A MODEL OF DYNAMICS PROFILE VARIATION, DEPENDING ON EXPRESSIVE INTENTION, IN PIANO PERFORMANCE OF CLASSICAL MUSIC

De Poli Giovanni, Rodà Antonio, Vidolin Alvise
Dipartimento di Elettronica e Informatica
Università di Padova – Via Gradenigo 6a – 35100 Padova - Italy
depoli@dei.unipd.it; ar@csc1.unipd.it; vidolin@dei.unipd.it

Abstract

We asked five pianists to perform several different versions of the same score, inspired by a set of sensorial and affective adjectives. An analysis of the dynamics profiles shows that notable differences can be recognized between the different versions of the same score. In spite of that, same relation can be found between the dynamics profile of the different musical interpretations. This feature allowed us to formalize the relation between the profiles by means of a limited number of parameters. All the performances were mapped into a two-dimensional space, the dynamics parametric space. The results show that most of the performances, inspired by the same adjective, are grouped together in the same region. Each region of the space can be, therefore, associated with a specific adjective. The model was applied on various piano scores. The results show that the model has good generalization attribute and can properly render the dynamics characteristics of performances on varying of performer's expressive intentions.

1 Introduction

It is known that several performances of the same score often differ significantly, in particular when the musicians are instructed to play it with different expressive intentions [1]. In this context, expressive intention is taken to mean the inspiration given to musician through adjectives in order to obtain different expressive performances. According to the played instrument, the performer can use various musical means (timing, dynamics, amplitude envelopes, vibrato, tongue etc.) to express his/her interpretation of the score. This work deals with dynamics profiles, i.e. the values of note intensity during the performance. The aim of this paper is the discussion of the following questions: 1) how do dynamics profile change when a musician is asked to play drawing inspiration from a particular expressive intention? 2) is there any common performance strategy if different musicians are inspired by the same expressive intention?

2 Model

We asked five pianists (called pianist A, B, C, D, and E) to play the first 16 bars of the second movement of Mozart's piano sonata K545. The musicians performed several different versions of this score, inspired by a set of sensorial and affective adjectives: natural (na), bright (br), dark (da), hard (ha), soft (so), heavy (he), light (li), passionate (pa), and flat (fl). All the pianists played the Yamaha Disklavier and the performances were recorded in MIDI format.

Figure 1 shows key-velocity values measured in the nine performances of a single pianist. Each curve (called dynamics profile) represents the set of values measured in a single performance. In order to simplify the discussion, we reported only the pianist A's data, even if the following comments are true also for the other pianists. Dynamics profiles allow us to know the exact course in time of key-velocity. Due to the large amount and variability of data, however, they don't allow an easy comparison among the musicians' performance strategies. To this end, it is necessary to define a model that allows a parametric description of the different performances. By means of the model parameters, it will be possible to highlight and compare the main expressive characteristics of the performances.

The model is based on the observation that the score structure suggests suitable behaviors to the player. In order to emphasize some elements of the music structure (i.e. phrases, accents, etc.), the musician changes dynamics by means of expressive patterns as crescendo, decrescendo, sforzando etc.; otherwise the performance would not sound musical. Many works analyzed the relation or, more correctly, the possible relations between music structure and dynamics [2], [3], [4], [5]. The fact that there are many different interpretations of the same score [6], however, shows that musician keeps many freedom degrees beyond this relation.

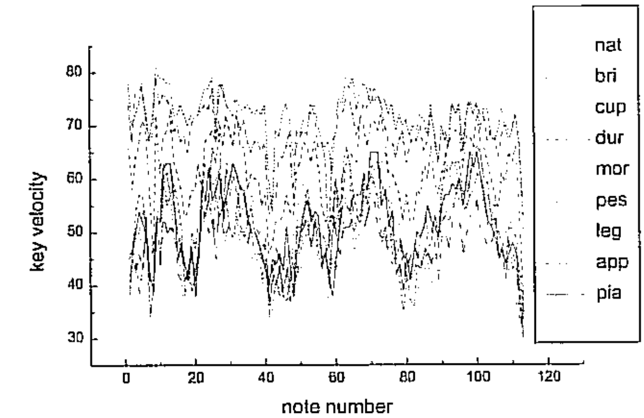

Fig 1: key-velocity in the nine performances of pianist A. The values were measured by means of Yamaha Disklavier.

	br	da	ha	so	he	li	pa	fl
na	0.55 \ 0.91	0.59 \ 0.87	0.45 \ 0.83	0.50 \ 0.82	0.28* \ 0.79	0.61 \ 0.87	0.58 \ 0.81	0.41 \ 0.84
br	-	0.35 \ 0.86	0.60 \ 0.85	0.42 \ 0.84	0.50 \ 0.79	0.51 \ 0.86	0.54 \ 0.82	0.45 \ 0.83
da	-	-	0.43 \ 0.82	0.52 \ 0.80	0.31 \ 0.77	0.54 \ 0.84	0.43 \ 0.80	0.51 \ 0.81
ha	-	-	-	0.38 \ 0.74	0.52 \ 0.86	0.42 \ 0.81	0.41 \ 0.79	0.45 \ 0.77
so	-	-	-	-	0.37 \ 0.62	0.57 \ 0.81	0.49 \ 0.75	0.42 \ 0.76
he	-	-	-	-	-	0.40 \ 0.76	0.27** \ 0.71	0.51 \ 0.74
li	-	-	-	-	-	-	0.61 \ 0.85	0.46 \ 0.84
pa	-	-	-	-	-	-	-	0.42 \ 0.69

Tab 1: minimum and maximum correlation coefficients calculated between the nine performances of each pianist ($p<0.001$ for all performances except * $p<0.003$ and ** $p<0.004$).

The hypothesis for the application of the model is that: when we ask to a musician to play in accordance with a particular expressive intention, he works on the available freedom degree, without destroying the relation between music structure and dynamics [7]. A proof of this hypothesis can be found in the dynamics profile of figure 1, where the structure of the score is the same for all the nine performances. If the relation between music structure and dynamics don't change, many common patterns could be observed among the profiles. To this end, the correlation coefficients between the different versions of each pianist were calculated.

Table 1 shows, for each pair of adjectives, the minimum and maximum correlation coefficients calculated for the five pianists. A significantly correlation can be noted between all the adjective ($p<0.004$). This result implies that all the profiles of each pianist have a similar shape, which we assume to be depending on music structure. Figure 2, in which dynamics profiles were normalized to zero mean and unitary variance, clarifies this observation. The relation between dynamics profiles and the main elements of music structure is particularly evident: for instance (see also figure 3) the musician emphasized with a *decrescendo* the end of the first *inciso* (bar 2), the first semi-phrase (bar 4), the first phrase (bar 8) and the period (bar 16).

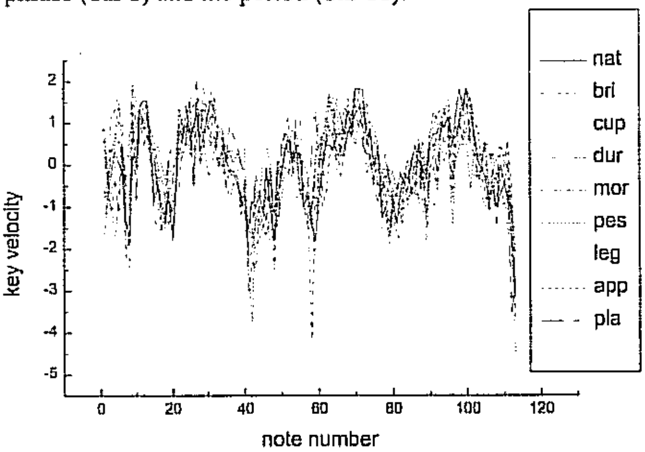

Fig 2: normalized dynamics profiles of pianist A.

The model exploits the idea that all the dynamics profiles can be obtained from an input profile (which agrees with the music structure) by means of some elementary transformations. We have now to define what is the input profile and which are the necessary transformations. We used, as input profile, the average of the dynamics profiles measured in the nine performances (figure 3). Since the mean calculation puts in evidence the performance common characteristics, which are supposed to depend by music structure, the average profile keeps intact the relation between structure and dynamics. Moreover, it represents the geometric center of gravity of the nine performances, which property we will discuss in the next paragraphs.

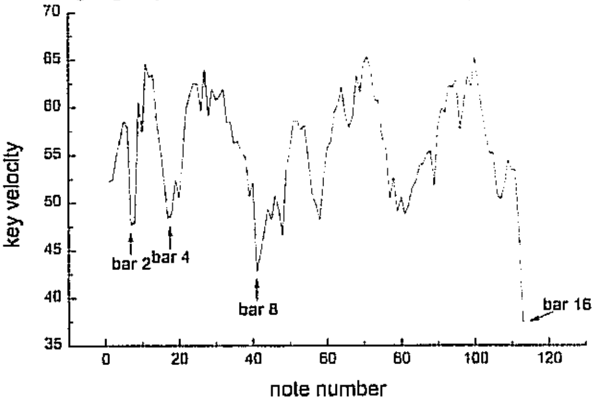

Fig 3: average profile of pianist A. It can be seen the relation between dynamics and the main elements of music structure

The transformations have to satisfy some conditions: 1) they have not to destroy the relation between structure and dynamics, 2) they have not to introduce too many parameters in order to not complicate the model unnecessarily. In order to represent the main characteristics of the performances, we used only two transformations: one shift and one expansion/compression of the values. The two above conditions are satisfied by a linear model, formally represented by the equation:

$$\tilde{y}_e(n) = k_e \cdot \bar{x} + m_e \cdot \left(x(n) - \bar{x}\right) \qquad \text{(Eq. 1)}$$

where $x(n)$ is the key-velocity of n-th note of the average profile, $\bar{x}$ is the mean of x, and are respectively the coefficients of shift and expansion/compression related to expressive intention e, $\tilde{y}_e(n)$ is the estimated key-velocity of the version related to expressive intention e. The parameters k_e and m_e, for each expressive intention, were estimated in order to minimize the square error $\sum_n (y_e(n) - \tilde{y}(n))^2$, where $y_e(n)$ is the key-velocity of the n-th note, measured in the performance inspired by the expressive intention e.

3 Results

An average profile for each pianist was been calculated and the model parameters of his nine versions were estimated. Two values (m_e and k_e) are associated to each performance. So we can map the performances in a two-dimensional space, called Dynamics Parametric Space (DPS), which axes are defined by the two model parameters.

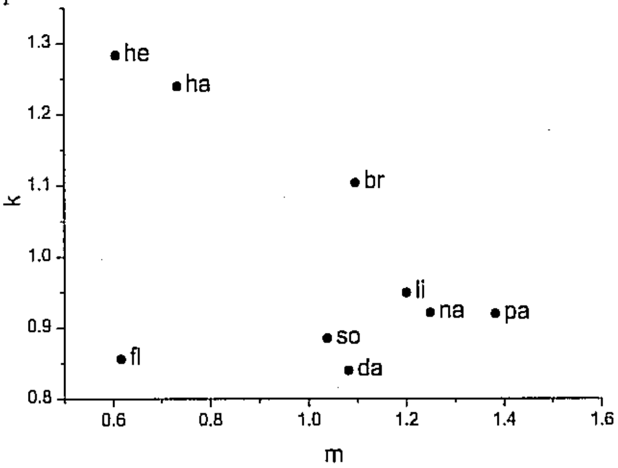

Fig 4: pianist A's performances, mapped in the Dynamics Parametric Space.

Figure 4 shows how the pianist A's versions are mapped in the DPS. By means of this space, we can easily obtain information about the musician's performance strategies: for instance the heavy version is characterized by higher key-velocity values (high k), in opposition to the dark version (low k); the passionate version is characterize by a large dynamics range (higher m), in opposite to the flat version (low m). So performances can be differentiated by means of two main characteristics: in this way, we answered to the first basic question.

Now we will discuss the second point, that is if there is any common strategy among the musicians. All the five pianists' performances were mapped in the DPS (figure 5).

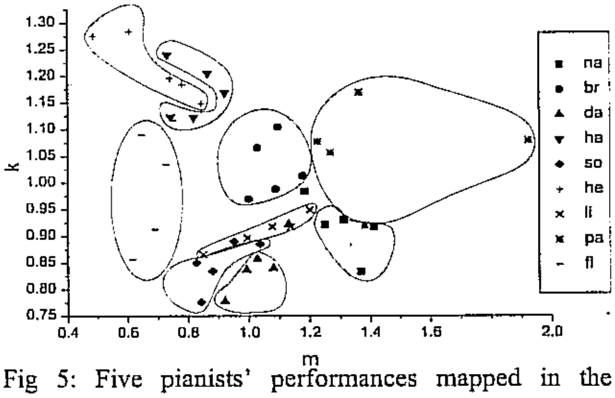

Fig 5: Five pianists' performances mapped in the Dynamics Parametric Space.

It can be seen that most performances, inspired by the same adjective, are grouped together in a region of the space. This fact signify that all the pianist have characterized, for instance, the soft version by means of a lower intensity (low k) and the bright version by means of intermediate values (k and m in the center of the DPS). These results suggest that there are some common strategies among the pianists, at least in relation to the proposed adjectives.

4 Discussion

In order to verify how the model works, the ratio between the variance accounted by the model and the total variance was calculated for each performance. Table 2 reports the mean, minimum, and maximum values, calculated among the performances of all the pianists. It can be seen that the mean variance accounted for by the model is about 67%, with a maximum value of above 90%. Only two performance have a variance below 50% (the hard and heavy version of pianist C).

	na	br	da	ha	so	he	li	pa	fl
Mean	78	70	67	61	65	53	73	73	60
Min	65	53	52	45	54	34	64	67	51
Max	90	91	86	83	78	73	86	79	77

Tab. 2: variance accounted by the model with two parameters (values are expressed as percentage of the total variance).

The values of table 2 are related to the two-parameters model (m and k). We developed a further analysis to test if both parameters are necessary. Table 3 reports the mean, minimum and maximum variance accounted for by a model, which have only the k parameter. It can be seen that, above all in the heavy and flat versions, the values are noticeably smaller. One performance has a negative value, which implies that in this case the model can't be apply. The second parameter allows a mean improvement of about 5%, with a maximum of 37%. This comments suggest that the two parameters are both necessary.

	na	br	da	ha	so	he	li	pa	fl
Mean	74	69	66	58	63	39	72	67	47
Min	60	53	52	40	52	-4	62	54	36
Max	88	90	85	82	78	64	85	75	62

Tab. 3: variance accounted for by the model with one parameter (values are expressed as perceptual of the total variance).

Another test of the model can be obtained by the production of computer performances, which key-velocity profiles are the estimated $\tilde{y}_e(n)$. So we can compare the original and computer generated performances and draw important observations about the model validity (analysis-by-synthesis method). The computer generated performances show that the model can well reproduce the global expressive characteristics of the original performance. In particular the expressive intentions, which characterize the original performance, are clearly recognizable. Some local characteristics, however, are not very well reproduced by the model. This observation can be set in a hierarchical view of the musical discourse [8]: the followed approach can catch expressive characteristics as far as phrase level, but not lower. Model's goal is not a complete treatment of musical interpretation, but a study of the general performance strategies of musicians. The model, however, can be used as a good basis in order to study and apply other models, which can catch more local characteristics [9].

Now we will try to clarify the sense and the use of the DPS. Some outcome can arise by the definition of the input profile (we chose the average profile). By means of simple calculations, it can be showed that the average profile is the geometric center of gravity of the performances mapped in the DPS. The numeric values in the DPS, therefore, can not be considered in an absolute sense, but they are relative to their center of gravity, i.e. their reciprocal position. For instance, we can say that the mean key-velocity difference between the light versions ($k \cong 0.8$) and the heavy versions ($k \cong 1.2$) is about 40%.

It is interesting to find out if the DPS can be as well used in an inverse way. That is, we want to verify if the DPS can suggest how whatever input profile have to be changed in order to communicate a certain expressive intention (e.g. harder, softer, etc.). The verification was obtained by means of analysis-by-synthesis method, using both K545 Sonata and other piano scores. First, we need a human performance of the score, by which the input profile can be drawn. Then we chose a point of the space that correspond to a certain expressive intention and his coordinates (m and k) are used as parameters of the equation 1. We did it for all the adjectives and we obtained performances that reflect, in a relative sense, the chosen expressive intentions.

The DPS was obtained (see above) using a set of 45 performances, so represent a kind of sampling of the space. What do intermediate points of the space mean? We hypothesize that they can be used as an interpolation of the original samples: i.e. the points between heavy and light versions would have intermediate expressive characteristics. Analysis-by-synthesis method was applied choosing intermediate points of the space: the computer-generated performances have intermediate characteristics and show that all the points of DPS have an expressive meaning. These results imply that DPS can be used in order to render a kind of morphing between expressive characteristics. Generally, during the same performance, a trajectory that moves from a region to another one of the DPS can be drawn. The parameters, in that case, are functions of time and the performance will be characterized by changeable expressive features.

5 Conclusions

Starting from piano performance analysis, a linear model of dynamics variations depending on expressive intentions was developed. This model can be applied both to performance analysis and to the field of automatic performance. In particular, it is possible to draw trajectories in the DPS, which allow to control continuously the dynamics characteristics of the computer-generated performances. Analysis-by-synthesis approach showed that a linear model could properly render expressive characteristics and the defined parameters are suitable to describe different performances of the same score.

Acknowledgment

This research was supported by Telecom Italia, under a research contract "Cantieri Multimediali".

References

[1] Canazza, S., De Poli, G., Roda', A., & Vidolin, A. (1997). Analysis and synthesis of expressive intentions in musical performance. In *Proceedings of the International Computer Music Conference 1997* (pp. 113-120). Tessaloniki: International Computer Music Association.

[2] Palmer, C. (1996). Anatomy of a performance: sources of musical expression. *Music Perception*, 13(3), 433-453.

[3] Todd, N. P. McA. (1992). The dynamic of dynamics: a model of musical expression. *Journal of Acoustical Society of America*, 91(6), 3540-3550.

[4] Friberg, A. (1991). Generative Rules for music performance: a formal description of a rule system. *Computer Music Journal*, 15(2), 56-71.

[5] Repp, B. H. (1990). Patterns of expressive timing in performances of a Beethoven minuet by nineteen famous pianists. *Journal of Acoustical Society of America*, 88(2), 622-641.

[6] De Poli, G., Rodà, A., & Vidolin, A. (1998). Note-by-note analysis of the influence of expressive intentions and musical structure in violin performance. Journal of New Music Research, Special Issue 1998.

[7] Repp, B. H. (1992). Diversity and commonality in music performance: An analysis of timing microstructure in Schumann's Träumerei. *Journal of Acoustic Society of America*, 92(5), 2546-2568.

[8] Lerdahl, F. & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge: The MIT Press.

[9] Windsor, W. L. & Clarke, E. F. (1997). Expressive timing and dynamics in real and artificial musical performances: using an algorithm as an analytical tool. *Music Perception*, 15(2), 127-152.