

Arm Gesture Recognition and Humanoid Imitation Using Functional Principal Component Analysis

Jacopo Aleotti, Alessandro Cionini, Luca Fontanili, Stefano Caselli
RIMLab - Robotics and Intelligent Machines Laboratory
Dipartimento di Ingegneria dell'Informazione, University of Parma, Italy
E-mail {aleotti,cionini,fontanili,caselli}@ce.unipr.it

Abstract—A method is proposed for gesture recognition and humanoid imitation based on Functional Principal Component Analysis (FPCA). FPCA is a statistical technique of functional data analysis that has never been applied before for humanoid imitation. In functional data analysis data (e.g. gestures) are functions that can be considered as observations of a random variable on a functional space. FPCA is an extension of multivariate PCA that provides functional principal components which describe the modes of variation in the data. In the proposed approach FPCA is used for both unsupervised clustering of training data and gesture recognition.

In this work we focus on arm gesture recognition. Human hand paths in Cartesian space are reconstructed from inertial sensors. Recognized gestures are reproduced by a small humanoid robot. The FPCA algorithm has also been compared to a state of the art algorithm for gesture classification based on Dynamic Time Warping (DTW). Results indicate that, in this domain, the FPCA algorithm achieves a comparable recognition rate while it outperforms DTW in terms of efficiency in execution time.

I. INTRODUCTION

Most modern sensors for motion analysis produce high-resolution data at high sampling rate that are stored in high-dimensional input vectors. These data have an intrinsic functional nature since curve fitting can be applied to represent them as continuous functions. This work explores the use of Functional Principal Component Analysis (FPCA) for unsupervised clustering of arm gestures and for gesture recognition. The proposed method has also been applied to humanoid robot imitation. FPCA is a technique that belongs to Functional Data Analysis (FDA) [24], which has not been considered before in robot imitation. Functional Data Analysis is a statistical methodology that enables quantitative analysis on continuous multidimensional data. In FDA, data samples (e.g. gestures), obtained from repeated observations, are analyzed jointly and are characterized by real-valued functions instead of vectors. FPCA is used to extract orthogonal functional principal components (FPCs) that reduce the dimensionality of the input data by projection on the FPCs. FPCA extends traditional multivariate Principal Component Analysis (PCA) and it provides numerous advantages: FPCA does not suffer from the curse of dimensionality, presents a better discriminatory power and handles missing or irregularly sampled curves with noisy observations. While PCA is a very well known method to decrease the size of the input data for fast classifications, FPCA has not been investigated in previous works.

Figure 1 gives an overview of the proposed gesture recognition approach. Training data are provided by several users from motion capture in the form of multiple arm gestures of different classes. Each gesture is defined by the hand path as a discrete time sequence in Cartesian space that is computed from inertial sensors. Discrete time sequences of gesture data are converted to training functional data by basis function expansions using B-splines (curve fitting). Then, FPCA is performed on all the training data to determine a finite set of functional principal components (FPCs) that explain the modes of variation in the data. Training functional data are also projected onto the FPCs, resulting in a finite set of functional principal components scores that compactly describe each input gesture (dimensionality reduction). Unsupervised clustering of the training data is then performed, in the space of the functional principal components scores, to organize the training set into disjoint classes of gestures and to generate a prototype gesture (exemplar) within each class.

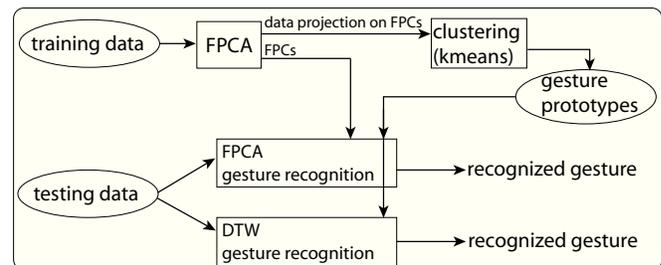


Fig. 1. Schematic diagram of the proposed gesture recognition method based on FPCA and comparison with DTW.

The computed FPCs and the gesture prototypes are used for recognition of new gestures. A gesture recognition algorithm has been developed by applying FPCA. Each new gesture to be recognized is first projected onto the FPCs and then it is categorized as belonging to the class of the closest prototype. Then, a small humanoid robot imitates the gesture by performing the trajectory of the recognized prototype. The proposed gesture recognition algorithm has also been compared with a high performance recognition algorithm based on Dynamic Time Warping (DTW). The developed DTW algorithm operates in the domain of the discrete time sequence data. Evaluation has pointed out that in the considered experimental protocol the two algorithms achieve a comparable recognition rate, but the proposed

FPCA method is more efficient in terms of execution time, which makes it especially suitable for gestures having long duration or high sampling-rate.

The paper is organized as follows. Section II reviews the state of the art regarding gesture recognition and imitation in humanoid robots. Section III describes the proposed method for gesture recognition based on FPCA, which is a general approach that does not depend on the experimental setup used for validation. Section IV illustrates the human arm motion capture system and it describes the solution of the correspondence problem for robot imitation. Section V presents the experimental evaluation of the approach. The paper closes in section VI summarizing the work.

II. RELATED WORK

This section presents a summary of previous works on humanoid imitation and gesture learning. As anticipated in the section I, to the best of our knowledge, no previous study has considered FPCA for gesture recognition and robot imitation. Many authors have investigated the use of Dynamic Time Warping for gesture recognition [33], [18], [2], [9], [1], [27], [10]. The standard template-based DTW algorithm for recognition, which will be recalled in section III, has been applied with several variants. DTW has proven user-independent and more robust than other statistical methods like Hidden Markov Models, that require a large number of training data and the selection of appropriate low level features [7], [18], [9].

In [14], [31] a mimesis learning approach was proposed using primitive symbols observation and Hidden Markov Models to generalize and synthesize motion patterns. The approach was extended in [17] to consider incremental learning of motion primitives. Inamura et al. [13] proposed a method for humanoid imitation of daily life behaviors based on attention points, specified by voice commands, such as motion constraints in manipulation tasks. In [11] a probabilistic approach was presented for recognition and reproduction of free-space movements for humanoid robots based on adaptation of motor primitives. Calinon et al. [5] developed a programming by demonstration model for stochastic recognition and humanoid arm gesture production. Shon et al. [28] proposed a nonlinear regression algorithm for mapping motion capture data to a humanoid robot using a latent variable space to reduce the high-dimensional observation space. In [12] parametric Hidden Markov Models have been investigated for recognition and generation of human movements that explicitly encode the goal of the actions such as reaching and pointing motions. Zinnen et al. [36] focused on the problem of recognizing gestures in continuous data streams using turning points to identify segments of interest in the human movements. Sigalas et al. [29] presented a method for hand gesture recognition based on neural networks classifiers and computer vision.

Other works have focused on the more specific problem of generalization and synthesis of robot gestures over multiple demonstrations [6], [19], [4], [32], [34]. In particular, in [6] a method was presented for extracting the goal of a task from

human demonstration and determining the best imitation strategy to satisfy the goal. Real-time strategies for humanoid robot imitation have been investigated in [23], [25], [15], [21], [16]. In particular, Menezes et al. [21] adopted a single camera system for motion capture and a particle filtering technique for imitation on a simulated humanoid robot. The related topic of motion retargeting, i.e. motion adaptation from one character to another with different kinematic structures, has been considered in computer vision and animation [20], [35], [26], [30], [8], [22].

III. METHOD

A. FPCA of human arm gestures

In this work a human arm gesture is characterized by the hand path in Cartesian space. The input training set G consists of N gestures g_i ($i=1, \dots, N$) that are recorded by different users. There are C classes of gestures c_h ($h=1, \dots, C$). Each gesture $g_i \in G$ belongs to one of the classes and it is defined by a triplet of sequences $g_i=(x_i(t_j), y_i(t_j), z_i(t_j))$ of Cartesian coordinates of the hand sampled at discrete times t_j with $j=1, \dots, m$. All discrete time sequences of each gesture are converted to functional data by curve fitting, e.g. the x -component $x_i(t)$ of gesture i is expressed in functional form as a linear combination of basis functions as follows

$$x_i(t) = \sum_{k=1}^K \lambda_{ik} \phi_k(t) \quad (1)$$

where $\phi_k(t)$ are the basis functions, K is the number of basis functions and λ_{ik} are the expansion coefficients. In this work a B-spline expansion has been adopted, which is suitable for analysis of human motion data. Hence, the input training set of gestures G can be expressed in its functional form as $G=\{(x_1(t), y_1(t), z_1(t)), \dots, (x_N(t), y_N(t), z_N(t))\}$. B-spline expansion is computed through a functional regression model, for each input data, that incorporates a roughness penalty [24]. The roughness penalty approach minimizes the ordinary sum of square error residuals plus a penalty term that imposes a smoothness constraint. The penalty term is the integrated squared second derivative of the estimated function, e.g. $\int [\ddot{x}(t)]^2 dt$ for the $x_i(t)$ component. Let $(\bar{x}_i(t), \bar{y}_i(t), \bar{z}_i(t))$ be the mean functions of gesture g_i . Functional data are centered by subtracting the mean functions. Let $(\nu_{xx}(s, t), \nu_{yy}(s, t), \nu_{zz}(s, t))$ be the covariance functions of the centered data, e.g. for the first component $\nu_{xx}(s, t)=(N-1)^{-1} \sum_{i=1}^N x_i(s)x_i(t)$. Let $(\nu_{xy}(s, t), \nu_{yz}(s, t), \nu_{xz}(s, t))$ be the cross-covariance functions, e.g. $\nu_{xy}(s, t)=(N-1)^{-1} \sum_{i=1}^N x_i(s)y_i(t)$.

The Functional Principal Component Analysis (FPCA) of the training motion data finds a finite set of R principal component functions that represent most of the variance of the input data. The functional principal components (FPCs) satisfy orthogonal conditions [24]. Each functional principal component ξ^r with $r=1, \dots, R$ is defined by a vector of weight functions $\xi^r=(\xi_x^r(t), \xi_y^r(t), \xi_z^r(t))$ solution of the

eigenequation system

$$V\xi^r = \rho^r \xi^r \quad (2)$$

where V is called the covariance operator and ρ^r is the eigenvalue. Equation 2 generalizes the standard eigenvalue equation of multivariate PCA and it is expanded as

$$\begin{aligned} \int \nu_{xx} \xi_x^r dt + \int \nu_{xy} \xi_y^r dt + \int \nu_{xz} \xi_z^r dt &= \rho^r \xi_x^r(s) \\ \int \nu_{xy} \xi_x^r dt + \int \nu_{yy} \xi_y^r dt + \int \nu_{yz} \xi_z^r dt &= \rho^r \xi_y^r(s) \\ \int \nu_{xz} \xi_x^r dt + \int \nu_{yz} \xi_y^r dt + \int \nu_{zz} \xi_z^r dt &= \rho^r \xi_z^r(s) \end{aligned} \quad (3)$$

In equation 3 the time dependence of the functions in the integrals has been omitted for simplicity. FPCs are also called eigenfunctions or harmonics. Each i -th trivariate gesture function $(x_i(t), y_i(t), z_i(t))$ can be projected onto all the R functional principal components to obtain a vector s_i of R functional principal components scores. The functional principal component score s_{ir} of the i -th trivariate gesture function projected onto the r -th functional principal component is given by

$$s_{ir} = \int x_i(t) \xi_x^r(t) dt + \int y_i(t) \xi_y^r(t) dt + \int z_i(t) \xi_z^r(t) dt \quad (4)$$

Thus, FPCA is used for reducing the dimensionality of functional data (gestures) by projecting them onto a finite number of functional principal components. The presented FPCA model can also be generalized and applied to n -dimensional input functions. FPCA enables unsupervised clustering of the vectors of functional principal components scores as explained in the following section.

B. Gesture clustering and recognition using FPCA

The proposed algorithm for classification of new gestures based on FPCA operates in the space (of size R) of the functional principal score vectors. Vectors s_i of training gestures are clustered in C classes of gestures. Unsupervised clustering is performed by the standard k-means algorithm. K-means partitions the training set into C classes and it extracts the cluster centers. Cluster centers are vectors of size R . The total number of gesture classes C is given as input.

After clustering, a prototype gesture (exemplar) is generated within each class. Prototype gesture p_{c_h} of class c_h ($h=1, \dots, C$) is generated as the cluster center using the l_2 -distance in the space of the functional principal score vectors. For each class a corresponding prototype gesture is also generated in the time domain, as the mean function of all the gestures of the class (which have been previously smoothed by the roughness penalty approach). The gesture prototype in the time domain is needed for both the DTW recognition algorithm and for gesture imitation. Of course there are other strategies that have been proposed in literature for prototype generation [6], [19], [4], [32], [34], which is not the core of this work. A pre-processing technique can also be applied for filtering out those gestures that are affected

by inconsistencies typical of human motion. This strategy has been explored in our previous work [3] and it is quite effective if the training set contains spurious trajectories.

The FPCs and the generated gesture prototypes are then used for recognition of new gestures. The proposed method for FPCA gesture recognition works as follows. Each new gesture g to be recognized is converted in its functional form and then it is projected onto the FPCs. Let s be the vector of scores of g . Gesture g is then categorized as belonging to the class c^* of the closest prototype using the l_2 -distance d between functional principal score vectors, i.e. $c^* = \operatorname{argmin}_{c_h} d(s, p_{c_h})$. The recognized arm gesture is then imitated by a small humanoid robot that reproduces the trajectory of the corresponding prototype.

C. Gesture recognition using dynamic time warping

The proposed FPCA gesture recognition method has been compared with a high-performance, state-of-the-art algorithm based on Dynamic Time Warping (DTW). DTW operates in the domain of the discrete time sequence data of Cartesian coordinates and it requires, as well as the FPCA method, the availability of prototype gestures. The standard DTW algorithm performs a nonlinear alignment between two time series data. The optimal alignment is found by minimizing a cumulative distance measure $DTW(\cdot)$ called cost function. Let $g=(x(t_j), y(t_j), z(t_j))$ be the new gesture sequence to be categorized (with $j=1, \dots, m$). Let also $p_{c_h}=(x_{p_{c_h}}(t_j), y_{p_{c_h}}(t_j), z_{p_{c_h}}(t_j))$ be the gesture prototypes (with $h=1, \dots, C$) of the classes characterized, as well, by their discrete time sequence data of length m . The gesture recognition algorithm based on DTW performs the alignment of gesture g with all the gesture prototypes and, then, gesture g is categorized as belonging to the class c^* of the closest prototype $c^* = \operatorname{argmin}_{c_h} DTW(g, p_{c_h})$. In particular, DTW finds the optimal alignment between g and a gesture prototype p_{c_h} by dynamic programming as the following equation $DTW(g, p_{c_h})=D(m, m)$ with

$$D(u, v) = d(u, v) + \min \left\{ \begin{array}{l} D(u-1, v), \\ D(u, v-1), \\ D(u-1, v-1) \end{array} \right\} \quad (5)$$

where d is a distance function between two elements of the two sequences that has been defined as

$$\begin{aligned} d(u, v) &= \\ &= \sqrt{(x(t_u) - x_{p_{c_h}}(t_v))^2 + (y(t_u) - y_{p_{c_h}}(t_v))^2 + (z(t_u) - z_{p_{c_h}}(t_v))^2} \end{aligned} \quad (6)$$

The two gesture recognition algorithms are compared in section V in terms of recognition rate and computational time.

IV. EXPERIMENTAL SETUP

The experimental setup for the evaluation of the proposed method consists of a motion capture system for recording human arm gestures and a small humanoid robot for imitation of the recognized gestures. Motion capture is performed by tracking the orientation of the upper right arm and the lower right arm using two Xsens MTx inertial motion sensors.

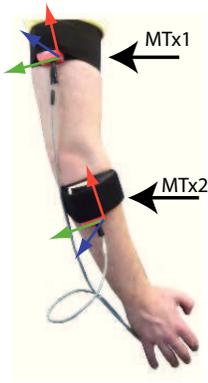


Fig. 2. Configuration of the inertial sensors for arm motion capture.

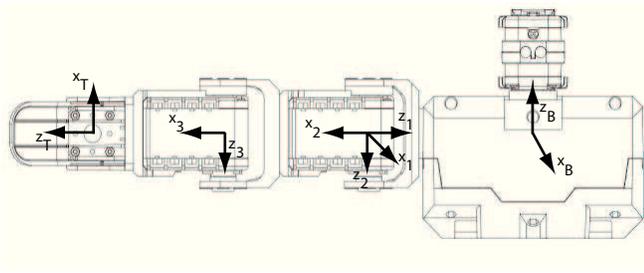


Fig. 3. Schematic diagram of the Bioloid right arm.

The Xsens MTx sensor incorporates a 3-axis accelerometer, gyroscope and magnetometer. The two sensors are attached to the user's body parts as shown in figure 2. A calibration phase is performed for each user, in which the lengths of the upper and lower arm are measured. Due to the absence of a third reference sensor (to be attached to the torso) users are asked not to rotate when performing a gesture and all measured joint angles are expressed with respect to a calibration pose. The calibration pose is recorded at the beginning of each session by asking the user to lay the arm along the body. In the current setup orientation data are sampled at $25Hz$. The Cartesian coordinates of the hand of the user are computed by forward kinematics.

A Robotis Bioloid robot is used for gesture imitation. The Bioloid is a small humanoid ($35cm$ tall, $1.95Kg$ of weight) with a total of 18 degrees of freedom. Figure 3 shows a schematic diagram of the Bioloid right arm with the joint reference frames. Joints are driven by Dynamixel AX-12 motors and synchronized via a control unit based on the Atmega128 microcontroller. Each robot arm has 3 degrees of freedom. In particular, there are two motors at the robot shoulder instead of the three degrees of freedom of the human shoulder. Therefore, the correspondence problem, i.e. the problem of mapping human motion to the robot, has been restricted to the imitation of the Cartesian hand path of the user by the tool point of the robot arm. For motor control, the reachable workspace of the human arm is scaled down to the robot arm workspace. Joint angles of the humanoid arm are computed through inverse kinematics.

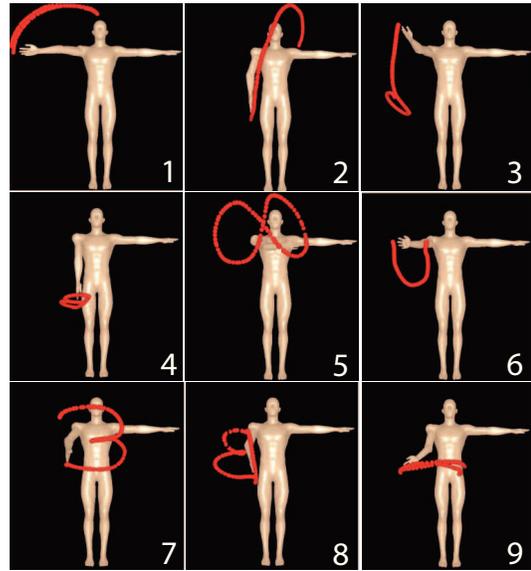


Fig. 4. Example gestures for each class in a real-time simulation environment.

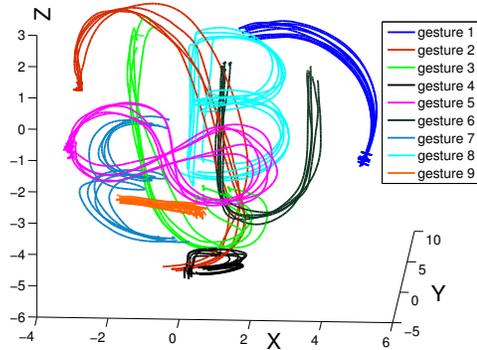


Fig. 5. 3D plot of a subset of gestures. Each class is assigned a color. 3D coordinates are in simulation units (actual length in meters is given by multiplying simulation units by a scale factor of 0.1125).

V. EVALUATION

The proposed method has been evaluated on a dataset of $C=9$ classes of gestures. The gesture classes are as follows: “1-vertical waving”, “2-swing sword”, “3-tennis forehand”, “4-horizontal B letter”, “5-infinite symbol”, “6-vertical U letter”, “7-vertical ϵ letter”, “8-vertical B letter”, “9-horizontal waving”. Figure 4 shows an example gesture for each class in a real-time simulation environment that has been developed as a graphical front-end.

Ten users have been recruited for building the training set (six men and four women, average age of 26 years). Each user performed 10 gestures for each class after a short trial session. Thus the training set contains a total of $N=900$ observations. Each gesture has a duration of $4s$, which consists of a discrete time sequence of $m=100$ samples (at $25Hz$). Figure 5 shows a 3D plot of a small subset of

FPC	% of variance explained
FPC1	48.72%
FPC2	29.65%
FPC3	9.78%
FPC4	8.30%
FPC5	1.77%
FPC6	1.47%
FPC7	0.15%
FPC8	0.09%
FPC9	0.07%

TABLE I

PERCENTAGE OF VARIANCE EXPLAINED BY THE FPCs

class	1	2	3	4	5	6	7	8	9
1	100								
2		100							
3			99	1					
4				100					
5					93		7		
6						100			
7							97	3	
8								100	
9							1		99

TABLE II

CONFUSION MATRIX OF TRAINING SET CLUSTERS

the training gestures. A test set of 495 gestures has been generated by eight users, four of which did not contribute to the generation of the training set. Moreover, to evaluate the robustness of the proposed algorithm 68 gestures of the test set were performed by deliberately moving the arm at non uniform speed with abrupt discontinuities. The variability in gestures pose in both the training and test sets has been estimated as $0.27m$ (by averaging per-class variances).

The functional principal component analysis of the training data was performed by computing the first $R=9$ functional principal components (FPCs). Table I reports the percentage of variance of the training set explained by the nine FPCs. It turns out that the first four FPCs explain about 96.45% of the variance of the input data. All the clusters, as explained in section III-B are automatically detected in the 9-dimensional space of the functional principal components scores by the unsupervised k-means algorithm. Table II reports the confusion matrix of the unsupervised clustering phase of the training data (ground truth is known). The clustering algorithm was able to organize the training data into 9 classes with only 12 errors (98.7%). Figure 6 shows a 3D plot where all the training gestures, in their functional form, have been projected onto the first three FPCs. The nine clusters are clearly visible even in the low dimensional latent space of the first three FPCs.

After clustering, prototype gestures are generated within each class. Figures 7 and 8 show all the clustered training gestures for two classes and the generated prototypes. Tables III and IV show the confusion matrices of the test set obtained by the FPCA and DTW recognition algorithms. Both algorithms achieve high and comparable recognition rates. The FPCA algorithm yields a total of 20 recognition errors over 495 tests (96% success rate) while the DTW

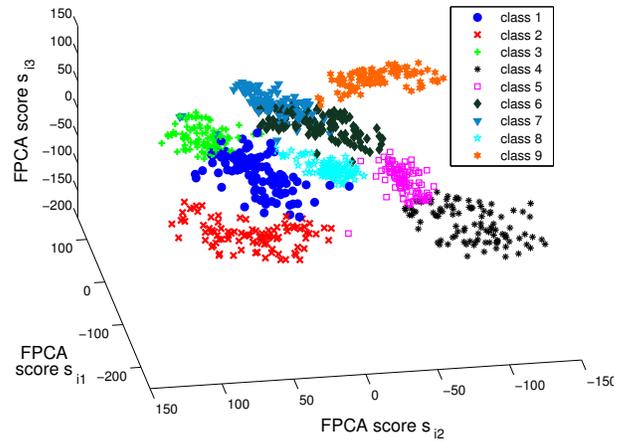


Fig. 6. Projection of the training data onto the first three FPCs. Axes represent the first three functional principal component scores s_{i1}, s_{i2}, s_{i3} of each gesture i .

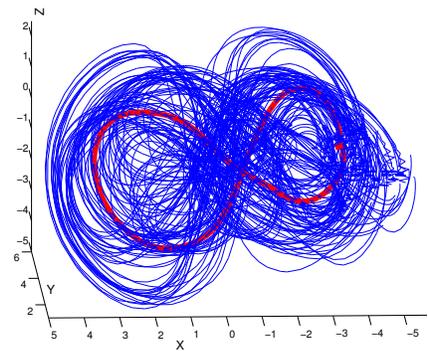


Fig. 7. All gestures clustered in class “5-infinite symbol” (blue curves) and the generated prototype gesture (red curve). 3D coordinates are in simulation units (actual length in meters is given by multiplying simulation units by a scale factor of 0.1125).

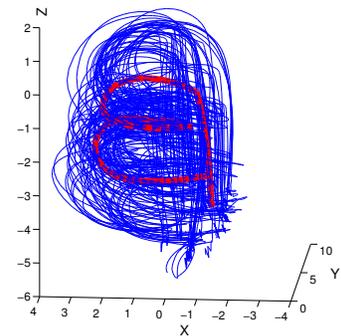


Fig. 8. All gestures clustered in class “8-vertical B letter” (blue curves) and the generated prototype gesture (red curve). 3D coordinates are in simulation units (actual length in meters is given by multiplying simulation units by a scale factor of 0.1125).

algorithm obtains a total of 10 recognition errors (98% success rate). It must be remarked that the test set contains

class	1	2	3	4	5	6	7	8	9
1	55								
2		52					1	2	
3			53					2	
4				55					
5					54		1		
6						54		1	
7						1	45	9	
8				2				53	
9				1					54

TABLE III
CONFUSION MATRIX OF TEST SET USING FPCA

class	1	2	3	4	5	6	7	8	9
1	55								
2		55							
3			55						
4				55					
5					55				
6						55			
7						3	49	3	
8				3		1		51	
9									55

TABLE IV
CONFUSION MATRIX OF TEST SET USING DTW

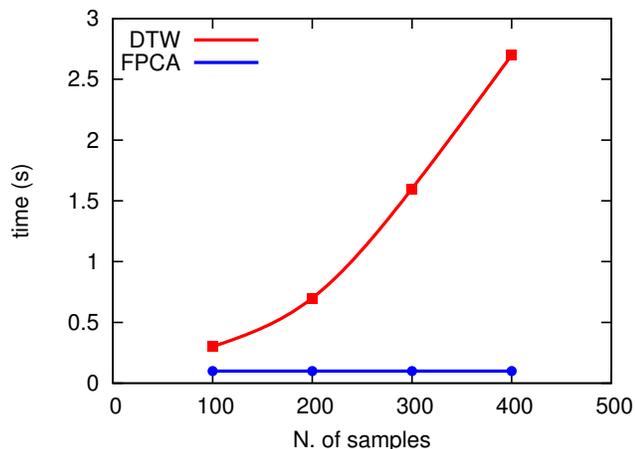


Fig. 9. Comparison of FPCA and DTW execution times for gesture recognition.

14% of gestures that have been performed at non-uniform speed, which did not affect the performance negatively, thus confirming the robustness of both recognition approaches.

The two algorithms have also been compared in terms of execution time by recognizing gestures performed with increasing durations, as shown in figure 9. The execution time of the recognition phase of the FPCA algorithm is rather independent from gesture duration. On the contrary, the execution time of the DTW algorithm has quadratic time complexity. Indeed, the high computational cost of the DTW is a known drawback of this popular algorithm. In particular, DTW requires about 2.5s to recognize a gesture containing 400 samples (16s of duration) while FPCA takes about 0.1s regardless of the number of samples. Experiments have

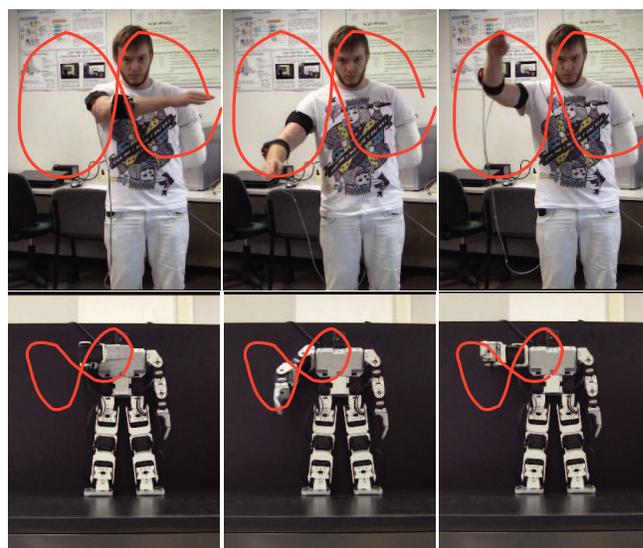


Fig. 10. Humanoid imitation experiment for a “infinite symbol” gesture (order from left to right).

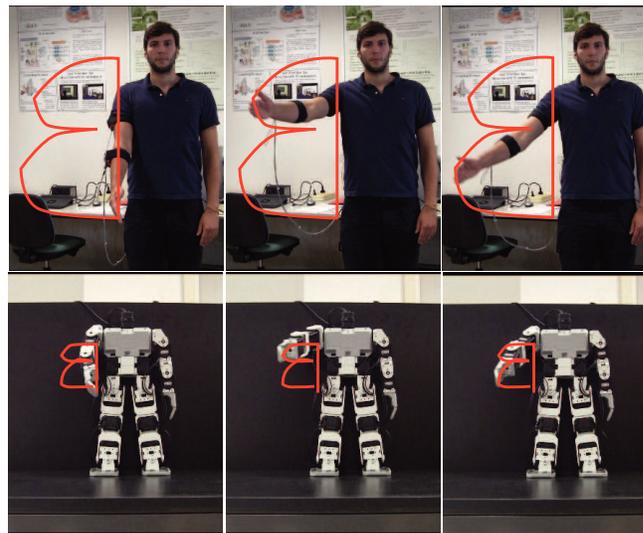


Fig. 11. Humanoid imitation experiment for a “vertical B letter” gesture (order from left to right).

been performed in MATLAB on an Intel core i7@2.80GHz. Figures 10 and 11 show the result of two robot imitation experiments for an “infinite symbol” gesture and a “vertical B letter” where the humanoid robot reproduces the prototype gesture of the recognized class.

VI. CONCLUSIONS

In this work a novel method for gesture recognition has been proposed. The approach is based on functional principal component analysis, a statistical tool that extends standard multivariate PCA when data consist of functional observations. FPCA extracts a finite set of functional principal components that form an orthogonal basis. Input data are projected onto functional principal components for dimension reduction. A gesture recognition algorithm has been developed in the latent space of functional principal components

scores. The recognition rate of the proposed FPCA algorithm is comparable to a state-of-the-art DTW algorithm. However, the proposed approach is more efficient in execution time and more appropriate for classification of long duration gestures because it does not suffer from the curse of dimensionality. Experiments of humanoid gesture imitation have also been reported. In future work the gesture recognition and imitation approach will be extended to bi-manual gestures.

REFERENCES

- [1] A. Akl, Chen Feng, and S. Valaee. A novel accelerometer-based gesture recognition system. *IEEE Transactions on Signal Processing*, 59(12):6197–6205, dec. 2011.
- [2] A. Akl and S. Valaee. Accelerometer-based gesture recognition via dynamic-time warping, affinity propagation, & compressive sensing. In *IEEE Intl Conference on Acoustics Speech and Signal Processing (ICASSP)*, pages 2270–2273, march 2010.
- [3] J. Aleotti and S. Caselli. Trajectory clustering and stochastic approximation for robot programming by demonstration. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1029–1034, aug. 2005.
- [4] S. Bitzer and S. Vijayakumar. Latent spaces for dynamic movement primitives. In *9th IEEE-RAS Intl Conference on Humanoid Robots*, pages 574–581, dec. 2009.
- [5] S. Calinon and A. Billard. Stochastic gesture production and recognition model for a humanoid robot. In *IEEE/RSJ Intl Conference on Intelligent Robots and Systems (IROS)*, volume 3, pages 2769–2774, sept.-2 oct. 2004.
- [6] S. Calinon, F. Guenter, and A. Billard. Goal-directed imitation in a humanoid robot. In *Proceedings of the IEEE Intl Conference on Robotics and Automation (ICRA)*, Barcelona, Spain, April 2005.
- [7] J. M. Carmona and J. Climent. A performance evaluation of HMM and DTW for gesture recognition. In *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, volume 7441 of *Lecture Notes in Computer Science*, pages 236–243. Springer Berlin Heidelberg, 2012.
- [8] B. Dariush, M. Gienger, A. Arumbakkam, C. Goerick, Youding Zhu, and K. Fujimura. Online and markerless motion retargeting with kinematic constraints. In *IEEE/RSJ Intl Conference on Intelligent Robots and Systems (IROS)*, pages 191–198, sept. 2008.
- [9] J. De Schutter, E. Di Lello, J.F.M. De Schutter, R. Matthysen, T. Benoit, and T. De Laet. Recognition of 6 dof rigid body motion trajectories using a coordinate-free representation. In *IEEE Intl Conference on Robotics and Automation (ICRA)*, pages 2071–2078, may 2011.
- [10] K. Dermitzakis, A.H. Arieta, and R. Pfeifer. Gesture recognition in upper-limb prosthetics: A viability study using dynamic time warping and gyroscopes. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 4530–4533, 30 2011-sept. 3 2011.
- [11] E. Drumwright and M.J. Mataric. Generating and recognizing free-space movements in humanoid robots. In *IEEE/RSJ Intl Conference on Intelligent Robots and Systems (IROS)*, volume 2, pages 1672–1678, oct. 2003.
- [12] D. Herzog, A. Ude, and V. Kruger. Motion imitation and recognition using parametric hidden Markov models. In *8th IEEE-RAS Intl Conference on Humanoid Robots*, pages 339–346, dec. 2008.
- [13] T. Inamura, N. Kojo, T. Sonoda, K. Sakamoto, K. Okada, and M. Inaba. Intent imitation using wearable motion capturing system with on-line teaching of task attention. In *Proceedings of the IEEE-RAS Intl Conference on Humanoid Robots*, Tsukuba, Japan, Dec. 2005.
- [14] T. Inamura, Y. Nakamura, H. Ezaki, and I. Toshima. Imitation and primitive symbol acquisition of humanoids by the integrated mimesis loop. In *Proceedings of the IEEE Intl Conference on Robotics and Automation (ICRA)*, Seoul, Korea, May 2001.
- [15] M. Ito and J. Tani. On-line imitative interaction with a humanoid robot using a mirror neuron model. In *Proceedings of the IEEE Intl Conference on Robotics and Automation (ICRA)*, New Orleans, LA, April 2004.
- [16] Seungsu Kim, Chang Hwan Kim, and Jong Hyeon Park. Human-like arm motion generation for humanoid robots using motion capture database. In *IEEE/RSJ Intl Conference on Intelligent Robots and Systems (IROS)*, pages 3486–3491, oct. 2006.
- [17] D. Kulic, W. Takano, and Y. Nakamura. Incremental learning, clustering and hierarchy formation of whole body motion patterns using adaptive hidden Markov chains. *The International Journal of Robotics Research*, 27(7):761–784, 2008.
- [18] Doo Young Kwon and M. Gross. A framework for 3d spatial gesture design and modeling using a wearable input device. In *11th IEEE Intl Symposium on Wearable Computers*, pages 23–26, oct. 2007.
- [19] Junghyun Kwon and F.C. Park. Natural movement generation using hidden Markov models and principal components. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 38(5):1184–1194, oct. 2008.
- [20] Wenyu Lu, Yan Liu, Jizhou Sun, and Libo Sun. A motion retargeting method for topologically different characters. In *Sixth Intl Conference on Computer Graphics, Imaging and Visualization (CGIV)*, pages 96–100, aug. 2009.
- [21] P. Menezes, F. Lerasle, J. Dias, and R. Chatila. A single camera motion capture system dedicated to gestures imitation. In *Proceedings of the IEEE-RAS Intl Conference on Humanoid Robots*, Tsukuba, Japan, Dec. 2005.
- [22] C. Ott, Dongheui Lee, and Y. Nakamura. Motion capture based human motion recognition and imitation by direct marker control. In *8th IEEE-RAS Intl Conference on Humanoid Robots*, pages 399–405, dec. 2008.
- [23] N.S. Pollard, J.K. Hodgins, M.J. Riley, and C.G. Atkeson. Adapting human motion for the control of a humanoid robot. In *Proceedings of the IEEE Intl Conference on Robotics and Automation (ICRA)*, Washington, May 2002.
- [24] J. Ramsay and B. W. Silverman. *Functional Data Analysis*. Springer Series in Statistics, 2005.
- [25] M. Riley, A. Ude, K. Wade, and C.G. Atkeson. Enabling real-time full-body imitation: a natural way of transferring human movement to humanoids. In *Proceedings of the IEEE Intl Conference on Robotics and Automation (ICRA)*, Taipei, Taiwan, Sept. 2003.
- [26] M. Salvati, Jeong Seung Zoo, N. Hashimoto, and M. Sato. "do like me", "do like this": Creating animations by teaching a virtual human. In *IEEE Intl Conference on Shape Modeling and Applications (SMI)*, pages 43–50, june 2007.
- [27] S. Sempena, N.U. Maulidevi, and P.R. Aryan. Human action recognition using dynamic time warping. In *Intl Conference on Electrical Engineering and Informatics (ICEEI)*, pages 1–5, july 2011.
- [28] A.P. Shon, K. Grochow, and R.P.N. Rao. Robotic imitation from human motion capture using gaussian processes. In *Proceedings of the IEEE-RAS Intl Conference on Humanoid Robots*, Tsukuba, Japan, Dec. 2005.
- [29] M. Sigalas, H. Baltzakis, and P. Trahanias. Gesture recognition based on arm tracking for human-robot interaction. In *IEEE/RSJ Intl Conference on Intelligent Robots and Systems (IROS)*, pages 5424–5429, oct. 2010.
- [30] W. Suleiman, E. Yoshida, F. Kanehiro, J.-P. Laumond, and A. Monin. On human motion imitation by humanoid robot. In *IEEE Intl Conference on Robotics and Automation (ICRA)*, pages 2697–2704, may 2008.
- [31] W. Takano and Y. Nakamura. Humanoid robot's autonomous acquisition of proto-symbols through motion segmentation. In *6th IEEE-RAS Intl Conference on Humanoid Robots*, pages 425–431, dec. 2006.
- [32] A. Thobbi and Weihua Sheng. Imitation learning of hand gestures and its evaluation for humanoid robots. In *IEEE Intl Conference on Information and Automation (ICIA)*, pages 60–65, june 2010.
- [33] M. Urban, P. Bajcsy, R. Kooper, and J.-C. Lementec. Recognition of arm gestures using multiple orientation sensors: repeatability assessment. In *7th Intl IEEE Conference on Intelligent Transportation Systems*, pages 553–558, oct. 2004.
- [34] A. Vakanski, I. Mantegh, A. Irish, and F. Janabi-Sharifi. Trajectory learning for robot programming by demonstration using hidden Markov model and dynamic time warping. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 42(4):1039–1052, aug. 2012.
- [35] Zhidong Xiao, J.J. Zhang, and S. Bell. Control of motion in character animation. In *Proceedings Eighth International Conference on Information Visualisation (IV 2004)*, pages 841–848, july 2004.
- [36] A. Zinnen and B. Schiele. A new approach to enable gesture recognition in continuous data streams. In *12th IEEE Intl Symposium on Wearable Computers*, pages 33–40, 28 2008-oct. 1 2008.