CLASSIFYING BACH'S HANDWRITTEN C-CLEFS

Masahiro Niitsuma **Yo Tomita**
School of Muisc and Sonic Arts, Queen's University, Belfast
mniitsuma01@qub.ac.uk y.tomita@qub.ac.uk

ABSTRACT

The aim of this study is to explore how we could use computational technology to help determination of the chronology of music manuscripts.

Applying a battery of techniques to Bach's manuscripts reveals the limitation in current image processing techniques, thereby clarifying future tasks. Analysis of C-clefs, the chosen musical symbol for this study, extracted from Bach's manuscripts dating from 1708–1748, is also carried out. Random forest using 15 features produces significant accuracy for chronological classification.

1. INTRODUCTION

In the development of western music, handwritten scores and parts have played a significant role even after the invention of making prints because they allowed composers to express their ideas in a personalized way. In manuscripts, the writer's intention is assumed to be present, and manuscripts are often the only surviving witness for them and their work, and for this reason, they should be analyzed with utmost care and attention.

Although optical music recognition (OMR) has been investigated actively for this, there has been little research investigating such aspects of music manuscripts beyond OMR. Enote history [1] [3, 4] and the researches by Fornes [10] are such examples, which deals with such as writer identification or how just a subtle change of handwriting could reveal the situation under which the writer was working.

This paper explores the analysis of Bach's C-clefs and we associates the image processing issues. C-clefs have been identified by Bach scholars as one of the most crucial criteria to date the manuscript. Musicologists such as Dadelsen [21] and Emery [9] claim that Bach's C-clef can be categorized into three or four groups and each group mainly appears in a specific period. Dadelsen applied this to identification of the chronological order of Bach's manuscripts. One of the weaknesses of their discussions seems to be the lack of any quantitative evaluation of their hypothesis. Their investigation is apparently supported by their deep background knowledge and experience, which cannot be easily emulated by computer. This lack of reproducibility of their research can be addressed in the musicology of the future. High reproducibility is in fact one of the biggest advantages of computational analysis.

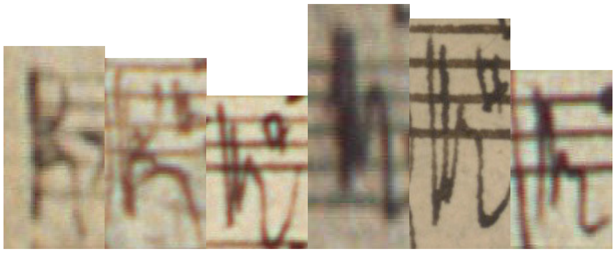
Figure 1 shows the C-clefs found in Bach's manuscripts arranged in a chronological order suggested by musicologists [14], which demonstrates that the shape of Bach's handwriting changed over time. Bach scholars investigate the issue of chronology by examining various types of evidence holistically. Evidence typically include watermarks, handwriting, a documented use of the manuscripts giving clues to specific dates, notational styles, and librettists. It seems risky, therefore, to draw a conclusion by contemplating only a single type of evidence such as C-clefs.

However, computational analysis can offer a totally objective and independent result, which can then be combined with other sources and knowledge such as the evidence mentioned above, which will hopefully lead to more reliable results. Can computational analysis offer the same conclusions as those arrived at by musicologists? The remainder of the paper is focused on this question by addressing the computational analysis of C-clefs.

2. IMAGE PROCESSING OF BACH'S MANUSCRIPTS

The extraction of C-clefs from the manuscripts requires accurate segmentation. However, the segmentation of old handwritten manuscripts proves to be a difficult task [6, 17]. The main difficulty seems to be caused by degradation such as show-through and bleed-through effects.

[1] Enote history is a name of the project which mainly concerns scriber identification in handwritten music manuscripts from the 18th century. This was achieved by the cooperation of several research institutes: the library of the university of Rostock, the department of musicology at the university of Rostock, the database research group at the department of computer science, and the Fraunhofer institute for computer graphics.



(a) 1708 (b) 1723 (c) 1733 (d) 1736 (e) 1739 (f) 1748

Figure 1. The C-clefs of Bach's handwriting in the order of chronology suggested by musicologists.

In addition, microfiche, the primary medium of the Bach's manuscripts in the study, gives the images in low-resolution, which creates further problems for image processing.

(a) Clear manuscript

(b) Obtained image by adaptive binarization and staff-line removal from (a)

(c) Unclear and degraded manuscript

(d) Obtained image by adaptive binarization and staff-line removal from (c)

Figure 2. Two results of staff-line removal; almost all the staff-lines are left in (d).

As text line localization is the essential part of the OCR process [12], staff line detection is one of the most difficult but important aspects of OMR, since staff lines, which are used to give meaning to certain symbols such as noteheads, prevent the segmentation of musical symbols. Although there are arguments about the necessity of the staff-line removal, most research regards it as essential. The volume of research dealing with staff-line detection and removal [5, 7, 8, 15, 16, 19] indicates the difficulty inherent in this process, especially in the case of handwritten music.

(a) (b) (c) (d) (e)

Figure 3. C-clefs cropped by the proposed method and prepared for feature extraction: (a)original clef; (b)binarization using Niblack's method; (c)line removal using Dalitz's method; (d)(e)other examples including irrelevant pixels.

We experimented with several staff-line removal methods implemented in Gamera [2], and we found Dalitz's method [7] effective although it has sometimes failed to find staff-lines, probably because it is sensitive to deformation. This happened especially when the staff-lines were curved or significantly thinner than usual. Figure 2 shows typical results of the staff-line removal. In addition to the difficulty inherent in the staff-line removal from the manuscripts, Bach's dense notation and the irrelevant pixels, which are most commonly resulted from the degradation of paper, cause touching symbols. Moreover, unclear and degraded manuscripts are often fragmented by binarization process. These problems make it difficult to automatically decide bounding box of each musical symbol.

As it still requires further work to resolve these difficulties, for the present study we decided to collect C-clefs by manually deciding the bounding box. Figure 3 shows the C-clef extracted by this method. This extraction is followed by both morphological operation and staff-line removal to procure clear image, in order to prepare for the feature extraction. This is shown in Figure 3(b) and (c).

3. EXPERIMENT

This section explores the classification of the C-clefs. Because there is a controversy among Bach scholars regarding both the authorship and chronology of C-clef forms, we

[2] Gamera is a toolkit for building document image recognition systems and cross platform library for the Python programming language. It provides a set of commonly needed functionality for document image analysis and allows for custom extensions as C++ or Python plugins and as toolkits. See http://gamera.informatik.hsnr.de/index.html for more detail.

have carefully selected the sample dataset from an undisputed portion of Bach's fair copies that date between 1708 and 1748. The detailed information of this is shown in Table 1. [3] We prepared two classification tasks using the same dataset: one is eight-class classification using the date proposed by Kobayashi as the label; the other is two-class classification which only distinguishes between A B C and D E F G H. This corresponds to determining if a certain clef was written before Bach arrived at Leipzig (i.e. May 1723) to assume his role as Thomas cantor as well as the director of music for the town, or after that date.

Feature selection is also an important factor for successful classification. For the present study, 15 features implemented in Gamera were used. Each feature is explained as follows [4]

- area

  The area of the bounding box.

- aspect ratio

  The aspect ratio of the bounding box.

- black area

  The number of black pixels.

- compactness

  The volume to surface ratio.

- moments

  The centre of gravity on x and y axis normalized by width and height.

- ncols feature

  The number of columns.

- nholes

  The averaged number of white runs not touching the border. This is computed both for each row and each column.

- nholes extended

  Divides the image into four strips and then does a nholes analysis on each of those strips. This is first done vertically and then horizontally, resulting in a total of eight feature values.

- nrows feature

  The number of rows.

- skelton features

  Generates a number of features based on the skeleton of an image.

- top bottom

  The first feature is the first row containing a black pixel, and the second feature is the last row containing a black pixel.

- volume

  The percentage of black pixels within the rectangular bounding box of the image.

- volume16regions

  Divides the image into a 4 x 4 grid of 16 regions and calculates the volume within each.

- volume64regions

  Divides the image into a 8 x 8 grid of 64 regions and calculates the volume within each.

- zenrike moments

  Computes the absolute values of the normalized zernike moments [18] up to order six.

In the experiment, the performance of random forest (RF), which worked the best in the preliminary experiment, was investigated using 10-fold cross-validation compared with other methods: support vector machine (SVM), bagging, and boosting [5]. RBF kernel was used as the kernel function of SVM, and this was automatically estimated from the result of a preliminary experiment. CART Algorithm was used as underlying classifiers for all the ensemble classifiers and the other parameters were set as default.

Table 2 shows the result of 10-fold cross-validation. The best accuracy for two-class classification was 89.95% obtained by random forest. This accuracy seems significant considering that the classification of Bach's handwriting has been attempted by only a few experts. The eight-class classification is far more complicated and thus extremely difficult even for human experts. The best accuracy 73.82% was achieved by RF for eight-class classification, which is in itself remarkable. While there is much room for improvement, these classifications may serve as a rough barometer for musicologists.

Tables 3 and 4 indicate the confusion matrix for the two-class and the eight-class classification. Note that this confusion matrix is the result of the classification using out of bag data, and its error rate tends to be higher than that of cross validation. In Table 3 , misclassification of B(during and

[3] See [13] and [14] for detailed discussion on the chronological issue of J.S.Bach's work.

[4] See http://gamera.sourceforge.net/doc/html/features.html#features for detailed information of the features.

[5] See [1, 2, 11, 20] for detailed explanation of each classifier used in the experiment. We used the implementation included in R package for all the classifiers. See http://www.r-project.org/ for more information about R.

Table 1. Data set used for the experiment.

ID	Name of the piece	BWV	Name of the source	Estimated date	Number of the clefs extracted
A	Cantata “Gott ist mein König”	BWV71	D-B, Mus. Ms. Bach P 45	1708	89
B	Alles mit Gott und nichts ohn’ Ihn	BWV1127	D-W, Ra B 24	1713	11
C	Inventions and Sinfonias	BWV772–801	D-B, Mus.ms. Bach P 610	1723	188
D	Sanctus	BWV232/III	D-B, Mus. ms. Bach P 13	1724	77
E	Magnificat	BWV243	D-B, Mus.ms. Bach P 39	1733	221
F	St Matthew Passion	BWV244	D-B, Mus.ms. Bach P 25	1736	633
G	Well-Tempered Clavier II, No. 10, 19, and 24	BWV879, 888, and 893	GB-Lbl, Add.MS. 35021	1739	69
H	Canonic Variations on Vom Himmel hoch	BWV769	D-B, Mus.ms. Bach P 271	1748	22

Table 2. Classification accuracy evaluated by 10-fold cross validation for the best four classifiers

	Two-class	Eight-class
Random forest	89.95%	73.82%
SVM	85.25%	72.36%
Bagging	88.35%	73.09%
Boosting	86.20%	60.35%

Table 3. Confusion matrix for two-class classification

	A	B	class.error
A	189	99	0.34
B	33	989	0.032

after Leipzig) to A(before Leipzig) is limited . The misclassification of A to B seems to be caused by a small sample size of A compared to B. In Table 4, the misclassification of classes such as A D and H into F is noticeable. This result implies that classes with a low sample size tends to be classified as the class with a large sample size such as F. It is interesting that class G is classified with fairly high accuracy although it has small sample size. This is probably because the deviation of the shapes of the C-clefs in G is small enough to achieve the high classification accuracy even with small sample data.

4. CONCLUSION AND FUTURE WORK

In this study, we proposed a new method to work out the chronology of music manuscripts by classifying the shape of C-clefs. Applying a battery of techniques to Bach’s manuscripts revealed the limitation of the current image processing techniques. The method of classifying C-clefs using 15 features and RF produced a result of 89.95% accuracy in two-class classification and 73.82% in eight-class classification.

The automatic collection of musical symbols from Bach’s manuscripts proves to be a challenging task, but is worth investigating further. The accuracy of C-clef classification can be improved by investigating the implementation of musical knowledge, and this should be explored in collaboration with musicologists. In this study, we assumed that all the clefs from the same page were written in the same period. However, there is deviation in shape even in the clefs on the same page and sometimes they looked as if they were added subsequently or even possibly by a different hand. For this level of analysis, it requires more sophisticated image processing techniques that are capable of handling more subtle changes in each music symbol.

Chronological identification is not a straightforward task. In contrast to OMR system, musicologists would take a complex approach, combining it with chronological, compositional, and notational information, placing them against the historical background of the source such as the situation under which the initial copying and revisions took place, the diplomatic polices that might reveal the purpose for which the score was made, and so on, to verify the initial hypothesis. It is hoped that quantification and statistical analyses such as what demonstrated in this paper will be perfected in future research, and that they are adopted by future musicologists to discover many more exciting facts hidden deep in the beautiful manuscripts of Johann Sebastian Bach.

5. ACKNOWLEDGEMENTS

This work is financially supported by the Nakajima foundation.

6. REFERENCES

[1] Leo Breiman. Bagging predictors. *Machine Learning*, 24:123–140, 1996.

[2] Leo Breiman. Random forests. *Machine Learning*, 45:5–32, 2001.

[3] I. Bruder. Integrating knowledge components for writer identification in a digital archive of historical music scores. *Proceedings of the Fourth ACM/IEEE Joint Conference on Digital Libraries*, page 397, 2004.

Table 4. Confusion matrix for eight-class classification

	A	B	C	D	E	F	G	H	class.error
A	26	0	3	2	3	55	0	0	0.71
B	0	7	1	0	0	3	0	0	0.36
C	0	0	160	1	17	5	5	0	0.15
D	2	0	7	8	21	39	0	0	0.90
E	0	0	18	1	146	56	0	0	0.34
F	17	0	4	5	35	572	0	0	0.096
G	0	0	2	0	0	0	67	0	0.029
H	0	0	0	0	0	22	0	0	1.0

[4] M. Bulacu and L. Schomaker. Text-independent writer identification and verification using textural and allographic features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(4):701–717, 2007.

[5] J. S. Cardoso, A. Capela, A. Rebelo, and C. Guedes. A connected path approach for staff detection on a music score. *Proceedings of the 15th IEEE International Conference on Image Processing*, pages 1005–1008, 2008.

[6] N. P. Carter. Segmentation and preliminary recognition of madrigals notated in white mensural notation. *Machine Vision and Applications*, 5(3):223–229, 1992.

[7] C. Dalitz, M. Droettboom, B. Pranzas, and I. Fujinaga. A comparative study of staff removal algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(5):753–766, 2008.

[8] Santos Cardoso dos, A. Capela, A. Rebelo, C. Guedes, and da Costa Pinto. Staff detection with stable paths. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(6):1134–1139, 2009.

[9] Walter Emery. The London autograph of the forty-eight. *Music and Letters*, 34(2):106–123, 1953.

[10] A. Fornes, J. Llados, G. Sanchez, and H. Bunke. Writer identification in old handwritten music scores. *Proceedings of the Eighth IAPR International Workshop on Document Analysis Systems*, pages 347–353, 2008.

[11] Yoav Freund and Robert E. Schapire. Experiments with a new boosting algorithm. *Proc. of the Thirteenth International Conference on Machine Learning*, pages 148–165, 1996.

[12] M. Y. Jaisimha. Model based restoration of document images for OCR. *Proceedings of the SPIE - The International Society for Optical Engineering*, 2660:297–308, 1996.

[13] Yoshitake Kobayashi. Zur Chronologie der Spätwerke Johann Sebastian Bachs. Kompositions- und Aufführungstätigkeit von 1736 bis 1750. *Bach-Jahrbuch*, 74:7–72, 1988.

[14] Yoshitake Kobayashi. *Die Kopisten Johann Sebastian Bachs: Katalog und Dokumentation*. Kassel: Bärenreiter, 2007.

[15] H. Miyao. Stave extraction for printed music scores. *Intelligent Data Engineering and Automated Learning - IDEAL 2002. Third International Conference (Lecture Notes in Computer Science Vol.2412)*, pages 562–568, 2002.

[16] A. Rebelo, A. Capela, J. F. P. da Costa, C. Guedes, E. Carrapatoso, and J. S. Cardoso. A shortest path approach for staff line detection. *Proceedings of 2007.AXMEDIS '07.Third International Conference on Automated Production of Cross Media Content for Multi-Channel Distribution*, pages 79–85, 2007.

[17] Florence Rossant. Robust and adaptive omr system including fuzzy modeling, fusion of musical rules, and possible error detection. *EURASIP Journal on Applied Signal Processing*, 2007(1), 2007.

[18] T. Obeidi S. Belkasim, E. Hassan. Explicit invariance of cartesian zernike moments. *Pattern Recognition Letters*, 28:1969–1980, 2007.

[19] Mariusz Szwoch. A robust detector for distorted music staves. *Proceedings of Computer Analysis of Images and Patterns.11th International Conference, CAIP 2005.*, pages 701–708, 2005.

[20] Vladimir Naumovich Vapnik. *The Nature of Statistical Learning Theory*. New York: Springer, 1995.

[21] Georg von Dadelsen. *Beiträge zur Chronologie der Werke Johann Sebastian Bachs*. Tübinger Bach-Studien, volume 4-5. Trossingen: Hohner, 1958.