

CONSTRUCTING PEDB 2nd EDITION: A MUSIC PERFORMANCE DATABASE WITH PHRASE INFORMATION

Mitsuyo Hashida

Soai University

hashida@soai.ac.jp

Eita Nakamura

Kyoto University

enakamura@sap.ist.i.kyoto-u.ac.jp

Haruhiro Katayose

Kwansei Gakuin University

katayose@kwansei.ac.jp

ABSTRACT

Music performance databases that can be referred to as numerical values play important roles in the research of music interpretation, the analysis of expressive performances, automatic transcription, and performance rendering technology. The authors have promoted the creation and public release of the CrestMusePEDB (Performance Expression DataBase), which is a performance expression database of more than two hundred virtuoso piano performances of classical music from the Baroque period through the early twentieth century, including music by Bach, Mozart, Beethoven and Chopin. The CrestMusePEDB has been used by more than fifty research institutions around the world. It has especially contributed to research on performance rendering systems as training data. Responding to the demand to increase the database, we have started a new three-year project to enhance the CrestMusePEDB with a 2nd edition that started in 2016. In the 2nd edition, phrase information that pianists had in mind while playing the performance is included, in addition to the performance data that contain quantitative data. This paper introduces an overview of the ongoing project.

1. INTRODUCTION

The importance of music databases has been recognized through the progress of music information retrieval technologies and benchmarks. Since the year 2000, some large-scale music databases have been created and have had a strong impact on the global research arena [1–4]. Meta-text information, such as the names of composers and musicians, has been attached to large-scale digital databases and has been used in the analysis of music styles, structures, and performance expressions, from the viewpoint of social filtering in MIR fields.

The performance expression data plays an important role in formulating impressions of music [5–8]. Providing a music performance expression database, especially describing deviation information from neutral expression, can be regarded as a research in sound and music community (SMC). In spite of there being many music researches

using music performance data, few projects have dealt with creation of a music performance database open to public.

In musicological analysis, some researchers constructed a database of the transition data of pitch and loudness and then use the database through statistical processing. Widmer *et al.* analyzed deviations in the tempi and dynamics of each beat from Horowitz’s piano performances [9]. Sapp *et al.*, working on the Mazurka Project [10], collected as many recordings of Chopin mazurka performances as possible in order to analyze deviations of tempo and dynamics by each beat in a similar manner to [9].

The authors have been promoting the creation and public release of the CrestMusePEDB (Performance Expression DataBase), which consists of more than two hundred virtuoso piano performances of classical music from the Baroque period through the early twentieth century, including music by Bach, Mozart, Beethoven, and Chopin [11]. The 1st edition of the CrestMusePEDB (ver.1.0-3.1) has been used musical data and has been used by more than fifty research institutions throughout the world. In particular, it has contributed to researches on the performance rendering systems as training data [12, 13].

The database is unique in providing a set of detailed data of expressive performances, including the local tempo for each beat and the dynamics, onset time deviation, and duration for every note. For example, the performance elements provided in the Mazurka Projects data [10] are beat-wise local tempos and dynamics and precise information with note-wise performance elements that cannot be extracted. In the MAPS database [14], which is widely used for polyphonic pitch analysis and piano transcription, performance data does not include any temporal deviations and thus cannot be thought of as realistic performance data in the aspect of musical expression. Such detailed performance expression data are crucial for constructing performance rendering systems and realistic performance models for analysis and transcription.

The size of the CrestMusePEDB 1st edition is not large enough, compared with other databases for computer science. Demand for the database has been increasing in recent years, particularly in the studies using machine learning techniques. In addition, data that explicitly describe the relationship between a performance and the musical structure intended by the performer has been required¹. Responding to these demands, we started a three-year project in 2016 to enhance the CrestMusePEDB in a 2nd edition,

¹ In many cases, the apex (the most important) note in a phrase is selected by the performer, and there may be the case that phrase sections are analyzed based on the performers own interpretation.

which is described in this paper.

2. CRESTMUSEPEDB 1ST EDITION

The 1st edition of the database CrestMusePEDB² aimed to accumulate descriptions of concrete performance expressions (velocity, onset timing, etc.) of individual notes as deviation information from ‘mechanical’ performance. It was focused on classical piano music from the Baroque period through the early twentieth century, including music by Bach, Mozart, Beethoven and Chopin. We chose 51 music pieces, including those often referred to by previous musical studies in the past couple of decades. We also chose various existing recordings by professional musicians. The database contains 121 performances played by one to ten players for each score.

The CrestMusePEDB 1st edition consisted of the following four kinds of component data.

PEDB-SCR (score text information): The score data are included in the database. Files in the MusicXML format and in the standard MIDI file (SMF) format are provided.

PEDB-IDX (audio performance credit): The catalogs of the performances from which expression data are extracted: album title, performer name, song title, CD number, year of publication, etc.

PEDB-DEV (performance deviation data): Outline of curves of tempo and dynamics, the delicate control of each note; deviation regarding starting time, duration, and dynamics from the tactus standard corresponding to the note. Performances from 1 to 10 performers were analyzed for each piece, and multiple deviation data were analyzed by different sound sources and provided for each performance. All data are described in the DeviationInstanceXML format [15, 16].

PEDB-STR (musical structure data): This contains the *estimated* information on a musical structure data set (hierarchical phrase structure and the top note of a phrase) from performances. The data are described in the compliant MusicXML format. The structure data corresponds with a performance expression data in PEDB-DEV. However, if multiple adequate interpretations exist in a piece, the multiple structure data are provided in the performance data.

PEDB-REC (original recordings): The recorded audio performance data were based on the PEDB-STR. Nine players provided 121 performances to compare different expressions from a similar phrase structure.

The primary data were given in SCR (score text information data) and DEV (performance deviation data) files. CrestMusePEDB does not contain any acoustic signal data (WAV files, AIFF files, MP3 files) except for PEDB-REC.

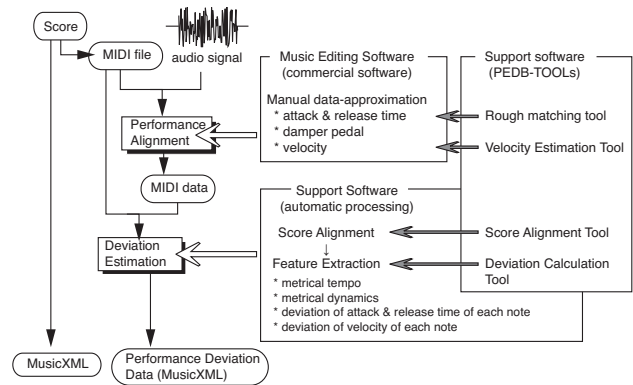


Figure 1. Outline of database generation (1st edition)

Instead, it contained the catalogs of the performances from which expression data were extracted.

Transcribing a performance audio signal into MIDI-level data was the core issue for constructing the CrestMusePEDB. Although we had improved the automation of the procedure, an iterated listening process by more than one expert with a commercial audio editor and original alignment software possessed higher reliability.

Fig. 1 illustrates the procedure for generating the PEDB-DEV information.

3. PEDB 2ND EDITION OVERVIEW

The 1st edition of the CrestMusePEDB has contributed to the SMC field, especially for performance rendering studies. The mainstream of the current performance systems refers to the existing performance data. Above all, systems based on recent machine learning techniques require large corpora. The size of the 1st edition CrestMusePEDB is not necessarily large, compared with the other databases published for the research of natural language processing or speech recognition. Demand for the database enhancement has been recently increasing.

Another expectation imposed on the performance database is the handling of information of the musical structure. Although virtuoso performances remain in the form of an acoustic signal, it is hard to find a material that shows the relationship between a performance and its musical structure that the performer intended. In many cases, we had no choice but to estimate the performers intention from the recorded performances.

Responding to these demands, we started a new three-year project in 2016, to enhance the database with a 2nd edition, with the goals of increasing the data size and providing structural information. One of the major problems with making the 1st edition was the workload required for the manual transcription processing of performances in the form of an acoustic signal. To solve this problem, we newly recorded performance data using YAMAHA Disklavier, with the cooperation of skillful pianists who have won prizes in piano competitions. This procedure enabled us to obtain music performance control data (MIDI) and acoustic data simultaneously.

To improve the efficiency of further analysis and utiliza-

² <http://crestmuse.jp/pedb/>

tion of the database, each note in the performance data should be given information of the corresponding note in its musical score. For this goal, the matching file is generated using our original score-performance alignment tool.

We released the beta version of the 2nd edition of the PEDB, which includes data of approximately 100 performances, at the end of May 2017, to meet users requirements regarding data format and data distribution methods. The beta version will include recorded MIDI files, score MIDI files, score files in the MusicXML format, musical structure data, and matching files. Performance in acoustic signals are also included as a reference material. Music structure data released in the beta version include phrase and sub-phrase, and apex notes in each phrase, which are obtained by interviewing the pianists, in a pdf format. Here, the apex note is the most important note from the pianists' perspective in each phrase. We are planning to discuss with the user group of the database the format of the machine-readable structure data and the deviation to be included in the next formal version with the user-group of the database.

4. PROCEDURE FOR MAKING THE DATABASE

4.1 Overview

The main goals of this edition are to enhance the performance data and to provide structural information paired with the performance data. The outline of database generation is shown in Fig. 2.

Musical structure differs depending on pianists' interpretation and even on the score version. Some of the musical works have multiple versions of the musical scores; such as the Paderewski Edition, the Henle Verlag, and the Peters Edition, e.g., Mozart's famous piano sonata in A Major K. 331.

Before the recording, the pianists were asked to prepare to play their most natural interpretation of the score that they usually use. Further they were requested to prepare additional interpretation such as by a different score edition, by a professional conductor suggests, and by over-expressed the pianists' interpretation. Pianists were requested to express the difference of these multiple interpretations regarding the phrase structure. After the recording, we interviewed the pianist on how (s)he tried to express the intended structure of the piece after listening to the recorded performances.

Through these steps, source materials for the database are obtained. Then, the materials are analyzed and converted to the data, as they can be referenced in the database. The main procedure of this stage is note-to-note matching between score data and the performance. Some notes in the performance may be erroneously played, and the number of notes played in trill is not constant. To handle such situations, we developed 'an original score-performance alignment tool based on a hidden Markov model (HMM) [17].

In the following subsections, the recording procedure and the overview of the alignment tool are described.

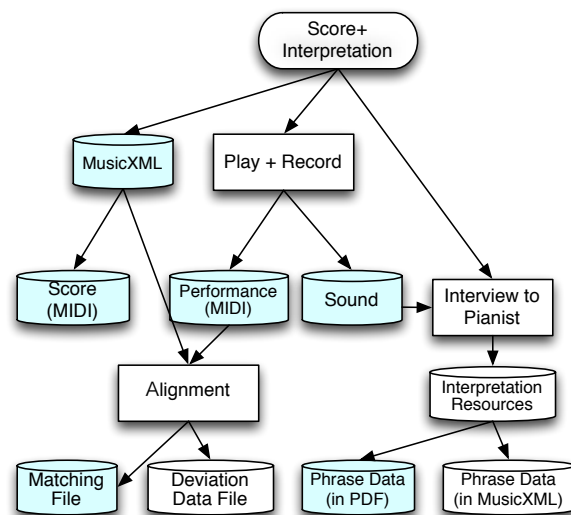


Figure 2. Outline of the database generation. Blue: data included in the beta version.

4.2 Recording

The key feature in the creation of the 2nd edition PEDB is that we can talk with the pianists directly about their performances. Before the recording procedure, we confirmed with each pianist that the recorded performance should clarify its phrase structure (phrase or sub-phrase) and its apex note, as the "interpretation" of the performance.

Pianists are asked to (1) play based on their own interpretation for all pieces and to (2) play with exaggerated expressions retaining the phrase structure for some pieces. In addition, for some pieces, pianists are asked to (3) play with the intension of accompanying a soloist or dancers. If there are different interpretations or score editions of one piece, they played with the both versions. For Mozart's piano sonata in A Major K. 331 and Beethoven's "Pathétique" Sonata, the 2nd movement, different versions of the scores have been provided to the pianists by the authors.

Recordings were done in several music laboratories or recording studios. As shown in Fig. 3, performances played with a YAMAHA Disklavier were recorded as both audio signals and MIDI data including controls of pedals, via ProTools. We also recorded the video for the interview process after recording.

4.3 Alignment Tool

After MIDI recordings are obtained, each MIDI signal is aligned to the corresponding musical score. To improve the efficiency, this is done in two stages: automatic alignment and correction by an annotator. In the automatic alignment stage, a MIDI performance is analyzed with an algorithm based on an HMM [17], which is one of the most accurate and fast alignment methods for symbolic music. A post-processing to detect performance errors, i.e., pitch errors, extra notes, and missing notes, is also included in this stage.

In the following stage of correction by an annotator, a vi-

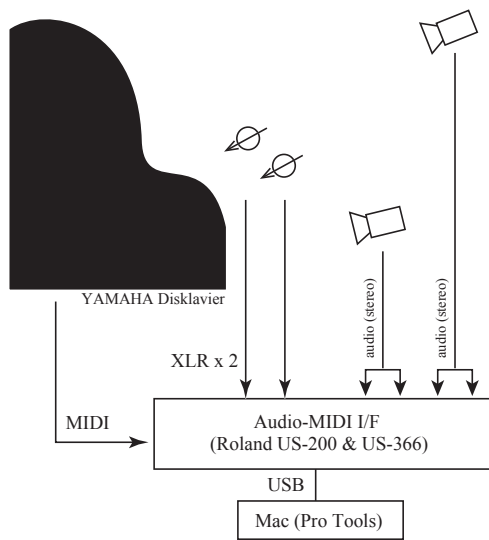


Figure 3. Technical setup for the recording

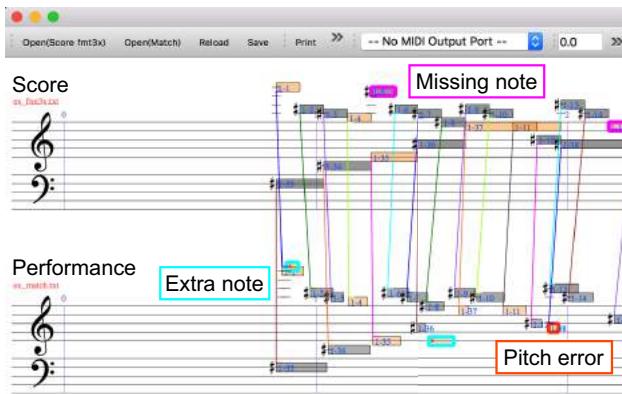


Figure 4. A visualization tool used to obtain the alignment information (see text).

sualization tool called *Alignment Editor* is used to facilitate the procedure. In the editor, the performance and score are represented as piano rolls, and the alignment information is represented as lines between the corresponding score notes and performed notes, as well as indications for performance errors (Fig. 4). On each illustrated note, an ID referring to the score notes is also presented, and the annotator can edit it to correct for the alignment. Upon saving the alignment information, the editor automatically reruns the performance error detection and updates the graphic.

5. PUBLISHING THE DATABASE

We released a beta version of the 2nd edition PEDB consisting of 103 performances of 43 pieces by two professional pianists at the end of May, 2017. Table 1 shows the list of the performances included in the beta version. As shown in this table, some of the pieces are played with more than one expression or interpretation.

This edition provides (1) recording files (WAV and MIDI), (2) score files (MusicXML and MIDI), (3) infor-



Figure 5. A sample of phrase structure data (W. A. Mozart's Piano Sonata K. 331, 1st Mov., Peters Edition.) Square bracket and circle mark denotes phrase/sub-phrase and apex, respectively.

mation regarding phrase and apex notes in phrases by in PDFs (Fig. 5) and (4) the alignment information by the original file format (*matching file format*) (see Fig. 2.)

In a matching file format, the recorded MIDI performance is represented as a piano roll. For each note, onset time, offset (key-release) time, pitch, onset velocity, and offset velocity are extracted and presented. In addition, the corresponding score note, the score time, and performance error information are provided for each performed note. To represent the performance error information, each performed note is categorized as a correct note, a pitch error, or an extra note, according to the results of the score-to-performance alignment. In the case of an extra note, no corresponding score note is given. The file also describes a list of missing score notes that have no corresponding notes in the performance.

As shown in Fig. 4, the durations of the performed notes are usually played shorter where the damper pedal is pressed. When pressed, the damper pedal sustains the sound of the notes until the pedal is released. It means that there are two interpretations for note offset (key-release) time: actual note-off time and pedal-release time. In a matching file format, we adopted the actual note-off time, as a offset (key-release) time. The pedal information is included in the recorded MIDI files.

The database is available from the PEDB 2nd Edition url³.

6. CONCLUDING REMARKS

In this paper, we introduced our latest attempt to enhance the PEDB and overviewed part of the database as a beta version to investigate the users requests for the database. Although the number of data currently collected in the beta version is small, we have completed the workflow for the database creation. In the coming years, we plan to increase the number of performance data to more than five hundred

³ <http://crestmuse.jp/pedb.edition2/>

Table 1. List of performances included in the beta version. *self*: the player's expression, *over*: over expression, *accomp.*: played as the accompaniment for solo instrument or chorus, *waltz*: focused on leading a dance, *Hoshina*: played along with a professional conductor's interpretation, *Henle* and *Peters*: used the score of Henle Edition and Peters Edition, and others: extra expressions via discussion with the authors and each player.

No.	Composer	Title	Performances			
			Player 1		Player 2	
			#	interpretation	#	interpretation
1	J. S. Bach	Invention No. 1	2	self / over	2	self / over
2	J. S. Bach	Invention No. 2	2	self / over	1	self
3	J. S. Bach	Invention No. 8	-	-	1	self
4	J. S. Bach	Invention No. 15	2	self / over	1	self
5	J. S. Bach	Wohltemperierte Klavier I-1, prelude	2	self / accomp.	-	
6	L. V. Beethoven	Für Elise	1	self	3	self / over / note_d
8	L. V. Beethoven	Piano Sonata No. 8 Mov. 1	1	self	1	self
9	L. V. Beethoven	Piano Sonata No. 8 Mov. 2	2	self / Hoshina	2	self / Hoshina
10	L. V. Beethoven	Piano Sonata No. 8 Mov. 3	1	self	1	self
7	L. V. Beethoven	Piano Sonata No. 14 Mov. 1	2	self / over	1	self
11	F. Chopin	Etude No. 3	2	self / over	1	self
12	F. Chopin	Fantaisie-Impromptu, Op. 66	2	self / over	1	self
15	F. Chopin	Mazurka No. 5	1	self	-	
13	F. Chopin	Mazurka No. 13	2	self / over	-	
14	F. Chopin	Mazurka No. 19	2	self / over	-	
16	F. Chopin	Nocturne No. 2	2	self / over	1	self
17	F. Chopin	Prelude No. 1	2	self / over	-	
18	F. Chopin	Prelude No. 4	1	self	1	self
19	F. Chopin	Prelude No. 7	1	self	1	self
20	F. Chopin	Prelude No. 15	1	self	1	self
21	F. Chopin	Prelude No. 20	1	self	-	
22	F. Chopin	Waltz No. 1	2	self / waltz	2	self / waltz
23	F. Chopin	Waltz No. 3	2	self / over	1	self
24	F. Chopin	Waltz No. 7	2	self / waltz	2	self / waltz
25	F. Chopin	Waltz No. 9	2	self / over	1	self
26	F. Chopin	Waltz No. 10	1	self	1	self
27	C. Debussy	La fille aux cheveux de lin	2	self / over	-	
28	C. Debussy	Rêverie	1	self	-	
29	E. Elgar	Salut d'amour Op. 12	2	self / accomp.	-	
30	G. Händel	Largo / Ombra mai fü	2	self / accomp.	-	
31	Japanese folk song	Makiba-no-asa	2	exp1 / exp2	-	
32	F. Liszt	Liebesträume	1	self	-	
33	F. Monpou	Impresiones intimas No. 5 "Pajaro triste"	1	self	-	
34	W. A. Mozart	Piano Sonata K. 331 Mov. 1	3	self / Henle / Peters	2	Henle / Peters
35	W. A. Mozart	Twelve Variations on "Ah vous dirai-je, Maman"	2	self / over	-	
36	T. Okano	Furusato	2	self / accomp.	2	self / accomp
37	T. Okano	Oboro-zuki-yo	2	self / accomp.	-	
38	S. Rachmaninov	Prelude Op. 3, No. 2	2	self / neutral	-	
39	M. Ravel	Pavane pour une infante défunte	1	self	-	
40	E. Satie	Gymnopédies No. 2	2	self / neutral	-	
41	R. Schumann	Kinderszenen No. 7 "Träumerei"	2	self / neutral	1	self
42	P. I. Tchaikovsky	The Seasons Op. 37b No. 6 "June: Barcarolle"	2	self / over	-	
Total:			70		31	

by adding new data to the previous version, considering the compatibility with the data-format of the PEDB 1st edition.

There is no other machine-readable performance database associated with musical structure. We hope that the database can be utilized for research in many research fields related to music performances. As future work, we would like to develop some applications, in addition to the data-format design, so the database can be used by researchers who are not familiar with information technology.

Acknowledgments

The authors are grateful to Dr. Y. Ogawa and Dr. S. Furuya for their helpful suggestion. We also thank Ms. E. Furuya

for her advice as a professional pianist, and Professor T. Murao for his cooperation. This work was supported by JSPS KAKENHI Grant Numbers 16H02917, 15K16054, and 16J05486. E. Nakamura is supported by the JSPS research fellowship (PD).

7. REFERENCES

- [1] J. S. Downie, "The music information retrieval evaluation exchange (mirex)," in *D-Lib Magazine*, 2006, p. Vol.12 No.12.
- [2] "http://staff.aist.go.jp/m.goto/rwc-mdb/," 2012 (last update).

- [3] D. McEnnis, C. McKay, and I. Fujinaga, "Overview of omen," in *Proc. ISMIR*, Victoria, 2006, pp. 7–12.
- [4] H. Schaffrath, "http://essen.themefinder.org/," 2000 (last update).
- [5] M. Senju and K. Ohgushi, "How are the player's ideas conveyed to the audience?" in *Music Perception*, vol. 4, no. 4, 1987, pp. 311–323.
- [6] H. Hoshina, *The Approach toward a Live Musical Expression: A Method of Performance Interpretation considered with energy*. Ongaku-no-tomo-sha, 1998, (written in Japanese).
- [7] N. Kenyon, *Simon Rattle: From Birmingham to Berlin*. Faber & Faber, 2003.
- [8] K. Stockhausen, *Texte zur elektronischen und instrumentalen Musik*, J. Shimizu, Ed. Gendai-shicho-shinsha, 1999, (Japanese translation edition).
- [9] G. Widmer, S. Dixon, S. Goebel, E. Pampalk, and A. Tobudic, "In research of the Horowitz factor," *AI Magazine*, vol. 24, no. 3, pp. 111–130, 2003. [Online]. Available: <https://www.aaai.org/ojs/index.php/aimagazine/article/view/1722/1620>
- [10] C. Sapp, "Comparative analysis of multiple musical performances," in *Proc. ISMIR*, Vienna, 2007, pp. 497–500.
- [11] M. Hashida, T. Matsui, and H. Katayose, "A new music database describing deviation information of performance expressions," in *Proc. ISMIR*, Kobe, 2008, pp. 489–494.
- [12] M. Hashida, K. Hirata, and H. Katayose, "Rencon Workshop 2011 (SMC-Rencon): Performance rendering contest for computer systems," in *Proc. SMC*, Padova, 2011.
- [13] H. Katayose, M. Hashida, G. De.Poli, and K. Hirata, "On evaluating systems for generating expressive music performance: the rencon experience," in *J. New Music Research*, vol. 41, no. 4, 2012, pp. 299–310.
- [14] V. Emiya, R. Badeau, and B. David, "Multipitch estimation of piano sounds using a new probabilistic spectral smoothness principle," in *IEEE TASLP*, vol. 18, no. 6, 2010, pp. 1643–1654.
- [15] T. Kitahara, "Mid-level representations of musical audio signals for music information retrieval," in *Advances in Music Information Retrieval*, Springer, vol. 274, 2010, pp. 65–91.
- [16] T. Kitahara and H. Katayose, "CrestMuse toolkit: A java-based framework for signal and symbolic music processing," in *Proc. Signal Processing (ICSP)*, Beijing, 2014.
- [17] E. Nakamura, N. Ono, S. Sagayama, and K. Watanabe, "A stochastic temporal model of polyphonic midi performance with ornaments," in *J. New Music Research*, vol. 44, no. 4, 2015, pp. 287–304.