

ISE 575/EE 675 – hw 3

Paper Report – Simon Dixon and Gerhard Widmer, “MATCH: A Music Alignment Tool Chest”

This paper outlines the implementation of MATCH, a tool whereby different performances of the same musical work can be compared, either offline or in real-time. The scope of MATCH extends primarily to the ability to index with greater accuracy works of music. Dixon gives the example of a listener who wants to hear a particular musical phrase played by different performers, without having to manually search through individual tracks in a compact disc. Thus it is of great use to students of musicology, as well as to music editors, for whom such audio splicing would be vital.

MATCH operates a dynamic time-warping algorithm which works faster than real-time (Dixon, 1). For every point in musical work A, a point in musical work B with the least ‘cost’ (least difference) is found. By continuing this process recursively, a ‘minimum cost path’ can be found in the space connecting the works. To decrease the complexity (and increase the efficiency) of this system, the points searched are restricted within a thicker ‘diagonal’ in this space (Dixon, 2-3). To do this, an ‘adaptive diagonal’ for the works is generated, which would change slope depending on the tempo differential at a given point (Dixon, 3).

The cost function between audio frames is obtained sorting the frequencies at a given frame into FFT (Fast Fourier Transform) bins. These bins are organized linearly at the low frequency range (up to 370Hz) and logarithmically beyond that. After this division, differences in respective bins between audio frames are used to establish a ‘cost’.

Using the Beatroot software, which automatically annotates beats, in conjunction with several multi-performed works, MATCH was shown to be successful in all but 2 out of 221 cases. Given possible errors or discretionary musical choices in performance, emphasis was placed on ‘score events’, rather than checking every individual note in musical samples. While the system was extended to orchestral music with some success, I imagine there might be some difficulty using MATCH with solo violin or cello music, where the onset time and pattern is under the complete discretion of the performer, and is sometimes very difficult to detect. Some vocal music might encounter the same problems.

There are some challenges with MATCH outlined by Dixon. These include the problem of normalization, with regard to volume, adjusted for salient parts of audio to match (pg. 5), and the difficulty in identifying the ‘first’ or ‘last’ notes – and hence audio frames – of works. A larger issue might be wider-spanning structural difference in performances (e.g. repeated sections), which could be solved by observing similarities at a higher-level understanding of the pieces (e.g. by understanding harmonic structure at some basic level).