The Empathy Machine
Generated music to augment empathic interactions

David Kadish, Nikolai Kummer, Aleksandra Dulic, and Homayoun Najjaran

University of British Columbia, Kelowna, Canada
dkadish@interchange.ubc.ca, nikolai.kummer@ubc.ca, aleksandra.dulic@ubc.ca,
homayoun.najjaran@ubc.ca

Abstract. The Empathy Machine is an interactive installation that augments a visitor's empathic sense during a social conversation. Empathy is a key component of interpersonal interactions that is often neglected by modern communication technologies. This system uses facial expression recognition to identify the emotional state of a user's conversation partner. It algorithmically generates emotional music to match the expressive state of the partner and plays the music to the user in a non-disruptive manner. The result is an augmentation of the user's emotional response to the emotional expression of their partner.

Keywords: emotional music synthesis, facial expression recognition, empathic interaction

1 Introduction

Empathy, the ability to share feelings with others, is an important part of interpersonal communication and plays an essential role in human relations. Many communication technologies such as telephones and e-mail inhibit one or more of the senses, becoming barriers between the communicators that reduce the opportunity for empathic interaction.

This paper presents a proposed demonstration of a system that augments empathy in live interpersonal interactions. The installation will allow a user to experience a conversation, while receiving direct feedback regarding the emotional state of the conversation partner in the form of music.

2 Empathy, Emotion and Music

The conceptual basis for the installation is rooted in the hypothesis that music can trigger an empathic response, similar to the response that occurs when one views an emotional expression. Empathy is the ability to identify with the feelings and emotions of another person or being. The neurological basis for empathy is the existence of so-called "mirror neurons" [4]. Mirror neurons are activated both when a subject sees an action or emotional display and when they perform that same activity. Emotional empathy, then, can manifest as a person feeling the

same emotion as someone they are observing simply by seeing an emotional gesture on the part of the observed.
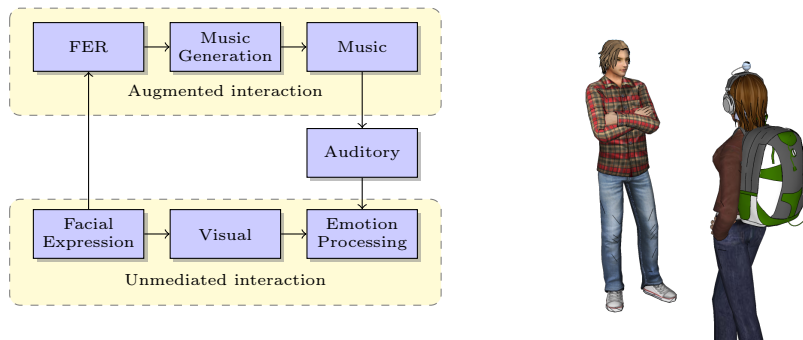
In addition to visual stimulation, music is often cited as a method of evoking an emotional response in a person[5]. In that way, music provides a secondary pathway to the emotional centres of the brain, which can be used to augment the emotional response derived from visual imput. The various properties of music that elicit or represent certain emotions have been studied by scientists since the 1930s[2].

3   Proposed Installation

3.1   Overview

The Empathy Machine augments a social interaction between a user and their conversation partner. The emotional state of the partner is assessed using automatic facial expression recognition (FER) software and the feedback to the user is played back to the user in the form of emotional music. The music is generated using an emotional music synthesis (EMS) system.

The user and their conversation partner sit at a table in the demonstration area. A webcam peers over the shoulder of the user at their partner. Video of the partner's face is input into the FER system, the expression is recognised, appropriate music is generated, and the result is played at background levels to the user through small speakers.

(a) System diagram showing how the Empathy Machine augments a user's empathetic understanding of a situation

(b) The Empathy Machine plays music to the user (green backpack), reflecting the expression of the partner (plaid shirt)

Fig. 1: Overviews of the Empathy Machine system

Table 1: Association between the emotions and musical features

	Tempo	Pitch	Rhythm Roughness	MIDI Instrument	Additional Parameters
Anger	Fast	High	Few repeated notes	#2 Piano	Stoccato, Phrygian mode
Fear	Fast	High		#100 FX 4 Atmosphere	Atonality
Happiness	Fast		Few repeated notes	#47 Orchestral Harp	Stoccato, Lydian mode
Sadness	Slow	Low	Repeated notes	#2 Piano	Legato

3.2   Facial Expression Recognition

The FER system is a modified version of the system described by Valstar and Pantic [7]. FaceTracker, an OpenCV-based algorithm for face registration and tracking based on [6], is used to track 66 points on a face across a series of frames.

The python machine learning package `scikit-learn` is used to perform classification on the features. For each facial action unit (AU) — the basic unit of facial motion — principle component analysis (PCA) is performed on the extracted features and a support vector machine (SVM) classifier is trained on the principle components to determine whether that AU is active. A second classifier is trained to map AU activation to the displayed emotion, using a multiclass one-versus-one SVM. This is used to distinguish between the display of one of four basic emotions and a neutral state.

3.3   Music Generation

The music used to represent the emotions is computer generated. Computer generated music lends itself to continuous generation and does not require an in-depth knowledge of musical composition. The music is generated using the AthenaCL algorithmic music generator. Algorithmic composition allows for the creation of non-repetitive music, thereby reducing strain on the user's ears. Musical parameters for the emotional music were selected from literature[1,3,8]. A summary of the parameters used for the music generation in the user study can be found in Table 1.

4   Future Development

The initial prototype of the Empathy Machine focuses on four emotions in a discrete sense. Future implementations will consider a more nuanced notion of degrees of emotion and will seek to add more basic emotions to the set of four. More effort will also be put into making the FER system more robust and to adding further complexity to the EMS system.

5   Conclusion

The use of music to create an emotional response is common in film and performance. The Empathy Machine represents the first attempt to use music generated in real-time to augment the emotional response to a live interaction. The

authors hope to provide an engaging demonstration of the possibilities that arise from the combination of real-time emotion detection and music generation.

References

1. Gabrielsson, A., Lindström, E.: The influence of musical structure on emotional expression. In: Juslin, P.N.E., Sloboda, J.A.E. (eds.) Music and emotion: Theory and research., pp. 223–248. Series in affective science, Oxford University Press, New York, New York, USA (2001)
2. Hevner, K.: Experimental Studies of the Elements of Expression in Music. The American Journal of Psychology 48(2), 246 (Apr 1936)
3. Juslin, P., Laukka, P.: Expression, Perception, and Induction of Musical Emotions: A Review and a Questionnaire Study of Everyday Listening. Journal of New Music Research 33(3), 217–238 (Sep 2004)
4. Preston, S.D., de Waal, F.B.M.: Empathy: Its ultimate and proximate bases. Behavioral and Brain Sciences 25(01), 1 (Jan 2003)
5. Rickard, N.S.: Intense emotional responses to music: a test of the physiological arousal hypothesis. Psychology of Music 32(4), 371–388 (Oct 2004)
6. Saragih, J., Lucey, S., Cohn, J.: Deformable model fitting by regularized landmark mean-shift. International Journal of Computer Vision pp. 1–16 (2011)
7. Valstar, M., Pantic, M.: Combined support vector machines and hidden markov models for modeling facial action temporal dynamics. In: Proceedings of the 2007 IEEE international conference on Human-computer interaction. pp. 118–127. Springer-Verlag (2007)
8. Wallis, I., Ingalls, T., Campana, E.: A Rule-Based Generative Music System Controlled by Desired Valence and Arousal. In: SMC 11 (2011)