A Domain-Specific Music Search Engine for Gait Training

Zhonghua Li and Ye Wang
School of Computing, National University of Singapore, Singapore
{lizhongh, wangye}@comp.nus.edu.sg

ABSTRACT

This paper demonstrates a domain-specific music retrieval system to help music therapists find appropriate music for Parkinson's disease patients in their gait training. Different from existing music search engines, this system incorporates multiple music dimensions (i.e., tempo, cultural style, and beat strength) required in gait training, and facilitates the searching process by allowing music retrieval directly on these dimensions. To support music search by tempo, a user-perception based method is also proposed to improve state-of-the-art tempo estimation algorithms. We conducted a user study and evaluated the system efficacy in searching music using different types of queries based on these music dimensions. Experimental results demonstrate the effectiveness and usability of our system in therapeutic gait training.

Categories and Subject Descriptors

H.3.3 [**Information Search and Retrieval**]: Information filtering, Query formulation; H.5.5 [**Sound and Music Computing**]: Signal analysis, synthesis, and processing

General Terms

Algorithms, Design, Experimentation, Human Factors

Keywords

domain-specific, music retrieval, gait training

1. INTRODUCTION

Millions of people across the globe are suffering from Parkinson's disease, and are progressively losing their motor control and physical movement. Music therapy research has indicated that using tempo-matched and patient-preferred auditory stimuli can help Parkinson's disease patients improve their walking in gait training. However, as patients' conditions vary significantly, searching for such music for each patient is very time-consuming using existing music search engines. We thus develop a music retrieval system for this particular domain by allowing music search directly using the desired music dimensions. In this demonstration, tempo, culture style and beat strength three music dimensions are integrated. Users can easily find music with a certain tempo, or with different culture styles or beat strength levels. Based on our previous work [5], this demo adopts a user-perception based tempo estimation algorithm, and also provides a tempo evaluation tool to help users check the tempi of the retrieved songs. To the best knowledge, this system is the first one to help music therapists to find suitable music for gait training.

MM'12, October 29–November 2, 2012, Nara, Japan.
ACM 978-1-4503-1089-5/12/10.

2. SYSTEM ARCHITECTURE

The architecture of our system is presented in Figure 1. This system consists of four main components. The *query*

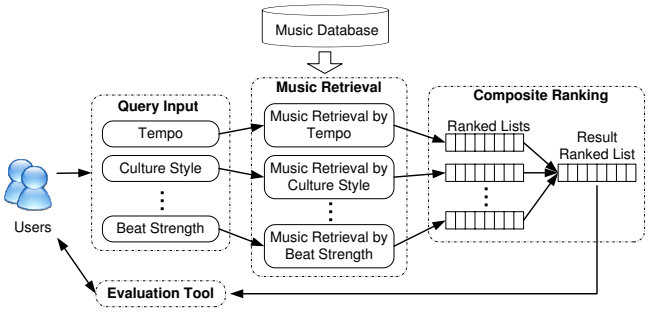

Figure 1: System architecture

input component supports different query methods on different music dimensions. For example, for tempo, users can directly input the desired tempo (in beat per minute). Music culture styles can be customized from a list of candidate culture styles. Beat strength can also be customized to different beat strength levels. Considering the heterogeneity of different dimensions, the *music retrieval* component analyzes and retrieves relevant music on each music dimension independently. The *composite ranking* component then combines the ranked lists from multiple dimensions into a result list. Finally, given the result song list, the *evaluation tool* component helps users to examine the relevance of these songs.

3. SYSTEM IMPLEMENTATION

3.1 Music Retrieval by Tempo

To achieve better tempo estimations, we propose a user-perception based method to correct the half/double octave errors made by most of the existing tempo estimation algorithms. Based on the collected music speed annotations, songs are firstly classified into fast and slow categories [1]. The tempo of a song estimated using existing algorithms is

then doubled or halved if this song is classified as fast/slow but with the estimated tempo lower/higher than a certain threshold. We evaluated four state-of-the-art tempo estimation algorithms on our database, and validated the effectiveness of the proposed octave error correction method. After tuning with our method, the algorithm of Klapuri et al. [2], which achieved the best evaluation performance is adopted. Given a query tempo, the distances between tempi of songs and this query tempo is measured using a Gaussian score [5].

3.2 Music Retrieval by Culture Style

An automatic classification approach is adopted to classify songs into different culture styles [6]. Both basic features (timbre, rhythm, and wavelet-based features) and musicology-based features are extracted, and support vector machine with radial basis kernel is used for classification. Given a query culture style, songs are ranked by the probability that they are classified as this culture style.

3.3 Music Retrieval by Beat Strength

Beat strength describes the prominence of beats within a music clip. We follow the implementation of pulse clarity in the MIRtoolbox [4], and adopt the best parameters evaluated by Lartillot et al. [3]. The estimated beat strength (real values between 0 and 1) is then linearly divided into weak, middle and strong three levels. Given a query beat strength level, songs are then ranked linearly based on their distances to the weakest, middle, and strongest point in these three levels.

3.4 Composite Ranking

The ranked lists retrieved from multiple dimensions are linearly fused into a result list with an equal weight. As stable songs are required in gait training, we further rank songs by their tempo stability [5].

3.5 Evaluation Tool

A web-based tempo evaluation tool is implemented in our system. While listening to a song, users can tap along with the song by clicking the mouse. All the tappings are recorded and the median time interval of these tappings is used to calculate tempo.

4. DEMONSTRATION AND EVALUATION

A snapshot of our system is illustrated as Figure 2. Users can directly search music by inputting the desired tempi, or customizing different culture styles or beat strength levels. A result list which aims to meet all the desired query dimensions is returned. Music can be played when clicking on its title. When using the tempo evaluation tool, both the number of taps and the tempo are presented.

To evaluate the system, we collected a music database of 787 songs from YouTube. These songs are of four culture styles, including Chinese, Indian, Malay and Western. We hired 11 subjects (5 amateur musicians and 6 non-musicians) to annotate each song with tempo, beat strength level and the perceived speed (fast/slow). We randomly sampled $\frac{2}{3}$ of the data for training culture style classifies and tunning thresholds in tempo estimation and beat strength analysis. The remaining songs severed for testing. We conducted a user study among 20 students and two music therapists to evaluate both the effectiveness and the usability of our system. The precision at rank 10 and mean average precision

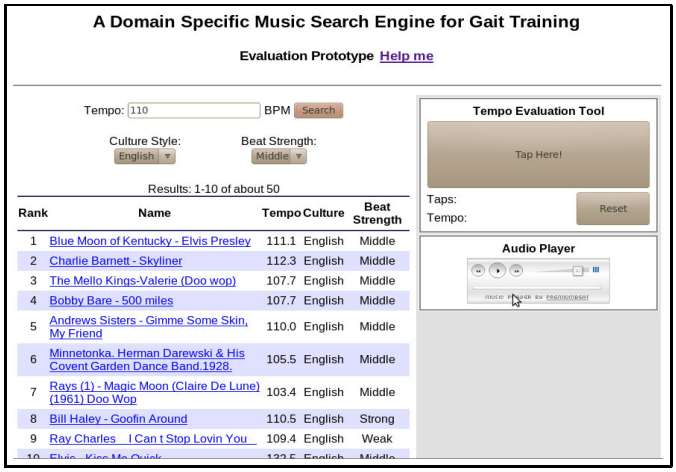

Figure 2: Snapshot of the system interface

Table 1: Precision at rank 10 (P@10) and mean average precision (MAP) for all query types

Query type	P@10	MAP	Query type	$P(10)$	MAP
Tempo (TP)	0.768	0.893	TP+BS	0.672	0.795
Culture style (CS)	1.000	1.000	CS+BS	0.837	0.892
Beat strength (BS)	0.811	0.842	TP+CS+BS	0.437	0.738
TP+CS	0.721	0.830	Overall	0.749	0.856

for all tested query types are calculated and presented in Table 1. To evaluate the usability of our system, subjects' responses were also collected and mapped to a score of $1 \sim 5$ representing "very unsatisfied" to "very satisfied". The mean scores of tempo, culture style, and beat strength are 4.09, 4.05, and 3.50, respectively. Both music therapists addressed the usefulness of our system in gait training.

5. CONCLUSIONS

In this demo, we present a domain-specific music search engine to help music therapists and Parkinson's disease patients find appropriate music for gait training. By supporting music search by the desired music dimensions and providing a tempo evaluation tool, our system saves great efforts of searching and manually checking the music. Currently, our system only integrates a few music dimensions. However, it exhibits the potential in exploring other music dimensions within this framework, and also provides some guidelines in designing music retrieval systems for other therapeutic applications.

6. REFERENCES

[1] J. Hockman and I. Fujinaga. Fast vs slow: Learning tempo octaves from user data. In *ISMIR*, pages 231–236, 2010.

[2] A. P. Klapuri, A. J. Eronen, and J. T. Astola. Analysis of the meter of acoustic musical signals. *IEEE Trans. on Speech and Audio Processing*, 14(1):342–355, 2006.

[3] O. Lartillot, T. Eerola, P. Toiviainen, and J. Fornari. Multi-feature modeling of pulse clarity: Design, validation, and optimization. In *ISMIR*, pages 521–526, 2008.

[4] O. Lartillot and P. Toiviainen. Mir in matlab (ii): A toolbox for musical feature extraction from audio. In *ISMIR*, pages 127–130, 2007.

[5] Z. Li, Q. Xiang, J. Hockman, J. Yang, Y. Yi, I. Fujinaga, and Y. Wang. A music search engine for therapeutic gait training. In *ACM Multimedia*, pages 627–630, 2010.

[6] Y. Liu, Q. Xiang, Y. Wang, and L. Cai. Cultural style based music classification of audio signals. In *ICASSP*, pages 57–60, 2009.