Conveying emotion through music

Generating affective background music using cellular automata

R.C. Paul
Human Media Interaction, EEMCS Faculty, University of Twente
P.O. Box 217, 7500 AE, Enschede, The Netherlands
r.c.paul@student.utwente.nl

ABSTRACT
Emotion adds affect and improves understanding in human-computer interaction. Emotion can not only be conveyed to the human user using visual cues like color or facial expression, but also using music. This paper describes development of CAM (Cellular Automata Music) that generates music using cellular automata and adds affect by modifying it according to a few simple rules that are based on detection of emotion in music. Input for the system is emotion which is given as a point on the two dimensional emotion space (2DES). The music is very basic because it lacks properties like patterns or chords, but it should be good enough to be used as background music. The system is evaluated with a user experiment in which participants were asked to indicate perceived emotions while listening to a few fragments. This article describes the state of art of music generation, detection of emotion in music and generating affective music and discusses the design and results of the experiment.

1. INTRODUCTION
Human-computer interaction involves humans and computers. Aspects of emotion can add affect and improves understanding in human-computer interaction, especially when a system is build for entertainment. In games, television series and especially movies, music plays an important role in setting the right ambiance and is used to induce emotion. It is very powerful; a car chase becomes much more exiting with appropriate music. When a slow romantic tune is played instead, ambiance will be completely different.

1.1 Motivation
Music in games cannot adapt to a given situation with ease. It is possible to record tracks for every unique setting and switch between them when the ambiance changes but a lot of tracks should be composed this way. This imposes the need for real-time generation of affective music.

Throughout this paper, the Virtual Storyteller is held as example project for which generation of affective music is useful. The Virtual Storyteller is a system that generates stories and tells them using a speaking embodied agent. I think that when these stories are accompanied by affective music, the Virtual Storyteller has more power in inducing emotion to the user which makes the stories more exiting.

1.2 Research questions
The main research question is the following: *How to generate music that induces emotion?* This is divided in three subquestions:

- What is the state of the art of music generation?
- What are relevant features for detecting emotion in music?
- How can we, using these features, generate affective music?

What I did is a user test combined with construction research. I investigated what possibilities exist for the Virtual Storyteller for example to use music as an (extra) channel through which emotion can be conveyed. I approached this problem by looking at the current technologies of detection of emotion in music because in comparison, literature about emotion in synthesis of music is almost non-existent. There are certain features that give a good indication of the mood or emotion of a musical piece and I used these features in a reverse way for the synthesis of music.

For the first subquestions, two small literature studies were conducted. One for the state of art of music generation (§2) and one for relevant features for detecting emotion in music (§3). After this, a music generation algorithm was set up using the information from the first activity (§4) and a user test was conducted (§5). In this test, participants were asked to indicate perceived emotions while listening to a few fragments. Results are presented in the end (§6).

2. STATE OF THE ART OF MUSIC GENERATION
Generating music can be as simple as playing some notes of random lengths at random positions in the time line. But that would not sound as music that way. What humans call music is hard to grasp for a computer. There are music theories, but not all of them are as formal as an algorithm

would like to have it. And it is arguable that music without expression or emotion is no music at all.

A lot of researchers [1, 6, 10, 11, 15, 21] use genetic algorithms for generation of music. A genetic algorithm tries to mimic evolution found in nature by utilizing its principles of mutation and fitness. See §2.1. Machine learning is also used [9, 13, 16], and Delgado et al. [7] use an expert system. In two approaches, cellular automata are being used [17, 8], see §2.2. [5] uses a theory about rhythm and metre.

2.1 Genetic algorithms

Alfonseca et al. [1] describe a genetic algorithm for music generation. The fitness function used is the NCD, the Normalized Compression Distance. There are different genetic operators (ways a piece of music is mutated). The article is strictly based on creating music that is like already existing music. The techniques in this paper are only interesting from the point of view of creating music that sounds like the music the guides are created from. From a emotional-music point of view, this introduces another level since there must be a mapping between emotions and the corresponding already existing pieces of music.

I-sound is work in progress and the article about it describes on a reasonably abstract level how Cruz et al. [6] are planning to build their system. Emotion is incorporated by using work of Lopes [12] on a theory about rhythm and metre based on the two basic rhythmic qualities of pulse salience and kinesis.

Khalifa et al. [10] describe a way of music synthesis using formal grammar. The goal is to create an unsupervised music composer based on patterns. A CFG based on music theory is used for this. The system created, the Autonomous Evolutionary Music Composer (AEMC) does not need musical or evolutionary input, hence it is autonomous and unsupervised.

Traditionally, fitness of a generated piece of music is evaluated by a human listener. Lo et al. [11] replaced this time consuming process with a trainable music evaluation system. This system can be trained on existing pieces and is implemented with an N-gram language model.

Ozcanet al. [15] describe components of genetic algorithms embedded in a tool named AMUSE. It is used to create improvised melodies given a harmonic context. The fitness function works on ten different melodic and rhythmic features.

Yee-King [21] describe work done on an automatic music improvised which attempts to follow an incoming audio stream. The fitness function of this genetic algorithm driven synthesis system compares the synthesized output with the input to match the incoming music.

2.2 Cellular Automata

In two approaches, cellular automata are being used [17, 8]. The concept of cellular automata is created by Stephen Wolfram and when generating music with them, their value is in the patterns they can produce. A cellular automaton (CA) consists of a number of cells next to each other, each having a binary value. The next state of the automaton is calculated from the value of the cell itself and its two neighbors and a Wolfram rule. This rule is a 8 bit number and each of the bits describe whether a cell should be switched on or of for each of the eight possible combinations of previous values of the cell itself and the two neighbors.

See Figure 1 for the cellular automaton built with Wolfram rule 60 [20] as example.

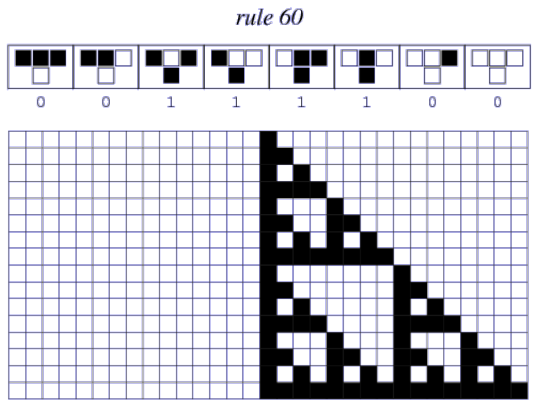

Figure 1: Example cellular automaton built with Rule 60

A cellular automaton is converted to music by first interpreting the rows to positive integers and then use these integers to select notes. This last step can be performed in various ways but the most common way is to just modulo this integer with the total number of available notes.

2.3 Other approaches

Campolongo and Vena [5] describe a unique way of selecting notes. After an introduction to parameters that can be derived from graphs (path length L and clustering coefficient C) and a definition of some specific graphs using these parameters, they describe a method for directly mapping music to graphs, or networks. On a low level, all vertices correspond to a note and all edges to to a 'follows'-relation. Using these graphs, Campolongo and Vena analyzed musical pieces of Bach and Mozart in terms of L and C. They also tried to generate music by traversal of specific graphs (imported or generated) and optimize networks by using genetic algorithms.

HYPERSCORE described by Bottoni et al. [2], allows users to virtually manipulate music as it is performed. This is done using a gestural interface and scaling along two axis of a geometrical representation of music.

For a more in depth review of literature, consult Lopez and Arcos [13] which is an overview of past efforts made in computer music under the roof of AI, divided over composition, improvisation and performance.

3. FEATURES FOR DETECTION

This section is about detection of emotion from music. Numerous papers are written on detecting emotion in music.

Variations exist in musical style, emotion representation and the way emotion is matched to low level features. Furthermore, software exists for extracting low level features; PsySound3 [4] which is a collection of Matlab scripts and MARSYAS [14].

Schubert [18] measured emotion in music in his PhD thesis and his work proved to be invaluable. Not much work had been done on analysis of emotion in music over time until then. Emotion is represented using the two dimensional model of valence and arousal, introduced by Thayer [19], the two dimensional emotion space (2DES). Valence is a scale from negative to positive and arousal from standby to energetic. See Figure 2.

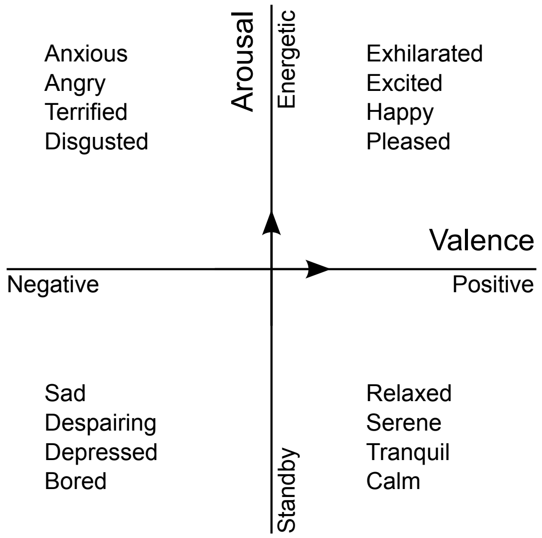

Figure 2: The two dimensional valence-arousal space.

The work includes an extensive analysis of the relation between music features and emotion shown on the 2DES. The analysis is based on literature. Authors do not always agree on the emotion in effect. Schubert measured continuous emotional responses to music, thereby validating his findings from literature.

Schubert selected features that are "perceptually and musically meaningful". A list of these features and the mapping to 2DES follows.

1. Dynamics. Loudness related. Works primarily on arousal. Loud music corresponds to an emotion with a high arousal and vice versa.
2. Mean pitch. Pitch related. It is the average pitch of a musical section. Most literature agrees on that a high pitch corresponds to more arousal, except Thayer [19] who states pitch has influence on the valence dimension.
3. Pitch range. Pitch related. It is the difference between the highest and the lowest pitch. The wider the range, the more arousal.
4. Variation in pitch. Pitch related. When a musical piece has no variation in its melodic line, all tones are the same. When variation is high, tones will vary a lot from the mean pitch. A large variation causes higher levels of arousal.
5. Melodic direction or contour. Pitch related. Can be seen as the change of melodic pitch. A positive direction is transformed to a positive influence on arousal.
6. Register. Pitch related. The octave in which a piece is played. A higher register means a more positive valence.
7. Mode. Pitch related. The mode of a piece of music is for example minor or major. Music performed in minor - intuitively - has a lower valence than when performed in major.
8. Timbre. Pitch related. The "sound" of an instrument or group of instruments. Think about bright and sharp vs. dark and soft. A complex feature with different definitions because the multiple levels of meaning.
9. Harmony. Pitch related. The combination of notes. No clear mapping to 2DES can be made because of the many possible combinations of musical notes.
10. Tempo. Duration related. This is measured in beats per minute. With a few beats per minute, the piece is considered slow. Tempo primarily works on arousal.
11. Articulation. Duration related. The "shape" of a sound. With this, notes can be long and connected or short and separated from each other. Most authors agree that this is of influence on valence. The more connected the notes are, the more negative is the valence. And vice versa.
12. Note onset. Duration related. This is the attack of the note. The portion of time between the start of the note and the moment the note is at full strength. A sort of fade-in time. The faster the onset, the higher the arousal and vice versa.
13. Vibrato. Duration related. A wave-like change of volume or pitch of a tone. No clear translation to a 2DES can be made because it is described by both depth and rate and can both be applicable to volume and pitch.
14. Rhythm. Duration related. Multiple definitions exist, no clear translation to a 2DES because there are too many different kinds of rhythms.
15. Metre. Duration related. The grouping of beats in music. The two most common are simple metres (divisible by two) and triple metres (divisible by three). The triple metres has a more positive influence on the arousal, but there is only one source in literature that suggests this.

4. GENERATING AFFECTIVE MUSIC

We have chosen to generate music with cellular automata because they are simple and fast and because they not need training, they can easily be adapted to be used in an affective environment. A downside is that most of the music generated by CAs is not very interesting to listen to actively, but that is not of a big concern because it is to be used in the background anyway. See Figure 3 for an overview of CAM. Emotion and a CA-rule are input and affective music is output.

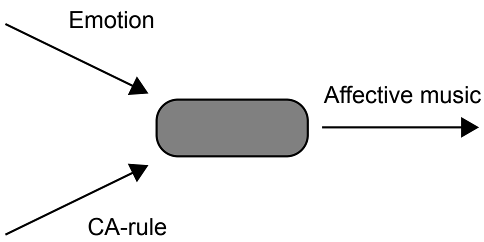

Figure 3: Overview of CAM with its inputs and output.

Not all music features identified by Schubert (§3) will be used for generating music with the CA. Features for which the transformation to a 2DES is not clear are useless for generating emotional music. These are timbre, harmony, vibrato, rhythm and metre. I'll also leave out register because I think it has too much influence on the sound of a generated piece of music. All positive valence music pieces would be generated a octave higher than negative valence music pieces. Furthermore, pitch already has been represented fairly well through mean pitch. Also, note that when a piece is played in a higher octave it also must have a higher mean pitch. But both features have a different transformation to a 2DES. A higher mean pitch means a higher arousal but a higher register can be interpreted as a more positive valence.

Because we have chosen for music generation by means of CA, we cannot alter variation in pitch and melodic direction without jeopardizing the nature of the CA generated nature because variation in pitch and melodic direction are dependent on the CA. When altering these, we would need to ignore all notes coming out of the CA. So these two features are left out and this leaves us with seven features: dynamics, mean pitch, pitch range, mode, tempo, articulation and note onset.

There are several issues to think about when incorporating these features in order to have a piece of music generated with a general emotion response that corresponds to the different 2DES transformation of all individual music features:

- Influence. If there would be two features, both working on the arousal dimension of the 2DES, do they both have the same influence? Most probably not. The problem is that the 2DES transformations described by Schubert are not quantified. They only apply to one of the two dimensions in one of the two directions. For the user test, this is not an issue because we will not test features separately from each other.
- Scale. Take the music feature tempo as example. What are the boundaries? Do we consider 60 beats per minute as slow or is 30 more appropriate? The same holds for the upper boundary and most of the other features. For this issue, I have estimated some defaults based on popular music.
- Linearity. Does a pitch increase in the lower regions of notes mean the same movement on the 2DES when this same increase is applied to some higher region? No meaningful answer can be given to these questions. For the user test, this is not an issue because we are only interested in whether there is a correlation or not.
- Monotonicity. There is a possibility that the mapping function is not monotonic. This means for example that maximum arousal using mean pitch is reached somewhere in between the minimum and maximum pitches. To keep the complexity of the test manageable, I assume all functions are monotonic.

Feature	Dimension	Scale
Dynamics	Arousal	Low volumes should be considerably lower than high volumes.
Mean pitch	Arousal	MIDI-note 50 ∼ 100
Pitch range	Arousal	10 ∼ 50 MIDI-notes
Mode	Valence	Major / minor
Tempo	Arousal	40 ∼ 150 beats per minute
Articulation	Valence	80% ∼ 0% silence
Note onset	Arousal	0% ∼ 100% of note length

Table 1: Mapping of how each feature will have its impact on the music itself. ∼ denotes a range.

Coding is how to generate music with the features encoded in them. See Table 1 for how each feature will have its impact on the music itself. Note that only two features have influence on the valence dimension.

See Figure 4 for when features are encoded. Mean pitch, pitch range and mode will be used to create a set of notes that the CA can pick, assuming each note has an equal probability to be picked. After the algorithm has done its work, dynamics, articulation and note onset are applied. And when we want to generate fragments of a limited length, tempo not only determines the speed in which notes are played, but also the number of notes the algorithm should generate.

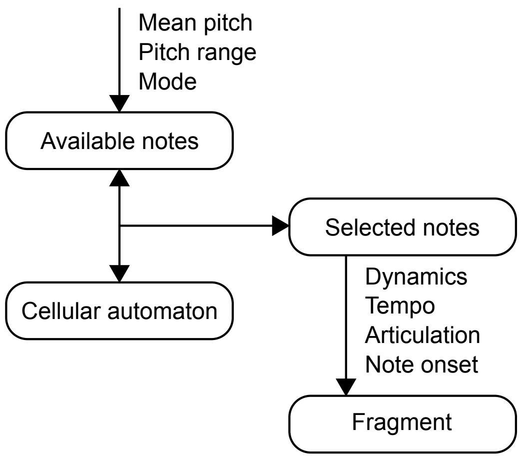

Figure 4: Not all features are encoded at the same time.

See Figure 5 for four visual representation of fragments that show pitches, lengths of notes and other features, placed inside a valence-arousal diagram. Note that these are the four extremes.

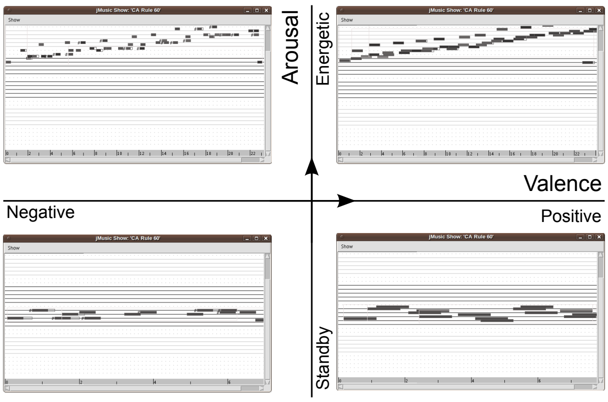

Figure 5: Visual representations of generated fragments placed in a valence-arousal diagram.

5. USER TEST DESIGN

For the test, 25 points will equally be divided over the two dimensional emotion space. For a 2DES with a range of $[-1.0 : 1.0]$, points are put on $(-0.8, -0.4, 0.0, 0.4, 0.8)$ for both axes. Furthermore, a selection of usable CA rules must be made. Some CAs result in fragments with only the same notes and some with only decreasing or increasing notes. See Figure 6 for a visualization of notes with an unacceptable repeating pattern.

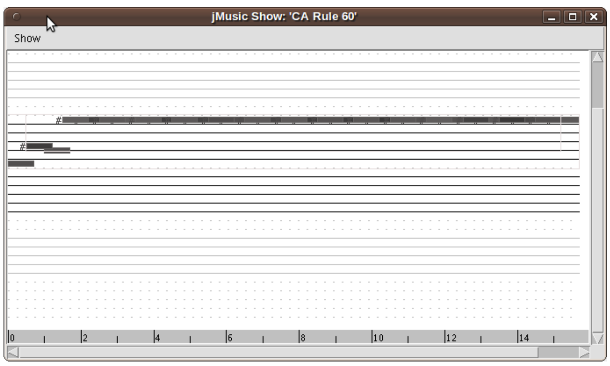

Figure 6: Rule 180 was not usable because after the fist few notes, the pitch did not change anymore.

Then, all combinations with valid rules and the 25 positions on the 2DES will be used to create a big pool of fragments of approximately 10 seconds so each of the positions on the 2DES has the same number of fragments. The user test itself is simple. Each participant will be asked to listen to 25 different fragments and to give a response using the Self Assessment Mannequin (SAM) [3]. Selection of fragments for the subject is random although it is ensured that the mean valence and mean arousal are in the $[-0.2 : 0.2]$ interval to prevent the possible case that a participants gets to hear significantly more fragments from one of the quadrants which would possibly influence the results. See Figure 7 for a screenshot of the Survey on Affective Music, the web application developed for this purpose.

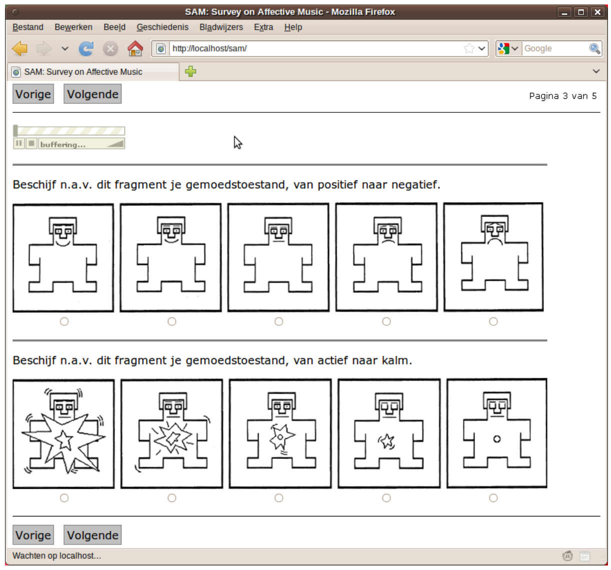

Figure 7: Screenshot of the Survey on Affective Music

Analysis consists of correlating input (valence and arousal for which each music fragment was generated) with output (valence and arousal the user has chosen). We will use the Pearson correlation coefficient. For both dimensions, input and output of the other dimension are correlated. So for valence -2 for example, all combinations of input arousal and output arousal for fragments with valence -2 are used for calculating this correlation. The same holds for valence -1, 0, 1, 2 and arousal -2, -1, 0, 1 and 2. We have chosen to split the results this way because it is the only way we could calculate Pearson's correlation coefficient.

The procedure of the test was as follows:

- select 25 fragments;
- describe the test and its purpose to the user;
- ask the user to fill in age, gender, musical preferences and whether he plays or played a musical instrument;
- show the user a dummy question with the fragment player and the SAM, hereby including information about the player, SAM itself, that the test should be done in moderate silence and that the fragment should be clearly audible;
- per fragment: play it and let the user choose a response with SAM;
- in the end: give the user the opportunity to fill in an e-mail address when he is interested in the outcome of this research.

88 rules were usable, so the total number of fragments generated was 2200. Invitations were sent to friends and family and placed on a digital learning environment. The test ran from 13th of January until the 25th of that same month, a total of 12 full days. 21 participants followed the test, of which 16 (76%) were male. Age: average = 24.81, standard

deviation = 4.56. 3 (14%) liked dance, 7 (33%) pop, 7 (33%) rock and 4 (19%) favored other music genres. None of the participants played an instrument.

6. RESULTS

See Figure 8 for 25 input-output matrices. Positions of the submatrices and the cells in them are positioned analogue to the 2DES diagram of Figure 2. The positions of the submatrices depict the input and the positions of numbers in the cells of the submatrices depict the output. The numbers themselves show how many times this position in the 2DES has been chosen. So, for all neutral fragments (0,0 on the 2DES), 2 responses were on $(1,1)$, 1 on $(-1,0)$, 3 on $(0,0)$, 2 on $(1,0)$, 2 on $(-1,-1)$, 3 on $(0,-1)$, 4 on $(1,-1)$ and 3 on $(0,-2)$. The way to look at this matrix is to see where the center of gravity is for each submatrix in respect to the position of the submatrix itself. When you look at the center of gravities of the slice arousal -2, you can see that most of them are also in arousal -2 or -1. When looking at the rows from top to bottom, the center of gravity is more or less also moving downwards. This effect seems less present when looking at columns from left to right or the other way around so this gives us an early indication on the performance of the system.

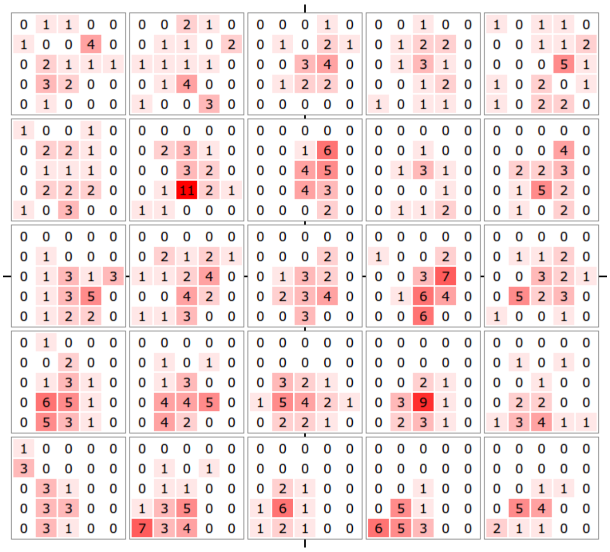

Figure 8: 25 input-output matrices, positioned analogous to the two dimensional emotion space.

We will calculate Pearson correlation coefficients for each of the input valences and arousals. We do this per 'slice'. A slice is one of the five rows or one of the five columns of the matrix of Figure 8. See Table 2 for the data for slice valence -2 for example. In this table, for all combinations of input-output arousals, the number of occurrences is counted. The Pearson correlation of this table is 0.233. With a 0.05 level of significance for this two-tailed test, the critical value is 0.188, so this correlation is significant.

See Table 3 for all correlation coefficients and whether they are significant. H_1 indicates whether the slice has a significant correlation or not. Remember that for correlation of an arousal slice, we actually measuring performance of the valence-dimension.

Slice	r	Pairs	Conf.	H_1
Valence -2	0.233	108	0.188	True
Valence -1	0.354	125	0.175	True
Valence 0	0.454	101	0.195	True
Valence 1	0.476	101	0.195	True
Valence 2	0.344	98	0.198	True
Arousal -2	0.063	96	0.200	False
Arousal -1	0.063	117	0.181	False
Arousal 0	-0.059	120	0.179	False
Arousal 1	0.218	106	0.190	True
Arousal 2	0.175	94	0.202	False

Table 3: Correlations for the full set of data.

Slice	r	Pairs	Conf.	H_1
Valence -2	0.187	83	0.215	False
Valence -1	0.276	91	0.205	True
Valence 0	0.445	75	0.226	True
Valence 1	0.434	76	0.225	True
Valence 2	0.345	74	0.228	True
Arousal -2	0.164	72	0.231	False
Arousal -1	0.040	86	0.211	False
Arousal 0	0.005	87	0.210	False
Arousal 1	0.282	87	0.210	True
Arousal 2	0.196	67	0.240	False

Table 4: Correlations for all male participants

To visualize the correlation between the predicted and observed values, see Figures 9 and 10 for contour plots of respectively the strongest and the weakest correlation. In the first plot, a south-west to north-east orientation is clearly visible which visually indicates a correlation.

See Table 4 for the results of all male participants. We see that slice valence -2 now does not give us a statistically significant correlation. We already saw the slice of the total was not very strong.

See Table 5 for the results of all female participants. We see that for four of the five valence slices, correlations are higher than those of the male participants, especially for the slice valence -2. Furthermore, arousal slices show an overall negative correlation, but none of those were significant.

Data split by musical preference is summarized in Table 6. Notably, slices arousal 1 and 2 perform very well under participants that have dance as their favorite music. The rest of the arousal slices also have higher correlations than any other musical preference.

7. CONCLUSION

We conclude that it is possible to generate affective music using CA and features that originally are used in detection of emotion in music. However, some problems arise on the arousal slices (correlations are much lower and only significant in one of the cases) so it is harder to induce valence. This can be caused by the fact that only two features work on the valence dimension (mode and articulation) but one

Predicted	-2	-2	-2	-2	-2	-1	-1	-1	-1	-1	0	0	0	0	0	1	1	1	1	1	2	2	2	2	2
Observed	-2	-1	0	1	2	-2	-1	0	1	2	-2	-1	0	1	2	-2	-1	0	1	2	-2	-1	0	1	2
Occurrences	**4**	6	4	3	1	9	**12**	5	2	1	5	9	**8**	1	0	4	6	3	**5**	2	1	5	5	5	**2**

Table 2: Data for the correlation coefficient for valence -2

Slice	r	Pairs	Conf.	H_1
Valence -2	0.584	23	0.412	True
Valence -1	0.605	28	0.373	True
Valence 0	0.535	26	0.387	True
Valence 1	0.659	23	0.412	True
Valence 2	0.209	24	0.403	False
Arousal -2	-0.104	23	0.412	False
Arousal -1	0.115	29	0.366	False
Arousal 0	-0.364	27	0.380	False
Arousal 1	-0.103	19	0.454	False
Arousal 2	0.138	26	0.387	False

Table 5: Correlations for all female participants

Slice	Pop	Rock	Dance	Other
Valence -2	0.216	**0.334**	0.216	**0.598**
Valence -1	0.284	0.239	**0.526**	**0.445**
Valence 0	0.319	**0.511**	-0.060	0.448
Valence 1	0.192	**0.636**	0.338	**0.754**
Valence 2	**0.479**	**0.524**	0.000	0.076
Arousal -2	-0.173	0.169	0.275	0.273
Arousal -1	-0.017	0.023	0.428	-0.011
Arousal 0	-0.007	-0.184	0.142	-0.097
Arousal 1	0.051	0.207	**0.599**	0.057
Arousal 2	0.305	0.164	**0.714**	-0.240

Table 6: Correlations for participants split by musical preference. Bold correlation coefficients depict that they are statistically significant.

could also argue that participants were emotionally not affected that much in contrast to the arousal.

An indication exists that females react better on arousal but worse on valence compared to males. Also, people who like dance seem to perform much better on the valence scale in comparison to people who like other music genres. But then again, they also perform less on most of the valence slices compared to the average.

The system that generates the music can be used real-time and continuous. This enables it to be used in a context of for example the Virtual Storyteller in which emotion can gradually and continuously be altered to keep up with the story.

Future work would include implementation in a real application for a more realistic evaluation. Also, it would be worthy to look into the influences of features on emotion individually.

8. REFERENCES

[1] M. Alfonseca, M. Cebrián, and A. Ortega. A simple genetic algorithm for music generation by means of algorithmic information theory. In *2007 IEEE Congress on Evolutionary Computation, CEC 2007*, pages 3035–3042, 2008.

[2] P. Bottoni, A. Labella, S. Faralli, M. Pierro, and C. Scozzafava. Interactive composition, performance and music generation through iterative structures. In *14th Annual ACM International Conference on Multimedia, MM 2006*, pages 189–192, Santa Barbara, CA, 2006.

[3] M. Bradley. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1):49–59, 1994.

[4] D. Cabrera, S. Furguson, F. Rizwi, and E. Schubert. Psysound3, `http://psysound.wikidot.com/`.

[5] G. Campolongo and S. Vena. Science of networks and music: A new approach on musical analysis and creation. In *EvoWorkshops 2006: EvoBIO, EvoCOMNET, EvoHOT, EvoIASP, EvoINTERACTION, EvoMUSART, and EvoSTOC*, volume 3907 LNCS, pages 642–651, Budapest, 2006.

[6] R. Cruz, A. Brisson, A. Paiva, and E. Lopes. I-sounds emotion-based music generation for virtual environments. In *2nd International Conference on Affective Computing and Intelligent Interaction, ACII 2007*, volume 4738 LNCS, pages 766–767, Lisbon, 2007.

[7] M. Delgado, W. Fajardo, and M. Molina-Solana. Inmamusys: Intelligent multiagent music system. pages –, E.T.S.I. Informa?tica y de Telecomunicacio?n, CPeriodista Daniel Saucedo

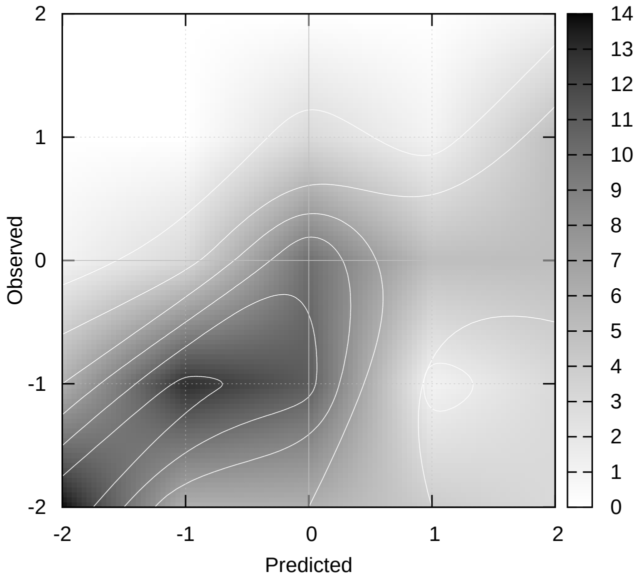

Figure 9: Contour plot for valence 1 ($r = 0.476$)

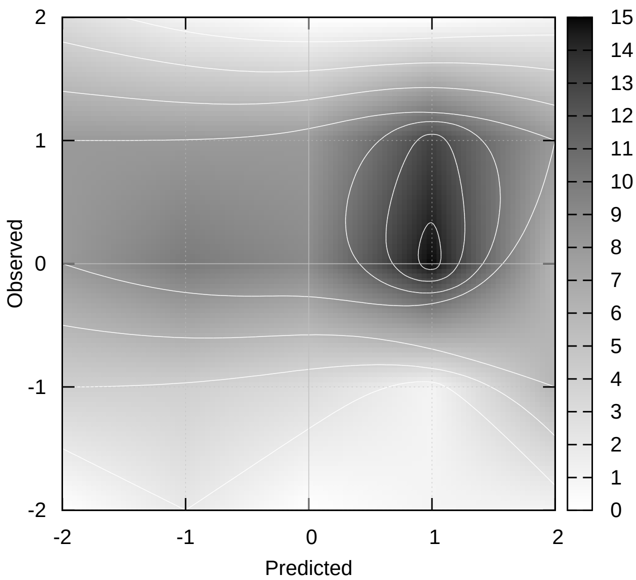

Figure 10: Contour plot for arousal 0 ($r = -0.059$)

Aranda, 18071 Granada, Spain.

[8] J. Flury and D. Bisig. Celerina - a generative music system using aesthetical reduction applied to simple cellular automata. In *FLAIRS 2006 - 19th International Florida Artificial Intelligence Research Society Conference*, volume 2006, pages 237–242, Melbourne Beach, FL, 2006.

[9] J. Franklin. Jazz melody generation using recurrent networks and reinforcement learning. *Int. J. on Artif. Intell. Tools*, 15(4):623–650, 2006.

[10] Y. Khalifa, B. Khan, J. Begovic, A. Wisdom, and A. Wheeler. Evolutionary music composer integrating formal grammar. In *9th Annual Genetic and Evolutionary Computation Conference, GECCO 2007*, pages 2519–2526, London, 2007.

[11] M. Lo and S. Lucas. Evolving musical sequences with n-gram based trainable fitness functions. In *2006 IEEE Congress on Evolutionary Computation, CEC 2006*, pages 601–608, Vancouver, BC, 2006.

[12] E. Lopes. *Just in Time: Towards a theory of rhythm and metre.* PhD thesis, University of Southampton, 2003.

[13] R. Lopez De Mantaras and J. Arcos. Ai and music: From composition to expressive performance. *AI Mag*, 23(3):43–57, 2002.

[14] S. Ness. Marysas, `http://marsyas.sness.net/`.

[15] E. Özcan and T. Erçal. A genetic algorithm for generating improvised music. In *8th International Conference on Artificial Evolution, EA 2007*, volume 4926 LNCS, pages 266–277, Tours, 2008.

[16] R. Ramirez and A. Hazan. A tool for generating and explaining expressive music performances of monophonic jazz melodies. *Int. J. on Artif. Intell. Tools*, 15(4):673–691, 2006.

[17] P. Reiners. Cellular automata and music, `http://www.ibm.com/developerworks/java/library/j-camusic/`.

[18] E. Schubert. *Measurement and time series analysis of emotion in music.* PhD thesis, School of Music & Music Education, University of New South Wales, Sydney, Australia, 1999.

[19] J. Thayer. *Multiple indicators of affective response to music.* PhD thesis, 1986. New York University.

[20] E. W. Weisstein. Rule 60. From MathWorld–A Wolfram Web Resource. `http://mathworld.wolfram.com/Rule60.html`.

[21] M. Yee-King. An automated music improviser using a genetic algorithm driven synthesis engine. In *EvoWorkshops 2007: EvoCOMNET, EvoFIN, EvoIASP, EvoINTERACTION, EvoMUSART, EvoSTOC and EvoTRANSLOG*, volume 4448 LNCS, pages 567–576, Valencia, 2007.