RM1 REAL TIME ONLINE ALIGNMENT ALGORITHM FOR MIREX 2017

Pedro Alonso
Interdisciplinary and Communication Department,
Polytechnic Univ. of Valencia
palonso@upv.es

Jose Ranilla, Raquel Cortina
Information Retrieval and Parallel Computing Department,
University of Oviedo
{raquel, ranilla}@uniovi.es

P. Vera-Candeas, F.J. Bris-Peñalver
Telecommunication Engineering Department,
University of Jaen
{pvera, fbris}@ujaen.es

ABSTRACT

This abstract describes a proposed score follower submitted to the MIREX 2017 Real-time Audio to Score Alignment (a.k.a. Score Following) evaluation task.

1. INTRODUCTION

Automatic accompaniment systems are defined as combinations of software tools and electronic devices that produce the generation of music in consonance with the performance.

This work is focused on systems based on following the score sheet of the known piece that is performed, where the intelligence of the system is given by synchronizing an audio recording of a musical piece with the corresponding music sheet. This is called audio to score alignment and it can be carried out offline or online.

Thus, we present a real-time score follower based on spectral factorization and online Dynamic Time Warping (DTW). The presented system has two separated stages, preprocessing and alignment. On the first one, we convert the score into a reference audio signal using a MIDI synthesizer software and we analyze the provided information in order to obtain the spectral patterns (i.e. basis functions) for each combination of the concurrent notes given at the score. These basis functions are learned from the synthetic MIDI signal using a method based on Alternated Non-Linear Least Squares (ANLS), where the gains are initialized with transcription obtained from the MIDI file. On the second stage, a real-time signal decomposition method with fixed basis functions per combination of notes is applied over the magnitude spectrogram of the input signal resulting in a distortion matrix that can be interpreted as the matching likelihood for each combination of notes at each frame. Finally, the score alignment is obtained using an on-line Dynamic Time Warping (DTW) over the distortion matrix in order to find the path with the minimum cost and then determine the states real duration.

2. SYSTEM DESCRIPTION

This proposal has the goal of produce a real time score follower able to fit in automatic music accompaniment systems.

2.1 Preprocessing Stage

2.1.1 States Definition

The aim of this stage is to compute the states and states sequence from the MIDI data. A state is defined as a combination of notes that occurs simultaneously in the ground truth transcription obtained from the MIDI file.

2.1.2 Basis Functions Learning

Once the states and the states sequence have been defined, the basis functions associated to each state are learned. To this end, we use a supervised method based on Alternating Non-Negative Least Squares (ANLS) method.

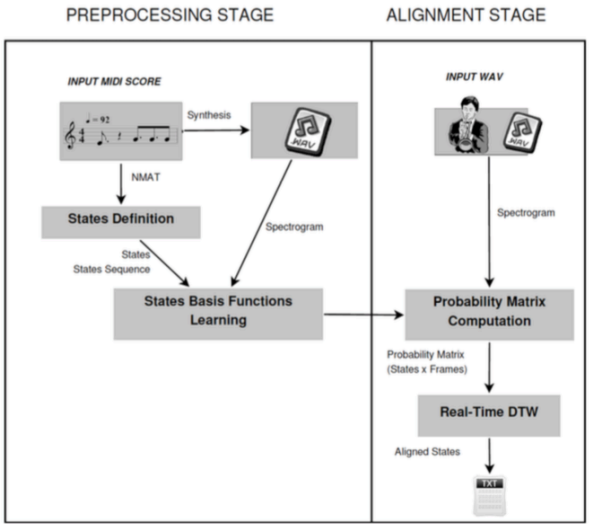

Figure 1. Real-Time Score Follower Block Diagram which this algorithm is based on.

2.2 Alignment stage

2.2.1 Probability Matrix Computation

The basis functions for each state are trained in advance using the MIDI data and kept fixed. Each basis function models the spectrum of a unique state. Now, the aim is to compute the gain matrix and the final cost matrix that

measures the likelihood between the estimated and the real spectrogram.

2.2.2 Real State Sequence Estimation by DTW

We used the following constrained DTW path, as explained in [1,2], is the accumulated cost value at the t-th frame and the m-th state at the sequence. Finally, DTW is used to match the score position with each input signal frame, performing an online alignment in order to reduce latency. The decision is made directly from the information contained into frame t.

Standard DTW assumes off-line search and the estimated path is obtained by backtracking of whole the signal. To extend DTW for the on-line search without backtracking, we simply select the reference state which has the smallest accumulated distance with the current performance frame

2. EVALUATION

Results are published in the MIREX 2017 results website.

3. REFERENCES

[1] J.J. Carabias, F.J. Rodríguez, P, Vera, "An audio to score alignment framework using spectral factorization and dynamic time warping", Proc. of the Music Information Retrieval Conference (ISMIR), Málaga, Spain, 2015.

[2] Pedro Alonso, P. Vera-Candeas, Raquel Cortina and José Ranilla, "An efficient musical accompaniment parallel system for mobile devices". *J. Supercomput,* 73, pp. 343–353, 2017.