Detecting an Offset-Adjusted Similarity Score based on Duchenne Smiles

Maximilian Henneberg
Chloe Eghtebas
Oliver De Candido
Technical University of Munich
Germany

Kai Kunze
Keio University
Japan

Jamie A Ward
Goldsmiths, University of London
UK

ABSTRACT

Detecting interpersonal synchrony in the wild through ubiquitous wearable sensing invites promising new social insights as well as the possibility of new interactions between humans-humans and humans-agents. We present the Offset-Adjusted SImilarity Score (OASIS), a real-time method of detecting similarity which we show working on visual detection of Duchenne smile between a pair of users. We conduct a user study survey (N = 27) to measure a user-based interoperability score on smile similarity and compare the user score with OASIS as well as the rolling window Pearson correlation and the Dynamic Time Warping (DTW) method. Ultimately, our results indicate that our algorithm has intrinsic qualities comparable to the user score and measures well to the statistical correlation methods. It takes the temporal offset between the input signals into account with the added benefit of being an algorithm which can be adapted to run in real-time will less computational intensity than traditional time series correlation methods.

CCS CONCEPTS

• **Human-centered computing** → *Social content sharing*; User studies; • **Mathematics of computing** → *Time series analysis*;

KEYWORDS

Interpersonal synchrony, Time Series Similarity

ACM Reference Format:
Maximilian Henneberg, Chloe Eghtebas, Oliver De Candido, Kai Kunze, and Jamie A Ward. 2023. Detecting an Offset-Adjusted Similarity Score based on Duchenne Smiles. In *CHI '23: ACM CHI Conference on Human Factors in Computing Systems, April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3544549.3585709

1 INTRODUCTION

Interpersonal synchrony is a measure of nonverbal coordination of gestures and physiological cues between two or more people over time [9]. Synchrony is continually investigated topic that is important metric that to looked into because it is a measurable social indicator that can enable many collaborative applications such as games [11] and social Virtual Reality (VR) [15] experiences. The ubiquity and number of sensors on wearable devices makes it promising to detect and study interpersonal synchrony in the wild of both the device wearer and those surrounding them [28]. By using body-worn cameras, like that of an Augmented Reality (AR) Head-Mounted Display (HMD), we can get information from people in the surrounding area, e.g., from body and facial signals, which can determine movement energy [26] and even extract heart rate [23] as valuable indicators of synchrony. Such social indicators can enable new applications, like tools to help people with autism pick up social cues [22], or to improve the collaboration between humans-humans or humans-robots [7] in industrial or business scenarios involving virtual assistants [25]. We built a similarity score based on facial data, specifically Duchenne smiles, to be used to study high level synchrony and to be incorporated into future applications that need such a social metric.

The nature of detecting synchrony in real-time in the wild is challenging. There exist many statistical tools to help detect coordination in time series [8]. Commonly used tools to detect synchrony are, e.g., Windowed Fourier Transform [18, Ch. 2], Granger Causality Analysis, Cross Wavelet coherence analysis [14], all of which can detect synchrony but are typically performed as a post analysis. We developed an algorithm, called Offset-Adjusted SImilarity Score (OASIS), to run on Action Unit (AU) AU12_r and AU6_c extracted from video data through OpenFace [3], which comprises the Duchenne smile [10]. Machine Learning (ML) approaches also exist to detect similarity on unstructured data but they require data for each modality of synchrony to train and are challenged by their lack of explainability [1, 24]. We get the AU time series signal from video data from each member in a pair as input to our OASIS algorithm, which then detects the time, shape, and value similarities of the smiles into a single similarity score. Our OASIS algorithm requires no prior trained model and is inspired by the Bag-of-Words (BoW) algorithm [21]. It uses the Symbolic Aggregate approXimation (SAX) algorithm [20] to represent the time series in a symbolic representation of "words" that we can ultimately use to find the similarly between the shape of two signals and lends itself to be human understandable. 1, shows an example of the similarity score calculated by OASIS. We see how the original Duchenne smile signals (labeled User 1 and User 2) are transformed into words, which are used to calculate the OASIS (*cf.* Section 4).

We compare the OASIS to traditional coordination analysis methods such as point-wise correlations and dynamically time-warped correlation methods as well as conduct an experiment to gather a user-based interoperability score to see how the OASIS differs from an "intuitive" similarity detection. Our score compares well to both correlative methods and has the ability to be adapted to

CHI '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9422-2/23/04.
https://doi.org/10.1145/3544549.3585709

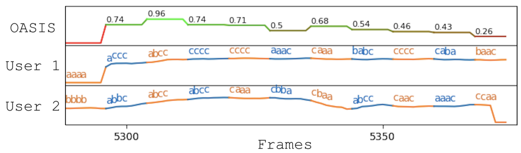

Figure 1: Offset-Adjusted SImilarity Score (OASIS) performed on two time series signals of the Duchenne smile detected by AU12_r and AU6_c.

run in real-time applications where performance is a big factor. An additional benefit of the OASIS to traditional correlation methods is that it has a noise cancellation smoothing effect for a less volatile coordination metric due to the nature of the SAX representation step, which our OASIS algorithm relies on. We also found when comparing the OASIS to how humans rate smile similarity, it works comparatively well and embodies an intuitive smile similarity like that which humans would detect. Although, this comparison lead us to find some false negatives which we must adapt the OASIS to as well as follow up tests to compare our algorithm's efficacy. We can use this as a foundation to continue to build more modalities of synchrony through wearable devices, into a singular score and are now one step closer to quantifying a nonverbal communication metric that can be used in applications for humans and computing systems alike.

2 RELATED WORK

Many statistical methods to detect coordination in pair-wise interactions exist, Cliff et. al created a taxonomy of such methods used in various fields [8]. Some of the methods that are commonly used for synchrony detection are windowed or take a more regional span of time into account to find correlation between two time series. Windowed methods take into account similarities in the signals based on a certain offset starting from statistical tools such as a rolling window Pearson correlation to more advanced, and computationally intensive, like DTW [5]. Additionally, measuring a frequency component is a common objective in synchrony detection where Windowed Fourier [18, Ch. 2] or Cross-Wavelet Transforms [14] are used. Cross-wavelet methods in particular operates over a wide range of time scales. Further, to avoid edge effects they are typically run over an entire time series in post-processing. This makes them unsuited for real-time analysis.

We want a real-time measure on unstructured video data which would require an online algorithm or a trained ML model. Autoencoders are an example of such an ML model that can work on unstructured time series data but are typically used for anomaly detection, see, e.g., [6]. When we train one autoencoder to predict the signal of another person, it can be used to detect synchrony [2]. One drawback of using ML-based methods to detect synchrony on time series data is their lack of interpretability [1, 24]. The BoW algorithm [21] was introduced to compare the similarity between long time series signals. This is done by transforming the time series signal into a higher level representation allowing for fast comparison. The similarity between time series signals is reduced to comparing the frequency of interpretable "words" (short patterns) in both signals. This allows one to compare signals on a structural level instead of comparing the shape of two time series data, e.g., using the Euclidean or the DTW distances, which is the inspiration for the OASIS.

3 PROBLEM FORMULATION

3.1 Collected Survey Data

Our proposed algorithm works on a pair of face frontal video data recordings or streams for unstructured conversations. Our OASIS algorithm can work online and does not require us to first train a model. All video streams are pre-processed (at 30 fps) through OpenFace [3] to extract the facial AUs of which AU6_c and AU12_r were used to create a single time series to represent the presence of a Duchenne smile [10] for each user. This signal is subsequently fed into our algorithm.

We developed our OASIS algorithm on an initial dataset, and wanted to further produce an equally comparable test set for a user-centric interoperability rating to additionaly evaluate OASIS algorithm's ability to detect smile synchrony to a human's ability. To collect an interoperability rating we hosted a survey with the three, three minute long video segments of an unstructured conversation over Zoom between the coauthors, answering five questions from a questionnaire (quesitons 4, 12, 16, 29, and 34 respectively) which tries to generate interpersonal closeness [4]. The order of the videos presented were randomized and the participants task was to press and hold a button, or their keyboard space bar, for the duration of anytime they found a similar "type" of smile between the co-authors in the video, which is the data we logged and used to determine if participants detected smile similarity. The type of smile was defined intuitively by asking the participants if the smile had similar "energy in which their smile is matching, if they are both showing teeth, and are both smiling with their eyes as well as their mouth". Video segments of high and low smile synchrony determined by the user score were chosen and had no audio, pause, or playback option. The survey experience lasted around 12 minutes to complete.

We ran the survey for two weeks, and had a total of 27 participants (m = 15, f = 12, d = 0) which we recruited via snowball sampling our immediate network of colleagues, friends, and family. Participants had average age of 33.37 years old (std = 13.6 years), 17 of which had completed a bachelors or masters program. Participants mostly lived in Germany (n = 20) but we also had participation from the US and Portugal (n = 2 respectively), Belgium, Russia, Korea (n = 1 respectively). Participants who completed at least one of the three videos in the survey were counted but only five completed the full user study. We think this was partly due to the fatiguing nature of the task or unclear instruction that there was more videos to label. Due to randomizing the order, the three videos had n=16, n=16, n=17 fully completed responses respectively.

3.2 Symbolic Representations of Signals

The BoW (bag-of-patterns) algorithm [21] inspired our proposed OASIS algorithm. To generate the "words" for the BoW, one first applies the SAX algorithm [20] to each frame of a sliding window which is passed over the whole time series signal. We use the idea of a symbolic representation of the time series signal in a sliding window to calculate the shape similarity between the two signals at each frame. The SAX algorithm creates this symbolic representation by first applying Piecewise Aggregate Approximation (PAA) [19]. This replaces the value of sub-sequences in the time series signal with their average value. The SAX algorithm then maps each average value to a symbolic value, e.g., "a", "b", or "c". It was shown in [20], that the distance between two symbolic representations lower bounds the distance between the original time series signals. This means that we can use the distance between the symbolic words as a lower bound on the shape similarity of the time series signals.

To calculate the symbolic representation of a time series signals, we need to define the following parameters: *windowLength* (ω), defines how many frames each window contains; *wordLength* (χ), defines how many symbols each window is mapped to (generally, $\omega/\chi \in \mathbb{Z}_{+}$); and, *numBins* ($\alpha$), defines the size of the alphabet or the number of symbols used to represent the time series signal (e.g., $\alpha = |\{\text{"a"}, \text{"b"}, \text{"c"}\}| = 3$). An upper limit on the number of potential words is α^{χ}.

4 THE OASIS ALGORITHM

Once we have a symbolic representation of our time series signals (using the SAX algorithm, *cf.* Section 3.2), our OASIS algorithm calculates the similarity score between the pair's signals and combines it into an overall score. Our OASIS algorithm requires a buffer of length $2\omega - 1$ frames to calculate the similarity between two signals. The OASIS is calculated via the combination of shape similarity (Φ), value similarity (Ψ), and temporal offset similarity coefficient (σ). These factors are multiplied to get

$$\kappa = \sigma \cdot \Phi \cdot \Psi \in [0, 1]. \tag{1}$$

Algorithm 1 Offset-Adjusted SImilarity Score (OASIS)

1: **Inputs**: Words $W_{a_0}, W_{b_0}, \ldots, W_{b_{\omega-1}}$, signals $S_{a_0}, S_{b_0}, \ldots, S_{b_{\omega-1}}$
2: **Initialize**: $\kappa \leftarrow 0$
3: $\theta_a \leftarrow \sum_{j=0}^{\omega} |S_{a_0}|/\omega$, $\theta_b \leftarrow \sum_{j=0}^{\omega} |S_{b_0}|/\omega$
4: **if** $\min(\theta_a, \theta_b) > \tau$ **then**
5:     **for** $\lambda = 0, \ldots, \omega - 1$ **do**
6:         $\sigma \leftarrow$ (2)
7:         **if** $\kappa \geq \sigma$ **then**
8:             **return** κ
9:         **end if**
10:         $\Phi \leftarrow$ (3), $\Psi \leftarrow$ (5)
11:         $\kappa \leftarrow \max(\kappa, \sigma * \Phi * \Psi)$
12:     **end for**
13: **end if**
14: **return** κ

Our OASIS algorithm is depicted in Figure 1. As inputs, we pass the symbolic words ($W_{a_0}, W_{b_0}, \ldots, W_{b_{\omega-1}}$), as well as the raw time series signals ($S_{a_0}, S_{b_0}, \ldots, S_{b_{\omega-1}}$) at various offset values. First, we check whether the signal energy is above a pre-defined threshold τ in Line 4. Next, to calculate (1), we fix the signal a_0 and pass over the $\omega - 1$ offset frames of the other signal b_λ (*cf.* Line 5). The offset coefficient (σ) is calculated as

$$\sigma = 1 - \frac{\lambda}{\omega}, \tag{2}$$

with the current offset $\lambda = 0, \ldots, \omega - 1$. This captures the time offset between the two signals, making our similarity score offset-adjusted. If the similarity score is larger than the given offset, we break the loop here since this is the most similar the two scores can be (*cf.* Line 7).

Next, we calculate shape similarity (Φ) as

$$\Phi = 1 - \left(\frac{\Delta_{a_0 b_\lambda}}{\chi}\right), \tag{3}$$

where $\Delta_{a_0 b_\lambda}$ is the character distance between the two words. To calculate this, we use the fixed word of the signal a_0 and compare this to the current offset word of signal b_λ. We then calculate the distance between each symbol in the words to each other and normalize the result, i.e.,

$$\Delta_{a_0 b_\lambda} = \sum_{i=0}^{\chi-1} \frac{\left|W_{a_{0_i}} - W_{b_{\lambda_i}}\right|}{\alpha - 1}, \tag{4}$$

where $W_{a_{0_i}}$ is the i-th symbol in the word a_0 and $W_{b_{\lambda_i}}$ is the i-th symbol in the offset word b_λ. The distance is calculated as the absolute difference of the Unicode character representations of the symbols. This is similar to template matching, see, e.g., [12].

Lastly, the value similarity (Ψ) is taken into account as

$$\Psi = 1 - \left(\frac{|\theta_{a_0} - \theta_{b_\lambda}|}{\omega}\right), \tag{5}$$

where $\theta_i = \sum_{j=0}^{\omega} |S_{i_j}|/\omega$ is the normalised square-root signal energy of the original time series signal S_i (non-symbolic representation), with $i = a_0, b_\lambda$. This is inspired by previous work using the energy of time series signals to define synchrony, see, e.g., [26]. Our proposed algorithm calculates (2), (3), (5) for each $\lambda = 0, \ldots, \omega - 1$ offset words of signal b_λ (*cf.* Lines 6, 10) unless the for-loop is prematurely stopped (*cf.* Line 7).

The similarity score is simultaneously calculated where the roles of signal a and signal b are reversed, i.e., the fixed word is b_0 and we loop over the buffered offset words of signal a_λ. By multiplying (2), (3), and (5) we have a working similarity score for a single modality. Finally, we take an average of the two similarity scores where signal a was the fixed word and where signal b was the fixed word. This is the final similarity score for the current frame. An implementation of the algorithm can be found here[1].

5 RESULTS

Figure 2 (below) shows the comparison of the original signals to the various similarity detection methods. All values in Figure 2 have been normalized between 0 and 1. The first row in Figure 2 shows the raw signal of the Duchenne smile of both users in dyad 1, i.e., the conversational pair in the first of the three videos. To detect a Duchenne smile, we used OpenFace [3] to get the signal of AU12_r if AU6_c was present. The second row shows a comparison of a rolling window Pearson correlation (ω = 8), followed by the third row of DTW over the entire signal. We get the DTW score by using the `accelerated_dtw` from the DTW python package [13] with the Euclidean distance measure. We plot the diagonal of the cost matrix inverted by subtracting by one and setting the values to 0 if either of the time series are also 0. The fourth row is the

[1] https://github.com/TUMFARSynchrony/synchrony-score

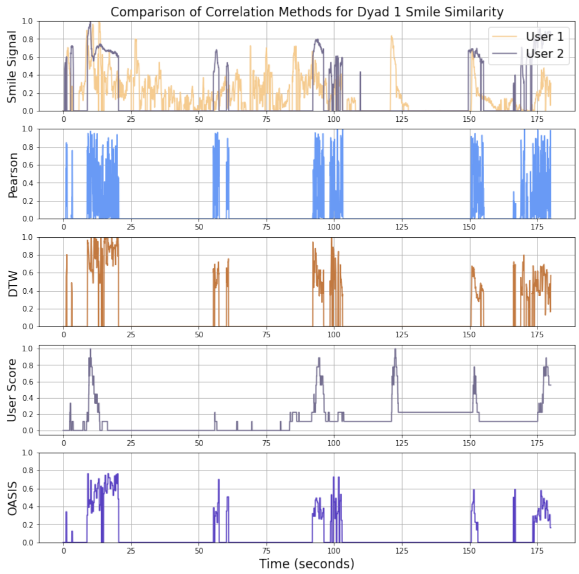

Figure 2: A comparison of similarity detection methods shown over the raw input signal of two users' Duchenne smiles (AU12_r when AU6_c was present). A rolling window Pearson correlation, DTW, user interoperability score, and the OASIS are shown.

user-based interoperability score, which was calculated by taking the normalized sum across when each participant detected smile similarity as outlined in 3.1. We take this to be a "ground truth" when evaluating the other similarity methods. Lastly, the fifth row is our OASIS algorithm run using $\alpha = 3$, $\omega = 8$, and $\chi = 4$ as input parameters chosen through an initial optimization on the data we used to develop our algorithm.

When comparing the placement of the user score to that of our OASIS algorithm, we see peaks at similar times. Chronologically, when comparing the OASIS around 90-95 seconds, the OASIS picks up on the diminished strength of similarity detected as shown by the user score, as compared to the other correlation methods. Next, looking at around 125 seconds, we have a false negative where the OASIS and the other correlative methods do not detect a smile similarity whereas the user score does. This is because the algorithms only works on Duchenne smile types. Looking at user 2's raw signal, no Duchenne smile was detected since OpenFace lost track of the eyes and therefore AU6_c. The user study participants might have also more widely interpreted smile similarity than a Duchenne smile, therefore having access to other signals that the OASIS did not consider. This is further indicated by the higher than 0 average of user score until the next peak at 150 seconds. At 150 seconds, the OASIS shape matches the user score rather closely. Lastly, at around 170 seconds, the users did not detect a smile similarity although the correlation methods and the OASIS did. Although the user score was detecting some similarity on average, this may have been too fast of a peak for participants to react and record to this strong instance of smile similarity. Finally, the OASIS also shows a more defined peak at this region because of its highly temporally adjusted factors.

	OASIS	DTW	Pearson	User
Area between curves [17]	22.29	25.81	23.93	N/A
Number of peaks	44	149	236	44

Table 1: The average area between the "ground truth" user score and the OASIS, DTW, and Pearson, and the average number of peaks.

To quantify the results observed in Figure 2, we calculate the average area between the curves [17] and the average number of peaks in the scores. We compare the area between our "ground truth" user score to each of the other scores. As we see in the first row of Table 1, the OASIS is the most similar to the user score with an average area between curves of 22.29 (lower is better). Moreover, in the second row of Table 1, we observe that the average number of peaks for the OASIS is the same as the user score. The number of peaks for the DTW and the Pearson scores are much larger than the user score. The number of peaks was calculated using the `find_peaks` function in SciPy signal processing toolbox [27]. This confirms the qualitative results we observe in Figure 2, i.e., the Pearson and the DTW scores fluctuate more than the user score and the OASIS.

6 DISCUSSION

From Section 5, we observe, both qualitatively and quantitatively, that our proposed OASIS algorithm is most similar to a user study score compared to two other measures of similarity. This could be due to the fact that the OASIS not only looks at the shape similarity between the users' signals, but it also takes the offset and value similarities into account. Moreover, since our OASIS algorithm has relatively low time-complexity (compared with DTW), it can be adapted to run real-time on wearable devices with comparable results.

A challenging aspect of using ML-based algorithms to define a synchrony score, e.g., using autoencoders [2], is their lack of transparency [1, 24]. However, our proposed OASIS algorithm is explainable by design since each part of the similarity score is interpretable. Moreover, the SAX representation of words directly indicates the shape of the underlying signal, making it directly interpretable by a human user.

One limitation we have is relying on the model performance of OpenFace for detecting facial AUs and markers whose false negatives are inherited by the similarity detection algorithms. Although, if these AUs are not detected, we argue that other similarity score algorithms and ML models will also not be able to detect any similarity. This issue can ultimately be counteracted by adding additional input signals or modalities involved in social interaction, e.g., other facial AUs, heart rate, body gestures, etc, to increase the robustness of the score. A benefit of our algorithm is also that it can be extended to work on more than one signal, such as signals directly collected and calculated in real-time on wearable devices (e.g., smart watches, HMDs, etc.) in the wild [16, 28]. An additional next step to incorporate into our OASIS algorithm is to expand on the different types of similarity we would like to detect. Adding long, medium, and short term buffers of the input signals can allow our OASIS algorithm to find similarities at different time scales.

Moreover, the SAX representations of these signals can be extended to accommodate different window lengths as well. Furthermore, our OASIS algorithm can be more specifically adapted to detecting synchrony, which is a subset of similarity detection that is predominately time dependant. We can incorporate a dissimilarity factor into the OASIS to define a true synchrony score through detecting dissimilarities along with similarities.

7 CONCLUSION

We develop an offset-adjusted similarity score that works online on two streamed time series signals. We compared our OASIS algorithim to the rolling window Pearson correlation and the DTW correlation methods, and conducted a survey to gather a user-based interoperability score for smile similarity detection. Given the benefits of having an online, explainable algorithm, it performs well when compared to statistical correlations as well as the collected user-based score. The OASIS can be used on low-power wearable devices enabling the study of high level synchrony trends in the wild. Lastly, this work is the seed to future work that ambitiously aims to quantify social nonverbal interactions for the evaluation and design of novel interactions in future ubiquitous AR/VR systems.

ACKNOWLEDGMENTS

We would like to thank Thad Starner for his invaluable input on evaluating our algorithm. Jamie Ward is funded by a Leverhulme supported Apex grant from The British Academy, Royal Academy of Engineering and Royal Society (APX\R1\201093).

REFERENCES

[1] 2016. *Explainable Artificial Intelligence (XAI)*. Technical Report. DARPA. https://www.darpa.mil/attachments/DARPA-BAA-16-53.pdf

[2] Shahin Amiriparian, Jing Han, Maximilian Schmitt, Alice Baird, Adria Mallol-Ragolta, Manuel Milling, Maurice Gerczuk, and Björn Schuller. 2019. Synchronization in interpersonal speech. *Frontiers in Robotics and AI* 6 (2019), 116.

[3] Brandon Amos, Bartosz Ludwiczuk, and Mahadev Satyanarayanan. 2016. *OpenFace: A general-purpose face recognition library with mobile applications*. Technical Report. CMU-CS-16-118, CMU School of Computer Science.

[4] Arthur Aron, Edward Melinat, Elaine N Aron, Robert Darrin Vallone, and Renee J Bator. 1997. The experimental generation of interpersonal closeness: A procedure and some preliminary findings. *Personality and social psychology bulletin* 23, 4 (1997), 363–377.

[5] Donald J Berndt and James Clifford. 1994. Using dynamic time warping to find patterns in time series.. In *KDD workshop*, Vol. 10. AAAI, 359–370.

[6] Raghavendra Chalapathy and Sanjay Chawla. 2019. Deep learning for anomaly detection: A survey. *arXiv preprint arXiv:1901.03407* (2019).

[7] Kishan Chandan, Vidisha Kudalkar, Xiang Li, and Shiqi Zhang. 2019. Negotiation-based human-robot collaboration via augmented reality. *arXiv preprint arXiv:1909.11227* (2019).

[8] Oliver M Cliff, Joseph T Lizier, Naotsugu Tsuchiya, and Ben D Fulcher. 2022. Unifying Pairwise Interactions in Complex Dynamics. *arXiv preprint arXiv:2201.11941* (2022).

[9] Emilie Delaherche, Mohamed Chetouani, Ammar Mahdhaoui, Catherine Saint-Georges, Sylvie Viaux, and David Cohen. 2012. Interpersonal synchrony: A survey of evaluation methods across disciplines. *IEEE Transactions on Affective Computing* 3, 3 (2012), 349–365.

[10] Paul Ekman, Richard J Davidson, and Wallace V Friesen. 1990. The Duchenne smile: Emotional expression and brain physiology: II. *Journal of personality and social psychology* 58, 2 (1990), 342.

[11] Cathy Mengying Fang, G. R. Marvez, Neska ElHaouij, and Rosalind Picard. 2022. Cardiac Arrest: Evaluating the Role of Biosignals in Gameplay Strategies and Players' Physiological Synchrony in Social Deception Games. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 240, 7 pages. https://doi.org/10.1145/3491101.3519670

[12] Jordan Frank, Shie Mannor, Joelle Pineau, and Doina Precup. 2012. Time series analysis using geometric template matching. *IEEE transactions on pattern analysis and machine intelligence* 35, 3 (2012), 740–754.

[13] Toni Giorgino. 2009. Computing and visualizing dynamic time warping alignments in R: the dtw package. *Journal of statistical Software* 31 (2009), 1–24.

[14] Aslak Grinsted, John C Moore, and Svetlana Jevrejeva. 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear processes in geophysics* 11 (2004), 561–566.

[15] Ihshan Gumilar, Amit Barde, Prasanth Sasikumar, Mark Billinghurst, Ashkan F. Hayati, Gun Lee, Yuda Munarko, Sanjit Singh, and Abdul Momin. 2022. Inter-Brain Synchrony and Eye Gaze Direction During Collaboration in VR. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 345, 7 pages. https://doi.org/10.1145/3491101.3519746

[16] Aman Gupta, Finn L Strivens, Benjamin Tag, Kai Kunze, and Jamie A Ward. 2019. Blink as You Sync: Uncovering Eye and Nod Synchrony in Conversation Using Wearable Sensing. In *Proceedings of the 23rd International Symposium on Wearable Computers* (London, United Kingdom) *(ISWC '19)*. Association for Computing Machinery, New York, NY, USA, 66–71. https://doi.org/10.1145/3341163.3347736

[17] Charles F Jekel, Gerhard Venter, Martin P Venter, Nielen Stander, and Raphael T Haftka. 2019. Similarity measures for identifying material parameters from hysteresis loops using inverse analysis. *International Journal of Material Forming* 12, 3 (2019), 355–378.

[18] Gerald Kaiser and Lonnie H Hudgins. 1994. *A friendly guide to wavelets*. Vol. 300. Springer.

[19] Eamonn Keogh, Kaushik Chakrabarti, Michael Pazzani, and Sharad Mehrotra. 2001. Locally adaptive dimensionality reduction for indexing large time series databases. In *Proceedings of the 2001 ACM SIGMOD international conference on Management of data*. 151–162.

[20] Jessica Lin, Eamonn Keogh, Stefano Lonardi, and Bill Chiu. 2003. A symbolic representation of time series, with implications for streaming algorithms. In *Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery*. 2–11.

[21] Jessica Lin, Rohan Khade, and Yuan Li. 2012. Rotation-invariant similarity in time series using bag-of-patterns representation. *Journal of Intelligent Information Systems* 39, 2 (2012), 287–315.

[22] Runpeng Liu, Joseph P Salisbury, Arshya Vahabzadeh, and Ned T Sahin. 2017. Feasibility of an autism-focused augmented reality smartglasses system for social communication and behavioral coaching. *Frontiers in pediatrics* 5 (2017), 145.

[23] Daniel McDuff, Christophe Hurter, and Mar Gonzalez-Franco. 2017. Pulse and vital sign measurement in mixed reality using a HoloLens. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*. 1–9.

[24] Christoph Molnar. 2020. *Interpretable machine learning*. Lulu. com.

[25] Shannon Monahan, Emmanuel Johnson, Gale Lucas, James Finch, and Jonathan Gratch. 2018. Autonomous agent that provides automated feedback improves negotiation skills. In *International Conference on Artificial Intelligence in Education*. Springer, 225–229.

[26] Fabian T Ramseyer. 2020. Motion energy analysis (MEA): A primer on the assessment of motion from video. *Journal of counseling psychology* 67, 4 (2020), 536.

[27] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. 2020. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* 17 (2020), 261–272. https://doi.org/10.1038/s41592-019-0686-2

[28] Jamie A Ward, Daniel Richardson, Guido Orgs, Kelly Hunter, and Antonia Hamilton. 2018. Sensing interpersonal synchrony between actors and autistic children in theatre using wrist-worn accelerometers. In *Proceedings of the 2018 ACM International Symposium on Wearable Computers*. 148–155.