A Data-driven Functional PCA Filter for Compensating the Effect of Sensor Position Changes in Motion Data

Roya Haratian
School of Electronic Engineering and Computer Science
Queen Mary University of London
London, UK

Chris Phillips
School of Electronic Engineering and Computer Science
Queen Mary University of London
London, UK

Tijana Timotijevic
School of Electronic Engineering and Computer Science
Queen Mary University of London
London, UK

ABSTRACT
Human body motion is typically captured by body area sensor networks. Accurate sensor placement with respect to anatomical landmarks is one of the main concerns in reliability of the motion capturing systems. Changes in position of sensors cause increased variability in motion data. Our goal is to isolate the characteristic features that represent the principle motion pattern. By using functional Principal Component Analysis (f-PCA) we compensate for the variation in data due to inadvertent movement of sensor placement. F-PCA is an effective tool for the study of human motion modeling by identifying hidden combinations and relationships between variables. The collected data from our experiment show differences between similar actions within different sessions of marker wearing. After applying f-PCA to the data, we show how the uncertainties due to sensor position changes can be compensated for.

Keywords
Motion capturing, measurement variability, functional Principal Component Analysis.

1. INTRODUCTION
On-body sensors can be used for capturing human movement. The motion of a body can be thought of as a collection of time series streams describing the joint angles which is called motion data. Motion data can be used in applications such as animation, sports biomechanics, rehabilitation, and so on. In many applications, the human body is approximated by a collection of articulated limbs that form a kinematic tree. Determining the anthropometry of the individual subject is called model calibration. Accurate sensor placement with respect to anatomical landmarks and location determination of joint centers with regard to these sensors are two important aspects of model calibration in motion capturing. There should be especial care to achieve sufficient accuracy to justify the results of motion capturing systems despite measurement variability.

Variability in movement patterns plays a fundamental role in motion analysis and its influence on the analysis of motion data should be taken into account. Variability and measurement errors in motion capturing via on-body sensors/markers can come from three primary sources: the technicians responsible for placing the sensors/markers, the measurement system, and the subject under evaluation. Variability is defined by the sum of variances from each independent source [1]. Sensor placement among technicians is the largest source of variability [2]. Particular care should be taken to ensure that sweating, rapid movements and the placement of markers during different trials and sessions on the subject's body, do not affect sensor/marker position according to the marker placement guidelines.

The reliability to measurements is directly affected by the sensor placement during different trials and sessions. If experimental errors conceal important motion deviations, meaningful information will be lost. On the other hand, if the limitations of the motion capturing methods are not understood, small deviations may be considered meaningful, thereby leading to over interpolation [3]. Every time that a subject tries to carry out the same movement, a certain amount of variation may be registered between the different sessions of marker wearing. Variability between sessions was found to be much higher than within-session variability due to the high potential for differences in the marker placement [1].

In clinical biomechanics, variability in motion data can be associated with arbitrary changes in position of sensors. Despite this, discovering the characteristics features that represent the main pattern(s) of motion is our concern. Standard data analysis techniques, which determine mean and standard deviation of time series data, summarize motion data in single patterns providing average behavior and show deviations as possible errors by standard deviation bands. Such techniques severely reduce features in the data, and may result in important information being discarded [4]. In particular, the results do not account for the information that may be inherent in all the variability apparent in the data. When different sessions of marker wearing are averaged, information can be lost and the average curve does not closely resemble any of the individual curves.

Multivariate statistical analysis has proved to be a powerful tool to eliminate collinearity and to facilitate analysis, presenting only the essential structures hidden in the data [5], but again the extent of data loss is a matter of concern. Among multivariate statistical techniques, linear transformations are computationally easier to perform and within linear transformations, the use of functional techniques may provide additional insight into differences in motion patterns. Functional principal component analysis (f-PCA) is an effective tool for the study of human movement in modeling motion curves by identifying hidden combinations and relationships between variables [6]. The basic philosophy of functional data analysis is the belief that the best unit of information is the entire observed function (a curve within a family of curves) rather than a string of numbers. It is assumed that data are supposed to have an underlying functional relationship governing them. It also allows extracting loadings and scores. Loadings are the correlation coefficients between the variables and the components. Scores are

BODYNETS 2012, September 24-26, Oslo, Norway

DOI 10.4108/icst.bodynets.2012.250004

the contributions of the principal components to each individual variable [4]. F-PCA is an extension of the traditional PCA, where the principal components are represented by functions rather than vectors.

PCA is a multivariate statistical technique. It can be used as a means of decorrelation by computing a new, much smaller set of uncorrelated variables, Principal Components – (PCs). Each new variable is a linear combination of the original ones. All the principal components are orthogonal to each other, so there is no redundant information. All remaining principal components are defined similarly, so that the lowest order components normally account for very little variance and can usually be ignored [7].

To reflect the true nature of motion data variability, we investigate the use of f-PCA for filtering and interpreting motion data, while accounting for their original variability. We compensate for uncertainties in the data, due to sensor position changes. Our approach is introduced in Section 2 by describing the experimental procedure, the data analysis and giving a brief introduction to f-PCA. Results and discussion are presented in section 3. Our paper is then concluded in Section 4.

2. METHOD

To discover the most significant feature(s) that represent the main motion pattern in spite of variability in motion data due to arbitrary changes in position of sensors, we evaluate similarities between repetitive actions across different sessions of marker wearing. To describe the compensation procedure, we first describe the process of motion data acquisition and then the data processing techniques used in our study. The experiment comprises subject preparation, the test procedure, and motion data capture. Then our data processing techniques are discussed which consists of time warping, standardization, and f-PCA.

2.1 Experiment

A commercial optical motion capture system, named Codamotion, has been used in our experiment. The sensors that are tracked by the Coda scanner units are small infra-red light emitting diodes. This system consists of twelve cameras operating at 200 Hz. Each Coda scanner unit contains three special cameras which detect infra-red pulses of light emitted by the Coda markers and locate the marker positions with very high resolution and linearity. The system should be calibrated before each experimental session according to the manufacturer guidelines. The calibrated system measures the positions of markers within a three dimensional co-ordinate system which is fixed in relation to the scanner unit [8].

There are several standard marker-sets for placing markers on the human body such as Cleveland Clinic, Saflo, Helen Hays, Codamotion, and so forth. For Bilateral Gait, the recommended Codamotion marker-set comprises a total of 22 standard markers was used in our experiment. The experimental procedure complies with university guidelines, approved by the local institutional review board. A full analysis of an individuals' motor behavior involves the evaluation of an appropriate number of individual repetitions. We asked a subject to participate in 10 marker-wearing sessions. There were inadvertent changes in position of markers for each session while following the standard marker set.

The process of measurements in each session involves instrumenting the legs and pelvis with active markers according to the Codamotion marker-set to perform the motion capture. Each session consists of 10 trials. The subject in each trial walks from the start point to an end point while the motion capture system captures the subject's motion. Each trial lasts for 6 seconds and the sampling frequency of the motion capturing system is 200 Hz. The subject was asked to walk at normal walking speed. This walking speed is maintained as far as possible whilst different marker placements are made across different sessions of marker wearing.

2.2 Data Analysis

Walking sequences are segmented into cycles. Each cycle includes two steps. We ask a subject to walk in a certain time interval and divide the walking actions into cycles. Each cycle is identified as the interval from toe-off to the following toe-off of the same foot or consecutive left/right heel contact depending on our chosen criteria. By using the vertical velocity changes of heel markers, consecutive right heel contacts determine the period of one stride. We use consecutive right heel contact in our experiment to separate each stride. Segmenting data into gait cycles almost always results in gait data cycles of differing lengths due to differences in motion speed.

In action recognition, identifying features during action sequences with different speeds or different numbers of samples in each cycle is an important issue. In such cases time normalization is necessary before or during the recognition process. Each cycle should be normalized so it is represented by the same number of samples. Linear time normalization and nonlinear time normalization using dynamic time warping are the most common technique that can be used for this purpose. Linear time normalization linearly converts the trajectory's time axis from the experimentally-recorded time units to an axis representing the gait cycle [9]. Dynamic time warping shifts the time index of each data point in a test trajectory to minimize the distance between the test and consensus trajectories. In general, time warping can be performed implicitly, i.e., by the resizing along the time axis of patterns that depict the evolution of a feature through time [10].

2.3 Functional Principal Component Analysis

To find the dominant modes of variation in the data, and tease apart deterministic and stochastic components of movement patterns, f-PCA would be a useful tool usually after subtracting the mean from each observation. It allows for separation of main and residual components within a data set. Viewing consistent features as coherent components imply the mechanisms generating these common structures follow deterministic rules otherwise they would not be consistent/coherent. In contrast, the residual components often contain a degree of randomness or stochasticity. f-PCA is an extension of the traditional multivariate PCA, where the principal components are represented by functions rather than vectors. Going from the multivariate PCA to the functional version, will result in eigenfunctions instead of eigenvectors and summations change into integrations. The upper limit number of principal components in the multivariate case is the number of variables, while in f-PCA the number of eigenfunctions is equal to minimum of K, which is the number of basis functions, and N which is the number of variables [11].

In first step, we fit a function to the data. To fit a function to our data as it is shown in (1), we use a set of functional buildings blocks $\emptyset_k$, $k = 1, 2, \ldots, K$, called basis functions which are combined linearly. That is, a function $x(t)$ defined in this way is expressed as follow, and called a basis function expansion.

$$x(t) = \sum_{k=1}^{K} c_k \emptyset_k \tag{1}$$

Parameters c_k , are the coefficients of the expansion. The matrix expression of *N* functions will be of the form $\mathbf{X}(t) = \mathbf{C}\,\emptyset(t)$, where $\mathbf{X}(t)$ is a vector of length *N* containing the function $x_i(t)$, and the coefficient matrix $\mathbf{C}$ has *N* rows and *K* columns. The sample variance-covariance function, $v(s,t)$ is defined as follows,

$$v(s,t) = N^{-1}\sum_{i}^{N} x_i(s)\,x_i(t). \qquad (2)$$

The functional eigenequation is

$$\int v(s,t)\xi(t)dt = \rho\xi(s), \qquad (3)$$

where ρ is eigenvalue and $\xi(s)$ is an eigenfunction of the variance-covariance function. Eigenfunction which is called principal component weight function, $\xi_1(s)$, can be found by (4).

$$\text{Maximize } \sum_i f_{i1}^2$$
$$\text{Subject to } \int \xi_1^2\,(s)ds = \left\|\xi_1\ \right\|^2 = 1, \qquad (4)$$

where the principal component score, is defined as

$$f_{i1}=\int \xi_1(\mathrm{s})\, x_i(\mathrm{s})ds. \qquad (5)$$

A non-increasing sequence of eigenvalues $\rho_1 \geq \rho_2 \geq \cdots \geq \rho_k$ can be constructed stepwise by requiring each new eigenfunction computed in step *l*, to be orthogonal to those computed on previous steps,

$$\int \xi_j(t)\,\xi_l(t)dt = 0, \qquad j = 1, \ldots, l-1$$
$$\int \xi_l^2(t)dt = 1. \qquad (6)$$

To separate components of movement patterns, f-PCA can be used especially when partitioning signals into deterministic and stochastic components, by subtracting either the one or the other from the signal. This can be regarded as filtering the noise or the common parts, respectively. As the effect of random changes in position of sensors is expected to randomness effect on the motion the motion data, to remove this effect from main and coherent component of movement, we partition the data into two elements, $\overrightarrow{\mathrm{x_i}}^{(\mathrm{global})}$ and $\overrightarrow{\mathrm{x_i}}^{(\mathrm{filtered})}$ which are shown in (7) .

$$\sum_{n=1}^{L<N} \xi_n(\mathrm{t})f_i^{(n)} + \sum_{n=L+1}^{N} \xi_n(\mathrm{t})f_i^{(n)} \qquad (7)$$

Where we assume the number of basis functions are more than the number of variables. The sum of the dominant principal components weight functions is given by $\overrightarrow{\mathrm{x_i}}^{(\mathrm{global})}$, so the resulting filter characteristic depends on the data. The number of modes that define the global pattern influences the filtered pattern. We apply this data-driven filter to the motion data to separate out the effect of random changes in sensors position and main motion pattern by keeping the dominant modes of variation in the data whilst considering the proportion of corresponding eigenvalues to the total variance.

3. RESULTS AND DISCUSSION

Kinematics variables for 10 motion-capture sessions were acquired. These variables are the angles of the pelvis, hip, knee, ankle, and foot in the X, Y, and Z planes respectively. Kinematics variables of each marker-wearing session were averaged over 10 trials to eliminate the effect of factors that are irrelevant to differences in position of sensors in each session. The cause of these factors could be different walking speed, different ways of walking because of the tiredness of the subject and so forth.

The kinematics variables are shown in Fig. 1. The differences which are shown between the variables of different sessions are due to inadvertent changes in position of sensors. The placement changes are random in all X, Y, and Z planes and they were made in radius of 2 cm from the correct place which simulates the effect of configuration errors during different sessions while following the standard marker set. We aim to eliminate the effect of these changes on the kinematics variables to obtain valid motion data. For example, in a clinical context, if a measurement made on a patient cannot be relied upon because of random errors then that measurement will not be useful [2].

After fitting Fourier basis functions into data of each session, we would have 15 lots of functions with 10 functions in each one, whereby 15 stands for the kinematics variables and 10 represents the number of sessions for each sequence with different sensor placement. Each lot contains one of the 15 kinematics variables in a cycle. By applying f-PCA on these lots, we obtain functional principal component functions for each of these kinematic variables. Then by keeping the most dominant mode of variation, and eliminating the rest, we try to retain the most important variations in data and eliminate the effect of inadvertent changes in the position of the sensors. After deleting the non-dominant mode of variation, we return the functions into initial domain by using the inverse f-PCA transform method.

Although the motion capture data are for the same actions, they are not the same because of changes in the position of sensors as we see in Fig. 1. After applying the proposed technique, we investigate if the effect of position changes in different sessions can be eliminated and the motion data made more consistent. By comparing the data that are shown in the upper and lower part of Fig. 1, before and after applying the data-driven f-PCA filtering, respectively, we see how similar the captured motion data are from different marker wearing sessions. Results show that by using the introduced technique, we can significantly reduce the effect of inadvertent changes in the position of the sensors on the captured data and extract the common mode of variation in the several sessions of marker wearing motion data.

4. CONCLUSION

Variability of kinematics measurements due to inadvertent sensor/marker placement changes was discussed in this paper. It is shown that there is measurement variability due to the failure to place markers/sensors accurately in motion capturing of the human body even following same placement protocol for each session. The variability may conceal important motion deviations and meaningful information can be lost. A f-PCA filter followed by other techniques to compensate for the effects of sensors position changes in motion data was applied on the motion data of the designed experiment. Results show differences between similar repetitive actions with random marker positional changes and how these variations can be compensated for by applying the signal processing techniques. By keeping the most dominant mode of variation in data the common motion pattern can be extracted from the data across multiple marker-wearing sessions. By using the data driven filter, we can isolate the salient movement pattern regardless of the variations that emerge in changes in position of sensors across sessions while following marker set up protocols.

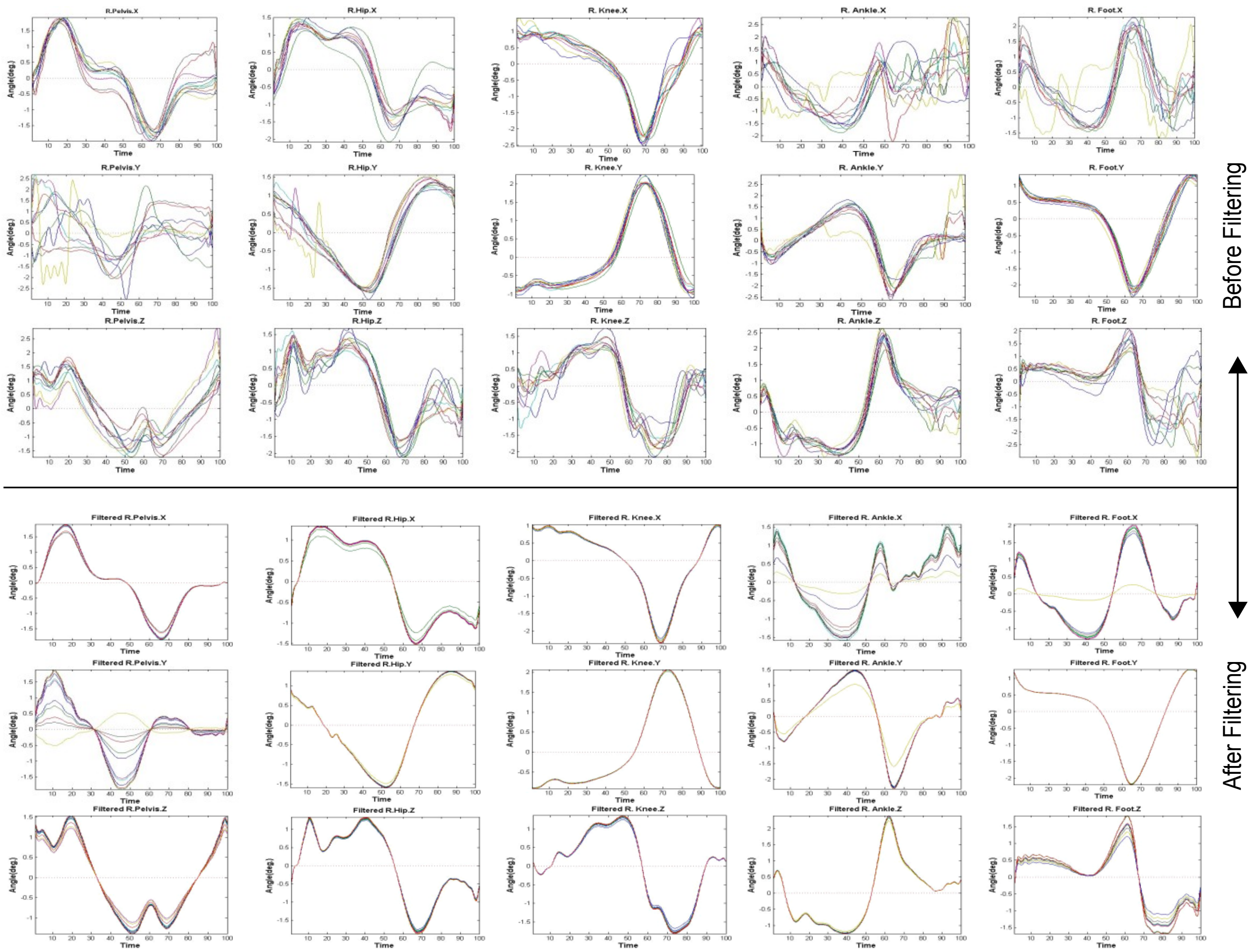

Figure 1: Kinematics variables of 10 marker wearing sessions, showing normalized angles of pelvis, hip, knee, ankle, and foot in X, Y, and Z planes before and after filtering

5. REFERENCES

[1] Gorton, G., Hebert, D., and Gannotti, M. 2009. Assessment of the kinematic variability among 12 motion analysis laboratories. *Journal of Gait and Posture*. 29 (2009), 398-402.

[2] Baker, R. 2006. Gait analysis methods in rehabilitation. *Journal of NeuroEngineering and Rehabilitation*. (2006).

[3] Schartz, M. , Trost, J., and Wervey, R. 2004. Measurment and management of error in quantitative gait data. *Journal of Gait and Posture*. 20 (2004), 196-203.

[4] Dona, G., Preatoni, E., Cobelli, C., Rodano, R., and Harrison, A. 2010. Application of functional principal component analysis in race walking: an emerging methodology. *Journal of Sport Biomechanics*. (2010), 284-301.

[5] Chau, T., Young, S., and Redekop, S. 2005. Managing variability in the summary and comparison of gait data. *Journal of Neuroengineering and Rehabilitation*. 2 (2005).

[6] Donoghue, O. A., Harrison, A. J., Coeffey,N., and Hayes, K. (2008). Functional data analysis of running kinematics in chronic achilles tendon injury. *Medicine and Science in Sports and Exercise*. 40 (7), 1323–1335.

[7] Daffertshofer, A., Lamoth, C., Meijer, O. and Beek, P. 2004. PCA in studying coordination and variability: a tutorial. *Journal of Biomechanics*. 19 (2004), 415-428.

[8] Codamotion user guid, Charnwood Dynamics Ltd, 2005.

[9] Boulgouris,N., Plataniotis, K., and Hazinakos, D. 2006. Gait recognition using linear time normalization. *Journal of Pattern Recognition*. 39 (2006), 969-979.

[10] Boulgouris,N., Plataniotis, K., and Hazinakos, D. 2006. Gait recognition using dynamic time warping. *Journal of Signal Processing*. (2004), 263-266.

[11] Ramsy, J.O., Hooker, G., and Grave, S. 2009. *Functional Data Analysis*, Springer.