USING GAMES TO COLLECT SEMANTIC ANNOTATIONS OF MUSIC

Douglas Turnbull[1], **Ruoran Liu**[1], **Luke Barrington**[2], **Gert Lanckriet**[2]
Dept. of Computer Science and Engineering[1]
Dept. of Electrical and Computer Engineering[2]
University of California, San Diego

ABSTRACT

TO BE WRITTEN LAST: (1) Develop *human computation* games that collect semantic labels (i.e., semantically relevant words) for music. (2) This technique used to collect semantic information about images. (3) Describe Listen Game as a multiplayer, online game with realtime feedback. (4) Produces high-quality semantic labels for music. (5) Using this data, we train a supervised multi-class labeling (SML) model and show that this model can annotate and retrieve music. The abstract should be placed at the top left column and should contain about 150-200 words.

1 INTRODUCTION

Collecting high-quality semantic annotations of music is often a difficult and time-consuming task. Examples of *semantic annotations* include chorus onset times [5], genre labels [12], or music similarity matrices [7]. In recent years, the Music Information Retrieval (MIR) community has focused on collecting standard data sets of such annotations for the purpose of system evaluation (e.g., MIREX competitions [2] , RWC Database [4]). However, these data sets are relatively small when compared to data sets that have been collected for other applied domains such as speech recognition [3], computer vision[1], and natural language processing [6, 8].

Traditionally, annotations have been collected by hand-labeling music [5, 12] , conducting surveys [16, 9, 11], or text-mining web-documents [10, 17]. The first two methods are not scalable since they can be time-cosuming and costly. The third methods generally produces low-quality data due to the collection documents that contain imprecise semantic information.

We propose the use of web-based games to collect semantic annotations of music. Recently, Von ahn et al. pioneered this technique by developing a suite of games (ESPGame [13], Peekaboom [15], Phetch [14]) which collect semantic information about images . These 'games with a purpose' provide an entertaining online interface for users who, though agreement with one another, produce high-quality semantic annotations as a bi-product of the game play. Von ahn refers to this data collection technique as "human computation" since the games harness the power of many human collaborators to solve problem that are beyond the reach of current Artificial Intelligence (AI) systems.

In this paper, we will describe *Listen Game*: a multi-player, web-based game that provides real-time feedback for players who actively labeling music with words. The purpose of this game is to collect word associations between audio content and words (i.e., semantic tokens). Our vocabulary of words consists of both preselected 'musically-relevant' words, such as words related to musical genre, instrumentation, or emotional content, and words that the users suggest while they are playing. The user's performance is tied to how much agreement there is across all players who are listening to the same song and have to label the song with the same small set of words. Over time, the game collect semantic word associations for a large set of songs. Note that these *word associations* are not binary labels, but rather real-valued weights that are proportional to the percentage of players who agree that a word does (or does not) describe a song. This is consistent with the notion that 'music is subjective': each listener has their own unique experience when listening to a song.

Our goal is to show that Listen Game is a scalable approach for collecting a large corpus of high-quality semantic word-song associations. In previous work [11], we presented a system that can automatically both *annotate* novel music with semantically meaningful words and *retrieve* relevant songs from a large database. Our system is based on a supervised multi-class labeling (SML) probabilistic model [1] and was training using the CAL500 data set [11]. This data set consists of 500 songs that have be recorded by 500 unique artists. Each song has annotated by a minimum of three people using a standard survey that asks the listener to a relate the song to 173 individual words. We will compared the data collected using Listen Game during a two-week pilot study to CAL500 data set. We will then show that the data collect Listen Game can be used to train our SML model.

2 COLLECTING MUSIC ANNOTATIONS

Discuss three techniques on collecting semantic word-song associations

2.1 Conducting Surveys

Pandora, Moodlogic, CAL 500

2.2 Mining the Web

Whitman, Berenzwig, Turnbull (Web2131) - detailed description of semantic weight matrix

2.3 Creating Games with Purpose

Human computation games of image - ESPGame to collect images, Peekaboom to collect segmentation, and fetch to understand how people query images. Comparison Minor Major - off-line game where are encourage to type words about music - like Open mind initiative.

3 LISTEN GAME

- Describe Listen Game - with detailed explanation of song clips, hierarchical vocabulary (semantic categories and words, best AND worst selection, scoring policy, freestyle round for growing vocabulary
- FIGURE: Screenshot of game
- Discuss why data is high quality - agreement amongst several independent users.
- Explain policy for converting annotations into semantic weight matrix.

4 SUPERVISED MULTICLASS LABELING MODEL

Discuss Supervised-Multiclass model - how it leads to automatic annotation and retrieval of music. Soft estimation, how to use semantic weight matrix

5 EVALUATION OF LISTEN500 DATA

Describe condition of pilot study - number of users, annotations, songs, exact method for converting annotations to semantic weight matrix

5.1 Direct comparison between Web 2131, CAL500 data and Listen500 data

TABLE: explicitly compared semantic weight matrices

5.2 SML Annotation and Retrieval Comparison

TABLE: standard IR metrics evaluated on SML model train using data from all three data sets.

6 FUTURE WORK

- Collect data more songs with growing vocabulary (LIST: good examples of novel words) - more users, more time
- Improved on existing game - spawn multiple (genre-specific) games, user-interface for visual impared-users,
- Create Novel Games
- Build a Search Engine direclty using semantic labels as indices or as training data for SML model - use search engine queries to feed game
- DO NOT DISCUSS (???) - personalization of search

7 REFERENCES

[1] G. Carneiro, A. B. Chan, P. J. Moreno, and N. Vasconcelos. Supervised learning of semantic classes for image annotation and retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(3):394–410, 2007.

[2] J. Stephen Downie. Mirex 2007 wiki. http://www.music-ir.org/mirexwiki/index.php.

[3] J.S. Garofolo, L.F. Lamel, W.M. Fisher, J.G. Fiscus, D.S. Pallett, , and N.L. Dahlgren. Darpa timit acoustic-phonetic continuous speech corpus. CD-ROM, 1993.

[4] M. Goto. Development of the RWC music database. In *International Congress on Acoustics*, pages 553–555, 2004.

[5] M. Goto. AIST annotation for RWC music database. In *ISMIR*, October 2006.

[6] D. Lewis. Reuters-21578 text categorization test collection, 1997.

[7] E. Pampalk, A. Flexer, and G. Widmer. Improvements of audio-based music similarity and genre classification. In *ISMIR 05*, pages 634–637, 2005.

[8] S. Roukos, D. Graff, , and D. Melamed. Hansard french/english. Linguistic Data Consortium, 1995.

[9] T. Sulzer. Moodlogic. http://www.moodlogic.com, 2007.

[10] D. Turnbull, L. Barrington, and G. Lanckriet. Modelling music and words using a multi-class naïve bayes approach. *ISMIR*, 2006.

[11] D. Turnbull, L. Barrington, D. Torres, and G. Lanckriet. Towards musical query-by-semantic description using the CAL500 data set. In *To appear in SIGIR '07*, 2007.

[12] G. Tzanetakis and P. R. Cook. Musical genre classification of audio signals. *IEEE Transaction on Speech and Audio Processing*, 10(5):293–302, 7 2002.

[13] L. von Ahn and L. Dabbish. Labeling images with a computer game. In *ACM CHI '04*, 2004.

[14] L. von Ahn, S. Ginosar, M. Kedia, R. Liu, and M. Blum. Improving accessibility of the web with a computer game. In *ACM CHI Notes*, 2006.

[15] L. von Ahn, R. Liu, and M. Blum. Peekaboom: A game for locating objects in images. In *ACM CHI '06*, 2006.

[16] T. Westergren. The music genome project. http://www.pandora.com, 2007.

[17] B. Whitman and D. Ellis. Automatic record reviews. *ISMIR*, 2004.