Using Optical Character Recognition to Process Handwritten Musical Notation

Tim Kang
University of California, San Diego
9500 Gilman Dr, La Jolla CA
tikang@ucsd.edu

Michael Perry
University of California, San Diego
9500 Gilman Dr, La Jolla CA
m2perry@ucsd.edu

Abstract

This project aims to use machine learning as a way to make the optical character recognition of handwritten modern musical notation more accurate.

1. Qualifications

Tim Kang is a junior at the University of California, San Diego. He worked last summer at the CALIT2 Immersive Visualization Lab and has extensive knowledge of music theory. Michael Perry is a senior at the University of California, San Diego. He has taken CSE150 "Programming Languages for AI" and has also travelled to Taiwan as part of the PRIME program where he did research on bee imaging.

2. Milestones

A brief list of four to six milestones, and deadlines to achieve the milestones.

2.1. Data Phase

Deadline (4-1-11): We will gather a sizeable amount of data so that a lot more data would not be required to be gathered later on. We need some data (sheet music) that our program will be able to read in. This is complicated by the fact that it should be handwritten. Computer typeset music still serves a purpose, however and can be used as early training.

2.2. Initial Implementation

Deadline (4-20-11): We will spend the next few weeks coding and implementing our program. Then we will finish running the data through the first prototype of the program and determine the problem areas. Change as needed.

2.3. Secondary Analysis

Deadline (5-10-11): Finish running the data through the second prototype. At this point, our program will have done a lot of learning on a large set already.

2.4. Final Analysis and Testing

Deadline (5-29-11): Starting testing and running data through the final prototype. At this point, it should be close to a working prototype. Make final adjustments and prepare to show to the public.

3. Division of Labor

Michael will contribute his AI and computer vision expertise. Both Tim and Michael will contribute to the programming. Tim will type up this document in LaTeX and look up relevant research papers.

4. A Few Questions

4.1. How to distinguish between the staff and actual music notation itself?

This will be probably one of our harder challenges. Unlike English characters, sheet music is written on a staff, which causes additional difficulties when trying to it digitally. We will need to find a way to somehow separate the notes from the staff. As this is a common problem, though, it probably has already been encountered and we might be able to find a solution by looking through some old research papers.

4.2. How to account for differences in the way that music notation is depicted by different persons?

Everybody has differences in their handwriting. This is especially true in the case of music notation. I've seen extremely bizarre depictions of bass clef in my time. Our program will probably figure out what is on the page using a statistical / probabilistic approach.

4.3. How to account for extraneous markings on the page that are not part of the score?

This is a problem also encountered in other forms of optical character recognition, but it is especially important in music. Missing a dot in an English sentence is not as important as missing a dot in a musical phrase.

5. Software

We will mainly use a subset of C++, along with the OpenCV library. It is possible that we might use other libraries or programming languages, as a convenience.

5.1. Datasets

We will mainly use existing sheet music to build up our data set. There is a large amount of sheet music that can be downloaded online, much of it handwritten. For example, one such site we could use is http://www.imslp.org.

References

[1] Arica, N., Yarman-Vural, F.: An overview of character recognition focused on off-line handwriting. IEEE Trans. Syst., Man, Cybern., Part C: Applica. Rev. 31(2), 216233 (2001). doi:10.1109/5326. 941845

[2] Bainbridge, D.: An extensible optical music recognition system. In: Nineteenth Australasian Computer Science Conference, pp. 308317 (1997)

[3] Bellini,P.,Bruno,I.,Nesi,P.:Optical music sheet segmentation.In: Proceedings of the 1st International Conference on Web Delivering of Music, pp. 183190 (2001)

[4] Blostein, D., Baird, H.S.: A critical survey of music image analy- sis. In: Baird Bunke, Y. (ed.) Structured Document Image Analy- sis, pp. 405434. Springer, Heidelberg (1992)

[5] Bojovic, M., Savic, M.D.: Training of hidden Markov models for cursive handwritten word recognition. In: ICPR 00: Proceedings of the International Conference on Pattern Recognition, p. 1973. IEEE Computer Society, Washington, DC, USA (2000)

[6] Capela, A., Rebelo, A., Cardoso, J.S., Guedes, C.: Staff line detection and removal with stable paths. In: Proceedings of the International Conference on Signal Processing and Multimedia Applications (SIGMAP 2008), pp. 263270 (2008). http://www.inescporto.pt/ jsc/publications/conferences/ 2008ACapelaSIGMAP.pdf

[7] Cardoso, J.S., Capela, A., Rebelo, A., Guedes, C.: A connected path approach for staff detection on a music score. In: Proceedings of the International Conference on Image Processing (ICIP 2008), pp. 10051008 (2008)

[8] Cardoso, J.S., Capela, A., Rebelo, A., Guedes, C., da Costa, J.P.: Staff detection with stable paths. IEEE Trans. Pattern Anal. Mach. Intell. 31(6), 11341139 (2009). doi:10.1109/TPAMI.2009.34

[9] Coasnon, B., Camillerapp, J.: Using grammars to segment and recognize music scores. In: Proceedings of DAS-94: International Association for Pattern Recognition Workshop on Document Analysis Systems, pp. 1527. Kaiserslautern (1993)

[10] Dalitz, C., Droettboom, M., Czerwinski, B., Fujigana, I.: Staff removal toolkit for gamera (20052007). http://music-staves. sourceforge.net

[11] Dalitz, C., Droettboom, M., Czerwinski, B., Fujigana, I.: A com- parative study of staff removal algorithms. IEEE Trans. Pattern Anal. Mach. Intell. 30, 753766 (2008)

[12] Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification (2nd Edn.). Wiley, New York (2000)

[13] Forns, A., Llads, J., Snchez, G.: Primitive segmentation in old handwritten music scores. In: Liu, W., Llads, J. (eds.) GREC, Lecture Notes in Computer Science, vol. 3926, pp. 279290.

[14] Fujinaga,I.:Staff detection and removal.In:George,S.(ed.)Visual Perception of Music Notation: On-Line and Off-Line Recogni- tion, pp. 139. Idea Group Inc, Hershey (2004)

[15] Kopec, G.E., Parc, P.A.C., Maltzcarnegie, D.A.: Markov source model for printed music decoding. J Electron Imaging, pp. 714 (1996)

[16] Mitobe, Y., Miyao, H., Maruyama, M.: A fast HMM algorithm based on stroke lengths for on-line recognition of handwritten music scores. In: IWFHR 04: Proceedings of the Ninth Inter- national Workshop on Frontiers in Handwriting Recognition, pp. 521526. IEEE Computer Society, Washington (2004). doi:10. 1109/IWFHR.2004.2

[17] Miyao, H., Nakano, Y.: Note symbol extraction for printed piano scores using neural networks. IEICE Trans. Inf. Syst. E79D, 548554 (1996)

[18] Miyao,H.,Okamoto,M.:Stave extraction for printed music scores using DP matching. J Adv. Comput. Intell. Intell. Inform. 8, 208 215 (2007)

[19] Ng, K.: Optical music analysis for printed music score and hand- written music manuscript.In:George,S.(ed.)Visual Perception of Music Notation: On-Line and Off-Line Recognition, pp. 108127. Idea Group Inc, Hershey (2004)

[20] Pugin, L.: Optical music recognition of early typographic prints using hidden Markov models. In: ISMIR, pp. 5356 (2006)

[21] Randriamahefa, R., Cocquerez, J., Fluhr, C., Pepin, F., Philipp, S.: Printed music recognition. In: Proceedings of the Second Inter- national Conference on Document Analysis and Recognition, pp. 898901 (1993). doi:10.1109/ICDAR.1993.395592

[22] Reed, K.T., Parker, J.R.: Automatic computer recognition of printed music. Proc. 13th Int. Conf. Pattern Recognit. 3, 803 807 (1996). doi:10.1109/ICPR.1996.547279