Guidelines for Modeling Expressivity in Virtual Patients for Healthcare Training

Wesley Ferreira de Ferreira, Luciana Nedel
AKCIT Lab, Institute of Informatics
Federal University of Rio Grande do Sul (UFRGS), Brazil
{wesley.ferreira, nedel}@inf.ufrgs.br

***Abstract*—Virtual patients are increasingly being used in healthcare education to support the development of communication and social skills. However, their effectiveness depends on the integration of expressive realism across verbal and non-verbal communication modalities. This paper proposes a set of guidelines for modeling emotional expressivity in virtual patients by combining facial expressions, body posture, and vocal prosody. Grounded in established affective science literature—including Ekman's basic emotions, the Facial Action Coding System (FACS), the Body Action and Posture (BAP) framework, and GeMAPS acoustic parameters—these guidelines synthesize expressive features for six core emotions: joy, sadness, anger, fear, disgust, and surprise. Our goal is to support the development of emotionally expressive virtual patients for medical simulations and virtual reality training.**

I. Introduction

The development of Virtual Humans is an important topic in simulations for healthcare training, especially in the increasing necessity of healthcare professionals forming and practicing before the contact with real patients [1]. Nowadays, standardized patients played by professional actors, are employed to foster the soft skills of healthcare professionals. However, this approach presents challenges in scalability and is associated with significant costs [2].

The use of simulations to improve social and behavioral skills in healthcare learning has become a relevant topic in recent years, and the new artificial intelligence (AI) tools have expanded the capacity of Virtual Humans to behave like real patients [3].

The modeling of virtual humans is challenging due to the complexity of real human behavior, especially in healthcare, where the nuances are important [4]. In this context, modeling virtual patients requires special attention in verbal and non-verbal aspects of communication and behavior. As well, the impact on the perceived emotion, and understanding of the specificity of modeling virtual patients against other virtual humans is a relevant and open area [5].

This work is part of a larger study that investigates the requirements for developing Virtual Patients to support healthcare students in learning social and communication skills, beginning with verbal communication [6]. The next phase of this study will focus on speech, facial, and body expressivity, starting with the design of a reference framework that captures key aspects of expressivity and the behavior of virtual patients.

Fig. 1. Representation of surprise, modeled in Unity UMA2, highlighting body posture and facial expression derived from the BAP and FACS frameworks.

II. Theoretical Foundation

A. Facial Expression

The design of a reference framework for the key aspects of expressivity and behavior in virtual patients must be grounded in literature from healthcare, human behavior, and computer science.

Paul Ekman conducted some of the most influential and well-founded studies on human emotions and facial expressivity. His work ranges from defending the existence of basic and universal emotions [7], to modeling emotions through facial expressions [8], and developing systematic guidelines for research in this area [9].

Another highly relevant work in the field of expression and emotions is the Geneva multimodal emotion portrayal (GEMEP) corpus [10].This work introduced a dataset of audio and video recordings of actors expressing 17 different emotions, and tested whether these emotions could be accurately recognized by others. The GEMEP work introduces the set of emotions listed in Table I.

Fig. 2. Illustration of three emotional states (fear, joy, and sadness) modeled in Unity's Multipurpose Avatar System (UMA2), applying the proposed guidelines for facial expressions and body posture integration.

TABLE I
GEMEP EMOTIONS SET.

Anger	Irritation	Contempt	Disgust
Sadness	Despair	Fear	Anxiety
Surprise	Interest	Admiration	Tenderness
Relief	Pleasure	Amusement	Joy
Pride			

The Facial Action Coding System (FACS) [8], proposed by Paul Ekman, is a widely accepted methodology for facial coding, offering a technically based and scientifically rigorous approach to measuring emotional expressions. This remains a reference in the study of emotions in the fields of social behavior, psychiatry, and artificial intelligence.

B. Body Action and Posture

The GEMEP corpus also include body movements and related them to the emotions subset in Table I. This work, developed to complement the FACS system [11], encoded 67 statistically relevant movements.

Subsequently, the same group presented a second study, reducing the set of body movements from 67 to 49 by applying principal component analysis (PCA) and identifying the emotion associated with each principal component, then segmenting the most relevant body movements for each emotion [12].

Considering the relevance of FACS, Dael's second work [12], Ekman's basic emotions [7], and the lack of exploratory studies using GEMEP directly to animation or modeling, we selected a subset of emotions and their associated facial expressions and body movements to propose clear and concise guidelines for animating virtual patients.

C. Speech Emotion Correlates

As important as facial and body expressivity, is speech expressivity. The Geneva minimalistic acoustic parameter set (GeMAPS) study [13], also using the GEMEP, evaluates the relevance and characteristics of speech and relates them to the emotions conveyed by the voice. The study indentified 18 vocal key aspects that complete and correct the transmission of emotions. Based on the well-known Ekman Emotions, we have extended our proposed guidelines to include speech characteristics.

III. Guidelines Proposal

Here we describe our proposed guidelines.

A. Facial Expression Guidelines based on FACS

The proposed guidelines for facial expression are the easiest among those presented. Based directly on FACS [8], we created a table relating each emotion with their facial expressions action units, and their descriptions. Table II presents these guidelines, while Figure 2 shows an illustration of the applied guidelines.

TABLE II
ACTION UNITS OF EACH EMOTION ACCORDING TO FACS FOR EACH EMOTION

Emotion	FACS Action Units Description
Joy	AU6: Cheek Raiser
	AU12: Lip Corner Puller (smile)
Sadness	AU1: Inner Brow Raiser
	AU4: Brow Lowerer
	AU15: Lip Corner Depressor
Anger	AU4: Brow Lowerer
	AU5: Upper Lid Raiser (wide eyes)
	AU7: Lid Tightener
	AU23: Lip Tightener
Fear	AU1+AU2: Inner and Outer Brow Raiser
	AU4: Brow Lowerer
	AU5: Upper Lid Raiser
	AU20: Lip Stretcher
	AU26: Jaw Drop
Surprise	AU1+2: Inner and Outer Brow Raiser
	AU5: Upper Lid Raiser
	AU26: Jaw Drop
Disgust	AU9: Nose Wrinkler
	AU10: Upper Lip Raiser
	AU16: Lower Lip Depressor

TABLE III
CORRELATION BETWEEN EKMAN EMOTIONS AND DAEL'S BEHAVIOR FACTORS

Cluster	Behavior Factor	Value	Ekman Emotion	Prop. in Emotion	Source / Note
1	Symmetrical up-down arm action (reduced)	–1.67	Anger	30%	Directly from BAP
1	Forward whole body movement (reduced)	–1.00	Anger	30%	Directly from BAP
1	Knee movement (reduced)	–0.95	Anger	30%	Directly from BAP
3	Head tilted lateral averted	–0.67	Fear	30%	Directly from BAP
3	Backward whole body movement	+1.46	Fear	30%	Directly from BAP
3	Backward whole body movement	-	Disgust	-	Mapped by behavioral correspondence
3	Head tilted lateral averted	-	Disgust	-	Mapped by behavioral correspondence
4	Illustrative action	+0.88	Anger	100%	Directly from BAP
4	Forward whole body movement	+1.96	Anger	100%	Directly from BAP
6	Arms held in front (reduced)	-	Disgust	-	Mapped by behavioral correspondence
6	Asymmetrical arm action (reduced)	-	Disgust	-	Mapped by behavioral correspondence
8	Symmetrical up-down arm action	+1.17	Joy	70%	Directly from BAP
8	Symmetrical up-down arm action	-	Surprise	-	Mapped by behavioral correspondence
9	Head tilted up averted with closed eyes	+2.07	Joy	40%	Directly from BAP
9	Head tilted up averted with closed eyes	-	Surprise	-	Mapped by behavioral correspondence
9	Asymmetrical arm action	+1.11	Joy	40%	Directly from BAP
10	Arms held in front (reduced)	–1.08	Sadness	60%	Directly from BAP
11	Head tilted lateral averted (reduced)	–1.04	Joy	100%	Directly from BAP
12	Knee movement	+1.77	Fear	60%	Directly from BAP
12	Knee movement	+1.77	Joy	30%	Directly from BAP
12	Knee movement	-	Surprise	-	Mapped by behavioral correspondence

TABLE IV
POSTURAL EMOTION GUIDELINES

Emotion	BAP Factor	Interpretation
Anger	Symmetrical up-down arm action (reduced)	Reduced symmetrical arm movement, indicating muscular tension and restrained aggression.
	Forward whole body movement	Forward torso lean, associated with readiness to confront or attack directly.
	Knee movement (reduced)	Limited leg movement, reinforcing bodily rigidity typical of controlled anger.
	Illustrative action	Emphatic, directional gestures used to convey assertiveness or dominant intent.
Fear	Head tilted lateral averted	Head turned and averted from stimulus, indicating visual avoidance and withdrawal.
	Backward whole body movement	Bodily retreat, expressing evasion or an attempt to distance from threat.
	Knee movement	Sudden leg activation, signaling readiness to flee or physical tension.
Joy	Symmetrical up-down arm action	Broad, rhythmic arm movements typical of enthusiasm and positive emotional engagement.
	Head tilted up with eyes closed	Head raised and eyes closed in an expression of pleasure, relaxation, or admiration.
	Asymmetrical arm action	Spontaneous, relaxed arm gestures often associated with playfulness or expressiveness.
	Knee movement	Energetic leg involvement reflecting joyful bodily dynamism.
	Head tilted lateral averted (reduced)	Lack of head or gaze aversion, indicating an open and receptive posture.
Sadness	Arms held in front (reduced)	Arms close to the body or lowered, suggesting emotional withdrawal and physical containment.
Disgust	Backward whole body movement (inferred)	Bodily retreat interpreted as rejection of the stimulus, similar to an aversive response.
	Head tilted lateral averted (inferred)	Head turned away in a gesture of visual rejection or disengagement.
	Arms held in front (reduced) (inferred)	Defensive or symbolically dismissive posture, expressing repulsion.
	Asymmetrical arm action (reduced) (inf.)	Inhibited gesture and physical discomfort, suggesting emotional blocking and implicit refusal.
Surprise	Symmetrical up-down arm action (inf.)	Sudden, wide arm movements as a reflexive response to an unexpected stimulus.
	Backward whole body movement (inf.)	Automatic bodily recoil in response to sudden input, associated with startle reaction.
	Head tilted up with eyes closed (inf.)	Head raised and eyes closed in an expression of awe or intense astonishment.
	Knee movement (inferred)	Rapid leg activation as an impulsive physical reaction to surprise.

B. Body Movements Guidelines based on BAP

Based on Dael's two works [11], [12], we produced a table presenting each emotion and its respective related relevant movements. The development of guidelines was based mainly on the second work [12], where they classify the 16 most relevant factors using PCA, which represent 73% of the variance. Table III correlates the emotional expressions and the cluster analysis presented in Dael's work [12].

Surprise and disgust are not presented in Dael's second work. To overcome this absence, we used a cross-relation with other references [14], [15] and mapped it to Dael's work.

Table IV presents the proposed guidelines based on the correlation of each emotion behavioral fact presented in Table III and the works from Dael et al. [11], [12], Coulson [14], and Meijer [15]. Figure 1 presents an illustration of the fear emotion.

C. Voice Parameters Guidelines based on GeMAPS

We applied directly the GeMAPS [13] to the RAVDESS [16], an open-source dataset similar to the harder-to-access GEMEP. Table V summarizes the 6 most relevant parameters in the voice, and their perceptual correlation with the perceived emotion compared to the neutral one.

TABLE V
PERCENTAGE CHANGE RELATIVE TO NEUTRAL BASELINE

Emotion	F0 mean (%)	F0 std (%)	Loudness (%)	Jitter (%)	Shimmer (%)	HNR (%)
Anger	20.7	24.6	365.5	46.8	1.7	-17.8
Disgust	5.9	33.0	71.3	47.0	2.0	-20.5
Fear	25.3	-7.6	208.0	28.3	-3.0	8.1
Sadness	7.7	2.2	34.3	30.3	1.9	2.0
Surprise	18.7	44.6	98.8	87.2	7.3	-15.9
Joy	18.4	12.5	165.1	22.4	-4.2	3.8

IV. CONCLUSION

This work presents a unified set of guidelines to model emotional expressivity in virtual patients, integrating facial expressions (FACS), body posture (BAP), and vocal cues (GeMAPS). By mapping these modalities to six core emotions, we provide a foundation for developing more expressive and realistic virtual patients for healthcare training. Future work includes validating these guidelines in user studies and expanding them to cover more complex affective states.

ACKNOWLEDGMENT

This work has been partially funded by the project SITE-VR: A Framework for Improving User Experiences in Immersive Training, supported by the Advanced Knowledge Center in Immersive Technologies (AKCIT), with financial resources from the PPI IoT/Manufatura 4.0 / PPI HardwareBR of the MCTI grant number 057/2023, signed with EMBRAPII. Thanks are also due to Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) through grants 421962/2023-2, 313108/2025-0, and to Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

REFERENCES

[1] D. A. Cook, P. J. Erwin, and M. M. Triola, "Computerized virtual patients in health professions education: a systematic review and meta-analysis," *Academic Medicine*, vol. 85, no. 10, pp. 1589–1602, 2010.

[2] J. Cleland, K. Abe, and J.-J. Rethans, "The use of simulated patients in medical education: Amee guide no 42," *Medical Teacher*, vol. 31, no. 6, pp. 477–486, 2009.

[3] P. Kenny, T. D. Parsons, J. Gratch, A. Leuski, and A. Rizzo, "Virtual humans for assisted health care training," in *Studies in health technology and informatics*, vol. 125. IOS Press, 2007, pp. 133–138.

[4] A. Rizzo and S. T. Koenig, "Is clinical virtual reality ready for primetime?" *Neuropsychology*, vol. 31, no. 8, pp. 877–899, 2017.

[5] D. Ring, I. Barry, J. Totzke, and B. Lok, "Designing virtual patients to elicit social and empathic responses from medical students," *Frontiers in Virtual Reality*, vol. 3, p. 846470, 2022.

[6] W. Ferreira, M. D. Negrão, A. Maciel, R. P. Torchelsen, and L. Nedel, "Llm vs. human driven conversation: A user study in virtual reality medical simulation," in *Proceedings of the Symposium on Virtual and Augmented Reality 2025 (SVR)*. Brazil: IEEE, 2025, to appear.

[7] P. Ekman and W. V. Friesen, "Constants across cultures in the face and emotion," *Journal of Personality and Social Psychology*, vol. 17, no. 2, pp. 124–129, 1971.

[8] ——, *Facial Action Coding System: A technique for the measurement of facial movement*. Palo Alto, CA: Consulting Psychologists Press, 1978.

[9] P. Ekman, W. V. Friesen, and P. C. Ellsworth, *Emotion in the Human Face: Guidelines for Research and an Integration of Findings*, 3rd ed. Dallas, TX: Malor Books, 2013.

[10] T. Bänziger and K. R. Scherer, "Introducing the geneva multimodal emotion portrayal (gemep) corpus," in *A Blueprint for Affective Computing: A sourcebook and manual*, K. R. Scherer, T. Bänziger, and E. Roesch, Eds. Oxford, UK: Oxford University Press, 2010, pp. 271–294.

[11] N. Dael, M. Mortillaro, and K. R. Scherer, "The body action and posture coding system (bap): Development and reliability," *Journal of Nonverbal Behavior*, vol. 36, pp. 97–121, 6 2012.

[12] ——, "Emotion expression in body action and posture." *Emotion*, vol. 12, pp. 1085–1101, 2012. [Online]. Available: https://doi.apa.org/doi/10.1037/a0025737

[13] F. Eyben, K. R. Scherer, B. W. Schuller, J. Sundberg, E. André, C. Busso, L. Devillers, J. Epps, P. Laukka, S. S. Narayanan *et al.*, "The geneva minimalistic acoustic parameter set (gemaps) for voice research and affective computing," *IEEE Transactions on Affective Computing*, vol. 7, no. 2, pp. 190–202, 2016.

[14] M. Coulson, "Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence," *Journal of Nonverbal Behavior*, vol. 28, no. 2, pp. 117–139, 2004.

[15] M. de Meijer, "The contribution of general features of body movement to the attribution of emotions," *Journal of Nonverbal Behavior*, vol. 13, no. 4, pp. 247–268, 1989.

[16] S. R. Livingstone and F. A. Russo, "The ryerson audio-visual database of emotional speech and song (ravdess)," *PLOS ONE*, vol. 13, no. 5, p. e0196391, 2018. [Online]. Available: https://doi.org/10.1371/journal.pone.0196391