Harnessing Machine Learning for interpersonal physical alignment

Roi Yozevitch
dept. of Electrical Engineering
Ariel University
Ariel, Israel

Hila Gvirts
dept. of Behavioral Sciences
Ariel University
Ariel, Israel

Ornit Apelboim
dept. of Behavioral Science
Ariel University
Ariel, Israel

Elhanan Mishraky
dept. of Computer Science
Ariel University
Ramat Gan, Israel

***Abstract*—This work presents a novel way to determine interpersonal physical synchrony state by inspecting hands' postures obtained from a unique $3D$ depth camera device named Leap-Motion Controller. Several ML methods are utilized such as SVM, shallow feed-forward ANN and XGBoot. We show that even a simple ANN can outperform XgBoost in simple classification tasks.**

I. Introduction

Machine learning techniques have shown great results in classification/prediction problems in numerous field. In this work, we harness those tools in the field of social neuroscience in general, and the sub-filed of interpersonal synchrony in particular. Assuming two people are synchronized, they also tend to align their physicality. In the context of this paper, we wish to measure this physical alignment by inspecting their hand gestures and measuring their similarity. As opposed to previous works which examined physical alignment via very specific (not natural) tasks [1], hand gestures are much more ecological and are much more understood in this synchrony context.

Hand gesture recognition [2] is attracting a growing interest due to its applications in many different fields, such as human-computer interaction, robotics, computer gaming, automatic sign-language interpretation and so on. The the Leap Motion controller [3] device has opened new opportunities for gesture recognition. Differently from the Microsoft Kinect [4], the Leap-Motion is explicitly targeted to hand gesture recognition and directly computes the position of the fingertips and the hand orientation. While to amount of information is limited compared to other depth cameras (e.g., Kinect), the extracted data is more accurate system with sub-millimeter accuracy. In contrast to standard multi-touch solutions, this above-surface sensor is discussed for use in realistic stereo 3D interaction systems [5]. The following figure shows both the controller itself and the $3D$ output of two hands simultaneously.

This work is part of a larger research that aim to to uncover the neural basis of alignment of two or more interacting individuals, known as 'interpersonal motor synchrony' [6].

Three different alignment states are defined:

In the first state, each participant, in her turn, is asked to move his hand freely over the Leap-Motion controller. We denote this as an "**Alone**" mode.

Fig. 1. The Leap Motion controller device

The second state is very similarly to the first. The only difference, however, is that the two participants can see each others movement during the experiment. We denote this as an "**Autonomous**" state.

The last state is different. In the last state, the two participants are asked to moved in synchronization. They are invited to interpret synchronization as they understand. We denote this state as "**synchronization**" state.

This work's purpose is to build an neural network that can classify between the Alone and the Autonomous modes of behaviour. This work is restricted to memory-less models and Recurrent Neural Networks(RNN [7]) models are still being tested.

There is a famous quote by scholar Andrew Ng that any task a person can do (compute) in one second of thought, an ANN can also achieve [8]. Empirical observation has taught us that while it is relatively an easy task to distinguish between autonomous and synchronized behaviours, it is much more complex to distinguish between autonomous and alone movements. This is mainly due to the fact that a synchronized movement is, by definition, synchronized. On the other hand,

978-1-5386-6378-3/18/$31.00 ©2018 IEEE

When people are instructed to move freely (autonomous behaviour), their respective movements are not aligned in any sense. Nevertheless, research shows that ”spontaneous synchronization” occasionally emerges in those scenarios [9]. It is the nature of this behaviour we would like to look into.

A. Paper Structure

The rest of the paper is organized as follows: the next section is devoted to previous works conducted both with the Leap-Motion controller.

Section III details the Experiment scenarios and the results of different ML models. In section IV, some interesting results are discussed. Finally, section V both summarizes and draws a vector for future work.

II. Previous Works

One of the first works conducted on the Leap motion was published in 2013 [10]. The authors try to utilized the Leap Motion controller for recognizing Australian Sign Language (Auslan). Back then, the algorithm couldn't handle perpendicular motions and scenarios where individual elements of the hands are brought together. Moreover, the authors used ANN to train their system. Alas, the system only learned simple individual movements. A year later, [11] have tried to recognize the American sign language via a K-fold SVM.

A. Our Contribution

To the best of our knowledge, this work is the first one to harness the power of ANN in the sub-field of social neuro-alignment. Even contemporary works, when utilize machine learning models, mainly work with simple models such as Support Vector Machine (*SVM*). While those models produce occasionally great results, they are limit in their capabilities [7]. As it turns out, even a shallow ANN can outperform *SVM* in our context.

III. Experiments and Results

The Leap-Motion controller reports the exact posture in each hand at $100Hz$. The data can be extracted via the official SDK. A python logger was implemented. Among others, the controller reports the position and velocity vectors (X-Y-Z), the yaw-pitch-roll values and the level in which the palm is open or closed (a real number $\in [0, 1]$). Those are the features for our feed-forward ANN.

The experiment was conducted as follows: Two participants sat in two sides of a table facing each-other. They were told to move in three different alignment states. As explained above, these states are:

1) An Alone Mode.
2) An autonomous mode
3) An intentional mode.

A picture from the experiment can be seen in figure 2. Five different couples were recorded. Since each person behaves differently, a subject independent classifier is mandatory. Thus, the system was trained with 4 couples and tested with the fifth.

The question is: is it possible to distinguish between the autonomous and the Alone modes of alignments?

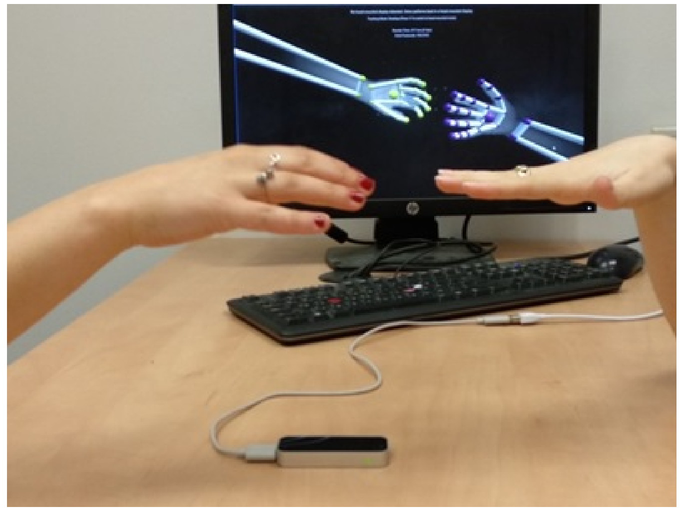

Fig. 2. A picture from the experiment. Two persons sit facing each other and move their hands freely over the Leap Motion controller.

A. Results

As explained above, three ML models have been tested. The first model is SVM (RBF kernel) since the lion share of the ML works in the field of social neuroscience utilize it as a classifier. The second model is Xgboost [12]. Recently, there is a growing interest in this classifier as a very efficient one. Finally, a shallow feed-forward ANN. The network has only two hidden layers and the number of trainable parameters is 373.

1) SVM: This model reached a total precision of 0.33 and a total recall of 0.58

2) XgBoost: This model reached a total precision of 0.74 and a recall of 0.69 and outperformed SVM.

3) ANN: As depicted in figure 3, the accuracy level obtained here was much higher and reached ≈ 0.94. This model outperformed the other models. While out of the scope of this work, one can speculate whether ANNs better describe neurophysiological mechanisms.

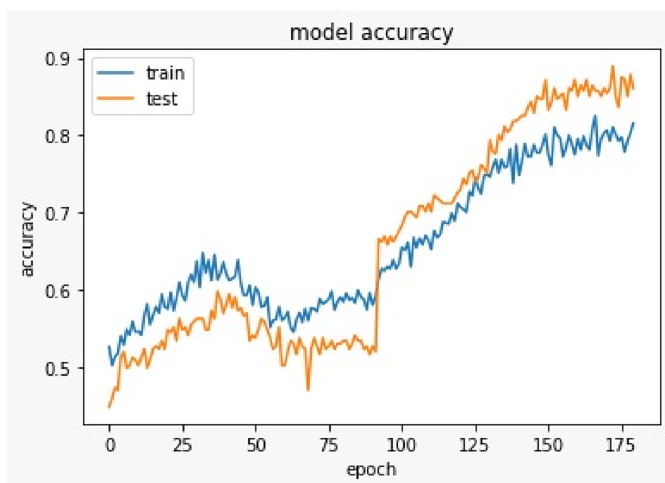

Fig. 3. Simple threshold for light sources detection

IV. Discussion

Figure 5 tells an interesting story. The X axis in the figure indicates the first participant's X position with respect to the controller (figure 4 and the Y axis indicates the second participant's X position. The lower the X value, the closer a person to a controller. While both the Alone and the spontaneous modes have no correlation, in the spontaneous mode, people tend move further from the controller and maintain a larger 'interpersonal distance'.

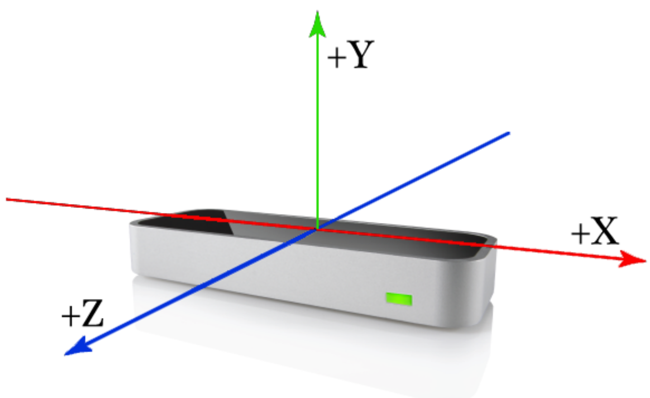

Fig. 4. The Leap Motion controller X-Y-X orientation.

Interestingly, theoretical approaches to model animal aggregations and collective behavior, suggest that individual actions are modelled as a combination of attraction or repulsion tendencies (e.g. collision avoidance [13]), alignment or behavior matching (e.g. velocity matching and synchronized motions; Reynolds [1987]). Notably, the concept of interpersonal space is related to both attraction and repulsion and the balance between the two determine the space between individuals. Recently, the authors of [14] conducted for the first time, an experimental study to examine whether the attraction rule govern spontaneous behaviour in humans. In accordance with the theoretical approaches for animal collective behavior, they found that individuals approached their neighbors spontaneously, if their neighbors' positions were visible. Our findings revel that humans, like animals, avoid collision (repulsion rule). Our findings are therefore complementary to those of [14], and together provide good initial empirical support for the possibility that the same rules that govern animals movement also govern humans.

While the correlation in both color is now high, one can see that the "ALONE" mode tends to be closer to the Leap-Motion's center (as can be seen in figure 5). In other words, this graph validates the correctness of the interpersonal distance' theory.

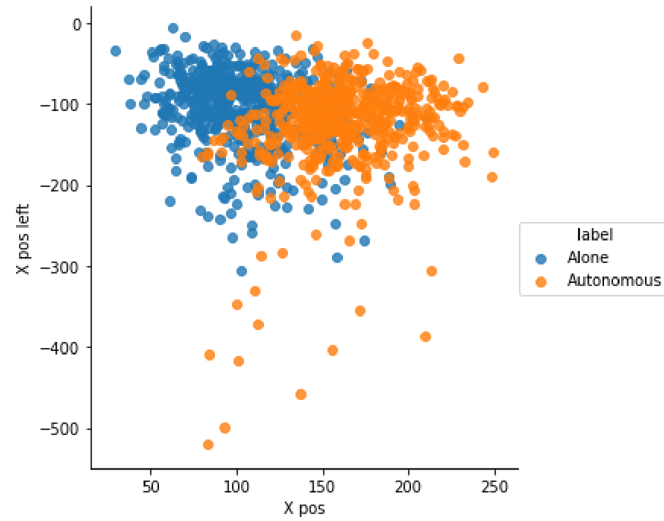

Fig. 5. X Position in two different mental states.

However, relying solely on the hand's X position yield much inferior results. on the other hand, since the model is robust (partly, due to the dropout regularization), omitting the X position value from the feature vector hardly changes the accuracy

V. Conclusions and Future Work

This work was an initial research towards learning the neurological patterns regrading emotional brain inter-person synchronization. Moreover, the very notion of synchrony cannot be reduced to a physical alignment with a simple similarity metrics (e.g., $L2$ norm). This is due to the fact that people are also considered in synchrony when they movements are exactly the opposite. In other words, if the moment A lifts his hand, B lower her hand and the moment A open his palm, B closes hers, we will label this behaviour as an intentional physical alignment that requires a more complex model. utilize a Recurrent-ANN (RNN) since the emotional state in time t is dependant of the emotional state at time $t-1$.

References

[1] Lior Noy, Erez Dekel, and Uri Alon. The mirror game as a paradigm for studying the dynamics of two people improvising motion together. *Proceedings of the National Academy of Sciences*, 108(52):20947–20952, 2011.

[2] Juan Pablo Wachs, Mathias Kölsch, Helman Stern, and Yael Edan. Vision-based hand-gesture applications. *Communications of the ACM*, 54(2):60–71, 2011.

[3] Leap Motion. Leap motion controller. *URl: https://www. leapmotion. com*, 2015.

[4] Zhengyou Zhang. Microsoft kinect sensor and its effect. *IEEE multimedia*, 19(2):4–10, 2012.

[5] Eugene Chng. New ways of accessing information spaces using 3d multitouch tables. In *Cyberworlds (CW), 2012 International Conference on*, pages 144–150. IEEE, 2012.

[6] Tanya L Chartrand and Jessica L Lakin. The antecedents and consequences of human behavioral mimicry. *Annual review of psychology*, 64:285–308, 2013.

[7] Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. *Deep learning*, volume 1. MIT press Cambridge, 2016.

[8] Andrew Ng. The state of artificial intelligence. https://www.youtube.com/watch?v=NKpuX$_y zdYs, December$2015.

[9] Jing Jiang, Bohan Dai, Danling Peng, Chaozhe Zhu, Li Liu, and Chunming Lu. Neural synchronization during face-to-face communication. *Journal of Neuroscience*, 32(45):16064–16069, 2012.

[10] Leigh Ellen Potter, Jake Araullo, and Lewis Carter. The leap motion controller: a view on sign language. In *Proceedings of the 25th Australian computer-human interaction conference: augmentation, application, innovation, collaboration*, pages 175–178. ACM, 2013.

[11] Ching-Hua Chuan, Eric Regina, and Caroline Guardino. American sign language recognition using leap motion sensor. In *Machine Learning and Applications (ICMLA), 2014 13th International Conference on*, pages 541–544. IEEE, 2014.

[12] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794. ACM, 2016.

[13] Iain D Couzin, Jens Krause, Richard James, Graeme D Ruxton, and Nigel R Franks. Collective memory and spatial sorting in animal groups. *Journal of theoretical biology*, 218(1):1–11, 2002.

[14] Michael Belz, Lennart W Pyritz, and Margarete Boos. Spontaneous flocking in human groups. *Behavioural processes*, 92:6–14, 2013.