Synthesis of Diverse Motions between Two Different Motion FPC Latent Spaces

Soya Shimizu[1], Ko Ayusawa[2,4], and Gentiane Venture[3,4]

***Abstract*—This paper introduces an innovative method for generating humanoid robot motion data using Functional Principal Component Analysis (FPCA). With increasing demand for robots to replace humans, there is a growing interest in humanoid robots. However, controlling complex humanoid robots, especially bipedal robots, remains challenging. Acquiring motion data for these robots is time-consuming and costly due to their high-dimensional complexity. Therefore, this paper proposes a solution for synthesis considering multiple latent spaces. By considering the intersection of two different spaces, synthesis while preserving the characteristics of the original motion is achieved. These innovations aim to establish a theoretical foundation for more efficient generation of motion data for humanoid robots.**

I. Introduction

Human motion presents rich exploration opportunities, especially in robotics. While humanoid robots aim for human-like appearances, many lack bipedal legs, relying on wheels instead. They also assist in evaluating exoskeletons for labor-intensive professions [1], [2]. Ongoing research explores deploying humanoid robots directly in hazardous or physically demanding tasks, marking a significant shift in robotics [3], [4].

Controlling humanoid robots is challenging due to factors like their height, freedom of movement, and balance issues. Various methods simplify their motion control, such as Choreonoid software and motion retargeting from human movements [5]–[11]. Direct Optimal Control (DOC) [12]–[16] and Inverse Optimal Control (IOC) help generate human-like movements [17]–[26], but current methods are labor-intensive and time-consuming. Streamlining these processes is necessary for future advancements in humanoid robot motion control.

To address these challenges, we introduced a sophisticated data synthesis methodology leveraging Functional Principal Component Analysis (FPCA) [27], [28]. FPCA condenses complex datasets into a lower-dimensional latent space, preserving fundamental motion characteristics for motion similarity assessment and efficient data synthesis. Other data reduction methods like autoencoders and reinforcement learning also exist, but they require learning processes and risk over-learning, unlike FPCA and PCA, which quickly reduce dimensionality without supervision.

In this study, FPC scores from the FPC space are treated as a single dataset representing motion characteristics, allowing for motion synthesis operations. While humanoid robots have previously performed motions using synthesized data [29], this thesis further explores the potential of motion synthesis using FPC scores from various perspectives.

[1]S. Shimizu is with Department of Mechanical Systems Engineering, Graduate School of Engineering, Tokyo University of Agriculture and Technology, Japan.

[2]K. Ayusawa is with the Human Augmentation Research Center, National Institute of Advanced Industrial Science and Technology (AIST), Japan. k.ayusawa@aist.go.jp

[3]G. Venture is with Department of Mechanical Engineering, The University of Tokyo, Japan. venture@g.ecc.u-tokyo.ac.jp

[4]K. Ayusawa and G. Venture are with CNRS-AIST JRL (Joint Robotics Laboratory), UMI3218/RL, Japan.

II. Motion Feature Extraction Using FPCA

In this chapter, we explain the method of transforming motion data $\boldsymbol{Q}$ into values in a low-dimensional space using FPCA. The motion data $\boldsymbol{Q}$ consists of trajectories of several variables, including joint angles $\boldsymbol{q}$, angular velocities $\dot{\boldsymbol{q}}$, and angular accelerations $\ddot{\boldsymbol{q}}$ of the robot.

A. Variables and B-Spline Functions for Motion Feature Extraction

A combination of B-spline functional bases is utilized to model the trajectory of joint angles, enhancing robustness against noise and reducing dimensionality. The functionally represented data is then utilized in FPCA as discussed in Section II-B. The variables $\boldsymbol{q}$, $\dot{\boldsymbol{q}}$, and $\ddot{\boldsymbol{q}}$ are represented as follows:

$$\begin{bmatrix} \boldsymbol{q} \\ \dot{\boldsymbol{q}} \\ \ddot{\boldsymbol{q}} \end{bmatrix} = \sum_{t=1}^{N_T} \sum_{i=1}^{N_B} \begin{bmatrix} b_{i,t} \\ \dot{b}_{i,t} \\ \ddot{b}_{i,t} \end{bmatrix} \cdot \sum_{i=1}^{N_B} \boldsymbol{w}_{qi} \Leftrightarrow \begin{bmatrix} \boldsymbol{q} \\ \dot{\boldsymbol{q}} \\ \ddot{\boldsymbol{q}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}_q \\ \dot{\boldsymbol{B}}_q \\ \ddot{\boldsymbol{B}}_q \end{bmatrix} \boldsymbol{W}_q \tag{1}$$

where, the weight vector $\boldsymbol{w}_{qi} \in \mathbb{R}^{N_J}$ represents the weights, while $b_{i,t}, \dot{b}_{i,t}$, and $\ddot{b}_{i,t} \in \mathbb{R}$ are B-spline basis functions, with N_B denoting their number. Matrices $\boldsymbol{B}_q$, $\dot{\boldsymbol{B}}_q$, and $\ddot{\boldsymbol{B}}_q \in \mathbb{R}^{N_T \times N_B}$ comprise these basis functions, with $\dot{\boldsymbol{B}}_q$ / $\ddot{\boldsymbol{B}}_q$ representing their first / second-order derivatives. The weight matrix $\boldsymbol{W}_q \in \mathbb{R}^{N_B \times N_J}$ captures the spline functions' weights. Equation (1) parameterizes trajectories $\boldsymbol{q}, \dot{\boldsymbol{q}}$, and $\ddot{\boldsymbol{q}}$ using coefficient parameters $\boldsymbol{W}_q$, effectively consolidating them into $\boldsymbol{W}_q$, leading to significant data reduction ($\boldsymbol{Q} \rightarrow \boldsymbol{W}_q$).

In the following subsection, we will further elaborate on methods to reduce the dimensionality of $\boldsymbol{W}_q$.

B. Conversion of Dataset into Low-Dimensional Latent Space Scores

Before delving into the main topic, it's crucial to discuss the vector $\boldsymbol{w} \in \mathbb{R}^{N_J \times N_B}$, obtained by concatenating the elements of the parameter matrices $\boldsymbol{W}_q$ into a single column. This discussion holds practical significance beyond theoretical

TABLE I
RMSE BETWEEN ORIGINAL ARM JOINT ANGLE TRAJECTORIES AND THOSE SYNTHESIZED USING DIFFERENT METHODS FOR COMBINING LEFT AND RIGHT ARM MOTIONS.

Combination		Synthesis method			
		Intersection [rad]		Linear [rad]	
Right	Left	Right	Left	Right	Left
Swing	Swing	0.2974	0.3947	0.2482	0.3789
Swing	Stretch	0.3101	0.3869	0.3450	0.3943
Stretch	Swing	0.4008	0.3697	0.4036	0.4099
Stretch	Stretch	0.3838	0.3782	0.3920	0.3006

abstraction, as $\boldsymbol{w}$ encapsulates the essence of motion data $\boldsymbol{Q}$ and forms the basis for applying the FPCA methodology.

Upon applying FPCA to the motion dataset, we obtain FPC scores. Given a motion dataset comprising k motion data $\boldsymbol{Q}_i(1 \leq i \leq k)$ and their corresponding $\boldsymbol{w}_i$ values, applying FPCA to the dataset $\boldsymbol{w}_{data} = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k]$ constructs the FPC space composed of FPC scores $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k]$. Each FPC score $\boldsymbol{x}_i$ can be expressed as follows:

$$\boldsymbol{x}_i = \boldsymbol{M}(\boldsymbol{w}_i - \overline{\boldsymbol{w}}) \Leftrightarrow \boldsymbol{w}_i = \boldsymbol{M}^{-1}\boldsymbol{x}_i + \overline{\boldsymbol{w}} \tag{2}$$

where, $\boldsymbol{M} \in \mathbb{R}^{N_{origin} \times N_{origin}}(N_{origin} = N_J \times N_B)$ is the conversion matrix calculated from FPCA, and $\overline{\boldsymbol{w}}$ refers to the mean value of the weights $\boldsymbol{w}_i$ depending on the given dataset. Based on (2), generating the motion data $\boldsymbol{Q}_i$ from FPC scores can be readily achieved from the FPC score $\boldsymbol{x}_i$. These scores contain the characteristics of the original data, and by appropriately combining them, it becomes possible to synthesize data while preserving the features of the original data [30].

III. Synthesis of Diverse Motions between Two Different Motion Latent Spaces

The FPC space allows blending motions via local and global spaces, where local spaces represent specific body parts' motions, and global spaces encompass the entire body's motion, influencing blending feasibility. Interpolating motions within the same local space yields blended motions, such as combining left and right lunge motions to generate new squat motions [31]. However, blending motions from different local spaces necessitates considering the global space. The concept of the "intersection point" bridges these spaces.

To comprehend the mathematical interpretation of motion synthesis with local and global spaces, consider the following. First, FPC scores in both local and global spaces are defined as follows:

$$\boldsymbol{M}_{\mathrm{G}}(\boldsymbol{w} - \overline{\boldsymbol{w}}_{\mathrm{G}}) = \boldsymbol{x}_{\mathrm{G}} \tag{3}$$

$$\begin{cases} \boldsymbol{w} = \boldsymbol{M}_{\mathrm{L1}}^{-1}\boldsymbol{x}_{\mathrm{L1}} + \overline{\boldsymbol{w}}_{\mathrm{L1}} \\ \boldsymbol{w} = \boldsymbol{M}_{\mathrm{L2}}^{-1}\boldsymbol{x}_{\mathrm{L2}} + \overline{\boldsymbol{w}}_{\mathrm{L2}} \end{cases} \tag{4}$$

Where, the subscript $_{\mathrm{G}}$ is utilized for global space values: the subscript $_{\mathrm{L1}}$ is utilized for the 1st motion data group, and $_{\mathrm{L2}}$ is utilized for the 2nd motion data group.

The relationship between the FPC scores in the three spaces illuminates the motion synthesis process: two local spaces and the global space. Local values such as $\boldsymbol{w}_{\mathrm{L1}}$, $\boldsymbol{w}_{\mathrm{L2}}$, $\boldsymbol{M}_{\mathrm{L1}}$, and $\boldsymbol{M}_{\mathrm{L2}}$ are transformed into global ones using (3) and (4) through the following formula:

$$\begin{cases} \boldsymbol{x}_{\mathrm{G}} = \boldsymbol{S}_{\mathrm{G/L1}}\boldsymbol{x}_{\mathrm{L1}} + \overline{\boldsymbol{x}}_{\mathrm{G/L1}} \\ \boldsymbol{x}_{\mathrm{G}} = \boldsymbol{S}_{\mathrm{G/L2}}\boldsymbol{x}_{\mathrm{L2}} + \overline{\boldsymbol{x}}_{\mathrm{G/L2}} \end{cases} \tag{5}$$

where,

$$\begin{cases} \boldsymbol{S}_{\mathrm{G/L1}} \triangleq \boldsymbol{M}_{\mathrm{G}}\boldsymbol{M}_{\mathrm{L1}}^{-1} \\ \boldsymbol{S}_{\mathrm{G/L2}} \triangleq \boldsymbol{M}_{\mathrm{G}}\boldsymbol{M}_{\mathrm{L2}}^{-1} \end{cases} \tag{6}$$

$$\begin{cases} \overline{\boldsymbol{x}}_{\mathrm{G/L1}} \triangleq \boldsymbol{M}_{\mathrm{G}}(\overline{\boldsymbol{w}}_{\mathrm{L1}} - \overline{\boldsymbol{w}}_{\mathrm{G}}) \\ \overline{\boldsymbol{x}}_{\mathrm{G/L2}} \triangleq \boldsymbol{M}_{\mathrm{G}}(\overline{\boldsymbol{w}}_{\mathrm{L2}} - \overline{\boldsymbol{w}}_{\mathrm{G}}) \end{cases} \tag{7}$$

Using these values enables to convert each local FPC scores into the corresponding global FPC score: $\boldsymbol{x}_{\mathrm{G}}^{(*)}$.

The solution to the simultaneous (5) yields the synthesized data, considering the original motion characteristics, represented by the intersection point $\boldsymbol{x}_{\mathrm{G}}$. Ultimately, the motion data $\boldsymbol{Q}$ is derived from $\boldsymbol{x}_{\mathrm{G}}^{(*)}$ using (1) and (2).

IV. Synthesizing Upper Body Motions in FPC Space

The proposed method has already synthesized motions involving both upper and lower body movements [30]. The study focused on synthesizing upper body motions by combining two types of prepared data: one involving swinging the arms forward (Swing) and the other involving extending the arms sideways from the body (Stretch). These movements allowed independent operation of the left and right arms. Joint angle data from both arms were used as the composite target, while data from only one arm were extracted for FPCA, resulting in four types of local spaces. Motion synthesis was performed by combining two of these local spaces. The study compared this method with linear blending and evaluated the preservation of original motion characteristics using root mean square error (RMSE) between original and synthesized motions.

The results are presented in Table I. For combinations such as Swing + Stretch, where different motions are performed by the left and right arms, using Intersection for synthesis tends to better preserve the original data's characteristics compared to linear blending, as indicated by the RMSE values. Conversely, for combinations like Swing + Swing, where both arms perform the same motion, linear blending shows lower RMSE values.

V. Conclusion

This paper discussed the synthesis of humanoid robot motion data using Functional Principal Component Analysis (FPCA) and the FPC space. The methodology involves considering multiple FPC spaces and finding their intersections. While the proposed method tends to better preserve the original motion's characteristics compared to linear blending in some combinations, this is not universally observed. To improve the synthesis process, further investigation into the reasons for these discrepancies is necessary. Future discussions will explore the possibility of calculating intersections of three or more spaces and creating finer-grained local spaces by segmenting body parts into smaller units.

References

[1] T. Ito, K. Ayusawa, E. Yoshida, and H. Kobayashi, "Simultaneous control framework for humanoid tracking human movement with interacting wearable assistive device," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 3604–3611, 2020.

[2] M. Dežman, S. Massardi, D. Pinto-Fernandez, V. Grosu, C. Rodriguez-Guerrero, J. Babič, and D. Torricelli, "A mechatronic leg replica to benchmark human–exoskeleton physical interactions," *Bioinspiration & Biomimetics*, vol. 18, no. 3, p. 036009, 2023.

[3] A. Adu-Bredu, G. Gibson, and J. W. Grizzle, "Exploring kinodynamic fabrics for reactive whole-body control of underactuated humanoid robots," *arXiv preprint arXiv:2303.04279*, 2023.

[4] P. Ferrari, L. Rossini, F. Ruscelli, A. Laurenzi, G. Oriolo, N. G. Tsagarakis, and E. Mingo Hoffman, "Multi-contact planning and control for humanoid robots: Design and validation of a complete framework," *Robotics and Autonomous Systems*, vol. 166, p. 104448, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0921889023000878

[5] S. Nakaoka, "Choreonoid: Extensible virtual robot environment built on an integrated gui framework," in *2012 IEEE/SICE International Symposium on System Integration (SII)*, 2012, pp. 79–85.

[6] H. Jeong, I. Lee, J. Oh, K. K. Lee, and J.-H. Oh, "A robust walking controller based on online optimization of ankle, hip, and stepping strategies," *IEEE Transactions on Robotics*, vol. 35, no. 6, pp. 1367–1386, 2019.

[7] T. Inamura and Y. Mizuchi, "Sigverse: A cloud-based vr platform for research on multimodal human-robot interaction," *Frontiers in Robotics and AI*, vol. 8, 2021. [Online]. Available: https://www.frontiersin.org/articles/10.3389/frobt.2021.549360

[8] K. Ayusawa and E. Yoshida, "Motion retargeting for humanoid robots based on simultaneous morphing parameter identification and motion optimization," *IEEE Transactions on Robotics*, vol. 33, no. 6, pp. 1343–1357, 2017.

[9] P. M. Viceconte, R. Camoriano, G. Romualdi, D. Ferigo, S. Dafarra, S. Traversaro, G. Oriolo, L. Rosasco, and D. Pucci, "Adherent: Learning human-like trajectory generators for whole-body control of humanoid robots," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 2779–2786, 2022.

[10] Q. Rouxel, K. Yuan, R. Wen, and Z. Li, "Multicontact motion retargeting using whole-body optimization of full kinematics and sequential force equilibrium," *IEEE/ASME Transactions on Mechatronics*, vol. 27, no. 5, pp. 4188–4198, 2022.

[11] D. Lim, D. Kim, and J. Park, "Online telemanipulation framework on humanoid for both manipulation and imitation," in *2022 19th International Conference on Ubiquitous Robots (UR)*, 2022, pp. 8–15.

[12] H. Bock and K. Plitt, "A multiple shooting algorithm for direct solution of optimal control problems," *IFAC Proceedings Volumes*, vol. 17, no. 2, pp. 1603–1608, 1984, 9th IFAC World Congress: A Bridge Between Control Science and Technology, Budapest, Hungary, 2-6 July 1984. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1474667017612059

[13] M. Posa, S. Kuindersma, and R. Tedrake, "Optimization and stabilization of trajectories for constrained dynamical systems," in *2016 IEEE Int. Conf. on Robotics and Automation*. IEEE, 2016, pp. 1366–1373.

[14] I. Maroger, O. Stasse, and B. Watier, "Walking human trajectory models and their application to humanoid robot locomotion," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 3465–3472.

[15] I. Maroger, N. Ramuzat, O. Stasse, and B. Watier, "Human trajectory prediction model and its coupling with a walking pattern generator of a humanoid robot," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 6361–6369, 2021.

[16] P. M. Wensing, M. Posa, Y. Hu, A. Escande, N. Mansard, and A. D. Prete, "Optimization-based control for dynamic legged robots," *IEEE Transactions on Robotics*, vol. 40, pp. 43–63, 2024.

[17] K. Mombaur, A. Truong, and J.-P. Laumond, "From human to humanoid locomotion—an inverse optimal control approach," *Autonomous robots*, vol. 28, pp. 369–383, 2010.

[18] K. Mombaur, J.-P. Laumond, and A. Truong, "An inverse optimal control approach to human motion modeling," in *Robotics Research: The 14th International Symposium ISRR*. Springer, 2011, pp. 451–468.

[19] I. Maroger, N. Ramuzat, O. Stasse, and B. Watier, "Human trajectory prediction model and its coupling with a walking pattern generator of a humanoid robot," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 6361–6369, 2021.

[20] N. Tsiantis, E. Balsa-Canto, and J. R. Banga, "Optimality and identification of dynamic models in systems biology: an inverse optimal control framework," *Bioinformatics*, vol. 34, no. 14, pp. 2433–2440, 03 2018. [Online]. Available: https://doi.org/10.1093/bioinformatics/bty139

[21] J. Colombel, D. Daney, and F. Charpillet, "On the reliability of inverse optimal control," in *2022 International Conference on Robotics and Automation (ICRA)*, 2022, pp. 8504–8510.

[22] D. Straub, M. Schultheis, H. Koeppl, and C. A. Rothkopf, "Probabilistic inverse optimal control with local linearization for non-linear partially observable systems," *arXiv preprint arXiv:2303.16698*, 2023.

[23] S. Albrecht, K. Ramírez-Amaro, F. Ruiz-Ugalde, D. Weikersdorfer, M. Leibold, M. Ulbrich, and M. Beetz, "Imitating human reaching motions using physically inspired optimization principles," in *2011 11th IEEE-RAS International Conference on Humanoid Robots*, 2011, pp. 602–607.

[24] J. Mainprice, R. Hayne, and D. Berenson, "Goal set inverse optimal control and iterative replanning for predicting human reaching motions in shared workspaces," *IEEE Transactions on Robotics*, vol. 32, no. 4, pp. 897–908, 2016.

[25] P. Englert, N. A. Vien, and M. Toussaint, "Inverse kkt: Learning cost functions of manipulation tasks from demonstrations," *The International Journal of Robotics Research*, vol. 36, no. 13-14, pp. 1474–1488, 2017. [Online]. Available: https://doi.org/10.1177/0278364917745980

[26] G. Gulletta, W. Erlhagen, and E. Bicho, "Human-like arm motion generation: A review," *Robotics*, vol. 9, no. 4, 2020. [Online]. Available: https://www.mdpi.com/2218-6581/9/4/102

[27] P. Besse and J. O. Ramsay, "Principal components analysis of sampled functions," *Psychometrika*, vol. 51, no. 2, pp. 285–311, 1986.

[28] G. M. James, T. J. Hastie, and C. A. Sugar, "Principal component models for sparse functional data," *Biometrika*, vol. 87, no. 3, pp. 587–602, 2000.

[29] S. Morishima, K. Ayusawa, E. Yoshida, and G. Venture, "Converting constrained whole-body human motions to humanoid using smoothing," *The Abstracts of the international conference on advanced mechatronics : toward evolutionary fusion of IT and mechatronics : ICAM*, vol. 2015.6, p. 314, 2015.

[30] S. Shimizu, K. Ayusawa, E. Yoshida, and G. Venture, "Whole-body motion blending under physical constraints using functional pca," in *2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)*, 2018, pp. 280–283.

[31] S. Morishima, K. Ayusawa, E. Yoshida, and G. Venture, "Whole-body motion retargeting using constrained smoothing and functional principle component analysis," in *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*. IEEE, 2016, pp. 294–299.