Score Following from Inaccurate Score and Audio Data using OMR and `music21`

Jordi Bartolomé Guillen
Universitat Politècnica de Catalunya
Barcelona 08034, Spain
jordi.bartolome.guillen@gmail.com

Michael Scott Cuthbert
M.I.T., Music and Theater Arts
Cambridge, Mass. 02139 USA
cuthbert@mit.edu

Score following has a long history as a method for precisely aligning rich audio data with a correctly notated "ground truth" musical score. The primary research problems to be solved in such cases are correct pitch and rhythm recognition from audio data (a field with a similarly rich history) and successful alignment of the recognized pitches with the correct data from the score. Such systems have proven successful and are now available commercially for microcomputer platforms and tablets such as the iPad. As successful as these systems have been, they also present a huge disadvantage. Only a very limited repertory of pieces have been encoded in computer readable formats that also present enough visual information to the performer such as dynamics, articulations, tempo markings, and correct enharmonics to be useful (e.g., no professional musician would ever perform from a MIDI transcription of a score). Commercial packages such as the Prodigy Digital Music Stand and the Tonara App for the iPad measure their total number of scores in the hundreds (Tonara, for instance, does not have a single piece for many common instruments such as the clarinet). Furthermore, the business models for most of these applications depend on buying the limited set of sheet music directly from the systems' creators; thus copyright considerations and marketability will always remain concerns (if you hold your breath waiting for your favorite Albanian pop leadsheet or flugelhorn concerto to appear on their music lists, you'll asphyxiate).

The ScoreFollower page turning system bundled with the free, open-source `music21` toolkit demonstrates a more plausible solution for automatic page turning on a much larger corpus of musical scores. By using images of professionally engraved scores—ubiquitous on internet sites such as IMSLP (www.imslp.org) or easily scanned by individual users—together with optical music recognition (OMR) software and a two-page format, ScoreFollower employs a loose tracking algorithm to determine with extremely high reliability approximately where the performer is within the score. Although OMR has been mentioned in passing as a potential tool for expanding repertories since the earliest page turning applications [3], to our knowledge, no one has built a system around its use before, even when OMR and page turning are discussed in close succession [2]. Even with OMR systems that extract only 90-95% of the pitches accurately and audio pitch detection algorithms (operating on built-in laptop microphones) that perform at a 75% level, ScoreFollower is able to tell the performer's rough position within a the score with enough accuracy so that in our tests it never missed a page turn. When the musician is determined to be securely on the last of the two pages, the first page is smoothly changed to the next page to be performed. As OMR and pitch detection accuracy improves, or simple pattern matching algorithms are replaced with more sophisticated ones that account for common errors in OMR and audio extraction, even extremely low quality scans and noisy environments will be opened up for page turning applications. Similarly, single-page views should become feasible by moving pages upwards as in a medieval scroll when the system is convinced that the performer no longer needs the notes at the top of the page. Underlying ScoreFollower is the `music21` python open-source toolkit, a flexible system that easily allows for adding additional musical intelligence (such as automatically detecting expected transpositions from instrument names).[1] The Peachnote[4] database of OMR versions of nearly all of IMSLP (by Vladimir Viro, who has consulted with the `music21` project) will soon make a vast number of scores available to performers and allow

curious musicians for the first time to perform their parts from full scores requiring too many page turns to be practical in paper formats. By showing the viability of messy data in score following, the authors hope to spur further research in OMR, automatic page turning, and using computer-displayed scores in live acoustic performance.

Further Details

The software preprocesses the score by dividing it into many small fragments whose size depends on the number of notes expected to be detected at a time. While playing, the performer's audio into is also divided into fragments. Each audio fragment is analyzed and the pitches and the rhythm are extracted. This audio-generated score is compared with the pitches and the rhythm extracted from the OMR printed score and, depending on the similarity, a probability of each segment being a match is computed. Usually, the fragment of the score with a high matching probability is chosen as the played fragment by the musician, thus letting the performer stop and start at any point on the page (e.g., go back and rehearse a passage several times); however in the case when several fragments are highly probable, the system should not merely choose the most closely matching fragment.

ScoreFollower also takes into account the position of the last matched fragment and taking tempo and other factors into consideration, it gives extra weight to fragments within the target zone. The alternation of recording, pitch extraction, segmentation, and comparison is performed rapidly enough that most of our tests were performed on a tiny netbook with a 1Ghz [[JORDI: Check please?]] Atom processor. To test the system, scores were performed on clarinet, piano (although we have not yet implemented any of the polyphonic pitch detection algorithms), and from a keyboard app on an iPhone played halfway across the room. Only when we tried singing the score with untrained voices were the errors in pitch detection so great as to occasionally lose track of where the performer was (giving higher weight to rhythmic matching could be a solution to this problem). Finally, the algorithm stops when the musician reaches the lasts notes of the score or it detects that he or she has conclusively stopped playing. Demonstrations of the system in English, Spanish, and Catalan are available at http://www.youtube.com/user/mscuthbert (A screenshot appears as Figure 1). The `music21` toolkit can be downloaded at web.mit.edu/music21 and started by running the graphicalInterfaceSF.py file in the audioSearch directory.

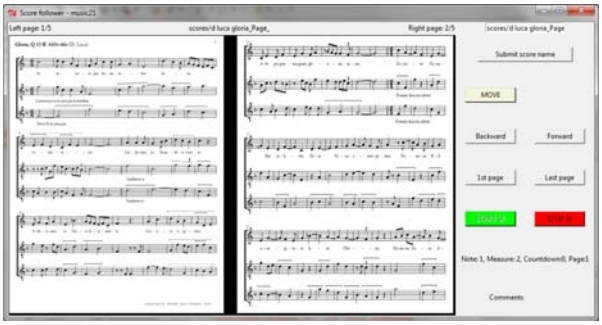

Figure 1: ScoreFollower user interface

Acknowledgements and References

`Music21` has been developed with a grant from the Seaver Institute.

[1] Cuthbert, Michael Scott and Christopher Ariza, "`music21`: A Toolkit for Computer-Aided Musicology and Symbolic Music Data," *Proc. Int. Symp. Music Inf. Retr.*, 2010: 637–42.

[2] Lin, Karen and Tim Bell. "Integrating Paper and Digital Music Information Systems," *Proc. Int. Symp. Music Inf. Retr.*, 2000.

[3] McPherson, John. "Page Turning: Score Automation for Musicians." Honours report, University of Canterbury, 1999. Citeseer: 10.1.23.3469, at p. 15.

[4] Viro, Vladimir. "Peachnote: Music Score Search and Analysis Platform," *Proc. Int. Symp. Music Inf. Retr.*, 2011. http://www.peachnote.com