A DYNAMIC ANALOGY BETWEEN INTEGRO-DIFFERENTIAL OPERATORS AND MUSICAL EXPRESSIVENESS

Giovanni De Poli, Antonio Rodà
Dept. Information Engineering
Padova University, Italy
giovanni.depoli@dei.unipd.it

Luca Mion
TasLab – Informatica Trentina
Trento, Italy
luca.mion@infotn.it

ABSTRACT

Music is often related to mathematics. Since Pythagoras, the focus is mainly on the relational and structural aspects of pitches described by arithmetic or geometric theories, and on the sound production and propagation described by differential equation, Fourier analysis and computer algorithms. However, music is not only score or sound; it conveys emotional and affective content. The aim of this paper is to explore a possible association between musical expressiveness and basic physical phenomena described by integro-differential operators.

1 INTRODUCTION

Intuitive awareness of the relationship between music and mathematics exists as early as Pythagoras, who made an attempt to investigate and to determine this relation through acoustics. Pythagoreans founded the quadrivium, the fourfold way of knowledge. They divided mathematical science in two parts: *how many* (i.e., discrete, quantity) and *how much* (i.e., continuous, magnitude). Each of these parts is further subdivided into either absolute (stable) or relative (in motion). Thus, the quadrivium consisted of arithmetic (discrete quantity which subsist in itself), music (discrete quantity which is related to another), geometry (continued magnitude immovable) and astronomy (continuous magnitude of self-motive nature). Since then, the link between music and mathematics has been very fruitful in defining harmony and consonance, scales and temperament (i.e., relation among pitches), or in describing the musical structures as symmetries, transpositions, etc. (i.e., a geometrical–spatial metaphor of melodic structure).

During the late sixteenth and early seventeenth centuries, music began to be recognised more as an art and to be treated as a language and analysed in expressive terms. During the same period, science was moving from theoretical to experimental. Attention to physical phenomena lead to the development of new mathematical theories to explain *change*, such as calculus and differential equations. Science started to investigate in mathematical terms the physical nature of music (i.e., sound production and acoustics). In the second half of the ninetieth century Helmholtz set the basis of scientific investigation of sound perception (i.e., psychoacoustics). If we look at the content of books on Music and Mathematics (e.g., [1]), we notice that the content includes mainly the relational and structural aspects of pitches, described using arithmetic or geometric theories, and sound production and propagation, described by differential equation, Fourier analysis and computer algorithms. Moreover recently music performance, which acts as mediation between composer and listener, is increasingly being studied in scientific terms, and mathematical models are being developed. The pioneering model [7] of musical expressiveness in music performance, based on kinematics, is particularly relevent for the aims of this paper. ,

However music is not only score or sound; it can convey also sensorial and/or affective contents. A piece of music can suggest a light or a heavy sensation, an happy or a sad emotion. Obviously, the fact that a listeners judges a piece of music as light does not mean that this is the content of that music. In fact, sensorial or affective adjectives can be considered as metaphors, by means of which a listener translates his physical and cognitive experience of music [4]. The aim of this paper is to start to explore if and how the aspects of musical expressiveness related to sensorial or affective experiences can be associated to basic physical phenomena described by integro-differential operators.

2 MATHEMATICS AND MUSIC EXPRESSIVENESS

In mathematics, an *operator* is a function which operates on (or modifies) another function. Often, an operator is a function acting on functions to produce other functions. Given a function $f(t)$, the simplest operators in mathematical analysis are the *differential* operator $\mathcal{D}$, which defines the derivative of a function,

$$\mathcal{D}f(t) = \frac{d}{dt}f(t), \tag{1}$$

SMC 2009, July 23-25, Porto, Portugal

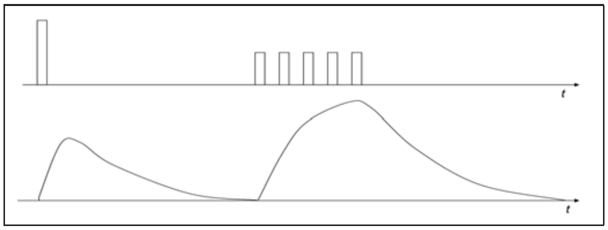

Figure 1. Analogy between time evolution of emotion and a ringing bell. Top: The pattern of strikes and intensity applied to a bell. Bottom: Intensity of the sound produced by the bell.

the *integral* operator $\mathcal{I}_0$, which defines the integral of a function with initial point $t = 0$

$$\mathcal{I}_0 f(t) = \int_0^t f(\tau) d\tau, \quad (2)$$

and the *proportional* operator, which scales the function by a constant k

$$\mathcal{K} f(t) = k f(t). \quad (3)$$

In the field of affective computing, Picard [6] outlines a signal representation for emotions and moods (emotional states lasting more than a couple of minutes) by a linear system preceded by a non-linearity. Picard uses the analogy of a ringing bell to illustrate the time course of emotion: both the bell sound and the emotional response, when triggered, have a fast rise time followed by a more gradual decay. The response sums if re-triggering occurs before the signal has completely decayed away. The intensity of the bell sound – and by analogy the intensity that a person might assign to their felt emotion – are shown in Fig. 1, when triggered by a single or repeated excitations. The influence of the person temperament is taken into account by the parameters of the linear model. Saturation and threshold effects are modelled by a smooth (sigmoid) non linear function applied to the inputs of the emotional system. The parameters of the function can be set according to the personality, mood, cognitive expectation, activation and arousal level of a person.

This analogy between time evolution of sound intensity and of emotion intensity suggested us to look for a possible generalization, in order to derive an analogy between integro-differential operators and musical expressiveness. In fact, the idea of using linear systems to model affective behaviour can be further developed. The basic building blocks of linear systems are the proportional, integral and derivative transformations, and they are described respectively by

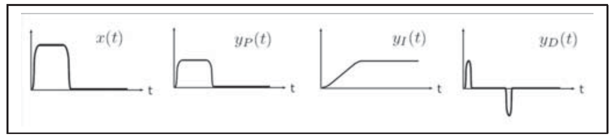

Figure 2. Functional transformation produced by proportional, integral and differential operators to a smoothed and large pulse $x(t)$ as input function.

the following equations:

$$y_P(t) = \mathcal{K}x(t) = k_P x(t) \quad (4)$$

$$y_I(t) = k_I \mathcal{I}_0 x(t) = k_I \int_0^t x(\tau) d\tau + y_I(0) \quad (5)$$

$$y_D(t) = k_D \mathcal{D} x(t) = k_D \frac{d}{dt} x(t) \quad (6)$$

where $x(t)$ is the input and $y_P(t)$, $y_I(t)$, $y_D(t)$ are the outputs of the basic blocks. Fig. 2 shows the behaviour of the proportional, integral and differential operators to a smoothed and large pulse $x(t)$ as input function.

3 EXPERIMENT

In order to verify if a metaphor based on $\mathcal{KID}$ operators can appropriately describe some expressive characteristics of music content, we conducted an experiment which investigates subject's associations between two sets of musical stimuli and three haptic attractors, that we assumed to be representatives of the three components of the $\mathcal{KID}$ metaphor.

3.1 Experiment design

Procedure. Participants were asked to listen to each musical excerpt and to associate it to one of the three attractors. Participants were allowed to listen to the excerpts and to test the attractors as many time as wished, and to change their responses until they were satisfied by their choices. The attractors induced the listeners to organize the musical excerpts on the basis of similarity criteria which depend on the characteristics of the attractors. Therefore, we can expect that different set of attractors will induce a different mental organization of musical excerpts. On the other hand, the features of the set of music excerpts can influence the results as well.

Materials. Two sets of musical stimuli were used in the experiment setup. The first set of musical excerpts comprise a subset of the performances of professional instrument players used in [5]: the theme from Händel's Sonata HWV 379 in E Minor Op. 1 No. 1 (Adagio) and the traditional song Twinkle Twinkle Little Star were played by flute and violin (Fig. 3). Each melody were played more times,

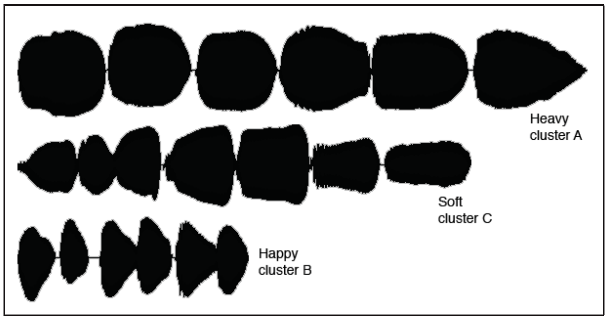

Figure 3. Simple musical stimuli belonging to different clusters: The first six notes of Twinkle Twinkle Little Star played by violin with respectively Heavy, Soft and Happy expressive intention.

with different expressive nuances, in order to convey expressive intentions Happy, Sad, Angry and Calm (affective space), Light, Heavy, Soft and Hard (sensorial space [3]) plus a Neutral performance. In total we took into account 2 (instruments) x 2 (pieces) x 9 (adjectives) = 36 examples with an average duration of 30s. These excerpts have a plain musical structure constituted by a tonal melody played by a single instrument, so we will refer to them as *simple musical stimuli*. They are oriented to study expressive aspects related to performance and they have already been analyzed by an acoustic point of view [5]. Moreover, the labels chosen to characterize the different performances allow a comparison between affective and sensorial space.

The second set is constituted by the data used in [2] to study the emotion communication in music: 27 musical stimuli extracted by recordings belonging to the Western music repertoire of the classic-romantic period. As these excerpts have a polyphonic structure with two or more instruments playing together, we will refer to them as *complex musical stimuli*. By means of a perceptual test Bigand and coworkers [2] found that these musical stimuli are organized along two dimensions, which were interpreted as associated to the emotional properties of valence and arousal, and are arranged in four groups, characterized by different levels of valence and arousal: a cluster with high arousal and high valence (HAHV); a cluster with high arousal and low valence (HALV); a cluster with low arousal and high valence (LAHV); a cluster with low arousal and low valence (LALV). These clusters roughly correspond to Happy, Angry, Calm and Sad adjectives respectively.

The set of attractors is composed by three haptic stimuli synthesized by means of a Phantom Omni haptic device [1], which simulates the basic effect of a mechanical mass–spring–damper system.

All the force feedback are omni-directional: the device reacts to the user's input in every points of the haptic sphere. Regarding the stimulus E (*elasticity*), the device generates a force feedback with intensity

$$f_E(t) = -K_{el} \cdot ||s(t) - s_0|| \tag{7}$$

where s_0 is the center of the haptic sphere, $s(t)$ is the position of the stylus at the instant t and K_{el} is the elasticity constant of the system. The stimulus F (*friction*) is characterized by a force feedback proportional to the velocity of the user's movement:

$$f_F(t) = -\eta_v \cdot v(t) \tag{8}$$

where $v(t)$ is the velocity of the stylus and η_v is the viscosity constant of the system. Finally, the stimulus I (*inertia*) simulates the interaction with an inertial mass m, moving in a field free of other (gravitas or magnetic) forces. The mass m is coupled to the stick, that we assume to have a negligible inertial mass. The intensity of the force follow the equation:

$$f_I(t) = -m \cdot a(t) \tag{9}$$

where $a(t)$ is the stylus acceleration. After several tests, we set $m = 0.5$ Kg, $K_{el} = 510$ N/m, and $\eta_v = 31.9$ Ns/m.

When using the dynamic analogy, force $f(t)$ is often subjectively considered as the cause and movement (velocity $v(t)$) as the effect. Thus, we are induced to associate Elastic attractor to $\mathcal{D}$ operator, Friction to $\mathcal{K}$ and Inertia to $\mathcal{I}_0$. This association constitutes the basis of our dynamic analogy.

A preliminary experiment was carried out with the aim of testing if the three haptic stimuli were perceived as different by the subjects. We presented in a random order nine stimuli, three equal stimuli for each type of feedback. Subjects were asked to group the stimuli according to their similarity. All the 20 subjects correctly grouped the stimuli in three categories, each one comprehending the three equal stimuli.

Apparatus. The sound files were represented on the computer screen by a visual interface implemented using the real-time sound synthesis environment PD (Pure Data). The interface consists on 3 buttons displayed on the top, associated to the 3 attractors, and on a set of buttons listed in column associated to the musical stimuli which are presented (in random order) to the participants, who were allowed to listen to the excerpts and to the attractors as many time as wished just by pressing the related button. Each of the buttons was also associated to a radio button where the participants could select one attractor only, by employing a three-alternative forced-choice (3AFC) method.

Participants. A total of 39 subjects participated to the experiment. Of these, 18 subjects (13 male, 5 female) were presented to the simple musical stimuli: 4 subjects had a professional musical training for 5 years at least and were referred as musicians (M); 4 subjects play an instrument and they are referred as amateurs (A); 10 subjects did not have any musical training and were referred as non-musicians

[1] http://www.sensable.com/haptic-phantom-omni.htm

stimuli	Friction	Elasticity	Inertia	resulting clusters
Angry	41	25	6	
Hard	40	17	15	A
Heavy	31	9	32	
Happy	10	60	2	B
Light	15	47	10	
Calm	13	5	54	
Sad	11	3	58	C
Soft	21	18	33	
Neutral	28	18	26	

Table 1. Contingency table of the subjects' responses in the experiment with the simple musical stimuli and resulting clusters.

(N). Participants were aged from 23 to 34 years (26 years average). The duration of the test was about 15 minutes. The other 21 subjects (7 male, 14 female) were presented to the complex musical stimuli: 3 subjects had a professional musical training for 5 years at least and were referred as musicians (M); 7 subjects play an instrument and they are referred as amateurs (A); 11 subjects did not have any musical training and were referred as non-musicians (N). Participants were aged from 22 to 53 years (30 years average). The duration of the test was about 20 minutes.

4 RESULTS

Simple musical stimuli. Subjects' responses were summarized into a two-way contingency table containing 9 rows (expressive intentions) and 3 columns (attractors Friction, Elasticity, and Inertia). Each cell of the table recorded the number of times (observed frequency) that each expressive intention was associated to an attractor (Tab. 1). Pearson's χ^2 test was used to compare observed frequencies with expected frequencies under the null hypothesis of independence, yielding a total value of $\chi^2 = 305.96$ ($df = 16$, $p < 0.001$), denoting strong evidence of a relation between expressive intentions and attractors. The analysis of each row, taken individually, shows that the association among expressive intentions and attractors is confirmed by Angry, Hard, Happy, Light, Sad, Calm ($\chi^2 > 21.9$, $df = 1$, $p < 0.001$), and weakly confirmed by Heavy ($\chi^2 = 5.0$, $df = 1$, $p < 0.05$). On the contrary, no significant relation among expressive intention and attractors has been found for Neutral and Soft.

We conducted two analysis to investigate the association between expressive intentions and attractors: a Simple Correspondence Analysis and a K-means clustering. The contingency table was submitted to Simple Correspondence Analysis in order to graphically represent the degree of association between expressive intentions and attractors according to their χ^2 distances. Since Simple Correspondence Analysis is applied on a 3-columns table (i.e., represented

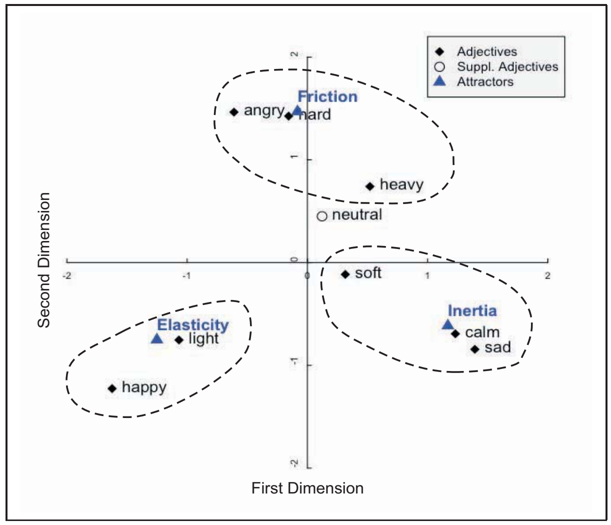

Figure 4. Bi-plot of the correspondence analysis on the experiment with simple musical stimuli.

by two degrees of freedom), we can derive two dimensions only (Fig. 4) with eigenvalues covering the total variance (first eigenvalue explains the 75.52% of the total inertia). The Neutral expression was considered as supplementary row (not used to perform the previous analysis) and its projection into the correspondence plot resulted close to the origin of the axis. We can see in Fig. 4 that the expressive intentions Angry-Hard-Heavy, Happy-Light, Calm-Sad-Soft are depicted close to attractors Friction, Elasticity and Inertia respectively. Then, we applied the K-means clustering (number of groups = 3) to the coordinates of the points in the Correspondence Plot in order to identify the stable groups. We did not take into account the Neutral intention. Three stable groups (see dashed lines in Fig. 4) were identified corresponding to the clusters: (A) Angry-Hard-Heavy, (B) Happy-Light and (C) Calm-Sad-Soft (see last column of Tab. 1). By grouping these expressive intentions according their cluster membership we found strong relation among clusters and attractors ($\chi^2 = 253.0$, $df = 4$, $p < 0.001$). Moreover, significant relation has been found between A cluster and F attractor, B cluster and E attractor and C cluster and I attractor ($\chi^2 > 65.5$, $df = 1$, $p < 0.001$).

It can be noticed that these clusters are consistent with the results of [5], where the same groups were obtained by clustering the same musical excerpts on the basis of their significant musical/acoustic features. It was found that cluster A is mainly related to Sound Pressure Level and some spectral cues, cluster B to Note Per Second and C to Attack time. These associations are confirmed by the canonical correspondence analysis on the data of our experiment.

Complex musical stimuli. Table 2 shows the subjects' responses with rows representing the 27 musical excerpts

clusters [2]	stimuli	Friction	Elasticity	Inertia
HALV	B12	7	13	1
	B16	9	8	4
	B17	4	15	2
	B18	9	8	4
	B25	9	5	7
	B26	12	5	4
	B27	5	8	8
HAHV	B10	9	9	3
	B11	12	3	6
	B13	9	12	0
	B14	7	12	2
	B15	6	14	1
	B22	7	11	3
	B23	7	11	3
	B24	6	13	2
LALV	B3	3	3	15
	B7	3	3	15
	B8	3	2	16
	B9	3	5	13
LAHV	B1	3	3	15
	B2	8	8	5
	B4	9	1	11
	B5	8	1	12
	B6	4	1	16
	B19	10	4	7
	B20	9	3	9
	B21	7	3	11

Table 2. Contingency table of the subjects' responses in the experiment with the complex musical stimuli.

and columns representing the 3 haptic attractors. The Pearson's χ^2 test denotes a strong relation between musical excerpts and attractors ($\chi^2 = 205.2, df = 52, p < 0.001$) and confirmed that subjects are able to distinguish the different haptic attractors and to use them to classify the musical excerpts.

The contingency table was submitted to Simple Correspondence Analysis in order to graphically represent the degree of association between musical stimuli and attractors (see Fig. 5). Then, we proceeded with a K-means analysis in order to identify clusters of stimuli. After several trials, we set the number of groups both to 3 and to 5 (respectively continuous and dashed lines in Fig. 5). Three of the five clusters include those stimuli which were associated to one of the haptic attractors. The other two clusters are composed by stimuli that subjects associated equally to two attractors: Elasticity - Friction and Inertia - Friction.

5 DISCUSSION

The results of the experiment with simple musical stimuli support a strong relation between the cluster Hard-Heavy-Angry and the Friction attractor, between the cluster Light-Happy and the Elasticity attractor, and between the cluster Sad-Calm-Soft and the Inertia attractor. Although from a semantic point of view these associations are not surprising (when I am angry I feel friction with someone; when I am happy I jump of the joy; when I am calm I move slow as an object with an high inertia), what is interesting is that the subjects were able to make this associations without any explicit semantic mediation, directly associating a musical stimulus to a haptic one.

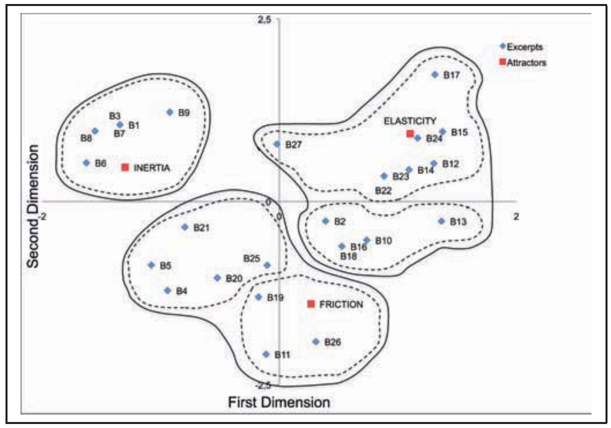

Figure 5. Correspondence analysis on experiment with complex musical stimuli.

In general, the subjects were able to consistently recognize common characteristics between musical stimuli and haptic attractors. Concerning the single expressive intentions, Neutral intention was not recognized as related to one single attractor, but the contingency table (Tab. 1) shows a balanced contribution of all the three attractors, as we could expect due to its meaning. It is interesting to note that, in some cases, the scores in the contingency table (Tab. 1) suggest the idea the three attractors Friction, Elasticity, and Inertia constitute a sort of basic components; the various expressive nuances can be represented as a combination of these components. E.g., Heavy performance was perceived as related not only to Friction, but also to Inertia, as we could expect. Moreover we can notice that the emotional response to a stimulus, as shown in Fig. 1, has a low pass characteristics: i.e. it can be modelled by a combination of $\mathcal{K}$ and $\mathcal{I}_0$ operators. This fact is in agreement with the projection of the Neutral expression in Fig. 4, which tends to be located between Friction and Inertia attractors.

The results of the experiment with the complex musical stimuli support a relation between the HAHV (Happy) cluster and the Elasticity attractor (confirmed by all the excerpts except for B10 and B11), and between the LALV (Sad) cluster and the Inertia attractor. On other cases, subjects responses are divided between two attractors: e.g., excerpts B4, B5, B20, and B21 are associated both to Inertia and Friction. This observation is coherent with the Fig. 5, where two clusters are composed by stimuli that subjects associated to two attractors: Elasticity-Friction and Inertia-Friction.

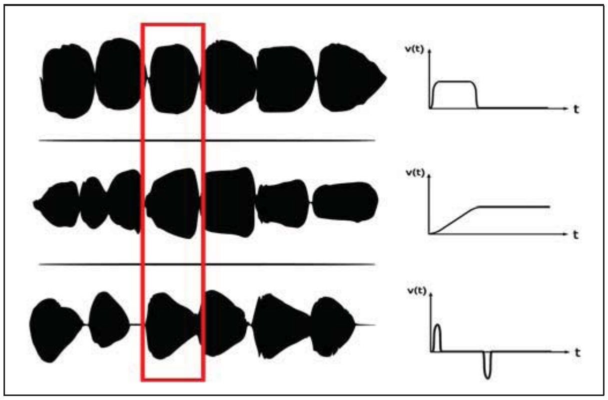

Figure 6. Comparison of sound envelope of expressive performances (Heavy, Soft and Happy) of Fig. 3 with the integro-differential operators $\mathcal{K}$, $\mathcal{I}_0$, $\mathcal{D}$ of Fig. 2.

With reference to our dynamic $\mathcal{KID}$ analogy, the cause-effect relation is represented by the mechanical admittance Y which mathematically describes the dynamic mapping and the qualitative behaviour from force $f(t)$ to velocity $v(t)$ by a linear combination of $\mathcal{K}$, $\mathcal{I}_0$, $\mathcal{D}$ operators or, physically, by a combination of friction, inertia and elasticity. We can distinguish resistive admittance which dissipates energy, from reactive impedance which stores energy. Ideal friction (operator $\mathcal{K}$) is a pure resistive admittance, while ideal inertia and elasticity (operators $\mathcal{I}_0$ and $\mathcal{D}$) are pure reactive admittances: in particular inertia stores kinematics energy and it opposes changes in movement, while elasticity stores potential energy and opposes changes in force.

In Fig. 6 we compare the sound intensity envelopes of 3 stimuli, belonging to the different clusters, with the velocity $v(t)$ resulting by applying the operator to a smoothed large force pulse $f(t)$ (see first graph in Fig. 2), or equivalently with the effect of using the force pulse as input to the corresponding haptic attractor. It can be seen that friction acts as a scaling factor of the input force and does not modify the shape of the input. The inertia (mass) tends to remain at its initial velocity (which is zero in the present example), then it grows progressively and remains constant when the input stops; the mass progressively augments its kinetics energy. The elasticity (spring) instead reacts immediately to the input variations; it stores potential energy which is used to oppose to force changes. From this qualitative description of the dynamic behaviour of the three basic elements we are induced to confirm the association of operator $\mathcal{K}$ to the cluster Hard/Heavy/Angry, $\mathcal{I}_0$ to cluster Sad/Calm/Soft, and $\mathcal{D}$ to cluster Light/Happy. Friction, elasticity and inertia are the basic properties of ideal mechanical systems and the dynamic behaviour of each real system depends by a weighted combination of friction and elasticity or of friction and inertia, where friction represents the quantitative aspect of the dynamical behaviour and elasticity/inertia represent the qualitative aspect of dynamical behaviour. The quantitative aspect can be associated to the vertical axis of Fig. 4 and 5, while the qualitative aspect can be associate to the horizontal axis.

Comparing these results we can notice that, although subjects are able to recognize the different haptic feedback, the $\mathcal{KID}$ metaphor seems to be more suitable for representing expressive cues in simple musical excerpts (where the expressive content is mainly related to performance cues) than in complex musical stimuli (where musical structure is more relevant). This result can be explained by the fact that music performance is more related to action–based aspects, whereas musical structure can involve aspects related to cognitive and/or cultural factors. Moreover, “real” music pieces usually are not characterized by a single expressive intention, but rather by a mixture of expressive nuances.

6 CONCLUSIONS

We proposed a dynamic analogy based on proportional operator $\mathcal{K}$, integral operator $\mathcal{I}_0$ and differential operator $\mathcal{D}$, and we carried out an experiment with the aim of investigating the relation between the $\mathcal{KID}$ metaphor and expressive music contents. The results let us hypothesize that relevant expressive characteristics of music can be associated to a weighted combination of quantitative and qualitative basic components described respectively by the $\mathcal{K}$ operator and by the opposition of $\mathcal{I}_0$ vs. $\mathcal{D}$ operators.

References

[1] D. Benson. *Music: A Mathematical Offering*. Cambridge University Press, 2006.

[2] E. Bigand, S. Vieillard, F. Madurell, J. Marozeau, and A. Dacquet. Multidimensional scaling of emotional responses to music: The effect of musical expertise and of the duration of the excerpts. *Cognition and Emotion*, 19(8):1113–1139, 2005.

[3] S. Canazza, G. De Poli, A. Rodà, and A. Vidolin. An abstract control space for communication of sensory expressive intentions in music performance. *Journal of the New Music Research*, 32(3):281–294, 2003.

[4] M. Leman. *Embodied Music Cognition and Mediation Technology*. The MIT Press, 2007.

[5] L. Mion and G. De Poli. Score-independent audio features for description of music expression. *IEEE Transactions on Speech, Audio and Language Processing*, 16(2):458–466, 2008.

[6] R. W: Picard. *Affective computing*. MIT Press, 1997.

[7] N. P. Todd. The kinematics of musical expression. *Journal of the Acoustical Society of America*, 97(3):1940–1949, 1995.