Why WoRMS?

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

***Abstract*—In this paper the organization of the first Workshop on Reading Music Systems (WoRMS) is motivated. The event is intended to become a forum for discussion and collaborative projects for all people interested in the optical music recognition and related applications. This first edition is held as a satellite event of the 19th International Society for Music Information Retrieval Conference. It is our hope that WoRMS helps to set the basis towards building a stable community around the field of music reading systems.**

***Index Terms*—Optical Music Recognition, Automatic Music Reading, Music Information Retrieval, Music Digital Libraries**

I. Introduction

Optical music recognition (OMR) is the research field that investigates how to make computers be capable of reading music. It is a very attractive field from different points of view. The obvious one is the hope of providing an efficient way of converting written music to a symbolically encoded format. On the other hand, the field is exciting for research from the perspective of computer science, machine learning and pattern recognition: while it might be true that the task of reading music computationally shares similarities with others (for instance, the optical recognition of text), music notation has enough nuances that make it appropriate to be considered as a specific field [1], and whose specific research cannot be easily generalized from other domains.

Despite all the above, OMR has experienced a slow progress. One of the main reasons for such a troublesome development is that there are few people devoted to working it out. Although the different perspectives of OMR make it really interesting for general research, it is difficult to find single specialists that cover all related areas (image processing, machine learning, computer music, and music notation in general). Perhaps, people who do not meet all these requirements feel some sense of vertigo before facing the problem, thereby choosing another field of research.

More importantly, we believe that there is a more relevant obstacle preventing a proper development of the field: an evident lack of communication in the community. Over the years, many techniques have been proposed to deal with OMR-related problems [2], but there are hardly re-uses of the work done previously by other researchers. This might be caused by the heterogeneity inherent to the field: publications in OMR are rather scattered. The purely methodological aspect is more related to machine learning, pattern recognition and document analysis — and there are many journals and conferences on these topics; however, the music information retrieval and digital libraries communities are the ones especially interested in OMR results, and so many publications have also fallen in this type of venues. Consequently, the different OMR authors rarely meet in person.

And it is for all the above, that the OMR community can hardly set the basis of the field so that researchers do not need to reinvent the wheel over and over again. There have been so many different approaches, data formats, evaluation criteria, and datasets that a more formal development of the field is somehow difficult to attain. That is why the organization of an event about optical recognition technologies for music notation becomes necessary.

II. A Workshop on Music Reading Systems

The 12th IAPR International Workshop on Graphics Recognition (GREC'17), co-located with the 14th IAPR International Conference on Document Analysis and Recognition (ICDAR'17), gathered a number of authors working in OMR (including the keynote!)[1]. In spite of some notable absences, it was a unique opportunity to hold a discussion group entirely dedicated to OMR [3]. However, it is important to note that this concentration of OMR researchers was thanks to the efforts of one of the organizers of GREC'17 (Alicia Fornés), who personally contacted many of these researchers to make contributions to the workshop.

During the discussion session, the latest contributions in the field were discussed: deep learning, as in other communities, provides a way of training powerful and accurate models, and conventional tasks (such as symbol classification or staff-line removal) no longer represent actual challenges [4], [5]. In fact, there is a current trend towards systems that decompose the recognition workflow in fewer stages, such as the direct detection of musical objects in the images [6] or end-to-end systems [7], [8]. In addition, public datasets recently released, such as MUSCIMA++ [9], PrIMuS [10] or DeepScores [11], allow researchers to perform comprehensive and reproducible experiments. Furthermore, given the complexity of the recognition task—which makes it difficult to think of achieving perfect results—another prominent line of current research is about

This work is supported by the Spanish Ministerio de Economía, Industria y Competitividad through Juan de la Cierva - Formación grant (Ref. FJCI-2016-27873).

[1] http://grec2017.loria.fr

interactive approaches [12], [13], where the human-machine interaction plays a key role in the workflow. Nevertheless, one issue upon which all the participants in the discussion agreed was the lack of adequate evaluation measures [14], [15], which indeed points to incomplete formulations of the OMR field in terms of input-output tasks.

The main conclusion of the discussion group was that OMR needs to properly define itself. This means finding proper formulations of its associated tasks, along with their corresponding evaluation criteria, that can be shared by the entire community. Hence, the idea of organizing an event exclusively dedicated to the technologies of reading systems for music notation naturally arose.

As mentioned on the website[2], the international Workshop on Reading Music Systems (WoRMS) is "*a novel workshop that tries to connect researchers who develop systems for reading music, such as in the field of Optical Music Recognition, with other researchers and practitioners that could benefit from such systems, like librarians or musicologists*".

It is worth emphasizing that the workshop is about systems that read documents of written music, which is not strictly exclusive of OMR: tasks such as score following or palaeographic analysis do have room in WoRMS, as long as the objects of study are related to this kind of documents. Furthermore, the workshop explicitly welcomes people without a technical background, yet interested in either using the technology or carrying out focused projects.

This first edition of the event is held as a satellite event of the 19th International Society for Music Information Retrieval Conference (ISMIR'18). Given that there exist several events about machine learning applied to document analysis (such as the International Conference on Document Analysis and Recognition and the International Conference on Frontiers in Handwriting Recognition), we believe that it is more interesting to get WoRMS closer to the community where applications may find real use, namely the Music Information Retrieval one. If the idea is to define the shape of the field, we definitely need to know how to formulate the tasks according to end users' needs, and therefore celebrating the event within ISMIR'18 may bring this type of audience as well.

III. Conclusions

Technology for the development of music reading systems is an exciting research avenue with a number of potential applications. However, the field might not have sufficient maturity due to the lack of a communicative community. This prevents the proper establishment of task formulations, evaluation protocols and generalizable techniques that can be exploited by other researchers.

The organization of WoRMS is a historic opportunity to set the basis towards an agreement on these issues, combining both the technical developments and the specific aspects of the different applications. We hope WoRMS will provide an opportunity to share ideas and discuss about current issues, with the aim of building a stable community around the field of music reading systems.

[2]https://sites.google.com/view/worms2018

Acknowledgements

I would like to thank my colleagues Jan Hajič jr. and Alexander Pacha for sharing the excitement of organizing this event. We do acknowledge the support from Ichiro Fujinaga and Philippe Rigaux, who helped us to make WoRMS become real.

References

[1] D. Bainbridge and T. Bell, "The challenge of optical music recognition," *Computers and the Humanities*, vol. 35, no. 2, pp. 95–121, 2001.

[2] A. Rebelo, I. Fujinaga, F. Paszkiewicz, A. R. S. Marçal, C. Guedes, and J. S. Cardoso, "Optical music recognition: state-of-the-art and open issues," *International Journal of Multimedia Information Retrieval*, vol. 1, no. 3, pp. 173–190, 2012.

[3] J. Calvo-Zaragoza, J. Hajič jr., and A. Pacha, "Optical music recognition discussion group at the graphics recognition workshop 2017," in *12th International Workshop on Graphics Recognition. Revised Selected Papers*, 2018, (in press).

[4] A.-J. Gallego and J. Calvo-Zaragoza, "Staff-line removal with selectional auto-encoders," *Expert Systems with Applications*, vol. 89, pp. 138–148.

[5] A. Pacha and H. Eidenberger, "Towards a universal music symbol classifier," in *Proceedings of the 12th IAPR International Workshop on Graphics Recognition*. IEEE Computer Society, pp. 35–36.

[6] A. Pacha, K.-Y. Choi, B. Coüasnon, Y. Ricquebourg, R. Zanibbi, and H. Eidenberger, "Handwritten music object detection: Open issues and baseline results," in *2018 13th IAPR Workshop on Document Analysis Systems (DAS)*, 2018, pp. 163–168.

[7] E. van der Wel and K. Ullrich, "Optical music recognition with convolutional sequence-to-sequence models," in *Proceedings of the 18th International Society for Music Information Retrieval Conference*, 2017, pp. 731–737.

[8] J. Calvo-Zaragoza and D. Rizo, "End-to-end neural optical music recognition of monophonic scores," *Applied Sciences*, no. 4, p. 606, 2018.

[9] J. Hajic and P. Pecina, "The MUSCIMA++ dataset for handwritten optical music recognition," in *14th IAPR International Conference on Document Analysis and Recognition*, 2017, pp. 39–46.

[10] J. Calvo-Zaragoza and D. Rizo, "Camera-PrIMuS: Neural end-to-end optical music recognition on realistic monophonic scores," in *19th International Society for Music Information Retrieval Conference*, 2018, (in press).

[11] L. Tuggener, I. Elezi, J. Schmidhuber, M. Pelillo, and T. Stadelmann, "DeepScores - A Dataset for Segmentation, Detection and Classification of Tiny Objects," *Computer Research Repository*, vol. abs/1804.00525, 2018.

[12] L. Chen, E. Stolterman, and C. Raphael, "Human-interactive optical music recognition," in *Proceedings of the 17th International Society for Music Information Retrieval Conference*, 2016, pp. 647–653.

[13] G. Vigliensoni, J. Calvo-Zaragoza, and I. Fujinaga, "An environment for machine pedagogy: Learning how to teach computers to read music," in *Joint Proceedings of the ACM IUI 2018 Workshops co-located with the 23rd ACM Conference on Intelligent User Interfaces*.

[14] D. Byrd and J. G. Simonsen, "Towards a standard testbed for optical music recognition: Definitions, metrics, and page images," *Journal of New Music Research*, vol. 44, no. 3, pp. 169–195, 2015.

[15] J. Hajič jr., J. Novotný, P. Pecina, and J. Pokorný, "Further steps towards a standard testbed for optical music recognition," in *Proceedings of the 17th International Society for Music Information Retrieval Conference*, 2016, pp. 157–163.