SHAPE FROM TEXTURE FOR OMNIDIRECTIONAL IMAGES

L. Jacques[⋆,∗]*, E. De Vito*[⋆]*, L. Bagnato*[⋆]*, P. Vandergheynst*[⋆]

(⋆) Ecole Polytechnique Fédérale de Lausanne (EPFL, LTS2)
CH-1015 Lausanne, Switzerland

(∗) Université catholique de Louvain (UCL, TELE)
B-1348 Louvain-la-Neuve, Belgium

{laurent.jacques, eugenio.devito, luigi.bagnato, pierre.vandergheynst}@epfl.ch

ABSTRACT

In this paper, we describe a method to infer 3-D scene information using a single view captured by an omnidirectional camera. The proposed technique is inscribed in the so called "Shape from Texture" problem : if the textures hold by 3-D surfaces respect some a priori models, the deformation due to their projection in the image contains both local information about surface depth and orientation.

To estimate this deformation, we adapt the work of Gårding and Lindeberg to the case of spherical images processing. The planar multiscale procedure allowing the definition of precise texture descriptors is here replaced by a multiscale representation compatible with the compactness of the sphere. More precisely, the multiscale representation is obtained by filtering the data by dilated copies of a mother function. The spherical dilation introduced is the *gnomonic* dilation, a simple variation of the stereographic dilation due to Antoine and Vandergheynst. This dilation has a simple interpretation in terms of projective geometry. It fits precisely the transformation that the apparent omnidirectional image of an object follows when the distance of this object to the sensor changes. A spherical texture descriptor, close to a deformation tensor, is then defined thanks to the use of simple filters that act as smoothed differential operators on the data.

Results are provided in the analysis of a synthetic example to illustrate the capacity of the proposed method.

1. INTRODUCTION

Recovery of 3-D scenes from vision sensors is generally performed using "shape cues" from the processing of the recorded data. It is generally accepted that efficient geometry assessments are performed by multiple sensors setup, from stereoscopic systems emulating biological vision, to more numerous sensors set with complex geometric treatments.

It is however possible to estimate 3-D scene geometry from a single view if the observed objects surfaces are textured and if the textures respect some basic stationary properties. Indeed, in that case, their projections on the image induces deformations altering this property. Assessing this deformation gives then clues on surface distance and orientation.

For that purpose, Lindeberg and Gårding have for instance designed a multiscale texture descriptor for planar camera images. This one is related to a first order integro-differential operator close to a perfect 2×2 deformation tensor. The scale at which this operator gives a maximal response and its eigenvalues at this scale contain both information on the distance and on the surface orientation.

LJ is funded by the Belgian National Funds for Scientific Research (FRS-FNRS). This work was realized when LJ was a visiting postdoc at EPFL.

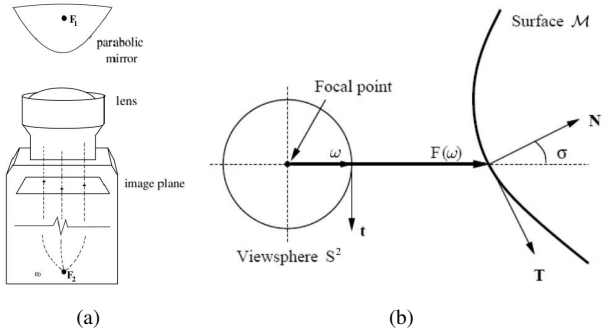


Figure 1: (a) A catadioptric imaging system. (b) Scene geometry [3].

There is an interest in developing the same theory for omnidirectional images such as those recorded by *catadioptric* cameras. In robotic systems for instance, the motion of an autonomous device can be driven by computer vision and by 3-D shape cues, such as texture deformations, obtained by image processing. However, planar cameras are severely limited in their field of view. In order to obtain an image of an entire scene and improve the localization awareness of the system, either multiple or rotating cameras must then be used [2].

An interesting alternative way to enhance the field of view is to use mirrors in conjunction with lenses. A parabolic catadioptric sensor is an example of such a system [1]. It is realized as a parabolic mirror which is placed in front of a camera approximating an orthographically projecting lens as depicted on Fig. 1(a). In such a case, the ray of light incident with the focus of the parabola is reflected to a ray of light parallel to the parabola axis. This construction is equivalent to a purely rotating perspective camera.

In [2], it is proved that there is an equivalence between any central catadioptric projection, like in the system above, and a composition of two conformal mappings on the sphere. The spherical coordinates of any incoming light rays can then be recovered through a simple inverse stereographic projection of the sensor images.

In the rest of this paper, we assume that this geometrical conversion has been realized and that the recorded image lives on the sphere S^2. In other words, we work with omnidirectional intensity image (gray level) $I : S^2 \to \mathbb{R}$, which associates to any point $\omega = (\theta, \varphi) \in S^2$ the intensity $I(\omega)$. Notice that $\theta \in [0, \pi]$ with $\theta = 0$ identified with the North

pole, and $\varphi \in [0, 2\pi)$.

2. SCENE DEFORMATION MODEL

We follow the mathematical framework defined by Gårding in [3] for the characterization of the projection from the real world scene on a 2-D image. Consider the *perspective mapping* from a surface $\mathscr{M} \subset \mathbb{R}^3$ in the 3-D scene to the unit view sphere S^2 with focal point (the center) assumed at origin (Fig. 1(b)). In each view direction $\omega \in S^2$ linked to the vector $\vec{p} \in \mathbb{R}^3$, we can define a local orthonormal system $(\vec{p}, \vec{t}, \vec{b})$. The vector $\vec{t}$ is the *tilt direction*, i.e. a vector parallel to the gradient of the distance $r(\omega) = r(\vec{p})$ from the focal point to $\mathscr{M}$ in direction ω, and $\vec{b} = \vec{p} \times \vec{t}$.

The perspective backprojection $F : S^2 \to \mathscr{M}$ is linked to the distance function r with $F(\omega) = r(\omega)\vec{p}$. We write hereafter $q = F(\omega) \in \mathscr{M}$. The derivative of F, i.e. its linear approximation $\mathrm{d}F_\omega$, is a mapping from the spherical tangent plane $T_\omega S^2$ on ω to the surface tangent plane $T_q\mathscr{M}$. We denote by $\vec{T}$ and $\vec{B}$ the normalized backprojection of $\vec{t}$ and $\vec{b}$ on $F(\omega)$. It is proved in [3] that the linear mapping $\mathrm{d}F_\omega$ from $T_\omega S^2$ in the base $(\vec{t}, \vec{b})$ to $T_q\mathscr{M}$ in the coordinates of $(\vec{T}, \vec{B})$ reads

$$\mathrm{d}F_\omega = \begin{pmatrix} r/\cos\sigma & 0 \\ 0 & r \end{pmatrix} = \begin{pmatrix} 1/m & 0 \\ 0 & 1/M \end{pmatrix}, \qquad (1)$$

where σ is the *slant* angle, i.e. the angle between the normal $\vec{N}$ of $\mathscr{M}$ on point $F(\omega)$ and the view direction $\vec{p}$.

The interpretation of the inverse eigenvalues $m < M$ of $\mathrm{d}F_\omega$ is the following : a unit circle drawn on $\mathscr{M}$ is mapped by F^{-1} on the view sphere S^2 to an ellipse of minor and major axis in direction $\vec{t}$ and $\vec{b}$ with lengths m and M respectively. In parallel, there is a relation between areas in $T_\omega S^2$ and in $T_q\mathscr{M}$: a unit area element in the later has an area of mM in $T_\omega S^2$. The mapping $\mathrm{d}F_\omega$ contains thus essential knowledge about the geometry of the surface $\mathscr{M}$.

3. MULTISCALE TEXTURE DESCRIPTOR

3.1 Planar Legacy

As explained in [3], the main principle of a "shape from texture" procedure is to estimate the mapping $\mathrm{d}F_\omega$ above from the only available information, i.e. from the recorded image intensities, using a priori surface texture models [4]. Lindeberg and Gårding introduced the (windowed) second moment matrix μ_I for flat gray level images $I(x)$ as the 2×2 tensor

$$\mu_I(y;t,s) \triangleq \int_{\mathbb{R}^2} (\nabla I(x,t))\,(\nabla I(x,t))^T\, w_s(y-x)\,\mathrm{d}x.$$

In that definition $I(x;t) = (g_t * I)(x)$ is a smoothed copy of I by a gaussian kernel $g_t(x) = (2\pi t^2)^{-1}\exp(-\|x\|^2/(2t^2))$ of *local scale* $t > 0$, ∇ is the usual 2-D gradient, and $w_s(x) = s^{-2}w(x/s)$ is a symmetrical window of *integration scale* $s > 0$. The explanation of the design and of the behavior of μ_I is the following. First, it is based on a scale-space extension of the original image, i.e. $I(x,t)$. This scale-space is a "probing" of the image content at different scale t. It has also the advantage of stabilizing the computation of the differential operator ∇, if computed for instance by finite difference.

Practically, the scale t, or *local scale*, has to be selected to the typical size of the basic pattern elements generating the texture. The integration scale has a different role. It gathers sufficient statistics of the texture gradient. In particular, it turns the rank-1 matrix $(\nabla I(x,t))\,(\nabla I(x,t))^T$ into a potentially rank-2 matrix μ_I. A possible choice for s is $s = \gamma t$, with γ set to 2 or 3, i.e. the texture is assumed made of elements (*texels*) gathered at γ times their typical size.

Interestingly, if the image domain undergoes some invertible linear transformation B, $I(x) = R(z)$ with $z = Bx$, it can be shown that $\mu_I(x;t,s) = B^T\mu_R(z;t,s)B$, with μ_R related both to the scale-space $R(z,t)$ of filter $g'_t(z) = (\det B)^{-1}g_t(B^{-1}z)$ and to the window $w'_s(z) = (\det B)^{-1}w_s(B^{-1}z)$. This simple fact creates a bridge between a texture on the surface and the projected texture on the planar sensor.

Indeed, if around $q = F(\omega)$ the object surface is locally well approximated by its tangent plane $T_q\mathscr{M}$, the texture pattern $\mathscr{T}(z)$ on $\mathscr{M} \simeq T_q\mathscr{M}$, with z expressed in the system $(\vec{T}, \vec{B})$, undergoes a linear transformation given by the inverse of $B_x = \mathrm{d}F_\omega\,\mathrm{d}G_x$, where $G : \mathbb{R}^2 \to S^2$ is the mapping projecting the planar sensor coordinates to the spherical domain, i.e. $\omega = G(x)$. If B_x is almost constant on small scales, μ_I is then conjugated by B_x to a similar second moment matrix $\mu_\mathscr{T}$ computed directly on $T_q\mathscr{M}$ from the painted texture pattern, i.e. $\mu_I(x;t,s) = B_x^T\mu_\mathscr{T}(z;t,s)B_x$.

3.2 Spherical extension using gnomonic dilation

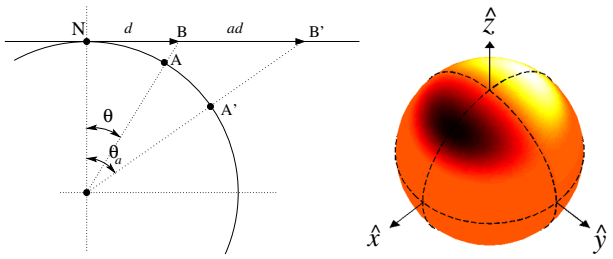
Since we are working here with image intensities I assumed to be directly available on the sphere, we need to define a specific spherical texture descriptor. Scale-space and gradient operator have thus to be replaced by their spherical counterpart.

Let us precise first some concepts proper to spherical data processing. The natural inner product between two functions u and v on the sphere reads $\langle u, v\rangle \triangleq \int_{S^2} u^*(\omega)v(\omega)\mathrm{d}\Omega$, where $(\cdot)^*$ denotes the complex conjugation and $\mathrm{d}\Omega = \sin\theta\,\mathrm{d}\theta\mathrm{d}\varphi$ is the usual rotation invariant measure on S^2. The L^2 norm of u is linked to this inner product by $\|u\|^2 \triangleq \langle u, u\rangle$. The set of finite energy functions on S^2, i.e. $L^2(S^2) = \{u : \|u\| < \infty\}$, is the natural Hilbert space of S^2 with the product $\langle\cdot,\cdot\rangle$.

Correlation between two functions u and v on the sphere can be defined by $(u \star v)(\rho) = \langle R_\rho u, v\rangle$ [5, 6], where R_ρ is the rotation operator described by the three Euler angles $\rho = (\varphi, \theta, \chi) = (\omega, \chi)$ thanks to the factorization $R_{\omega,\chi} = R^{\hat{y}}_\varphi R^{\hat{z}}_\theta R^{\hat{y}}_\chi$, where $R^{\hat{k}}_\nu$ is a rotation of angle ν around the axis $\hat{k} \in \mathbb{R}^3$.

Filter dilations on S^2 have to be introduced carefully due to the compactness of this space. We use a variant of the *stereographic dilation* [7, 6] that we call the *gnomonic dilation*. It is represented by the operator $D_t : L^2(S^2_N) \to L^2(S^2_N)$, with S^2_N the North hemisphere[1] of S^2. This later is defined by $[D_t u](\omega) = \lambda(t,\theta)^{1/2}\,u(\theta_{t^{-1}}, \varphi)$, with $\tan\theta_t = t\tan\theta$. Geometrically, this dilatation amounts to project u in the tangent plane to the North pole using the *gnomonic* (or central) projection Π, dilates there the projected function with the usual planar dilation d_t such that $d_t u(x) = t^{-1}u(x/t)$,

[1] This dilation does not act injectively on the whole sphere. This has no effect in our application since, as described after, the projective geometry of the 3-D scene is compatible with this hemispheric restriction.



(a) (b)

Figure 2: (a) Point A of latitude θ is sent to B in the tangent plane to the North pole N. Point B, at distance d from the origin, is sent to B' at distance ad by planar dilation of factor a. B' is then projected back to the sphere on A' at latitude θ_a. (b) The filter $\psi_x = \Pi^{-1}\frac{\partial}{\partial x}g$.

and finally projects back with Π^{-1} the result on the sphere, i.e. $D_t = \Pi^{-1} d_t \Pi$. Fig. 2(a) explains this dilation geometrically. The projection $\Pi : L^2(\mathbb{R}^2) \to L^2(S^2_N)$ is an isometry. i.e. it preserves the L^2 norm of the function after projection. Mathematically, it is such that $[\Pi u](r,\varphi) = (1+r^2)^{-3/4} u(\tan^{-1} r, \varphi)$ for $u \in L^2(S^2_N)$ and $r = \tan\theta$.

The factor $\lambda(t,\theta) \triangleq t^{-2}\big(\frac{1+\tan^2\theta}{1+t^{-2}\tan^2\theta}\big)^{3/2}$, named *cocycle*, allows the preservation of the L^2 norm, i.e. $\|D_t u\| = \|u\|$. A L^1 normalized variant $\tilde{D}_t$ of this dilation operator exists simply by replacing $\lambda^{1/2}$ by λ in the dilation definition so that $\|\tilde{D}_t u\|_1 = \|u\|_1 \triangleq \int_{S^2} |u(\omega)| \mathrm{d}\Omega$. From their definitions, for small angles, gnomonic dilations D_t and $\tilde{D}_t$ converge to their planar dilations counterparts (*Euclidean asymptotic property* [7]).

Thanks to all these concepts, we are ready now to propose the following texture descriptor,

$$\mu_I(\omega';t,s) \;=\; \int_{S^2} [JJ^T](\omega,t)\; [R_{\omega',0}\tilde{D}_s w](\omega)\; \mathrm{d}\Omega. \qquad (2)$$

The L^1 normalized dilation $\tilde{D}_s$ has been selected for the scaling of the window w to preserve the statistical meaning of the integration scale s. The choice of an L^2 normalized dilation in the definition of J is explained in Section 4.

The vector $J = (J_\theta, J_\varphi)^T$ is the vector formed by the components

$$\begin{array}{rcl} J_\theta(\omega,t) & \triangleq & \big((D_t\psi_x) \star I\big)(\omega,0) \\ J_\varphi(\omega,t) & \triangleq & \big((D_t\psi_y) \star I\big)(\omega,0). \end{array} \qquad (3)$$

The vector J is the spherical counterpart of the smoothed gradient vector ∇I of the planar texture descriptor. It is linked to two particular filters $\psi_x(\omega)$ and $\psi_y(\omega)$ corresponding to the gnomonic backprojection of the x and y derivatives of a Gaussian g in the tangent plane to the North pole [6], i.e. $\psi_x = \Pi^{-1}\frac{\partial}{\partial x}g$ and $\psi_y = \Pi^{-1}\frac{\partial}{\partial y}g$. The first filter is presented on Fig. 2(b). Since these filters are moved after dilation[2] by D_t on the point ω in (3), it is easy to prove that they are oriented according to the vectors $\vec{e}_\theta = (\sin\theta)^{-1}[(\vec{\omega}^T\hat{z})\,\vec{\omega} - \hat{z}]$ and $\vec{e}_\varphi = (\sin\theta)^{-1}\vec{\omega}\times\hat{z}$ respectively. Since $(\vec{e}_\theta, \vec{e}_\varphi)$ is an orthonormal basis of $T_\omega S^2$, J is thus a smoothed measure of the gradient of I on $T_\omega S^2$.

[2]This simple fact makes our analysis invariant under rotation of the data.

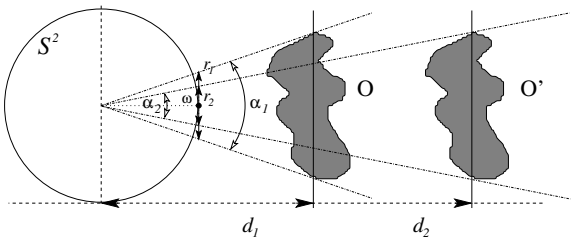


Figure 3: Gnomonic dilation and projective geometry.

To go further in this interpretation, recall that the planar gradient ∇ is related to directional derivatives of any differentiable function u through $\frac{\partial u}{\partial \vec{e}_\alpha} = \vec{e}_\alpha^T \nabla u$, where $\vec{e}_\alpha = (\cos\alpha, \sin\alpha)^T$ is a unit vector in the direction α. The vector J possesses a similar property. If $\psi_\alpha \triangleq \Pi^{-1}\frac{\partial g}{\partial \vec{e}_\alpha} = \Pi^{-1}\vec{e}_\alpha^T\nabla g$, since Π and D_t are radial, $J_\alpha(\omega,t) \triangleq (D_t\psi_\alpha * I)(\omega,0) = \vec{e}_\alpha^T J(\omega,t)$. This observation is related to the extension of *steerability* on the sphere, thanks to the radiality of the projection Π [8, 6].

Let us conclude this section by noticing that the gnomonic dilation has a nice compatibility with the projective geometry linking any 3-D scene to its spherical apparent image. Indeed, by construction, the gnomonic dilation is nothing but the transformation that obeys the omnidirectional image of an object when only the distance of this latter to the sensor is changing. Figure 3 helps visualize this principle. Let us take a thin object O at distance d_1 from the view sphere S^2. This object spans a solid angle α_1 on S^2 and has an apparent width r_1 in the tangent plane $T_\omega S^2$, i.e. $r_1 = \tan\alpha_1$. The same object placed at distance d_2 and denoted now by O' corresponds to an apparent width $r_2 = \tan\alpha_2 = (d_1/d_2)\tan\alpha_1 = (d_2/d_1)r_1$, and α_2 is therefore the gnomonic dilation of α_1 by a factor $d_1/d_2 < 1$.

It seems then clear that if I and I' are the images of O and O' respectively, some "characteristic" (gnomonic) scale of the elements of I around direction ω has to be multiplied by the same factor d_1/d_2 to correspond to the characteristic scale of I' around the same direction. We use this principle in the next section to develop a surface depth estimator.

4. AUTOMATIC SCALE SELECTION AND DEPTH ESTIMATION

In [9], when $s = \gamma t$, it is proved that the determinant or the trace of the planar texture descriptor reaches a maximum at characteristic texture scales if the derivatives are normalized according to $\nabla' = t\nabla$. This effect is then used to select automatically the scale $t = t_{\max}(x)$, where $t_{\max}$ is the scale where either $\det\mu_I$ or $\mathrm{trace}\,\mu_I$ is maximum.

We use here the same scale selection using only the determinant. However, in (2), the derivative renormalization is not required since we select an L^2 normalization of the filters in J. This choice, at least for small scales, amounts to multiply an L^1 normalization by t, having the same effect than the planar derivative normalization. Our texture descriptor is thus fully equivalent to the planar one with renormalization of its derivatives.

Consequently, thanks to the heuristic argument given in Section 3.2 linking gnomonic dilation and local depth, we

use

$$\hat{r}(\omega) \triangleq t_{\max}^{-1}(\omega),$$

as an estimator of the depth $r(\omega)$. We define also the *characteristic second moment matrix* as $\mu_I(\omega) \triangleq \mu_I(\omega, t_{\max}(\omega), \gamma t_{\max}(\omega))$.

Notice that the estimator $\hat{r}$ is biased as soon as the object surface is not perpendicular to the view direction, i.e. for $\sigma \neq 0$. Indeed, in that case, the isotropic averaging realized implicitely by μ_I is not realized on the same texture elements since, as seen in the next section in equation (4), the surface slanting induces anisotropic deformations of the apparent texture. A better estimator would have thus to correct this anisotropic effect. However, the price of that correction would be a higher computational time since extra parameters must be introduced in the texture descriptor. For that reason, we will keep $\hat{r}$ remembering that it corresponds to a first order guess of the true depth r.

5. TEXTURE MODELS AND SLANT ESTIMATION

As shown above, our texture descriptor is asymptotically equivalent to the planar texture descriptor of [9] for small scales. Indeed, since the supports of $D_t\psi_x$ and $D_t\psi_y$ are well localized around their centrum, in that case the gnomonic dilation is close to planar dilation thanks to its Euclidean asymptotic property described in Section 3.

Using the notations introduced in Section 2, if we assume that the differential map dF_ω is roughly constant at the typical scales of the texture variations, then

$$\mu_I(\omega) \quad = \quad \mathrm{d}F_\omega^T \, \mu_{\mathscr{T}}(q) \, \mathrm{d}F_\omega, \tag{4}$$

where $q = F(\omega)$. With this assumption, the two texture models described in [9], namely *weakly isotropic* (WI) and *constant area* (CA) hypothesis, can be used.

Weakly Isotropic Textures : This hypothesis assumes that the texture displays roughly the same behavior on $\mathscr{M}$ whatever is the point and the direction where we look at it. This leads to assume $\mu_{\mathscr{T}}(q) = c\,\mathrm{Id}$ for a certain constant $c > 0$. In that case, (4) leads to $\mu_I(\omega) \propto \mathrm{d}F_\omega^T \mathrm{d}F_\omega$. Eigenvalues of $\mathrm{d}F_\omega$ are then proportional to the square root of the eigenvalues $0 < \lambda_2 < \lambda_1$ of $\mu_I(\omega)$. Therefore, the slant estimator of the WI hypothesis is $\cos\hat{\sigma}_{\mathrm{iso}}(\omega) \triangleq \sqrt{\lambda_2(\omega)/\lambda_1(\omega)}$.

Textures of Constant Area : For this case, the texture is assumed to present a less restrictive behavior : the local size of the texture elements, and thus also $\det\mu_{\mathscr{T}}$, is almost constant on $\mathscr{M}$. Therefore, from (4), $\det\mu_I(\omega) = \det\mu_{\mathscr{T}}(q)(\det\mathrm{d}F_\omega)^2$, and $mM = \left[\det\mu_{\mathscr{T}}(q)/\det\mu_I(\omega)\right]^{1/2}$. Defining $A_I = (\det\mu_I)^{-1/2}$, the CA hypothesis implies then $\frac{\nabla A_I}{A_I} = \frac{\nabla(mM)}{mM} = -\tan\sigma \begin{pmatrix} 3 + r\kappa_t/\cos\sigma \\ r\tau \end{pmatrix}$, where the last geometric equality is proved in [3] and where the gradient is the spherical gradient. The two quantities κ_t and τ are the normal curvature and the torsion of the surface in direction $\vec{T}$ respectively. If those are very small, e.g. for very smooth surfaces, the CA slant estimator is $|\tan\hat{\sigma}_{\mathrm{ca}}| \triangleq \frac{1}{3}\frac{\|\nabla A_I\|}{A_I}$.

6. ALGORITHM AND IMPLEMENTATION

Our algorithm estimates shape informations from textures, i.e. distance r and slant angle σ from the texture descriptor μ_I. The integration window is set to $w = \Pi^{-1}g$ where g is the planar Gaussian of Section 3.1, while the integration scale $s = \gamma t$, with $\gamma = 3$. The algorithm is summarized by the following steps :

1. Fix a range of scale $[t_0, t_1] \subset \mathbb{R}_+^*$ where to take t.
2. Compute $\mu_I(\omega; t, \gamma t)$, $\hat{r}(\omega)$ and $\mu_I(\omega)$.
3. Using $\mu_I(\omega)$, compute $\hat{\sigma}_{\mathrm{iso}}$ and $\hat{\sigma}_{\mathrm{ca}}$ according to the texture model selected.

Some explanations have to be provided about the computation of μ_I, i.e. about the computation of correlations on S^2 between a signal I and filters ψ_x, ψ_y and w in a discretized formalism. We follow the numerical procedure explained in [5].The intensity I is sampled on an equiangular grid $\mathscr{G}$ of $N \times N$ pixels (here $N = 512$). Functions on the sphere can be decomposed in the orthornormal basis of Spherical Harmonics (SH) that plays the same role than the Fourier basis on flat spaces. Band-limited functions on the sphere, i.e. such that SH coefficients are zero after a cutoff frequency, can be sampled on an equiangular grid without loss of information. For isotropic filters, a correlation/convolution theorem exists expressing correlation as the inverse SH transform (ISHT) of the product of the SH transforms of the signal and of the filter. If the filter is anisotropic (e.g. ψ_x and ψ_y), this product is actually a matrix product and the ISHT has to be replaced by an inverse Wigner transform (IWT). On equiangular grids SHT and ISHT can be computed very quickly, i.e. in $O(N^2 \log^2 N)$ using the S2Kit[3] C library. IWT can be also computed rapidly using the SOFT[3] C libary. Additional speedup are available thanks to the special structure of ψ_x and ψ_y (i.e. in $\cos\varphi$ or $\sin\varphi$). These filters are indeed *steerable* of order 1 [8, 5] and this property makes the correlation computed by the matrix product of the Wigner transforms limited to few non-zero matrix slices.

7. EXPERIMENT

As a simple experiment, we study the recovery of the geometry of a cube painted with a black and white checkerboard (Fig. 4(b)). This object is seen from inside by a perfect omnidirectional camera placed on one of the median lines exactly at half distance between the cube center and one of its face (Fig. 4(a)). This non-symetric situation improved the challenge of the estimation compared to that of a camera placed at the origin. The omnidirectional view is presented in Fig. 4(c). This synthetic scene and its omnidirectional view have been generated using Blender[4].

Figure 4(d) presents the estimation $\hat{r}(\omega)$ of the depth in every direction ω. Without calibration, this estimation is of course valid up to an unknown multiplicative constant. It seems clear that the center of the faces are correctly localized. As described in Section 4, the bias of $\hat{r}$ induces errors as soon as the slant in non trivial. The estimator has also some trouble to estimate depth close to face junction since two different texture behaviors are mixed in these regions due to the filters size.

Figures 5(a-d) present the results of the slant estimations. The absolute value of the true slant is displayed in Fig. 5(a). The results of the WI and CA slant estimation (Figs 5(b) and

[3] SOFT and S2Kit are based on [10], see `http://www.cs.dartmouth.edu/~geelong/`

[4] `http://www.blender.org`.

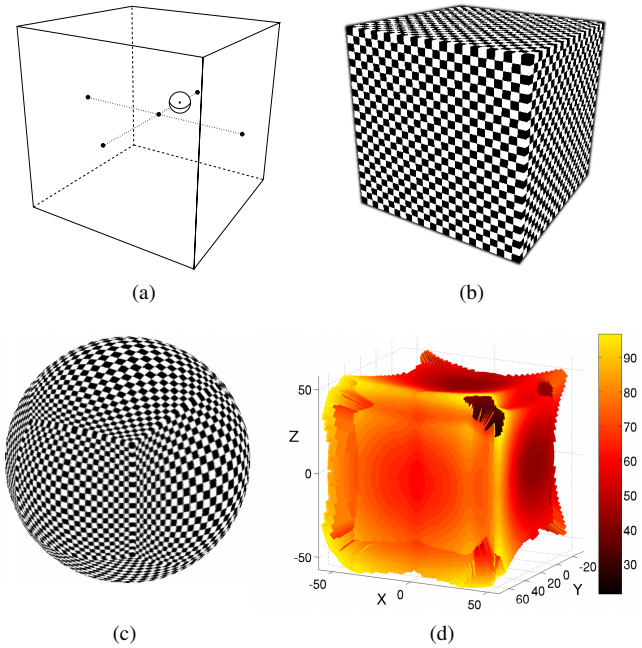


Figure 4: (a) 3-D scene sketch. The small sphere represents the camera location. (b) Blender perspective view of the cube. (c) Omnidirectional vision of the camera (d) Estimated distance.

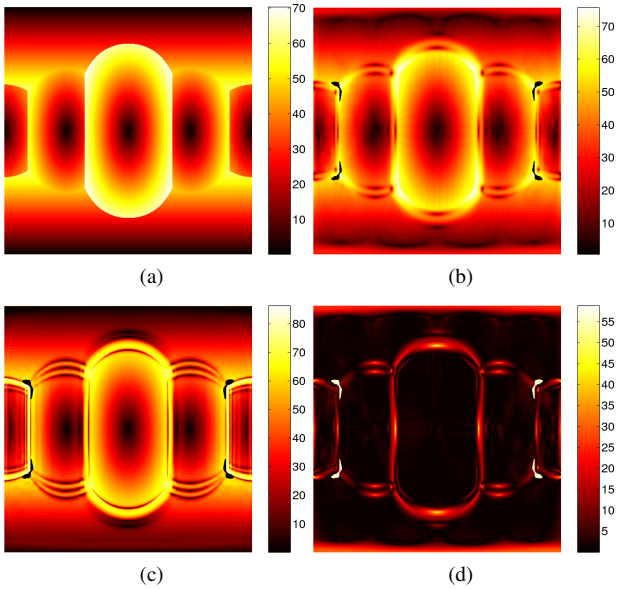


Figure 5: Absolute value of the true slant of the cube (on a (θ, φ) map) according to the scene geometry. For each figure, the horizontal axis refers to $\varphi \in [0, 2\pi)$, $\theta \in [0, \pi)$ is the vertical axis, and the colorbar is in degrees. (b) Slant estimated by the WI hypothesis. (c) Slant estimated by the CA hypthesis. (d) Absolute difference between true slant and WI slant.

5(c)) show that in both situations the accuracy is fairly good except again on the face junction areas, as shown in the difference map between WI slant and the ground truth. Careful inspections of the relative errors between WI and CA slants and the true slant show that the CA hypothesis works better when the true slant is small, while globally the MSE between WI slant and the ground truth is slightly lower than for CA slant.

8. CONCLUSION

This paper has explained how to obtain 3-D shape information from textures in the particular context of omnidirectional image processing. The presented results demonstrate the efficiency of our mathematical tools in a controlled synthetic situation. Experiments on real world images have still to be realized. Our paper however proves that 3-D shape cues from textures can be derived from data delivered by omnidirectional vision systems. A realistic and efficient system providing a complete 3-D scene recovery from real omnidirectional views would have to combine our texture treatment to other indications, e.g. cues from object occlusions, bluring due to the link between depth and out-of-focus object, motion parallax, stereo vision cues, ...

REFERENCES

[1] I. Tosic, I. Bogdanova, P. Frossard, and P. Vandergheynst, "Multiresolution Motion Estimation for Omnidirectional Images," in *EUSIPCO*, 2005.

[2] C. Geyer and K. Daniilidis, "Catadioptric projective geometry," *International Journal of Computer Vision*, vol. 45, no. 3, pp. 223–243, December 2001.

[3] J. Garding, "Shape from texture for smooth curved surfaces," in *Computer Vision - ECCV '92. Second European Conference on Computer Vision Proceedings*. 1992, pp. 630–8, Springer-Verlag Berlin, Germany.

[4] H. Fuhr, TU Munchen, G.W.L.H. Munich, A. Sinica, T. Taipei, B. Torresani, and F. Marseille, "Shape from texture using continuous wavelet transform," *Proceedings of SPIE, the International Society for Optical Engineering*, vol. 4119, pp. 93–107, 2000.

[5] Y. Wiaux, L. Jacques, and P. Vandergheynst, "Fast directional correlation on the sphere with steerable filters," *Astrophys. J.*, vol. 652, pp. 820, 2006.

[6] Y. Wiaux, L. Jacques, and P. Vandergheynst, "Correspondence principle between spherical and Euclidean wavelets," *Astrophys. J.*, vol. 632, pp. 15, 2005.

[7] J.-P. Antoine and P. Vandergheynst, "Wavelets on the 2-sphere : a group theoretical approach," *Appl. Comput. Harmon. Anal.*, vol. 7, pp. 1–30, 1999.

[8] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 9, pp. 891–906, 1991.

[9] T. Lindeberg and J. Garding, "Shape from texture from a multi-scale perspective," in *Computer Vision, 1993. Proceedings., Fourth International Conference on*, 1993, pp. 683–691.

[10] D.M. Healy, D.N. Rockmore, PJ Kostelec, and S. Moore, "FFTs for the 2-Sphere-Improvements and Variations," *Journal of Fourier Analysis and Applications*, vol. 9, no. 4, pp. 341–385, 2003.