DISSEMINATION LEVEL: PUBLIC

Social Interaction and Entrainment using Music PeRformancE

SIEMPRE

D1.4 Models and algorithms for analysis of creative social interaction (ver.2)

Version	*Edited by*	*Changes*
1	UNIGE (Camurri,Glowinski, Gnecco)	
2	QUB (Jaimovich)	
3	IIT (Badino, d'Ausilio)	
4	UPF (Papiotis)	

Table of Contents

1. INTRODUCTION

This deliverable is an update of the deliverable D1.3, which describes the first versions of models and algorithms for analysis of creative social interaction.
A main objective of this project is the development of computational models and algorithms for the analysis of the social signals the project addresses, e.g., synchronization and leadership. However, models and algorithms are often discussed together with other issues, e.g., experimental set-up, feature extraction. Indeed, the modeling and algorithmic issues are often highly technical, and in many kinds of report it is likely to be seen as detail that should be put to one side in order to convey the broad picture. The aim of this deliverable is to provide a place where research on models, algorithms, and techniques can be properly recorded and considered without cluttering reports that focus on other levels.

2. TECHNIQUES FOR ANALYSIS OF STRING QUARTET

2.1 Computational analysis of interdependence

This section describes the procedure we have followed to measure the difference in the behavior between two performance conditions of the string quartet. Such performance conditions were characterized by different interpretative conditions: in the first condition (A), musicians play the music piece in normal condition, like in a concert. In the second condition (B), the first violin modified his usual interpretation by adding rhythmic and dynamic changes unexpected to the other musicians (e.g., playing *forte* where the written agogics is *piano*, speeding up when a *rallentando* is requested). The other members of the quartet were not aware of these new versions before playing.
In a string quartet, the musicians influence each other's performance in order to achieve a common musical goal, not only in terms of timing but also in other dimensions of the performance such as dynamics, intonation, and timbre. In this sense, joint performance in an ensemble goes beyond the concept of strict temporal synchronization and closer to a concept of multi-dimensional interdependence among musicians. In the following chapter we describe our methodology on analyzing interdependence in a string quartet regarding intonation and dynamics, and the obtained results.

2.1.1 Data acquisition

During the recordings we simultaneously capture two types of data: the sound produced by the musicians via piezoelectric pickups and conventional microphones, as well as the sound-producing gestures of each musician via a wired motion tracking system. Although only audio features are used in this particular study, the low-level features extracted from the motion capture data are used in the data post-processing steps (namely the score-to-performance alignment).
Both audio and motion capture data are acquired simultaneously and synchronized in real time using a master clock generator as well as a linear timecode audio signal. Individual audio for each musician is captured through the use of piezoelectric pickups attached to the bridge of the instrument. The overall sound of the ensemble is captured using a cardioid medium diaphragm condenser microphone, as well as a binaural stereo recording dummy head; the binaural stereo recording is then used to set the gain of each individual pickup signal so that the audio level of each instrument corresponds to the overall acoustic result. For each individual pickup signal, we extract two audio features: the fundamental

frequency (F0) as an estimation of pitch (using the YIN [1] algorithm), and the root mean square audio energy as an estimation of audio intensity.
Besides the audio signals, instrumental - i.e. sound-producing - gestures are also acquired through the use of a wired MOCAP system, as detailed in [2]. By analyzing the raw MOCAP data, a set of bowing features is extracted; such features include bow transversal velocity, bow pressing force, bow-bridge distance, bow tilt & inclination, et cetera.

2.1.2 Data processing & feature extraction

Score alignment
For every recording, a semi-automatic alignment between the performance and the music score is performed using a dynamic programming approach, a variation of the well-known Viterbi algorithm. This approach focuses into three main regions of each note: the note body and two transition segments (onset and offset). Different costs are computed for each segment, using features extracted by the audio as well as the bowing features. Finally, the optimal note segmentation is obtained so that a total cost (computed as the sum of the costs corresponding to the complete sequence of note segments) is minimized. This method, which can be seen in more detail in [2], has so far provided robust results that only in few occasions require manual correction.
Extraction of intonation adjustments
As a descriptor of intonation, the pitch contour of each musician's individual audio recording is obtained by applying a logarithmic transformation to the extracted fundamental frequency (akin to pitch cent conversion). In order to extract the adjustments of each musician's intonation, we consider the aligned score as 'reference pitch', i.e. perfect, non-adjusted intonation according to the equal temperament system; this representation of the performance-aligned score is then subtracted from the obtained pitch contour (or the warped pitch contour for the solo condition). This final feature is our estimation of intonation adjustment, with positive values indicating a note played slightly sharper and negative values indicating a note played slightly flatter than equal tempered intonation, respectively.
Estimation of dynamics intensity
For each musician's individual audio recording, the Root Mean Square (RMS) Energy of the audio signal is computed in the time domain over a sliding window of 40 milliseconds. The values of RMS energy are converted to a logarithmic scale, in order to obtain an estimation of loudness that is closer to human auditory perception (as per Fechner's law). Finally, the log-RMS energy of the signal is smoothed using a median filter with a window of 300 milliseconds.

2.1.3 Interdependence measures

Four methods for measuring interdependence have been evaluated and will be briefly presented; for an excellent literature review on these methods we redirect the reader to [3]. Since the methods we introduce originate from different research fields, they vary in their complexity and required background knowledge. For that reason, we do not include a full mathematical formulation for each method; we refer the reader to the original publications where each method has been introduced for a complete description. We opted to divide the methods based on whether they are most suited for dependences of a linear or nonlinear nature. For each category, there is one symmetric and one directional method; the difference between the two being that a directional method can also assess the direction of influence between two interacting timeseries besides the strength of the interdependence.
Linear methods
Pearson product-moment correlation coefficient. This is the most common method utilized for quantifying the linear dependence between timeseries x and y; its output $p_{x,y}$ between timeseries x and y ranges from $p_{x,y} = -1$, i.e. complete linear inverse correlation between timeseries x and y, to $p_{x,y} = 1$, i.e. complete linear direct correlation between timeseries x and y, with a value $p_{x,y} = 0$ suggesting an

absence of linear dependence between timeseries x and y. Since linear correlation is calculated pairwise and is symmetric, we calculate the correlation coefficient for each one of the 6 possible pairs between the four musicians.

Granger causality. In studying the relationship between variables, it is often useful to assess the directionality of that relationship; besides the overall degree of interdependence, it is also important to draw conclusions about the direction of influence, i.e. whether variable 1 is influencing variable 2 more than the opposite. A method that is capable of giving such an estimate is Granger causality, a statistical concept that was first applied to Econometrics (see [4]) and more recently to Neuroscience. It poses the hypothesis that if timeseries x causes timeseries y, then past values of x should significantly help in predicting future values of y as opposed to simply using past values of y to predict its own future; this is assessed through the use of a multivariate vector autoregressive modeling (or linear regression analysis, depending on the approach). In this study, we used a freely available MATLAB toolbox by [5]. It must be noted that there exist approaches that modify Granger causality to incorporate nonlinear properties of signals; however, we have not yet tested them and are not included in this study. Since Granger causality is a directional measure, the connectivity analysis for four musicians yields 12 pairwise causality assessments $G_{x,y}$ (three causal pairs for each variable) as well as the total causal density $G_{density}$ of the ensemble, a bounded value between 0 and 1, with 0 signifying a complete lack of causal interactivity; we use the latter value as an estimation of the total amount of causal interactivity sustained by the network of musicians.

Nonlinear methods

Mutual Information. Mutual Information is a non-directional measure originating from the field of Information theory. It is not a nonlinear method per se, being based on the concept of entropy as proposed by Shannon in the 1950s and therefore dealing with reduction of information rather than the linearity of the data. For a pair of timeseries x and y, mutual information measures the difference between two types of joint entropy; the joint entropy of the two variables as measured from the data, and the joint entropy of the two variables as if they were independent. If they are indeed independent, $MI_{x,y}$ is zero. Otherwise, $MI_{x,y}$ is a positive, non-bounded value that represents the amount of information that one gains about x by knowing the outcome of y. Similar to the case of linear correlation, the analysis of a network of four musicians yields 6 pairwise values of Mutual Information. An average of these six values is used as an estimation of the overall Mutual Information-based interdependence of the quartet.

Nonlinear coupling coefficient. There exists a variety of nonlinear interdependence measures that quantify the signature of directional couplings between two timeseries x and y; it is assumed that the processes behind the timeseries are characterized by separate deterministic dynamics which both exhibit an independent self-sustained motion. Assuming the existence of directionality, i.e. the dynamics of one process driving the dynamics of the other, directional couplings can in principle be detected by quantifying the probability with which close states of the driven dynamics are mapped to close states of the driving dynamics. The state space of x (i.e. the set of all its possible states with each state corresponding to a unique point in the space of that set) is reconstructed using the method of state space embedding; then, a number of spatially nearest neighbors are selected for each point x_n in the state space (excluding temporal neighbors through the use of a given threshold). Finally, the squared mean Euclidean distance from the nearest k neighbors of x_n is calculated, along with the y-conditioned squared mean Euclidean distance (by replacing the nearest neighbors of x_n by the equal time partners of the closest neighbors of x_n). It has been shown that, when $x \rightarrow y$ coupling occurs, there is increased probability that close states of y are mapped to close states of x. Several available measures based on the above paradigm exist; of these we use the measure *L*, which was recently shown to be of higher sensitivity and specificity for directional couplings than previous approaches. The output of $L_{x,y}$ is a bounded value between 0 and 1, with 0 signifying a complete lack of interdependence. For a more in-depth explanation of this particular method as well as a proper mathematical formulation, we direct the

reader to [6] where the method was originally introduced. As is the case for Granger causality, 12 pairwise calculations of coupling $L_{x,y}$ are carried out. The average value of the set of 12 coupling values is used as an estimation of the overall Nonlinear coupling-based interdependence of the quartet.

2.1.4 Results

As mentioned above, we have applied the above methodology on two sets of experimental data, each one focusing on a different aspect of music performance. The first experiment is focused on Intonation, while the second one focuses on Dynamics. Both experiments are based on an exercise handbook for string quartets (Heimann, 1995), specifically designed to assist in improving the ensemble's capabilities for collaborative expression. Each exercise consists of short, simple musical tasks whose challenge lies in achieving overall synchrony rather than correctly performing one's individual task successfully. Depending on the exercise, the score can also have extra instructions/annotations on what is the specific goal that must be achieved by the quartet.
We record the musicians' performance in two experimental conditions: *solo* and *ensemble*. In the first condition (*solo*), each musician must perform their part alone without having access to the full ensemble score nor the instructions that accompany the exercise. In this way we wish to eliminate any type of external influence on the performance, be it restrictions imposed by other voices of the ensemble or instructions by the composer that are not in relation to the individual score of the performer. Following the solo recordings of each quartet member, the group of musicians is provided with the full ensemble score plus the composer instructions; they are then left to rehearse for a short period (~10 minutes) until they are able to fulfill the requirements of the exercise. Following the rehearsal, the quartet is recorded in the second experimental condition (*ensemble*) performing the exercise as a group.
In the case of exercises with annotations on the score, as in the Intonation exercise, we split the *ensemble* condition in two sub-conditions: *ensemble1*, where the quartet rehearses and performs the exercise without the annotations, and *ensemble2*, where the quartet rehearses the exercise again with the annotations added to the score and then records the final take.
Results from the Intonation experiment
Correlation values per musician pairs as well as average correlation values can be seen in Table 1:

Musician pair	**Pearson correlation, intonation**		
	ensemble2	**ensemble1**	**solo**
V1, V2	0.315	0.297	0.098
V1, VLA	-0.045	0.037	-0.350
V1, CLO	-0.210	-0.365	-0.086
V2, VLA	-0.212	0.021	0.090
V2, CLO	-0.421	-0.508	-0.010
VLA, CLO	-0.100	-0.145	-0.099
Average	**-0.112**	**-0.110**	**-0.059**

Table 1. Correlation coefficient of intonation adjustments per musician, for all three experimental conditions.
One can definitely observe similarities among the two ensemble set-ups as well as dissimilarity between the solo and ensemble set-ups; it can also be observed that the correlation values for the two ensemble set-ups is sporadically higher than the solo set-up. Nevertheless, this is not particularly consistent (such as the case of the violin1/viola pair in the solo condition); this keeps us still cautious against making an assumption regarding interdependence from the correlation values.

For the rest of the interdependence methods, we performed a sliding window analysis of interdependence; there are three main reasons behind this choice. First, as we are studying variables that change with time it is natural that the amount of interdependence will also vary; something which can potentially reveal the role of the musical score in the performance, as the collaborative task that must be jointly carried out by the musicians. Second, by windowing the signals by the average note length, we can reduce possible non-linearities and non-stationarities in our data, thus making the interdependence measures more reliable. Finally, we can deal with a smaller amount of data at a time, which removes the need to downsample our signals in order to cope with memory requirements. The analysis window was set equal to the average note length.
The mean value of interdependence across all analysis windows for each method following this analysis can be seen in Table 2 :

Method	**Interdependence strength, intonation**		
	ensemble 2	**ensemble 1**	**solo**
NC	0.334	0.307	0.166
MI	0.574	0.521	0.224
CD	0.034	0.025	0.024

Table 2. Nonlinear coupling (NC), Mutual Information (MI) and Causal Density (CD) of intonation adjustments, for all three experimental conditions.
It can be seen that the nonlinear methods are successful at separating the solo and two ensemble set-ups. We ran a 1-way ANOVA test to quantify the separation between the three set-ups: the results showed that both the Nonlinear coupling coefficient as well as Mutual Information could successfully separate the interdependence means of one set-up from the other at a confidence level larger than 95%, with the solo set- up showing significantly less interdependence than the two ensemble set-ups for both interdependence methods. One the other hand, Granger causality failed to provide significant separation for the three experimental set-ups; although we hypothesize that this is due to the nonlinearity of our data, it is necessary to investigate towards the validation of this hypothesis in future work.
Results from the Dynamics experiment
Correlation values per musician pairs as well as average correlation values can be seen in Table 3:

Musician pair	**Pearson correlation, dynamics exercise 1**		**Pearson correlation, dynamics exercise 2**	
	ensemble	**solo**	**ensemble**	**solo**
V1, V2	0.9267	0.8806	0.9427	0.9615
V1, VLA	0.9207	0.8719	0.9042	0.8895
V1, CLO	0.9291	0.8658	0.9081	0.8595
V2, VLA	0.9094	0.8880	0.9210	0.9346
V2, CLO	0.9002	0.8844	0.9247	0.9140
VLA, CLO	0.9345	0.9584	0.9212	0.9400
Average	**0.9237**	**0.8825**	**0.9211**	**0.9243**

Table 3. Correlation coefficient of log-RMS per musician, for both exercises and experimental conditions.

It can be seen by the above table that correlation is not capable of showing a significant difference between the ensemble and solo conditions. Besides that, it is clear that all of the studied features are very correlated.
For the rest of the interdependence methods, we performed a sliding window analysis of interdependence, in this particular case using a short (1 second) as well as a long (5 seconds) analysis window. The mean value of interdependence for each method across all analysis frames can be seen in Table 4:

Method	**Interdependence strength, dynamics exercise 1**		**Interdependence strength, dynamics exercise 2**	
	ensemble	**solo**	**ensemble**	**solo**
NC, short window	0.8276	0.7438	0.9333	0.8589
NC, long window	0.7441	0.6212	0.8257	0.6943
MI, short window	0.7149	0.6398	0.9123	0.7680
MI, long window	0.7801	0.7463	1.2372	0.9763
CD, short window	0.0046	0.0011	0.0063	0.0018
CD, long window	0.0059	0.0044	0.0041	0.0040

Table 4. Nonlinear coupling (NC), Mutual Information (MI) and Causal Density (CD) of log-RMS, for both exercises and experimental conditions.
It can be observed that the interdependence strength is consistently higher for all ensemble conditions – with some reservation for the case of Granger Causality where the separation between ensemble and solo is small. However, a 1-way ANOVA analysis of the interdependence strength of each analysis window failed to provide significant separation between the ensemble and solo conditions at a 95% confidence level. Another observation is that the larger analysis window provides better separation between the two experimental conditions for the nonlinear coupling coefficient; a logical outcome given the fact that the calculation of the nonlinear coupling coefficient relies on past values in order to assess the coupling strength.
It is evident by looking at the two last tables is that values of interdependence are generally high, compared to those obtained for the intonation analysis; an important difference between the two cases being that in the intonation case the analyzed feature is not pitch itself but pitch adjustments, which to a certain degree eliminates the effect of the score on the interdependence strength. For the current case of dynamics, merely the existence of synchronized crescendi as well as simultaneous changes in dynamics is bound to increase the overall interdependence strength, making significant separation between solo and ensemble difficult.
Separating the effect of the score is a feasible task for the case of intonation, since the 'reference pitch' of each note is already known from the score. In the case of dynamics however, there is no absolute value for dynamics intensity and therefore no objective reference through which 'dynamics adjustments' can be estimated and removed from our features.
In order to reduce the effect of the underlying score as much as possible, we employed a rough 'dynamics adjustments' extraction method which is outlined as follows: for every note in the musical score, we subtract a linear trend from the log-RMS feature; this way, the note-to-note changes in

dynamics are greatly reduced, making temporal fluctuations of dynamics within each note's boundaries much more prevalent. It must be noted however, that this affects the studied features in a biased way since the removed linear trend does not necessarily coincide with the 'reference' dynamics value for each note, which remains undefined. Mean interdependence values for the above scenario are shown in Table 5:

Method	**Interdependence strength, dynamics exercise 1 (score-independent)**		**Interdependence strength, dynamics exercise 2 (score-independent)**	
	ensemble	**solo**	**ensemble**	**solo**
NC	0.6427	0.4698	0.7302	0.5186
MI	0.3076	0.0896	0.6272	0.0288
CD	0.0209	0.0202	0.0159	0.0163

Table 5. Score-independent Nonlinear coupling (NC), Mutual Information (MI) and Causal Density (CD) of log-RMS, for both exercises and experimental conditions.

Two observations can be made from the above table. First, the overall strength of interdependence has been reduced. Second, the separation between ensemble and solo is much larger, again with the exception of Granger Causality. An 1-way ANOVA analysis of the interdependence strength for each note did show significant separation at a 95% confidence level, for the Nonlinear coupling as well as the Mutual Information interdependence methods.

The above can point to two main conclusions; first, that the musical score is indeed a very important factor in a musical ensemble's interdependence. Second, that the non-linear interdependence methods are not only capable at detecting higher interdependence strength for the ensemble condition as compared to the solo condition, but also at quantifying the overall strength of interdependence.

2.2 Extraction of an expressive visual group feature for the discrimination of two performance conditions by a string quartet

This section describes the procedure we have followed to measure the difference in the behavior between two performance conditions of the string quartet. Such performance conditions were characterized by different interpretative conditions: in the first condition (A), musicians play the music piece in normal condition (like a concert one). In the second condition (B), the first violin modified usual interpretation by adding rhythmic and dynamic changes unexpected to the other musicians (e.g., playing *forte* where the written agogics is *piano*, speeding up when a *rallentando* is requested). The other members of the quartet were not aware of these new versions before playing.

Condition A (normal) Condition B (unusual)

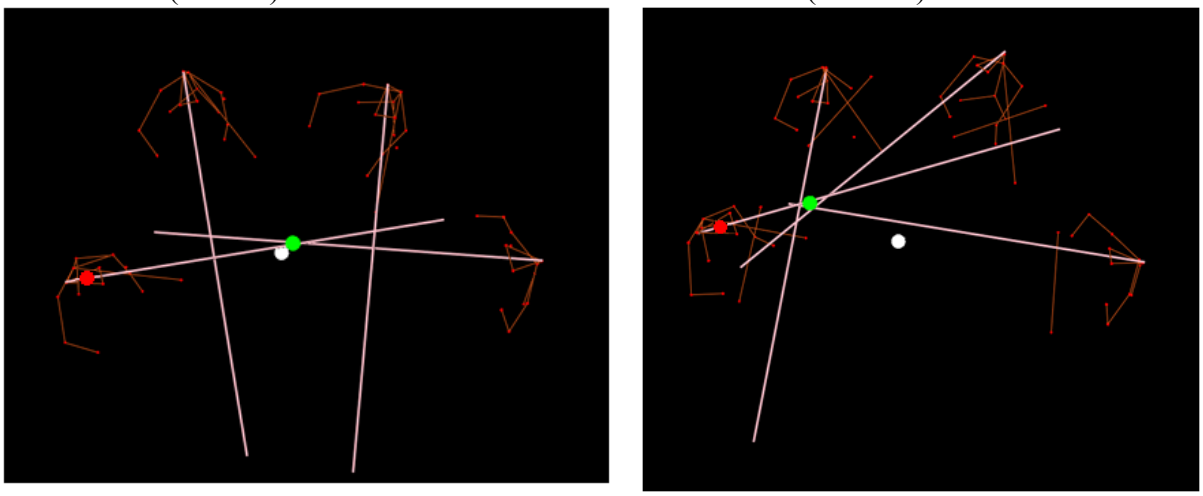

Figure shows a snapshot of the head's markers positions of the players when condition A (normal) and condition B (unusual) are tested, respectively. White half lines refer to the direction of head and green dot correspond to the point of convergence of the four half-lines. One can observe that in normal condition heads converge toward the string quartet "ear" (Glowinski et al. 2012) (white dot in the picture above), i.e. the subjective center of the string quartet defined by the musicians themselves. Whereas in the condition B (unusual), heads converge towards the first violinist (red dot)

The following is a description of the algorithm to determine the position of the "point of convergence" of the string quartet.

First, we numbered the music players from 1 to 4, so player 1, 2, 3, 4 denote, respectively, the 1st violin player, the 2nd violin player, the viola player, and the cello player.

For each frame k ($k = 1, \ldots, N_{frames}$) of a recording:

1. For each music player i ($i = 1, \ldots, 4$), the current position vector $\mathbf{p}_i^{(k)}$ in the horizontal plane of the barycenter of the head of the musician is computed as the mean of the position vectors of the three markers located on the head of the musician. Then, the current direction $\mathbf{d}_i^{(k)}$ projected onto the horizontal plane of the head of the musician is computed. Each $\mathbf{d}_i^{(k)}$ is defined as the unit vector connecting the barycenter of the head of the musician i to the point located in the middle of the line between the two other markers above his eyes.

2. For each music player i ($i = 1, \ldots, 4$), the half line $\mathbf{HL}_i^{(k)}$ starting from the point $\mathbf{p}_i^{(k)}$ and with direction $\mathbf{d}_i^{(k)}$ is computed, i.e., the set of all the points with position vectors
$$\mathbf{p}_i^{(k)} + t\,\mathbf{d}_i^{(k)},$$
where t is the maximum distance between any couple of musicians.

3. For each pair (i,j) of musicians ($i, j=1,\ldots4$, $i \neq j$), the position vector $\mathbf{c}_{i,j}^{(k)}$ of the intersection between the two half lines $\mathbf{HL}_i^{(k)}$ and $\mathbf{HL}_j^{(k)}$ is computed. n ($n \leq 6$) of pairwise intersections are computed: $\mathbf{c}_{1,2}^{(k)}$, $\mathbf{c}_{1,3}^{(k)}$, $\mathbf{c}_{1,4}^{(k)}$, $\mathbf{c}_{2,3}^{(k)}$, $\mathbf{c}_{2,4}^{(k)}$, $\mathbf{c}_{3,4}^{(k)}$.

4. If all 6 pairwise intersections exist, the position vector $\mathbf{c}^{(k)}$ of the point of convergence is computed as the mean of the position vectors $\mathbf{c}_{1,2}^{(k)}$, $\mathbf{c}_{1,3}^{(k)}$, $\mathbf{c}_{1,4}^{(k)}$, $\mathbf{c}_{2,3}^{(k)}$, $\mathbf{c}_{2,4}^{(k)}$, $\mathbf{c}_{3,4}^{(k)}$, i.e.,

$$\mathbf{c}^{(k)} = \frac{\mathbf{c}_{1,2}{}^{(k)} + \mathbf{c}_{1,3}{}^{(k)} + \mathbf{c}_{1,4}{}^{(k)} + \mathbf{c}_{2,3}{}^{(k)} + \mathbf{c}_{2,4}{}^{(k)} + \mathbf{c}_{3,4}{}^{(k)}}{6}.$$

The difference between the string quartet *ear* and **c** is currently used as a measure of dominance: in the figure above (right, condition B), the first violin is dominant over the others, while condition A correspond to an equilibrium of relations between musicians. A submitted paper includes more details and variations of the algorithm.

2.3 Measuring leadership using granger causality

This section describes the Granger Causality (G-causality) -based analysis carried out for the study of leadership patterns within a quartet presented at the workshop on Social Behavior in Music at ACM ICMI 12 (Glowinski et al., 2012).
The analysis of leadership patterns was based on the concept of Driving Force already described in the deliverable D1.3, section 3.1, equations 6 and 7. The computation of the Driving Force, which is based on the G-causality method, was carried out on the time series of Euclidian distances between the musician heads and the "ear" of the quartet. G-causality was computed every second on 3-second sliding windows. The length of the "history" in the Auto-Regressive models was computed at each window using the Akaike information criterion.
For each musician *i* we first computed the sum of all his Driving Force (henceforth the Driving Force of a musician) on all the other musician j ($\sum_{j\neq i} DF_{i\to j}$). We then sorted the four summed driving forces in ascending order. Finally, we fit the four data points with a second order polynomial. Considering the polynomial in descending power, the first two coefficients give us information about the distribution of the leadership (see figure 1). For example, if both coefficients are very small the leadership will be equally distributed within the quartet, whereas if the first coefficient has a large positive value the leadership is concentrated on one single musician.

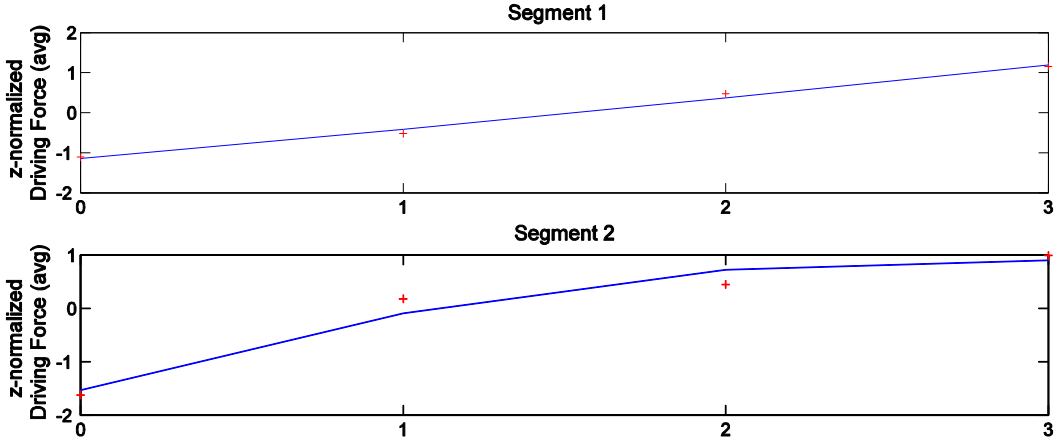

Figure shows the Leadership distribution in two different segments of a performance. The red crosses represent the $\sum_{j\neq i} DF_{i\to j}$ of each musician i (numbered from 0 to 3). The solid line is the 2-nd order polynomial fitting of leadership distribution values. Note that the number associated to a musician in two different segments can be different.}

The Driving Force of each musician ($\sum_{j \neq i} DF_{i \to j}$) was also used to identify the musician that most influenced the other musicians by counting the number of times $\sum_{j \neq i} DF_{i \to j}$ was positive.

2.3.1 Granger Causality-based inter-musician communication within a quartet

The goal of this work was to study the relation between the quantity of information flow among musicians (inter-musician communication) and : i) the complexity in the coordination among musicians ii) unexpected (by all musicians but the first violin) rhythmic and dynamical changes to the score.
Inter-musician communication was computed using conditional G-causality as described in the deliverable 1.3, section 3.1, equation 8 (in that document inter-musician communication is referred to as inter-violinist interaction). Again, the G-causality method was applied to the time series of the Euclidian distances between the musician heads and the "ear" of the quartet. Conditional G-causality was computed every 500 milliseconds (i.e., 2Hz "sampling frequency") on 3-second sliding windows.
We then had to align the inter-musician communication over time and the annotation of both complexity and the unexpected changes on the actual performances. Since complexity and unexpected changes were annotated on the audio signal of the performance which is sampled at much higher frequency than the "sampling" frequency of the inter-musician communication (2Hz) we oversampled and then interpolated the inter-musician communication. Interpolation was carried out by inserting zeros into the new (value missing) samples and then applying a low-pass filter.
Finally we compared:
- the average inter-musician communication within complex intervals vs. the average inter-musician communication within non-complex intervals of a same take.
- the average inter-musician communication within score intervals where unexpected changes occurred vs. the average inter-musician communication within the same score intervals but without any unexpected change in the actual performance.
In the first case the comparison was performed within the same take, while in the latter case we used two different takes (where the same score was executed), one where unexpected changes occurred and a corresponding one where no unexpected change occurred.

2.4 Reference

1. de Cheveigne, A. & Kawahara, H. (2002). YIN, a fundamental frequency estimator for speech and music. *Journal of the Acoustical Society of America*, 111(4):1917–1930.
2. Maestre, E. (2009). Modeling instrumental gestures: an analysis/synthesis framework for violin bowing. *PhD thesis, Universitat Pompeu Fabra.*
3. Pereda, E., Quiroga, R. Q., & Bhattacharya, J. (2005). Nonlinear multivariate analysis of neurophysiological signals. *Progress in Neurobiology*, 77(1-2):1–37.
4. Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37(3):424–438.
5. Seth, A. K. (2010). A MATLAB toolbox for Granger causal connectivity analysis. *Journal of Neuroscience Methods*, 186(2):262–273.
6. Chicharro, D. & Andrzejak, R. G. (2009). Reliable detection of directional couplings using rank statistics. *Physical Review E - Statistical, Nonlinear and Soft Matter Physics*, 80(2 Pt 2):026217.

3. TECHNIQUES FOR ANALYSIS OF ORCHESTRA

3.1 Extraction of an expressive visual group feature for the discrimination of two performance conditions by an orchestra

This section describes the procedure we have followed to measure the difference in the behavior between two performance conditions of the first violin section of an orchestra. Such performance conditions were characterized by different eye contact between the first violin section and the second violin one, and between the first section and the conductor (cf. D2.2)

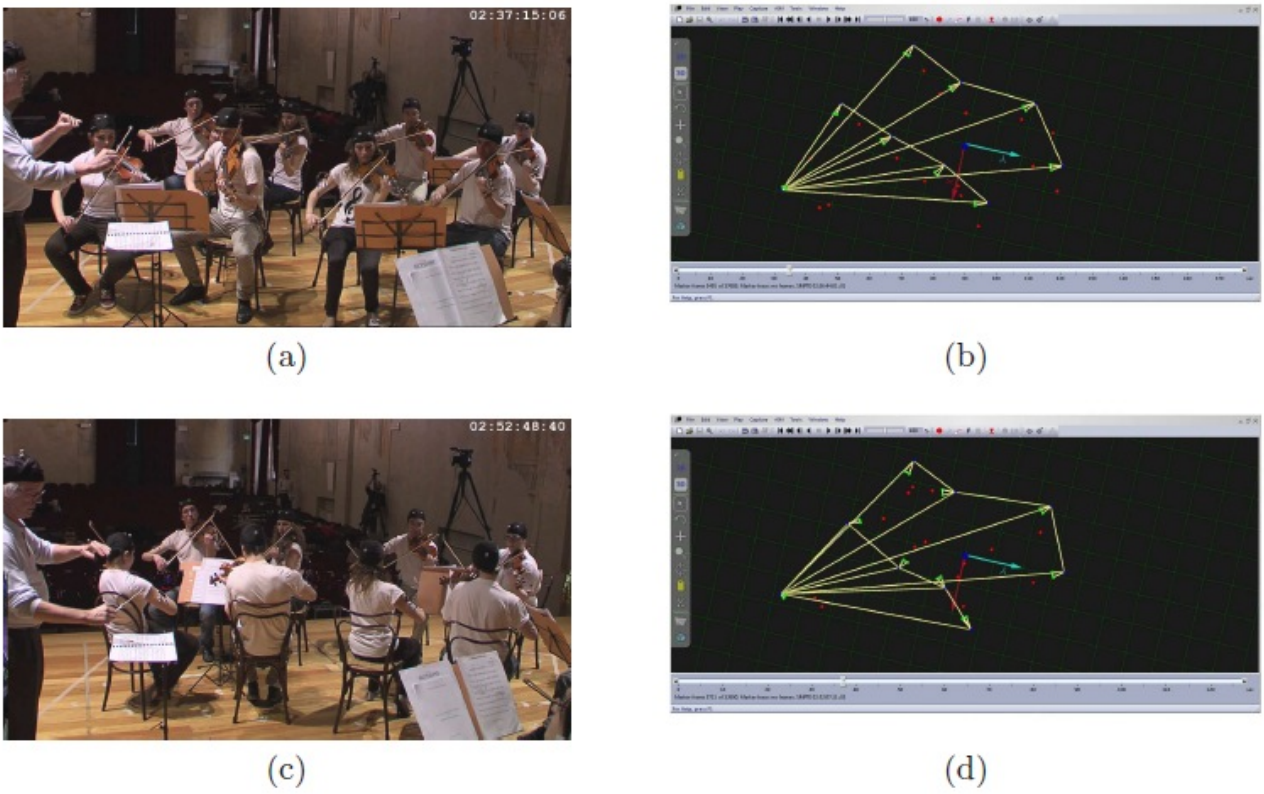

Figure shows Panels (a) and (c) show the players and the conductor when condition A and condition B are tested, respectively. Panels (b) and (d) show a snapshot of the head's markers positions of the players and the conductor when condition A and condition B are tested, respectively. Triangles correspond to positions and directions of heads. The red points are the unlabeled markers (mainly associated with the bows), cf. D2.2.

Movement data were collected by using a Qualisys motion capture system equipped with 7 cameras, integrated with the EyesWeb XMI platform (www.eyesweb.org) for obtaining synchronized multimodal data. A reduced data set, made of 12 recordings, was extracted from the collected data, and a group feature *f*, related to the movement of the heads of the musicians of the first section, was automatically computed.

The following is a description of how such group feature *f* was computed.

For each frame k ($k = 1, \ldots, N_{frames}$) of a recording

1. For each violinist i ($i = 1, \ldots, 4$) in the first section, the current direction $\mathbf{d}_i^{(j)}$ in the horizontal plane of the head of the musician and the corresponding sample mean direction $\bar{\mathbf{d}}_I$ are computed. Each $\mathbf{d}_i^{(j)}$ is defined as the unit vector connecting the marker on the back of the neck of the musician i to the point located in the middle of the line between the two other markers above his/her eyes, whereas $\bar{\mathbf{d}}_i$ is obtained by averaging each component of $\mathbf{d}_i^{(j)}$ with respect to j and normalizing the obtained vector.

2. For each violinist i in the first section, the oriented angle $\theta_i^{(j)}$ between the two vectors $\bar{\mathbf{d}}_i$ and $\mathbf{d}_i^{(j}$ is computed. Then, the vector $\mathbf{b}^{(j)} \in \mathbb{R}^4$ with components $\theta_1^{(j)}$, $\theta_2^{(j)}$, $\theta_3^{(j)}$, $\theta_4^{(j)}$, and its sample mean $\bar{\mathbf{b}}$ with respect to the frames and its sample covariance matrix are computed.
$$sample\_cov(\mathbf{b}) := \frac{1}{N_{frames} - 1}\left(\mathbf{b}^{(j)} - \bar{\mathbf{b}}^{(j)}\right)\left(\mathbf{b}^{(j)} - \bar{\mathbf{b}}^{(j)}\right)^T \in \mathbb{R}^{4\times 4}.$$
In case of missing data (e.g., an unlabeled head marker or an undetected one), the corresponding frames are discarded and $\bar{\mathbf{b}}$ and $sample\_cov(\mathbf{b})$ are evaluated using only the remaining frames (thus reducing the effective value of N_{frames}).

3. For each of the two conditions, the sample mean $\bar{\mathbf{b}}$ and the sample covariance matrix $sample\_cov(\mathbf{b})$ were averaged over the six repetitions of the same music piece (three for each conductor). The next two tables show the resulting average of $sample\_cov(\mathbf{b})$ under conditions A and B, respectively.

	θ_1	θ_2	θ_3	θ_4
θ_1	0.0188	-0.0050	0.0103	0.0069
θ_2	-0.0050	0.0695	-0.0051	0.0338
θ_3	0.0103	-0.0051	0.0547	0.0143
θ_4	0.0069	0.0338	0.0143	0.0829

Table 1. Average sample covariance matrix (rad^2) of the feature vector **b** under condition A.

	θ_1	θ_2	θ_3	θ_4
θ_1	0.2156	0.0221	0.0103	-0.0402
θ_2	0.0221	0.1407	-0.0908	-0.0077
θ_3	0.0103	-0.0908	0.25477	0.0425
θ_4	-0.0402	0.0077	0.0425	0.2680

Table 2. Average sample covariance matrix (rad^2) of the feature vector b under condition B.

4. Finally, we defined the group feature f as the sum of the diagonal elements (trace) of the average of $sample\ cov(\mathbf{b})$. This is a group feature since it takes into account the behavior of all the violinists of the first section.

The obtained value of the group feature f (and its empirical standard deviation) was about 0.2259 rad^2 (0.2191 rad^2, respectively) for condition A and 0.8790 rad^2 (0.3916 rad^2, resp.) for condition B. So, in a sense, larger deviations from the mean directions seemed to be associated with condition B (in which the violinists of the first section were not able to see the conductor) with respect to condition A (in

which they were able to see the conductor). This behavior in condition B may be motivated by the absence of a reference point (the conductor) to look at in such a situation.

According to a further inspection of the available data, this change in the relative size of the movements of the heads when passing from condition A to condition B seemed not to depend on the conductor, although different absolute sizes of the movements of the heads of the musicians were observed for the two conductors.

Possible future extensions of the analysis of the results of this experiment involve:

- computation of the feature *f* (and other features) at a local scale, not only at a global scale (i.e., on the entire recording), as made so far. In this way one could be able to take into account other aspects of the experiment, e.g., the movement of the baton of the conductor (not examined in the present work), and to detect possible differences in the behavior of the musicians during different musical phrases;
- investigation of the position of the "point of convergence" of the first violin section under both conditions A and B, and its distance from the conductor under condition A;
- investigation of the interaction between the two violin sections, under both conditions A and B;
- study of the influence of the conductor/music piece on the behavior of the musicians;
- use of other statistical techniques in the data analysis, e.g., ANOVA and Granger's causality.

4. TECHNIQUES FOR ANALYSIS OF AUDIENCE

4.1 Extraction of Features from Physiological Signals

In this section a new approach on the design of algorithms for bio-inspired systems for audience measurements is presented. The extensive physiological database collected during the *Emotion in Motion* (EiM) series of experiments (Jaimovich et al. in press; Jaimovich et al. 2012) has allowed the design and calibration of automatic feature extraction tools, with special consideration to automatic artifact detection and removal techniques. In particular, these algorithms focus on two channels of physiology measured from audiences within the SIEMPRE context, electrodermal activity (EDA) and heart rate (HR).
Two tools were developed in MATLAB to extract features from the physiological signals described above: *EDAtool* and *HRtool*. The extraction of features includes detection and removal of artifacts and abnormalities in the data. The output from both tools delivers the processed feature vectors (e.g. phasic EDA or HR) as well as a confidence index (Q) or accuracy of the input signal, which is defined as the percentage of the signal that did not present artifacts. This value can be utilized later to remove cases from the analysis that fall below a specified confidence threshold. The EiM data has been used to both design and evaluate these algorithms.

4.1.1 EDAtool

EDAtool is a function developed to pre-process the EDA signal. Its processing includes the removal of electrical noise and the detection and measurement of artifacts. Additionally, it separates the EDA signal into phasic and tonic components. The processing of the EDA signal has several stages, which

are detailed in the following sections, and illustrated in Figures below. Additionally, EDAtool is available online.[1]

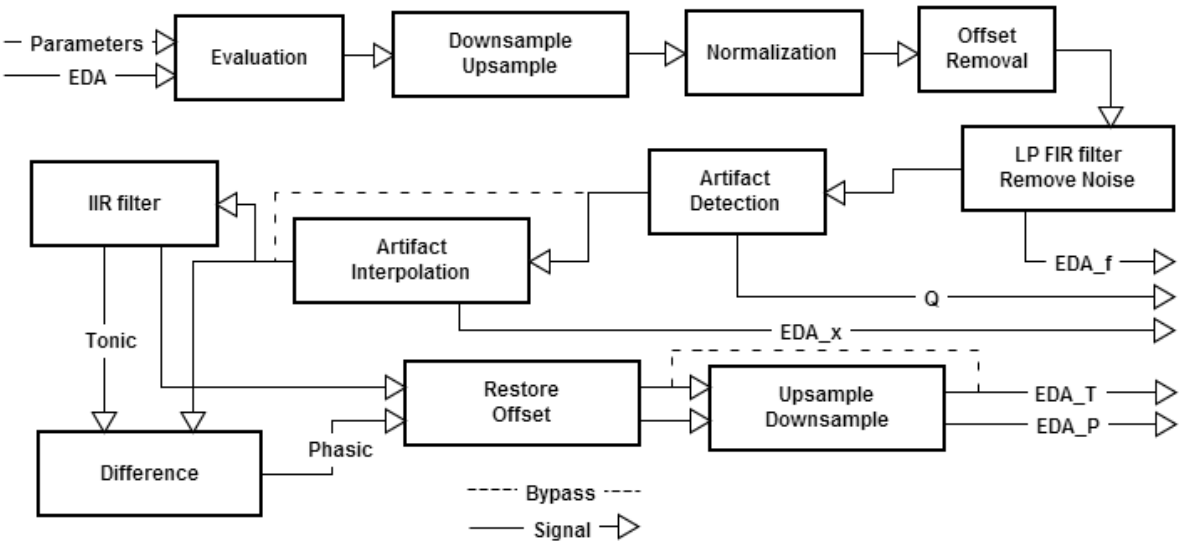

Figure shows Block diagram of the EDAtool processing stages.

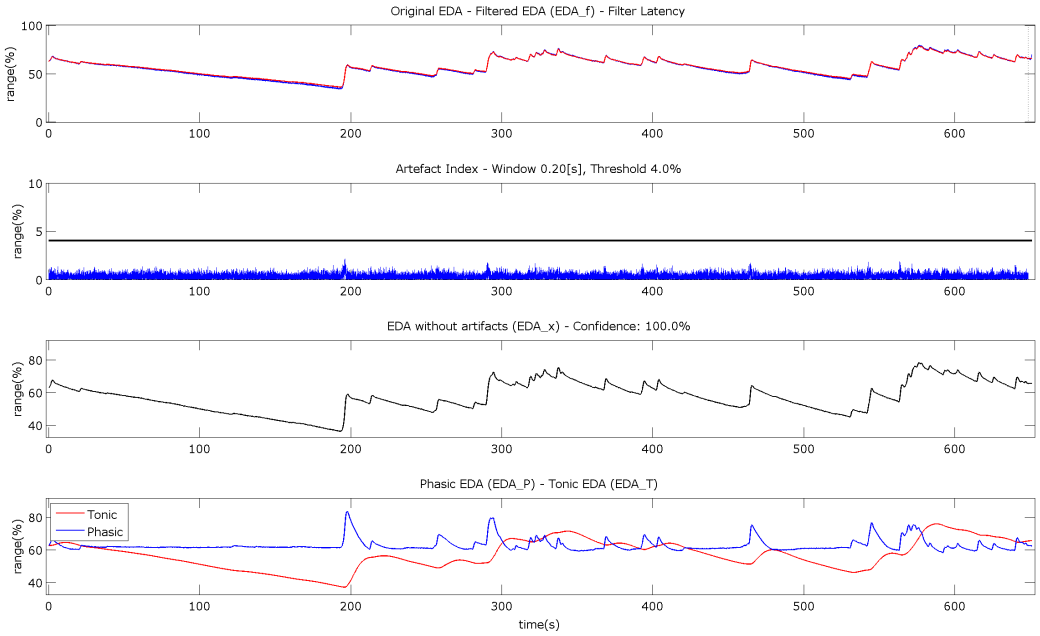

Figure shows the Stages of the EDAtool on an electrodermal signal. The top plot shows the original (blue) signal and the low-passed filtered signal (red) without electrical noise. The next plot shows the artifact detection method, which identifies abrupt changes in the signal. The 3rd plot from the top shows the filtered signal with the artifacts removed. The bottom plot shows the phasic (blue) and tonic (red) components of the signal.

1.1.1 Revision of Input Parameters

The function has several input variables: the raw EDA signal as a vector, the sample rate (SR) of the recorded signal, and options to adjust the parameters of the EDA processing. These parameters are described in Table 3.

[1] http://www.musicsensorsemotion.com/2012/06/21/edatool/

Table 3. Parameters for *EDAtool.*

PARAMETER	DESCRIPTION
RangeMin	Lowest EDA sensor value (e.g. open circuit).
RangeMax	Highest EDA sensor value (e.g. closed circuit).
WindowSize	Size of artifact window in seconds.
Threshold	Threshold for detecting artifacts within window, measured in percentage of the sensor's range.
InterOption	Option to perform linear interpolation between artifacts.
ResampleOption	Option to resample the output vectors to the original SR.
Debug	Option to print and plot processing results.

The first section of the EDA algorithm simply checks that all the input variables have been entered correctly, and that they are within the proper ranges and formats (e.g. duration of EDA must be longer than the filters' impulse response times, etc.).

1.1.2 Pre-processing

Initially, the EDA signal is resampled to 50[Hz], using decimation or interpolation, depending on the original SR entered to the tool. This standardization is implemented in order to keep the filter coefficients constant for every signal, with a bandwidth that allows the detection of artifacts. Subsequently, any sample outside the sensor's range (specified in the input parameters) is limited, and counted as non-valid for accuracy purposes. Next, the start value of the signal is saved, and then the signal is shifted to start at zero, in order to be filtered without DC component. Finally, the signal is normalized in accordance to the range of the EDA sensor specified by the user.

4.1.2 Removal of Electrical Noise

The next section in the processing script removes any noise that is outside the EDA spectrum, typically electrical noise. For this, a 224 order low-pass FIR filter is used, with a cut-off frequency of 0.5[Hz] (stop frequency of 1[Hz] at -60[dB]), see Figure below the latency produced by the filter is compensated by shifting the EDA vector samples. This filtered EDA vector is delivered as the first output of the EDAtool function (see Figure below).

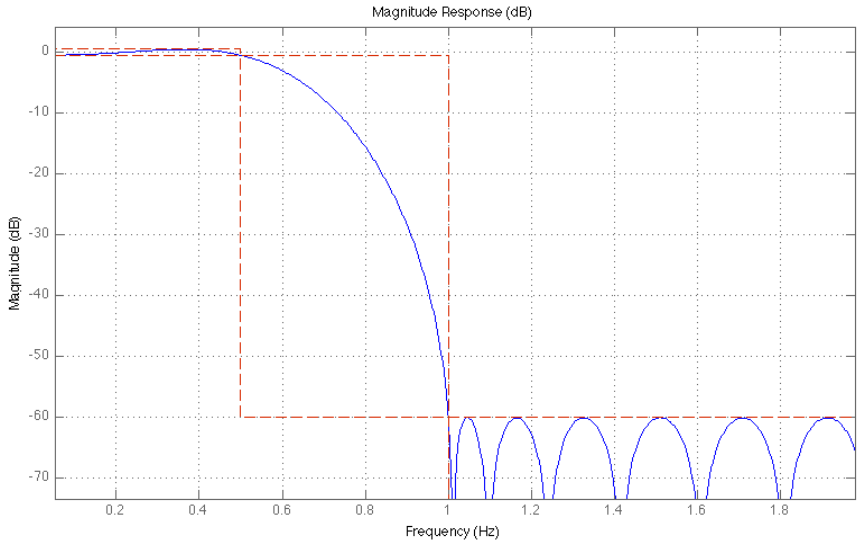

Figure shows the Magnitude response of the noise removal EDA filter.

4.1.3 Artifact Detection and Removal

EDA artifacts are typically caused by problems with the electrode-to-skin connection, which results in a discontinuity or rapid change in the conductivity measured by the sensor (see Figure below). Boucsein (Boucsein 2012, p.140) gives a thorough overview of the physiologically and recording based artifact causes. He argues that due to the multiple origins of artifacts (e.g. loss of contact, change in contact pressure, subject's temperature, respiration induced SCRs, etc.), artifacts should be identified and corrected manually by researchers, yet this requires a significant amount of time for databases such as the EiM, and is not feasible for real-time applications. For this reason, automatic detection and removal of EDA artifacts induced by motion has been developed as part of the *EDAtool*. With regard to other source of artifacts, such as respiration induced SCRs or temperature, this needs to be controlled with additional sensors (e.g. thermistors), and are not currently considered in this tool.

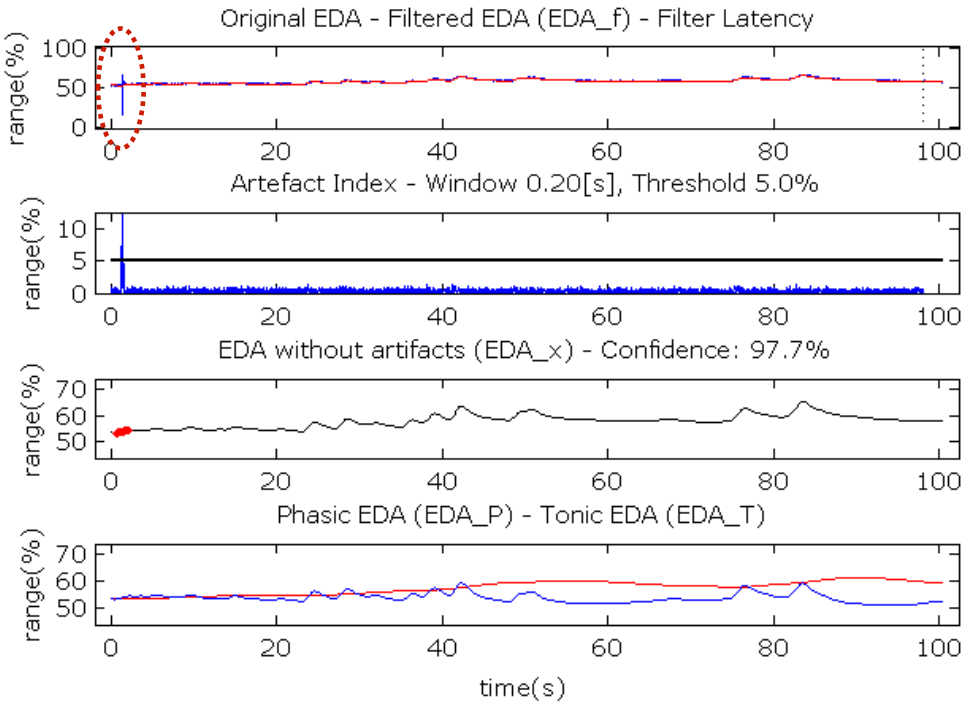

Figure shows Example of EDA signal with artifact.

The detection of artifacts consists of evaluating the gradient between the edges of a sliding window against a fixed threshold. Because the changes in EDA are slower than changes produced by artifacts, it is possible to separate both events with this method. The size of the sliding window, as well as the threshold value, is specified by the user in the input parameters (see Table 3). If an artifact is detected, the algorithm replaces the adjacent samples with non-valid values using a 1.5[s] window, centered on the artifact.

A confidence index (Q) is obtained by calculating the ratio of non-valid values versus valid samples in the signal. After evaluating the values of Q against multiple examples of the EiM database, another variable was factored into the confidence index. It was noticed that on certain cases, the baseline of the EDA signal was shifted significantly before and after an artifact (see Figure above). This may be due, for example, to the electrode not making a good connection initially, and after an adjustment (which triggers an artifact) the signal settles to the correct baseline. Instead of artificially correcting this difference, it was decided to consider this factor in the calculation of Q by decreasing its value proportionally to the difference between the EDA signal before and after the artifact.

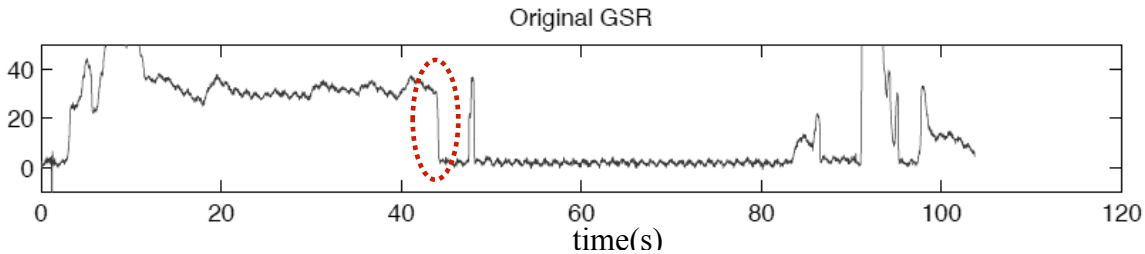

Figure shows an example of a baseline shift in EDA before and after an artifact (circled).

In summary, the confidence index is calculated based on the number of samples outside the sensor's range, the samples replaced by artifacts, and the amount of base shift before and after every artifact.

4.1.4 Artifact Interpolation

If requested, the EDAtool interpolates between the detected artifacts. For this, the algorithm searches for any non-valid sample in the signal, identified during the previous sections, and performs a linear interpolation between the values before and after the non-valid sample.

4.1.5 Tonic and Phasic EDA

After testing several Finite Impulse Response (FIR) filters to separate the EDA signal into its tonic and phasic components, it was observed that the filter orders required to obtain the desired signals were producing extremely long latencies (over 15 seconds) with high processing demand. This created a problem for the EiM database analysis, which meant the loss of significant sections of the EDA vector, and also for any potential real-time applications that require faster processing times.
The solution was then to design an Infinite Impulse Response (IIR) Butterworth low-pass filter with a cut-off frequency of 0.001[Hz] (stop frequency of 1[Hz] at -60[dB]). Even though this option involves the inherent stability and phase (time dispersion) issues associated with IIR filters (Mitra 2001), the filter coefficients were manually calibrated against the EiM database in order to obtain stable results. Moreover, this filter is executed after the artifact and noise removal stages; hence it is applied to EDA signals with a similar frequency content, which will produce a similar dispersion from the IIR filter. Figure shows a comparison between the magnitude response of the IIR filter utilized, versus two FIR filters; one of the same order as the IIR and the other with similar results but higher order. Notice how neither FIR filter has the sharp response of the IIR filter, which was required to separate the tonic and phasic components from EDA.

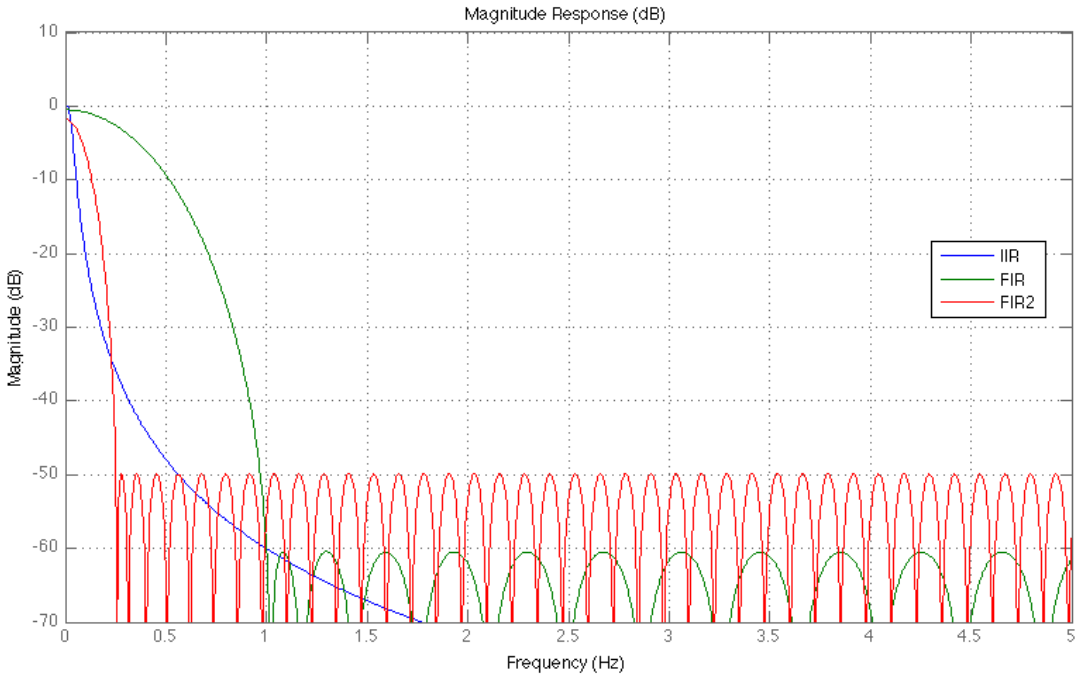

Figure shows the Comparison of magnitude response between IIR filter used in tonic and phasic separation, versus two FIR filters with similar parameters.

The tonic component is then extracted from the output of the IIR filter, while the phasic component is obtained from the difference between the original signal (supplied to the IIR filter) and the tonic component (see Figure above, bottom plot).

4.1.6 EDAtool outputs

The final section of EDAtool returns each EDA vector to the original baseline (stored in the pre-processing section), and resamples each vector to the original SR (if requested by user). Originally, the output signals were all set to start in zero, but the relationship between the EDA baseline and the amplitude of phasic changes confirmed the importance of evaluating the changes in the EDA signal with respect to its absolute level (Jaimovich et al. in press).
In summary, the EDAtool delivers five different outputs, described in Table 4 and Figure above.

Table 4. Outputs of the *EDAtool.*

VARIABLE	DESCRIPTION
EDA_f	Low-pass filtered EDA, without electrical noise
EDA_x	EDA_f without artifacts (removed or interpolated)
EDA_P	Phasic component of EDA (EDR)
EDA_T	Tonic component of EDA (EDL)
Q	Confidence index for EDA (0-100)

4.1.7 EDAtool Parameters

The EDAtool parameters were calibrated using EDA signals from the EiM database. The range was set at the values of opened and closed circuit configurations of the BioControl sensor used in our lab.[2] The size of the artifact window, as well as the threshold level were chosen after reviewing hundreds of

[2] http://infusionsystems.com/catalog/product_info.php/products_id/203

signals, manually identifying artifacts and changing these parameters until they were recognized by the *EDAtool*. If an artifact was found, a 2.5s duration window was chosen to remove the EDA section, to then apply a linear interpolation between the edges of the artifact.

4.2 HRtool

HRtool is a function developed to convert the data from an Electrocardiogram (ECG) or Pulse Oximetry (POX) signal into a HR vector. The function identifies heart pulses in the input signal discarding artifacts and ectopic beats, to then generate a two dimensional matrix with the BPM value of each pulse and its relative location in time. The HRtool function and documentation are available online.[3]

HRtool takes five input variables: The ECG or POX signal, the SR of this signal, a debug option, a parameter vector and the type of input signal (e.g. POX or ECG). It is important to point out that *HRtool* checks for a TTL pulse signal (e.g. 0s and 1s) when the POX option is selected (this is due to the output of the Mediaid sensors used in our lab[4]). If an actual pulse plethysmography signal is used, then the ECG option will automatically detect the peaks and extract HR. The parameter vector contains the minimum and maximum acceptable heart rate values, the maximum change ratio between two consecutive beats, and the threshold for beat detection; defined as the ratio between the beat values and the standard deviation of the ECG signal within a moving window. It outputs a vector with the HR values and the time in seconds where they occurred, relative to the start of the vector. Additionally, it outputs the mean HR and confidence index Q.

After an initial section that checks that the parameters are entered correctly and that the length of the input signal and SR are over certain minimum values, the function is divided in three processing steps (see Figures and description below).

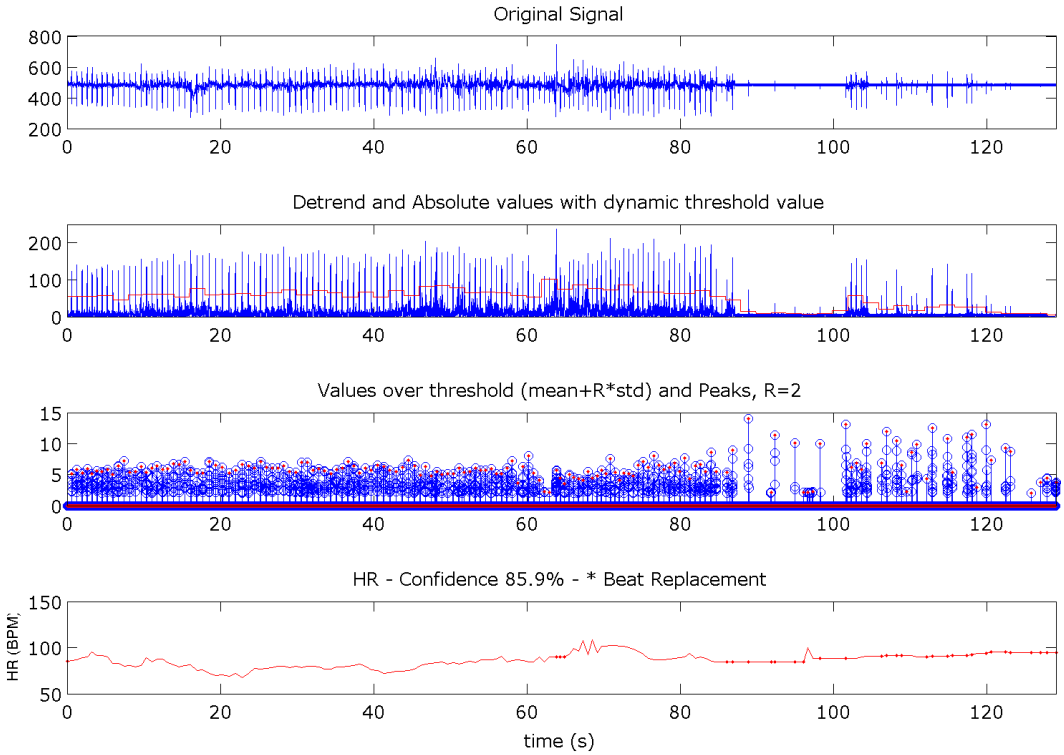

Figure shows Stages of the HRtool on an ECG signal. The top plot shows the original ECG signal. The two middle plots show the peak detection stages, with a dynamic threshold. The bottom plot shows the

[3] http://www.musicsensorsemotion.com/2012/06/21/hrtool/
[4] http://www.mediaidinc.com/OemSolutions.html

final HR vector, with the resulting replacements of values that are outside the specified ranges (marked as dots in the plot).

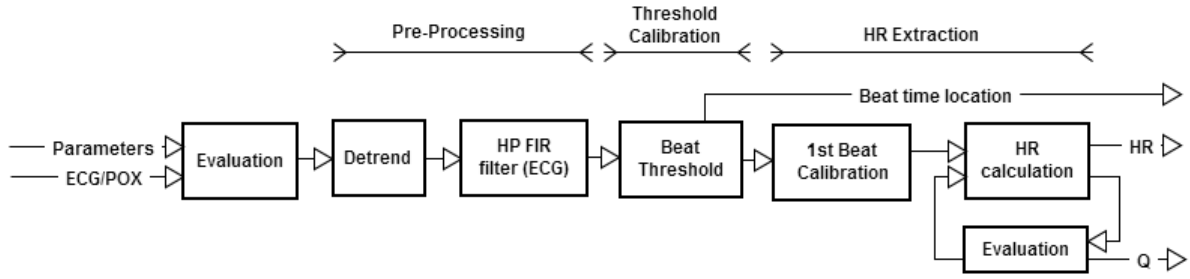

Figure shows HRtool block diagram.

4.2.1 Pre-Processing

The first step for either signal, ECG or POX, is to remove any baseline trends in the signal. This is accomplished by subtracting the best straight-line fit trend from the vector.
For ECG, the signal is filtered using an FIR high-pass Kaiser Window, with cut-off frequency of 3[Hz] (stop frequency of 2[Hz] at -60[dB]). The filter coefficients are calculated each time the function is called, and they depend on the SR of the signal. After filtering, the ECG signal is time shifted to compensate for the filter's latency.

4.2.2 Threshold Calibration

For TTL POX signals, the algorithm computes a threshold for peak detection by calculating the signal's standard deviation plus a 20%. Zeros replace values below this threshold and the remaining samples are substituted by ones.
The ECG signal is scanned by a moving window of 2 seconds duration. The signal for each window is z-normalized and the threshold is calculated by adding the mean value of the ECG signal in the window, with the product of the standard deviation and the threshold entered by the user. This allows having a dynamic threshold across the duration of the signal, as can be seen in **Error! Reference source not found.**, 2nd plot from the top.

4.2.3 HR Extraction

Before extracting the HR for the entire vector, the algorithm computes an average HR value from the first 10 seconds in the signal, which is used as the initial HR of the vector.
In order to find peaks in the ECG signal, *HRtool* creates a sliding window of a size equivalent to the minimum beat-to-beat distance obtained from the maximum HR value entered by the user. The algorithm then searches for the highest value in the window (the R wave) and saves its position, to then calculate the interval with the subsequent peak value. For TTL POX signals, the first sample above the threshold is utilized.
After measuring the RR intervals between pulses and calculating each corresponding HR value in BPM, the algorithm evaluates the HR vector replacing any values that are outside the ranges entered by the user (e.g. values within the maximum and minimum HR range, and within the maximum change ratio between two consequent pulses).
Finally, the confidence index Q is obtained by calculating the ratio of replaced beats against normal beats. In the example signal shown in Figure below, accuracy is at 85.9%, which falls below the acceptance tolerance of 90% used in the experiments analyzed in EiM, and would be discarded as a valid case.

4.2.4 HRtool Parameters

The HRtool parameters were set with the following boundaries for the processing of the EiM signals: Any RR intervals lower than 50[BPM] or higher than 130[BPM] were discarded. The maximum allowed change for a new beat with respect to the previous one was set to 20%. Regarding the decision on these parameters, the literature states that up to this date there is no clear standardization for short-term measurements of heart rate variability (Nunan et al. 2010), and neither for which method should be used for interpreting ectopic beats and artifacts (Salo et al. 2001). Moreover, the majority of the methods proposed in the literature use statistics extracted from the total duration of the recorded data to identify abnormal beats (Kleiger et al. 2005), which is incompatible with the real-time implementations intended to be used in SIEMPRE. Nevertheless, there has been a thorough manual examination of the HR feature extraction tool, screening hundreds of different cases presented in the EiM database, in order to test the correct application of the above parameters.
Normal sinus rhythm is defined between 60 and 100[BPM], with variations of less than 10% between beats. Yet, the parameters were set in values with higher tolerance levels than the ones defined, due to the amount of signals with accuracy over 90% decreased considerably. Moreover, setting the parameters to comply with normal sinus rhythms does not account for the possibility that music might produce cardiac rhythms outside the normal standards.

4.3 Autonomic response to music listening

As mentioned in D1.3, techniques used in to measure autonomic response to music listening are mainly consolidated techniques for data analysis, such Principal Component Analysis, Factor Analysis, correlation. An experiment was conducted to investigate the temporal structure of the GEMS dimensions based on the dynamic emotional judgments, using temporal Principal Component Analysis (tPCA). A computer interface was used (cf. Deliverables D 1.1 and D2.2) and 36 musical excerpts – 4 pieces by GEMS dimensions- were evaluated dynamically. A series of analyses using z-scores were performed and the first observation was that individuals show high agreement between them on emotions expressed by music (see Figure below).

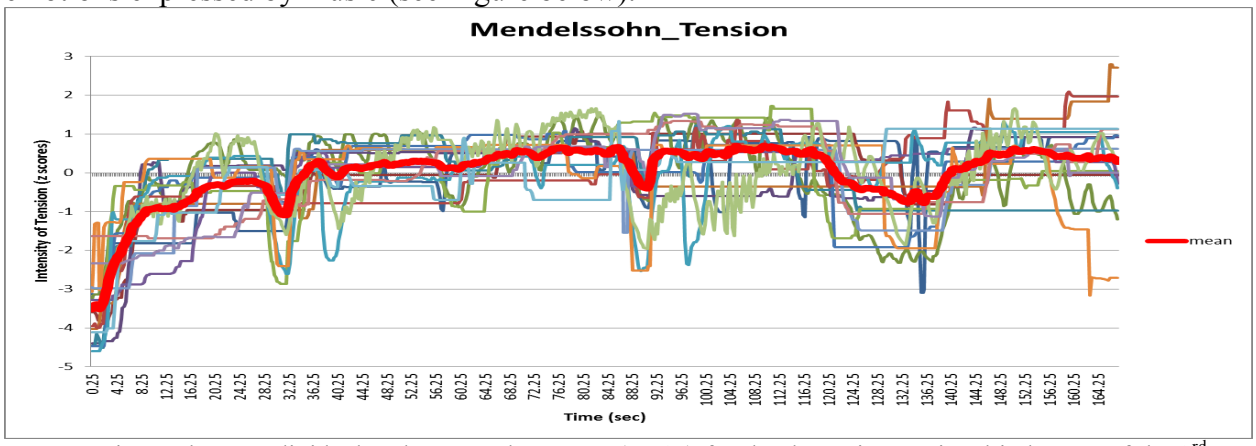

Figure shows Individual and averaged z-scores (N=17) for the dynamic emotional judgment of the 3rd movement of the Octet for string in E-moll by F. Mendelssohn judged on the dimension of « Tension» (duration: 2'47'). This figure illustrates the great agreement between participants regarding the emotion expressed by music. The lines represent the normalized participant's judgments. The average is represented by the red line. The y-axis represents the intensity of the emotion expressed by music (here, Tension) and the x-axis, the time. We observe that this musical excerpt rises rapidly in power then drops slightly to finally ascend. We note that the expression of the emotion in this musical excerpt does not represent a straight line along time and that the

In order to evaluate the GEMS temporal dimensionalities, a temporal principal component analysis (tPCA) on the four musical excerpts for each dimension has been performed. The inflexion of the Eigen values function shows the existence of five components (Figure below).

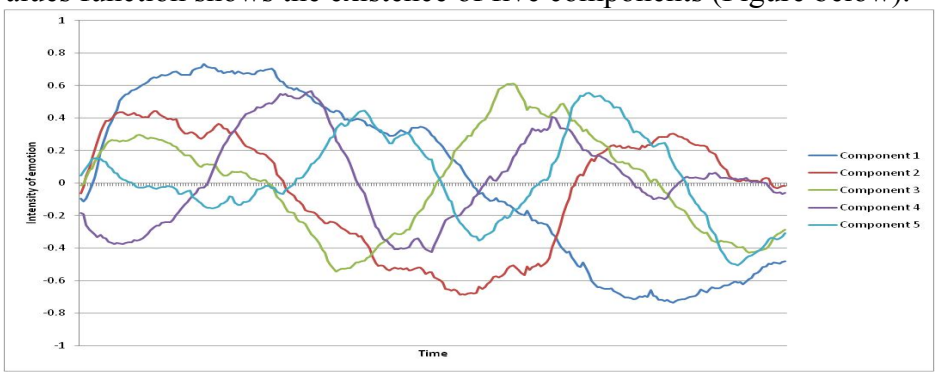

Figure shows five components for the temporal structure of the GEMS dimensions expressed by music revealed by the tPCA.

Component 1 clearly represents the emotional pattern of the Joyful activation and the Power dimensions, while component 2 is more representative of the emotional pattern of the Transcendence dimension, but also share some characteristics with the Power dimension (Figures below).

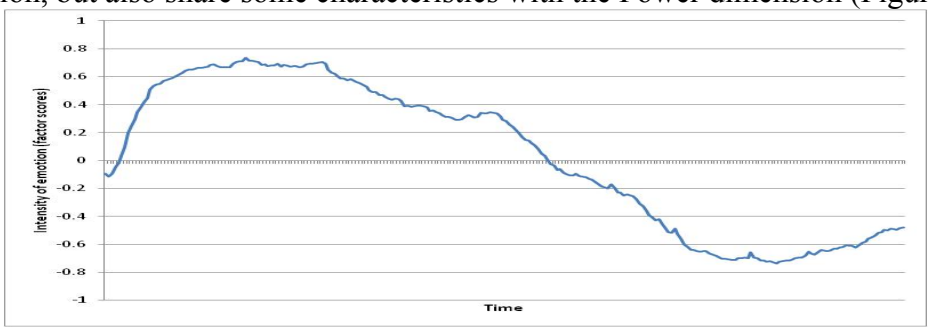

Figure shows the temporal structure of the emotion expressed by music for the first component (e.g. Power and Joyful Activation).

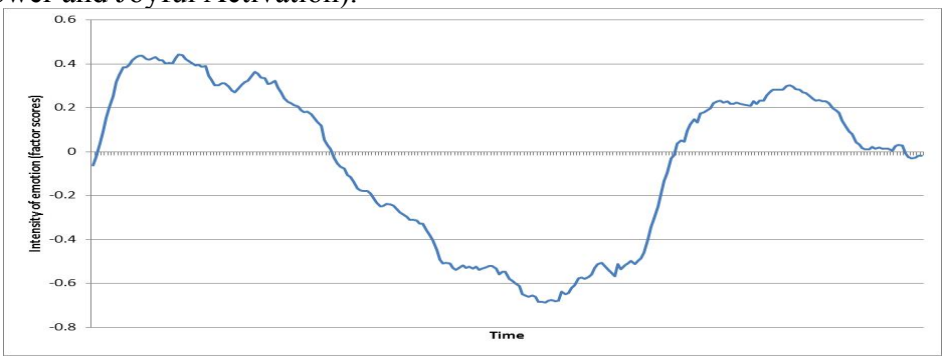

Figure shows the Temporal structure of the emotion expressed by music for the second component (e.g. Transcendence and Power dimensions).

The only component that differentiates the Power and the Transcendence dimensions is the component 1. The most representative GEMS dimension for the component 4 is Wonder and the Tension dimension is characterized by the component fourth but less by the fifth, which are more representative of the Sadness and Tenderness dimensions (Figure below).

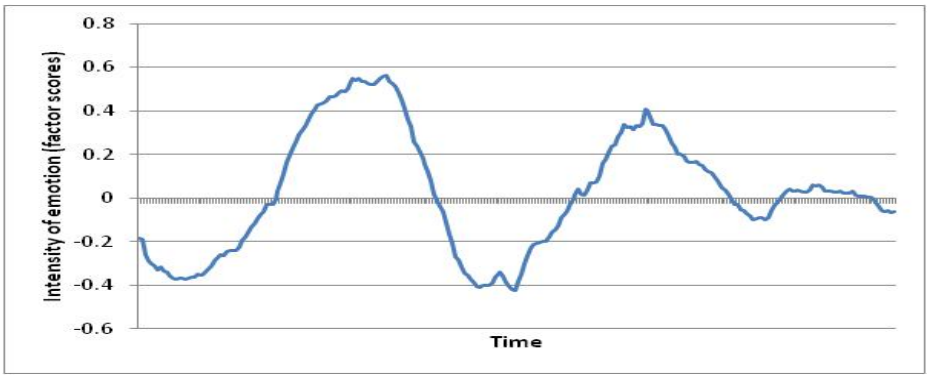

Figure shows Temporal structure of the emotion expressed by music for the fourth component (e.g. Wonder and Tension).

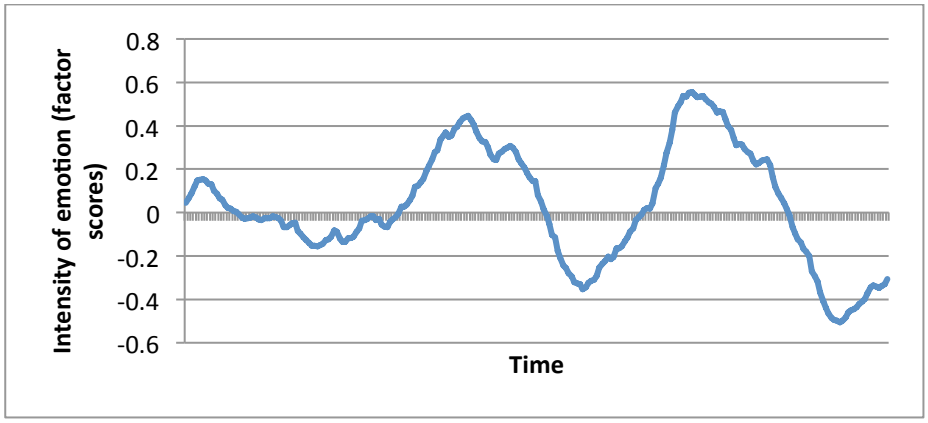

Figure shows Temporal structure of the emotion expressed by music for the fifth component (e.g. Sadness and Tenderness).

The tPCA reveals that the GEMS dimensions do not seem to have the same temporal emotional patterns across the different musical excerpts and therefore highlights the importance of capturing the emotional time course to better understand the emotions expressed by music. These analyses also reveals that is possible to group some GEMS dimensions on the same component (e.g. Power and Transcendence) but demonstrates also that there is clear different temporal pattern between other dimensions (e.g. Wonder and Sadness), reflecting the complexity and the specificity of emotions expressed by music.

1.2 References

1. Boucsein, W., 2012. *Electrodermal Activity*, New York: Springer.
2. Jaimovich, J. et al., 2012. The Emotion in Motion Experiment: Using an Interactive Installation as a Means for Understanding Emotional Response to Music. In *Proceedings of the 12th International Conference on New Interfaces for Musical Expression, 21-23 May*. New Interfaces for Musical Expression (NIME). University of Michigan, Ann Arbor, pp. 457–458.
3. Jaimovich, J., Coghlan, N. & Knapp, R.B., in press. Emotion in Motion: A Study of Music and Affective Response. In *Post-Proceedings of the 9th International Symposium on Computer Music Modeling and Retrieval (CMMR) Music and Emotions*. Symposium on Computer Music Modeling and Retrieval. Queen Mary University of London: Springer; Lecture Notes in Computer Science.
4. Kleiger, R.E., Stein, P.K. & Bigger Jr., J.T., 2005. Heart Rate Variability: Measurement and Clinical Utility. *Annals of Noninvasive Electrocardiology*, 10(1), pp.88–101.

5. Mitra, S.K., 2001. *Digital Signal Processing: A Computer-Based Approach*, New Deli, India: McGraw-Hill Higher Education.
6. Nunan, D., Sandercock, G.H.R. & Brodie, D.A., 2010. A Quantitative Systematic Review of Normal Values for Short-Term Heart Rate Variability in Healthy Adults. *Pacing and Clinical Electrophysiology*, 33(11), pp.1407–1417.
7. Salo, M.A., Huikuri, H.V. & Seppanen, T., 2001. Ectopic Beats in Heart Rate Variability Analysis: Effects of Editing on Time and Frequency Domain Measures. *Annals of Noninvasive Electrocardiology*, 6(1), pp.5–17.

5. PUBLICATIONS

1. A. Camurri, A. D'Ausilio, L. Fadiga, D. Glowinski, G. Gnecco, M. Sanguineti, G.Varni, G. Volpe, "Analysis of music ensemble performance as a test-bed for social interaction: Methods from Operations Research and preliminary results", AIRO WINTER 2013, Champoluc, 28 gennaio-01 febbraio 2013.
2. B. Bortz, S. Salazar, J. Jaimovich, R. B. Knapp, G. Wang (2012) ShEMP: A Mobile Framework for Shared Emotion, Music, and Physiology, ACM-ICMI SBM workshop, Santa Monica, California, USA. PDF
3. D.Glowinski L.Badino A.D'Ausilio, A.Camurri, L.Fadiga (2012) Analysis of Leadership in a String Quartet, ACM-ICMI SBM workshop, Santa Monica, California, USA. PDF
4. D.Glowinski, K.Torres-Eliard, C.Chiorri, A.Camurri, D.Grandjean (2012) Can naïve observers distinguish a violinist's solo from an ensemble performance? A pilot study", ACM-ICMI SBM workshop, Santa Monica, California, USA. PDF