A web-based game for collecting music metadata

MICHAEL I MANDEL AND DANIEL P W ELLIS
{mim,dpwe}@ee.columbia.edu · LabROSA · Columbia University, New York, NY

1 Overview: majorminer.com

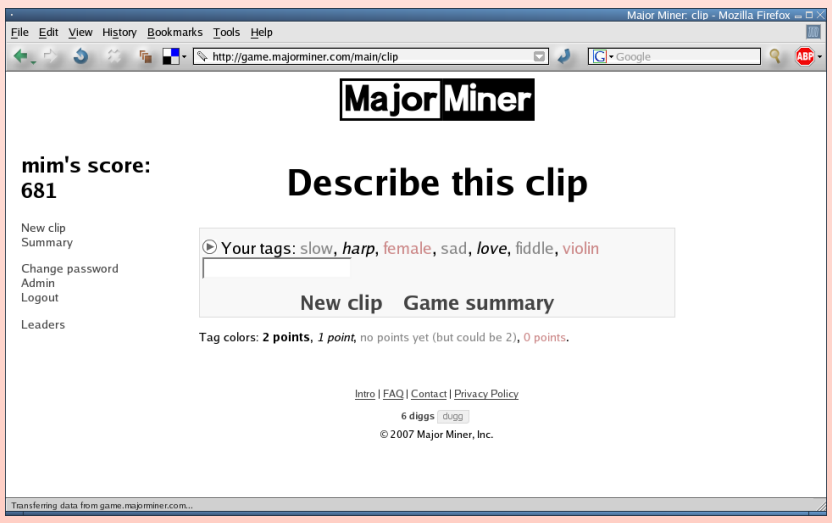
- A web-based game that makes collecting descriptions of musical excerpts fun, easy, useful, and objective
- Players describe 10 second **clips** of songs and score points when their descriptions match those of other players
- The rules were designed to encourage players to be thorough
- The clip length was chosen to make judgments more objective and specific

2 Motivation: words help people find music

- Automatic music descriptions would help people find and browse new music
- Descriptions of clips are most likely to be about the sound itself, as opposed to the social context, the band's reputation, etc.
- Sound-based automatic analysis tools are best suited to describing clips
- We are collecting descriptions of clips for training such tools

3 Related work: games and music-text connection

- Whitman and Ellis (2004) train a system to associate noun phrases and adjectives from reviews with audio. Based on (Whitman and Rifkin, 2002).
- Turnbull et al. (2006) use naïve Bayes classifier to annotate and retrieve music
- Ellis et al. (2002) describe "Musicseer" game for collecting artist similarity data
- The "ESP Game" (von Ahn and Dabbish, 2004) popularized the idea of allowing any response as long as it could be verified
- Check out the other music-text games at ISMIR (Turnbull et al., 2007; Law et al., 2007)



Screen shot of a game in progress, tagging a clip

4 Rules of the game

- A player listens to one randomly selected 10-second song excerpt (**clip**)
- She describes the clip with a word or phrase (**tag**)
- If her tag is the first to agree with another player's tag on that clip, she scores *1 point* and the other player scores **2 points**
- If two other players have already agreed on that tag, she scores 0 points
- She can tag each clip as many times as she desires
- After finishing that clip, she can see one other player's tags for it

5 Design considerations

- The scoring rules were designed to encourage players to describe clips thoroughly; to be original, yet relevant
- The clip selection procedure was designed to make the game fun for both new and veteran players
- All aspects of the game were designed to avoid cheating, collusion, or other manipulations of scores or data

Selecting clips: based on the player's experience level

- New players are given clips already seen by others so they can score immediately
- Experienced players are given fresh clips so they can be first to use many tags
- The probability of giving a player a fresh clip is the ratio of the number of clips seen by that player to the number seen by any player

Revealing labels: to habituate new players

- Once a player is done with a clip, she can see how one other player has described it
- Those tags belong to the first player to see the clip, which remains the same no matter how many subsequent players see it
- The tags are likely to be those of an experienced player, so habituate new players

References

D. Ellis, B. Whitman, A. Berenzweig, and S. Lawrence. The quest for ground truth in musical artist similarity. In *Proc ISMIR*, pages 170–177, 2002.

E. L. M. Law, L. von Ahn, R. B. Dannenberg, and M. Crawford. Tagatune: A game for music and sound annotation. In *Proc ISMIR*, 2007.

D. Turnbull, L. Barrington, and G. Lanckriet. Modeling music and words using a multi-class naïve bayes approach. In *Proc ISMIR*, pages 254–259, 2006.

D. Turnbull, R. Liu, L. Barrington, and G. Lanckriet. A game-based approach for collecting semantic annotations of music. In *Proc ISMIR*, 2007.

L. von Ahn and L. Dabbish. Labeling images with a computer game. In *Proc SIGCHI conference on Human factors in computing systems*, pages 319 – 326, 2004.

B. Whitman and D. Ellis. Automatic record reviews. In *Proc ISMIR*, pages 86–93, 2004.

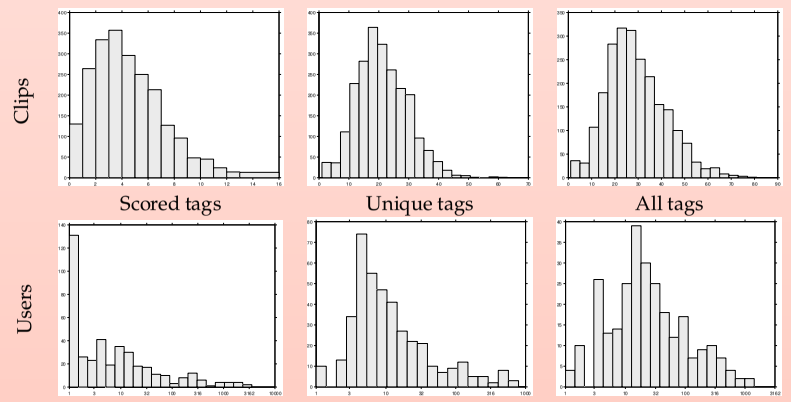
B. Whitman and R. Rifkin. Musical query-by-description as a multiclass learning problem. In *IEEE Workshop on Multimedia Signal Processing*, pages 153–156, 2002.

6 Data collected: popular tags

Tags 1–30	Verified	Tags 31–60	Verified	Interesting Tags	Verified
drums	879	ambient	56	violin	20
guitar	803	saxophone	54	sample	18
male	703	solo	51	love	18
rock	640	distortion	50	live	17
electronic	468	punk	48	harmony	17
pop	466	drum	48	u2	15
synth	457	quiet	47	chorus	15
bass	399	sax	46	sad	14
female	336	silence	44	horn	14
dance	310	strings	42	heavy	14
techno	234	keyboard	39	dark	14
piano	171	trumpet	35	talk	13
electronica	163	organ	35	experimental	13
synthesizer	152	end	34	intro	12
rap	149	loud	33	hard	12
slow	145	acoustic	33	echo	12
vocal	141	trance	31	duet	12
voice	136	noise	31	cure	12
hip hop	128	metal	30	woman	11
vocals	126	folk	30	orchestra	11
jazz	126	club	30	noisy	11
beat	107	hiphop	29	depeche mode	11
fast	102	indie	28	scratching	10
80s	101	repetitive	27	new age	10
instrumental	90	horns	27	brass	10
drum machine	85	industrial	26	talking	9
british	77	ballad	26	spoken	9
country	71	jazzy	24	smiths	9
house	61	80's	24	rave	9
soft	60	reverb	22	radiohead	9

7 Future work

- Train models to automatically describe music
- Analyze similarity between clips, players, words
- Combine audio- and word-based similarities to improve both
- Build anchor space features from classification on a subset of these tags
- Use automatic descriptions to help determine structure of songs



Histograms of user and clip statistics