

Rotation Estimation from Spherical Images

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

Abstract

Robotic navigation algorithms increasingly make use of the panoramic field of view provided by omnidirectional images to assist with localization tasks. Since the images taken by a particular class of omnidirectional sensors can be mapped to the sphere, the problem of attitude estimation arising from 3D rotations of the camera can be treated as a problem of estimating rotations between spherical images. Recently it has been shown that direct signal processing techniques are effective tools in handling rotations of the sphere, but are limited when the signal is altered by larger rotations of omnidirectional cameras. We present an effective solution to the attitude estimation problem under large rotations. Our approach utilizes a Shift Theorem for the Spherical Fourier Transform to produce a solution in the spectral domain.

1 Introduction

The longstanding problem of estimating the motion of a camera between a pair of images has been approached from numerous angles. For sufficiently small motions (differential motion), the computation of optical flow has been used to extract motion parameters, whereas for larger motions it has been the tracking of features or points in the images. These algorithms were first developed for conventional perspective images and later exported and refined to new image modalities, such as panoramic images.

Unlike conventional imaging, the panoramic view of an omnidirectional sensor ensures that scene structures will remain within the camera's field-of-view through a considerable amount of motion. This scene-persistence property which has made panoramic sensors appealing for mobile navigation tasks has also led to the development of localization algorithms based on the global structure of a scene. Appearance-based localization [4] is effective given a large database of reference views with PCA. Rotations around only the optical axis of the panoramic system have also been studied [8]. These techniques are both applied directly to panoramic images.

It is well known that images obtained from a single-viewpoint catadioptric system can be mapped to the sphere. Such a mapping, as described in [3], allows for the use of spherical signal analysis in the study of omnidirectional images obtained with a parabolic mirror [6]. It has been shown that 3D rotations can be estimated from the Fourier descriptors of the spherical images. These solutions are direct estimations of the rotation parameters using a Shift Theorem relating the spectra of rotated images [1, 5].

In general, rotations of a camera will capture new scene portions while relinquishing others. The magnitude of these signal inconsistencies depend on the camera's field-of-view (e.g. a camera that images the entire sphere does not exhibit such effects). Since omnidirectional images represent a hemispherical field of view, the effect of these inconsistencies is negligible when compared with a traditional narrow-fov camera. Only for significantly large rotations does this become an issue for the global estimation algorithms.

In this paper we present a novel method for rotation estimation from spherical images. We devise a two-step algorithm which first treats the problem as a global correlation based matching estimation. The second step of refining the estimate is based on a two-dimensional search. We compare this second step with the direct minimization described in [5]. Our incremental approach proves resilient to the signal inconsistencies inherent in large rotations of omnidirectional sensors.

2 The Spherical Fourier Transform

We begin our exploration of spherical harmonic analysis with some introductory definitions. We associate a point on the unit sphere with the unit vector $\omega(\theta, \phi) = (\cos \phi \sin \theta, \sin \phi \sin \theta, \cos \theta)^T$, where θ and ϕ represent the angles of colatitude and longitude, respectively. The rotation of a function defined on S^2 is given by an element $R(\alpha, \beta, \gamma)$ of the rotation group $SO(3)$, where α , β , and γ are the ZYZ Euler angles.

As the angular portion of the solution to Laplace's equation in spherical coordinates, the spherical harmonic functions Y_m^l form a complete orthonormal basis over the unit

sphere:

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

where $P_m^l(\cos(\theta))$ are associated Legendre polynomials. Thus, for any function $f(\omega) \in L^2(S^2)$, we have a Spherical Fourier Transform (SFT) given as

$$\begin{aligned} f(\omega) &= \sum_{l \in \mathbb{N}} \sum_{|m| \leq l} \hat{f}_m^l Y_m^l(\omega) \\ \hat{f}_m^l &= \int_{\omega \in S^2} f(\omega) \overline{Y_m^l(\omega)} d\omega \end{aligned} \quad (1)$$

As images, our spherical functions exist on a uniformly sampled angular grid, and so we must use the following theorem from [2] to compute the integral in (2):

Theorem 2.1 (Driscoll and Healy) *Let $f(\omega)$ be a band-limited function, such that $\hat{f}_m^l = 0, \forall l \geq B$, where B is the function bandwidth. Then for each $|m| \leq l < B$,*

$$\hat{f}_m^l = \frac{\sqrt{2\pi}}{2B} \sum_{j=0}^{2B-1} \sum_{k=0}^{2B-1} a_j^{(B)} f(\theta_j, \phi_k) \overline{Y_m^l(\theta_j, \phi_k)}$$

where the function samples $f(\theta_j, \phi_k)$ are chosen from the equiangular grid: $\theta_j = j\pi/2B, \phi_k = k\pi/B$, and the weights $a_j^{(B)}$ are the solutions to the system:

$$\sum_{j=0}^{2B-1} a_j^{(B)} \sqrt{\frac{2l+1}{2}} P_0^l(\cos(\theta_j)) = \delta_{l,0}, \quad \forall l < 2B$$

Two important properties of the spherical harmonic functions Y_m^l are:

$$Y_m^l(R^{-1}\eta) = \sum_{|k| \leq l} Y_k^l(\eta) U_{km}^l(R) \quad (3)$$

$$Y_m^l(\beta, \alpha) = \sqrt{\frac{2l+1}{4\pi}} U_{m0}^l(R_{\alpha, \beta, \gamma}) \quad (4)$$

The $(2l+1) \times (2l+1)$ matrices U^l are the irreducible unitary matrix representations of the transformation group $SO(3)$, whose elements are given by

$$U_{mk}^l(R) = e^{-im\alpha} P_{mk}^l(\cos \beta) e^{-ik\gamma}, \quad (5)$$

where P_{mk}^l are the generalized Legendre polynomials.

From (3) we obtain a Shift Theorem relating coefficients of rotated functions [5]:

$$h(\omega) = f(R^{-1}\omega) \Leftrightarrow \hat{h}_m^l = \sum_{|k| \leq l} \hat{f}_k^l U_{mk}^l(R) \quad (6)$$

3 Rotation Estimation

In this section we will develop the full estimation algorithm. We begin with a summary of the method presented in [5]. Readers are referred to the full text for complete details.

3.1 Direct Nonlinear Estimation (DNE)

From our selection of ZYZ Euler angles to parameterize $SO(3)$ we have $R(\alpha, \beta, \gamma) = R_2(\beta + \pi, \frac{\pi}{2}, \gamma + \frac{\pi}{2}) R_1(\alpha + \frac{\pi}{2}, \frac{\pi}{2}, 0)$, and from the unitarity of the representations U^l we have $U^l(R_2 R_1) = U^l(R_2) U^l(R_1)$. If we expand (6) using this identity, we obtain

$$\begin{aligned} \hat{h}_m^l &= e^{-im(\alpha + \frac{\pi}{2})} \sum_{|p| \leq l} e^{-ip(\gamma + \frac{\pi}{2})} \hat{f}_p^l \\ &\quad \cdot \sum_{|k| \leq l} P_{pk}^l(0) P_{km}^l(0) e^{-ik(\beta + \pi)} \end{aligned} \quad (7)$$

From (7) we can see that \hat{h}_0^l is independent of the first Euler angle α . For each pair $(\hat{f}_0^l, \hat{h}_0^l)$ we have one equation in two unknowns, β and γ . A non-linear system in these unknowns can be minimized efficiently using any of a number of quasi-Newton methods. Once we have estimates for β and γ , we proceed to estimate α . Again, we refer to (7) to see that if β and γ are known (or without loss of generality, zero), we have the following relationship between coefficients:

$$\hat{h}_m^l = e^{-im\alpha} \hat{f}_m^l \quad (8)$$

With only one unknown, our estimate for α is obtained by minimizing the following for any $l \geq 1$:

$$\sum_{1 \leq p \leq l} \sum_{1 \leq m \leq p} ((e^{-im\alpha} \hat{f}_m^p - \hat{h}_m^p)^2 = 0$$

3.2 North Pole Motion (NPM)

The Shift Theorem (6) shows us that the U^l matrix representations of the rotation group $SO(3)$ are the spectral analogue to 3D rotations. As vectors in \mathbb{R}^3 are rotated by orthogonal matrices, the $(2l+1)$ -length complex vectors \hat{f}^l , comprised of all coefficients of degree l , are transformed by the unitary matrices U^l . This allows us to rewrite (6) as a linear transformation of the spectrum by a semi-infinite block diagonal matrix:

$$\begin{pmatrix} \hat{h}^1 \\ \hat{h}^2 \\ \vdots \\ \hat{h}^l \end{pmatrix} = \begin{pmatrix} U^1 & & & \emptyset \\ & U^2 & & \\ & & \ddots & \\ \emptyset & & & U^l \end{pmatrix} \begin{pmatrix} \hat{f}^1 \\ \hat{f}^2 \\ \vdots \\ \hat{f}^l \end{pmatrix}. \quad (9)$$

If we substitute (4) into the central column and row of every matrix U^i of the system (9), we obtain for every degree l :

$$\sqrt{\frac{2l+1}{4\pi}} \hat{f}_0^l = \sum_{m=-l}^l Y_m^l(\beta, \alpha) \hat{h}_m^l \quad (10)$$

$$\sqrt{\frac{2l+1}{4\pi}} \hat{h}_0^l = \sum_{k=-l}^l Y_k^l(\beta, \gamma) (-1)^k \overline{\hat{f}_k^l} \quad (11)$$

The previous relations give a harmonic description of the North Pole transformations for both the original and rotated images. Let $\omega = R \cdot \eta$ and $\omega' = R^T \cdot \eta$ be the unit vectors into which the North Pole η is transformed by the rotation R and its inverse R^T . Then, due to the North Pole's invariance under rotations about the Z-axis, we have $\omega = (\beta, \alpha)$ and $\omega' = (-\beta, -\gamma)$. Expanding $h(\omega)$ and $f(\omega')$ in spherical harmonics (1), we obtain directly the relations (10,11).

According to (10,11), for every degree $l \geq 0$, we can estimate the Euler angles as a solution of the following minimization problem:

$$(\beta, \alpha)_l = \min_{\theta, \phi} \left\{ \left| ISFT\{H^l\} - \sqrt{\frac{2l+1}{4\pi}} \hat{f}_0^l \right|^2 \right\} \quad (12)$$

$$(\beta, \gamma)_l = \min_{\theta, \phi} \left\{ \left| ISFT\{F^l\} - \sqrt{\frac{2l+1}{4\pi}} \hat{h}_0^l \right|^2 \right\} \quad (13)$$

where H^l and F^l are the $(l+1)^2$ -length vectors:

$$H^l = [\emptyset \quad \dots \quad \emptyset \quad \hat{h}_{-l}^l \quad \dots \quad \hat{h}_0^l \quad \dots \quad \hat{h}_l^l]^T$$

$$F^l = [\emptyset \quad \dots \quad \emptyset \quad (-1)^{-l} \overline{\hat{f}}_{-l}^l \quad \dots \quad \hat{f}_0^l \quad \dots \quad (-1)^l \overline{\hat{f}}_l^l]^T$$

In general, estimating a 3D rotation requires minimizing a cost function of 3 variables (Euler angles). With (12) and (13), we have reduced the number to two (Fig.1).

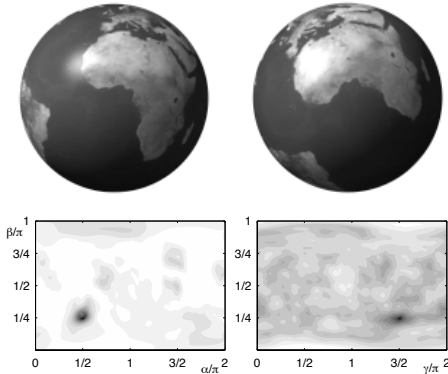


Figure 1. Two spherical images related by the 3D rotation $R(\alpha = \pi/2, \beta = \pi/4, \gamma = 3\pi/2)$. On the bottom the two maps (β, α) and (β, γ) used for the estimation, obtained according to (12) and (13).

3.3 Normalized Cross Correlation (NCC)

The first assumption made by the two previous algorithms is that the spherical images being analyzed are in fact identical up to a rotation. In practical implementations

however, there are a number of reasons this assumption may be violated. For example, a catadioptric imaging system consisting of a telecentric lens and a parabolic mirror produces a hemispherical image. A general rotation of the mirror about its focal point induces a signal change in addition to a spherical rotation. See Figure 2 for an example of a spherical image from a catadioptric system. For very large rotations, this signal change degrades the relationship (6) to the point where the parameter estimations described in sections (3.1) and (3.2) will prove ineffective. In such cases, a correlation-based matching analysis will provide a more robust estimation.

If we define the normalized cross-correlation of two spherical images $f(\omega)$ and $h(\omega)$ as

$$K(R) = \int_{\omega \in S^2} F(\omega) \overline{\Lambda_R H(\omega)} d\omega, \quad R \in SO(3), \quad (14)$$

then the R for which $K(R)$ takes its maximum value is the rotation that relates $f(\omega)$ with $h(\omega)$. Here $F(\omega)$ and $H(\omega)$ are normalized images ($F(\omega) = f(\omega) - \bar{f}$). If we substitute (1) and (6) into (14), we have:

$$K(R) = \int_{\omega \in S^2} \sum_l \sum_m \sum_p f_m^l \overline{\hat{h}}_p^l \overline{U}_{pm}^l(R) \quad (15)$$

Remembering that any rotation can be sundered into successive rotations, we can write

$$K(R) = \sum_l \sum_{mpk} \hat{f}_m^l \overline{\hat{h}}_p^l P_{pk}^l(0) P_{km}^l(0) e^{-i(p\gamma' + k\beta' + m\alpha')} \quad (16)$$

where $\gamma' = \gamma + \frac{\pi}{2}$, $\beta' = \beta + \pi$, $\alpha' = \alpha + \frac{\pi}{2}$. The 3D FFT of $K(R)$ is then given by [7]:

$$\hat{K}_{abc} = \sum_l \hat{f}_a^l \overline{\hat{h}}_b^l P_{bc}^l(0) P_{ca}^l(0).$$

Thus, \hat{K} is a matrix with size $(2L+1)^3$, whose elements can be computed directly from the SFT coefficients of f and h . The 3D-IFFT of \hat{K} restores our correlation function $K(R)$. The resolution of $K(R)$ depends on the number of coefficients calculated: K is accurate up to $(360/(2L+1))^\circ$. Since the resulting exhaustive search for the maximum value in our rotation space is 3 dimensional, we use a coarse correlation grid which provides us with an initial estimate for the direct minimization techniques described earlier.

Our final estimation algorithm consists of two steps. First, a coarse correlation grid is computed yielding an initial estimate of the rotation. This provides us with an estimate of the regions in both images which constitute a signal change. This unmatched signal content is then removed from each image by setting the greyscale value of the unmatched pixels to 0. In the second stage, one of the two minimization techniques described in sections (3.1) and (3.2) is applied to obtain a final estimate.

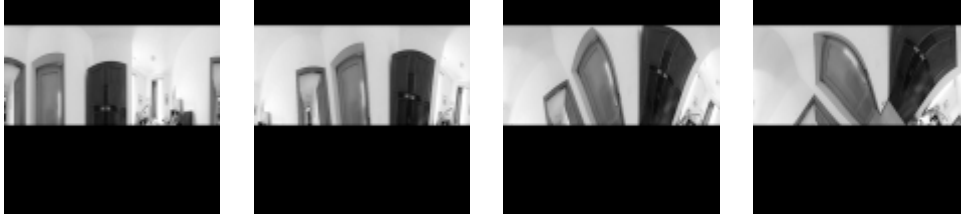


Figure 2. Four catadioptric images (I_0, I_1, I_2, I_3), mapped on the (θ, ϕ) -plane, obtained by artificial progressive rotation.

4 Experiments

We label the full algorithms of section (3.3) as **NCC+DNE** or **NCC+NPM**, depending on which estimation is used in the second stage. We compare four different techniques (**DNE**, **NPM**, **NCC+DNE**, and **NCC+NPM**). Figure 2 shows the first set of spherical test images, and Table 1 shows the results of each algorithm as the rotation around the Y-axis (which contributes the signal change) increases. We also compare using real rotated images the two full algorithms in the presence of small camera translations. The complete rigid motion occurs in the plane perpendicular to the optical axis, which allows for an accurate measurement of the ground truth for the rotational component. Results are shown in the last row of Table 1.

		α	β	γ
I_1 ($45^\circ, 12^\circ, 45^\circ$)	DNE	29.27°	3.67°	42.8°
	NPM	40.32°	5.12°	32.26°
	NCC+DNE	49.0°	10.2°	43.41°
	NCC+NPM	47.82°	9.52°	46.78°
I_2 ($45^\circ, 36^\circ, 45^\circ$)	DNE	—	∞	∞
	NPM	43.88°	13.24°	57.38°
	NCC+DNE	41.19°	30.5°	50.77°
	NCC+NPM	46.28°	33.17°	43.20°
I_3 ($45^\circ, 60^\circ, 45^\circ$)	DNE	—	∞	∞
	NPM	38.28°	24.46°	48.36°
	NCC+DNE	43.52°	51.76°	47.34°
	NCC+NPM	44.23°	55.91°	47.82°
R+T $R : (-5^\circ, 0^\circ, 0^\circ)$	NCC+DNE	-0.31°	-1.5°	-3.6°
	NCC+NPM	-2.45°	1.98°	-1.86°

Table 1. Estimation results. DNE: non linear minimization, NPM: North Pole motion, NCC: Correlation maximization (used in the full algorithm). A value of ∞ means the nonlinear minimization did not converge.

5 Conclusion

This work addressed the problem of estimating 3D spherical rotations by analyzing the transformations induced in the spectral domain. We proposed a two step method: first a coarse correlation matching was computed to locate regions of signal change, and then a final estimation was performed by techniques which analyze the spectral transformation associated with a rotation. Estimation performances on artificial as well as real rotations were presented, and the results clearly indicate that the full algorithm is necessary to accurately estimate large rotations. Although our approach is shown to be robust to the existence of small translations, it remains to extend the algorithm to include general 3D rigid motions.

References

- [1] G. Burel and H. Henoco. Determination of the orientation of 3d objects using spherical harmonics. *Graphical Models and Image Processing*, 57:400–408, 1995.
- [2] J. Driscoll and D. Healy. Computing fourier transforms and convolutions on the 2-sphere. *Advances in Applied Mathematics*, 15:202–250, 1994.
- [3] C. Geyer and K. Daniilidis. Catadioptric projective geometry. *International Journal of Computer Vision*, 43:223–243, 2001.
- [4] M. Jogan and A. Leonardis. Robust localization using eigenspace of spinning-images. In *IEEE Workshop on Omnidirectional Vision, Hilton Head, SC, June 12*, pages 37–46, 2000.
- [5] A. Makadia and K. Daniilidis. Direct 3d rotation estimation from spherical images via a generalized shift theorem. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages 217–224, 2003.
- [6] S. Nayar. Catadioptric omnidirectional camera. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages 482–488, Puerto Rico, June 17-19, 1997.
- [7] L. Sorgi and K. Daniilidis. Template gradient matching in spherical images. In *Electronic Imaging 04*, San Jose, January 18-22, 2004.
- [8] 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.