The quest for believability: exploring FACS adaptations for emotion facial expressions in virtual humans

The Quest for Believability: Exploring FACS Adaptations for Emotion Facial Expressions in Virtual Humans

Stefano Calzolari, Francesco Strada, Andrea Bottino
Politecnico di Torino, Control and Computer Engineering Department
Torino, 10129, Italy. Email(s): {name.surname}@polito.it

***Abstract*—In interactive computer graphics, Facial Action Coding System (FACS) has been adapted to enhance the emotional expressiveness of Virtual Humans (VHs) by associating certain Action Units (AUs) with corresponding facial blendshapes. In this way, animators can (theoretically) recreate any human emotion on a VH's face with precision and flexibility. However, conveying realistic and believable emotional expressions with this approach comes with some challenges. In particular, given a set of AUs representing a particular emotion, it is not straightforward to define the correct set of blendshape weights that can render the same realistic and believable emotion on all VHs, as even small differences in weight values can drastically change the perceived emotion. This complexity raises several critical questions such as: is there for each emotion a *universal set* of blendshape weights that can effectively convey that emotion across all VHs? How can this set be found? If such a universal set proves elusive, can optimal combinations be identified for specific subgroups based on specific facial features such as men and women? Answering these questions is critical to understanding the general applicability of FACS-based facial emotion coding, which allows designers and animators to easily develop VHs that are able to interact with users in a way that is both emotionally rich and authentic. This paper explores these issues through a preliminary investigation aimed at defining realistic representations of *happiness*.**

***Index Terms*—FACS coding, emotional expressiveness, virtual humans, facial blendshapes**

I. Introduction

Virtual Humans (VH) are virtual characters that act and look like humans in a virtual environment [1]. Their application spans various fields, including medical education [2], e-learning and gaming, and the roles assumed range from patient simulators [3], [4] to interactive characters [5], [6]. This versatility underlines their importance in creating immersive, interactive environments that enable complex simulations and interactions [1].

The effectiveness of VHs in simulating human-like behavior and emotional expression is critical for user interaction and engagement [7], [1].The representation of believable emotions (where "believable" here means that they are truthful in the eyes of the user [7]) enriches the user experience and contributes to the development of compelling and immersive experiences [8], [9]. However, due to the complexity of emotions and the influence of various factors on their subjective perception (such as cultural background, personal experiences and contextual information), achieving authentic emotional expressions is a challenge. This requires advanced technical methods for accurate representation with VHs, especially for facial expressions, the primary medium for emotion communication [10], [8].

The Facial Action Coding System (FACS) is an important tool for understanding human facial expressions [11]. FACS decomposes each facial expression into combinations of Action Units (AUs), each representing a specific element of the expression and associated with a quantifiable intensity $I \in [0, 1]$. By identifying and manipulating the intensity of these AUs, FACS provides a methodical way to accurately analyze and recreate a wide range of emotional expressions.

In the context of VHs, FACS becomes a suitable option for the representation of emotions. This process involves translating FACS codes into the weights of specific blendshapes, predefined 3D models of atomic facial movements that can be blended to varying degrees to achieve the desired emotion. When mapping AUs to blendshapes, an important reference for the creation of emotions that can be expressed by VHs is the result of psychological research that has identified six basic emotions (*anger*, *disgust*, *fear*, *joy*, *sadness* and *surprise*) that are universally and innately expressed by humans regardless of gender, ethnicity or other facial characteristics [12]. The coding of these emotions in the context of FACS is further elaborated in EMotional FACS (EMFACS) [13].

However, EMFACS outlines AU combinations for basic emotions without detailing intensity levels, posing a challenge in fine-tuning VH blendshape weights for generating believable emotions. Prior FACS applications in VHs [3], [14], [15] assumed its universal applicability across models, yet, to our knowledge, there is no evidence in the literature. As a matter of fact, slight blendshape adjustments might critically affect emotion perception, indicating a potential need for model-specific tuning (Fig. 1).

The nuances and gaps in EMFACS prompt researchers and practitioners to ask whether there is a universal set of blendshape weights for effectively conveying each emotion across all VHs, what methods are available to identify such a set, or whether it is possible to find optimal combinations for specific subgroups. Answering these questions is important to assess the broad applicability of FACS-based coding of facial emotions.

Our preliminary investigation aims to explore the feasibility of such a standardized approach for coding facial emotions

Fig. 1: Four examples of the same FACS encoding on different characters, 2 from MetaHuman (top) and 2 from Rocketbox (bottom).

in VHs. In doing so, we focus on identifying a universally applicable set of weighted AUs for the accurate representation of *happiness*. When a single combination is impractical, we try to find a subset of weights that effectively embody that emotion. The rationale of focusing on a single emotion in our study is due to the extensive resources required for a comparative analysis of multiple emotions, as explained below. This approach therefore allows us to evaluate the effectiveness and challenges of the methodology and to lay the foundation for future research on different emotional expressions.

The proposed method involves the analysis of 100 characters of different gender and resolution, each with nine different weighting combinations for the AUs associated to *happiness*, as well as a neutral expression. For each character, the images showing these variations were comparatively analyzed to rank the AU weightings from highest to lowest effectiveness in expressing *happiness*. The ranking was done with a group of volunteers who rated the pairs of images using a web application. Volunteers indicated which of the two images most convincingly expressed *happiness*, contributing to a voting score for each image. The final score reflects the frequency with which a particular image was chosen over another in these pairings. Finally, we used these scores to create a global ranking to determine the most effective AU encoding for each character and then analyze the universality of each encoding across different character sets.

The preliminary results show that some FACS encoding better represent the happiness emotion across the whole dataset but characters specific facial features have to be considered. In particular there seem to be significant differences across the two characters' source. This is probably due to the different level of visual fidelity of the sets.

II. Background

The use of FACS to enhance the emotional expressiveness of VHs has been unfolded in various studies. First attempts, such as [15], pioneered the integration of FACS with blendshape-based techniques to animate VHs and highlighted the technical feasibility of this approach. However, these early studies primarily demonstrated implementation methods without addressing the believability of the expressed emotions, as they assumed that the applicability of FACS to VHs reflected their effectiveness in humans. This assumption persisted in later studies [14], which aimed for more realistic simulations without critically evaluating the authenticity of the emotional expressions.

Similar conclusions were reached by [16], which provided a comprehensive overview of methods for representing facial expressiveness, including FACS, but did not question the basic assumption of their direct applicability to VHs. In parallel, the study on emotional expressions of VHs in [17] hinted at the potential for emotional contagion between users and VHs, but did not explicitly use FACS. This study highlights the diversity of human smiles and suggests a nuanced range of emotional expressions beyond the original scope of FACS.

The exploration of emotional contagion with a group of FACS-implemented VHs in [18] indicated the potential of FACS for emotional expressiveness. Nevertheless, questions about the specific application of FACS and the characteristics of the VHs used remain unanswered. Similarly, the investigation of pain expressions among VHs proposed in [3] revealed that higher intensities of AUs used to define the expression correlate with increased perception of pain expressed by the face, suggesting the applicability of FACS. However, this paper does not investigate whether these results can be generalized to other characters and emotions.

These studies show that a growing body of work demonstrates the technical integration of FACS for animating VHs, but there is still a significant gap in understanding the universality and believability of these expressions. The question of whether there is a universal set of AU weights that can effectively convey each basic emotion across all VHs remains unanswered. This gap emphasizes the need to further investigate how FACS can be adapted to account for the variability of facial features across different VHs and how these adaptations affect the perceived authenticity of emotional expressions.

Our work aims to address these gaps through an initial exploration of the potential of a standardized FACS application for different VH models to provide a methodological basis for believable emotional expressions.

III. Methods

In the following, we will use the term ***expression*** to refer to the vector of AU weights defining a facial expression. Rendered ***portraits*** were used to represent these encodings on a set of virtual characters, with each portrait representing a

Fig. 2: Examples of visual artifacts on a Rocketbox character rendered with all AUs' weights to 1.

pair of *(character, expression)*. The characters are of different gender, age and different 3D mesh resolutions.

The focus of our research is on capturing and analyzing user perceptions of these portraits, i.e. their ability to convey a realistic and believable representation of emotion, and then infer the generalization properties of these expressions across characters. In the following, we describe how the portrait dataset was created (Section III-A) and how the ratings and ranking of portraits and expressions were collected (Section III-B).

A. Portrait Database

The dataset for our research includes a collection of portraits of different characters expressing *happiness* with different expressions.

The character dataset consists of 100 models from two different sets: 50 from the Rocketbox [19] set and 50 from the MetaHumans [20] library. The models from both libraries are equipped with a set of facial blendshapes associated one-to-one with FACS AUs. We chose these character libraries because they differ significantly in resolution and quality. Rocketbox characters have a lower-resolution mesh, making them less detailed but also requiring fewer computational resources. MetaHumans are known for their high-quality and high-resolution representations. They are ideal for applications that require human-like characters but they poses a challenge in terms of computational requirements, which can limit their use in real-time applications and in large-scale simulations that require numerous characters. The selected characters are further subdivided by gender (27 male and 23 females for Rocketbox, 25 males and 25 females for MetaHumans) and cover different age groups and ethnicities. In the following, the term ***character group*** refers to the subgroups defined in the whole set of characters.

According to EMFACS, *happiness* involves AU6 (cheek raiser) and AU12 (lip corner puller) [13]. To investigate how different combinations of blendshape weights affect the representation of *happiness* on VHs, we chose three different and discrete intensity levels for each AU, namely 0.4, 0.675 and 0.95. Values below 0.4 were excluded as they produced expressions that were too similar to a neutral state, while the maximum intensity of 1 was omitted due to the generation of rendering artifacts on some low-resolution characters used in our experiments (Fig. 2). These numbers result in nine different combinations of AU intensities, to which we added a neutral face encoding with all AU weights set to 0. This serves as a control value (i.e. to check if the ratings are consistent, as the neutral expression should be rated lower than any other expression).

To generate portraits, we developed a rendering pipeline in Unreal Engine 5 that receives as input a *(character, expression)* pair and outputs the corresponding portrait rendered under uniform conditions of lighting, camera parameters and background environment. The data set includes 1000 portraits (i.e. 100 different characters representing 10 expressions).

B. Voting App

The ranking procedure in our study was designed to evaluate and compare expressions according to their ability to portray the emotion realistically and believably, using a method similar to a round robin tournament. In this approach, each expression is pitted against all other expressions of the same character in a series of comparisons. Each time an expression A is favored over another expression B, this is recorded as a win for expression A and a loss for expression B. The data collected is the number of wins and losses for each expression of each character.

To derive a meaningful and comparable metric from these data, we first compute a normalized ranking score $r_{c,i}$ for each character c and each expression i. This value is computed as follows:

$$r_{c,i} = \frac{w_{c,i} - l_{c,i}}{w_{c,i} + l_{c,i}} \tag{1}$$

where $w_{c,i}$ and $l_{c,i}$ are the number of wins and losses for expression i of character c, and the score is a value in $[0, 1]$ independently of the actual number of votes received.

The accuracy $A_{i,g}$ of an expression i is calculated based on the top-k accuracy for the expression i relative to the group g. This is determined by the number of characters for which the expression i ranks within the top k positions of the ranking list, divided by the total number of characters in the group. This metric aims to measure how universally effective an expression is within a given character group by quantifying its ability to be ranked highly (i.e. be considered very realistic) among the various characters within that group.

The total number of comparisons an individual character must make to complete a tournament round can be computed as follows. If p is the number of expressions of a character, the number of "matches" involving all character expressions in a single tournament is $T = p \cdot (p-1)/2$. With n characters, a total of $n \cdot T$ votes must therefore be collected from users. It is obvious that the more tournaments are completed, the more accurate the expression ranking will be.

The voting app presents two clickable images and asks the user to select the portrait that best matches the *happiness* emotion and then confirm their choice by clicking the *NEXT* button (Fig. 3). The confirmation step has been added so that the user can check their choice before casting their vote. Once the vote has been cast, the user is presented with another pair

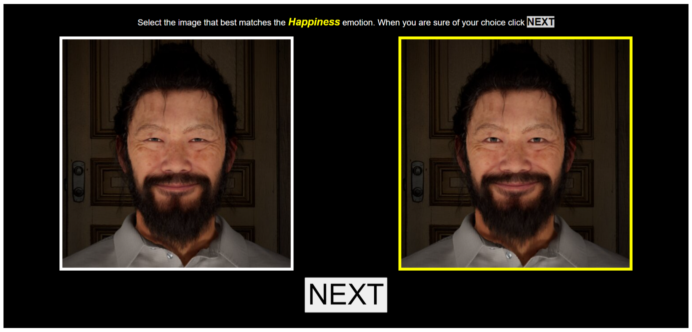

Fig. 3: A sample view of the web app.

of portraits selected to guarantee an even distribution of their comparisons.

IV. Experimental results

In this section, we present the experiment we conducted to evaluate the applicability of FACS encoding for the expression of emotions in VHs and discuss the results. The rest of the section is organized as follows: subsection IV-A presents in details the experimental protocol used to conduct the data analysis and subsequent conclusions drawing. Subsection IV-B analyzes the results by progressively deepen the focus from the whole character dataset to its specific characters' groups.

A. Experimental protocol

The study involved 169 participants who cast approximately 9,000 votes, corresponding to two full rounds of portrait comparisons since for *happiness* we have $p = 10$, $T = 45$ and $n = 100$ and a total of 4,500 comparisons for a single tournament round. After collecting user preferences, we calculated the normalized ranking scores $r_{c,i}$ for each expression i of each character c and the corresponding accuracy of each expression across different VH groups. Finally, we analyzed these accuracies. Given the limited diversity of ethnicities and the lack of reference values for the age of the characters available, we identified groups characterized by two variables: Gender (male, M, female, F) and Resolution (low resolution, L, for Rocketbox, and high resolution H, for MetaHumans).

B. Results

We begin by analyzing the overall results, i.e. the results obtained with a single group (ALL) in which all characters (regardless of their gender and resolution) are grouped together. Before starting the discussion of the results, we present the ablation experiments for the hyperparameter k, i.e. the number of top expressions considered in our accuracy calculations. Table I shows the accuracy for the ALL group with $k \in [1, 3]$. It can be seen that all k values lead to similar relative rankings of the expressions. This consistency led us to the decision to use $k = 3$ for all our analyses based on the following two considerations. First, choosing $k = 3$ allows us to map a broader range of user preferences by identifying expressions that, while not always the first choice, are highly competitive in terms of believability. Second, a higher k value increases the robustness of the analysis by mitigating individual perceptual variability, ensuring more stable and reliable conclusions for a diverse user base.

Fig. 4: Examples of top scoring expressions for different characters (top row: MetaHuman; bottom row: Rocketbox)

Overall Results. From the results in Table I, we can draw the following observations. First, expressions can be divided into three distinct groups based on their accuracy: low ($A_{i,ALL}) \in [0, 0.1]$), medium ($A_{i,ALL}) \in [0.1, 0.6]$) and high ($A_{i,ALL}) \in [0.6, 1]$). The existence of these groups suggests that the expressions in the high accuracy group are generally more effective at conveying *happiness*, while other expressions do so only to a lesser extent or not at all (low and medium accuracy groups). Second, among the three expressions with the highest accuracy (w7, w8, w9), none clearly stands out as the most effective expression. This observation suggests that (i) there may not be a one-size-fits-all solution for conveying *happiness* in VHs, and (ii) there may be a universal set of expressions that can believably convey *happiness* (Fig. 4).

In all the three high ranking expressions the strong emphasis on the action of smiling (AU12, "lip-pulling") is central to the communication of *happiness*. In fact, the credibility of the *happiness* expression decreases when the intensity of AU12 is reduced. Interestingly, the expression w9, which maximizes the weighting of AU12 and AU6 ("cheek raiser") among the three, does not outperform w7 and w8 in terms of accuracy (Table I). These nuanced differences in the AU weights of the top expressions reveal a critical balance between the intensity of the emotional expression and its believability.

Gender. When examining the most effective expressions for conveying *happiness* across genders (w7, w8 and w9), the minimal differences in accuracy per expression between male and female VHs – with a maximum delta of only 0.08 – emphasize that these expressions are gender agnostic. This "gender neutrality" simplifies the design process by allowing creators to apply a consistent set of expressions to a variety of character models without the need to adjust for gender. Although differences in accuracy can be observed for other expressions (particularly w4 and w5), their lower overall effectiveness negates the relevance of these differences in developing an overall strategy for designing emotionally expressive VHs.

Resolution. The comparison between low and high-resolution models reveals a significant impact of visual fidelity in the believability of emotional expressions. The advanced

TABLE I: Accuracy of Happiness expression in different groups of characters, categorized by overall accuracy. The ALL columns focus on the variations at $k = 1$, $k = 2$ and $k = 3$ across all characters, while the following columns present the accuracy for characters grouped by gender, resolution and specific subgroups. The best results in **bold**, the second ranked results in underlined, the third ranked results in *ITALIC*.

	AUs weights		ALL			Gender		Resolution		Subgroups			
expression	**AU12**	**AU6**	**k = 1**	**k = 2**	**k = 3**	**Male**	**Female**	**Low**	**High**	**LM**	**LF**	**HM**	**HF**
w0	0.000	0.000	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
w1	0.400	0.400	0.00	0.00	0.01	0.02	0.00	0.02	0.00	0.04	0.00	0.00	0.00
w2	0.400	0.675	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
w3	0.400	0.950	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
w4	0.675	0.400	0.01	0.13	0.25	0.31	0.17	0.41	0.10	0.57	0.26	0.08	0.12
w5	0.675	0.675	0.04	0.10	0.20	0.17	0.25	0.35	0.06	0.26	0.43	0.08	0.04
w6	0.675	0.950	0.03	0.07	0.19	0.23	0.19	0.04	0.32	0.09	0.00	0.36	0.28
w7	0.950	0.400	**0.43**	**0.66**	**0.83**	0.75	0.83	**0.91**	*0.76*	**0.83**	**1.00**	*0.68*	0.84
w8	0.950	0.675	*0.24*	0.57	0.81	**0.85**	**0.85**	0.74	**0.88**	0.74	0.74	**0.96**	*0.80*
w9	0.950	0.950	0.25	*0.47*	*0.71*	*0.67*	*0.71*	*0.52*	0.88	*0.48*	*0.57*	0.84	**0.92**

rendering features of MetaHuman models, including detailed skin textures and dynamic wrinkle maps for each character (Fig. 5) enhance the realism of facial expressions. Conversely, Rocketbox characters rely on simpler geometric deformations to represent facial expressions. Consequently, it can be observed that the accuracy decreases as the weight of AU6 increases. This can be explained by the fact that the resulting geometric deformation of the cheeks may appear exaggerated or unnatural due to the lack of fine-grained visual cues such as wrinkles or subtle skin stretching, ultimately reducing the believability of the expression. Moreover, the lack of detail in the facial mesh may result in different FACS encodings producing visually similar, if not almost identical, results (Fig. 6), blurring to some extent the differences between various emotional intensities and leading to a homogenization of perceived expressions. In general, the results suggest that expressions w7, w8 and w9 are generally effective for both types of resolution, albeit with varying degrees of believability influenced by the graphical accuracy of the models. The challenge is therefore to adapt the design of the facial expressions to the possibilities and limitations of the character's resolution.

Subgroups. The detailed analysis of the subgroups (FL, ML, FH, MH) does not provide any new insights beyond those gained from the discussions on gender and resolution confirming that gender results in minimal differences in expression effectiveness, with the resolution dimension having a stronger influence.

C. Limitations

While our work provides valuable insights into the use of EMFACS coding to convey realistic emotions in VHs, it presents some limitations that create opportunities for future research. First, our preliminary results focus on *happiness* and their generalizability to other emotions is unknown. Furthermore, the group of characters should be expanded to analyze the effects of variables such as ethnicity, age, and other demographic characteristics, which was not possible with our character pool. Second, the method used in this study does not scale well with larger numbers of AUs and discrete weight values, as this exponentially increases the number of portraits to compare. Future research could benefit from using active sampling methods [21], [22], [23] to reduce the number of pairwise comparisons required. Third, the current analysis only includes static images and thus neglects the dynamic dimension of emotion expressions, which contributes significantly to the perception of their authenticity and intensity. Fourth, our group of volunteers consisted mainly of students and researchers from our university. In this sample, perceptions of the believability of emotional expressions may be influenced by cultural and demographic factors [24], [25]. Therefore, future studies should diversify the group of subjects to reduce potential biases related to these factors. Finally, this study relies only on explicit measurements, i.e. the participants' choice between two images), and to gain a more nuanced comprehension of participants' perceptions, future studies may incorporate implicit metrics such as monitoring pupil dilata-

Fig. 5: Example of MetaHuman's wrinkle maps variation across two different expressions (Left: w1; Right: w9)

Fig. 6: Examples of different expressions (from left to right, w7, w8, w9 and the character's optimal expression w4) resulting in perceptive similar portraits when applied to a low-resolution character.

tion, skin conductance or heart rate.

V. Conclusions

This research is a pilot study to investigate the use of FACS for the realistic expression of facial emotions with VHs. The results hints to the existence of a universal set of expressions that are effective with characters of different gender and resolution. Although the scope of the study is limited, the same research framework can be used to analyze different emotions and characters. Future work will focus on extending our research along the lines described in the Limitations (Section IV-C).

References

[1] B. Lugrin, C. Pelachaud, and D. Traum, Eds., *The Handbook on Socially Interactive Agents: 20 Years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 1: Methods, Behavior, Cognition*, 1st ed. New York, NY, USA: Association for Computing Machinery, 2021, vol. 37.

[2] F. Strada, A. Bottino, F. Lamberti, G. Mormando, and P. L. Ingrassia, "Holo-blsd–a holographic tool for self-training and self-evaluation of emergency response skills," *IEEE Transactions on Emerging Topics in Computing*, vol. 9, no. 3, pp. 1581–1595, 2019.

[3] G. Carbonell and J. Schild, "Conceptual Design of Emotional and Pain Expressions of a Virtual Patient in a Virtual Reality Training for Paramedics," in *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. Christchurch, New Zealand: IEEE, Mar. 2022, pp. 222–226.

[4] A. Robb, R. Kopper, R. Ambani, F. Qayyum, D. Lind, L.-M. Su, and B. Lok, "Leveraging Virtual Humans to Effectively Prepare Learners for Stressful Interpersonal Experiences," *IEEE Trans. Visual. Comput. Graphics*, vol. 19, no. 4, pp. 662–670, Apr. 2013.

[5] E. Battegazzorre, F. Strada, L. De Francesco, A. Mazza, O. Dal Monte, and A. Bottino, "Influencer based virtual reality stroop task," *IEEE Access*, 2023.

[6] A. Hartholt, S. Mozgai, E. Fast, M. Liewer, A. Reilly, W. Whitcup, and A. S. Rizzo, "Virtual Humans in Augmented Reality: A First Step towards Real-World Embedded Virtual Roleplayers," in *Proceedings of the 7th International Conference on Human-Agent Interaction*. Kyoto Japan: ACM, Sep. 2019, pp. 205–207.

[7] J. Bates *et al.*, "The role of emotion in believable agents," *Communications of the ACM*, vol. 37, no. 7, pp. 122–125, 1994.

[8] R. Hortensius, F. Hekele, and E. S. Cross, "The Perception of Emotion in Artificial Agents," *IEEE Trans. Cogn. Dev. Syst.*, vol. 10, no. 4, pp. 852–864, Dec. 2018.

[9] A. Chirico and A. Gaggioli, "When virtual feels real: Comparing emotional responses and presence in virtual and natural environments," *Cyberpsychology, Behavior, and Social Networking*, vol. 22, no. 3, p. 220–226, Mar. 2019.

[10] P. M. Niedenthal and M. Brauer, "Social Functionality of Human Emotion," *Annu. Rev. Psychol.*, vol. 63, no. 1, pp. 259–285, Jan. 2012.

[11] P. Ekman and W. V. Friesen, "Facial action coding system," *Environmental Psychology & Nonverbal Behavior*, 1978.

[12] P. Ekman *et al.*, "Basic emotions," *Handbook of cognition and emotion*, vol. 98, no. 45-60, p. 16, 1999.

[13] W. V. Friesen, P. Ekman *et al.*, "Emfacs-7: Emotional facial action coding system," *Unpublished manuscript, University of California at San Francisco*, vol. 2, no. 36, p. 1, 1983.

[14] Y. Tisserand, R. Aylett, M. Mortillaro, and D. Rudrauf, "Real-time simulation of virtual humans' emotional facial expressions, harnessing autonomic physiological and musculoskeletal control," in *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*. Virtual Event Scotland UK: ACM, Oct. 2020, pp. 1–8.

[15] M. H. Alkawaz, D. Mohamad, A. H. Basori, and T. Saba, "Blend Shape Interpolation and FACS for Realistic Avatar," *3D Res*, vol. 6, no. 1, p. 6, Mar. 2015.

[16] M. S. M. Rahim, A. H. Basori, S. M. Saadi, A. E. Rad, and A. Rehman, "Emotional Facial Expression and Tears Simulation: An Analysis & Comparison of Current Approaches," *3D Res*, vol. 6, no. 2, p. 14, Jun. 2015.

[17] J. Tsai, E. Bowring, S. Marsella, W. Wood, and M. Tambe, "A study of emotional contagion with virtual characters," in *Intelligent Virtual Agents: 12th International Conference, IVA 2012, Santa Cruz, CA, USA, September, 12-14, 2012. Proceedings 12*. Springer, 2012, pp. 81–88.

[18] M. Volonte, Y.-C. Hsu, K.-Y. Liu, J. P. Mazer, S.-K. Wong, and S. V. Babu, "Effects of Interacting with a Crowd of Emotional Virtual Humans on Users' Affective and Non-Verbal Behaviors," in *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. Atlanta, GA, USA: IEEE, Mar. 2020, pp. 293–302.

[19] M. Gonzalez-Franco, E. Ofek, Y. Pan, A. Antley, A. Steed, B. Spanlang, A. Maselli, D. Banakou, N. Pelechano, S. Orts-Escolano, V. Orvalho, L. Trutoiu, M. Wojcik, M. V. Sanchez-Vives, J. Bailenson, M. Slater, and J. Lanier, "The rocketbox library and the utility of freely available rigged avatars," *Frontiers in Virtual Reality*, vol. 1, Nov. 2020.

[20] E. Games, "MetaHuman, Realistic Person Creator - Unreal Engine," https://www.unrealengine.com/en-US/metahuman, [Accessed 02-02-2024].

[21] A. Mikhailiuk, C. Wilmot, M. Perez-Ortiz, D. Yue, and R. Mantiuk, "Active Sampling for Pairwise Comparisons via Approximate Message Passing and Information Gain Maximization," Apr. 2020, arXiv:2004.05691 [cs, stat].

[22] J. Li, R. Mantiuk, J. Wang, S. Ling, and P. L. Callet, "Hybrid-MST: A Hybrid Active Sampling Strategy for Pairwise Preference Aggregation," *Advances in neural information processing systems*, vol. 31, 2018.

[23] S. Ling, J. Li, A. F. Perrin, Z. Li, L. Krasula, and P. L. Callet, "Strategy for Boosting Pair Comparison and Improving Quality Assessment Accuracy," Oct. 2020, arXiv:2010.00370 [cs].

[24] J. B. Engelmann and M. Pogosyan, "Emotion perception across cultures: the role of cognitive mechanisms," *Frontiers in Psychology*, vol. 4, 2013.

[25] N. Haug, L. Geyrhofer, A. Londei, E. Dervic, A. Desvars-Larrive, V. Loreto, B. Pinior, S. Thurner, and P. Klimek, "Ranking the effectiveness of worldwide covid-19 government interventions," *Nature Human Behaviour*, vol. 4, no. 12, p. 1303–1312, Nov. 2020.