall focus on the case of planar surfaces, keeping in mind the obvious extensions to surfaces which are "approximately planar", and therefore may be locally accurately approximated by planes. We shall also limit our analysis to the case of "stationary" textures (see Ref. 1 and references therein for discussion of the isotropic case).

Stationarity (as well as isotropy) may also be imposed in different ways. Namely, deterministic or stochastic assumptions may be made. In the first case simple stationary models involve sinusoids and linear combinations of sinusoids. In the second case, the model involves (second order) stationary processes.

- *Simple model: Sinusoids:* we assume models for the images of the form

$$g(\underline{x}) = A(\underline{x})\,\cos(\underline{k}.\underline{x} + \varphi)\ , \tag{1}$$

  where A is a (slowly varying) real amplitude function. Under the deformations induced by geometry (see below), such a model transforms into an image of the form

$$I(\underline{x}) = A_I(\underline{x})\,\cos\left(\underline{\Omega}(\underline{x}).\underline{x} + \varphi\right)\ , \tag{2}$$

  where $\underline{\Omega}$ may be interpreted as a local frequency, which carries information about the deformation map.

*More generally, all surfaces which may be written as homogeneous spaces for a motion group.

- *Linear combinations of sinusoids:* Deterministically regular textures are more adequately described by harmonic-type models of the form

$$g(\underline{x}) = \sum_{n=1}^{N} A_n(\underline{x}) \cos(\underline{k}_n.\underline{x} + \varphi_n) \ , \tag{3}$$

  where the A_n's are slowly varying amplitudes, and the φ_n's are phase shifts. Under perspective projection, such models transform into linear combinations of "locally periodic" components of the form (2). Such deformed models are called *locally harmonic*.

- *Stationary processes:* imposing stationarity in a stochastic sense amounts to consider a random field $\{X_{\underline{x}}, \underline{x} \in \mathbb{R}^2\}$ over some probability space $(\mathcal{A}, \mathcal{F}, \mathbb{P})$, and assume that its distribution (more precisely, the distributions of all finite random vectors $(X_{\underline{x}_1}, \dots X_{\underline{x}_n})$) is translation invariant. One generally concentrates on the subclass of second order random fields (i.e. random fields such that $\mathbb{E}\left\{|X_{\underline{x}}|^2\right\} < \infty$ for all $\underline{x}$), and only assumes second order stationarity: for all $\underline{x}, \underline{y} \in \mathbb{R}^2$, $\mathbb{E}\left\{X_{\underline{x}+\underline{u}}\right\} = \mathbb{E}\left\{X_{\underline{x}}\right\}$ and $\mathbb{E}\left\{X_{\underline{x}+\underline{u}}\overline{X}_{\underline{y}+\underline{u}}\right\} = \mathbb{E}\left\{X_{\underline{x}}\overline{X}_{\underline{y}}\right\}$ for all $\underline{u} \in \mathbb{R}^2$. It is well known that such second order stationary random fields admit a spectral representation. In particular, the covariance function C_X takes the form

$$C_X(\underline{x}) := \mathbb{E}\left\{X_{\underline{x}}\overline{X}_{\underline{0}}\right\} - |\mathbb{E}\left\{X_{\underline{0}}\right\}|^2 = \frac{1}{4\pi^2} \int_{\mathbb{R}^2} e^{i\underline{k}.\underline{x}} \, d\nu_X(\underline{k}) \ ,$$

  where ν_X is the spectral measure of X. Notice that the case when ν_X is discrete is the stochastic version of the harmonic models above.

  Under perspective projection, a stationary random field transforms into a – usually nonstationary – random process, and its variations contain information about the geometry of the projection.[4]

We shall here essentially focus on the harmonic models, and refer to Ref. 3 for an elaboration of the random situation.

3. THE PROJECTION MODEL

3.1. Perspective Projection Geometry

The general geometrical framework for the shape from texture model has been developed by J. Gårding.[1] We use here a simplified formalism, derived and used in Refs 5 and 6, adapted to the estimation of planar surfaces.

We consider a reference coordinate system (x_i, y_i, z) (with respect to a basis $(\underline{e}_1, \underline{e}_2, \underline{n})$), and the image plane is taken parallel to the xOy plane, at distance f (the *focal length*). $\underline{n}$ is therefore the normal to the image plane. The object surface S is assumed to be (or locally approximated by) a plane, and we denote by $\underline{N}$ the surface normal. Then we have

$$\Sigma = \{\underline{x}, \langle \underline{x}, \underline{n} \rangle = f\} \ , \qquad \text{and} \qquad S = \{\underline{x}, \langle \underline{x}, \underline{N} \rangle = z\} \ ,$$

for some fixed positive number z. The *slant angle* σ is defined as the angle between the image and the object normals:

$$\langle \underline{n}, \underline{N} \rangle = \cos\sigma \ . \tag{4}$$

The perspective projection $P : S \to \Sigma$, is given by

$$P\underline{x} = \frac{f}{\langle \underline{x}, \underline{n} \rangle} \, \underline{x} \ . \tag{5}$$

Let $\underline{t}$ denote the orthogonal projection of the object normal $\underline{N}$ onto the image plane Σ, and let $\underline{b}$ be the vector in the image plane, such that $(\underline{t}, \underline{b}, \underline{n})$ is an oriented orthonormal basis. $(\underline{t}, \underline{b})$ induces a coordinate system on

Σ, where the origin $\underline{0}_I$ is given by the point where the image normal meets the image plane. $\underline{t}$ is called the *tilt vector*, and the angle between $\underline{t}$ and the unit vector $\underline{e}_1$ is the *tilt angle*, i.e.

$$\underline{t} = \underline{e}_1 \cos\tau + \underline{e}_2 \sin\tau \ . \tag{6}$$

Let dP^{-1} denote the Jacobian of $P^{-1} : \Sigma \to S$ at the image origin and set

$$\underline{T} = \frac{dP^{-1}\underline{t}}{\|dP^{-1}\underline{t}\|} \ , \qquad \text{and} \qquad \underline{B} = \frac{dP^{-1}\underline{b}}{\|dP^{-1}\underline{b}\|} \ .$$

$(\underline{t}, \underline{b}, \underline{n})$ and $(\underline{T}, \underline{B}, \underline{N})$ are orthonormal bases. Picking $P^{-1}(\underline{0}_I)$ as origin, we obtain thus a coordinate system on S. Given a point of coordinates (x, y) (wrt $(\underline{t}, \underline{b})$) in the image plane, its inverse image by P in the object plane has the $(\underline{T}, \underline{B})$-coordinates

$$\begin{pmatrix} x_S \\ y_S \end{pmatrix} = \frac{z}{\cos\sigma} \frac{1}{f + x\tan\sigma} \begin{pmatrix} 1 & 0 \\ 0 & \cos\sigma \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \ . \tag{7}$$

Notice the role of the focal length f; notice also that the renormalized focal length $f' = f/\tan\sigma$ introduces itself naturally. When f' is large enough, the deformation map may be reasonably considered linear, as long as the coordinates x and y in the image plane are small compared to f'. This is the starting point of the approximations we develop below.

Remarks 1. Our analysis puts a certain emphasis on the correct choice of image origin, somewhat contrary to the attitude taken by other authors.[1,2,4] The latter assume the image to be given on a sphere, which has the advantage that the local geometry of the imaging process does not depend on the position in the image. The transfer from image sphere to image plane is then performed by use of the Gaze transform,[1] or simply omitted, which is justified whenever the focal length is large enough.
Since we consider only a single planar surface, we can afford to be more precise. Note however that whenever there is more than one surface normal to be estimated, the following arguments and formulas will have to be adjusted to the position in the image for which the surface normal is to be calculated. This might complicate the calculations considerably.

3.2. Perspective Projection of Harmonic Textures

We now limit our investigations to the special case of harmonic textures, as in (1) and (3), and we examine the effect of perspective-induced distortion on such models. Consider the case of a periodic texture pattern on the object plane, and denote by $\underline{k}_S = \begin{pmatrix} k_1 \\ k_2 \end{pmatrix}$ the corresponding vector of frequencies, expressed in the $(\underline{T}, \underline{B})$ basis. Then the observed vector of frequencies in the $(\underline{t}, \underline{b})$ basis reads

$$\underline{k} = \frac{z}{f + x\tan\sigma} \begin{pmatrix} k_1 \sec\sigma \\ k_2 \end{pmatrix}$$

The phase of the distorted image is

$$\phi(x, y) = \frac{z}{f + x\tan\sigma} \ (k_1\, x \sec\sigma + k_2\, y)$$

and its gradient has coordinates

$$\partial_x\phi(x, y) = \frac{z\cos\sigma\,(-k_1\, f + k_2 y \sin\sigma)}{-\,(f\cos\sigma)^2 - 2\, f x \cos\sigma \sin\sigma - x^2 + (x\cos\sigma)^2} \ ; \qquad \partial_y\phi(x, y) = \frac{z k_2}{f + x\tan\sigma} \tag{8}$$

Introducing reduced variables $u = \frac{x\tan\sigma}{f}$, $v = \frac{y\tan\sigma}{f}$, $\rho = \frac{k_2\cos\sigma}{k_1}$ and $\gamma = \frac{f\cos\sigma}{z k_1}$, we obtain the following closed form

$$|\nabla\phi(x, y)|^{-1} = |\gamma| \, \frac{(u+1)^2}{\rho\sqrt{(u+1)^2 \,+\, (v - 1/\rho)^2}} \ . \tag{9}$$

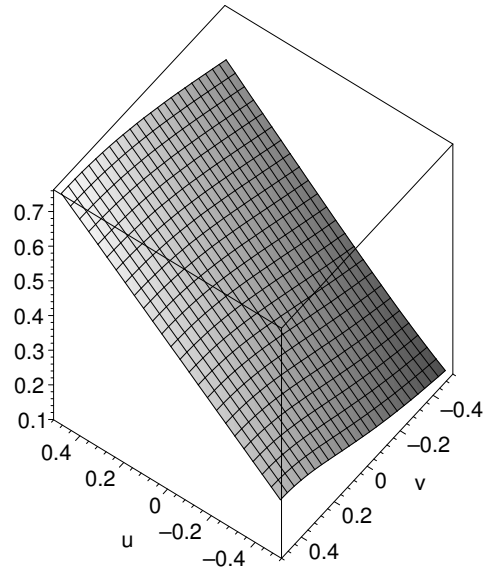

Figure 3. Example of the behavior of the local scales given in (10), with $\gamma = 1$ and $\rho = 2$.

For values of u and v sufficiently small (and far away from -1 and $1/\rho$ respectively), the behavior of the scale $|\nabla\phi(x,y)|^{-1}$ turns out to be fairly close to linear (see **Figure 3** for an example), which suggests to use linear approximation. The introduced variables once again show that the renormalized focal length f' is a natural parameter in this problem.

A first order Taylor development of (9) yields the following expression

$$|\nabla\phi(x,y)|^{-1} \approx \frac{|\gamma|}{\sqrt{\rho^2+1}} \left(1 + \frac{(\rho^2+2)}{(\rho^2+1)}\, u + \frac{\rho}{(\rho^2+1)}\, v\right) \tag{10}$$

As argued in Ref. 6, the "perceived tilt" often corresponds to frequency changes in the tilt direction, i.e. to the situation $v \approx 0$. The approximation above allows us to make the necessary corrections in the case $v \neq 0$, as we shall see in the next section.

But first let us take a closer look at the angle $\alpha(\underline{x})$ of $\Omega(\underline{x})$ with $\underline{t}$. From (8) we obtain

$$\tan\alpha(\underline{x}) = -\frac{(u+1)\,\rho}{\rho\, v - 1} \quad .$$

Hence, using the relation $\tan(\theta_1 - \theta_2) = \dfrac{\tan\theta_1 - \tan\theta_2}{1 + \tan\theta_1 \tan\theta_2}$, we find that

$$\alpha(\underline{x}) - \alpha(\underline{0}) = \operatorname{atan}\left(\frac{u + \rho v}{v - \rho u - \rho - \rho^{-1}}\right) \quad , \tag{11}$$

and since u, v are small (and $\rho + \rho^{-1} \geq 2$), we see that we may safely assume the frequency angle to be constant.

4. SURFACE ORIENTATION FROM LOCAL SCALES

Let us now explain how the various formulas derived in the last section can be used for the estimation of slant and tilt. Assume for a moment that we already obtained estimates of $|\nabla\phi(x_i,y_i)|^{-1}$, at various positions (x_i, y_i)

around the image origin. Here the coordinates relate to the basis $(\underline{e}_1, \underline{e}_2)$, which is usually different from $(\underline{t}, \underline{b})$. Under the first order assumption and after suitable normalization, we have that

$$|\nabla\phi(x_i, y_i)|^{-1} \approx 1 + Ax_i + By_i \quad ,$$

where the vector $(A, B)^t$ is (up to normalization) an estimate of the gradient $\underline{t}_0$ of $|\nabla\phi|^{-1}$ at the origin. However, equation (10) from the last section is not directly applicable, since it uses the $(\underline{t}, \underline{b})$-coordinates, which are not available at this stage. One way around this consists in estimating an additional datum, which is the direction of local frequency in the image origin: Let Ω_0 denote the phase at $\underline{0}$, and consider the angle β between Ω_0 and $\underline{t}_0$. It turns out that there is a way to compute tilt from Ω_0 and β, and then slant from tilt and the length of $\underline{t}_0$.

To see this, let us compute the various quantities in the $(\underline{t}, \underline{b})$-basis: The angle α_0 between Ω_0 and $\underline{t}$ is given by

$$\tan\alpha_0 = \frac{k_2}{k_1 \sec\sigma} = \rho \quad .$$

On the other hand, if τ_0 denotes the angle between $\underline{t}_0$ and $\underline{t}$, we find by equation (10) that

$$\tan\tau_0 = \frac{\rho}{\rho^2+2} \quad , \qquad \text{and hence} \qquad \beta = \operatorname{atan}\left(\frac{\rho}{\rho^2+2}\right) - \operatorname{atan}\rho \quad .$$

Taking the tangent on both sides and simplifying, we obtain $\tan\beta = -\frac{\rho}{2}$, or

$$\alpha_0 = -\operatorname{atan}(2\tan\beta) \quad . \tag{12}$$

This yields α_0, and since we can estimate Ω_0 in arbitrary coordinates, we gain access to $\underline{t}$.

It remains to compute σ from the length of $\underline{t}_0$. Again from equation (10) we see that in tilt coordinates

$$\|\underline{t}_0\| = \frac{\tan\sigma}{f}\sqrt{1 + \frac{3}{\rho^2+1}} \quad , \tag{13}$$

which is readily solved for σ, knowledge of f provided. In any case we can estimate the normalized focal length.

5. CONTINUOUS 2D WAVELET TRANSFORM

The two-dimensional continuous wavelet transform (CWT for short) was introduced by R. Murenzi.[7] It is an extension of the one-dimensional transform, incorporating (two-dimensional) translations, scalar rescalings and rotations. We give here a brief account of its main properties.

5.1. Definition

Given $\psi \in L^2(\mathbb{R}^2)$, associate with it the family of shifted, scaled and rotated copies $\psi_{\underline{b}s\theta}$, defined by

$$\psi_{\underline{b}s\theta}(\underline{x}) = \frac{1}{s^2}\,\psi\left(r_{-\theta}.\frac{\underline{x}-\underline{b}}{s}\right) \quad , \tag{14}$$

where $r_{-\theta}$ denotes a rotation by the angle $-\theta$. Given any two-dimensional signal $f \in L^2(\mathbb{R}^2)$, its CWT is the function of 3 variables $\mathcal{W}_\psi f$ given by

$$\mathcal{W}_\psi f(\underline{b}, s, \theta) = \langle f, \psi_{\underline{b}s\theta}\rangle = \frac{1}{s^2}\int_{\mathbb{R}^2} f(\underline{x})\overline{\psi}\left(r_{-\theta}.\frac{\underline{x}-\underline{b}}{s}\right)\,d\underline{x} \ . \tag{15}$$

It is well known[7] that any $f \in L^2(\mathbb{R}^2)$ has a continuous expansion with respect to the family $\psi_{\underline{b}s\theta}, \underline{b} \in \mathbb{R}^2, s \in \mathbb{R}^+, \theta \in [0, 2\pi[$, provided ψ satisfies some admissibility condition (essentially stating that ψ is sufficiently oscillatory). Since we do not make use of such expansions, we shall not elaborate on this point here, and rather focus on the specific properties of the CWT, in particular the covariance properties (see subsection 5.2 below).

CWT has several important advantages over the multiresolution wavelet transform, which turn out to be crucial for application to shape from texture. The main advantage is the built-in covariance of CWT, which we discuss in more detail below. CWT is also easy to design, and the admissible wavelets are easily generated (see subsection 5.3 below) and much less constrained than multiresolution wavelets (for example, one may use wavelets ψ with exponential decay, such that $\hat{\psi}$ has exponential decay too). CWT also offers much freedom in the choice of scales: for the method we describe in this paper, scale redundancy is an essential requirement. Finally, the possibility of varying simultaneously and independently scale and rotation angle allows us to use CWT as band-pass filtering, and therefore to filter locally periodic components in distorted signals.

As an illustration, let us briefly describe the behavior of CWT on locally harmonic signals, of the form (2). Assuming that both the amplitude A and the local frequency Ω are slowly varying, we obtain directly from a first order Taylor expansion

$$\mathcal{W}_\psi I(\underline{b}, s, \theta) \approx \frac{1}{2} A_I(\underline{b})\, \overline{\hat{\psi}}(sr_{-\theta}\Omega(\underline{b}))e^{i(\Omega(\underline{b}).\underline{b}+\varphi)} \ .$$

Assuming that $\hat{\psi}$ has sharp localization near the frequency $\underline{k} = \underline{k}_0$ (like the Morlet wavelets, see below), the wavelet transform $\mathcal{W}_\psi I$ is localized near a "ridge surface" of equation $s = s(\underline{b}), \theta = \theta(\underline{b})$, with $s(\underline{b})r_{-\theta(\underline{b})}\Omega(\underline{b})) = \underline{k}_0$, i.e. $s(\underline{b})^{-1}$ and $\theta(\underline{b})$ are the polar coordinates of the local frequency vector associated with the image, with respect to an appropriate reference frame. Such properties have been used for shape from texture in Ref. 6.

Remarks 2. Let us mention for the record that the CWT may be extended to the case of random signals. More precisely, given a second order random process $\{X_{\underline{x}}, \underline{x} \in \mathbb{R}^2\}$, it may be shown (see Ref. 8 for the one-dimensional case; the two-dimensional situation may be handled in the same way) that under some mild assumptions on the wavelet ψ, the CWT $\mathcal{W}_\psi X$ is a second order random field; if in addition X is second order stationary, with spectral measure $d\nu_X(\underline{k})$, the random fields $\underline{b} \to \mathcal{W}_\psi X(\underline{b}, s, \theta)$ are second order stationary for all s, θ.

5.2. Covariance Properties

The main property for our purpose is the covariance with respect to translations, rotations, and rescalings. Namely, let $f \in L^2(\mathbb{R}^2)$, and let $g \in L^2(\mathbb{R}^2)$ be defined by

$$g(\underline{x}) = f\left(r_{-\theta_0}.\frac{\underline{x}-\underline{b}_0}{s_0}\right) \ ; \qquad \text{then} \qquad \mathcal{W}_\psi g(\underline{b}, s, \theta) = \mathcal{W}_\psi f\left(\frac{\underline{b}-\underline{b}_0}{s_0}, \frac{s}{s_0}, \theta-\theta_0\right) \ .$$

We shall mostly be concerned with the translation and rescaling covariance. Translation covariance implies that the CWT may be regarded as a continuous family of linear filters. Rescaling covariance will play an essential role in the shape from texture problem we study, and will be the key of the approximations involved in our methods.

Remarks 3. Let X be a second order stationary random field, with spectral measure $d\nu_X$. As a consequence of stationarity, $\mathbb{E}\left\{|\mathcal{W}_\psi X(\underline{b}, s, \theta)|^2\right\}$ does not depend on $\underline{b}$. In addition, $\mathbb{E}\left\{|\mathcal{W}_\psi X(\underline{b}, s, \theta)|^2\right\} = \int_{\mathbb{R}^2} |\hat{\psi}(sr_{-\theta}\underline{k})|^2\, d\nu_X(\underline{k})$.

5.3. Examples

Different classes of analyzing wavelets have been proposed and studied in the literature. The choice of an analyzing wavelet is often a trade-off between computational efficiency and time-frequency localization (Most often, wavelets which are well localized in both sides of the Fourier transform are not naturally associated with fast algorithms.) For our purpose, the main requirements are space and frequency localization. In the examples presented in this paper, we shall use the so-called two-dimensional Morlet wavelet, defined by

$$\psi(\underline{x}) = \exp\left\{-\frac{|\underline{x}|^2}{2\sigma^2}\right\} \exp\left\{i\underline{k}_0.\underline{x}\right\} + \epsilon(\underline{x}) \ , \tag{16}$$

where ϵ is a correction function, introduced so that ψ has vanishing integral[†]. σ and $\underline{k}_0$ are parameters, which control the bandwidth and the intrinsic frequency of the wavelet respectively. ψ has exponential decay, as well

[†] ϵ is generally numerically small, and may be neglected when $\underline{k}_0$ is large enough. In addition, the zero integral property is required for the inversion of the CWT, which we don't use here.

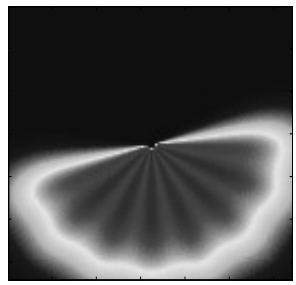

Figure 4. Filter bank corresponding to the wavelet transform.

as $\hat{\psi}$. The Fourier transform

$$\hat{\psi}(\underline{k}) = \frac{2\pi}{\sigma^2} \exp\left\{-\sigma^2 \frac{|\underline{k} - \underline{k}_0|^2}{2}\right\} + \hat{\epsilon}(\underline{k})$$

is essentially concentrated near $\underline{k} \approx \underline{k}_0$, so that the Fourier transforms of scaled and rotated copies $\widehat{\psi_{\underline{b}s\theta}}$ of ψ are essentially concentrated near $\underline{k} \approx r_\theta.\underline{k}_0/s$.

An example of Morlet wavelets (more precisely, the transfer functions of the filter bank) is displayed in **Fig.** 4. Alternatives include the two-dimensional Cauchy wavelets and generalizations (where the emphasis is put on the control of angular frequency localization), extensively studied in Ref. 9.

6. ESTIMATING LOCAL SCALES FROM WAVELET COEFFICIENTS

6.1. Wavelet Coefficients and Local Scale

After the basic properties of the CWT have been established, let us now describe how wavelet coefficients can be used to estimate local distortion. Assuming that the locally harmonic model applies to the image I, we obtain from the above formulas that

$$\mathcal{W}_\psi I(\underline{x}, s, \theta) \approx \frac{1}{2} A_I(\underline{x}) \overline{\hat{\psi}}(s r_{-\theta} \Omega(\underline{x})) e^{i(\Omega(\underline{x}).\underline{x} + \phi)} \quad ,$$

where A_I and Ω are slowly changing functions. From now on we assume A_I to be constant and drop it from our discussion.[‡] As we have seen in Section 3, we have for $\underline{x}$ close to the origin that

$$\Omega(\underline{x}) \approx \nabla\phi(x, y) \approx s_I(\underline{x})^{-1} \Omega(\underline{0}) \quad , \tag{17}$$

where

$$s_I(\underline{x}) = \frac{|\nabla\phi(0,0)|}{|\nabla\phi(x,y)|} \quad .$$

In order to get rid of the phase terms, we pass to the square modulus and obtain

$$|\mathcal{W}_\psi I(\underline{0}, s, \theta)|^2 \approx |\mathcal{W}_\psi I(\underline{x}, s_I(\underline{x}) s, \theta)|^2 \quad . \tag{18}$$

6.2. Estimation Procedure

We propose to use the following approach for the estimation of $s_I(\underline{x})$: We define $F_{\underline{x}}(s, \theta) := |\mathcal{W}_\psi I(\underline{x}, s, \theta)|^2$ and $F_{\underline{0}}$ analogously. First we pick a (compact) set $U \subset \mathbb{R}^+$ of "relevant" scales. Then we introduce a function Φ_U, by

$$\Phi_U(\tilde{s}) := \int_U \int_0^{2\pi} |F_{\underline{0}}(s, \theta) - F_{\underline{x}}(\tilde{s} s, \theta)| d\theta \frac{ds}{s} \quad .$$

[‡]This amounts to ignoring shading effects in the shape from texture problem. At least practically these effects can be easily dealt with by a simple normalization step, see Subsection 7.3.

Φ_U punishes the failure of the relation

$$F_{\underline{0}}(s,\theta) \approx F_{\underline{x}}(\tilde{s}s,\theta)$$

on $U \times [0, 2\pi]$. From relation (18) we expect that $\Phi_U(s_I(\underline{x}))$ is small, and we thus propose taking a minimizer of Φ_U as an estimate of $s_I(\underline{x})$.

If we restrict the search to a compact set V, we are guaranteed the existence of a minimizer in V (Φ_U is continuous). We shall not discuss uniqueness of the minimizer; the following discussion applies to any chosen minimizer. Let us stress that picking V amounts to making some a priori assumptions on the variations of local scale, i.e., mainly on the size of slant.

6.3. Precision of the estimation algorithm

On what properties of the wavelet coefficients does the performance of the estimation procedure depend? We claim that two aspects need to be considered:

- The accuracy, with which (18) holds;
- The behaviour of the wavelet coefficients on the set $U \times [0, 2\pi]$.

The relevance of the first point is obvious; this was the initial motivation for the algorithm. It clearly depends on the validity of the first order approximation on the (effective) supports of the wavelets involved. It is usually guaranteed by two facts: Above all the assumption that the distortion is reasonably well-behaved, i.e., the renormalized focal length is big. Also, by choosing the scales occurring in U, we have control over the supports of the wavelets. However the latter argument is only of limited use: As we shall see below, a certain range of scales is needed for our algorithm. Moreover, choosing arbitrarily small scales hardly makes sense when we are dealing with images of finite resolution.

Before we discuss the second point, let us simplify the argument by dropping the rotation parameter and considering $F_{\underline{x}}$ and $F_{\underline{0}}$ as functions of scale only. Accordingly Φ_U will only be defined by integration over scale. The adaptation of the following to the appropriate setting is somewhat lengthy, but more or less straightforward.

In order to get an idea of what kind of condition we are looking for, let us assume that the function $F_{\underline{x}}(\cdot)$ is constant on VU. Then Φ_U is constant on V, hence minimizing Φ_U does not give any information at all, regardless of the degree of precision with which (18) holds.

This example illustrates that an additional assumption is necessary. It seems to be desirable to pick U located around some kind of peak, i.e., an area where $F_{\underline{0}}$ is significantly larger than outside. Then one expects the estimation procedure to look for similar peaks of $F_{\underline{x}}$. The following result shows that this strategy works. It allows an estimate of the precision of the algorithm under quite general conditions, i.e., whenever the average of $F_{\underline{0}}$ inside U is bigger than twice the supremum outside (in a suitable neighborhood of U).

Theorem. Let $V = [r^{-1}, r]$, for some $r > 1$. Let $U = [t_1, t_2]$ be such that $t_2/t_1 \geq r^2$. Suppose that the precision of relation (18) is controlled by some constant $\mathcal{P}$, i.e., suppose that

$$|F_{\underline{0}}(s) - F_{\underline{x}}(s_I(\underline{x})s)| \leq \mathcal{P} \quad ,$$

for all $s \in V^2U = [t_1 r^{-2}, t_2 r^2]$. Assume moreover that

$$\mathcal{C}_U := \frac{1}{2}\,\frac{1}{\log(t_2/t1)} \int_{t_1}^{t_2} F_{\underline{0}}(s)\frac{ds}{s} - \sup\{F_{\underline{0}}(s) : s \in [t_1 r^{-1}, t_2] \cup [t_2, t_2 r]\} > 0 \quad .$$

If $s_I(\underline{x}) \in V$, then a minimizer $\tilde{s}_0$ of Φ_U in V fulfills

$$|\log(s_I(\underline{x})^{-1}\tilde{s}_0)| \leq \frac{2\mathcal{P}\log(t_2/t_1)}{\mathcal{C}_U} \quad .$$

Sketch of Proof. We start out the proof by introducing the mapping

$$\widetilde{\Phi}_U : \tilde{s} \mapsto \int_{t_1}^{t_2} |F_{\underline{0}}(\tilde{s}s) - F_{\underline{0}}(s)| \frac{ds}{s} \quad ,$$

which does not depend on $F_{\underline{x}}$. The key observation is that since $F_{\underline{x}}$ is approximately a shifted (in scale) copy of $F_{\underline{0}}$, Φ_U is approximately a shifted copy of $\widetilde{\Phi}_U$. More precisely, we have the estimate

$$\begin{aligned} \left| \Phi_U(\tilde{s}) - \widetilde{\Phi}_U(s_I(\underline{x})^{-1}\tilde{s}) \right| &= \left| \int_{t_1}^{t_2} |F_{\underline{0}}(s) - F_{\underline{x}}(\tilde{s}s)| - |F_{\underline{0}}(s_I(\underline{x})^{-1}\tilde{s}s) - F_{\underline{0}}(s)| \frac{ds}{s} \right| \\ &\leq \int_{t_1}^{t_2} |F_{\underline{0}}(s_I(\underline{x})^{-1}\tilde{s}s) - F_{\underline{x}}(\tilde{s}s)| \frac{ds}{s} \leq \mathcal{P} \log(t_2/t_1) \quad , \end{aligned}$$

where we applied the assumption on $\mathcal{P}$ (note that $s_I(\underline{x})^{-1}\tilde{s}s \in V^2U$). Moreover, if $\tilde{s}_0$ is a minimizer,

$$\Phi_U(\tilde{s}_0) \leq \Phi_U(s_I(\underline{x})) \leq \mathcal{P} \log(t_2/t_1) \quad ,$$

again by the assumption on $\mathcal{P}$, whence finally

$$\widetilde{\Phi}_U(s_I(\underline{x})^{-1}\tilde{s}_0) \leq 2\mathcal{P} \log(t_2/t_1) \quad .$$

On the other hand, under the assumption $\mathcal{C}_U > 0$, a purely measure-theoretic argument (see Ref. 10) allows the lower estimate

$$\widetilde{\Phi}_U(s_I(\underline{x})^{-1}\tilde{s}_0) \geq \mathcal{C}_U |\log(s_I(\underline{x})^{-1}\tilde{s}_0)| \quad ,$$

which finishes the proof. □

On an informal level, the theorem allows us to rather easily judge or predict the performance of the algorithm on a given image. The first order model will be accurate whenever the local scales do not change too rapidly. Hence, by simply looking at the image, we usually know when to expect a small $\mathcal{P}$ as required in the theorem. Moreover, by graphically displaying the wavelet coefficients $F_{\underline{0}}$, one can verify whether they are well localized inside the set U.

Whether the theorem can be useful for precise error estimates is another matter. Theoretically, the estimation of $\mathcal{P}$ poses no problem (just apply Taylor's formula several times), but whether this can be done in the form of a concrete a priori estimate for a given image is not clear to us. The situation for $\mathcal{C}_U$ is a bit different: We are free in choosing U, and one might consider taking the candidate U for which $\mathcal{C}_U$ – which is easy to compute – is maximal.

6.4. Generalizations

The extraction of local scales and the subsequent estimation of the surface orientation was already proposed for images with a single frequency component in Ref. 6. In this case, computing the ridges seems to be the most natural approach to the extraction of local scales. However, we saw that the behaviour of local scales only depends on the frequency angle and not on the length of the original frequency vector, and that implies that the local scale model is appropriate whenever there is a predominant frequency **direction**. This covers, for instance, the case of several distinct harmonic components pointing the same way. Now, for this more general model, minimizing Φ_U may well be simpler: There might no longer be a distinct ridge, and if there are several ridges to be found, separating them is not always trivial. By comparison, the notion of concentration used in the theorem is more flexible.

Moreover the estimation procedure can be adapted to accommodate more general distortions: Whenever there exist distinct directions in the pattern, the distortion effects are no longer well described by a simple scaling. In the general case [§], $s_I(x)$ has to be replaced by a distortion matrix $D_I(x)$ (essentially the jacobian of P^{-1}, normalized to give the identity at the image origin). Assuming that $D_I(x)$ is well approximated by some

[§] Details concerning the following discussion can be found in Refs. 1, 3, 4.

$s_I(\underline{x})r_{\theta_I(\underline{x})}$, we could once again use the covariance property of the wavelet transform to obtain a relation of the form

$$|\mathcal{W}_\psi I(\underline{0}, s, \theta)|^2 \approx |\mathcal{W}_\psi I(\underline{x}, s_I(\underline{x})s, \theta_I(\underline{x}) + \theta)|^2 \quad ,$$

which could be exploited by an algorithm similar to the one we introduced. (The search would then run through both scales and rotations.) Unfortunately the distortion matrices arising in the shape from texture problem cannot be well described by a scaling and a rotation. However, the particular structure of the D_I suggests a different procedure, which we intend to study further: Estimate scalings for each direction separately (using the above algorithm), and from these estimate the eigenvectors and -values of $D_I(\underline{x})$. The estimation of surface orientation from these matrices could, for instance, be attained by a suitable adaptation of techniques developed in Ref. 11.

Finally let us notice that a side effect of using the wavelet coefficient square modulus for the algorithm is that it allows one to cover stochastic texture models as well. For if g is a stationary process, the wavelet coefficients will be mean zero random variables, and instead of the coefficients themselves (or their expectation) it is their variance which displays the interesting behaviour with respect to distortions (see Ref. 3 for a more detailed exposition). In practice, where we are given a single realization of a process, estimating wavelet coefficient variances is performed by smoothing the coefficient square modulus, which may be justified under suitable ergodicity assumptions.[4]

7. IMPLEMENTATION DETAILS

Given a textured image with emergent frequency direction, the algorithm we propose consists of the following steps:

1. Compute the wavelet coefficient square modulus $F(\underline{x}, r_j, \theta_k)$, for a chosen set $r_0, \ldots, r_{n_s}$ of scales and a chosen set $\theta_1, \ldots, \theta_{n_a}$ of angles.

2. (Optional) Smooth the F coefficients.

3. Pick an image patch Q, usually centered around the point $\underline{x}_0$ where the image normal meets the image plane, (i.e., the origin of the coordinate system used in Section 3). For each $\underline{x}_i = (x_i, y_i) \in Q$, estimate the local scale $s_i = s_I(\underline{x}_i)$, using the algorithm from Section 6.

4. Fit a plane to the points $(\underline{x}_i, s_i)$ in three-space. Hence the local scales are approximated (in the image coordinate system centered around $\underline{x}_0$) by $s_i \approx 1 + Ax_i + By_i$. Then the vector (A, B) is an estimate of the gradient of the local scales in the origin, so the angle between (A, B) and the first image coordinate is the first tilt estimate τ_0.

5. Estimate the direction of local frequency in the origin $\underline{x}_0$ from the coefficients $F(\underline{x}_0, \cdot, \cdot)$. This gives the angle β between projected frequency and first tilt estimate.

6. From β and τ_0 compute true tilt τ. From τ and the length of the vector (A, B) compute slant.

7.1. Computing Wavelet Coefficients

The continuous wavelet transform can be viewed as the output of a filterbank applied to the image, and it is a standard technique to use FFT for the implementation of the filterbank. This approach is simple to implement and allows arbitrary resolution in scales and angles, which is crucial for our application.

We discretize the dilations in a standard way: Pick a smallest scale r_0 and fix the number n_o of octaves, as well as the number n_v of voices per octave. We then let $r_j := 2^{j/n_v}$, where j ranges from 0 to $n_s = n_v n_o$. The scales are thus logarithmically equidistant, which is the natural discretization for integration against the measure $\frac{ds}{s}$.

In our implementation we used the Morlet wavelet. Since the images are real-valued, we only consider angles between 0 and 180 degrees; this gives $\theta_k = 180 * (k-1)/n_a$, with k going from 0 to $n_a - 1$.

It is clear that the resolution of angles and scales affects the performance of the shape-from-texture algorithm: The resolution in scale has to be large enough, so as to be sensible to the distortions arising in the image. The estimation procedure in Step 4 inherits the resolution from the scales used for calculating the wavelet transform: The estimated local scales will be integer powers of $2^{1/n_v}$. So one should not pick n_v too small. Similar reasoning

applies to the angle: The number of angles directly influences the precision of the estimation of the frequency angle in Step 5. In fine-tuning the wavelet transform one should see to it that the Morlet wavelet parameters (in particular, its concentration in space or frequency) are adjusted to the discretization. What is needed is a good tiling of the frequency plane, in the sense that the whole plane is covered (otherwise the wavelet coefficients might fail to catch important frequency components of the image), but the covering should not be too redundant, so as to guarantee a sharp localization of energy.

7.2. Smoothing

As we mentioned in the last section, smoothing the coefficients can be considered as a means of dealing with stochastic texture models: In this case we have to estimate the expectation of the wavelet coefficient square modulus from a single realization of the process. Using (possibly weighted) moving averages for this task can be justified by ergodicity arguments. This point has been investigated extensively in Ref. 4.

7.3. Estimating Local Distortion

Here we implement a discrete version of the minimization procedure described in section 6. We used a straightforward approach, approximating integrals by Riemann sums (legitimate since wavelet transforms are smooth functions.)

To be more precise, let $F_0(j,k) = |\mathcal{W}_\psi I(\underline{0}, r_j, \theta_k)|^2$ and $F_i(j,k) = |\mathcal{W}_\psi I(\underline{x}_i, r_j, \theta_k)|^2$. The set V of possible candidates is assumed to cover two octaves, i.e., $V = [1/2, 2]$. We further restrict the search by considering only powers of $2^{1/n_v}$; this way we avoid interpolating coefficients. For U we pick the maximal possible range, that is, $U = [2r_0, r_{n_s}/2]$. In short, we compute the function

$$\Phi(\ell) = \sum_{j=n_v}^{n-n_v} \sum_{k=1}^{n_a} |F_0(j,k) - F_i(j+\ell,k)| \ , \qquad \text{for} \quad \ell = -n_v, \ldots, n_v \quad ,$$

and for the local scale we obtain $s_i = 2^{\ell_0/n_v}$, where ℓ_0 is the minimizer of Φ.

The evaluation of Φ might be modified to deal with the amplitude function A_I, responsible for modelling shading effects. A common and necessary assumption for the shape from texture problem is the so-called "scale separation principle",[1,2] stating that the variations due to shading effects happen on a much larger scale than the variations of the texture. Hence, at least for reasonably small scales, A_I can be assumed to enter the wavelet coefficients F_0 and F_i in a multiplicative way. Thus normalizing the coefficients before computing Φ cancels A_I.

The choices of Q, V and n_v limit the range of distortions which may be detected. If Q is chosen too small and if slant is small, then the distortion is below the threshold $2^{1/n_v}$, for every $\underline{x}_i \in Q$, hence it goes unnoticed. If, on the other hand, slant is big, and Q is chosen too big, then, outside a certain smaller set $\widetilde{Q} \subset Q$, the local distortion will be outside the prescribed range of candidates. Hence the estimate is not reliable in this case. A way around this might consist in choosing Q adaptively: Start out with a very small Q and increase it until the detected local scales cover a certain range of values.

7.4. Fitting Local Scales

Here we use simple linear regression. It might be useful to consider something more robust to deal with outliers.

7.5. Estimating Local Frequency Direction

The estimation of local frequency from wavelet or Gabor coefficients is by now a standard procedure. We have chosen a rather simple method: Given $F_0(j,\ell)$, we let $G(\ell) := \sum_j F_0(j,\ell)$, which is a measure of energy concentration in the direction $\alpha_\psi + \theta_\ell$. Here α_ψ is the center frequency of the wavelet ψ; for the Morlet wavelet α_ψ is the angle belonging to the vector $\underline{k}_0$. The assumption that the texture has a dominant frequency component implies that G has a single maximum at some position ℓ_0, and in the worst case a similarly big value at one of the neighboring angles. We estimate the frequency angle by taking the weighted sum of the corresponding angle and its neighbours,

$$\alpha_0 = \alpha_\psi + \frac{\sum_{i=-1}^{1} \theta_{k_0+i} G(\ell_0+i)}{\sum_{i=-1}^{1} G(\ell_0+i)} \quad ,$$

where we use $G(0) = G(n_a)$ and $G(n_a+1) = G(1)$, when needed.

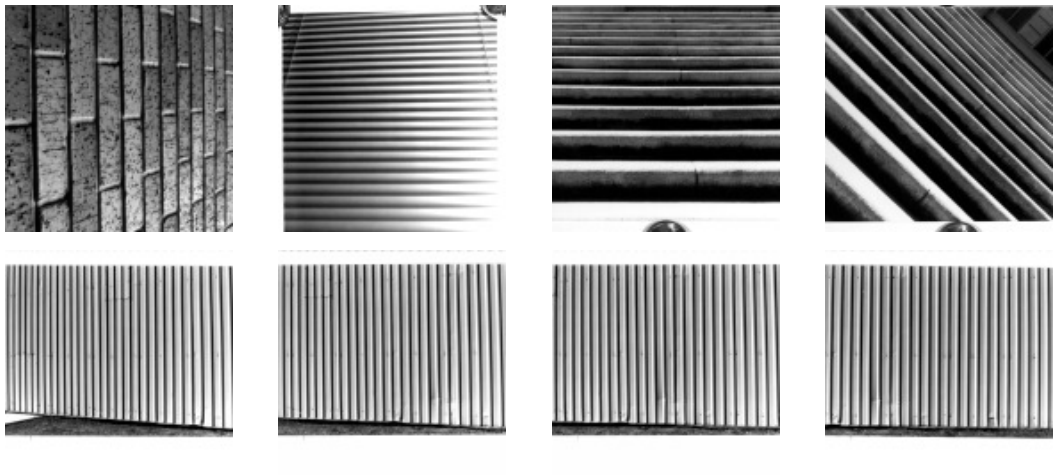

Figure 5. Images from Super and Bovik: brick wall (*brick*), venetian blind (*blind*), steps-1 (*steps1*), steps-2 (*steps2*), wall-4 (*wall4*),wall-3 (*wall3*), wall-2 (*wall2*),wall-1 (*wall1*).

7.6. From Local Scales to Slant and Tilt

This is a straightforward application of equations (12) and (13): Letting $\beta := \tau_0 - \alpha_0$, we obtain

$$\rho = -2\tan\beta \quad , \qquad \tau = \alpha_0 + \mathrm{atan}\rho \quad , \qquad \sigma = \mathrm{atan}\left(f\sqrt{\frac{A^2+B^2}{1+\frac{3}{\rho^2+1}}}\right) \quad .$$

8. TEST RESULTS AND CONCLUSIONS

The algorithm described in the previous section has been tested on several sets of synthetic and real images. We present below a set of results which illustrate the behavior of the proposed method in a number of situations of interest. More detailed and systematic results will be presented elsewhere.

We shall mainly focus on a set of images considered first by Super and Bovik[5, 12] and studied also by Hwang *et al.*.[6] For those images, the tilt direction coincides with the direction of frequency changes, so that the correction brought by Eq. (12) does not improve the resolution. In fact, in all cases the estimates of the frequency direction (i.e., the result of Step **5** of the algorithm) were quite precise. As a consequence, the angle β between the first tilt estimate and the estimated frequency direction is essentially just the error committed in Steps **3** and **4**, and (12) further increases this error (for small β, it more or less doubles it).

Nevertheless, we obtain in such a case results of a precision comparable to those of Refs. 5, 6, 12. We display in **Table 1** the results (slant $\hat{\sigma}$ and tilt $\hat{\tau}$) obtained with our method, and compare them with the results of Super and Bovik ($\hat{\sigma}_{SB}$ and $\hat{\tau}_{SB}$), and Hwang *et al.* ($\hat{\sigma}_{HLC}$ and $\hat{\tau}_{HLC}$; only one of the two slant estimates given in Ref. 6 is reproduced here). Let us stress that the results were obtained by processing the images directly, without any image-specific pre-processing or post-processing. The parameter values for the results in **Table 1** are as follows: We used a two-dimensional Morlet wavelet, with 8 different directions (8 angles), and 40 scales: 5 octaves (i.e. powers of two) and 8 voices (scales) per octave, i.e. all together 320 different filters. The corresponding filter bank is displayed in **Fig. 4**. We calculated the local scales for a rectangle around the image center, using normalization and smoothing by moving averages. The venetian blind with its rather strong variations in shading is a good illustration of the effect of normalization: Without this step, the algorithm gave a tilt estimate of 173.1 degrees, which is dramatically wrong. The case of the brick wall is similar, here the error increased from 2.37 to 18.88 degrees.

The results are of quality comparable with those of Refs. 5, 6, 12, even though we may notice a consistent underestimation of slant. The results for *wall1* are consistent (the slant being zero), while the tilt estimate for *wall2* is quite poor. In that case, the slant is rather small, so there is hardly any variation in local scale. The

image	true τ	$\hat{\tau}$	$\hat{\tau}_{SB}$	$\hat{\tau}_{HLC}$	true σ	$\hat{\sigma}$	$\hat{\sigma}_{SB}$	$\hat{\sigma}_{HLC}$
brick	0	2.37	1.1	-1.61	55	48.57	52.3	53.34
blind	90	97.26	90.9	92.98	40	34.79	32.1	41.81
steps1	90	90.70	88.6	91.45	70	67.71	68	68.61
steps2	45	44.87	45.1	46.02	70	66.25	71	68.05
wall4	180	177.60	179.5	179.51	30	26.45	26.4	25
wall3	180	176.48	178.9	-178.05	20	18.58	16.7	19.47
wall2	180	149.00	177.3	-174.35	10	6.97	9.7	11.51
wall1	-	-	-	-	0	-	1.0	0

Table 1. Results for slant and tilt estimates from Super and Bovik images.

image	true τ	$\hat{\tau}$	$\hat{\tau}_{SB}$	$\hat{\tau}_{HLC}$	true σ	$\hat{\sigma}$	$\hat{\sigma}_{SB}$	$\hat{\sigma}_{HLC}$
wall4	180	177.15	179.5	179.51	30	26.55	26.4	25
wall3	180	177.27	178.9	-178.05	20	17.05	16.7	19.47
wall2	180	170.38	177.3	-174.35	10	9.89	9.7	11.51

Table 2. Results for slant and tilt estimates from Super and Bovik images, obtained with better frequency resolution.

results may be improved by increasing the scale resolution, as shown in **Table 2**. In the latter case, 16 voices per octave were used instead of 8, with a better frequency resolution for the filters (the effective support of the wavelets in the Fourier domain -in the radial direction- being halved.)

The second set consists of (extremely simple) artificial 2D signals, which aim at demonstrating the effect of the tilt correction (with respect to the method of Ref. 6). $\hat{\tau}$ and $\hat{\sigma}$ are respectively the tilt and slant estimates obtained by our method, and $\hat{\tau}_{HLC}$ is the estimate without correction, i.e. in the spirit of Ref. 6. The results, given in **Table 3** show clearly the improvement introduced by our approach, which is significant as soon as the tilt direction is not close to the frequency change direction. More detailed results will be presented elsewhere.

The approach developed in the present paper, limited to the case of planar surfaces, may be extended to non-planar surfaces without deep modifications, as long as the curvatures and the torsion are small enough so that planar approximations may be used (see nevertheless the discussion at the end of section 3.1.) In the general case, variations of the shape operator have to be taken into account, which makes the situation much more tricky, as may be seen in Refs 2 and 4.

ACKNOWLEDGMENTS

Hartmut Führ would like to thank the Laboratoire d'Analyse, Topologie et Probabilités in Marseille, and in particular the group SYSDYS, for their hospitality. He acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG) under the contract Fu 402/1. W.L. Hwang and B. Torrésani acknowledge support from Centre National de la Recherche Scientifique (CNRS, France) and National Science Council (NSC, Taiwan).

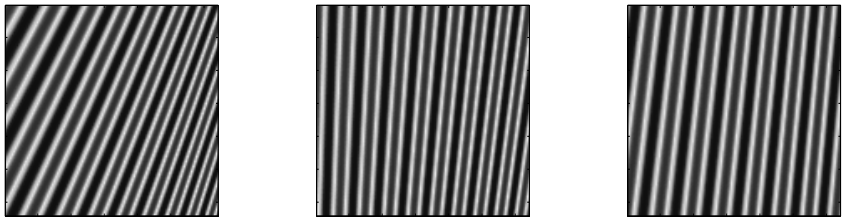

Figure 6. Artificial deformed sinusoidal images(F1, F2 and F3).

image	true τ	$\hat{\tau}$	$\hat{\tau}_{HLC}$	true σ	$\hat{\sigma}$
F1	0	-1.00	10.38	65	64.81
F2	40	41.3	24.78	45	46.96
F3	-10	-10.95	-3.09	30	28.68

Table 3. Results for slant and tilt estimates from artificial images.

REFERENCES

1. J. Gårding, "Shape from texture for smooth curved surfaces in perspective projection", *J. Math. Imaging and Vision* **2**, pp. 329-352, 1992.
2. J.A. Malik and R. Rosenholz, "Computing local surface orientation and shape from texture for curved surfaces", *Int. J. Computer Vision* **23**, pp. 149-168, 1997.
3. M. Clerc and S. Mallat, "Shape from texture and shading with wavelets", in *Dynamical Systems, Control, Coding, Computer Vision*, Progress in Systems and Control Theory, vol. 25, pp. 393-417, Birkhäuser 1999.
4. M. Clerc, *Analyse par ondelettes de processus localement dilatés, et application au gradient de texture*, Thesis (1999).
5. B.J. Super and A.C. Bovik, "Planar surface orientation from texture spatial frequencies", *Pattern Recognit.* **28**, pp. 729-743,1995.
6. W.-L. Hwang, C.-S. Lu and P.-C. Chung, "Shape from texture: Estimation of planar surface orientation through the ridge surfaces of continuous wavelet transform", *IEEE Trans. IP* **7**, pp. 773-780, 1998.
7. R. Murenzi, *Ondelettes Multidimensionnelles et applications à l'analyse d'images*, PhD Thesis, Louvain la Neuve, Belgium (in french).
8. S. Cambanis and C. Houdré, "On the Continuous Wavelet Transform of Second Order Processes. *IEEE Trans. Inf. Th.,* **41** (1995), 628–642
9. J.P. Antoine, R. Murenzi and P. Vandergheynst, Directional wavelets revisited: Cauchy wavelets and symmetry detection in patterns, *Appl. and Comp. Harm. Anal.* **6** (1999), 314-345
10. H. Führ, "Estimating local distortion with wavelets", Preprint.
11. J. Gårding, "Surface orientation and curvature from differential texture distortion", in *Proc. 5th International Conference on Computer Vision*, (Cambridge, MA), pp. 733–739, 1995.
12. B.J. Super and A.C. Bovik, "Shape from texture using local spectral moments", *IEEE Trans. PAMI* **17**, pp. 333-343, 1995.