McGill

Schulich School of Music
of McGill University
555 Sherbrooke Street West
Montreal, Quebec H3A 1E3

Schulich School of Music
École de musique Schulich

École de Musique Schulich
de l'Université McGill
555, rue Sherbrooke Ouest
Montréal (Québec) H3A 1E3

(514) 398-4535
(514) 398-8061
www.mcgill.ca/music

4 June 2019

To Whom It May Concern:

This is my review of the doctoral thesis by Jan Hajič jr. entitled "Optical Recognition of Handwritten Music Notations" at Charles University. The thesis provides excellent overview of the history and the current state of research in optical music recognition (OMR) followed by several new results in the research domain.

Perhaps the most important new scientific result is the design of the directed acyclic graph (DAG) called MuNG for common Western music notation (CWMN) and in particular for the handwritten music. Much of the thesis deals with this handwritten CWMN, which has not been researched as well as printed CWMN in the area of OMR.

The MUSICMA++ dataset would be valuable for training and evaluating other OMR models. This model applies to both printed and handwritten music notation. The emphasis on handwritten music is important because not only OMR research has been concentrated on printed music but, as with text, the recognition of handwritten music is much more difficult because of the wide variety of handwritten styles.

For application in other areas, the method, described in Section 6.4, of finding the objective metric to calculate the "cost to correct" using human subjects would be useful for creation of digital diplomatic facsimile of manuscripts used in other disciplines of historical documents. This evaluation method is important for comparing different OMR and other document recognition strategies and also to estimate the cost of OMR and other digitization projects.

Corrections:

p. 10: The definition of the "step" should be improved, for example, the semitone is undefined.
The definition of the "octave" should be improved, for example, there should be a mention of 12 semitones.
In footnote 2, there should be a reference to support the statements.

p. 11: "e.g." → "e.g.,"

p. 20: The definitions of "replayability" and "reprintability" are not well defined. Later in p. 87 (of the PDF file, or 1:13), another attempt at the definitions are offered but I find them insufficient and the purpose of defining them are also unclear. In general, music scholars do not need to have the scores converted to audio for the purpose of their research. In some cases, it may be useful for the general public to hear automated MIDI transcriptions of the score but they are usually of poor musical quality.

p. 24: "esp." should be spelled out to "especially"

p. 26: The dissertation by Nicholas Carter should be listed.

p. 27: "John Ashley et al." should be "Bourgoyne et al."

p. 28: There is a poor page break.

p. 31: Content missing in footnote 10.

p. 32: The following sentence needs further explanation since the evaluation is not clearly defined in this context: "Furthermore, evaluating against more complex formats is notoriously problematic." Also, the succeeding paragraph also requires more background explanation.

p. 33: "recurent" → "recurrent"

p. 34: "e.g." → "e.g.,"

p. 38: Links at the bottom of the page are incorrectly linked in the PDF version.

p. 39: "voice..." → "voice, etc."
Writing style too informal: "We of course propose that the answer is – no."
Fig. 4.1 is referenced here but is placed on p. 41.

p. 44: Does "netto" mean "total"?

p. 45: "Seeing as" → "Seeing that"

p. 54: This sentence seems out of place: "A separate chapter are precedence edges."

p. 55: "it" → "is"
Too informal: "worth its salt"

p. 60: "severy" → "severely"
"esp." → "especially"
I don't understand why the "dependency graph-based methodology" is especially good "for manuscript recognition." The graph-based method is just as applicable for printed music.

p. 61: Unless "an appropriate graph alignment algorithm is developed," the usefulness of MuNG is limited except for the alignment of its MIDI output and the value of using the graph cannot be evaluated. Note that the MIDI output can be generate without creating the graph.

1:10: Fig. 7 should have clefs.

1:27–33: In the Reference section, there are several entries (about 30) with inconsistent capitalization of paper titles. There are two incomplete entries (a chapter in a book), and a few misspellings.

p. 171–185: In the Bibliography, the capitalization of paper titles is inconsistent.

p. 189: "a Practical" → "A Practical"
"jr.:" → "jr."

p.190: "How current optical music recognition systems are becoming useful for digital libraries" → "How Current Optical Music Recognition Systems are Becoming Useful for Digital Libraries"

"Creating annotated resources for polarity classification in Czech" → "Creating Annotated Resources for Polarity Cslassification in Czech"

Inconsistent usage of Oxford commas (missing a comma): p. 11 (octave), (1/32); p. 16 (1/8); p. 39 (signature), (flags); p. 43 (staffspaces), (OS X); p. 52 (staffspaces).

Finally, I believe that the Jan Hajič jr. has the ability for creative scientific work and with the proper revisions suggested above, this thesis is ready to proceed to the defense.

Sincerely,

Ichiro Fujinaga, Ph.D.
Associate Professor
Schulich School of Music
McGill University