Group Theoretic Approach to Recovery of Omnidirectional Images

Amal Aafif[1] Robert Boyer[2]

[1,2]Department of Mathematics, Drexel University, Philadelphia, PA

Abstract

Omnidirectional images have become important in the field of computer vision as they offer larger and more complex views of the world. Image processing of such images requires an approach that accounts for symmetries of the image and the possible transformations and deformations of objects within the image. Group representation theory already provides the framework for a general Fourier transform and its inverse over the 2D Euclidean motion group SE(2). We found that the operator-valued Fourier transform on the motion group reduces for a two-dimensional image to a standard Fourier transform in polar form. This observation allows bypassing the complicated inversion formula for non-commutative groups. In this paper, a numerical implementation of the SE(2) Fourier transform is used to recover omnidirectional images from the polar frequency domain. Further, we show that a simple splitting of the transform before inversion significantly improves image reconstruction.

Keywords: omnidirectional vision, image processing, non-commutative harmonic analysis

1. Background

Applications in computer vision have expanded as larger images can now be stored, processed and analyzed quickly and efficiently. Autonomous robot navigation, automated surveillance and medical imaging can all benefit from wider fields of view with minimal distortion. Omnidirectional images are captured using rotating cameras, multiple cameras, or mirror-camera systems to create a full 360º view. In mirror-camera systems, a camera is pointed at a conic mirror, allowing the camera to view all directions around it at once. The resulting picture is a 2D image of a 360º view of the camera surroundings. Depending on the type of mirror used, such images can be distorted with varying degrees and un-warping techniques have been developed to create a typical panoramic view. These techniques, however, cannot eliminate distortion completely and do not preserve relative locations of objects in the room. For instance, a 4 sided room and its contents could appear as objects along a single "long" wall. For this reason, we prefer to work directly with the original omnidirectional image as opposed to a panoramic image.

In pattern recognition and feature identification, we expect objects in an image to be transformed under rotation, translation and scaling. Objects in omnidirectional images can also be further deformed by the mirror. To describe such transformations, techniques involving generalized Fourier transforms have been explored. Motion invariants based on non-commutative harmonic analysis on the Euclidean motion group were presented in [3] and used in [2] for template matching of black objects in a white background under rotation, translation and scaling. In [1,7] an extensive overview of the use of non-commutative harmonic analysis in engineering applications is presented. A description of a non-commutative approach to inverting the Radon transform with Wiener filtering is described in [6]. We explore a simpler implementation of the inverse Fourier transform over the Euclidean motion group to determine how accurately it can recover the original image.

2. Euclidean Motion Group

Euclidean motion group, SE(2), is a connected component of the group of rigid body motions, or all one-to-one distance preserving maps from the plane to itself. SE(2) is a solvable, non-commutative Lie group given by the semi-direct product of $\mathbf{R}^2$ with SO(2), the special orthogonal group of proper rotations on the plane. The transformations $g \in SE(2)$ are denoted by $g = (z, k)$ where z is in polar coordinates (ρ, θ) and k is rotation in the form $(\cos \omega, \sin \omega)$.

We identify the 2D image plane as a function on the homogeneous space SE(2)/SO(2) with SE(2) acting transitively upon it. Any transformations or deformations on the image are group actions. The next step is to determine the general Fourier transform for this group as described by representation theory. Since the motion groups are semi-direct products of the rotation group with the Euclidean space, they are non-commutative. As a consequence, the Fourier

transform is no longer scalar valued as in the Euclidean case. The infinite-dimensional irreducible unitary representations π_λ, with $\lambda > 0$, of SE(2) are realized on $L^2(S^1)$ and given by

$$[\pi_\lambda(g)f](\alpha) = \exp(i\lambda\Re(\alpha^{-1}z))f(k^{-1}\alpha),$$

where $f \in L^2(S^1), z \in \mathbf{C}$, and $k, \alpha \in S^1$. The operator-valued Fourier transform $F(\varphi,\lambda)$ for $\varphi \in L^1(\mathrm{SE}(2))$ is a bounded operator on $L^2(S^1)$ given by

$$[F(\varphi,\lambda)f](\alpha) =$$
$$\int_{\mathbf{C}}\int_{S^1} \varphi(z,k)\exp(i\lambda\Re(\alpha^{-1}z))f(k^{-1}\alpha)dk\,d\alpha.$$

We will identify an integrable function φ on the plane with the corresponding function on SE(2) so that its values are independent of k; that is, $\varphi(z,k) = \varphi(z)$ for all $k \in S^1$.

The first important consequence for all such functions is that the null space of $F(\varphi,\lambda)$ consists exactly of those functions whose integral over S^1 is 0. In particular, $F(\varphi,\lambda)$ is a rank-one operator completely determined by its value on the vector $e_0 \in L^1(S^1)$ which is the function that is identically 1. The second consequence is that the function $F(\varphi,\lambda)e_0(k_0)$ is nothing more than the usual 2D Euclidean Fourier transform of φ evaluated at the point on the plane $(\lambda\cos\theta_0, \lambda\sin\theta_0)$ where $k_0=(\cos\theta_0, \sin\theta_0)$.

The general inversion formula for the operator-valued Fourier transform is

$$\varphi(g_0) = \int_0^\infty Tr[\pi_\lambda(g_0) * F(\varphi,\lambda)]\lambda\, d\lambda$$

where Tr is the trace operator and $*$ is the conjugate transpose. A direct implementation of this formula is very difficult because it requires the evaluation of the trace of the infinite-dimensional operators $\pi_\lambda(g_0)*F(\varphi,\lambda)$. Fortunately, when φ comes from a function on the plane, the inversion formula can be avoided altogether because the $F(\varphi,\lambda)$ reduces to the Euclidean Fourier transform. In [6], a similar reduction is used by treating $F(\varphi,\lambda)e_0$ as a rank-one operator but the authors used a Fourier series expansion rather than the explicit connection between the SE(2) transform and the usual Fourier transform in polar form.

3. Method and Results

The images considered in this work are gray-scale images taken by a camera-mirror system using a spherical mirror. All the information is contained within a certain radius from the center of the image. The outer edges are usually white or black and contain no information. Spherical mirrors do not preserve lines and areas so objects closest to the center of the image are distorted minimally. Figure 1(a) contains a simulated checkerboard wall as viewed by a camera coupled with a spherical mirror while figure 1(b) is a demonstration of the camera-mirror system's ability to capture a full 360º view.

The full SE(2) Fourier transform applied to an image on the plane reduces to the usual 2D Fourier transform in polar coordinates. This eliminates the non-commutativity of the problem and theoretically we can directly invert the polar Fourier transform to recover the original image.

We found it difficult to reconstruct the original image from the Fourier transform in polar coordinates. The center of a polar image is over-determined by many pixels while the corners are reduced to very few pixels. Consequently, the polar-to-rectangular conversion relies heavily on interpolation towards the edges. While the image can be reconstructed through direct Fourier inversion, the polar-to-rectangular conversion introduces a significant amount of noise or "grayness" in the image (figure 2(a)). This problem is somewhat remedied by increasing the number of points in the polar image to produce more information at the edges, even up 1000 radii × 1000 angles for a 300 × 300 pixel image in rectangular coordinates. Figure 2(b) shows the improvement in reconstructed image clarity obtained by increasing the density of the polar image. We also note that reconstructed image is rotated 180º, an artifact perhaps due to rotational aliasing. A more rigorous theoretical analysis may help identify the source of this artifact.

Regardless of the image used, a polar density above 1000 × 1000 points did not produce a recognizable image (Figure 3). In fact, the results in Figure 3 resemble those found with using very low polar densities (100 × 100, 200 × 200). This leads us to believe that the aliasing might be periodic. Due to the increase in computation time required for interpolation, we were unable reconstruct images from very large polar FFTs (above 2000 × 2000 pixels).

Since the polar-to-rectangular conversion clearly plays a role in image reconstruction and the interpolation into rectangular coordinates generally produces accurate points near the center of the image, it was reasonable to split the image into parts to obtain several "regions" of accurate points throughout the image. Without regard to the image content, the Fourier transform image was split into 4 equal parts and then 16 equal parts before performing the polar-to-rectangular conversion. By using only 500 × 500 polar points, we obtained a greatly enhanced

reconstruction of the omnidirectional image (Figure 4), which is comparable to and even clearer than the result in figure 3(b). A certain amount of "grayness" still remains in the reconstruction – due to noise in Fourier transform from the interpolation, as shown in the error map in Figure 5(a). This bisection method decreased the average error by 29%. The reconstruction with 500 × 500 polar points with bisection is comparable to performing the same experiment with direct Fourier inversion with 1000 × 1000 polar points with a decrease in error by 11%.

Fig. 1: (a) A simulated checkerboard wall viewed by camera looking at spherical mirror
(b) Image captured using spherical mirror. The camera is the black object at the center. Image courtesy of http://cmp.felk.cvut.cz/demos/Omnivis/index.html

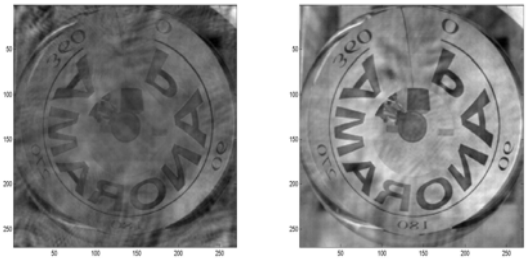

Fig. 2: (a) Reconstruction of panorama image using 500 × 500 polar points and (b) 1000 × 1000 polar points. Reconstruction is rotated by 180 degrees.

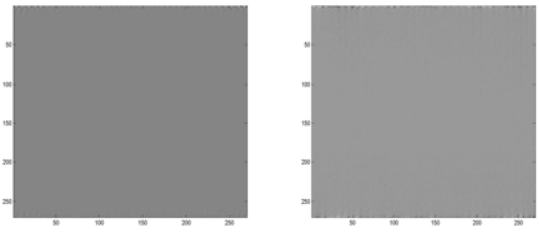

Fig. 3: (a) Reconstruction of panorama image using 1100 × 1100 polar points and (b) 1500 × 1500 polar points.

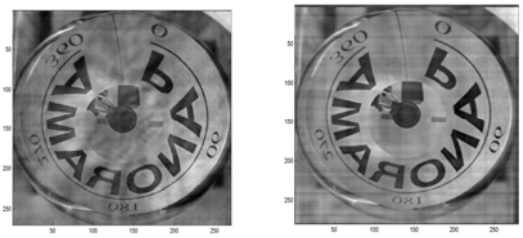

Fig. 4: Reconstruction of the panorama image using 500 × 500 polar points and splitting the image into (a) 4 parts and (b) 16 parts.

A more adaptive approach to splitting the polar FFT image could be developed in the future to determine if some regions in the image warrant more subdivisions than others. The reconstruction of the simulated checkerboard yielded similar results in image enhancement though the error map in figure 5(b) contains fewer regions of low error. Such images could further benefit from an additional non-commutative filtering step. A standard Wiener filter did not improve the results.

4. Conclusions and Future Work

Numerical implementation of non-commutative harmonic analysis techniques in computer vision applications is still limited. On the other hand, the Euclidean motion group, both in 2 and 3 dimensions is well suited for analyzing omnidirectional images. We reduce the problem of Fourier inversion over a non-commutative group to one over the usual Euclidean plane. While polar-to-rectangular conversion and direct Fourier inversion produce relatively low image quality results, a simple bisection method quickly reduces some of the error without significantly changing the computational complexity. Image content could be taken into consideration to develop an adaptive partitioning of the Fourier transform before rectangular conversion and Fourier inversion. A non-commutative Weiner filtering technique, described in [6], could also be explored to improve performance. Finally, these non-commutative frequency-based methods could be further exploited in template matching.

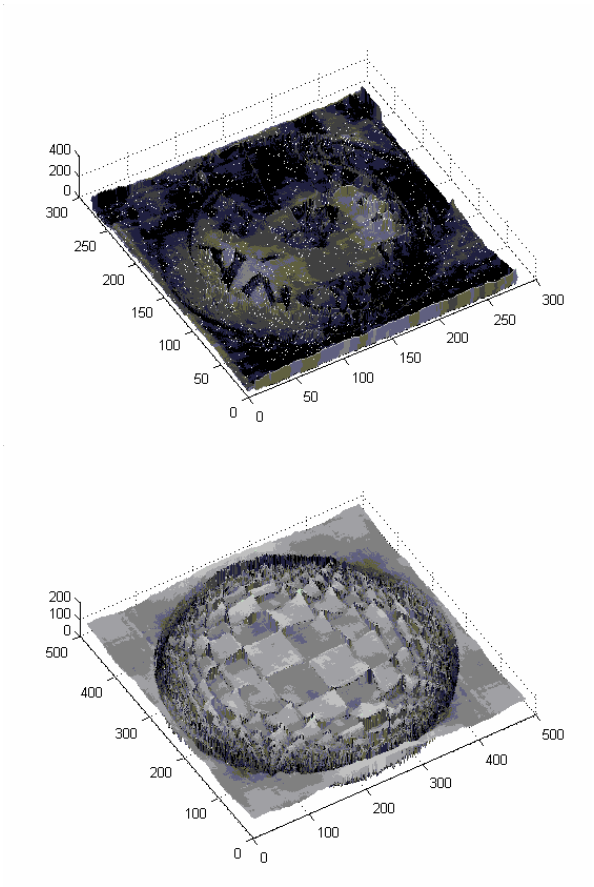

Figure 5: (a) Error map for reconstructed panorama image in figure 4(b). (b) Error map for reconstructed checkerboard image using 500×500 polar points and 16 parts. Error ranging from low to high is represented by gray levels ranging from black to white.

5. References

[1] G.S. Chirikjian and A.B. Kyatkin, *Engineering Applications in Non-commutative Harmonic Analysis: with Emphasis on Rotation and Motion Groups,* CRC Press Online, 2000.

[2] H. Fonga, "Pattern recognition in gray-level images using Fourier analysis," *Pattern Recognition Letters 17* (14) , pp.1477-1489, 1996.

[3] J-P. Gauthier, G. Bornard, and M. Silbermann, "Motions and Pattern Analysis: Harmonic Analysis on Motion Groups and their Homogeneous Spaces," *IEEE Transactions on Systems, Man, and Cybernetics Vol. 21, No. 1,* pp. 159-172. 1991.

[4] B. Hall, *Lie Groups, Lie Algebras and Representations: An Elementary Introduction*, Springer-Verlag, New York, NY, 2003.

[5] J. Turski, "Geometric Fourier Analysis of the Conformal Camera for Active Vision*," SIAM Review 46* (2), pp. 230-255, 2004.

[6] C.E. Yarman and B.Yazici, "Radon Transform Inversion via Wiener Filtering over the Euclidean Motion Group," *Proceedings of the International Conference on Image Processing, Volume 2*, pp. II- 811-14 vol.3, Sept. 2003.

[7] B. Yazici, "Stochastic deconvolution over groups," *IEEE Transactions on Information Theory 20*(3), pp. 494-510, 2004.

[8] M. Zribi, "Description of three-dimensional gray-level objects by the harmonic analysis approach," *Pattern Recognition Letters 23*(1-3): pp. 235-243, 2002.