

Music Mood Player Implementation Applied In Daycare Using Self Organizing Map Method

Kadek Cahya Dewi¹, Luh Arida Ayu Rahning Putri²
Department of Computer Science, Udayana University
Kampus Bukit Jimbaran, Denpasar 80361, Bali, Indonesia
Email: cahya.dewi@cs.unud.ac.id¹, luh.arida@cs.unud.ac.id²

Abstrak. *Implementasi Music Mood Player Yang Diaplikasikan Pada Daycare Menggunakan Metode Self Organizing Map. Musik adalah seni, hiburan dan aktivitas manusia yang melibatkan suara-suara yang teratur. Musik berkaitan erat dengan psikologi manusia. Sepotong musik sering dikaitkan dengan kata sifat tertentu seperti senang, sedih, romantis, dsb. Keterkaitan antara musik dengan mood tertentu ini telah banyak digunakan dalam berbagai kesempatan oleh manusia, untuk itu klasifikasi musik berdasarkan keterkaitannya dengan emosi tertentu menjadi penting. Daycare merupakan salah satu lembaga yang memanfaatkan musik sebagai terapi atau sarana pendukung dalam kegiatan pengasuhan anak. Penelitian ini fokus pada implementasi music mood player menggunakan Self Organizing Map yang diaplikasikan pada Daycare. Fitur yang digunakan sebagai ciri adalah rhythm pattern dari musik tersebut. Parameter mood didapatkan dari Robert Thayer's energy-stress model yang terdiri dari exuberance / gembira, contentment / rileks, anxious / cemas dan depression. Sistem diuji dengan lagu dari berbagai genre dan mood hasil klasifikasi dibandingkan dengan mood dari pakar psikologi anak. Mood lagu dari sistem dapat diset secara otomatis disesuaikan dengan aktivitas pada daycare.*

Kata Kunci: *Music Information Retrieval, Klasifikasi Mood, Klasifikasi Musik, Self Organizing Map, Rhythm Patterns.*

Abstract. Music is an art, entertainment and human activity that involve some organized sounds. Music is closely related to human psychology. A piece of music often associated with certain adjectives such as happy, sad, romantic and many more. The linkage between the music with a certain mood has been widely used in various occasions by people, there for music classification based on relevance to a particular emotion is important. Daycare is one example of an institution that used music as therapy or tools of support in each of its parenting activities. This research concerns in implementation of a music mood player using Self Organizing Map applied at the Daycare. The features that are used on this music mood player are rhythm patterns of the music. The mood parameters that used in this system is based on Robert Thayer's energy-stress model which are exuberance / happy, contentment / relax, anxious and depression. The system is tested using a set of songs with various genres and the classification results are compared with the mood obtained by child psychology expert. The system can be set automatically according to the activities at daycare.

Keywords: Music Information Retrieval, Mood Classification, Music Classification, Self Organizing Map, Rhythm Patterns.

1. Introduction

Music is an art, entertainment and human activities that involve the voices of regular.

Music is all the possibilities that could happen to the voices / sounds and silence to be organized into a series of meaningful hearing. The meaning is not acquired verbal meaning, but the aural sense. The meaning of aural harmony means the perceived sound when listening to music. For example when listening a traditional song 'Yamko Rambe Yamko' which originated from Irian Jaya, listeners may not understand the intent of the song because they do not know the language that used in the lyrics or do not understand what instruments are played. However, when we hear the music, it is very possible as a matter of making nice to hear, making enthusiastic, sadness, grief or perhaps touching. Another example of the aural sense is when listening to an instrumental song.

Music is closely related to human psychology. A piece of music often associated with certain adjectives such as happy, sad, romantic, etc. The linkage between the music with a certain mood has been widely used in various occasions by people. For example in a musical movie, music is used to reinforce the atmosphere of the specific scene, the dramatic music is used for background suspense scene, music scene eager for war, the fun music is used as a background scene of humor, etc. In addition to support smart parenting programs will be very useful if the mothers know and understand the classification of music based on mood parameters, so they can do song selection in accordance with the desired conditions / moods. For example when the children wake up just choose the happy song (full of spirit). There are many more examples that cannot be mentioned here.

A number of researches on music classification based on mood have been conducted. Feng et.al. classified music based on Dixon's beat detection [Feng et.al, 2003]. Li and Ogihara classified music using Support Vector Machines (SVM) [Li and Ogihara, 2004]. Leman et.al. based his classification on three level analysis: subjective judgments to manual-based musical analysis to acoustical-based feature analysis [Leman et.al., 2004]. Other researchers include Wang et.al. which uses Support Vector Machines (SVM) [Wang et.al., 2004], Wiczorkowska et.al. with K-Nearest Neighbor [Wiczorkowska et.al.] and Baum with Naïve Bayes, Random Forest and Support Vector Machines [Baum, 2006].

This paper will discuss the classification of music based on relevance to a particular emotion/mood. It will compare the rhythms pattern of music. All feature sets are based on fluctuation of modulation amplitudes in psycho acoustically transformed spectrum data, namely the combination of rhythm patterns, rhythm histograms and statistical spectrum descriptors of the music. The system is tested using a set of songs with various genres and the classification results are compared with the mood obtained by child psychology expert. Clustering and classification method obtained by Self Organizing Map method.

The research looked at the phenomenon at the daycare that uses music as a companion in childcare in order to balance the right brain and left brain. As seen in one of the Daycare in Denpasar that manage the type of music adjusted to the child's activities on the schedule of activities from 7 am to 5 pm in order to use music as the support tools on parenting. The type of music is refers to the aural sense or mood of the music. The system will be able to set automatically the mood on the specific time based on the schedule, due to set the best mood for the activities.

2. Analysis and Design System

The software that was built for this research is called Music Mood Player. The emotion model that used in the system is the Robert Thayer's energy-stress model [Meyers, 2007] which consists of exuberance / happy, contentment / relax, anxious / anxiety and depression. Music Mood Player allows users to perform clustering and classification based on the song's mood. The users can also use the Music Mood Player to play songs and manage the playlist.

The users can perform clustering and classification on the data track. The system will perform feature extraction process before clustering and classification process. Classification process can be implemented if the clustering has been carried out. In addition this system also allows the user to manage the music player like play, stop and pause music and manage the

playlist like show playlist, add, remove songs and setting automatic the playlist. The concept map of the Music Mood Player is shown in Figure 1.

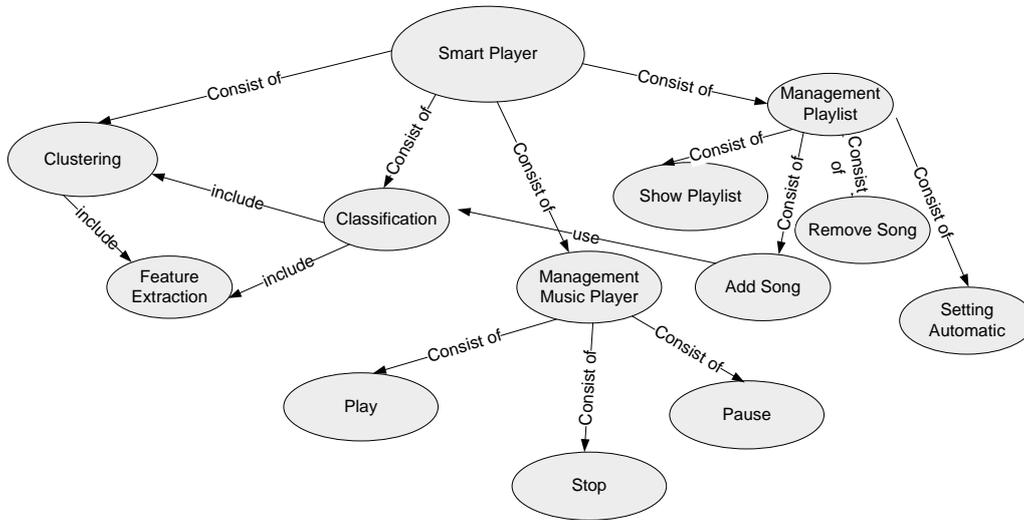


Figure 1. Concept Map of Music Mood Player

As shown in Figure 1, Music Mood Player consists of several main processes. The feature extraction process is the process to get the pattern of a song. The following is the feature extraction process in accordance with Rauber’s research [Rauber et.al., 2002] and the block diagram of audio extraction is shown by Figure 2.

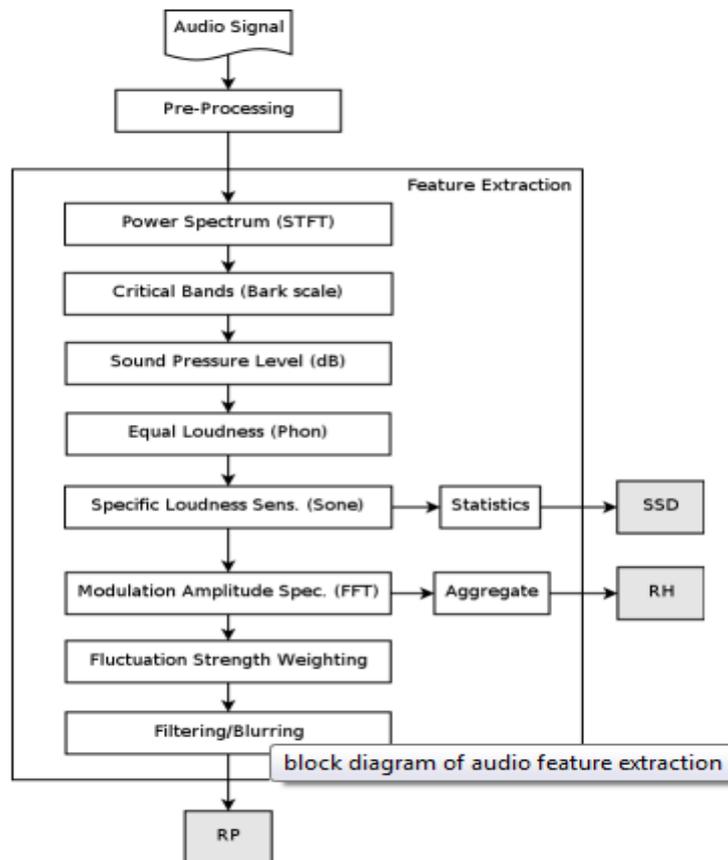


Figure 2. Block Diagram Of Audio Feature Extraction

Preprocessing: (1) Audio decoding. Converting audio format into Pulse Code Modulation (PCM) form. (2) Audio quality reduction. Audio quality is reduced from stereo to mono (sound exchange). Decreasing the number of channels from two to one and is done by taking the average of the two channels. Audio sampling rate is changed from 44 KHz to 11 KHz. (3) Splitting music into segments with each segment of size 6s in. The time 6s is deemed to have enough to get the impression from the style of a piece of music.

Feature Extraction for Rhythm Patterns: (1) Performing FFT transform into a spectrogram. The FFT window size used is 256 samples to meet the 23ms sampling of mp3 (253 samples). (2) Bundle the frequencies into 24 critical frequency bands to meet the Bark scale. Bark scale ranges from 1 to 24 Barks. (3) Calculate the spectral masking effects with spreading function. (4) Transform into decibel, to form a base 10 logarithmic scale. (5) Calculate the equal loudness level in Phon. Forty Phon = 40 db-SPL tone at 1 kHz frequency. (6) Calculate loudness sensation. One SONE = 1 kHz tone at 40 db-SPL. (7) Perform SONE transformation into rhythm patterns with the FFT. (8) Limit the amplitude modulation to 60. It means that for every 24 critical bands 60 values is obtained for modulation frequencies between 0 to 10Hz. This results in 1440 values representing the fluctuation strength. (9) Filtering rhythm patterns using the Gaussian and the gradient.

During feature extraction for the Rhythm Patterns, it was computed a Statistical Spectrum Descriptor (SSD) for the 24 critical bands. From the SONE representation of the spectrum (Sonogram), we compute the following statistical moments for each critical band: mean, median, variance, skewness, kurtosis, min-value and max-value, resulting in a 168-dimensional feature vector. Contrary to the Rhythm Patterns and the SSD, this feature set does not contain information per critical band. The magnitudes of each modulation frequency bin of all 24 critical bands are summed up in order to form a histogram of modulation magnitude per modulation frequency. This feature set contains 60 attributes, according to modulation frequencies between 0.168 and 10 Hz.

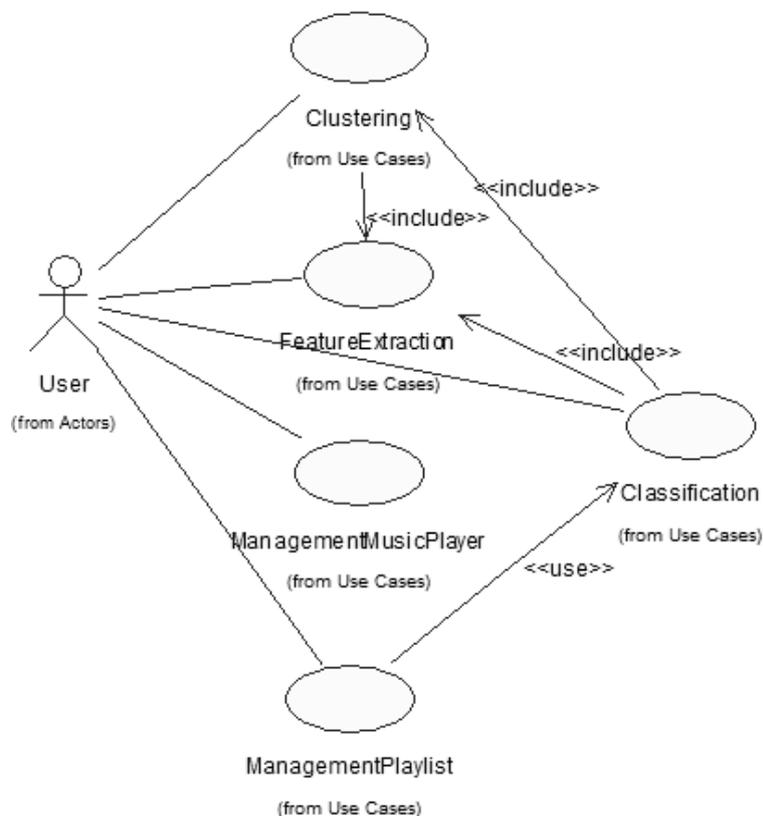


Figure 3. Use Case Diagram of the *Music Mood Player*

Clustering is done by Self Organizing Map method and is carried out after the feature extraction process. The SOMeJB toolbox is used for clustering and visualization of the clustering results. Here are the parameters required for clustering: (1) The input is a 2-dimensional matrix $n \times d$, where n is the number of data and d is the data dimensions resulted from feature extraction. (2) The output layer is a 2-dimensional matrix with the number of map units equals the number of terms / words on the Thayer emotion model. (3) Calculate the Euclidian distance. (4) Models the emotion as Thayer's emotion model which consists of exuberance, contentment, anxious and depression. The Thayer's emotion terms are exuberant, triumphant, carefree, anxious, frantic, terror, content, serene, ominous, and depression. (5) Visualization of the clustering results with smoothed data histogram.

Classification is carried out using the Self Organizing Map. Classification is carried out after the feature extraction and clustering. Classification with the SOM method is performed by entering testing data on the network and labeling moods according to the results of the clustering. Classification process is done by SOMeJB Toolbox. The results of the classification using the SOM methods were compared with the results obtained by child psychology expert.

Figure 3 shows the use case diagram which consists of one actor (the user) and five use cases (Feature Extraction, Clustering, Classification, Management Music Player and Management Playlist). The implementation of the use cases are described in the class diagram which captures the structure of all classes that make up the system. Figure 4 is a class diagram of the Music Mood Player which consists of four classes, one class boundary and three class control.

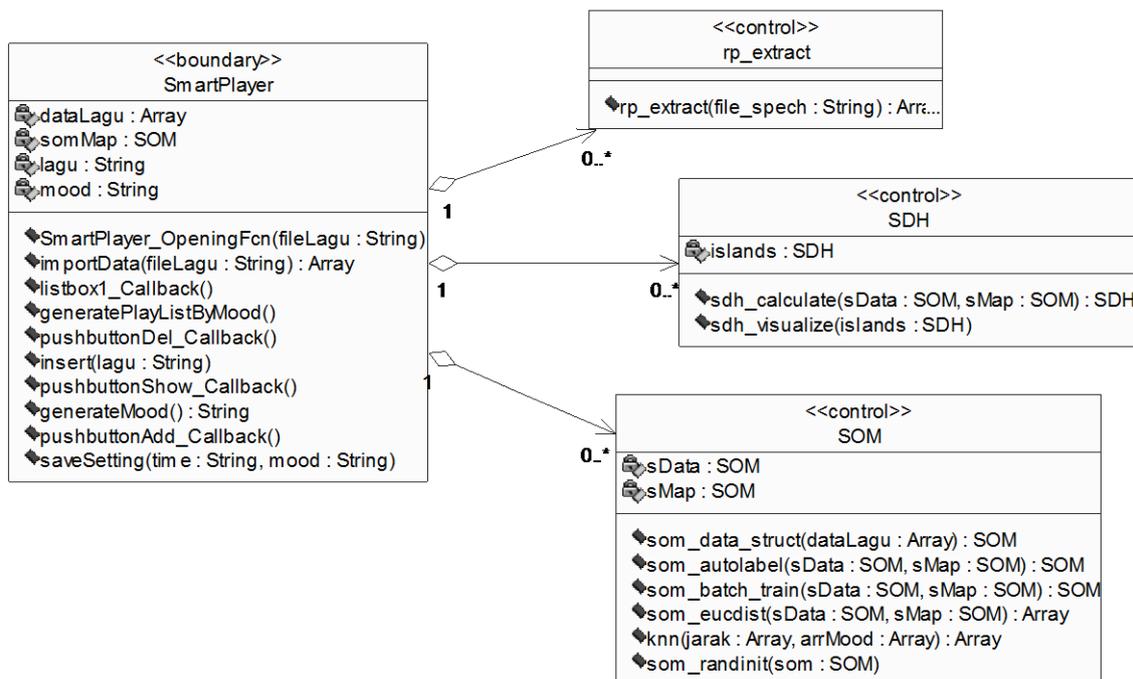


Figure 4. Class Diagram of the Music Mood Player

The system was built using four toolboxes, Windows Media Player ActiveX Control toolbox, RP_Extract toolbox, SOM_toolbox and SOMeJB Toolbox as shown in Figure 5. Application layer is the layer that handles the interaction with the user. Windows Media Player ActiveX Control is used which handles the music player and the playlist management of the Music Mood Player. From the application layer, the toolboxes on the process layer are called by the Music Mood Player from its function. Those are the RP_Extract toolbox (for feature extraction), SOMeJB toolbox to handle clustering function and SOM toolbox for dealing with classification and calls the ODBC driver on the data layer as a connection to the database.

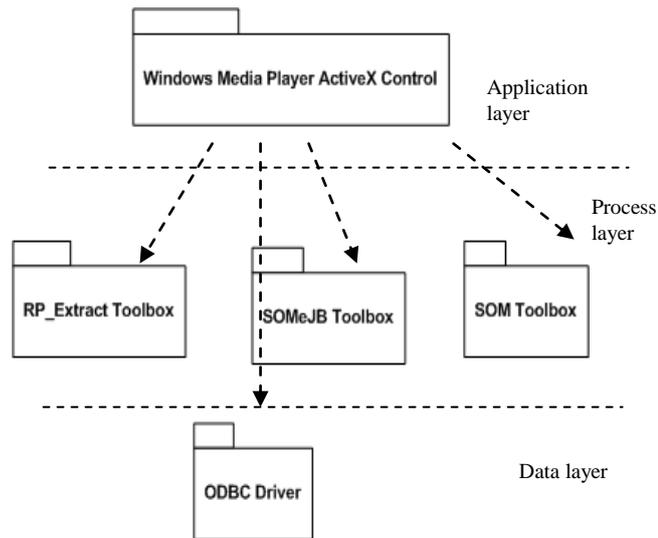


Figure 5. Architecture of the *Music Mood Player*

In Music Mood Player song data are stored in the database using two tables namely Song and MoodSetting tables as shown in Figure 6. Table Song has four fields. Those are *kode* as the primary key, *pathLagu* describing the address where the song is stored on the storage, *title* which describes the song's title and *mood* that describes the mood of the song. Table MoodSetting only has 2 fields, namely *time* that describes the time and *mood* that describes what kind of mood will be set for a specific time.

Song	
PK	<u>kode</u>
	pathSong title mood

MoodSetting	
PK	<u>time</u>
	mood

Figure 6. The Song data of the *Music Mood Player*

3. Research Results and Discussion

The Music Mood Player is tested with 135 songs in both Indonesian and English languages, and there are also some songs taken from the soundtrack of the children movies. 120 songs of the 135 songs are used for clustering process and are used as the basis of classification and as the training data on the SOM method. Fifteen other songs are used in classification process as the testing data. Feature extraction can be done in two ways, extraction with a number of songs at once and extraction of an individual song. Feature extraction involving many songs is performed at the initial data clustering. Feature extraction for one song is performed at the time of classification. Each song results in a feature vector of size 1 x 1440 as mentioned earlier. The extracted rhythm pattern is then stored in song.txt and the song list is stored in songlist.txt.

Clustering process was executed by 120 songs and hence the SOM data is 120x1440 and SOM label is 120x1. With the map size of 2x5 that produces 10 map units the SOM final quantization error is 1,011 and the final topographic error is 0,625. The clustering results are visualized in the form of islands of music as shown in Figure 7 where the islands are labeled with mood. The songs are placed according to the level of sound similarity and islands formed in accordance with the cluster density. Islands of music using a color code with color levels

range from dark blue (deep sea) to light blue (shallow water) to yellow (beach) to dark green (forest) to light green (hills) to gray (rocks) and white (mountain). Each island which is marked by passage of land / forest represents a different style of music (different cluster). Mountains indicate the center of the cluster, so if there is an island with more than one mountain, it can be said as a style in a style (sub-style).

The results of islands of music with Smoothed Data Histograms (SDH) strongly depend on the value of 'spread' attribute on SDH. In this research the spread is selected as 1. Spread parameters on the islands of music shows membership degree. The result shows that there are 2 main cluster with sub-cluster (the mountain) for each cluster which indicates the existence of differences in style / sub-cluster on the island.

The results of the classification using SOM methods were compared with the results obtained by child psychology expert as shown in Table 1. The testing result gives 86.67% accuracy (compared with the results of the child psychology expert). The classification results shown that the same song with different arrangements can be classified into a different mood. Like the song of Ibu Pertiwi with 3 different arrangements which are classified into 3 different moods. The song "Ibu Pertiwi mentah.mp3" with the mood of depression, "Ibu Pertiwi.mp3" with the mood of contentment and "Kulihat Ibu Pertiwi.mp3" with the mood of exuberance.

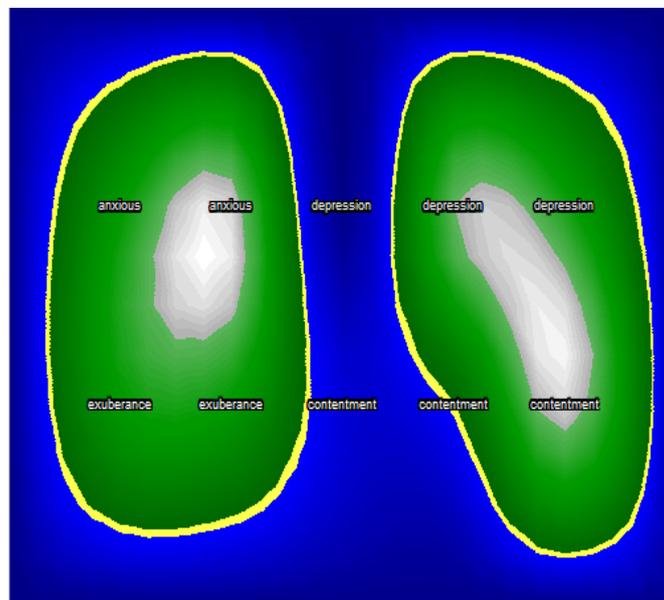


Figure 7. Islands of Music

The management of music player and management of playlists are tested by displaying the songs in a playlist according to the mood, add songs, delete songs and show song's description of the selected songs and play songs in a playlist as shown in Figure 8. Figure 8 will also show automatic setting of the Music Mood Player. User can choose the mood of music by checking the mood checkbox. To play the music just double click the song on the playlist or click play button after choosing the song. Add song button  function is for adding a song to the playlist. If the song was not already classified, the system will classify the song into specific mood first and then add the song to the playlist. Delete songs button  will remove selected song from playlist. When user choose one song from playlist, the song description will show on the Song Description panel. The system is able to automatically play songs in a specific mood based on time that was already set before. User can set the mood on specific time from Automatic Setting panel.



Figure 8. Music Mood Player

Table 1. Comparison of classification results

No	Song's Title	Mood SOM	Mood Psychology
1	01 Armada Masa Depan.MP3	1	1
2	03. Believe - Melody.mp3	1	1
3	04. Wonderful Me.mp3	1	1
4	07. Believe - Liana & Alexa.mp3	1	1
5	11 SpongeBob Patrick & The Monster - Now That Were Men.mp3	3	3
6	Bob Tutupoli-Widuri.mp3	4	2
7	Medley Ulang Tahun.mp3	1	1
8	Tasya - _Di Stasiun_ & _Kereta Apiku.mp3	1	1
9	The Lion King -02- I Just cant wait to be king.MP3	2	1
10	01. Menghujam Jantungku.mp3	1	1
11	01 Walau Habis Terang.MP3	1	1
12	08. SHE'S THE ONE - ROBBIE WILLIAMS.mp3	2	2
13	05 Tak Bisakah.MP3	1	1
14	Melly Goeslaw feat. Baim - Catatanku (OST Sinetron Buku Harian Baim).mp3	4	4
15	Lifehouse - You And Me	2	2

Table info : 1 : Exuberance, 2 : contentment, 3 : anxious, 4 : depression

4. Conclusions

The research results show that the implementation of Music Mood Player Applied In Daycare could be developed based on Self Organizing Map. In this research, music is classified based on sound similarity obtained from the rhythm patterns based on mood parameters according to Thayer's theory. Clustering results using Self-Organizing Map method is influenced by the number of map units used. Classification of music based on mood parameters by the method of Self Organizing Map with 120 songs has reached 86.67% accuracy. Songs with different arrangements can be classified into a different mood. The system could be set automatically to play songs in a specific mood on the specific time depend on the Daycare activities schedule.

References

- Baum, D., 2006, EmoMusic - Classifying Music According to Emotion, *Proceedings of the 7th Workshop on Data Analysis (WDA2006)*, Kosice Slovakia.
- Baum, D., Rauber, A., 2006, Emotional Descriptors for Map-based Access to Music Libraries, *Proceedings of the 9th International Conference on Asian Digital Libraries*, Kyoto Japan.
- Cunningham, P and Delany, S. J., 2007, k-Nearest Neighbour Classifiers, *technical Report UCD-CSI-2007-4*, Dublin.
- Feng, Y., Zhuang, Y., and Pan, Y., 2003, Music information retrieval by detecting mood via computational media aesthetics, *Proceedings of the IEEE/WIC International Conference on Web Intelligence*, Washington USA.
- Leman, M., Vermeulen, V., Voogdt, L. D., and Moelants, D., 2004, Using audio features to model the affective response to music, *Proceedings of the International Symposium on Musical Acoustics*, Nara Japan.
- Li, T. and Ogihara, M., 2004, Content-based music similarity search and emotion detection, *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, -.
- Lu, L., Liu, D., and Zhang, H., 2006, Automatic mood detection and tracking of music audio signals, *Proceedings of the IEEE Transactions on Audio, Speech and Language Processing*, -.
- Meyers, O., 2007, *A Mood-Based Music Classification and Exploration System*, Master of Science in Media Arts and Sciences, Massachusetts Institute Of Technology, United States.
- Pampalk, E., 2001, *Islands of Music: Analysis, Organization and Visualization of Music Archives*, Department of Software Technology and Interactive Systems, Vienna University of Technology, Vienna.
- Pampalk, E., Rauber, A., Merkl, D., 2002, Content-based Organization and Visualization of Music Archives, *Proceedings of ACM Multimedia*, Juan-les-Pins France.
- Pampalk, E., Rauber, A., Merkl, D., 2002, Using Smoothed Data Histograms for Cluster Visualization in Self-Organizing Maps, *Proceedings of the Intl Conf on Artificial Neural Networks (ICANN 2002)*, Madrid Spain.
- Rauber, A., Pampalk, E., Merkl, D., 2002, Using Psycho-Acoustic Models and Self-Organizing Maps to Create a Hierarchical Structuring of Music by Musical Styles, *Proceedings of the 3rd International Conference on Music Information Retrieval (ISMIR 2002)*, Paris France.
- Siang, J J, 2005, *Jaringan Syaraf Tiruan & Pemogramannya Menggunakan Matlab*, ANDI, Yogyakarta.
- Skowronek, J., McKinney, M. F., and Van de Pa, S., 2006, Ground truth for automatic music mood classification, *Proceedings of the 7th International Conference on Music*

Information Retrieval, Victoria Canada.

- Tellegen, A., Watson, D. and Clark, L., 1999, On the dimensional and hierarchical structure of affect, *Psychological Science*,-.
- Tolos, M., Tato, R. and Kemp, T., 2005, Mood-based navigation through large collections of musical data, *Consumer Communications and Networking Conference*, Las Vegas USA.
- Tzanetakis, G. and Cook, P., 1999, Marsyas: a framework for audio analysis, *Organized Sound*, 4(3):169–175.
- Wang, M., Zhang, N. and Zhu, H., 2004, User-adaptive music emotion recognition, *Proceedings of the International Conference on Signal Processing*, Istanbul Turkey.
- Wieczorkowska, A. , Synak, P., Lewis, R., and Ras, Z., Extracting emotions from music data, *Proceedings of the 15th International Symposium on Methodologies for Intelligent Systems*, USA.