

Composer attribution by quantifying compositional strategies

Peter van Kranenburg

Department of Information and Computing Sciences

Utrecht University

petervk@cs.uu.nl

Abstract

Taking a theory of musical style developed by Leonard B. Meyer as a starting point, an experiment is described in which statistical pattern recognition algorithms are used to characterize a particular musical style with respect to other styles. The resulting description can be used in authorship discussions. In the current study, a number of disputed organ works from the Bach catalogue is used to illustrate the possibilities of this approach.

Keywords: Musical Style, Pattern Recognition, Classical Music, Composer Attribution, Johann Sebastian Bach.

1. Introduction

In order to describe a musical style or differences between styles or the historical development of certain styles, a theory of style is necessary. This applies to “traditional” descriptions of musical style as well as studies in which tools and algorithms from information technology are used.

In [5], Leonard Meyer develops a theory of musical style that can be used as starting point for studies in which statistical pattern recognition algorithms are used to study and compare musical styles. Meyer defines (musical) style as follows: *Style is a replication of patterning, whether in human behavior or in the artifacts produced by human behavior; that results from a series of choices made within some set of constraints.*

Without repeating patterns, there would be no style at all. The constraints are important for they shape a musical style by allowing certain patterns and disallowing others. Meyer distinguishes three levels in these constraints: Laws, rules and strategies. Laws are universal constraints, e.g., One cannot ask a piccolo to play a contra G. The second level, the *rules* are intracultural constraints. It is in the rules that music from the Renaissance differs from music from the Baroque. The third level, the *strategies* are constraints the composer subjects himself to, within the rules of a certain cultural established style. Thus it is in the strategies that the music of G.F. Handel differs from the music of G.Ph. Telemann. Strategies reside on conscious as well as

on unconscious levels. Certain patterns are ingrained during the training and development of a composer and are not replicated consciously.

In the second part of his book, Meyer applies his theory to nineteenth century western classical music. He addresses some general patterns that recur in many compositions from that age and connects these patterns to the underlying romantic esthetic and ideology. In doing so, he is forced to limit himself to proof by example. For a more profound evaluation of musical styles, it would be necessary to make extensively use of all available data (i.c. everything in all considered scores). For achieving this, statistical pattern recognition algorithms can be of great use. As Meyer himself states: “Since all classification and all generalization about stylistic traits are based on some estimate of relative frequency, statistics are inescapable.” ([5], p. 64).

2. A Pattern Recognition Approach

Meyer’s theory offers a foundation for the design of experiments in which algorithms from statistical pattern recognition are used. The features that will represent (parts of) compositions can be allied with the replicated patterns that are mentioned in Meyer’s definition. Assuming that for a certain musicological problem the scores involved are electronically available, a major task will be the extraction of the feature values from those scores. From the perspective of “traditional” style analysis, large-scale features are more interesting than small-scale features, e.g., in order to determine the way in which a certain composition resembles a sonata-form, a global overview of the entire composition is necessary. From the perspective of algorithmic extraction, small-scale features are more interesting, because the algorithms to extract them are less complicated and the results less ambiguous. It is, for example, not clear how to quantify the extent to which a composition resembles a certain sonata-form, but it is much less difficult to determine the proportion of parallel thirds with respect to all interval successions in the composition. So we need small-scale patterns, which can be easily detected and counted, and of which we have many.

With the limitation of the previous section in mind, a set of twenty features is designed. The smallest scale in a score is that of the relation of a single note to the other notes around it. When a single note is part of a voice in a polyphonic composition, it is more independent than when

it is part of a chord. Because of this, most features quantify aspects of the relations between the different voices, which means that only polyphonic compositions can be represented with the designed feature set. Since we will use this representation for studying authorship of organ fugues, this is not a problem. There are also some other features in the set, that describe more global characteristics. The features are described in [1]. Here a list of them is provided:

- | | | | |
|--------------------|-----------------|--------------------|--------------------|
| 1. StabTimeslice | 6. PitchEntropy | 11. PartAugFourths | 16. PartOctaves |
| 2. DissPart | 7. VoiceDensity | 12. PartDimFifths | 17. ParThirds |
| 3. BeginBarDiss | 8. PartSeconds | 13. PartFifths | 18. ParFourth |
| 4. SonorityEntropy | 9. PartThirds | 14. PartSixths | 19. ParSixths |
| 5. HarmonyEntropy | 10. PartFourth | 15. PartSevenths | 20. StepSuspension |

By measuring all these features, compositions are represented as vectors in a 20-dimensional space. To such a data set various kinds of pattern recognition algorithms can be applied.

3. Organ Fugues ascribed to J.S. Bach

As a pilot experiment, a data set is assembled with 16 fugues for organ that are listed in the catalogue of compositions of Johann Sebastian Bach ([7]). Of six of these fugues the authorship has been disputed. Also five fugues of his eldest son, Wilhelm Friedemann Bach, and eight of his most important student, Johann Ludwig Krebs, are incorporated. So we have a three-class data set.¹ Each composition is segmented using a segmenting method described in [1], so each composition is represented by a "cloud" of points.

The Fisher-transformation (described in [8], p. 145ff) can be used to project the data points onto a two-dimensional space in such a way that the classes are optimal separated (figure 1). This projection shows that the compositions of each composer do form a cluster. So it can serve as a reference for classifying compositions which might be composed by be one of the three involved composers.

Figure 1 indicates where the data points of the disputed fugues are projectet. Some interesting observations can be made. The F minor fugue BWV 534, is projected among the fugues of J.L. Krebs. This fugue has been ascribed to W.F. Bach ([3]). With the current result, that ascription can be rejected. An ascription to J.L. Krebs seems more likely. A suggested composer for BWV 536 is J.P. Kellner ([4]). If this is true, Kellners style resembles more the style of J.S. Bach than that of the other two composers. BWV 537 is said to be composed partly by J.S. Bach (bar 1–40) and partly by J.L. Krebs ([6]). The first part is projected among the works of J.S. Bach indeed. The second part however, is outside of both the Bach-region and the Krebs-region. The ending of the fugue is in the region between J.S. Bach and Krebs. This does not fully support the hypothesis, but it shows that a large part of the fugue is not Bach-like. Also Bach's authorship of the fugue in C minor, BWV 546, has

¹ The dataset is available from: <http://www.musical-style-recognition.org>.

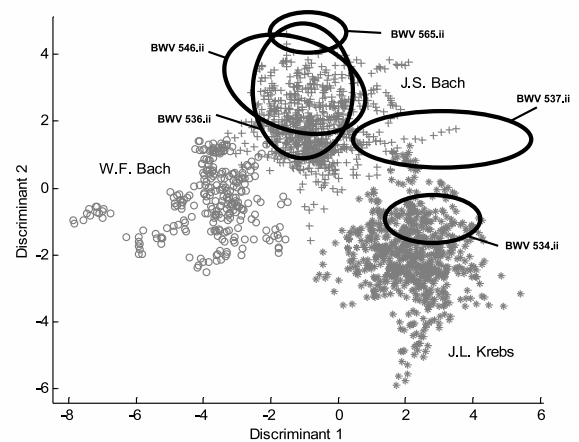


Figure 1. Projection of disputed fugues on top of the the compositions of J.S. Bach (+), W.F. Bach (o) and J.L. Krebs (*).

been doubted ([2]). The current evaluation shows us that, with respect to the styles of W.F. Bach and J.L. Krebs, this fugue has the characteristics of the style of J.S. Bach. The most famous organ work in existence, the toccata and fugue in D minor, BWV 565, is not projected among the other compositions of Bach. This confirms the doubts expressed in [9].

4. Conclusion

Although the current results don't offer enough evidence to draw conclusions about the authorship of the involved compositions, it is clear that the proposed method is very helpful in finding hypotheses about differences in personal styles and thus for studying authorship problems.

References

- [1] E. Backer and P. van Kranenburg, "On musical stylometry — a pattern recognition approach", in *Pattern Recognition Letters*, 26 (2005), 299–309.
- [2] W. Breig, "Versuch einer Theorie der Bachschen Orgelfuge", in *Die Musikforschung* 48 (1995), 14–52.
- [3] P. Dirksen, "Het auteurschap van Praeludium en fuga in f (BWV 534)", in *Het Orgel* 96 (2000), nr. 5, 5–14.
- [4] D. Humphreys, "A Bach Polyglot—the A major Prelude & Fugue BWV 536", in *The Organ Yearbook XX* (1989), 72–87.
- [5] L.B. Meyer, *Style and Music—Theory, History, and Ideology*, Chicago, 1989.
- [6] J. O'Donnell, "Mattheson, Krebs and the Fantasia & Fugue in C minor BWV 537", in *The Organ Yearbook XX* (1989), 88–95.
- [7] W. Schmieder, *Thematisch-systematisches Verzeichnis der musikalischen Werke von Johann Sebastian Bach. Bach-Werke-Verzeichnis—2. überarbeitete und erweiterte Ausgabe*, Wiesbaden, 2¹⁹⁹⁰.
- [8] A. Webb, *Statistical Pattern Recognition*, Chichester, 2²⁰⁰².
- [9] P. Williams, "BWV 565: a toccata in D minor for organ by J. S. Bach?", *Early Music* 9 (1981), 330–337.