Mood Cloud : A Real-Time Music Mood Visualization Tool

Cyril Laurier and Perfecto Herrera

Music Technology Group
Universitat Pompeu Fabra
Ocata 1, 08003 Barcelona, Spain
{claurier, pherrera}@iua.upf.edu

Abstract. We present Mood Cloud, an application of automatic music mood prediction from audio content. While playing a song, we visualize in real-time the prediction probabilities of five mood categories : "happy", "sad", "aggressive", "relax" and "party". Each mood is represented by a colored bar graph with text, dynamically resized according to the mood probability. The resulting application is a dynamic visualization of the mood predictions, demonstrating the performance of current state of the art techniques in Music Information Retrieval and especially in automatic mood classification.

1 Introduction

In the past few years, research in Music Information Retrieval (MIR) has been very active to produce automatic classification methods. These algorithms aim to ease the process of annotating data, as we are faced with a large and increasing amount of digital music. In this paper we address the task of classifying music by mood from audio content. This problem has inspired several relevant studies in the MIR field[1–8]. In 2007, the MIREX[1] contest has for the first time organized the "Audio Music Mood Classification" task. This demonstrates the interest of the MIR community about this topic. Indeed, this kind of categorization is particularly pertinent because of its link with one of the primary motive to listen to music: to feel emotions[9]. Although automatic music mood classification seems a difficult task, recent studies have shown that it is possible to a certain extent. It can give satisfying results if we consider a few simple categories and if we check for valid agreements between people. The goal of this Mood Cloud project is to create a meaningful mood visualization tool in real-time. It aims to prove the potential of current state of the art algorithms to predict the mood in music. In the remainder of this paper, we will introduce the Mood Cloud application and explain the methods employed to build the database of examples (ground truth), extract the audio features and classify the data. We will detail the concept of the application, present screenshots and finally conclude with ideas for future works.

[1] The Music Information Retrieval Evaluation eXchange is a community-based formal evaluation framework, see[10] for more details.

2 Mood Cloud

In this section, we describe the methods used to create the ground truth, extract the features and compute the classifiers. Then we detail how the visualization tool has been designed and implemented, lastly we display screenshots of the application.

2.1 Ground Truth

In order to make our application aware about mood concepts, we need to train it with examples. The database used for this project is made of mainstream popular music. The mood categories we have chosen are : “happy”, “sad”, “aggressive”, “relax” and “party”. For each mood category we have generated a synonym set using Wordnet[2] and looked for the songs mostly tagged with these terms in the last.fm[3] website. We asked annotators to validate this selection by listening to playlists made with these tracks. We considered a song to be valid if the category was confirmed by, at least, one listener, as the pre-selection from last.fm granted that the song was likely to deserve that tag. We asked the annotators to concentrate on audio as much as possible in order to avoid a bias due to the lyrics. The resulting database is composed of 1300 songs divided between 5 categories of interest plus their complementary categories (“not happy”, “not sad”, “not aggressive” , “not relaxed” and “not party”). Each mood class is seen as a binary problem for which we obtained an equal distribution of instances.

2.2 Feature Extraction and Classification

The first step to classify the music from acoustical information is to extract audio features. We have computed different kind of features: timbral (for instance MFCC, spectral centroid), rhythmic (for example tempo, onset rate), tonal (like Harmonic Pitch Class Profiles) and temporal descriptors. All these descriptors are standard and derived from recent research in Music Information Retrieval[11]. For each song, the frame-basis extracted features were summarized with their average and variance. In order to classify the data we have selected an algorithm called Support Vector Machines (SVM) known to be efficient and to provide relatively convincing results. In our case it is even more relevant because our problems are binary (for instance “happy” versus “not happy”) and the SVM are designed to solve binary problems. We computed the classifiers with the LibSVM library[12]. In table 1, we report the mean accuracies using the SVM algorithm for each category and its complementary on 10 runs of 10-fold cross-validation. With a satisfying accuracy of 86.22% in average, we consider that these models can be exploited in our context. As we wanted a continuous value and not only a binary decision, we used the probability value provided by the SVM algorithm.

[2] http://wordnet.princeton.edu

[3] http://www.last.fm, Last.fm is a music recommandation website that has a large community very active in associating labels (tags) with music they listen to.

Category vs Not Category	SVM Accuracy
Aggressive vs Not Aggressive	94.63% (4.06)
Happy vs Not Happy	75.45% (8.54)
Sad vs Not Sad	82.87% (8.22)
Relax vs Not Relax	91.57% (5.17)
Party vs Not Party	86.59% (6.53)

Table 1. Mean accuracy in percentage of the SVM algorithm for each category, in 10 runs of 10-fold cross validation. In parenthesis is the standard deviation.

2.3 The Mood Cloud Application

This visualization tool pre-computes the mood evolution by means of a supervised learning approach, using a feature set designed for this task and a SVM algorithm as described in section 2.2. To estimate the mood probabilities, we use SVM models trained on our ground truth data. The predictions are computed within windows of several seconds to show the evolution of the mood prediction. While playing a music, bar graphs representing each mood are resized according to the predicted probability value on the current audio segment. To enhance the user's experience, the color of each mood bar was inspired from Bresin's results[13]. In table 2, we show the color associated with each emotion (when it matched with one of our categories). For the "relax" category we have chosen a color close to the "sad" and "tenderness" proposed colors. For the "party" category, we have elected a color close to "happiness" and "contentment".

Emotion	Hue value (Color)	Category in Mood Cloud
Anger	0 (Red)	Aggressive
Happiness	0.167 (Yellow)	Happy
Sadness	0.75 (Violet)	Sad

Table 2. Hue values related to Emotion categories (from Bresin[13]) and associated category in Mood Cloud.

Technical details: The Mood Cloud application is divided in two parts. The first part is the processing module that extracts the features and outputs probabilities for each segments. It uses the LibSVM library with SVM models precomputed on our ground truth. The processing module is made in Python and C++ and runs under Linux, however the source code is cross-platform. The second part is the visualization module, created with Adobe Flash[4]. It can be run on any platform if an Internet browser with Flash player is installed. The interaction between both modules is achieved via XML.

[4] http://www.adobe.com/products/flash

2.4 Screenshots of the Mood Cloud Application

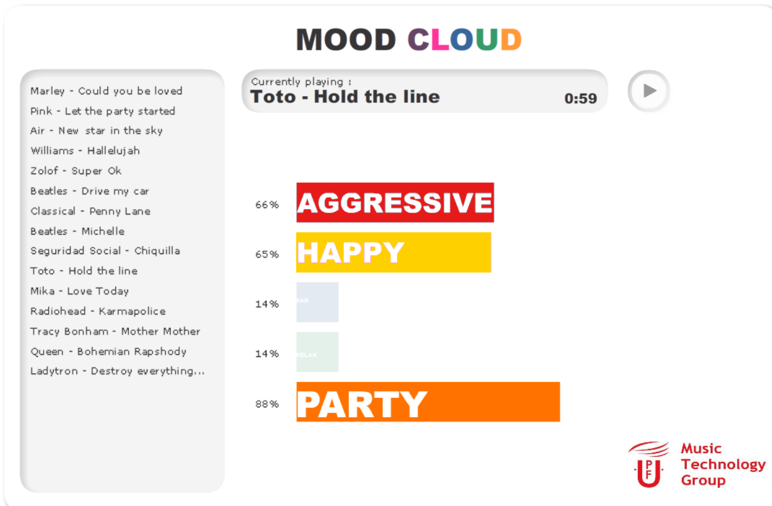

Fig. 1. Screenshot of Mood Cloud for the song "Hold the line" by Toto.

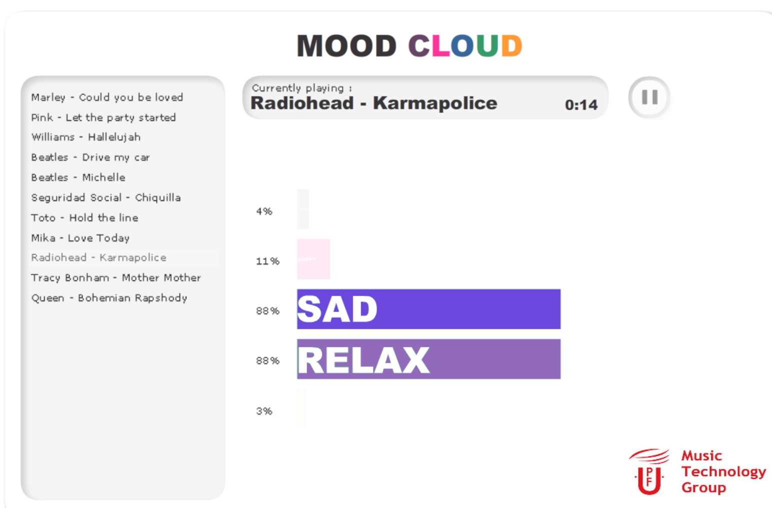

Fig. 2. Screenshot of Mood Cloud for the song "Karmapolice" by Radiohead.

3 Conclusion

The Mood Cloud application provides an intuitive and understandable visualization of automatic mood classification algorithms. This tool helps to understand the potential and the limitations of such techniques. A future work would be to integrate it as a plug-in for music players like Skowronek in[6]. The next step would be to allow the user to tune the algorithm to his own perception using relevance feedback. It would create a personalized version of the classifiers and mood concepts. We could also let the user define his own categories. Finally we can consider to have a model inspired by psychological studies like for instance the theory of expectation from Huron[14].

References

1. Li, M., Ogihara, T.: Detecting emotion in music Proceedings of the International Symposium on Music Information Retrieval Baltimore, USA (2003)
2. Feng, Y., Zhuang, Y., Pan, Y. : Music information retrieval by detecting mood via computational media aesthetics Proceedings of the 2003 IEEE/WIC International Conference on Web Intelligence Washington, DC, USA (2003)
3. Lu, L., Liu, D., Zhang, H.-J.: Automatic mood detection and tracking of music audio signals, IEEE Transactions on Audio, Speech, and Language Processing, vol.14, no.1, pp.5-18, (Jan. 2006)
4. Shi, Y.-Y., Zhu, X., Kim, H.-G., Eom, K.-W.: A tempo feature via modulation spectrum analysis and its application to music emotion classification Proceedings of the 2006 IEEE ICME, pp.1085-1088 Toronto, Ontario, Canada (2006)
5. Mandel, M., Poliner, G, and Ellis, D.: Support vector machine active learning for music retrieval. Multimedia Systems, vol.12-1, (Aug. 2006)
6. Skowronek, J., McKinney, M., van de Par, S.: A Demonstrator for Automatic Music Mood Estimation. Proceedings of 8th International Conference on Music Information Retrieval. Vienna, Austria (2007)
7. Sordo, M., Laurier, C., Celma, O.: Annotating Music Collections: How content-based similarity helps to propagate labels Proceedings of 8th International Conference on Music Information Retrieval. Vienna, Austria (2007)
8. Yang, Y.-H. , Lin, Y.-C., Su, Y.-F., Chen, H.H.: A regression approach to music emotion recognition IEEE Transactions on Audio, Speech, and Language Processing, vol.16, no.2, pp.448-457, (Feb. 2008)
9. Juslin, P.N., Laukka, P.: Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening Journal of New Music Research, vol. 33, no. 3 (2004)
10. Downie, J.S., The music information retrieval evaluation exchange (MIREX), D-Lib Magazine vol. 12-12 (2006)
11. Laurier, C. , Herrera, P.: Audio Music Mood Classification Using Support Vector Machine. MIREX Audio Music Mood Classification contest. ISMIR, Vienna (2007)
12. Chang, C.-C., Lin, C.-J.: LIBSVM: a library for support vector machines. Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm/ (2001).
13. Bresin, R.: What is the color of that music performance? Proceedings of the International Computer Music Conference. Barcelona (2005)
14. Huron, D.: Sweet Anticipation, Music and the Psychology of Expectation. MIT Press (2008)