Real-time Accompaniment Using Lyrics-Matching QBH

Panagiotis Papiotis and Hendrik Purwins

Music Technology Group, Universitat Pompeu Fabra
Roc Boronat, 138
08018 Barcelona, Spain
panos.papiotis@gmail.com
hendrik.purwins@upf.edu

Abstract. Query-by-Humming (QBH) is an increasingly popular technology that allows users to browse through a song database by singing/humming a part of the song they wish to retrieve. Besides these cases, QBH can be also used in applications such as Score Alignment and Real-Time Accompaniment, to locate the exact position of the soloist within the piece. We propose a system that initially pinpoints a singer's location within a song using a multi-similarity QBH approach, and then aligns an existing audio recording of the accompaniment by modifying its tempo and dynamics in real time. Preliminary results for our QBH algorithm show that our approach can achieve 58.5% Top-1 accuracy in locating the phrase that contains the exact lyrics – an improvement of 170% over the basic pitch contour method.

Keywords: Query by Humming, Real-time Accompaniment, Interactive Music Systems, Singing Voice

Introduction

Real-time Accompaniment systems such as Score Following/Alignment [1] attempt to align an existing score to the performance of a soloist in an online manner, very much like a human accompanist would align and adapt his/her performance to match that of the soloist. In this context, a capable QBH system would be invaluable; it would automatically locate the soloist's position within the musical piece, providing the starting point for the alignment process. The soloist could stop and start again from a different point in the piece, without having to manually adjust the starting position. In the special case of the singing voice, a useful feature that could be utilized to discriminate between identical melodies are the piece's lyrics; with the possible exception of the chorus, every melodic line is coupled with a different lyric. Therefore, if one can match the lyrics of the query to the lyrics of the reference, the exact location of the soloist within the piece would be pinpointed with absolute accuracy.

Melody and Lyrics Matching by Dynamic Time Warping

We implemented a QBH system for singing voice, which calculates the similarity between a sung query and every phrase within a vocal recording of the reference song. Using Dynamic Time Warping (DTW), we apply a three-phase similarity measurement for Pitch contour, Mel Frequency Cepstral Coefficients (MFCCs) and RMS Energy envelope respectively. Our algorithm combines these three methods to

retrieve the starting position of the matching phrase within the reference recording, in a manner similar to [2].

Preliminary results on a dataset of seven songs for a total of 114 queries can be seen in Table 1.

Table 1. Average and random accuracy for our algorithm.

Song title	Accuracy	Random accuracy
She's leaving home	0.61	0.076
Butterflies and hurricanes	0.66	0.055
Nude	0.61	0.036
Bohemian Rhapsody	0.57	0.049
A day in the life	0.45	0.041
All the small things	0.6	0.134
Message in a bottle	0.6	0.066
Overall	**0.585**	**0.0471**

The average query length is 6.1 seconds, while the average response time is 0.8 times the duration of the query.

The real-time alignment of the accompaniment audio is currently under development; it consists of utilizing the warp matrix produced by the DTW algorithm to adjust the time-stretch ratio, and comparing the RMS envelope of the reference and the soloist's voice in real time to modify the dynamics. An audiovisual demonstration of our system is available on the Internet for evaluation purposes, on the address provided in [3]. The method is explained in more detail in [4].

References

1. A. Cont: "ANTESCOFO: Anticipatory Synchronization and Control of Interactive Parameters in Computer Music", Proceedings of International Computer Music Conference (ICMC), August 2008, Belfast, Ireland.
2. L.Wang, S.Huang, S.Hu, J.Liang and B.Xu: "An Effective and Efficient Method for Query by Humming System Based on Multi-Similarity Measurement Fusion", IEEE International Conference on Audio, Language and Image Processing, Shanghai, 2008.
3. QBH Lyrics Alignment - http://www.youtube.com/user/qbhlyricsalignment
4. P. Papiotis, H. Purwins (2010). A Lyrics-Matching QBH System with Applications in Real-Time Accompaniment. Sound and Music Computing Conference.