Fast Eigenspace Decomposition of Images of Objects With Variation in Illumination and Pose

Randy C. Hoover, *Member, IEEE*, Anthony A. Maciejewski, *Fellow, IEEE*, and Rodney G. Roberts, *Senior Member, IEEE*

***Abstract*—Many appearance-based classification problems such as principal component analysis, linear discriminant analysis, and locally preserving projections involve computing the principal components (eigenspace) of a large set of images. Although the online expense associated with appearance-based techniques is small, the offline computational burden becomes prohibitive for practical applications. This paper presents a method to reduce the expense of computing the eigenspace decomposition of a set of images when variations in both illumination and pose are present. In particular, it is shown that the set of images of an object under a wide range of illumination conditions and a fixed pose can be significantly reduced by projecting these data onto a few low-frequency spherical harmonics, producing a set of "harmonic images." It is then shown that the dimensionality of the set of harmonic images at different poses can be further reduced by utilizing the fast Fourier transform. An eigenspace decomposition is then applied in the spectral domain at a much lower dimension, thereby significantly reducing the computational expense. An analysis is also provided, showing that the principal eigenimages computed assuming a single illumination source are capable of recovering a significant amount of information from images of objects when multiple illumination sources exist.**

***Index Terms*—Eigenspace decomposition, Fourier transform, illumination variation, pose estimation, spherical harmonics.**

I. Introduction

APPEARANCE-BASED methods represent one computationally efficient approach for dealing with the online classification of 3-D objects from 2-D images under a wide range of illumination conditions. Appearance-based methods take advantage of the fact that a set of highly correlated images can be represented by a much smaller subspace [1]. The most popular appearance-based methods include principal component analysis (PCA), Fisher's linear discriminant analysis (FLDA), and locally preserving projections (LPP) [1]–[13].

PCA involves finding a linear subspace (referred to as the eigenspace) that maximizes the variance between images in a training set. The high-dimensional image can be projected onto the principal components (referred to as eigenimages), and classification can be carried out by performing a nearest neighbor search in the eigenspace [2]–[6]. An alternate approach for classification (particularly when dealing with multiple classes) is to use a class-specific linear projection such as FLDA [7]. Unlike PCA, rather than finding a linear subspace to maximize the total variance in the image data, FLDA determines the linear subspace that maximizes the variance between different classes in the data while minimizing the variance within each class. A common issue associated with FLDA is the small-sample-size problem. Fortunately, this issue can be overcome by using an intermediate space that is computed using the principal components of the image data [7]–[10].

Both PCA and FLDA find a linear subspace for classification using the global information contained in the data set. LPP and orthogonal LPP (OLPP), on the other hand, attempt to find a linear subspace for classification that accounts for local characteristics of the data as well. The process proceeds by constructing a similarity matrix to account for local information in the image data and a density matrix to measure the local density around each image. The subspace is then computed by solving a generalized eigenvalue problem involving the graph Laplacian. Unfortunately, similar to FLDA, the image data need to be projected onto an intermediate PCA space before the eigenvalue problem can be solved [11]–[13].

Note that all of these techniques require the computation of the principal components before classification can be performed. Computing the principal components of a large set of images is prohibitively expensive and thereby discourages the use of PCA-based techniques in real-world applications. Reducing the computational burden associated with computing the principal components has been addressed using several different approaches based on either iterative power methods, conjugate gradient algorithms, or eigenspace updating [14]–[16]. A fundamentally different approach was proposed by Chang *et al.* [17], where the authors show that the Fourier transform can be used to approximate the desired subspace dimension, as well as the principal eigenimages, if the image data set is correlated in one dimension. This result has recently been extended to correlations in higher dimensions that are due to a change in orientation (assuming constant lighting conditions)

Manuscript received June 30, 2009; revised November 27, 2009 and March 11, 2010; accepted May 12, 2010. Date of publication July 12, 2010; date of current version March 16, 2011. Portions of this paper were presented in part at the IEEE International Conference on Systems, Man, and Cybernetics, San Antonio, TX, October 11–14, 2009. This paper was recommended by Associate Editor L. Wang.

R. C. Hoover is with the Electrical and Computer Engineering Department, South Dakota School of Mines and Technology, Rapid City, SD 57701 USA (e-mail: randy.hoover@sdsmt.edu).

A. A. Maciejewski is with the Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523-1373 USA (e-mail: aam@colostate.edu; www.engr.colostate.edu/~aam).

R. G. Roberts is with the Department of Electrical and Computer Engineering, College of Engineering, Florida Agricultural and Mechanical University–The Florida State University, Tallahassee, FL 32310-6046 USA (e-mail: rroberts@eng.fsu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMCB.2010.2052920

by using spherical harmonics and Wigner-D functions in place of the Fourier transform [18]–[21].

It has been shown empirically that the set of images of a convex largely diffuse object under a wide range of illumination conditions and a fixed pose approximately lies within a 9-D linear subspace [22], [23]. This result was analytically verified in [24] and [25] by expanding the Lambertian kernel in a series of spherical harmonics. The principal eigenimages of the set of images can then be calculated by evaluating a set of spherical harmonics at the surface normals of the object and applying standard ray-tracing techniques. In [26], the authors present a method to efficiently compute the principal eigenimages of an image data set when variations in both illumination (from a *single* distant point light source) and pose exist. The method is based on first reducing the dimensionality of the data due to variations in illumination and a fixed pose by using a truncated series of spherical harmonics, generating a set of "harmonic images." The authors then show that the eigenspace decomposition of the entire data set (variations in illumination and pose) can be efficiently computed by applying Chang's eigenspace decomposition algorithm to the resulting set of harmonic images.

The contributions of the current work are threefold: First, we extend the analysis in [26] by showing that the set of (orthogonalized) harmonic images are very close approximations to the true eigenspace decomposition as computed using standard PCA techniques. Second, we provide an analysis of how effective the principal eigenimages computed assuming a *single* distant point light source are at recovering information from images of the object when *multiple* distant point light sources exist. Finally, we show that the principal components of the entire set of images (variations in illumination and pose) can be efficiently computed by projecting the set of harmonic images onto a Fourier basis using Chang's algorithm. The pose of the object can then be estimated using either PCA, FLDA, LPP, or OLPP, all of which require the principal components of the image data to be computed.

The remainder of this paper is organized as follows. In Section II, the fundamentals needed to apply an eigenspace decomposition to a related image data set are reviewed, much of which are discussed in [17]. Section II also gives a brief overview of PCA, FLDA, LPP, and OLPP, as well as eigenspace estimation and quality measures that can be used to evaluate two different subspaces [27]. In Section III, we give a brief introduction to spherical harmonics and discuss the problem of reducing the dimensionality of the image data in the illumination dimension. This section also presents an analysis on the effects that multiple illumination sources have when using a subspace generated assuming a single illumination source (SS). In Section IV, a brief overview of Chang's eigenspace decomposition algorithm is provided. We then use this to develop a computationally efficient algorithm for estimating the principal eigenimages of a data set due to variations in illumination and pose in Section V. Section VI provides an evaluation of the accuracy and computational expense associated with using the proposed algorithm as compared to the true eigenspace decomposition. Finally, some concluding remarks are provided in Section VII.

TABLE I
NOMENCLATURE USED THROUGHOUT THIS PAPER

$\mathcal{X}$:	image matrix
X:	image data matrix
μ:	mean image of X
$\hat{X}$:	"unbiased" (mean subtracted) image data matrix
h:	number of pixels in one column of each image
v:	number of pixels in one row of each image
m:	total number of pixels ($h \times v$) in each image
n:	total number of images in an image data set
S^2:	the surface of the sphere (2-sphere)
α_i:	angle of longitude
β_i:	angle of co-latitude
r:	r^{th} pose of the object
a:	number of poses of the object
b:	number of illumination directions
$\boldsymbol{\xi}_i$:	unit vector parametrization of S^2
U:	matrix of left singular vectors of $\hat{X}$
Σ:	diagonal matrix of singular values of $\hat{X}$
V:	matrix of right singular vectors of $\hat{X}$
$\mathbf{u}_i$:	the i^{th} column of U
U_k:	matrix consisting of the first k columns of U
$\tilde{U}$:	matrix of estimated left singular vectors of $\hat{X}$
$\tilde{U}_k$:	matrix consisting of the first k columns of $\tilde{U}$
$\tilde{\mathbf{u}}_i$:	estimate of $\mathbf{u}_i$
ρ:	energy recovery ratio
E_o:	user specified energy recovery due to a change in orientation
E:	user specified energy recovery
SC:	subspace criterion
$Y_{p,q}$:	spherical harmonic of degree p and order q
$\mathcal{I}_r$:	set of images of an object under illumination variation at pose r
μ_r:	mean image of $\mathcal{I}_r$
$\boldsymbol{f}(\boldsymbol{\xi}_i, r)$:	image of the object under illumination direction $\boldsymbol{\xi}_i$ at pose r
$\boldsymbol{f}^r_{p,q}$:	harmonic image of degree p and order q
H:	real Fourier matrix
$\hat{X}_{p,q}$:	matrix containing harmonic images of degree p and order q across all poses
$Z^{p,q}_{j_i}$:	projection of the matrix $\hat{X}_{p,q}$ onto the first j_i Fourier harmonics
$\bar{X}$:	matrix containing all nine of the matrices $Z^{p,q}_{j_i}$

II. PRELIMINARIES

A. Mathematical Description[1]

This work considers grayscale images that are described by an $h \times v$ array of square pixels with intensity values normalized between zero and one (Table I). Thus, an image is represented by a matrix $\mathcal{X} \in [0,1]^{h \times v}$. The image $\mathcal{X}$ is then "row scanned" to obtain the *image vector* $\boldsymbol{f}$ of length $m = hv$, 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 X, where, typically, $m > n$ with fixed n [17].

In this paper, we consider sets of images of a known (rigid) object under different pose and illumination conditions. The objects are placed at the center of an *illumination sphere*, while the camera is moved to discrete locations on a line of constant colatitude. At each of the discrete locations, images of the object are captured under a dense but finite set of illumination conditions generated from a single distant point light source arriving from a distinct illumination direction. Under these

[1]The mathematical description of the images used in this work is very similar to that used in [21]; however, we repeat this presentation here to make this paper self-contained.

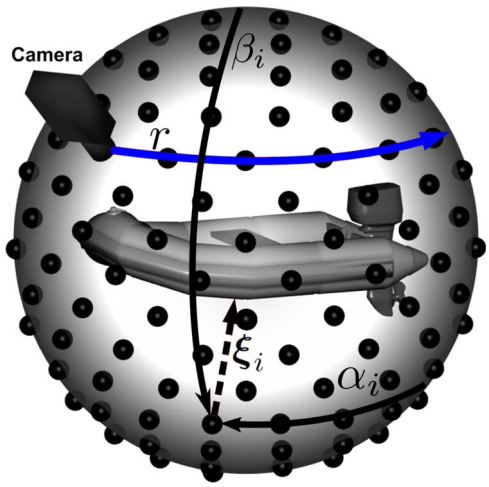

Fig. 1. Graphical depiction of the proposed method of acquiring images from a dense set of illumination conditions at each pose. The object is placed at the center of the illumination sphere with the camera moving along a line of constant colatitude. The black dots on the sphere represent different illumination directions. As the camera moves along the line of colatitude, an image of the object is captured under each of the distinct illumination conditions.

conditions, the image vector can be parameterized by $\boldsymbol{f} = \boldsymbol{f}(\boldsymbol{\xi}_i, r)$, where $r \in \{0, \ldots, a-1\}$ represents the rth pose of the object and $\boldsymbol{\xi}_i, i \in \{0, \ldots, b-1\}$, is the unit vector pointing at the angle of colatitude $\beta_i \in (0, \pi)$, measured down from the upper pole, and the angle of longitude $\alpha_i \in [0, 2\pi)$, which is the parameterization of the ith direction of the point light source at each pose. An example of this procedure is shown in Fig. 1.

We define

$$\mathcal{I}_r = \{\boldsymbol{f}(\boldsymbol{\xi}_i, r) | i = 0, \ldots, b-1\} \tag{1}$$

as the set of images of an object under the illumination directions defined by $\boldsymbol{\xi}_i$ at pose r. Using this notation, the image data matrix X can be constructed as

$$X = [\mathcal{I}_0, \mathcal{I}_1, \ldots, \mathcal{I}_{a-1}] \tag{2}$$

where the first b columns of X correspond to images of the object at a single pose under b different illumination conditions. The average image vector $\mu \in \mathbb{R}^m$ 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.

B. Pose Estimation via Classification

If we are interested in estimating the pose of the object from a new input image $\boldsymbol{f}_{\text{new}}$, the problem then becomes that of determining which of the sets $\mathcal{I}_r$ the new input image belongs to. This can be formulated as a multiclass classification problem where the sets $\mathcal{I}_r$ correspond to a single class and (in the current work) there are $r = 0, \ldots, a-1$ unique classes. One approach to solving this problem would be to simply use a nearest neighbor classifier in the image space. The new input image $\boldsymbol{f}_{\text{new}}$ would be compared to each image in the learning set $\hat{X}$, and classification can be carried out by finding the closest matching image in the image space. Unfortunately, this approach has several drawbacks in terms of computational expense, large storage requirements, and robustness to unknown imaging conditions. Several approaches have been developed to overcome these drawbacks, the three most popular of which are discussed in the following paragraphs.

1) PCA: To overcome the computational expense and storage requirements associated with classifying a single image in the image space, techniques such as PCA can be used to reduce the dimensionality of the search space. PCA attempts to find a linear subspace that maximizes the variance between each image in the learning set $\hat{X}$. The general problem is formulated as follows: Find the matrix $U \in \mathbb{R}^{m \times k}$ with $\mathbf{u}_i$ orthonormal such that the determinant of $U^T S_T U$ is maximized, where $S_T = \hat{X}\hat{X}^T$ is referred to as the total scatter matrix. It can be shown that the matrix U that maximizes the determinant of $U^T S_T U$ is the k eigenvectors of S_T associated with the largest eigenvalues [1]. While PCA is useful for dimensionality reduction, it has been shown that discriminant analysis typically outperforms PCA when dealing with multiclass classification [8], [28]–[30].

2) Discriminant Analysis: An alternate approach to maximizing the total scatter of the learning samples is to use class-specific linear dimensionality reduction such as FLDA [7], [8]. Unlike PCA, FLDA attempts to find a linear subspace that minimizes the within-class scatter while maximizing the between-class scatter. The idea is that the set of images $\mathcal{I}_r$ lies close to a low-dimensional linear subspace and is therefore linearly separable [8], [22], [31], [32]. Unfortunately, in most practical applications, $n < m$, which implies that the within-class scatter matrix is singular. To overcome this issue, the within-class scatter matrix is typically projected onto an intermediate subspace using PCA [8], [28].

3) LPP: Both PCA and FLDA attempt to find a linear subspace for dimensionality reduction using global information contained in the learning samples. LPP and OLPP, however, attempt to account for local structure as well. LPP is based on finding a subspace that preserves local distance in the image space by using the graph Laplacian [11]–[13]. Unfortunately, as with FLDA, singularities arise in the construction of the new basis vectors when $n < m$. Therefore, similar to FLDA, PCA is used by first projecting the learning samples onto the principal components [11]–[13]. It is important to note that all of these classification techniques require the computation of the principal components of the image data matrix $\hat{X}$.

C. Eigenspace Estimation

One approach to computing the principal components of $\hat{X}$ is to utilize the singular-value decomposition (SVD). The *thin* SVD of $\hat{X}$ is given by $\hat{X} = U\Sigma V^T$, where $U \in \mathbb{R}^{m \times n}$ is composed of the eigenvectors of the total scatter matrix S_T. The columns of U, denoted $\mathbf{u}_i$, $i = 1, \ldots, n$, are also referred to as the left singular vectors or the eigenimages of $\hat{X}$. Unfortunately, most SVD algorithms require on the order of $\mathcal{O}(mn^2)$ operations to compute, which is computationally prohibitive when m and n are large. To avoid the direct computation of the principal components, several approaches have been proposed to estimate them based on either iterative power

methods, conjugate gradient algorithms, eigenspace updating, or matrix subsampling [14]–[16], [27], [33], [34].

A fundamentally different approach to estimating the principal components was proposed by Chang *et al.* [17], where the authors 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 Fourier harmonics (refer to Section IV). In [18]–[21], several extensions to Chang's algorithm were developed to account for correlation in higher dimensions. The premise behind these extensions were to replace the Fourier harmonics with spherical harmonics and Wigner-D functions. Finally, in [26], Hoover *et al.* present an efficient technique to compute the principal components of an image data set when variations in both illumination and pose exist. The technique is based on expanding the set of images of an object under a fixed pose and a wide range of illumination conditions (assuming a single distant point light source) into a series of spherical harmonics using the spherical harmonic transform (SHT). This expansion results in a set of harmonic images. Chang's algorithm can then be applied to the set of harmonic images to compute the eigenspace of the entire image data set (variations in both illumination and pose).

In the current work, we extend the results obtained in [26] by providing an analysis of why a low-dimensional set of harmonic images is sufficient to capture most of the variance of an image sequence generated from a fixed pose and a wide range of illumination conditions. We show that, after orthogonalization, the set of harmonic images are very good approximations to the first few eigenimages as computed by the direct SVD. We then extend this analysis by showing that the low-dimensional subspace computed for variations in illumination generated from a *single* illumination source is sufficient to recover a significant amount of information when the object is being illuminated by *multiple* illumination sources. In fact, it is shown that the major effect of multiple illumination sources is a higher probability that local specularities will be illuminated. We then show that the principal components of the entire data set can be estimated by applying Chang's algorithm to the resulting set of harmonic images. Classification can then be accomplished by applying any of the techniques described in Section II-B. Because we are estimating the principal components rather than explicitly computing them, two quality measures are defined in the next section to evaluate the accuracy of the estimates [27].

D. Quality Measures[2]

1) Energy Recovery Ratio: True and approximated eigenimages of $\hat{X}$ can also be compared in terms of their capability of recovering the amount of the total energy in $\hat{X}$. The "energy recovery ratio," denoted ρ, is defined as [17]

$$\rho(\hat{X}, \tilde{U}_k) = \frac{\sum_{i=1}^{k} \left\| \tilde{\mathbf{u}}_i^T \hat{X} \right\|^2}{\|\hat{X}\|_F^2}. \tag{3}$$

Note that, if the $\tilde{\mathbf{u}}_i$'s are orthonormal, $\rho \leq 1$.

[2]The quality measures used in this work are the same as those described in [21].

2) SC: True eigenimages give an optimum energy recovery ratio; therefore, it is possible that more estimated eigenimages are required to achieve the same energy recovery ratio. Thus, another measure used in this study is the degree to which estimated eigenimages span the subspace of the first k^* true eigenimages. This measure is referred to as the subspace criterion (SC) given by

$$SC = \sqrt{\frac{1}{k^*} \sum_{i=1}^{k} \sum_{j=1}^{k^*} \left(\tilde{\mathbf{u}}_i^T \mathbf{u}_j \right)^2}. \tag{4}$$

III. Spherical Harmonics

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). They have been used for solving partial differential equations in spherical geometry for weather and climate models [35], geophysics [36], [37], and quantum mechanics [38], [39], as well as a host of other related applications [40]. Over the last decade, spherical harmonics have been gaining popularity in the computer vision and computer graphics arena. Spherical harmonics have been applied to several computer vision applications with unknown lighting [24], [25], [41]–[43], as well as 3-D model retrieval [44], [45], 3-D shape descriptors [46], and pose estimation [18]–[21]. Spherical harmonics have also been applied to rotation estimation and convolution of spherical images [47].

Spherical harmonics, denoted $Y_{p,q}$, are the angular solutions to Laplace's equation in spherical coordinates and have the factorization

$$Y_{p,q}(\alpha, \beta) = \kappa_{p,q} P_{p,q}\left(\cos(\beta)\right) e^{iq\alpha} \tag{5}$$

where $P_{p,q}(\cos(\beta))$ is the associated Legendre polynomial of degree p and order q and $\kappa_{p,q}$ is a normalization constant, which is equal to

$$\kappa_{p,q} = \sqrt{\left(\frac{2p+1}{4\pi}\right) \frac{(p-|q|)!}{(p+|q|)!}}. \tag{6}$$

Using the aforementioned normalization, the spherical harmonics satisfy the condition

$$\int_0^{2\pi} \int_0^{\pi} Y_{p_1,q_1} \left(Y_{p_2,q_2}\right)^* \sin(\beta)\, d\beta\, d\alpha = \delta_{p_1 p_2} \delta_{q_1 q_2} \tag{7}$$

where the superscript * is the complex conjugate and δ_{ij} is the Kronecker delta, i.e., they form an orthonormal basis for S^2. As a result, any square integrable function $f(\alpha, \beta) \in L^2(S^2)$, where L^2 is the Hilbert space of square integrable functions, may be projected onto this basis as

$$f(\alpha, \beta) = \sum_{p=0}^{\infty} \sum_{|q| \leq p} f_{p,q} Y_{p,q}(\alpha, \beta) \tag{8}$$

with the expansion coefficients $f_{p,q}$ computed as

$$f_{p,q} = \int_0^{2\pi} \int_0^{\pi} f(\alpha, \beta) Y_{p,q}(\alpha, \beta)^* \sin(\beta)\, d\beta\, d\alpha. \tag{9}$$

Unfortunately, computing the harmonic coefficients by evaluating the integrals in (9) is prohibitively expensive. In order to reduce this computational expense, the integrals in (9) need to be approximated by finite sums, and the development of a discrete SHT is needed. This topic has been addressed in various ways dating back to the 1800s [48] and is the subject of the next section.

B. Discrete SHT

In the development of a discrete SHT, a first step is deciding the best discretization of the sphere to define the sampling pattern. Three popular discretizations are commonly used when performing spectral analysis on the surface of the sphere. In [49] and [50], Swarztrauber and Spotz proposed a method for computing the discrete SHT using the Gauss–Legendre tessellation, as well as an efficient method for computing the quadrature weights and points [51]. Alternatively, Driscoll and Healy [52] and Healy *et al.* [53] proposed a method for computing the discrete SHT using an equiangular grid of Chebyshev nodes. In [54], Górski *et al.* propose the hierarchical equal-area isolatitude pixelization (HEALPix) which has the advantage that the samples have equal area weighting over S^2, and as a result, they do not oversample the polar regions. In [55], the authors compared the three tessellations and determined that, for the application of eigenspace decomposition, the HEALPix discretization performed the best in terms of angular resolution in sampling and estimation of the eigenspace. Therefore, in the current work, the HEALPix discretization is used to define the sampling pattern over S^2.

Using the HEALPix discretization, a real-valued band-limited function $\boldsymbol{f}(\boldsymbol{\xi}_i)$ whose domain is $L^2(S^2)$ can be represented by its discrete spherical harmonic expansion as

$$f(\boldsymbol{\xi}_i) = \sum_{p=0}^{p_{\max}} \sum_{|q|\le p} f_{p,q} Y_{p,q}(\boldsymbol{\xi}_i). \tag{10}$$

Recall that $\boldsymbol{\xi}_i$, $i \in \{0, \ldots, b-1\}$, is the unit vector pointing at the angle of colatitude $\beta_i \in (0, \pi)$, measured down from the upper pole, and the angle of longitude $\alpha_i \in [0, 2\pi)$, which is the parameterization of the sphere in spherical coordinates. In (10), it is assumed that the signal power for $p > p_{\max}$ is insignificant, and $p_{\max}$ is chosen to prevent aliasing. The expansion coefficients are calculated using

$$f_{p,q} = \frac{4\pi}{n} \sum_{i=0}^{n-1} f(\boldsymbol{\xi}_i) Y_{p,q}(\boldsymbol{\xi}_i) \tag{11}$$

where $Y_{p,q}(\boldsymbol{\xi}_i)$ is the real-valued spherical harmonic defined by

$$Y_{p,q}(\boldsymbol{\xi}_i) = \begin{cases} \sqrt{2}\kappa_{p,q} \cos(q\alpha_i) P_{p,q}(x), & \text{if } q > 0 \\ \sqrt{2}\kappa_{p,q} \sin(|q|\alpha_i) P_{p,|q|}(x), & \text{if } q < 0 \\ \kappa_{p,0} P_{p,0}(x), & \text{if } q = 0 \end{cases} \tag{12}$$

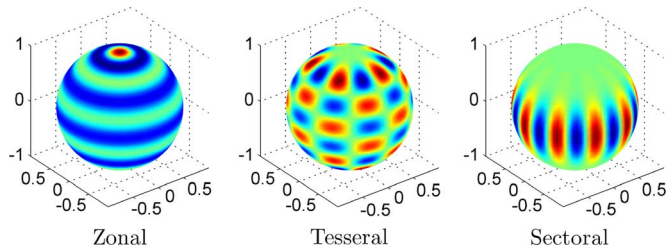

Fig. 2. Real spherical harmonics $Y_{p,q}(\cdot)$ for $p = 8$. The left plot is for $q = 0$ and is referred to as the zonal harmonic, the center plot is for $q = p/2$ and is referred to as a tesseral harmonic, and the right plot is for $q = p$ and is referred to as a sectoral harmonic.

and $P_{p,0}(x) = P_p(x)$ is the Legendre polynomial of degree p with $x = \cos(\beta_i)$. In the development of the S^2 discrete SHT, real spherical harmonics are used because the functions that we are dealing with (namely, images) are real valued. Examples of the real spherical harmonics projected onto the sphere for $p = 8$ and three different values of q are shown in Fig. 2.

C. Spherical Harmonic Images

In the context of the current work, the set of images of an object under variations in illumination conditions and constant pose can be considered to be a (multivalued) band-limited function defined on $L^2(S^2)$. If, however, we consider a single pixel of each image in the set $\mathcal{I}_r$ and reparameterize $f(\boldsymbol{\xi}_i)$ and $f_{p,q}$ in (10), the discrete SHT can be constructed as

$$f(\boldsymbol{\xi}_i, r) = \sum_{p=0}^{p_{\max}} \sum_{|q|\le p} f_{p,q}^r Y_{p,q}(\boldsymbol{\xi}_i) \tag{13}$$

where $f(\boldsymbol{\xi}_i, r) \in [0, 1]$ is a single pixel of the image data vector $\boldsymbol{f}(\boldsymbol{\xi}_i, r)$ at pose r, $Y_{p,q}(\boldsymbol{\xi}_i)$ is the real spherical harmonic of degree p and order q defined in (12), $f_{p,q}^r$ is the corresponding harmonic coefficient at pose r, and $p_{\max} = (\sqrt{3b}/2) - 1$. The harmonic coefficients $f_{p,q}^r$ are calculated using

$$f_{p,q}^r = \frac{4\pi}{b} \sum_{i=0}^{b-1} f(\boldsymbol{\xi}_i, r) Y_{p,q}(\boldsymbol{\xi}_i). \tag{14}$$

Recall that $\{\boldsymbol{f}(\boldsymbol{\xi}_i, r) | i = 0, \ldots, b-1\}$ is the set $\mathcal{I}_r$ which corresponds to images of the object under the illumination directions defined by $\boldsymbol{\xi}_i$ at pose r. If all m pixels of this set are expanded using (14), then $\boldsymbol{f}_{p,q}^r \in \mathbb{R}^{m \times 1}$ represents a harmonic image of degree p and order q at pose r. Our goal is to verify that, for most objects, orthonormalizing the set of harmonic images provides a good approximation to the eigenimages as computed by using the SVD directly.

To illustrate this for a simple object, a set of images of a Lambertian sphere was collected under 768 different illumination directions defined by the HEALPix spherical discretization. The energy recovery ratio defined in (3) was then computed, and it was determined that the first five eigenimages are capable of recovering over 99.8% of the energy of the set of images. To understand why this is the case, the spherical harmonic power spectra $\|\boldsymbol{f}_{p,q}\|$ were also computed for the Lambertian sphere, the results of which are shown in Fig. 3.

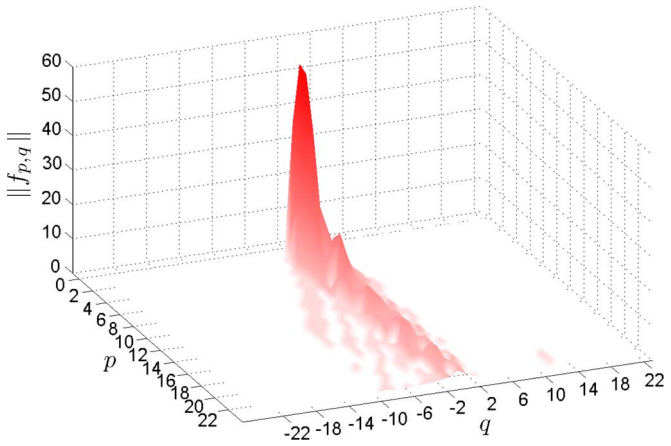

Fig. 3. Spherical harmonic power spectra $\|\boldsymbol{f}_{p,q}\|$ of the test sphere. As can be seen from the figure, most of the energy is concentrated around the low-frequency spherical harmonics. Note that the value r in $\|\boldsymbol{f}_{p,q}\|$ is omitted because the sphere looks the same from every pose.

Notice that the majority of the energy in the power spectra is concentrated around the low-frequency spherical harmonics ($p \leq 2$), which implies that, for the Lambertian sphere, the set $\mathcal{I}_r$ is extremely band limited. This further supports the claim that the set of images of an object under a constant pose and variation in illumination conditions can be well represented by a low-dimensional linear subspace [22], [24], [25]. Although images of a Lambertian sphere under illumination variation represent an admittedly ideal condition, in the next section, we show that arbitrary objects (non-Lambertian and nonconvex) are still well represented by a low-dimensional set of harmonic images.

D. SS

Prior to analyzing the effects of multiple illumination sources, we first consider the case where an arbitrary object is illuminated by a single distant point light source. Our goal is to verify that, for most objects, orthonormalizing the set of harmonic images provides a good approximation to the eigenimages as computed by using the SVD directly.

To illustrate this, computer-aided design (CAD)-generated ray-traced images of 20 different objects of varying surface material properties were captured from 90 different poses and 48 different light source directions at each pose. An example image of each object is shown in Fig. 4 (the CAD models were provided by Legaz [56]). For each of the 90 different poses, the harmonic transform in (14) was used to reduce the dimensionality of the data from 48 images to 9, 16, 25, and 36 harmonic images, i.e., $p = 2$, 3, 4, and 5, respectively. The harmonic images were then orthonormalized, and the energy recovery ratio defined in (3) was used to compute how much energy each of the four subspaces is capable of recovering at each pose. The minimum amount of energy recovered across all 90 poses for all 20 objects is shown in Fig. 5. Notice that, with the exception of objects 17, 18, and 20, over 95% of the energy is recovered by the 9-D subspace for all 90 poses. Furthermore, adding additional harmonic images does not significantly increase the amount of energy recovered. To evaluate how well the low-dimensional set of orthonormalized harmonic images estimates the true eigenimages as computed by the SVD, the first 9, 16, 25, and 36 true eigenimages were also computed for each of the 90 different poses. Fig. 6

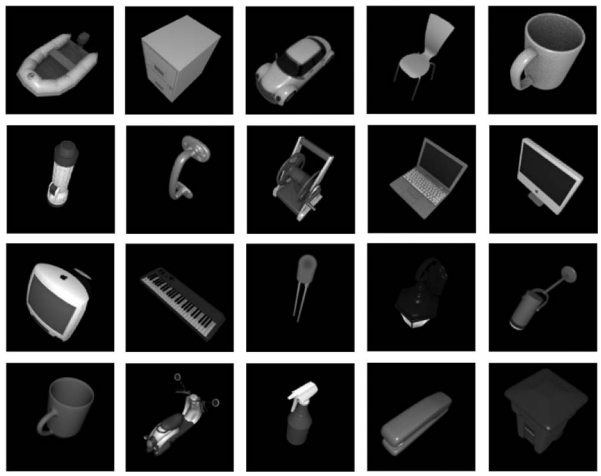

Fig. 4. Ray-traced CAD models courtesy of Legaz [56]. Each object is sampled, as discussed in Section II, at a resolution of 128 × 128. The objects are ordered from left to right and then top to bottom.

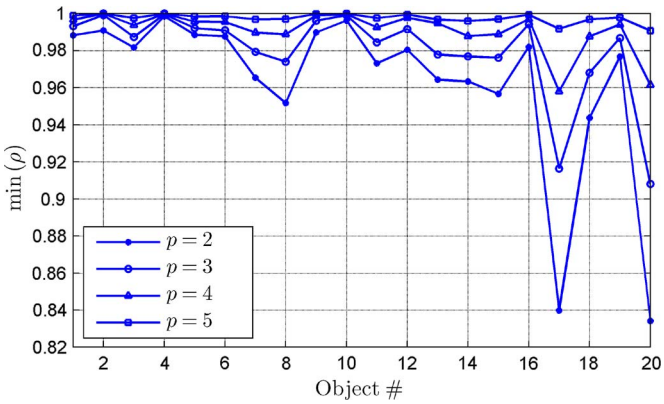

Fig. 5. Minimum amount of energy recovered by 9, 16, 25, and 36 harmonic images for each object under all 90 test poses. With the exception of objects 17, 18, and 20, over 95% of the energy is recovered by the 9-D linear subspace for all 90 poses.

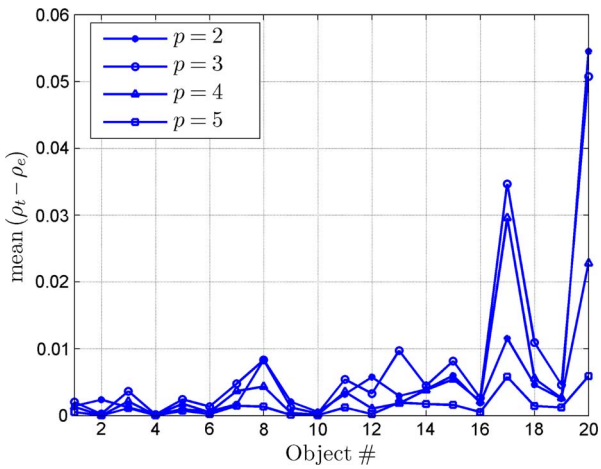

Fig. 6. Difference in energy recovered by the true 9-, 16-, 25-, and 36-dimensional subspace as computed using the SVD and the set of harmonic images as computed by the SHT for each object averaged over all 90 test poses.

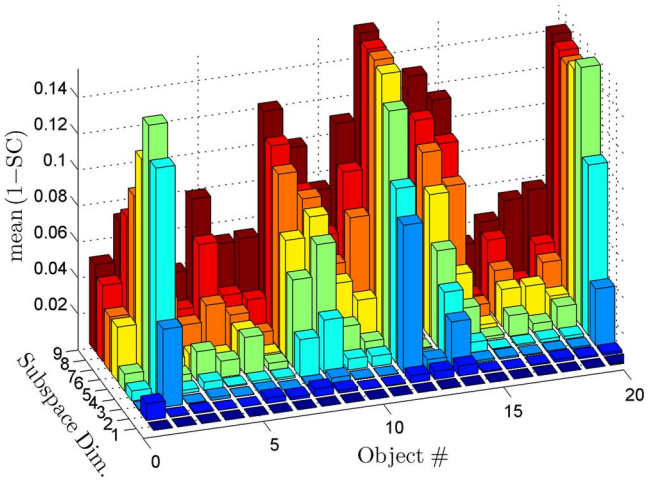

Fig. 7. SC averaged across all 90 test poses for each of the 20 objects in Fig. 4 using a 9-D subspace.

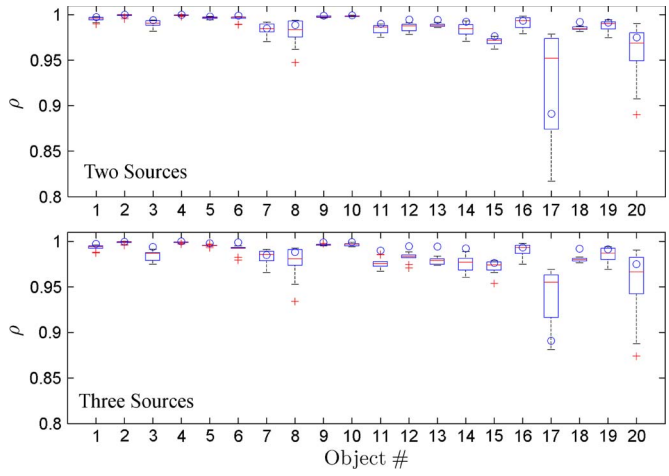

Fig. 8. Distribution of the energy recovery for (top) two and (bottom) TS. The box represents the interquartile region with the bar representing the median of the data when only an SS is present. The plus signs represent outliers in the data, and the circles represent the median energy recovery when only an SS is considered. The basis used for recovery was computed from the orthonormalized harmonic images using the first nine harmonics with a single illumination direction.

shows the average difference between the energy recovered by the true eigenimages and the harmonic images across all 90 poses. As can be seen from the figure, again, with the exception of objects 17, 18, and 20, there is less than 1% difference in energy recovered by the true eigenimages versus the harmonic images as computed using (14).

Finally, the SC defined in (4) was used to evaluate how well the 9-D subspace computed using the orthonormalized harmonic images spans the subspace constructed from the first nine eigenimages as computed by the SVD. Fig. 7 shows $1 - SC$ averaged across all 90 poses for each of the 20 objects in Fig. 4. As can be seen from the figure, the harmonic images span over 85% of the same space as the first nine true eigenimages.

It is important to note that we lose very little information about the image data by truncating the SHT. Instead, we have performed a change of basis that effectively condenses the information in each of the sets $\mathcal{I}_r$, allowing us to reduce the dimension of the space due to variations in illumination. At this point, we could simply perform an SVD on the remaining set of harmonic images and obtain very close approximations to the principal eigenimages of the original set $\hat{X}$. However, as we show in Section IV, the dimensionality of the data can be further reduced by performing another change of basis on the harmonic images when a change in pose is considered.

E. Multiple Illumination Sources

We now turn our attention to an investigation of variation due to multiple distant point light sources and fixed pose. For this evaluation, images of each of the objects in Fig. 4 were captured from 10 different poses and 48 different illumination conditions at each pose. The 48 different illumination conditions first consisted of an SS, then two illumination sources (DS), and finally three illumination sources (TS). The procedure for capturing images of the object under different illumination conditions using multiple illumination sources is as follows: First, TS were placed at random locations around the sphere. With the camera stationary, the first source was illuminated, while the other two were not, and an image of the object was captured. The second illumination source was then also illuminated, and a second image of the object was captured. Finally, the third illumination source was illuminated, and the final image was captured. This process is repeated 48 times at each of the ten poses. We then used the sampling procedure shown in Fig. 1 at each of the ten poses and computed the 9-D set of harmonic images using (14) for each pose. Each set of harmonic images was then orthonormalized. Our interest is in how effective this subspace is at recovering information from an image data set of the object from the same pose but significantly different illumination directions/conditions. To this end, we used the energy recovery ratio defined in (3) to compute how much of the total energy could be recovered for each of the three sets of image data (SS, DS, and TS). The results are shown in Fig. 8. The top plot in Fig. 8 shows the distributions of the energy recovered for each of the 20 objects across all ten poses when DS are present, and the bottom plot shows the same distribution when TS are present. One thing to notice from the figure is that, even though the median energy recovered for a single source is typically higher than that of multiple sources, for most of the objects, it is only slightly higher. Furthermore, with the exception of objects 17 and 20, the distributions remain fairly tight across all ten poses.

Some of the reasons for the drop in energy recovered when there are multiple illumination sources become clear when examining Fig. 9. The top row in Fig. 9 shows how well the low-dimensional subspace computed from an SS can reconstruct an image of object 1 in Fig. 4 under random illumination directions from TS. The first image in the top row shows the image to be reconstructed, while the remaining images in the top row show the absolute difference between the reconstructed image and the original image. As can be seen from the figure, the major sources of reconstruction error using the 9-D subspace are from specularities in local regions of the object. Notice that, for object 1 under the current viewing conditions (both viewing direction and illumination), very few local specular spikes exist. As a result, the reconstruction appears to be fairly accurate.

Fig. 9. Reconstruction error of a single image of (top) object 1 and (bottom) object 20 from Fig. 4 under three different illumination sources. The basis images used for reconstruction were computed using the SHT of the images generated assuming an SS and the HEALPix distribution. The first image in each pair is the original image. The remaining images show the absolute difference between the reconstructed image and the original image.

The second sequence of images is that of object 20 from Fig. 4. The bottom row provides the analogous information for object 20 from Fig. 4; however, as can be seen from the figure, for this object, the local specular regions are much larger. As a result, the low-dimensional subspace has difficulty recovering the specular spikes. Furthermore, there does not appear to be a significant improvement in reconstruction when more than a 5-D subspace is used, i.e., the reconstruction error changes very little when the subspace used for reconstruction increases beyond five dimensions. This shows that one of the major effects of illuminating object 20 from multiple directions is that there is an increased probability that several of the local specularities will be illuminated in a single image. The effect of cast and attached shadows, however, is reduced due to multiple illumination sources. This result can be deduced from the fact that the probability that the shadow from an SS will be illuminated by another illumination source is significantly increased. This suggests that recognition and pose estimation of largely diffuse objects with complex geometry may actually benefit when multiple illumination sources are present; however, this is not analyzed in this paper.

To evaluate quantitatively how well the 9-D subspace is capable of reconstructing each of the 480 (10 poses and 48 illumination directions at each pose) test images for each object in Fig. 4, an alternate metric is used. If each of the row-scanned images $\boldsymbol{f}$ in $\hat{X}$ is treated as a single point in an m-dimensional space, then the Euclidean norm between the true image $\boldsymbol{f}_t$ and the reconstructed image $\boldsymbol{f}_r$ provides a metric for determining how good the image reconstruction is using the low-dimensional subspace. Fig. 10 shows the distribution of $\|\boldsymbol{f}_t - \boldsymbol{f}_r\|_2$ for all 480 test images under all three illumination conditions (SS, DS, and TS). One item of interest in Fig. 10 is that the median reconstruction error is very similar regardless of how many illumination sources are present; however, the distribution of the reconstruction error is much larger when multiple sources are present. This, again, implies that there is a higher probability of illuminating a local specular region when more than one illumination source is present.

The aforementioned analysis verifies both qualitatively and quantitatively that the dimensionality of the data due to variations in illumination and constant pose can be reduced by using the discrete SHT. We now turn our attention to dimensionality due to a change in pose. The first step toward this goal is to discuss one of the fastest known algorithms for computing the eigenspace decomposition of a one-dimensionally correlated set of images under constant illumination conditions. This is the subject of the next section.

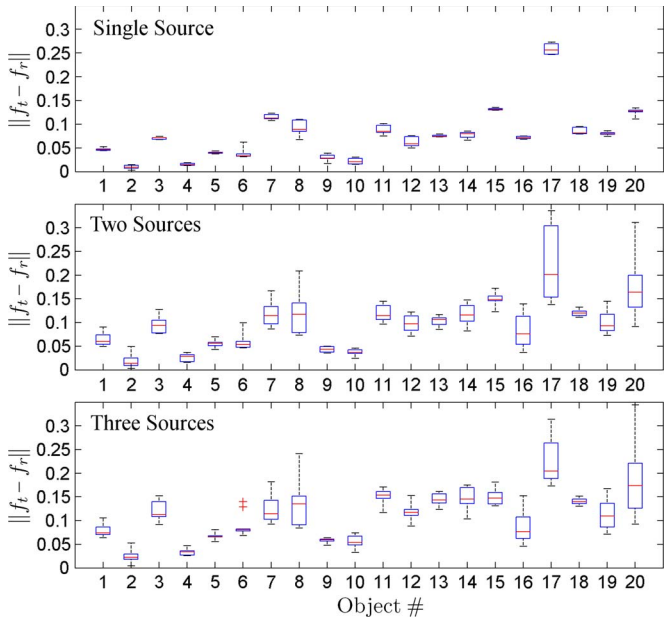

Fig. 10. Distribution of $\|\boldsymbol{f}_t - \boldsymbol{f}_r\|_2$ for all 480 test images of each object using (top) an SS, (center) DS, and (bottom) TS.

IV. Overview of Chang's Algorithm

A. Introduction

This section provides an overview of one of the fastest known algorithms for estimating the first k eigenimages of an image data set generated due to variations in pose under nonvarying illumination conditions, the details of which can be found in [17]. Consider capturing images of the object, as discussed in Section II, under constant illumination conditions. Capturing images in this manner results in a one-dimensionally correlated image data matrix X. Consider the special case where the image data matrix X is constructed such that the image $\boldsymbol{f}_{i+1}$ is obtained from $\boldsymbol{f}_i$ by a planar rotation of $2\pi/n$; then, the correlation matrix $X^T X$ is given by

$$X^T X = \begin{bmatrix} \boldsymbol{f}_1^T \boldsymbol{f}_1 & \boldsymbol{f}_1^T \boldsymbol{f}_2 & \cdots & \boldsymbol{f}_1^T \boldsymbol{f}_n \\ \boldsymbol{f}_2^T \boldsymbol{f}_1 & \boldsymbol{f}_2^T \boldsymbol{f}_2 & \cdots & \boldsymbol{f}_2^T \boldsymbol{f}_n \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{f}_n^T \boldsymbol{f}_1 & \boldsymbol{f}_n^T \boldsymbol{f}_2 & \cdots & \boldsymbol{f}_n^T \boldsymbol{f}_n \end{bmatrix}. \tag{15}$$

It can be shown [17] that $X^T X$ is a *circulant* matrix with circularly symmetric rows. For this special case, the eigendecomposition of $X^T X$ is given by the discrete Fourier transform (DFT), i.e.,

$$X^T X = H D H^T \tag{16}$$

where D is the $n \times n$ matrix given by

$$D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \tag{17}$$

and H is the real DFT matrix defined as

$$H = \sqrt{\frac{2}{n}} \begin{bmatrix} \frac{1}{\sqrt{2}} & c_0 & -s_0 & c_0 & -s_0 & \cdots \\ \frac{1}{\sqrt{2}} & c_1 & -s_1 & c_2 & -s_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots \\ \frac{1}{\sqrt{2}} & c_{n-1} & -s_{n-1} & c_{2(n-1)} & -s_{2(n-1)} & \cdots \end{bmatrix} \tag{18}$$

with $c_k = \cos(2\pi k/n)$ and $s_k = \sin(2\pi k/n)$.

Using the aforementioned development, an unordered SVD of X can be obtained for this special case by letting $V = H$, i.e., the right singular vectors of X in this case are given by pure sinusoids of frequencies that are multiples of $2\pi/n$ rad. The left singular vectors (eigenimages) can then be obtained by computing $XH = U\Sigma$, which can be computed efficiently using fast Fourier transform (FFT) techniques [17].

B. Chang's Eigendecomposition Algorithm

While the aforementioned analysis does not hold for arbitrary image data sets, it has been shown in [17] that the analytical expressions for planar rotations serve as a good approximation for the eigendecomposition of image data sets correlated in one dimension. In general, for image sequences correlated in one dimension, the following two properties can be observed [17], [27].

1) The right singular vectors of X are well approximated by sinusoids of frequencies that are multiples of $2\pi/n$ rad, and the power spectra of the right singular vectors consist of a narrow band around the corresponding dominant harmonics.
2) The dominant frequencies of the power spectra of the (ordered) singular vectors increase approximately linearly with their index.

These two properties indicate that the right singular vectors of an image data set correlated in one dimension are approximately spanned by the first few low-frequency harmonics. Therefore, by projecting the image data set X onto these first few low-frequency harmonics and computing the eigendecomposition in the spectral domain, the computational expense associated with computing the SVD can be significantly reduced.

Chang's algorithm makes use of the aforementioned two properties to estimate the subspace dimension k as well as the principal eigenimages $\tilde{U}_k$ of the image data matrix X to obtain a desired energy recovery ratio. It was shown in [17] that, if the power spectra of the first j right singular vectors of X are restricted to the band $[0, 2\pi j/n]$, then for $\rho(X^T, H_j) \geq E$, the quantity $\rho(X, \tilde{U}_k)$ will exceed E for some $k \leq j$, where H_j is the matrix containing the first j columns of H and E is a user-specified value. This inequality shows that the energy recovery ratio as computed using the first few low-frequency harmonics of H provides a lower bound on the energy recovery ratio as computed using the estimated eigenimages. Furthermore, this bound is shown to be extremely tight in most cases [17], with a tight upper bound given by the energy recovery ratio as computed by the "true" eigenimages. In other words, the first k estimated eigenimages $\tilde{U}_k$ of the matrix product XH_j are shown to be very good estimates of U_k.

V. Fast Eigenspace Decomposition Algorithm

A. Algorithm Development

Our objective is to estimate the first k principal eigenimages $\tilde{U}_k$ of $\hat{X}$ such that $\rho(\hat{X}, \tilde{U}_k) \geq E$, where E is the user-specified energy recovery ratio. To this end, we make use of two observations; the first is that reducing the dimensionality of the data due to a change in illumination can be efficiently done using the analysis provided in Section III, resulting in a set of harmonic images $\boldsymbol{f}^r_{p,q}$ at each pose. Note that each harmonic image corresponds to a spherical harmonic of degree p and order q at each of the a poses. Therefore, each set of harmonic images corresponding to a given value of p and q across all r can be concatenated to form the matrix

$$\hat{X}_{p,q} = \left[\boldsymbol{f}^0_{p,q}, \boldsymbol{f}^1_{p,q}, \ldots, \boldsymbol{f}^{a-1}_{p,q}\right]. \tag{19}$$

Furthermore, because the harmonic expansion is truncated, there will be nine such matrices in total, with each of size $\mathbb{R}^{m \times a}$.

Each of the nine matrices in (19) now only contains variations due to a change in pose for a given spherical harmonic coefficient and thus is correlated in a single dimension. Therefore, the second observation that can be made is that the dimensionality of the data due to a change in orientation can be reduced by applying the results observed by Chang *et al.* to each of the nine matrices $\hat{X}_{p,q}$. In other words, we can assume that the right singular vectors of $\hat{X}_{p,q}$ are well approximated by a few low-frequency Fourier harmonics, and the FFT can be used to determine $\rho(\hat{X}^T_{p,q}, H_{j_i}) \geq E_o$ for each of the nine p, q combinations, where E_o is a user-specified value for the energy recovery along the orientation (pose) dimension for each of the nine harmonics. Notice that j_i, $i = 1, 2, \ldots, 9$, corresponds to the number of Fourier harmonics required for the ith (p, q) combination to achieve the user-specified energy recovery ratio E_o. Let $Z^{p,q}_{j_i}$ denote the matrix $\hat{X}_{p,q} H_{j_i} \in \mathbb{R}^{m \times j_i}$ for each (p, q) combination, and construct the reduced order matrix

$$\bar{X} = \left[Z^{0,0}_{j_1}, Z^{1,-1}_{j_2}, Z^{1,0}_{j_3}, Z^{1,1}_{j_4}, Z^{2,-2}_{j_5}, \ldots, Z^{2,2}_{j_9}\right] \tag{20}$$

that effectively recombines the image data due to variations in both illumination and pose into a single matrix. The entire algorithm is summarized as follows:

Eigendecomposition Algorithm Summary

1) Use the SHT to compute the matrices $P_r = [\boldsymbol{f}^r_{0,0}, \boldsymbol{f}^r_{1,-1}, \boldsymbol{f}^r_{1,0}, \boldsymbol{f}^r_{1,1}, \boldsymbol{f}^r_{2,-2}, \ldots, \boldsymbol{f}^r_{2,2}]$ for all r.
2) Construct the matrices $\hat{X}_{p,q}$ by concatenating each of the harmonic images $\boldsymbol{f}^r_{p,q}$ in P_r for each r as shown in (19).
3) For each of the nine matrices $\hat{X}_{p,q}$, determine the smallest number j_i such that $\rho(\hat{X}^T_{p,q}, H_{j_i}) \geq E_o$, where E_o is the user-specified energy recovery ratio in the orientation

(pose) dimension and $i = 1, 2, \ldots, 9$ corresponds to the ith matrix $\hat{X}_{p,q}$.

4) Let $Z_{j_i}^{p,q}$ denote the matrix $\hat{X}_{p,q}H_{j_i}$, and construct the matrix $\bar{X}$ defined in (20). Note that the matrices $Z_{j_i}^{p,q}$ can be efficiently computed using the FFT.
5) Compute the SVD of $\bar{X} = \tilde{U}\tilde{S}\tilde{V}$.
6) Return $\rho(\hat{X}, \tilde{U}_k) \geq E$, where E is the user-specified energy recovery ratio.

B. Algorithm Discussion

The aforementioned algorithm takes advantage of the fact that most SVD algorithms require $\mathcal{O}(mn^2)$ operations to compute the full SVD of $\hat{X}$; this is computationally prohibitive when n is large (which is the case when we consider variations in both illumination and pose). Therefore, using the analysis provided in Section III, the dimensionality of the data due to a change in illumination conditions can be efficiently reduced by doing a change of basis using the discrete SHT defined in (14) to obtain the matrices $\hat{X}_{p,q}$. Each of the matrices $\hat{X}_{p,q}$ now only contains variation due to a change in pose for each p, q combination. Therefore, another change of basis is performed on each of the matrices $\hat{X}_{p,q}$ by using the FFT. Based on the observations made in [17], using the Fourier basis, only a few low-frequency Fourier harmonics are required to account for most of the information in $\hat{X}_{p,q}$. Therefore, the matrices $Z_{j_i}^{p,q}$ are constructed by truncating the FFT after a user-specified amount of energy is recovered.

At this point, it is important to note that, for most objects, we will have lost very little information by performing the two changes of basis. Therefore, we construct the matrix $\bar{X}$ and compute SVD$(\bar{X})$ to obtain the principal components $\tilde{U}_k$ which result in excellent estimates of U_k. The computational expense associated with the computation of $\bar{X}$ scales to $\mathcal{O}(mc^2)$, where c is the number of columns in $\bar{X}$ and has an upper bound of $9a = 810$, which is significantly less than the $n = 4320$ columns contained in $\hat{X}$.[3]

VI. Experimental Results

A. Test Data

The proposed algorithm detailed in Section V was tested on each of the objects in Fig. 4. Recall that each of the objects was sampled at a resolution of 128 $\times$ 128 from 90 different poses under 48 different light source locations at each pose. To accurately represent real objects using CAD models, the reflectance model used accounts for material properties, such as surface roughness and surface hardness, and incorporates a mix of diffuse and specular reflection using the Cook–Torrance reflectance model [57]. The mean image was then subtracted to construct the image data matrix $\hat{X}$. The parameters used in the algorithm were $E_o = 0.95$ for the orientation reduction and $E = 0.8$ for the total energy recovered. The true SVD of the image data matrix $\hat{X}$ was also computed using MATLAB for comparison. The quality measures outlined in Section II were used to evaluate the accuracy of the estimated subspace.

[3]An alternate approach to computing the SVD of $\bar{X}$ would be to apply the techniques described in [33] and [34] and subsample the columns of $\bar{X}$. This was not done in this paper but is left for a future research direction.

TABLE II
SUBSPACE DIMENSION k AND THE TIME REQUIRED TO ESTIMATE THE FIRST k LEFT SINGULAR VECTORS FOR EACH OBJECT TO MEET THE USER-SPECIFIED ENERGY RECOVERY RATIO $E = 0.8$. THE RESULTS ARE COMPARED AGAINST THE TRUE SVD USING MATLAB. THE TABLE ALSO SHOWS THE COLUMN DIMENSION OF $\bar{X}$

Object	Dim. k		Time [min.]		Col. Dim. of $\bar{X}$
	True	Proposed	True	Proposed	
1	17	17	31.274	0.111	378
2	9	9	25.528	0.070	162
3	13	13	32.342	0.116	379
4	15	15	29.955	0.137	474
5	10	10	31.564	0.076	229
6	14	15	30.954	0.181	576
7	16	17	27.874	0.099	239
8	31	31	31.551	0.152	446
9	19	19	30.842	0.162	502
10	14	14	31.597	0.154	448
11	22	22	31.736	0.117	356
12	20	20	31.825	0.188	561
13	8	8	21.117	0.114	254
14	12	12	30.830	0.107	270
15	23	23	30.776	0.153	472
16	27	27	21.272	0.109	249
17	196	217	22.857	0.183	552
18	20	20	15.501	0.093	173
19	25	25	21.489	0.152	439
20	33	46	21.433	0.099	209
Mean			27.616	0.129	368
Min.			15.501	0.070	162
Max.			32.342	0.188	576

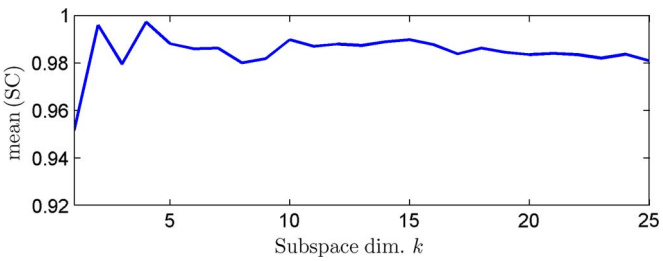

Fig. 11. SC as a function of the subspace dimension k averaged across all objects in Fig. 4.

B. Performance and Computational Savings

Table II shows the required subspace dimension k and the time required to estimate the first k left singular vectors $\tilde{U}_k$ for each object in Fig. 4 to meet the user-specified energy recovery ratio $E = 0.8$. 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 $\tilde{U}_k$ are very good estimates of U_k at significant computational savings. Table II also shows the column dimension of $\bar{X}$ in step 5 of the proposed algorithm. Note that, for the current test data, the number of columns in $\hat{X}$ is 4320, whereas for all 20 objects in Fig. 4, the number of columns in $\bar{X}$ never exceeds 576, thus resulting in significant computational savings.

Fig. 11 shows the SC as a function of the subspace dimension k averaged across all objects in Fig. 4. As can be seen from the figure, the estimated left singular vectors $\tilde{U}_k$ computed by

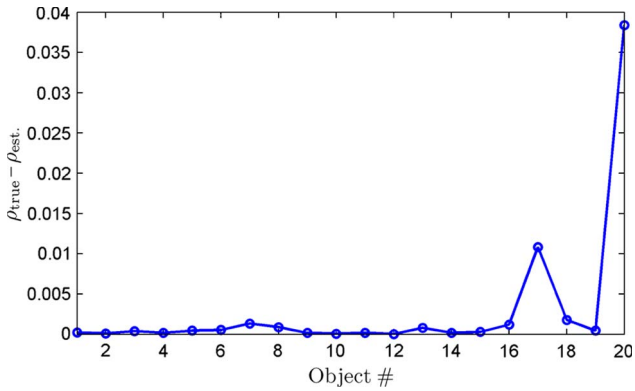

Fig. 12. Difference in energy recovered by the true SVD and the proposed algorithm for all objects in Fig. 4. The subspace dimension used for the calculation is listed in column 2 in Table II.

the proposed algorithm are very good approximations to the true left singular vectors as computed by the direct SVD in terms of spanning the same subspace as $\hat{X}$. Fig. 12 shows the difference in energy recovered by the true SVD and the proposed algorithm for all objects in Fig. 4. As can be seen from the figure, with the exception of objects 17 and 20, there is less than 0.25% difference in how much energy the proposed algorithm is capable of recovering compared to the true SVD. The subspace dimension used for the energy calculation is outlined in column 2 in Table II. As is apparent from Table II and Figs. 11 and 12, the estimates $\tilde{U}_k$ using the proposed algorithm are very good approximations to the left singular vectors U_k at significant computational savings.

VII. Discussion

This paper has presented an algorithm to efficiently estimate the eigenspace decomposition of an image data set generated due to a change in both illumination and pose. The algorithm is based on using the SHT to reduce the dimensionality of the data due to a change in illumination conditions, generating a set of harmonic images. It was shown that, after orthogonalization, the set of harmonic images are very good approximations to the true eigenimages as computed using the SVD. It was then shown that the harmonic images can be projected onto a few low-frequency Fourier harmonics for data reduction due to a change in pose. The algorithm was tested on several objects under a wide range of illumination conditions and different orientations. In addition to significant computational savings as compared to directly calculating the SVD, it has been shown that the estimated eigenimages are very close to the true eigenimages. This was true even for the case of multiple illumination sources. Finally, the proposed approach has the additional advantage of reduced memory storage requirements.

References

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

[2] L. Sirovich and M. Kirby, "Low-dimensional procedure for the characterization of human faces," *J. Opt. Soc. Amer.*, vol. 4, no. 3, pp. 519–524, Mar. 1987.

[3] M. Kirby and L. Sirovich, "Application of the Karhunen–Loeve procedure for the characterization of human faces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 1, pp. 103–108, Jan. 1990.

[4] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cogn. Neurosci.*, vol. 3, no. 1, pp. 71–86, Mar. 1991.

[5] M. H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 1, pp. 34–58, Jan. 2002.

[6] Y. Pang, D. Tao, Y. Yuan, and X. Li, "Binary two-dimensional PCA," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 38, no. 4, pp. 1176–1180, Aug. 2008.

[7] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Ann. Eugenics*, vol. 7, pp. 179–188, 1936.

[8] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.

[9] J. Yang and J. Y. Yang, "Why can LDA be performed in PCA transformed space?" *Pattern Recognit.*, vol. 36, no. 2, pp. 563–566, Feb. 2003.

[10] W. Zuo, D. Zhang, J. Yang, and K. Wang, "BDPCA plus LDA: A novel fast feature extraction technique for face recognition," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 36, no. 4, pp. 946–953, Aug. 2006.

[11] M. Belkin and P. Niyogi, "Laplacian eigenmaps and spectral techniques for embedding and clustering," in *Proc. Adv. Neural Inf. Process. Syst. 14*, 2001, vol. 14, pp. 585–591.

[12] X. He, S. Yan, Y. Hu, P. Niyogi, and H.-J. Zhang, "Face recognition using Laplacianfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 7, no. 3, pp. 328–340, Mar. 2005.

[13] D. Cai, X. He, J. Han, and H.-J. Zhang, "Orthogonal Laplacianfaces for face recognition," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3608–3614, Nov. 2006.

[14] 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. Acoust., Speech, Signal Process.*, vol. 37, no. 10, pp. 1550–1556, Oct. 1989.

[15] 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.

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

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

[18] R. C. Hoover, A. A. Maciejewski, and R. G. Roberts, "Aerial pose detection of 3-D objects using hemispherical harmonics," in *Proc. IEEE SSIAI*, Santa Fe, NM, Mar. 2008, pp. 157–160.

[19] 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 *Proc. IEEE Int. Conf. Robot. Autom.*, Pasadena, CA, May 2008, pp. 993–998.

[20] R. C. Hoover, A. A. Maciejewski, and R. G. Roberts, "Pose detection of 3-D objects using images sampled on $SO(3)$, spherical harmonics, and Wigner- D matrices," in *Proc. IEEE Conf. Autom. Sci. Eng.*, Washington, DC, Aug. 2008, pp. 47–52.

[21] R. C. Hoover, A. A. Maciejewski, and R. G. Roberts, "Eigendecomposition of images correlated on S^1, S^2, and $SO(3)$ using spectral theory," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2562–2571, Nov. 2009.

[22] R. Epstein, P. Hallinan, and A. Yuille, "5 $\pm$ 2 eigenimages suffice: An empirical investigation of low-dimensional lighting models," in *Proc. IEEE Workshop Phys.-Based Vis.*, Cambridge, MA, Jun. 1995, pp. 108–116.

[23] S. K. Nayar and H. Murase, "Dimensionality of illumination in appearance matching," in *Proc. IEEE Int. Conf. Robot. Autom.*, Minneapolis, MN, Apr. 1996, pp. 1326–1332.

[24] R. Ramamoorthi, "Analytic PCA construction for theoretical analysis of lighting variability in images of a Lambertian object," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 10, pp. 1322–1333, Oct. 2002.

[25] R. Basri and D. Jacobs, "Lambertian reflectance and linear subspaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 2, pp. 218–233, Feb. 2003.

[26] R. C. Hoover, A. A. Maciejewski, R. G. Roberts, and R. P. Hoppal, "An illustration of eigenspace decomposition for illumination invariant pose estimation," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, San Antonio, TX, Oct. 2009, pp. 3515–3520.

[27] 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 Process.*, vol. 15, no. 8, pp. 2376–2387, Aug. 2006.

[28] A. M. Martinez and A. C. Kak, "PCA versus LDA," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 2, pp. 228–233, Feb. 2001.
[29] K. Delac, M. Grgic, and S. Grgic, "Independent comparative study of PCA, ICA, and LDA on the FERET data set," *Int. J. Imag. Syst. Technol.*, vol. 15, no. 5, pp. 252–260, 2006.
[30] J. Yang, D. Zhang, and J.-Y. Yang, "Constructing PCA baseline algorithms to reevaluate ICA-based face recognition performance," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 37, no. 4, pp. 1015–1021, Aug. 2007.
[31] P. Hallinan, "A low-dimensional representation of human faces for arbitrary lighting conditions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Seattle, WA, Jun. 1994, pp. 995–999.
[32] P. N. Belhumeur and D. J. Kriegman, "What is the set of images of an object under all possible illumination conditions?" *Int. J. Comput. Vis.*, vol. 28, no. 3, pp. 245–260, Jul. 1998.
[33] C. K. I. Williams and M. Seeger, "Using the Nyström method to speed up kernel machines," in *Proc. NIPS*, 2000, pp. 682–688.
[34] A. Deshpande, L. Rademacher, S. Vempala, and G. Wang, "Matrix approximation and projective clustering via volume sampling," *Theory Comput.*, vol. 2, no. 1, pp. 225–247, 2006.
[35] P. N. Swarztrauber, "The vector harmonic transform method for solving partial differential equations in spherical geometry," *Mon. Weather Rev.*, vol. 121, pp. 3415–3437, May 1993.
[36] M. A. Wieczorek and F. J. Simons, "Localized spectral analysis on the sphere," *Geophys. J. Int.*, vol. 162, no. 3, pp. 655–675, Sep. 2005.
[37] F. J. Simons, F. A. Dahlen, and M. A. Wieczorek, "Spatiospectral localization on a sphere," *SIAM Rev.*, vol. 48, no. 3, pp. 504–536, 2006.
[38] D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii, *Quantum Theory of Angular Momentum*. Hackensack, NJ: World Scientific, 1988.
[39] G. B. Arfken, *Mathematical Methods for Physicists*. New York: Academic, 1970.
[40] G. S. Chirikjian and A. B. Kyatkin, *Engineering Applications of Noncommutative Harmonic Analysis: With Emphasis on Rotation and Motion Groups*. Boca Raton, FL: CRC Press, 2001.
[41] L. Qing, S. Shan, and W. Gao, "Eigen-harmonic faces: Face recognition under generic lighting," in *Proc. IEEE Int. Conf. Autom. Face Gesture Recog.*, Seoul, Korea, May 2004, pp. 296–301.
[42] S. Romdhani, J. Ho, T. Vetter, and D. J. Kriegman, "Face recognition using 3-D models: Pose and illumination," *Proc. IEEE*, vol. 294, no. 11, pp. 1977–1999, Nov. 2006.
[43] L. Zhang and D. Samaras, "Face recognition from a single training image under arbitrary unknown lighting using spherical harmonics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 3, pp. 351–363, Mar. 2006.
[44] B. Bustos, D. A. Keim, D. Saupe, T. Schreck, and D. V. Vranic, "Feature-based similarity search in 3-D object databases," *ACM Comput. Surv.*, vol. 37, no. 4, pp. 345–387, Dec. 2005.
[45] D. Saupe and D. V. Vranic, "3-D model retrieval with spherical harmonics and moments," in *Proc. DAGM-Symp. Pattern Recog.*, Munich, Germany, Sep. 2001, pp. 392–397.
[46] D. V. Vranic, "An improvement of rotation invariant 3D-shape based on functions on concentric spheres," in *Proc. IEEE Conf. Image Process.*, Barcelona, Spain, Sep. 2003, pp. 757–760.
[47] A. Makadia and K. Daniilidis, "Rotation recovery from spherical images without correspondences," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 7, pp. 1170–1175, Jul. 2006.
[48] N. Sneeuw, "Global spherical harmonic analysis by least-squares andnumerical quadrature methods in historical perspective," *Geophys. J. Int.*, vol. 118, no. 3, pp. 707–716, Jan. 1994.
[49] 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.
[50] P. N. Swarztrauber and W. F. Spotz, "Generalized discrete spherical harmonic transforms," *J. Comput. Phys.*, vol. 159, no. 2, pp. 213–230, Apr. 2000.
[51] P. N. Swarztrauber, "On computing the points and weights for Gauss–Legendre quadrature," *SIAM J. Sci. Comput.*, vol. 24, no. 3, pp. 945–954, 2002.
[52] J. R. Driscoll and D. M. Healy, Jr., "Computing Fourier transforms and convolutions on the 2-sphere," *Adv. Appl. Math.*, vol. 15, no. 2, pp. 202–250, Jun. 1994.
[53] D. M. Healy, Jr., D. Rockmore, P. Kostelec, and S. Moore, "FFTs for the 2-sphere-improvements and variations," *J. Fourier Anal. Appl.*, vol. 9, no. 4, pp. 341–385, Jul. 2003.
[54] 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," *Astrophys. J.*, vol. 622, no. 2, pp. 759–771, Apr. 2005.
[55] 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 *Proc. IEEE SSIAI*, Santa Fe, NM, Mar. 2008, pp. 41–44.
[56] K. Legaz, Kator Legaz: 3-D Model Database for Blender, 2007. [Online]. Available: http://www.katorlegaz.com/3d_models/index.php
[57] Blender: Open Source 3-D Modeling Application, 2009. [Online]. Available: http://www.blender.org/

Randy C. Hoover (S'06–M'09) received the B.S. degree in electrical engineering and the M.S. degree in measurement and control engineering from Idaho State University, Pocatello, in 2002 and 2004, respectively, and the Ph.D. degree in electrical engineering from Colorado State University, Fort Collins, in 2009.

From 2003 to 2004, he was a National Science Foundation Fellow. He is currently an Assistant Professor with the Electrical and Computer Engineering Department, South Dakota School of Mines and Technology, Rapid City. His research interests are in the areas of computer vision, robotics, and control theory.

Anthony A. Maciejewski (F'05) received the B.S.E.E., M.S., and Ph.D. degrees from The Ohio State University, Columbus, in 1982, 1984, and 1987, respectively.

From 1988 to 2001, he was a Professor of electrical and computer engineering with Purdue University, West Lafayette, IN. He is currently the Head of the Department of Electrical and Computer Engineering, Colorado State University, Fort Collins.

Rodney G. Roberts (SM'02) received the B.S. degree in electrical engineering and the B.S. degree in mathematics from the Rose–Hulman Institute of Technology, Terre Haute, IN, in 1987 and the M.S.E.E. and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, IN, in 1988 and 1992, respectively.

From 1992 to 1994, he was a National Research Council Fellow with Wright-Patterson Air Force Base, Dayton, OH. Since 1994, he has been with the College of Engineering, Florida Agricultural and Mechanical University–The Florida State University, Tallahassee, where he is currently a Professor of electrical and computer engineering. His research interests are in the areas of robotics and image processing.