Revised Selected Papers

Accademia Musicale Studio Musica
Michele Della Ventura, *editor*

2020

Proceedings of the
International Conference on

# New Music Concepts Inspired Education and New Computer Science Generation

Vol. 7

STUDIO MUSICA

**Accademia Musicale Studio Musica**

# International Conference on New Music Concepts Inspired Education and New Computer Science Generation

Proceeding Book
Vol. 7

Accademia Musicale Studio Musica
Michele Della Ventura
Editor

Printed in Italy
First edition: March 2020


www.studiomusicatreviso.it
Accademia Musicale Studio Musica – Treviso (Italy)
ISBN: 978-88-944350-3-0

# Preface

This volume of proceedings from the conference provides an opportunity for readers to engage with a selection of refereed papers that were presented during the International Conference on New Music Concepts, Inspired Education and New Computer Science Generation. The reader will sample here reports of research on topics ranging from a diverse set of disciplines, including mathematical models in music, computer science, learning and conceptual change; teaching strategies, e-learning and innovative learning, neuroscience, engineering and machine learning.

This conference intended to provide a platform for those researchers in music, education, computer science and educational technology to share experiences of effectively applying cutting-edge technologies to learning and to further spark brightening prospects. It is hoped that the findings of each work presented at the conference have enlightened relevant researchers or education practitioners to create more effective learning environments.

This year we received 57 papers from 19 countries worldwide. After a rigorous review process, 24 paper were accepted for presentation or poster display at the conference, yelling an acceptance rate of 42%. All the submissions were reviewed on the basis of their significance, novelty, technical quality, and practical impact.

The Conferece featured three keynote speakers: Prof. **Giuditta Alessandrini** (Università degli Studi Roma TRE, Italy), Prof. **Renee Timmers** (The University of Sheffield, UK) and Prof. **Axel Roebel** (IRCAM Paris, France).

I would like to thank the Organizing Committee for their efforts and time spent to ensure the success of the conference. I would also like to express my gratitude to the program Committee members for their timely and helpful reviews. Last but not least, I would like to thank all the authors for they contribution in maintaining a high-quality conference and I hope in your continued support in playing a significant role in the Innovative Technologies and Learning community in the future.

March 2020

Michele Della Ventura

## Conference Chair

*Michele Della Ventura*, Accademia Musicale Studio Musica, Treviso, Italy

## Keynote Speakers

*Giuditta Alessandrini,* Università degli Studi Roma TRE, Italy
*Renee Timmers,* The University of Sheffield, UK
*Axel Roebel,* IRCAM Paris, France

## International Scientific Committee

Patricia Alessandrini, Goldsmiths, University of London, UK
Jioanne Armitage, University of Leeds, UK
Suzanne Aspden, Faculty of Music, University of Oxford, UK
Jean-Julien Aucouturier, IRCAM, Paris, France
Per Bloland, Miami University, Ohio, USA
Jeffrey Boehm, Bath Spa University, UK
David Carabias Galindo, University of Segovia, Spain
Marko Ciciliani, University for Music and Performing Arts Vienna, Austria
Sally Jo Cunningham,University of Waikato, New Zealand
Ching-Hua Chuan, University of North Florida, U.S.A.
Darryl N. Davis, University of Hull, UK
Marlo De Lara, University of Leeds, UK
Elga Dorner, Central European University, Budapest, Hungary
Simon Emmerson, De Montfort University, Leicester, UK
Travis Garrison, University of Central Missouri, USA
Inés María Monreal Guerrero, University of Valladolid, Spain
Duncan Williams, University of Plymouth, UK
Andrew Hankinson, Bodleian Libraries, University of Oxford, UK
Joseph Hyde, Bath SPA University, UK
Wladyslaw Homenda, Warsaw University of Technology, Poland
Orestis Karamanlis, Bournemouth University, UK
Alexandros Kontogeorgakopoulos, Cardiff Metropolitan University, UK
Steven Jan, University of Huddersfield, UK
Tae Hong Park, New York University Steinhardt, USA
Rudolf Rabenstein, University Erlangen-Nuremberg, Erlangen, Germany
Silvia Rosani, Goldsmiths, University of London, UK
Robert Rowe, New York University, USA
Nikos Stavropoulos, Leeds Beckett University, UK
Jacob David Sudol, Florida International University, U.S.A.
Eva Zangerle, University of Innsbruck, Austria

# Contents

## *Learning Tools, Leraning Technologies, Learning Practices*



## *Culture and Music*



## *Poster presentation*

# New Music Concepts

# Automatic Identification of Melody Tracks of Piano Sonatas using a Random Forest Classifier

Po-Chun Wang, Alvin W. Y. Su

SCREAM Lab., Department of CSIE,
National Cheng-Kung University,
Tainan, Taiwan
F74046153@gs.ncku.edu.tw

**Abstract.** In this paper, an identification method of melody tracks of classical piano sonatas is presented. The tracks which are regarded as 'melody lead' are important cues in music interpretation when symbolic sheet music is concerned, especially when computer synthesis of emotional and expressive music is desired. In this work, four new features are proposed. Combined with five conventional features, there are nine features to be extracted from a standard MIDI file. Then, random forest classifier is applied to determine whether a measure is 'melody-like' or 'accompaniment-like'. There are 8 manually annotated classical piano sonatas used to validate the proposed method. Over 90% accuracy is achieved and is 6% higher than the previous work in this art.

**Keywords.** melody finding, music analysis, music perception, symbolic representation

## 1 Introduction

Expression is usually highly related to the notations on the score such as intensity, articulation or pedal use. 'Melody lead' was also considered as another expressive strategy independent of the above [13, 14]. The strategy provides an interpretation method that doesn't require expression notations which may be absent in many sheet music. The expression strategy of melody lead is a dominating cue in multi-voiced music [6]. A melody lead voice is usually played louder and precedes the others. It helps listeners identify the melody line in multi-voiced music environment [7]. This information is especially important when expressive synthesis of music is desired [2][10].

Synthesis of a piano sonata is a practical example. Piano sonatas are sonatas written for a solo piano. They usually consist of two to four movements. The first movement is usually written in sonata form. In piano sonata, music is divided into two parts: melody and accompaniment. They are usually assigned to the left hand and the right hand, respectively. Previous work showed that melody lead was found to increase with expressiveness [3]. It was also shown that melody lead is caused by dynamic differentiation in skilled piano performance [8]. Therefore, it is important to identify the melody tracks.

The melody track identification problem was studied in two aspects: audio and symbolic. In digital sound domain, melody lines were extracted from .wav files [1, 4, 12, 17]. Such studies using symbolic formats of music data is less seen.

In symbolic domain, the melody line seldom switches between tracks for most homophony music. However, melody exists in two or more tracks in many music forms.

Tang et al. [18] proposed a statistical method on selection of candidate melody tracks. For each track, music was described in a sequence of features including AvgVel (average velocity), PMRatio (polyphonic ratio), SilenceRatio, Range and TrackName. Uitdenbogerd and Zobel [19] developed four algorithms for melody line detecting. Li Liu and Cai Junwei [9] extracted melody line based on their melody similarity theory.

When the melody appears in more than one tracks, the identification becomes more complicated. It can be examined as a melody/accompaniment classification problem [5, 16]. Friberg [5] proposed twelve features, including five pitch features, two IOI features, articulation features, timbre feature, and so on. A Support Vector Machines (SVM) was applied to classify these two classes. In David Rizo's [16] work, melody tracks were judged by empirical experiences and features of each track. Five features with twenty descriptions were extracted. They were in category of track information, pitch, pitch intervals, note durations and syncopation. For each track, the probability of being melody or accompaniment was judged by a random forest classifier.

Previous works have been tested for genres of music such as pop, jazz and classical, but none of them has ever been tested for piano sonatas. In most piano sonatas, the melody usually alternates between tracks. Sometimes the melody can also exist in both tracks.

Methods such as Friberg's [5] required additional information such as expression notations. Therefore, Rizo's work [16] is used for comparison because it can recognize melody(s) of multi-track standard MIDI files. The accuracy of this work is 6 percent higher than Rizo's when the first movement of the eight manually annotated classical piano sonatas used in this work are tested.

The rest of the paper is divided into three main sections. First, the methodology on identifying and selecting the melody tracks is introduced. Next, a computer simulation is performed by using eight piano sonatas. The method proposed in [16] is also tested. Finally, conclusion and future work are given.

## 2 Method

In this work, each measure is treated as a segment. The proposed feature vector $\vec{v}$ is described by

$$\vec{v} = [FR,\ SI,\ LCR,\ SDD,\ SDI,\ NR,\ DNC,\ CN,\ RP]. \tag{1}$$

These nine features are respectively Floating Rate (*FR*), Significant Interval difference (*SI*), Level Crossing Rate (*LCR*), Standard Deviation of Duration (*SDD*) and Interval (*SDI*), Note Rate (*NR*), Distinct Note Count (*DNC*), Chord Number (*HN*) and Repeated Pattern (*RP*). The detail of each feature is shown below. $P_i$ represents the i-th pitch. $S$ represents the total note number of a segment. $D$ represents the interval sets where

$$D = \{P_{i+1} - P_i \mid 0 < i \leq S\}. \tag{2}$$

In the following context, the notation $\|\cdot\|$ represents the size of the set.

**Features**
**Floating Rate** (*FR*): This is the average of pitch differences in a segment.

$$FR = \frac{1}{S}\sum_{i=0}^{S-1}(P_{i+1} - P_i) \tag{3}$$

**Significant Interval difference** (*SI*): This shows the significant interval change in a segment. Intervals greater than three semitones are counted.

$$SI = \|\{x \in D \mid x \geq 3\}\| \tag{4}$$

**Level Crossing Rate** (*LCR*): This describes the number of notes crossing mean value $M$, which $M = \frac{1}{S}\sum_{i=0}^{S-1} P_i$.

$$LCR = \|\{P_i \mid (P_{i+1} - M)(P_i - M) < 0\}\|, \quad 0 \leq i \leq S-1 \tag{5}$$

**Standard Deviation of Duration** (*SDD*): $M_L$ represents the mean of the durations of the notes in the segment, where $L_i$ is the duration of the i-th note.

$$M_L = \frac{1}{S}\sum_{i=0}^{S-1} L_i \tag{6}$$

$$SDD = \sqrt{\frac{1}{S}|L_i - M_L|^2}, \qquad 0 \leq i \leq S-1. \tag{7}$$

**Standard Deviation of Interval** (*SDI*): M is the mean of MIDI pitch numbers, where $M = \frac{1}{S}\sum_{i=0}^{S-1} P_i$.

$$SDI = \sqrt{\frac{1}{S}|P_i - M|^2}, \qquad 0 \le i \le S-1. \tag{8}$$

**Note Rate** (*NR*): NR is the number of notes per segment.

$$NR = \frac{total\ note\ number}{unit\ time}. \tag{9}$$

**Distinct Note Count** (*DNC*): *DNC* is the amount of all MIDI pitch numbers in a segment.

**Harmonic interval Number** (*HN*): A harmonic interval event is identified when two or more than two notes appear simultaneously. *HN* counts the number of such events in a segment.

**Repeated Pattern** (*RP*): *RP* is designed to detect the Alberti bass. It is a style of accompaniment, which usually has repeated interval sets in a segment.

Fig. 1. The $D$ of these notes is $[7, -3, 3, -7, 7, -3, 3]$.

Fig. 2. The $D$ of these notes is $[-1, -2, 7, 3, 2, -7, -2]$.

*RP* is calculated in the following steps. First, two vectors in (10) and (11) are considered.

$$u_1[n] = \begin{cases} P_{n+1} - P_n, & 0 \le n < S \\ 0, & otherwise \end{cases} \tag{10}$$

$$u_2[n] = \begin{cases} P_{n+1} - P_n, & 0 \le n < j \\ 0, & otherwise \end{cases}, \text{ for } 1 < j \le \frac{1}{2} S. \tag{11}$$

For all j, cross-correlation between $u_1$ and $u_2$ is computed and resampled.

$$c_j[n] = \sum_{m=0}^{j} u_2[m]\, u_1[n+m] \ . \tag{12}$$

$$C_j[n] = c_j[i \times n]. \tag{13}$$

Thus, RP is computed as follow:

$$RP = min\left(\left\{ \left| \sum_{k=0}^{n} \frac{C_j[k]}{v_2^2[k]} - 1 \right| \;\middle|\; 0 \le n \le \frac{T}{j} \right\}\right) \tag{14}$$

where $1 < j \le \frac{1}{2}S$. When $u_1$ is the perfect repetition of $u_2$, $RP = 0$.

Taking Fig. 1 as an example, $RP$ = 0.76, while $RP$ of Fig. 2 is 2.04. It is obvious that Fig. 1 has more $RP$ potential than the Fig. 2.
This vector $u_1$ is seen as the input to the classifier. A track will be identified as either 'melody-like' or 'accompaniment-like' through a random forest classifier.

**The random forest classifier** [11]
A random forest classifier is an ensemble of decision trees. Trees are weighted and trained by various sub-samples of dataset. It reduces the over-fitting problem on decision tree classifier. In our work, the Scikit-learn [15] package is used to build a random forest classifier. It contains 50 trees and the Gini impurity is considered on nodes splitting. The result could be seen as the probability of containing melody-line in a segment in our case.

# 3 Results

**Datasets**

Eight classical piano sonatas list in Table I are used. Only the first movement of each piano sonata from four classical period composers is selected. Three professional musicians/composers from National Taiwan Normal University, Taiwan were invited to annotate the scores.

The scores are separated into left and right-hand tracks. For each measure, the track(s) containing perceived melody are identified and the rest of tracks are considered as accompaniment. An example is shown in Fig. 3.

TABLE I: INFORMATION OF THE PIANO SONATAS.

| | Score | Composer | Measures | Time signatures |
|---|---|---|---|---|
| b_4_1 | piano sonata No.4 in E-flat major Op.7 | Beethoven | 361 | 6/8 |
| b_20_1 | piano sonata No.20 in G major Op.49, No.2 | Beethoven | 122 | 4/4 |
| c_40_1 | piano sonata in G major, Op.40, No.1 | Clementi | 209 | 4/4 |
| c_47_1 | piano sonata in B flat, Op.47, No.2 | Clementi | 132 | 4/4 |
| h_38_1 | keyboard sonata No.38 in F major, Hob.XVI:23 | Haydn | 128 | 2/4 |
| h_50_1 | keyboard sonata No.50 in D major, Hob.XVI:37 | Haydn | 103 | 4/4 |
| m_7_1 | piano sonata No.7 in C major, K.309 | Mozart | 155 | 4/4 |
| m_16_1 | piano sonata No.16 in C major, K.545 | Mozart | 73 | 4/4 |

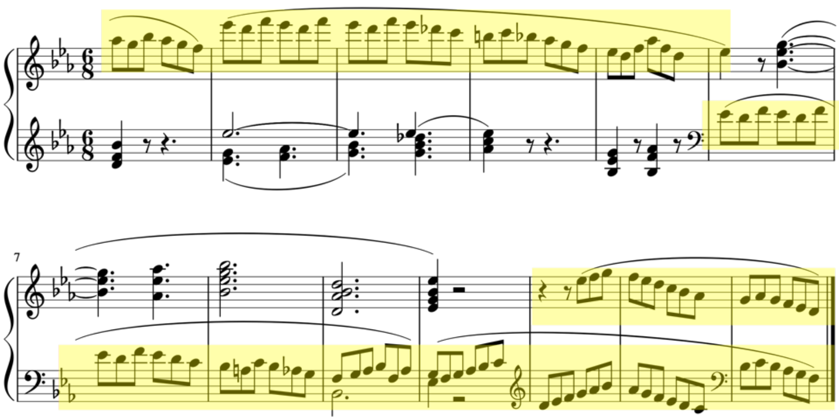


Fig. 3. The yellow parts are annotated by the scorers as 'melody'. It is noted that both tracks of the last three measures are annotated as 'melody'. The rest of the score is annotated as 'accompaniment'.

## Experiment

80% of the measures of the sonatas are randomly split as the training set, and the other 20% are used as the test set. There are 2052 (1283*2*0.8) measures used as the training data. The random forest classifier takes the features vector $\vec{v}$ of a segment described in the previous section and returns the probability of being a melody-like segment. In this work, a segment is equal to a measure. In the experiment, precision, recall and F-measure are shown. True-positive (TN) is the percentage of melody-like segments successfully classified. False-positive (FP), true-negative (TN) and false-negative (FN) are defined accordingly. The accuracy is calculated by (15).

$$accuracy = \frac{TN + TP}{TN + TP + FN + FP} \tag{15}$$

The results are compared with Rizo's [16] method. Table II shows the comparison of the accuracy of each method. Table III shows the comparison of F-measure.

TABLE II: THE COMPARISON OF ACCURACY.

| score | accuracy (This work) | accuracy (Rizo) |
|---|---|---|
| b_4_1 | 0.84 | 0.84 |
| b_20_1 | 0.94 | 0.92 |
| c_40_1 | 0.9 | 0.87 |
| c_47_1 | 0.85 | 0.87 |
| h_38_1 | 0.83 | 0.83 |
| h_50_1 | 0.86 | 0.64 |
| m_7_1 | 0.94 | 0.82 |
| m_16_1 | 0.97 | 0.87 |
| average | 0.9 | 0.83 |

TABLE III: THE COMPARISON OF F1-SCORE.

| score | accuracy (This work) | accuracy (Rizo) |
|---|---|---|
| b_4_1 | 0.86 | 0.85 |
| b_20_1 | 0.94 | 0.92 |
| c_40_1 | 0.9 | 0.87 |
| c_47_1 | 0.87 | 0.87 |
| h_38_1 | 0.87 | 0.85 |
| h_50_1 | 0.91 | 0.73 |
| m_7_1 | 0.94 | 0.85 |
| m_16_1 | 0.97 | 0.88 |
| average | 0.91 | 0.85 |

The average accuracy and F1-score of this work are about 6% higher than Rizo's. The worst F1-score of this work happens in b_4_1, which is 86%. The worst case of Rizo's work happens in h_50_1, which is only 73%.
**Discussion**

In this paper, new features such as *RP*, *HN*, *SI*, *LCR* are proposed in this work. Table IV shows that the average accuracy without four new features is decreased by 1.3%.

TABLE IV: THE COMPARISON OF REMOVING FEATURES.

| Features | accuracy |
|---|---|
| $\vec{v}$ without *RP, HN, SI, LCR* | 88.7% |
| $\vec{v}$ | 90% |

Though there are five features that are also used in [16], it is noted that these features are modified in this work. For example, *DNC* is the combination of the "Track Information" category and the "Pitch" category, and *NR* contains more than two descriptions in "Note durations" including the longest duration and the mean duration. These modifications also account for the improvement over the method proposed in [16].

## 4 Conclusion

In this work, the melody tracks of each measure are identified in piano sonatas. Nine features including $FR, SI, LCR, SDD, SDI, NR, DNC, CN, RP$ are employed. The scope of melody track identification is narrowed from a song to a measure because piano sonatas aren't homophonic. For each measure, a random forest classifier is used to classify it into a 'melody-like' class or an 'accompaniment-like' class. The experiment shows that this work is about 6% more accurate than the method in [16].

In the future, the first attempt is to add more piano sonatas in the training set. In addition, it is desired that the segment size can be further reduced to a beat. Furthermore, it is expected to implement the identification method on different types of music, such as a string quartet and a symphony.

## References

[1] Berenzweig, Adam L., and Daniel PW Ellis. "Locating singing voice segments within music signals." Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No. 01TH8575). IEEE, 2001.

[2] Bresin, Roberto, and Anders Friberg. "Synthesis and decoding of emotionally expressive music performance." IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No. 99CH37028). Vol. 4. IEEE, 1999.
[3] Bruno H. Repp. (1996). "Patterns of note onset asynchronies in expressive piano performance." Journal of the Acoustical Society of America, 100, 3917–3932.
[4] Eggink, Jana, and Guy J. Brown. "Extracting Melody Lines From Complex Audio." ISMIR. 2004.
[5] Friberg, Anders, and Sven Ahlbäck. "A method for recognising the melody in a symbolic polyphonic score." Proceedings of the 9th International Conference on Music Perception and Cognition (ICMPC), Bologna, Italy. 2006.
[6] Goebl, Werner, and Richard Parncutt. "Asynchrony versus intensity as cues for melody perception in chords and real music." Proc. of the 5th Triennial ESCOM Conference, Hanover, Germany. 2003.
[7] Goebl, Werner. "Melody lead in piano performance: Expressive device or artifact?." The Journal of the Acoustical Society of America 110.1 (2001): 563-572.
[8] Goebl, Werner. "Skilled piano performance: Melody lead caused by dynamic differentiation." Proc. of the 6th Int. Conf. on Music Perception and Cognition. 2000.
[9] Li, Liu, et al. "Melody extraction from polyphonic MIDI files based on melody similarity." 2008 International Symposium on Information Science and Engineering. Vol. 2. IEEE, 2008.
[10] Li, Pei-Ching, et al. "Analysis of Expressive Musical Terms in Violin Using Score-Informed and Expression-Based Audio Features." ISMIR. 2015.
[11] Liaw, Andy, and Matthew Wiener. "Classification and regression by randomForest." R news 2.3 (2002): 18-22.
[12] Marolt, Matija. "Audio melody extraction based on timbral similarity of melodic fragments." EUROCON 2005-The International Conference on" Computer as a Tool". Vol. 2. IEEE, 2005.
[13] Palmer, Caroline. "Anatomy of a performance: Sources of musical expression." Music perception: An interdisciplinary journal 13.3 (1996): 433-453.
[14] Palmer, Caroline. "Music performance." Annual review of psychology 48.1 (1997): 115-138.
[15] Pedregosa, Fabian, et al. "Scikit-learn: Machine learning in Python." Journal of machine learning research 12.Oct (2011): 2825-2830.
[16] Rizo, David, et al. "A Pattern Recognition Approach for Melody Track Selection in MIDI Files." ISMIR. 2006.
[17] Salamon, Justin, et al. "Mc and challenges." IEEE Signal Processing Magazine 31.2 (2014): 118-134
[18] Tang, Michael, Yip Chi Lap, and Ben Kao. "Selection of melody lines for music databases." Proceedings 24th Annual International Computer Software and Applications Conference. COMPSAC2000. IEEE, 2000.

[19] Uitdenbogerd, Alexandra L., and Justin Zobel. "Manipulation of music for melody matching." Proceedings of the sixth ACM international conference on Multimedia. 1998.

This book presents a collection of selected papers that present the current variety of all aspect of the research at a high level, in the fields of music, education and computer science. The book meets the growing demand of practitioners, researchers, scientists, educators and students for a comprehensive introduction to key topics in these fields. The volume focuses on easy-to-understand examples and a guide to additional literature.

Michele Della Ventura, editor

**New Music Concepts, Inspired Education, Computer Science**

Revised Selected Papers

ISBN: 978-88-944350-3-0

9 788894 435030

www.studiomusicatreviso.it