

Eigengestures for natural human computer interface

Piotr Gawron Przemysław Głomb
Jarosław Adam Miszczak Zbigniew Puchała *

May 6, 2011

Abstract

We present the application of Principal Component Analysis for data acquired during the design of a natural gesture interface. We investigate the concept of an eigengesture for motion capture hand gesture data and present the visualisation of principal components obtained in the course of conducted experiments. We also show the influence of dimensionality reduction on reconstructed gesture data quality.

1 Introduction

Human-computer interface (HCI) which uses gestures promises to make certain forms of user interfaces more effective and subjectively enjoyable. One of important problems in creating such interface is the selection of gestures to recognize in the system. It has been noted [13] that choosing gestures that are perceived by users as natural is one of decisive factors in interface and interaction performance. At the same time, a large amount of research is focused on fixed movements geared towards efficiency of recognition, not interaction [13].

We view the analysis of natural gestures as a prerequisite of constructing an effective gestural HCI. As a tool for this task, it is natural to use **Principal Component Analysis** (PCA) [7]. PCA has been successfully applied for analysis and **feature extraction** *i.e.* of faces (the famous ‘eigenface’ approach [12]). For human motion, PCA has been

*Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, Bałtycka 5, 44-100 Gliwice, Poland, {gawron,przemg,miszczak,z.puchala}@iitis.pl

found to be a useful tool for dimensionality reduction (see *i.e.* [14]). Eigengestures appear in a number of publications, *i.e.* [10], where they are used as input for motion predictor. In [15] they are used for synthesis of additional training data for HMM. In [3] eigengesture projection is used for real-time classification. We argue, however, that the eigenvectors of human gestures—especially hand gestures—should be investigated beyond the effect they have in improving data processing (*i.e.* classification score); the structure of the decomposition may lead to important clues for data characteristics, as it has been the case for images [6]. To the best of authors’ knowledge, this is still a research field with limited number of contributions: in [16] eigen-decomposition of 2D gesture images is only pictured without discussion, whereas in [17] a basic analysis is done only for whole body gestures; main eigenvector are identified with deictic (pointing) gestures.

The main contribution of this work is application of PCA to analysis of the data representing human hand gestures obtained using motion capture glove. We show the influence of dimensionality reduction on reconstructed signal quality. We use the notion of eigengesture to the collected data in order to visualize main features of natural human gestures.

This article is organized as follows. Section 2 presents the experiment methodology; the sample set of gestures, acquisition methods, participants and procedure. Section 3 details the application of PCA to motion capture gesture data. Section 4 presents discussion the computed principal components. Section 5 presents visualization of eigengestures. Last section presents concluding remarks.

2 Method

For our experiment, we used base of 22 different type of gestures, each type represented by 20 instances — 4 people performing the gestures, each of them made the gesture 5 times (three with normal speed, then one fast following with one slow execution). The gestures are detailed in table 1. For discussion on gesture choice the reader is referred to [4].

The gestures were recorded with DG5VHand motion capture glove [1], containing 5 finger bend sensors (resistance type), and three-axis accelerometer producing three acceleration and two orientation readings. Sampling frequency was approximately 33 Hz.

The participants for the experiments were chosen from the employees of Institute of Theoretical and Applied Informatics of the Polish Academy of Sciences. Four males were instructed which gestures were going to be performed in the experiments and were given instructions

Table 1: The gesture prepared with the proposed methodology

	Name	Class ^a	Motion ^b	Comments
1	<i>A-OK</i>	symbolic	F	common ‘okay’ gesture
2	<i>Walking</i>	iconic	TF	fingers depict a walking person
3	<i>Cutting</i>	iconic	F	fingers portrait cutting a sheet of paper
4	<i>Shove away</i>	iconic	T	hand shoves away imaginary object
5	<i>Point at self</i>	deictic	RF	finger points at the user
6	<i>Thumbs up</i>	symbolic	RF	classic ‘thumbs up’ gesture
7	<i>Crazy</i>	symbolic	TRF	symbolizes ‘a crazy person’
8	<i>Knocking</i>	iconic	RF	finger in knocking motion
9	<i>Cutthroat</i>	symbolic	TR	common taunting gesture
10	<i>Money</i>	symbolic	F	popular ‘money’ sign
11	<i>Thumbs down</i>	symbolic	RF	classic ‘thumbs down’ gesture
12	<i>Doubting</i>	symbolic	F	popular (Polish?) flippant ‘I doubt’
13	<i>Continue</i>	iconic ^c	R	circular hand motion ‘continue’, ‘go on’
14	<i>Speaking</i>	iconic	F	hand portraits a speaking mouth
15	<i>Hello</i>	symbolic ^c	R	greeting gesture, waving hand motion
16	<i>Grasping</i>	manipulative	TF	grasping an object
17	<i>Scaling</i>	manipulative	F	finger movement depicts size change
18	<i>Rotating</i>	manipulative	R	hand rotation depicts object rotation
19	<i>Come here</i>	symbolic ^c	F	fingers waving; ‘come here’
20	<i>Telephone</i>	symbolic	TRF	popular ‘phone’ depiction
21	<i>Go away</i>	symbolic ^c	F	fingers waving; ‘go away’
22	<i>Relocate</i>	deictic	TF	‘put that there’

^a We use the terms ‘symbolic’, ‘deictic’, and ‘iconic’ based on McNeill & Levy [9] classification, supplemented with a category of ‘manipulative’ gestures (following [11])

^b Significant motion components: T-hand translation, R-hand rotation, F-individual finger movement

^c This gesture is usually accompanied with a specific object (deictic) reference

how the gestures should be performed. A training session was conducted before the experiment.

During the experiment, each participant was sitting at the table with the motion capture glove on his right hand. Before the start of the experiment, the hand of the participant was placed on the table in a fixed initial position. At the command given by the operator sitting in front of the participant, the participant performed the gestures. Each gesture was performed three times at the natural pace. Additionally, each gesture was made once at a rapid pace and once at a slow pace. Gestures number 2, 3, 7, 8, 10, 12, 13, 14, 15, 17, 18, 19 are periodical and in their case the single performance consisted of three periods. The operator decided about the end of data acquisition.

Input data consist of sequences of vectors $g_{t_n} \in \mathbb{R}^{10}$, $n \in \{1, N_i\}$ which are state vectors of the measurement device registered in subsequent moments t_n of time. The time difference $t_{n+1} - t_n$ is almost constant and approximately 30 ms. Each registered gesture forms a matrix $G_i \in M_{N_i, 10}(\mathbb{R})$. Acquisition time for every gesture is different, therefore the number of samples N_i depends on the sample. We acquired $K = 22$ different gestures, which are repeated $L = 20$ times.

3 Data processing

Our chosen statistical tool was **Principal Component Analysis** (PCA). It has been successfully applied in the domain of signal processing to various datasets such as: human faces [12], mesh animation [2].

3.1 Principal Component Analysis

For the sake of consistency we start by recalling basic facts concerning Singular Value Decomposition (SVD) [5] and **Principal Component Analysis** (PCA) [7].

Let $A \in \mathbb{M}_{m,n}$ has rank $k \leq m$. Then there exist orthogonal matrices $U \in \mathbb{M}_m$ and $V \in \mathbb{M}_n$ such that $A = U\Sigma V^T$. The matrix $\Sigma = \{\sigma_{ij}\} \in \mathbb{M}_{m,n}$ is such that $\sigma_{ij} = 0$, for $i \neq j$, and $\sigma_{11} \geq \sigma_{22} \geq \dots \geq \sigma_{kk} > \sigma_{k+1,k+1} = \dots = \sigma_{qq} = 0$, with $q = \min(m, n)$.

The numbers $\sigma_{ii} \equiv \sigma_i$ are called *singular values*, *i.e.* non-negative square roots of the eigenvalues of AA^T . The columns of U are eigenvectors of AA^T and the columns of V are eigenvectors of $A^T A$.

Principal Component Analysis allows us to convert a set of observations of correlated variables into the so-called *principal components*, *i.e.* a set of values of uncorrelated variables.

Formally, the i -th principal component is the i -th column vector of the matrix $V_{:,i} \times \sigma_{ii}$ obtained as a SVD of the data matrix. In order

to perform PCA on the data acquired in different units, the data need to be unified to a common units. In our case, the initial vector of data is transformed by the *studentisation*, *i.e.* by subtracting the empirical mean and dividing by the empirical standard deviation.

3.2 Organisation of data

As the input of the algorithm we have $K \times L$ matrices G_i . Each matrix represents a single realisation of a gesture. In order to perform PCA, the data are transformed in the following way:

1. Re-sampling: for every signal indexed by $s \in \{1, \dots, S = 10\}$: $G_{t_n, s} \rightarrow G'_{t'_n, s}$, where t_n indexes time samples of the gesture as acquired from the capture device, $t'_n \in \{1, \dots, N = 20\}$ using linear interpolation.
2. Arranging into the tensor:

$$T_{k, l, t'_n, s} = (G'_{t'_n, s})_{k, l},$$

where $k \in \{1, \dots, K = 22\}$ denotes number of the gesture type and $l \in \{1, \dots, L = 10\}$ denotes individual realisation of a gesture.

3. Double integration of signal from accelerometers to transform acceleration into position variable.
4. Centring and normalisation: for every signal s : $T'_{k, l, :, s} = (T_{k, l, :, s} - \overline{T_{:, :, :, s}}) / \sigma(T_{:, :, :, s})$.

The data are arranged into a matrix $X_{(k, l), (t_n, s)} = T'_{k, l, t_n, s}$ whose columns consists of vectorised distinct realisations of gestures. Such a matrix is then feed into SVD algorithm.

A sample of our data is visualised in Fig. 1 a) which presents the re-sampled, centred and rescaled second realisation of the *Cutting* gesture described in our data tensor by sub-matrix $T'_{3, 2, :, :}$.

[width=]pics/origvsgdenoised

Figure 1: A sample of the gestures dataset. The data are normalised and centred. Single realisation of *Cutting* gesture. Upper plot bending of fingers: T — thumb, I — index, M — middle, R — ring, L — little; lower plot: dashed line — palm roll, dotted line — palm pitch, $X Y Z$ — palm position in space. a) original data, b) approximation reconstructed using only 20 first principal components.

4 Application of PCA to data exploration

One of the typical applications of PCA to the analysis of the data obtained from the experiment is to reduce their dimensionality. Fig. 2 shows mean quality of the approximation of the original dataset in function of the number of principal components used to reconstruct the dataset. The distance in the Figure is scaled so that the approximation using only the first principal component gives 1. It can be easily seen that the dataset can be efficiently approximated using low rank approximation.

[width=0.45]pics/approximation-quality

Figure 2: Relative Euclidean distance between the dataset and its approximation obtained using first n principal components.

A comparison of original data sample *vs* its low rank approximation is shown In Fig. 1, sub-plot a) shows original data for *Cutting* gesture and sub-plot b) shows the same data reconstructed using only first 20 principal components.

5 Visualization of eigengestures

The coordinates in which the eigengestures are obtained are artificial. To create a visualisation one needs to change the coordinates to suited for the *hand presentation model*.

The change of coordinates is obtained by affine transformation acting independently on each dimension (sensor data). The parameters of these transformations (scales and translations) are obtained in the following way:

- The scale factor for each sensor (dimension) is a quotient of 0.05 – 0.95 quantile dispersion of this sensor data and the dispersion of the sensor in the eigengesture.
- The translation is calculated in such way that each visualised eigengesture has unified starting position.

In Fig. 3 the first two eigengestures (principal components) are shown. The first eigengesture looks very natural and resembles gestures commonly used by humans in process of **communication**. We found that higher eigengestures do not look very natural especially because of the negative values of the finger bends. Due to orthogonality of left singular vectors obtained from SVD it can not be expected

[width=0.75]pics/eigengestures

Figure 3: Visualization of two first eigengestures (principal components). On top: normalized and centred plots of signals in time. Upper plot bending of fingers: T — thumb, I — index, M — middle, R — ring, L — little; lower plot: dashed line — palm roll, dotted line — palm pitch, $X Y Z$ — palm position in space. At the bottom: shapes of hands in selected time moments. View is from the perspective of a person performing the gesture. For the sake of the clarity of the picture space position of the palm is omitted.

that eigengestures will be similar to natural gestures performed by humans.

We observed that time plots of eigengestures around number 100 and higher are very noisy.

6 Conclusions and future work

In this work our goal was to explore the space of human gestures using **Principal Component Analysis**. Visualisation of eigengestures is a tool that allows us to understand better the dataset we acquired during the experiment. We have shown that natural human gestures acquired with use of motion capture device can be efficiently approximated using 50 to 100 coefficients. Additionally we have identified the principal component of the gestures dataset.

Future work will consist of application of the obtained results to analysis of quality of gesture recognition. We will compare PCA with Higher Order Singular Value Decomposition [8] as data dimensionality reduction techniques.

Acknowledgements

This work has been partially supported by the Polish Ministry of Science and Higher Education projects NN516405137 and NN519442339. We would like to thank Paweł Kowalski and Sebastian Opozda for providing and adapting the model of the hand. We would like to thank Dr Ryszard Winiarczyk for encouraging us to publish this work.

The final publication is available at www.springerlink.com.

References

- [1] DG5 VHand 2.0 OEM Technical Datasheet. Technical report, DGTech Engineering Solutions, November 2007. Release 1.1.
- [2] M. Alexa and W. Müller. Representing animations by principal components. *Comput. Graph. Forum*, 19(3):411–418, 2000.
- [3] H. Birk, T.B. Moeslund, and C.B. Madsen. Real-time recognition of hand alphabet gestures using principal component analysis. In *10th Scandinavian Conference on Image Analysis*, 1997.
- [4] P. Głomb, M. Romaszewski, S. Opozda, and A. Sochan. Choosing and modeling gesture database for natural user interface. In *GW 2011: The 9th International Gesture Workshop “Gesture in Embodied Communication and Human-Computer Interaction”*, 2011. Accepted for publication.
- [5] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, 3rd edition, 1996.
- [6] A. Hyvärinen, J. Hurri, and P.O. Hoyer. *Natural Image Statistics: A probabilistic approach to early computational vision*. Springer-Verlag New York Inc, 2009.
- [7] I. T. Jolliffe. *Principal Component Analysis*. Springer Series in Statistics. Springer-Verlag, 2nd edition, 2002.
- [8] T.G. Kolda and B.W. Bader. Tensor decompositions and applications. *SIAM review*, 51(3):455–500, 2009.
- [9] D. McNeill. *Hand and Mind: What Gestures Reveal about Thought*. The University of Chicago Press, 1992.
- [10] M. Nakajima, S. Uchida, A. Mori, R. Kurazume, R. Taniguchi, T. Hasegawa, and H. Sakoe. Motion prediction based on eigen-gestures. Technical Report PRMU2006 130-160, Institute of Electronics, Information and Communication Engineers, 2006.
- [11] F. Quek, D. McNeill, R. Bryll, S. Duncan, X.F. Ma, C. Kirbas, K.E. McCullough, and R. Ansari. Multimodal human discourse: gesture and speech. *ACM Trans. Comput.-Hum. Interact.*, 9:171–193, 2002.
- [12] M. Turk and A. Pentland. Eigenfaces for recognition. *J. Cognitive Neurosci.*, 3(1):71–86, 2011.
- [13] A. Wexelblat. Research challenges in gesture: Open issues and unsolved problems. In Ipke Wachsmuth and Martin Fröhlich, editors, *Gesture and Sign Language in Human-Computer Interaction*, volume 1371 of *LNCS*, pages 1–11. Springer Berlin / Heidelberg, 1998.

- [14] K. Witte, H. Schobesberger, and C. Peham. Motion pattern analysis of gait in horseback riding by means of principal component analysis. *Hum. Mov. Sci.*, 28(3):394 – 405, 2009.
- [15] H.D. Yang, A.Y. Park, and S.W. Lee. Gesture spotting and recognition for human–robot interaction. *Robotics, IEEE Transactions on*, 23(2):256–270, 2007.
- [16] M. Yao, X. Qu, Q. Gu, T. Ruan, and Z. Lou. Online PCA with adaptive subspace method for real-time hand gesture learning and recognition. *Wseas Transactions On Computers*, 9(6), 2010.
- [17] J.R. Zhang, K. Guo, C. Herwana, and J.R. Kender. Annotation and taxonomy of gestures in lecture videos. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, pages 1 –8, 2010.