An Analysis of Sphere Tessellations for Pose Estimation of 3-D Objects Using Spherically Correlated Images*

Randy C. Hoover and Anthony A. Maciejewski
Dept. of Electrical and Computer Eng.
Colorado State University
Fort Collins, CO 80523-1373, USA
Email: {hooverr, aam}@colostate.edu

Rodney G. Roberts
Dept. of Electrical and Computer Eng.
Florida A & M - Florida State University
Tallahassee, FL 32310-6046, USA
Email: rroberts@eng.fsu.edu

Abstract

Eigendecomposition is a common technique used for pose detection of three-dimensional (3-D) objects from two-dimensional (2-D) images. It has been shown in previous work that the eigendecomposition can be estimated using spherical sampling in conjunction with the Spherical Harmonic Transform. The issue then becomes deciding on the best tessellation of the sphere to define the sampling pattern. In this paper we evaluate three popular tessellations and compare and contrast their computational performance, as well as their estimation accuracy for the eigendecomposition of this spherical data set.

1. Introduction

Object recognition and pose detection of three-dimensional (3-D) objects from two-dimensional (2-D) images are important issues in computer vision. Subspace methods, also referred to as eigenspace methods, principal component analysis, or the Karhunen-Loeve transformation [1, 2], represent one computationally efficient approach for dealing with this class of problems and have been used for a variety of applications in this area. In [3, 4] it is shown that a set of highly correlated images can be approximately represented by a small set of eigenimages. Once the principal eigenimages of an image data set have been determined, using these eigenimages is very computationally efficient for the on-line classification of 3-D objects. Unfortunately, the off-line calculation for determining the appropriate subspace dimension, as well as the principal eigenimages themselves is computationally expensive.

This drawback has been addressed using several different approaches [5–11]. In [10], Chang *et al.* showed that if the image data set was correlated in one dimension, then the FFT may be used to approximate the desired subspace dimension, as well as the principal eigenimages. In [11] the authors extended this concept to correlation in two-dimensions by using the *Spherical Harmonic Transform* (SHT) in place of the FFT to capture the frequency information of this data set. The issue is then in deciding the "best" tessellation of the sphere such that a representative set of samples is achieved. In this work we consider three different tessellations (see Fig. 1), namely, the Gauss-Legendre (GL) grid [12], an equi-angular (EA) grid [13], and the HEALPix grid [14]. Once the image data set for a particular object is sampled on these grids, the eigendecomposition is estimated based on the SHT. The eigendecompositions for the three tessellations are then compared for computational efficiency and accuracy of the estimation.

Figure 1. This figure shows the three different tessellations analyzed in this paper. The Gauss-Legendre (left), equi-angular (center), and HEALPix (right).

2. Preliminaries

2.1. The SVD of Correlated Images

In this work, a gray-scale image is described by an $h \times v$ array of square pixels with intensity values normalized between 0 and 1 [10]. Thus, an image will be represented by a matrix $\mathcal{X} \in [0,1]^{h \times v}$. Because sets of related images are considered in this paper, the *image vector* $\boldsymbol{f}$ of length $m = hv$ is obtained by "row-scanning" an image into a column vector, i.e., $\boldsymbol{f} = \text{vec}(\mathcal{X}^T)$. The *image data matrix*

*This work was supported in part by the Missile Defense Agency under contract no. HQ0006-05-C-0035. Approved for Public Release 07-MDA-3067 (18 JAN 08).

of a set of images $\mathcal{X}_1, \ldots, \mathcal{X}_n$ is an $m \times n$ matrix, denoted X, and defined as $X = [\boldsymbol{f}_1, \cdots, \boldsymbol{f}_n]$, where typically $m > n$ with fixed n. Because we will be sampling images on the unit sphere, it should be noted that $\boldsymbol{f} = \boldsymbol{f}(\boldsymbol{\xi}_p)$ where $\boldsymbol{\xi}_p$, $p \in \{0, \ldots, n-1\}$ is the unit vector pointing at the angle of co-latitude $\theta_p \in [0, \pi]$ measured down from the z axis, and the angle of longitude $\phi_p \in [0, 2\pi)$, which is the parameterization of the sphere in spherical coordinates.

The singular value decomposition (SVD) of X is given by $X = U\Sigma V^T$, where $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are orthogonal, and $\Sigma = [\Sigma_d \ \mathbf{0}]^T \in \mathbb{R}^{m\times n}$ where $\Sigma_d = \text{diag}(\sigma_1, \cdots, \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$ and $\mathbf{0}$ is an n by $m-n$ zero matrix. The columns of U, denoted $\mathbf{u}_i, i = 1, \ldots, m$, are referred to as the left singular vectors or eigenimages of X, while the columns of V, denoted $\mathbf{v}_i, i = 1, \ldots, n$ are referred to as the right singular vectors of X. In practice, the left singular vectors $\mathbf{u}_i$ are not known or computed exactly, and instead estimates $\tilde{\mathbf{u}}_1, \ldots, \tilde{\mathbf{u}}_k$ which form a k-dimensional basis are used [10]. The measures we will use for quantifying the accuracy of these estimates for each tessellation are described in the following subsections.

2.2. Residue Between Subspaces

The possibility that the data matrix $B \in \Re^{m\times k}$ can be rotated into the data matrix $A \in \Re^{m\times k}$ is explored [15] by solving the problem

$$\Delta = \min_{Q} ||A - BQ||_F \tag{1}$$

where $||\cdot||_F$ represents the Frobenius norm, $Q \in \Re^{k\times k}$ is an orthogonal matrix, and Δ is the residue.

The Q that minimizes $||A - BQ||_F$ in (1) can be calculated as follows:

- Form the matrix $C = B^T A$.
- Compute the SVD of C, i.e., $C = U_c \Sigma_c V_c^T$.
- Find the orthogonal matrix $Q = U_c V_c^T$.

The residue Δ after evaluating (1) using the above Q will satisfy

$$\Delta^2 = \text{tr}(A^T A) + \text{tr}(B^T B) - 2\sum_{i=1}^{k} \sigma_{ci} \tag{2}$$

where σ_{ci} is the i^{th} singular value of C. The smaller the residue Δ, the closer A and B are to representing the same subspace. To determine the rotation between two sets of eigenimages, let

$$\begin{aligned} A &= U_k, \\ B &= \tilde{U}_k, \end{aligned} \tag{3}$$

where U_k and $\tilde{U}_k$ are the matrices containing the first k true and approximated eigenimages as their columns, respectively.

2.3. Energy Recovery Ratio

True and approximated eigenimages of X can also be compared in terms of their capability of recovering the amount of the total energy in X. The "energy recovery ratio" denoted ρ, is defined as

$$\rho(X, \tilde{\mathbf{u}}_1, \ldots, \tilde{\mathbf{u}}_k) = \frac{\sum_{i=1}^{k} ||\tilde{\mathbf{u}}_i^T X||^2}{||X||_F^2} \tag{4}$$

where $||\cdot||_F$ denotes the Frobenius norm. Note that if the $\tilde{\mathbf{u}}_i$ are orthonormal, $\rho \leq 1$.

2.4. Subspace Criterion

True eigenimages give an optimum energy recovery ratio in (4), therefore, it is possible that more approximated eigenimages are required to achieve the same energy recovery ratio. Hence, another measure used in this study is the degree to which approximate eigenimages span the subspace of the first k^* true eigenimages, which will be referred to as the subspace criterion, γ, given by

$$\gamma = \sqrt{\frac{1}{k^*} \sum_{i=1}^{k} \sum_{j=1}^{k^*} (\tilde{\mathbf{u}}_i^T \mathbf{u}_j)^2}. \tag{5}$$

3. SHT Algorithm

In this section, we briefly give an overview of the SHT algorithm developed in [11]. For this analysis, we are using CAD generated ray-traced images, examples of which are shown in Fig. 2, (the CAD models were provided by [16]). Each of the objects are sampled according to one of the above tessellations, to produce the image data matrix X. We then compute the matrix F whose i^{th} row is the SHT of the i^{th} row of X, denoted from this point forward as SHT(X). Because of the spherical correlation, the principal eigenimages $\tilde{\mathbf{u}}_1, \ldots, \tilde{\mathbf{u}}_k$ of SVD(F) serve as excellent estimates to those of X. The following algorithm is then used to estimate the desired subspace based on a user specified energy recovery ratio μ and "bandwidth" B.

SHT Eigendecomposition Algorithm

1. Form the matrix F which is the SHT(X).
2. Form the matrix H whose columns are the ordered columns of F in descending order according to their norm.
3. Set $q = (B)^2[1 - (1/2)^{N+1}]$ with $N = 0$ initially.

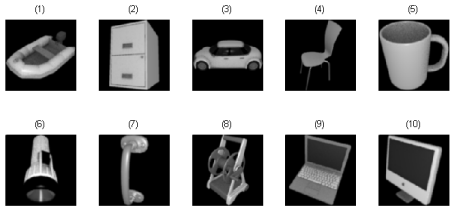

Figure 2. Ray-traced CAD models courtesy of Kator Legaz [16]. Each object is sampled as discussed above at a resolution of 128 $\times$ 128.

4. Construct the matrix H_q which is the matrix consisting of the first q columns of H.
5. Compute SVD(H_q).
6. If $\rho(X, \tilde{\mathbf{u}}_1, \ldots, \tilde{\mathbf{u}}_q) < \mu$. Let $N = N + 1$ and repeat Steps 3 through 6.
7. Return $\tilde{\mathbf{u}}_1, \ldots, \tilde{\mathbf{u}}_k$ such that $\rho(X, \tilde{\mathbf{u}}_1, \ldots, \tilde{\mathbf{u}}_k) \geq \mu$. Note that $k \leq q$.

The computational savings in the SHT algorithm is a result of the fact that the transform is lossy. The quantification of this loss depends on the tessellation used and is outlined in Table 1. Using the HEALPix tessellation the number of harmonic images can be up to 3/4 the number of samples; however, for a consistent comparison, we truncate the HEALPix algorithm to produce the same number of harmonic images as the other two tessellations. While this loss may be viewed as a drawback in terms of image reconstruction, this is actually an advantage in terms of computational complexity. Furthermore, because subspace methods typically work well with a much smaller subspace dimension, and as a result less recoverable energy, this method has shown significant computational savings with little loss in performance.

Table 1. Sample information for each of the three tessellations.

	GL	EA	HEALPix
# Samples	512	1024	768
# Harmonics	256	256	256

4. Analysis of the Results

The above algorithm was tested on all of the objects shown in Fig. 2. All of the images were both scale and intensity normalized to create the image data matrix X. Finally, the matrix F was computed condensing the image data matrix to 256 harmonic images. The direct SVD(X) is also computed for a ground truth comparison.

Table 2 shows the computation time required to calculate the required subspace dimension k, as well as estimate the left singular vectors $\tilde{\mathbf{u}}_1, \ldots, \tilde{\mathbf{u}}_k$ for each of the three tessellation to exceed the user specified energy recovery $\mu = 0.9$. As can be seen from the table, all three tessellations are comparable in terms of computational complexity. This is derived from the fact that the resulting number of columns are the same in each tessellation, the difference in times are a result of added structure embedded by the SHT. The performance of each tessellation is also tested for the ability to estimate the true left singular vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k$. Figure 3 shows the difference between the subspace dimension as computed by the SVD(X) and all three tessellations for an energy recovery ratio of $\mu = 0.9$. As can be seen from the figure, the HEALPix sampling grid gives a very good estimate of the subspace dimension as compared to the true SVD. This estimate is extremely important in that the online computation time is dependent on this being as small as possible.

Table 2. Computation time required for each of the three tessellations to meet the user specified energy recovery ratio $\mu = 0.9$.

Computation Time [sec.]			
	GL	EA	HEALPix
mean	5.05	6.47	5.67
min	4.56	5.52	4.98
max	6.13	7.94	6.79

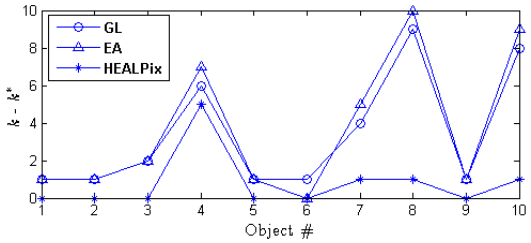

Figure 3. Difference between the subspace dimension computed by the SVD(X) and that computed by the SHT algorithm for each of the three tessellations to meet the user specified energy recovery ratio $\mu = 0.90$.

As mentioned in Section 2.3, the true SVD(X) gives the optimal subspace; therefore, the quality of the estimated subspace is also important. The quality measures outlined in Section 2 were computed for all three tessellations and averaged over all objects in Fig. 2. The results are depicted in Fig. 4. As can be seen from the figure, all three tessellations provide a very good estimate of the true eigendecomposition, with the HEALPix tessellation giving the best performance.

5. Discussion

In [11] the authors propose a method for 3-D pose estimation using spherical sampling and the Spherical Har-

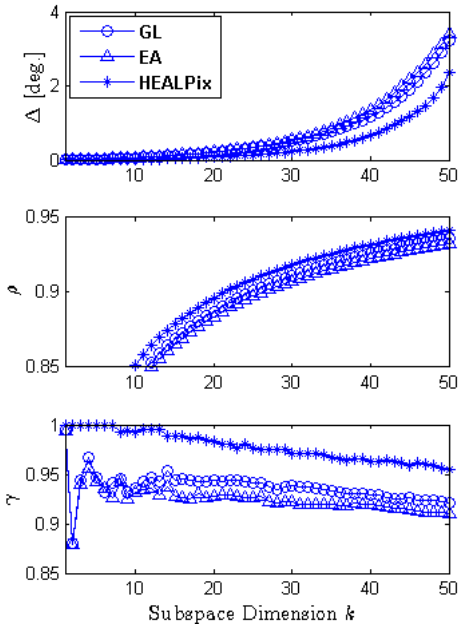

Figure 4. Quality measures outlined in Section 2 computed for all three tessellations and averaged over all objects in Fig. 2.

monic Transform. The issue then becomes deciding on the "best" tessellation of the sphere to use such that a representative discretization is achieved. Three popular tessellations were compared in this paper, namely, the Gauss-Legendre grid, the equi-angular grid, and the HEALPix grid. The tessellations were then evaluated based on their ability to estimate the required subspace based on a user specified energy recovery ratio. The computation times for each tessellation were computed, and the quality of each estimation was evaluated based on the quality measures outlined in Section 2. These results were then compared to the results obtained by computing the eigendecomposition via the direct SVD. While the HEALPix tessellation outperforms the other two tessellations in terms of quality of estimation, there is no exact quadrature for the SHT transform as there is in the other two. This inaccuracy in the integration over S^2 can be remedied using an iterative quadrature correction; however this increases the time required to compute the transform. It has also been shown in [17] that for arbitrary functions defined on S^2, the performance and computation times in computing both the forward and inverse transform using the HEALPix tessellation is nearly equivalent to using the EA tessellation. This is a function of the application however and the results presented here suggest that for the pose estimation problem, the HEALPix tessellation outperforms the other two. This result is based on the fact that for a similar number of samples, the HEALPix tessellation gives better angular resolution and does not bias the SVD by oversampling the polar regions.

References

[1] K. Fukunaga, *Introduction to Statistical Pattern Recognition*. London, U.K.: Academic, 1990.

[2] J. J. Gerbrands, "On the relationships between SVD, KLT and PCA," *Pattern Recognition*, vol. 14, no. 1-6, pp. 375–381, 1981.

[3] H. Murase and S. K. Nayar, "Visual learning and recognition of 3-D objects from appearance," *Int. J. Comp. Vis.*, vol. 14, no. 1, pp. 5–24, Jan. 1995.

[4] H. Murakami and V. Kumar, "Efficient calculation of primary images from a set of images," *IEEE Trans. PAMI*, vol. 4, no. 5, pp. 511–515, Sept. 1982.

[5] R. Haimi-Cohen and A. Cohen, "Gradient-type algorithms for partial singular value decomposition," *IEEE Trans. PAMI*, vol. PAMI-9, no. 1, pp. 137–142, Jan. 1987.

[6] X. Yang, T. K. Sarkar, and E. Arvas, "A survey of conjugate gradient algorithms for solution of extreme eigen-problems for a symmetric matrix," *IEEE Trans. ASSP*, vol. 37, no. 10, pp. 1550–1556, Oct. 1989.

[7] C. R. Vogel and J. G. Wade, "Iterative SVD-based methods for ill-posed problems," *SIAM J. Sci. Comput.*, vol. 15, no. 3, pp. 736–754, May 1994.

[8] S. Chandrasekaran, B. Manjunath, Y. Wang, J. Winkeler, and H. Zhang, "An eigenspace update algorithm for image analysis," *CVGIP: Graphic Models and Image Proc.*, vol. 59, no. 5, pp. 321–332, Sept. 1997.

[9] K. Saitwal, A. A. Maciejewski, R. G. Roberts, and B. Draper, "Using the low-resolution properties of correlated images to improve the computational efficiency of eigenspace decomposition," *IEEE Trans. Image Proc.*, vol. 15, no. 8, pp. 2376–2387, Aug. 2006.

[10] C. Y. Chang, A. A. Maciejewski, and V. Balakrishnan, "Fast eigenspace decomposition of correlated images," *IEEE Trans. Image Proc.*, vol. 9, no. 11, pp. 1937–1949, Nov. 2000.

[11] R. C. Hoover, A. A. Maciejewski, and R. G. Roberts, "Pose detection of 3-D objects using S^2-correlated images and discrete spherical harmonic transforms," accepted to appear in *IEEE Int. Conf. Robot. Automat.*, Pasedena, CA, May 19-23, 2008.

[12] P. N. Swarztrauber, "Computing the points and weights for Gauss-Legendre quadrature," *SIAM J. on Sci. Comp.*, vol. 24, no. 3, pp. 945–954, 2002.

[13] D.M. Healy Jr, D.N. Rockmore, P.J. Kostelec, and S. Moore, "FFTs for the 2-sphere-improvements and variations," *J. of Fourier Anal. and App.*, vol. 9, no. 4, pp. 341–385, Jul. 2003.

[14] Górski, K. M. Hivon, E. Banday, A. J. Wandelt, B. D. Hansen, F. K. Reinecke, and M. Bartelmann, "HEALPix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere," *The Astrophysical Journal*, vol. 622, pp. 759–771, Apr. 2005.

[15] G. H. Golub and C. F. V. Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins, 1996.

[16] K. Legaz, "3-D model database for use with blender (web)," www.katorlegaz.com/3d_models/index.php, 2007.

[17] Y. Wiaux, L. Jacques, P. Vielva, and P. Vandergheynst, "Fast directional correlation on the sphere with steerable filters," *The Astrophysical Journal*, vol. 652, p. 820, 2006.