4th IEEE Conference on Automation Science and Engineering
Key Bridge Marriott, Washington DC, USA
August 23-26, 2008

Pose Detection of 3-D Objects Using Images Sampled on $SO(3)$, Spherical Harmonics, and Wigner-D Matrices

Randy C. Hoover*, Anthony A. Maciejewski*, and Rodney G. Roberts+

*Dept. of Electrical and Computer Eng.
Colorado State University
Fort Collins, CO 80523-1373, USA
Email: {hooverr, aam}@colostate.edu

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

***Abstract*— Determining the pose of three-dimensional objects from two-dimensional images has become an important issue in industrial automation applications. Eigendecomposition represents one computationally efficient method for dealing with this class of problems. One major drawback of using the eigendecomposition technique is the expensive off-line computation required to calculate the optimal subspace. This off-line computational expense may preclude the use of eigendecomposition for industrial applications where time is of the essence. In this work, we address this issue by proposing a computationally efficient algorithm for estimating the eigendecomposition to a user specified accuracy. In particular, we sample the rotation group $SO(3)$ in a manner that allows us to take advantage of the correlation in $SO(3)$ and transform the data from the spatial domain to the spectral domain. We then present an algorithm to estimate the eigendecomposition in this domain, thus relieving the computational burden. Experimental results are presented to compare the proposed algorithm to the true eigendecomposition, as well as assess the computational savings.**

I. Introduction

Over the last several decades, pose detection of three-dimensional (3-D) objects from two-dimensional (2-D) images has become an important issue in vision-based automation applications. Specific examples include automated assembly and inspection, robotic welding, human-machine interaction and PCB inspection. (Refer to Table 2 in [1] for numerous applications.)

Subspace methods, also referred to as eigenspace methods, principal component analysis, or the Karhunen-Loeve transformation [2], [3], represent one computationally efficient approach for dealing with object detection, pose estimation, as well as other vision-based problems. In [4], [5] 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 off-line computational burden is a major deterrent for industrial applications due to the length of time required for accurate classification of any particular object. This drawback has been addressed using several different approaches based on either iterative power methods, conjugate gradient algorithms, or eigenspace updating [6]–[8]. A fundamentally different approach was proposed by Chang *et al.* [9] where the authors show that the FFT may be used to approximate the desired subspace dimension, as well as the principal eigenimages, if the image data set is correlated in one dimension. A modified version of Chang's algorithm has also been applied to images characterized by three parameters in which the images were captured from a spherical patch above the object [10]. In [11], the authors present a method to extend Chang's algorithm when the image data set is correlated in two dimensions by using the *Spherical Harmonic Transform* (SHT) in place of the FFT. Finally, in [12] it is shown that the computational efficiency of Chang's algorithm can be further increased by using the low resolution properties of the image data set.

In this paper, we extend the work of [9] and [11] using theory from spectral analysis. In particular, we sample the rotation group $SO(3)$ in a manner that allows us to take advantage of the correlation in $SO(3)$ and transform the data from the spatial domain to the spectral domain. We then propose an algorithm to estimate the eigendecomposition in this domain thus reducing the computational burden. In the next section, we explain the fundamentals needed to apply an eigendecomposition to a related image data set, much of which is discussed in [9]. In Section III, we give a brief introduction to Spherical Harmonics and Wigner-D matrices, and the $SO(3)$ FFT. In Section IV we propose a computationally efficient algorithm for estimating the principal eigenimages of the image data set. Experimental results are given in Section V, with conclusions and future work outlined in Section VI.

II. Preliminaries

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. 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

This work was supported in part by the Missile Defense Agency under contract no. HQ0006-05-C-0035. Approved for public release 08-MDA-3382 (May 20, 2008)

$m = hv$ is obtained by "row-scanning" an image into a column vector, i.e., $\boldsymbol{f} = \text{vec}(\mathcal{X}^T) \in \mathbb{R}^{hv\times 1}$. The *image data matrix* 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 $SO(3)$, it should be noted that $n = ab$, where a is the number of samples defined on the sphere, and b is the number of planar rotations captured at each sample. The image vector is then $\boldsymbol{f} = \boldsymbol{f}(\boldsymbol{\xi}_{p_r})$ where $\boldsymbol{\xi}_{p_r}$, $p \in \{0, \ldots, a-1\}$, is the unit vector pointing at the angle of co-latitude $\beta_p \in [0, \pi]$ measured down from the upper pole, and the angle of longitude $\alpha_p \in [0, 2\pi)$, which is the parameterization of the sphere in spherical coordinates. In $\boldsymbol{\xi}_{p_r}$, the value $r \in \{0, \ldots, b-1\}$ is the r^{th} planar rotation $\gamma_r \in [0, 2\pi)$ at sample p. The average image vector is then subtracted from the image data matrix X to generate the zero mean image data matrix $\hat{X}$, which has the interpretation of an "unbiased" image data matrix.

The *thin* singular value decomposition (SVD) of $\hat{X}$ is given by $\hat{X} = \hat{U}\hat{\Sigma}\hat{V}^T$ where $\hat{U} \in \mathbb{R}^{m\times n}$ and $\hat{V} \in \mathbb{R}^{n\times n}$ are orthogonal, and $\hat{\Sigma} \in \mathbb{R}^{n\times n}$ where $\hat{\Sigma} = \text{diag}(\hat{\sigma}_1, \cdots, \hat{\sigma}_n)$ with $\hat{\sigma}_1 \geq \hat{\sigma}_2 \geq \cdots \geq \hat{\sigma}_n \geq 0$. The columns of $\hat{U}$, denoted $\hat{\mathbf{u}}_i, i = 1, \ldots, n$, are referred to as the left singular vectors or eigenimages of $\hat{X}$, while the columns of $\hat{V}$, denoted $\hat{\mathbf{v}}_i, i = 1, \ldots, n$ are referred to as the right singular vectors of $\hat{X}$. In practice, the left singular vectors $\hat{\mathbf{u}}_i$ are not known or computed exactly, and instead estimates $\tilde{\hat{\mathbf{u}}}_1, \ldots, \tilde{\hat{\mathbf{u}}}_k$, denoted $\tilde{\hat{U}}_k$ that form a k-dimensional basis, are used. The accuracy of a practical implementation of subspace methods then depends on three factors: the properties of $\hat{X}$, the dimension k, and the quality of the estimates $\tilde{\hat{\mathbf{u}}}_i$. One measure we use to quantify the quality of the estimated eigenimages is the "energy recovery ratio," ρ, and defined as

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

where $||\cdot||_F$ denotes the Frobenius norm, and $\tilde{\hat{\mathbf{u}}}_i$ is the i^{th} column of $\tilde{\hat{U}}_k$. The energy recovery ratio is a measure used to quantify how much energy a particular set of basis vectors can recover from the original data matrix $\hat{X}$. A second measure we use for quantifying the accuracy of the estimated eigenimages is the change in ρ, i.e.,

$$\Delta\rho(\hat{X}, \tilde{\hat{U}}_k) = \rho(\hat{X}, \tilde{\hat{U}}_k) - \rho(\hat{X}, \tilde{\hat{U}}_{k-1}) \tag{2}$$

that quantifies how much additional energy is recovered by adding the k^{th} basis vector to the subspace [9].

The principal calculation required with subspace methods is the precomputation of estimates of the left singular vectors $\tilde{\hat{U}}_k$ of the $m \times n$ matrix $\hat{X}$. This is a very computationally expensive operation when m and n are large. Reducing this computational expense by exploiting correlation between images has been the topic of much previous research [6]–[12]. In [9], Chang *et al.* showed that if the image data matrix was correlated in one dimension, then the right singular vectors are approximately spanned by a few low frequency harmonics. As a result, the FFT of the image data matrix may be used to approximate the desired subspace dimension k, as well as the principal eigenimages $\hat{U}_k$. This result has been extended to correlation in two dimensions by using the SHT in place of the FFT [11]. In the current work we extend the approach in [9] and [11] to the full rotation group ($SO(3)$). By extending the SHT to the full rotation group, fully general 3-D pose estimation is made possible using the Wigner-D rotation matrices. A brief introduction to the SHT and Wigner-D matrices is presented in the next section.

III. Spherical Harmonics and Wigner-D Matrices

A. Introduction

Spherical harmonics have been applied to a variety of problems that arise on the surface of the unit sphere (denoted as the 2-sphere or S^2) [11]. In recent years, due to the ever increasing amount of digital data, the development of a fast discrete SHT has been a growing area of research [13]–[15][1].

The first step in computing a discrete SHT is to determine how the sphere should be discretized. Three popular discretizations are commonly used, namely, the Gauss-Legendre grid [13], [17], an equi-angular grid of Chebyshev nodes [14], [18], and the Hierarchical Equal Area isoLatitude Pixelization (HEALPix) grid [15]. In [19], the authors compared the three above mentioned discretizations and it was determined that for this particular application, the HEALPix discretization performed the best in terms of better angular resolution in sampling and better estimation of the eigenspace. Therefore, in this paper we use the HEALPix discretization to define the sampling pattern over the sphere's surface. (Refer to Fig. 1 for an example of the HEALPix sampling pattern.)

B. Discrete Spherical Harmonic Transform on $SO(3)$

A real valued band-limited function $f(\boldsymbol{\xi}_{p_0})$ whose domain is $L^2(S^2)$ may be represented by its spherical harmonic expansion as

$$f(\boldsymbol{\xi}_{p_0}) = \sum_{l=0}^{l_{\max}} \sum_{|m|\leq l} f_l^m Y_l^m(\boldsymbol{\xi}_{p_0}) \tag{3}$$

where $Y_l^m(\boldsymbol{\xi}_{p_0})$ is the spherical harmonic of degree l and order m, and f_l^m is the corresponding Fourier coefficient. Note that $\boldsymbol{\xi}_{p_0}$ represents a point on S^2, and can be rotated via an element of the rotation group. The rotation group is the set of real 3×3 orthogonal matrices of determinant +1, which define proper rotations about the origin of $\mathbb{R}^3$. In spectral theory, it is often the convention to define these matrices using standard $z - y - z$ Euler rotation matrices where the z-axis is the upper pole. Therefore, any rotation $g(\alpha, \beta, \gamma) \in SO(3)$ can be written as

$$g(\alpha, \beta, \gamma) = R_z(\alpha) R_y(\beta) R_z(\gamma) \tag{4}$$

[1]This topic has been addressed in various ways dating back to the 1800's [16].

where $R_z(\alpha)$ and $R_y(\beta)$ represent a rotation about the z-axis by α radians, and a rotation about the y-axis by β radians respectively. Given any $g \in SO(3)$, we define the rotation of the function $f(\boldsymbol{\xi}_{p_0}) \in L^2(S^2)$ by the linear operator Λ_g such that

$$\Lambda_g f(\boldsymbol{\xi}_{p_0}) = f(g^{-1}\boldsymbol{\xi}_{p_0}). \tag{5}$$

Applying the operator Λ_g to the function $f(\boldsymbol{\xi}_{p_0})$ is simply a rotation of the function restricted to the surface of the sphere [20]. The effect this has on the function in the spectral domain (i.e. the effect on the harmonic coefficient f_l^m) can be deduced from the fact that rotated versions of the spherical harmonics are simply linear combinations of harmonics of the same degree. That is,

$$\Lambda_g Y_l^m(\boldsymbol{\xi}_{p_0}) = \sum_{|k|\leq l} Y_l^k(\boldsymbol{\xi}_{p_0}) D_{lm}^k(\boldsymbol{\xi}_{p_r}) \tag{6}$$

where $D_{lm}^k(\boldsymbol{\xi}_{p_r})$ is the $(2l+1)\times(2l+1)$ Wigner-D matrix.

Using (6), it can be shown that a function $f(\boldsymbol{\xi}_{p_r}) \in L^2(SO(3))$ may be represented by its harmonic expansion using Wigner-D matrices [20], [21]. That is,

$$f(\boldsymbol{\xi}_{p_r}) = \sum_{l=0}^{l_{\max}} \sum_{|m|\leq l} \sum_{|m'|\leq l} f_{mm'}^l D_{mm'}^l(\boldsymbol{\xi}_{p_r}) \tag{7}$$

where $f(\boldsymbol{\xi}_{p_r}) \in [0,1]$ is a single pixel of the image data vector $\boldsymbol{f}(\boldsymbol{\xi}_{p_r})$. Once again we remind the reader that $\boldsymbol{\xi}_{p_r}$, $p \in \{0,\ldots,a-1\}$ is the unit vector pointing at the angle of co-latitude $\beta_p \in [0,\pi]$ measured down from the upper pole, and the angle of longitude $\alpha_p \in [0,2\pi)$, which is the parameterization of the sphere in spherical coordinates. In $\boldsymbol{\xi}_{p_r}$, the value $r \in \{0,\ldots,b-1\}$ is the r^{th} planar rotation $\gamma_r \in [0,2\pi)$ at sample p. In the above equation, it is assumed that the signal power for $l > l_{\max}$ is insignificant, and $l_{\max}$ is chosen such that aliasing does not occur. The expansion coefficients $f_{mm'}^l$ are then calculated using

$$f_{mm'}^l = \frac{4\pi}{a} \sum_{p=0}^{a-1} \sum_{r=0}^{b-1} f(\boldsymbol{\xi}_{p_r}) D_{mm'}^{l*}(\boldsymbol{\xi}_{p_r}). \tag{8}$$

In the above equations, with the unit vector $\boldsymbol{\xi}_{p_r}$ parameterized by $(\alpha_p, \beta_p, \gamma_r)$,

$$D_{mm'}^l(\alpha_p,\beta_p,\gamma_r) = e^{-im\alpha_p} d_{mm'}^l(\beta_p) e^{-im'\gamma_r} \tag{9}$$

where $d_{mm'}^l(\beta_p)$ is known as Wigner's (small) d-matrix defined by

$$d_{mm'}^l(\beta_p) = \sqrt{\tfrac{(l+m')!(l-m')!}{(l+m)!(l-m)!}} \left(\sin\tfrac{\beta_p}{2}\right)^{m'-m} \left(\cos\tfrac{\beta_p}{2}\right)^{m+m'} \times P_{(m'-m,m+m')}^{(l-m')}(\cos\beta_p) \tag{10}$$

and $P_{(\cdot,\cdot)}^{(\cdot)}(x)$ is a Jacobi polynomial. For computational convenience, the d-matrices may be computed quickly using a three term recurrence relationship [20], [22], [23].

Because $f(\boldsymbol{\xi}_{p_r})$ is a real valued function, it is more convenient to work with the real valued Wigner-D matrices denoted here as $\Delta_{mm'}^l$. The construction of rotation matrices in the basis of real spherical harmonics is discussed in [23]–[25], and can be defined as

$$\begin{aligned}\Delta_{mm'}^l = {} & \operatorname{sign}(m')\Phi_m(\alpha_p)\Phi_{m'}(\gamma_r)\tfrac{d_{|m'||m|}^l + (-1)^m d_{|m|(-|m'|)}^l}{2} \\ & -\operatorname{sign}(m)\Phi_{-m}(\alpha_p)\Phi_{-m'}(\gamma_r)\tfrac{d_{|m'||m|}^l - (-1)^m d_{|m|(-|m'|)}^l}{2}\end{aligned} \tag{11}$$

where

$$\Phi_m(x) = \begin{cases} \sqrt{2}\cos(mx) & \text{if } m > 0 \\ 1 & \text{if } m = 0 \\ \sqrt{2}\sin(|m|x) & \text{if } m < 0 \end{cases}, \tag{12}$$

and β_p has been omitted from the Wigner-d matrices for notational convenience. The $SO(3)$ FFT may then be computed using (8) and replacing $D_{mm'}^{l*}(\boldsymbol{\xi}_{p_r})$ with $\Delta_{mm'}^l(\boldsymbol{\xi}_{p_r})$.

IV. Fast Eigendecomposition Algorithm

Our objective is to estimate the first k principal eigenimages $\tilde{\hat{U}}_k$ of $\hat{X}$ such that $\Delta\rho(\hat{X}, \tilde{\hat{U}}_k) \leq \epsilon$, where ϵ is the user specified change in energy. The first step in computing the principal eigenimages, is to construct the image data matrix $\hat{X}$. The approach taken here is to consider the object placed at the center of an imaginary unit sphere (Fig. 1) and to sample $SO(3)$ appropriately. The mean image is then subtracted from each sampled image to form the image data matrix $\hat{X}$. As mentioned in Section III-A, we use the HEALPix sampling pattern for the discretization of the sphere.

Using the HEALPix sampling pattern is based on subdividing the sphere using the parameter N_{side}, resulting in $a = 12N_{\text{side}}^2$ sample points on the sphere [15]. At each of the $12N_{\text{side}}^2$ sample points, b planar rotated images of the object are captured by rotating the camera through an angle γ_r. The number of planar rotated images captured depends on the angular resolution of the $12N_{\text{side}}^2$ samples defined on S^2. The angular resolution is denoted θ_{pix} and calculated as $\theta_{\text{pix}} = \sqrt{\frac{3}{\pi}}\frac{60^o}{N_{\text{side}}}$. To maintain homogeneous sampling on $SO(3)$, $b = \lfloor(360/\theta_{\text{pix}})\rfloor$, where $\lfloor(\cdot)\rfloor$ is the highest integer less than or equal to $(\cdot)$. Finally, to prevent aliasing, $l_{\max} = 3N_{\text{side}} - 1$ is used in the forward transform. Because the HEALPix sampling pattern is isolatitudinal, the

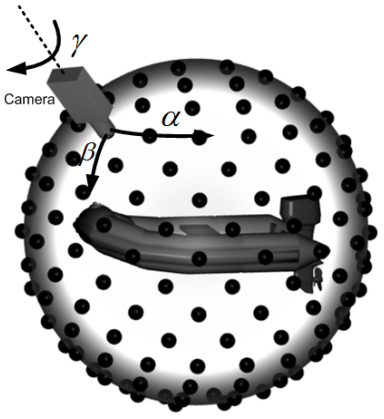

Fig. 1. Image acquisition for the training image data matrix $\hat{X}$, where b planar rotated sample images are taken at each black dot on the sphere.

Fig. 2. Ray-traced CAD models courtesy of Kator Legaz [26]. Each object is sampled as discussed above at a resolution of 128 × 128. Each of the above images are both scale and intensity normalized. The objects are ordered from left to right, then top to bottom.

computation of the Wigner-d matrices (which is the most computationally expensive portion of the $SO(3)$ FFT) is minimal.

Once the image data matrix $\hat{X}$ has been constructed, we compute the matrix F whose i^{th} row is the $SO(3)$ FFT of the i^{th} row of $\hat{X}$, denoted from this point forward as $SOFT(\hat{X})$. This can be computed quickly using the methods described in [20], however for small $l_{\max}$ the computational savings of this method are slim. Alternatively, $SOFT(\hat{X})$ can be cast as the matrix multiplication

$$F = \hat{X}Z, \tag{13}$$

where $\hat{X} \in \mathbb{R}^{m\times n}$ is the image data matrix with the image mean removed, and the images ordered in terms of $\boldsymbol{\xi}_{p_r}$ as

$$\begin{aligned}\hat{X} \;=\; & [\boldsymbol{f}(\boldsymbol{\xi}_{0_0}), \boldsymbol{f}(\boldsymbol{\xi}_{0_1}), \cdots, \boldsymbol{f}(\boldsymbol{\xi}_{0_{b-1}}),\\ & \boldsymbol{f}(\boldsymbol{\xi}_{1_0}), \boldsymbol{f}(\boldsymbol{\xi}_{1_1}), \cdots, \boldsymbol{f}(\boldsymbol{\xi}_{1_{b-1}}), \cdots,\\ & \boldsymbol{f}(\boldsymbol{\xi}_{a-1_0}), \boldsymbol{f}(\boldsymbol{\xi}_{a-1_1}), \cdots, \boldsymbol{f}(\boldsymbol{\xi}_{a-1_{b-1}})].\end{aligned} \tag{14}$$

The matrices $\Delta^l_{mm'}(\cdot) \in \mathbb{R}^{(2l+1)\times(2l+1)}$ may be row scanned such that each row is concatenated to form the row vector $\boldsymbol{\delta}^l(\cdot) = \text{vec}(\Delta^l_{mm'}(\cdot)) \in \mathbb{R}^{1\times(2l+1)^2}$ for any given l. Using this notation, the matrix Z can be constructed as

$$Z = \begin{bmatrix} \boldsymbol{\delta}^0(\boldsymbol{\xi}_{0_0}) & \cdots & \boldsymbol{\delta}^{3N_{\text{side}}-1}(\boldsymbol{\xi}_{0_0}) \\ \boldsymbol{\delta}^0(\boldsymbol{\xi}_{0_1}) & \cdots & \boldsymbol{\delta}^{3N_{\text{side}}-1}(\boldsymbol{\xi}_{0_1}) \\ \vdots & \vdots & \vdots \\ \boldsymbol{\delta}^0(\boldsymbol{\xi}_{a-1_{b-1}}) & \cdots & \boldsymbol{\delta}^{3N_{\text{side}}-1}(\boldsymbol{\xi}_{a-1_{b-1}}) \end{bmatrix} \tag{15}$$

where the rows of Z are ordered in terms of $\boldsymbol{\xi}_{p_r}$ such that the matrix product $F = \hat{X}Z$ makes sense. Note that in computing $SOFT(\hat{X})$ the matrix Z may be pre-computed for several different values of $l_{\max}$ and stored for later use.

We now propose an algorithm for estimating the first k principal eigenimages $\tilde{U}_k$ of $\hat{X}$ such that $\Delta\rho(\hat{X}, \tilde{U}_k) \leq \epsilon$, where ϵ is the user specified change in energy.

EIGENDECOMPOSITION ALGORITHM

1) Form the matrix F which is the $SOFT(\hat{X})$.
2) Form the matrix H whose columns are the ordered columns of F in descending order according to their norm.
3) Set $q = \lfloor N_{\text{side}}(36N^2_{\text{side}} - 1)[1 - (1/2)^{N+1}] \rfloor$, with N=0 initially.
4) Construct the matrix H_q which is the matrix consisting of the first q columns of H.
5) Compute SVD(H_q) = $\tilde{\hat{U}}_q \tilde{\hat{\Sigma}}_q \tilde{\hat{V}}_q^T$. (The key observation here is that H_q contains q columns, which is considerably less than the n columns of $\hat{X}$.)
6) If $\Delta\rho(\hat{X}, \tilde{\hat{U}}_q) > \epsilon$. Let $N = N + 1$ and repeat Steps 3 through 6. Because the SVD of H_q is already available, the eigenspace can simply be updated by modifying the algorithm outlined in [8].
7) Return $\tilde{\hat{U}}_k$ such that $\Delta\rho(\hat{X}, \tilde{\hat{U}}_k) \leq \epsilon$. Note $k \leq q$.

Even though the image data matrix $\hat{X}$ is correlated in three-dimensions (i.e., correlated in α, β, and γ), most of the energy of $\hat{X}$ is contained in the low frequency harmonics of F. A typical example of this is shown in Fig. 3, which is the $SO(3)$ power spectra of object (15) from Fig. 2 for degrees $l = 2, 3, 4, 5$. As is apparent from the figure, as the degree l increases, the $SO(3)$ power spectra becomes significantly attenuated. The key advantage of the proposed algorithm follows from the fact that the principal eigenimages $\tilde{\hat{U}}_k$ of the SVD of the lower frequency harmonics of F serve as excellent estimates to those of $\hat{X}$ at a significant computational savings.

It is important to note that the full SVD($\hat{X}$) requires on the order of mn^2 flops. Therefore computing the SVD on a

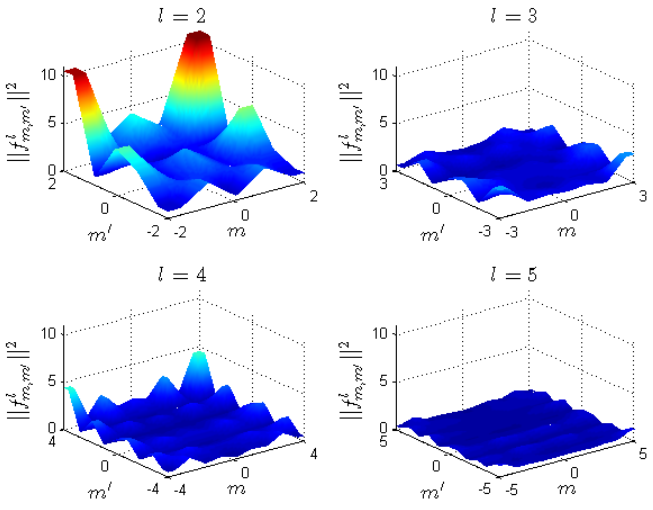

Fig. 3. This figure shows the power spectra $||\boldsymbol{f}^l_{mm'}||^2$ of object (15) from Fig. 2 for degrees $l = 2, 3, 4, 5$. As can be seen from the figure, most of the energy is concentrated around the lower frequency harmonics, i.e. those with lower values of l. (We omit $l = 0$ as this represents the mean value and is significantly higher in magnitude.)

fewer number of columns significantly reduces the computational expense. It is difficult to evaluate the computational complexity of the above algorithm due to the binary split used to determine k. If $N = 0$ however (step 3 of the algorithm), the complexity is on the order of mq^2 flops where $q = \lfloor \frac{N_{\text{side}}}{2}(36N_{\text{side}}^2 - 1) \rfloor$. This is the cost of computing the SVD(H_q) in step 5 of the algorithm. In our experience, for the objects listed in Fig. 2, we have never needed to use more that the first split ($N = 0$) in step 3 of the algorithm for accurate pose estimation.

V. Experimental Results

A. Test Data

The proposed algorithm detailed in Section IV was tested on the ray-traced CAD objects shown in Fig. 2 (the CAD models were provided by [26]). The parameter $N_{\text{side}} = 5$ was used, resulting in $b = 30$ and $12bN_{\text{side}}^2 = 9000$ images per object. The images were then both scale and intensity normalized to create the image data matrix X. The "unbiased" image data matrix $\hat{X}$ was then constructed by subtracting the mean image from the image data matrix X. Finally, the matrix F was computed condensing the image data set from 9000 images to 4495 harmonic images. The true SVD($\hat{X}$) was also computed for a ground truth comparison.

B. Performance and Computational Savings

Fig. 4 shows a plot of the energy recovery ratio ρ [top] and the change in energy recovered $\Delta\rho$ [bottom] as a function of the subspace dimension k averaged across all objects in Fig. 2. As can be seen from the figure, the estimated left singular vectors $\tilde{\hat{U}}_k$ computed by the proposed algorithm are very good approximations to the true left singular vectors as computed by the direct SVD in terms of being able to recover the energy in $\hat{X}$. This data is for the first split $N = 0$, i.e., step 3 of the proposed algorithm, if the entire $N_{\text{side}}(36N_{\text{side}}^2 - 1)$ harmonic images are used, the difference becomes indistinguishable. Note that when computing quantities dealing with energy (i.e. ρ and $\Delta\rho$), the energy remaining after the mean image is subtracted is the quantity computed.

Table I shows the required subspace dimension k, the amount of energy recovered at this subspace dimension, and the time required to estimate the first k left singular vectors $\tilde{\hat{U}}_k$ for each object to meet the user specified change in energy $\epsilon = 0.01$. This result is compared to the true SVD as computed by MATLAB. As is apparent from the table, using the proposed algorithm, the left singular vectors $\tilde{\hat{U}}_k$ are very good estimates of $\hat{U}_k$ at a significant computational savings. Again, only the first split $N = 0$ (step 3) of the algorithm is was used to compute this data, resulting in an average speed-up factor of over 40.

C. Error Analysis

Based on the data provided in Table I, and Fig. 4, it is obvious that the proposed algorithm is capable of estimating

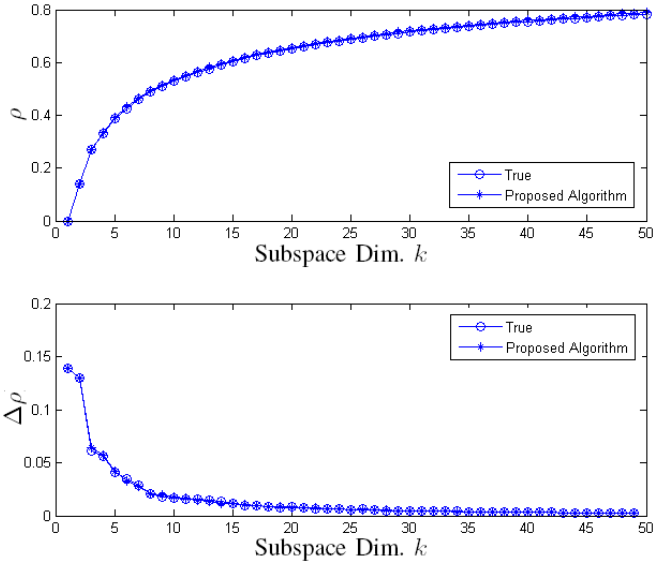

Fig. 4. This figure shows the energy recovery ratio ρ [top] and change in energy $\Delta\rho$ [bottom] vs. subspace dimension k averaged across all 20 objects in Fig. 2.

TABLE I

THIS TABLE SHOWS THE REQUIRED SUBSPACE DIMENSION k, THE AMOUNT OF ENERGY RECOVERED AT THIS SUBSPACE DIMENSION, AND THE TIME REQUIRED TO ESTIMATE THE FIRST k LEFT SINGULAR VECTORS FOR EACH OBJECT TO MEET THE USER SPECIFIED CHANGE IN ENERGY $\epsilon = 0.01$. THE RESULTS ARE COMPARED AGAINST THE TRUE SVD USING MATLAB.

Object	k	Time [hours]		Energy ρ [%]	
		True	Proposed	True	Proposed
1	19	3.550	0.082	66.99	66.94
2	17	3.432	0.076	68.39	68.22
3	16	3.544	0.083	63.80	63.42
4	16	3.558	0.124	54.68	54.30
5	14	3.589	0.076	63.78	63.53
6	14	3.322	0.076	76.66	76.33
7	19	2.424	0.092	60.40	60.13
8	16	3.688	0.101	40.65	39.83
9	16	3.678	0.084	60.12	58.83
10	21	3.634	0.086	60.05	59.63
11	20	3.584	0.083	64.76	64.44
12	15	3.598	0.096	57.61	56.21
13	10	3.403	0.102	50.65	48.31
14	15	3.627	0.091	64.35	64.25
15	16	3.393	0.093	66.43	64.62
16	14	3.881	0.075	70.77	70.62
17	17	3.617	0.095	60.02	59.80
18	15	3.599	0.087	68.43	67.72
19	8	3.542	0.079	71.32	69.46
20	9	4.039	0.077	64.23	63.00
mean		3.53	0.088	62.70	61.98
max		4.039	0.124	40.65	39.83
min		2.424	0.075	76.66	76.33

the required subspace dimension and left singular vectors of $\hat{X}$ at a significant computational savings. However, because the $SO(3)$ FFT is lossy, it is important to try to quantify the amount of energy lost. To do this, all $N_{\text{side}}(36N_{\text{side}}^2 - 1)$ harmonic images were used to estimate the left singular vectors of $\hat{X}$. Using these estimates, the energy recovery ratio $\rho(\hat{X}, \tilde{\hat{U}}_{N_{\text{side}}(36N_{\text{side}}^2-1)})$ was computed for each object in

Fig. 2. The top plot in Fig. 5 shows the maximum amount of energy recovered per object. As seen in the figure, over 99.7% of the energy in $\hat{X}$ is recoverable for all objects when using the proposed algorithm. The bottom plot in Fig. 5 shows the maximum difference in energy recovery per object if only the first $N_{\text{side}}(36N_{\text{side}}^2-1)$ true left singular vectors are used to recover the same energy, i.e., the bottom plot shows $\max(\rho(\hat{X},\hat{U}_k) - \rho(\hat{X},\tilde{\hat{U}}_k))$ for all $k \leq N_{\text{side}}(36N_{\text{side}}^2 - 1)$ for each object. As seen in the figure, the maximum error one could expect to see is less than 3% across the entire subspace for the objects shown in Fig. 2.

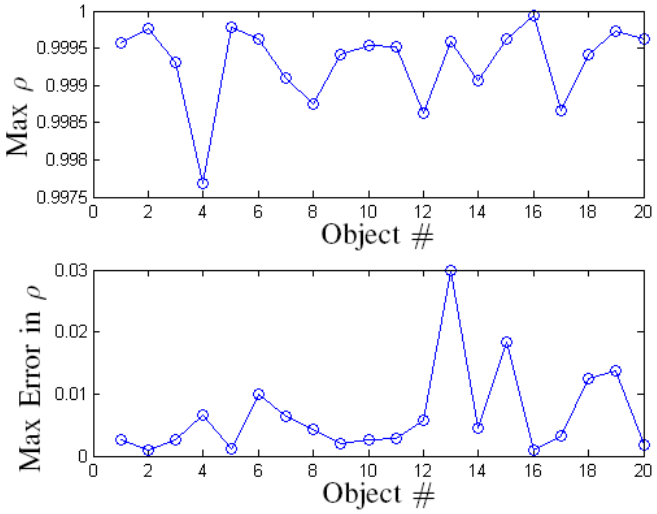

Fig. 5. This figure shows the maximum achievable energy recovery ratio ρ for each object in Fig. 2 [top], and the maximum difference in energy recovery between the first $N_{\text{side}}(36N_{\text{side}}^2 - 1)$ true left singular vectors of SVD($\hat{X}$) and the estimated left singular vectors $\tilde{\hat{U}}_{N_{\text{side}}(36N_{\text{side}}^2-1)}$ as computed by the proposed algorithm [bottom].

VI. Conclusions and Future Work

We have illustrated a computationally efficient algorithm for estimating the eigendecomposition of images correlated in $SO(3)$ by using the discrete SHT and Wigner-D matrices. The algorithm was tested on a variety of 3-D objects with images captured from different vantage points around the sphere. In addition to significant computational savings as compared to the direct SVD approach, we have shown that the estimated eigenimages are very close to the true eigenimages as computed by the direct SVD. We have also shown that the error associated with using the proposed algorithm is minimal with respect to the computational savings. Future work will focus on validating the proposed algorithm on true 3-D objects rather than CAD models and an implementation of a this technique for a real-time visual servo control system.

References

[1] E. N. Malamasa, E. G. Petrakisa, M. Zervakisa, L. Petit, and J.-D. Legat, "A survey on industrial vision systems, applications and tools," *Image and Vision Comp.*, vol. 21, no. 2, pp. 171–188, 2003.

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

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

[4] 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.

[5] 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.

[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] 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.

[10] K. Saitwal, A. A. Maciejewski, and R. G. Roberts, "Computationally efficient eigenspace decomposition of correlated images characterized by three parameters," accepted to appear in *Patt. Anal. and Apps.*, 2008.

[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," in *IEEE Int. Conf. Robot. Automat.*, May 19-23, 2008, pp. 993–998.

[12] 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.

[13] P. N. Swarztrauber and W. F. Spotz, "Generalized discrete spherical harmonic transforms," *J. of Comp. Phys.*, vol. 159, no. 2, pp. 213–230, Apr. 2000.

[14] D. M. Healy Jr., D. Rockmore, P. 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, July 2003.

[15] K. M. Górski, E. Hivon, A. L. Banday, B. D. Wandelt, F. K. Hansen, M. 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.

[16] N. Sneeuw, "Global spherical harmonic analysis by least squares and numerical quadrature methods in historical perspective," http://www.geomatics.ucalgary.ca/~sneeuw/publ/neumann.ps, 2007.

[17] P. N. Swarztrauber, "On the spectral approximation of discrete scalar and vector functions on the sphere," *SIAM J. Numer Anal.*, vol. 16, no. 6, pp. 934–949, Dec. 1979.

[18] J. R. Driscoll and D. M. Healy Jr., "Computing Fourier transforms and convolutions on the 2-sphere," *Adv. App. Math*, vol. 15, no. 2, pp. 202–250, 1994.

[19] R. C. Hoover, A. A. Maciejewski, and R. G. Roberts, "An analysis of sphere tessellations for pose estimation of 3-D objects using spherically correlated images," in *IEEE SSIAI*, March 24-27, 2008, pp. 41–44.

[20] P. J. Kostelec and D. N. Rockmore, "FFTs on the Rotation Group," Santa Fe Institute Working Papers Series Paper #03-11-060, 2003.

[21] A. Makadia and K. Daniilidis, "Rotation recovery from spherical images without correspondences," *IEEE Trans. PAMI*, vol. 28, no. 7, pp. 1170–1175, July 2006.

[22] D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii, *Quantum Theory of Angular Momentum*. New Jersey, USA: World Scientific, 1988.

[23] M. Blanco, M. Florez, and M. Bermejo, "Evaluation of the rotation matrices in the basis of real spherical harmonics," *Journal of Molecular Structure: THEOCHEM*, vol. 419, pp. 19–27(9), 8 December 1997.

[24] J. Ivanic and K. Ruedenberg, "Rotation matrices for real spherical harmonics. Direct determination by recursion," *J. Phys. Chem*, vol. 100, pp. 6342–6347, 1996.

[25] T. Risbo, "Fourier transform summation of Legendre series and D-functions," *J. Phys. Chem*, vol. 100, pp. 6342–6347, 1996.

[26] K. Legaz. (2007) Kator Legaz: 3-D model database for Blender. [Online]. Available: http://www.katorlegaz.com/3d_models/index.php