Omnidirectional image processing

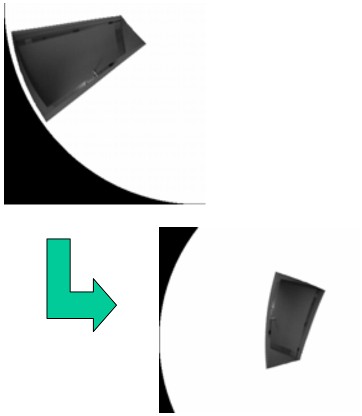
Kostas Daniilidis
GRASP Laboratory
University of Pennsylvania

With contributions of Ameesh Makadia, Thomas Buelow, Daniel Rudoy, and Lorenzo Sorgi

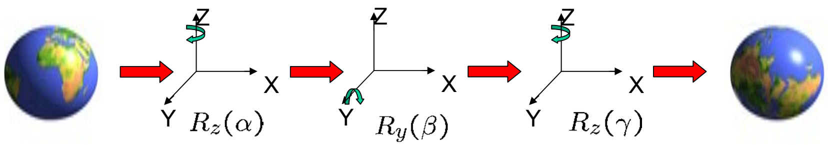
Image processing in perspective images

- Images obtained through perspective projection undergo local mappings:
  - Translations
  - Similitude
  - Affine
  - Projective (Collineations).
- This assumption is implicit in template matching and filtering.

Template deformation in an omni-image is not covered by any of these mappings



Original image $\xrightarrow{\text{calibration}}$ Spherical image



The rotation of a function $f(\eta)$ by an element $g \in SO(3)$ is defined with the operator Λ_g as $\Lambda_g f(\eta) = f(g^{-1}\eta)$

The integration of a function $f(\eta) \in L^2(S^2)$ is defined as

$$\int_{\eta \in S^2} f(\eta) d\eta = \int_0^{2\pi} \int_0^{\pi} f(\theta, \phi) sin(\theta) d\theta d\phi$$

How does convolution look like on the sphere?

- What is the "shift" in the convolution?
- It is a 3D-rotation acting as an operator:

$$(f * h)(\eta) = \int_{g \in SO(3)} f(gn) h(g^{-1}\eta)\, dg, \quad \eta \in S^2$$

North pole

$$\eta := (\cos(\varphi)\sin(\vartheta),\ \sin(\varphi)\sin(\vartheta),\ \cos(\vartheta)),$$

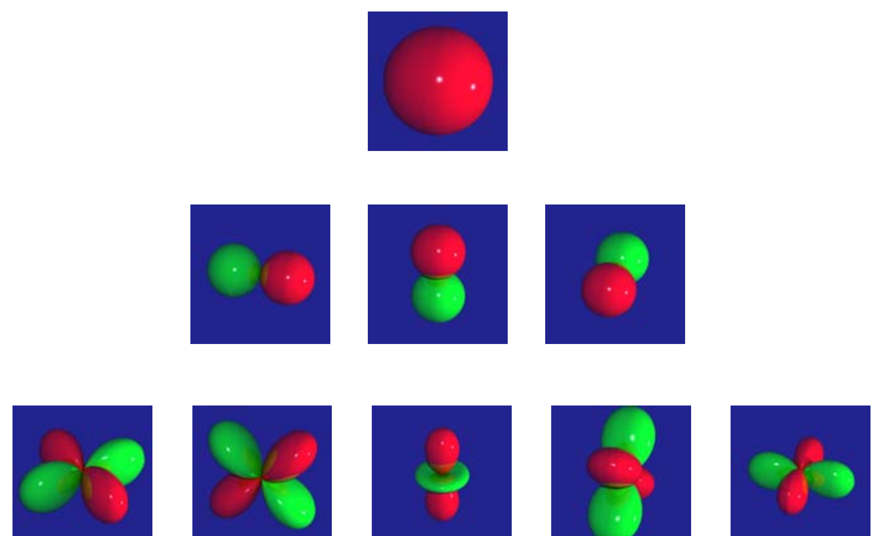
What about a Fourier transform on the sphere?

Look for a decomposition of functions on the sphere into subspaces invariant under SO(3): Eigenfunctions of the Laplace equation, the spherical harmonics

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

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

Spherical Harmonics



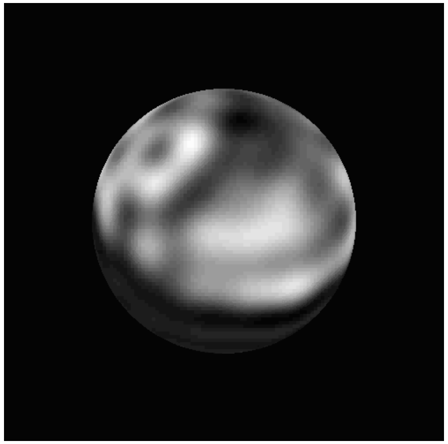
Spherical Harmonic Transform

$$f(\theta,\phi) = \sum_{l\in\mathbb{N}} \sum_{|m|\le l} \hat{f}_{lm} Y^l_m(\theta,\phi)$$

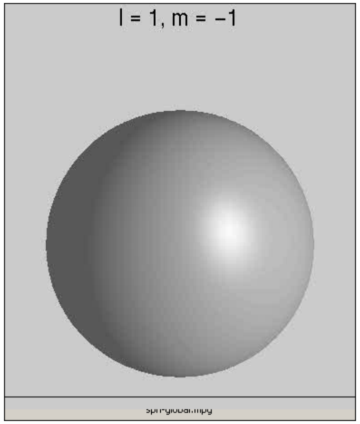
$$\hat{f}_{lm} = \int_{\eta\in S^2} f(\eta)\overline{Y^l_m}(\eta)d\eta$$

The $(2l+1)$ $\hat{f}_{lm}$ are the spherical harmonic coefficients of degree l.

Reconstruction with Spherical Harmonics



Spherical range images



Some facts on groups and homogeneous spaces

Lie group is a group with its elements on a smooth manifold and the group operation and inversion being smooth maps.

Examples: The real line and the circle with the addition operation are Lie groups. Also real square invertible matrices GL(n), the rotation groups SO(2) and SO(3), and the Lorentz groups SO(3,1).

A group G is acting on a space X when there is a map G x X -> 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.

For example, the isometry group SE(2) acts on the plane R^2. The rotation group SO(3) can act on the sphere 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**.

Let us think of images as homogeneous spaces G/H with the group G acting on them (G is a Lie group and for any $x, y \in G/H$ exists a g such that $y = gx$):

- **Planar pictures are defined on $SE(2)/SO(2)$ with rigid motions acting on them.**
- **Spherical images are defined on $SO(3)/SO(2)$ with 3D-rotations acting on them.**

A representation of a group G is a homomorphism $T : G \rightarrow GL(V)$.

A representation is unitary if for all $g \in G$

$$(T(g)v, T(g)w) = (v, w) \quad \forall v, w \in V.$$

A representation T is reducible if there is a proper subspace W of V which is invariant under T. Otherwise, T is irreducible.

Let T be representation of group G in vector space V. Then T is irreducible if and only if the only $A : V \rightarrow V'$ satisfying $T(g)A = AT(g)$, $\forall g \in G$ are $A = \lambda I$.

If the acting group is unimodular and locally compact, the group has an irreducible representation $U(g,p)$.

The Fourier transform of a function on the homogeneous space G/H exists:

$$F(p) = \int_{\eta \in G/H} f(\eta) U(g^{-1}, proj(p)) d\eta.$$

In the case of the sphere $S^2 = SO(3)/SO(1)$

$$\hat{f}^l_m = \int_{\eta \in G/H} f(\eta) U^l_{m0}(\eta) d\eta$$

where U^l_{mn} the irreducible unitary representation of $SO(3)$.

SO(3) irreducible unitary representation

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

Framework for image processing in various domains

- 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.

Problem 1: Rotation estimation

Sphere SO(3)/SO(2)

Problem: Compute the rotation of a spherical image directly from its spherical harmonic coefficients (no correspondence).

Current methods: Iterative closest point gradient decent minimization or hierarchical flow algorithms.

Shift Theorem

$$\Lambda_g Y^l_m(\eta) = \sum_{|n|\le l} U^l_{mn}(g) Y^l_n(\eta)$$

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

$$^{l} = \begin{bmatrix} U^l_{-l,-l} & U^l_{-l,-l+1} & \cdots & U^l_{-l,l} \\ U^l_{-l+1,-l} & & & \vdots \\ \vdots & & \ddots & \vdots \\ U^l_{l,-l} & U^l_{l,-l+1} & \cdots & U^l_{l,l} \end{bmatrix} \quad Y^l = \begin{bmatrix} Y^l_{-l} \\ \vdots \\ Y^l_l \end{bmatrix} \hat{f}^l = \begin{bmatrix} \hat{f}_{l,-l} \\ \vdots \\ \hat{f}_{l,l} \end{bmatrix}$$

$$\Lambda_g Y^l = U^l(g) Y^l \quad \Lambda_g \hat{f}_l = U^l(g)^T \hat{f}_l$$

Shift Theorem

$$\hat{f}^g_l \equiv \Lambda_g \hat{f}_l = U^l(g)^T \hat{f}_l$$

$$\hat{f}^g_{lm} = \sum_{|p|\le l} U^l_{pm} \hat{f}_{lp}$$

Image Invariants

$$\Lambda_g \hat{f}_l = U^l(g)^T \hat{f}_l$$

Since U^l is a unitary matrix, its rows form a unitary basis. The rows (and columns) have length 1 and their Hermitian inner product is zero. Thus, a transformation by a unitary matrix does not affect a vector's length.

$$K_l(f(\eta)) = \sum_{|m|\le l} \overline{\hat{f}_{lm}} \hat{f}_{lm}$$

Convolution Theorem (scanned from Driscoll-Healy-94)

THEOREM 1. *For functions f, h in $L^2(S^2)$, the transform of the convolution is a pointwise product of the transforms*

$$(f * h)\hat{}(l,m) = 2\pi\sqrt{\frac{4\pi}{2l+1}}\,\hat{f}(l,m)\hat{h}(l,0)$$

Approach

Problem: Determine if two spherical images A and B are related by a rotation $g(\alpha, \beta, \gamma)$, and if so, what are α, β, and γ.

We can use our invariant function $K_l(f(\eta))$ to determine if two images are identical up to a rotation.

We extract Euler angles of rotation from the Shift Theorem.

$$U^l \text{ gives } U^l(g_1 g_2) = U^l(g_1)U^l(g_2)$$

$$g(\alpha, \beta, \gamma) = g_1(\alpha + \tfrac{\pi}{2}, \tfrac{\pi}{2}, 0)g_2(\beta + \pi, \tfrac{\pi}{2}, \gamma + \tfrac{\pi}{2})$$

$$\Lambda_{g_2 g_1} \hat{f}_l = (U^l(g_1))^T (U^l(g_2))^T \hat{f}_l$$

$$\hat{f}^g_{lm} = e^{-im(\gamma+\frac{\pi}{2})} \sum_{|p|\leq l} e^{-ip(\alpha+\frac{\pi}{2})} \hat{f}_{lp} \sum_{|k|\leq l} P^l_{pk}(0) P^l_{km}(0) e^{-ik(\beta+\pi)}$$

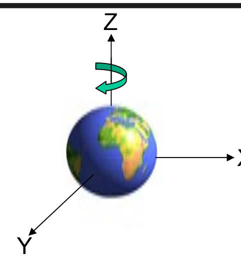
Parameter Estimation

To begin, examine two simple cases:

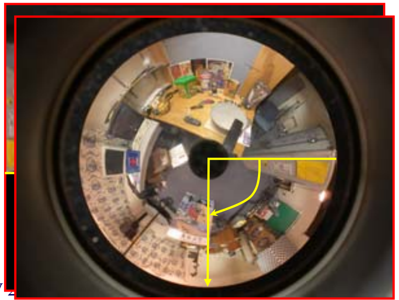
1. Estimate rotation around Z-axis
   - beta and either alpha or gamma are known
   - beta is zero: solution is not unique. Assume only alpha needs estimating.
2. Estimate rotation around Y-axis
   - alpha and gamma are known

Estimating Rotation Around Z-axis

Without loss of generality, we assume that only alpha needs to be estimated.

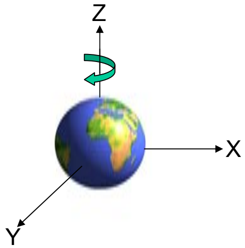


A rotation alpha is equivalent to…



- A translation along Phi in the Theta-Phi plane.
- A rotation around the origin in the original catadioptric image plane.

Estimating Rotation Around Z-axis

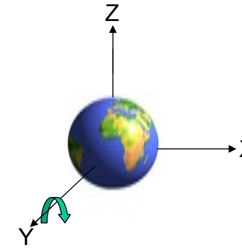


Re-examining the shift property we see that the angle alpha appears only once.

$$\hat{f}^{g}_{lm} = e^{-im(\gamma+\frac{\pi}{2})} \sum_{|p|\le l} e^{-ip(\alpha+\frac{\pi}{2})} \hat{f}_{lp} \sum_{|k|\le l} P^{l}_{pk}(0) P^{l}_{km}(0) e^{-ik(\beta+\pi)}$$

We can generate an over-constrained system using multiple coefficients with m>0

Estimating Rotation Around Y-axis



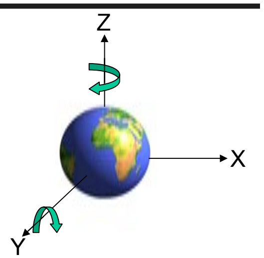
Without loss of generality, we assume that only beta is nonzero (apply known alpha and gamma rotations to images prior to estimation).

Rewriting the shift property we get

$$\Lambda_g \hat{f}_{lm} = \sum_{|p|\le l} e^{-ip\beta} C^{l}_{mp}$$

$$C^{l}_{mp} = e^{-im(\frac{\pi}{2})}\left(\sum_{|l|\le l} e^{-ik(\frac{\pi}{2})} \hat{f}_{lk} P^{l}_{kp}(0) P^{l}_{pm}(0) e^{-ik\pi}\right)$$

Estimating All Parameters

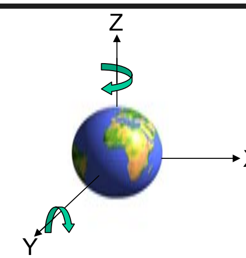


Estimation is done in two steps

1. Generate estimates for beta and gamma.

2. Use beta and gamma as input to solving for alpha, which we already know how to calculate.

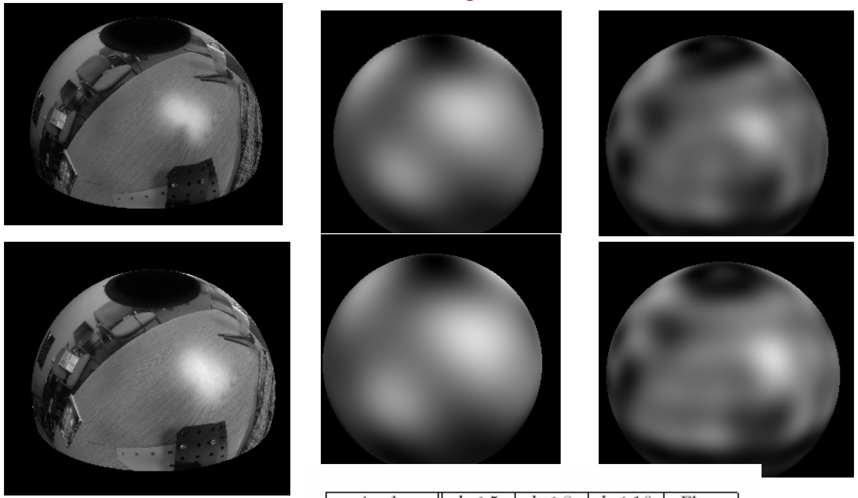
Estimating beta and gamma



The first rotation of alpha is not reflected in the coefficients f_l0

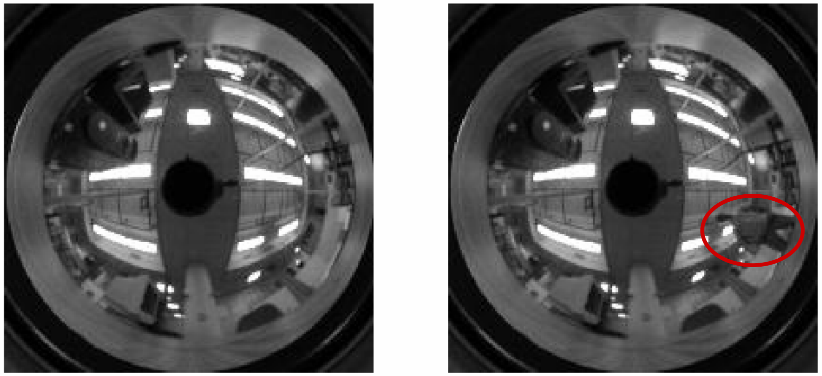
Using only the equations for the coefficients f_l0, we get an over-constrained system for the two unknowns beta and gamma

Estimation from very few coefficients!



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

Resistant to clutter



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

6%, 10%, and 13% clutter

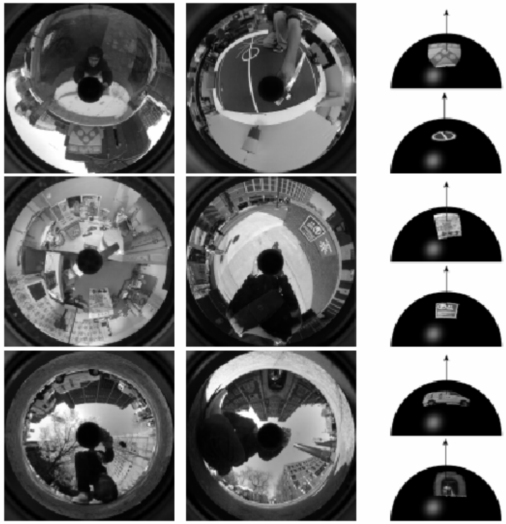
Problem 2: Template matching

Given $f(\eta), h(\eta) \in L^2(S^2)$, the correlation between $f(\eta)$ and $h(\eta)$ is defined as

$$g(\alpha, \beta, \gamma) = \int_{S^2} f(\eta)\Lambda(R)h(\eta)\ d\eta$$

Correlation can be obtained from the spherical harmonics $\hat{f}^l_m$ and $\hat{h}^l_m$ via the 3-D Inverse Discrete Fourier Transform as

$$g(\alpha, \beta, \gamma) = IDFT\{\sum_l \hat{f}^l_m \overline{\hat{h}^l_k} U^l_{m,h}(\pi/2) U^l_{h,k}(\pi/2)\}.$$



Harmonic analysis

- **Global shape descriptors (moment, Fourier-descriptors) of the 60's-80's have been abandoned because of occlusions.**
- **Omnidirectional images give you large closed areas persistent in images (many appearance based techniques)**
- **Classical Fourier can not be applied anyway due to the new deformations.**
- **Let us re-think Fourier-transforms!**

References - Books

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

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

W. Miller, *Topics in Harmonic Analysis*, manuscript in the web math.umn.edu, 2002

W. Rossman, *Introduction to Lie Groups,* Oxford University Press, 2002.

References – Papers

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

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.

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

A. Makadia, K. Daniilidis, Direct 3D rotation estimation via a generalized shift theorem, *Computer Vision Pattern Recognition,* 2003