Direct 3D-Rotation Estimation from Spherical Images via a generalized shift theorem

Ameesh Makadia and Kostas Daniilidis
Department of Computer and Information Science
University of Pennsylvania

Abstract

Omnidirectional images arising from 3D-motion of a camera contain persistent structures over a large variation of motions because of their large field of view. This persistence made appearance-based methods attractive for robot localization given reference views. Assuming that central omnidirectional images can be mapped to the sphere, the question is what are the underlying mappings of the sphere that can reflect a rotational camera motion. Given such a mapping, we propose a systematic way for finding invariance and the mapping parameters themselves based on the generalization of the Fourier transform. Using results from representation theory, we can generalize the Fourier transform to any homogeneous space with a transitively acting group. Such a case is the sphere with rotation as the acting group. The spherical harmonics of an image pair are related to each other through a shift theorem involving the irreducible representation of the rotation group. We show how to extract Euler angles using this theorem. We study the effect of the number of spherical harmonic coefficients as well as the effect of violation of appearance persistence in real imagery.

1. Introduction

Research in localization or estimation of camera displacement has been dominated by the detection and tracking of landmarks or salient image features in images. In the vast majority of discrete motion approaches, the features are points. Only in the case of differential motion, appearing in the image as optical flow, direct approaches have been employed using spatiotemporal image derivatives only. The latter approaches have been quite successful in case of affine or projective transformations in significant areas of images and in particularly in the production of mosaics. Though designed for image velocities, their hierarchical implementations can also handle larger motions.

Recent interest from the areas of robotics and graphics to omnidirectional or panoramic sensors has given rise to a new group of approaches for localization which are based on the global appearance of the scene. Panoramic mosaics or omnidirectional images contain significant scene content with persistent appearance. Image content can change only due to change of viewpoint or (dis-)appearing components but not because of change in viewing direction. The fact that important structures remain visible enables global navigation tasks to be performed image-based. Appearance-based localization [18, 16] has been shown to be successful using a large set of reference views and applying PCA on them. Pajdla and Hlavac [29] studied the rotation just around the optical axis of panoramic systems.

It is worth, revisiting the problem of image transformations now that the scope involves several image modalities and that the image deformations are more general. Regarding conventional imagery, it is basic knowledge that we can compute affine invariants and that we can compute an affine transformation from combinations of image moments [15, 22] or from Fourier descriptors [30]. Little work has been done on the computation of 3D-rotations from area-based features [17, 28, 24]. Though, there has been much work on projective invariants based on points or curves, there is hardly work on projective image invariants.

Based on harmonic analysis [13, 7, 20] we propose a framework for studying image deformations applicable not only in the plane but also in other domains like the sphere. This framework involves three steps:

- Identify the domain of definition of the signal as a homogeneous space and the group acting on it.
- Check whether an irreducible unitary representation exists for the acting group. Compute the generalized Fourier transform of the image.
- Compute the transformation (group action) from a generalized shift theorem. Compute invariants from the magnitude of the Fourier coefficients.

In this paper, we apply this "recipe" on spherical images arising from omnidirectional cameras with acting group the group of rotations SO(3). As a matter of fact there is no simpler action we can model on the sphere, because even the equivalent of a plane shift is a rotation on the sphere.

Our approach can be used not only for camera motion, but also for template matching and target tracking on the sphere. The Fourier transform on the sphere is the well known spherical harmonic transform and the Euler angles of the group action are computed from the generalized shift theorem.

The recipe is known from abstract harmonic analysis and generalizes to non-commutative groups. Chirikjian and Kyatkin [7] have published a very good reference involving engineering applications. A very good introduction can be found in [9]. Previous work in computer vision on spherical harmonics addresses range images [3, 19, 6].

The contribution of this paper to computer vision is twofold:

- We present a new methodology on image processing of spherical images regarding rotations. Recall that even simple image shifts on omnidirectional images are actually rotations. Note that the sphere is the most convenient domain to work with. If we use tangential planes then we cannot capture larger deformations.

- As an example we show direct rotation estimation from the spherical harmonic coefficients. This involves a novel derivate of the spherical harmonic shift theorem that allows us to express all Euler angles in exponentials. Attitude estimation from global features can be regarded as a matching problem and can be tuned by choosing any of the high energy spherical coefficients we prefer.

How far can such an approach go beyond the sphere and a 3D-rotation acting on the sphere we do not know. We know that several results are applicable on the effect of 3D-rotations on the perspective plane. We discuss these future questions in the conclusion.

The closest approaches to ours are all approaches that compute collineations directly from image intensities [4, 23, 25] or their derivatives [21]. General differential motions of an calibrated camera are estimated in [11] and of an uncalibrated camera in [5]. The disadvantage of existing direct methods is that they rely on a series of iterations involving a warping of the image at every step and in principle they are iterative closest point algorithms. Our method would be comparable to the derivative-based methods when applied hierarchically but our harmonic decomposition allows us to use any of the energy components we want. We indeed implement and compare with our approach a simple version of estimation of small rotational motion from the brightness change constraint equation.

We will start the paper with the math preliminaries presenting comprehensively the spherical Fourier transform. For a superficial survey of the terms used, the reader is referred to the appendix. Then we present the shift theorem and a novel corollary useful for our algorithm in the subsequent section. Experiments are conducted with real images, with artificial or real motion, and our estimation is compared to a conventional flow technique. We conclude with an outlook and a short description of the biological relevance.

2 The Spherical Fourier Transform

The treatment of spherical harmonics is based on [10, 1] We parameterize points on the unit sphere S^2 using traditional spherical coordinates, with θ representing the angle of colatitude ($\theta \in [0, \pi]$), and ϕ the angle of longitude ($\phi \in [0, 2\pi)$). Thus, any point $\eta(\theta, \phi) \in S^2$ is uniquely written as the unit vector

$$\eta(\theta, \phi) = (\cos\phi \sin\theta, \sin\phi \sin\theta, \cos\theta).$$

We will write $f(\eta)$ for any function whose domain is restricted to the unit sphere.

The rotation of a point $\eta \in S^2$ is performed by an element g of the rotation group $SO(3)$. In this paper we will parameterize $SO(3)$ with ZYZ Euler Angles, for reasons that will be evident later. That is to say, any rotation $g(\alpha, \beta, \gamma) \in SO(3)$ can be written as

$$g(\alpha, \beta, \gamma) = R_z(\gamma) R_y(\beta) R_z(\alpha)$$

where $R_z(\alpha)$ and $R_y(\beta)$ represent respectively a rotation around the z-axis by α radians and a rotation around the y-axis by β radians. For any $g \in SO(3)$ and function $f(\eta)$, we define the rotation of $f(\eta)$ with the operator Λ_g such that

$$\Lambda_g f(\eta) = f(g^{-1}\eta)$$

We also define the integration of a function $f(\eta) \in L^2(S^2)$ as

$$\int_{\eta \in S^2} f(\eta) d\eta$$

where $d\eta = \sin\theta d\theta d\phi$ is the rotation-invariant volume measure on the sphere.

Having introduced the notation for the integration and rotation of a function $f(\eta) \in L^2(S^2)$, we proceed to introduce the expansion of functions on the sphere.

In traditional Fourier analysis, a basis for all functions on the line is generated by examining the eigenfunctions of the Laplacian restricted to the circle. Similarly, on the sphere, the eigenfunctions of the spherical Laplacian (the laplacian in $\mathbb{R}^3$ restricted to the sphere) are the spherical harmonic functions $Y_{lm} : S^2 \mapsto \mathbb{C}$. These eigenfunctions form an eigenspace of harmonic homogeneous polynomials which have dimension $2L + 1$. Thus, the $2L + 1$ spherical

harmonics for each $L \geq 0$ form a basis for any $f(\eta) \in S^2$. The $(2l+1)$ spherical harmonics of degree l are given as

$$Y_m^l(\theta,\phi) = (-1)^m \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} P_m^l(\cos\theta) e^{im\phi}$$

where P_m^l are the associated Legendre Functions

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

and the normalization factor is chosen to satisfy the orthogonality relation

$$\int_{\eta \in S^2} Y_{lm}(\eta) Y_{l'm'}(\eta) d\eta = \delta_{mm'}\delta_{ll'},$$

where δ_{ab} is the Kronecker delta function.

Since spherical harmonics provide an orthonormal basis for $L^2(S^2)$, any function $f(\eta) \in L^2(S^2)$ can be expanded in this basis:

$$\begin{aligned} f(\theta,\phi) &= \sum_{l\in\mathbb{N}} \sum_{|m|\leq l} \hat{f}(l,m) Y_l^m(\theta,\phi) \\ \hat{f}(l,m) &= \int_{\eta\in S^2} f(\eta) \bar{Y}_l^m(\eta) d\eta \end{aligned}$$

The f_{lm} are commonly referred to as the spherical harmonic coefficients of $f(\eta)$.

2.1 The Shift Theorem

The shift theorem for functions on the line examines the effect on the Fourier space of translations in the function space. Analogously, for $f(\eta) \in L^2(S^2)$, we would like to study what effect a rotation in the function space has on the function's representation in the frequency space. We first note that a rotated spherical harmonic of degree l can be written as a linear combination of spherical harmonics of the same degree:

$$\Lambda_g Y_{lm}(\eta) = \sum_{|n|\leq l} U_{mn}^l(g) Y_{ln}(\eta),$$

where the U_{mn}^l are the irreducible unitary representations of the rotation group $SO(3)$:

$$U_{mn}^l(g(\gamma,\beta,\alpha)) = e^{-im\gamma} P_{mn}^l(\cos(\beta)) e^{-in\alpha}.$$

The P_{mn}^l are generalized associated Legendre polynomials which can be calculated efficiently using recurrence relations.

Each spherical harmonic coefficient f_{lm}^g of a rotated function $\Lambda_g f(\eta)$ is a linear combination of the coefficients of the original function $f(\eta)$:

$$\begin{aligned} \Lambda_g \hat{f}_{lm} &= \int_{\eta'\in S^2} f(\eta') \bar{Y}_{lm}(g\eta') d(\eta'), \eta' = g^{-1}\eta \\ &= \int_{\eta'\in S^2} f(\eta') \sum_{|p|\leq l} U_{mp}^l(g^{-1}) \bar{Y}_{lp}(\eta') d(\eta') \\ &= \sum_{|p|\leq l} U_{mp}^l(g^{-1}) \int_{\eta'\in S^2} f(\eta') \bar{Y}_{lp}(\eta') d(\eta') \\ &= \sum_{|p|\leq l} \hat{f}_{lp} U_{pm}^l(g). \end{aligned}$$

Suppose we are given two spherical images, I_f and I_g, where I_g is related to I_f by some unknown rotation $R \in SO(3)$. We can use equation (1) to solve for R.

2.2 Image Invariants

We would like to recognize properties of functions that remain invariant under rotations. Specifically, we would like to identify real-valued functions $K(f(\eta))$ such that $K(f(\eta)) = K(\lambda_g f(\eta))$ for any $g \in SO(3)$. The shift theorem allows us to easily identify image invariants. We define the invariant function $K_l(f(\eta)$ as

$$K_l(f(\eta)) = \sum_{|m|\leq l} \bar{f}_{lm} f_{lm}$$

where f_{lm} are the coefficients of $f(\eta)$. If we let f_l be the $(2l+1)\times 1$ column vector of all coefficients of degree l we easily verify that $K(f(\eta))$ is indeed invariant under rotations of $f(\eta)$:

$$\begin{aligned} K_l(\Lambda_g f(\eta)) &= \sum_{|m|\leq l} \Lambda_g \bar{f}_{lm} \Lambda_g f_{lm} \\ &= \sum_{|m|\leq l} (\sum_{|p|\leq l} \bar{f}_{lp} \bar{U}_{pm}^l)(\sum_{|p|\leq l} f_{lp} U_{pm}^l) \\ &= (\bar{f}_l)^T \bar{U}^l (U^l)^T f_l \\ &= K(f(\eta)) \end{aligned}$$

3. Simplifying rotation estimation

Image invariants offer an efficient and accurate technique to determine if any two functions on the sphere are rotated relative to each other. Once it has been verified that two functions $f(\eta), h(\eta) \in L^2(S^2)$ are indeed related by a rotation ($\exists g \in SO(3)$ such that $f = \Lambda_g h$), we would like to obtain the Euler angles of the rotation. The method we present here is to utilize the shift theorem (1) in order to retrieve the parameters of the rotation.

If we examine (1), we see that the Euler angle corresponding to a rotation about the y-axis, β, is contained

within the generalized associated Legendre polynomials P^l_{mn}. For the purposes of estimation, the nature of these polynomials will make it quite computationally difficult to accurately extract β. First, we recognize that any $g \in SO(3)$ can be written as $g(\alpha,\beta,\gamma) = g_1(\alpha + \frac{\pi}{2}, \frac{\pi}{2}, 0) g_2(\beta + \pi, \frac{\pi}{2}, \gamma + \frac{\pi}{2})$. The unitarity of the representations U^l gives

$$U^l_{mn}(g_1 g_2) = \sum_{|p|\leq l} U^l_{mp}(g_1) U^l_{pn}(g_2)$$

We can rewrite (1) as

$$\begin{aligned}\Lambda_{g(\alpha,\beta,\gamma)} f_{lm} &= \sum_{|p|\leq l} f_{lp} \sum_{|k|\leq l} U^l_{pk}(g_1(\alpha + \frac{\pi}{2}, \frac{\pi}{2}, 0)) \cdot \\ &\quad U^l_{km}(g_2(\beta + \pi, \frac{\pi}{2}, \gamma + \frac{\pi}{2}))\end{aligned}$$

to obtain the following:

Corollary 1 *If f_{lm} and f^g_{lm} are coefficients before and after the $SO(3)$ action g, they are related as follows*

$$\begin{aligned}f^g_{lm} &= e^{-im(\gamma+\frac{\pi}{2})} \sum_{|p|\leq l} e^{-ip(\alpha+\frac{\pi}{2})} f_{lp} \cdot \\ &\quad \sum_{|k|\leq l} P^l_{pk}(0) P^l_{km}(0) e^{-ik(\beta+\pi)} f_{lp}\end{aligned}$$

It is important to note that the polynomials $P^l_{mn}(0)$ can now be calculated easily, and that all three Euler angles now appear only in exponential terms. We can take advantage of this corollary to produce the following lemma:

Lemma 2 *If f_{lm} and f^g_{lm} are coefficients before and after the $SO(3)$ action g, then for any $l > 0$, l even, we can write a linear system of the form $Ax = b$ in terms of exponentials of the Euler angle β.*

Proof: We rewrite Corollary 1 as

$$\begin{aligned}\Lambda_g f_{lm} &= \sum_{|p|\leq l} e^{-ip\beta} C^l_{mp} \\ C^l_{mp} &= e^{-im(\gamma+\frac{\pi}{2})} \cdot \\ &\quad (\sum_{|l|\leq l} e^{-ik(\alpha+\frac{\pi}{2})} f_{lk} P^l_{kp}(0) P^l_{pm}(0) e^{-ik\pi})\end{aligned}$$

Notice that for an even l, we have if $|m|$ odd, $|p|$ even, then $C^l_{mp} = -C^l_{(-m)p}$. If $|m|$ even, $|p|$ even, then $C^l_{mp} = C^l_{(-m)p}$. With this we can reduce the size of our equations to get

$$f^g_{lm} + (-1)^m f^g_{l(-m)} = \sum_{|p|\leq l, podd} e^{-ip\beta}(C^l_{mp} + (-1)^m C^l_{(-m)p}$$

The number of terms in the summation is l, and so if we create such an equation for each $1 \leq m \leq l$, we are left with a linear system of the traditional form $Ax = b$, where A is an $l \times l$ matrix independent of β, x is a vector dependent only on the exponentials of β, and b is a vector dependent only on the rotated coefficients. Remark: A is full-rank if there is sufficient spatial orientation in the images. An aperture theorem would inform us of the conditions on the orientations that would lead to singularities in A.

3.1 Estimation Algorithm

We introduce a method for estimating the rotation g between any pair of functions $f(\eta), h(\eta) \in L^2(S^2)$. We begin by solving a few simple cases first.

3.1.1 Rotation about Z

The first case we will examine is one where the second Euler angle β is zero. This is equivalent to having only a rotation around the Z-axis, and so without loss of generality we can also assume α is zero and thus we only need to estimate γ. We can rewrite Corollary 1 as

$$\begin{aligned}\Lambda_g f_{lm} &= e^{-im(\gamma+\frac{\pi}{2})} C_{lm} \\ C_{lm} &= \sum_{|p|\leq l} e^{-ip(\frac{\pi}{2})} f_{lp} \cdot \\ &\quad \sum_{|k|\leq l} P^l_{pk}(0) P^l_{km}(0) e^{-ik(\pi)} f_{lp}\end{aligned}$$

Since we have only one unknown, we clearly have an overconstrained problem in γ. So, for any $l \geq 1$ we choose, we have the following minimization problem:

$$\sum_{1\leq p\leq l} \sum_{0\leq m\leq l} (e^{-im(\gamma+\frac{\pi}{2})} C_{pm} - \Lambda_g f_{pm})^2 = 0$$

This system can be minimized easily using nonlinear techniques such as a quasi-newton approach like Broyden's method. This method can also be trivially adapted to the case where α and β are known.

3.1.2 Estimating β

The next case we study is one where only the rotation around the Y-axis needs to be estimated. In this case, we assume α and γ are known. However, Lemma 2 presents us with a system that is linear in exponentials of β, so we can simply use a non-linear minimization technique as described earlier to form an estimate for β.

3.1.3 Estimating α, β, and γ

If all three Euler angles are unknown, we will use a full nonlinear global minimization approach. One simplification we

will take is that we will not minimize with respect to all three unknowns simultaneously.

The first step is to minimize for α and β. For any coefficient f_{lm} with $m = 0$, we see that Corollary 1 has no dependence on γ. Thus, we have an overconstrained system in the two unknowns α and β. We form this system from all equations (derived from Corollary 1) for f^g_{p0}, $p \leq l$, for some $l > 1$. We can use Broyden's method to minimize for α and β. Once we have estimates for α and β, the second step is to estimate γ. This method was described in section 3.1.1

4 Experiments

In this section we present some results obtained from the implementation of our rotation estimation algorithms. In the case where all three Euler angles are estimated, we use an optical-flow based technique for comparison purposes.

4.1 Spherical Images

Catadioptric systems with a unique effective viewpoint have been proven to be convex reflective surfaces of revolution with a parabolic or hyperbolic profile. Geyer and Daniilidis [14] showed that such projections are equivalent with a projection on the sphere followed by a projection from a point on the sphere axis to the plane. In the parabolic case, the second projection is a stereographic projection from the sphere to the catadioptric plane (also the image plane):

$$\begin{aligned} u &= \cot\frac{\theta}{2}\cos\phi \\ v &= \cot\frac{\theta}{2}\sin\phi. \end{aligned}$$

Given a calibrated camera and catadioptric image $I(u, v)$ we define its inverse stereographic mapping onto the sphere as

$$I_S(\theta, \phi) \stackrel{def}{=} I(\cot\frac{\theta}{2}\cos\phi, \cot\frac{\theta}{2}\sin\phi).$$

This mapping allows us to interpolate an image on the sphere given a catadioptric image. The range of this mapping is only limited by the field of view of the original catadioptric system, and so to fully image the sphere a 360° field of view catadioptric system would be required.

4.2 Image generation

To obtain our spherical images, we used a catadioptric system consisting of a Nikon Coolpix 995 digital camera along with a parabolic mirror attachment produced by Remote Reality. The field of view of the mirror used is 212°. The size of the original catadioptric images was 2048×1536 pixels without compression, and the parabolic mirror filled up a region of approximately 1400×1400 pixels. The images are mapped to the sphere by interpolating onto the θ-ϕ plane, where sampling is uniform. Figure 1 shows a sample catadioptric image obtained from a parabolic mirror and its corresponding projection onto the sphere.

Figure 1: On the left is a parabolic catadioptric image. In the middle is the spherical image represented on the θ-ϕ plane, and on the right is the image displayed on the sphere.

4.3 Discrete Spherical Harmonic Transform

Suppose we are given samples of a function $f(\eta)$. We would still like to be able to calculate the Spherical Harmonic coefficients of the original function using only the samples. Driscoll and Healy have shown that if a band-limited function $f(\eta)$ has band limit b ($f_{lm} = 0, l > b$), then the function can be recovered exactly from the samples, provided the samples are uniform in θ and ϕ and there are at least $2b$ samples in each. This is important because it allows us to compute the Spherical Harmonic coefficients using only the samples of the original function:

$$f_l m = \frac{\pi^2}{2b^2}\sum_{j=0}^{2b}\sum_{k=0}^{2b-1} f(\eta_{jk})\bar{Y}_{lm}(\eta_{jk}), l \leq b \; and \; |m| \leq l$$

where $\eta_{jk} = \eta(\theta_j, \phi_k)$, and $\theta_j = \frac{\pi j}{2b}$, $\phi_k = \frac{\pi k}{b}$. These coefficients can be computed efficiently (with an FFT along ϕ). For more information, readers are referred to Driscoll and Healy [10, 1].

4.4 Comparison method

We also implement a rotation-estimation based on optical flow for comparison purposes. To calculate the flow on the sphere, we use the technique described in [8]. The flow equation is

$$I_\theta \dot{p}_\theta + \frac{1}{\sin\theta} I_\phi \dot{p}_\phi = -I_t,$$

where I_θ, I_ϕ, and I_t are the spatial (in the θ,ϕ directions), and temporal image derivates. The flow vector on the plane tangent to the sphere is $(\dot{p}_\theta, \dot{p}_\phi, 0)^T$. If p is a point on the sphere and $\dot{p}$ is the flow at point p in cartesian coordinates, then we have $\dot{p} = \omega \times p$, where ω is the angular velocity about the rotation axis. After calculating the flow on the

sphere, we have an overconstrained system for the unknown ω. If we assume the the flow is small, we have $R = I + \hat{\omega}$, where $R = g(\alpha, \beta, \gamma)$ and $\hat{\omega}$ is the skew-symmetric matrix corresponding to ω.

4.5 Results

We first present the results of our estimation algorithm applied to artificially rotated spherical images. Figure 2 gives an example of a spherical image and its synthetically rotated counterpart. The first test we perform is the estimation of all thee Euler angles using the technique described in section 3.1.3. In this case the rotation is known, and so we can easily verify the accuracy of our estimation. We tested the algorithm with three different degrees of Spherical Harmonics (i.e. the system minimized only consisted of harmonics up to a certain degree). Table 1 presents the estimated rotation angles.

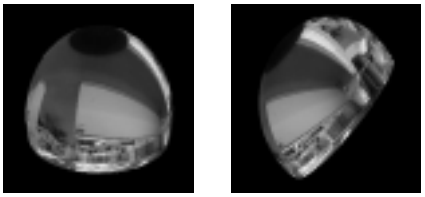

Figure 2: Rotated Images. On the left is the spherical image generated by the catadioptric system. On the right is the same image synthetically rotated by $g(\alpha, \beta, \gamma) = g(60^\circ, 45^\circ, 270^\circ)$.

Angle	$l \leq 5$	$l \leq 8$	$l \leq 16$
$\alpha = 60^\circ$	60.12°	59.98°	60.22°
$\beta = 45^\circ$	45.06°	45.03°	45.13°
$\gamma = 25^\circ$	24.91°	25.12°	25.10°

Table 1: Results from running non-linear minimization over all three angles for artificially rotated images. l value implies that no coefficients with degree greater than l were used.

We now examine the effect on our estimation technique when we violate the assumption of appearance persistence, which is modeled with clutter in the scene. Figure 3 shows two similar images, with the only difference being a human being has entered the scene in the cluttered image. We attempt to estimate the rotation between the original reference frame and a cluttered image which has been rotated relative to the reference frame. Table 2 displays some results for different amounts of clutter in the images.

In addition to clutter, we also study the estimation method's response to scalings of the image. Table 3 shows how accurately the rotation is estimated when the scale factor of 10% is used.

Figure 3: Clutter Images. On the left is a reference frame, and on the right is the same image with a human in the scene.

6% 10% 13%

Figure 4: Clutter images thresholded to display quantity of clutter. Percentages refer to the number of pixels affected.

We now present some results from images generated by physical rotations of the mirror. Since the parabolic mirror does not image an entire sphere, rotations of the mirror naturally reflect appearance of clutter in the scene. Rotating the mirror about its optical axis is equivalent to having only one non-zero Euler angle (Figure 5). We estimate this single angle using the technique described in section 3.1.1, and some results are shown in Table 4.

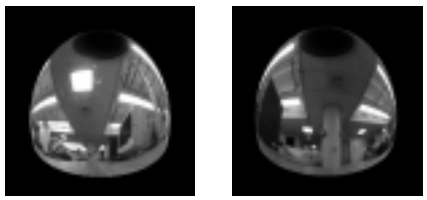

Figure 5: Rotated Images. Two spherical images from a sequence of catadioptric images rotated relatively about the mirror's optical axis.

We now attempt to estimate all three Euler angles from rotated image generated by physical rotations of the parabolic mirror. We use directly the technique explained in section 3.1.1. Ground truth was calculated with physical measurements and verified by fitting the global rotation to images. Results are in Table 5.

It is important to note that for a small rotation in real data, the Euler angle representing a rotation about the Y-axis (β) is underestimated. We can attribute this in part to the nature of the catadioptric camera setup. The bar that fixes the camera to the parabolic mirror is coincident with the camera's optical axis. This results in a black region around the north pole in all catadioptric images. For any

Angle	$l \leq 8$	$l \leq 16$	Flow	$l \leq 8$	$l \leq 16$	Flow	$l \leq 8$	$l \leq 16$	Flow
$\alpha = 15°$	$14.96°$	$14.96°$	$14.88°$	$14.57°$	$14.83°$	$14.76°$	$14.19°$	$14.19°$	$14.45°$
$\beta = 13.8°$	$13.87°$	$14.03°$	13.88	$13.87°$	$13.81°$	13.90	$13.96°$	$13.96°$	13.98
$\gamma = 12.8°$	$13.01°$	$12.89°$	$12.94°$	$13.11°$	$13.11°$	$13.41°$	$13.74°$	13.68	13.50

Table 2: Results from clutter estimation. The first column is the ground truth. Columns 2-4 are results of each technique for the first cluttered image (6% clutter). Columns 5-8 are for 10% clutter, and the remaining columns are the results on the image with 13% clutter.

Angle	$l \leq 5$	$l \leq 8$	$l \leq 16$
$\alpha = 18°$	$18.64°$	$18.46°$	$18.91°$
$\beta = 16.3°$	$17.61°$	$17.00°$	$17.15°$
$\gamma = 15°$	$14.63°$	$14.89°$	$14.77°$

Table 3: Results from running non-linear minimization over all three angles for images with scale variation.

Angle	$l \leq 5$	$l \leq 8$	$l \leq 16$
$\gamma = 30°$	$29.96°$	$29.99°$	$30.02°$
$\gamma = 36°$	$36.06°$	$35.93°$	$35.81°$
$\gamma = 45°$	$44.12°$	$44.65°$	$44.71°$

Table 4: Results from estimating γ as described in 3.1.1. Three different rotations were tested.

rotation about the Y-axis, this region remains unchanged. A similar effect occurs also at the border of the mirror. For example, regardless of rotation, the region imaged on the sphere does not change, which would not be the case if the entire sphere were imaged. Also, in comparison, calculation of the optical flow on the sphere is rather computationally expensive, whereas very few Spherical Harmonic coefficients need to be calculated in the approach presented here.

5. Conclusion

There is significant evidence from biological findings and in particularly in insects [12, 27] that motion perception and navigation are based on global matched filters implemented by neurons of almost spherical receptive fields. Our approach is not a model of such an architecture but provides the tools for analyzing whether it is possible. In particularly, it naturally contains the notion of scale in the order of the spherical harmonic coefficients. The results are very encouraging in particularly regarding the number of coefficients necessary to recover the unknown attitude. We might speculate that such coefficients are global encoders of the image and that rotations can be estimated from a simple look-up table associating coefficient combinations with angles.

In this article, we have treated only the pure rotation case. Though the main apparent motion in omnidirectional

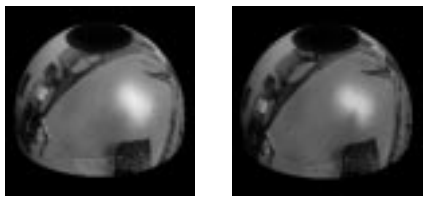

Figure 6: Images of real rotated images on the sphere (small rotation).

Angle	$l \leq 5$	$l \leq 8$	$l \leq 16$	Flow
$\alpha = -2.4°$	$-2.2°$	$-2.3°$	$-2.2°$	$-2.7°$
$\beta = 9.3°$	$7.2°$	$7.6°$	$7.3°$	$8.1°$
$\gamma = 2.2°$	$2.5°$	$2.4°$	$2.7°$	$1.9°$

Table 5: Rotation estimates for real data with small rotation.

images is due to rotation and our algorithm is robust to existence of small translation, it is interesting to extend it to general motions which still are group actions. Such ones have to obey to a global depth description, a good example would be mappings on the sphere analogous to collineations. We believe that configurations involving omnicams on air vehicles or mobile robots could thus be covered so that localization happens just appearance-based using some reference views.

A Informal math exposition

This section is not meant as a formal enumeration of the assumptions we make. It is a rather intuitive description of what is required to apply the recipe of a generalized Fourier transform. We assume that the reader is familiar with the concept of a group. Let us live with an intuitive definition of a *Lie group* that its elements are on smooth manifold and that the group operation and the inversion are smooth maps [2]. [1]The real line and the circle are Lie groups with respect to addition and well known matrix Lie groups are the general linear group of square invertible matrices, the rotation groups SO(n), and the Lorentz group.

[1]The only condition we miss from the formal definition that it has to be a topological group.

A group G is *acting on* a space X when there is a map $G \times X \rightarrow X$ such that the identity element of the group leaves X as is and a composition of two actions has the same effect as the action of the composition of two group operations (associativity). For example, the isometry group $SE(2)$ acts on the plane $\mathbb{R}^2$. The rotation group $SO(3)$ can act on the sphere $\mathbb{S}^2$. The set of all $gx \in X$ for any $g \in G$ is called the *orbit* of x. If the group possesses an orbit, that means for any $a, b \in X$, $ga = b$ for a $g \in G$, then the group action is called *transitive*. For example, there is always a rotation mapping one point on the sphere to another. If a subgroup H of G fixes a point $x \in X$ then H is called the *isotropy* group. A typical example of an isotropy group is the subgroup $SO(2)$ of $SO(3)$ acting on the north-pole of a sphere.

A space X with a transitive Lie group action G is called *homogeneous space*. If the isotropy group is H, it is denoted with G/H. The plane $\mathbb{R}^2$ is the homogeneous space $SE(2)/SO(2)$. The sphere $\mathbb{S}^2$ is the homogeneous space $SO(3)/SO(2)$.

Images are usually defined on homogeneous spaces and their deformations are the group actions. The question is now, for which groups does a Fourier transform exist? The answer requires us first to be able to integrate on the group and on the homogeneous space and second to find the Fourier basis analogous to $e^{ix\omega}$ on the real line. For the integration, we ask the group to be *unimodular*. We are not going into the issues of Haar measure existence, but we will just mention the fact that SO(3) is compact and that compact groups are unimodular. The Fourier series of a group exists if the group possesses an *irreducible unitary representation* where irreducible can be intuitively explained as non having trivial invariant subspaces [26]. Among other properties, if considered as a basis it is orthogonal. We are going to use the irreducible representation of $SO(3)$ which is the basis of the generalized Legendre polynomials $P^l_{mn}(\cos x)$.

References

[1] G.B. Arfken and H.J. Weber. *Mathematical Methods for Physicists*. Academic Press, 1966.

[2] A. Baker. *Matrix Groups*. Springer-Verlag, 2002.

[3] D. Ballard and C. Brown. *Computer Vision*. Prentice-Hall, Englewood-Cliffs, NJ, 1982.

[4] J.R. Bergen, P. Anandan, K.J. Hanna, and R. Hingorani. Hierarchical model-based motion estimation. In *Proc. Second European Conference on Computer Vision*, pages 237–252, Santa Margerita, Italy, May 23-26, G. Sandini (Ed.), Lecture Notes in Computer Science 588, Springer-Verlag, Berlin et al., 1992.

[5] T. Brodsky and C. Fermuller. Self-calibration from image derivatives. *International Journal of Computer Vision*, 48:91–114;, 2002.

[6] Th. Bulow. Spherical diffusion for 3d surface smoothing. In *Proc. 1st International Symposium on 3D Data Processing, Visualization, and Transmission*, Padova, Italy, 2002.

[7] G.S. Chirikjian and A.B. Kyatkin. *Engineering Applications of Noncommutative Harmonic Analysis: WIth Emphasis on Rotation and Motion Groups*. CRC Press, 2000.

[8] K. Daniilidis, A. Makadia, and T. Bülow. Image processing in catadioptric planes: Spatiotemporal derivatives and optical flow computation. In *Workshop on Omnidirectional Vision, Copenhagen, June 22*, pages 3–12, 2002.

[9] A. Deitmar. *A first course in harmonic analysis*. Springer Verlag, 2001.

[10] J.R. Driscoll and D.M. Healy. Computing fourier transforms and convolutions on the 2-sphere. *Advances in Applied Mathematics*, 15:202–250, 1994.

[11] C. Fermuller and J. Aloimonos. Direct perception of three-dimensional motion from patterns of visual motion. *Science*, 270:1973–1976, 1995.

[12] M.O. Franz and H.G. Krapp. Wide-field, motion- sensitive neurons and matched filters for optic flow fields. *Biological Cybernetics*, 83:185–197, 2000.

[13] J.-P. Gaulthier, G. Bornard, and M. Silberman. Motion and pattern analysis: harmonic analysis on motion groups and their homogeneous spaces. *IEEE Trans. Systems, Man, and Cybernetics*, 21:159–172, 1991.

[14] C. Geyer and K. Daniilidis. Catadioptric projective geometry. *International Journal of Computer Vision*, 43:223–243, 2001.

[15] E. Hall. *Computer Image Processing and Recognition*. Academic Press, 1979.

[16] A. Kelly. Mobile robot localization from large-scale appearance mosaics. *Intern. Journal of Robotics Research*, 19:1104–1125, 2000.

[17] R. Lenz. Rotation-invariant operators. In *Proc. Int. Conf. on Pattern Recognition*, pages 1130–1132, Paris, France, Sept. 28-3, 1986.

[18] A. Leonardis and M. Jogan. Robust localization using eigenspace of spinning-images. In *IEEE Workshop on Omnidirectional Vision, Hilton Head, SC, June 12*, pages 37–46, 2000.

[19] Art Matheny and Dmitry B. Goldgof. The use of three and four-dimensional surface harmonics for rigid and nonrigid shape recover y and representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 17(10):967–981, 1995.

[20] M.I. Miller and L. Younes. Group actions, homeomorphisms, and matching: A general framework. *International Journal of Computer Vision*, 41:61–84, 2001.

[21] S. Negahdaripour and B.K.P. Horn. Direct passive navigation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 9:168–176, 1987.

[22] J. Sato and R. Cipolla. Extracting group transformations from image moments. *Computer Vision and Image Understanding*, 73:29–42, 1999.

[23] H. S. Sawhney and R. Kumar. True multi–image alignment and its application to mosaicing and lens distortion correction. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages 450–456, Puerto Rico, June 17-19, 1997.

[24] J. Segman, J. Rubinstein, and Y. Zeevi. The canonical coordinates method for pattern deformation: Theoretical and computational considerations. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 14:1171–1183, 1992.

[25] H.-Y. Shum and R. Szeliski. Systems and experiment paper: construction of panoramic image mosaics with global and local alignment. *International Journal of Computer Vision*, 36:101–130, 2000.

[26] B. Simon. *Representations of finite and compact groups*. Grad, Studies in Math. AMS, 1996.

[27] M.V. Srinivasan. Ants match as they march. *Nature*, 392:660–661, 2001.

[28] M. Tanaka. On the representation of the projected motion group in 2+1d. *Pattern Recognition Letters*, 14:871–678, 1993.

[29] T.Pajdla and V.Hlavac. Zero phase representation of panoramic images for image based localization. In *Int. Conf. on Computer Analysis of Images and Patterns*, pages 550–557, Ljubljana, Slovenia, Sept. 1-3, 1999.

[30] C.T. Zahn and R.Z. Roskies. Fourier descriptors for plane close curves. *IEEE Trans. Computers*, 21:269–281, 1972.