We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Application of Linear and Nonlinear Dimensionality Reduction Methods

Ramana Vinjamuri[1,4], Wei Wang[1,4], Mingui Sun[2] and Zhi-Hong Mao[3]
[1]*Department of Physical Medicine and Rehabilitation*
[2]*Department of Neurological Surgery*
[3]*Department of Electrical and Computer Engineering*
[4]*Center for Neural Basis of Cognition*
University of Pittsburgh, Pittsburgh, PA
USA

1. Introduction

Dimensionality reduction methods have proved to be important tools in exploratory analysis as well as confirmatory analysis for data mining in various fields of science and technology. Where ever applications involve reducing to fewer dimensions, feature selection, pattern recognition, clustering, dimensionality reduction methods have been used to overcome the curse of dimensionality. In particular, Principal Component Analysis (PCA) is widely used and accepted linear dimensionality reduction method which has achieved successful results in various biological and industrial applications, while demanding less computational power. On the other hand, several nonlinear dimensionality reduction methods such as kernel PCA (kPCA), Isomap and local linear embedding (LLE) have been developed. It has been observed that nonlinear methods proved to be effective only for specific datasets and failed to generalize over real world data, even at the cost of heavy computational burden to accommodate nonlinearity.

We have systematically investigated the use of linear dimensionality reduction methods in extracting movement primitives or synergies in hand movements in Vinjamuri et al. (2010a;b; 2011). In this chapter, we applied linear (PCA and Multidimensional Scaling (MDS)) and nonlinear (kPCA, Isomap, LLE) dimensionality reduction methods in extracting kinematic synergies in grasping tasks of the human hand. At first, we used PCA and MDS on joint angular velocities of the human hand, to derive synergies. The results obtained indicated ease and effectiveness of using PCA. Then we used nonlinear dimensionality reduction methods for deriving synergies. The synergies extracted from both linear and nonlinear methods were used to reconstruct the joint angular velocities of natural movements and ASL postural movements by using an l_1-minimization algorithm. The results suggest that PCA outperformed all three nonlinear methods in reconstructing the movements.

2. Synergies

The concept of synergies (in Greek *synergos* means working together) was first represented numerically by Bernstein Bernstein (1967). Although synergies were originally defined by Bernstein as high-level control of kinematic parameters, different definitions of synergies exist

and the term has been generalized to indicate the shared patterns observed in the behaviors of muscles, joints, forces, actions, etc. Synergies in hand movements especially present a complex optimization problem as to how the central nervous system (CNS) controls the hand with over 25 degrees of freedom(DoF) (Mackenzie & Iberall (1994)). Yet, the CNS handles all the movements effortlessly and at the same time dexterously. Endeavoring to solve the DoF problem, many researchers have proposed several concepts of synergies such as the following:

(i) Postural synergies: In Jerde et al. (2003); Mason et al. (2001); Santello et al. (1998; 2002); Thakur et al. (2008); Todorov & Ghahramani (2004), researchers found that the entire act of grasp can be described by a small number of dominant postures, which were defined as postural synergies.

(ii) Kinematic synergies: Studies in Grinyagin et al. (2005); Vinjamuri et al. (2007) expressed the angular velocities of finger joints as linear combinations of a small number of kinematic synergies, which were also angular velocities of finger joints but were extracted from a large set of natural movements. Kinematic synergies are not limited to hand movements. In (d'Avella et al. (2006)), d'Avella et al. reported that kinematic synergies were found in tracking 7 DoF arm movements.

(iii) Dynamic synergies: Dynamic synergies were defined as stable correlations between joint torques that were found during precision grip movements in Grinyagin et al. (2005).

The above classification was already presented in Vinjamuri et al. (2010b). In addition to synergies proposed in postures, and kinematics which are of relevance to the current study, synergies were also proposed in muscle activities d'Avella et al. (2006).

2.1 What are temporal postural synergies?

The synergies presented in this chapter are purely kinematic synergies. These are the synergies derived from angular velocities of the finger joints of the human hand collected during grasping tasks. For example, in the Fig. 1 two synergies ($\mathbf{s}^1$, $\mathbf{s}^2$) combine using a weighted linear combination ($w_1\mathbf{s}^1 + w_2\mathbf{s}^2$) to achieve a grasping hand movement. w_1 and w_2 represent weights of control signals. Each row of a synergy corresponds to the angular velocity profile of a finger joint; For example, the first synergy represents the synchronous large movement of first joint and medium movement of the second joint followed by a small movement of the third joint. In this example, $\mathbf{s}^1$ (blue) and $\mathbf{s}^2$ (brown) form a weighted ($w_1 = w_2 = 0.5$) combination to result in the aggregate movement(black) on the right hand side. For illustration purposes, only 3 of 10 joints of the hand are shown in the synergies and the reconstructed movement. Also shown in the figure are the hand postures of the reconstructed movement across time. As these synergies preserve both the temporal structure and the postural information these are termed as *temporal postural synergies* (Vinjamuri et al. (2010a;b)).

2.2 Applications of synergies

In our attempt to apply linear and nonlinear dimensionality reduction methods to solve the problem of extraction of synergies, let us first know how these synergies are being used in the real world applications in the areas of prosthesis and rehabilitation.

(i) Prosthetics: Apart from neuro-physiological significance, synergies are viewed to be crucial design elements in future generation prosthetic hands. Biologically inspired synergies have already taken prime place in artificial hands (Popovic & Popovic (2001)). Synergies based
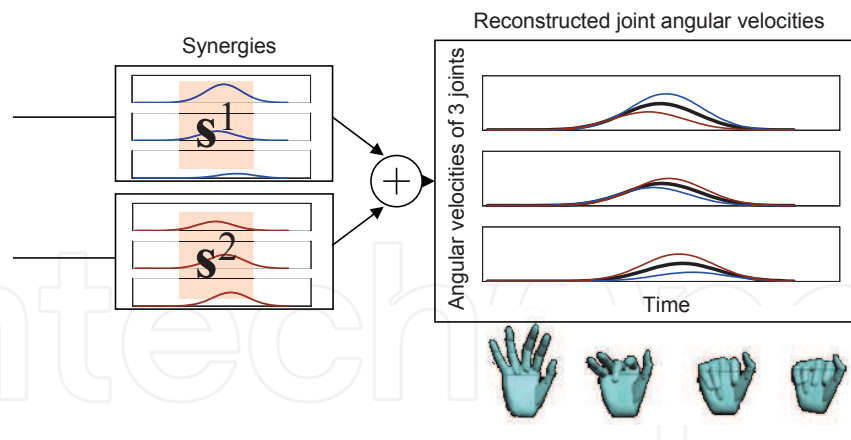
Fig. 1. Two distinct synergies ($\mathbf{s}^1$, $\mathbf{s}^2$) use a weighted linear combination ($w_1\mathbf{s}^1 + w_2\mathbf{s}^2$) to achieve a grasping hand movement. w_1 and w_2 represent weights of control signals. Each row of a synergy corresponds to the angular velocity profile of a finger joint; For example, the first synergy represents the synchronous large movement of first joint and medium movement of the second joint followed by a small movement of the third joint. In this example, $\mathbf{s}^1$ (blue) and $\mathbf{s}^2$ (brown) form a weighted ($w_1 = w_2 = 0.5$) combination to result in the aggregate movement(black) on the right hand side. For illustration purposes, only 3 of 10 joints of the hand are shown in the synergies and the reconstructed movement. Also shown in the figure are the hand postures of the reconstructed movement across time. Adapted from Vinjamuri et al. (2011)

on the principles of data reduction and dimensionality reduction, are soon to find place in tele-surgery and tele-robotics (Vinjamuri et al. (2007)). Synergies are projected to be miniature windows to provide immense help in next generation rehabilitation. Recently our group has demonstrated a synergy based brain machine interface where two control signals calculated from the spectral powers of the brain signals controlled two synergies, that commanded a 10 DoF virtual hand (Vinjamuri et al. (2011)). This showed promising results for controlling a synergy-based neural prosthesis.

(ii) Diagnostics: Applying similar concepts of synergies on the hand movements of the individuals with movement disorders, the sources that contain the tremor were isolated. Using blind source separation and dimensionality reduction methods, the possible neural sources that contained tremor were extracted from the hand movements of individuals with Essential Tremor (Vinjamuri et al. (2009)). This led to an efficient quantification of tremor.

(iii) Robotics: Biologically inspired synergies are being used in balance control of humanoid robots (Hauser et al. (2007)). Based on the principle that biological organisms recruit kinematic synergies that manage several joints, a control strategy for balance of humanoid robots was developed. This control strategy reduced computational complexity following a biological framework that central nervous system reduces the computational complexity of managing numerous degrees of freedom by effectively utilizing the synergies. Biologically inspired neural network controller models (Bernabucci et al. (2007)) that can manage ballistic arm movements have been developed. The models simulated the kinematic aspects, with bell-shaped wrist velocity profiles, and generated movement specific muscular synergies for the execution of movements.

(iv) Rehabilitation: Bimanual coordination is damaged in brain lesions and brain disorders Vinjamuri et al. (2008). Using a small set of modifiable and adjustable synergies

tremendously simplifies the task of learning new skills or adapting to new environments. Constructing internal neural representations from a linear combination of a reduced set of basis functions might be crucial for generalizing to novel tasks and new environmental conditions (Flash & Hochner (2005); Poggio & Bizzi (2004)).

2.3 Extraction of synergies

Synergies or movement primitives are viewed as small building blocks of movement that are present inherently within the movements and are shared across several movements. In other words, for example, in a set of hundred grasping movements, there might be a five or six synergies that are shared and common across all the movements. So it is to say that these hundred hand movements are composed of synergies. How do we decompose these hundred hand movements to a few building blocks of movement? This is the problem we are trying to solve.

In order to extract these primitives, several methods have been used. Several researchers view this as a problem of extracting basis functions. In fact, PCA can be viewed as extracting basis functions that are orthogonal to each other. Radial basis functions were also used as synergy approximations. Gradient descend method and non-negative matrix factorization methods (d'Avella et al. (2003)), multivariate statistical techniques (Santello et al. (2002)) were used in extracting the synergies. Different from the above interpretations of synergies, Todorov & Ghahramani (2004) suggested that synergistic control may not mean dimensionality reduction or simplification, but might imply task optimization using optimal feedback control.

In the coming sections we will use linear and nonlinear dimensionality reduction methods in extracting the synergies.

2.4 Dimensionality reduction methods for extracting synergies

In the previous section, we listed different methods used to extract the synergies. In this section these methods were limited to dimensionality reduction methods as these are of relevance to this chapter.

Based on the principal component analysis, Jerde et al. (Jerde et al. (2003)) found support for the existence of postural synergies of angular configuration. The shape of human hand can be predicted using a reduced set of variables and postural synergies. Similarly, Santello et al. (1998) showed that a small number of postural synergies were sufficient to describe how human subjects grasped a large set of different objects. Moreover, Mason et al. (2001) used singular value decomposition (SVD) to demonstrate that a large number of hand postures during reach-to-grasp can be constructed by a small number of principal components or eigen postures.

With PCA, Braido & Zhang (2004) examined the temporal co-variation between finger-joint angles. Their results supported the view that the multi-joint acts of the hand are subject to stereotypical motion patterns controlled via simple kinematic synergies. In the above mentioned study of eigen postures, Mason et al. (2001) also investigated the temporal evolutions of the eigen postures and observed similar kinematic synergies across subjects and grasps. In addition, kinematic synergies have been observed in the spatiotemporal coordination between thumb and index finger movements and co-ordination of tip-to-tip finger movements (Cole & Abbs (1986)).

Another concept of synergies was proposed by d'Avella et al. (2003). Although their work was not directly related to the hand movements, they investigated the muscle synergies of frogs during a variety of motor behaviors such as kicking. Using a gradient descent method, they decomposed the muscle activities into linear combinations of three task-independent time-varying synergies. They also observed that these synergies were very much related to movement kinematics and that similarities existed between synergies in different tasks.

3. Preparing the hand kinematics for dimensionality reduction

In this section, we first recorded the joint angles when ten subjects participated in an experiment of reaching and grasping tasks while wearing a dataglove. Then we transformed the recorded joint angles into joint angular velocities and further preprocessed it to prepare datasets to be used as inputs to the dimensionality reduction methods.
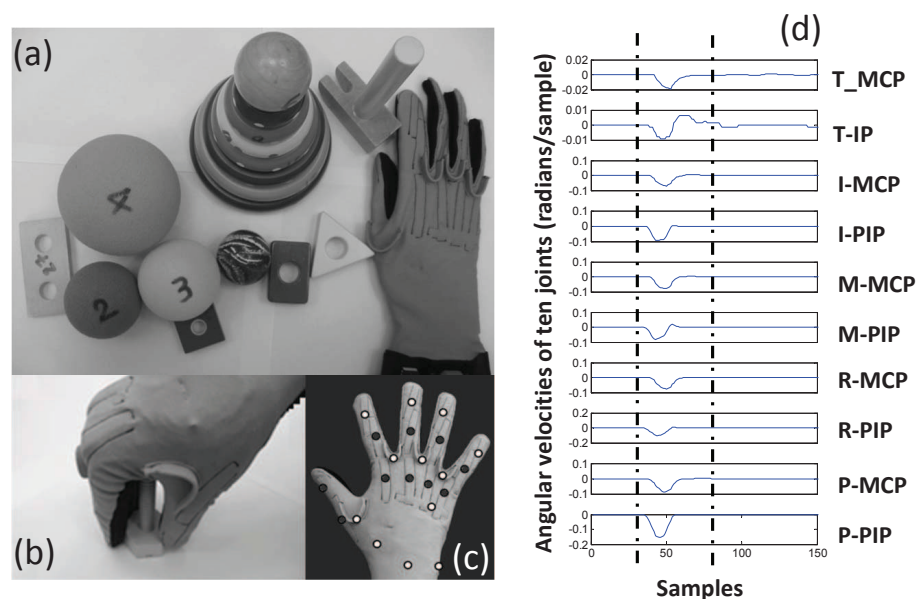

Fig. 2. (a) Objects grasped by the subjects. (b) An example of a rapid grasp of a wooden toy nut. (c) Sensors of the CyberGlove used for analysis (dark circles.) (d) A sample of rapid movement profile (finger-joint-angular-velocity profile). Onset and end of movements are marked in the figure. Abbreviations: T, thumb; I, index finger; M, middle finger; R, ring finger; P, pinky; CMC, carpometacarpal joint; MCP, metacarpophalangeal joint; IP, interphalangeal joint; PIP, proximal interphalangeal joint; DIP, distal interphalangeal joint. The dotted lines showed the onset and end of movement.

3.1 Experiment

The experimental setup consisted of a right-handed CyberGlove (CyberGlove Systems LLC, San Jose, CA, USA) equipped with 22 sensors which can measure angles at all the finger joints. For the purpose of reducing computational burden, in this study we only considered 10 of the sensors which correspond to the metacarpophalangeal (MCP) and interphalangeal (IP) joints of the thumb and the MCP and proximal interphalangeal (PIP) joints of the other four fingers as shown in Fig. 2(c). These ten joints can capture most characteristics of the hand in grasping tasks.
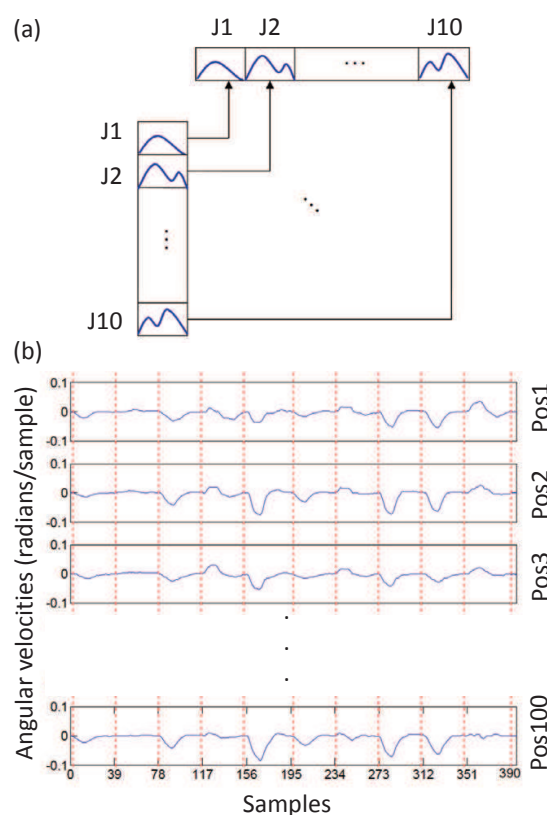
Fig. 3. Angular-velocity matrix. (a) Cascading angular-velocity profiles of 10 joints to form a row of the angular-velocity matrix. (b) Each row of the angular-velocity matrix represents a grasping task. In each row the angular-velocity profiles of 10 joints are separated by dotted red lines. Hundred such tasks put together is an angular-velocity matrix.

A typical task consisted of grasping the objects of various shapes and sizes as shown in Fig. 2(a). Objects (wooden and plastic) of different shapes (spheres, circular discs, rectangles, pentagons, nuts, and bolts) and different dimensions were used in the grasping tasks and were selected based on two strategies. One was gradually increasing sizes of similar shaped objects, and the other was using different shapes. Start and stop times of each task were signaled by computer-generated beeps. In each task, the subject was in a seated position, resting his/her right hand at a corner of a table and upon hearing the beep, grasped the object placed on the table. At the time of the start beep hand was in rest posture, and then the subject grasped the object and held it until the stop beep. Between the grasps, there was enough time for the subjects to avoid the affects due to fatigue on succeeding tasks. The experiment was split into two phases, training phase and testing phase, the difference in these two being the velocity of grasps and types of grasps.

3.2 Training

In the training phase, subjects were instructed to rapidly grasp 50 objects, one at a time. This was repeated for the same 50 objects, and thus the whole training phase obtained 100 rapid grasps. Only these 100 rapid grasps were used in extracting synergies.

3.3 Testing

In the testing phase, subjects were instructed to grasp the above 50 objects naturally (slower than the rapid grasps) then repeat the same again. So far the tasks involved only grasping action. To widen the scope of applicability of the synergies, subjects were also asked to pose 36 American Sign Language (ASL) postures. Here subjects started from an initial posture and stopped at one ASL posture. These postures consisted of 10 numbers (0-9) and 26 alphabets (A-Z). Note that these movements are different from grasping tasks. This is the testing phase which consisted of 100 natural grasps and 36 ASL postural movements. The synergies were derived from the hand movements collected in the training phase using linear and nonlinear dimensionality reduction methods. Then they were used in the reconstruction of movements collected during the testing phase.

3.4 Preprocessing

After obtaining the joint angles at various times from the rapid grasps, angular velocities were calculated. These angular velocities were filtered from noise. Only the relevant projectile movement (about 0.45 second or 39 samples at a sampling rate of 86 Hz) of the entire angular-velocity profile was preserved and the rest was truncated (Fig. 2(d)).

Next an angular-velocity matrix, denoted V, was constructed for each subject. Angular-velocity profiles of the 10 joints corresponding to one rapid grasp were cascaded such that each row of the angular-velocity matrix represented one movement in time. The matrix consisted of 100 rows and $39 \times 10 = 390$ columns:

$$V = \begin{bmatrix} v_1^1(1) & \cdots & v_1^1(39) & \cdots & v_{10}^1(1) & \cdots & v_{10}^1(39) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ v_1^g(1) & \cdots & v_1^g(39) & \cdots & v_{10}^g(1) & \cdots & v_{10}^g(39) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ v_1^{100}(1) & \cdots & v_1^{100}(39) & \cdots & v_{10}^{100}(1) & \cdots & v_{10}^{100}(39) \end{bmatrix} \quad (1)$$

where $v_i^g(t)$ represents the angular velocity of joint i ($i = 1, ..., 10$) at time t ($t = 1, ..., 39$) in the g-th rapid-grasping task ($g = 1, ..., 100$). An illustration of this transformation was shown in the Fig. 3.

4. Linear dimensionality reduction methods

In this section we derived synergies using two unique linear dimensionality reduction methods, namely, PCA and MDS. The angular-velocity matrix computed in preprocessing was used as input to these methods. Linear methods are easy to use and demand less computational power when compared to nonlinear methods, hence this first exercise.

4.1 Principal component analysis

The winning advantage of PCA is less time for computation and equally effective results when compared to gradient descent methods(Vinjamuri et al. (2007)). PCs are essentially the most commonly used patterns across the data. In this case, PCs are the synergies which are most commonly used across different movements. Moreover these PCs when

graphically visualized revealed anatomical implications of physiological properties of human hand prehension(Vinjamuri et al. (2010a;b)).

There are several ways to implement PCA. Two most widely used methods were shown below. First method has three steps: (1) Subtract mean from the data (2) Calculate covariance matrix (3) Compute eigen values and eigen vectors of covariance matrix. Principal components are eigen vectors. Second method uses singular value decomposition (SVD). Third method is a function readily available in Statistics Tool Box of MATLAB which essentially implements first method.

```
load hald; %Load sample dataset in MATLAB
Data = ingredients';
%Fetch dimensions of the data (M Dimensions x N observations)
[M,N] = size(Data);
MeanofData = mean(Data,2); %Calculate Mean of the Data
Data = Data - repmat(MeanofData,1,N); %Subtract Mean from the Data
Covariance = 1 / (N-1) * Data * Data'; %Calculate the covariance of the Data
[pc, latent] = eig(Covariance); %Find the eigenvectors and eigenvalues
latent = diag(latent); %Take only diagonal elements that are eigen values
[rd, rindices] = sort(-1*latent); %Sort them in decreasing order
latent = latent(rindices); %Extract eigen values corresponding to indices
pc = pc(:,rindices); %Extract principal components corresponding to indices
score = (pc' * Data)'; %Project the Data into PC space

load hald; %Load sample dataset in MATLAB
Data = ingredients';
%Fetch dimensions of the data (M Dimensions x N observations)
[M,N] = size(Data);
MeanofData = mean(Data,2); %Calculate Mean of the Data
Data = Data - repmat(MeanofData,1,N); %Subtract Mean from the Data
Y = Data' / sqrt(N-1); %Normalized Data
[u,S,pc] = svd(Y); %Peform SVD
S = diag(S); %Extract diagonal elements corresponding to eigen values
latent = S .* S; %Calculate the eigen values
score = (pc' * Data)'; %Project the Data into PC space

load hald; %Load sample dataset in MATLAB
[pc,score,latent,tsquare] = princomp(ingredients); %Peform PCA
```

Here PCA using SVD Jolliffe (2002) was performed on the angular-velocity matrix V of each subject:

$$V = U\Sigma S \tag{2}$$

where U is a 100-by-100 matrix, which has orthonormal columns so that $U'U = I_{100\times100}$ (100-by-100 identity matrix); S is a 100-by-390 matrix, which has orthonormal rows so that $SS' = I_{100\times100}$; and Σ is a 100-by-100 diagonal matrix: $\text{diag}\{\lambda_1, \lambda_2, ..., \lambda_{100}\}$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{100} \geq 0$. Matrix V can be approximated by another matrix $\tilde{V}$ with reduced rank m by replacing Σ with Σ_m, which contains only the m largest singular values, i.e., $\lambda_1, ..., \lambda_m$ (the other singular values are replaced by zeros). The approximation matrix $\tilde{V}$ can be written in a more compact form:

$$\tilde{V} = U_m \, \text{diag}\{\lambda_1, ..., \lambda_m\} \, S_m \tag{3}$$

where U_m is a 100-by-m matrix containing the first m columns of U and S_m is a m-by-390 matrix containing the first m rows of S. Denoting $W = U_m \, \text{diag}\{\lambda_1, ..., \lambda_m\}$, we have

$$V \approx \tilde{V} = W S_m. \tag{4}$$

Then each row of S_m is called a *principal component* (PC), and W is called the weight matrix.

For easy comparison, let us name the elements of S_m in a way similar to (1):

$$S_m \equiv \begin{bmatrix} s_1^1(1) & \cdots & s_1^1(39) & \cdots & s_{10}^1(1) & \cdots & s_{10}^1(39) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ s_1^m(1) & \cdots & s_1^m(39) & \cdots & s_{10}^m(1) & \cdots & s_{10}^m(39) \end{bmatrix} \tag{5}$$

and name the elements of W in the following way:

$$W = \begin{bmatrix} w_1^1 & \cdots & w_m^1 \\ \vdots & \vdots & \vdots \\ w_1^g & \cdots & w_m^g \\ \vdots & \vdots & \vdots \\ w_1^{100} & \cdots & w_m^{100} \end{bmatrix}. \tag{6}$$

According to (4), each row of V can be approximated by a linear combination of m PCs, and according to (4), (1), (5), and (6), we have

$$v_i^g(t) \approx \sum_{j=1}^{m} w_j^g s_i^j(t) \tag{7}$$

for $i = 1, ..., 10$, $g = 1, ..., 100$, and $t = 1, ..., 39$.

Thus the above SVD procedure has found a solution to the synergy-extraction problem: The angular-velocity profiles (obtained by rearranging all joints row-wise for the PCs)

$$\begin{bmatrix} s_1^j(1) & \cdots & s_1^j(39) \\ s_2^j(1) & \cdots & s_2^j(39) \\ \vdots & \vdots & \vdots \\ s_{10}^j(1) & \cdots & s_{10}^j(39) \end{bmatrix}, \; j = 1, ..., m$$

can be viewed as synergies. According to (4) or (7), these synergies can serve as "building blocks" to reconstruct joint-angular-velocity profiles of hand movements.

To decide m, the number of PCs or synergies that we want to use in reconstruction of the testing movements, we consider the accuracy of approximation in (4) or (7). The approximation accuracy can be measured by an index defined as

$$\frac{\lambda_1^2 + \lambda_2^2 + \cdots + \lambda_m^2}{\lambda_1^2 + \lambda_2^2 + \cdots + \lambda_{100}^2}.$$
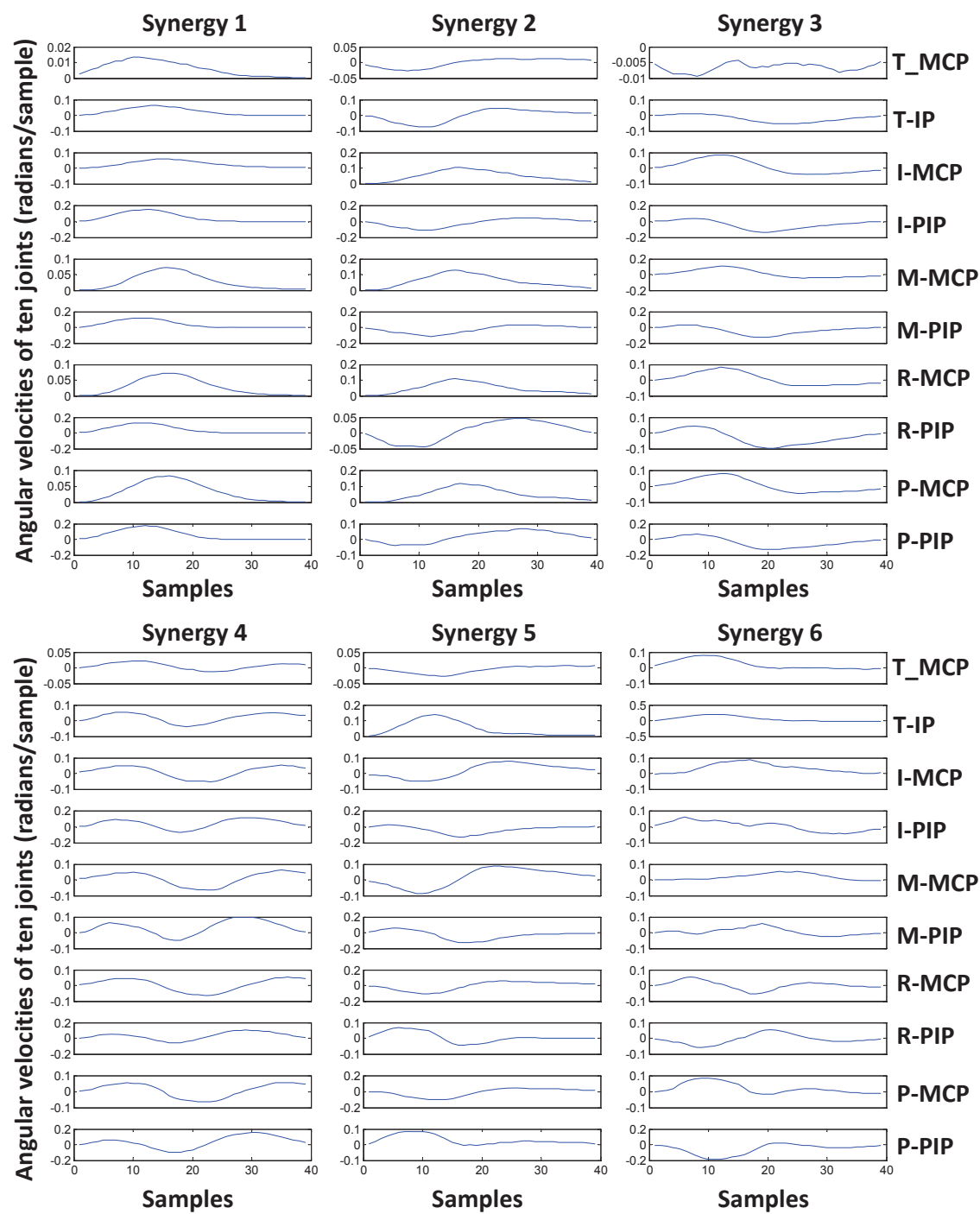
Fig. 4. Six kinematic synergies obtained for subject 1 using PCA. Each synergy is about 0.45 s in duration (39 samples at 86 Hz). Abbreviations: T, thumb; I, index finger; M, middle finger; R, ring finger; P, pinky finger; MCP, metacarpophalangeal joint; IP, interphalangeal joint; PIP, proximal IP joint.

The larger this index is, the closer the approximation is. This index also provides indication of the fraction of total variance of the data matrix accounted by the PCs. To ensure satisfactory approximation, the index should be greater than some threshold. In this study, we used 95% as the threshold (a commonly used threshold Jolliffe (2002)) to determine the number of PCs or synergies (i.e. m). With this threshold we found the six synergies can account for 95% of variance in the postures. Fig. 4 shows six kinematic synergies obtained for subject 1 using PCA.

4.2 Multidimensional scaling

Classical Multidimensional Scaling (MDS) can still be grouped under linear methods. This was introduced here to the reader to give a different perspective of dimensionality reduction in a slightly different analytical approach when compared to PCA discussed previously. The two methods PCA and MDS are unique as they perform dimensionality reduction in different ways. PCA operates on covariance matrix where as MDS operates on distance matrix. In MDS, a Euclidean distance matrix is calculated from the original matrix. This is nothing but a pairwise distance matrix between the variables in the input matrix. This method tries to preserve these pairwise distances in a low dimensional space, thus allowing for dimensionality reduction and preserving the inherent structure of the data simultaneously. PCA and MDS were compared using a simple example in MATLAB below.

```
load hald; %Load sample dataset in MATLAB
[pc,score,latent,tsquare] = princomp(ingredients); %Peform PCA

D = pdist(ingredients); %Calculate pairwise distances between ingredients
[Y, e] = cmdscale(D); %Perform Classical MDS
```

score in PCA represented the data that was projected in the PC space. Compare this to Y calculated in MDS. These are same. Similarly in place of sample dataset when the posture matrix V was used as input to MDS, it yielded the same synergies as PCA. This was introduced here because we build upon this method for the nonlinear methods coming up in the next section.

5. Nonlinear dimensionality reduction methods

So far, we have investigated the use of linear dimensionality reduction methods (PCA and MDS) in extracting synergies. In this section we used nonlinear dimensionality reduction methods for the same purpose. The motivation to explore nonlinear methods was that physiologists who studied motor control have propounded that there were inherent nonlinearities in the human motor system. By using nonlinear methods we could probably achieve improved precision in reconstruction of natural movements. The nonlinear methods applied in this chapter are Isomap, local linear embedding (LLE), and kernel PCA (kPCA). The first two methods Isomap and LLE, are built on the framework of classical multidimensional scaling discussed in the previous section. kPCA is built on the framework of PCA.

5.1 Isomap

Isomap is similar to PCA and MDS. Although Isomap does linear estimations in the data point neighborhoods, the synergies extracted are nonlinear because these small neighborhoods are stitched together without trying to maintain linearity.

The following were the steps involved in estimating nonlinear synergies using Isomap:

1. Define neighbors for each data point

2. Find D, a matrix of inter-point distances

3. Find eigenvectors of $\tau(\text{D})$, where $\tau(\text{D}) = -HSH/2$, $S_{ij} = (D_{ij})^2$ and $H_{ij} = d_{ij} - 1/N$, where N is number of data points and d is Kronecker delta function.

In PCA we estimated the eigen values and eigen vectors of covariance of the data. Similarly, here, we took a nonlinear approach to preserve inter-point distances on the manifold. The matrix D is similar to covariance matrix in PCA. D can actually be thought of as the covariance matrix in higher dimensions. Since in an N-dimensional space, the dimensions are the data points, the covariance for a particular pair of dimensions is the distance between the data points that define those dimensions.

Although this method looks linear like PCA, the source of nonlinearity is the method in which inter-point distances are calculated. For Isomap, we do not use the Euclidean distances between the points. If we use, it becomes classical MDS discussed in previous section. Rather, we use those distances only for points considered neighbors. The rest of inter-point distances are calculated by finding the shortest path through the graph on the manifold using Floyd's algorithm (Tenenbaum et al. (2000)). The goal of the Isomap is to preserve the geodesic distances rather than the euclidian distances. Geodesic distances are calculated by moving along the approximate nonlinear manifold with given data point and interpolation between them.

We used drtoolbox in MATLAB by van der Maaten et al. (2009) to perform Isomap on the angular-velocity matrix to extract synergies. Fig. 5showed the top six synergies extracted using this method. Similar to PCA, all the nonlinear methods also yield the nonlinear synergies in descending order of their significance. The synergies extracted using this method had more submovements when compared to those in PCA.

5.2 Local Linear Embedding

Locally Linear Embedding (LLE) as the name suggests, tries to find a nonlinear manifold by stitching together small linear neighborhoods (Roweis & Saul (2000)). This is very similar to Isomap. The difference between the two algorithms is in how they do the stitching. Isomap does this by doing a graph traversal by preserving geodesic distances while LLE does it by finding a set of weights that perform local linear interpolations that closely approximate the data.

The following were the steps involved in estimating nonlinear synergies using LLE:

1. Define neighbors for each data point

2. Find weights that allow neighbors to interpolate original data accurately

3. Given those weights, find new data points that minimize interpolation error in lower dimensional space

We used drtoolbox in MATLAB by van der Maaten et al. (2009) to perform LLE on the angular-velocity matrix to extract synergies. Fig. 6showed the top six synergies extracted using this method. Similar to Isomap, in this method also we found more submovements in synergies than those from PCA.
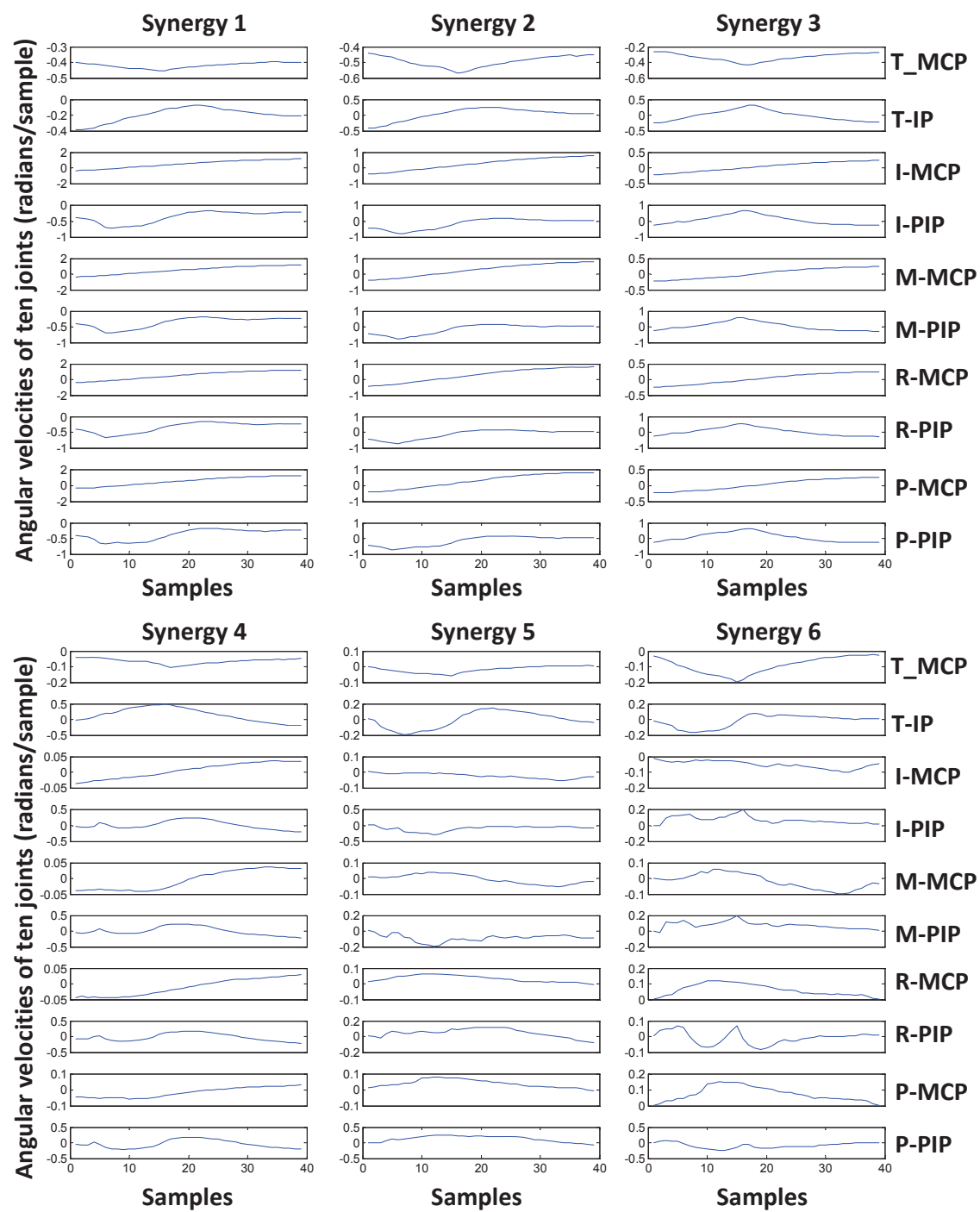
Fig. 5. Six kinematic synergies obtained for subject 1 using Isomap. Each synergy is about 0.45 s in duration (39 samples at 86 Hz). Abbreviations: T, thumb; I, index finger; M, middle finger; R, ring finger; P, pinky finger; MCP, metacarpophalangeal joint; IP, interphalangeal joint; PIP, proximal IP joint.
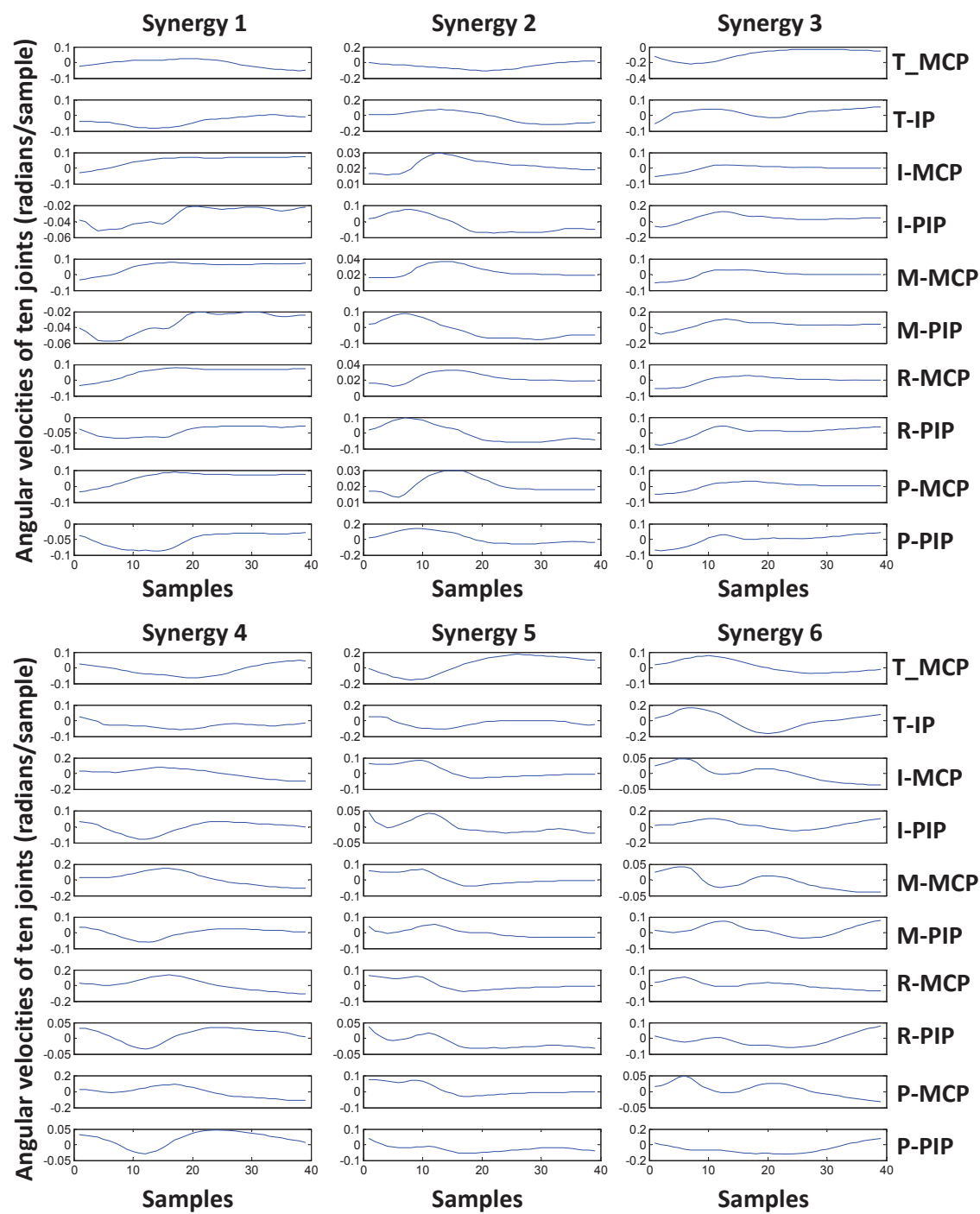
Fig. 6. Six kinematic synergies obtained for subject 1 using LLE. Each synergy is about 0.45 s in duration (39 samples at 86 Hz). Abbreviations: T, thumb; I, index finger; M, middle finger; R, ring finger; P, pinky finger; MCP, metacarpophalangeal joint; IP, interphalangeal joint; PIP, proximal IP joint.

5.3 Kernel PCA

Kernel PCA (kPCA) is an extension of PCA in a high-dimensional space (Scholkopf et al. (1998)). A high-dimensional space is first constructed by using a kernel function. Instead of directly doing a PCA on the data, the kernel based high dimensional feature space is used as input. In this chapter, we have used a gaussian kernel function. Kernel PCA computes the principal eigenvectors of the kernel matrix, rather than those of the covariance matrix. A kernel matrix is similar to the inner product of the data points in the high dimensional space that is constructed using the kernel function. The application of PCA in the kernel space provides Kernel PCA the property of constructing nonlinear mappings. We used drtoolbox in MATLAB by van der Maaten et al. (2009) to perform kPCA on the angular-velocity matrix to extract synergies. Fig. 7showed the top six synergies extracted using this method. These synergies were similar to those obtained from PCA.

6. Reconstruction of natural and ASL movements

The synergies extracted from linear and nonlinear dimensionality reduction methods were used in reconstruction of natural movements. l_1-norm minimization was used to reconstruct natural and ASL movements from the extracted synergies. This method with illustrations was already presented in Vinjamuri et al. (2010a). We have included a brief explanation here for readability and for the sake of completeness. We ask the readers to refer Vinjamuri et al. (2010a) for further details.

Briefly, these were the steps involved in l_1-norm minimization algorithm that was used for reconstruction of natural and ASL movements. Let us assume for a subject m synergies were obtained. The duration of the synergies is t_s samples ($t_s = 39$ in this study). Consider an angular-velocity profile of the subject, $\{\mathbf{v}(t), t = 1, ..., T\}$, where T ($T = 82$ in this study) represents the movement duration (in samples). This profile can be rewritten as a row vector, denoted $\mathbf{v}_{\text{row}}$:

$$\mathbf{v}_{\text{row}} = [v_1(1), ..., v_1(T), ..., v_{10}(1), ..., v_{10}(T)].$$

Similarly, a synergy $\mathbf{s}^j(\cdot)$ can be rewritten as the following row vector:

$$[s_1^j(1), ..., s_1^j(t_s), 0, ..., 0, ..., s_{10}^j(1), ..., s_{10}^j(t_s), 0, ..., 0].$$

We add $T - t_s$ zeros after each $s_i^j(t_s)$ ($i = 1, ..., 10$) in the above vector in order to make the length of the vector the same as that of $\mathbf{v}_{\text{row}}$. If the synergy is shifted in time by t_{jk} ($t_{jk} \leq T - t_s$) samples, then we obtain the following row vector:

$$[0, ..., 0, s_1^j(1), ..., s_1^j(t_s), 0, ..., 0, ...,$$
$$0, ..., 0, s_{10}^j(1), ..., s_{10}^j(t_s), 0, ..., 0]$$

with t_{jk} zeros added before each $s_i^j(1)$ and $T - t_s - t_{jk}$ zeros added after each $s_i^j(t_s)$.

Then we construct a matrix as shown in Fig. 8 consisting of the row vectors of the synergies and all their possible shifts with $1 \leq t_{jk} \leq T - t_s$.

With the above notation, we are trying to achieve a linear combination of synergies that can reconstruct the velocity profiles as in the following equation.

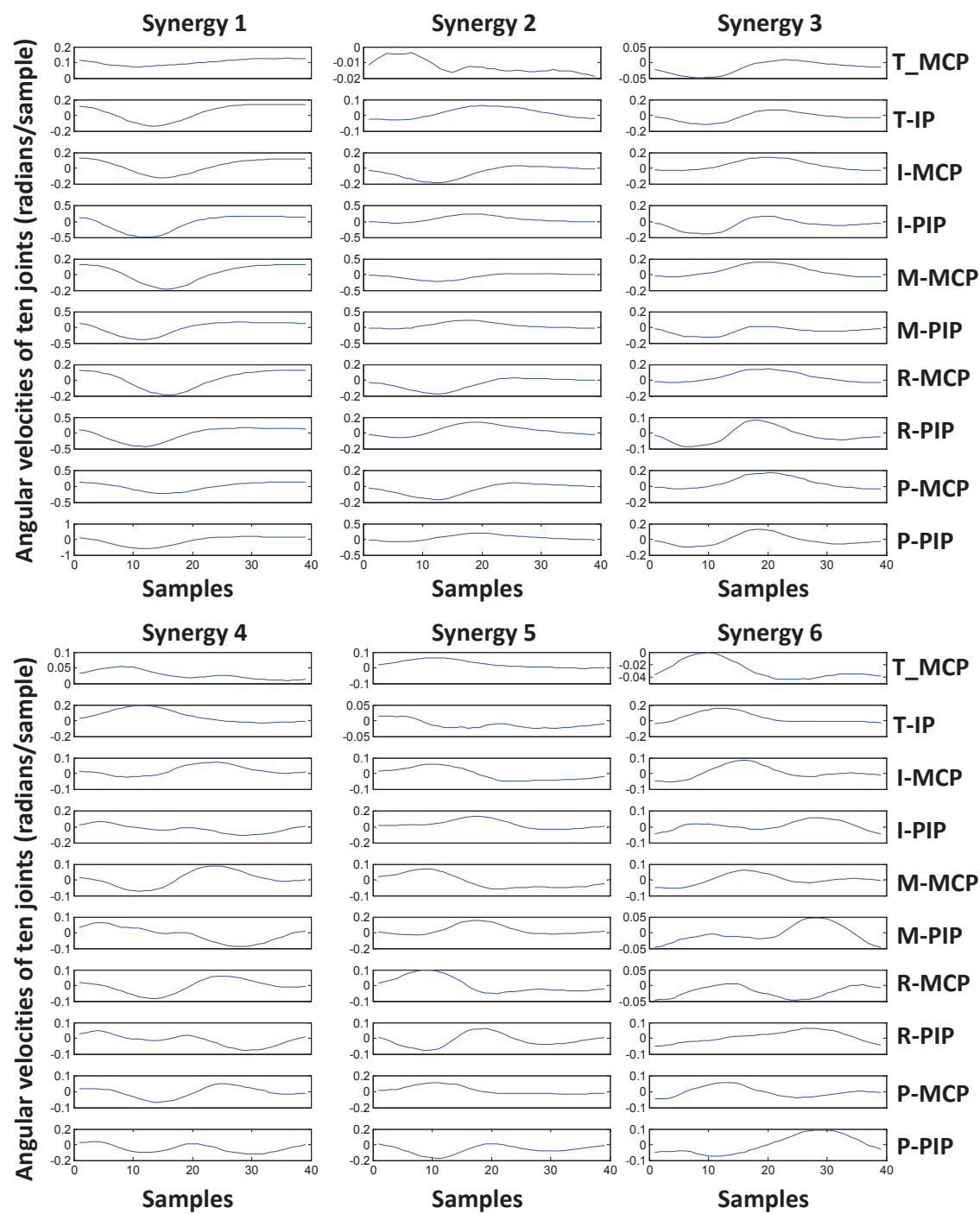$$\mathbf{v}_{\text{row}} = \mathbf{c}B \tag{8}$$

Fig. 7. Six kinematic synergies obtained for subject 1 using kPCA. Each synergy is about 0.45 s in duration (39 samples at 86 Hz). Abbreviations: T, thumb; I, index finger; M, middle finger; R, ring finger; P, pinky finger; MCP, metacarpophalangeal joint; IP, interphalangeal joint; PIP, proximal IP joint.
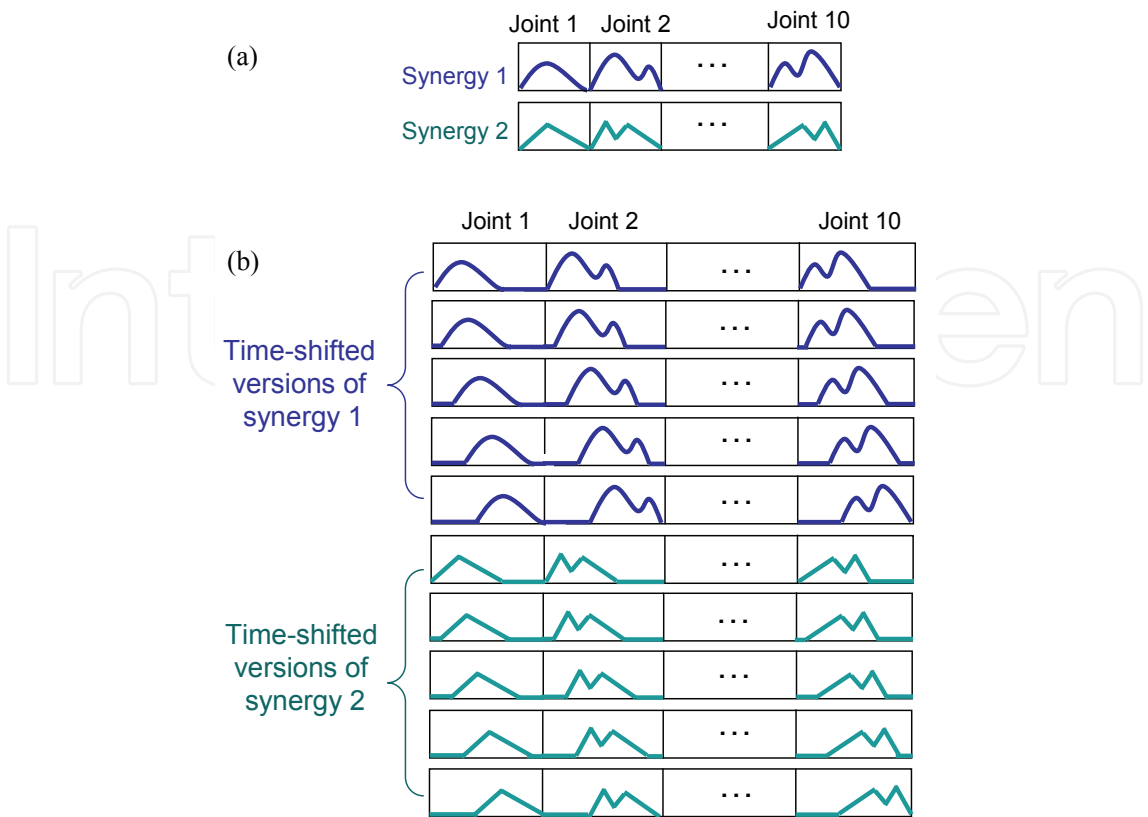
Fig. 8. (a) Synergies were rearranged as row vectors (b) A template matrix was formed by adding the time-shifted versions of synergies along the rows. Adapted from Vinjamuri et al. (2010a).

where $\mathbf{c}$ denotes

$$[0, ..., c_{11}, ..., 0, ..., c_{1K_1}, ..., 0, ..., c_{m1}, ..., 0, ..., c_{mK_m}, ..., 0]$$

with nonzero values c_{jk} appearing at the $(T - t_s + 1)(j - 1) + t_{jk}$-th elements of $\mathbf{c}$. The matrix B (shown in Fig. 8(b)) can be viewed as a bank or library of template functions with each row of B as a template. This bank can be overcomplete and contain linearly dependent subsets. Therefore, for a given movement profile $\mathbf{v}_{\text{row}}$ and an overcomplete bank of template functions B, there exists an infinite number of $\mathbf{c}$ satisfying (8).

We hypothesize that the strategy of central nervous system for dimensionality reduction in movement control is to use a small number of synergies and a small number of recruitments of these synergies for movement generation. Therefore, the coefficient vector $\mathbf{c}$ in (8) should be sparse, i.e., having a lot of zeros and only a small number of nonzero elements. Therefore, we seek the sparsest coefficient vector $\mathbf{c}$ such that $\mathbf{c}B = \mathbf{v}_{\text{row}}$.

The following was optimization problem that was used in selection of synergies in reconstruction of a particular movement.

$$\text{Minimize} \quad \| \mathbf{c} \|_1 + \frac{1}{\lambda} \| \mathbf{c}B - \mathbf{v}_{\text{row}} \|_2^2 \tag{9}$$

where $\| \cdot \|_2$ represents the l_2 norm or Euclidean norm of a vector and λ is a regulation parameter.

Using the above optimization algorithm, the synergies extracted from four methods (PCA, Isomap, LLE, and kPCA) were used in reconstruction of natural movements and ASL postural movements. The reconstruction errors were calculated using the methods in Vinjamuri et al. (2010a). Figures 9 and 10 showed the comparison between the four dimensionality reduction methods. Fig. 9 showed the reconstruction errors for 100 natural movements and Fig. 10 showed the reconstruction errors for 36 ASL postural movements for all four methods. It is observed that PCA still has the best overall performance when compared with the novel nonlinear methods. Fig. 11 showed one of the best reconstructions by all four methods. Tables 1 and 2 summarize the reconstruction results obtained for all ten subjects.
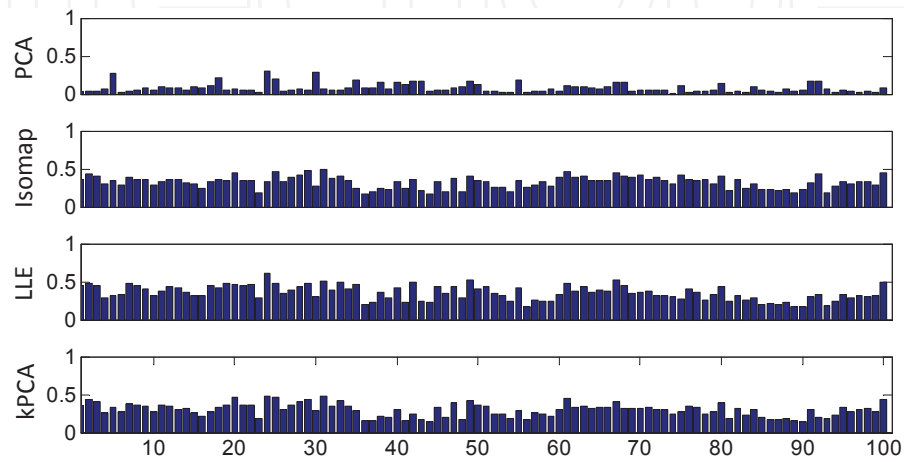

Fig. 9. The reconstruction errors for 100 natural movements with four dimensionality reduction methods PCA, Isomap, LLE, and kPCA. PCA outperformed other three nonlinear methods.
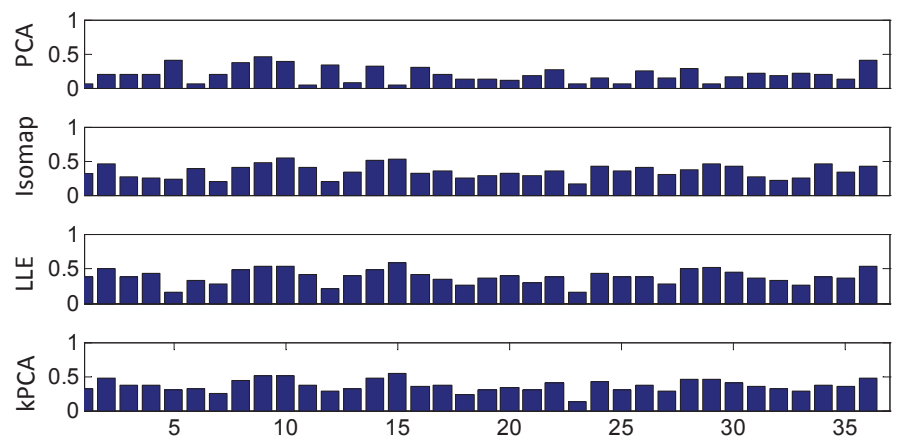

Fig. 10. The reconstruction errors for 36 ASL postural movements with four dimensionality reduction methods PCA, Isomap, LLE, and kPCA. All methods performed poorly in reconstruction of ASL movements when compared to natural movements. PCA performed better than other three nonlinear methods.

7. Summary

In this chapter we applied linear and nonlinear dimensionality reduction methods to extract movement primitives or synergies from rapid reach and grasp movements. We then used

these synergies to reconstruct natural movements that were similar to activities of daily living. To broaden the applicability of synergies we also tested them on ASL postural movements which are different movements when compared to natural reach and grasp movements. We employed four types of dimensionality reduction methods: (1) PCA and MDS (2) Isomap (3) LLE (4) kPCA. PCA is a well known linear dimensionality reduction method. Two widely used PCA implementations (Covariance and SVD) were presented and relevant MATLAB codes were provided. Classical MDS is very similar to PCA but operates on a distance matrix. This was introduced as Isomap and LLE both work on a similar framework. Isomap and LLE are both neighborhood graph based nonlinear dimensionality reduction methods. The difference between Isomap and LLE is that the former is a global dimensionality reduction technique where as the former as the name suggests is a local linear technique. kPCA is similar to Isomap as it is a global dimensionality reduction method, but the uniqueness of the method is in using kernel tricks to transform the input data to higher dimensional space.

Subject	PCA	Isomap	LLE	kPCA
1	0.0797 ± 0.0596	0.3240 ± 0.0769	0.3544 ± 0.0954	0.2954 ± 0.0851
2	0.0845 ± 0.0781	0.3780 ± 0.0519	0.3998 ± 0.0765	0.3444 ± 0.0658
3	0.0678 ± 0.0396	0.3180 ± 0.0554	0.3234 ± 0.0834	0.2654 ± 0.0767
4	0.0590 ± 0.0432	0.3680 ± 0.0663	0.3879 ± 0.0645	0.3214 ± 0.0873
5	0.0997 ± 0.0831	0.2640 ± 0.0734	0.2950 ± 0.0790	0.2230 ± 0.0667
6	0.0661 ± 0.0976	0.3240 ± 0.0545	0.3754 ± 0.0531	0.2875 ± 0.0931
7	0.0598 ± 0.0542	0.4140 ± 0.0787	0.4344 ± 0.0632	0.3895 ± 0.0696
8	0.0732 ± 0.0348	0.3540 ± 0.0989	0.3954 ± 0.0854	0.3123 ± 0.0555
9	0.0814 ± 0.0212	0.2290 ± 0.0823	0.3100 ± 0.0991	0.2074 ± 0.0651
10	0.0883 ± 0.0443	0.2490 ± 0.0665	0.2780 ± 0.0799	0.2189 ± 0.0799

Table 1. Mean reconstruction errors ($\pm$ standard deviation) in natural movements

Subject	PCA	Isomap	LLE	kPCA
1	0.2449 ± 0.0986	0.3480 ± 0.0988	0.3867 ± 0.1052	0.3654 ± 0.0865
2	0.2782 ± 0.0915	0.3280 ± 0.0891	0.3661 ± 0.0963	0.3291 ± 0.0745
3	0.2677 ± 0.0866	0.3110 ± 0.0764	0.3527 ± 0.0881	0.3321 ± 0.0666
4	0.2321 ± 0.0645	0.3450 ± 0.0662	0.3667 ± 0.1123	0.3540 ± 0.0831
5	0.2899 ± 0.0713	0.3780 ± 0.0843	0.3889 ± 0.0792	0.3624 ± 0.0934
6	0.2569 ± 0.0456	0.3200 ± 0.0894	0.3457 ± 0.0856	0.3654 ± 0.0678
7	0.2319 ± 0.0521	0.3890 ± 0.0663	0.3999 ± 0.0996	0.3889 ± 0.0610
8	0.2789 ± 0.0486	0.3671 ± 0.0713	0.3897 ± 0.1152	0.3764 ± 0.0885
9	0.2999 ± 0.0816	0.3335 ± 0.0465	0.3867 ± 0.1672	0.3478 ± 0.0585
10	0.2569 ± 0.0936	0.3146 ± 0.0548	0.3547 ± 0.0082	0.3354 ± 0.0635

Table 2. Mean reconstruction errors ($\pm$ standard deviation) in ASL postural movements

Thus the reader was given the opportunity to sample different varieties of dimensionality reduction methods. Quantitative and qualitative comparison of the results obtained from reconstruction follows but the verdict is that PCA outperformed the nonlinear methods employed in this chapter.

The results from the reconstructions reveal that nonlinear techniques do not outperform the traditional PCA for both natural movements as well as ASL postural movements. The
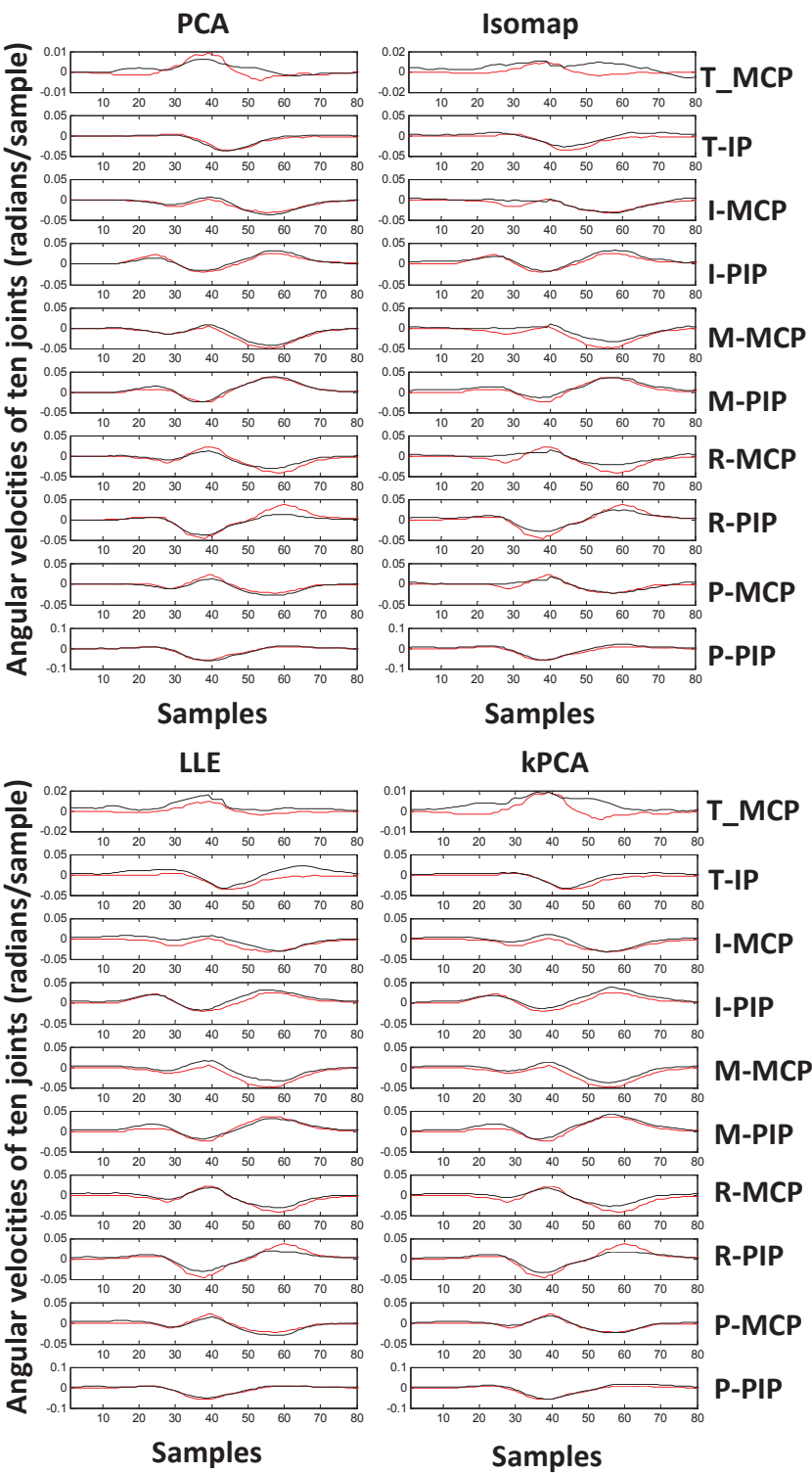
Fig. 11. An example reconstruction (in black) of a natural movement (in red) for task 24 when subject 1 was grasping an object.

reconstruction errors were more for ASL postural movements when compared to those of natural movements for all methods. The reconstruction errors were in general larger for Isomap and LLE when compared with PCA and kPCA and of course PCA had outstanding performance for more than 90% of the tasks. van der Maaten et al. (van der Maaten et al. (2009)) also found that nonlinear methods performed well on specific data sets but could not perform better than PCA for real world tasks. For example, for the Swiss roll data set that contains points that lie on a spiral like two dimensional manifold within a three dimensional space, several nonlinear techniques such as Isomap, LLE were able to find the two dimensional planar embedding, but linear techniques like PCA failed to find so. The reasons for two nonlinear methods Isomap and LLE to perform poorly in this study, might be that they relied neighborhood graphs. Moreover LLE might have been biased to local properties that do not necessarily follow the global properties of high dimensional data. It was surprising to see that Isomap, being a global dimensionality reduction technique performed poorly when compared to LLE for natural movements. kPCA performed better than Isomap and LLE, but kPCA does suffer from the limitation of selection of ideal kernel. The selection of gaussian kernel in this study might not be favorable in extracting the kinematic synergies in this study. In conclusion, although there are numerical advantages and disadvantages with both linear and nonlinear dimensionality reduction methods, PCA seemed to generalize and perform well on the real world data.

8. Acknowledgments

9. References

Bernabucci, I., Conforto, S., Capozza, M., Accornero, N., Schmid, M. & D'Alessio, T. (2007). A biologically inspired neural network controller for ballistic arm movements, *Journal of NeuroEngineering and Rehabilitation* 4(33): 1–17.

Bernstein, N. (1967). *The Co-ordination and Regulation of Movements*, Pergamon Press, Oxford, UK.

Braido, P. & Zhang, X. (2004). Quantitative analysis of finger motion coordination in hand manipulative and gestic acts, *Human Movement Science* 22: 661–678.

Cole, K. J. & Abbs, J. H. (1986). Coordination of three-joint digit movements for rapid finger-thumb grasp, *Journal of Neurophysiology* 55(6): 1407–1423.

d'Avella, A., Portone, A., Fernandez, L. & Lacquaniti, F. (2006). Control of fast-reaching movements by muscle synergy combinations, *Journal of Neuroscience* 26(30): 7791–7810.

d'Avella, A., Saltiel, P. & Bizzi, E. (2003). Combinations of muscle synergies in the contruction of a natural motor behavior, *Nature Neuroscience* 6(3): 300–308.

Flash, T. & Hochner, B. (2005). Motor primitives in vertebrates and invertebrates, *Current Opinion in Neurobiology* 15(6): 660–666.

Grinyagin, I. V., Biryukova, E. V. & Maier, M. A. (2005). Kinematic and dynamic synergies of human precision-grip movements, *Journal of Neurophysiology* 94(4): 2284–2294.

Hauser, H., Neumann, G., Ijspeert, A. J. & Maass, W. (2007). Biologically inspired kinematic synergies provide a new paradigm for balance control of humanoid robots, *Proc. IEEE-RAS 7th Intl. Conf. Humanoid Robots*, Pittsburgh, PA, USA.

Jerde, T. E., Soechting, J. F. & Flanders, M. (2003). Biological constraints simplify the recognition of hand shapes, 50(2): 565–569.

Jolliffe, I. T. (2002). *Principal Component Analysis, 2nd Ed.*, Springer, New York, NY, USA.

Mackenzie, C. L. & Iberall, T. (1994). *The Grasping Hand (Advances in Psychology)*, North-Holland, Amsterdam, Netherlands.

Mason, C. R., Gomez, J. E. & Ebner, T. J. (2001). Hand synergies during reach-to-grasp, *Journal of Neurophysiology* 86(6): 2896–2910.

Poggio, T. & Bizzi, E. (2004). Generalization in vision and motor control, *Nature* 431: 768–774.

Popovic, M. & Popovic, D. (2001). Cloning biological synergies improves control of elbow neuroprostheses, 20(1): 74–81.

Roweis, S. & Saul, L. (2000). Nonlinear dimensionality reduction by locally linear embedding, *Science* 290(5500): 2323–2326.

Santello, M., Flanders, M. & Soechting, J. F. (1998). Postural hand synergies for tool use, *Journal of Neuroscience* 18(23): 10105–10115.

Santello, M., Flanders, M. & Soechting, J. F. (2002). Patterns of hand motion during grasping and the influence of sensory guidance, *Journal of Neuroscience* 22(4): 1426–1435.

Scholkopf, B., Smola, A. . J. & Muller, K.-R. (1998). Nonlinear component analysis as a kernel eigenvalue problem., *Neural Computation* 10(5): 1299–1319.

Tenenbaum, J. . B., Silva, V. D. & Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction, *Science* 290: 2319–2323.

Thakur, P. H., Bastian, A. J. & Hsiao, S. S. (2008). Multidigit movement synergies of the human hand in an unconstrained haptic exploration task, *Journal of Neuroscience* 28(6): 1271–1281.

Todorov, E. & Ghahramani, Z. (2004). Analysis of the synergies underlying complex hand manipulation, *Proc. 26th Annual International Conference of the IEEE EMBS*, San Francisco, CA, USA, pp. 4637–4640.

van der Maaten, L. J. P., Postma, E. O. & van den Herik, H. J. (2009). Dimensionality reduction: A comparative review, *Tilburg University Technical Report, TiCC-TR* 2009-005.

Vinjamuri, R., Crammond, D. J., Kondziolka, D. & Mao, Z.-H. (2009). Extraction of sources of tremor in hand movements of patients with movement disorders, *IEEE Transactions on Information Technology and Biomedicine* 13(1): 49–59.

Vinjamuri, R., Mao, Z.-H., Sclabassi, R. & Sun, M. (2007). Time-varying synergies in velocity profiles of finger joints of the hand during reach and grasp, *Proc. 29th Annual International Conference of the IEEE EMBS*, Lyon, France, pp. 4846–4849.

Vinjamuri, R., Sun, M., Chang, C.-C., Lee, H.-N., Sclabassi, R. J. & Mao, Z.-H. (2010a). Dimensionality reduction in control and coordination of human hand, *IEEE Transactions on Biomedical Engineering* 57(2): 284–295.

Vinjamuri, R., Sun, M., Chang, C.-C., Lee, H.-N., Sclabassi, R. J. & Mao, Z.-H. (2010b). Temporal postural synergies of the hand in rapid grasping tasks, *IEEE Transactions on Information Technology and Biomedicine* 14(4): 986–994.

Vinjamuri, R., Sun, M., Crammond, D., Sclabassi, R. & Mao, Z.-H. (2008). Inherent bimanual postural synergies in hands, *Proc. 30th Annual International Conference of the IEEE EMBS*, Vancouver, British Columbia, Canada, pp. 5093–5096.

Vinjamuri, R., Weber, D., Mao, Z. H., Collinger, J., Degenhart, A., Kelly, J., Boninger, M., Tyler-Kabara, E. & Wang, W. (2011). Towards synergy-based brain-machine interfaces, *IEEE Transactions on Information Technology and Biomedicine* 15(5): 726–736.

Principal Component Analysis

Edited by Dr. Parinya Sanguansat

This book is aimed at raising awareness of researchers, scientists and engineers on the benefits of Principal Component Analysis (PCA) in data analysis. In this book, the reader will find the applications of PCA in fields such as image processing, biometric, face recognition and speech processing. It also includes the core concepts and the state-of-the-art methods in data analysis and feature extraction.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Ramana Vinjamuri, WeiWang, Mingui Sun and Zhi-Hong Mao (2012). Application of Linear and Nonlinear Dimensionality Reduction Methods, Principal Component Analysis, Dr. Parinya Sanguansat (Ed.), ISBN: 978-953-51-0195-6, InTech, Available from: http://www.intechopen.com/books/principal-component-analysis/application-of-linear-and-nonlinear-dimensionality-reduction-methods