A STUDY OF HUMAN MOOD TAGGING OF MUSICAL PIECES

Ian Kaminskyj[1] *and Alexandra L Uitdenbogerd*[2]

[1]Monash University, Wellington Rd, Clayton, 3168, Australia,

[2]RMIT University, GPO Box 2476V, Melbourne, 3001, Australia

ABSTRACT

We conducted a survey in which participants were required to label the mood conveyed within a variety of musical pieces. Two different representations of mood were used, the 2D emotion space as well as updated Hevner mood labels. The results show that survey responses using the two mood representations were both consistent as well as sensible. In terms of music piece characteristics that influenced participant's responses, it has been shown that the intensity/energy, tempo and beat strength consistently influenced participant's mood responses while tonality and pitch did not. Finally, the survey has raised many important questions relating to labeling musical pieces with mood, including the handling of a musical piece conveying more than one mood simultaneously, as well as a musical piece that conveys rapid mood changes.

1. INTRODUCTION

There has been recent attention focused on developing systems to automatically determine the mood conveyed in a musical piece [1, 2, 8] in order to allow users to find music based on its perceived mood. When testing their developed systems, however, these researchers have used the results of human mood tagging of musical pieces performed either by themselves or a small selection of individuals. The issue of how representative these human mood tags were to the broader population, however, was never addressed.

The method to best represent mood within a musical piece also still remains to be determined. Should it be via a list of predefined mood labels, for example, as specified by Schubert's updated Hevner mood labels [6], or via a 2D emotion space (2DES) as specified by Russell [4] or Thayer [1].

To try to answer these questions, participants were asked to complete a survey in which they were required to label the mood conveyed within a variety of musical pieces using both representations to see (a) how consistently the participants labeled the mood of musical pieces, (b) how well they used the two representations for illustrating the mood of musical pieces, and (c) how consistent the results were using the two representation methods. A final aim of the survey was to determine whether there existed any characteristics within a musical piece, such as tempo and intensity that consistently influenced respondent's responses.

2. METHOD

2.1 Survey Participants and Conditions

All participants in the survey were attendees of the 2007 Australasian Computer Music conference. Most participants are involved in computer music composition or data sonification of some kind. A presentation was first made to attendees about the Mood Juke Box project [7], with a brief explanation of the updated Hevner mood labels and the 2DES representation of mood. A 2DES graph showing happy, angry, calm and sad in each of the quadrants was kept visible to participants as they listened to the excerpts and filled in the questionnaire. The questionnaire itself consisted of an item number, the mood labels in their 9 clusters, and a 2DES graph for users to (a) circle the best label describing the perceived mood of the music, and (b) indicate the best location in the 2DES graph to match the mood.

Each piece was played twice, with the exception of the first piece, which was played four times. The entire survey was completed in under 20 minutes.

2.2 The Music

The music presented to participants largely consisted of 10 second excerpts from pieces of music in the collection of 128 pieces used for earlier work in music genre classification (henceforth referred to as the *Mitri collection*) [3]. Only one piece was selected per artist. Most of the collection could be considered to be from popular music, with a predominance of rock and electronic aspects to it. There were, however, three "classical" music excerpts. In addition to the 17 pieces selected from the Mitri collection, 2 pieces from the Magnatune website (one rock, one electronic) were used, and one from Uitdenbogerd's personal collection (a performance of a plainsong composition by Bingen).

Pieces were selected to cover a range of tempi, strongly percussive beats and an absence of percussion, major and minor tonality as well as pitchless (rap). Some selections were included due to the use of distorted instrument sounds. There was also a diversity of pitch ranges from low to high in the main melodic component (where present) of the excerpts. The selections were made by Uitdenbogerd and, while the process was systematic to a considerable degree, and the classical vocal selection from her personal collection was based on 2 random numbers to select a CD and track from a CD jukebox, the collection cannot be considered completely free of subjectivity.

Excerpts from the Mitri collection as well as the Bingen piece were all 10 seconds long, commencing at the 1 minute mark in the original recording. Due to an error during experiment execution, the two Magnatune pieces were played from the beginning for 10 - 15 seconds.

3. RESULTS

Thirty participants took part in the survey comprising 20 pieces. Of these, several respondents either (a) did not submit responses to all pieces, (b) provided more than one Consequently, only the results for 14 pieces were used, for 23 participants for the updated Hevner mood labels part of mood label per piece, (c) filled out the survey form incorrectly, or (d) did not complete the 2DES response. the survey, and 21 participants for the 2DES part of the survey. Table 1

provides a summary of results obtained for both the updated Hevner mood labels and the 2DES. Figure 1 shows how the quadrants referred to in the tables and text are defined. Table 3 highlights the characteristics of the musical pieces used and how these compare with the respondent's responses on the 2DES.

	Musical Piece													
Mood label groups	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Group1–Happy (Q1)		**43**	13		4		13			9		4	4	
Group2-Merry (Q1)		9	**26**		22		**35**	4	9	**26**		13	9	
Group3-Calm (Q4)	17			30	4	26		**35**	26	9	**39**		**39**	**70**
Group4-Dreamy (Q4)	26					**30**	4	13	22	13	35			
Group5-Sad (Q3)	**48**	4	17	17	9	26	13	26	**35**	**26**	4	4	13	26
Group6-Sacred (Q4)	4	17	13	**43**		4	9	4		4		13	17	4
Group7-Tragic (Q3)	4			4		4			9	4			9	
Group8-Angry (Q2)		9	17		**48**	9	4	13		4	13	17		
Group9-Exciting (Q1)		17	13	4	13		22	4		4	9	**48**	9	
2DES parameter														
Quad-1	5	**71**	**33**	10	14		**71**	5	5	**38**		**62**	**29**	5
Quad-2		14	29		**62**		10	5	5	5		29	10	
Quad-3	**71**		10	19	14	14	14	29	38	19	**57**		24	24
Quad-4	24	5	5	**52**		**57**		**38**	**43**	10	24		19	**52**
Origin			5			10			5	10	5			
+Valence			5	5	5	10		10		5	5		5	
-Valence						10							5	
+Arousal		10	14		5		5	5		5	5	10		
-Arousal				14				10	5	10	5		10	19

LEGEND:
Group1 adjectives - Bright, Cheerful, Happy, Joyous; Group2 adjectives - Humorous, Light, Lyrical, Merry, Playful; Group3 adjectives - Calm, Delicate, Graceful, Quiet, Relaxed, Serene, Soothing, Tender, Tranquil; Group4 adjectives - Dreamy, Sentimental; Group5 adjectives - Dark, Depressing, Gloomy, Melancholy, Mournful, Sad, Solemn; Group6 adjectives - Heavy, Majestic, Sacred, Serious, Spiritual, Vigorous
Group7 adjectives - Tragic, Yearning; Group8 adjectives - Agitated, Angry, Restless, Tense
Group9 adjectives - Dramatic, Exciting, Exhilarated, Passionate, Sensational, Soaring, Triumphant
2DES – two-dimensional emotion space
Quad-1 – quadrant 1, Quad-2 – quadrant 2, Quad-3 – quadrant 3, Quad-4 – quadrant 4
Origin – response at origin, +Valence – response on +ve, Valence axis; -Valence – ditto,
+Arousal – response on +ve, Arousal axis; -Arousal - ditto

Table 1: Summary of responses obtained for both the updated Hevner mood labels and the 2DES

4. DISCUSSION

Table 1 shows the correlation between the updated Hevner mood (UHM) labels and the 2DES. If we highlight the dominant UHM group together with the dominant 2DES quadrant for each piece in Table 1, we can observe how the survey respondents have related the two modes of mood description to describe the same mood conveyed in each piece.

Table 2 summarises how the UHM label groups map onto the four 2DES quadrants. Although for most pieces, a clear mapping exists, for pieces 11 and 13, no such mapping is clearly shown. Piece 13 has caused problems with other aspects of the survey (see Table 3 below), so perhaps it must have somehow confused the survey respondents leading to inconsistent survey results.

For piece 11, why UHM groups 3 and 4 (calm and dreamy) have mapped onto quadrant 3 (low arousal, negative valence) is unclear. Finally, based on Schubert's research [3], the mappings shown in Table 2 make sense. One would expect that UHM groups happy, merry and exciting should map onto the high arousal, positive valence quadrant. Similarly, it makes sense that UHM groups calm, dreamy and sacred belong in the low arousal, positive valence quadrant. It is not surprising that the angry UHM group

Group	2DES quadrant
1-Happy	1 (piece 2)
2-Merry	1 (piece 3, 7 and 10)
3-Calm	4 (piece 8 and 14)
4-Dreamy	4 (piece 6 (in combination with Group3-Calm))
5-Sad	3 (piece 1)
6-Sacred	4 (piece 4)
7-Tragic	3 (piece 9 (in combination with Group5-Sad to provide the second highest rated 2DES quadrant))
8-Angry	2 (piece 5)
9-Exciting	1 (piece 12)

Table 2: UHM label groups vs 2DES Quadrants

	Musical Piece													
Piece parameter	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Intensity	l	h	m	l	m	l	m	l	l	m	l	h	l	l
Tempo	s	f	m	m	f	m	m	m	s	m	s	f	m	s
Tonality	M	m	M	m	np	M	m	m	m	M	M	M	m	M
Pitch	h	h	l	h	m	h	mi	l	m	m	m	m	hi	mi
Beat Strength	l	h	h	l	h	l	h	m	l	m	l	h	l	l
2DES parameter														
Quad-1	5	71	33	10	14		71	5	5	38		62	29	5
Quad-2		14	29		62		10	5	5	5		29	10	
Quad-3	71		10	19	14	14	14	29	38	19	57		24	24
Quad-4	24	5	5	52		57		38	43	10	24		19	52
Origin			5			10			5	10	5			
+Valence			5	5	5	10		10		5	5		5	
-Valence						10							5	
+Arousal		10	14		5		5	5		5	5	10		
-Arousal				14				10	5	10	5		10	19

LEGEND:
Intensity: h – high, m – medium, l – low; Tempo: f – fast, m – medium, s - slow
Tonality: M – major, m – minor, np – no pitch; Pitch: h – high, m – medium, l – low
Beat Strength: h – high, m – medium, l – low; 2DES – 2D emotion space
Quad-1 – quadrant 1, Quad-2 – quadrant 2, Quad-3 – quadrant 3, Quad-4 – quadrant 4
Origin – response at origin; +Valence – response on +ve Valence axis; -Valence – ditto
+Arousal – response on +ve Arousal axis; -Arousal – ditto

Table 3: Characteristics of the musical pieces versus respondent's responses on the 2DES

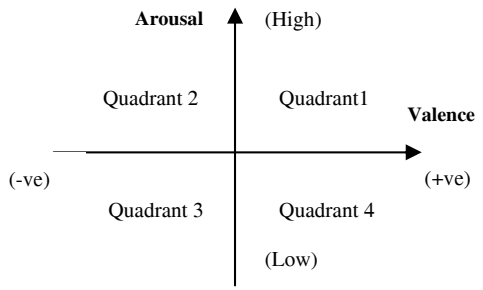

Figure 1: The 2D emotion space referred to in our paper

should map onto the high arousal, negative valence quadrant. Finally, mapping the sad, and to some extent, the tragic, UHM groups onto the low arousal, negative valence quadrant is also entirely sensible. Table 1, therefore shows that the survey respondent's results between the UHM list and the 2DES are entirely consistent and sensible. This adds further credibility to the validity of the survey results.

The results shown in Table 3 support the following relationships between the characteristics of a musical piece and the 2DES:

1. Pieces of high intensity/energy (eg. 2 and 12) map onto quadrants 1 and 2 as well as the positive Arousal axis, while pieces of low intensity/energy (eg. 1, 4, 6, 8, 9, 11 and 14) map onto quadrants 3 and 4 as well as the negative Arousal axis. Pieces of medium intensity/energy are generally more spread out across the 2DES (eg. 3 and 10). This confirms that intensity/energy indicates the arousal value within a piece. The only exception to this was 13.
2. Pieces of fast tempo (eg. 2, 5, 12) map onto quadrants 1 and 2 as well as the positive Arousal axis, while pieces of slow tempo (eg. 1, 9, 11 and 14) map onto quadrants 3 and 4 as well as the negative Arousal axis. Pieces of medium tempo are generally more spread out across the 2DES (eg. 3, 7, 8, 10 and 13). This supports the notion that tempo is highly correlated with intensity/energy, whereby fast pieces tend to comprise higher amounts of intensity/energy than slow pieces.
3. There does not appear to be any recognizable trend for pieces of either major or minor tonality with respect to the 2DES.
4. There does not appear to be any recognizable trend for pieces of either high, low or no pitch with respect to the 2DES.
5. Pieces of high beat strength (eg. 2, 3, 5, 7 and 12) map onto quadrants 1 and 2 as well as the positive Arousal axis, while pieces of low beat strength (eg. 1, 4, 6, 9, 11 and 14) map onto quadrants 3 and 4 as well as the negative Arousal axis. Pieces of medium beat strength are generally more spread out across the 2DES (eg. 8 and 10). The only exception to this was again, 13.

Looking at the survey responses in Table 3, there do not appear to be any pieces that clearly fall into either the positive (quadrants 1 and 4) or negative (quadrants 2 and 3) valence axis areas of the

2DES. On this basis, unlike the arousal axis, it makes it quite difficult to determine which characteristics of a musical piece clearly denote its valence.

Piece 13 was J.S. Bach's Concerto No. 4 in G Major – Andante. It is possible that being the only classical piece in the survey (apart from the Bingen piece) may explain the unusual inconsistent results that were experienced with it.

Many issues were raised by the execution of the experiment, and by the participants themselves.

A limitation of our experiment is that there is significant similarity in the grouping of pieces according to intensity/energy and beat strength. It would be difficult to determine whether both predict mood, or whether just one of these features does.

An important issue that has implications for future research into music mood classification is the existence of "mood counterpoint", that is, more than one mood being present in a piece of music simultaneously. One piece in the collection of excerpts suggested two moods: one energetic mood in the introductory riff, and a gentler mood in the melodic instrument that entered several seconds later above the existing riff. Existing work has examined moods that occur serially throughout a piece, but we are unaware of work on mood counterpoint. A related comment was the possibility of two moods existing simultaneously, such as a piece being "calmly angry".

It was strongly suggested that an important factor in perception of emotion or mood in music is the dynamic, that is, the loudness or volume over time. This certainly seems to be born out by studies of emotion variation during a musical work (for example [5]). It is possible that the volume at which the music was presented (fairly low), may have affected the choice of mood by participants.

At least one participant objected strongly to the idea of reducing music to specific categories, based on the principle of limiting the way that listeners respond to music. There was also discomfort in being required to fill out a "forced choice" questionnaire that had no options for multiple or null answers (though participants could leave a particular question blank, or cease participation at any time).

One participant who was a computer music composer, stated that they hadn't really thought of music in terms of emotion. They chose to not complete the survey for this reason. It was suggested by another attendee that many members of the computer music community practiced "reduced listening", which is a method of listening that focuses on the sound itself and not on any associated meaning[1]. Sound engineers may also practise this type of listening more readily than semantic or casual listeners.

The cognitive load for participants was very high. One person stated that 10 seconds was far too short to perceive mood. There was also not much time for participants to decide on the mood for each piece.

The perception of mood in music is likely to be (sub-)culturally specific.

While possibly not an issue for this experiment per se, it was noted that a person's reaction to a given piece of music varies over time with familiarity.

5. CONCLUSION

Participants in our survey were required to label the mood conveyed within a variety of musical pieces. Two different representations of mood were used, the 2DES as well as UHM labels. The results show that survey responses using the two mood representations were both consistent as well as sensible. In terms of music piece characteristics that influenced participant's responses, it was shown that the intensity/energy, tempo and beat strength consistently influenced participant's mood responses while tonality and pitch did not. Finally, the survey has raised many important questions relating to labeling musical pieces with moods.

Future work will entail (a) closer examination of the mood counterpoint concept and how it can best be evaluated using human mood tagging surveys and automated mood classification systems, (b) expanding the survey to cover a much wider range of participants and musical pieces by going on-line, (c) determining whether beat strength and energy/intensity are separate factors for predicting arousal, and (d) developing a deeper understanding of the mapping of the UHM labels onto the 2DES.

6. REFERENCES

1. Feng Y., Zhuang Y. and Pan. Y. "Music information retrieval by detecting mood via computational media aesthetics", In J. Liu, editor, Proc. IEEE Intl. Conf. on Web Intelligence, pp. 235-241, Washington, USA, October. IEEE, IEEE Computer Society, 2003.

2. Liu, D., Lu, L. and Zhang, H.-J. "Automatic mood detection and tracking of music audio signals", IEEE Trans. ASSP, 14(1):5-18, January, 2006.

3. Mitri, G., Uitdenbogerd, A. L. and Ciesielski. V. "Automatic music classification problems", Proc. Australasian Computer Science Conference, 2004.

4. Schubert, E. "Measuring emotion continuously: Validity and reliability of the two dimensional emotion space", Aust. Jnl. Psychology, 51:154-165, 1999.

5. Schubert, E. "Modeling perceived emotion with continuous musical features", Music Perception, 21(4):561-585, 2004.

6. Schubert, E. "Update of Hevner's adjective checklist", Perceptual and Motor Skills, 96:1117-1122, 2003.

7. Uitdenbogerd, A. and Kaminskyj, I. "The Mood Juke Box project: classifying and retrieving music by its perceived emotional content", Proc. Australasian Comp. Music Conf., Canberra, Jun. 2007

8. Yuan-Yuan Shi, Xuan Zhu, Hyoung-Gook Kim, Ki-Wan Eom, "A tempo feature via modulation spectrum analysis and Its application to music emotion classification" ICME 2006, p1085-8, 2006

[1] (http://this.is/herrahelviti/writings/makingStrange_essay.pdf).