Euclid meets Fourier: Applying harmonic analysis to essential matrix estimation in omnidirectional cameras

Christopher Geyer, *Shankar Sastry* and *Ruzena Bajcsy*
{cgeyer,sastry,bajcsy}@eecs.berkeley.edu
University of California, Berkeley

Abstract. By combining notions from geometry, signal processing and harmonic analysis, we propose a new method for the estimation of the motion between two omnidirectional cameras. We show that a densely sampled likelihood function can be obtained on the space of essential matrices via a convolution of two signals. The first signal expresses the epipolar geometry of two views, and the second signal encodes the similarity of intensities (or some other measure) between a pixel in one image and a pixel in another image. The proposed method is analogous to a Hough or Radon transform on the space of essential matrices, and is a first step to integrating signal processing and geometry. For computational reasons, we are not aware of researchers attempting a Hough transform on the space of essential matrices, so we are not aware of similar work. Nevertheless, there are some similarities between the proposed method and the recent work of Makadia and Daniilidis [1] and Wexler et al. [2]. In the former case the authors propose rotation estimation using a shift theorem in $SO(3)$, and the latter investigates the estimation of arbitrary epipolar geometries. The breakthrough in this paper is that we can efficiently compute the convolution using spherical and rotational harmonic representations of the signals. Estimation using the proposed method has several advantages: we can automatically represents ambiguities; we are able to estimate multiple motions; and we obtain a framework which can take into account arbitrary, non-Gaussian sensor noise models such as simple blob correspondence.

§ 1 Introduction

The problem of recovering the motion between two cameras is an old and important problem in computer vision. Its solution has applications in robotics, virtual reality and movie-making. In robotics, for example, digital cameras offer a passive way to correct and update odometry with respect to a world reference frame. Most methods for egomotion or structure-from-motion estimation in two views can be put into two categories: (1) small time-scale, small motion estimation utilizing optical flow or otherwise linearization of the brightness constancy constraint, including direct methods; or (2) discrete, wide baseline motion estimation by fitting an essential or fundamental matrix to point correspondences. This paper falls into the latter category, in the sense that we are interested in discrete, possibly large, motions. However, we aim, like some of the former methods, to use all of the information in contained in the signal to perform the estimation.

The spirit of two papers pervade this one. The first is recent work by Wexler at al. [2] on the non-parametric and featureless estimation of arbitrary epipolar geometries. In comparison, we exploit the symmetries of the epipolar geometry of calibrated cameras to regularize the estimation. The second is the work of Makadia and Daniilidis [1] where the authors propose a method of rotation estimation by taking advantage of a shift theorem for the spherical harmonic representation of functions on the sphere.

In this paper we combine these two approaches into one. We propose to develop an analog to the Hough transform for the space $\mathcal{E}$ of essential matrices. The transform defined here is not performed on a per-point basis as is done in voting schemes [3]. Instead, any point correspondence are represented using a function defined on the space of all point pairs, $S^2 \times S^2$—the direct product of the viewing sphere with itself—which is peaked at point correspondences. Though potentially any function encoding a likelihood of correspondence could be used. The transform is computed by convolving said function with an epipolar mask, equivalently, the characteristic function of the set of points obeying a canonical epipolar constraint. This transform can be computed efficiently using spherical and rotational harmonics. An advantage of this formulation is that the result can be interpreted as a likelihood function evaluated discretely on $\mathcal{E}$. It is trivial to find a global maximum of the function, and furthermore ambiguities are automatically exhibited in the likelihood.

What can this paper contribute to the enormous body of literature relating to relative pose estimation? First, we believe new problems are driving new paradigms for structure-from-motion problems. For example, given a distributed, wireless network of cameras, what is the minimum amount of information that is necessary to transmit between cameras to recover the relative positions? In the absence of reliable correspondences, how can we approach these problems? In this paper we argue for a signal processing approach and propose to develop tools to jointly analyze signals and geometry.

What follows is a review of two-view geometry and the typical paradigm for structure-from-motion estimation. We then show that a Hough transform on $\mathcal{E}$ is equal to a convolution, and then develop the tools to efficiently evaluate such a convolution. We end with a description of an implementation and several experiments.

§ 2 Two-view geometry

This section contains a brief review of two-view geometry, we refer to [4] for further reference. We assume familiarity with camera matrices and some projective geometry. Let P and Q be the projection matrices induced by two calibrated perspective cameras. If $p = PX$ and $q = QX$, then it is known that the pair (p, q) is constrained to lie in a proper subset of $\mathbb{P}^2 \times \mathbb{P}^2$—the direct product of the projective plane with itself. In particular if we assume that $P = (I, 0)$ and $Q = (R, t)$ for some $R \in \mathrm{SO}(3)$ and $t \in \mathbb{R}^3$, then there is a bilinear constraint on p and q, namely:

$$q^T E p = 0 \tag{1}$$

where E is known as an *essential matrix*, and is equal to $\hat{t}R$, where $\hat{t}$ is the skew symmetric matrix such that $\hat{t}s = t \times s$ for all s. Equation (1) is known as the *epipolar constraint.*

If $\mathcal{E}$ denotes the set of all essential matrices, then it is known that a given 3×3 matrix is an element of $\mathcal{E}$ if and only if its two non-zero singular values are equal. Consequently any essential matrix E is at most pre- and post-multiplications by special orthogonal matrices from the *canonical essential matrix* $E_0 = \hat{z}$, i.e.:

$$E = U E_0 V^T \tag{2}$$

where $U, V \in \mathrm{SO}(3)$. This demonstrates that $\mathrm{SO}(3) \times \mathrm{SO}(3)$ parameterizes $\mathcal{E}$. An explicit parameterization can be obtained by using ZYZ-Euler coordinates to parameterize each rotation matrix, i.e., $R(\theta, \phi, \psi) = e^{\theta\hat{z}} e^{\phi\hat{y}} e^{\psi\hat{z}}$. Then a 6-tuple $(\theta_1, \phi_1, \psi_1, \theta_2, \phi_2, \psi_2)$ yields:

$$\underbrace{e^{\theta_1\hat{z}} e^{\phi_1\hat{y}} e^{\psi_1\hat{z}}}_{U} E_0 \underbrace{e^{-\psi_2\hat{z}} e^{-\phi_2\hat{y}} e^{-\theta_2\hat{z}}}_{V^T}.$$

This follows from the rules $\hat{x}^T = -\hat{x}$ and $\left(e^A\right)^T = e^{A^T}$. Since E_0 commutes with rotations about the z-axis, i.e., $e^{\psi_1\hat{z}} E_0 e^{-\psi_2\hat{z}} = e^{(\psi_1-\psi_2)\hat{z}} E_0$ the parameterization is redundant. Thus we can reduce the number of parameters by one, considering only the angle $\delta = \psi_1 - \psi_2$.

For a given essential matrix E, equation (1) defines a subset in the space of image pairs (p, q):

$$\mathcal{P}_E = \left\{(p, q) \in S^2 \times S^2 : q^T E p = 0\right\}.$$

It is reasonable to substitute $S^2 \times S^2$ for $\mathbb{P}^2 \times \mathbb{P}^2$, where S^2 is the unit sphere in $\mathbb{R}^3$, as long as we maintain the equivalence relation induced by $\mathbb{P}^2$. However, this precaution is not necessary and later we will show that cheirality ambiguities can be resolved automatically. Continuing, by equation (2), there is a relationship between any $\mathcal{P}_E$ and $\mathcal{P}_{E_0}$, namely

$$\mathcal{P}_E = (U^{-1}, V^{-1}) \cdot \mathcal{P}_{E_0} ,$$

denoting the pairwise multiplication of every element of $\mathcal{P}_{E_0}$ by (U^{-1}, V^{-1}). Consequently the set of subsets $\{\mathcal{P}_E\}$ can be indexed by 5-tuples $(\theta_1, \phi_1, \delta, \theta_2, \psi_2)$.

§ 3 Essential matrix estimation

Motion estimation is now equivalent to finding an essential matrix E such that $\mathcal{P}_E$ contains a given set of point pairs $\{x_i = (p_i, q_i)\}$. Most methods to estimate E rely on regression, wherein one generally supposes that the data set $\{x_i\}$ satisfies an implicit equation $f_\xi(x) = 0$, where f_ξ is taken from some family indexed by $\xi \in \mathcal{X}$. An estimate $\hat{\xi}$ is obtained from the data by minimizing the cost function $\sum_i \rho(f_\xi(x_i))$, over ξ, where for example $\rho(x) = x^2$. In this case $f(p, q) = q^T E p$ and $\mathcal{X} = \mathcal{E}$.

An alternative, not yet used for essential matrix estimation, is the Hough transform, which was motivated by the problem of finding subatomic particle tracks in particle acceleration experiments [5]. The idea has been generalized to the aforementioned regression problem as follows: (1) begin with an accumulator image in the parameter space $\mathcal{X}$; (2) for every data point x_i increment those bins corresponding to parameters $\xi \in \mathcal{X}$ such that $f_\xi(x) < \varepsilon$; (3) ideally, the true model exhibits itself as a peak in the accumulator, and the location of this peak is the estimate.

Because the space of essential matrices is five dimensional, the accumulator for a Hough transform on a discretization of $\mathcal{E}$ can be quite large. In addition, computation is costly since the transform requires that for each data point one determines those bins that need to be incremented. We presume the computational costs, storage requirements, and the existence of other robust methods are reasons for the understandable lack of interest in a Hough transform on the space of essential matrices. The following sections, however, show that the computational costs can be reduced and some of these objections can be overcome.

§ 4 A Hough transform on $\mathcal{E}$

The goal of this section is to show that a Hough transform on $\mathcal{E}$ can be written as a kind of convolution. We suppose that (p_i, q_i) is a list of point correspondences and that h is their Hough transform on $\mathcal{E}$. In particular, for every essential matrix E we expect $h(E)$ to equal the number of pairs satisfying $|p_i^T E q_i| < \varepsilon$ for some ε. Let $\mathcal{P}_{E,\varepsilon}$ be the set of point pairs in $S^2 \times S^2$ which satisfy this inequality. Recall that $\mathcal{P}_E$ is the set of pairs exactly satisfying the epipolar constraint and therefore equals the intersection of all $\mathcal{P}_{E,\varepsilon}$.

In order to count the number of point pairs lying within a given $\mathcal{P}_{E,\varepsilon}$ we integrate the product $f_{E,\varepsilon} \cdot g$, where $f_{E,\varepsilon}$ is the characteristic function of $\mathcal{P}_{E,\varepsilon}$ and g is a sum of Dirac delta functions centered at point correspondences. In particular, we let $f_{E,\varepsilon}(p,q) = 1$ if $(p,q) \in \mathcal{P}_{E,\varepsilon}$ and 0 otherwise, and g is defined to be:

$$g(p,q) \;=\; \textstyle\sum_{i=1}^{n} \delta\big[(p,q) - (p_i,q_i)\big] \;\approx\; \sum_{i=1}^{n} e^{-\lambda\left[\left(cos^{-1}p^T p_i\right)^2 + \left(cos^{-1}q^T q_i\right)^2\right]}. \quad (3)$$

The product $f_{E,\varepsilon} \cdot g$ effectively masks out all point pairs lying outside of $\mathcal{P}_{E,\varepsilon}$; its integral counts only those point pairs lying within $\mathcal{P}_{E,\varepsilon}$, and therefore,

$$h(E) \;= \int\limits_{S^2\times S^2} g(p,q) f_{E,\varepsilon}(p,q)\, dp\, dq \;= \int\limits_{S^2\times S^2} g(p,q) f_{E_0,\varepsilon}(V^{-1}p, U^{-1}q)\, dp\ dq, \qquad (4)$$

for *any* $U, V \in \mathrm{SO}(3)$ such that $E = UE_0V^T$. Since any pair of rotations (U,V) determines an essential matrix, we can redefine h so as to be a function of the pair. Recall, though, that using a pair of matrices to parameterize $\mathcal{E}$ introduces a one-parameter redundancy. In this case, the redundancy exhibits itself because of the invariance of f_{E_0} to identical rotations about the z-axis, i.e., $f_{E_0,\varepsilon}(e^{\theta\hat{z}}p, e^{\theta\hat{z}}q) = f_{E_0,\varepsilon}(p,q)$ for all θ and $p, q \in S^2$.

For now, ignore the redundancy and notice that equation (4) is the convolution of f_{E_0} and g, where in the usual case what would be the translation of the kernel we have pair-wise left-multiplication by (U^{-1}, V^{-1}). Thus the convolution takes two functions on $S^2 \times S^2$ and yields a function on $\mathrm{SO}(3) \times \mathrm{SO}(3)$. How can we compute this convolution efficiently? In the next two sections we show that using harmonic representations of f_{E_0} and g allow efficient calculation of such a convolution.

§ 5 Rotational harmonics

Fast convolution algorithms for $2\,\pi$-periodic functions are well-known, and often used in signal processing. We aim in this section to relate some of the corresponding results for functions on SO(3), the space of rotations. To begin with, note that the usual Fourier series representation is simply a projection onto the orthonormal basis $\{e_k(\theta) = e^{-i\theta k}\}_{k\in\mathbb{Z}}$. Each e_k satisfies $e_k(\theta+\phi) = e_k(\theta)e_k(\phi)$ and is therefore a homomorphism from $[0, 2\pi)$ to SU(1)—the set of unit complex numbers. It should be easy to convince oneself that the homomorphism property gives rise to the shift and convolution theorems for Fourier series. In fact, there is a similar set of basis functions for functions for SO(3), sharing the homomorphism property and giving rise to its own shift and convolution theorems.

In particular, the corresponding basis for functions on SO(3) is generated by the following family of functions written in terms of ZYZ Euler coordinates:

$$D^l_{m,n}(e^{\theta\hat{z}} e^{\phi\hat{y}} e^{\psi\hat{z}}) = e^{-im\theta} P^l_{m,n}(\cos\phi) e^{-in\psi},$$

where l is non-negative, $|m|$ and $|n| \leqslant l$, and the $P^l_{m,n}(x)$ are called *associated Legendre polynomials*. The $D^l_{m,n}$ are known as the Wigner D-functions and references on their properties can be found in [6] and [7]. Fig. 1 shows contour plots

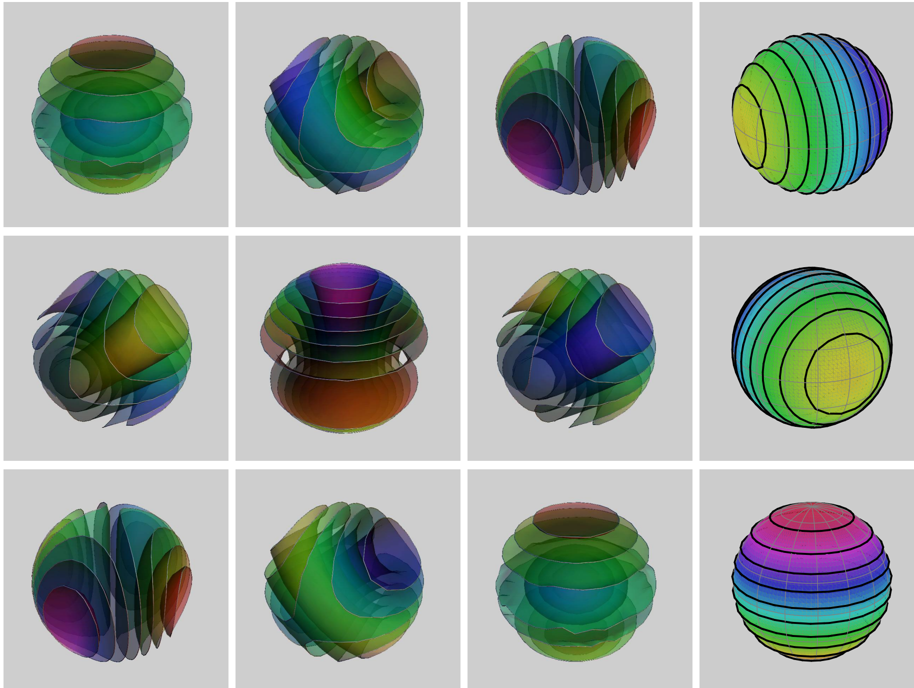


Fig. 1. Three left columns: spatial contour plots of the real parts of first degree basis functions for $L^2(\mathrm{SO}(3))$ in log-space, i.e., $D^l_{m,n}(e^{\hat{p}})$ for $m = -1, 0, 1$, $n = -1, 0, 1$ and $|p| \leq \pi$. Right column: contour plots of real and imaginary parts of first degree spherical harmonic basis functions, described in §6.

of these basis functions for $l = 1$. The properties of the functions $D^l_{m,n}$ are more apparent if we collect them all in a matrix for constant l. In particular, for each l we can form the $(2\,l + 1) \times (2\,l + 1)$ matrix

$$D_l(R) = \begin{pmatrix} D^l_{-l,-l}(R) & \cdots & D^l_{-l,0}(R) & \cdots & D^l_{-l,l}(R) \\ \vdots & \ddots & \vdots & & \vdots \\ D^l_{0,-l}(R) & \cdots & D^l_{0,0}(R) & \cdots & D^l_{0,l}(R) \\ \vdots & & \vdots & \ddots & \vdots \\ D^l_{l,-l}(R) & \cdots & D^l_{l,0}(R) & \cdots & D^l_{l,l}(R) \end{pmatrix}.$$

One can show, see [6] for example, that the resulting mapping from the space of rotations to a subset of $(2\,l + 1) \times (2\,l + 1)$ matrices is a homomorphism. In particular, for all R and S in SO(3) one has: (a) $D_l(R^T) = D_l(R)^\dagger$; (b) $D_l(R \cdot S) = D_l(R) \cdot D_l(S)$; and (c) $D_l(R) \cdot D_l(R)^\dagger = I$, where $X^\dagger = \bar{X}^T$. Thus D_l is a homorphism from SO(3) to SU$(2\,l + 1)$.

In continuation with the analogy with e_k, the *rotational harmonic coefficients* of some $f : \mathrm{SO}(3) \to \mathbb{C}$ are the projections to the set of orthonormal functions

$\{D^l_{m,n}\}$:

$$f^l_{m,n} = \int_0^{2\pi}\int_0^{\pi}\int_0^{2\pi} f(\theta,\phi,\psi) D^l_{m,n}(\theta,\phi,\psi)^\dagger \cos\phi \, d\theta \, d\phi \, d\psi \, .$$

We use the measure $dR = \cos\phi \, d\theta \, d\phi \, d\psi$, which is known as the Haar measure [?,?] and has the property that $\int f dR$ is invariant to pre- or post-rotation of f. For convenience we write f_l to denote the matrix of degree-l coefficients. The homomophism property of the D_l gives rise to shift and convolution theorems. For example, if we define shift operators

$$(\Lambda_R f)(S) \;=\; f(RS) \quad \text{and} \quad (\Gamma_R f)(S) \;=\; f(SR) \, ,$$

then by the homomorphism property, a constant $D_l(R)$ factors and the shifted coefficients obey:

$$(\Lambda_R f)_l \;=\; D_l(R^{-1}) \cdot f_l \quad \text{and} \quad (\Gamma_R f)_l \;=\; f_l \cdot D_l(R^{-1}) \, ,$$

i.e., matrix multiplication by $D_l(R^{-1})$ on the left or right for each individual l-degree matrix of coefficients. Furthermore, if we define the convolution of two functions $f, g : \mathrm{SO}(3) \to \mathbb{C}$ to be:

$$(f \star g)(R) = \int f(S) g(S^{-1}R) dR \tag{5}$$

then one can show that

$$(f \star g)_l = f_l \cdot g_l,$$

which is again matrix multiplication of the degree-l coefficients yielding the equivalent convolution result for rotational harmonics.

We have implemented a fast, discrete rotational harmonic transform (RHT) consisting of a FFT in the θ and ψ coordinates, evaluated at $\theta_i, \psi_i = (i + \frac{1}{2})\pi/l$, $0 \leqslant i < l$, followed by a projection to stored vectors of $p^l_{m,n} = [P^l_{m,n}(\cos\phi_i)]$ evaluated at $\phi_i = (i + \frac{1}{2})\pi/2l$, $0 \leqslant i < l$. In total this requires $O(l^4)$ computation and $O(l^4)$ storage. The inverse transform has the same costs. In [?,?] the authors prove that such a scheme is exact (in exact arithmetic) and that there is a corresponding notion of a band-limited function in which an L-band-limited function satisfies $f_l = 0$ for all $l \geqslant L$.

§ 6 Spherical harmonics and convolution on the sphere

One possible definition of the sphere is that it is a quotient space of SO(3). Consequently, spherical harmonics are a special case of the rotational harmonics. The quotient is obtained by identifying S^2 with an equivalence class on SO(3), where $R \sim S$ if and only if $Rz = Sz$, with $z = (0,0,1)^T$. The point on the sphere corresponding to some R is the image $p = Rz$. Since $Rz = z$ if and only if $R = e^{\theta \hat{z}}$, we find that $R \sim S$ if and only if $S = Re^{\theta \hat{z}}$. For any function f

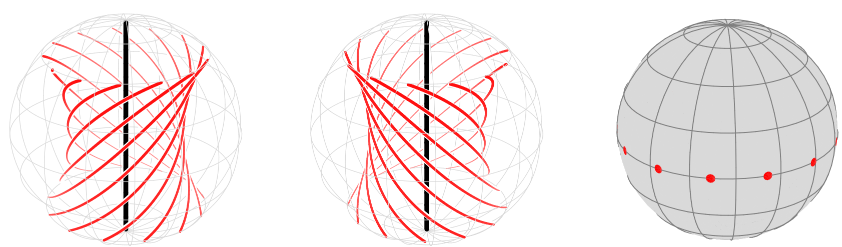
defined on SO(3)/ $\sim$, we expect that $f(R) = f(S)$ for all $R \sim S$. This condition is equivalent to the requirement that $f(R)$ be equal, almost everywhere, to its average over its equivalence class:

$$f(R) = \frac{1}{2\pi} \int f(R\, e^{\theta \hat{z}})\, dR\,. \tag{6}$$

This puts a condition on f_l, in particular that $f_l = f_l \cdot (\int \frac{1}{2\pi}\, D^l(e^{\theta\hat{z}})\, d\theta)$. Since $D^l_{m,n}(e^{\theta\hat{z}}) = \delta_{m,n} e^{-im\theta}$, the integral produces a matrix consisting of a single one in the middle column and row, yielding the equation

$$f_l = f_l \cdot \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix},$$

thereby constraining all but f_l's middle column to be zero. We can therefore treat f_l as a single column vector. The spherical harmonic transform (SHT) is then the projection of functions on $S^2 \cong$ SO(3)/ $\sim$ onto the basis $\{D^l_{m,0}(R)\}$.



Left cosets Right cosets Correspondence with sphere

Fig. 2. Left and right cosets of SO(3) in log-space. Each curve represents a coset corresponding to a point on the sphere. The z-axis corresponds to the subgroup $\{e^{\theta\hat{z}}\} \simeq SO(2)$, and, in the left-most diagram, the curves are examples of left cosets, $\{R \cdot e^{\theta\hat{z}}\}$, which happen to correspond to points lying on the equator of the sphere shown on the far right. The middle diagram shows examples of right cosets which twist in a direction opposite that of the left cosets.

We can define a second equivalence class on SO(3) which would yield an equally suitable represention of S^2. In particular, suppose $R \approx S$ if and only if $z^T R = z^T S$. The equivalence classes of $\approx$ are related to those of $\sim$ by transposition, or equivalently, inversion. In particular, $R \sim S$ if and only if $R^T \approx S^T$. Fig. 2 shows the difference between left and right cosets in SO(3) in logarithmic coordinates. By similar arguments one finds that for a g defined on SO(3)/ $\approx$, g_l contains a single non-zero *row*, and is therefore equivalent to a row vector. If f and g are defined on SO(3)/ $\sim$ and SO(3)/ $\approx$, respectively, and if they are equal when their respective domains are identified with the sphere, then $f_l = g_l^\dagger$.

We can now define a convolution for the sphere, taking two functions $f, g : S^2 \to \mathbb{R}$ and yielding a function $h : \mathrm{SO}(3) \to \mathbb{R}$ equal to the dot product of f and $(g \circ R^{-1})$ for all R. We use the following definition of convolution:

$$h(R) = \int f(p) g(R^{-1} p)\, dp.$$

If $f : \mathrm{SO}(3)/\sim \to \mathbb{R}$ and $g : \mathrm{SO}(3)/\approx \to \mathbb{R}$ then this convolution can be rewritten as follows:

$$h(R) = \frac{1}{2\pi} \int f(S) g(S^{-1} R)\, dS$$

which is effectively equation (5). Division by 2π accounts for the size of the equivalence classes in SO(3). Then, by the convolution result,

$$h_l = f_l g_l.$$

Note that f_l is a column vector and g_l is a row vector, so that h_l is an outer product. If f_l and g_l are canonical spherical harmonic coefficients (where by default we assume that $S^2 \cong \mathrm{SO}(3)/\sim$, and so both are column vectors) then $h_l = f_l \cdot g_l^{\dagger}$.

§ 7 Implementation and experiments

We now have the necessary machinery to implement a Hough transform using a fast convolution. The inset on the next page describes an algorithm to compute a likelihood evaluated on a discrete subset of $\mathcal{E}$ from a set of points correspondences. Because of space constraints, some details are omitted and will be made available as a technical report.

To motivate further study, we include three synthetic experiments. In the first case we wish to determine how well the likelihood is defined in the presence of ambiguity. In particular, we will let our point correspondences be projections of points infinitely far away. Thus, in the rotational component, we hope for a well-defined mode. In the translational component, however, we expect a flat, mode-less distribution. The upper-left diagram of Fig. 3 shows the marginal likelihood function of the rotation component in log-space. The value that we expect is indicated by the (red) coordinate axes. The resulting mode is indeed centered near this ideal value. On the right side of Fig. 3, we plot the marginal likelihood of the translation which is relatively flat.

Next, we perform a similar experiment though without ambiguity, and evaluate the resulting likelihood. We generate twenty-five randomly distributed points such that the depth is on the order of the distance between the viewpoints. The resulting likelihoods in the rotational and translational components are shown on the left and right sides of the middle row of Fig. 3, respectively. Finally, in the third experiment we demonstrate estimation with two motions. The results are shown on the bottom row of Fig. 3. Though not as well-defined as before, two peaks do evidence themselves and are located near the true values indicated by the (red and green) coordinate axes in the marginal of rotation, and the arrows in the marginal of translation.

Algorithm 1. Global likelihood computation on $\mathcal{E}$ given a set of putative point correspondences (p_i, q_i).

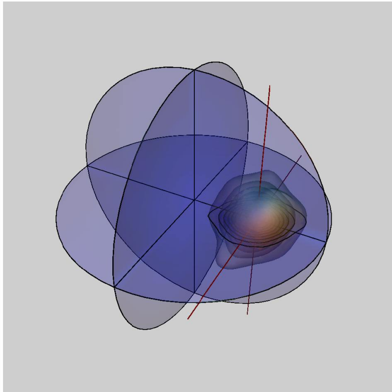
1. Prepare the canonical epipolar mask $f(\theta_1, \phi_1, \theta_2, \phi_2)$, and correspondence likelihood $g(\theta_1, \phi_1, \theta_2, \phi_2)$ as defined in equation (3). Discretely evaluate each at $\theta_1, \theta_2 = \pi k/L$ and $\phi_1, \phi_2 = \pi(k + 1/2)/2L$, where $k = 0, \ldots, 2L - 1$.
2. Compute the fast, discrete, L-bandlimited, separable SHT's of f and g using the separable basis $\{D^{l_1}_{m_1,0}(\theta_1, \phi_1, 0) D^{l_2}_{m_2,0}(\theta_2, \phi_2, 0)\}$, yielding entries $f^{l_1,l_2}_{m_1,m_2}$ and $g^{l_1,l_2}_{m_1,m_2}$ for $0 \leq l_1, l_2 < L$, $-l_1 \leq m_1 \leq l_1$ and $-l_2 \leq m_2 \leq l_2$.
3. The rotational harmonic coefficients of g are computed from $h^{l_1,l_1}_{m_1,n_1,m_2,-n_1} = f^{l_1,l_2}_{m_1,n_1} \left(g^{l_1,l_2}_{m_1,-n_1}\right)^\dagger$. Since $g(\theta_1, \phi_1, \psi_1, \theta_2, \phi_2, \psi_2)$ is independent of $\psi_1 + \psi_2$, all other terms are zero.
4. Compute discrete samples of $h(\theta_1, \phi_1, \delta, \theta_2, \phi_2)$ by applying the discrete inverse RHT in two passes to the entries $h^{l_1,l_2}_{m_1,n_1,m_2,-n_1}$, with modifications to account for zero off-diagonal terms.
5. Each $g(U, V)$ is analogous to a sum of normally distributed random variables; therefore, to emulate a likelihood, apply a saturation to g, i.e., $p(U, V) = \operatorname{erf}\left((g(U, V) - \mu)/(\sqrt{2}\sigma)\right)^k$ where $k > 0$, $\mu = Eg$, and $\sigma = \operatorname{Var}(g)^{1/2}$.
6. Optionally, to compute marginals of p, revert to the frequency representation by computing forward transforms. One can show that the unique decomposition $UE_0V^T = R\hat{t}$ yields marginals p_t and p_R, in translation and rotation respectively, whose respective spherical and rotational coefficients are: $(p_t)^l_m = p^{0,l}_{0,0,m,0}$ and $(p_R)^l_{m,n} = \sum_j (-i)^{j+2n} p^{l,l}_{m,j,-n,-j}$.

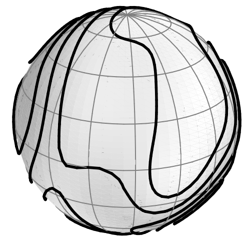
§ 8 Discussion

In this paper we have presented new techniques for motion estimation in omnidirectional cameras. We propose a transform analogous to the Hough transform for estimating one or more motions capable of handling ambiguity in motion estimation.

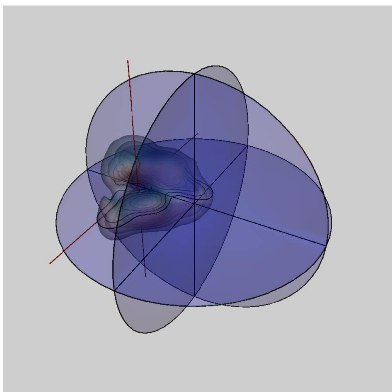
There are several avenues of research that hope to investigate in the future. First, this is a first pass at an implementation and we hope that improvements can be made. Second, note that computing the Hough transform in the manner proposed allows substituting the correspondence likelihood function g with any signal we can imagine. In particular, can we choose more vague correspondence functions, taking into account sensor noise and mismatch? Suppose for example, that we view objects which are constant color and for which it is difficult to extract features. If we are presented with a "blob," can we use such objects as a basis for correspondence? Can the method be used to reduce a dependence on features? One of the goals of this community has been to unite signal processing and geometry. We hope that this work can contribute to this program.

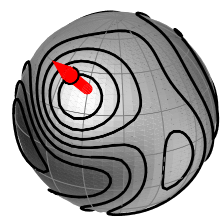
Acknowledgements The authors are grateful for the generous support of DARPA DAAD-19-02-1-0383 and NSF IIS-0122599 grants.



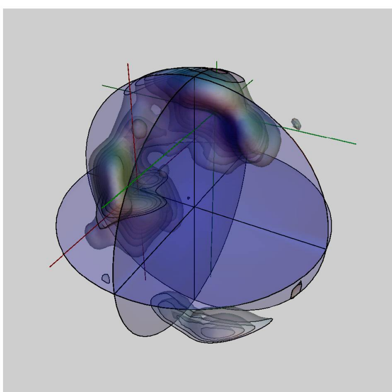


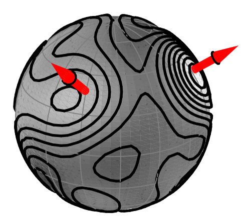
Experiment 1. We let $E = e^{-\pi\hat{x}/2}\hat{z}$ and simulate points infinitely far away. As expected, the resulting marginal in translation is flat, though the mode is well-defined in the rotational marginal.





Experiment 2. We suppose $E = e^{\pi\hat{x}/2}\widehat{(1,1,1)}$ and simulate points at a depth on the order of the distance between the viewpoints. Marginals in rotation and translation are peaked at the true values.





Experiment 3. Suppose points obey either $E_1 = e^{\pi\hat{x}/2}\widehat{(1,1,1)}$ or $E_2 = e^{\widehat{(0,1,2)}}\widehat{(-1,1,1)}$. Two peaks arise close to true values in both marginal likelihoods.

Fig. 3. Three experiments showing the applyication of a Hough transform to randomly generated point correspondences. In the first and second cases, twenty-five points were randomly generated. In the last case, twenty-five points were allowed per motion.

References

1. Makadia, A., Daniilidis, K.: Direct 3d-rotation estimation from spherical images via a generalized shift theorem. In: Madison, WI, June 18-20. (2003) 217–224, vol. II

2. Wexler, Y., Fitzgibbon, A., Zisserman, A.: Learning epipolar geometry from image sequences. In: Madison, WI, June 18-20. (2003) 209–216, vol. II
3. Tang, C., Medioni, G., Lee, M.: N-dimensional tensor voting and applications to epipolar geometry. IEEE Trans. Pattern Analysis and Machine Intelligence **23** (2001) 829–844
4. Hartley, R., Zisserman, A.: Multiple View Geometry. Cambridge Univ. Press (2000)
5. Hough, P.V.C.: Method and means for recognizing complex patterns. United States Patent No. 3,069,654 (1962)
6. Talman, J.: Special Functions: A Group Theoretic Approach. W. A. Benjamin, Inc., New York (1968)
7. Chirikjian, G., Kyatkin, A.: Engineering Applications of Noncommutative Harmonic Analysis: With Emphasis on Rotations and Motion Groups. CRC Press (2000)