Towards the Analysis of Movement Variability in Human-Humanoid Imitation Activities

Leave Authors Anonymous
for Submission
City, Country
e-mail address

Leave Authors Anonymous
for Submission
City, Country
e-mail address

ABSTRACT

In this paper, we present preliminary results for the analysis of movement variability in order to quantify face-to-face human-humanoid imitation activities. We applied the state space reconstruction's theorem to test our hypothesis where participants, even performing the same arm movement, presented minor difference in the way they moved. With this in mind, we asked eighteen participants to copy NAO's arm movements while we collected data from inertial sensors attached to the participants' wrist and estimated the head pose using the OpenFace framework. With the proposed metric, we found that two participants were moving their arms asymmetrically while others move their arms symmetrically. We also showed that two participants where moving their head even when NAO's head was static. Although the work is in its early stage, the results are promising for applications in rehabilitation, sport science, entertainment or education.

ACM Classification Keywords

I.2.9. Robotics: Sensors; G.3. PROBABILITY AND STATISTICS: Time series analysis

Author Keywords

Human-Robot Interaction; Human-Humanoid Imitation; Wearable Inertial Sensors; State Space Reconstruction; Nonlinear dynamics; Dynamics Invariants

INTRODUCTION

Movement variability is an inherent feature within a person and between persons movements [7]. Recently, Herzfeld et al. [5] conducted experiments to state that movement variability is not only noise but a source of movement exploration which at certain point of the exploration such variability is becoming a source of movement exploration. With this in mind, we have found that there is little research in the area of human-robot interaction that is focused on the quantification of movement variability.

HAI'17, October 17–20, 2016, Bielefeld, Germany

ISBN 123-4567-24-567/08/06...$15.00

DOI: http://dx.doi.org/10.475/123_4

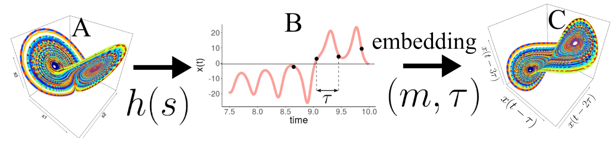

Figure 1. State Space Reconstruction. A. $M-$**dimensional state space** $s(t)$**; B.** $1-$**dimensional measurement time-series** $x(t)$**; and C.** $N-$**dimensional reconstructed state space** $v(t)$ **where** $M \geq N$ **(adapted from [9]).**

The paper is divided into an intuitive explanation of the state space reconstruction, experiment design, results and conclusions.

METHOD

State Space Reconstruction's Theorem

The purpose of the State Space Reconstruction's Theorem, also known as time-delay embedding, is to reconstruct the topological properties of an unknown $M-$dimensional state space $s(t)$ from a $1-$dimensional measurement $x(t)$ in order to reconstruct an $N-$dimensional embedding space (Figure 1). The time-delay embedding assumes that the time-series is a sequence $x(t) = h(s(t))$, where $h : S \rightarrow \mathbb{R}^M$ is a measurement function on the unknown dynamical system, being $x(t)$ observable. Thus, the time delay reconstruction in m dimensions with a time delay τ is defined as: $\bar{x}(t) = (x(t), x(t-\tau), ..., x(t-(m-1)\tau))$. Then a further transformation is considered, e.g. PCA, in order to reduce the dimensionality of the $m-$dimensional reconstructed state space to a $k-$dimensional space [11]. For this work, we assume that the signal, $x(t)$, is produced by some time-varying system in our case the time-series are produced by the linear acceleration of the inertial sensors. We assume that the time-series, over some time period, exhibit a systematic variability between and within persons movements. What we do not know is how reliable the quantification methods for movement variability are and how the levels of imitation of a given range of movement variability can be established.

Determining the embedding parameters (m and τ)

Although State Space Reconstruction's Theorem has been used extensively in gait recognition and walking, running and cycling activities [4, 10], the computation of the minimal embedding parameters largely depend on the structure of the time-series (amplitude, frequency, nonlinearity). To which, we however first compute the minimal embedding parameters using the Cao's algorithm [3] and the mutual information and then we manually increase the dimensionality of the reconstructed state space until the attractor is untangled.

EXPERIMENT DESIGN

Measuring Movement

To understand the movement variability of the participants, we use four Wearable Inertial Sensors SEN-10736 SparkFun 9DOF RAZOR with triple-axis accelerometer and triple-axis gyroscope (Figure 3A). The data were transmitted via RN42 bluetooth module for which we set a sampling rate of 50 Hz and collected using ROS [8].

Time-series from the Accelerometer Sensor

The sequences $(a_x(n), a_y(n), a_z(n))$ are the raw data collected from the triaxial accelerometer sensor. Then, for instance, the time-series $a_x(n)$ with a length of N samples is employed to get the embedded state space matrix, Ea_x, with m rows and $N-(m-1)\tau$ columns. PCA is then applied to reduce de dimensionality of the data choosing the first two components of the rotated data in order to reconstruct the state space.

Head Pose Estimation

Estimating head pose in human-robot interactions is an active area of research where challenges like real-time tracking, the use of less invasive equipment or the long-time preparation of calibration techniques of the motion capture systems remain to be solved. Recently, Lemaignan et al. [6] proposed a head pose estimator using a monocular RGB webcam which is able to track a head with rotations up to $\pm 40°$ horizontally and $\pm 30°$ vertically. However, OpenFace, a fully open source real-time facial behaviour analysis, non only provides head pose (orientation and motion) but also a state-of-the-art performance in facial landmark motion, facial expressions, and eye gaze [1]. We therefore select the OpenFace because of the simple set up, less invasive and the features for face behaviour, besides it can operate with a simple webcam in real-time (Figure 3B).

EXPERIMENT

Hypothesis

In our previous experiments of a face-to-face human-humanoid imitation activity [12], we applied the State Space Reconstruction's Theorem to quantify the level of imitation for horizontal and vertical upper arm movements. In such experiment, we observed in the recorded videos that the effects like boredom, fatigue or level of engagement play an important role in the influence that each persons moves. With this in mind, we hypothesised that not only the inertial sensors attached to the body can provide information about movement variability but also the face expressions and head pose estimation which, we believe, will lead us to get better understanding of the movement variability in human-to-humanoid activities and therefore create more reliable metrics to quantify such variability.

Participants and Procedure

To test our hypothesis, we collected data from eighteen healthy participants: eight male participant (age 18 ± 3.43) and ten female (age 18 ± 0.43) in which inertial sensors were attached to the wrist of both the participant and the humanoid robot, and put the webcam in front of the participant for the head pose estimation (Figure 3A). In the experiment, participant were asked to imitate NAO's upper arm movements as shown in Figure 3.

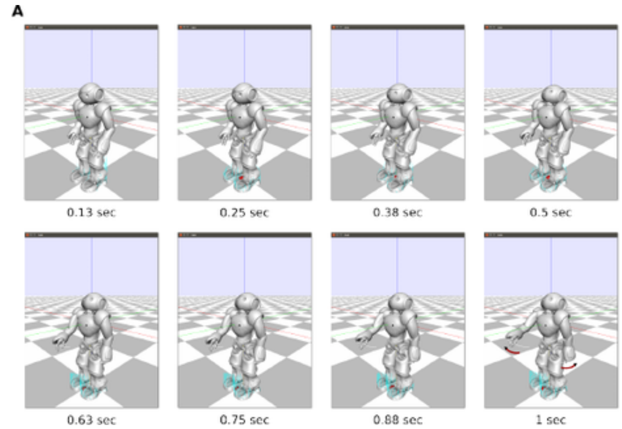

Figure 2. NAO's arm movements

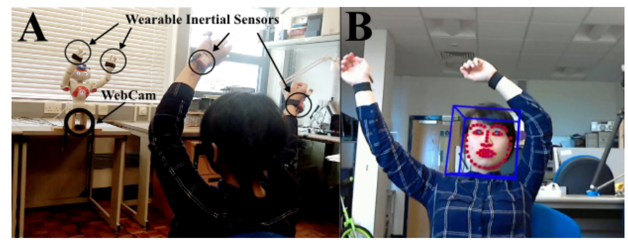

Figure 3. A. Experimental set-up: face-to-face imitation with NAO humanoid robot; B. Head pose estimation with OpenFace [1]

RESULTS

In Figure 4A, we only presented time-series for the participant 12 for the sensors attached to the right and left wrist (red and blue). With this in mind, we showed that the asymmetry of the signals which is reflected in p12-sLeft and p12-sRight in (Figure 4B) as well as in the error plot (Figure 4C). For the error bars of the OpenFace, you can notice the well distributed data in the interquartile range for participants p04 and p17 (Figure 4C) which means that those participants were moving their heads as their hands. The expanded distribution of the interquartile range for participant p14 means that p14 head's movement was quite static.

CONCLUSION

Having proposed the use of the state space reconstruction to analyse human-humanoid imitation activity, several questions remain to be investigated such as the understanding of variability of emotions and motions in one-to-one or one-to-many human-humanoid interactions. Considering the fact that the robot's head was static in the activity, we observed that the interquartile range of the proposed metric for participants p04 and p17 is an indication of the movement of their heads as a tendency for the arm movements. We believe that such

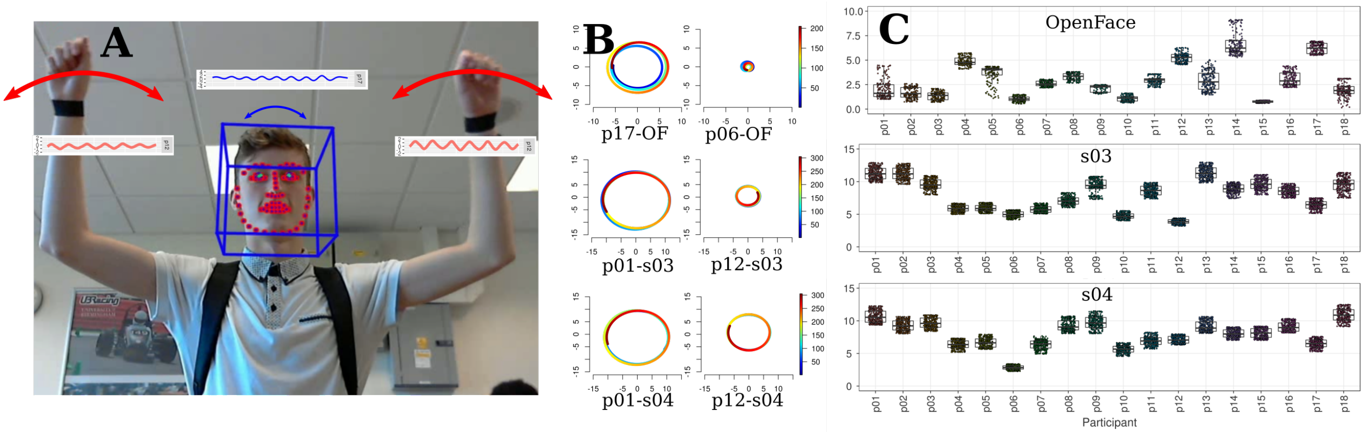

Figure 4. A. time-series for the inertial sensors $a_x(n)$ and the Head Pose Estimation in the T_x axis; B. State Space Reconstruction with $m = 100$ and $\tau = 4$ for participants 17 and 16 for the head pose estimation and participants 01 and 12 for sensor 03 and 04; C. Error bars for the head pose estimation, sensor 03 and sensor 04 for the eighteen participants.

behaviour requires further investigation for which the motor experience affects the visual sensitivity of human action [2].

In future experiments, there are four areas that we intend to investigate: (a) performance of experiments of not only one-to-one interaction but two-humans to one-humanoid and three-humans to one-humanoid interactions; (b) exploration of complex movements which can be performed by both persons and NAO; (c) data collection from a wider range of individuals (different gender, age and state of health) and from additional inertial sensors attached to the body; and (d) application of deep learning techniques to automatically classify the movement variability.

ACKNOWLEDGMENTS

XXX gratefully acknowledges XXX for funding his doctoral studies at University of XXX.

REFERENCES

1. Tadas Baltrušaitis, Peter Robinson, and Louis-Philippe Morency. 2016. OpenFace: an open source facial behavior analysis toolkit. In *IEEE Winter Conference on Applications of Computer Vision*.
2. R. Blake and M. Shiffrar. 2007. Perception of Human Motion. In *Annual Review of Psychology*, Vol. 58. Annual Reviews, 47–73.
3. Liangyue Cao. 1997. Practical method for determining the minimum embedding dimension of a scalar time series. *Physica D: Nonlinear Phenomena* 110 (1997), 43–50.
4. Jordan Frank, Shie Mannor, and Doina Precup. 2010. Activity and Gait Recognition with Time-Delay Embeddings Time-Delay Embeddings. *AAAI Conference on Artificial Intelligence* (2010), 407–408.
5. David J Herzfeld and Reza Shadmehr. 2014. Motor variability is not noise, but grist for the learning mill. *Nat Neurosci* 17, 1 (2014), 149–150.
6. S. Lemaignan, F. Garcia, A. Jacq, and P. Dillenbourg. 2016. From real-time attention assessment to "with-me-ness" in human-robot interaction. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 157–164.
7. K.M. Newell and D.M. Corcos. 1993. *Variability and Motor Control*. Human Kinetics Publishers.
8. Morgan Quigley, Ken Conley, Brian P. Gerkey, Josh Faust, Tully Foote, Jeremy Leibs, Rob Wheeler, and Andrew Y. Ng. 2009. ROS: an open-source Robot Operating System. In *ICRA Workshop on Open Source Software*.
9. Juan-Carlos Quintana-Duque. 2012. Non-linear Dynamic Invariants Based on Embedding Reconstruction of Systems for Pedaling Motion. In *9. Symposium der dvs-Sektion Sportinformatik, Univ. Konstanz.*
10. Albert Samà, Francisco J. Ruiz, Núria Agell, Carlos Pérez-López, Andreu Català, and Joan Cabestany. 2013. Gait identification by means of box approximation geometry of reconstructed attractors in latent space. *Neurocomputing* 121 (2013), 79–88.
11. L. C. Uzal, G. L. Grinblat, and P. F. Verdes. 2011. Optimal reconstruction of dynamical systems: A noise amplification approach. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 84, 1 (2011).
12. Author X, Author XX, and Author XXX. 2017. The X Title. In *Booktitle X*.