



DEGREE PROJECT IN MATHEMATICS,  
SECOND CYCLE, 30 CREDITS  
*STOCKHOLM, SWEDEN 2017*

# **Modeling Natural Human Hand Motion for Grasp Animation**

**JOHANNES JEPPSSON**

**KTH ROYAL INSTITUTE OF TECHNOLOGY  
SCHOOL OF ENGINEERING SCIENCES**



# **Modeling Natural Human Hand Motion for Grasp Animation**

**JOHANNES JEPPSSON**

Degree Projects in Mathematical Statistics (30 ECTS credits)  
Degree Programme in Applied and Computational Mathematics (120 credits)  
KTH Royal Institute of Technology year 2017  
Supervisor at Gleechi: Dan Song  
Supervisor at KTH: Timo Koski  
Examiner at KTH: Timo Koski

*TRITA-MAT-E 2017:53*  
*ISRN-KTH/MAT/E--17/53--SE*

Royal Institute of Technology  
*School of Engineering Sciences*  
**KTH SCI**  
SE-100 44 Stockholm, Sweden  
URL: [www.kth.se/sci](http://www.kth.se/sci)

## Abstract

This report was carried out at Gleechi, a Swedish start-up company working with implementing hand use in Virtual Reality. The thesis presents hand models used to generate natural looking grasping motions. One model were made for each of the thirty-three different grasp types in Feix's *The GRASP Taxonomy*.

Each model is based on functional principal components analysis which was performed on data containing recorded joint angles of grasping motions from real subjects. Prior to functional principal components analysis, dynamic time warping was performed on the recorded joint angles in order to put them on the same length and make them suitable for statistical analysis. The last step of the analysis was to project the data onto the functional principal components and train Gaussian mixture models on the weights obtained. New grasping motions could be generated by sampling weights from the Gaussian mixture models and attaching them to the functional principal components.

The generated grasps were in general satisfying, but all of the thirty-three grasps were not distinguishable from each other. This was most likely caused by the fact that each degree of freedom was modelled in isolation from each other, so that no correlation between them was included in the model.



# Modellering av Naturliga Handrörelser för Greppanimationer

Johannes Jeppsson

## Abstract

Denna rapport utfördes på Gleechi, ett svenskt start-up företag som jobbar med att implementera handrörelser i Virtual Reality. Uppsatsen presenterar statistiska modeller för att generera handrörelser som utför olika typer av grepp och som ser naturliga ut. En modell skapades för alla trettiofyra grepp i Feixs *The GRASP Taxonomy*.

Varje modell bygger på funktionell principalkomponentsanalys som utfördes på data innehållande inspelade vinklar från fingerleder från personer som utförde olika grepp på föremål. Innan funktional principalkomponentanalys utfördes så genomfördes dynamic time warping på datan för att få de inspelade greppen på samma längd och göra den lämplig för statistisk analys. Det sista steget i analysen var att projicera ned datan på principalkomponenterna och träna Gaussian mixture models på vikterna som erhöles. Nya grepp kunde då genereras genom att dra vikter från Gaussian mixture models och skapa linjärkombinationer med de funktionella principalkomponenterna.

De genererade greppen var generellt sett tillfredställande, men alla trettiofyra grepp typer var inte särskiljbara från varandra. Detta berodde med största sannolikhet på att varje frihetsgrad modellerades isolerat från de andra så att ingen korrelation mellan dem var inkluderad i modellen.



## Acknowledgements

I would like to thank Timo Koski for his mathematical insight, his ability to calm me down when I was sure everything was doomed, and his entertaining stories.

I would like to thank Gabriel Isheden for sharing his exceptional ability of analyzing problems and for his kindness of taking me under his wings.

I would like to thank Sixten Matz for his fantastic ideas and clear-sight.

I would like to thank Henriette Matz for being such a fantastic support and team player.

Finally I would like to thank Gleechi for their hospitality.

Stockholm, July 2017

Johannes Jeppsson



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Background . . . . .	1
1.2	Related Work . . . . .	2
1.3	Outline . . . . .	7
	<b>Abbreviations</b>	<b>9</b>
<b>2</b>	<b>Model</b>	<b>11</b>
2.1	Choice of model . . . . .	11
2.2	Description of model . . . . .	12
<b>3</b>	<b>Theory</b>	<b>13</b>
3.1	Dynamic Time Warping for One-dimensional Sequences . . . . .	13
3.2	Dynamic Time Warping for Multivariate Sequences . . . . .	14
3.3	Principal Components Analysis . . . . .	14
3.4	Functional Principal Components Analysis . . . . .	15
<b>4</b>	<b>Method</b>	<b>17</b>
4.1	Implementation . . . . .	17
4.2	Classification . . . . .	21
<b>5</b>	<b>Results</b>	<b>23</b>
5.1	Result of implemented techniques . . . . .	23
5.2	Generating grasps . . . . .	31
5.3	Classification . . . . .	36
<b>6</b>	<b>Summary</b>	<b>41</b>
<b>A</b>	<b>Generated grasps</b>	<b>43</b>



# Chapter 1

## Introduction

This degree project aims to use a data-driven approach to generate natural looking hand motions in the process of reaching to grasp objects in the context of Virtual Reality applications.

Realistic hand grasp animation is an important aspect for an immersive Virtual Reality experience, as grasping objects is one of the most important ways humans interact with the world. This report focuses on the generation of the hand and finger motions in the process of grasping objects, but not the process of establishing a realistic grip on the object. This degree project was carried out at Gleechi; a Swedish start-up company producing software for realistic hand interaction in real-time; <http://www.gleechi.com>.

The task of this project was to use a data-driven technique for generating natural looking human hand motion in the context of grasping objects in order to investigate if this could be a viable approach for future use.

### 1.1 Background

Humans are very good at perceiving details in hand gestures, e.g. Jörg et al. found in [18] that hand and finger motion is important for conveying the intended meaning of a scenario. Being able to produce accurate and realistic hand motions is therefore important for increasing the feeling of presence in using Virtual Reality.

Generating realistic hand motions in Virtual Reality is however a difficult task. And there are several obstacles connected with generating accurate and realistic hand movements.

#### 1.1.1 High Degree of Freedom

The human hand itself is a very complex organ that consists of twenty-seven bones with interlinking joints. Any mechanical model attempting to represent the hand with its dynamic properties is bound to be complex.

Animating hand motion is thus a very difficult task since there are a high number of interdependent degrees of freedom to be controlled for.

The fact that the hand has many degrees of freedom must be taken into account in the data collection process, i.e. the more dimensions that are taken into account, the more data needs to be collected to account for the extra dimensions.

### **1.1.2 Naturalness**

Generating hand motions that look realistic to humans pose another demand on the generation process. Movements that for some reason are perceived as artificial will get in the way of the Virtual Reality experience. In continuous use, it might also be necessary for movements performing a similar task to exhibit some variation, in order to not look artificial. Naturalness is the benchmark of generated motions, but it is also the hardest to pinpoint.

### **1.1.3 Data driven**

There are hand models that use databases as starting points and search them for similar motions that might be configured for the task at hand. However, the aim of this thesis is to produce a model that can sample data without using a database. This means that a probabilistic distribution that reflects hand motions must be estimated.

### **1.1.4 Real-time generation**

It is the goal of Gleechi to obtain a model that is able to sample new motions in real-time. This puts an upper bound on how complicated the model can be in terms of computer power. Since the task of this thesis is exploratory in its nature, this aspect is not of immediate importance. But when choosing a model it should be considered whether it is possible to make real-time simulations eventually.

### **1.1.5 Non-linearity**

Hand motions are intrinsically non-linear, and thus a model for generating hand motions should preferably also be able to generate non-linear motions.

## **1.2 Related Work**

There have been several different approaches on how to go about generating natural-looking hand motions in Virtual Reality. The most prominent methods, which have been considered in this thesis, can be categorized

into techniques based on principal component analysis (PCA), Gaussian process latent variable machines (GPLVMs) and restricted Boltzmann machines (RBMs) respectively.

### **1.2.1 Grasp taxonomies**

An important aspect of analyzing grasping of objects has been the effort of creating grasp taxonomies, such as the Cutkosky grasp taxonomy [2]. For instance Feix et al. in has provided a list in [32] of thirty-three different grasps, when only considering the hand and not the object shape and size. Figure 1.1 below contains images of the thirty-three grasp types of the Feix taxonomy. Taxonomies are considering the final hand posture when a grip is established, but by that virtue they are also important for how reaching to grasp motions differ. Taxonomies provide the foundation for differentiating between grasp types.

1	Large Diameter		12	Precision Disk		23	Adduction Grip	
2	Small Diameter		13	Precision Sphere		24	Tip Pinch	
3	Medium Wrap		14	Tripod		25	Lateral Tripod	
4	Adducted Thumb		15	Fixed Hook		26	Sphere 4 Finger	
5	Light Tool		16	Lateral		27	Quadpod	
6	Prismatic 4 Finger		17	Index Finger Extension		28	Sphere 3 Finger	
7	Prismatic 3 Finger		18	Extension Type		29	Stick	
8	Prismatic 2 Finger		19	Distal Type		30	Palmar	
9	Palmar Pinch		20	Writing Tripod		31	Ring	
10	Power Disk		21	Tripod Variation		32	Ventral	
11	Power Sphere		22	Parallel Extension		33	Inferior Pincer	

Figure 1.1: The thirty-three grasp types of the Feix grasp taxonomy.

### 1.2.2 PCA based approaches

PCA based techniques arise mainly from the effort to tackle the many degrees of freedoms in the hand.

In [3], Santello et al. explored the idea of lack of individuation in finger movements. They asked five right-handed subjects to grasp each of fifty-seven imaginary objects five times. They recorded the final posture of the hand with fifteen degrees of freedom and then performed discriminant analysis, regression analysis, and principal components analysis. They found that the first two principal components accounted for more than eighty percent of the variance in hand shapes.

A follow up study on Santello et al. [3] was later performed by Dai et al. [26]. They performed PCA and found that eighty-three percent of the variance was captured by the first three principal components. After this they performed functional principal components analysis on the projected data and found that the first two functional components accounted for about ninety-five percent of variation in the grasping motion. The subjects performed fifteen different grasps from the Cutkosky grasp taxonomy, and they suggest four different groups of motion based on k-means-clustering.

Ciocarlie et al. [9] used the first two principal components for grasp synthesis of four different hand models. In order to find a conforming grasp, they formulated an energy function that described the distance of previously selected contact points on the hand model to the object. The energy function was then minimized using Simulated Annealing. As a feasibility test, they took the grasp found by minimizing the energy function and use inverse kinematics to find an arm position. If an arm position was found without any collision, the configuration was chosen. Using the first two principal components, the resulting grasp did not always conform to the surface of the object. In those cases they closed each finger until it fully enclosed the object.

Amor et al. [13] collected a wide range of possible hand shapes, on which they then performed PCA to achieve low-dimensional grasp space. To discriminate to hand positions that are anatomically feasible, they learnt a Gaussian Mixture Model (GMM) on the PCA space. The number of Gaussians used in the GMM was chosen using the Bayesian Information Criterion. To synthesize new grasps, they stipulated a grasp metric based on the distance of the sensors to the object, an estimate of the stability of the grasp, and a penalty value. To find a suitable grasp they minimized the metric using Dynamic Hill Climbing technique. The grasp found was accepted if it also satisfied that the probability of the grasp was above a certain threshold, according to the GMM.

Zhao et al. [27] divided a grasping hand motion into reaching, closing and manipulating phases. They obtained grip-dependent probabilistic motion models for closing motion. Training data was labeled with the type of

grasp used (e.g. pinch grasp) and the final database included motion capture data from hands grasping ten different objects using ten different grip modes. To process the data, dynamic time warping was used to warp the geometric and temporal training data respectively into comparable time series of equal length. FPCA was then applied to the warped time series to obtain a low-dimensional, parametric representation for closing motion. Furthermore, they learn a joint probability distribution to model the correlation between the geometric and timing variations. The prior distributions were modeled using a Gaussian mixture model. The parametric representation and the joint probability distribution define a generative model for closing-phase motion synthesis. Synthesizing new motion was done by sampling the joint probability distribution and inserting the samples into the parametric representation. For synthesizing the reaching motion, they searched for a similar motion from the database and confined it to the first frame of the closing motion using smoothing techniques.

In [31] Du et al. extended FPCA by scaling, a method they call SFPCA. They applied FPCA on weighted data, where the weights were found such that they minimized the error in Euclidean joint space.

### 1.2.3 RBMs

RBMs are a form of neural networks that have been used extensively in modeling locomotion.

In [11] Taylor et al. introduced the Conditional Restricted Boltzmann Machine (CRBM) to model different styles of human motion. The CRBM is a two-layer neural network that extends the Restricted Boltzmann Machine (RBM) by taking into consideration the temporal aspect of the data. This is done by also taking as input the previous  $n$  configurations, making the input have autoregressive connections. In their experiment they included the previous three time steps. The model was then trained by a method called contrastive divergence. The trained CRBM is a probability density model of sequences, thus the CRBM can be used to synthesize new motions by initializing with the first  $n$  time steps. Sampling from the joint distribution was done by performing alternating Gibbs sampling. The resulting model was able to model different styles of motion that were present in the training data. To generate a motion of a specific style, the model was initialized with the first  $n$  time steps of that style, e.g. jogging or walking.

In [17] Taylor and Hinton presented an extension to the CRBM, called the Factored CRBM. The factored CRBM aims to improve the ability to blend motion styles or to transition smoothly between them, and it also lowers the computational load from  $O(N^3)$  to  $O(N^2)$ . For purposes of comparison, they also used a two-layer CRBM which successfully models ten different styles of walking. The factored CRBM was also able to model the ten different motion styles, as well as blending and transitioning between styles.

In [19] Taylor et al. presented another extension to CRBM that includes latent style variables, which they call Implicit Mixture of CRBMs (imCRBM). imCRBMs can be trained unsupervised, in which case the model detects atomic motion primitives in the data.

Chiu and Marsella [20] made use of the property of stacking RBMs. They introduced a hierarchical factored conditional Restricted Boltzmann Machine (HFCRBM) which consisted of two layers. The first layer is a restricted CRBM which extract patterns of the motion and the second layer is an FCRBM. By this structure they are able to generate motion styles as well as blending of styles with better performance than a single FCRBM.

In [21] Chiu and Marsella modified the HFCRBM by stacking an FCRBM on top of a CRBM. The modified HFCRBM was used to generate a gesture motion from speech.

#### 1.2.4 GPLVMs

GPLVMs is a dimensionality reduction technique that generalizes PCA to represent the data with Gaussian Processes, this provides for non-linear mapping from the latent space to the data space [5]. GPLVMs has been used to model human poses in [6].

In [8] Wang et al. introduced Gaussian Process Dynamical Models (GPDMs), which extend GPLVMs with an autoregressive component which allows to model the temporal aspect. Wang et al. later used GPDMs with style-specific factors to model human locomotion of different styles in [12] and [14].

Chiu et al. generated gestures based on speech labels in [29] with the use of GPDMs. In their approach, each point in the reduced space corresponds to a motion frame in the observed space; to generate gestures, they sampled a trajectory in the reduced space and mapped it back to the observed space. In particular they interpolated trajectories to generate smooth transitions between gestures connected to different speech labels.

### 1.3 Outline

In chapter two, the choice of model is presented, specifically in section 2.2 the model for generating hand motions is described. In chapter three, the theoretical background for dynamic time warping, principal component analysis and functional principal component analysis is presented. In chapter four, the methodology of the analysis is discussed. In chapter five, the results of the analysis are presented.



# Abbreviations

DoF	Degree of freedom
DTW	Dynamic time warping
FPCA	Functional principal component analysis
FPCs	Functional principal components
GMM	Gaussian mixture model
LOOCV	Leave-one-out cross-validation
PCA	Principal component analysis
PCs	Principal components



# Chapter 2

## Model

In this chapter, the choice of using a model based on FPCA is argued for on the basis of the literature study. Next, the general outline of the model is presented.

### 2.1 Choice of model

The main reason a model based on GPLVMs was not chosen is that they only provide a mapping from latent space to observed space, and no mapping in the opposite direction. While this is what is desired in generation of grasps, it is not suitable for classification purposes; i.e. given an observed grasp, find the mapping to latent space.

CRBMs have the advantage that all DoFs come to use, so that no dimension reduction is needed. However, CRBMs studied in the literature review are not deterministic in the sense that they are not guaranteed to stay within the type they are initialized in. This makes CRBMs problematic for generating grasps of specific types.

FPCA does not show the obstacles presented for GPLVMs or CRBMs above which is why it is the chosen methods of this thesis.

With respect to the obstacles for generating hand motions presented in the previous chapter, FPCA is a good candidate. First, FPCs are able to describe non-linear motions. Second, only the FPCs are needed to be kept in the computer memory, so real-time generation is most likely possible. Third, it is a data driven approach which fits into the objectives of Gleechi. Finally, in order to handle the high amount of DoFs, FPCA has been applied individually to each DoF. This approach is obviously a simplification of the hand motions since correlation between joint angles are not accounted for by this approach.

## 2.2 Description of model

In order to apply FPCA, the first step was to divide the data by grasp type, and within each grasp type divide the data by DoF, so that FPCA could be applied to each DoF separately. The second step was to get the data on the same length; for this purpose DTW was applied to all grasps within a grasp type. DTW resulted in two datasets; one containing the warped motions, and the other containing the time warping functions.

Next, FPCA was applied to each DoF for each grasp type. Each DoF,  $\theta_i(t)$ , describes a one-dimensional angle in a finger joint, and the change over time in the angle could thus be modeled by a linear combination of the mean motion and the FPCs as

$$\theta_i(t) = p_{i,0}(t) + \alpha_{i,1}p_{i,1}(t) + \cdots + \alpha_{i,k}p_{i,k}(t) \quad (2.1)$$

$$= \mathbf{p}_i^T(t)\boldsymbol{\alpha}_i, \quad (2.2)$$

where  $\mathbf{p}_i(t)$  contains the FPCs and of which  $p_{i,0}(t)$  represents the mean function.

Then FPCA was performed on the time warping functions,  $\tau_i(t)$ , as well. Temporal variations could be modeled similarly as

$$\tau_i(t) = q_{i,0}(t) + \beta_{i,1}q_{i,1}(t) + \cdots + \beta_{i,l}q_{i,l}(t) \quad (2.3)$$

$$= \mathbf{q}_i^T(t)\boldsymbol{\beta}_i, \quad (2.4)$$

where  $\mathbf{q}_i(t)$  contains the FPCs and of which  $q_{i,0}(t)$  represents the mean function.

In the analysis, two FPCs was used for joint angles and one FPC was used for time warping functions, hence  $k$  was set to 2 and  $l$  was set to 1.

The end result was one model for each grasp type of Feix's grasp taxonomy, thirty-three models in total. A parametric model for generating motions for each grasp type was obtained by

$$\Theta(t) = \begin{pmatrix} \mathbf{p}_1^T(\mathbf{q}_1^T(t)\boldsymbol{\beta}_1)\boldsymbol{\alpha}_1 \\ \vdots \\ \mathbf{p}_{16}^T(\mathbf{q}_{16}^T(t)\boldsymbol{\beta}_{16})\boldsymbol{\alpha}_{16} \end{pmatrix}, \quad (2.5)$$

where variations of the grasp type were obtained by providing different weights,  $\boldsymbol{\alpha}_i$  and  $\boldsymbol{\beta}_i$ .

The last step of the model was to train GMMs to be able to estimate the weights  $\boldsymbol{\alpha}_i$  and  $\boldsymbol{\beta}_i$ . This was done by projecting the recorded grasps in the data onto the FPCs and train the GMMs on the projected weights. One GMM was trained for DoF.

For generating new grasps, estimate weights,  $\hat{\boldsymbol{\alpha}}_i$  and  $\hat{\boldsymbol{\beta}}_i$ , were sampled from the GMMs and inserted into the parametric model 2.5.

# Chapter 3

## Theory

In this section, the theory of DTW and FPCA, which are central to the model, will be briefly explained.

### 3.1 Dynamic Time Warping for One-dimensional Sequences

DTW [1, 24] is a technique for measuring similarity between two temporal sequences. The intuition is that the sequences are aligned by locally stretching and shrinking until they are as similar as possible.

In its basic form, DTW takes two sequences,  $\{x_i\}_{i=1}^M$  and  $\{y_j\}_{j=1}^N$ , and finds a new set of indices  $\{m_k\}_{k=1}^L$  and  $\{n_k\}_{k=1}^L$  of same length  $L$  such that

$$\sum_{l=1}^L d(x_{m_l}, y_{n_l}), \quad (3.1)$$

is minimized for some distance function  $d(\cdot, \cdot)$  (e.g. Euclidean distance). The two sequences  $\{x_{m_k}\}_{k=1}^L$  and  $\{y_{n_k}\}_{k=1}^L$  are referred to as the *warped* sequences, and the indices  $\{m_k\}_{k=1}^L$  and  $\{n_k\}_{k=1}^L$  are referred to as the *warping paths*.

The warping paths must satisfy the following conditions

1.  $1 \leq m_k \leq M, 1 \leq n_k \leq N$
2.  $m_1 = n_1 = 1$  and  $m_L = M, n_L = N$ ,
3.  $(m_{k+1} - m_k, n_{k+1} - n_k) = (0, 1), (1, 0)$  or  $(1, 1)$ .

The first condition assures that the warping paths are defined on the same interval as the original sequences. The second condition assures that the boundary conditions are honored, i.e. the warped sequences have the same starting and ending point as the original sequences. The third and final condition assures that at least one of the warping paths are increasing all

the time in order to avoid that the warping paths result only in repeating the indices that produce the smallest distance.

Let us reformulate DTW slightly to link it to the multivariate case. Let  $\mathbf{x} \in \mathbb{R}^M$  and  $\mathbf{y} \in \mathbb{R}^N$  represent the sequences to align. Finding the warping paths is equivalent to finding matrices  $\mathbf{W}_x \in \mathbb{R}^{M \times L}$  and  $\mathbf{W}_y \in \mathbb{R}^{N \times L}$  with columns containing only zeros except for a single one such that the conditions above are fulfilled and

$$d(\mathbf{x}\mathbf{W}_x, \mathbf{y}\mathbf{W}_y) \tag{3.2}$$

is minimized.

### 3.2 Dynamic Time Warping for Multivariate Sequences

In the multivariate case, the multivariate sequences to be aligned can be represented by matrices. Let us also allow for there to be  $N$  different sequences to align. For an informative approach to multivariate DTW [25] can be consulted.

Let the original sequences be represented by a set of  $d$ -dimensional matrices  $\{\mathbf{X}_i\}_{i=1}^N$ ,  $\mathbf{X}_i \in \mathbb{R}^{d \times n_i}$ , where  $n_i$  represents the length of matrix  $\mathbf{X}_i$ . As in the one-dimensional case the objective is to find a set of matrices  $\{\mathbf{W}_k\}_{k=1}^N$ , with columns containing only zeros except for a single one such that

$$\sum_{i=1}^N \sum_{j=1}^N d(\mathbf{X}_i\mathbf{W}_i, \mathbf{X}_j\mathbf{W}_j), \tag{3.3}$$

is minimized. The warped sequences are represented by the set  $\{\mathbf{X}_k\mathbf{W}_k\}_{k=1}^N$  which consists of  $d$ -dimensional matrices of equal length  $L$ . The warping paths are represented by the set  $\{\mathbf{W}_k\}_{k=1}^N$ .

### 3.3 Principal Components Analysis

Principal components analysis (PCA) is a technique that converts a set of observations into a set of linearly uncorrelated variables. PCA is commonly used for dimensional reduction.

Given  $N$  observations of a  $d$ -dimensional variable represented in a matrix  $\mathbf{X} \in \mathbb{R}^{N \times d}$ , PCA finds the orthogonal basis in  $\mathbb{R}^d$ , where each basis vector, in descending order, explains as much variance as possible in the sampled data  $\mathbf{X}$ .

What PCA does in particular is to find the eigenvectors of the sample covariance matrix  $\frac{1}{N}\mathbf{X}^T\mathbf{X}$  (assuming that  $\mathbf{X}$  has zero mean). The largest

eigenvalue corresponds to the basis vector that explains the most variance in the collected data. Thus choosing any subset of the eigenvectors corresponding to the largest eigenvalues is going to span a subspace of  $\mathbb{R}^d$  that explains the most variance possible in  $\mathbf{X}$  for that amount of basis vectors.

### 3.4 Functional Principal Components Analysis

Function principal component analysis (FPCA) is a generalization of PCA in the Hilbert space  $L_2(\mathcal{T})$ . Any of [34, 30, 28] can be consulted, where the last gives a very good explanation of the more general case.

Let  $L_2(\mathcal{T})$  be a Hilbert space of square integrable functions with respect to Lebesgue measure  $dt$  on an interval  $\mathcal{T} = [a, b]$ ,  $a < b$ . The inner product on  $L_2(\mathcal{T})$  is defined as

$$\langle f, g \rangle = \int_{\mathcal{T}} f(t)g(t)dt. \quad (3.4)$$

Further, each element  $f \in L_2(\mathcal{T})$  has a continuous mean function  $\mu(t) = \mathbb{E}[f(t)]$ , and a continuous covariance function  $K_f(s, t) = \text{Cov}(f(s), f(t))$ .

For an arbitrary random variable  $X \in L_2(\mathcal{T})$ , define  $Z$  as

$$Z(t) = X(t) - \mu(t). \quad (3.5)$$

Since  $Z$  is square integrable, has zero mean, is defined over a closed and bounded interval and has a continuous covariance function, the Karhunen-Loève theorem implies that  $Z$  has the representation

$$Z(t) = \sum_{r=1}^{\infty} \xi_r \phi_r(t), \quad (3.6)$$

where  $\{\phi_r\}_{r=1}^{\infty}$  is an orthonormal basis of  $L_2(\mathcal{T})$  and  $\xi_r = \langle Z, \phi_r \rangle$ .

Further, the orthonormal basis  $\{\phi_r\}_{r=1}^{\infty}$  are the eigenfunctions of the linear operator  $T_{K_Z}$  formed by the covariance kernel of  $Z$ , defined as

$$T_{K_Z}f = \int \text{Cov}(Z(s), Z(\cdot))f(t)dt \quad (3.7)$$

$$= \langle K_Z(s, \cdot), f \rangle. \quad (3.8)$$

That is to say that  $T_{K_Z}$  admits eigenfunctions  $\phi_r \in L_2(\mathcal{T})$  satisfying

$$T_{K_Z}\phi_r = \lambda_r\phi_r(s). \quad (3.9)$$

Using (3.5) and (3.6) gives a representation of  $X$  expressed in the basis  $\{\phi_r\}_{r=1}^{\infty}$  as

$$X(t) = \mu(t) + \sum_{r=1}^{\infty} \xi_r \phi_r(t). \quad (3.10)$$

Here  $\phi_r$  is referred to as the  $r$ :th FPC. The coefficients  $\{\xi_r\}_{r=1}^{\infty}$  in (3.10) minimize the expression

$$\|X - \mu - \sum_{r=1}^{\infty} \xi_r \phi_r(t)\|. \quad (3.11)$$

Assuming that the eigenvalues  $\{\lambda_r\}_{r=1}^{\infty}$  are expressed in non-increasing order,  $X$  can be readily approximated by

$$X(t) \approx \mu(t) + \sum_{r=1}^p \xi_r \phi_r(t), \quad (3.12)$$

for some finite  $p \in \mathbb{N}$ .

FPCA is a useful technique to obtain a small dimension space which captures much of the variability in the data. The first FPC accounts for the most variation, while the second FPC accounts for the largest variation orthogonal to the first FPC, and so on. Much of the variation in the data can be captured by only using a few FPCs as in (3.12).

# Chapter 4

## Method

In this chapter, the procedure in each step of the method is described in more detail. First, the treatment of data is described, then how the time warping was performed followed by the application of FPCA and PCA to the treated data, and lastly how the classification was performed.

### 4.1 Implementation

The purpose of the analysis was to create generative models for natural hand motion as described in chapter two.

In order to accomplish this goal, the data was grouped by grasp types after which DTW was performed. Next, FPCA was performed on the data within each grasp type. Lastly one GMM was trained on the FPC weights for each DoF. In order to generate motions, weights could be sampled from the GMMs.

An alternative approach was tested where additionally PCA was performed in each time step after DTW. FPCA was then performed on the PCA weights instead. One GMM was trained on the FPC weights for each degree of freedom as in the first approach. This second approach reduces the DoFs and the computational load, but in turn it also contains less information than the first model and was expected to be less accurate.

Finally, a classification approach was taken to compare the models' ability to distinguish between different grasp types. The classification was done by simply calculating the the  $L_2$ -distance between a grasp and its projection in each model. The grasp was considered to belong to the type which accounts for the lowest distance. For comparison, a similar classification approach was taken where PCA was performed on the last frame and the Euclidean distance was calculated of the last frame instead.

### 4.1.1 Data

In essence, hand motion data are joint angle values in the form of time series.

In this report, the data used was from the HUST dataset [33]. The HUST dataset contains recorded data from thirty subjects performing grasps from Feix’s grasp taxonomy [32]. Images of the thirty-three different grasp types in Feix’s taxonomy is found in figure 1.1 above.

### HUST dataset

The HUST dataset contains data from thirty subjects performing all thirty-three grasp types from Feix’s grasp taxonomy. Each grasp was performed on three objects of different sizes and shapes. For each object, the grasp was performed three times to depress random error. This resulted in ninety grasps for each object, 270 grasps for each grasp type, and 8910 grasps in total in the HUST dataset.

The data was recorded using sixteen degrees of freedom, and was recorded at fifty hertz. Each grasp was saved as a comma-separated file with sixteen columns representing degrees of freedom, and succeeding rows representing measured joint angles at an interval of 0.02 seconds (50 Hz).

The HUST dataset only provides angles for each performed grasp. Knowledge about e.g. hand sizes, object sizes, distance to object or definition of zero angles are not provided. This lack of information made the visualization hard to make exact, since the hand model had to be tuned on how the final configuration was supposed to look. Another reason for difficulty in tuning the visual model was that the two angles CMC1 and ABD1 were not orthogonal.

### 4.1.2 Time Warping

For DTW, the software written in MATLAB provided by [25] was used.

In general the recorded samples had different length, i.e. the number of frames for each recorded grasping differed. In order to be able to perform FPCA and PCA in each time frame, the grasps had to be aligned in a way that made comparison at each frame meaningful.

The grasps were grouped by grasp type and then further grouped by which of the three objects was grasped. This further grouping was needed in order for the time warping software to be able to handle the amount of data.

Before warping, any recorded grasp with significantly longer frame count than the rest was removed. Typically ten recorded grasps for each grasp type were removed in this manner.

For each grasp type, the mean frame count was used as the target length for warping. All three objects were warped separately to this length, and the warped grasps were stored together for further analysis.

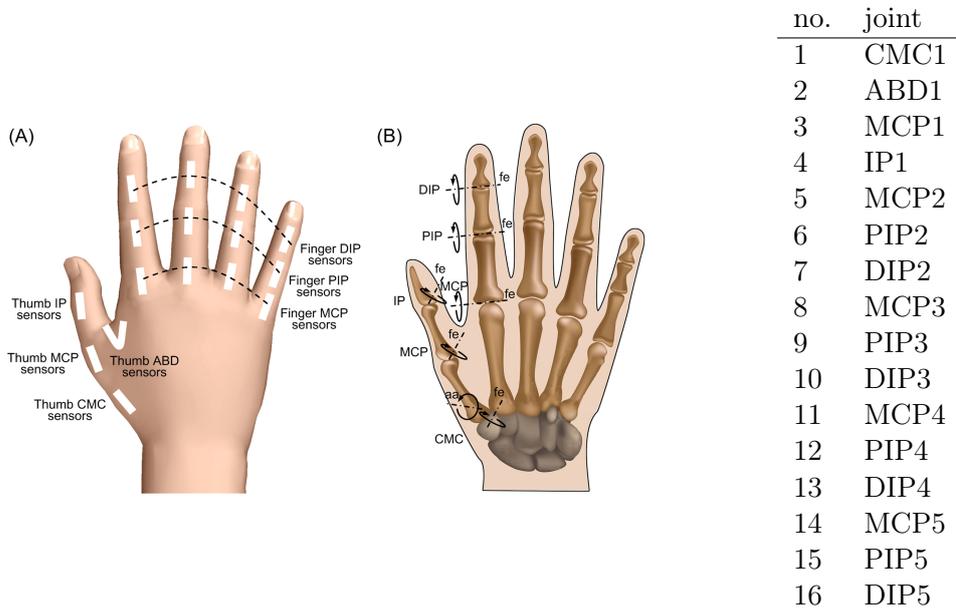


Figure 4.1: Description of the sixteen DoFs used in the HUST dataset. The numbering of joints (e.g. PIP4) is referring to the corresponding finger; 1 = thumb, 2 = index finger, 3 = middle finger, 4 = ring finger, 5 = little finger.

The DTW procedure resulted in the warped grasps as well as their warping paths, where the latter were mappings between warped frames and original frames. Thus statistical analysis could be carried out on the warped grasps as well as the warping paths separately.

### 4.1.3 FPCA implementation

From DTW, a collection of warped grasps of the same length for each grasp type was provided. To get data which was suitable for FPCA, the warped data was divided by DoF. So in each grasp type, the data was divided into sixteen datasets, each one describing realizations of one degree of freedom.

FPCA was performed separately on each DoF. This means that for each grasp type, FPCA was performed sixteen times, once for each DoF. The first two FPCs were used for each DoF. For FPCA, the R package `fda.usc` provided by [23] was used.

### FPCA on Warping Paths

FPCA was also carried out on the warping paths of each grasp type. For this analysis only one FPC was used since it preserved one hundred percent of the variation. Thus each warping path could be represented by the linear combination of the mean function and the first FPC of its grasp type.

### FPCA with preceding PCA

An alternative approach was attempted, wherein FPCA was preceded by PCA. The purpose of this approach was to further reduce the dimensionality of the data, resulting in less variables to control for.

After DTW, PCA was performed in each successive frame. The first PCs of every frame were concatenated together in chronological order, and the same thing was done to the second PCs and so on. The purpose of this procedure was to create a time dependent function of how the grasp changed over time with lower dimensionality. FPCA was then applicable to the weights of the concatenated PCs.

Given a discrete measurement of a grasp in  $T$  time frames,  $\mathbf{g} = \{g_t\}_{t=1}^T$ ,  $g_t \in \mathbb{R}^{16}$ , the first step was to carry out PCA in each time frame,  $t$ , so that

$$g_t \approx \sum_{i=1}^d \alpha_{i,t} \mathbf{p}_{i,t}, \quad (4.1)$$

where  $\mathbf{p}_{i,t}$  is the  $i$ :th PC at frame  $t$ ,  $d$  is the number of PCs and  $\alpha_{i,t}$  is the projection of  $g_t$  onto  $\mathbf{p}_{i,t}$  (assuming that  $\mathbf{p}_{i,t}$  is chosen so that  $\|\mathbf{p}_{i,t}\|_2 = 1$ ).

The next step was to concatenate each PC's projections as  $\alpha_i = \{\alpha_{i,t}\}_{t=1}^T$  to get the evolution of the projections over time. Then  $\alpha_i$  was approximated as a continuous function and FPCA was performed on it, so that

$$\alpha_i(\tau) \approx \sum_{j=1}^k \beta_{i,j} \mathbf{f}_{i,j}(\tau), \quad (4.2)$$

where  $\mathbf{f}_{i,j}$  is the  $j$ :th FPC for  $\alpha_i$ ,  $k$  is the number of FPCs and  $\beta_{i,j}$  is the projection of  $\alpha_i$  onto  $\mathbf{f}_{i,j}$  (assuming that  $\|\mathbf{f}_{i,j}\|_{L_2} = 1$ ).

In summary, combining (4.1) and (4.2) gives

$$g_t \approx \sum_{i=1}^d \left( \sum_{j=1}^k \beta_{i,j} \mathbf{f}_{i,j} \right) (t) \mathbf{p}_{i,t}. \quad (4.3)$$

#### 4.1.4 GMMs

From FPCA, each DoF of the warped grasp could be approximately represented by the mean function and a linear combination of FPCs. For each DoF, two FPCs were used and hence two weights were needed for representation. The way in which new grasps was generated was through sampling FPC weights from GMMs.

In order to train GMMs for this purpose, all grasps was projected on the FPCs of their own grasp type in order to obtain their FPC weights. For each DoF, the set of all two-tuples containing the two weights for a projected DoF were used to train a two dimensional GMM with two components. When the GMMs were trained, weights could be sampled from them.

To generate new grasps, weights were sampled from the GMMs and used to make linear combinations of FPCs as described in (2.5).

For training GMMs, the R package `mixtools` provided by [15] was used. In this package, the expectation-maximization (EM) algorithm is used to find the parameters for normal distributions contained in the GMMs.

#### 4.1.5 Visualization

For visualizing grasping motions, the page <http://www.mymodelrobot.appspot.com/> was used [35]. To use this site a model had to be defined, which was provided by Gleechi. Grasping motions could then be visualized by providing comma separated files with angles of each joint defined in the model.

## 4.2 Classification

For evaluating how well each model was at representing its specific type of grasp, a classification scheme was performed in which all unwarped grasp were fed into each model to determine how well that model was able to represent the grasp.

When a grasp was fed to the model, the first step was to warp the grasp to the warped time of the model. The warping was done by the mean function obtained from FPCA on warping paths.

Next, the warped grasp was projected on the FPCs and the  $L_2$  distance was calculated between the warped grasp,  $\mathbf{g}$ , and its projection,

$$\|\mathbf{g} - \sum_{i=1}^k \langle \mathbf{g}, \mathbf{f}_k \rangle \mathbf{f}_k\|_{L_2}, \quad (4.4)$$

where  $\mathbf{f}_k$  is the k:th FPC of the model.

For comparison, a classification based on PCA of the last frame of each grasp type was performed. Similarly to the first approach, the Euclidean distance between the last frame of each grasp and its projection was calculated.



# Chapter 5

## Results

In this chapter the results of the analysis are presented and discussed.

In the first section the results of each individual implemented technique is presented, then in the next section the results from generating grasps are presented, finally the results of the classification is presented.

### 5.1 Result of implemented techniques

In this section, the result of each implemented technique is presented. An overlook of time warping and GMM training can be found in table 5.1 below, of which the details will be further discussed in the following sections.

	Object	Target frames	Warping	EM-algorithm
1	Large Diameter	43		
2	Small Diameter	41	linear	
3	Medium Wrap	42		
4	Adducted Thumb	40	linear	
5	Light Tool	38		
6	Prismatic 4 Finger	42		
7	Prismatic 3 Finger	44		
8	Prismatic 2 Finger	40		
9	Palmar Pinch	32		Fail on DoF 13
10	Power Disk	47	linear	
11	Power Sphere	43	linear	
12	Precision Disk	44		
13	Precision Sphere	41		
14	Tripod	39		
15	Fixed Hook	37		
16	Lateral	37	linear	
17	Index Finger Extension	40		
18	Extension Type	39		
19	Distal Type	54		
20	Writing Tripod	39	linear	
21	Tripod Variation	44	linear	
22	Parallel Extension	34		Fail on DoF 9, 10, 13
23	Adduction Grip	42		
24	Tip Pinch	36		
25	Lateral Tripod	38		
26	Sphere 4 Finger	42		
27	Quadpod	41		
28	Sphere 3 Finger	46		
29	Stick	41	linear	
30	Palmar	37		
31	Ring	39		Fail on dof 10, 13, 15, 16
32	Ventral	37		
33	Inferior Pincer	40		Fail on dof 10, 13, 14, 15

Table 5.1: Summary of warping and GMM training for each model. In the column *Warping* it is stated if linear warping had to be used as described in section 5.1.1. In the column *EM-algorithm* it is stated if problems occurred when training the GMMs due to EM-algorithm diverging as described in section 5.1.4.

### 5.1.1 Time warping

Time warping was performed on each grasp type with target length set to the mean frame length of data of that grasp type.

The warping was not always satisfactory for all grasps. This manifested itself as a repetition of a single time frame in the warped grasp, as can be seen in fig 5.1. Any grasp that was warped in this manner was removed. The amount of unsatisfactory warps was typically one or two for each object.

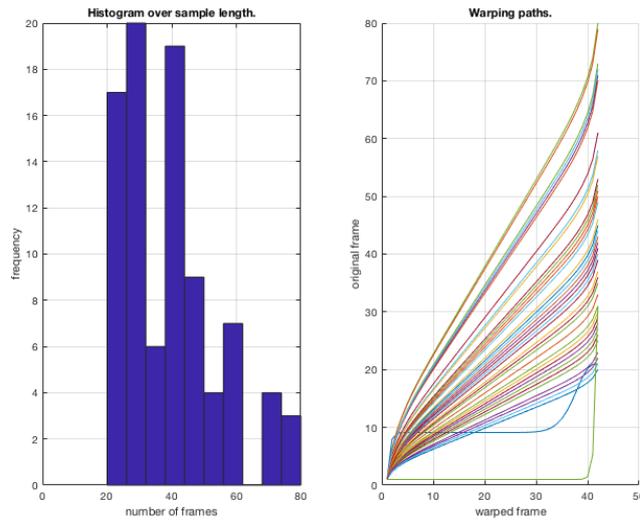


Figure 5.1: Example from time warping of *medium wrap*. The left subfigure shows a histogram of sample lengths before warping. The right subfigure shows warping paths, where two paths can be seen to be stationary in long intervals. In this example all warping paths are defined between frame zero and forty-two

In eight cases out of thirty-three, the warping software failed altogether and was unable to produce results. In these cases, a simpler form of warping was used instead, where the warping paths were linear. An example of this can be seen in fig 5.2. The models in which linear warping had to be used were *small diameter*, *adducted thumb*, *power disk*, *power sphere*, *lateral*, *writing tripod*, *tripod variation* and *stick*.

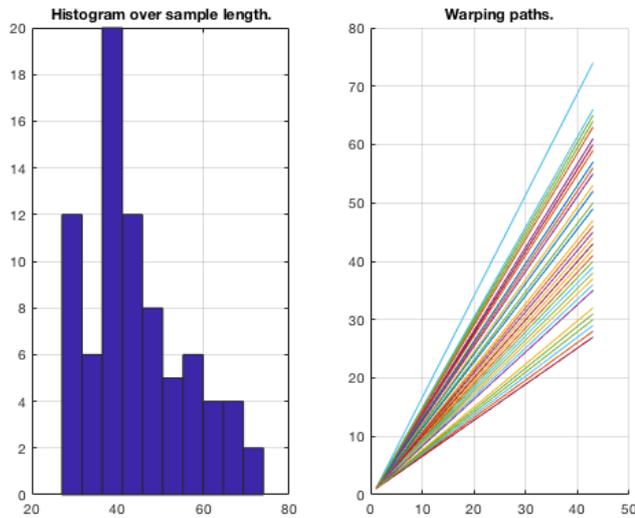


Figure 5.2: Example from time warping of *power sphere* when linear warping paths were used. The left subfigure shows a histogram of sample lengths before warping. The right subfigure shows warping paths, which can be seen to be linear. In this example all warping paths are defined between frame zero and forty-three

### 5.1.2 FPCA with preceding PCA

This approach did not work well, and was abandoned after initial results showed none to little resemblance with the grasp they were supposed to mimic.

The most obvious reason for the failure of this approach was the extra loss of information from the dimension reduction step in PCA.

Initial visualization of PCs showed them to be jerky in time. Since PCA in one frame was done in isolation from the rest, there was no guarantee for smoothness in transition with respect to neighbouring PCs.

Another source for jerkiness was that the PCs were switching sign in succeeding frames as can be seen in fig 5.3. This was expected since the sign of the PCs was arbitrary. The smoothness could be improved by manually switching sign on some PCs; this was not affecting the analysis since the negative of a vector is linearly dependent on the vector itself. Even though this correction was made, the results were still unsatisfactory.

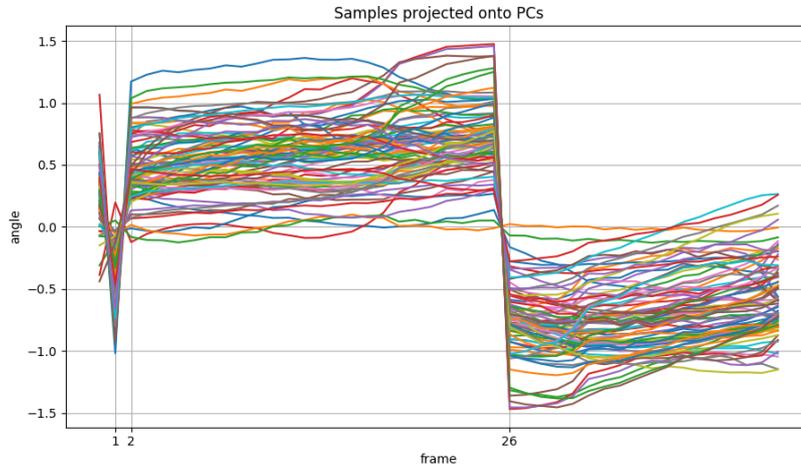


Figure 5.3: Projections reveal sign changes in principal components at successive frames.

### 5.1.3 FPCA

Considering all analyses, FPCA with two FPCs was able to explain between 86.00 and 99.89 percent of the variance in the data, with a mean at 95.90 percent. Thus the FPCs was successful in spanning a large subspace of the data and consequently had the capacity of making good approximations of joint angles. This means that successful grasp generation was most likely not severely restricted by FPCs' inability to reproduce certain grasp configurations lying outside FPC-space.

Variance explained by FPCs is presented in more detail in table 5.2 below.

An example of FPCA visualization together with data is presented in fig 5.4. And an example of data expressed as linear combination of FPCs is presented in fig 5.5. In these figures it can be seen that FPCs gives reasonably good approximations of the curves.

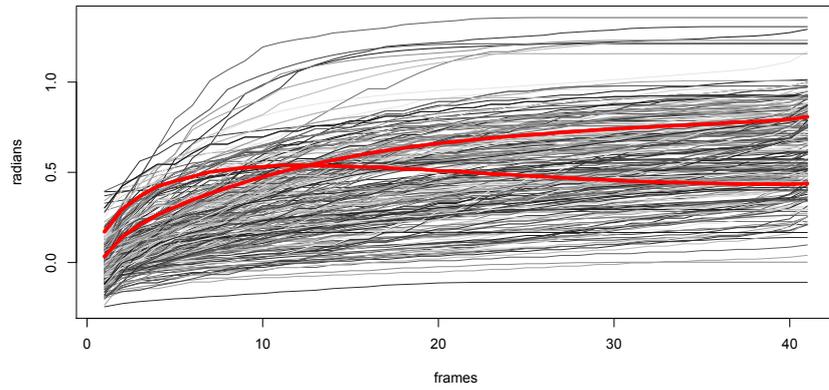


Figure 5.4: Example of samples of an angle and result of FPCA. Samples are in grey and mean function added to FPCs are in red.

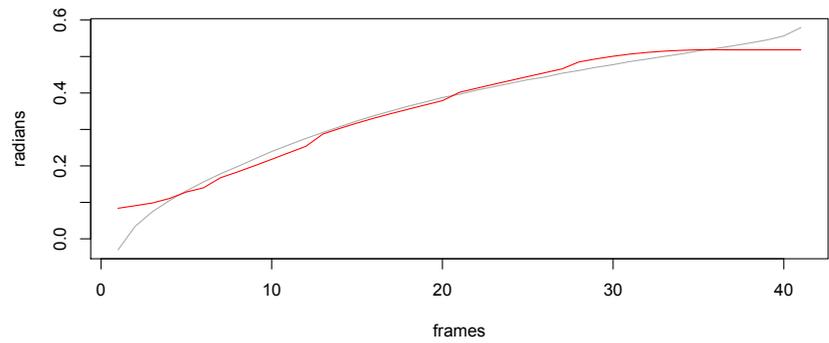


Figure 5.5: Example of a sample and its projection on FPCs. Sample is in grey and its projection is in red.

	Model 1	Model 2	Model 3	Model 4	Model 5
Min	93.00	90.69	91.13	86.97	92.28
Median	95.03	92.62	94.17	93.21	95.92
Mean	95.20	93.26	94.13	93.48	95.84
Max	99.04	99.03	98.40	98.58	98.71
	Model 6	Model 7	Model 8	Model 9	Model 10
Min	96.60	95.36	95.68	97.76	86.00
Median	97.49	97.72	98.05	98.96	92.55
Mean	97.63	97.33	97.80	98.70	92.03
Max	98.79	98.66	98.83	99.34	97.56
	Model 11	Model 12	Model 13	Model 14	Model 15
Min	89.82	92.83	95.77	94.77	93.40
Median	93.58	96.84	97.75	97.86	96.39
Mean	93.27	96.62	97.55	97.69	96.25
Max	98.57	98.02	99.12	98.93	98.89
	Model 16	Model 17	Model 18	Model 19	Model 20
Min	91.64	91.34	95.65	90.38	89.39
Median	96.28	94.79	96.91	95.31	94.79
Mean	95.75	95.20	96.96	95.20	94.01
Max	97.82	99.18	98.74	97.26	97.29
	Model 21	Model 22	Model 23	Model 24	Model 25
Min	87.43	93.98	93.88	96.97	94.85
Median	91.90	97.49	96.73	98.66	97.44
Mean	92.22	97.19	96.64	98.48	97.44
Max	96.90	99.36	99.43	99.22	98.89
	Model 26	Model 27	Model 28	Model 29	Model 30
Min	92.97	95.43	93.53	86.56	92.94
Median	95.19	98.12	96.45	94.59	95.34
Mean	95.39	97.94	96.21	94.22	95.65
Max	98.31	99.03	97.94	98.67	98.98
	Model 31	Model 32	Model 33		
Min	90.47	91.10	94.33		
Median	96.62	95.79	98.22		
Mean	96.02	95.65	97.77		
Max	99.76	99.02	99.89		

Table 5.2: Variance explained by FPCs in each model averaged over DoF.

### 5.1.4 GMMs

Finding proper parameters for GMMs was in some cases unsuccessful, due to the EM-algorithm diverging. In these cases a single two-dimensional Gaussian distribution was used instead, with parameters sample mean and sample covariance. The cases in which the EM-algorithm failed can be seen in table 5.1.

An example of how the probability distributions obtained for GMMs could look is presented in fig 5.6.

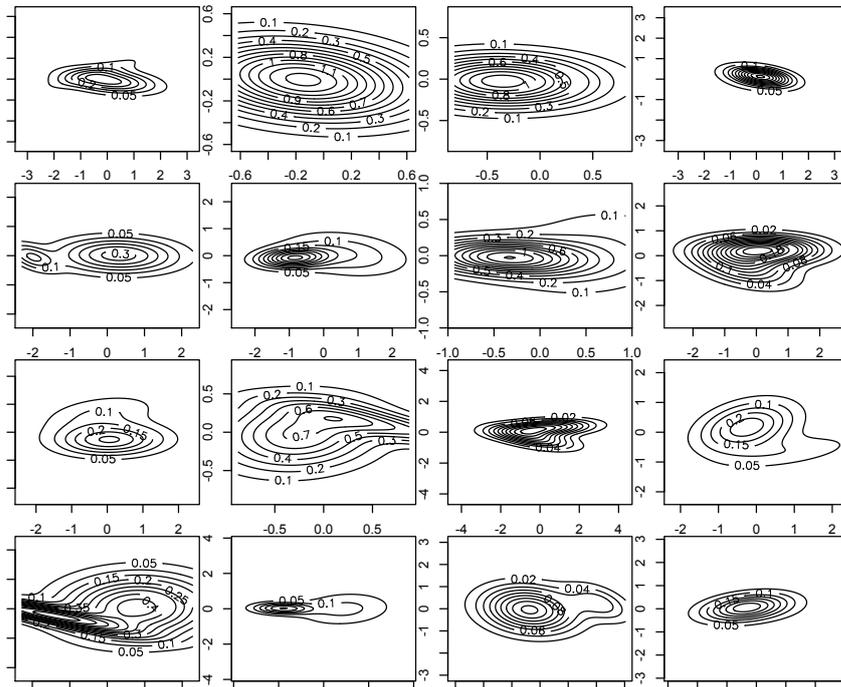


Figure 5.6: Example of probability distribution of GMMs. This plot shows probability distribution function level curves of GMMs for grasp type *writing tripod*. First DoF is shown in top left corner and following DoFs proceeds from left to right.

The probability distributions had means centered at zero. Weights set to zero corresponded to linear combinations of FPCs consisting only of the mean function.

From inspection of the probability distributions, it could be seen that the range of values likely to be obtained from the GMMs were relatively large, which indicated some variation in the training data and consequently variation in the generated weights.

Since the variance of weight values were relatively large, the FPCs were thus significant in the linear combinations for generated grasps; contrary to

weights being very small and generated grasps being mostly governed by the mean functions.

One likely source for this variation is that hand sizes and shapes differ, producing different sets of angles for the same task. Since the data set contains data from thirty different subjects, it is very likely that some sets of angles acted in unison in this way. Taking angle values from significantly different hand shapes might also produce incorrect hand configurations for the task to some degree.

In light of this, the absence of correlation between DoFs in the model was likely to have an effect on the generated data.

## 5.2 Generating grasps

The grasps generated by the models were in general similar to the grasp types they were supposed to mimic; they were able to capture distinct features. However, more subtle features were not captured. For example, this was the case for grasps that are only distinguished by how many fingers are in contact with the object, e.g. grasp types *prismatic 4 finger*, *prismatic 3 finger* and *prismatic 2 finger*. For comparison, appendix A contains images of the last frame of generated grasps next to grasps from the training data of each type.

In general, there were two issues with the grasps generated by the models. First, fingers did not move in unison, resulting in configurations where they were not equally bent. Second, in many cases generated grasps looked similar to generated grasps of other types, which made it hard to identify the specific grasp. The second issue is most likely a consequence of the first since fingers not moving in unison made some distinguishable features disappear. This was for example the case for grasp types *prismatic 4 finger*, *prismatic 3 finger* and *prismatic 2 finger* mentioned above.

As was mentioned earlier, it is important to note that the space represented by joint angles are not in one-to-one correlation with the hand configuration space when considering many subjects. This is due to differences in hand shapes and sizes, i.e. when two individuals perform the same type of grasp, the joint angles are not guaranteed to be the same. Thus it was not expected that joint angles alone could sufficiently describe a hand configuration. In light of this, the fact that each DoF was modeled separately is most likely an important factor for the fingers not moving in unison and for many grasp types being indistinguishable. This means that correlation between DoFs was not included in the model, and in extension that individual-specific variations were not accounted for.

In table 5.3, the mean  $L_2$ -distances between grasps and its projections are presented. In each row, the values have been divided by the frame length of the model to facilitate comparison between rows. This table gives an indication of how well the different models were performing. In twenty-

eight cases the model had its lowest mean distance for its own grasp type. In fourteen cases the grasp type had its lowest mean distance for its own model. And in twelve cases it happened simultaneously that the model had its lowest mean distance for its own grasp type and that the grasp type had its lowest mean distance for its own model. Worth noting is that nine grasp types had their lowest mean distance for model nine, *palmar pinch*, while sixteen models were not the minimum mean distance of any grasp type.

These observations indicate that the set of models does not amount to a very good distinction between grasps. From visual inspection, there were several groups of grasps that were hard to discern from one another. One such group of grasps that was not easily distinguishable from one another consists of grasp types *adducted thumb*, *light tool* and *fixed hook*. Another group consists of types *prismatic 4 finger*, *prismatic 3 finger*, *prismatic 2 finger* and *writing tripod*. A third group consists of grasp types *precision disk*, *precision sphere* and *tripod*.

Model	Grasp type																																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33			
1	7.34	11.62	11.16	12.33	13.88	9.73	9.72	11.02	12.81	14.54	10.30	9.26	10.65	10.79	12.90	14.10	12.67	9.96	8.23	12.37	12.07	12.07	9.51	11.38	10.69	9.37	10.86	8.87	18.40	13.46	10.34	15.92	10.05			
2	12.22	11.80	11.59	10.39	14.01	14.52	16.93	16.82	18.54	14.50	10.98	12.81	13.32	18.30	13.70	13.96	14.31	14.37	14.67	17.08	17.08	18.01	14.03	12.31	17.58	12.65	18.92	11.59	17.03	13.52	14.86	15.19	15.07			
3	11.77	11.95	10.24	12.90	12.96	15.46	15.63	16.90	18.48	14.64	12.62	12.31	16.72	15.87	12.26	17.09	13.43	13.31	12.65	15.83	15.67	16.76	12.50	17.29	15.78	11.64	17.21	10.79	15.94	12.64	13.93	14.22	13.82			
4	11.01	11.49	10.22	11.62	14.35	14.94	15.68	17.43	14.86	12.41	11.55	15.65	14.86	10.66	16.07	12.50	12.30	10.54	14.97	14.67	15.96	11.13	16.15	14.58	11.19	15.92	10.15	14.49	11.55	13.66	13.02	13.18				
5	13.02	13.39	12.26	16.13	8.91	15.27	16.04	16.01	16.62	13.85	14.24	11.61	16.68	14.97	10.33	10.48	10.15	8.16	7.69	9.98	11.11	9.77	7.45	7.11	7.04	7.84	6.70	7.77	13.73	11.19	10.06	12.32	9.04			
6	9.20	12.34	10.37	13.88	10.77	6.09	6.32	6.75	7.60	14.33	12.44	7.77	7.52	7.29	11.39	10.78	11.33	8.49	8.17	9.96	10.81	10.34	8.15	7.80	7.51	8.66	7.33	8.36	14.62	12.46	10.38	13.09	9.57			
7	9.55	13.07	11.24	13.60	12.19	6.79	6.49	7.08	8.00	15.07	12.44	7.77	7.52	7.29	11.39	10.78	11.33	8.49	8.17	9.96	10.81	10.34	8.15	7.80	7.51	8.66	7.33	8.36	14.62	12.46	10.38	13.09	9.57			
8	10.89	14.84	12.62	15.23	12.88	7.16	7.04	6.99	8.81	16.43	14.16	8.47	7.86	7.48	12.50	11.06	12.17	9.03	9.00	10.03	11.81	11.25	8.69	7.86	7.66	9.42	7.56	9.04	15.03	14.28	11.69	13.49	10.85			
9	8.04	12.70	11.42	13.79	11.82	6.53	6.09	6.09	6.23	14.40	11.03	6.00	6.70	6.10	11.40	10.93	11.19	7.56	6.75	8.83	11.25	7.39	6.62	6.47	6.62	7.23	6.17	7.52	16.63	12.61	8.02	14.41	6.77			
10	15.75	15.12	14.42	22.28	14.03	20.99	21.17	23.06	24.25	10.74	14.14	15.65	22.77	21.55	13.64	25.16	16.59	19.48	11.51	23.55	24.63	23.08	14.64	23.67	21.56	12.85	23.42	13.59	16.22	14.58	16.85	15.67	15.53			
11	9.69	12.33	10.63	13.57	13.23	12.81	12.57	14.51	16.58	12.76	6.53	10.84	13.82	13.94	12.63	15.80	12.77	12.32	9.03	14.56	14.60	15.61	11.45	14.82	13.68	9.91	14.30	9.46	17.54	13.03	11.51	15.40	11.45			
12	8.44	12.33	10.95	14.64	12.37	6.98	7.26	7.85	8.32	12.99	10.12	9.16	7.03	7.37	11.68	12.45	11.02	8.76	6.93	11.27	12.28	9.97	7.24	7.91	7.67	7.13	7.20	6.70	16.94	12.83	9.06	14.45	7.90			
13	8.58	12.62	10.90	14.26	11.89	5.88	6.10	6.51	6.89	13.77	10.86	6.70	7.19	9.86	10.85	9.02	6.70	7.19	9.86	10.85	9.02	6.70	6.34	6.52	7.48	5.70	7.66	15.91	12.00	9.07	13.70	8.03				
14	9.94	13.78	11.86	14.94	11.89	6.49	6.64	6.65	7.75	15.04	10.84	7.41	6.72	6.32	11.58	10.30	10.84	8.18	8.00	9.91	11.16	10.18	7.57	7.06	6.90	8.47	6.47	8.22	14.09	13.03	10.85	12.30	9.74			
15	12.64	12.97	11.78	13.34	9.61	15.69	16.28	16.70	18.07	14.04	13.85	12.19	17.32	15.79	9.15	18.09	12.86	13.73	10.43	17.34	17.14	17.76	10.72	17.57	15.17	11.20	17.41	10.81	12.46	10.59	14.73	11.75	13.36			
16	13.04	14.33	12.23	14.87	12.08	8.87	8.87	8.88	10.89	17.26	16.61	9.59	9.37	8.76	11.49	6.53	10.64	8.86	11.43	10.85	11.17	14.21	9.67	9.17	8.69	10.95	8.94	10.55	11.85	11.44	13.26	11.10	12.99			
17	12.99	15.40	14.36	16.14	14.39	14.09	14.24	14.76	16.19	15.35	14.14	11.78	15.00	14.20	13.72	16.93	12.32	11.17	16.43	15.70	15.90	11.70	15.72	13.90	11.82	15.40	11.51	15.26	14.04	16.06	11.47	14.86				
18	10.59	13.91	12.33	13.16	13.36	9.46	9.61	9.96	10.96	16.53	13.83	9.39	9.87	9.55	12.69	12.46	11.57	8.10	11.81	11.01	11.45	9.28	10.02	9.74	10.27	10.03	9.79	15.77	12.42	11.59	13.50	11.39				
19	11.08	14.48	12.79	16.77	12.21	10.37	10.68	11.42	12.43	15.11	13.74	10.23	10.87	10.88	12.13	15.28	13.20	11.71	7.91	14.71	15.51	13.46	10.16	11.61	11.32	10.04	10.89	9.64	15.39	13.80	11.87	14.20	10.68			
20	12.44	14.81	12.51	15.17	13.44	9.32	9.05	9.66	11.77	16.97	15.46	9.82	9.78	9.60	12.57	11.73	11.98	9.44	10.91	9.89	10.30	13.70	10.38	9.99	9.29	11.08	9.56	10.69	13.86	12.69	13.61	12.39	13.41			
21	14.07	16.69	14.28	16.11	15.15	11.33	11.02	11.83	14.45	18.40	16.89	11.56	11.99	11.89	14.40	14.31	13.28	10.84	11.95	12.08	10.97	15.42	11.85	12.04	11.36	12.81	11.86	12.41	15.64	14.48	15.45	14.06	15.55			
22	9.14	16.48	15.63	14.34	16.87	9.20	9.54	9.26	8.25	18.19	12.68	8.26	8.09	8.53	15.96	15.68	13.96	6.79	7.20	14.16	11.41	5.54	7.72	8.44	9.11	10.30	8.58	10.14	23.93	17.14	9.21	18.78	8.47			
23	9.48	12.44	10.97	12.78	10.65	8.56	8.96	9.00	9.97	13.86	11.45	8.07	9.15	8.70	10.51	11.42	10.88	8.87	7.57	11.35	11.69	11.15	7.08	9.34	8.58	8.60	8.90	8.25	13.83	11.54	10.44	12.03	9.25			
24	9.17	13.24	11.40	14.31	12.07	6.07	6.43	6.43	6.73	15.30	12.48	6.80	6.25	6.37	11.63	10.86	11.49	7.94	7.61	10.74	11.22	9.24	7.90	6.95	6.80	8.20	6.19	8.34	15.95	12.67	9.08	14.25	8.24			
25	10.12	13.50	11.49	13.90	11.79	6.68	6.50	6.53	8.13	15.25	13.15	7.46	7.13	6.67	11.08	9.64	10.27	7.80	8.33	6.70	10.55	7.64	7.01	6.46	8.58	6.67	8.46	13.06	12.10	11.36	11.21	10.48				
26	10.18	12.49	11.11	16.61	11.37	10.14	10.48	11.84	12.22	12.45	11.48	8.62	11.15	10.78	11.24	14.96	12.48	11.76	8.04	14.56	16.24	13.38	9.19	12.20	11.20	7.65	6.94	7.65	9.94	11.33	7.96	15.33	12.87	11.26	13.92	9.83
27	9.19	12.75	10.92	14.53	14.49	5.94	6.04	6.44	7.07	14.08	12.05	6.49	5.89	6.12	10.89	10.65	10.57	7.94	7.39	9.62	10.90	9.72	7.22	6.49	6.46	7.65	6.94	7.56	14.31	11.87	9.93	12.51	8.87			
28	11.42	13.36	11.87	16.91	11.83	11.56	11.81	12.76	13.52	14.54	13.44	9.96	12.50	11.82	11.91	15.83	12.91	11.97	9.45	15.39	16.48	14.20	10.36	13.30	12.27	9.35	12.57	8.90	14.86	13.46	12.63	13.43	11.54			
29	14.49	14.95	13.74	16.50	11.81	15.19	15.54	16.66	17.30	16.70	16.85	13.00	17.04	16.35	11.84	18.15	13.41	13.29	12.85	17.58	16.66	17.30	12.07	16.77	15.13	12.57	17.17	11.74	11.53	12.98	10.49	15.09	12.17	14.25		
30	13.28	13.17	12.17	14.06	11.68	15.94	16.29	16.83	18.23	14.69	14.31	12.75	17.34	16.05	11.04	17.32	12.85	13.31	11.85	16.84	16.25	17.62	11.38	17.38	15.33	12.18	17.15	11.63	12.98	10.49	15.09	12.17	14.25			
31	8.93	13.24	12.53	15.85	15.39	9.59	9.46	10.01	10.48	15.32	10.61	8.17	8.92	9.62	14.57	15.22	15.78	9.93	7.87	13.87	13.86	10.20	8.95	9.57	9.63	9.08	9.24	9.11	22.07	14.92	6.77	19.37	7.05			
32	13.00	14.85	14.03	14.98	12.12	14.73	15.19	15.47	16.52	15.79	14.87	12.21	16.36	14.88	11.89	17.72	10.53	12.48	11.10	17.26	15.87	15.81	10.89	16.42	14.62	11.55	16.72	11.18	12.11	12.59	16.44	9.43	15.23			
33	8.09	12.63	11.75	15.73	13.83	7.69	7.89	8.14	7.79	14.17	10.15	6.59	6.95	7.45	13.30	13.94	14.54	8.87	6.61	12.83	13.06	8.72	7.22	7.37	7.88	7.43	7.13	7.86	20.71	14.49	6.73	18.31	9.05			

Table 5.3: Table showing mean of  $L_2$ -distances of grasp types and its projections. The values was multiplied by a factor of hundred and divided by model frame length,  $100 \cdot \|g - \text{proj}_i g\|/T_i$  for the purposes of easier reading and comparison between cells respectively. Rows represent statistical model and columns represent grasp type. The cell defined by e.g. row five, column four shows the mean  $L_2$  distance when all grasps of type four was fed into the model created for grasp type five. Comparing this column with the cell defined by row four, column five shows that the classification is not symmetric. The lowest value for each statistical model has been highlighted. The lowest value for each grasp type has been framed.

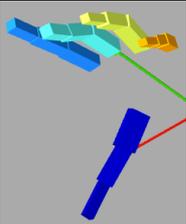
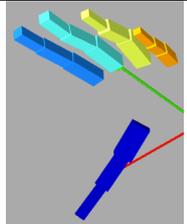
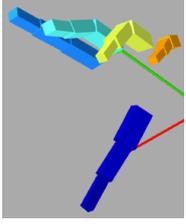
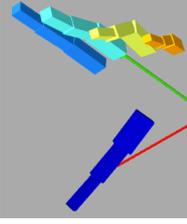
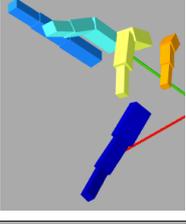
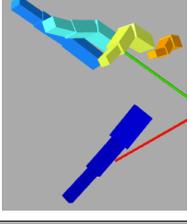
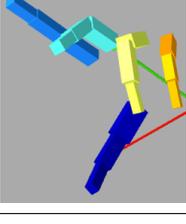
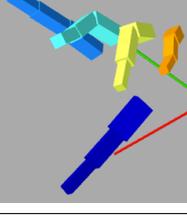
### 5.2.1 Temporal comparison

Disregarding how well the grasp type was captured by the model, the temporal aspect was in general well captured in the models. There was no significant issues in this aspect, like e.g. fluctuations or jerkiness. The generated grasps lacked a little bit in smoothness compared to the training data.

Two grasp types have been selected for illustrating temporal comparison.

The first case, which is shown in table 5.4, shows grasp type *index finger extension*, in which the grasp type was successfully modelled. This case fits well in the overall description above, since there is a distinctive feature which was well captured, namely the extended index finger.

The other case, which is shown in table 5.5, shows grasp type *lateral tripod*, in which the grasp type was not successfully modelled. This case gives an example of how small variations in bending between fingers was not well captured.

	Data sample	Generated sample
1		
2		
3		
4		

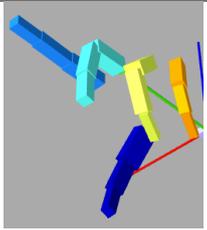
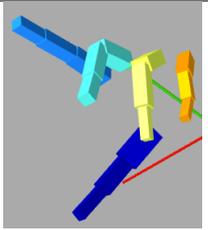
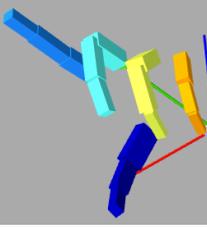
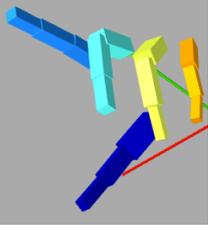
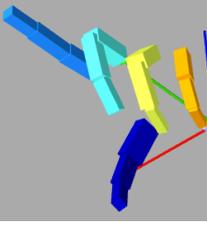
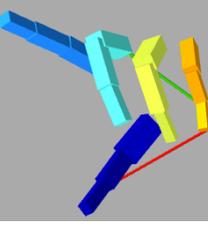
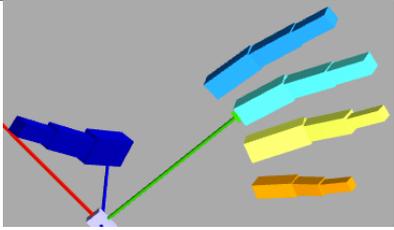
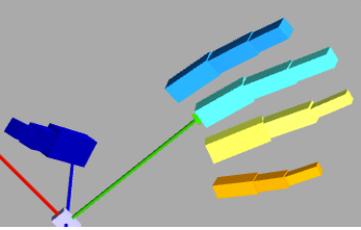
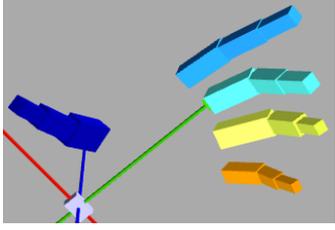
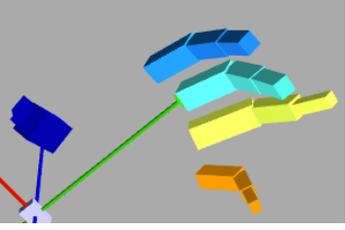
5		
6		
7		

Table 5.4: Temporal comparison between data sample and generated sample of grasp type *index finger extension*. In this comparison it can be seen that the generated sample did well in capturing the distinct feature of the grasp.

	Data sample	Generated sample
1		
2		

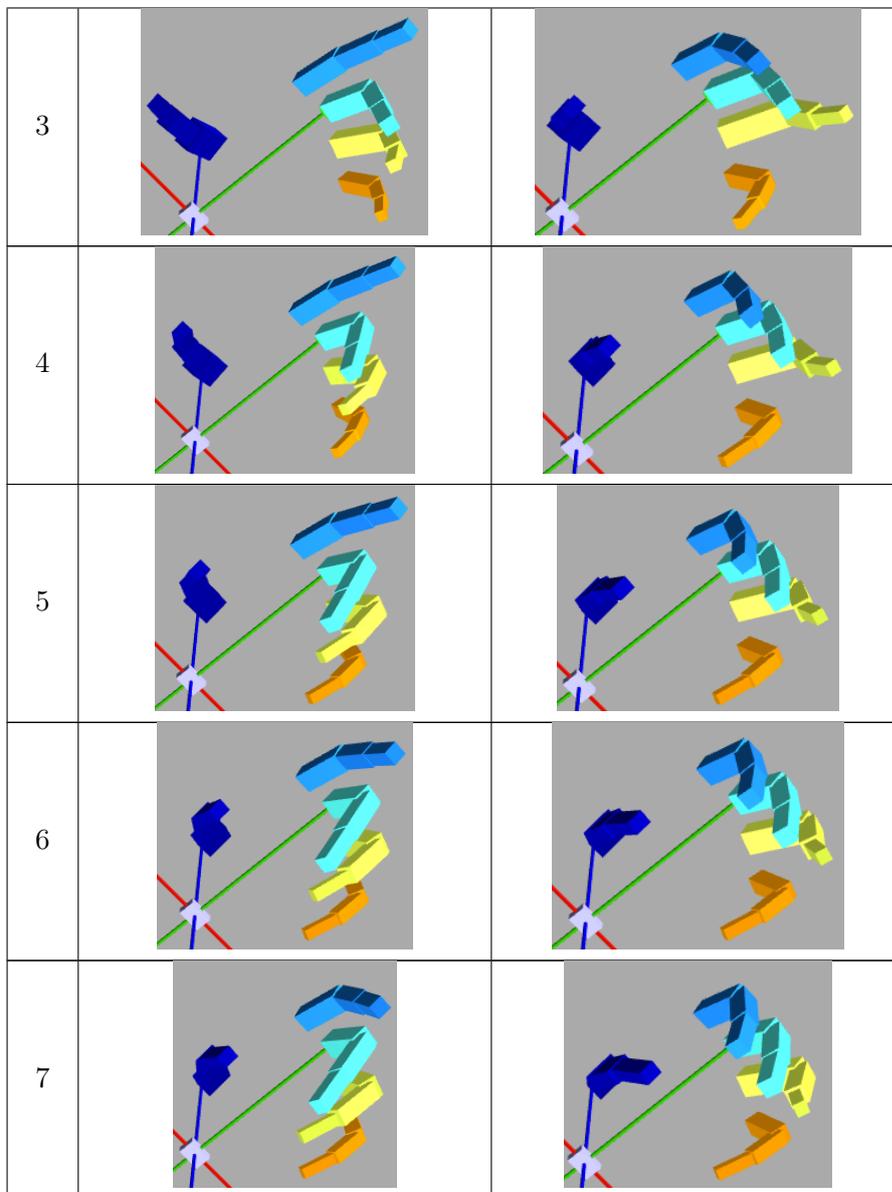


Table 5.5: Temporal comparison between data sample and generated sample of grasp type *lateral tripod*. In this comparison it can be seen that the generated grasp failed to capture the variations in bending between fingers.

### 5.3 Classification

In this section the results of the two classification approaches using FPCA and last frame PCA are presented.

### 5.3.1 FPCA

Using the models for classification was not successful. For some grasp types, grasps were more often classified outside their own type than not. It was however expected that the classification would perform badly for two reasons. First, in the classification procedure, all DoFs were weighted equally, so that the overall hand posture was considered for classification. This approach was clearly not proficient at identifying distinctive features of each grasp. Second, the FPCA-based classification used the whole temporal movement of the hand, but grasp types are identified by the final configuration only and much of the initial movement was similar for all the grasps. Thus the identifying posture was a smaller part of the movement. The confusion matrix for classification is presented in table 5.6 below.

The worst case was grasp type *stick*, which was only correctly classified six times. The most successful case was grasp *parallel extension*, with two-hundred and fifteen correctly classified grasps. These results are reasonable when comparing to the other grasps; *parallel extension* is very distinct in the sense that there are no other grasps in which almost all fingers are fully extended, thus creating a big difference in terms of joint angles. Grasp type *stick* on the other hand has a similar configuration to many other grasp types, e.g. *light tool* and *fixed hook*.

As in table 5.3, some models are overrepresented in classification, most prominently model nine, *palmar pinch*, which nine grasp types were most often classified as. There were however less grasp types more often classified outside their own model in the confusion matrix compared to how many grasp types had their lowest mean distance outside their own model.

In general, the confusion matrix in table 5.6 shows a similar pattern as in table 5.3 above. Grasp types classified in models correspond in general to grasp types having low mean  $L_2$ -distances in that model compared to others. This is expected since the the  $L_2$ -distance between grasps and their projections was used in both cases. Numbers were not divided by frame length in the confusion matrix as in table 5.3.

Model	Grasp type																																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33		
1	155	24	8	7	0	2	6	1	0	0	57	2	1	0	0	1	0	2	3	1	0	0	7	0	0	1	0	0	2	2	0	0	0	0	
2	3	61	14	5	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2	0	0	0	0		
3	2	18	81	12	3	0	0	0	0	7	10	0	0	0	0	1	0	0	0	1	1	0	0	1	0	0	0	0	2	2	0	0	0		
4	2	6	11	63	0	0	0	1	0	0	2	0	0	0	6	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0		
5	0	2	4	9	140	0	0	0	1	17	3	0	0	0	53	5	0	0	19	1	3	0	18	2	1	16	0	25	38	23	0	5	0		
6	8	0	0	0	0	27	26	7	2	0	1	6	9	9	1	0	1	7	0	1	0	1	5	6	1	5	0	0	1	0	0	1	0		
7	4	0	0	0	0	3	20	4	1	0	0	1	0	0	0	1	0	1	0	1	0	0	0	0	1	0	1	0	1	0	0	0	0		
8	2	5	4	7	3	8	20	53	1	3	1	0	0	9	6	1	3	1	7	4	5	0	2	4	21	0	2	0	8	0	0	4	0		
9	26	8	13	2	15	88	60	81	183	6	9	59	83	94	3	19	9	12	65	21	3	29	52	113	56	35	93	43	8	3	22	2	37		
10	0	0	0	0	0	0	0	0	0	120	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
11	4	16	13	0	0	0	0	0	0	3	61	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	3	0	0	0	0	0		
12	4	0	0	0	2	3	0	0	0	3	2	43	11	3	0	0	1	0	1	2	0	0	0	5	0	15	3	5	0	0	0	1	0		
13	2	0	0	0	0	6	2	2	1	0	0	22	55	4	0	0	0	0	0	1	0	1	1	1	1	1	16	2	0	1	0	0	0		
14	4	0	3	0	1	8	11	18	1	0	0	5	13	48	0	13	1	6	13	7	0	13	2	14	0	20	7	0	1	0	3	0	0		
15	1	32	35	82	48	0	5	1	0	29	7	0	0	0	133	4	23	30	4	7	0	13	0	1	11	0	6	59	60	1	0	3	0		
16	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	46	0	1	0	2	0	1	0	0	0	0	0	2	1	0	0	0	0		
17	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	166	1	0	0	0	5	0	0	0	0	0	0	0	4	1	0	19	0	0	
18	7	12	6	10	1	6	5	12	2	1	3	2	2	2	8	31	19	95	6	40	80	2	17	7	11	0	1	12	10	14	3	11	0	1	
19	1	2	1	0	1	1	1	0	0	3	1	0	0	2	0	1	0	0	26	0	1	0	0	0	0	2	1	1	1	2	0	0	0	1	0
20	0	2	3	3	0	0	0	1	0	0	1	0	0	1	2	0	2	0	0	0	0	0	0	0	0	0	0	0	1	3	0	0	1	0	0
21	0	1	4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	10	0	0	0	0	0	0	0	2	1	0	0	0	0	
22	2	1	2	4	1	5	12	6	6	0	2	4	5	4	0	4	49	4	14	19	215	2	1	2	1	1	1	2	0	1	3	4	1	0	0
23	4	0	1	4	3	9	6	8	6	0	5	4	6	6	5	8	4	8	1	6	2	0	64	4	4	5	1	7	6	10	0	7	0	0	
24	4	4	1	11	7	28	27	18	25	1	0	25	29	20	4	31	0	19	18	27	18	1	15	69	30	4	17	4	13	6	8	3	5	0	
25	4	2	1	9	2	7	26	26	3	1	0	5	2	33	5	49	8	19	10	55	52	0	27	10	87	4	20	3	12	6	0	9	0	0	
26	10	22	14	10	4	10	4	1	1	31	40	12	30	21	0	11	1	3	6	10	2	0	4	2	9	7	63	0	2	1	0	0	0	0	
27	2	0	0	0	1	23	14	8	0	0	0	5	3	0	0	2	0	0	0	0	0	0	0	0	10	1	60	1	0	0	0	0	0	0	
28	0	5	7	2	1	0	1	0	0	0	3	0	0	0	0	0	3	8	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	13	8	6	5	1	1	0	0	19	6	0	0	0	14	7	0	0	5	2	6	5	0	4	0	1	1	0	7	14	102	0	1	0	0
31	0	2	0	0	0	0	0	0	0	2	8	0	0	0	0	1	0	0	1	0	1	0	0	2	0	0	3	0	0	0	0	156	0	53	0
32	1	4	2	5	1	0	0	0	0	1	0	0	0	2	1	6	30	1	4	1	0	0	2	0	0	1	1	46	1	0	171	0	53	0	
33	6	0	6	1	5	10	9	6	8	2	12	9	6	5	2	8	0	1	17	1	0	15	13	1	11	0	8	4	7	57	0	149	0	0	

Table 5.6: Confusion matrix for FPCA.

### 5.3.2 Last frame PCA

PCA-based classification performed much better than FPCA-based classification. This was expected since only the final position of the hand was considered, which is also the position that identifies the grasp type. Five PCs were used for the PCA-based classification. A confusion matrix of the classification is found in table 5.8 below.

In table 5.7 the cumulative sum of variance explained by using up to all sixteen PCs is shown. It can be seen that using five PCs amount to explaining more than eighty percent of the variance for most grasp types.

Grasp type	Number of PCs															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0.555	0.695	0.755	0.806	0.848	0.881	0.912	0.936	0.955	0.967	0.976	0.984	0.990	0.994	0.997	1.000
2	0.363	0.576	0.675	0.754	0.811	0.860	0.898	0.921	0.944	0.964	0.974	0.982	0.989	0.994	0.997	1.000
3	0.384	0.599	0.687	0.748	0.802	0.846	0.883	0.918	0.941	0.961	0.972	0.981	0.989	0.994	0.998	1.000
4	0.295	0.471	0.598	0.684	0.753	0.810	0.859	0.903	0.927	0.948	0.961	0.974	0.983	0.990	0.995	1.000
5	0.327	0.611	0.718	0.781	0.834	0.880	0.918	0.944	0.961	0.972	0.980	0.986	0.992	0.995	0.998	1.000
6	0.478	0.619	0.696	0.765	0.830	0.882	0.908	0.931	0.949	0.964	0.976	0.984	0.990	0.994	0.997	1.000
7	0.412	0.615	0.721	0.784	0.833	0.876	0.907	0.934	0.953	0.968	0.978	0.986	0.991	0.994	0.997	1.000
8	0.451	0.683	0.760	0.825	0.862	0.891	0.918	0.938	0.956	0.968	0.977	0.985	0.992	0.995	0.998	1.000
9	0.543	0.692	0.807	0.862	0.896	0.921	0.941	0.955	0.966	0.975	0.983	0.988	0.993	0.997	0.998	1.000
10	0.337	0.540	0.648	0.735	0.802	0.854	0.897	0.922	0.942	0.962	0.972	0.980	0.987	0.994	0.997	1.000
11	0.444	0.635	0.719	0.782	0.826	0.866	0.900	0.926	0.947	0.961	0.971	0.980	0.987	0.993	0.997	1.000
12	0.309	0.562	0.678	0.751	0.814	0.853	0.890	0.923	0.942	0.960	0.974	0.982	0.990	0.994	0.997	1.000
13	0.392	0.583	0.687	0.773	0.826	0.865	0.899	0.928	0.948	0.965	0.973	0.981	0.988	0.993	0.997	1.000
14	0.436	0.626	0.747	0.803	0.849	0.883	0.910	0.931	0.951	0.963	0.975	0.984	0.989	0.993	0.997	1.000
15	0.504	0.753	0.806	0.849	0.889	0.914	0.939	0.955	0.967	0.975	0.982	0.987	0.992	0.995	0.999	1.000
16	0.389	0.567	0.674	0.751	0.814	0.854	0.889	0.918	0.942	0.958	0.970	0.979	0.986	0.992	0.997	1.000
17	0.315	0.536	0.649	0.754	0.811	0.851	0.890	0.918	0.943	0.961	0.973	0.983	0.988	0.993	0.997	1.000
18	0.491	0.668	0.754	0.814	0.862	0.898	0.929	0.947	0.962	0.972	0.982	0.988	0.992	0.996	0.998	1.000
19	0.401	0.656	0.752	0.804	0.847	0.880	0.910	0.932	0.950	0.963	0.973	0.981	0.987	0.993	0.997	1.000
20	0.332	0.533	0.644	0.730	0.788	0.832	0.872	0.904	0.931	0.950	0.966	0.978	0.987	0.993	0.997	1.000
21	0.311	0.508	0.634	0.723	0.781	0.830	0.875	0.907	0.938	0.956	0.972	0.981	0.988	0.995	0.998	1.000
22	0.496	0.615	0.717	0.789	0.839	0.874	0.901	0.924	0.941	0.954	0.966	0.976	0.985	0.990	0.996	1.000
23	0.427	0.602	0.693	0.766	0.825	0.875	0.908	0.935	0.953	0.967	0.977	0.985	0.991	0.994	0.998	1.000
24	0.583	0.749	0.837	0.873	0.899	0.924	0.943	0.955	0.966	0.975	0.983	0.989	0.993	0.996	0.999	1.000
25	0.350	0.623	0.716	0.786	0.832	0.872	0.897	0.920	0.939	0.956	0.970	0.981	0.987	0.993	0.997	1.000
26	0.377	0.629	0.716	0.786	0.827	0.867	0.902	0.927	0.949	0.965	0.975	0.983	0.990	0.995	0.998	1.000
27	0.343	0.574	0.674	0.755	0.815	0.855	0.890	0.920	0.944	0.960	0.973	0.982	0.988	0.993	0.997	1.000
28	0.398	0.663	0.740	0.797	0.842	0.877	0.907	0.935	0.954	0.967	0.977	0.985	0.992	0.996	0.998	1.000
29	0.376	0.579	0.671	0.738	0.805	0.855	0.888	0.920	0.940	0.957	0.970	0.980	0.988	0.993	0.998	1.000
30	0.487	0.653	0.761	0.830	0.877	0.911	0.931	0.950	0.965	0.974	0.981	0.988	0.992	0.996	0.998	1.000
31	0.272	0.505	0.648	0.735	0.804	0.848	0.886	0.908	0.928	0.948	0.965	0.978	0.986	0.993	0.997	1.000
32	0.297	0.571	0.657	0.731	0.788	0.835	0.872	0.903	0.927	0.946	0.963	0.975	0.985	0.991	0.996	1.000
33	0.389	0.544	0.659	0.749	0.800	0.849	0.885	0.909	0.929	0.948	0.965	0.977	0.987	0.992	0.997	1.000

Table 5.7: Table showing the cumulative sum of variance explained by PCs for last frame PCA of each grasp type.

Model	Grasp type																																								
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33								
1	191	24	20	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
2	12	148	28	2	2	5	9	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
3	11	42	166	3	5	0	1	0	0	0	0	0	0	0	0	11	0	1	4	1	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
4	1	1	7	168	27	0	0	0	0	0	0	0	0	0	39	0	0	4	1	1	0	0	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
5	0	0	11	5	145	0	0	0	1	0	0	0	0	0	15	0	0	0	0	0	0	0	8	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
6	8	2	0	0	0	159	14	0	0	0	0	0	0	2	1	0	3	16	0	0	0	2	1	0	3	3	5	0	0	0	0	0	0	0	0	0	0	0	0		
7	1	0	0	0	0	4	173	18	8	2	2	2	0	0	0	2	0	2	0	2	5	0	0	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
8	0	1	0	0	0	0	0	0	0	2	139	8	0	0	0	0	0	10	15	2	0	2	0	2	15	21	1	0	0	0	0	0	0	0	0	0	0	0	0		
9	3	0	0	0	0	1	0	0	1	0	0	0	0	20	0	0	0	0	0	2	0	0	0	0	7	7	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
10	0	0	0	0	0	2	5	0	0	2	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
11	2	1	0	0	0	19	7	3	0	0	5	8	134	14	11	0	0	0	0	0	0	0	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
12	0	2	0	0	0	4	2	0	0	4	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
13	0	0	0	0	0	9	2	0	0	9	2	0	0	6	22	34	134	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
14	0	0	0	1	0	0	8	1	14	3	0	0	0	0	0	0	0	4	28	2	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
15	0	0	0	0	44	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
16	0	0	2	0	0	17	0	0	1	1	0	0	0	0	0	1	196	5	0	0	0	0	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	228	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
18	5	4	3	10	1	0	1	0	3	0	0	0	0	0	0	0	144	1	15	15	11	9	5	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
19	0	3	4	1	1	10	2	15	10	0	1	0	0	3	23	1	1	3	9	90	3	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
20	1	1	0	2	0	0	0	2	2	0	0	0	0	0	0	6	0	8	8	110	37	0	11	4	34	0	0	8	5	0	1	14	0	6	0	0	0	0	0	0	
21	8	1	1	0	0	1	1	8	0	0	1	8	0	0	0	0	27	1	9	145	0	11	3	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
22	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
23	0	0	0	4	8	5	10	4	0	0	3	0	0	2	4	5	0	2	2	1	1	0	173	3	10	2	0	3	4	5	0	0	0	0	0	0	0	0	0	0	
24	0	0	1	0	0	0	0	0	0	0	3	45	0	0	0	0	0	1	17	2	0	8	136	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
25	0	0	4	0	0	3	4	16	13	0	0	0	0	0	24	0	2	2	20	21	10	0	2	109	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	3	0	0	0	0	3	0	0	0	0	0	12	29	24	28	0	0	3	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
27	10	0	0	0	0	6	24	8	3	3	2	6	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
28	1	7	1	2	1	0	18	10	0	0	0	0	0	0	0	0	0	1	2	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
29	1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	7	0	4	0	19	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
30	0	0	0	6	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
31	0	0	2	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
32	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 5.8: Confusion matrix for last frame PCA classification using five PCs. Rows represent the type PCA was performed on, columns represent grasp type fed into the model.

# Chapter 6

## Summary

In this report a statistical approach was taken to generate natural looking grasp motions in a Virtual Reality context.

Training was made on data from the thirty-three grasp types of Feix’s grasp taxonomy, and one model was built for each grasp type. The data consisted of recordings of sixteen joint angles measured at fifty hertz. Grasping motions of thirty subjects were included in the data.

Each model was created by first applying DTW to the data, then FPCA was conducted, and finally GMMs were trained in order to be able to generate grasps. In every model, each DoF was modeled separately.

The models were in general good at capturing distinctive features of grasp types, but failed to capture more subtle features. The failing manifested itself as an inability to produce finger movements moving in uniform. This made grasp types with subtle differences hard to distinguish and made grasps defined by how many fingers were in contact with the object in many cases incorrect.

The underlying reason for failure to capture subtle features was likely due to the fact that hand sizes and shapes differ, producing different sets of angles for the same task between subjects. Since FPCA resulted in a mean 95.9 percent explained of the variability in the data, successful grasp generation was most likely not restricted by FPCs’ inability to reproduce certain grasp configurations lying outside FPC-space. Instead, the fact that each DoF was modelled separately was most likely the reason for failed grasp generation. Since each DoF was modeled in isolation, no correlation between DoFs was included in the models, thus differences in hand shapes and sizes were not accounted for.

For future models, some correlation between DoFs should be included in the model to be able to generate fingers moving in unison and generate satisfying grasps for grasps types with subtle differences from each other.

A classification scheme was conducted where  $L_2$ -distances was calculated between a grasp and its projection on the FPCs in each model. For com-

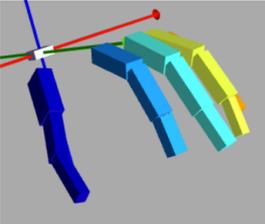
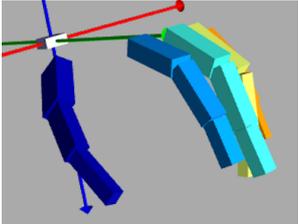
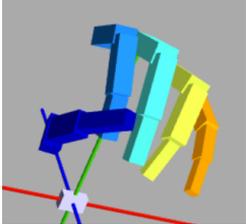
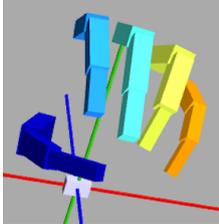
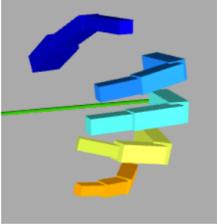
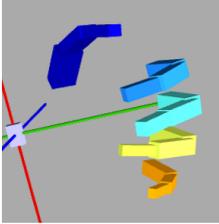
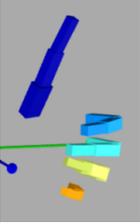
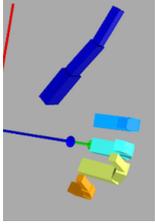
parison the Euclidean distance between the last frame of a grasp and its projection on the first five PCs of each model was conducted. The PCA-based scheme clearly outperformed the FPCA-based scheme, which did not perform well.

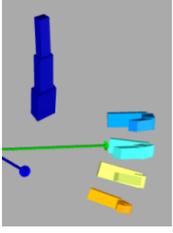
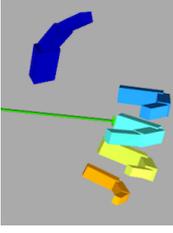
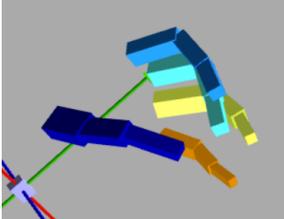
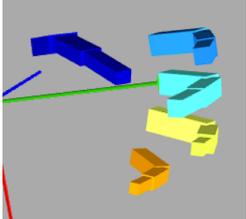
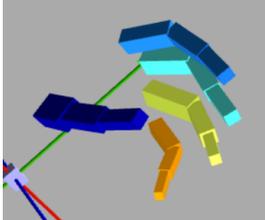
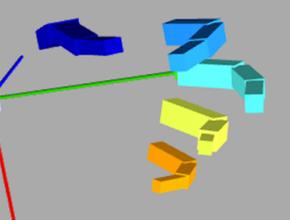
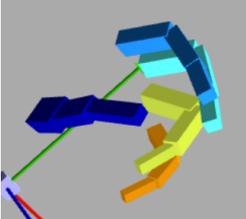
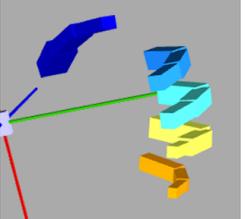
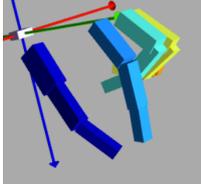
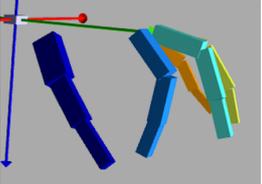
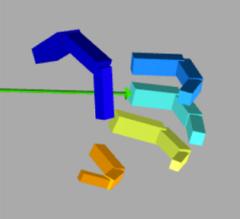
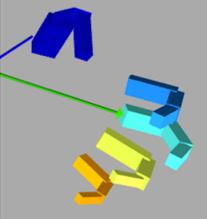
It was however not surprising that the classification based on FPCA did not work well for two reasons. First, all DoFs were weighted equally, thus the overall hand posture was considered for classification instead of weighing the distinctive DoFs more than the less distinctive DoFs. Second, the whole temporal movement of the hand was included in the FPCs, but grasp types are identified by the final position only, and much of the initial movements were similar for many grasps. Thus the identifying posture was a smaller part of the movement.

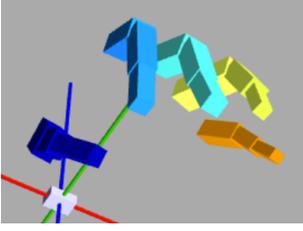
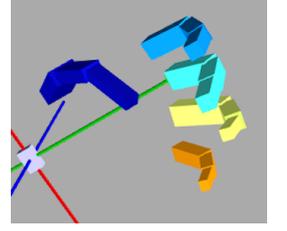
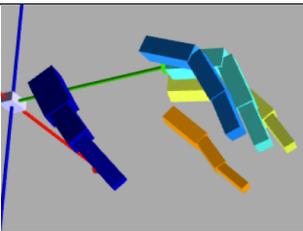
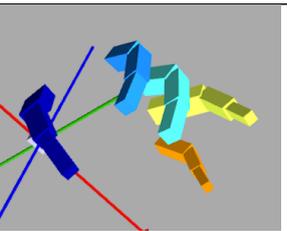
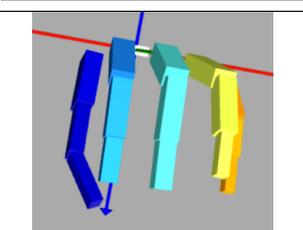
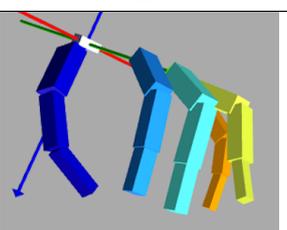
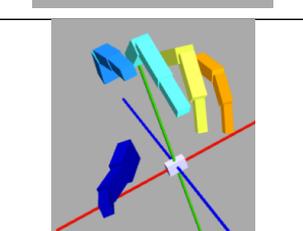
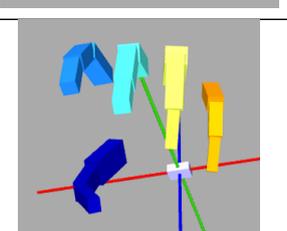
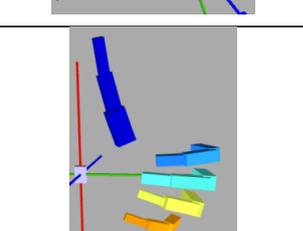
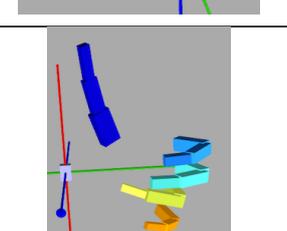
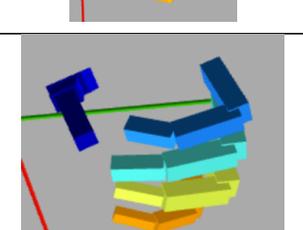
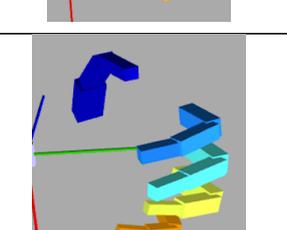
To remedy this failure, a future classification based on FPCA could benefit from using larger weights on DoFs important for the grasp type. Another improvement could be to use the final posture for classification, which was shown in the PCA-based classification to be a more favorable approach.

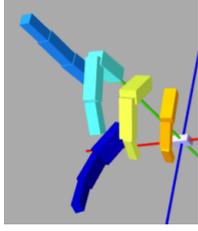
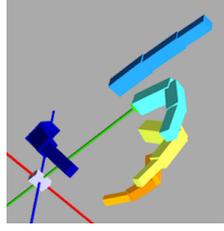
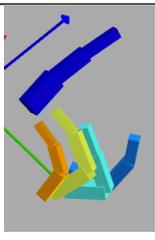
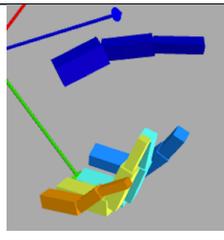
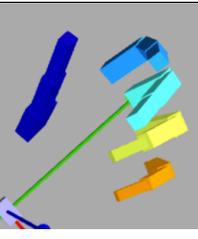
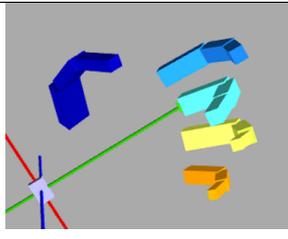
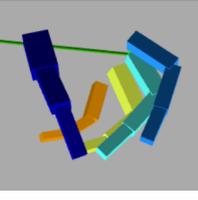
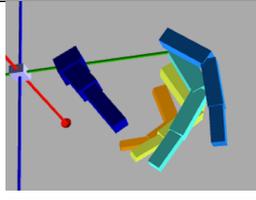
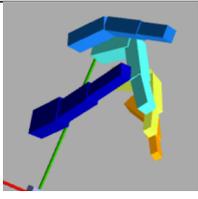
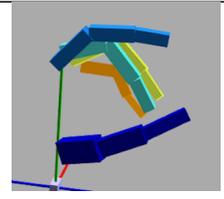
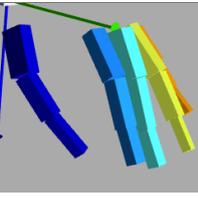
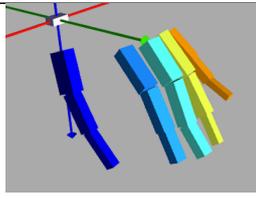
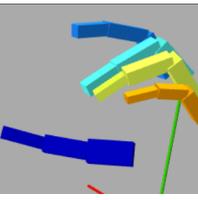
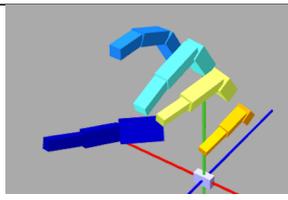
# Appendix A

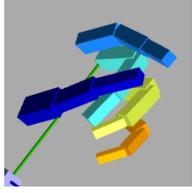
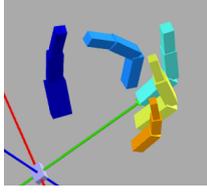
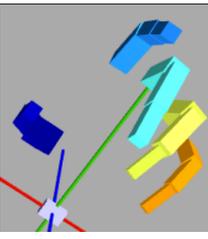
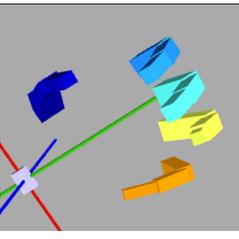
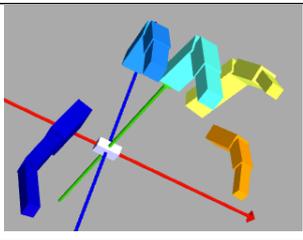
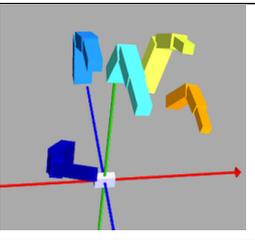
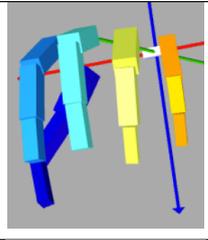
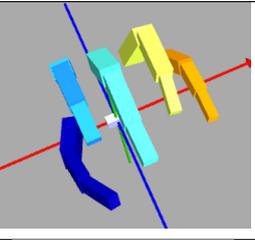
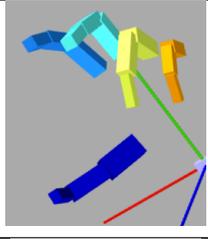
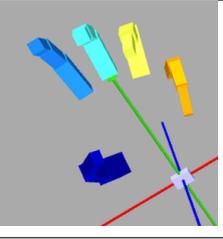
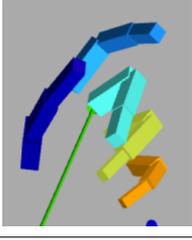
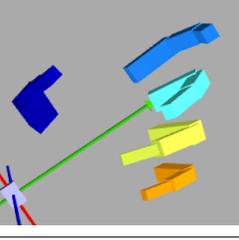
## Generated grasps

	Grasp type	Data sample	Generated sample
1	Large diameter		
2	Small diameter		
3	Medium wrap		
4	Adducted thumb		

5	Light tool		
6	Prismatic 4 finger		
7	Prismatic 3 finger		
8	Prismatic 2 finger		
9	Palmar pinch		
10	Power disk		

11	Power sphere		
12	Precision disk		
13	Precision sphere		
14	Tripod		
15	Fixed hook		
16	Lateral		

17	Index finger extension		
18	Extension type		
19	Distal type		
20	Writing tripod		
21	Tripod variation		
22	Parallel extension		
23	Adduction grip		

24	Tip pinch		
25	Lateral tripod		
26	Sphere 4 finger		
27	Quadpod		
28	Sphere 3 finger		
29	Stick		

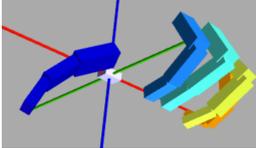
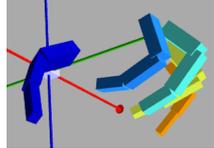
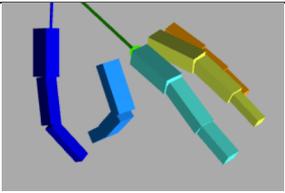
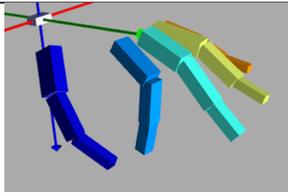
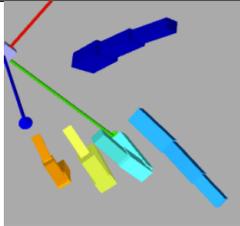
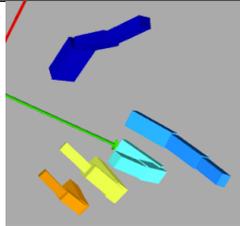
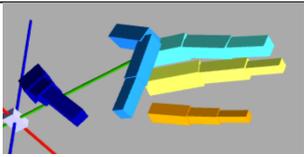
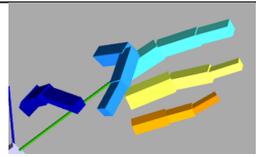
30	Palmar 4 finger		
31	Ring		
32	Ventral		
33	Inferior pincer		

Table A.1: Comparison of training data and generated samples. Middle column contains images of the last frame of training data. Rightmost column contains images of the last frame of generated grasps.

# Bibliography

- [1] Hiroaki Sakoe and Seibi Chiba. “Dynamic programming algorithm optimization for spoken word recognition”. In: *IEEE transactions on acoustics, speech, and signal processing* 26.1 (1978), pp. 43–49.
- [2] M. R. Cutkosky. “On grasp choice, grasp models, and the design of hands for manufacturing tasks”. In: *Robotics and Automation, IEEE Transactions on* 5.3 (1989), pp. 269–279.
- [3] Marco Santello, Martha Flanders, and John F. Soechting. “Postural Hand Synergies for Tool Use”. In: *The Journal of Neuroscience* 18.23 (Dec. 1998), pp. 10105–10115. URL: <http://www.jneurosci.org/content/18/23/10105.abstract>.
- [4] N. Somia et al. “The Initiation and Sequence of Digital Joint Motion”. In: *Journal of Hand Surgery* 23.6 (1998), pp. 792–795.
- [5] Neil Lawrence. “Gaussian Process Latent Variable Models for Visualisation of High Dimensional Data”. In: *In NIPS*. 2003, p. 2004.
- [6] Keith Grochow et al. “Style-based inverse kinematics”. In: *ACM transactions on graphics (TOG)*. Vol. 23. 3. ACM. 2004, pp. 522–531.
- [7] Raquel Urtasun et al. “Priors for people tracking from small training sets”. In: *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*. Vol. 1. IEEE. 2005, pp. 403–410.
- [8] Jack Wang, Aaron Hertzmann, and David M Blei. “Gaussian process dynamical models”. In: *Advances in neural information processing systems*. 2006, pp. 1441–1448.
- [9] Matei Ciocarlie, Corey Goldfeder, and Peter Allen. “Dexterous Grasping via Eigengrasps: A Low-Dimensional Approach to a High-Complexity Problem”. In: (2007). URL: <http://www.coreygoldfeder.com/papers/RSS07.pdf>.
- [10] Ying Li, Jiaxin L. Fu, and Nancy S. Pollard. “Data-Driven Grasp Synthesis Using Shape Matching and Task-Based Pruning.” In: *IEEE Trans. Vis. Comput. Graph.* 13.4 (2007), pp. 732–747. URL: <http://dblp.uni-trier.de/db/journals/tvcg/tvcg13.html#LiFP07>.

- [11] Graham W. Taylor, Geoffrey E. Hinton, and Sam T. Roweis. “Modeling Human Motion Using Binary Latent Variables.” In: *NIPS*. Ed. by Bernhard Schölkopf, John C. Platt, and Thomas Hofmann. MIT Press, 2007, pp. 1345–1352. ISBN: 0-262-19568-2. URL: <http://dblp.uni-trier.de/db/conf/nips/nips2006.html#TaylorHR06>.
- [12] Jack M. Wang, David J. Fleet, and Aaron Hertzmann. “Multifactor Gaussian Process Models for Style-content Separation”. In: *Proceedings of the 24th International Conference on Machine Learning*. ICML '07. Corvallis, Oregon, USA: ACM, 2007, pp. 975–982. ISBN: 978-1-59593-793-3.
- [13] Heni B. Amor et al. “Grasp synthesis from low-dimensional probabilistic grasp models”. In: *Computer Animation and Virtual Worlds* 19.3-4 (2008), pp. 445–454. ISSN: 1546-427X. URL: <http://dx.doi.org/10.1002/cav.252>.
- [14] Jack M Wang, David J Fleet, and Aaron Hertzmann. “Gaussian process dynamical models for human motion”. In: *IEEE transactions on pattern analysis and machine intelligence* 30.2 (2008), pp. 283–298.
- [15] Tatiana Benaglia et al. “mixtools: An R Package for Analyzing Finite Mixture Models”. In: *Journal of Statistical Software* 32.6 (2009), pp. 1–29. URL: <http://www.jstatsoft.org/v32/i06/>.
- [16] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*. Second Edition. Springer-Verlag New York, 2009.
- [17] Graham W. Taylor and Geoffrey E. Hinton. “Factored Conditional Restricted Boltzmann Machines for Modeling Motion Style”. In: *Proceedings of the 26th Annual International Conference on Machine Learning*. ICML '09. Montreal, Quebec, Canada: ACM, 2009, pp. 1025–1032. ISBN: 978-1-60558-516-1. URL: <http://dx.doi.org/10.1145/1553374.1553505>.
- [18] Sophie Jörg, Jessica Hodgins, and Carol O’Sullivan. “The Perception of Finger Motions”. In: *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization*. APGV '10. Los Angeles, California: ACM, 2010, pp. 129–133. ISBN: 978-1-4503-0248-7.
- [19] Graham W. Taylor et al. “Dynamical binary latent variable models for 3D human pose tracking.” In: *CVPR*. IEEE Computer Society, 2010, pp. 631–638. ISBN: 978-1-4244-6984-0. URL: <http://dblp.uni-trier.de/db/conf/cvpr/cvpr2010.html#TaylorSFH10>.
- [20] Chung-Cheng Chiu and Stacy C. Marsella. “A style controller for generating virtual human behaviors”. In: *The Tenth International Conference on Autonomous Agents and Multiagent Systems*. Taipei, Taiwan, May 2011.

- [21] Chung-Cheng Chiu and Stacy C. Marsella. “How to train your avatar: A data driven approach to gesture generation”. In: *The 11th International Conference on Intelligent Virtual Agents (IVA 2011)*. Reykavik, Iceland, Sept. 2011.
- [22] Henning Hamer et al. “Data-driven animation of hand-object interactions.” In: *FG*. IEEE, 2011, pp. 360–367. URL: <http://dblp.uni-trier.de/db/conf/fgr/fg2011.html#HamerGUG11>.
- [23] Manuel Febrero-Bande and Manuel Oviedo de la Fuente. “Statistical Computing in Functional Data Analysis: The R Package *fda.usc*”. In: *Journal of Statistical Software* 51.4 (2012), pp. 1–28. URL: <http://www.jstatsoft.org/v51/i04/>.
- [24] Jeremy Kun. *Dynamic Time Warping for Sequence Comparison*. <https://jeremykun.com/2012/07/25/dynamic-time-warping/>. Blog. 2012.
- [25] Feng Zhou and Fernando De la Torre Frade. “Generalized Time Warping for Multi-modal Alignment of Human Motion”. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2012.
- [26] Wei Dai, Yu Sun, and Xiaoning Qian. “Functional analysis of grasping motion.” In: IEEE, 2013, pp. 3507–3513.
- [27] Wenping Zhao et al. “Robust realtime physics-based motion control for human grasping.” In: *ACM Trans. Graph.* 32.6 (2013), 207:1–207:12. URL: <http://dblp.uni-trier.de/db/journals/tog/tog32.html#ZhaoZMC13>.
- [28] Jeng-Min Chiou, Yu-Ting Chen, and Ya-Fang Yang. “Multivariate functional principal component analysis: A normalization approach”. In: *Statistica Sinica* (2014), pp. 1571–1596.
- [29] Chung-Cheng Chiu and Stacy C. Marsella. “Gesture Generation with Low-Dimensional Embeddings”. In: *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems*. Paris, France, May 2014, pp. 781–788.
- [30] Honglang Wang. *Functional Principal Component Analysis*. [Online; accessed 18-August-2017]. 2014. URL: <https://honglangwang.wordpress.com/2014/12/20/functional-principal-component-analysis/>.
- [31] Han Du et al. “Scaled Functional Principal Component Analysis for Human Motion Synthesis”. In: *Proceedings of the 9th International Conference on Motion in Games*. MIG ’16. Burlingame, California: ACM, 2016, pp. 139–144. ISBN: 978-1-4503-4592-7.
- [32] T. Feix et al. “The GRASP Taxonomy of Human Grasp Types”. In: *Human-Machine Systems, IEEE Transactions on* 46.1 (2016), pp. 66–77.

- [33] Ming-Jin Liu et al. “Biomechanical characteristics of hand coordination in grasping activities of daily living”. In: *PloS one* 11.1 (2016), e0146193.
- [34] Wikipedia. *Functional principal component analysis* — *Wikipedia, The Free Encyclopedia*. [Online; accessed 18-August-2017]. 2017. URL: [https://en.wikipedia.org/w/index.php?title=Functional\\_principal\\_component\\_analysis&oldid=783422668](https://en.wikipedia.org/w/index.php?title=Functional_principal_component_analysis&oldid=783422668).
- [35] Igor Zubrycki and Grzegorz Granosik. “Fundamentals of Robotics with MyModelRobot”. In: (). URL: <http://www.mymodelrobot.appspot.com/>.



TRITA -MAT-E 2017:53  
ISRN -KTH/MAT/E--17/53--SE