How Current Optical Music Recognition Systems Are Becoming Useful for Digital Libraries

Jan Hajič jr.
Charles University, Czech Republic
hajicj@ufal.mff.cuni.cz

Marta Kolárová
Charles University, Czech Republic
marta.kolarova@ff.cuni.cz

Alexander Pacha
TU Wien, Austria
alexander.pacha@tuwien.ac.at

Jorge Calvo-Zaragoza
Universitat Politècnica de València, Spain
jcalvo@prhlt.upv.es

ABSTRACT

Optical Music Recognition (OMR) promises to make large collections of sheet music searchable by their musical content. It would open up novel ways of accessing the vast amount of written music that has never been recorded before. For a long time, OMR was not living up to that promise, as its performance was simply not good enough, especially on handwritten music or under non-ideal image conditions. However, OMR has recently seen a number of improvements, mainly due to the advances in machine learning. In this work, we take an OMR system based on the traditional pipeline and an end-to-end system, which represent the current state of the art, and illustrate in proof-of-concept experiments their applicability in retrieval settings. We also provide an example of a musicological study that can be replicated with OMR outputs at much lower costs. Taken together, this indicates that in some settings, current OMR can be used as a general tool for enriching digital libraries.

CCS CONCEPTS

• **Information systems → Music retrieval**; *Image search*; • **Applied computing → Digital libraries and archives**; *Document searching*; **Graphics recognition and interpretation**;

KEYWORDS

Optical Music Recognition, Music Information Retrieval, Symbolic Music Search, Music Digital Libraries, Digital Musicology

ACM Reference Format:

Jan Hajič jr., Marta Kolárová, Alexander Pacha, and Jorge Calvo-Zaragoza. 2018. How Current Optical Music Recognition Systems Are Becoming Useful for Digital Libraries. In *Proceedings of Digital Libraries for Musicology (DLfM'18).* ACM, New York, NY, USA, Article 4, 5 pages. https: //doi.org/10.1145/3273024.3273034

1 INTRODUCTION

Optical Music Recognition (OMR), the field of computationally reading music notation in documents, is long known to hold significant promise for music libraries. The ability to search in vast archives of musical manuscripts using their content rather than solely their metadata would open entirely new avenues of large-scale research in digital musicology. A large number of compositions have never been recorded or digitized before; most of them probably exist only as manuscripts, since typesetting music has historically been a costly endeavor. As OMR is a cost-effective alternative to arrive at a structured encoding, it is, therefore, a key to significantly diversify the digitally available sources both to the general and professional audience. This applies to works from the 20$^{\text{th}}$ and 21$^{\text{st}}$ centuries as well: many are currently collecting dust in composers' private collections because there are insufficient resources to typeset them.

OMR is also known not to work very well [5, 19], and existing methods are rarely applicable beyond specific collections of scores. However, we believe recent OMR advances (e.g., [3, 16]) warrant revisiting this assertion. The contribution of this paper is to provide evidence for digital librarians and musicologists that current approaches to OMR can make it applicable in the following downstream scenarios: content-based retrieval, especially at the page level, of handwritten scores; melodic similarity matching; and digital musicology studies based on data aggregation.

Furthermore, the current OMR state of the art relies purely on supervised machine learning. Therefore, rather than demonstrating the use of an OMR system within a specific project (e.g., [6, 8]), our paper can be interpreted to set general expectations on the performance of the given OMR methods across analogous application scenarios, as long as comparable training data is available. The OMR methods are independent from specific use-cases, to the point where one can follow a "cookbook" to apply them to a new collection; costs are mainly shifted onto manual supervised data acquisition, which is a standardized, predictable task that does not require competitive computer vision expertise.

2 RELATED WORK

The PROBADO project [7], the Levy Collection [6], the OMRAS project [8], the digital version of the Liber Usualis [1, 22] within the SIMSSA project [12], the RISM project[1], the RILM project[2] and more recently PatternFinder [13] reflect the ongoing effort to create digital libraries of a large body of music and enable searching and indexing those collections. These systems feature powerful engines to evaluate a range of queries in an extensive database of symbolic music, e.g., searching by melody, by interval or looking

[1]http://opac.rism.info/metaopac/
[2]http://www.rilm.org/

DLfM '18, September 28, 2018, Paris, France

ACM ISBN 978-1-4503-6522-2/18/09.
https://doi.org/10.1145/3273024.3273034

up a particular note-sequence with optional wildcards. This power is enabled once a symbolic representation of the music is available — and without OMR, obtaining this representation has to be done manually, which is expensive and time-consuming.

Attempts have been made to use generic OMR to extract the requisite symbolic representation of music directly from musical score images, but the OMR component proved to be a weak point. In [10, 11], the authors describe how to match scanned sheet music to audio recordings automatically with an OMR algorithm doing the initial sheet music transcription. However, the evaluated algorithms produced such a large number of errors, that a subsequent correction was required before being able to match the OMR output to the audio representation. Similarly, in [2] the authors describe how to match musical themes from multiple sources using OMR, but the OMR output contained too many errors, and the authors had to resort to a drastically simplified representation, practically discarding note durations, clefs, and even absolute pitches.

OMR accuracy has been a significant bottleneck in the further development of similar applications to the extent that more success has been achieved by retrieving raw score images rather than their structured encodings [18] – but this does not provide the structured encodings that enable further processing and research.

3 OMR SYSTEMS

We showcase two OMR systems representing the current state of the art for obtaining the musical content from sheet music images. Both systems output a MIDI representation corresponding to the music score in the input image. First, we use a traditional full-pipeline system [16], applied to retrieval. Second, we use an end-to-end system [3]. Both systems are based on supervised learning with generic neural networks.

The full-pipeline approach (FP-OMR) builds off of the traditional OMR stages [20]. However, the detection method (U-Nets for semantic segmentation and a Connected Components detector [16]) jointly performs segmentation and classification on the input image directly, without removing the staff lines. Notation assembly is also performed with a machine learning method, as the Notation Graph representation [15] allows decomposing this step into a series of local decisions. MIDI is then inferred deterministically from the notation graph.[3] The advantages of this system are its applicability to arbitrarily complex music (given corresponding training data), the possibility of exporting the results in a rich representation such as MEI from the Notation Graph,[4] and its ability to operate on manuscripts, since the statistical methods can deal with the topological uncertainties of handwriting. Its disadvantages are that the symbol detection network is sensitive to low-level properties of the training dataset, thus requiring separate training sets for every source of data, and that the notation assembly model is currently underdeveloped.

The end-to-end approach (E2E-OMR) uses a convolutional recurrent neural network (CRNN) that is capable of providing the sequence of music symbols from the image of a single staff [3]. The term end-to-end signifies that the model is trained to directly produce the correct sequence of musical events, *without* providing geometric information of where each symbol is located. Although this reduces the effort when creating the ground-truth data, the CRNN design is so far inherently limited to single-staff, monophonic music. The system has only been trained on born-digital printed scores, but with artificial distortions to simulate more realistic score images.

[3]Implemented in https://github.com/hajicj/muscima.

[4]Theoretically. Only MIDI export is currently implemented.

4 RETRIEVAL EXPERIMENTS

We define several retrieval tasks over a small test collection, evaluated with common retrieval metrics. The similarity between two MIDI files is computed using Dynamic Time Warping over the pitch sequences (discarding durations, which are still too unreliable), similar to [2]. We assume a human user will verify retrieved items from a ranked list and stop when the first non-duplicate score is returned. For this, we return Mean Average Precision (MAP@k), where k is the number of duplicates for a given page in the collection (in our case, MAP@49).

As the retrieval collection, we use CVC-MUSCIMA [9]. This dataset contains 20 distinct pages of music, each copied by 50 people, for a total of 1000 images. Since the individual pieces exhibit a significant amount of variability, using the entire collection would make the problem extremely easy. For that reason, we select a *confuse-retrieval* subset of 7 pages. While decreasing the collection size would typically improve retrieval performance, in this case, the remaining 13 pieces are so distinctive that including them would make the collection *less* challenging. One advantage of this dataset is that we know in advance how many copies of each page exist in the database, so the experiments in this section can thus be seen as indicators of the general ability of the OMR system to deal with retrieving manuscripts with different handwriting styles.

We prepare all the MIDI representations of the score images in this section with the full-pipeline system (FP-OMR), as it is capable of dealing with entire pages instead of just individual staffs. We investigate page retrieval when querying with full pages (e.g., searching for copies of a piece) or just with snippets (searching for pages using individual staffs).

4.1 Page Queries

Musical manuscripts were often manually copied; in large collections and across collections, there may be duplicates of the same music that are accidentally kept as a separate composition. One might want to discover such copies automatically. This is the first retrieval task we simulate.

In principle, this task is easy once OMR systems achieve results somewhat above a random baseline. The collection is quite small – 350 pages in total. In the MIDI representation, pages are long sequences in a very sparse space, so any minimally robust similarity score should yield good results.[5] Page retrieval is therefore a natural starting point for demonstrating whether OMR systems are useful for anything at all: if an OMR system fails on this task, it can hardly be expected to be useful anywhere else. So far, we are not aware of any OMR system that can handle handwritten music scores even with remote success.

[5]One does not even necessarily have to use the musical content of the scores to match them, given a smart enough algorithm dealing with the score images. However, "smart enough" may be daunting, as one would have to contend with different handwriting styles, segmentations of scores into staffs and pages, etc.

	MAP@1	MAP@10	MAP@49
Page queries, OMR2OMR	1.0	1.0	0.998
Page queries, cross-modal	1.0	1.0	0.998
Snippet queries, OMR2OMR	0.928	0.834	0.763
Snippet queries, cross-modal	0.606	0.610	0.577

Table 1: Results for page retrieval using page queries and snippet queries under two modalities: using OMR for creating the database and the query (OMR2OMR) or just for the database (cross-modal) and query with ground-truth MIDI.

	Spearman		Pearson	
Query	M_1	M_2	M_1	M_2
OMR2OMR	94.0	96.9	96.4	97.0
cross-modal	93.8	97.1	97.0	97.1

Table 2: Average Spearman's and Pearson's correlation coefficients (in %) for the similarity between the original MIDI file and the MIDI file generated by the OMR system. M_1 and M_2 refers to *ShapeH* and *Time* symbolic similarity functions, respectively, from Urbano's *MelodyShape* library.

4.2 Snippet Queries

One might want to search not only using entire pieces, but also with shorter segments. We imagine musicologists, e.g., tracing the genealogy of a musical thought throughout a substantial body of work, or looking for musical citations across a geographic area. Here, the query is much shorter, and therefore OMR mistakes matter proportionally more.

4.3 Evaluation and Results

Both tasks are evaluated in two modalities: when the database and the query are created using the same OMR system (OMR2OMR), and when only the database is created by the OMR system and the queries are taken from the ground truth MIDI (cross-modal: simulating searching a sheet music database with, e.g., a keyboard capture sequence). If both the query and the database are processed with the same OMR system, some of the system's limitations may cancel out (e.g., ignoring key signatures), whereas when querying a sheet music database with MIDI from a different source, these limitations come to light.

The retrieval results are shown in Table 1. The FP-OMR system can deal with manuscripts of CVC-MUSCIMA well enough to retrieve copies of the same score reliably. When snippets are used as queries, the applicability of the system would depend on the specific scenario; the results in Table 1, row 3 indicate that the OMR system will be better suited in situations that require precision rather than recall. In the cross-modal setting, the simplifications made by [16] render the system practically useless at the granularity of individual staffs.

5 SYMBOLIC MUSIC SIMILARITY

Besides content-based retrieval, one may have various other reasons to compute similarity over symbolic representations of music [13, 14, 17, 21]. As we cannot realistically evaluate OMR systems in all these settings, we can instead try to measure how the errors made by OMR systems influence the behavior of the standard symbolic similarity metrics.

Figure 1: Sample from PrIMuS dataset, synthetically distorted to resemble non-ideal sheet conditions.

We use Urbano's *MelodyShape* library[6] as the battery of standard metrics, available for measuring symbolic music similarity. Specifically, we consider *ShapeH* (M_1) and *Time* (M_2) similarity functions [24], as these ranked top in previous editions of the MIREX Symbolic Melodic Similarity challenge.[7]

The data used for this experiment corresponds to a subset of the PrIMuS dataset [4], which contains synthetically rendered scores of real music incipits from the RISM database. An incipit is the opening sequence of a song and can be used for the identification of a musical work. Therefore, they represent suitable musical elements for showcasing OMR-based search. We specifically consider the partition of images that have been distorted to resemble difficult conditions that might appear in some real cases [3]. An example from this collection is shown in Fig. 1.

The experiment considers the similarity between an incipit that acts as a query, and each sample of two sets of 1500 incipits: the real (ground truth) MIDI files and the MIDI files generated by the E2E-OMR system. For evaluation, Spearman's and Pearson's correlation coefficients are computed between the similarities obtained from the same query in both datasets. While Spearman's coefficient measures only whether the relationship is monotonous, Pearson's coefficient also measures if the relationship is linearly correlated. The higher these correlation coefficients are, the more smoothly an OMR system can substitute human input, to provide MIDI in applications where the given similarity function is used. A total of 1000 such queries were made, and the averaged coefficients are reported. In addition, we study the same two modalities as before: in the first one, the query is the MIDI output of the OMR system that read the image (OMR2OMR); in the other one, the query is taken from the ground-truth MIDI representation (cross-modal).

The results of this experiment are provided in Table 2. In most cases, the correlation coefficients are higher than 95 %. Reflecting the observations in [24], OMR errors perturb M1 more with respect to the rank-aware Spearman's correlation. Considering the high figures of the Pearson's coefficient, reorderings caused by OMR mistakes are likely to occur for samples that are very similar anyway. With respect to the M_2 metric, fewer reorderings are observed, while rank-unaware correlations remain the same.

[6] https://github.com/julian-urbano/MelodyShape
[7] See http://www.music-ir.org/mirex/wiki/2015:Symbolic_Melodic_Similarity_Results for further details.

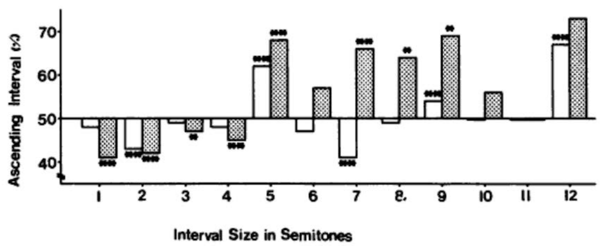

Figure 2: Original figure from [25] summarizing their quantitative results. Note that this figure compares data for two stylistically different datasets: western music (white), and folk tunes (gray). The authors were looking for how the difference between ascending/descending interval distributions could be used to distinguish melodies originating from the music of the respective styles.

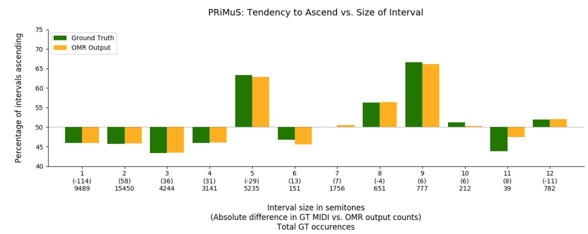

Figure 3: Ascending vs. descending tendency by interval size in semitones, comparing ground truth MIDI and OMR outputs on the monophonic PrIMuS dataset. The dataset is monostylistic; colors differentiate the method of MIDI acquisition. We observe comparable tendencies to the Composer data (white bars in Fig. 2).

6 CASE STUDY FOR DIGITAL MUSICOLOGY

So far, we have shown the extent to which current state-of-the-art OMR can enhance a digital library's indexing and search. In this section, we illustrate how OMR systems can be useful to musicologists when working with such an enriched collection.

As a model for a musicological investigation that could plausibly benefit from OMR, we use the work of Vos and Troost [25]. Its authors propose characterizing the Western classical music genre based on the joint distribution of interval sizes and direction (ascending vs. descending), and compare these against a corpus of non-artificial music: both quantitatively and in a perceptual experiment.[8] We re-trace the quantitative portion of their work, showing that in this data aggregation scenario, the OMR systems would lead the researcher to propose the same hypothesis while obviating the need for manual data entry and checking. The errors that OMR introduces are offset by the vastly greater scale at which data aggregation is enabled, compared to manual data entry.

The quantitative findings of [25] are summarized in Fig. 2 and reproduced in Fig. 3. We compute the same distribution from the PrIMuS set of incipits [4]. This dataset is stylistically the same as the Dictionary of Musical Themes (DMT), although it is not limited to themes. The ascending/descending interval distributions are shown in Fig. 3. We found that for no interval size the balance between its ascending and descending instances was significantly different between the ground truth MIDI and OMR outputs (two-tailed binomial test at levels 0.05 and 0.01, following [25]).

Comparing the figures 2 and 3, one could discover the same trends. There are meaningful differences for the fifth and the octave, which may be because the DMT only contains prominent melodies, while PrIMuS data contains all incipits, including those from middle voices. However, our point is rather that one can see the same trend both in the ground truth MIDI and OMR outputs, indicating that the manual labor of data acquisition in [25] can be avoided using OMR without substantially putting the conclusions into question.

[8] This study has been cited over 180 times and is used, e.g., to illustrate the functionality of the MIDItoolbox software [23].

7 CONCLUSIONS

We have attempted to show on several scenarios that recent advances in OMR state of the art have, to an extent, made OMR a more relevant technology. We believe these advances, especially given the underlying generic machine learning methodology, have implications for designing and enriching digital collections of sheet music. Being aware of these advances can be valuable for various stakeholders such as librarians and musicologists.

The showcased methods still have inherent limitations. Chiefly, learning does not transfer easily between datasets. The currently best-performing methods require re-training for each archive, even though the notation style may be the same. This implies that for every use-case, manual annotations will be necessary, and it is difficult to estimate in advance how much data will be enough. Furthermore, the systems are still not accurate enough to provide functionality such as playback or structured encoding. Beyond sufficient accuracy, further concerns also remain before "full-text" search in music can be done at a truly massive scale – efficient representations of music notation and its indexing, multimedia linking (such as lyrics alignment), and user interface design.

Overall, we conclude that the current state of the art in OMR enables (1) adding content-based similarity and retrieval functionality to music score image databases, especially for use-cases that do not require fine granularity, (2) applications based on symbolic melodic similarity, (3) research in digital musicology that builds on aggregating massive amounts of data and quantitative conclusions. The experiments in this paper should be considered as supporting evidence for these conclusions. We hope that the interested reader will find the reported capabilities of state-of-the-art OMR worth considering.

ACKNOWLEDGMENTS

Jan Hajič jr. acknowledges support by the Czech Science Foundation grant no. P103/12/G084, Charles University Grant Agency grants 1444217 and 170217, and by SVV project 260 453; Marta Kolárová is supported by Charles University Grant Agency grant 1444217.

REFERENCES

[1] Andrew Hankinson, John Ashley Burgoyne, Gabriel Vigliensoni, Alastair Porter, Jessica Thompson, Wendy Liu, Remi Chiu, and Ichiro Fujinaga. 2012. Digital Document Image Retrieval Using Optical Music Recognition. In *Proceedings of the 13th International Society for Music Information Retrieval Conference*, Fabien Gouyon, Perfecto Herrera, Luis Gustavo Martins, and Meinard Müller (Eds.). 577–582.

[2] Stefan Balke, Sanu Pulimootil Achankunju, and Meinard Müller. 2015. Matching Musical Themes based on noisy OCR and OMR input, In 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings* 2015-August (2015), 703–707. DOI: http://dx.doi.org/10.1109/ICASSP.2015.7178060

[3] Jorge Calvo-Zaragoza and David Rizo. 2018. Camera-PrIMuS: Neural End-to-End Optical Music Recognition on Realistic Monophonic Scores. In *19th International Society for Music Information Retrieval Conference*. (in press).

[4] Jorge Calvo-Zaragoza and David Rizo. 2018. End-to-End Neural Optical Music Recognition of Monophonic Scores. *Applied Sciences* 4 (2018). DOI: http://dx.doi.org/10.3390/app8040606

[5] L.iang Chen and Kun Duan. 2016. MIDI-assisted egocentric optical music recognition, In 2016 IEEE Winter Conference on Applications of Computer Vision (WACV). *2016 IEEE Winter Conference on Applications of Computer Vision, WACV 2016* (2016). DOI: http://dx.doi.org/10.1109/WACV.2016.7477714 cited By 0; Conference of IEEE Winter Conference on Applications of Computer Vision, WACV 2016 ; Conference Date: 7 March 2016 Through 10 March 2016; Conference Code:121834.

[6] G. Sayeed Choudhury, M. Droetboom, Tim DiLauro, Ichiro Fujinaga, and Brian Harrington. 2000. Optical Music Recognition System within a Large-Scale Digitization Project. In *1st International Symposium on Music Information Retrieval*. 1–6.

[7] Jürgen Diet and Frank Kurth. 2007. The Probado Music Repository at the Bavarian State Library.. In *ISMIR*. 501–504.

[8] Matthew J. Dovey. 2004. Overview of the OMRAS project: Online music retrieval and searching. *Journal of the American Society for Information Science and Technology* 55, 12 (2004), 1100–1107. DOI: http://dx.doi.org/10.1002/asi.20063

[9] Alicia Fornés, Anjan Dutta, Albert Gordo, and Josep Lladós. 2012. CVC-MUSCIMA: a ground truth of handwritten music score images for writer identification and staff removal. *International Journal on Document Analysis and Recognition (IJDAR)* 15, 3 (2012), 243–251. DOI: http://dx.doi.org/10.1007/s10032-011-0168-2

[10] C. Fremerey, D. Damm, F. Kurth, and M. Clausen. 2009. Handling Scanned Sheet Music and Audio Recordings in Digital Music Libraries. In *Proceedings of the International Conference on Acoustics NAG/DAGA*. 1–2.

[11] Christian Fremerey, Meinard Müller, Frank Kurth, and Michael Clausen. 2008. Automatic mapping of scanned sheet music to audio recordings. In *Proceedings of the 9th International Society for Music Information Retrieval Conference (ISMIR 2008)*. 413–418. http://ismir2008.ismir.net/papers/ISMIR2008_116.pdf

[12] Ichiro Fujinaga, Andrew Hankinson, and Julie E. Cumming. 2014. Introduction to SIMSSA (Single Interface for Music Score Searching and Analysis). In *Proceedings of the 1st International Workshop on Digital Libraries for Musicology*. ACM, 1–3. DOI: http://dx.doi.org/10.1145/2660168.2660184

[13] David Garfinkle, Claire Arthur, Peter Schubert, Julie Cumming, and Ichiro Fujinaga. 2017. PatternFinder: Content-Based Music Retrieval with Music21. In *Proceedings of the 4th International Workshop on Digital Libraries for Musicology (DLfM '17)*. ACM, New York, NY, USA, 5–8. DOI: http://dx.doi.org/10.1145/3144749.3144751

[14] Joe George and Lior Shamir. 2014. Computer analysis of similarities between albums in popular music. *Pattern Recognition Letters* 45 (2014), 78–84.

[15] Jan jr. Hajič and Pavel Pecina. 2017. The MUSCIMA++ Dataset for Handwritten Optical Music Recognition. *Proceedings of the 14th IAPR International Conference on Document Analysis and Recognition* (2017).

[16] Jan Hajič jr., Matthias Dorfer, Gerhard Widmer, and Pavel Pecina. 2018. Towards Full-Pipeline Handwritten OMR with Musical Symbol Detection by U-Nets. In *19th International Society for Music Information Retrieval Conference*. (in press).

[17] Alan Marsden. 2012. Interrogating Melodic Similarity: A Definitive Phenomenon or the Product of Interpretation? *Journal of New Music Research* 41, 4 (2012), 323–335.

[18] Matthias Dorfer, Andreas Arzt, and Gerhard Widmer. 2017. Learning Audio-Sheet Music Correspondences for Score Identification and Offline Alignment. In *Proceedings of the 18th International Society for Music Information Retrieval Conference, ISMIR 2017, Suzhou, China, October 23-27, 2017*, Sally Jo Cunningham, Zhiyao Duan, Xiao Hu, and Douglas Turnbull (Eds.). 115–122. https://ismir2017.smcnus.org/wp-content/uploads/2017/10/32_Paper.pdf

[19] Alexander Pacha and Horst Eidenberger. 2017. Towards Self-Learning Optical Music Recognition. In *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*. 795–800. DOI: http://dx.doi.org/10.1109/ICMLA.2017.00-60

[20] Ana Rebelo, Ichiro Fujinaga, Filipe Paszkiewicz, Andre R.S. Marcal, Carlos Guedes, and Jaime S. Cardoso. 2012. Optical music recognition: state-of-the-art and open issues. *International Journal of Multimedia Information Retrieval* 1, 3 (March 2012), 173–190. DOI: http://dx.doi.org/10.1007/s13735-012-0004-6

[21] David Rizo. 2010. *Symbolic music comparison with tree data structures*. Ph.D. Dissertation. Universidad de Alicante.

[22] Jessica Thompson, Andrew Hankinson, and Ichiro Fujinaga. 2011. Searching the Liber Usualis: Using CouchDB and ElasticSearch to Query Graphical Music Documents. In *Proceedings of the 12th International Society for Music Information Retrieval Conference*. http://ismir2011.ismir.net/latebreaking/LB-10.pdf

[23] P. Toiviainen and T. Eerola. 2016. MIDI toolbox 1.1. https://github.com/miditoolbox/. (2016).

[24] Julián Urbano. 2013. *MIREX 2013 Symbolic Melodic Similarity: A Geometric Model supported with Hybrid Sequence Alignment*. Technical Report. Music Information Retrieval Evaluation eXchange.

[25] Piet G. Vos and Jim M. Troost. 1989. Ascending and Descending Melodic Intervals: Statistical Findings and their Perceptual Relevance. *Music Perception* 6, 4 (1989), 383–396.