

# i.MTV - An Integrated System for MTV Affective Analysis

Shiliang Zhang<sup>1,2</sup>, Qingming Huang<sup>1,2</sup>, Qi Tian<sup>3</sup>, Shuqiang Jiang<sup>1</sup>, Wen Gao<sup>1,4</sup>

<sup>1</sup> Key Lab of Intell. Info. Process., Inst. of Comput. Tech., Chinese Academy of Sciences, Beijing 100190, China

<sup>2</sup> Graduate University of Chinese Academy of Sciences, Beijing 100080, China

<sup>3</sup> Department of Computer Science, University of Texas at San Antonio, TX 78249, U.S.A.

<sup>4</sup> Institute of Digital Media, Peking University, Beijing 100871, China

{slzhang, qmhuang, sqjiang, wgao}@jdl.ac.cn, qitian@cs.utsa.edu

## ABSTRACT

In modern time, MTV has become an important favorite pastime to people because of its conciseness, convenience and the ability to bring both audio and visual experiences to audiences. It has become an significant task to develop new techniques for natural, user-friendly, and effective MTV access. In this demo, an integrated system (*i.MTV*) is constructed for MTV Affective Analysis, Visualization, Retrieval, and User Profile Analysis. We not only perform the effective MTV affective analysis, but also propose novel Affective Visualization techniques to make the abstract affective states intuitive and friendly to users. Based on the affective analysis and visualization, MTV affective retrieval and management are achieved. Furthermore, novel methods are proposed for user affective preferences analysis and MTV recommendation.

## Categories and Subject Descriptors

I.2.10 [Artificial Intelligence]: Vision and Scene Understanding – *Video Analysis*.

**General Terms:** Algorithms, Experimentation, Performance

**Keywords:** Affective Content Analysis, Dimensional Affective Model, Affinity Propagation, Affective Visualization

## 1. INTRODUCTION

Similar as music and movie are important favorite pastimes, MTV (Music TV) is an important entertaining media form to people. Furthermore, in recent years, the amount of MTV data has been increasing at an astonishing speed and mobile sets which can play both video and music have become more and more ubiquitous. However, the traditional methods of MTV classification and retrieval based on Artist, Album, and Title are still not intelligent enough and have many limitations when people want to manage or retrieve MTVs using semantic and abstract concepts.

Combining psychological theories and computer science, affective content analysis can analyze the content of multimedia data and recognize the emotion and affective states contained. Consequently, affective content analysis is more consistent with human perception mechanisms and the applications based on it could be more friendly, usable and natural to human. Utilizing the advantages presented by affective analysis, an integrated MTV management system: *i.MTV* is built.

As shown in Figure 1, the framework of *i.MTV* consists of three modules. In the Affective Analysis module, affective modeling is

realized based on the extracted affective features and personalized weights learned from the interactive weight adjustment. The computed affective state of each MTV is noted as a two-dimensional vector (Arousal and Valence [1] standing for the intensity and type of the affective state respectively).

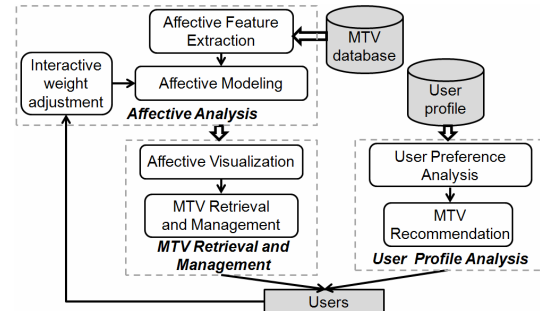


Figure 1. The framework for *i.MTV*

In MTV Retrieval and Management module, novel Affective Visualization techniques are proposed to make affective state intuitive, and thus enable effective and user-friendly MTV retrieval and management. Finally in the User Profile Analysis module, a very recent and effective clustering method: Affinity Propagation [2] is adopted to learn users' preferences from their play histories. With the learned preferences, MTV recommendation is carried out.

## 2. DESCRIPTION OF *i.MTV*

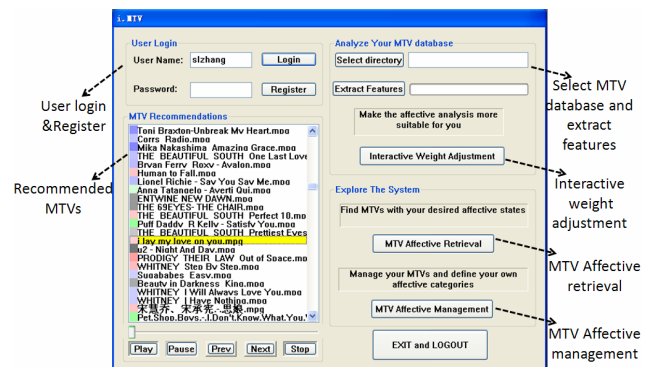
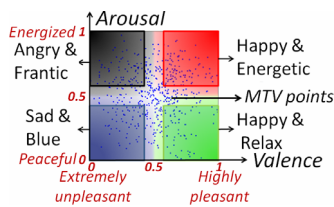


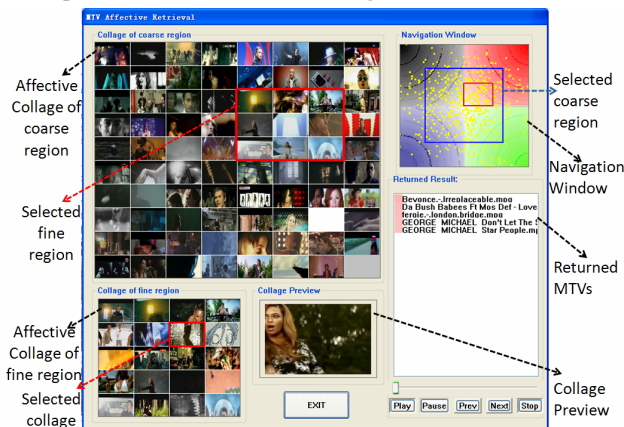
Figure 2. The user interface

In *i.MTV*, after a user logging in, the recorded profile will be loaded and analyzed to capture current user's affective preferences; meanwhile the personalized weights and affective features will also be loaded. Consequently, the affective state and recommendation rate of each MTV in the database will be computed. As shown in Figure 2, MTVs will be sorted by their recommendation rates and listed in "MTV Recommendation" with colors representing their affective states after user logging in.



**Figure 3. MTV points in Dimensional Affective Space**

In the psychological points of view, human emotion can be denoted as a two-dimensional continuous affective space and different regions represent different affective states and emotions [1]. Based on this principle, MTVs could be visualized as points in the 2-dimensional Arousal-Valence space (Figure 3). Thus the affective states of MTVs can be intuitively sensed according to their spatial positions. Referring to the framework in [3], we propose Affective Visualization combining Dimensional Affective Space and Video Collage [4]. Consequently, the proposed Affective Visualization is effective for both affective state representation and MTV browsing.

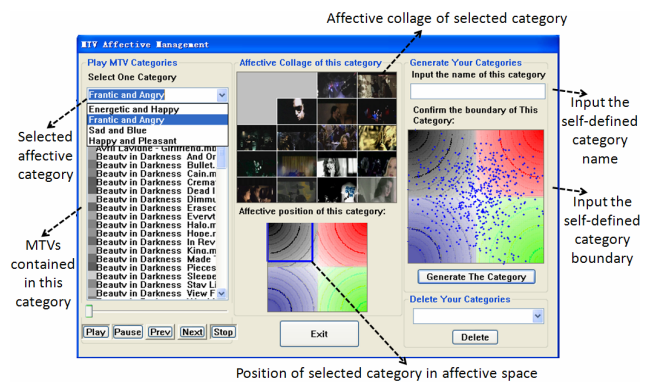


**Figure 4. The interface for MTV affective retrieval**

Figure 4 illustrates the interface for Affective Visualization based MTV affective retrieval. The Navigation Window is constructed based on the Affective Space for three purposes:

- Mapping MTVs into MTV points in the A-V space to give user an intuitive visualization about the affective distribution of his/her MTV collection.
- User can drag mouse and confirm the coarse region from the Navigation Window as the coarse affective query.
- User is informed the positions of the selected coarse and fine regions in affective space.

After selecting a coarse region, Affective Collage is adopted to provide more information to users about the MTVs in the coarse region. As Figure 4 shows, Affective Collage consists of the video key frames extracted from MTVs in the  $10 \times 8$  regions which divide the selected coarse region. With the help of the content and relative position of Affective Collage, users can drag the mouse and select a fine region which they are more interested in. The collages of the selected fine region will then be represented to enable users choose collages in finer scales and preview the selected collages in the Collage Preview Window. Through selecting and browsing collages in the fine region, users can confirm their desired collages. MTVs in the regions covered by these collages will be returned to users.



**Figure 5. The interface for affective category management**

It is convenient for users to define affective categories *themselves* so they can select their desired categories to enjoy. As shown in Figure 5, in *i.MTV*, users can customize categories' names and confirm their boundaries by specifying affective regions in Affective Space. For example, user can define a category "Frantic and Angry" which contains MTVs in the up-left part of the affective space. User can even define a category named "the feeling like I just get up" which contains MTVs in the center part of affective space. After the categories being generated, users can select their desired MTVs to enjoy. The Affective Collage and affective position of the selected category will also be presented. Since the affective categories are defined by users, the categories would be more consistent with users' personal tastes and habits.

### 3. CONCLUSIONS

We have demonstrated *i.MTV*, a prototype system for MTV Affective Analysis, Visualization, Retrieval and User Profile Analysis. As the first of its kind in MTV affective analysis, it provides significant promising applications to MTV lovers. For example, *i.MTV* can be transplanted into music players and mobile phones to improve the entertaining functions. *i.MTV* can also be developed into an internet MTV search engine to enable efficient and user-friendly MTV search. Our future work includes the improvement of current affective modeling and construction of a network based *i.MTV*.

### 4. ACKNOWLEDGEMENT

This work was supported in part by National Natural Science Foundation of China under Grant 60702035 and 60773136, in part by National Hi-Tech Development Program (863 Program) of China under Grant 2006AA01Z117. The authors also gratefully acknowledge the support of K. C. Wong Education Foundation, Hong Kong.

### 5. REFERENCES

- [1] H. Schlosberg. Three dimensions of emotion Psychol. Rev, vol. 61, no. 2, pp. 81-88, Mar. 1954.
- [2] B. J. Frey and D. Dueck. Clustering by passing messages between data points. SECINCE, vol. 315:972-976, Feb. 2007.
- [3] G. P. Nguyen and M. Worring. Optimization of interactive visual similarity based search. ACM Transactions on Multimedia Computing, Communications, and Applications, vol.4, Jan. 2008.
- [4] X. L. Liu, T. Mei, X. S. Hua, B. Yang, and H. Q. Zhou. Video collage. In ACM International Conference on Multimedia, pp. 461-462, 2007.