2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) | 979-8-3503-7449-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/VRW62533.2024.00123

Towards Understanding Personality Expression via Body Motion

Sinan Sonlu * Bilkent University
Yalım Doğan † Bilkent University
Arçin Ülkü Ergüzen ‡ Bilkent University
Musa Ege Ünalan § Bilkent University
Serkan Demirci ¶ Bilkent University
Funda Durupinar ‖ University of Massachusetts Boston
Uğur Güdükbay ** Bilkent University

Abstract

This work addresses the challenge of data scarcity in personality-labeled datasets by introducing personality labels to clips from two open datasets, ZeroEGGS and Bandai, which provide diverse full-body animations. To this end, we present a user study to annotate short clips from both sets with labels based on the Five-Factor Model (FFM) of personality. We chose features informed by Laban Movement Analysis (LMA) to represent each animation. These features then guided us to select the samples of distinct motion styles to be included in the user study, obtaining high personality variance and keeping the study duration and cost viable. Using the labeled data, we then ran a correlation analysis to find features that indicate high correlation with each personality dimension. Our regression analysis results indicate that highly correlated features are promising in accurate personality estimation. We share our early findings, code, and data publicly.

Index Terms: Computing methodologies—Artificial intelligence—Computer vision—Activity recognition and understanding; Computing methodologies—Computer graphics—Animation—Motion processing

1 Introduction

Applications involving intelligent assistants benefit from understanding what happens in a scene and how it occurs. Recognizing human psychological states can significantly improve human-computer interaction. For example, if a system knows whether a user is introverted or extroverted, it can respond more accurately to their queries. Similarly, visual assistance systems can offer more customized information to visually impaired individuals by interpreting the style of people's movements in their environment.

Body language provides subtle cues about a person's psychological self, necessitating close analysis. Our movements and posture reflect our emotions and personality; for instance, no two hand waves are identical. While apparent personality might not always reflect self-reported or actual personality traits, people generally make accurate judgments about others' personalities [11, 23]. These judgments often draw on voice [22], appearance [15], body language [26], and facial expressions [3]. These features are detectable in video input but can be influenced by external factors such as environment and lighting. In contrast, animation data offers a cleaner way to interpret the expressed personality.

*e-mail: sinan.sonlu@bilkent.edu.tr
†e-mail: yalim.dogan@bilkent.edu.tr
‡e-mail: ulku.erguzen@bilkent.edu.tr
§e-mail: ege.unalan@bilkent.edu.tr
¶e-mail: serkan.demirci@bilkent.edu.tr
‖e-mail: funda.durupinarbabur@umb.edu
**e-mail: gudukbay@cs.bilkent.edu.tr

This work-in-progress aims for content-independent personality recognition in short animation clips based on full-body motion cues in 3D animation without relying on facial and language cues. One challenge in data-driven personality recognition is the need for datasets with personality labels. Existing personality-annotated motion datasets [7, 4, 10, 24] are limited in providing full-body joint data and action variety. To address this, via a user study, we introduce personality labels to clips from two open datasets, ZeroEGGS [12] and Bandai [18]. Both datasets provide full-body animations and portray various actions. The labels are based on the Five-Factor Model (FFM) of personality [6], which inspects the individual under five orthogonal dimensions: openness, conscientiousness, extraversion, agreeableness, and neuroticism. We provide the data labels and our code for the analysis publicly [1].

We use Laban Motion Analysis (LMA)-inspired descriptive features to analyze movement cues related to apparent personality. LMA offers a way to identify movement style with its Effort qualities, which describe the dynamic aspects of movement in four dimensions: Space, Weight, Time, and Flow. Due to its success in the literature in conveying personality in animation [8, 29], recognizing personality [9], emotions [31, 5], and actions [25], LMA provides a solid basis to select movement features. We divide the animation clips from the datasets into unique, short segments based on LMA features to assess the viewers' immediate perceptions of personality.

We determine the features that influence apparent personality through correlation analysis between our LMA-driven motion features and the user-annotated personality labels of the samples. We utilize these parameters in a regression model to determine the FFM personality labels of input motion. Unlike existing work, we focus on predicting personality independent of the motion content so that the resulting system will generalize better. We discuss our early results in this work-in-progress article.

2 Related Work

Previous work focuses on different input features for predicting personality from video or animation. Locomotion speed and style can predict robot personality [2]. Humans can interpret simple motion cues of basic shapes with different personality traits, and these features can be used for assessing the personality of the movement [19]; similarly, in videos, the personality can be predicted using minimal information [20]. Gesturing [21] and hand motion [32] are essential in accurately assessing personality. Facial expressions are also influential on personality perception [16], automatic recognition systems often utilize facial landmarks [30, 28] and a combination of different audiovisual features [17]. It is also possible to analyze the speaking language to predict personality [14].

Existing human motion datasets that include personality labels offer limited motion and variety. For example, the UDIVA dataset [24] includes motion capture of the upper body of sitting individuals while interacting with objects on the table. The First Impressions dataset [10] includes close-up videos of the individuals showing only the face and shoulders with limited motion since they mainly talk to the camera. The AMIGOS dataset [4] records the upper bodies of

[1] `https://github.com/sinansonlu/animation-personality`

979-8-3503-7449-0/24/$31.00 ©2024 IEEE
DOI 10.1109/VRW62533.2024.00123

people watching and reacting to videos where the movement is minimal. Personality in a nonsocial context dataset [7] includes daily life activities within an environment but has reduced joint accuracy due to overlapping scene objects. On the other hand, predicting motion personality and generating motion with personality requires a good variety of animation samples in terms of motion content and style. We conducted a user study to label selected animation samples in terms of their personality traits.

3 Method

This work aims to find the personality information in motion style. Previous work establishes a mapping between personality and motion using procedural animation techniques [8] and comparative analysis. In this work, we wanted to establish ground truth values on unaltered, real-life motion capture data. To this end, we conducted a user study to annotate existing motion capture datasets, ZeroEGGS [12] and Bandai [18], which include a good variety of animation content and style for single-person animations. We present the study and further data analysis in this section.

3.1 Data

The ZeroEGGS [12] dataset contains 67 sequences of monologues performed in 19 motion styles, where each style has 3 to 10 minutes of animation; the Bandai [18] dataset contains 20 different animation content in 15 different acted styles. We preferred to focus on short motion sequences for the user study and the subsequent analyses. Bandai samples are sufficiently short, but ZeroEGGS samples are 3 to 10-minute animations per style. Thus, we prepared short animation segments of 3 to 10 seconds, resulting in 1291 samples, which is reduced to 100 with our selection process described in the following subsection. ZeroEGGS segments of each style differ in numbers, yet they are balanced based on our motion features. We excluded the locomotion alternatives (walk-right, walk-left, and walk-back) and long dance animations from the Bandai dataset, resulting in 128 samples. A total of 228 animations are labeled in our user study.

3.2 Sample Selection

To keep the study duration and cost manageable and avoid redundancy, we selected the animations expressing the most distinct personality traits and, thus, the most unique LMA features. We selected 21 parameters, such as distances or angles formed by different joints, following earlier research utilizing LMA for emotion analysis [31,5], to describe each sample. We constructed a vector of K per parameter for each animation, where each element represents the average value of that parameter for the corresponding period. For example, for $K = 5$, we have five averages for each $N/5$ frames of an animation of N frames. Consequently, a $21 \times K$ matrix summarizes the LMA features of each animation. By using different values of K, we can represent the motion at different levels of detail.

The Euclidean distance between two such matrices averaged over different K values was used to measure dissimilarity between the two animations. Animation pairs that differ on various detail levels have high dissimilarity (see Figure 1). We started with 1291 animation segments from ZeroEGGS and followed a procedural approach to reduce the animation count to 100. This procedure starts by identifying the two animations that are most similar in style. It then calculates the average dissimilarity of these two animations compared to the other samples. The animation that is more similar to the remaining samples is removed. This process leaves us with samples with the maximum dissimilarity among themselves, indicating that they likely represent distinct styles.

3.3 Feature Selection

Bandai animations include 3D rotations of 22 joints and the translation of the root joint, resulting in 69 parameters per frame. ZeroEGGS animations also include finger joints and more spine bones,

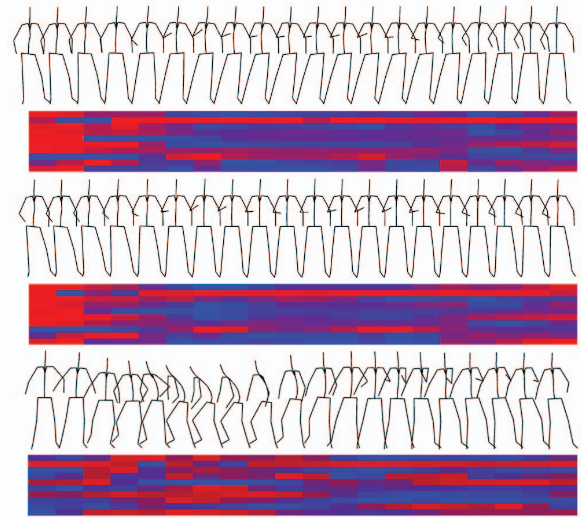

Figure 1: Top 10 rows of the feature matrices of two similar animations (top two pair) and a dissimilar third one.

which makes a total of 50 joints, resulting in 153 parameters per frame. Not all input features are meaningful for personality analysis and can be combined. For example, the rotations of the individual spine bones would contribute to the overall inclination of the upper body and thus can be reduced into a single element. Moreover, for certain joints, the small changes in consecutive frames are less significant than the overall change of the motion; hence, we can use the sum over a couple of frames to keep the time dimension limited.

Previous research suggests that personality expression in motion is related to the interaction between different joints rather than the movements of individual joints [27]. For example, the linear distance between the hands can better reflect extraversion than the individual rotations of both hands, arms, and shoulders. Furthermore, LMA Efforts can guide the selection of motion parameters relevant to personality as shown in the literature [8,29]. Thus, we calculate the following parameters for the joints:

1. *Distance:* The Euclidean distance between joint pairs. Horizontal pairs are associated with Space Effort regarding direction, and vertical pairs are associated with Weight Effort regarding strength.
2. *Speed:* The Euclidean distance of joints between consecutive frames, associated with Time Effort regarding urgency.
3. *Acceleration:* The change of joint speed between consecutive frames. It is associated with Flow Effort regarding control.
4. *Angle:* The positive angle formed by any joint trio.
5. *Volume:* The volume of the Convex Hull of a subset of joints, associated with Space Effort regarding spatial intention.

3.4 Study Design

We implemented a tool for rating the animation files using Unity; a screenshot of the WebGL application is shown in Figure 2. The participants were presented with randomly selected unlabeled animations after entering email addresses and demographic information into the system. The animation was looped until the participants submitted their responses. Sliders below the 3D model helped the user rotate the model and go to a specific animation time. While moving the sliders, the animation was paused. We used an abbreviated form of Ten-Item Personality Inventory [13] where each personality trait appears with its polar opposite on a 7-point Likert scale. After submitting the personality values for the current animation, a new one appeared until the participant rated predefined numbers of samples.

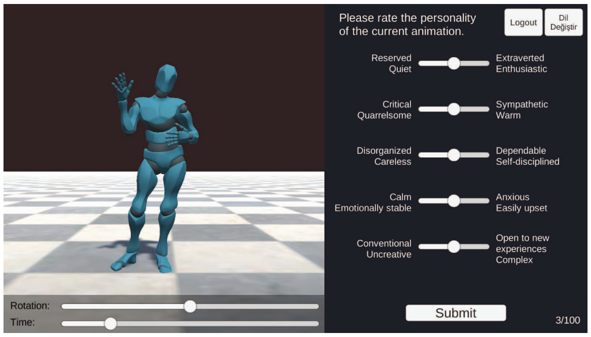

Figure 2: Screenshot from the user study for personality labeling.

We performed studies for the selected animations of Bandai (128 samples) and ZeroEGGS (100 samples) datasets. Participants were recruited via email invitations and the crowdsourcing service, *Prolific*. The users invited by email participated in the study voluntarily, rating any number of samples they could. The users from Prolific were paid based on the median time they spent on the task, which included rating 10 unlabeled animation samples. A total of 181 unique users participated in the Bandai study, where each sample was rated by at least 15 and at most 20 different users. The unique user count for the ZeroEGGS study was 109, with each sample rated by 15 different users. We calculated the average user rating to determine the ground truth personality values per sample. We omitted the ratings submitted in less time than the duration of the animation.

4 Analysis

We analyzed the personality labels of the annotated samples to determine relevant input features. Each personality dimension is expressed in the $[-3,3]$ range, where 0 corresponds to being neutral for that dimension. To understand the relationship of each parameter set with each personality dimension, we first ran a correlation analysis, calculating the Pearson correlation values between each calculated parameter and the personality factor. Table 1 shows the highest correlation coefficients achieved for parameters in the corresponding group with each personality trait. We observe that distance, angle, and volume-based features are generally superior, and acceleration highly correlates with conscientiousness.

Table 1: Pearson correlation (r) and p values for the best parameters per category and factor. * indicates $p < .002$, ** indicates $p < .001$.

Parameter	O	C	E	A	N
$r_{Distance}$	$.548^{**}$	$.359^{**}$	$.599^{**}$	$-.474^{**}$	$-.387^{**}$
r_{Speed}	$.210$	$-.470^{**}$	$.450^{**}$	$-.328^{**}$	$.454^{**}$
$r_{Acceleration}$	$.265^{*}$	$-.523^{**}$	$.502^{**}$	$-.473^{**}$	$.503^{**}$
r_{Angle}	$-.568^{**}$	$.410^{**}$	$.671^{**}$	$-.555^{**}$	$.423^{**}$
r_{Volume}	$.414^{**}$	$-.420^{**}$	$.608^{**}$	$-.543^{**}$	$.308^{**}$

Following a greedy approach, we also investigated the combinations of the calculated parameters per group. We experimented with different combinations of coefficients by multiplying them with -1, 0, or 1 and summing them up to find the combinations that yielded the highest correlations (Table 2). We see that distance and angle-based parameters are more descriptive than the parameters of the other categories; we chose 23 input features based on the results of this correlation study. Please refer to our code for these features.

Next, we performed regression experiments for estimating personality given input motion. We organized the data in windows of size

Table 2: Pearson correlation coefficients (r) for the best combinations in each parameter group. All the p values are < 0.001.

Parameter	O	C	E	A	N
$r_{Distance}$	.693	.545	.826	-.653	-.582
r_{Speed}	-.265	.524	-.535	.355	-.479
$r_{Acceleration}$	.292	.556	-.520	.473	-.550
r_{Angle}	.854	.826	.932	.823	-.803
r_{Volume}	.714	.678	.766	.747	-.548
$r_{Overall}$	.906	.914	.970	-.910	-.905

30 by sliding them one frame at a time. Each subsample has 23 input features and an output vector of size 5. Each output vector element represents one personality factor in the range $[-3,3]$, corresponding to the TIPI scores obtained from the user study.

We used 30% of the samples for testing, keeping the ones originating from the same animation clip before splitting on the same side. For example, if a long animation was split into two parts, both were used for training or testing. After normalizing the data, we experimented with different hidden unit sizes for a two-layer sequential model. We used Adam optimizer with a learning rate of 0.0001 and batch size of 32. We computed Mean Square Error (MSE) for loss; hidden units utilized Exponential Linear Unit (ELU) activation. We used L1, L2, and Batch normalization, initializing hidden layers with HeNormal [2]. We trained each model for 500 epochs and averaged the minimum MSE over five runs. We observe the best combination as $512-256$ hidden units, resulting in an MSE of 0.530 over the range $[-3,3]$. The results indicate that the selected input features can provide a sufficiently accurate estimate of personality even with a simple, fully connected network.

5 Conclusion and Future Work

We report our progress towards understanding how personality is expressed through motion by LMA-inspired movement features and a user study. We propose an elimination process to identify motion samples with unique features for labeling, which helps reduce the budget without losing variety. The resulting labels can be utilized in supervised learning for predicting personality and generating personality-enriched animations.

The current work employs a simple feed-forward network for personality prediction. In the future, we plan to explore more sophisticated models, such as Graph Convolutional Networks, which show promising results in recognizing personality in 3D animation data with the help of LMA-based features [9]. Future work can analyze estimation performance with limited animation data; for example, using only hand motions can effectively predict personality [1].

The limited sample size is a problem for successful training; we overcome this issue by reducing the feature space using mid-level features. An alternative solution can be data augmentation by introducing controlled noise to data to increase sample count. LMA can offer a way to alter the samples without changing the action content of motion. Previous work shows that Laban Efforts significantly correlate with apparent personality [8]; adjusting animations considering LMA helps alter personality perception in conversational agents [29]. The animation data can be augmented by modifying the bone rotations according to LMA, providing controlled variation without drastically altering their motion and action type.

Acknowledgments

This research is supported by The Scientific and Technological Research Council of Turkey (TÜBİTAK) under Grant No. 122E123.

[2] https://www.tensorflow.org/api_docs/python/tf/keras/initializers/HeNormal

REFERENCES

[1] A. Adkins, A. Normoyle, L. Lin, Y. Sun, Y. Ye, M. Di Luca, and S. Jörg. How important are detailed hand motions for communication for a virtual character through the lens of charades? *ACM Transactions on Graphics*, 42(3):1–16, 2023.

[2] A. Agnihotri, A. Chan, S. Hedaoo, and H. Knight. Distinguishing robot personality from motion. In *Companion of the ACM/IEEE International Conference on Human-Robot Interaction*, pp. 87–89, 2020.

[3] J.-I. Biel, L. Teijeiro-Mosquera, and D. Gatica-Perez. FaceTube: predicting personality from facial expressions of emotion in online conversational video. In *Proceedings of the 14th ACM International Conference on Multimodal Interaction*, ICMI '12, pp. 53–56, 2012.

[4] J. A. M. Correa, M. K. Abadi, N. Sebe, and I. Patras. AMIGOS: A dataset for affect, personality and mood research on individuals and groups. *IEEE Transactions on Affective Computing*, 12(2):479–493, 2021.

[5] S. Dewan, S. Agarwal, and N. Singh. Laban movement analysis to classify emotions from motion. In *Tenth International Conference on Machine Vision (ICMV 2017)*, vol. 10696, pp. 717–724. SPIE, 2018.

[6] J. M. Digman. Personality structure: Emergence of the five-factor model. *Annual review of psychology*, 41(1):417–440, 1990.

[7] D. Dotti, M. Popa, and S. Asteriadis. Behavior and personality analysis in a nonsocial context dataset. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, CVPRW '18, pp. 2354–2362, 2018.

[8] F. Durupinar, M. Kapadia, S. Deutsch, M. Neff, and N. I. Badler. Perform: Perceptual approach for adding OCEAN personality to human motion using Laban Movement Analysis. *ACM Transactions on Graphics*, 36(1):Article no. 6, 16 pages, 2016.

[9] Z. Erkoç, S. Demirci, S. Sonlu, and U. Güdükbay. Skeleton-based personality recognition using Laban Movement Analysis. In *Understanding Social Behavior in Dyadic and Small Group Interactions*, vol. 173, pp. 74–87. Proceedings of Machine Learning Research, 2022.

[10] H. J. Escalante, H. Kaya, A. A. Salah, S. Escalera, Y. Güçlütürk, U. Güçlü, X. Baró, I. Guyon, J. C. J. Junior, M. Madadi, et al. Modeling, recognizing, and explaining apparent personality from videos. *IEEE Transactions on Affective Computing*, 13(2):894–911, 2020.

[11] D. C. Funder. Accurate personality judgment. *Current Directions in Psychological Science*, 21(3):177–182, 2012.

[12] S. Ghorbani, Y. Ferstl, D. Holden, N. F. Troje, and M.-A. Carbonneau. Zeroeggs: Zero-shot example-based gesture generation from speech. *Computer Graphics Forum*, 42(1):206–216, 2023. doi: 10.1111/cgf.14734

[13] S. Gosling, P. J. Rentfrow, and W. Swann. A very brief measure of the big five personality domains. *Journal of Research in Personality*, 37:504–528, 2003.

[14] L. Hickman, R. Saef, V. Ng, S. E. Woo, L. Tay, and N. Bosch. Developing and evaluating language-based machine learning algorithms for inferring applicant personality in video interviews. *Human Resource Management Journal*, In press.

[15] Y. Hu, C. J. Parde, M. Q. Hill, N. Mahmood, and A. J. O'Toole. First impressions of personality traits from body shapes. *Psychological Science*, 29(12):1969–1983, 2018.

[16] J. Hyde, E. J. Carter, S. Kiesler, and J. K. Hodgins. Evaluating animated characters: facial motion magnitude influences personality perceptions. *ACM Transactions on Applied Perception*, 13(2):Article no. 81, 17 pages, 2016.

[17] O. Kampman, E. J. Barezi, D. Bertero, and P. Fung. Investigating audio, video, and text fusion methods for end-to-end automatic personality prediction. In I. Gurevych and Y. Miyao, eds., *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pp. 606–611. Association for Computational Linguistics, Melbourne, Australia, July 2018. doi: 10.18653/v1/P18-2096

[18] M. Kobayashi, C.-C. Liao, K. Inoue, S. Yojima, and M. Takahashi. Motion capture dataset for practical use of AI-based motion editing and stylization. *arXiv:2306.08861 [cs.CV]*, 2023.

[19] M. Koppensteiner. Perceiving personality in simple motion cues. *Journal of Research in Personality*, 45(4):358–363, 2011.

[20] M. Koppensteiner. Motion cues that make an impression: Predicting perceived personality by minimal motion information. *Journal of Experimental Social Psychology*, 49(6):1137–1143, 2013.

[21] M. Koppensteiner and K. Grammer. Motion patterns in political speech and their influence on personality ratings. *Journal of Research in Personality*, 44(3):374–379, 2010.

[22] P. McAleer, A. Todorov, and P. Belin. How do you say 'hello'? personality impressions from brief novel voices. *PLOS One*, 9(3):e90779, 2014.

[23] C. Nater and E. Zell. Accuracy of social perception: An integration and review of meta-analyses. *Social and Personality Psychology Compass*, 9(9):481–494, 2015.

[24] C. Palmero, J. Selva, S. Smeureanu, J. C. J. Junior, A. Clapés, A. Moseguí, Z. Zhang, D. Gallardo, G. Guilera, D. Leiva, et al. Context-aware personality inference in dyadic scenarios: Introducing the UDIVA dataset. In *Proceedings of the IEEE Winter Conference on Applications of Computer Vision Workshops*, WACVW '21, pp. 1–12, 2021.

[25] Z. Ramezanpanah, M. Mallem, and F. Davesne. Human action recognition using laban movement analysis and dynamic time warping. *Procedia Computer Science*, 176:390–399, 2020.

[26] M. Romeo, D. Hernández García, T. Han, A. Cangelosi, and K. Jokinen. Predicting apparent personality from body language: benchmarking deep learning architectures for adaptive social human–robot interaction. *Advanced Robotics*, 35(19):1167–1179, 2021.

[27] H. J. Smith and M. Neff. Understanding the impact of animated gesture performance on personality perceptions. *ACM Transactions on Graphics*, 36(4):Article no. 49, 12 pages, 2017.

[28] S. Song, S. Jaiswal, E. Sanchez, G. Tzimiropoulos, L. Shen, and M. Valstar. Self-supervised learning of person-specific facial dynamics for automatic personality recognition. *IEEE Transactions on Affective Computing*, 2021.

[29] S. Sonlu, U. Güdükbay, and F. Durupinar. A conversational agent framework with multi-modal personality expression. *ACM Transactions on Graphics*, 40(1):Article no. 7, 16 pages, 2021.

[30] H.-Y. Suen, K.-E. Hung, and C.-L. Lin. TensorFlow-based automatic personality recognition used in asynchronous video interviews. *IEEE Access*, 7:61018–61023, 2019.

[31] A. Truong, H. Boujut, and T. Zaharia. Laban descriptors for gesture recognition and emotional analysis. *The Visual Computer*, 32:83–98, 2016.

[32] Y. Wang, J. E. F. Tree, M. Walker, and M. Neff. Assessing the impact of hand motion on virtual character personality. *ACM Transactions on Applied Perception*, 13(2):Article no. 9, 23 pages, 2016.