Toward empirical analysis for stylistic expression in piano performance

Yu-Fen Huang[1] and Li Su[1]

Music & Culture Technology Lab (MCTL), Institute of Information Science,
Academia Sinica, Taiwan
yfhuang@iis.sinica.edu.tw
lisu@iis.sinica.edu.tw

Abstract. In the performance of Western art music, musicians apply various strategies to manipulate the performed sound, and communicate their musical interpretations via these subtle acoustic variations. It is a common practice for musicians to use typical conventions to express each compositional style (e.g. Baroque, Classical, or Romantic compositions). However, such stylistic expressive conventions has yet been fully discussed in previous research. In this initial foray, we systematically compare the expressive strategies for different piano compositions. A series of piano performance are recorded with a controlled experimental setting (3 compositions $\times$ 8 pianists $\times$ 3 repeated trials = 72 recordings), and expressive acoustic elements are derived using Music Information Retrieval techniques. In our analysis, we reveal that expressive manners in music performance exhibit stable and systematic features corresponding to each music composition, and those stylistic trends serve as empirical observations for typical performance conventions in different music styles.

Keywords: expressiveness, performance style, piano performance, computational musicology, Music Information Retrieval

1 Introduction

In the past decades, the way how music audiences approach, appreciate, and get to understand music has been evolved with the revolution of digital technologies. From attending physical concerts, purchasing audio/video medium (e.g. CD, DVD), to getting access to large amount of digitized performance recordings via online streaming services, audiences embrace the opportunities to explore the variety of music performance. In the context of Western art music, musicians have the privilege to interpret the written composition, and to communicate their understanding of the music piece via intricate variations in their performance (e.g. micro-timing, dynamic, timbral, and articulation arrangement). Audiences also enjoy the process to contemplate and compare diverse artistic variations in different performance versions, and through which process to discover potential insights for classical repertoire. The artistic, expressive variations in music performances therefore serve as an essential communicative vehicle to deliver musical ideas in the cultured convention.

Recently, systematic studies for large-scale music performance corpus are advanced with the facilitation of Music Information Retrieval (MIR) techniques. Computational models have been established to map general connections between musical attributes (e.g. note length, note pitch, phrase position) and expressive variations in performance [12] [16]. From musicological aspect, scholars focus on specific genres (e.g. Baroque music) and examine how particular aesthetic styles in music are shaped by artistic variations in performance [9]. Musicians also contribute their creative and unique interpretations in the process of performance execution. [18] [33]. The execution of music performance is therefore a complex and interactive process across aforementioned factors - the composition, stylistic convention, and individual musicians. The relationships between these aspects and how they interact together, however, have yet been fully discussed in previous research, particularly with systematic analysis of individual factors.

In this study, we aim to provide context-valid observations in terms of the interactions between different factors leading to musical expressiveness. A piano performance corpus is collected under a controlled experimental setting, and Music Information Retrieval (MIR) techniques are applied to retrieve expressive variations in tempo and dynamics. The performance variations are analyzed in conjunction with compositional elements through statistical and time-series methods. We identify important factors to induce unique music expression, and then subsequently investigate how those factors interact in different performing contexts. The contribution of this work is threehold:

- To compile a new piano performance dataset with controlled experimental design for comparison;
- To identify different key factors affecting music expression in individual scenario via statistical analysis;
- To provide empirical observations of how different factors interact together in a time-series process.

In the next section, we will discuss previous studies regarding musical expressiveness. The data collection and data processing procedure of this study will be reported in Section 3. In Section 4, we will investigate stylistic expressive trends found in individual compositions, and distinctive expressions bound for different compositional elements. In Section 5, our analysis results will be discussed in conjunction with findings in previous research, and we therefore suggest that our analysis can be implemented as an empirical approach to describe systematic variations for different expressive styles in music performance.

2 Related work

The expressiveness in music performance is shaped by complex interactions among diverse factors. In the context of Western art music, composers follow conventional rules to construct the melodic, rhythmic, and harmonic configurations of music (*compositional factor*) [1] [23], and each composer's work would exhibit distinctive character according to the composer's preference (*stylistic factor*) [6]. During the performance process, musicians have their unique fashion to communicate personal musical interpretations, and control the variations in performed acoustic sound (e.g., micro-timing, dynamic, and articulation variations) (*musician factor*) [25] [26]. GERMS model systematically categorized different origin of musical expressiveness including: generative rules, emotional expression, random variations, motion principles, and stylistic unexpectedness [19]. We will review related works regarding three different origin of musical expressiveness in this section.

2.1 Compositional factors in music performance

In previous studies, rule-based models are established to describe the connection between compositional elements and expressive variations. The KTH model combines generative rules in melodic, harmonic, rhythmic, and phrase aspects to predict the timing, dynamics, and articulation execution in performance [12]. For piano performance, rule-based models and linear Gaussian models can be applied to jointly predict the tempo, dynamic, and articulation variations in performance according to multiple attributes in melodic, rhythmic, and harmonic aspects of performance [11]. Based on large amounts of jazz performance data, inductive logic rules for expressive elements (note onset deviation, dynamic variation, and ornamentation) are found in jazz music [14]. It is also found that tempo and dynamic variations interact together in music performance, and the tempo-loudness trajectory is an effective description to illustrate distinctive features of performance style [4].

Another cluster of studies apply machine learning approach to explore potential relationships between compositional elements and expressive variations. The connection between expressive tempo variations and musical phrase is mapped using Gaussian Mixture Models (GMMs)[24]. Models with transitional hidden state are applied to predict expressive variations according to score-informed attributes. For instance, Hidden Markov Model (HMM) and Hierarchical HMM are used to estimate the expressive variations in piano performance [17]. Conditional random fields (CRFs) are applied to predict expressive elements based on melodic and harmonic components [20]. Feed Forward Neural Networks (FFNNs) are used to predict the dynamic variations based on local-level score-informed attributes including the pitch, duration, and the note's relative interval with neighboring notes [3]. Linear and non-linear models for musical expression are systematically evaluated in [2], and it is concluded that compared to linear models, non-linear models have better performance to estimate the tempo and dynamic changes in music performance.

2.2 Stylistic factors in music performance

Computational models are also built to explore specific performance styles. Restricted Boltzmann Machines (RBMs) are capable of predicting expressive accentuations in piano performance [36]. For solo violin performance, long-term dynamic variations can be successfully modelled using Random Forest, k-nearest neighbors (k-NN), and Support Vector Machine (SVM) [15]. For string quartet, the timing deviation, dynamic level, and the extent of vibrato in performance can be estimated based on melodic (e.g. relative interval) and rhythmic (e.g. metrical hierarchy) descriptors using model trees, k-NN, and SVM [21]. For jazz music, Decision Tree, SVM, and Neural Network (NN) are developed to formulate the stylistic deviations in jazz performance, and the improvised embellishments can be predicted from attributes including the chord type, note duration, and phrase [14].

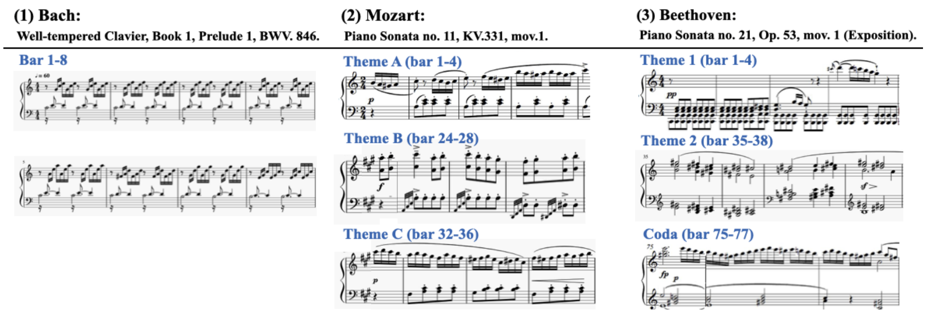


Fig. 1. The repertoire for data collection: three piano solo works by Bach, Mozart and Beethoven.

2.3 Musician factors in music performance

Musicians have their personal expressive manners in music performance. In order to compare different performance versions for the same composition, entropy-based deviation measures are used to describe expressive timing patterns in individual performance versions [25]. Hierarchical clustering is also an useful implement to distinguish different trends of performing styles for orchestral works [26]. For violin performance, individual violinists have their own expressive strategies to convey melodic patterns and phrase structure in the composition [29]. For jazz music, different performance styles by individual musicians can be successfully distinguished based on their intra-note features (e.g. note's attack level, sustain duration, amount of legato, spectral centroid, spectral tilt), inter-note features (e.g. relative pitch and duration to the neighboring notes), and note-to-note transition (pitch contour) [34].

The style of music performance is highly idiosyncratic according to the music genre, the music composition, and the musician. In particular, in Western art music, it is a common practice for pianists to play compositions in Baroque, Classical, and Romantic period following distinctive conventions. Aforementioned studies tends to explore a specific aspect of musical expression (i.e. either compositional, stylistic, or musician factor alone). Yet in music performance, the interactive, dynamical process among different factors work together to shape diverse variations. Based on the foundation of previous research, this work collects a series of piano performance data, and systematically analyze how individual factors interplay together to shape the overall musical expressiveness.

3 Method

In order to systematically examine how different factors affect the expressive variations in piano performance, a series of piano performance data are collected in this study. In this section, we report the procedure for data collection, and the data processing methods to extract expressive variations from recorded performance audio.

3.1 Data collection

Eight pianists are recruited to participate the recording sessions. The recruited participants are graduate/undergraduate students majoring in piano in music department at university (male = 4, female = 4). We reached out the participant pool via personal contact of music department staff. Participants are all right-handed, with the average age of

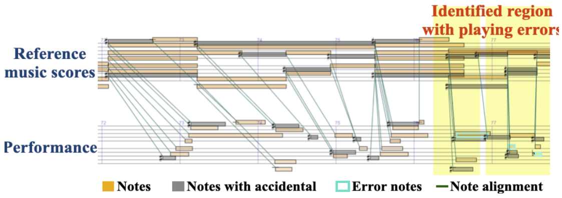


Fig. 2. The automatic audio-to-score alignment process to extract note timing from performance.

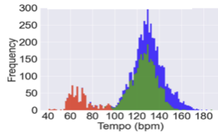
21.63 (SD = 0.70). They have learnt piano for 14.88 years in average (SD = 2.42), and practice for 3.75 hours in average per day at the time of data collection (SD = 1.84).

In order to compare different performance styles, piano solo works written by different composers are selected as materials (Fig. 1). Repeated measures design is used to observe if pianists show stable features when performing the same piece of music multiple times. The three music pieces selected for data collection are Bach Well-tempered Clavier, Book 1, Prelude 1, BWV.846; Mozart Piano Sonata no. 11, KV. 331, mov.1, and Beethoven Piano Sonata no. 21, Op. 53, mov.1 (bar 1-86, Exposition). In our experiment, each pianist played each music piece for 3 times in a random order, which resulted in 72 performance recordings (3 music pieces x 3 performances x 8 pianists). The recording sessions took place at the Motion Analysis Lab (National Yang Ming Chiao Tung University, Taiwan), where the performances were recorded on Yamaha digital piano P-115. In order to accurately align each notes in the performance with the music scores in the subsequent stage of data analysis, the performance is recorded as both midi and audio formats. During recording sessions, participants' body were also attached to optical markers to record 3-d motion capture data for their performance body movement, which will be further analyzed elsewhere.

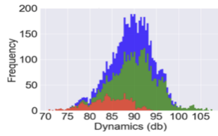
3.2 Data processing

The dynamic and tempo variations in piano performance are the two expressive features to be analyzed in our subsequent investigation. In order to derive our target features, the first step is to perform audio-to-score alignment and to obtain the time stamp for each note in the performance. The audio-to-score alignment for polyphonic music, particularly with casual playing errors and asynchrony between two hands in piano performance, has been considered as a challenging task in Music Information Retrieval [5] [13] [30]. We applied an automatic alignment method to align recorded midi files and music scores files, in which Viterbi algorithms are used to exclude playing error regions and alignment errors in a pre-alignment process, and then hidden Markov models are applied to divide notes playing by two hands and accurately re-align each note based on the merged information [31] (see Fig. 2).

In order to associate expressive variations with the overall compositional structures, we analyze the dynamic and tempo variations at bar-level instead of instant dynamics and tempo. For tempo variations, the timing for each beat is extracted from the audio-to-score alignment data. The note information (e.g. the note pitch, duration, bar and beat position) is extracted from xml files using Python library Music21 [7], and such information from music scores serves as the reference to locate corresponding onset timing for each beat in the performance. In case of absent note on downbeat, linear interpolation is performed based on the timing of neighboring beats; in case of the asynchrony



Tempo statistics				K-W test	
Composition	Mean	SD	Range	Composition factor	Pianist factor
Bach	69.70	10.93	57.22	p< .001 All pairs are significant	p<.001 Not significant pairs: P1-P2, P1-P8, P3-P5, P3-P7, P5-P7
Mozart	126.77	10.85	96.27		
Beethoven	136.30	14.58	124.67		



Dynamics statistics				K-W test	
Composition	Mean	SD	Range	Composition factor	Pianist factor
Bach	84.23	4.54	24.87	p< .001 All pairs are significant	p<.001 Not significant pairs: P1-P3, P2-P3, P2-P8, P4-P7, P5-P8
Mozart	90.97	4.77	31.82		
Beethoven	89.26	4.34	24.47		

Fig. 3. The data distribution and statistics for two expressive elements (tempo (upper), dynamics (lower)) in piano works written by Bach (red), Mozart (green), and Beethoven (blue).

between two hands, the note with the lowest register is taken as the reference. The ***tempo*** per bar is then defined as the average bpm (beat per minute) per bar. For dynamic data, the decibel is computed from the input audio (sampling rate = 22050 Hz) using Python library Librosa [28]. In order to eliminate the disturbance of local noise, moving average (window size = 220 samples, roughly 0.01 seconds) is used to smooth the original data. Since each note has a natural ADSR (attack-decay-sustain-release) curve, and the maximum volume on the attack is the main feature concerned, the ***dynamic*** level per bar is defined as the maximum decibel within a bar duration. Conventional music analysis is performed on the music scores, and structural features of music works including the harmonic progression, phrase and sectional boundaries are analyzed to be compared with features extracted from performance audio.

4 Analysis results

Through the data collection and processing procedure, a series of piano performance data are collected, and expressive tempo and dynamic variations are extracted. In this section, we report our analyses and observations for stylistic expression in three aspects: 1) the general expressive manners, 2) the interaction between different compositional elements and expressive features, and 3) the time-series expressive trends found in individual compositions, which can be regarded as individual stylistic expressive strategies attached to the composition.

4.1 General expressive manners

In our analysis, it appears that pianists use different strategies to express each composition. In Fig. 3, different tendencies of pianists' expressive manners can be observed in the distributions and statistics of tempo (the upper panel) and dynamics (the lower panel). In order to distinguish the influence from compositions versus from pianists, We perform statistical tests on two factors (composition and performer). Since the distributions of both tempo and dynamic data violate the assumption of homogeneity in Levene's test, non-parametric tests (Kruskal-Wallis tests) are performed instead of regular ANOVA, and Bonferroni correction is applied to post hoc analysis [10]. Regarding the tempo data, the statistical analyses yield significant differences between all three compositions, while the performances for Mozart's and Beethoven's pieces have more similar average tempo (126.77 and 136.30 bpm) compared to Bach's piece (69.70 bpm). Comparing the performances for Mozart's and Beethoven's compositions, Beethoven's composition possesses higher variations (SD = 14.58, range = 124.67) than Mozart's work (SD = 10.85, range = 96.27). It is worth noted that according to the statistic analyses on the pianist factor, some pianists have distinctive expressive strategies, which can be distinguished from other pianists' expressive trend.

4.2 Interactions between compositional and expressive elements

Pianists perform individual compositions with diverse expressive manners, and they may use different expressive strategies to communicate each compositional element. In this section, we further analyze the interaction between different compositional elements and expressive features. We focus on the phrase, section, and harmony aspects of composition, and manually-annotate five different features for each musical bar in each composition: 1) section boundary (on section boundary/ non-boundary), 2) phrase boundary (on phrase boundary/ non-boundary), 3) section position (in the first /middle/ last one third of a section), 4) phrase position (in the first/ middle/ last one third of a phrase), 5) harmony (I/ V/ other types of chord). We contemplate both the boundary and relative position for section/phrase, since in our preliminary observation, we found that musicians tend to show different manners at section/phrase boundaries, and their expressive tendencies also vary when they initiate a new section/phrase versus when they are approaching the end of section/phrase. For the harmonic aspect, we only compare three types of chord to simplify the analysis process, since it is not straight forward to observe the overall general tendency in the comparisons of many groups (e.g. comparing all 7 degrees of chord lead to C_2^7 = 21 combinations). In many chord types, we choose tonic and dominant chords to analyze, considering that those two chords take essential position in Western tonal music and are often used to signify structural location in music (authentic or half cadence).

For statistic analyses, we take the expressive measurements (tempo/ dynamics) in each bar, and then split the data into groups according to their compositional elements. Fig. 4 (A0 and B0) shows all the 26 comparison groups for statistic analysis. We first perform normality tests for all groups, and subsequently carry out homogeneity tests for three-group comparisons (section position, phrase position, and harmony type). For groups violating the normality or homogeneity assumptions, the non-parametric counterpart is performed instead of parametric test (i.e. t-test or Mann-Whitney U test for two-group comparisons; one-way ANOVA or Kruskal-Wallis test for three-group comparisons). For three-group comparisons, we further carry out post hoc tests to compare different combinations. Aforementioned procedure is performed three times for the three compositions individually.

In Fig. 4, the general expressive tendencies (column 0) and differences between compositions (column 1 - 3) can be observed. For tempo variations (Fig. 4, A0), musicians generally incline to slow down at section and phrase boundaries, as well as at the bars with tonic and dominate chords, since those chords may coincide with cadence. But different expressions emerge when comparing three compositions. For the section position, in Bach's and Mozart's music piece (Fig. 4, A1 and A2), musicians' tempo variation exhibits a U-shape curve, in which they tend to perform with faster tempi at the beginning and end of section, and slightly slow down in the middle of section, whereas in Beethoven's music piece (Fig. 4, A3), musicians' tempo curve tilts toward

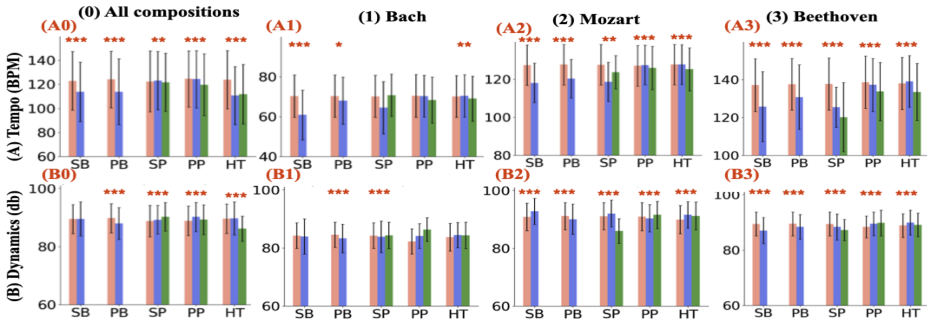


Fig. 4. The statistic analyses for 2 expressive elements (tempo (row A), dynamics (row B)), with 5 compositional elements in 3 music pieces (Bach (column 1), Mozart (column 2), Beethoven (column 3)), including the means (bars), standard deviations (error bars), and significance level (stars). Comparison groups are: section boundary (SB), phrase boundary (PB) (non-boundary (pink), boundary (blue)), section position (SP), phrase position (PP) (the first 1/3 (pink), the middle 1/3 (blue), the last 1/3 (green)), harmony types (HT) (others (pink), I (blue), V (green)).

the end of section, in which they apply a ritardando to highlight the end of section. For dynamic variations (Fig. 4, B0), it is shown that pianists tend to perform with softer dynamic levels when they are approaching the end of section or phrase. The softer dynamic level sometimes incorporate with ritardando to be used as the expressive strategy to shape 'the sense of direction toward the end of phrase/ section' in performance. We can observe that most of comparisons yield significant difference between groups, which indicates that musicians generally jointly use the combination of different expressive variations (tempo and dynamics) to deliver compositional traits in music, whereas in Bach's music piece, the section, phrase, and harmonic structures sometimes are not manifest in expressive variations.

4.3 Stylistic time-series expressions in compositions

In addition to the comparison between different compositional elements, in music performance, both expressive variations and compositional elements are revealed during the course of time. We therefore take a step further in this section to discuss the time-series connection between expressive and compositional elements. In Fig. 5, the time-series curve of average tempo (Row A) and dynamics (Row B) per bar (for all trials performed by all pianists) are aligned with musical elements in compositions including phrases (Row C), sectional boundaries (Row D), and harmonic progression (Row E). In our analysis, the expression curves show that pianists tend to adopt different patterns of variation in their performance to convey distinctive traits in each composition.

In Bach's work, pianists perform with a steady tempo, except an obvious ritardandos indicating the cadence at the end of the piece (Fig. 5, 1A). The dynamic variation in Bach shows distinct features corresponding to the phrase structure and the harmonic progression (Fig. 5, 1B). For the phrase structure, the dynamic curve exhibits an inverted U-shape matching with phrase boundaries per 4 to 6 bars, which shows that pianists tend to perform a crescendo for the first half of the phrase, and then perform a decrescendo for the second half of the phrase. For the harmonic progression, louder performance dynamics are applied to emphasize harmony with higher tension such as

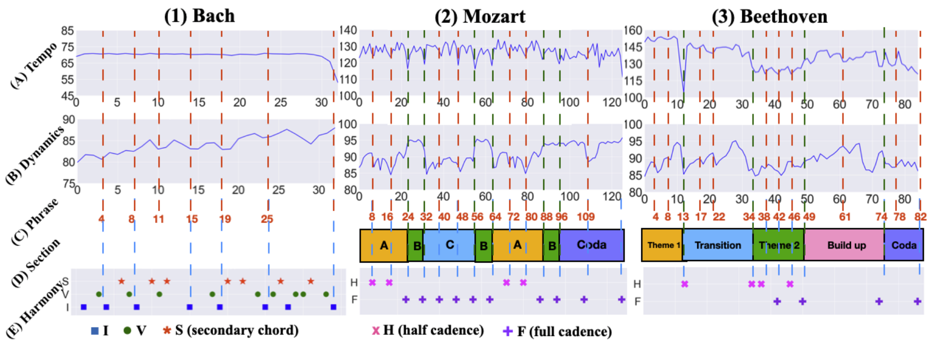


Fig. 5. Time-series trends of tempo and dynamic variations in piano performance. The expressive curves of mean tempo (bpm) (Row A) and the mean dynamics (db) (Row B), aligned with compositional elements including phrases (Row C), sectional boundaries (Row D), and harmonic progression (Row E) in Bach's (Column 1), Mozart's (Column 2), and Beethoven's (Column 3) compositions.

secondary chords (red markers), whereas softer dynamics associate with tonic chord (blue markers), which often coincides with the boundary of phrase and represents the release of harmony tension.

In Mozart's composition, the tempo and dynamic curves in pianists' performances exhibit different traits compared to Bach's work. In contrast with the smooth tempo curve in Bach's work, pianists apply tempo variations to express the phrase structure when performing Mozart's work, in which their tempi tend to slow down at phrase boundaries (Fig. 5, 2A). For dynamic variations, in Bach's work, pianists use dynamic variations (crescendo-decrescendo patterns) to express phrase structure, whereas dynamic variations in Mozart's work serve as the means to convey higher-level music structure (Fig. 5, 2B). In Mozart's work, the regions with relatively louder dynamic levels often coincide with the appearance of theme B. In Mozart's this composition, theme A and theme C possess contrasting characters compared to theme B. Theme A and theme C mostly consist of rapid sixteenth notes, whereas the main components of theme B are unison chords played by both hands simultaneously. It would be a natural practice for pianists to perform theme B with a louder dynamic level in this case. An interesting observation is that given the contrasting dynamic levels between different themes (theme A, C versus theme B), the dynamic variations in Mozart's work still reflect the harmonic progression at the local-level. As shown in (Fig. 5, 2B), within individual themes, the valleys of dynamic curve at local regions often coincide with the release of harmonic tension, such as half cadences (pink markers) or full cadences (purple markers). Those observations indicate that the dynamic curve in Mozart's work is formed by complex interactions between diverse musical components, including the theme arrangement and the harmonic progression.

In Beethoven's composition, pianists accentuate the musical structure using different strategies compared to the previous two compositions. It appears that pianists tend to focus on the higher-level structure of music rather than local-level details in this composition, and employ combinative strategies to emphasize their interpretation of the overall musical structure. For the tempo variation, Bach's work has a smooth tempo curve, and in Mozart's work, pianists apply tempo variations (accelerando-ritardando pattern) to express local phrase boundaries. On the other hand, in Beethoven's composition, obvious valleys in tempo curve tend to coincide with the end of structural sections rather than local phrase boundaries (Fig. 5, 3A), which indicates that pianists employ a noticeable ritardando to signify the end of the section. The dynamic variations in Beethoven's work exhibit multi-layered musical features in the composition (Fig. 5, 3B). The global trend in dynamic variation shows inverted U-shapes corresponding to the high-level section structure of different themes, which suggests that pianists' performances exhibit the crescendo-decrescendo global pattern for each structural section. In addition, the dynamic curve within local regions still reflects detailed local-level features in the composition, in such a way that occasional fluctuations with limited range match with phrase boundaries, and the curve valleys are usually consistent with the locations with lower harmonic tension (half or full cadences).

To summarize general time-series trends observed in the tempo and dynamic variations, pianists employ diverse strategies to communicate the harmonic, phrase, and sectional structure in the three music compositions. Pianists generally utilize dynamic variations (crescendo-decrescendo pattern) to convey the phrase and harmonic structure in Bach's work. In contrast, in Mozart's work, dynamic variations are the means to communicate higher-level sectional structure rather than local phrase boundaries, and the phrase structure is more manifest in the tempo variation curve (ritardando at phrase end). In Beethoven's composition, tempo and dynamic variations exhibit complex influences from diverse musical features. The sectional structure is evident in both tempo (ritardando at section end) and dynamic variations (crescendo-decrescendo pattern), and the dynamic fluctuations are affected by global features in sectional structure, as well as by local features in phrase and harmony.

5 Discussion

In the previous section, we reported our findings regarding tempo and dynamic variations in performances of three piano pieces, and how pianists apply different expressive variations to communicate distinctive musical structures in the composition. We will further incorporate our findings with previous research in this section.

According to our analysis, musical phrase appears to be one of the main components for musicians to express in their performance. Previous research reported that the arching pattern in tempo curve [9] [35] and dynamic variations [15] are attached to phrase formation. In our analysis, we further found that pianists apply diverse strategies to express phrase structure when they are performing different compositions. For instance, the dynamic variations indicate the phrase structure in Bach's composition (crescendo-decrescendo pattern per phrase), whereas tempo variations are mostly used to express phrase in Mozart's composition (slow down at phrase end). Different expressive strategies also reflects diverse compositional characters in these two music pieces. Bach's composition holds an invariant rhythmic pattern, which is expressed by a stable tempo in pianists' performances. In contrast, pianists are more likely to emphasize the dynamic change in order to highlight the tension-release process for the harmonic progression in Bach's composition, and such harmonic progression usually conforms with phrase structure (e.g. cadence at the end of phrase). The compositional structure in Baroque period mostly focuses on the development of short motives, whereas compositions in Classical period emphasize clear formation of phrase. In Mozart's composition, pianists therefore apply the accelerando-ritardando pattern in the tempo curve to shape the direction of the phrase.

Regarding the combination of local music elements and the global structure of music, previous studies suggest that the expressive manner in music performance is affected by both local elements (e.g. melodic peak, rhythmic grouping) [32] [22] and global structure (e.g. sectional arrangement) [9]. In our analysis, we found that pianists apply different strategies to stress local and global elements. For instance, Beethoven's composition exhibits an interesting combination of local and global factors, in which the general curve in both tempo and dynamic variations remain mostly consistent with global sectional arrangement, while the variations still show small-range fluctuations corresponding to local phrase boundaries. Compared to Bach's and Mozart's works, Beethoven's composition has sophisticated theme transformation accompanied by frequent modulation, and the manifest harmonic tension build-up process is one of the key features in Beethoven's compositions. Pianists may therefore manipulate both tempo and dynamic variations in their performance to communicate this important structural character.

It emerges from our analysis that in piano performance, expressive variations in tempo and dynamics exhibit systematic variations consistent with musical structures. Such systematic variations can be regarded as typical components to shape distinctive performance style, in which we generally expect that pianists should apply different expressive conventions when they are performing different styles of music in Western art music (e.g. Baroque, Classical, Romantic compositions). Our analysis method and results can serve as empirical means and provide observations for diverse performance styles in Western art music. Our current analysis is limited to piano performances for several selected compositions, and this analysis procedure can be further applied to the investigation for wider range of repertoire and for different instrument's performances.

6 Conclusion

In this paper, we show that pianists' performances exhibit systematic expressive variations corresponding to diverse compositional styles in Western art music. We collected 72 piano performance recordings for three compositions, and derived expressive variations in tempo and dynamics using automatic audio-to-score alignment and MIR techniques. Statistical and time-series analyses are performed to clarify the relationship between different compositional and expressive element, as well as their time-series connections during the course of performance. It is found that pianists apply stylistic expressive variations to communicate musical components at both global (e.g. sectional arrangement, harmonic progression) and local (e.g. phrase boundaries) levels, and they choose different expressive strategies according to distinctive traits of each composition. We suggest that those systematic variations in expressive elements constitute the core of distinctive performance style, and the complex interaction among diverse expressive elements (e.g. tempo and dynamic variations) at multi-layered musical structures (local and global levels) can compose an empirical approach to describe and compare idiosyncratic music performance styles.

References

1. Aldwell, E., Schachter, C., Cadwallader, A.: Harmony and Voice Leading. Cengage Learning (2018)
2. Cancino-Chacón, C.E., Gademaier, T., Widmer, G., Grachten, M.: An evaluation of linear and non-linear models of expressive dynamics in classical piano and symphonic music. Machine Learning, 106(6), 887-909 (2017)
3. Cancino-Chacón, C.E., Grachten, M.: An evaluation of score descriptors combined with non-linear models of expressive dynamics in music. In: Proceedings of International Conference on Discovery Science (2015)
4. Cancino-Chacón, C.E., Grachten, M., Goebl, W., Widmer, G.: Computational models of expressive music performance: A comprehensive and critical review. Frontiers in Digital Humanities, 5, 25 (2018)
5. Chen, C. T., Jang, J. S. R., Liou, W.: Improved score-performance alignment algorithms on polyphonic music. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (2014)
6. Cook, N.: A Guide to Musical Analysis. Oxford University Press, USA (1994)
7. Cuthbert, M.S., Ariza, C.: Music21: A toolkit for computer-Aided musicology and symbolic music data. In: Proceedings of International Society for Music Information Retrieval Conference (2010)
8. Dodson, A.: Expressive asynchrony in a recording of Chopin's Prelude No. 6 in B Minor by Vladimir de Pachmann. Music Theory Spectrum, 33(1), 59-64 (2011)
9. Fabian, D.: A Musicology of Performance: Theory and Method Based on Bach's Solos for Violin. Open Book Publishers (2015)
10. Field, A.: Discovering Statistics Using IBM SPSS Statistics (4th edition). Sage (2013)
11. Flossmann, S., Grachten, M., Widmer, G.: Expressive performance rendering with probabilistic models. In: A. Kirke, E. R. Miranda (Eds.), Guide to Computing for Expressive Music Performance, pp. 75-98, Springer (2013)
12. Friberg, A., Bresin, R., Sundberg, J.: Overview of the KTH rule system for musical performance. Advances in Cognitive Psychology, 2(2-3), 145-161 (2006)
13. Gingras, B., McAdams, S.: Improved score-performance matching using both structural and temporal information from MIDI recordings. Journal of New Music Research, 40(1), 43-57 (2011)
14. Giraldo, S., Ramirez, R.: A machine learning approach to ornamentation modelling and synthesis in jazz guitar. Journal of Mathematics and Music, 10(2), 107-126 (2016)
15. Giraldo, S., Waddell, G., Nou, I., Ortega, A., Mayor, O., Perez, A., Ramirez, R.: Automatic assessment of tone quality in violin music performance. Frontiers in Psychology, 10, 334 (2019)
16. Grachten, M., Widmer, G.: Linear basis models for prediction and analysis of musical expression. Journal of New Music Research, 41(4), 311-322 (2012)
17. Grindlay, G., Helmbold, D.: Modeling, analyzing, and synthesizing expressive piano performance with graphical models. Machine Learning, 65(2), 361-387 (2006)
18. Héroux, I.: Creative processes in the shaping of a musical interpretation: a study of nine professional musicians. Frontiers in Psychology, 9, 665 (2018)
19. Juslin, P.N.: Five facets of musical expression: A psychologist's perspective on music performance. Psychology of Music, 31(3), 273-302 (2003)
20. Kim, T. H., Fukayama, S., Nishimoto, T., Sagayama, S.: Statistical approach to automatic expressive rendition of polyphonic piano music. In A. Kirke, E.R. Miranda (Eds.), Guide to Computing for Expressive Music Performance, pp. 145-179, Springer (2013)
21. Kosta, K., Ramírez, R., Bandtlow, O. F., Chew, E.: Mapping between dynamic markings and performed loudness: a machine learning approach. Journal of Mathematics and Music, 10(2), 149-172 (2016)
22. Leech-Wilkinson, D.: Cortot's Berceuse. Music Analysis, 34(3), 335-363 (2015)
23. Lerdahl, F., Jackendoff, R.: A Generative Theory of Tonal Music. MIT Press (1983)
24. Li, S., Dixon, S., Plumbley, M.D.: Clustering expressive timing with regressed polynomial coefficients demonstrated by a model selection test. In: Proceedings of International Society for Music Information Retrieval Conference (2017)
25. Liem, C. C., Hanjalic, A.: Expressive timing from cross-performance and audio-based alignment patterns: An extended case study. In: Proceedings of International Society for Music Information Retrieval Conference (2011)
26. Liem, C. C., Hanjalic, A.: Comparative analysis of orchestral performance recordings: An image-based approach. In: Proceedings of International Society for Music Information Retrieval Conference (2015)
27. Llorens, A.: Brahmsian articulation: Ambiguous and unfixed structures in op. 38. Music Theory Online, 27(4) (2021)
28. McFee, B., Raffel, C., Liang, D., Ellis, D. P., McVicar, M., Battenberg, E., Nieto, O.: Librosa: Audio and music signal analysis in python. In: Proceedings of the 14th Python in Science Conference (2015)
29. Molina-Solana, M., Lluis-Arcos J., Gomez, E.: Identifying violin performers by their expressive trends. Intelligent Data Analysis, 14(5), 555-571 (2010)
30. Nakamura, E., Ono, N., Saito, Y., Sagayama, S.: Merged-output hidden Markov model for score following of MIDI performance with ornaments, desynchronized voices, repeats and skips. In: Proceedings of International Computer Music Conference (2014)
31. Nakamura, E., Yoshii, K., Katayose, H.: Performance error detection and post-processing for fast and accurate symbolic music alignment. In: Proceedings of International Society for Music Information Retrieval Conference (2017)
32. Ornoy, E., Cohen, S.: Analysis of contemporary violin recordings of 19th century repertoire: Identifying trends and impacts. Frontiers in psychology, 9, 22-33 (2018)
33. Payne, E.: Creativity beyond innovation: Musical performance and craft. Musicae Scientiae, 20(3), 325-344 (2016)
34. Ramirez, R., Maestre, E., Serra, X.: Automatic performer identification in commercial monophonic jazz performances. Pattern Recognition Letters, 31(12), 1514-1523 (2010)
35. Spiro, N., Gold, N., Rink, J.: Musical motives in performance: A study of absolute timing patterns. In E. Chew, G, Assayag, J. B. Smith (Eds.), Mathemusical conversations: Mathematics and computation in music performance and composition, pp. 109-128. World Scientific (2016)
36. Van Herwaarden, S., Grachten, M., De Haas, W. B.: Predicting expressive dynamics in piano performances using neural networks. In: Proceedings of The Conference of International Society for Music Information Retrieval (2014)
37. Windsor, W. L., Desain, P., Penel, A., Borkent, M.: A structurally guided method for the decomposition of expression in music performance. The Journal of the Acoustical Society of America, 119(2), 1182-1193 (2006)