SVD-based Feature Extraction from Time-series Motion Data and Its Application to Gesture Recognition

Isao Hayashi
Kansai University
Ryozenji-cho, Takatsuki
Osaka 569-1095, Japan
ihaya@cbii.kutc.kansai-u.ac.jp

Yinlai Jiang
The University of Electro-Communications
Chofugaoka, Chofu,
Tokyo 1828585, Japan
jiang@hi.mce.uec.ac.jp

Shuoyu Wang
Kochi University of Technology
Tosayamada, Kami
Kochi 7828502, Japan
wang.shuoyu@kochi-tech.ac.jp

ABSTRACT
Singular value decomposition is used to extract features from time-series motion data. A matrix consisting of the time-series data is decomposed into left singular vectors which represent the patterns of the motion and singular values as a scalar, by which each corresponding left singular vector affects the matrix. Gesture recognition using the extracted features suggest the effectiveness of the method.

Categories and Subject Descriptors
I.5.2 [**Pattern Recognition**]: Design Methodology—*Feature evaluation and selection*; G.1.3 [**Numerical Analysis**]: Numerical Linear Algebra—*Singular value decomposition*

General Terms
Algorithms

Keywords
Feature extraction, SVD, Gesture recognition.

1. INTRODUCTION
The information of motion, such as trajectory, speed, and acceleration, has been measured to analyze and evaluate motion. Time-series data analysis is usually necessary to extract features from the measured data. In this paper, a method to extract embodied knowledge of body motion from time-series data by using singular value decomposition (SVD) [1] is introduced and testified. In this method, the left and right singular vectors and the singular values are decomposed from a Hankel matrix defined from the data measured with sensors [2, 3]. Since the left singular vector represents the characteristics of the matrix and the singular value represents the strength of the corresponding left singular vector, the left singular vector with a bigger singular value is considered as a representative pattern of the time-series data.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
BICT 2014, December 01-03, Boston, United States
Copyright © 2015 ICST 978-1-63190-053-2
DOI 10.4108/icst.bict.2014.257811

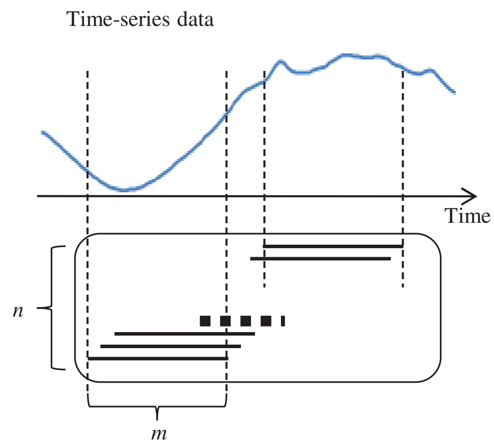

Figure 1: Design of Matrix $M_X^{i,G}$

2. FEATURE EXTRACTION USING SVD
Suppose that w points $(P_1, P_2, \cdots, P_w)$ of the body are measured while a person is performing a motion since the motion is usually measured by multiple sensors. On point P_i, the measured data series of motion G is denoted as $\tau^{i,G}$. The data series of $\tau^{i,G}$ consists of three-dimensional data $(X^{i,G}, Y^{i,G}, Z^{i,G})$. From the time-series data $\tau^{i,G} = (X^{i,G}, Y^{i,G}, Z^{i,G})$, n vectors by m data sampling are extracted by overlapping and the matrices $M_X^{i,G}, M_Y^{i,G}$, and $M_Z^{i,G}$ are constructed as a collective of the measurement data on the X, Y, and Z coordinates of the motion, respectively. Fig. 1 shows a design for constructing the matrix $M_X^{i,G}$. The matrices $M_X^{i,G}$ are described as $M_X^{i,G} = (X_1^{i,G}, X_2^{i,G}, ..., X_n^{i,G})^T$, where $X_p^{i,G} = (x_{p,1}^{i,G}, x_{p,2}^{i,G}, \cdots, x_{p,m}^{i,G}),\ p = 1, 2, \cdots, n$, and x is a datum on the X coordinate. We define $Y_p^{i,G}$ and $Z_p^{i,G}$ in the same way. For any m-by-n matrix M, a factorization of M is M=$U\Sigma V$, where U=(u_1, u_2, ..., u_m) contains the left singular vectors of M, V=(v_1, v_2, ..., v_n) contains the right singular vectors of M, and the matrix Σ is an m-by-n diagonal matrix with nonnegative real singular values on the diagonal.

Take an m-by-n matrix $M_k^{i,G},\ k = \{X, Y, Z\}$ as the general format of $M_X^{i,G}, M_Y^{i,G}, M_Z^{i,G}$, which is composed by overlapping the subsets of measurement data. The SVD of the matrix $M_k^{i,G}$ is $M_k^{i,G} = U_k^{i,G}\Sigma_k^{i,G}\{V_k^{i,G}\}^T$, where the matrix $\Sigma_k^{i,G}$ is an m-by-n

diagonal matrix whose diagonal entries are the singular values of $M_k^{i,G}$. The matrix $U_k^{i,G}$ contains the left singular vectors of $M_k^{i,G}$ and the matrix $V_k^{i,G}$ contains the right singular vectors of $M_k^{i,G}$. Intuitively, the left singular vectors in $U_X^{i,G}$ form a set of patterns of $M_X^{i,G}$ and the diagonal values in matrix $\Sigma_X^{i,G}$ are the singular values, which can be considered as scalars, by which each corresponding left singular vectors affect the matrix $M_X^{i,G}$.That is, the greater the singular value is, the more dominant the corresponding pattern is.

3. GESTURE RECOGNITION

We conducted a gesture recognition experiment to demonstrate the effectiveness of feature extraction using the left singular vectors. Five kinds of hand gestures, CH (Come here), GA (Go away), GR (Go right), GL (Go left), and CD (Calm down) which are commonly used in daily life, were performed by two subjects, SW and ST. The gestures were performed in a $50cm \times 50cm \times 50cm$ cubic space, whose zero point and coordinate system are shown in Fig. 2. The motions of the hand gestures were measured with Movetr/3D and GE60/W (Library, Tokyo, Japan).Five markers were measured: P_1 on the tip of the thumb, P_2 on the tip of the middle finger, P_3 on the tip of the little finger, P_4 on the thumb side of the wrist and P_5 on the little finger side of the wrist. One gesture was executed

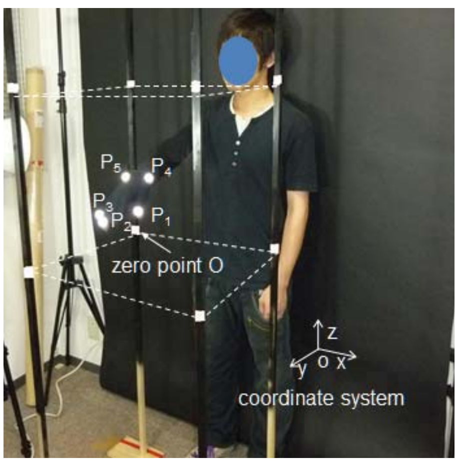

Figure 2: Environment for Measurement of Gestures

9 times by each subject. Data of the first five executions were used as the acquisition of patterns of the gesture. Data of the last four times were used to distinguish the gesture.

The observed data series are divided into training data series and checking data series.We developed a gesture recognition method using the left singular vectors extracted from the time-series data of gesture motion based on the similarity between gestures. The similarity is defined by the total absolute differential of the left singular vectors between the training data and the checking data at the same order.The estimation is defined by counting the number of minimal similarity values on the markers and the most counted gesture is output as the recognition result.

Table 1 shows the large counting number of minimized similarity values at each marker. The results are shown in the form of a/b, where a is the number of counted times of the most counted gesture and b is the name of the gesture. Over 80% accuracy was obtained at all the markers. Especially, results of the first marker P_1 got the

Table 1: Recognition Accuracy

Sub./Ges.	Marker				
TW	P_1	P_2	P_3	P_4	P_5
CH	4/GA	4/GA	4/CD	4/GA	4/GA
GA	8/GA	9/GA	9/GA	8/GA	8/GA
GR	7/GR	8/GR	8/GR	6/GR	5/GR
GL	5/GL	4/GL	4/GL	5/GL	5/GL
CD	6/CD	5/GA	5/GA	7/CD	8/CD
ST	P_1	P_2	P_3	P_4	P_5
CH	8/CH	9/CH	11/CH	9/CH	7/CH
GA	7/GA	9/GA	9/GA	9/GA	5/CD
GR	9/GR	10/GR	11/GR	6/GR	9/GR
GL	7/GL	8/GL	10/GL	5/GL	8/GL
CD	4/CD	5/GL	5/GL	5/CD	6/CD
Accuracy (%)	93.85	80.28	81.58	93.75	86.15

highest accuracy of 93.85%. Since P_1 measures the time series at the tip of the thumb and is largely related to movement of the thumb, it is suggested that the motion of the thumb is important in gesture recognition. The recognition results suggest that the left singular vectors extracted from the time-series data can be used as features to distinguish gestures.

In order to further study the recognition results, a questionnaire was carried out.8 respondents watched the gesture videos and evaluated the similarity between each two gestures. The score was given between 1 and 10 points, with 1 being not similar and 10 being very similar. Table 2 shows the average evaluation results. Apparently, the gestures are evaluated most similar to themselves. However, different gestures such as Come Here and Calm Down were evaluated quite similar. These results are consistent with the recognition results to a certain degree.

Table 2: Results of the questionnaire

	CH	GA	GR	GL	CD
CH	7.9	6.5	1.8	1.7	7.3
GA	6.7	9.1	1.6	1.6	6.4
GR	1.8	1.5	8.5	5.5	1.8
GL	1.6	1.5	5.7	8.0	1.7
CD	7.0	5.7	1.8	1.8	7.8

4. ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Numbers 25750255, 24300203, and the Canon Foundation.

5. REFERENCES

[1] Skillicorn, D. B. 2007. *Understanding Complex Datasets: Data Mining with Matrix Decompositions*, Florida, USA: Chapman and Hall/CRC, .

[2] Hayashi, I. Jiang, Y. and Wang, S. Y. 2011 “Acquisition of Embodied Knowledge on Gesture Motion by Singular Spectrum Analysis," *J. Adv. Comput. Intelli. Intelli. Inform.*, 15, 8(Nov. 2011), 1011-1018.

[3] Jiang, Y. Hayashi, I. and Wang, S. Y. 2014 “Knowledge Acquisition Method based on Singular Value Decomposition for Human Motion Analysis," *IEEE Trans. Knowl. Data Eng.*, 26 (2014). DOI=http://dx.doi.org/10.1109/TKDE.2014.2316521.(to appear)