Towards Automatic Classification of Sheet Music

Giuseppe De Pasquale[1], Blerina Spahiu[2][0000−0002−6958−8215], Pietro Ducange[3][0000−0003−4510−1350], and Andrea Maurino[2][0000−0001−9803−3668]

[1]Leonardo S.p.A. Sistemi di Difesa, via Monterusciello, 75, 80078, Pozzuoli, Napoli
`giuseppe.depasquale@leonardocompany.com`
[2]University of Milano-Bicocca, Viale Sarca, 336 - 20126 Milan, Italy
`blerina.spahiu|andrea.maurino@unimib.it`
[3]University of Pisa, Largo Lucio Lazzarino, Pisa, Italy
`pietro.ducange@unipi.it`

Abstract. Automatic music classification has been of interest since digital data about music became available within the Web. For this task, different automatic classification approaches have been proposed but all existing approaches are based on the analysis of sounds. To the best of our knowledge, there is no automatic solution that considers only the sheet music for classification. Therefore, within the following study, we introduce a machine-learning based approach in order to assign an author to new sheet music. Different features, that best represent the style of a writer has been extracted, and are given in input for training to a kNN algorithm. In addition, the article discusses the results and cases when the classifier fails to assign the right author.

Keywords: sheet music classification · machine learning · feature extraction.

1 Introduction

Sometimes, for a trained ear, it can be easy to identify the author of an unheard music played on the piano. Often, this happens when someone tries to identifying a famous author, such as Mozart, Beethoven or Chopin, of whom many concerts, sonatas or other compositions have already been heard before. When an author has his own specific style, it is also possible to recognize him even though we are listening to a song never heard before. Of course, there are authors who have similar styles, especially those belonging to the same school as the German or Austrian school, the Russian school, the Italian school, the French school, etc. This paper aims at characterising the piano writing of three authors, namely Chopin, Beethoven and Mozart. There exists several similarities between these authors, since the German Beethoven was certainly influenced by the music of the Austrian composer Mozart. Beethoven moved to Vienna at 22 years old and trained together with Mozart and Haydn.

The scope of this paper is to investigate if the writing of Chopin, Beethoven and Mozart, their way of composing, and part of their scores can be used to classify new ones using machine learning techniques. There exist a bunch of works in classifying authors and genres, such as those discussed in [8, 13, 11, 17, 7, 9, 20], which have the aim of classifying music, classical music in particular, through the recording of a few minutes of sound, their spectral analysis and the extraction of suitable features. Such approaches are at an advanced stage in terms of sampling capacity, ability to make Fourier transforms with many points, computational skills and advanced algorithmic techniques. Moreover, such approaches can count on increasingly powerful processors with larger and larger memories capacity. The goal of these approaches is to obtain features that can characterize different song writers so that it can be easy to classify new musical scores that do not have an author. Such features convey the sequence of notes and chords, their duration and intensity, and breaks in time or frequency.

With sophisticated means it is possible to create faithfully scores starting by a piece of sound played in piano. The score of a song has all this information, thus allowing us to focus directly on the extraction of the features and their classification, leaving out the decoding phase of the sound of the musical piece.

In summary, the aim of this paper is to fill the gap and present an initial work toward an automatic classification of musical scores for classical music writers. Our contributions are two fold and can be summarized as follows: (i) to identify a set of features that best represent the style of a music composer; and (ii) to extract such features used for classification.

This paper is organized as follows: preliminaries in understanding the main concepts and definitions in music scores are given in Section 2 while the state-of-the-art is presented in Section 3. A general overview of the workflow and the analysis of the main features that better characterize the style of a composer is given in Section 4. The results of the classification approach and their analysis are presented in Section 5 while conclusions end the paper in Section 6.

2 Preliminaries

In this section we first introduce some preliminary definitions and concepts needed to understand and describe our approach.

Musical notation or semiography: is the system of signs (staff, notes, keys, etc.) used to express music graphically.

Staff: is a bundle of five horizontal and equidistant straight lines, used in musical writing to indicate, in conjunction with the clef, the height of each note.

Clef: The clef is a graphic sign that is placed at the beginning of the musical staff to indicate the position of the note to which it corresponds, thus determining the name, position and pitch of all the other marked notes of the same staff. It can also be marked inside the body itself, when the height of the notes requires a change. In a staff two types of clefs are used: Treble clef (also called G Clef because of the staff line which the clef wraps around (see first staff in Fig. 2))

and Bass clef (also called F clef because of the staff line between two dots of the clef is F (see second staff in Fig. 2)).

Notes - Rests: Musical notes mean the series of seven types of sounds: C, D, E, F, G, A and B. Graphically the notes are distinguished from the place they occupy in the staff according to the reading clef. Octaves is the distance between one note (like C#) and the next note bearing its same name (the next C# that's either higher or lower). In terms of physics, an octave is the distance between one note and another note that's double its frequency.

Beats - Measure bar - Tempo: The notes of a musical composition are divided, according to a metricorhythmic scheme, called tempo, into many distinct and consecutive groups, called *beats*, by means of vertical separation lines, called measurement bars. The bar is the set of musical values (figure and accents) between two measure bars. Tempo is the metric-rhythmic pattern of the bars, intended as a unit of duration and accentuation.

Tonality: With tonality we refer to that particular constraint that elevates the sounds of one scale, according to the place it occupies in the order of the scale itself. **Scale:** is a gradual succession of sounds contained within an octave. Scales can be distinguished by the number of sounds and, above all, by their reciprocal distance relationships. The scale is divided into: diatonic (made up of tones and semitones like the succession of the seven natural sounds) and chromatic (consists of the succession of all 12 semitones contained in the octave).

Interval: The interval is the tonal distance that passes between two sounds. Intervals can be distinguished in harmonic (those from simultaneous sounds) and melodic (those with subsequent sounds).

Melody: Melody is any rhythmic sequence of single notes, capable of expressing a musical thought.

Harmony: is the science of chords. A chord is the union of different sounds, merged into a single acoustic expression. Chords are formed, basically, by superimposing two or more sounds of the same scale.

3   Related Work

To the best of our knowledge, there is no previous work that deals with extracting features from staff for the classification of musical artists. Related published work on features extraction can be grouped into three categories: (i) through audio signals; (ii) through optical music recognition; and (iii) through musical space representation.

3.1   Features extraction through audio signals

The idea behind [19] is that a particular musical genre is characterized by statistical properties related to the instrumentation, the rhythm of its content. For this reason they propose the extraction of a set of features to represent the structure and musical instrumentation, as well as a new set of features to represent the rhythmic structure. The performance of these feature sets was assessed

through the training of classifiers for the recognition of statistical models that use real-world audio collections.

Authors in [6] propose an approach for the classification of music into three major categories: rock, classical and jazz through the use of spectrograms. Humans are much better at distinguishing small pitch changes at low frequencies than at high frequencies. This scale makes the features more similar to what humans perceive. Such algorithm use the texture-of-texture models to generate feature vectors out of spectrograms. These features are capable to profile the frequency power of a sound as music progresses. Finally, an attempt is made to classify the data generated using a variety of classifiers.

The work discussed in [11] proposes a two-step methodology to classify music. First, various sources of musical instruments are identified and separated in the audio signal. Then, features are extracted from the separated signals that correspond to distinct musical instruments. Afterwards, the extracted timbre, rhythm and intonation features from the identified instruments are used to classify music clip. This procedure is similar to a human listener who is able to determine the genre of a musical signal and, at the same time, distinguish a number of different musical instruments in a complex sound music.

The goal of the contribution in [18] is to identify the most likely pianist, given a set of performances of the same piece performed by different pianists. For the classification, a set of simple features is proposed which characteristics best a musical performer: timing (variations in tempo), dynamics (variations in loudness), and articulation (the use of overlaps and pauses between successive notes). Experiments show that by using machine learning it is possible for a machine to distinguish music performers (pianists) on the basis of their performance style.

The melody extraction algorithms proposed in [16, 4] aim to produce a sequence of frequency values corresponding to the tones of the dominant melody from a musical recording. The term melody is intended as the single (monophonic) sequence of tones that a listener would reproduce if asked to whistle or sing a piece of polyphonic music. The melody extraction algorithms are commonly evaluated by humans by comparing the estimated intonation sequences with respect to the real one. For the evaluation, pieces of classical music, recordings of people singing along with the music were collected. Experiments show that there are several difficulties in extracting the melody for this particular repertoire. However, the density and the complexity of the notes were identified as the most relevant features for the classification.

3.2   Feature extraction through optical music recognition

For the classification of music genres, the work discussed in [5] proposes an approach that converts an audio signal into spectrograms. Features are then extracted from the time-frequency images which are used to model musical genres in a classification system. The features are based on a local binary pattern, a structural operator that has been successful in recent research on image classification. The experiments were performed with two well-known data sets: the Latin music database (LMD) and ISMIR 2004. The results obtained by analysing

the extracted features always exceed those based on audio content only. Using the LMD data set, the accuracy of the approach is approximately 82.33%. while for the ISMIR 2004 database, the best result obtained is around 80.65%.

Authors of [9] propose an automatic system for classifying musical genres based on a strategy for selecting local features using the self-adaptive algorithm Armony Search (SAHS). Five acoustic characteristics (i.e. intensity, pitch, timbre, tonality and rhythm) are extracted. Finally, each SVM (one-to-one vector machine) classifier is fed with the corresponding local feature set and the majority voting method is used to classify each music recording.

The contribution in [14] presents a new and effective approach for the automatic recognition of the musical genre based on the fusion of different sets of features. Both acoustic and visual features are considered and merged into a final set. Evaluations show that such approach achieves an accuracy for the classification comparable or even better than other approaches of the state-of-the-art. The aim of this approach is to classify music from its representation as spectrogram and proposes a set of descriptors and classifiers to maximize the performance that could be obtained from the visual characteristics.

3.3 Feature extraction through musical space representation

This category aims to perform coherent analysis of the structure of the musical piece by representing notes in the plane that reveals affinities and structures between notes [3]. Tonnez is the classical tool that represent harmonic relationships between chords using a two-dimensional lattice [2].

4   The Proposed Approach

In the following, we describe the workflow of the proposed approach. First, we describe the general workflow starting from the input file and every step until the classification algorithm. Second, we describe in details the feature extraction process which we apply for our task in assigning the author of a composition. Third, we describe the classification algorithm.

4.1 Workflow

In Fig. 1 is given the general approach for music classification starting from music sheets. As a requirement for this research, we propose a new approach written ad hoc to extract features from scores. Since optical reading and interpretation of the scores are not subject of this research, we use piano pieces by the three authors downloaded from the Web in $midi$[1] format (step 1). Subsequently, using the MuseScore[2] notation and musical composition software, the downloaded files can be transformed into *.xml* format (step 2). Such format obtains a representation on the staff in a faithfully way for various compositions. It should be noted

---

[1] http://www.music.mcgill.ca/ ich/classes/mumt306/StandardMIDIfileformat.html
[2] https://musescore.org/

that the staves obtained in this way are not exactly the same as the originals, but faithfully represent the succession of notes and rests of the original compositions, keeping their duration and speed. Such transformation conserves melody, harmony, time and rhythm of a piece, which are the real content of the piece and all characteristics that distinguish one composition from the other. Next, we extract a set of features that could best represent a compositions and that could better allow to identify different music writers (step 4). This step was supported by an in-depth research of similar studies, of data sets already available and of algorithms already used for the classification of songs or musical styles. Actually, in this preliminary work, we divided each staff into *periods* composed of 8 measures and for each of them we extract a set of features. Finally, we classify each period (step 5), assigning to it a specific author (step 6).
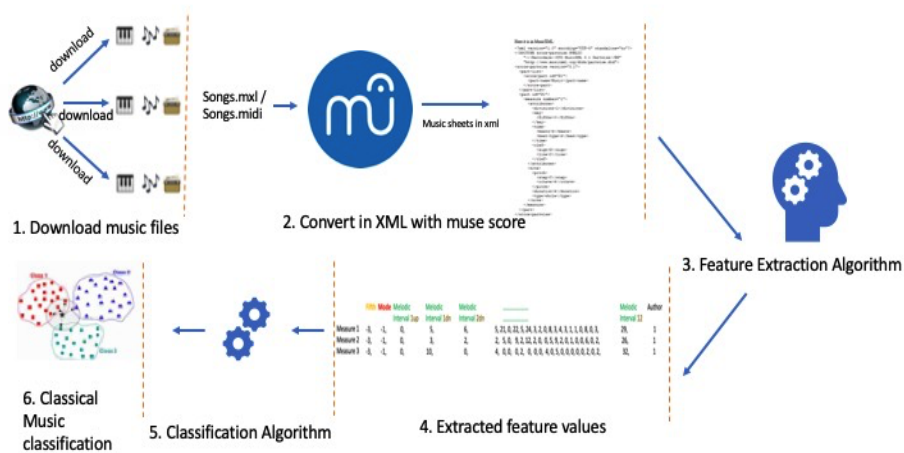

**Fig. 1.** General workflow for the approach

4.2 Feature Extraction

The task of identifying which are the best features, from musical sheets, that can distinguish one author from another is very difficult. A piano composition is formed by the succession of single notes and chords of different duration performed at different speeds and articulated through the synchronization of the right hand and left hand of the performer. The characteristic of a composition is the combination of melody and harmony that express the sensitivity of the authors and their inspiration supported by an excellent technique. Taking into consideration the staff of a composition, we identified several features as important for characterizing a music composition. As stated before, in this preliminary work, we divided each staff into *periods* composed of 8 measures. For each period, we extract a collection of features grouped into 4 macro-features, as shown in Fig. 2 and discussed in the following:

- **Notes:** Musical notes means the series of 7 types of sounds: C,D, E, F, G, A and B. For each note, we extract its occurrence within the period. Moreover,

we extract also the occurrence of notes with accidentals (flat ♭ and sharp ♯). The number of equal notes used by the composer can be a useful feature to characterize a composition. In this study all equal notes, even belonging to different octaves are considered as the same note. Moreover, the occurrence of each note is calculated without considering its value, namely its relative duration. Thus, in total, we can extract 12 features regarding the notes macro-features.

– **Rests:** Rests represent period of silence in a music composition. For each period, we count the number of rests, without considering its relative duration.

– **Tonality:** this macro-feature includes both the *fifth* and the *mode* of the stuff. The fifth is a set of flats ♭ or sharps ♯ that are generally written immediately after the clef. It is described by an integer feature, defined in $[-7, 7]$, where the negative values regard flats and positive values regard sharps. The modulus represents the number of symbols. Given a certain number of symbols, two specific tonalities, also called keys, may be associated to the composition. Also the value zero can describe two possible keys, namely C major or A minor. Mode can have two values $+ 1$ for major mode and $- 1$ for minor mode.Indeed, we recall that the *circle of fifths* is a graphic representation of the possible tonalities where the outer circle describes major mode while the inner circle describes the minor mode. It is known that generally a minor mode is more suitable for sad and melancholic music than a major mode that is more suitable for a cheerful and joyful work.

– **Interval:** An interval represents the distance in semitones between two notes. There exists two types of intervals: *Melodic* intervals (from subsequent notes) and *Harmonic* intervals (from simultaneous notes which form a chord). Regarding melodic intervals, passing from one note to another, we may identify ascending or descending intervals, if we have an increasing or a decreasing of the frequency of the note, respectively. In order to reduce the number of intervals, distances are calculated modulo 12. Since the maximum number of semitones in an octave is 12, we can identify in total 23 intervals for the melodic intervals. Indeed, we consider 11 ascending and 11 descending intervals and a neutral interval, considering that the 12th ascending interval coincides with the 12th descending interval. The melodic interval between two chords is calculated between the two highest notes of the two chords. For each interval, we calculate its frequency in the period. In the case of harmonic intervals, we calculate the distances in semitones between two notes, starting from the lowest note, thus we can extract 12 intervals. We recall that for each chord, we calculate the distances between all the possible combination of two notes that form the chord. In conclusion, we can extract 23 and 12 features for the melodic and harmonic intervals, respectively.

In the following, we show an example of how the values for a set of macro-features such as *Tonality & Melodic Intervals* are extracted for Chopin's prelude n. 20 op. 28. Each staff is divided into phrases composed of 8 measures. Each measure is represented with a row of values as in Fig. 3. The prelude under

Fig. 2. Features considered for extraction from staff
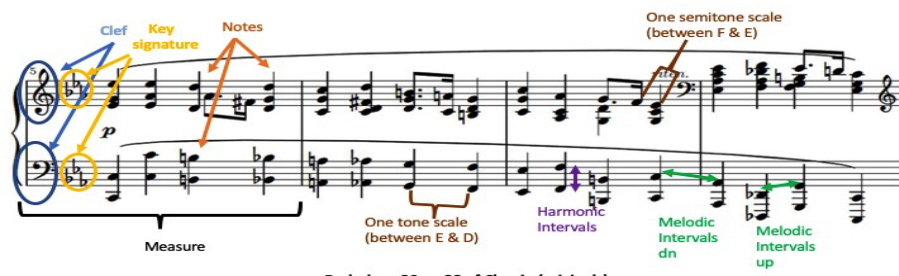
	Clef	Key Sig.	Melodic Interval 1up	Melodic Interval 1dn	Melodic Interval 2dn	....	Melodic Interval 12	Author
Measure 1	-3,	-1,	0,	5,	6,	5, 21, 0, 22, 5, 24, 3, 2, 0, 8, 3, 4, 3, 1, 1, 0, 8, 0, 3,	29,	1
Measure 2	-3,	-1,	0,	3,	2,	2, 5, 0, 9, 2,12, 2, 0, 0, 5, 9, 2, 0, 1, 0, 0, 6, 0, 2,	26,	1
Measure 3	-3,	-1,	0,	10,	0,	4, 0, 0, 0, 2, 0, 0, 0, 4, 0, 5, 0, 0, 0, 0, 0, 2, 0, 2,	32,	1

examination is a prelude in C minor composed of 26 measures (13 for each hand) consisting of 13 measures in the G key (right hand) and 13 in the F key (left hand). Since we have divided the staff into periods of 8 measures, we get $|26/8|$ periods, i.e. 3 instances to be classified. As we are considering periods of 8 measures, the remaining two measures are not considered for the classification. In Fig. 2, we show just an extract of the prelude, containing just a portion of two periods, for the left and for the right hands, respectively. The first feature that we extract is the *Fifth*, which is the same for each period. In this case, the value of this feature is equal to -3, namely 3 flats in key, that represent the key Eb major or C minor. The value of the mode allows us to discriminate between the two possible keys. In this case, the vale of the feature *mode* is equal to -1, thus the key is C minor. Finally, for each of the 23 melodic intervals, we extract a value which represents its occurrence in the period. In the example, for the first period, we find 0 intervals where notes have ascended by one semitone ($MelodicInterval_{1up}$), while there are 5 intervals where notes have descended by 1 seminote ($MelodicInterval_{1dn}$).

	Fifth	Mode	Melodic Interval 1up	Melodic Interval 1dn	Melodic Interval 2dn	....	Melodic Interval 12	Author
Period 1	-3,	-1,	0,	5,	6,	5, 21, 0, 22, 5, 24, 3, 2, 0, 8, 3, 4, 3, 1, 1, 0, 8, 0, 3,	29,	1
Period 2	-3,	-1,	0,	3,	2,	2, 5, 0, 9, 2, 12, 2, 0, 0, 5, 9, 2, 0, 1, 0, 0, 6, 0, 2,	26,	1
Period 3	-3,	-1,	0,	10,	0,	4, 0, 0, 2, 0, 0, 0, 4, 0, 5, 0, 0, 0, 0, 0, 2, 0, 2,	32,	1

Fig. 3. Feature values for Tonality & Melodic Intervals for the Prelude n. 20 op 28 of Chopin (Fig. 2)

4.3 Classification approaches

Once we have extracted a set of features for describing a period of a staff, we can classify it in order to identify its author. The classification problem has been widely studied in the database [12], data mining [15], and information

retrieval communities [10]. Typically, a classification model is in charge of assign a class, among a pre-defined set of classes, to a new input pattern described as a vector in a specific multi-dimensional space. The parameters of the classifier can be learnt during a supervised learning stage adopting a labeled training set. As a preliminary experiment, we choose to test the simple kNN classification algorithm. In this case, the parameters of the model are all the instances of the training set. Given a new pattern, the kNN classifier finds the nearest k patterns (nearest neighbors) to the one to be classified [1]. The classifier then assigns to the new pattern the most frequent class, among the k nearest neighbors.

5 Experimental Analysis

In this section we describe the results of the experiments carried out for the classification of new musical sheets. Firstly, we provide a general overview of the data used for the experiment. Secondly, we describe the results of the experiment for different feature combination and finally we provide a discussion of the results.

5.1 Data Corpus

In the following experiments, we considered 131 music sheets of Chopin (61), Beethoven (29) and Mozart (41). All music sheets are converted into xml format (see Fig. 1) and after the feature extraction process, as discussed in Section 4, we obtained a complete data corpus consisting of 6063 instances (corresponding to 6063 *periods*), including 1919 from Chopin, 2010 from Beethoven and 2134 from Mozart. Actually, in this preliminary work, we classify each single instance rather than the entire music sheet.

5.2 Experimental results

In this section, we report the results for the experiments for classical music classification between three authors. In order to identify which are the features that mostly characterize a composer, we carried out a number of experiments considering different subsets of features, among the ones described in Section 4. For each experiment, we carried out a 10 fold cross-validation. We recall that in this preliminary experiments, we classify single periods extracted from music sheets, assigning to each of them a specific author.

We considered two categories of experiments. For the first category, we build models to classify periods into one of the three authors, namely Chopin, Mozart and Beethoven. As regards the second category, we consider just two authors for the classification task, namely Mozart and Beethoven. We skipped Chopin in the second category since Mozart and Beethoven share several melodic and harmonic aspects. Indeed, they acted in an overlapped period during the last years of the 18th Century. Thus, distinguishing between these two authors may be more challenging.

Table 1 shows the results of the classification task for five experiments, for each category, adopting the kNN as classification algorithm (k=5, using Euclidean distance). In the first column, we show the different combinations of macro-features that characterized a specific experiment. The second and third columns regard, respectively, the first (Complete Data Corpus) and the second category (Partial Data Corpus) of experiments. These columns summarize the accuracy on the test set, expressed in terms of the average values of the percentage of correctly classified periods. As regards the combination of macro-features, we experimentally verified that the Tonality should be always included for achieving meaningful results. For the sake of brevity, we skipped to show the results achieved without considering this macro-feature.

Table 1. Period classification accuracy for different macro-feature combinations

Combination of Macro-features	Accuracy (%) Complete Data Corpus	Accuracy (%) Partial Data Corpus
Tonality & Harmonic Intervals	75.12	79.55
Tonality & Melodic Intervals	**82.96**	**83.78**
Tonality & Notes	77.86	80.82
Tonality & Intervals	**84.18**	**84.67**
Tonality & Notes & Rests	81.25	83.11

5.3 Results discussion

We can observe that the highest accuracy is achieved if the melodic intervals are included among the features describing each phrase. Indeed, in these cases, the classifiers achieves an accuracy up to 84.18% and 84.67%, for the complete and the partial data corpus, respectively. Good accuracy levels are achieved if Tonality, Notes and Rests are considered as macro-features for classification. These results confirms that the melodic aspect is what helps to better discriminate a composition author.

In Table 2, we show the average F-measure calculated for each class, namely for each author. This indicator allows us to understand how good is the recognition of each single author by the classifier: the higher the value, the better the recognition. As expected, considering the complete data corpus, since the style of Chopin is very different from that of Beethoven and Mozart, his periods are always better recognized than the other two authors. As regards Mozart,

Table 2. Average F-measure for each author

	Complete Data Corpus			Partial Data Corpus	
Combination of Macro-features	Chopin	Beethoven	Mozart	Beethoven	Mozart
Tonality & Harmonic Intervals	0.81	0.66	0.78	0.77	0.81
Tonality & Melodic Intervals	0.87	0.78	0.83	0.83	0.84
Tonality & Notes	0.84	0.71	0.78	0.80	0.81
Tonality & Intervals	0.89	0.79	0.83	0.84	0.85
Tonality & Notes	0.87	0.75	0.80	0.82	0.83

we can see that he is always better recognized than Beethoven. Finally, as expected, considering the partial data corpus, the recognition capability of both Beethoven and Mozart increases, and the gap between the recognition between this two authors decreases.

6   Conclusions

This study has been inspired by the following statement: in some cases it is possible, for a trained ear, to recognize an author by listening to one of his/her composition that the user has never listened before. The idea is that, the information that characterizes an author, is all contained in the scores of his compositions and it is not necessary to listen to their songs to recognize them.

In this work, we have discussed a framework for the automatic classification of musical sheets. Specifically, we have adopted as inputs midi files of classical music compositions of three famous composers, namely Chopin, Mozart and Beethoven. Then, adopting an open source program, we have transformed the midi files into XML files which codify the music sheets. Each music sheet has been divided into periods of 8 measures and from each of them we have extracted a set of macro-features. Finally, we have experimented different combinations of macro-features for classifying the periods into three classes, namely the three authors, using the kNN algorithm. Preliminary results have shown that good classification accuracy can be achieved if the features which describe the tonality and the harmonic and melodic intervals are adopted.

The proposed framework may be used for recognizing the author of a musical piece, even in digital format or in forms of musical sheets, found after years for which the composer is unknown.

As future works, we plan to extract other features and to experiment additional classifiers that might help us to improve the classification accuracy. Furthermore, we plan to extend our data corpus by including more authors and more musical scores.

Acknowledgments

References

1. S. D. Bay.   Combining Nearest Neighbor Classifiers Through Multiple Feature Subsets. In *Proceedings of the Fifteenth International Conference on Machine Learning (ICML)*, volume 98, pages 37–45. Morgan Kaufmann Publishers Inc., 1998. Citeseer, doi=10.1.1.52.455.
2. T. Bergstrom, K. Karahalios, and J. C. Hart. Isochords: visualizing structure in music. In *Proceedings of Graphics Interface 2007*, pages 297–304, 2007.

3. L. Bigo, M. Giraud, R. Groult, N. Guiomard-Kagan, and F. Levé. Sketching sonata form structure in selected classical string quartets. 2017.

4. J. Bosch and E. Gómez. Melody extraction in symphonic classical music: a comparative study of mutual agreement between humans and algorithms. In *Proc. 9th Conference on Interdisciplinary Musicology–CIM14, Berlin, Germany*, 2014.

5. Y. M. Costa, L. Oliveira, A. L. Koerich, F. Gouyon, and J. Martins. Music genre classification using lbp textural features. *Signal Processing*, 92(11):2723–2737, 2012.

6. H. Deshpande, R. Singh, and U. Nam. Classification of music signals in the visual domain. In *Proceedings of the COST-G6 Conference on Digital Audio Effects*, pages 1–4, 2001.

7. M. Genussov and I. Cohen. Musical genre classification of audio signals using geometric methods. In *2010 18th European Signal Processing Conference*, pages 497–501. IEEE, 2010.

8. T. George, E. Georg, and C. Perry. Automatic musical genre classification of audio signals. In *Proceedings of the 2nd international symposium on music information retrieval, Indiana*, 2001.

9. Y.-F. Huang, S.-M. Lin, H.-Y. Wu, and Y.-S. Li. Music genre classification based on local feature selection using a self-adaptive harmony search algorithm. *Data & Knowledge Engineering*, 92:60–76, 2014.

10. T. Joachims. *Learning to classify text using support vector machines: Methods, theory and algorithms*, volume 186. 2002. Kluwer Academic Publishers Norwell, 10.1162/089120103322753374.

11. A. Lampropoulos, P. Lampropoulou, and G. Tsihrintzis. Musical genre classification of audio data using source separation techniques. In *Proc. IEEE 5th EURASIP Conference on Speech and Image Processing, Multimedia Communications and Services*. Citeseer, 2005.

12. T.-S. Lim, W.-Y. Loh, and Y.-S. Shih. A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms. *Machine learning*, 40(3):203–228, 2000. Kluwer Academic Publishers, https://doi.org/10.1023/A:1007608224229.

13. M. McKinney and J. Breebaart. Features for audio and music classification. 2003.

14. L. Nanni, Y. M. Costa, A. Lumini, M. Y. Kim, and S. R. Baek. Combining visual and acoustic features for music genre classification. *Expert Systems with Applications*, 45:108–117, 2016.

15. T. N. Phyu. Survey of classification techniques in data mining. In *Proceedings of the International MultiConference of Engineers and Computer Scientists*, volume 1, pages 18–20, 2009. Citeseer, doi=10.1.1.148.6474.

16. J. Salamon, E. Gómez, D. P. Ellis, and G. Richard. Melody extraction from polyphonic music signals: Approaches, applications, and challenges. *IEEE Signal Processing Magazine*, 31(2):118–134, 2014.

17. J. G. sArnal Barbedo and A. Lopes. Automatic genre classification of musical signals. *EURASIP Journal on Advances in Signal Processing*, 2007(1):064960, 2006.

18. E. Stamatatos and G. Widmer. Automatic identification of music performers with learning ensembles. *Artificial Intelligence*, 165(1):37–56, 2005.

19. G. Tzanetakis and P. Cook. Musical genre classification of audio signals. *IEEE Transactions on speech and audio processing*, 10(5):293–302, 2002.

20. C. Weiß, M. Mauch, S. Dixon, and M. Müller. Investigating style evolution of western classical music: A computational approach. *Musicae Scientiae*, 23(4):486–507, 2019.