January 2001

Surface Representations Using Spherical Harmonics and Gabor Wavelets on the Sphere

Thomas Bülow
University of Pennsylvania

Kostas Daniilidis
University of Pennsylvania, kostas@cis.upenn.edu

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-01-37.

Surface Representations Using Spherical Harmonics and Gabor Wavelets on the Sphere

Abstract
In this paper we present a new scheme for the representation of object surfaces. The purpose is to model a surface efficiently in a coarse to fine hierarchy. Our scheme is based on the combination of spherical harmonic functions and wavelet networks on the sphere. The coefficients can be estimated from scattered data sampled from a star-shaped object's surface. Spherical harmonic functions are used to model the coarse structure of the surface, while spherical Gabor wavelets are used for the representation of fine scale detail. Theoretical background on wavelets on the sphere is provided as well as a discussion of implementation issues concerning convolutions on the sphere. Results are presented which show the efficiency of the proposed representation.

Keywords
surface geometry, wavelets, spherical Gabor filters

Comments
University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-01-37.

Surface Representations using Spherical Harmonics and Gabor Wavelets on the Sphere

Thomas Bülow and Kostas Daniilidis

GRASP Laboratory
Department of Computer and Information Science
University of Pennsylvania, Philadelphia, USA
{thomasbl, kostas}@grasp.cis.upenn.edu

Technical Report: MS-CIS-01-37

Abstract

In this paper we present a new scheme for the representation of object surfaces. The purpose is to model a surface efficiently in a coarse to fine hierarchy. Our scheme is based on the combination of spherical harmonic functions and wavelet networks on the sphere. The coefficients can be estimated from scattered data sampled form a star-shaped object's surface. Spherical harmonic functions are used to model the coarse structure of the surface, while spherical Gabor wavelets are used for the representation of fine scale detail. Theoretical background on wavelets on the sphere is provided as well as a discussion of implementation issues concerning convolutions on the sphere. Results are presented which show the efficiency of the proposed representation.

Keywords: Surface Geometry, Wavelets, spherical Gabor filters.

This work was partly supported by the German Research Association (Deutsche Forschungsgemeinschaft – DFG) under the grant Bu 1259/2-1.

1 Introduction

Surface and 3D object representations have attracted intense research efforts since the beginnings of computer vision and graphics as well as in computational approaches in related disciplines like photogrammetry, geography, seismology, and many other areas. Three-dimensional representations are applied in rendering, design, construction, shape recovery, tracking, segmentation, recognition, and transmission of scene objects. Such representations can be parametric like generalized cylinders and superquadrics [13], globally or locally deformable models [17], implicit surfaces [24], or volumetric. In the graphics community, polyhedral meshes are popular because of the advance in graphics hardware. In shape models and CAGD function bases like splines or wavelets are used. Any attempt to review the work over three decades would yield an incomplete and unfair document. So, we refer to an older [3] and a most recent textbook [23].

The goal of our approach is to establish representations efficient for transmission of dynamic 3D objects. In this paper we concentrate on static representations but we propose representations appropriate for moving objects. The requirements for the representation we seek are: **(1)** The data are irregularly sampled 3D measurements possibly corrupted by noise and outliers. As opposed to computer generated models they are obtained from sensors: laser scanners or stereo vision systems. **(2)** Local changes and deformations should not have a global effect, in either case of motion or just data corruption. **(3)** The approximation should be appropriate for **transmission**: the resulting code should be short and it should include multiple resolutions in a coarse to fine hierarchy to facilitate progressive transmission. **(4)** The approximation should be invariant with respect to motions, in this case we address only rotations, so that tracking reduces to finding the change in the approximation parameters and not in re-fitting the model.

Due to the multi-resolution and the locality requirement, wavelet families are the first coming in question. They have been proven successful in image (JPEG2000) as well as in geometry compression [21]. However, shifting or rotating the data change the representation entirely.

In this paper, we propose networks of Gabor functions on the sphere. Wavelet networks can satisfy the invariance property, they are still a multi-resolution representation, and such approximations are not affected by local changes. Wavelet networks [26, 4] and radial basis functions [19] have been introduced for the approximation of continuous functions [26]. Recently Gabor wavelet networks have been used for representing objects in images [14]. Like truncated discrete wavelet decompositions wavelet networks approximate functions by a linear superposition of wavelets. However, wavelet networks allow the wavelet parameters (position, scale, orientation) to be optimized instead of taking values sampled from a regular grid in space and frequency.

We will treat object surfaces as functions defined on the sphere. To make this approach feasible we have to restrict ourselves to star-shaped objects. Star-shaped objects contain a point in there interior such that each ray originating from this point intersects the object surface exactly once. While this allows us to treat the surface as a function defined on the sphere, we are well aware that it also imposes a severe restriction on the class of objects representable by the proposed scheme. However, this is a question of segmentation which we do not address yet.

The use of spherical harmonics for the representation of objects was first proposed by Schudy and Ballard [22, 3]. They model the dynamic heart volume by spherical harmonics with periodically time-varying coefficients. Spherical harmonics have later been used to compress coarse scale head models [10]. Higher dimensional surface harmonics have been used for static and dynamic object representations [16].

Multiresolution representations of surfaces have first been proposed in [9]. Based on this work a discrete spherical wavelet transform was introduced [21]. This construction works on a semi-regular triangulation of the sphere. The sphere is parameterized by an indexing scheme of the grid vertices.

Since we are interested in invariances of the representation against rotations we turn our attention to methods using continuous parameterizations of the sphere. There are two different approaches to Gabor filters on the sphere. The first one is based on the definition of Gabor functions as windowed harmonic functions [11]. Since on S^2

this approach leads to windowed spherical harmonics, the characteristics of Gabor filters as directional filters is not carried over from the plane to the sphere. This can be achieved by using a stereographic mapping of planar directional filters to the sphere [8]. The foundation of this approach lies in a group theoretical approach to wavelet transforms on the sphere [2].

Our algorithm proceeds in two steps as follows: As described above the data is assumed to be given as a scalar function defined on the unit sphere. In the first step, to represent its *coarse* structure, this function is expanded into a series of spherical harmonic functions. We refrain from representing *fine* detail using spherical harmonics since these functions are globally supported. Fine detail will rather be represented by local functions on the sphere. This is accomplished in the second step using spherical Gabor filters. These filters are optimized on the residual function remaining after the first step. A non-linear minimization yields the optimal offsets, orientations, and scales for the Gabor wavelet network. Starting values for the iterative minimization are established through convolution with Gabor filters with a fixed set of scales.

Though the original contribution of the paper is in the use of Gabor wavelet networks as an efficient surface representation, we believe that presenting the properties of convolution on the sphere is of independent interest to the graphics and vision community.

2 Mathematical Preliminaries

In this section we summarize some facts about spherical harmonics functions which we will use in the rest of this article. As an overall reference on this subject we refer to [6, 5]. In the following we parameterize the unit sphere S^2 embedded in $\mathbb{R}^3$ using standard spherical coordinates. Thus, an element of $\eta \in S^2$ will be written as

$$\eta := (\cos(\varphi)\sin(\vartheta),\ \sin(\varphi)\sin(\vartheta),\ \cos(\vartheta)), \tag{1}$$

with $\varphi \in [0, 2\pi)$ and $\vartheta \in [0, \pi]$.

The spherical harmonic functions $Y_{lm} : S^2 \to \mathbb{C}$ are defined as the everywhere regular eigenfunctions of the spherical Laplace operator. These functions constitute a complete orthonormal system of the

space of square integrable functions on the sphere $L^2(S^2)$.

$$Y_{lm}(\eta) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}} P_l^m(\cos(\vartheta)) e^{im\varphi}, \tag{2}$$

With $l \in \mathbb{N}$ and $|m| \leq l$. Here P_l^m denote the associated Legendre functions

$$P_l^m(x) = \frac{(-1)^m (1-x^2)^{m/2}}{2^l l!} \frac{d^{l+m}}{dx^{l+m}} (x^2-1)^l. \tag{3}$$

Any function on the sphere can be expanded into spherical harmonics:

$$f = \sum_{l \in \mathbb{N}} \sum_{|m| \leq l} f_{lm} Y_{lm}. \tag{4}$$

The coefficients f_{lm} can be extracted from a given function f as follows

$$f_{lm} = \int_{\eta \in S^2} f(\eta)\ \bar{Y}_{lm}(\eta)\ d\eta. \tag{5}$$

For the surface element on the sphere we use the shorthand notation $d\eta := \sin(\vartheta)\, d\vartheta\, d\varphi$. If $f_{lm} = 0$ for all $l > L$, f is called band-limited with band-width L. The set of coefficients f_{lm} is called the spherical Fourier transform of f.

If we deal with a real function f we can expand it as well using real coefficients only:

$$f = \sum_{l \in \mathbb{N}} \left[a_{l0}\, Y_{l0}^R + \sum_{m=1}^{l} (a_{lm}\, Y_{lm}^R + b_{lm}\, Y_{lm}^I) \right], \tag{6}$$

with

$$Y_{lm}^I = \sqrt{2}\, \Im(Y_{lm}),\ Y_{l0}^R = \Re(Y_{l0}),\ \text{and}\ Y_{lm}^R = \sqrt{2}\, \Re(Y_{lm}),\ \text{for}\ m > 0, \tag{7}$$

and

$$a_{lm} := \int_{\eta \in S^2} f(\eta)\, Y_{lm}^R(\eta)\, d\eta, \quad b_{lm} := \int_{\eta \in S^2} f(\eta)\, Y_{lm}^I(\eta)\, d\eta. \tag{8}$$

Rotations in $\mathbb{R}^3$ will be parameterized by Euler angles such that any $g \in SO(3)$ will be written as

$$g(\gamma, \beta, \alpha) = R_z(\gamma) R_y(\beta) R_z(\alpha), \tag{9}$$

where R_y and R_z denote rotation about the y-, and z-axis, respectively. In matrix notation R_y and R_z take the form

$$R_y(\beta) = \begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{bmatrix}, \quad R_z(\alpha) = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{10}$$

Rotating a function $f \in L^2(S^2)$ will be performed by the operator $\Lambda(g)$ which is defined by

$$\Lambda(g)f(\eta) := f(g^{-1}\eta). \tag{11}$$

An important property of spherical harmonics is that the subspaces

$$L^2(S^2) \supset Y_l := \{Y_{lm} |\ |m| \leq l\}, \quad l \in \mathbb{N} \tag{12}$$

are invariant under the action of $SO(3)$. Thus, a rotated spherical harmonic function can be expressed by

$$\Lambda(g)Y_{lm}(\eta) = \sum_{|n| \leq l} U^l_{mn}(g) Y_{ln}(\eta), \tag{13}$$

where the matrix $(2l+1) \times (2l+1)$ matrix $U^l_{mn}(g)$ is given by

$$U^l_{mn}(g(\gamma, \beta, \alpha)) = e^{-im\gamma} P^l_{mn}(\cos(\beta)) e^{-in\alpha}. \tag{14}$$

The explicit form of the generalized associated Legendre polynomials P^l_{mn} can be found in [5]. Finally, we note that

$$U^l_{mn}(g_2 g_1) = \sum_{|k| \leq l} U^l_{mk}(g_2) U^l_{kn}(g_1). \tag{15}$$

3 Spherical Wavelet Networks

3.1 Wavelet Networks

Wavelet networks have emerged as a tool for function approximation from the combination of wavelet decompositions with ideas from the theory of neural networks. Networks of radial basis functions

for function approximation have been investigated in [19]. Wavelet theory investigates families of functions

$$\Phi = \{g_{mn}(x) = a_0^{-m/2} g(a_0^{-m}x - nb_0) | m, n \in \mathbb{Z}\} \tag{16}$$

which are derived from a single function $g \in L^2(\mathbb{R})$, satisfying $\int g(x)\, dx = 0$, by dilations and translation of fixed step size. An important result is that there exist values $a_0 > 1$ and $b_0 > 0$ for a given function g such that the family Φ constitutes a frame for $L^2(\mathbb{R})$. That is numbers $A > 0$ and $B < \infty$ exist such that

$$A\|f\|^2 \leq \sum_{m,n\in\mathbb{Z}} |\langle g_{mn}, f\rangle|^2 \leq B\|f\|^2 \tag{17}$$

for all $f \in L^2(\mathbb{R})$. The frame property guarantees that f can be reconstructed in a stable way from the coefficients $c_{mn} := \langle g_{mn}, f\rangle$. In practice a finite number of wavelets can be used for the approximation of f in a truncated wavelet decomposition

$$f(x) \approx \sum_{k=1}^{N} w_k \phi_k(x), \quad \phi_k \in \Phi. \tag{18}$$

The idea behind wavelet networks developed by Zhang and Benveniste[1] [26] is that a superposition of N wavelets should give a better approximation if the translation and dilation parameters were not restricted to a discrete lattice. They propose to approximate a function as

$$f(x) \approx \tilde{f}_{\boldsymbol{\theta}}(x) = \sum_{k=1}^{N} w_k g\left(\frac{x - b_k}{a_k}\right), \quad a_k \in \mathbb{R}^+, b_k \in \mathbb{R}. \tag{19}$$

where $\boldsymbol{\theta}$ is the vector containing all free parameters

$$\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_N), \qquad \text{with } \theta_i = (w_i, a_i, b_i). \tag{20}$$

The parameter vector $\boldsymbol{\theta}$ has to be adjusted in order to minimize the objective function

$$E_f(\boldsymbol{\theta}) := \|f - \tilde{f}_{\boldsymbol{\theta}}\|^2. \tag{21}$$

[1] Zhang and Benveniste introduced wavelet networks for approximating n-dimensional functions. In order to keep notation simple we sketch the main idea only for $n = 1$.

In wavelet networks for the approximation of n-dimensional functions other parameters, especially rotations, can be taken into account in order to increase flexibility [26, 14]. Recently, wavelet networks using odd 2D Gabor filters have been used for the representation of face images and head pose estimation [14].

Discrete families of Gabor functions are known to constitute frames [7]. The frame bounds are also known in the case that an additional discrete rotation parameter is introduced [15].

3.2 Wavelets on the Sphere

In order to apply wavelet-networks to functions on the sphere, we first have to define the set of wavelet-functions on the sphere. This involves the definition of a mother-wavelet $\Psi \in L^2(S^2)$ and a set of transforms on the sphere which generate a function space by acting on Ψ. These transforms should correspond to translations, rotations and dilations in $\mathbb{R}^2$ which are used in order to define planar CWT's [1] and wavelet-networks [26]. The set of motions on the sphere (corresponding to translations and rotations in $\mathbb{R}^2$) is easily identified with the group of rotations in 3D space $SO(3)$. A more serious problem arises with the definition of dilations on the sphere. The naïve approach of defining a dilation about the north-pole by $D_a : S^2 \to S^2, (\varphi, \vartheta) \mapsto (\varphi, a\vartheta)$ is doomed to fail since it is either not surjective ($a < 1$) or not injective ($a > 1$). For $a < 1$ this could be remedied by assigning some value, say 0, to the sector $\{(\varphi, \vartheta) | \vartheta > a\pi\}$. However, $a > 1$ leads to "overlapping" around the south-pole, so this dilation is not well-defined. Recently, the following definition of dilation on the sphere has been proposed [2]: Dilating a function on the sphere about the north-pole is performed by subsequently (1) projecting the function stereographically from the south-pole to the plane tangent to S^2 at the north-pole, (2) dilate the mapped function within the tangent plane, and (3) map the dilated function back to the sphere by inverse stereographic projection. Along the same line the construction of a mother wavelet on the sphere can be performed [8]. A mother wavelet is defined in $\mathbb{R}^2$ and mapped by inverse stereographic projection to the sphere. We will follow this approach here.

The stereographic projection $\Pi : S^2 \backslash \{(0,0,-1)\} \to \mathbb{R}^2, (\varphi, \vartheta) \mapsto (x(\varphi,\vartheta), y(\varphi,\vartheta))$ is given by

$$x(\varphi,\vartheta) = \frac{2\sin(\vartheta)\cos(\varphi)}{1+\cos(\vartheta)}, \quad y(\varphi,\vartheta) = \frac{2\sin(\vartheta)\sin(\varphi)}{1+\cos(\vartheta)} \tag{22}$$

Lets consider a three-parameter family of odd Gabor function in $\mathbb{R}^2$:

$$G(x,y;\sigma_x,\sigma_y,\omega) := \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\left[x^2/\sigma_x^2 + y^2/\sigma_y^2\right]\right) \sin(\omega x). \tag{23}$$

Here σ_x and σ_y denote the standard deviation of the Gaussian envelope function and ω is the spatial frequency.

We define the spherical Gabor function G_S as

$$G_S(\eta;\sigma_x,\sigma_y,\alpha,\omega,\eta_0) := \Lambda(g(\varphi_0,\vartheta_0,\alpha))G(x(\varphi,\vartheta),y(\varphi,\vartheta);\sigma_x,\sigma_y,\omega) \tag{24}$$

Here $x(\varphi,\vartheta)$ are $y(\varphi,\vartheta)$ are defined as in (22). Note, that x and y are not defined for $\vartheta = \pi$. Since $\sqrt{x^2+y^2}$ goes to infinity for $\vartheta \to \pi$ and our Gabor filters go to zero for $\sqrt{x^2+y^2}$ goes to infinity, we set

$$G_S(\eta;\sigma_x,\sigma_y,\alpha,\omega,\eta_0) = 0, \text{ for } \eta = \eta_0 \tag{25}$$

The rotation $g(\varphi_0,\vartheta_0,\alpha)$ has the following effect. First the function is rotated about the north-pole by α. This has the same effect as a rotation about the origin in $\mathbb{R}^2$ prior to inverse stereographic projection would have. The remaining rotations by ϑ_0 and φ_0 take the north-pole to

$$\eta_0 := (\cos(\varphi_0)\sin(\vartheta_0),\ \sin(\varphi_0)\sin(\vartheta_0),\ \cos(\vartheta_0)), \tag{26}$$

An example of a spherical Gabor filter is shown in Fig. 1. The set of Gabor functions G_S constitutes a wavelet family only if all functions in the set can be derived from a single mother wavelet by dilations and motions on the sphere. In order to achieve this, the parameters σ_x, σ_y, and ω must be coupled

$$c_1 := \frac{\sigma_x}{\sigma_y}, \quad c_2 := \sigma_x\omega \tag{27}$$

where c_1 and c_2 are constants. c_1 is the aspect ratio of the Gabor function and c_2 determines the number of oscillations of the *sin* function that fall into the Gaussian window. Freezing c_1 and c_2 reduces the number of free parameters of a spherical Gabor filter to 4.

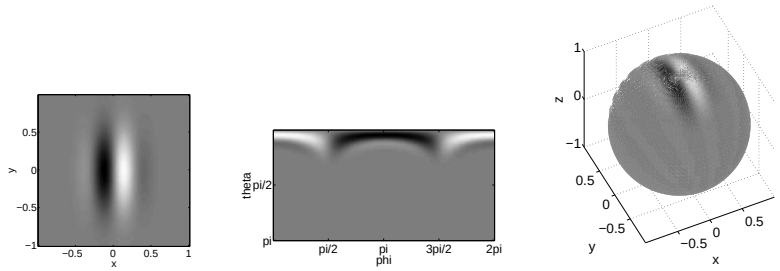

Fig. 1. Left: A planar Gabor filter G with $\sigma_x = 0.3$, $\sigma_y = 0.5$, and $\omega = 10$. **Middle and Right:** The filter G_S as obtained by inverse stereographic projection of G. No rotation is applied ($g = Id$).

3.3 Spherical Wavelet Networks

In this section we will use the idea of wavelet networks as introduced in Sect. 3.1 and combine it with Gabor wavelets on the sphere as defined in (23) and (24). The aim is to approximate a given function $f \in L^2(S^2)$ by $\tilde{f}$

$$\tilde{f}_{\boldsymbol{\theta}}(\eta) = \sum_{i=1}^{N} w_i G_S(\eta; \sigma_{x\,i}, \sigma_{y\,i}, \alpha_i, \omega_i, \eta_{0\,i}), \tag{28}$$

where $\boldsymbol{\theta}$ is the vector containing the parameters of the wavelet network

$$\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_N), \qquad \text{with } \theta_i = (w_i, \sigma_{x\,i}, \sigma_{y\,i}, \alpha_i, \omega_i, \eta_{0\,i}). \tag{29}$$

The objective function which is going to be minimized is again

$$E_f(\boldsymbol{\theta}) = \|f - \tilde{f}_{\boldsymbol{\theta}}\|^2. \tag{30}$$

Optimization Procedure. The functions $\tilde{f}_{\boldsymbol{\theta}}$ constitute a model for functions in $L^2(S^2)$. This model is non-linear in the model-parameters $\boldsymbol{\theta}$. In order to minimize the objective function (30) we use the Levenberg-Marquart algorithm which is a standard method for non-linear model fitting [20]. Instead of optimizing the whole vector $\boldsymbol{\theta}$ at once, we optimize the Gabor wavelets separately. Let $\tilde{f}_{(\theta_1,\ldots,\theta_{n-1})}$ be the optimized approximation of f using the first $n-1$ wavelets. The nth wavelet is then optimized in order to minimize

$$\|f - \tilde{f}_{(\theta_1,\ldots,\theta_{n-1})} - w_n G_S(\eta; \sigma_{x\,n}, \sigma_{y\,n}, \alpha_n, \omega_n, \eta_{0\,n})\|^2 \tag{31}$$

As mentioned in Sect. 3.2 c_1 and c_2 representing the aspect ratio and the number of osciallations of the functions should be kept constant for the filters to be called wavelets. In order to allow for more flexibility, we will allow these values to vary in a given range. In our experiments we use the ranges

$$1 \leq c_1 \leq 3, \quad 1 \leq c_2 \leq 4 \tag{32}$$

As a further constraint we allow σ_x to vary only within an interval $[\sigma_x^i/2, 2\sigma_x^i]$, where σ_x^i is the initial value.

Initialization. In [14] a pyramid like initialization is proposed: Wavelet positions are initialized on regular grids on different scales. Starting from a coarse grid for large scale wavelets, finer grids are chosen for small scale wavelets. This method cannot be extended directly to spherical wavelets, since there is no regular grid on the sphere. We compare two different initializations here. In both schemes the initial scale, aspect ratio and frequency will be set to fixed, predefined values. (1) The initial wavelet positions are drawn randomly from a uniform distribution on the sphere. The initial values for the orientation α will be drawn from a uniform distribution on the interval $[0, \pi)$. (2) Another possible way to initialize the positions is data-dependent. Given a function $d_f : S^2 \to [0, 1]$. We draw random initial positions $\eta^i \in S^2$ from a probability distribution proportional to d_f.

We select n initial positions as follows. Draw a random η from the uniform distribution on the sphere. Draw another number ξ from the uniform distribution on the interval $[0, 1]$. If $\xi \geq d_f(\eta)$ add η to the set of initial wavelet position and update d_f. Repeat until n initial positions have been selected.

We will now present the function d_f used in our experiments and explain the "update d_f" step in the above algorithm. We will make use of the relation between the wavelet transform and wavelet networks which was mentioned in Sect. 3.1. Convolving the function $f \in S^2$ with a Gabor filter G_S will lead to high filter responses at positions which are likely to be close to optimal wavelet positions. The convolution will be defined as

$$(f * h)(g) = \int_{\eta \in S^2} (\Lambda(g) f(\eta)) h(\eta)\, d\eta, \quad g \in SO(3). \tag{33}$$

For a given $G_S = G_S(\eta; \sigma_x, \sigma_y, 0, \omega, 0)$ we have

$$(G_S * h)(g(\varphi, \vartheta, \alpha)) = \langle G_S(\cdot; \sigma_x, \sigma_y, \alpha, \omega, \eta(\varphi, \vartheta)), h \rangle. \tag{34}$$

We define d_f as

$$d_f(\eta(\varphi, \vartheta)) := \left[\int_{\alpha \in [0,\pi)} |(G_S * h)(g(\varphi, \vartheta, \alpha))| \, d\alpha \right]^2. \tag{35}$$

The initial orientation α_i at position $\eta(\varphi, \vartheta)$ is determined as

$$\alpha_i = \arg \max_{\alpha \in [0,\pi)} |(G_S * h)(g(\varphi, \vartheta, \alpha))|. \tag{36}$$

In the "update d_f" step we subtract from d_f a Gaussian function with standard deviation σ_x centered at η. This reduces the chance of selecting another initial position too close to η.

3.4 Combining Spherical Harmonics and Gabor Wavelets

Our motivation for the use of Gabor filters as opposed to spherical harmonics was that the energy of Gabor filters is locally concentrated while the energy of each spherical harmonic function is distributed all over the sphere. Fine local detail is therefore more appropriately represented by Gabor filters. The spherical harmonics with low l, however, represent low-frequency components and thus model the coarse structure of the object well. We therefore propose to combine the two approaches in the following way. Let f be the function to be approximated. The coarse structure of f is given by

$$f_c = \sum_{l=0}^{L} \sum_{|m| \leq l} f_{lm} Y_{lm}. \tag{37}$$

for some small value of L. In our experiments we used $L = 5$. The wavelet network is then optimized on the residual $f_{res} := f - f_c$. Since rotated spherical harmonics are related to an unrotated basis by the matrices U^l as shown in (13) we can use this representation for pose estimation as well. It follows from (5), (13), and the unitarity of the matrices U^l, that the spherical Fourier coefficients of an object

in two different orientations (related by the rotation $g \in SO(3)$) are related by

$$(\Lambda(g)f)_{lm} = \sum_{|n|\leq l} U^l_{nm}(g) f_{ln} \tag{38}$$

Given $(\Lambda(g)f)_{lm}$ and f_{ln} for a given value l this allows us to recover g. This means that a rotation of the object can be recovered from the object's spherical Fourier transform.

4 Implementation

4.1 Converting 3D Data to Functions on the Sphere

The data is usually given as a set of points $\{\boldsymbol{x}_i\}$ sampled from the object surface. We attempt to find the desired interior point (the origin) by matching a sphere to the data-points. We minimize the mean square algebraic distance of the data points to the sphere [27]:

$$\sum_{i=1}^{n} [a_1(x_i^2 + y_i^2 + z_i^2) + a_2 x_i + a_3 y_i + a_4 z_i + a_5]^2. \tag{39}$$

It is well known that the solution is the singular vector corresponding to the smallest singular value of the resulting data matrix. Once a sphere is fitted we move the coordinate origin to the center of the sphere and consider the distances of the data points from the origin as scalars given on the sphere. We attempt to approximate these scattered data points by a scalar function on the sphere.

4.2 Discrete Spherical Harmonic Transform

To speed up computations we will interpolated the data on a regular grid in the following. It is possible to uniformly sample the sphere by inscribing a regular polyhedron into the sphere and identify its vertices with the sample locations. Unfortunately, the regular polyhedron with the highest number of vertices is the dodecahedron which has 20 vertices [12]. Thus, instead of sampling the sphere uniformly we decide to sample the (ϑ, φ)-plane regularly.

$$\left\{ (\vartheta_j, \varphi_k) = \left(\frac{\pi j}{M-1}, \frac{2\pi k}{N} \right) \middle| \, j,k \in \mathbb{Z}, 0 \leq j < M, 0 \leq k < N \right\} \tag{40}$$

If scattered data is given we will interpolate it to obtain values on these grid points.

There exist fast algorithms for computing the spherical harmonic transform (see [18] and references given therein). Even without using sophisticated algorithms, a speedup can be achieved by observing the separability of the discrete spherical Fourier transform.

$$f_{lm} = \sum_{j=0}^{M} \sum_{k=0}^{N} f(\vartheta_j, \varphi_k)\, e^{im\varphi_k} P_l^m(\cos(\vartheta_j)) \sin(\vartheta_j) \tag{41}$$

This allows for the use of an FFT algorithm in the φ_k coordinate. We used this latter method in our experiments.

4.3 Convolution on the Sphere

During the initialization procedure of a spherical wavelet network we have to perform the convolution of two functions given on the sphere. We will present a fast method to perform this convolution.

For functions defined on $\mathbb{R}^n$, the *convolution theorem* states that convolution in the spatial domain corresponds to multiplication in frequency domain. Thanks to FFT algorithms this can considerably speed up practical implementations for convolutions of discrete signals. There exists an analogous theorem for convolutions on the sphere.

Theorem 1 (Spherical Convolution). *Let $f, h \in L^2(S^2)$ and $g \in SO(3)$ then*

$$(f * h)(g) = \sum_{l \in \mathbb{N}} \sum_{|m| \le l} \sum_{|n| \le l} U_{nm}^l(g) f_{ln} \bar{h}_{lm} \tag{42}$$

where, f_{ln} and h_{lm} are the spherical Fourier transforms of f and h.

Proof. Using definition in (33) and the fact that the Y_{lm} constitute an orthonormal basis of $L^2(S^2)$ we obtain

$$\begin{aligned} \int_{\eta \in S^2} (\Lambda(g) f(\eta)) \bar{h}(\eta)\, d\eta &= \int_{\eta \in S^2} \sum_{l \in \mathbb{N}} \sum_{|m| \le k} (\Lambda(g) f)_{lm} Y_{lm}(\eta) \sum_{k \in \mathbb{N}} \sum_{|n| \le k} \bar{h}_{kn} \bar{Y}_{kn}(\eta)\, d\eta \\ &= \sum_{l \in \mathbb{N}} \sum_{|m| \le l} (\Lambda(g) f)_{lm} \bar{h}_{lm}. \end{aligned} \tag{43}$$

The coefficients $(\Lambda(g)f)_{lm}$ are related to the f_{lm} by (38): Inserting this results into (43) completes the proof. □

Theorem 1 allows to calculate the convolution of two functions on the sphere from their spherical Fourier transform. However evaluating (42) directly is very time-consuming since for each value of $g \in SO(3)$ the matrices $U^l_{mn}(g)$ must be calculated up to some maximum value of l.

We show an elegant solution to this problem which is a special case of a method recently introduced by Wandelt et al. [25]. First let us notice that all $g \in SO(3)$ can be expressed in the form

$$g(\gamma, \beta, \alpha) = g(\gamma + \pi/2, \pi/2, 0)g(\beta + \pi, \pi/2, \alpha + \pi/2) =: g_2 g_1. \quad (44)$$

Using this factorization in (42) yields

$$\begin{aligned}(f * h)(g) &= \sum_{l\in\mathbb{N}} \sum_{|m|\le l} \sum_{|n|\le l} U^l_{nm}(g_2 g_1) f_{ln} \bar{h}_{lm} \\ &= \sum_{l\in\mathbb{N}} \sum_{|m|\le l} \sum_{|n|\le l} \sum_{|k|\le l} f_{lk} U^l_{kn}(g_2) U^l_{nm}(g_1) \bar{h}_{lm}. \qquad (45)\end{aligned}$$

The factorization introduced in (45) has the effect that all three Euler angles occur in the complex exponential functions rather than in the polynomials P^l_{mn} (see Eq. (14)). Combining these results yields

$$\begin{aligned}(f * h)(g) &= \sum_{lmnk} f_{lk} e^{-ik(\gamma+\pi/2)} P^l_{kn}(0) e^{-in(\beta+\pi)} P^l_{nm}(0) e^{-im(\alpha+\pi/2)} \bar{h}_{lm} \\ &= \sum_{lmnk} (-1)^k (-i)^{n+m} f_{lk} P^l_{kn}(0) P^l_{nm}(0) e^{-i(k\gamma+n\beta+m\alpha)} \bar{h}_{lm}. \quad (46)\end{aligned}$$

The inverse Fourier transform of (46) is found as

$$\begin{aligned}&\frac{1}{(2\pi)^3} \int_{[0,2\pi]^3} (f * h)(g) e^{i(\gamma K + \beta N + \alpha M)} d\alpha \, d\beta \, d\gamma \\ &= \sum_{l\in\mathbb{N}} (-1)^n (-i)^{k+m} f_{lK} P^l_{KN}(0) P^l_{NM}(0) \hat{h}_{lM} =: \sum_{l\in\mathbb{N}} (f * h)^l_{KMN}. \quad (47)\end{aligned}$$

Here, $(f * h)^l_{KMN}$ is defined to be zero if any of $|K|$, $|M|$, or $|N|$ is greater than l. In practice we will assume f and h to be band-limited with band-width L such that we have a finite sum in (47). From (47) we can derive $(f * h)(g)$ using a 3D inverse FFT algorithm.

5 Results

We demonstrate results on range data of human heads acquired by a laser scanner. We first show an example of data interpolated on a (φ, ϑ) grid (see Fig. 2). We also show the probability distribution d_f (35) and initial wavelet positions drawn from this distribution. Next we show the results for a combination of spherical harmonics

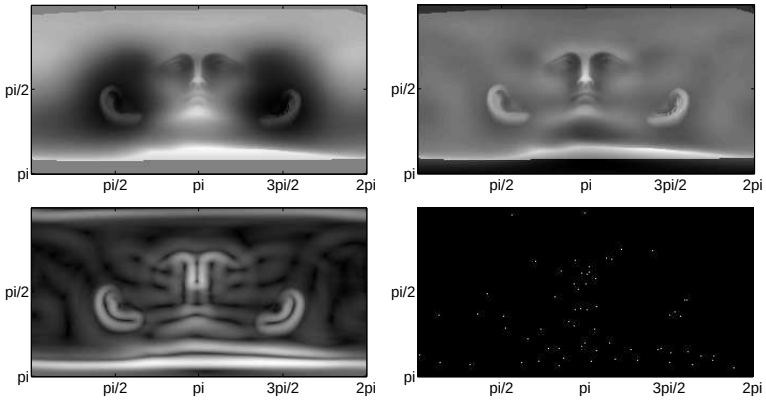

Fig. 2. Top row. **Left**: A sample data set interpolated on a regular grid in the (φ, ϑ)-plane. **Right**: The residual image after subtraction of spherical harmonic reconstruction up to $L = 5$. **Bottom row**. **Left**: The function d_f (35) for $\sigma_x = 0.1$, $c_1 = 1.5$, and $c_2 = 2$. **Right**: 64 initial positions drawn corresponding to initialization method (2) described in Sect. 3.3.

and a wavelet network. The data used in the first set of experiments consisted of a set of 225.896 vertices. The data was interpolated on a 158×313 regular (φ, ϑ) grid. Figure 3 shows the reconstruction from the interpolated data and three approaches to data reduction. (1) Triangulation and decimation of the original data.[2] A decimation to 6374 (corresponding to 19122 floating point numbers (three coordinates per vertex)) vertices is shown. (2) The interpolated data was expanded into a spherical harmonic series up to $l = 35$, which corresponds to 1296 real coefficients. (3) A combination of a spherical harmonic expansion up to $l = 5$ and wavelet network containing 168 wavelets (corresponding to 1176 floating point numbers). The wavelets were initialized on the scales $\sigma_x = 0.4$ (24 wavelets),

[2] For decimation we used the commercial Cyberware software, which is part of the Cyberware 3D scanner system. See `http://www.cyberware.com`.

$\sigma_x = 0.2$ (48 wavelets), and $\sigma_x = 0.1$ (96 wavelets) as described as method (2) in Sect. 3.3. Figure 4 shows the same methods as Fig. 3

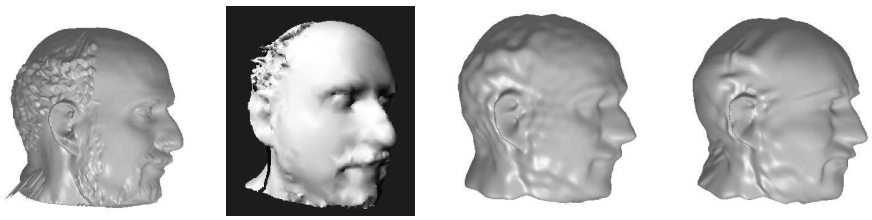

Fig. 3. From left to right: Original data, interpolated on the grid. Decimated data. Spherical harmonic expansion. Spherical harmonics and wavelet network. Number of parameters to store (from left to right): 49454, 19122, 1296, and 1176. (See text for details.)

for another data set. The original data was again interpolated on a 158×313 grid. The decimated triangulation contains 3671 vertices (11013 numbers). The spherical harmonic expansion contains vertices up to $l = 40$ (1681 coefficients). The wavelet network contains 238 wavelets (1702 parameters). All coefficients and parameters are treated as floating point numbers. No quantization scheme has been applied yet. It can be seen that both the spherical harmonic expan-

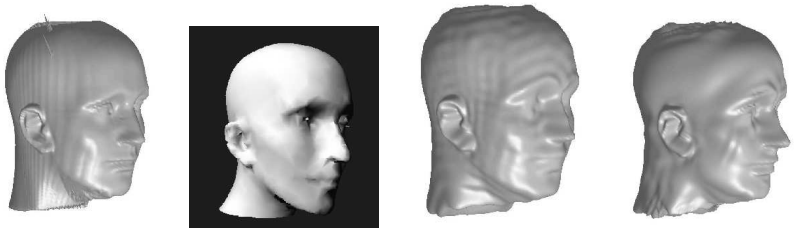

Fig. 4. Like Fig. 3. Number of parameters to store (from left to right): 49454, 11013, 1681, and 1702. (See text for details.)

sion and the combination with wavelet networks need much less coefficients than a decimation of the triangulated data. While smooth surfaces appear "wavy" in the purely spherical harmonic expansion, wavelet networks concentrate detail at real object features.

6 Conclusion

In this paper, we proposed a new approximation method for range data using bases of spherical Gabor functions. Unlike, other wavelet bases defined on a fixed grid in space, scale, and orientation, our method finds the optimal values given an a priori number of basis wavelets. We introduced the theory how to define and dilate wavelets on the sphere. The properties of convolution on the sphere were introduced and a first approximation is obtained by projection through convolution. This approximation serves as an initialization for the non-linear minimization with respect to the wavelet parameters.

We tested our method on human head data obtained with a laser scanner and obtained a very promising data reduction with a visually compelling effect. We compared the results with a commercial polygonal mesh simplification program and found out that our approximation is less susceptible to noise but also details. Our future work will be on updating the spherical Gabor network during motion. Rigid motions can be estimated by updating the center of the sphere and using (38) to compute the rotation.

References

1. J. Antoine and R. Murenzi. Two-dimensional directional wavelets and the scale-angle representation. *Signal Processing*, 52, 1996.
2. J. Antoine and P. Vandergheynst. Wavelets on the 2-sphere: A group-theoretical approach. *Appl. Comput. Harmon. Anal.*, 7, 1999.
3. D.H. Ballard and C.M. Brown. *Computer Vision*. Prentice-Hall Inc., 1982.
4. Ch. Bernard, S. Mallat, and J-J. Slotine. Wavelet interpolation networks. In *European Symposium on Artificial Neural Networks*, Bruges, Belgium, 1998.
5. G.S. Chirikjian and A.B. Kyatkin. *Engineering Applications of Noncommutative Harmonic Analysis*. CRC Press, 2001.
6. R. Courant and D. Hilbert. *Methods of Mathematical Physics*, volume 1. Interscience Publishers, Inc., New York, 1953.
7. I. Daubechies. The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. on Information Theory*, 36(5):961–1004, Sept. 1990.
8. L. Demanet and P. P. Vandergheynst. Directional wavelets on the sphere. Technical Report R-2001-2, Signal Processing Laboratory (LTS), Swiss Federal Institute of Technology/Lausanne (EPFL), 2001.
9. T.D. DeRose, M. Lounsbery, and J. Warren. Multiresolution analysis for surfaces of arbitrary topological type. Technical Report 93-10-05, Dept. of Comp. Sci. and Eng., University of Washington, 1993.
10. Sarp Ertürk and Tim J. Dennis. Efficient representations of 3d human head models. In *BMVC 99*, pages 329–339, 1999.

11. W. Freeden and U. Windheuser. Combined spherical harmonic and wavelet expansion — a future concept in earth's gravitational determination. *Applied and Computational Harmonic Analysis*, 4:1–37, 1997.
12. D. Hilbert and S. Cohn-Vossen. *Geometry and the Imagination.* Chelsea Publishing Company, 1952.
13. A. Jaklic, A. Leonardis, and F. Solina. *Segmentation and Recovery of Superquadrics.* Kluwer Academic Publishers, Utrecht, 2000.
14. V. Krüger, S. Bruns, and G. Sommer. Efficient head pose estimation with gabor wavelet networks. In *Proc. British Machine Vision Conference*, Bristol, Sept. 12-14 2000.
15. T.S. Lee. Image representation using 2d gabor wavelets. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18(10):959–971, Oct. 1996.
16. Art Matheny and Dmitry B. Goldgof. The use of three and four-dimensional surface harmonics for rigid and nonrigid shape recovery and representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 17(10):967–981, 1995.
17. D. Metaxas and D. Terzopoulos. Dynamic 3d models with local and global deformations: Deformable superquadrics. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13:703–714, 1991.
18. M.J. Mohlenkamp. A fast transform for spherical harmonics. *Journal of Fourier Analysis and Applications*, 5:159–184, 1999.
19. T. Poggio and F. Girosi. Networks for approximation and learning. *Proc. of the IEEE*, 78(9):1481–1497, 1990.
20. W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C.* Cambridge University Press, 2nd edition, 1992.
21. P. Schröder and W. Sweldens. Spherical wavelets: Efficiently representing functions on the sphere. *Computer Graphics Proceedings (SIGGRAPH 95)*, pages 161–172, 1995.
22. R.B. Schudy and D. H. Ballard. Towards an anatomical model of heart motion as seen in 4-d cardiac ultrasound data. In *6th Conf. on Computer Applications in Radiology and Computer-Aided Analysis of Radiological Images*, June 1979.
23. L.G. Shapiro and G.C. Stockman. *Computer Vision.* Prentice Hall, 2001.
24. G. Taubin. Estimation of planar curves, surfaces, and nonplanar space curves defined by implicit equations with applications to edge and range image segmentation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1991.
25. B.D. Wandelt and K.M. Gorski. Fast convolution on the sphere. *Phys Rev D63, 123002/1-6*, 2001.
26. Q. Zhang and A. Benveniste. Wavelet networks. *IEEE Trans. on Neural Networks*, 3(6):889–898, Nov. 1992.
27. Z. Zhang. Parameter-estimation techniques: a tutorial with application to conic fitting. *Image and Vision Computing*, 15:59–76, 1997.