# A Symmetry Preserving Singular Value Decomposition *

Mili I. Shah and Danny C. Sorensen

October 2, 2005

**Abstract**

A reduced order representation of a large data set is often realized through a principle component analysis based upon a singular value decomposition (SVD) of the data. The left singular vectors of a truncated SVD provide the reduced basis. In several applications such as facial analysis and protein dynamics, structural symmetry is inherent in the data. Typically, reflective or rotational symmetry is expected to be present in these applications. In protein dynamics, determining this symmetry allows one to provide SVD major modes of motion that best describe the symmetric movements of the protein. In face detection, symmetry in the SVD allows for more efficient compression algorithms. Here, we present a method to compute the plane of reflective symmetry or the axis of rotational symmetry of a large set of points. Moreover, we develop a symmetry preserving singular value decomposition (SPSVD) that best approximates the given set while respecting the symmetry.

Interesting subproblems arise in the presence of noisy data or in situations where most, but not all, of the structure is symmetric. An important part of the determination of the axis of rotational symmetry or the plane of reflective symmetry is an iterative re-weighting scheme. This scheme is rapidly convergent in practice and seems to be very effective in ignoring outliers (points that do not respect the symmetry).

## 1 Introduction

Determining symmetry within a collection of spatially oriented points is a problem that occurs in many fields including molecular biology and face recognition analysis. In these applications, large amounts of data are generally collected, and it is desirable to approximate this data with a compressed representation. In some applications, the data is known to obey certain symmetry conditions, and it is profitable to preserve such symmetry in the compressed approximation.

Taking advantage of symmetry leads to better modeling of physical processes as well as more efficient storage and computational schemes.

For a given set of points $\mathcal{S} = \{\mathbf{x}_j : 1 \leq j \leq m\}$ in $n$-dimensional space, we form an $n \times m$ matrix $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m]$. The truncated singular value decomposition (SVD) provides a low rank approximation to $\mathbf{X}$ and therefore also to the data set $\mathcal{S}$. If $\mathbf{USV}^T = \mathbf{X}$ is an SVD of $\mathbf{X}$, then it is well known that the best rank $k$ approximation to $\mathbf{X}$ (in both the 2-norm and Frobenius norm) is given by $\mathbf{X}_k = \mathbf{U}_k \mathbf{S}_k \mathbf{V}_k^T$, where $\mathbf{U}_k, \mathbf{V}_k$ represent the leading $k$ columns of $\mathbf{U}, \mathbf{V}$ and $\mathbf{S}_k$ represents the leading $k \times k$ principal submatrix of $\mathbf{S}$. Here, we are concerned with preserving symmetry relations present in the set $\mathcal{S}$ and hence in the matrix $\mathbf{X}$. In particular, we desire the best low rank approximation $\mathbf{X}_k$ that also exhibits the same symmetries as the matrix $\mathbf{X}$. This is accomplished by providing a symmetry preserving singular value decomposition (SPSVD).

We concentrate on determining two types of symmetry: rotational and reflective. The computational schemes for calculating the best symmetric approximation of a given set involve two steps for each case. For reflective symmetry, the first step is to obtain the normal to an approximate plane of reflective symmetry, where the normal is defined to be the unit vector perpendicular to a hyperplane for which the given set can be split into two mirror image sets. For rotational symmetry, we first determine an approximate axis of rotational symmetry about which the given set can be rotated ($2\pi/k$ degrees in 3D) and returned to the same set. Then, in the second step, we find the best approximation to the given set that has the appropriate symmetries with respect to the approximate plane of symmetry or axis of rotation with the aid of the SPSVD.

For practical applications, we must consider noisy data sets. Thus, we need to construct a normal vector or axis of rotation that diminishes the effects of outliers. This is accomplished by creating an iterative re-weighting scheme that minimizes deviation from symmetry in a weighted Frobenius norm. With our weighted normal or axis of rotation, we build our SPSVD that preserves the respective symmetries as in the non-weighted scheme.

We also provide a means to compute just the dominant portion (leading $k$ terms) of the SPSVD that is well suited to large scale computation. This computation only requires matrix-vector products involving the point set represented as a matrix. The ARPACK software [8] can be used in this large scale case. The computation is no more expensive than constructing the leading terms of the SVD of the full set of points without the symmetry constraint. Computational examples involving the backbone of the HIV-1 protease molecule are presented here. These examples provide trajectories that result in matrices of dimension 9,000 by 10,000. The computations were performed on a multiprocessor cluster using the parallel P_ARPACK version of ARPACK.

There has been considerable research in the area of symmetry detection. Atallah [1] constructs an order $n \log n$ algorithm that determines the line of reflective symmetry of a perfectly symmetric planar object by reducing the system to a permutation problem. Optimizing a coefficient of symmetry is employed by Marola to determine an axis of symmetry for planar images [9]. Zabrodsky et al. [16] employ a continuous symmetry measure and apply it

to finding reflective and rotational symmetries in chemistry. Kazhdan extends this idea to 3D objects by creating a continuous 2D function that measures the invariance of an object with respect to reflective symmetry about each plane that goes through the object's center of mass [4].

Many papers use the following fundamental properties of symmetry, which can be found in [15, 10, 11], to determine reflective and rotational symmetry. In this literature, the term "principal axes" refer to the eigenvectors of the correlation matrix $\mathbf{X}\mathbf{X}^T$ of the set of points, i.e., the left singular vectors of $\mathbf{X}$. The observation is that:

- Any plane of symmetry of a body is perpendicular to a principal axis.
- Any axis of rotational symmetry of a body is a principal axis.

Minovic et al. start with this idea and build an octree representation to find symmetries of a 3D object [12]. Sun and Sherrah [14] begin by looking at the extended Gaussian image of an object, and then search along the principal axes for the strongest symmetry measure. O'Mara and Owens [13] also search for the principal axis with the largest symmetry measure. However, their symmetry measure is more refined, since it takes into effect intensity values. Colliot et al. [3] extend O'Mara and Owens' research by starting with the highest symmetry measure principal axis. Then, they optimize the axis of symmetry using the Nelder-Mead downhill simplex method. They apply this method to facial recognition and brain scan applications.

The idea of a symmetric approximation to a set of data points has come up in partial differential equations and in face detection. Aubry et al. prove that any truncated approximation to a dynamical system must maintain its respective symmetries. They derive a method of truncation, based on Proper Orthogonal Decomposition, that obeys the symmetries of the original infinite-dimensional system [2]. Smaoui and Armbruster present a way to symmetrize the eigenmodes of the Karhunen-Loeve basis in a computationally efficient matter [2]. Kirby and Sirovich [5, 6] present a symmetric approximation based on taking the average of the even and odd (correctly oriented) symmetric faces. We prove here that taking the average gives the best symmetric approximation (in the Frobenius norm) to the original data set, and we generalize this result to give the best symmetric approximation to a set that possesses $k$-fold rotational symmetry.

The folding method is employed by Zabrodsky et al. [17] to calculate the best symmetric approximation to a set. This method produces an approximation that is equivalent to ours. However, our proof indicates how to calculate an SPSVD that gives the best low rank symmetric approximation to a set efficiently for large scale matrices.

This paper is organized as follows. Section 2 defines perfect reflective and rotational symmetry. Finding an optimal hyperplane of reflective symmetry for noisy data is developed and analyzed in section 3; while choosing the axes of rotational symmetry for noisy data is discussed in section 4. Finally, section 5 develops an SPSVD that best approximates the given data set and provides an

algorithm for directly computing the best low rank symmetry preserving approximation in a way that is suitable for large scale computation. Computational results are presented in section 6.

Throughout the discussion, $\| \cdot \|$ shall denote the 2-norm and $\| \cdot \|_F$ shall denote the Frobenius norm.

# 2 Perfect Symmetry

In this section, we lay out the basic defining properties of reflective and rotational symmetry. We also give analytic specifications of the normal to a plane of reflection and the axis of rotational symmetry when the given data set possesses exact symmetry relations.

## 2.1 Reflective Symmetry

Recall that a hyperplane $\mathcal{H}$ is specified by a constant $\gamma$ and a vector $\mathbf{w}$ via $\mathcal{H} := \{\mathbf{x} : \gamma + \mathbf{w}^T \mathbf{x} = 0\}$. The vector $\mathbf{w}$ is called the normal to the plane. We say that a set of points $\mathcal{S} \subset \mathbb{R}^n$ is *reflectively symmetric with respect to the hyperplane* $\mathcal{H}$ if for every point $\mathbf{s} \in \mathcal{S}$, there exists a point $\hat{\mathbf{s}}$ such that $\hat{\mathbf{s}} = \mathbf{s} + \tau \mathbf{w}$ for some scalar $\tau$ with $\mathbf{s} + \frac{\tau}{2}\mathbf{w} \in \mathcal{H}$. It is easily shown that the center $\mathbf{c} \equiv \frac{1}{m} \sum_{\mathbf{s} \in \mathcal{S}} s$ of the point set lies in the plane of symmetry, where $m$ is the number of elements in $\mathcal{S}$. A simple rigid translation of the point set will allow us to assume that the center is at the origin $\mathbf{c} = 0$ and hence also that $\gamma = 0$. These assumptions will be made throughout this discussion. For simplicity, we shall also assume that no points of $\mathcal{S}$ lie in the plane of symmetry.

The following lemma is an immediate consequence of the fact that for each $\mathbf{s} \in \mathcal{S}$ there is a reflected point $\hat{\mathbf{s}} = \mathbf{s} + \tau \mathbf{w} \in \mathcal{S}$.

**Lemma 2.1** *A set $\mathcal{S}$ is reflectively symmetric with respect to a hyperplane $\mathcal{H}$ with unit normal $\mathbf{w}$ if and only if*

$$\mathcal{S} = (\mathbf{I} - 2\mathbf{w}\mathbf{w}^T)\mathcal{S}.$$

**Lemma 2.2** *If $\mathcal{S}$ is reflectively symmetric about $\mathcal{H}$, then the center $\mathbf{c} \in \mathcal{H}$.*

If $\mathcal{S}$ is reflectively symmetric about $\mathcal{H}$, we can arrange the points of $\mathcal{S}$ into two sets represented as two $n \times \frac{m}{2}$ dimensional matrices $\mathbf{X}_0$ and $\mathbf{X}_1$ such that

$$\mathbf{X}_0 = (\mathbf{I} - 2\mathbf{w}\mathbf{w}^T)\mathbf{X}_1.$$

Moreover, there is no loss of generality in assuming that $\mathbf{w}^T \mathbf{X}_0 > 0$ and that $\mathbf{w}^T \mathbf{X}_1 < 0$ (elementwise).

## 2.2 Rotational Symmetry

We say that a set of points $\mathcal{S} \subset \mathbb{R}^n \bigcap \mathbf{q}^\perp$ is *k-fold rotationally symmetric about an axis* $\mathbf{q} \in \mathbb{R}^n$ if there exist an $n \times n$ orthogonal matrix $\mathbf{R}(\mathbf{q})$ such that for

every point $\mathbf{s} \in \mathcal{S}$, there are $k - 1$ distinct points $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_{k-1} \in \mathcal{S}$ with $\mathbf{R}(\mathbf{q})^j \mathbf{s} = \mathbf{s}_j$ for $j = 1, 2, \ldots, k - 1$. We call $\mathbf{q}$ the rotational axis of symmetry and $\mathbf{R}(\mathbf{q})$ the rotation matrix.

**Lemma 2.3** *A set $\mathcal{S}$ is $k$-fold rotationally symmetric with respect to a rotational axis $\mathbf{q}$ if and only if for $j = 1, 2, \ldots, k - 1$*

$$\mathcal{S} = \mathbf{R}(\mathbf{q})^j \mathcal{S} = (\mathbf{I} - \mathbf{Q}\mathbf{G}\mathbf{Q}^T)^j \mathcal{S},$$

*where $\mathbf{Q} \in I\!\!R^{n \times (n-1)}$ with $[\mathbf{q}, \mathbf{Q}]$ forming an orthogonal matrix, and $\mathbf{I} - \mathbf{G} \in I\!\!R^{(n-1) \times (n-1)}$ is a rotation with $(\mathbf{I} - \mathbf{G})^k = \mathbf{I}$.*

Note, $(\mathbf{R}(\mathbf{q}))^k = (\mathbf{I} - \mathbf{Q}\mathbf{G}\mathbf{Q}^T)^k = \mathbf{I}$, and for $n = 3$, the matrix $\mathbf{I}_2 - \mathbf{G}$ is a $2 \times 2$ plane rotation through an angle of $\theta = 2\pi/k$ degrees.

If $\mathcal{S}$ is $k$-fold rotationally symmetric about $\mathbf{q}$, we can arrange the points of $\mathcal{S}$ into $k$ sets represented as matrices $\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_{k-1}$ such that

$$\mathbf{X}_j = (\mathbf{I} - \mathbf{Q}\mathbf{G}\mathbf{Q}^T)^j \mathbf{X}_0$$

for $j = 1, 2, \ldots, k - 1$. Again, we will assume that the center $\mathbf{c}$ of the data is at the origin. This can always be attained in general by a simple uniform translation of all the points of $\mathcal{S}$.

# 3 Optimal Value of Reflective w

Generally, in practice, the given set $\mathcal{S}$ is not exactly symmetric with respect to any particular plane. However, we may think of calculating a $\mathbf{w}$ that does the best possible job of specifying a plane that separates $\mathcal{S}$ into two sets $\mathbf{X}_0$ and $\mathbf{X}_1$ (again represented as matrices) that are "nearly" symmetric with respect to the plane.

It is possible to find an initial separation of $\mathcal{S}$ into $\mathbf{X}_0$ and $\mathbf{X}_1$ that are paired to be nearly symmetric with respect to a plane determined by a calculated $\mathbf{w}$. Methods for this are discussed in [10]. However, for this discussion, we shall assume that a partitioning of $\mathcal{S}$ into $\mathbf{X}_0$ and $\mathbf{X}_1$ is given such that the columns of the two matrices are correctly paired.

The specification of $\mathbf{w}$ may be expressed as an optimization problem

$$\min_{\|\mathbf{w}\|=1} \{\|\mathbf{X}_0 - \mathbf{W}\mathbf{X}_1\|_F : \mathbf{W} = \mathbf{I} - 2\mathbf{w}\mathbf{w}^T\}. \tag{1}$$

**Lemma 3.1** *The solution $\mathbf{w}$ to the minimization problem (1) is the unit eigenvector corresponding to the smallest eigenvalue of the symmetric indefinite matrix*

$$\mathbf{M}_0 = \mathbf{X}_0 \mathbf{X}_1^T + \mathbf{X}_1 \mathbf{X}_0^T.$$

**Proof.**

$$
\begin{aligned}
\|\mathbf{X}_0 - \mathbf{W}\mathbf{X}_1\|_F^2 &= tr\{(\mathbf{X}_0 - \mathbf{X}_1)(\mathbf{X}_0 - \mathbf{X}_1)^T\} + 4tr\{\mathbf{w}\mathbf{w}^T\mathbf{X}_1(\mathbf{X}_0 - \mathbf{X}_1)^T\} \\
&\quad + 4tr\{(\mathbf{w}\mathbf{w}^T\mathbf{X}_1)(\mathbf{w}\mathbf{w}^T\mathbf{X}_1)^T\} \\
&= tr\{(\mathbf{X}_0 - \mathbf{X}_1)(\mathbf{X}_0 - \mathbf{X}_1)^T\} + 4\mathbf{w}^T\mathbf{X}_1(\mathbf{X}_0 - \mathbf{X}_1)^T\mathbf{w} \\
&\quad + 4\mathbf{w}^T(\mathbf{X}_1\mathbf{X}_1^T)\mathbf{w} \\
&= tr\{(\mathbf{X}_0 - \mathbf{X}_1)(\mathbf{X}_0 - \mathbf{X}_1)^T\} + 4\mathbf{w}^T(\mathbf{X}_1\mathbf{X}_0^T)\mathbf{w} \\
&= tr\{(\mathbf{X}_0 - \mathbf{X}_1)(\mathbf{X}_0 - \mathbf{X}_1)^T\} + 2\mathbf{w}^T(\mathbf{X}_1\mathbf{X}_0^T + \mathbf{X}_0\mathbf{X}_1^T)\mathbf{w}
\end{aligned}
$$

where we have used $\mathbf{w}^T\mathbf{w} = 1$ and that $tr\{\mathbf{A}\mathbf{B}\} = tr\{\mathbf{B}\mathbf{A}\}$.

Clearly this quantity is minimized when $2\mathbf{w}^T(\mathbf{X}_1\mathbf{X}_0^T + \mathbf{X}_0\mathbf{X}_1^T)\mathbf{w}$ is minimized, and this occurs precisely when $\mathbf{w}$ is the (unit norm) eigenvector corresponding to the smallest eigenvalue of the symmetric matrix

$$
\mathbf{M} = \mathbf{X}_1\mathbf{X}_0^T + \mathbf{X}_0\mathbf{X}_1^T.
$$

∎

A weighting can be introduced into the minimization problem (1) which gives a way to de-emphasize anomalies in the supposed symmetry relation. In this case, we must solve

$$
\min_{\|\mathbf{w}\|=1} \{\|(\mathbf{X}_0 - \mathbf{W}\mathbf{X}_1)\mathbf{D}\|_F : \mathbf{W} = \mathbf{I} - 2\mathbf{w}\mathbf{w}^T\}, \tag{2}
$$

where $\mathbf{D}$ is a diagonal weighting matrix.

**Lemma 3.2** *The solution $\mathbf{w}$ of to the minimization problem (2) is the unit eigenvector corresponding to the smallest eigenvalue of the symmetric indefinite matrix*

$$
\mathbf{M}_D = \mathbf{X}_0\mathbf{D}^2\mathbf{X}_1^T + \mathbf{X}_1\mathbf{D}^2\mathbf{X}_0^T. \tag{3}
$$

**Proof.** Similar to Lemma 3.1. ∎

We have devised an iterative re-weighting scheme to construct a $\mathbf{D}$ that diminishes the influence of outliers in the SPSVD. Given a guess $\mathbf{z}$ to the normal vector $\mathbf{w}$, the basic idea is to weight the $i$-th column of $\mathbf{X}_0 - \mathbf{W}\mathbf{X}_1$, i.e. $\mathbf{x}_i^{(0)} - (\mathbf{I} - 2\mathbf{w}\mathbf{w}^T)\mathbf{x}_i^{(1)}$, by the reciprocal of the norm of $\mathbf{x}_i^{(0)} - (\mathbf{I} - 2\mathbf{z}\mathbf{z}^T)\mathbf{x}_i^{(1)}$, where $\mathbf{z}$ is a unit vector. The motivation is to penalize (give the smallest weight) to the pairs $\mathbf{x}_j^0$, $\mathbf{x}_j^1$ that are farthest from being symmetric with respect to $\mathbf{z}$.

Let us define

$$
F(\mathbf{z}, \mathbf{w}) = \sum_{i=1}^{m} \left(\frac{f_j(\mathbf{w})}{f_j(\mathbf{z})}\right)^2 = \|(\mathbf{X}_0 - \mathbf{W}\mathbf{X}_1)\mathbf{D}(\mathbf{z})\|_F^2,
$$

where $f_j(\mathbf{z}) = \|\mathbf{x}_j^{(0)} - (\mathbf{I} - 2\mathbf{z}\mathbf{z}^T)\mathbf{x}_j^{(1)}\|$ and $\mathbf{D}(\mathbf{z}) = \text{diag}\{f_j(\mathbf{z})^{-1}\}$. To find the optimal normal with respect to this weighting, we choose $\mathbf{w}$ as the point

that minimizes $\|(\mathbf{X}_0 - \mathbf{W}\mathbf{X}_1)\mathbf{D}(\mathbf{z})\|_F$, as described in Lemma 3.2. Then, the approximate $\mathbf{w}$ associated with this weighting solves

$$\min_{\|\mathbf{w}\|=1} F(\mathbf{z}, \mathbf{w}). \tag{4}$$

This suggests an iterative re-weighting scheme that will adjust the vector $\mathbf{z}$ to optimally diminish the effect of outliers; begin with an initial guess $\mathbf{z}_0$ and iterate

$$\mathbf{z}_{k+1} = \arg \min_{\|\mathbf{w}\|=1} F(\mathbf{z}_k, \mathbf{w}), \quad k = 0, 1, 2, \ldots \tag{5}$$

until $\|\mathbf{z}_{k+1} - \mathbf{z}_k\|$ is sufficiently small. Upon convergence, this fixed point iteration will solve the following max-min problem

$$\max_{\|\mathbf{z}\|=1} \left\{ \min_{\|\mathbf{v}\|=1} F(\mathbf{z}, \mathbf{v}) \right\} \tag{6}$$

as the following lemma indicates.

**Lemma 3.3** *If $\mathbf{v} = \mathbf{z}$ is a fixed point of the minimization problem (4), then $\mathbf{z}$ is a solution to the max-min problem (6), and $F(\mathbf{z}, \mathbf{v}) = m$.*

**Proof.** Given $\mathbf{z}$, $\|\mathbf{z}\| = 1$,

$$\min_{\|\mathbf{v}\|=1} \sum_{j=1}^{m} \left( \frac{f_j(\mathbf{v})}{f_j(\mathbf{z})} \right)^2 \le \sum_{j=1}^{m} \left( \frac{f_j(\mathbf{z})}{f_j(\mathbf{z})} \right)^2 = m.$$

Hence,

$$\max_{\|\mathbf{z}\|=1} \left\{ \min_{\|\mathbf{v}\|=1} F(\mathbf{z}, \mathbf{v}) \right\} \le m.$$

If $\mathbf{v} = \mathbf{z}$, then $F(\mathbf{z}, \mathbf{v}) = F(\mathbf{z}, \mathbf{z}) = m$. Therefore, any fixed point of the minimization problem (4) is a solution to the max-min problem (6). ∎

We've shown in the above lemma that a fixed point of iteration (5) solves the max-min problem (6). Now, we will show the existence of a fixed point to the iteration (5) in Lemma 3.4.

**Lemma 3.4** *There is a point $\mathbf{z}_*$ of unit norm such that*

$$\mathbf{z}_* = \arg \min_{\|\mathbf{w}\|=1} F(\mathbf{z}_*, \mathbf{w}).$$

**Proof.** Let $\mathbf{M}_i = \|\mathbf{x}_i^{(0)} - \mathbf{x}_i^{(1)}\|^2 \mathbf{I} + 2(\mathbf{x}_i^{(0)}\mathbf{x}_i^{(1)^T} + \mathbf{x}_i^{(1)}\mathbf{x}_i^{(0)^T})$. For a given $\mathbf{z}$, any $\mathbf{w}$ that solves

$$\min_{\|\mathbf{w}\|=1} F(\mathbf{z}, \mathbf{w}) = \min_{\|\mathbf{w}\|=1} \sum_{i=1}^{m} \frac{\mathbf{w}^T \mathbf{M}_i \mathbf{w}}{\mathbf{z}^T \mathbf{M}_i \mathbf{z}}$$

7

will also solve

$$\min_{\|\mathbf{w}\|=1} \Phi(\mathbf{z})F(\mathbf{z},\mathbf{w}) = \min_{\|\mathbf{w}\|=1} \sum_{i=1}^{m} \phi_i(\mathbf{z})\mathbf{w}^T \mathbf{M}_i \mathbf{w},$$

where $\Phi(\mathbf{z}) = \prod_{i=1}^{m} \mathbf{z}^T \mathbf{M}_i \mathbf{z}$, and $\phi_i(\mathbf{z}) = \prod_{\substack{j=1 \\ j \neq i}}^{m} \mathbf{z}^T \mathbf{M}_j \mathbf{z}$. The function $\Phi(\mathbf{z})$ restricted to the unit $n$-sphere is a continuous function on a compact set. Therefore, $\min_{\mathbf{z}} \Phi(\mathbf{z}) = \Phi(\mathbf{z}_*)$ is attained at some point $\mathbf{z} = \mathbf{z}_*$ on the unit sphere.

From Lagrange Theory, we see that

$$\nabla\Phi(\mathbf{z}_*) = 2 \sum_{i=1}^{m} \phi_i(\mathbf{z}_*)\mathbf{M}_i \mathbf{z}_* = 2\mathbf{z}_* \lambda,$$

or, if we denote $\mathbf{M}(\mathbf{z}) = \sum_{i=1}^{m} \phi_i(\mathbf{z})\mathbf{M}_i$,

$$\mathbf{M}(\mathbf{z}_*)\mathbf{z}_* = \mathbf{z}_* \lambda.$$

Now, it is straightforward to show that an eigenvector corresponding to the smallest eigenvalue of $\mathbf{M}(\mathbf{z}_*)$ is also an eigenvector corresponding to the smallest eigenvalue of $\mathbf{M}_D$ in equation (3) with $\mathbf{D} = \mathbf{D}(\mathbf{z}_*)$. Therefore, it is sufficient to show that $\lambda$ is the smallest eigenvalue of $\mathbf{M}(\mathbf{z}_*)$ to show that $\mathbf{z}_*$ is a fixed point. The following argument will establish this.

Due to the KKT first and second order necessary conditions, for all $\mathbf{w}$ such that $\mathbf{w}^T \mathbf{z}_* = 0$, we must have

$$\mathbf{w}^T \nabla\Phi(\mathbf{z}_*) = \mathbf{w}^T \mathbf{M}(\mathbf{z}_*)\mathbf{z}_* = 0,$$

and

$$\mathbf{w}^T \left(\nabla^2\Phi(\mathbf{z}_*) - 2\lambda\mathbf{I}\right)\mathbf{w} \geq 0. \tag{7}$$

Now,

$$\nabla^2\Phi(\mathbf{z}) = 2\sum_{i=1}^{m} \phi_i(\mathbf{z})\mathbf{M}_i + 2\sum_{i=1}^{m} \mathbf{M}_i \mathbf{z}\nabla\phi_i(\mathbf{z})^T$$

and

$$
\begin{aligned}
\nabla\phi_i(\mathbf{z}) &= \nabla\left(\prod_{\substack{j=1 \\ j\neq i}}^{m} \mathbf{z}^T \mathbf{M}_j \mathbf{z}\right) \\
&= \nabla\left(\frac{\Phi(\mathbf{z})}{\mathbf{z}^T \mathbf{M}_i \mathbf{z}}\right) \\
&= \frac{1}{\mathbf{z}^T \mathbf{M}_i \mathbf{z}}\nabla\Phi(\mathbf{z}) - \frac{2\Phi(\mathbf{z})}{(\mathbf{z}^T \mathbf{M}_i \mathbf{z})^2}\mathbf{M}_i \mathbf{z}.
\end{aligned}
$$

Therefore,

$$\mathbf{w}^T \nabla\phi_i(\mathbf{z}_*) = -\frac{2\Phi(\mathbf{z}_*)}{(\mathbf{z}_*{}^T \mathbf{M}_i \mathbf{z}_*)^2}\mathbf{w}^T \mathbf{M}_i \mathbf{z}_*. \tag{8}$$

Substituting expression (8) into the formula for $\mathbf{w}^T \left( \nabla^2 \Phi(\mathbf{z}_*) - 2\lambda \mathbf{I} \right) \mathbf{w}$ in the second order necessary conditions (7) gives

$$
\begin{aligned}
0 & \leq 2\mathbf{w}^T \mathbf{M}(\mathbf{z}_*)\mathbf{w} - 4 \left( \frac{\mathbf{w}^T \mathbf{M}_i \mathbf{z}_*}{\mathbf{z}_*{}^T \mathbf{M}_i \mathbf{z}_*} \right)^2 \Phi(\mathbf{z}_*) - 2\lambda \\
& \leq 2(\mu - \lambda),
\end{aligned}
$$

where $\mu = \mathbf{w}^T \mathbf{M}(\mathbf{z}_*)\mathbf{w}$. Thus, $\lambda \leq \mu$ for any eigenvalue $\mu$ of $\mathbf{M}(\mathbf{z}_*)$. Since $\lambda$ is the smallest eigenvalue of $\mathbf{M}(\mathbf{z}_*)$ we have established that a constrained minimizer $\mathbf{z}_*$ of $\Phi(\mathbf{z})$ satisfies $\mathbf{z}_* = \arg\min_{\|\mathbf{w}\|=1} F(\mathbf{z}_*, \mathbf{w})$. ∎

**Remark:** We have assumed in Lemma 3.4 that $\Phi(\mathbf{z}) \neq 0$. This is a reasonable assumption, since the only way $\Phi(\mathbf{z}) = 0$ is if $\|\mathbf{x}_i^{(0)}\| = \|\mathbf{x}_i^{(1)}\|$ for some pair $(\mathbf{x}_i^{(0)}, \mathbf{x}_i^{(1)})$. Since we are dealing with noisy sets, it is unlikely that these norms are precisely equal in practice. Nevertheless, we are considering equivalent reformulations that avoid this difficulty altogether.
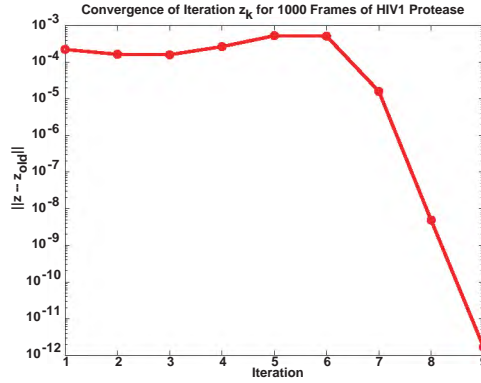


Figure 1: Convergence of 1000 frames of HIV-1 Protease using iteration (5).

Convergence of the iterates $\mathbf{z}_k$ produced by (5) is yet to be proven. However, the convergence history shown in Figures 1 and 2 is typical, and iteration (5) seems to be convergent in practice. Lemma 3.5 does at least establish that the sequence $\Phi(\mathbf{z}_k)$ is monotonically decreasing and convergent.

**Lemma 3.5** *The sequence $\Phi(\mathbf{z}_k)$, with $\mathbf{z}_k$ produced by iteration (5), is convergent.*

**Proof.** In the proof of Lemma 3.4, we show that a constrained minimizer $\mathbf{z}_*$ of

$$
\Phi(\mathbf{z}) = \prod_{i=1}^{m} \mathbf{z}^T \mathbf{M}_i \mathbf{z} = \prod_{i=1}^{m} \|\mathbf{x}_i^{(0)} - (\mathbf{I} - 2\mathbf{z}\mathbf{z}^T)\mathbf{x}_i^{(1)}\|^2
$$

9

is a fixed point to iteration (5). If we can show $\Phi(\mathbf{z}_k)$, where $\mathbf{z}_k$ satisfies iteration (5), is a monotonically decreasing function, we'll have proven that the sequence $\Phi(\mathbf{z}_k)$, with $\mathbf{z}_k$ produced by iteration (5), is convergent. Notice,

$$\frac{\Phi(\mathbf{z}_{k+1})}{\Phi(\mathbf{z}_k)} = \prod_{i=1}^{m} \frac{\|\mathbf{x}_i^{(0)} - (\mathbf{I} - 2\mathbf{z}_{k+1}\mathbf{z}_{k+1}^T)\mathbf{x}_i^{(1)}\|^2}{\|\mathbf{x}_i^{(0)} - (\mathbf{I} - 2\mathbf{z}_k\mathbf{z}_k^T)\mathbf{x}_i^{(1)}\|^2}.$$

And, $\mathbf{z}_{k+1}$ is chosen such that it minimizes the optimization problem (4); thus

$$\sum_{i=1}^{m} \frac{\|\mathbf{x}_i^{(0)} - (\mathbf{I} - 2\mathbf{z}_{k+1}\mathbf{z}_{k+1}^T)\mathbf{x}_i^{(1)}\|^2}{\|\mathbf{x}_i^{(0)} - (\mathbf{I} - 2\mathbf{z}_k\mathbf{z}_k^T)\mathbf{x}_i^{(1)}\|^2} \leq \sum_{i=1}^{m} \frac{\|\mathbf{x}_i^{(0)} - (\mathbf{I} - 2\mathbf{z}_k\mathbf{z}_k^T)\mathbf{x}_i^{(1)}\|^2}{\|\mathbf{x}_i^{(0)} - (\mathbf{I} - 2\mathbf{z}_k\mathbf{z}_k^T)\mathbf{x}_i^{(1)}\|^2} = m.$$

Since the geometric mean never exceeds the arithmetic mean,

$$\left[\prod_{i=1}^{m} \frac{\|\mathbf{x}_i^{(0)} - (\mathbf{I} - 2\mathbf{z}_{k+1}\mathbf{z}_{k+1}^T)\mathbf{x}_i^{(1)}\|^2}{\|\mathbf{x}_i^{(0)} - (\mathbf{I} - 2\mathbf{z}_k\mathbf{z}_k^T)\mathbf{x}_i^{(1)}\|^2}\right]^{(1/m)} \leq \frac{1}{m}\sum_{i=1}^{m} \frac{\|\mathbf{x}_i^{(0)} - (\mathbf{I} - 2\mathbf{z}_{k+1}\mathbf{z}_{k+1}^T)\mathbf{x}_i^{(1)}\|^2}{\|\mathbf{x}_i^{(0)} - (\mathbf{I} - 2\mathbf{z}_k\mathbf{z}_k^T)\mathbf{x}_i^{(1)}\|^2} \leq 1.$$

Thus,

$$\prod_{i=1}^{m} \frac{\|\mathbf{x}_i^{(0)} - (\mathbf{I} - 2\mathbf{z}_{k+1}\mathbf{z}_{k+1}^T)\mathbf{x}_i^{(1)}\|^2}{\|\mathbf{x}_i^{(0)} - (\mathbf{I} - 2\mathbf{z}_k\mathbf{z}_k^T)\mathbf{x}_i^{(1)}\|^2} \leq 1.$$

Hence, $\Phi(\mathbf{z}_k)$ is a monotonically decreasing sequence that is bounded below and is therefore convergent. ∎
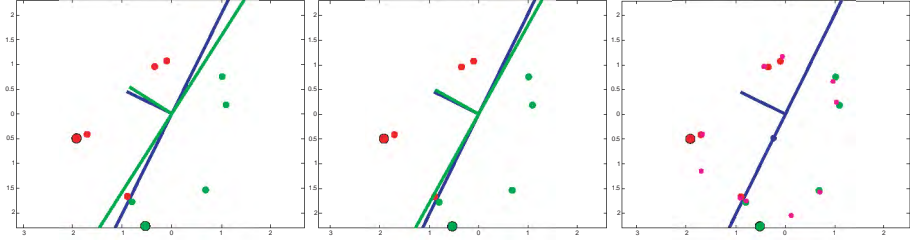


Figure 2: Iterations showing that our weighting is a good choice. Notice how as the iterations progress the normal converges to the correct solution, even in the presence of outliers (larger dots). The smaller dots in the last frame show our best symmetric approximation to the original data set.

We have compared the convergence of iteration (5) to a fixed point with the modified compass search method [7] on an equivalent optimization problem:

$$\min_{\|\mathbf{z}\|=1} \|\mathbf{z} - \mathbf{v}\|, \tag{9}$$

where, as before, $\mathbf{v}$ is the eigenvector associated with the smallest eigenvalue of equation (3) with $\mathbf{D} = \mathrm{diag}(f_j(\mathbf{z})^{-1})$. We see that, in general, iteration (5) converges faster and more efficiently when compared to the compass search method. Also, more accurate results are usually obtained with iteration (5).

10

# 4    Optimal Value of Rotational Axis q

Recall that for a perfectly rotationally symmetric set,

$$\mathbf{X}_j = (\mathbf{I} - \mathbf{Q}\mathbf{G}\mathbf{Q}^T)^j \mathbf{X}_0, \tag{10}$$

where the columns of $[\mathbf{q}, \mathbf{Q}]$ form an orthogonal set. This specification suggests a means to compute the axis of rotation.

**Lemma 4.1** *Suppose $\mathbf{X}_0$ has rank $n$ and that $\mathbf{G}$ is nonsingular. Then, $\mathbf{q}$ is an axis of rotational symmetry if and only if*

$$\mathbf{q}^T[(k-1)\mathbf{X}_0 - \sum_{j=1}^{k-1} \mathbf{X}_j] = 0. \tag{11}$$

**Proof.** First, note that if $\mathbf{q}$ is an axis of rotational symmetry, then $\mathbf{q}^T\mathbf{Q} = 0$ must hold, and thus,

$$\mathbf{q}^T\mathbf{X}_j = \mathbf{q}^T(\mathbf{I} - \mathbf{Q}\mathbf{G}\mathbf{Q}^T)^j\mathbf{X}_0 = \mathbf{q}^T\mathbf{X}_0, \quad \text{for} \quad j = 1, 2, \ldots, k,$$

which implies equation (11) must hold.

From equation (10),

$$
\begin{aligned}
\mathbf{X}_j &= (\mathbf{I} - \mathbf{Q}\mathbf{G}\mathbf{Q}^T)^j\mathbf{X}_0 \\
&= (\mathbf{q}\mathbf{q}^T + \mathbf{Q}(\mathbf{I} - \mathbf{G})\mathbf{Q}^T)^j\mathbf{X}_0 \\
&= (\mathbf{q}\mathbf{q}^T + \mathbf{Q}(\mathbf{I} - \mathbf{G})^j\mathbf{Q}^T)\mathbf{X}_0.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\sum_{j=1}^{k-1}\mathbf{X}_j &= ((k-1)\mathbf{q}\mathbf{q}^T + \mathbf{Q}(\sum_{j=1}^{k-1}(\mathbf{I} - \mathbf{G})^j)\mathbf{Q}^T)\mathbf{X}_0 \\
&= ((k-1)\mathbf{q}\mathbf{q}^T - \mathbf{Q}\mathbf{Q}^T)\mathbf{X}_0 = k\mathbf{q}\mathbf{q}^T\mathbf{X}_0 - \mathbf{X}_0,
\end{aligned}
$$

since $(\mathbf{I} - \mathbf{G})^k = \mathbf{I}$ implies $\sum_{j=1}^{k-1}(\mathbf{I} - \mathbf{G})^j = -\mathbf{I}$ when $\mathbf{G}$ is nonsingular. From this, it follows that

$$(k-1)\mathbf{X}_0 - \sum_{j=1}^{k-1}\mathbf{X}_j = k(\mathbf{I} - \mathbf{q}\mathbf{q}^T)\mathbf{X}_0.$$

Now, suppose $\hat{\mathbf{q}}$ is any unit vector that satisfies equation (11) (in place of $\mathbf{q}$). Since $\mathbf{X}_0$ is full rank and $\hat{\mathbf{q}}$ satisfies equation (11),

$$0 = \hat{\mathbf{q}}^T[(k-1)\mathbf{X}_0 - \sum_{j=1}^{k-1}\mathbf{X}_j] = k\hat{\mathbf{q}}^T(\mathbf{I} - \mathbf{q}\mathbf{q}^T)\mathbf{X}_0$$

implies that $\hat{\mathbf{q}} = \mathbf{q}(\hat{\mathbf{q}}^T\mathbf{q})$. Since both $\mathbf{q}$ and $\hat{\mathbf{q}}$ are unit length, it follows from Cauchy-Schwarz that $\hat{\mathbf{q}} = \pm\mathbf{q}$.  ∎

**Remark:** In $I\!\!R^3$ the only way $\mathbf{G}$ can be singular is if it is identically $\mathbf{0}$, and since we are assuming many points, it is also not unreasonable to assume $\mathbf{X}_0$ has full rank.

This exact condition is satisfying, but in general, we are not given a perfectly symmetric data set $\mathcal{S}$. Therefore, we need to be able to specify an approximate rotational axis $\mathbf{q}$ that best fits the data. To this end, we shall assume a partitioning of $\mathcal{S}$ into $\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_k$ such that the columns of the matrices are correctly paired. Then, we can formulate the optimization problem

$$\min_{\|\mathbf{q}\|=1} \{\|\mathbf{q}^T[(k-1)\mathbf{X}_0 - \sum_{i=1}^{k-1}\mathbf{X}_i]\|_F\} \tag{12}$$

to specify our approximate rotational axis of symmetry $\mathbf{q}$. Of course, we can characterize $\mathbf{q}$ as follows:

**Lemma 4.2** *The solution $\mathbf{q}$ to the minimization problem (12) is the unit eigenvector corresponding to the smallest eigenvalue of $\mathbf{MM}^T$, where*

$$\mathbf{M} = (k-1)\mathbf{X}_0 - \sum_{i=1}^{k-1}\mathbf{X}_i. \tag{13}$$

Note that this characterization provides a computational mechanism that is robust in the presence of noise. An alternate specification of $\mathbf{q}$ suggested by Minovic et al. is to consider the principal axis of the inertia matrix (correlation matrix) associated with the distinct eigenvalue for an initial guess to the rotational axes of symmetry. The motivation for this is that with exact symmetry the inertia matrix will have a distinct eigenvalue of multiplicity one and another eigenvalue of multiplicity $n-1$. However, in the presence of noise, this criteria may fail, since the eigenvalues are likely to be perturbed into a distinct set making it impossible to distinguish the "correct" eigenvalue. For example, consider the following 4-fold perfectly rotationally symmetric data set with respect to $\mathbf{q} = [1, \ 0, \ 0]^T$:

$$\begin{pmatrix} 1 & 4 & 0 & 1 & 4 & 0 & 1 & 4 & 0 & 1 & 4 & 0 \\ 0 & 1 & 4 & 0 & 0 & 1 & 0 & -1 & -4 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & -1 & 4 & 0 & 0 & -1 & 0 & 1 & 4 \end{pmatrix}$$

with eigenvalues 34.6667, 36, 36 after centering. In this case, we can clearly distinguish the distinct eigenvalue. However, if we perturb the data to the following data set,

$$\begin{pmatrix} 1 & 4 & 0.01 & 1 & 4 & 0.01 & 1 & 4 & 0.01 & 1 & 4 & 0.01 \\ 0.1 & 1 & 4 & .1 & 0 & 1 & 0.1 & -1 & -4 & 0.1 & 0 & 1 \\ 0.1 & 0.01 & 1 & 0.01 & -1 & 4 & 0.01 & 0 & -1 & 0.01 & 1 & 4 \end{pmatrix},$$

then the eigenvalues of the centered data become approximately 34.4870, 35.998, 36.0755. In this case, calculating the correct axis of rotation becomes impossible for the Minovic algorithm. But, our algorithm constructs it correctly.

As with reflective symmetry, we can introduce a weighting scheme that minimizes the influence of outliers in the supposed rotational symmetry relation:

$$\min_{\|\mathbf{q}\|=1} \{\|\mathbf{q}^T[(k-1)\mathbf{X}_0 - \sum_{i=1}^{k-1}\mathbf{X}_i]\mathbf{D}\|_F\} \tag{14}$$

where $\mathbf{D}$ is a diagonal weighting matrix. If such a weighting has been specified, then:

**Lemma 4.3** *The solution to the optimization problem (14) is the unit eigenvector $\mathbf{q}$ corresponding to the smallest eigenvalue of $\mathbf{M}\mathbf{D}^2\mathbf{M}^T$, where $\mathbf{M}$ is defined as in equation (13).*

As in reflective symmetry, we have developed an iterative re-weighting scheme to specify the weighting matrix $\mathbf{D}$ of the minimization problem (14) that effectively diminishes the influence of outliers in the final SPSVD approximation. Given a guess $\mathbf{z}$ of unit length, the $i$-th column of $\mathbf{M}$ is weighted by $g_i(\mathbf{z})^{-1}$, where $g_i(\mathbf{z}) = \|\mathbf{z}^T\left[(k-1)\mathbf{x}_i^{(0)} - \sum_{j=1}^{k}\mathbf{x}_i^{(j)}\right]\|$. If we define

$$G(\mathbf{z}, \mathbf{q}) = \sum_{i=1}^{m}\left(\frac{g_i(\mathbf{q})}{g_i(\mathbf{z})}\right)^2 = \|\mathbf{q}^T[(k-1)\mathbf{X}_0 - \sum_{i=1}^{k-1}\mathbf{X}_i]\mathbf{D}(\mathbf{z})\|_F^2,$$

then the approximate $\mathbf{q}$ associated with this weighting solves

$$\min_{\|\mathbf{q}\|=1} G(\mathbf{z}, \mathbf{q}). \tag{15}$$

The motivation for this is to put greater weight on points that are more symmetric with respect to $\mathbf{z}$ than points that are not. Then, $\mathbf{q}$ is constructed to have the optimal normal with respect to the weighting as described in Lemma 4.3. If $\mathbf{q}$ is not acceptable, then $\mathbf{z} \leftarrow \mathbf{q}$, and the process is repeated until an acceptable $\mathbf{q}$ is found. This suggests an iterative re-weighting. Given an initial guess $\mathbf{z}_0$ to the axis of rotation, we iterate:

$$\mathbf{z}_{k+1} = \arg\min_{\|\mathbf{q}\|=1} G(\mathbf{z}_k, \mathbf{q}) \tag{16}$$

until $\|\mathbf{z}_{k+1} - \mathbf{z}_k\|$ is under a predetermined tolerance. A fixed point of iteration (16) is the solution to the following max-min problem

$$\max_{\|\mathbf{z}\|=1}\left\{\min_{\|\mathbf{q}\|=1} G(\mathbf{z}, \mathbf{q})\right\}. \tag{17}$$

as the next lemma suggests.

**Lemma 4.4** *If $\mathbf{q} = \mathbf{z}$ is a fixed point of the iteration (16), then $\mathbf{q}$ is a solution to the max-min problem (17), and $G(\mathbf{z}, \mathbf{q}) = m$.*

**Proof.** The proof is essentially the same as the proof of Lemma 3.3. ∎

Moreover, we have the following

**Lemma 4.5** *There exists a fixed point to iteration (16).*

**Proof.** The proof is essentially the same as the proof of Lemma 3.4. ∎

We have also compared iteration (16) with the modified compass search method on the equivalent optimization problem:

$$\min_{\|\mathbf{z}\|=1} \|\mathbf{z} - \mathbf{q}\| \tag{18}$$

where $\mathbf{q}$ is the eigenvector associated with the smallest eigenvalue of $\mathbf{MD}^2\mathbf{M}^T$ with $\mathbf{D} = \text{diag}(g_j(\mathbf{z})^{-1})$. We have observed that iteration (16) is generally more efficient and produces more accurate fixed point solutions when compared to the compass search method.

# 5   Best Symmetric Approximation to a Set

To find the best reflective or rotational symmetric approximation to a set, we can take advantage of the following theorem. For reflective symmetry $\mathbf{R} = \mathbf{W}$ and $\mathbf{W}^2 = \mathbf{I}$, and in the case of rotational symmetry $\mathbf{R} = \mathbf{R}(\mathbf{q})$ and $\mathbf{R}(\mathbf{q})^k = \mathbf{I}$.

**Theorem 5.1** *If*

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_0 \\ \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_{k-1} \end{pmatrix}$$

*where*

$$\mathbf{R}^{k-i}\mathbf{X}_i = \mathbf{X}_0 + \mathbf{E}_i,$$

*and $\mathbf{R}^k = \mathbf{I}$, then*

$$\min_{\widehat{\mathbf{X}}_{j+1}=\mathbf{R}\widehat{\mathbf{X}}_j\, j=0,1,\ldots,k-2} \left\| \begin{pmatrix} \mathbf{X}_0 \\ \vdots \\ \mathbf{X}_{k-1} \end{pmatrix} - \begin{pmatrix} \widehat{\mathbf{X}}_0 \\ \vdots \\ \widehat{\mathbf{X}}_{k-1} \end{pmatrix} \right\|_F^2 = \frac{1}{k}\sum_{i=0}^{k-1}\sum_{j=i+1}^{k-1} \|\mathbf{E}_j - \mathbf{R}^{j-i}\mathbf{E}_i\|_F^2$$

*and the SVD*

$$\mathbf{USV}^T = \begin{pmatrix} \widehat{\mathbf{X}}_0 \\ \vdots \\ \widehat{\mathbf{X}}_{k-1} \end{pmatrix}$$

14

*satisfies*

$$\mathbf{U} = \frac{1}{\sqrt{k}} \begin{pmatrix} \mathbf{U}_0 \\ \vdots \\ \mathbf{U}_{k-1} \end{pmatrix}, \quad \mathbf{S} = \sqrt{k}\mathbf{S}_0, \quad \mathbf{V} = \mathbf{V}_0,$$

*where*

$$\mathbf{U}_j = \mathbf{R}^j \mathbf{U}_0, \quad \text{for} \quad j = 0, 1, 2, \ldots, k-1,$$

*and*

$$\mathbf{U}_0 \mathbf{S}_0 \mathbf{V}_0^T = \frac{1}{k}(\mathbf{X}_0 + \mathbf{R}^{k-1}\mathbf{X}_1 + \mathbf{R}^{k-2}\mathbf{X}_2 + \cdots + \mathbf{R}\mathbf{X}_{k-1}).$$

**Proof.** The proof will consist of a sequence of straightforward lemmas. We begin by assuming that we have perfect symmetry.

**Lemma 5.2** *Suppose* $\mathbf{E}_j = 0$ *for all* $j = 0, 1, 2, \ldots, k-1$ *and let*

$$\begin{pmatrix} \mathbf{X}_0 \\ \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_{k-1} \end{pmatrix} = \begin{pmatrix} \mathbf{U}_0 \\ \mathbf{U}_1 \\ \vdots \\ \mathbf{U}_{k-1} \end{pmatrix} \mathbf{S}\mathbf{V}^T \tag{19}$$

*be the short form SVD of* $\mathbf{X}$*. Then,*

$$\mathbf{U}_i = \mathbf{R}^i \mathbf{U}_0$$

*where* $i = 0, 1, \ldots, k-1$*.*

**Proof.** From (19), we have

$$\mathbf{U}_j = \mathbf{X}_j \mathbf{V}\mathbf{S}^{-1}$$

*where* $\mathbf{U}_0^T\mathbf{U}_0 + \mathbf{U}_1^T\mathbf{U}_1 + \cdots + \mathbf{U}_{k-1}^T\mathbf{U}_{k-1} = \mathbf{I}$*. Thus,*

$$\mathbf{U}_j = \mathbf{X}_j \mathbf{V}\mathbf{S}^{-1} = \mathbf{R}^j \mathbf{X}_0 \mathbf{V}\mathbf{S}^{-1} = \mathbf{R}^j \mathbf{U}_0.$$

∎

Therefore, when $\mathbf{R}$ is known, the SVD of a perfectly symmetric set may be efficiently computed by just taking the SVD of $\mathbf{X}_0$ and putting $\mathbf{U}_j = \mathbf{R}\mathbf{U}_{j-1}$, $1 \leq j \leq k-1$. Combining this fact with the following lemma leads to an algorithm for calculating the best low rank approximation to a matrix that preserves symmetry.

**Lemma 5.3** *Let* $\mathbf{X}_0 = \mathbf{U}_0 \mathbf{S}_0 \mathbf{V}_0^T$ *be the short form SVD of* $\mathbf{X}_0$ *where* $\mathbf{U}_0^T\mathbf{U}_0 = \mathbf{V}_0^T\mathbf{V}_0 = \mathbf{I}$*. Then,*

$$\begin{pmatrix} \mathbf{X}_0 \\ \vdots \\ \mathbf{X}_0 \end{pmatrix} = \mathbf{U}\mathbf{S}\mathbf{V}^T$$

*is the SVD of the composite matrix, where*

$$\mathbf{U} = \frac{1}{\sqrt{k}} \begin{pmatrix} \mathbf{U}_0 \\ \vdots \\ \mathbf{U}_0 \end{pmatrix}, \quad \mathbf{S} = \sqrt{k}\mathbf{S}_0, \quad \mathbf{V} = \mathbf{V}_0.$$

**Proof.** Clearly, $\mathbf{U}^T\mathbf{U} = \mathbf{I}$, and

$$\begin{pmatrix} \mathbf{X}_0 \\ \vdots \\ \mathbf{X}_0 \end{pmatrix} = \begin{pmatrix} \mathbf{U}_0 \\ \vdots \\ \mathbf{U}_0 \end{pmatrix} \mathbf{S}_0\mathbf{V}_0^T = \frac{1}{\sqrt{k}} \begin{pmatrix} \mathbf{U}_0 \\ \vdots \\ \mathbf{U}_0 \end{pmatrix} \sqrt{k}\mathbf{S}_0\mathbf{V}_0^T$$

$$= \mathbf{U}\mathbf{S}\mathbf{V}^T,$$

which is indeed the SVD. ∎

We are now ready to give the best low rank approximation that preserves symmetry for a noisy data set.

**Lemma 5.4** *Let* $\widehat{\mathbf{Z}} = \frac{1}{k}(\mathbf{Z}_0 + \mathbf{Z}_1 + \cdots + \mathbf{Z}_{k-1})$. *Then,* $\mathbf{Z} = \widehat{\mathbf{Z}}$ *solves*

$$\min_{\mathbf{Z}} \left\| \begin{pmatrix} \mathbf{Z}_0 \\ \vdots \\ \mathbf{Z}_{k-1} \end{pmatrix} - \begin{pmatrix} \mathbf{Z} \\ \vdots \\ \mathbf{Z} \end{pmatrix} \right\|_F^2.$$

**Proof.** Consider

$$\left\| \begin{pmatrix} \mathbf{Z}_0 \\ \vdots \\ \mathbf{Z}_{k-1} \end{pmatrix} - \begin{pmatrix} \mathbf{Z} \\ \vdots \\ \mathbf{Z} \end{pmatrix} \right\|_F^2 = \|\mathbf{Z}_0 - \mathbf{Z}\|_F^2 + \|\mathbf{Z}_1 - \mathbf{Z}\|_F^2 + \cdots + \|\mathbf{Z}_{k-1} - \mathbf{Z}\|_F^2,$$

and note that

$$\|\mathbf{Z}_i - \mathbf{Z}\|_F^2 = tr(\mathbf{Z}_i^T\mathbf{Z}_i) - 2tr(\mathbf{Z}_i^T\mathbf{Z}) + tr(\mathbf{Z}^T\mathbf{Z})$$

for $i = 0, 1, 2, \ldots, k - 1$. Therefore,

$$\left\| \begin{pmatrix} \mathbf{Z}_0 \\ \vdots \\ \mathbf{Z}_{k-1} \end{pmatrix} - \begin{pmatrix} \mathbf{Z} \\ \vdots \\ \mathbf{Z} \end{pmatrix} \right\|_F^2 = tr(\sum_{i=0}^{k-1} \mathbf{Z}_i^T\mathbf{Z}_i) - 2tr(\sum_{i=0}^{k-1} \mathbf{Z}_i^T\mathbf{Z}) + (k)tr(\mathbf{Z}^T\mathbf{Z}).$$

However,

$$-2tr(\sum_{i=0}^{k-1}\mathbf{Z}_i^T\mathbf{Z}) + (k)tr(\mathbf{Z}^T\mathbf{Z}) = -2tr(\frac{1}{\sqrt{k}}(\sum_{i=0}^{k-1}\mathbf{Z}_i)^T\sqrt{k}\mathbf{Z}) + tr((\sqrt{k}\mathbf{Z})^T(\sqrt{k}\mathbf{Z}))$$

$$= -tr(\frac{1}{\sqrt{k}}\sum_{i=0}^{k-1}\mathbf{Z}_i^T\frac{1}{\sqrt{k}}\sum_{i=0}^{k-1}\mathbf{Z}_i) + tr(\frac{1}{\sqrt{k}}\sum_{i=0}^{k-1}\mathbf{Z}_i^T\frac{1}{\sqrt{k}}\sum_{i=0}^{k-1}\mathbf{Z}_i)$$

$$-2tr(\frac{1}{\sqrt{k}}(\sum_{i=0}^{k-1}\mathbf{Z}_i)^T\sqrt{k}\mathbf{Z}) + tr((\sqrt{k}\mathbf{Z})^T(\sqrt{k}\mathbf{Z}))$$

$$= -\frac{1}{k}tr(\sum_{i=0}^{k-1}\mathbf{Z}_i^T\sum_{j=0}^{k-1}\mathbf{Z}_j) + \|\frac{1}{\sqrt{k}}\sum_{i=0}^{k-1}\mathbf{Z}_i - \sqrt{k}\mathbf{Z}\|_F^2.$$

The fact that $tr\mathbf{Z}_i^T\mathbf{Z}_j = tr\mathbf{Z}_j^T\mathbf{Z}_i$ and some tedious book keeping will show

$$tr(\sum_{i=0}^{k-1}\mathbf{Z}_i^T\mathbf{Z}_i) - \frac{1}{k}tr(\sum_{i=0}^{k-1}\mathbf{Z}_i^T\sum_{j=0}^{k-1}\mathbf{Z}_j) = \frac{k-1}{k}tr(\sum_{i=0}^{k-1}\mathbf{Z}_i^T\mathbf{Z}_i) - \frac{2}{k}\sum_{i=0}^{k-1}\sum_{j=i+1}^{k-1}tr(\mathbf{Z}_i^T\mathbf{Z}_j)$$

$$= \frac{1}{k}\sum_{i=0}^{k-1}\sum_{j=i+1}^{k-1}\|\mathbf{Z}_i - \mathbf{Z}_j\|_F^2.$$

Hence,

$$\left\|\begin{pmatrix} \mathbf{Z}_0 \\ \vdots \\ \mathbf{Z}_{k-1} \end{pmatrix} - \begin{pmatrix} \mathbf{Z} \\ \vdots \\ \mathbf{Z} \end{pmatrix}\right\|_F^2 = \frac{1}{k}\sum_{i=0}^{k-1}\sum_{j=i+1}^{k-1}\|\mathbf{Z}_i - \mathbf{Z}_j\|_F^2 + k\left\|\frac{1}{k}\sum_{i=0}^{k-1}\mathbf{Z}_i - \mathbf{Z}\right\|_F^2$$

$$\geq \frac{1}{k}\sum_{i=0}^{k-1}\sum_{j=i+1}^{k-1}\|\mathbf{Z}_i - \mathbf{Z}_j\|_F^2$$

with equality if and only if

$$\mathbf{Z} = \widehat{\mathbf{Z}} = \frac{1}{k}\sum_{i=0}^{k-1}\mathbf{Z}_i.$$

∎

These lemmas establish Theorem 5.1, since solving

$$\min_{\widehat{\mathbf{X}}_{j+1}=\mathbf{R}\widehat{\mathbf{X}}_j}\left\|\begin{pmatrix} \mathbf{X}_0 \\ \vdots \\ \mathbf{X}_{k-1} \end{pmatrix} - \begin{pmatrix} \widehat{\mathbf{X}}_0 \\ \vdots \\ \widehat{\mathbf{X}}_{k-1} \end{pmatrix}\right\|_F^2$$

is equivalent to solving

$$\min_{\widehat{\mathbf{X}}_0}\left\|\begin{pmatrix} \mathbf{X}_0 \\ \mathbf{R}^{k-1}\mathbf{X}_1 \\ \vdots \\ \mathbf{R}\mathbf{X}_{k-1} \end{pmatrix} - \begin{pmatrix} \widehat{\mathbf{X}}_0 \\ \widehat{\mathbf{X}}_0 \\ \vdots \\ \widehat{\mathbf{X}}_0 \end{pmatrix}\right\|_F^2$$

17

because

$$
\begin{pmatrix}
\mathbf{I} & & & \\
& \mathbf{R}^{k-1} & & \\
& & \ddots & \\
& & & \mathbf{R}
\end{pmatrix}
$$

is unitary. Therefore, by Lemma 5.4, $\widehat{\mathbf{X}}_0 = \frac{1}{k}\sum_{i=0}^{k-1}\mathbf{R}^{k-i}\mathbf{X}_i$, and

$$
\min_{\widehat{\mathbf{X}}_j = \mathbf{R}^j\widehat{\mathbf{X}}_0}\left\|\begin{pmatrix}\mathbf{X}_0 \\ \vdots \\ \mathbf{X}_{k-1}\end{pmatrix} - \begin{pmatrix}\widehat{\mathbf{X}}_0 \\ \vdots \\ \widehat{\mathbf{X}}_{k-1}\end{pmatrix}\right\|_F^2 = \frac{1}{k}\sum_{i=0}^{k-1}\sum_{j=i+1}^{k-1}\|\mathbf{E}_j - \mathbf{R}^{j-i}\mathbf{E}_i\|_F^2
$$

where $\mathbf{R}^{k-i}\mathbf{X}_i = \mathbf{X}_0 + \mathbf{E}_i$.  ∎

# 6   Algorithms and Computational Results

The algorithmic structure for both the reflective and rotational SPSVD is the same. It consists of two major steps

1. Determine the normal $\mathbf{w}$ or the axis $\mathbf{q}$ for reflective or rotational symmetry respectively.

2. Compute the standard SVD

$$
\mathbf{U}_0\mathbf{S}_0\mathbf{V}_0^T = \frac{1}{k}(\mathbf{X}_0 + \mathbf{R}^{k-1}\mathbf{X}_1 + \mathbf{R}^{k-2}\mathbf{X}_2 + \cdots + \mathbf{R}\mathbf{X}_{k-1})
$$

   where $\mathbf{R}$ is a reflector determined by $\mathbf{w}$ or a rotation about the axis determined by $\mathbf{q}$.

We seek the dominant (largest) singular values and this can be done in a straightforward manner using the ARPACK software on a serial computer or P_ARPACK on a parallel system. Of course, one might question the use of ARPACK on dense problems. However, the timings shown in Figure 3 clearly verify that it is computationally more efficient to calculate only the leading $k$ terms (singular values) using ARPACK instead of computing all of the singular values and then discarding $n-k$ of them for large scale matrices. One may either specify $k$ or utilize a restarting scheme to adjust $k$ until $\sigma_k \geq tol * \sigma_1 > \sigma_{k+1}$. The important computational point is that only matrix-vector products of the form

$$
\mathbf{u} = \frac{1}{k}(\mathbf{X}_0 + \mathbf{R}^{k-1}\mathbf{X}_1 + \mathbf{R}^{k-2}\mathbf{X}_2 + \cdots + \mathbf{R}\mathbf{X}_{k-1})\mathbf{v},
$$

are required and this is slightly less work than is needed to compute the corresponding standard SVD of $\mathbf{X}$ without the symmetry constraint.
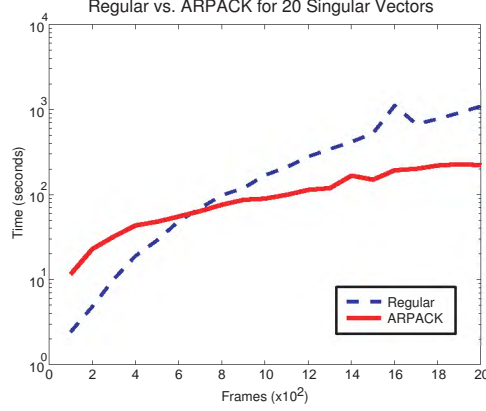
Figure 3: Comparison of calculating the largest 20 singular vectors of an HIV1 Protease trajectory using ARPACK and a dense SVD solver.

### 6.0.1 SPSVD in Protein Dynamics

Given a dynamical system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \;\; \mathbf{x}(0) = \mathbf{x}_0$, there are well known techniques for dimension reduction based upon the Gramian of the trajectory $\{\mathbf{x}(t), t \geq 0\}$. The technique is known as Proper Orthogonal Decomposition in computational fluid dynamics, as Karhunen-Loeve decomposition in face recognition and detection, and as Principle Component Analysis in molecular dynamics. For a system with $n$-dimensional state vectors, the Gramian

$$\mathcal{P} = \int_0^\infty \mathbf{x}(\tau)\mathbf{x}(\tau)^T d\tau$$

is an $n \times n$ symmetric positive (semi-)definite matrix (assuming it exists). The eigensystem of $\mathcal{P}$

$$\mathcal{P} = \mathbf{U}\mathbf{S}^2\mathbf{U}^T$$

provides an orthogonal basis via the columns of $\mathbf{U}$, and in this basis we have the representation

$$\mathbf{x}(t) = \mathbf{U}\mathbf{S}\mathbf{v}(t)$$

with the components of $\mathbf{v}(t)$ being mutually orthogonal $\mathcal{L}_2(0, \infty)$ functions. If the diagonal elements of the positive semidefinite diagonal matrix $\mathbf{S}$ decay rapidly (assuming they are in decreasing order), then a reduced basis representation of the trajectory may be obtained by discarding the trailing terms and considering the approximation $\mathbf{x}_k = \mathbf{U}_k\mathbf{S}_k\mathbf{v}_k(t)$ where the subscript $k$ denotes the leading $k$ columns and/or components. This is usually approximated using snapshots consisting of values $\mathbf{x}(t_j)$ of the trajectory at discrete time points and forming the $n \times m$ matrix

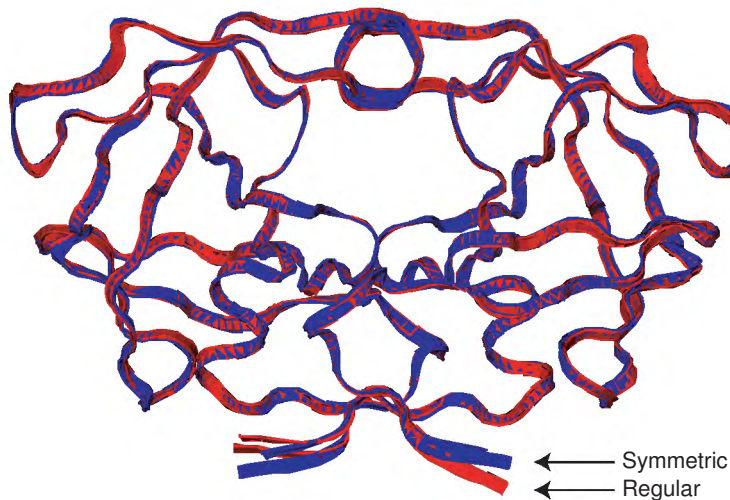$$\mathbf{X} = [\mathbf{x}(t_1), \mathbf{x}(t_2), \ldots, \mathbf{x}(t_m)].$$

19

Figure 4: Comparison of SVD vs. SPSVD. Notice the nice fit for all but the indicated region and its symmetric counterpart.

The SVD of $\mathbf{X}$ provides

$$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^T \approx \mathbf{U}_k\mathbf{S}_k\mathbf{V}_k^T,$$

where

$$\mathbf{U}^T\mathbf{U} = \mathbf{V}^T\mathbf{V} = \mathbf{I}_n \quad \mathbf{S} = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_n)$$

with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$. This is a direct approximation to the continuous derivation if we consider

$$\mathcal{P} \approx \frac{1}{m}\mathbf{X}\mathbf{X}^T = \frac{1}{m}\sum_j \mathbf{x}(t_j)\mathbf{x}(t_j)^T,$$

where the approximation to $\mathcal{P}$ is given by a quadrature rule. Here, we are concerned with introducing symmetry constraints into this approximation when appropriate. In molecular dynamics, there is often a known spatial structural symmetry for the state variables, and the purpose of the constrained SVD approximation developed here is to impose such symmetry constraints on the approximate trajectory through the SPSVD.

This method has been implemented using P_ARPACK on a Linux cluster with 6 dual-processor nodes consisting of 1600MHz AMD Athlon processors with 1GB RAM per node and a 1GB/s Ethernet connection. The method was applied to compute the leading 20 symmetric major modes for a HIV-1 protease molecule. The molecule consists of 3120 atoms, and hence, the state has 9360 degrees of freedom. The molecular dynamics trajectory consisted of 10000 time steps (snapshots). This resulted in

1. The first 20 symmetric singular vectors took 244 secs.
   This includes axis of rotation determination.

2. The first 20 standard singular vectors took 118 secs.

This may seem contradictory to the claim that the SPSVD should be as efficient as regular SVD. However, the need to compute the axis of rotation significantly adds to the run time. If more singular vectors are computed, the SPSVD indeed runs faster than regular SVD.

1. The first 50 symmetric singular vectors took 312 secs.
   This includes axis of rotation determination.

2. The first 50 standard singular vectors took 390 secs.

These computations were done for both reflective and rotational symmetry with essentially the same computational time. The computation of the reflective normal or the axis of rotation was included in both SPSVD approximations. As this normal/axis determination is quite demanding, these computations indicate that obtaining the leading terms of the SVD is comparable for both the symmetry preserving and standard SVD cases. Moreover, both are well suited to the large scale setting when P_ARPACK is used.

It turns out that HIV-1 protease has a 2-fold rotational symmetry and this aspect is preserved while providing good approximations to the full trajectory as can be seen in Figure 4. Additional visualizations are available at the web site http://www.caam.rice.edu/~sorensen/ under "recent talks".

### 6.0.2 Face Recognition

Generalizations of techniques described here can be used to orient faces once the plane of symmetry has been found. Once the correct orientation is attained, the SPSVD can find the best symmetric approximation to the face.

We notice that a face seems to have reflective symmetry through the vertical mid-line of the face (through the center of the eyes, middle of the nose, etc.). Therefore, if a face is correctly oriented, we have a reflectively symmetric data set of intensity values. The left half of the face forms $\mathbf{X}_0$, while the right half gives us $\mathbf{X}_1$. Note that the columns of $\mathbf{X}_1$ will have to be in reverse order to maintain correctly paired data points with relation to $\mathbf{X}_0$. Then, using SPSVD, we know that our best symmetric approximation will be formed by taking the average of the intensity levels of the left and right half of the face, i.e., the best symmetric approximation

$$\mathbf{S} = [\mathbf{A} \ \hat{\mathbf{A}}],$$

where $\mathbf{A} = \frac{1}{2}(\mathbf{X}_0 + \mathbf{X}_1)$ and $\hat{\mathbf{A}}$ is the matrix $\mathbf{A}$ with its columns in reverse order.

The SPSVD was applied to a series of newly synthesized, laser-scanned (Cyberware TM), 256×256 gray scaled pixel heads without hair. The face database was provided by the Max-Planck Institute for Biological Cybernetics in Tuebingen, Germany. An example of one of the faces and its symmetric counterpart

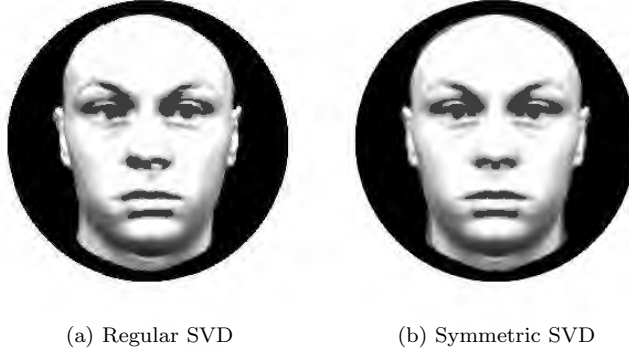(a) Regular SVD      (b) Symmetric SVD

Figure 5: Comparison of SVD verses SPSVD on faces.

can be seen in Figure 5. The SPSVD gives a good approximation to the original head, while the storage is essentially cut in half. We should also note that the sudden decrease of the singular values in the SPSVD occurs at an index that is approximately half that of the regular SVD (Figure 6). This suggests that a lower rank approximation from the SPSVD could give a better approximation to the original data set when compared to a regular low rank SVD approximation.

# 7 Conclusion

This paper has described a mathematical formulation of a symmetry preserving singular value decomposition which has led to practical (parallel) algorithms suitable for large scale computation. Criteria and methods were given for the calculation of reflective normal and rotational axis of symmetry of objects in $I\!R^n$ that are able to overcome problems with noisy data and outliers. The resulting technique is able to compute the best low rank symmetry preserving approximation to a given set.
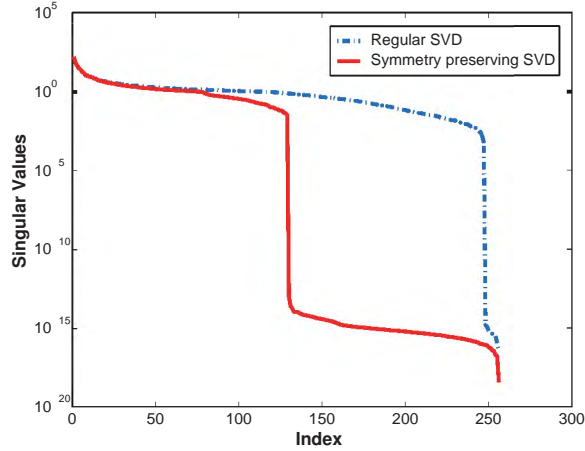
# 8 Acknowledgment

Figure 6: Singular values of SVD and SPSVD.

# References

[1] M. Atallah, *On symmetry detection*, IEEE Trans. Comput, C-34 (1985), pp. 663–636.

[2] N. Aubry, W.-Y. Lian, and E. S. Titi, *Preserving symmetries in the proper orthogonal decomposition*, SIAM J. Sci. Comput., 14 (1993), pp. 483–505.

[3] O. Colliot, A. V. Tuzikov, R. M. Cesar, and I. Bloch, *Approximate reflectional symmetries of fuzzy objects with an application in model-based object recognition*, Fuzzy Sets and Systems, 147 (2004), pp. 141–163.

[4] M. Kazhdan, B. Chazelle, D. Dobkin, T. Funkhouser, and S. Rusinkiewics, *A reflective symmetry descriptor for 3D models*, Algorithmica, 38 (2003), pp. 201–225.

[5] M. Kirby and L. Sirovich, *Application of the Karhunen-Loeve procedure for the characterization of human faces*, J. Opt. Soc. Amer. A, 4 (1987), pp. 519–524.

[6] ———, *Low-dimensional procedure for the characterization of human faces*, IEEE Trans. Pattern Anal. Mach. Intell., 12 (1990), pp. 103–108.

[7] T. G. Kolda, R. M. Lewis, and V. Torczon, *Optimization by direct search: new perspectives on some classical and modern methods*, SIAM Review, 45 (2003), pp. 385–482.

[8] R. Lehoucq, D. Sorensen, and C. Yang, *ARPACK Users Guide: Solution of Large Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods*, SIAM Publications, Philadelphia, 1998.

[9]  G. Marola, *On the detection of the axes of symmetry of symmetric and almost symmetric planar images*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2 (1989), pp. 104–108.

[10]  P. Minovic, S. Ishikawa, and K. Kato, *Three dimensional symmetry identification part I :theory*, Memoirs of the Kyushu Institute of Technology, 21 (1992), pp. 1–16.

[11]  ———, *Three dimensional symmetry identification part II :general algorithm and its application to medical images*, Memoirs of the Kyushu Institute of Technology, 21 (1992), pp. 17–26.

[12]  P. Minovic, S. Ishikawa, and K. Kato, *Symmetry identification of a 3D object represented by octree*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 15 (1993), pp. 507–514.

[13]  D. O'Mara and R. Owens, *Measuring bilateral symmetry in digital images*, in TENCON Digital Signal Processing Applications, 1996, pp. 151–156.

[14]  C. Sun and J. Sherrah, *3D symmetry detection using the extended gaussian image*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 19 (1997), pp. 164–169.

[15]  K. R. Symon, *Mechanics*, Addison-Wesley, Philippines, 1971.

[16]  H. Zabrodsky, S. Peleg, and D. Avnir, *Continuous symmetry measures*, J. American Chem. Soc, 114 (1992), pp. 7843–7851.

[17]  ———, *Symmetry as a continuous feature*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(12) (December 1995), pp. 1154–1166.