

Signature Random Fields for Accommodating Illumination Variability

Matthew Cooper*

Keith Doolittle

Michael I. Miller

FX Palo Alto Laboratory
3400 Hillview Ave. Bldg. 4
Palo Alto, CA 94304

Center for Imaging Science
Johns Hopkins University
Baltimore, MD 21218

Center for Imaging Science
Johns Hopkins University
Baltimore, MD 21218

Abstract

In this paper, we document an extension to traditional pattern-theoretic object templates to jointly accommodate variations in object pose and in the radiant appearance of the object surface. We first review classical object templates accommodating pose variation. We then develop an efficient subspace representation for the object radiance indexed on the surface of the three dimensional object template. We integrate the low-dimensional representation for the object radiance, or signature, into the pattern-theoretic template, and present the results of orientation estimation experiments. The experiments demonstrate both estimation performance fluctuations under varying illumination conditions and performance degradations associated with unknown scene illumination. We also present a Bayesian approach for estimation accommodating illumination variability.

1 Introduction

In this paper, we consider the pattern-theoretic approach to Bayesian object recognition broadly outlined in [1]. Our contribution is the extension of the rigid-body object templates that accommodate pose variability to additionally accommodate variability in the radiant appearance of the object surface. This includes the variability induced by changes in scene illumination, a notorious obstacle to object recognition using video imaging systems.

Our work builds on the application of Karhunen-Loève (KL) methods in computer vision, e.g. [2, 3], automatic target recognition [4], and neuroanatomy, [5]. We assume a Lambertian surface, and we model the diffuse radiance of the object surface, henceforth denoted the object signature, as a random field. This random field is indexed on a discrete set \mathcal{L} of spatial locations on the object surface, geometrically specified by the three-dimensional object template (e.g. CAD model of object geometry). As detailed in [3, 6], the space of images of an object under

varying illumination is accurately approximated by low dimensional linear subspaces. We extend these results to three dimensional rigid-body object templates. For simplicity, we model the signature as a scalar¹ random field with dimensionality $L = |\mathcal{L}|$. For detailed geometric object models, L is on the order of 1000. The KL expansion of the signature random field is an orthonormal basis of object signatures, $\{\Phi_i(l) : i = 1, \dots, L, l \in \mathcal{L}\}$. We represent the object signature efficiently via the C dimensional vector of expansion coefficients $C \ll L$ on to the dominant KL basis signatures; throughout, $C = 10$.

2 Extending Pattern Theoretic Representations

First we review traditional pattern-theoretic rigid-body templates for accommodating the variability of position and orientation of objects. We then expand these representations to accommodate changes in object appearance associated with illumination variability.

2.1 Accommodating Pose Variability

In the context of rigid body object recognition, templates are constructed corresponding to CAD representations of the surface manifolds of the rigid objects. Denote such an ideal template as $I \equiv \{I(l), l \in \mathcal{L}\}$, \mathcal{L} the space indexing location on the object. Geometric variation is introduced via the rigid motions of translation and rotation. The translation group accommodates variability in the position of objects, while the rotation group accommodates variability in object orientation. Figure 1 shows a geometric object template for a coffee cup at two different orientations.

For ground-based scenes, we use the axis-fixed rotation group, identified with $\text{SO}(2)$ and parameterized by rotation angle, and translations in the plane, \mathbb{R}^2 ; transformations

¹Equivalently, we assume that brightness corresponds to a single wavelength; extensions to vector-valued (bandpass filter-based) colorspace are straightforward.

*Corresponding author: cooper@pal.xerox.com .

are of the form, $s = (O, a)$. Composing these variations, the transformation group for ground-based objects is the two-dimensional Special Euclidean group $\mathbf{SE}(2)$. Elements of $\mathbf{SE}(2)$ are composed according to $(O, a) \circ (O', a') = (OO', a + Oa')$. The deformable template over which inference occurs is the orbit under $\mathcal{S} = \mathbf{SE}(2)$:

$$\mathcal{I} \equiv \{I(O, a) : (O, a) \in \mathcal{S}\}. \quad (1)$$

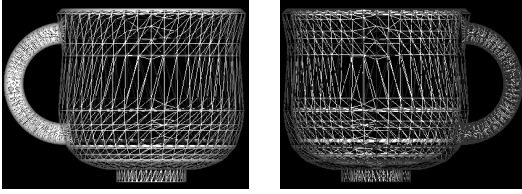


Figure 1: Object templates for a coffee cup visualized at two different orientations. The geometric object template consists of 3749 vertices.

2.2 Efficient Representation of Object Signature

An additional source of variability in imagery of objects whose pose parameters are known is variation in the radiant intensity of the object surface. We denote this class of variability in object appearance as signature variability. Throughout, we assume a Lambertian surface [7]. In the context of video imaging systems, a complete description of the signature of an object specifies the diffuse radiance as a function of position on the object surface. This surface can be decomposed into polygonal facets constituting a CAD model representation of the surface. Define the signature, $T \equiv \{T(l) : l \in \mathcal{L}\}$ as the radiant intensity indexed over the surface lattice \mathcal{L} given by the template. Incorporating signature information, the template becomes the set of all objects generated under the group action with superimposed scalar signature T . The orbit under the group action becomes

$$\mathcal{I} \equiv \{I(O, a, T) : (O, a, T) \in \mathbf{SE}(2) \times \mathbb{R}^L\}. \quad (2)$$

Representing the object signature increases the dimensionality of the parameterization from 3 ($\mathbf{SE}(2)$) to $L + 3$. We want to minimize this increase. Define the signature random field T to be a scalar-valued Gaussian random field representing the object radiant intensity with mean μ_T and covariance \mathbf{K}_T . Our approach is to statistically characterize signature variation via empirical covariances. KL expansion of the covariance operator of the random field T gives a complete orthonormal set of eigenvectors $\{\Phi_i\}$ defined on

the object surface. Expand T :

$$T(l) = \mu_T(l) + \sum_{i=1}^L \alpha_i \Phi_i(l). \quad (3)$$

The eigenvectors satisfy

$$\lambda_i \Phi_i(l) = \int_{\mathcal{L}} \mathbf{K}_T(l, l') \Phi_i(l') d\nu(l'), \quad (4)$$

where $d\nu(\cdot)$ denotes the measure on the surface manifold \mathcal{L} . In general L is large, so we choose a C term subspace $C \ll L$ to represent T optimally in a minimum-mean-squared-error sense [5].

We have generated databases of object signatures by varying the illumination direction. We construct an N element database of object signatures, using graphics software to compute the radiance at the lattice points on the object template's surface. We uniformly sample illumination directions by calculating points on the surface of a triangulated unit sphere. We then compute the KL expansion of the signature random field directly from the empirical covariance. The sample covariance, $\hat{\mathbf{K}}_T$, has elements

$$\hat{\mathbf{K}}_T(l, l') = \frac{1}{N} \sum_{k=1}^N (T_k(l) - \bar{T}(l))(T_k(l') - \bar{T}(l')) , \quad (5)$$

where the sample mean is

$$\bar{T}(l) = \frac{1}{N} \sum_{k=1}^N T_k(l) . \quad (6)$$

The sample covariance is expanded into an eigenfunction/eigenvalue representation numerically via Singular Value Decomposition (SVD) as detailed in [4]. Performing the KL expansion of the empirical covariance matrix is equivalent to the statistical technique of principal components analysis [8]. Hence the Φ_i are the signature principal components.

2.3 Results

Figure 2 shows visualizations of the first three basis functions from the KL expansion of the signature random field for a sphere. Intuitively, the lighting directions exhibited by the three signatures for the sphere roughly correspond to a basis for \mathbb{R}^3 . Figure 3 shows two views each of Φ_1, Φ_2 , and Φ_3 computed for a teapot in the top, middle, and bottom rows, respectively. Figure 4 shows the corresponding power spectra of the KL expansions, indicating that most of the variability of the two signature random fields is concentrated in the first several basis functions. In both cases, the first three basis functions dominate the expansions, as

predicted in [6]. Substituting the C -dimensional coefficient representation for the object signature into (2), we have

$$\mathcal{I} \equiv \underbrace{\{I(O, a, \bar{\alpha})\}}_{I(s)} : (O, a, \bar{\alpha}) \in \underbrace{\mathbf{SE}(2) \times \mathbb{R}^C}_s. \quad (7)$$

We shall assume the identification $I(s) \leftrightarrow s = (O, a, T)$, so that an element of the Special Euclidean group and a signature determine the object configuration completely.

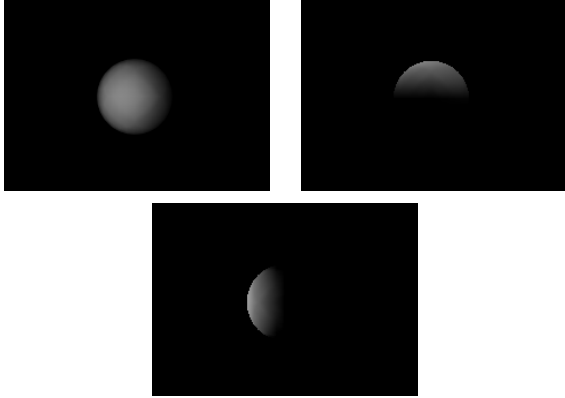


Figure 2: Renderings of the first three signature principal components, $\Phi_1(l)$, $\Phi_2(l)$ (top row), $\Phi_3(l)$, (bottom row) for a sphere.

3 Estimation Experiments

We have completed performance analysis for the problem of object orientation estimation, quantifying performance variations and degradations due to illumination variability.

3.1 MMSE Orientation Estimation

A key advantage of the Bayesian approach to object recognition is the ability to quantitatively assess performance. We focus on the problem of object orientation estimation and present results for the teapot of Figure 3. We consider only rotation about the vertical axis, and parameterize orientation using a finite set $\mathcal{O} \subset \mathbf{SO}(2)$, the two by two orthogonal matrix group. Our estimation algorithm is the minimum-mean-squared-error (MMSE) estimator with respect to the Hilbert-Schmidt matrix norm [9]. The estimator is a functional of the Bayesian posterior for the parameters describing the appearance of the object given the video image of the scene. In particular, given an image observation i^D ,

$$\begin{aligned} O_{HS}(i^D) &= \underset{o \in \mathbf{SO}(2)}{\text{ArgMin}} \int_{\mathbf{SO}(2)} \|o - o'\|^2 \pi(o' | i^D) d\mu(o') \\ &= \frac{A}{|A|}, \end{aligned} \quad (8)$$

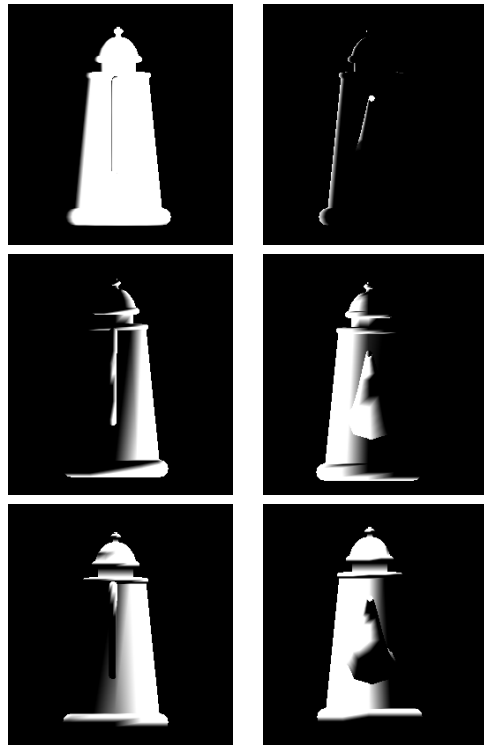


Figure 3: Renderings of the first three signature principal components, $\Phi_1(l)$ (top), $\Phi_2(l)$ (middle), $\Phi_3(l)$ (bottom), for a teapot. Two renderings are used to view the signature principal components for the teapot over its surface.

where $|\cdot|$ denotes the matrix determinant, and A is

$$A = \int_{\mathbf{SO}(2)} o \pi(o | i^D) d\mu(o). \quad (9)$$

We evaluate the estimator and the associated minimum-mean-squared estimation error numerically using \mathcal{O}

$$\hat{A} = \sum_{o_m \in \mathcal{O}} o_m \pi(o_m | i^D), \quad (10)$$

and numerically approximate (8) as

$$\hat{O}_{HS} = \frac{\hat{A}}{|\hat{A}|}. \quad (11)$$

3.2 Bayes Posterior

Bayesian recognition is founded on the posterior density for the parameters describing the object's appearance in the three dimensional scene given image data. Given an image i^D of the object, the joint posterior for the object pose, $a \in \mathbf{SE}(2)$ and signature $\bar{\alpha} \in \mathbb{R}^C$, is

$$\pi(a, \bar{\alpha} | i^D) = \frac{1}{Z(i^D)} L(i^D | a, \bar{\alpha}) \pi(a) \pi(\bar{\alpha}). \quad (12)$$

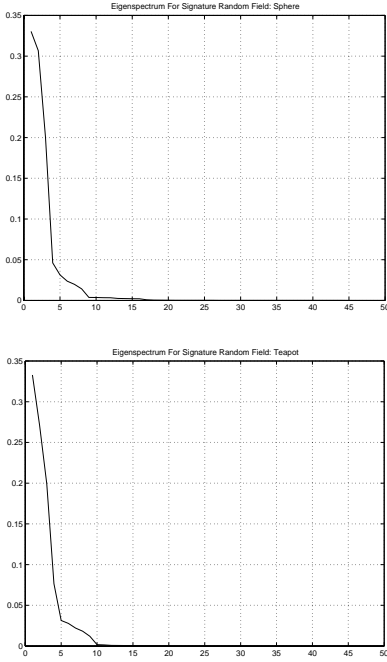


Figure 4: The normalized power spectrum ($\frac{\lambda_k}{\sum_{k=1}^L \lambda_k}$) of the KL expansions for the sphere (top) and the teapot (bottom). The x-axis indexes the eigenvectors of the empirical covariance (5).

$L(i^D|a, \bar{\alpha})$ is a pixel-based data likelihood for the imaging sensor. $\pi(a)$ is the prior on the pose parameters, assumed to be uniform in static recognition scenarios. We model the expansion coefficients as independent zero-mean Gaussian random variables with variances given by the eigenvalues of the signature random field’s empirical covariance. $Z(i^D)$ is the appropriate normalizer.

3.3 Experimental Results

To assess performance we calculate the minimum-mean-squared orientation estimation error. Denoting the space of image observations as \mathcal{I}^D , we plot the following error,

$$\begin{aligned}
 \mathcal{R} &= \int_{\mathbf{SO}(2) \times \mathcal{I}^D} \|O_{HS}(i^D) - o\|^2 d\mu(o, i^D) \quad (13) \\
 &= \int_{\mathbf{SO}(2)} \underbrace{\int_{\mathcal{I}^D} \|O_{HS}(i^D) - o\|^2 d\mu(i^D|o)}_{\mathcal{R}(o)} d\mu(o) \\
 &\simeq \hat{\mathcal{R}} = \frac{1}{M} \sum_{m=1}^M \underbrace{\frac{1}{D} \sum_{j=1}^D \|\hat{O}_{HS}(i_j^D) - o_m\|^2}_{\hat{\mathcal{R}}(o_m)}, \quad (14)
 \end{aligned}$$

where $M = |\mathcal{O}|$. Per (13), the estimation error is the expected error over the joint space of orientations and image observations. For numerical evaluation, we approximate the equivalent nested expectation. In the inner sum of (14), the inner integral is approximated by Monte Carlo sampling over the space of images of the object at orientation o_m . The outer expectation is then approximated using the discrete set \mathcal{O} . Notice that for each $o_m \in \mathcal{O}$, a set of images $\{i_j^D : j = 1, \dots, D\}$ are randomly sampled from \mathcal{I}^D . For each i_j^D , the MMSE estimator of (11) is calculated. Thus, the error \mathcal{R} depends directly on the posterior model employed in (10), as will be further demonstrated below.

In Figure 5, $\hat{\mathcal{R}}$ is evaluated over a range of image noise levels. The x-axis shows the standard deviation of the additive Gaussian image noise. The curves show estimator performance under each of the illumination conditions rendered below the plot. For these calculations, the illumination conditions are assumed known *a priori*; i.e. the employed posterior in (14) is $\pi(o_m|i^d, \alpha_o)$ where α_o denotes the observed object signature. Intuitively, varying illumination impacts the accuracy of the estimate of the teapot’s orientation.

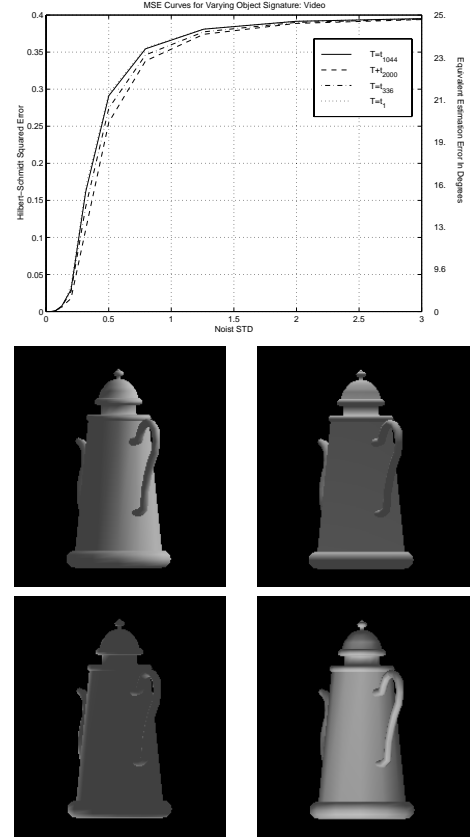


Figure 5: The plot shows MMSE variations due to illumination variability.

In Figure 6, we assess performance loss associated with unknown scene illumination. Here we focus on the error measure $\hat{\mathcal{R}}(30^\circ)$ as defined in the inner sum of (14). The solid curves in both panels indicate ideal performance conditioning on the observed illumination conditions, or equivalently, the observed object signature, α_o . For these curves, the posterior model employed is $\pi(o_m|i^d, \alpha_o)$. α_o is given by the signature rendered in the middle right panel of Figure 5 in the performance curves in the left panel of Figure 6. In the performance curves in the right panel, α_o is given by the signature rendered in the bottom right panel of Figure 5.

The remaining curves compare two approaches to estimation in unknown illumination conditions. The dashed curves (without x-marks) correspond to performance conditioning on fixed, mismatched assumptions of the object signature. In other words, these curves are based on posterior models assuming illumination conditions that are inconsistent with the observed scene: $\pi(o_m|i^D, \alpha_m)$ where $\alpha_m \neq \alpha_o$. The solid curves with x-marks indicate performance using a posterior model for the object signature that marginalizes the nuisance parameters of object signature. Specifically, for $o_m \in \mathcal{O}$,

$$\pi(o_m|i^D) = \int_{\mathbb{R}^C} \pi(o_m, \bar{\alpha}|i^D) d\bar{\alpha} \quad (15)$$

We evaluate the integral of (15) via Monte Carlo integration. For each $o_m \in \mathcal{O}$, a set of illumination conditions is randomly sampled $\{\bar{\alpha}_p : p = 1, \dots, P\}$. We assume as before that the coefficients of the object signature, $\bar{\alpha} \in \mathbb{R}^C$ are Gaussian distributed. (15) is evaluated for each o_m to construct a posterior mass function on \mathcal{O} . For estimation accommodating illumination variability, this posterior model is used to evaluate the MMSE estimator in (10). This variation on the MMSE Hilbert-Schmidt estimator is detailed in [13, 14].

4 Related Work

In contrast to traditional image-based techniques, e.g. [2, 10], we employ a complete three-dimensional model to describe the object's geometry and thus its surface. Our parametric representation for the object pose and surface radiance is separable. We apply dimension reduction only to the set of object signatures, indexed originally on the object template's entire surface, regardless of any viewing geometry by which portions of the object surface are occluded.

Belhumier, *et al.*, have employed subspace approximations to illumination cones to efficiently represent the space of images of faces under varying illumination in fixed pose [3]. They have employed KL methods to develop low-dimensional object representations accommodating illumination variability. In [11], they consider recognition under

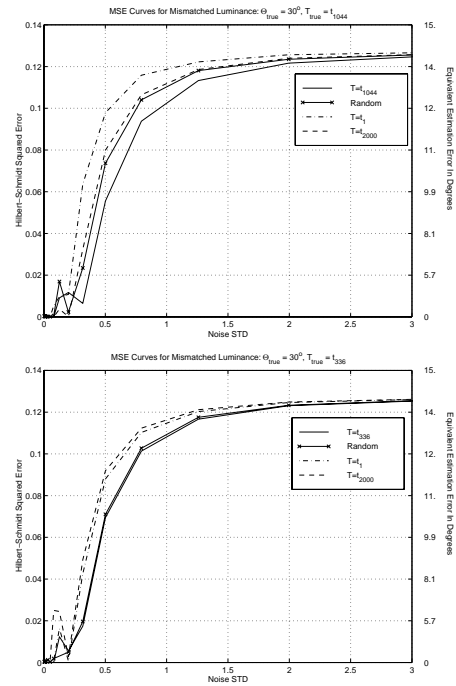


Figure 6: The two panels show pose estimation performance conditioning on correct, randomly sampled, and incorrect illumination information.

variable pose and illumination by defining a pose-indexed family of illumination cones for each face (i.e. object class). The approaches taken here and in [11] are similar in spirit as both employ low-dimensional parametric object representations from which a rich set of images of the object under varying poses and illumination conditions can be generated. However, the representation of (7) for the space of object signatures is founded on the three-dimensional object template; hence, our representation for object signature is independent of the object pose. Furthermore, we model the variability of the object radiance directly on the object surface, applying the KL expansion to a random field defined on the discretized object surface rather than on the image plane.

The method documented herein was originally developed in the context of forward-looking infrared (FLIR) imaging, where the object surface varies in appearance according to its emitted thermal radiation. Thus the space of object signatures corresponds to the set of thermal profiles of the object surface. Orientation estimation experiments were performed using both maximum *a posteriori* [12] and MMSE [13] estimation techniques accommodating the variability in the object thermal signature. An information-theoretic analysis quantifying information loss regarding the object pose that was due to thermal signature variation

was reported in [4].

5 Conclusion

In this paper, we extend traditional pattern-theoretic object templates to jointly accommodate variations in pose and surface radiance. We have introduced a low-dimensional representation for the set of radiant signatures of a three-dimensional object template by modeling the object signature as a random field indexed on the object surface. When the object's surface is Lambertian, the space of signatures will be accurately represented by as few as three signature principal components. The representation is based on the KL expansion of the signature random field, and is pose-independent.

This extension of template representations enables recognition of objects under varying illumination while retaining low dimensionality. The extended object template's dimensionality increases from 3 for pose variability ($SE(2)$), to $3 + C \simeq 13$ for both pose and signature variability. We have also documented a set of orientation estimation experiments demonstrating both performance variations and degradations due to varying illumination conditions. Finally, we have documented an extension to the the MMSE orientation estimator accommodating unknown scene illumination. This adaptation of MMSE estimation outperformed estimators conditioning on incorrect scene illumination, and approached ideal performance in our experiments.

Acknowledgment

This research was done while M. Cooper was with the Center for Imaging Science at The Johns Hopkins University.

References

- [1] M. Miller, U. Grenander, J. O'Sullivan, and D. Snyder. Automatic Target Recognition Organized Via Jump-Diffusion Algorithms. *IEEE Trans. on Image Processing*, **6**(1):157-174, 1997.
- [2] M. Turk and A. Pentland. Eigenfaces for Recognition. *J. Cognitive Neuroscience* **3**(1):71-96, 1991.
- [3] P. Belhumeur and D. Kriegman. What is the Set of Images of an Object under All Possible Illumination Conditions? *Intl. J. Computer Vision*, **28**(3):1-16, 1998.
- [4] M. Cooper and M. Miller. Information Measures for Object Recognition Accommodating Signature Variability. *IEEE Trans. on Information Theory*, **46**(5):1896-1907, 2000.
- [5] S. Joshi, M. Miller, and U. Grenander. On the Geometry and Shape of Brain Sub-Manifolds. *Intl. J. of Pattern Recognition & Artificial Intelligence*, **11**(8):1317-1343, 1997.
- [6] S. K. Nayar and H. Murase, Dimensionality of Illumination Manifolds in Eigenspace, CUCS-21-94, Technical Report, Dept. of Computer Science, Columbia University, 1994.
- [7] D. Forsyth and J. Ponce. *Computer Vision - A Modern Approach*, in preparation, Prentice-Hall, <http://www.cs.berkeley.edu/~daf/book.html>.
- [8] R. Muirhead. *Aspects of Multivariate Statistical Theory*. J. Wiley & Sons, 1982.
- [9] U. Grenander, M. Miller, and A. Srivastava. Hilbert-Schmidt Lower Bounds For Estimators on Matrix Lie Groups. *IEEE Trans. on Pattern Analysis & Machine Intelligence*, **20**(8):1-13, August, 1998.
- [10] B. Moghaddam and A. Pentland. Probabilistic Visual Learning for Object Representation. *IEEE Trans. on Pattern Analysis & Machine Intelligence*, **19**(7):696-710, 1997.
- [11] A. Georghiadis, P. Belhumeur, and D. Kriegman. From Few to Many: Illumination Cone Models for Face Recognition Under Variable Lighting and Pose. *IEEE Trans. on Pattern Analysis & Machine Intelligence*, **23**(6):643-660, 2001.
- [12] M. Cooper, A. Lanterman, S. Joshi, and M. Miller. Representing The Variation of Thermodynamic State Via Principal Components Analysis. In *Proceedings of the Third Workshop On Conventional Weapon ATR*. November 1996, pp. 481-490.
- [13] M. Cooper, U. Grenander, M. Miller, and A. Srivastava. Accommodating Geometric and Thermodynamic Variability for Forward-Looking Infrared Radar Systems. *Proc. SPIE*, **3070**:162-72, 1997.
- [14] M. Cooper. *Information Measures for Object Recognition Accommodating Signature Variability*. D.Sc. Thesis, Dept. of Electrical Engineering, Washington University, St. Louis, MO, 1999.