Multifunctional Principal Component Analysis for Human-like Grasping

Marco Monforte, Fanny Ficuciello, and Bruno Siciliano

PRISMA Lab, Department of Electrical Engineering and Information Technology
University of Naples Federico II
Via Claudio 21, 80125, Naples, Italy

Abstract. In this paper, a method to derive the synergies subspace of an anthropomorphic robotic arm–hand system is proposed. Several human demonstrations of different objects grasping are measured using the Xsens MVN suite and then mapped to a seven Degree-of-Freedom (DoF) robotic arm. Exploiting the anthropomorphism of the kinematic structure of the manipulator, two Closed-Loop Inverse Kinematics (CLIK) algorithms are used to reproduce accurately the master's movements. Once the database of movements is created, the synergies subspace are derived applying the Multivariate Functional Principal Component Analysis (MFPCA) in the joint space. A mean function, a set of basis functions for each joint and a pre-defined number of scalar coefficients are obtained for each demonstration. In the computed subspace each demonstration can be parametrized by means of a few number of coefficients, preserving the major variance of the entire movement. Moreover, a Multilevel Neural Networks (MNNs) is trained in order to approximate the relationship between the object characteristics and the synergies coefficients, allowing generalization for unknown objects. The tests are conducted on a setup composed by a KUKA LWR 4+ Arm and a SCHUNK 5-Finger Hand, using the Xsens MVN suite to acquire the demonstrations.

1 INTRODUCTION

In order to execute complex tasks and interact deftly with the environment, robots are becoming more and more sophisticated and endowed with great manipulation capabilities. This dexterity requires a high number of DoFs of the arm–hand system, along with high sensorimotor and reasoning skills. On the other hand, this implies the need of greater computational capabilities, more powerful hardware and different control strategies to speed up and simplify the execution of these complex tasks.

Despite a classical model-based method [1], in the latest years robotics and neuroscience researchers have focused their efforts on bio-inspired systems and joined their efforts in order to try to understand and recreate particular aspects, structures and/or behaviours of biologic systems. One of the main sources of inspiration is, of course, the human being [2], [3].

Santello *et al* [4] and later Mason *et al* [5] show that human hand movements during grasping tasks are computed in a space of highly reduced dimensions with respect to the joint space.

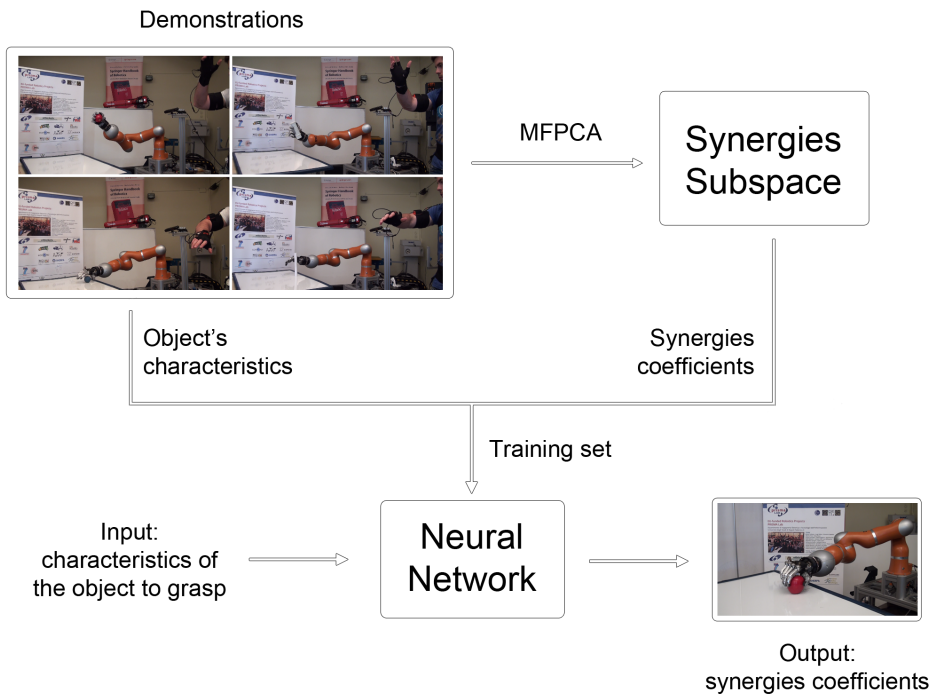

Fig. 1. Diagram of the methodological approach

This reduced subspace is called synergies subspace. A synergy mapping method from a human hand to a robotic one is proposed in [6], operating in the task space and with the usage of a virtual sphere. In [7] a Neural Network (NN) is trained with the object's features and the synergies coefficients in order to generalize the grasping strategy.

Grasping tasks do not rely only on the hand, but also on the arm configuration. A possible way to generalize the arm motion could be the Dynamical Movement Primitives (DMPs) [8], as explored in [9], [10]. Another possibility is the MFPCA. With the aim of extending the work done in [7] and in [3], we propose to obtain the synergies subspace of a robotic arm–hand system with the MFPCA technique, which is an extension of the Functional Principal Component Analysis (FPCA) to the multivariate case. FPCA, that allows computing the dominant modes of variation in functional data, can be seen as the extension of the Principal Component Analysis (PCA) to the functional case. The motivation to investigate the MFPCA for dimensionality reduction is to preserve as

much information as possible contained in the motion coordination. Instead of final configuration considered in [7], [3], here the demonstrations are time series.

To this purpose, reach-to-grasp movements are acquired by teleoperating the robotic system. The grasped objects have different dimensions. On this set of demonstrations, the MFPCA is applied in order to link each object to a fixed number of scalar coefficients that reproduce the grasping movement.

Finally, MNNS are trained with pairs of object's characteristics and scalar coefficients in order to try to generalize the behaviour to new objects. A schematic recap is in Fig. 1. The scope of this work is to compute the synergies subspace of the arm–hand system and to evaluate if MNNs, successfully tested on the hand, are good enough to find the function that maps the object characteristics into the MFPCA coefficients. In future works, a Reinforcement Learning (RL) approach could be introduced to explore the space around the final configuration computed by the MNN.

The paper is organized as follows: Section 2 describes the experimental setup; Section 3 describes the technical approach adopted for the human motion mapping on the robotic arm, for the synergies subspace derivation and for the MNN training; Section 4 illustrates the mapping procedure using the Xsens MVN; in Section 5 the MFPCA is briefly described together with the synergies subspace computation procedure; in Section 6 the training of the MNN is analyzed along with the results; finally, Section 7 provides conclusions and future developments.

2 EXPERIMENTAL SETUP

The arm–hand robotic system used to acquire the demonstrations is composed by a SCHUNK 5-Finger Hand (S5FH) and a KUKA Lightweight Robot (LWR) 4+ arm. The S5FH is an underactuated anthropomorphic robotic hand with 20 DoFs actuated by 9 motors. Its mechanical design presents an example of hand synergies through mechanical couplings, to reduce the weight and the dimensions. The KUKA robotic arm has 7 DoFs like the human arm. The redundant joint allows changing the internal configuration of the arm while keeping the end-effector in a fixed pose. The Robot Operating System (ROS) is used to control the S5FH through an SVH Driver developed by Forschungszentrum Informatik (FZI). This driver for the low-level interface allows an easy control using a C++ customized library. The KUKA LWR 4+ is controlled using the Fast Research Interface (FRI) library, developed at Stanford University. The motion capture tool used to acquire the demonstrations is the Xsens MVN package, composed by the Xsens suite and the MVN software. The Xsens suite does not need cameras, emitters or markers but consists essentially of 17 MTx sensors. After a calibration process, it is possible to obtain position and orientation of the body parts where the sensors are attached. Each MTx includes a 3D gyroscope, a 3D accelerometer and a 3D magnetometer. The MVN software, instead, receives the data from the suite and reproduces the human movements in a simulator. Moreover, the network streamer can be used to send the data received to third

applications using UDP/TCP-IP packets. Finally, Matlab NN Toolbox has been used to train the MNNs.

Fig. 2. SCHUNK 5-Finger Hand, KUKA LWR 4+ and Xsens suite

3 TECHNICAL APPROACH

In order to replicate human manipulation skills, artificial limbs for robotics and prosthetics require a biomechanical human-like structure with high number of DoFs. On the other hand, recent advances in neuroscience research have shown that the human Central Nervous System (CNS) coordinates these DoFs to generate complex tasks by means of a synergistic organization at postural, muscular and neural level. Using a limited set of motion patterns, named postural synergies, a wide range of grasping tasks can be achieved, reducing substantially the dimension of the robotic hand control problem. The main objective of this work is to apply this concept of dimensionality reduction to a robotic arm, generalizing it to the entire movement and not only looking at the final pose. We have investigated the MFPCA technique and to this purpose, we firstly need a set of demonstrations. Thus, the first problem is how to create this dataset.

To provide the robotic system with human-like movements, the demonstrations are obtained by teleoperating the systems. The main advantages of this technique is to be fast and general: if we need further demonstrations, we just execute them, recording the variables of interest while performing the task. Imitation learning can be achieved in different ways, the main methodologies are two: teleoperation, using tracking systems like vision, exoskeletons or other wearable motion sensors, and kinesthetic teaching, where the robot is physically moved by the human to execute the task. Since we deal with high-DoF platforms, the latter method is not really practicable: the human master should move too many joints at the same time, resulting in discontinuous movements and jerky trajectories. Therefore, in this work the system is teleoperated using a motion tracking suite. However, in order to reproduce the demonstrations as close as possible,

teleoperation needs a mapping of human motion to the robotic system. For this purpose, two Closed-Loop Inverse Kinematics (CLIK) algorithms are used to control the KUKA arm. The first CLIK algorithm controls the first three joints (shoulder) by giving to the robotic arm the same orientation of the master. The second CLIK algorithm controls the S5FH palm orientation with respect to the robotic forearm using the last three joints of the robot. Moreover, the null space of the second CLIK Jacobian, that relates the velocities of the last four joints to the Cartesian space velocities, is used to reproduce the angle between the human arm and forearm. Once the mapping problem is solved, several demonstrations are acquired to cover a complete grasp taxonomy [6].

Table 1. Objects used for the demonstrations

	Object	**Diameter [cm]**	**Height [cm]**
Balls	Bottle cap	2.7	2.7
	Table tennis ball	4.4	4.4
	Plastic plum	5.2	5.2
	Racquetball	5.5	5.5
	Plastic peach	5.9	5.9
	Tennis ball	6.4	6.4
	Blue ball	6.9	6.9
	Plastic orange	7.1	7.1
	Plastic apple	7.4	7.4
	Red ball	8.8	8.8
	Softball	9.6	9.6
Cylinders	N°1	1.2	21
	N°2	1.8	21
	N°3	2.4	21
	N°4	3.3	21
	N°5	4.9	18.2
	N°6	5.1	16
	N°7	5.4	23
	N°8	6	21
	N°9	6.6	21
	N°10	7	21.0
	Chips tube	7.5	21.0

Robot joint trajectories are recorded during teleoperated reach and grasp tasks of different objects (Table 1). The considered objects are “balls” and “cylinders” of different dimension grasped with precision and/or a power grasp, depending on whether the object allows both or not.

Once the dataset is created, Dynamic Time Warping technique [11] is applied for better data conditioning, and a subsampling is carried out in order to have all the time series as similar as possible and with same length.

Finally, we have a database of 38 demonstrations with the same length. Afterwards, MFPCA technique is applied and a set of scalar coefficients parametrizing the demonstrations are obtained.

The pairs constituted by object characteristics and related synergies coefficients have been used to train an MNN using Matlab Toolbox. The goal is to generalize to novel objects the synergy-based grasping motion obtained from the MFPCA.

4 HUMAN MOTION MAPPING

Regarding the human hand motion mapping to the S5FH, the work already done in [8] is exploited. A Jacobian matrix of the synergies is computed and used into a CLIK algorithm to map human hand grasping configurations to the robotic hand. Further details of the mapping and the results can be found [9].

To map the human motion to the robotic arm, we recall that the MTx sensor on the Xsens suite provides position and orientation of the human limbs whith respect to the global frame of the mo-cap system. Moreover, the KUKA arm presents 7 DoFs as the human arm. Thus the developed method to map the motion involves two CLIK algorithms that take the arm and the hand orientations as reference input respectively.

We look at the robotic arm like two separate kinematic structures: the first one made by three joints (shoulder) and the second one made by four joints (wrist) with one redundant DoF (elbow). To reproduce at best the human movement, the angle between the arm and the forearm is computed with geometric properties and assigned to the redundant joint, exploiting the null space of the Jacobian matrix of the second kinematic structure.

With the described mapping algorithm, we are able to teleoperate the KUKA arm, as can be seen in Fig. 3. 38 demonstrations are performed and the seven joint trajectories are recorded along the entire task execution.

5 SYNERGIES SUBSPACE COMPUTATION

The PCA technique is a tool for multivariate data analysis and dimensionality reduction, extended to functional data from 1950, as Functional Principal Component Analysis (FPCA). Functional data can be seen as the realization of a stochastic process that, under mild assumptions, can be expressed as a sequence of uncorrelated random variables, called Functional Principal Components (FPCs). The i-th realization of the stochastic process X has the form:

$$X_i(t) = \mu(t) + \sum_{k=1}^{\infty} \xi_{ik}\varphi_k(t) \quad t \in T \tag{1}$$

Fig. 3. Snapshots of performed demonstrations

where $\mu(t)$ is the mean function of the dataset, φ_k is the k-th functional principal component and ξ_{ik} is the k-th coefficient of the respective FPC that parametrize the i-th realization.

The basic idea is to map the initial domain of the dataset into a subspace that allows parametrizing each realization with a reduced number of scalar parameters. The major advantage of these techniques is that they preserve major variance with respect to other transformations like Fourier or Wavelet.

In the PCA each input realization is a vector of scalar components and the dimension is finite. The mean value is a vector and the eigenvalues/eigenvectors of the entire dataset have a finite number of components. In the FPCA each realization is a time series, hence the dimension is infinite. The mean value is a mean function and the number of eigenvalues/eigenfunctions of the entire dataset is infinite (Tab. 2).

The FPCA can be extended to the multivariate case as Multifunctional Principal Component Analysis (MFPCA). Each observation has now $n \geq 2$ functions, possibly defined on different domains $T_1, ...T_n$. It has the form of a vector X such that:

$$X_i(\mathbf{t}) = (X_i^{(1)}(t_1), ..., X_i^{(n)}(t_n)) \in \mathbb{R}^n \tag{2}$$

with $\mathbf{t} = (t_1, ..., t_n) \in T = T_1 \times ... \times T_n$.

The decomposition of a realization of the stochastic process can be thus extended as:

$$X_i(\mathbf{t}) = \mu(\mathbf{t}) + \sum_{k=1}^{\infty} \xi_{ik}\varphi_k(\mathbf{t}) \quad \mathbf{t} \in T \tag{3}$$

By truncation of the sum to m components, we are capable of parametrizing each realization with m scalar coefficients, also called *scores*.

Table 2. Comparison between PCA and FPCA techniques

Element	In PCA	In FPCA
Data	$X \in \mathbb{R}^p$	$X \in L^2(T)$
Dimension	$p < \infty$	∞
Mean	$\mu = \mathrm{E}(X)$	$\mu(t) = \mathrm{E}(X(t))$
Covariance	$\mathrm{Cov}(X) = \Sigma_{p\times p}$	$\mathrm{Cov}(X(s), X(t)) = G(s,t)$
Eigenvalues	$\lambda_1, \lambda_2, ..., \lambda_p$	$\lambda_1, \lambda_2, ...$
Eigenvectors/Eigenfunctions	$\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_p$	$\varphi_1(t), \varphi_2(t), ...$
Inner Product	$\langle \mathbf{X}, \mathbf{Y} \rangle = \Sigma_{k=1}^{p} X_k Y_k$	$\langle \mathbf{X}, \mathbf{Y} \rangle = \int_T X(t)Y(t)dt$
Principal Components	$z_k = \langle X - \mu, \boldsymbol{v}_k \rangle, k = 1, 2, ..., p$	$\xi_k = \langle X - \mu, \varphi_k \rangle, k = 1, 2, ...$

After the data conditioning, we have now the matrix $\mathbf{D} = \{\mathbf{d}_i \mid i = 1, ..., 38\}$ of the demonstrations. Each $\mathbf{d_i}$ is a reaching and grasping execution in different conditions (different objects or same object with different grasp types) and is composed by seven functions representing the values of the KUKA arm joints during the demonstration of the task:

$$\mathbf{d}_i = \left(d_i^{(1)}(t), ..., d_i^{(7)}(t)\right) \quad t \in T \tag{4}$$

The MFPCA is applied on $\mathbf{D}$ by setting the desired number of FPCs $m = 4$ (see [11] for further details on the procedure). We obtain mean functions $\mu^{(j)}(t)$ and m eigenfunctions $\varphi_1^{(j)}, ..., \varphi_m^{(j)}$ for each joint ($j = 1, ..., 7$). The m eigenfunctions represent the basis of the subspace for the relative joint. Finally, we obtain a matrix of scores $\Xi \in \mathbb{R}^{38\times m}$ where each row parametrizes a demonstration. It is important to highlight that the k-th score modulates the correspondent k-th FPC for each joint. In other words, the coefficient the same for all the seven joints, while the eigenfunctions (and thus the basis of the subspace) are different for each of them.

Reconstructing the x, y and z trajectories (Fig. 4) we can see that 2 FPCs represent quite well the original data. Yet, if we rely only on MNNs to grasp the objects, we will need to be more precise. Therefore, the usage of 3 or 4 FPCs will improve performances. Another important advantage is the filtering effect of the MFPCA: the reconstructions eliminate the noise acquired while performing the demonstration.

6 NEURAL NETWORK TRAINING

To confer autonomy to the system, MNNs can be a suitable choice due to their ability to learn mapping functions. Moreover, the usage of synergies reduces the search space of the learning algorithm, implying also a simplification of the

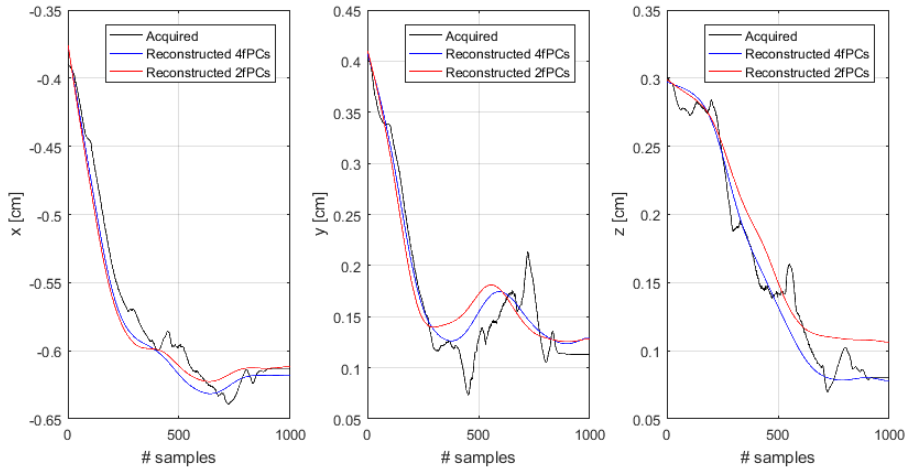

Fig. 4. x, y and z coordinates reconstructed using 2 FPCs (red) and 4 FPCs (blue).

architecture, especially regarding the number of hidden layers and number of neurons.

Two MNNs are used in this work, one for the hand and one for the arm. Their architecture is experimentally chosen by trying different combinations of hidden levels and neurons and analyzing the corresponding performance in terms of Mean Squared Error (MSE). We found out that a feedforward structure with 2 hidden layers and 10 neurons for each is the suitable choice for the particular application.

To learn the function that relates the object shape and grasp type with the arm-hand grasping action, MNNs need a training dataset constituted by pairs of object geometric features and scores obtained from the MFPCA. The input parameters are the *height*, *diameter* and an information on the *grasp type* (power or precision). Since MMNs inputs must have the same dimensions for all the objects, the *height* is set equal to the *diameter* for spherical shaped objects. The output are the three synergy coefficients that determine the configuration of the SCHUNK 5-Finger Hand and the 4 coefficients of the KUKA arm that determine the reaching movement towards the object, respectively for each MNN. It is well known that neural networks performance increase when decrease the dimension of the space where is defined the function to be learned. Therefore, for the arm we decided to test the net with 4 FPCs since they can reconstruct the motion with a lower error with respect to the use of the first two FPCs, meanwhile the dimension of the search space is reduced from 7 to 4 with a good compromise between performance and generalization capabilities. The goal now is to experimentally evaluate if the use of MNNs based on FPCs can generalize grasps of new objects with respect to the database used for training.

The first evaluation is made by comparing the waveforms obtained from the Multilevel Neural Network with the reconstruction shown in Fig. 4 using 4 eigenfunctions.

The comparison is performed on a grasp example contained in the database and is shown in Fig. 5. As we expected, the combination of eigenfunctions computed with the Multilevel Neural Network provides a higher reconstruction error for each grasp in the database. This is the drawback of the generalization process provided by the net. To evaluate the ability of the MNN to generalize the grasps, we have selected two objects with geometrical features closed to objects contained in the database and two objects different from the examples contained in the database.

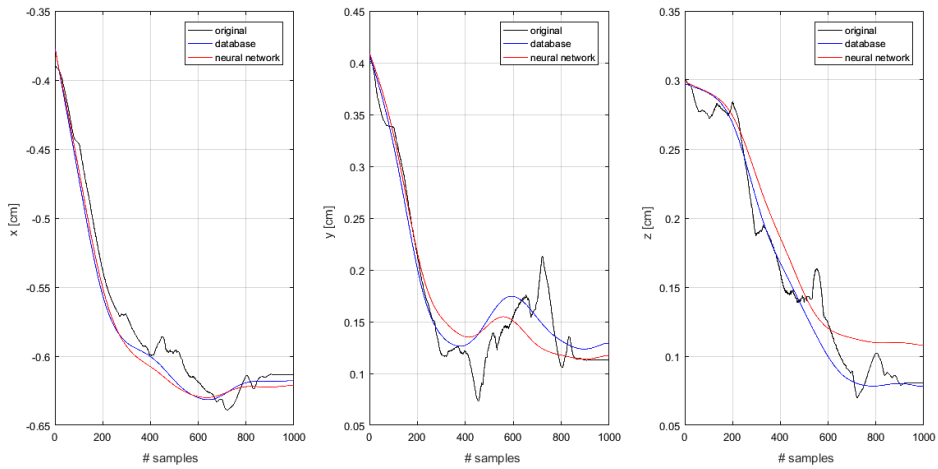

Fig. 5. Comparison between the reconstruction with 4 FPCs of Fig. 4 and the reconstruction with 4 FPCs obtained from the neural network.

From experiments represented in Fig. 6, we can state that the MNNs are not good enough to realize grasps of new objects but need to be integrated with strategies based on trial-and-error.

Indeed, from an error analysis performed on a wide set of objects, the average of the errors between the hand palm posture planned by the MNNs and the hand posture measured from the human demonstration is:

$$\mathbf{e}_p = \begin{bmatrix} 0.99 \; 0.93 \; 1.64 \end{bmatrix} \tag{5}$$

$$\mathbf{e}_o = \left(\begin{bmatrix} 0.65 \; 0.51 \; 0.55 \end{bmatrix}, 0.20944 \right) \tag{6}$$

where $\mathbf{e}_p$ is the position error expressed in [cm] and $\mathbf{e}_o$ is the orientation error with the axis-angle system (angle in radians).

Fig. 6. Examples of grasping actions obtained using the MFPCA outputs. Figures on the top represent grasps of objects closed to ones contained in the dataset. Figures on the bottom represent grasps of objects with dimensions that are more different with respect to the dataset ones.

Despite the pose error that doesn't always ensure successful grasps, the hand for each grasp is very close to the target, this makes the coefficients provided by the MNNs a good initialization for a reinforcement learning algorithm.

7 CONCLUSIONS AND FUTURE WORK

Multifunctional Principal Component Analysis method to derive the synergies subspace of a robotic system with a high number of DoFs is proposed in this paper. MFPCA technique is an extension of the well-known PCA to the multivariate functional data. Several demonstrations of reach-to-grasp tasks for objects with different dimensions are acquired. The acquisitions are carried out using a motion capture system and a mapping of the human movement to the robotic arm. MFPCA is applied to the collected data by defining a set of FPCs for each joint and a set of scalar coefficients, each one parametrizing a demonstration. By considering the first four FPCs for the arm, it is shown that the reconstructed motion provides a very good approximation of the original functional data.

Afterwords, we used MNNs to generalize the grasps. The training procedure is carried out using the Matlab NN Toolbox and pairs of objects geometric features and relative FPCs coefficients. We have observed that the MNNs it is not sufficient to plan new object grasps. Nevertheless, even if the grasp fails, the hand is close to the correct position. This is a necessary condition for further improvements of the system using a reinforcement learning strategy. Future works, will

develop solutions to explore the space around the initial configuration provided by the MNNs, in order to look for a correct final configuration that allows the system to grasp the object.

ACKNOWLEDGMENTS

The research leading to these results has been partially supported by the RoDyMan project, funded by the European Union (EU) Seventh Framework Programme (FP7/2007-2013) under ERC AdG-320992, and partially by MUSHA project, National Italian Grant under Programma STAR Linea 1. The authors are solely responsible for the content of this paper, which does not represent the opinion of the EU, and the EU is not responsible for any use that might be made of the information contained therein.

References

1. F. Ficuciello, R. Carloni, L.C. Visser, S. Stramigioli, "Port-Hamiltonian modeling for soft-finger manipulation", in *Proc. IEEE/RSJ Internationl Conference on Intelligent Robots and Systems*, Taipei, Taiwan, pp. 4281–4286, 2010.
2. F. Ficuciello, G. Palli, C. Melchiorri, B. Siciliano, "Planning and control during reach to grasp using the three predominant UB Hand IV postural synergies", in *Proc. IEEE International Conference on Robotics and Automation*, St. Paul MN, USA, pp. 2255–2260, 2012.
3. F. Ficuciello, D. Zaccara, B. Siciliano, "Synergy-based policy improvement with path integrals for anthropomorphic hands", in *Proc. IEEE/RSJ Internationl Conference on Intelligent Robots and Systems*, Daejeon, Korea, pp. 1940–1945, 2016.
4. M. Santello, M. Flanders, J. Soechting, "Postural hand synergies for tool use", *Journal of Neuroscience*, 18(23), 10105–10115, 1998.
5. C. Mason, J. Gomez, T. Ebner, "Hand synergies during reach-to-grasp", *Journal of Neurophysiology*, 86(6), 2896–2910, 2001.
6. G. Gioioso, G. Salvietti, M. Malvezzi, D. Prattichizzo, "Mapping synergies from human to robotic hands with dissimilar kinematics: An object based approach", in *Proc. IEEE International Conference on Robotics and Automation, Workshop on Manipulation Under Uncertainty*, Shanghai, China, 2011.
7. F. Ficuciello, G. Palli, C. Melchiorri, B. Siciliano, "Postural synergies and neural network for autonomous grasping: A tool for dextrous prosthetic and robotic hands", in *Proc. International Conference on NeuroRehabilitation*, Toledo, Spain, 2012.
8. A.J. Ijspeert, J. Nakanishi, H. Hoffmann, P. Pastor, S. Schaal, "Dynamical movement primitives: Learning attractor models for motor behaviors", *Neural Computation*, 25(2), 328–373, 2013.
9. P. Pastor, H. Hoffmann, T. Asfour, S. Schaal, "Learning and generalization of motor skills by learning from demonstrations", in *Proc. IEEE International Conference on Robotics and Automation*, Kobe, Japan, pp. 763–768, 2009.
10. A. Gams, A. Ude, "Generalization of example movements with dynamic systems", in *9th IEEE–RAS International Conference on Humanoid Robots*, Paris, France, pp. 28–33, 2009.

11. T. Giorgino, "Computing and visualizing dynamic time warping alignments in R: The dtw package", *Journal of Statistical Software*, 31(7), 1–24, 2009.
12. F. Ficuciello, A. Federico, V. Lippiello, B. Siciliano, "Synergies evaluation of the SCHUNK S5FH for grasping control", in *Proc. 15th International Symposium on Advances in Robot Kinematics*, Grasse, France, 2016.
13. I.T. Jolliffe, *Principal Component Analysis*, Springer, 2002.
14. C. Happ, S. Greven, "Multivariate functional principal component analysis for data observed on different (dimensional) domains", *Journal of the American Statistical Association*, doi: 10.1080/01621459.2016.1273115, 2017.