Perception-Driven Affect: Enhancing Virtual Human Realism Through Environmental Sound

Valentino Clerx, Olaf Janssen
Fontys University of Applied Sciences, ICT

Introduction

Sound strongly influences human emotion and behavior, yet most virtual humans remain insensitive to their acoustic environment. While modern language models generate realistic dialogue, they lack internal emotional continuity and perception-driven mood change, resulting in emotionally static agents.

Psychoacoustic research shows that sound characteristics such as loudness, sharpness, and roughness have measurable effects on affective state. Dimensional emotion models, particularly the Pleasure–Arousal–Dominance (PAD) model, allow these influences to be represented as continuous emotional shifts. This research presents a Psychoacoustic Perception System that maps acoustic parameters onto the PAD model, enabling environmentally grounded and dynamically evolving emotional behavior in virtual humans.

Methods

Emotional Mapping

A study by Balev (2025) presents a software framework for dynamic mood regulation in virtual empathic agents, aimed at producing more realistic and believable interactions. The framework integrates affective modeling using the Pleasure–Arousal–Dominance (PAD) model (Russell, 1980) to represent the agent's internal emotional state.

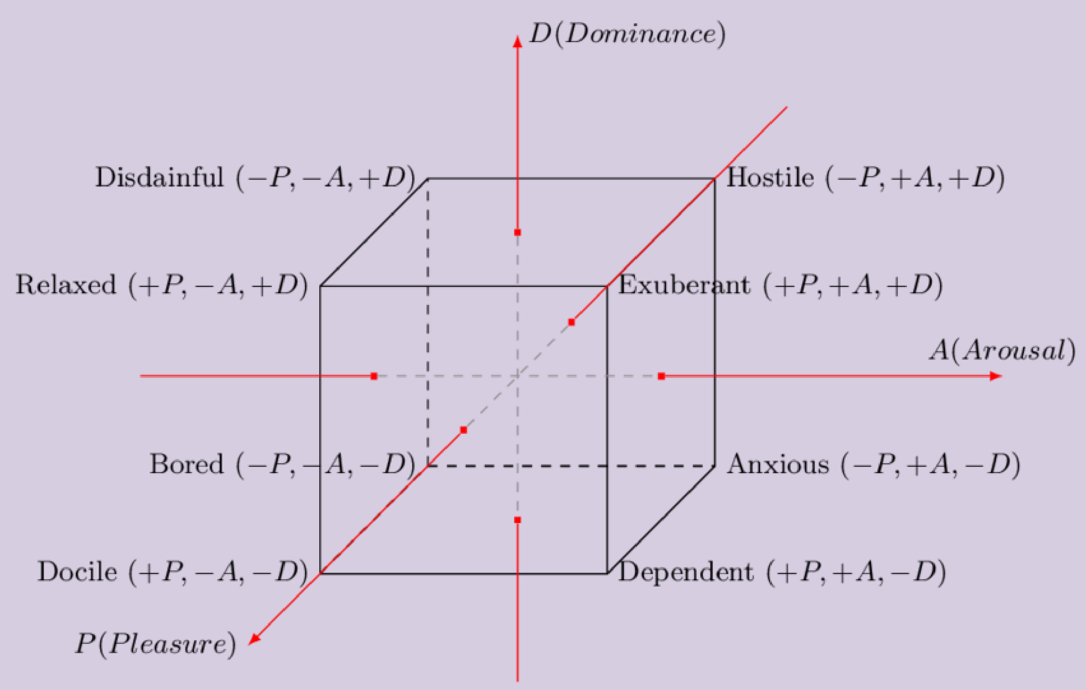

Figure 1 *Pleasantness, Arousal, Dominance Model (Mehrabian and Russell, 1974)*

Psychoacoustic Analysis

Psychoacoustics is the scientific discipline that investigates how sound is perceived by the human peripheral auditory system. The field examines the quantitative relationships between the objective physical properties of sound waves and the subjective psychological auditory sensations they evoke (Fastl and Zwicker, 2007; Moore, 2012).

Parameter	Concise Definition
Loudness (N)	Perceived intensity of sound (quiet to loud).
Sharpness (S)	High-frequency weighting of the sound spectrum.
Roughness (R)	Perception of rapid amplitude modulations (15–300 Hz).
Fluctuation (F)	Perception of slow temporal changes (up to 4 Hz).
Tonality (T)	Ratio of tonal components to broadband noise.
Impulsiveness (I)	Measure of sudden, transient sound content.

Table 1 *Psychoacoustic Parameters (Yang, 2025)*

Predicting Emotional Response

Yang (2025) provides a systematic analysis of how objective psychoacoustic parameters, representing auditory sensations, influence human emotional responses. By using a derivative of Russel's PAD model, using only the Pleasantness and Arousal axis, Yang was able to correlate psychoacoustic parameters to quadrants on these scales, leading to further correlation with certain emotions.

Sound Parameter(s)	Primary Emotional Dimension(s) Influenced	Emotional Effect / Correlation
Loudness (N), Roughness (R)	Relaxation–Stress axis (Quadrants II & IV); Overall Arousal	Higher values correlate with **stress, irritation, annoyance**
Impulsiveness (I), Sharpness (S)	Boredom–Excitement axis (Quadrants I & III)	Higher values perceived as **interesting, exciting**
Fluctuation (F), Tonality (T), Tonality Frequency (TF)	Arousal; Boredom–Excitement axis	Greater fluctuation/tonal content increases **activation**

Table 2 *Psychoacoustic influence on emotional quadrants (Yang, 2025)*

Computing Emotional Values

Using the psychoacoustic parameter values, linear regression can be used to calculate and predict emotional values (Yang, 2020). Using dedicated formulas like *Relaxing = -0.79·log(N_range) + 3.06*, emotional values can be calculated and plotted over the PAD derivative. Ultimately leading to integration into Balev's internal affection system model.

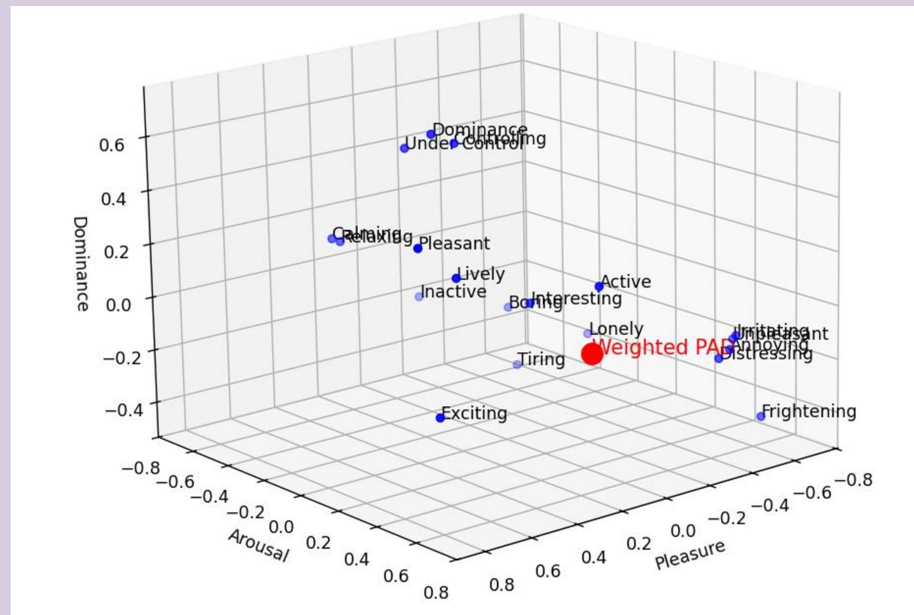

Figure 2 *Emotions plotted on the PAD model*

Measuring Perceived Realism

The system was evaluated through technical validation of psychoacoustic metrics and a human-centered A–B–C experiment comparing an acoustically responsive virtual human with a sensory-isolated control. User perception was assessed using the Godspeed Questionnaire to determine whether acoustic grounding improves perceived realism.

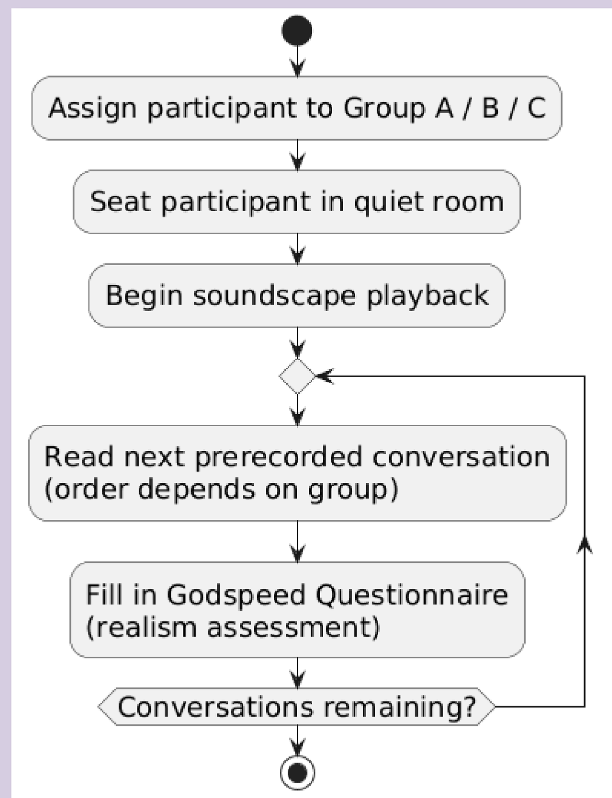

Figure 3 *User Test Procedure*

Results

The system captured relative acoustic differences (Camden vs. Marchmont) and mapped them to PAD values (Camden: -0.12, 0.48, -0.15; Marchmont: 0.35, -0.22, 0.08), enabling mood adaptation. In a user study, Michael (mood + perception) scored highest on Godspeed dimensions—Consciousness 4.07, Humanlikeness 3.92—with statistically significant differences across agents ($\chi^2(2)=8.64$, p=.013), showing enhanced realism and emotional presence.

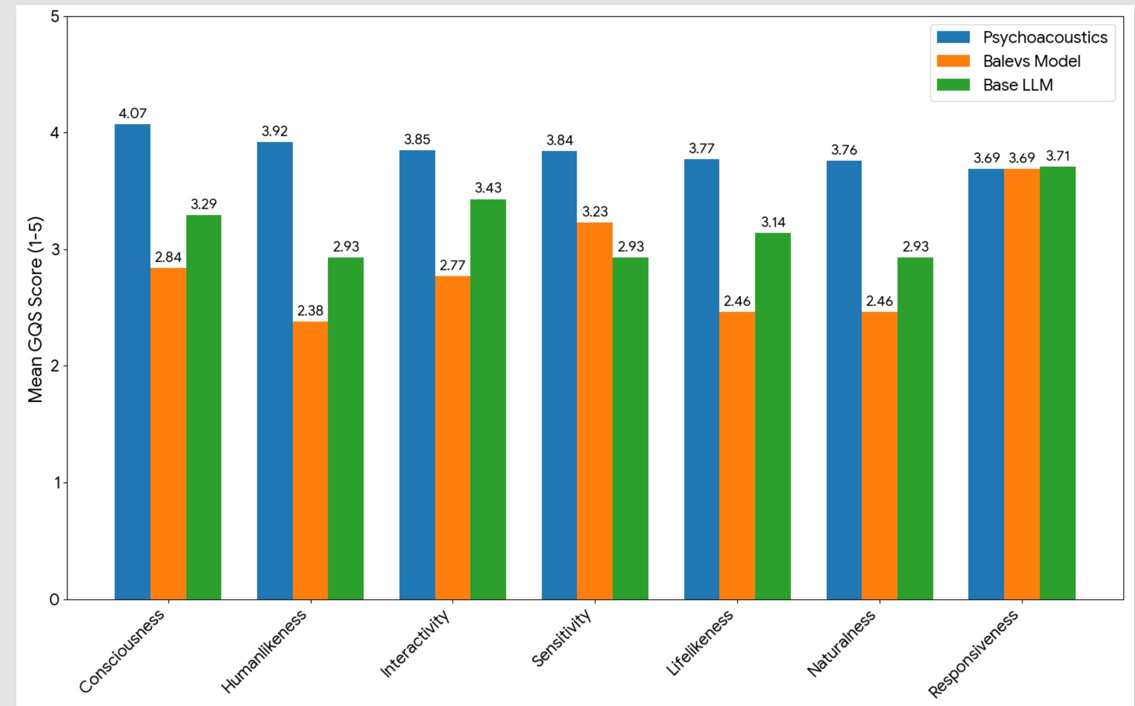

Figure 4 *Survey results and scores*

Discussion

The agent with the psychoacoustic perception system scored highest across all Godspeed dimensions—Consciousness 4.07, Humanlikeness 3.92, Interactivity 3.85—outperforming the internal-mood-only agent (Liam) and the base LLM (Eric). Liam showed moderate improvement over Eric, particularly in responsiveness and interactivity, but remained less emotionally convincing.

These results indicate that grounding a virtual human's affective state in environmental sound enhances emotional presence, coherence, and context-aware mood shifts. Differences were statistically significant ($\chi^2(2)=8.64$, p=.013). Limitations include the small sample size (n=13), a single soundscape, and simplified PAD-based modeling, suggesting results are indicative trends. Nevertheless, sensory-driven affective modulation improves perceived realism and lifelike behavior.

Sources

Balev, K. (2025). A Software Framework for Internal Dynamic Mood Regulation in Emphatic Virtual Humans.

Bartneck, C. (2023). Godspeed Questionnaire Series: Translations and usage. In *Social Robotics* (pp. 13–19). DOI:10.1007/978-3-030-89738-3_24-1

Fastl, H., & Zwicker, E. (2007). *Psychoacoustics: Facts and models* (3rd ed.). Springer-Verlag.

Yang. (2025). How auditory sensations affect human emotional responses to acoustic environments. Journal of Environmental Management. https://doi.org/10.1016/j.jenvman.2025.124742