Reinforcement Learning of Listener Response for Mood Classification of Audio

Jack Stockholm
School of Interactive Arts and Technology
Simon Fraser University
Surrey, BC, Canada
jack_stockholm@sfu.ca

Philippe Pasquier
School of Interactive Arts and Technology
Simon Fraser University
Surrey, BC, Canada
pasquier@sfu.ca

Abstract

This paper describes a method of applying a reinforcement learning artificial intelligence to categorize audio files by mood based on listener response during a performance. The system discussed is implemented in a performance art environment designed to present the moods of multiple participants simultaneously in a room via a diffusion of representative audio samples.

1 Introduction

There has been much interest in artificial intelligence music systems which can perform audio in response to user requests based on broad categories such as genre or mood. Many use categorization data which is collected from listeners during playback of the audio to make their decisions. This process is complicated by the fact that individual music or audio files may be classified differently by different individuals. Likewise, a single piece of audio may be classified differently by the same individual on different occasions [1].

This paper documents a simple reinforcement learning method for the classification and performance of audio files by mood which accounts for inconsistency in listener response. This method is currently implemented in a net art installation, titled Eavesdropping, which mixes audio representations of the moods of several participants in a room to increase connectedness and to elicit empathy among participants. Each participant indicates his mood on one of several laptop computers in the room and the system selects an audio sample from a pre-loaded set of audio files to match his mood [2, 3]. The audio is then played back to the participant from his laptop, at which point the participant can indicate whether the audio matches the mood indicated or not. This response is stored in a reinforcement table for each audio file to improve subsequent selections.

2 Mood Classification

Moods in this system are represented by a circumplex ordering of affect around two dimensions [4, 5]. On one dimension, valence relates to the pleasantness of an affective experience, on the other, arousal relates to the perception of arousal associated with such an experience. Therefore audio with a depressed or gloomy mood will have a low valence and low arousal, and audio with an angry mood will have a low valence and high arousal. Likewise, audio with a contented or satisfied mood exhibits a high valence and low arousal while audio which is ecstatic or exultant exhibits a high valence and high arousal [6]. Note that while the term mood is often used interchangeably with emotion, in this context mood is defined as having a longer duration than more the episodic emotion [7].

The mood matrix is presented to participants via a simple grid with a moveable dot which can be dragged to a location to indicate mood [8]. Figure 1 shows the Eavesdropping performance interface with a mood selected at 1x1 on the grid. Position is evaluated where the center of the dot falls between the hash marks on the graphical grid. Note that simplified terms have been used to identify the axes in order to be understandable to a wider audience, with ‘pleasure’ used for ‘valence’ and ‘energy’ used for ‘arousal’. This graphical map offers several advantages over common approaches to mood classification in audio using discrete adjective descriptors. First, the model is quite simple to understand for participants as it only applies two scales and therefore requires minimal time investment versus reading through and selecting from a dictionary of mood adjectives. Second, adjective descriptors have been found to have a variety of meanings across a range of participants [9]. The use of

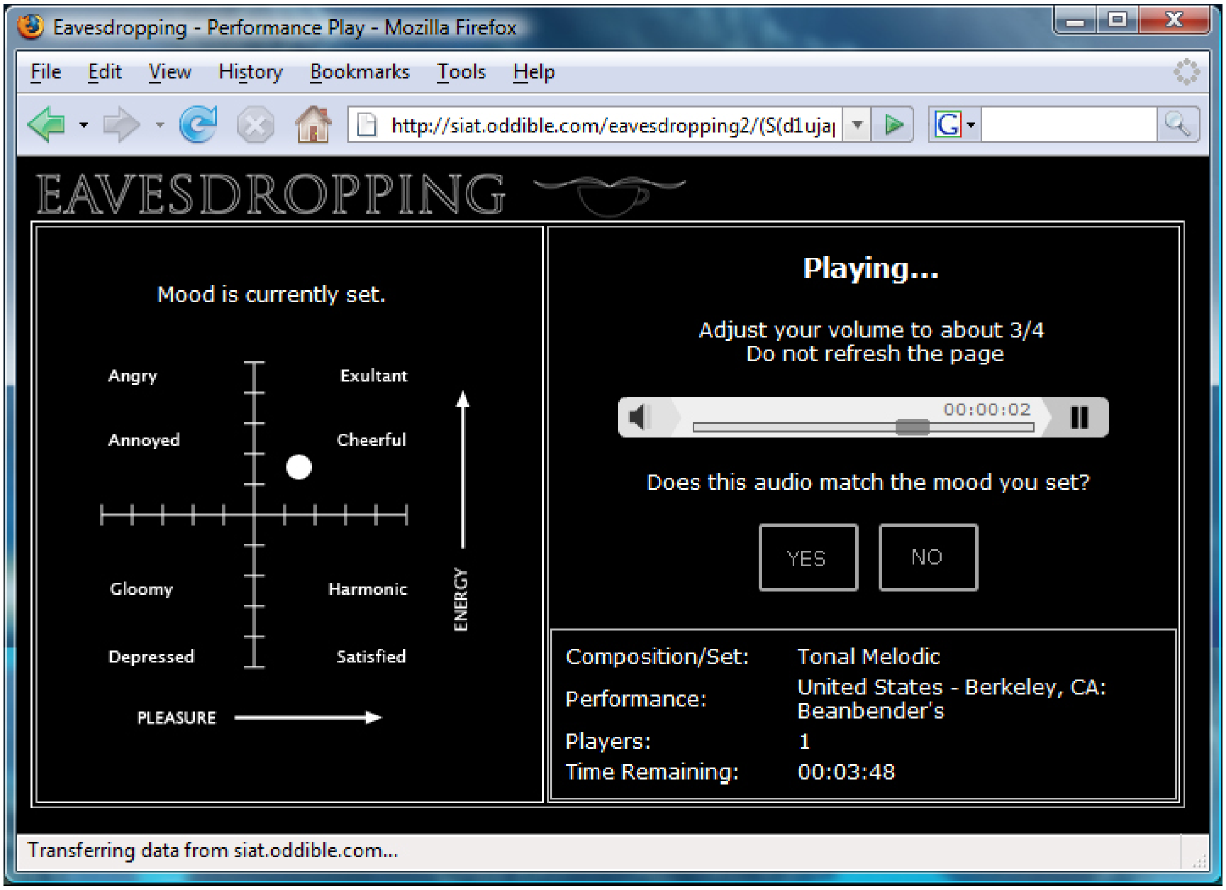

Figure 1. The Eavesdropping performance interface showing the mood matrix and the mood evaluation question.

a mood map with labels minimizes confusion between terms.

When the system begins playing audio to match a participant's indicated mood, the mood matrix screen is accompanied by a simple pair of Yes / No buttons (seen on the right side of Figure 1) asking the participant whether the audio matches their mood or not. This response is recorded by the reinforcement learning system to improve mapping of moods to audio files.

3 Motivations

The Eavesdropping project was motivated by the intention to produce an engaging web audio art environment that can: raise awareness, increase connectedness and facilitate interaction, between networked participants in the same physical space [10, 11]. This is accomplished through a system which mimics the social acoustic ecology and auditory gesture in public spaces [12, 13] to increase shared experiences by capturing mood data from participants on the network and mapping moods to audio files for playback.

Social spaces are alive with audio cues to the dispositions of their participants projected through culturally reinforced sounds such as coughing to indicate irritation or laughing loudly to attract attention. Eavesdropping is a web-based system intended for performance in a networked, social environment such as a café where several laptop users are gathered. The system encourages participants to share their moods with each other via audio selected by the Eavesdropping system and played at each participant's laptop. It is intended that the sharing of audio moods and the awareness that others are engaged with the same interactive interface will increase connectedness and empathy among participants.

3.1 Personal Association

The performance interface is designed to engage participants by giving them agency over the audio presented during a performance and to associate that audio to them personally. Early pilot versions of the system did not request participants to enter their mood but merely performed mood-associated audio from each participant's laptop. Issues with user agency manifest both immediately during the performance as well as in participant frustration expressed during subsequent question and answer sessions. During the performance, participants engaged in all sorts of actions which clearly expressed their intent to be involved in the performance. Some people turned their laptops around to face the majority of other participants. Others raised their laptops above their heads as if to be heard. Some opened multiple browser windows to the system so that their computers were playing multiple sessions. Still others opened music players on their machines and contributed outside sources of audio to the mix.

In question and answer sessions after the performances, many participants suggested that though they felt an association with the music from their laptops via proximity, they wanted more control. Despite the fact that an audience is accustomed to passive listening when playing radios or mp3 players or even music from laptops, once the intent is that their laptops are performing for the rest of the room, participants want agency. Likewise, participants have no issues with control when listening to a live performance by others, but when the performance uses their own laptop as the instrument, participants demand control over their machines.

3.2 Participant Agency

The resolution we employed for these issues was to a) allow participants to input their mood into the system such that the audio playing would represent their mood, and to b) allow them to validate whether that audio was effective at representing their moods. This attempts to address prior user issues in several ways. First, we suspect that by associating the audio directly with participants' moods that this will increase the affinity between participants and the audio as well as increase connectedness among participants with each other. Second, giving each participant a minimal interface which requests involvement during the performance we suspect will both satisfy participants' need for control as well as maintain attention to the current performance without participants seeking to intervene with outside actions or becoming distracted by the interface itself. Additionally, we suspect that enabling the participant to validate whether the mood playing from his laptop represents his mood will disarm any potential embarrassment resulting from the audio association, because the participant can claim that the audio representation is inaccurate. User studies evaluating these assumptions are planned for summer 2009.

3.3 Mood Representation Accuracy

Critical to utilizing mood-associated audio to increase connectedness between audience participants is to establish that the mapping between audio files and moods is accurate. It is the purpose of this paper to detail the second generation of a reinforcement learning system which acknowledges that different users may classify the same audio differently, and that an individual user may classify the same piece of audio differently at different times. It is expected that while there will always be variations in participant response to the audio file-to-mood mapping, that as the number of reinforcements increases (via learning phase or performances), that audio files will converge on accurate mood representation maps. The following section details a system which has shown in pilot evaluations to converge on accurate mood mappings.

4 Machine Learning for Moods

The mood classification system detailed here functions similarly to a supervised learning model in which the machine is making choices to match an expected response and the participant indicates whether the response is accurate, thus guiding the system to converge on the correct answers. However, in traditional supervised learning the participant is considered an expert with the correct answer. In the case of matching audio to moods there are factors which undermine the participant's correctness. First, the system is designed to function for a variety of participants and different participants classify the mood of audio differently based on personal preferences. Second, mood classification is often relative to the current mood of the participant; individual participants may classify the same audio differently depending on their mood [1].

4.1 Q-Learning and Mood Values

The algorithm managing the mood representation data based on participant responses is a variation of the popular Q-learning reinforcement learning [14, 15]. In this case, a Q-value table is utilized to record responses as well as for choosing an audio file to represent a specific mood.

Each audio file is associated with a 5*5 data table to store the utility values that have been learned for that file for each possible mood. The two mood characteristics are each rounded to discrete integers on a scale of 1 to 5 offering 25 possible moods in the table. When an audio file is first added to the system, this table is pre-seeded with very small random values (meaning that this file is nearly equally appropriate for any mood). The higher the Q-value at a specific mood-coordinate for a file, the more likely is that file to get selected to represent that mood. Values in the mood table can range from 0.0 to 1.0. Each time a mood is reinforced for a given file (j), the system also stores the frequency count (nj) to track how many reinforcements a particular file has received.

4.2 Mood Reinforcement

During the performance a participant's Yes or No response to the question of whether the audio matches their mood determines if a file will be positively or negatively reinforced to represent the mood the participant has indicated. In this case we simply update the Q-value (Q(j)(x, y)) for the current mood (x, y) and the current file (j) by adding the value of the learning rate (α) multiplied by the reward value (R) to its existing value as seen in (1).

$$Q_{(j)}(x,y) = Q_{(j)}(x,y) + \alpha R \tag{1}$$

At present the learning rate is set to a constant, 0.1, and the reward value has been set to 1 for positive reinforcements and -1 for negative reinforcements. Given that the range of values for each mood falls between 0.0 and 1.0, each file can reach its maximum value with ten consecutive positive reinforcements. The reinforcement maximum is capped at 1.0 rather than rescaling values. Once the maximum has been reached, subsequent positive reinforcements merely increment the reinforcement count.

4.3 Exploitation and Exploration

Selection of files to represent a participant's mood utilizes a system that takes into account the fact that in a learning-based model, the best fit for a mood may not be the file that has the highest Q-value for that specific mood. There may be suitable files that have been less reinforced (and thus probably less used) that are worth exploring. In general, reinforcement learning faces the problem of balancing exploitation and exploration. This trade-off, classic in artificial intelligence and machine learning, is about choosing at any given point whether to exploit the file that has the highest degree of confidence to represent a specific mood (in this cases the highest Q-value) or exploring files for which the real Q-value is less known.

A pure exploitation strategy formula as presented in (2) would select the file (j) which has the highest Q-value (Q) at the mood location specified (x, y), pondered by its "confidence" (the ratio between the number of reinforcements received for j(nj) and the total number of feedbacks available for the given audio set so far (N).

$$\mathrm{j} \leftarrow \underset{\mathrm{j}}{\operatorname{argmax}} \frac{\mathrm{n_j}}{\mathrm{N}} \mathrm{Q_j}(\mathrm{x},\mathrm{y}) \tag{2}$$

In order to balance exploitation and exploration, a Softmax selection policy is utilized to vary the selection probabilities as a graded function of estimated value. In our case, we utilized the Gibbs measure, or Boltzmann distribution, which is commonly used in machine learning exploration. It chooses file (j) with the following probability (3):

$$\frac{e^{\mathrm{Q_j}(\mathrm{x},\mathrm{y})/T}}{\sum_{x,y} e^{\mathrm{Q_j}(\mathrm{x},\mathrm{y})/T}} \tag{3}$$

The greedy action is still given the highest value but others are weighted according to their Q-values. T represents a positive parameter called temperature with high values for T causing all actions to be nearly equally probable and low values for T causing a greater difference in selection probability for files whose values differ. For our purposes we set T to 0.4. With this value, exploration is significant to allow for fluctuations in participant response while still exploiting known values enough to ensure that users would hear audio appropriate to their mood selection. Because we wanted the system to remain adaptive to variations in participant response there is no mechanism to decrease T over time as is common.

Initial pilots have indicated that this Softmax action selection method will cause convergence of the learning system toward the true Q-values (under the assumption that these exist and are static) and thus an optimal mapping.

5 Evaluation and Conclusion

The reinforcement learning strategy described in this paper is currently implemented in the net art project, Eavesdropping, accessible online at www.oddible.com/eavesdropping [16]. User studies will be performed during summer 2009 utilizing online surveys issued after each performance as well as analysis of the data generated by the audio selection and reinforcement learning system.

User studies will evaluate the success of the interface, the sense of connectedness among participants, and the accuracy of the audio to represent participant mood. The latter will be compared to the user response data from the reinforcement learning system as a means of validation. There are several survey instruments that have been utilized in the field of computer supported collaborative work (CSCW) as well as in online connectedness studies in ubiquitous computing that show promise to evaluate this system without having to reinvent a new instrument [17, 18].

The audio selection and reinforcement learning system records all user interaction as discrete records. It records the intended mood, the audio file selected for that mood, the current number of reinforcements for that audio file, the current Q-value for that audio file at that mood, and the current time within the performance. Evaluation of the reinforcement learning system will explore whether audio file mappings converge to static values. Due to the extensive amount of data collected, it will be interesting to correlate other data points, such as whether certain performances yielded more positive or negative reinforcements than others, whether mood mappings for certain audio files converge and others do not, and whether certain moods of participants cause them to indicate more positive or negative reinforcements.

6 Acknowledgements

This research is partially funded by a commission from Rhizome.org and the New Museum of Contemporary Art in New York, with additional resources provided by Simon Fraser University. Eavesdropping was also featured in the 404 Festival of Electronic Art in Trieste, Italy in 2008. This system is online at http://www.oddible.com/cafe.

7 References

[1] S. Knobloch and D. Zillmann. (2002, Mood management via the digital jukebox. *J. Commun. 52(2),* pp. 351-366.

[2] J. Stockholm and P. Pasquier. Eavesdropping: Audience interaction in networked audio performance. Presented at MM '08: Proceeding of the 16th ACM International Conference on Multimedia.

[3] J. Stockholm. (2008, 12/01). Eavesdropping: Network mediated performance in social space. *Leonardo Music Journal* -pp. 55-58.

[4] J. A. Russell. (1980, A circumplex model of affect. *Journal of Personality & Social Psychology 39(6),* pp. 1161.

[6] R. E. Thayer, *The Biopsychology of Mood and Arousal.* Oxford University Press, 1989.

[7] L. A. Feldman. (1995, Valence focus and arousal focus: Individual differences in the structure of affective experience. *Journal of Personality & Social Psychology 69*pp. 153.

[8] C. Beedie, P. Terry and A. Lane. (2005, Distinctions between emotion and mood. *Cognition & Emotion 19(6),* pp. 847.

[9] M. Eladhari, R. Nieuwdorp and M. Fridenfalk, "The soundtrack of your mind: Mind music - adaptive audio for game characters," in *ACE '06: Proceedings of the 2006 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology,* 2006, pp. 54-61.

[10] D. Liu, L. Lu and H. Zhang. Automatic mood detection from acoustic music data. Presented at Proceedings of the International Symposium on Music Information Retrieval (ISMIR'03).

[11] P. Dourish and S. Bly, "Portholes: Supporting awareness in a distributed work group," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 1992, pp. 541-547.

[12] P. Markopoulos, N. Romero, J. v. Baren, W. IJsselsteijn, B. d. Ruyter and B. Farshchian, "Keeping in touch with the family: Home and away with the ASTRA awareness system," in *CHI '04 Extended Abstracts on Human Factors in Computing Systems,* 2004, pp. 1351-1354.

[13] B. Truax, *Acoustic Communication.* London: Ablex Publishing, 2001.

[14] B. Truax, *Handbook of Acoustic Ecology.* London: Cambridge Street Publishing, 1999.

[15] R. S. Sutton. (1998, Reinforcement learning: An introduction. *Neural Networks, IEEE Transactions on 9(5),* pp. 1054-1054.

[16] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach.* New Jersey: Prentice-Hall, 1995, pp. 932.

[17] J. Stockholm, "Eavesdropping Website," vol. 2008, January 26, 2008. 2008.

[18] J. van Baren, W. A. IJsselsteijn, N. Romero, P. Markopoulos and B. de Ruyter, "Affective benefits in communication: The development and field-testing of a new questionnaire measure," in *PRESENCE 2003,* 2003.

[19] A. L. Gonzales. Validation of an inventory of social connectedness. Presented at GROUP '09: Proceedings of the ACM 2009 International Conference on Supporting Group Work.