SIEMPRE

DISSEMINATION LEVEL

Social Interaction and Entrainment using Music PeRformancE

SIEMPRE

D1.3 Models and algorithms for analysis of creative social interaction

Version	*Edited by*	*Changes*
1	UNIGE	Initial draft; Quartet and Orchestra
2	IIT	Audience and Orchestra
3	QUB	Revision and integration of new contributes
4	UPF	Quartet
4	UNIGE-CH	Audience, thermocamera and neural responses to music
5	UNIGE	Revision

TABLE OF CONTENTS

1. INTRODUCTION 3

2. TECHNIQUES FOR ANALYSIS OF STRING QUARTET 3

2.1 EXPRESSIVE CUES OF AN ENSEMBLE 3
2.2 PLAYING ALONE OR IN ENSEMBLE 5
2.3 THE IMPORTANCE OF SCORE-PERFORMANCE ALIGNMENT FOR THE ESTIMATION OF MEANINGFUL DESCRIPTORS OF MUSIC PERFORMANCE 9

3. TECHNIQUES FOR ANALYSIS OF ORCHESTRA 12

3.1 MEASURING INTER-MUSICIAN GRANGER CAUSAL RELATIONS 12
3.2 EVALUATING LEADERSHIP WITH THE KURAMOTO MODEL 15
3.3 EVALUATING LEADERSHIP IN MUSIC ENSEMBLE PERFORMANCE VIA MEASURES OF NETWORK CENTRALITY AND GAME THEORY 22

4. TECHNIQUES FOR ANALYSIS OF AUDIENCE 26

4.1 AUTONOMIC RESPONSE TO MUSIC LISTENING 26
4.2 MEASURING AUDIENCE TEMPERATURE VARIATIONS 27

5. CONCLUSIONS 30

6. PUBLICATIONS 31

1. INTRODUCTION

SIEMPRE addresses problems at many different levels – from the construction of appropriate scenarios, to the choice of instrumentation, to effective storage of high volumes of data, to the choice of inferential statistics that are suitable for showing that features of a situation depend on each other or influence each other.

A central issue for the 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 Expressive cues of an ensemble

Group gesture features
Research of UNIGE in the second year includes the study of global kinematic features of the group. For example, the picture below shows the MoCap data of a string quartet (Quartetto di Cremona recordings performed in July and September 2011): in particular the features *occupation volume* of the trunk of each musician and the vectors of the direction of each head (where the face of each musician is directed at each moment): these vectors contribute to the definition of a polygon (see picture), whose dynamics and shape may be related to joint action. Besides the heads, other vectors are extracted from different body layers: shoulders, hips, feet. EyesWeb implementation of algorithms computing these features is in progress (a paper is in preparation). This research will continue in the third year.

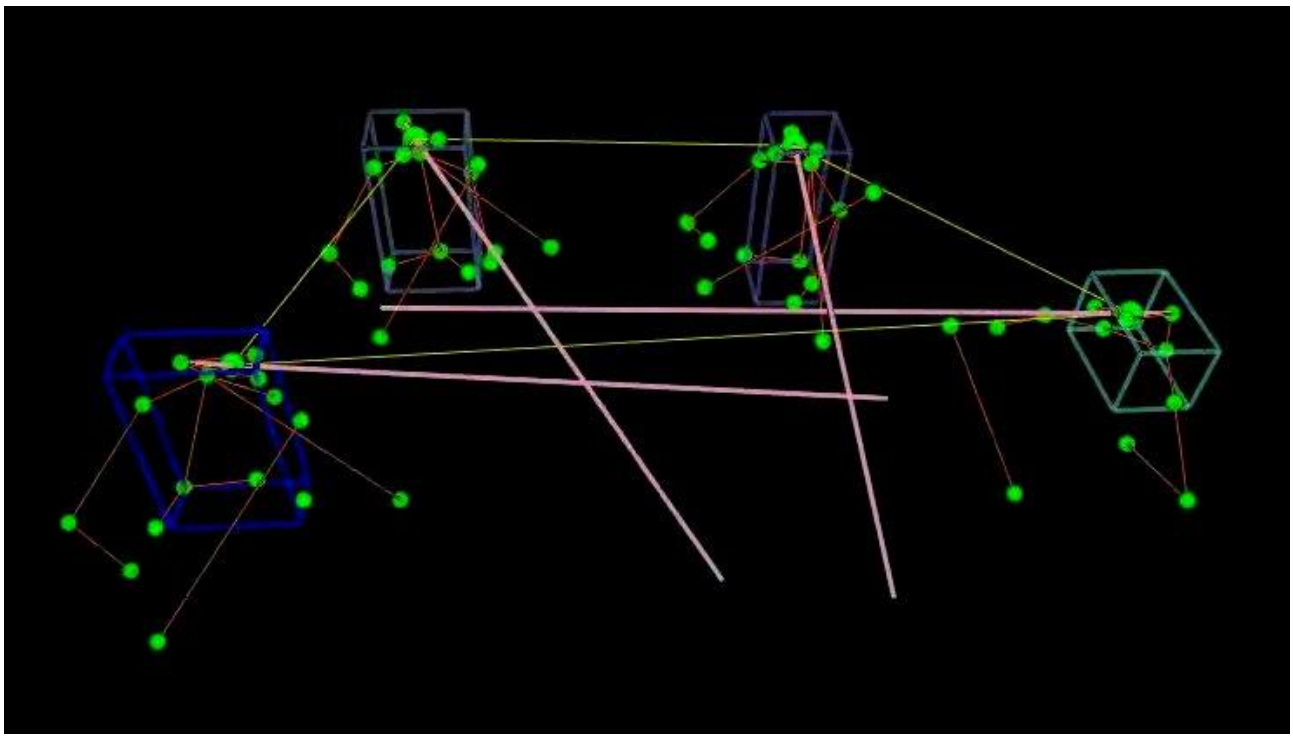


Figure 1 Real-time visualisation of global kinematic features computed in EyesWeb

Individual gesture features

Other features are under investigation on the movement of the head in the quartet experiment solo Vs joint performance. Figure 2 shows the trajectories of the heads of musicians while performing a musical phrase. (Paper submitted)

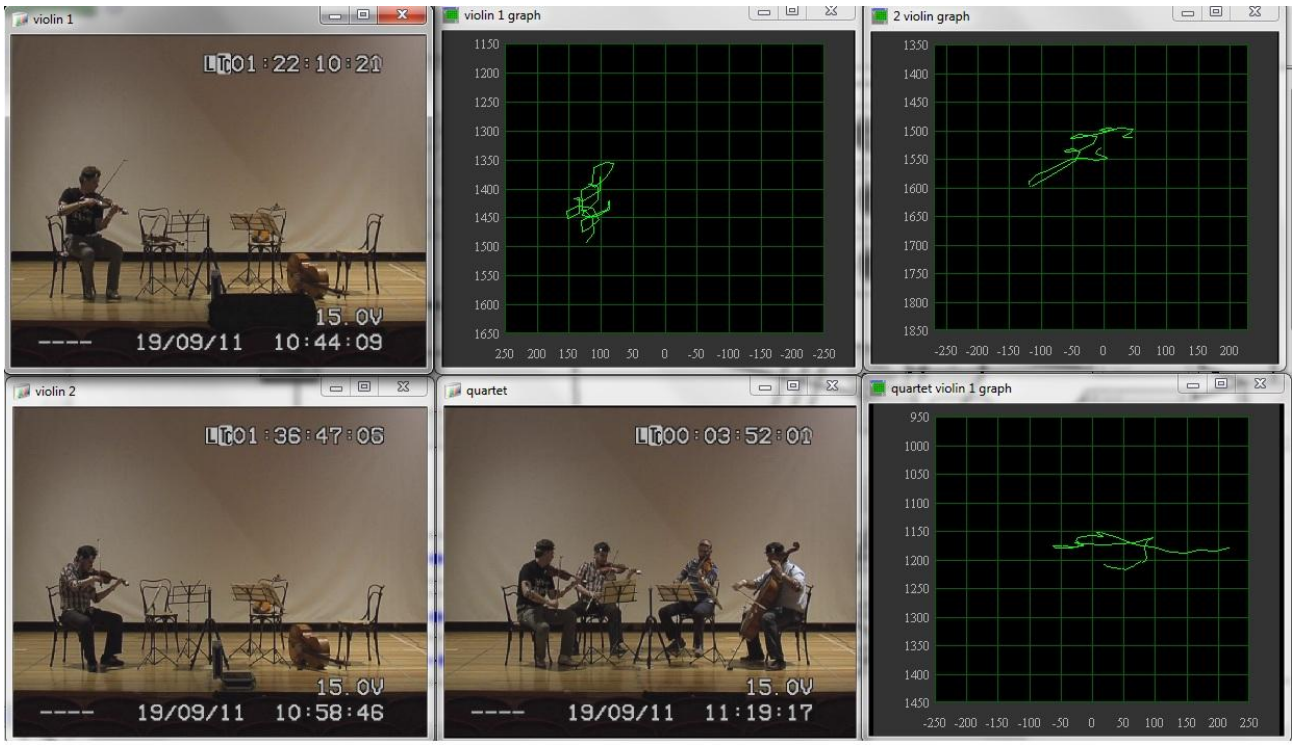


Figure 2 Trajectories of musician's heads during performance

2.2 Playing alone or in ensemble

We propose a system capable to distinguish between violinists' roles in solo and ensemble performances depending on the regularity of their head movements. Following the conceptual framework introduced in **Deliverable D1.2**, this system, illustrated in Figure 3, is composed by three components. In Solo and Ensemble Analysis modules we analyze two violin players playing in a solo and as members of a string quartet, respectively. Violinists head movements are obtained by a motion capture system, and we analyze the regularity of head movements using a measure of entropy (SampEn).

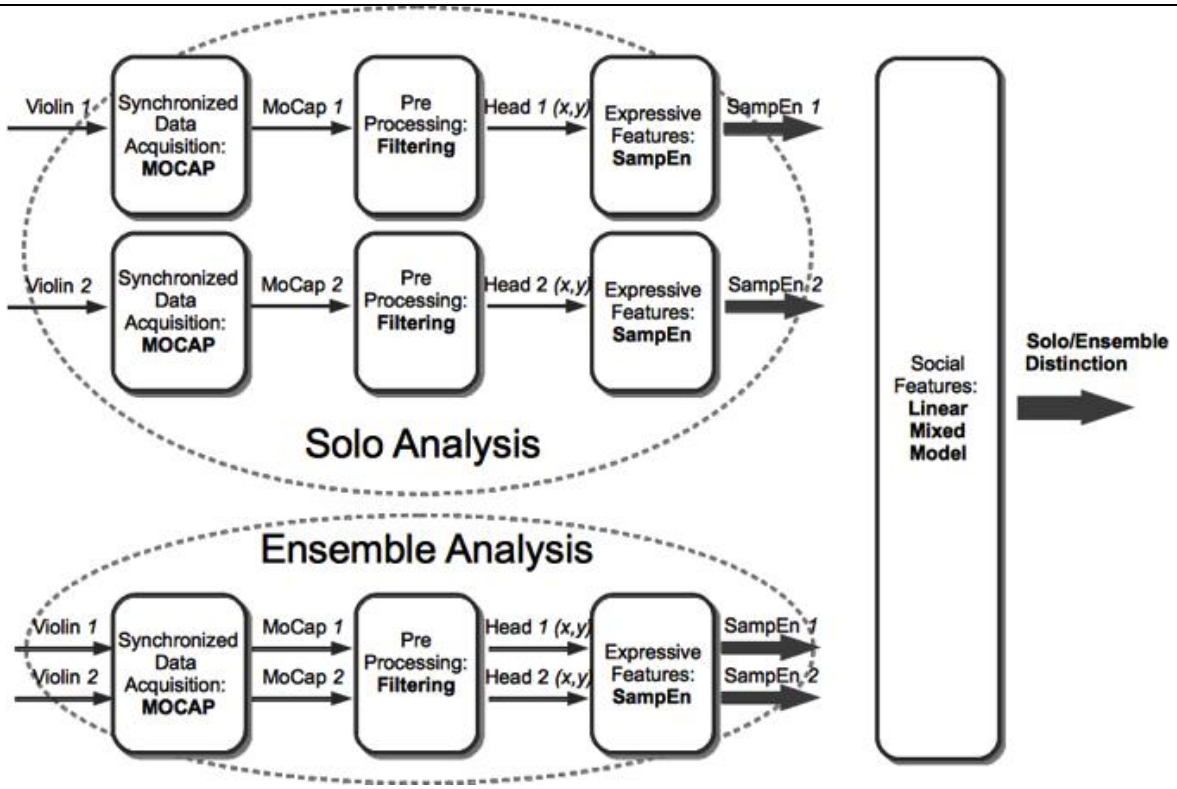


Figure 3 System Architecture

Data Acquisition (MOCAP)
Musicians' behavior is captured by means of the Qualisys Motion Capture system (www.qualisys.com). Upper part of Figure 4 shows the 3D points corresponding to the markers placed on the musicians' joints, displayed in the picture below in Figure 4. A total number of 16 mocap Qualysis markers have been defined for the analysis of the movement of musicians following preliminary series of experiment described in **Deliverable D2.1**: 13 markers are placed on the main joints of the upper-body parts and are used to create a body-segment model of the musician, based on the Dempsters anthropometric reference [2]. The remaining 3 markers are used to monitor the violin body and the bow (one marker on the scroll of the violin and two on the top and bottom of the bow). The SIEMPRE real-time applications based on EyesWeb XMI software platform [1] synchronizes Qualisys MoCap data together with video and audio data.

Figure 4 The string quartet Quartetto di Cremona during arehearsal at Casa Paganini, Genoa. Above the 3D visualizationof the Qualisys MoCap data, below, the musicianswearing the on-body sensors (mocap reflective markersand physiological sensors

Pre-Processing: Filtering

We compute the position of each musician's head center of gravity (COG) starting from three markers placed on the musician's head, two on the front and one in the back (see Figure 4). Each musician's head x and y coordinates are roto-translated with respect to the musician's body to obtain the head displacement in the forward/backward and right/left (lateral) directions. Following the recommendations in [26], analysis was conducted on the increment of the roto-translated position time series.

Sample Entropy (SampEn)

The analysis of head's movement is performed by computing the Sample Entropy (SampEn) measure, a non-linear technique initially developed by [9] and improved by [4] to quantify the behavior regularity. The main difference between this measure and traditional time and frequency domain techniques (e.g., spectral analysis) is that SampEn considers the recent movement history. For example, suppose that one swings his head forward/backward in a periodic way to support a rhythmic pulse (see Figure 5, blue line), and then he suddenly increases his head excursion at the beginning of a more animated musical phrase. SampEn (see Figure 5, red line) distinguishes this sudden change in motion whereas a traditional entropy approach will consider each frame as a separate event and compute an average value of these events entropy, ignoring the dynamic of the input signal.

SampEn (Sample Entropy) was developed within the conceptual framework of the Kolmogorov-Sinai (KS) entropy for time series analysis where entropy is defined as a quantity measuring the mean rate of new information production [9]. Pincus [7] and Richman & Moorman [9], compute the K-S entropy for real-world, noisy time series of finite length. High values of SampEn indicate disorder, smaller values indicate greater regularity. SampEn has been applied to a variety of physiological data (heart rate, EMG, see [10] for a review). Most recent applications deal with behavioral data (e.g., investigating postural control mechanisms) and some specifically address affective and social dynamics [3], [5].

Given a standardized one-dimensional discrete time series of length *N*, $X = \{x_1, \ldots, x_i, \ldots, x_N\}$:

1) construct vectors of length m (similarly to the time delay embedding procedure [6,11]):

$$u_{i(m)} = \{x_i, \ldots, x_{i+m-1}\}, 1 \leq i \leq N - m \qquad (1)$$

2) compute the correlation sum $U_i^m(r)$ to estimate similar subsequences (or *template vectors*) of length *m* within the time series:

$$U_i^m(r) = \frac{1}{(N-m-1)} \sum_{i=1, i \neq j}^{N-m} \theta(r - \|u_i(m) - u_j(m)\|_\infty) \qquad (2)$$

where $u_i(m)$ and $u_j(m)$ are the template vectors of length *m* formed from the standardized time series, at time *i* and *j* respectively, *N* is the number of samples in the time series, *r* is the tolerance (or radius), θ is the Heaviside function, and $\| \|_\infty$ is the maximum norm defined by $max_{0 \leq k \leq m-1} |x_{j+k} - x_{i+k}|$

3) calculate the average of U_i^m, i.e., the probability that two vectors will match in the m-dimensional reconstructed state space

$$U^m(r) = \frac{1}{(N-m)} \sum_{i=1}^{N-m} U_i^m(r) \qquad (3)$$

4) set *m = m + 1* and repeat steps 1-4

5) calculate the sample entropy of X_n

$$SampEn(X_n, m, r) = -ln \frac{U^{m+1}(r)}{U^m(r)} \qquad (4)$$

Equation 4 corresponds to the conditional probability that *m*-length subsequences in the time series remain similar (as defined by Equation 3) when one more point (*m + 1*) is added to those subsequences. SampEn computes the negative natural logarithm of this conditional probability. Small values of SampEn indicate regularity.

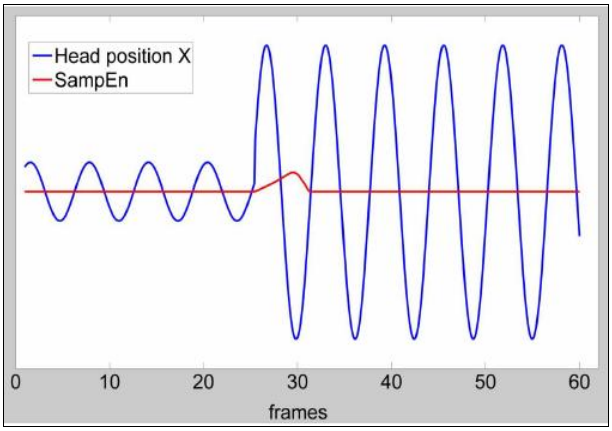


Figure 5: Sampen Example

Linear Mixed Model

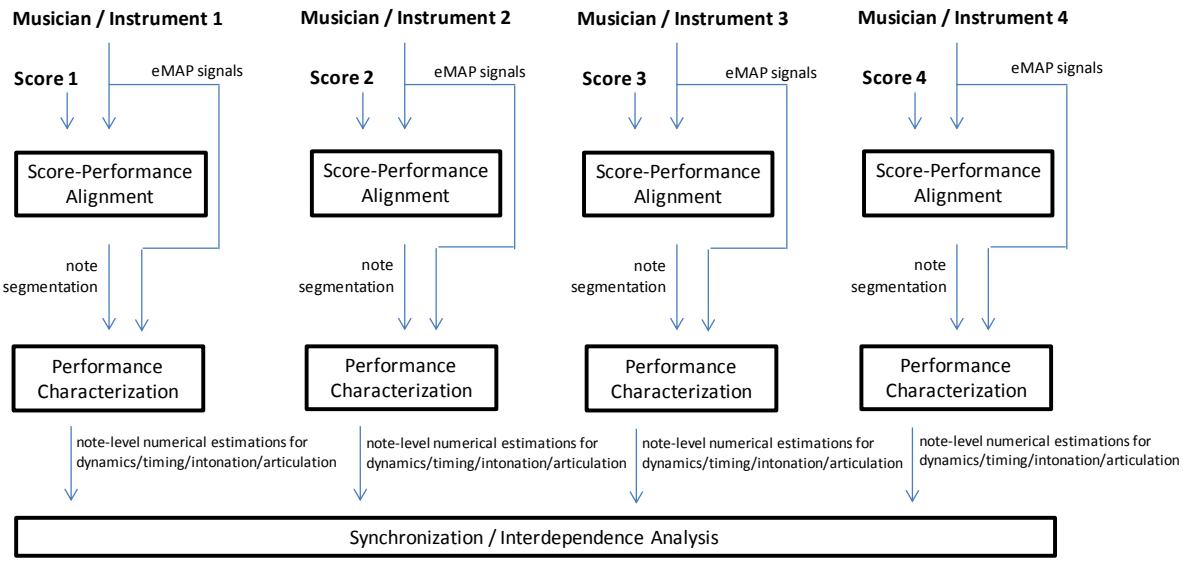
Starting from the SampEn values, we evaluate the effects of Musician (first violin Vs. second violin) and Condition (Solo Vs. Ensemble) on head movement regularity: for example, is the regularity of first violin's head movement significantly higher in the solo performance than in the ensemble performance? A Linear Mixed Model (LMM) is implemented to estimate *fixed* effects of each predictor (Musician and Condition) and *random* effect (variations observed for each musician taken individually over the different blocks of the experiment) [12]. The model is similar in many respects to ordinary multiple regression, with the difference that it allows correlation between the observations. This model further extends repeated measures models to allow an unequal number of repetitions, e.g., unbalanced experimental data where a different number of performances are recorded for each musician.

References

[1] A. Camurri, P. Coletta, G. Varni, and S. Ghisio. Developing multimodal interactive systems with EyesWeb XMI. In Proceedings of the 7th international conference on New interfaces for musical expression, pages 305–308. ACM Press New York, NY, USA, 2007.
[2] W.T. Dempster and G.R.L. Gaughran. Properties of body seg- ments based on size and weight. American Journal of Anatomy, 120(1):33–54, 1967.
[3] D. Glowinski, P. Coletta, G. Volpe, A. Camurri, C. Chiorri, and A. Schenone. Multi-scale entropy analysis of dominance in social creative activities. In Proc. of the Intl ACM Conf on Multimedia, pages 1035–1038. ACM, 2010.
[4] RB Govindan, JD Wilson, H. Eswaran, CL Lowery, and H. Preißl. Revisiting sample entropy analysis. Physica A: Statistical Mechanics and its Applications, 376:158–164, 2007.
[5] J. Kim and E. Andre´. Four-Channel Biosignal Analysis and Feature Extraction for Automatic Emotion Recognition. Biomedical Engineering Systems and Technologies, pages 265–277, 2008.
[6] NH Packard, JP Crutchfield, JD Farmer, and RS Shaw. Geometry from a time series. Physical Review Letters, 45(9):712–716, 1980.
[7] S.M. Pincus. Approximate entropy as a measure of system complexity. Proceedings of the National Academy of Sciences of the United States of America, 88(6):2297, 1991.
[8] S. Ramdani, B. Seigle, J. Lagarde, F. Bouchara, and P.L. Bernard. On the use of sample entropy to analyze human postural sway data. Medical engineering & physics, 31(8):1023–1031, 2009.
[9] J.S.RichmanandJ.R.Moorman.Physiologicaltime-seriesanalysis using approximate entropy and sample entropy. American Journal of Physiology- Heart and Circulatory Physiology, 278(6):H2039, 2000.
[10] A.J.E. Seely and P.T. Macklem. Complex systems and the technology of variability analysis. Crit Care, 8(6):R367–84, 2004.
[11] F. Takens. Detecting strange attractors in turbulence. Dynamical systems and turbulence, Warwick 1980, pages 366-381, 1980
[12] R.A.McLean,W.L.Sanders,andW.W.Stroup.Aunifiedapproach to mixed linear models. American Statistician, pages 54–64, 1991

2.3 The importance of score-performance alignment for the estimation of meaningful descriptors of music performance

In order to extract exact timings, intonation, dynamics levels, and articulation degrees from a musician's performance, note-level analysis of extracted data needs to be conducted. When available, the techniques employed for estimating numerical values for each individual musician's performance dimensions (timing, intonation, dynamics, and articulation) during a performance make necessary to carry out score-performance alignment. Those descriptors are estimated at note level, allowing for high granularity in the analysis. Such temporal granularity (defined by the notes in a musical score to be performed) is going to enable the analysis of synchronization and inter-dependence between musicians, and therefore provide means for studying social interaction. The overall process is illustrated in the following figure for the scenario of string quartet.



Schematic representation of the methodology for the analysis of synchronization/interdependence among musicians (in the string quartet scenario) via estimating numerical descriptors for performance (musical execution) parameters after note-level score-performance alignment.

Method for aligning score to recorded performance data

A very important step in analyzing music performance is quantifying the differences between the performance of the musician and the musical score. Since a real performance can never be a perfect execution of the score, such a comparison requires that the nominal score is somehow temporally aligned to the performance first. This means that for every note in the score, its onset and offset time must be detected in the recorded performance; while in a short recording this can be done by manual inspection, recordings with large amounts of data such as the ones analyzed in this project make the manual annotation of each note onset and offset time an extremely arduous and time-consuming task. For this reason, it is imperative that such a procedure is carried out automatically via a computational method. Two different methods of performance-to-score alignment have been implemented and used thus far; the first method utilizes only the audio signal and is more suited for a basic, simplistic analysis, while the second method utilizes both the audio signal as well as the instrumental gesture features and yields a far more accurate analysis.

In the first method, the score is converted to an audio format using an open-source MIDI synthesizer. Having both the score and the recorded performance represented as audio, we extract Chroma features from both signals. Chroma features are a representation for music audio in which the entire sound spectrum is projected onto a 12-bin histogram, with each bin representing the 12 distinct semitones (or

chroma) of the musical octave – thus describing the tonal progression of an audio file. The above features are then given as an input to a dynamic programming routine (dynamic time warping), which calculates the optimal alignment between the score and the recorded performance. While this method is capable of performing a 'rough' alignment (useful for tasks such as analyzing the role of the score structure in the performance), its accuracy does not permit analyzing note-level events.
For the second method, we exploit the recorded motion capture data and its derived performance descriptors (bow position, bow transversal velocity, bow pressing force etc.). In this method bow direction changes as well as more subtle measurements, such as an estimation of the applied bow force, help in detecting note change events, combined with features extracted from the audio (such as the estimated fundamental frequency and the root mean square energy of the recorded sound). These features are given as an input to a dynamic programming routine (the Viterbi algorithm), which is capable of performing a much more accurate alignment with an error margin of a few milliseconds.
It must be noted that both methods are semi-automatic; i.e. the output must be always checked manually in order to verify that the alignment is correct. In general, the usage of the first method is proposed for cases where detailed instrumental gesture data does not exist, or when a large error margin is acceptable.

Analysis of timing synchronization in string quartet performance
In ensemble performances audio-visual cues helps musicians to reach a perfect synchronization among the parts. To understand how those clues affect performance timing we need to estimate musical tempo independently to other factors.

The score-performance alignment obtained by applying the methods described above can be used to quantify an overall performance tempo in beats per minute (bpm) by computing a ratio of total number of beats and performance duration in minutes. More importantly, we can derive a tempo evolution in time (tempo curve) by quantifying tempo in consecutive time-lagged windows. Even when tempo is perceived as steady, the analysis of onset time sequences reveals a significant variability on the duration of performed notes. This means, in particularly, that there is generally a significant difference between the expected on time onset (on time with respect to a locally stable tempo) and the performed onset. These discrepancies of onset times are part of any human performance and should not be considered as performance errors. It is our objective to use such timing deviations as a central part of our study of interdependence in ensemble performance (more particularly in the string quartet scenario).

We divide the temporal analysis of a performance in two parts:

- Macro-Tempo (or just 'tempo'): the tempo experienced by a listener in a relatively long period of time (up to about 10 s), it is thus considered to change rather slowly (the evolution of tempo in time is referred as tempo curve)
- Micro-tempo (or just timing), comprising of slight anticipation of note events followed by a deferral of subsequent events in a way that the result does not contribute to macro-tempo.

In the string quartet scenario, four musicians are playing their individual parts as part of an ensemble. Macro/micro-tempo are respectively estimated (a) for each individual musician, by means of performing the analysis on each individual performance, by attending to individual score-performance alignment; and (b) for the ensemble as a whole, by performing the analysis on a derived joint score and joint performance. The joint score is obtained by treating the score score as a temporal succession of chords, thus gathering together simultaneous notes played by the different voices. The joint performance can be derived by averaging the onset/offset of each simultaneous note on the performance, thus finding a mean onset/offset time for each chord.

In addition, outperforming the same analysis on individual parts played solo enables not only to objectively compare it with the ensemble in terms from the perspective of macro-tempo, but also from the micro-tempo point of view, since we can abstract micro-tempo in the two cases by relating it to the bottom line tempo curve.

Analyzing the way musicians (both individually and jointly) keep their tempo has great potentialities for understanding the entraining process governing the performance. Discrepancies among individual/joint tempo curves can help also identifying if there is, at each moment, a clear leader, driving the performance. Also, discrepancies among musicians in terms of micro-tempo provide substantial clues for understanding the communication process in several musical contexts. A preliminary case of study has been accepted for CMMR12 (see reference section) where tempo and timing of a musical excerpt for string quartet have been compared for ensemble and solo cases.

Protocol for the extraction of inter-dependence measures from a string quartet ensemble

As mentioned before, our method of analysing music performance in an ensemble is to quantify the differences of each musician's performance from the corresponding score. Since there are many aspects of the performance that can be studied, and since one of the main characteristics of the studied phenomenon is its variability along time, our analysis yields a set of time series for each musician which describes the performance in terms of timing (tempo), dynamics intensity, intonation, timbre et cetera.

The temporal evolution of these time series is driven by several underlying processes, one of them being the inter-dependence between the members of the ensemble. We employ several computational methods capable of measuring interdependence between time series, each one suited for different types of data.

One way of categorizing such methods is whether they are suited for linear or non-linear dependencies between the time series. Another way is whether they are symmetric or directional, i.e. not only characterizing the strength of the inter-dependence but also detecting which time series is causing/influencing which. The methods currently employed are the following:

- Pearson product-moment correlation coefficient (linear, symmetric)
- Mutual Information (non-linear, symmetric)
- Granger causality (linear, directional)
- Non-linear coupling coefficient (non-linear, directional)

It must be pointed out that in a music piece, the relationships between musicians and their functional roles within the ensemble (leader, follower) are susceptible to changes according to the structure of the composition. For that reason, we also place our focus on the variation of inter-dependence strength and direction along time. This aspect is here studied by applying the above methods on overlapping segments of the performance, in order to capture this particular evolution along time and music structure.

3. TECHNIQUES FOR ANALYSIS OF ORCHESTRA

3.1 Measuring inter-musician Granger causal relations

In this section we first describe the Granger causality method and then describe two Granger causality-based metrics we proposed to measure the conductor's driving force on the violinist and the inter-violinist interaction strength respectively.

Modeling human motion using linear predictors

Granger causality, in its standard and linear formulation, is based on (linear) Autoregressive Models (AR). AR models belong to the family of the Linear Dynamical Systems, which has been extensively used in modeling human motion [3-6]. An AR(k) model of a time series x is defined as:

$$(1) \quad x(t) = \sum_{j=1}^{l} a_j x(t-j) + \varepsilon_R(t)$$

where $x(t)$ is the value of the time series x at time t, l is the order of the model (i.e., the length of the history observed in the model), a_j ($j = 1, \dots, l$) are the weights for the history (the model parameters), and $\varepsilon_R(t)$ is the residual (prediction error).

There are two widely used criteria for selecting the optimal order of a linear predictor (i.e., the order that guarantees the best goodness of fit of the model): the Akaike's Information Criterion (AIC) [7] and the Schwarz's Bayesian Information Criterion (BIC) [8]. The parameters a_j can be computed by using Ordinary Least Squares.

Since Granger causality is based on AR models the validity of the inferred causal relations depends on the validity of the AR models (more specifically of the unrestricted AR models, see next section). To assess the validity of an AR model different tests can be carried out, ranging from tests of the non-correlation of the residuals to tests of the goodness-of-fit of the model (for example, the goodness-of-fit can be measured as the sum of squares of the residuals).

Granger causality

A time series X is said to "Granger cause" a time series Y, if the past values of X provide statistically significant information to predict the next value of Y [1]. The prediction is computed using AR models. Two AR models are required: an unrestricted AR model where the history of all time series is assumed to contribute to the prediction of the current value of a time series; and a restricted AR model where the time series whose causality value (on the other time series) is computed is excluded from the history. Given two time series X and Y, the unrestricted model is defined as:

$$(2) \quad \begin{aligned} x(t) &= \sum_{j=1}^{l} a_{U,j} x(t-j) + \sum_{j=1}^{l} b_{U,j} y(t-j) + \varepsilon_U(t) \\ y(t) &= \sum_{j=1}^{l} c_{U,j} x(t-j) + \sum_{j=1}^{l} d_{U,j} y(t-j) + \eta_U(t) \end{aligned}$$

While the restricted model is defined as:

$$(3) \quad \begin{aligned} x(t) &= \sum_{j=1}^{l} a_{R,j} x(t-j) + \varepsilon_R(t) \\ y(t) &= \sum_{j=1}^{l} d_{R,j} y(t-j) + \eta_R(t) \end{aligned}$$

Then the magnitude of the causality from X to Y and from Y to X can be measured respectively as:

$$(4) \quad \mathcal{F}_{x \to y} = \ln \frac{\mathrm{H}_R}{\mathrm{H}_U}, \qquad \mathcal{F}_{y \to x} = \ln \frac{\mathrm{E}_R}{\mathrm{E}_U}$$

where E and H are the model error variances:

(5) $$E_R = var(\varepsilon_R(t)),\ E_U = var(\varepsilon_U(t)),$$
$$H_R = var(\eta_R(t)),\ H_U = var(\eta_U(t))$$

Once the Granger causality values have been computed, we need to test their statistical significance, i.e., we need to infer the significant causal relations. A significance test can be done by carrying out an F-test of the null hypothesis that the model parameters referring to the time series of which we compute the "causal strength" (on the other time series) are all zero (e.g., parameters $b_{U,j}$ in model (2) to test the significance of $\mathcal{F}_{y\to x}$). When more than two time series are analyzed some corrections (e.g., the Bonferroni correction) are applied to the F-test.

When the interaction of more than two time series is addressed, repeated pair-wise Granger causality computations can lead to misleading results. To avoid that, a simple extension of Granger causality, sometimes referred to as Conditional Granger causality, has been proposed by Ding et al., 2006 [9]. Suppose we have three time series X, Y and Z, then the Conditional Granger causality from Y to X given Z is defined as the log ratio of the error variance of the restricted model where only Y is excluded from the history (when modeling X) and the variance of the unrestricted model, where the history of all time series X, Y and Z is included.

The Granger causality analysis, including AR model validation and statistical tests of causal interactions, were carried out by using the "Granger Causality Connectivity Analysis" MatLab toolbox [2].

Non-linear Granger causality

Granger causal relations might be erroneously inferred (or ignored) when the linearity assumptions of the linear AR models are wrong, i.e., when there are significant non-linear interdependencies between the observed times series. Several solutions to extend Granger causality to the non-linear case have been proposed [10,11]. We implemented the non-linear Granger causality method proposed by Ancona et al., 2004 [12] and based on kernel-based AR models where kernels are radial basis functions

Conductor-to-Violinist driving force

The Conductor-to-Violinist driving force was computed as follows.

At each observation window we tested whether the Granger causality values in each conductor-violinist pair were statistically significant. Then we computed the driving force, in a given piece of music, of the conductor (C) on violinist V_i as:

(6) $$\frac{\sum_j^{N_W} S(\mathcal{F}_{C_j\to V_{i,j}}) - S(\mathcal{F}_{V_{i,j}\to C_j})}{N_W}$$

where the C_j and $V_{i,j}$ are respectively the slices of the conductor and of the V_i violinist time series at window j, N_w is the overall number of observation windows and

(7) $$S(\mathcal{F}_{x\to y}) = \begin{cases} 1 \text{ if } \mathcal{F}_{x\to y} \text{is significant} \\ 0 \qquad\qquad \text{otherwise} \end{cases}$$

As mentioned above, a free parameter of the analysis we carried out was the length of the observation. Different window sizes result in different N_w values and different driving force values. The results reported are obtained with a 5-second window, as this window size was the size that produced the best goodness of fit (in terms of average squared residuals) of the (unrestricted) AR model.

Note that according to our definition the conductor's driving force mainly depends on the number of times the conductor significantly exerts his influence on the violinist rather than on the magnitude of the conductor's influence. Finally, in this experiment and in experiment 2 (see below) the AR model order was fixed and set to 10. A constant model order was used to guarantee that the autoregressive models relied on the same 'history' independently of the conductor so that no bias due to different histories could affect the comparison between the two conductors. The fixed model order was selected using the following method. We first used AIC and BIC to compute model orders for each observation

window (within each execution). At each observation window the two model orders selected by the AIC and the BIC criterions are the orders that minimize AIC and BIC respectively. The average model order (averaged over each piece of music) ranged between 2 and 10. The highest value (10) was then chosen as the fixed model order used for all the autoregressive models. Within the 2-10 range, the maximum value 10 is the one by which the complexity of the model is largest (and so, in theory, its goodness of fit on the training data) and it is motivated by the fact that by using that value we never miss any relevant history at any observation window (assuming that AIC and BIC capture all and only the relevant history). When comparing the different model orders in the 2-10 range the model order 10 produced the best goodness of fit, as expected.

Inter-violinist interaction

The inter-violinist interaction strength was defined as:

$$(8) \qquad \frac{\sum_{k}^{N_W}\sum_{i}\sum_{i\neq j} S(\mathcal{F}_{V_{i,k}\to V_{j,k}|\,C})}{(N_W/(N_v(N_v-1))}$$

where N_v is the number of violinists and $N_v(N_v - 1)$ is a normalization term. Note that in this case the use of the Conditional Granger causality is mandatory. Using a non-conditional Granger causality would mean ignoring the influence of the conductor and misinterpret it as influence from one musician to the other, so, e.g., simple delays between two violinists would be erroneously interpreted as causal relations.

REFERENCE

(1) Granger CWJ (1969) Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37: 424-438.

(2) Friston K (2011) What Is Optimal about Motor Control? Neuron 72: 488-98.

(3) Del Vecchio, D., Murray, R. M. & Perona, P. Decomposition of Human Motion into Dynamics Based Primitives with Application to Drawing Tasks. *Automatica* **39** (2003).

(4) Lu, C. & Ferrier, N. J. Repetitive motion analysis: segmentation and event classification. *IEEE Trans. Pattern Anal. Machine Intell.* **26**, 258-263 (2004).

(5) Bissacco, A. Modeling and learning contact dynamics in human motion. Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society, 2005.

(6) Bissacco, A. & Soatto, S. Classifying Gaits Without Contact Forces. Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society, 2006.

(7) Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **19**, 716-723 (1974).

(8) Schwartz, G. Estimating the Dimension of a Model. *Ann. Stat.* **5**, 461-464 (1978).

(9) Ding, M., Chen, Y. & Bressler, S. L. in *Handbook of Time Series Analysis*, (Eds. Schelter, S., Winterhalder, N. & Timmer, J.) 437-460 (Wiley, Wienheim, 2006).

(10) Freiwald, W. A. *et al*., Testing non-linearity and directedness of interactions between neural groups in the macaque inferotemporal cortex *J. Neurosci. Meth.* **94**, 105-119 (1999).

(11) Chen, Y. H., Rangarajan, G., Feng, J. F. & Ding, M. Z. Analyzing multiple nonlinear time series with extended Granger causality *Phys. Lett. A* **324**, 26-35 (2004).

(12) Ancona, N., Marinazzo, D. & Stramaglia, E. Radial basis function approach to nonlinear Granger causality of time series. *Phys. Rev. E* **70**, 056221 (2004).

3.2 Evaluating leadership with the Kuramoto model

Background

This research stems from the work by Boerner and colleagues (Boerner, 2004) on the leadership of orchestra conductors and on the work by Yoshida and colleagues (Yoshida, 2002) on soft entrainment. Whereas the former addresses ensemble playing from the point of view of entrainment, the latter faces the affective component.

The Boerner model

Boerner and colleagues showed that leadership of orchestra conductors is a combination of authority and charisma. Leadership is goal-directed. The goal is artistic quality. Cooperation is a precondition for artistic quality, i.e., to reach the goal, and musicians' skill and motivation enhance it. In the special case of orchestra, leadership success and effectiveness to get cooperation can be expressed in terms of entrainment of the music ensemble: "The decisive task of the conductor is the centralized co-ordination of orchestra playing"; "The specific nature of leadership success in the orchestra requires precise ensemble playing by the musicians" (Boerner, 2004). This applies within sections, between sections, and in relations to any other performer involved.

Following these results, given an archive of orchestra performances, possible differences in the effectiveness of the leadership between the performances can be in principle indirectly measured from the degree of entrainment that the orchestra reaches. That is, provided the same starting reference level of skill and motivations of the players, a measure of entrainment of orchestra playing reflects the ability of the conductor to carry out her task and to reach her goal. Therefore the degree of entrainment among the players is a crucial quality in orchestral performance.

Soft Entrainment

Yoshida and colleagues claim that, in real music performance, players cannot indeed completely reach the maximum level of entrainment. Instead, the alternation of high and low level of entrainment (*soft entrainment*) is deemed to make a performance more vivid and impressive because the unconscious process can create spontaneous, active feelings in the music. This is a well-known issue in automatic music synthesis, where MIDI files perfectly reproduced by a computer are usually not able to convey affective content suitably. Rather, deviations from the perfection of the regular performance are needed to convey emotion. Players exploit such a variation in the level of entrainment for expressing their artistic feeling and a good balance between high and low level of entrainment produces effective ensembles. In a musical phrase, soft entrainment follows a pattern such that entrainment has two major peaks, the lowest at the beginning of the phrase and the highest at its end (Yoshida, 2002).

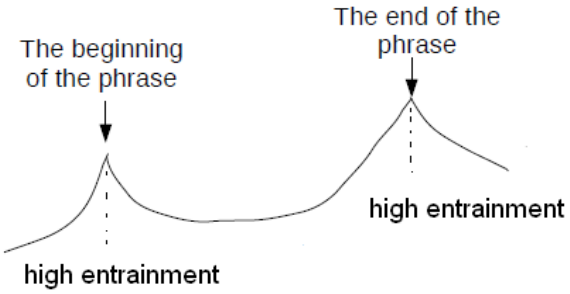


Figure 6 The pattern of soft entrainment: high entrainment occurs at the beginning and at the end of the musical phrase.

The Kuramoto Model

Playing in an orchestra requires cooperation due to a shared goal, division of roles and monitoring of progresses. This clearly implies coordination of movement making a collective rhythm emerge. Each player has her own rhythm, but through the interaction with the other players belonging to her section or to other sections, or with the conductor, she modifies this rhythm toward the collective rhythm. *Phase Synchronisation* is the simplest detectable form of entrainment among two or more signals, and in particular motion signals. Phase synchronisation in presence of rhythmic interacting units is an issue that was already faced by using the Kuramoto model.

The classical Kuramoto model

The Kuramoto model is a mathematical model broadly used to study the emergence of synchronous behaviour in a population of coupled phase oscillators. The original model describes the dynamics of every oscillator taking into account both the individual parameters of each single oscillator and the influence of the other oscillators linked by some connective arrangements with it. Given a population of N oscillators, having phase $\theta_i(t)$ and natural frequencies ω_i distributed with a given probability density $g(\omega)$, the dynamics of each oscillator is generally governed by (Acerbon, 2005):

$$\dot{\theta}_i(t) = \omega_i + \sum_{j=1}^{N} K_{ij} sin(\theta_j(t) - \theta_i(t)) \quad i = 1, ...N,$$

where $\theta'_i(t)$ is the rate of change of phase; K_{ij} and $\theta_j(t)$-$\theta_i(t)$ are the coupling strength and the phase difference between the i^{th} and j^{th} oscillators, respectively; $sin(\theta)$ is the selected coupling function.

On the one hand each oscillator tends to run independently at its own frequency. On the other hand the coupling strength tends to synchronise it to all the others. When the coupling is sufficiently weak, the oscillators run incoherently, whereas beyond a certain threshold collective synchronisation emerges spontaneously (Acebron, 2005).
Kuramoto ran his model under the following assumptions: nearly identical oscillators, mean-field (i.e., *all-to-all* coupling among oscillators and $K_{ij} = K/N$ for each i,j), time-independent coupling, a very large number of oscillators, and unimodality and symmetry of the probability density of the natural frequencies $g(\omega)$. These assumptions allow for writing the model in the previous equation more conveniently as:

$$\dot{\theta}_i(t) = \omega_i + \frac{K}{N} \sum_{j=1}^{N} sin(\theta_j(t) - \theta_i(t)) \quad i = 1, ...N,$$

This model was successfully exploited in many different contexts ranging from physics to neurology. However, in order to make the Kuramoto model more suitable to study synchronisation in realistic situations such as, for example, neuronal networks, some generalisations were needed.

Generalising the Kuramoto Model

Many generalisations of the classical Kuramoto model were proposed in the literature. These generalisations mainly concern (i) the coupling function, (ii) the connectivity arrangement, and (iii) the time-variation of the parameters in the model. As for the coupling function, Kuramoto chose the *sin* function in the formulation of his model. Nevertheless, in laboratory simulations already, the use of more complex periodic coupling functions allowed probing the phenomenon in a deeper and more

precise way. The simplest proposed approach consists of expressing the coupling function $h(\theta)$ as a Fourier series (Daido,1996):

$$h(\theta) = \sum_{k=1}^{\infty} (h_k^s sin(k\theta) + h_k^c cos(k\theta))$$

where k is the number of the series terms, and h_k^s and h_k^c the series coefficients.

Concerning connectivity arrangements, evidently real world networks of oscillators typically show a more complex structure than the all-to-all coupling assumed in the original Kuramoto model. In order to allow the study of such kinds of networks, each oscillator is represented as a node and a $N \times N$ adjacency matrix describing all the connections is incorporated in the model formulation. The coupling between nodes i and j in the network is thus expressed as $C_{ij} = a_{ij}\,\sigma_{ij}$, being σ_{ij} the coupling strength and $A=a_{ij}$ the adjacency matrix (e.g., see (Arenas, 2008) (Wang, 2009)).
As for time-variation of model parameters, this namely includes the time-variation of the natural frequency of the oscillators (i.e., $\omega_i=\omega_i(t)$) and the time-variation of the coupling strength (i.e., $K_{ij} = K_{ij}(t)$). Cumin and Unsworth (2007) investigated the effect of time-varying natural frequencies and coupling strengths with simulation studies, and, for example, the effect of a negative coupling strength, showing that increasing connectivity with negative constant coupling reduces synchronisation.

Measuring entrainment by the order parameters

In the classical Kuramoto model, the collective dynamics of the whole population of oscillators is measured by the complex order parameter (Acebron, 2005):

$$re^{I\psi} = \frac{1}{N}\sum_{j=1}^{N} e^{I\theta_j}$$

where $I=\sqrt{-1}$, $0 \leq r \leq 1$ measures the phase coherence of the population, and ψ is the average phase. The complex order parameter can be interpreted as the collective rhythm produced by the whole population (Strogatz, 2000). If $r \approx 1$ all oscillators are phase locked and act as a single giant oscillator. If $r \approx 0$ they add incoherently and no common rhythm is produced.

The theoretical studies proposing generalisations of the Kuramoto model often introduce modified order parameters coping with the generalisations included in the model. For example, the generalisation of the model to a coupling function expressed as a Fourier series is accompanied by the definition of an order function, which is a generalisation of the Kuramoto order parameter (Daido, 1996). Nevertheless, the literature includes many studies where the classical Kuramoto order parameter is used, even when a generalised model is adopted.
Restrepo and colleagues (2005) proposed an order parameter that can be used when the population of oscillators is arranged in a network with a general connectivity (i.e., the assumption of all-to-all coupling is not required) and the coupling terms between pairs of oscillators may be different and negative, meaning that the coupling may even drive the oscillators to be out of phase. Such an order parameter is defined as:

$$r = \frac{\sum_{i=1}^{N} r_i}{\sum_{i=1}^{N} d_i}$$

where r_i is a positive real valued local order parameter defined by:

$$r_i e^{I\psi_i} = \frac{1}{N}\sum_{j=1}^{N} C_{ij} e^{I\theta_j}$$

and

$$d_i = \sum_{j=1}^{N} C_{ij}$$

The interpretation of the Restrepo's order parameter is similar to the classical one (e.g., see (Wang, 2009). When the oscillators are not synchronised, the order parameter *r* has near-zero values; whereas if the oscillators are fully synchronised, *r* becomes near to unity.

The Syncronisation Index (SI)

When there are two or more interacting populations, an overall order parameter has to be computed by combining together the order parameters *r* of each single population. The overall order parameter *R* is defined as the time-average of the single order parameters *r*. However *R* is not a good measure of synchronisation when the populations are phase-locked with a not in-phase locking. To tackle this problem, Kiss et al. (2008) conceived two further overall parameters: (i) R_{mod} a modified version of *R*, and (ii) the Synchronisation Index (*SI*).

SI has the edge over R_{mod} to drop off quickly to zero when the groups lack synchronisation. Due to this good sensitivity to interpopulation synchronisation *SI* was chosen for this study.
SI is defined as follows:

$$SI = \frac{\sigma(\langle r_1 \rangle + \langle r_2 \rangle)}{2}$$

where σ measures the phase synchrony between the two groups and is computed as $\sigma = 1\text{-}S/S_{max}$,

with *S* being the Shannon entropy of the cyclic phase distribution and S_{max} the maximal entropy. σ has an approximately unitary value for phase-locked groups and a close to zero value for phase-drifting.

The metric of effectiveness of leadership

As counted in Background, a measure of effectiveness of orchestra leadership has to take into account both the overall level of entrainment and the occurrence of soft entrainment. It should be noted that this would be a measure of the effects of the conduction, being rather independent on the causes (i.e., on which specific means the conductor uses to perform her task successfully).
Here, a metric of effectiveness of leadership (*EoL*) is described. This metric simply combines entrainment and soft entrainment by using a weighted sum of these as follows:

$$EoL = \alpha\frac{\sum_{j=1}^{n\_ph} \overline{SI}_{thr}}{n\_ph} + \beta\frac{\sum_{j=1}^{n\_ph} soft\_ent}{n\_ph}$$

where the two terms of the sum are the number of times (phrases) in which *SI* averaged along a phrase overcomes a prefixed threshold *t* of entrainment, and soft entrainment occurs in a phrase of the excerpt, respectively. The overall number of phrases is *n_ph*. The coefficients α and β range from 0 to

1 and allow to weight the entrainment and soft entrainment terms.

Entrainment and effectiveness of leadership
EoL was computed on the orchestra data from the first series of experiments (see D4.1 for details). The generalised Kuramoto model, already successfully fitted, was (see also D4.1):

$$\dot{\theta}_i(t) = \omega_i + \sum_{j=1}^{N} C_{ij}[sin(\theta_j(t) - \theta_i(t)) + cos(\theta_j(t) - \theta_i(t))] \quad i = 1, ...N$$

By using this model, entrainment was computed using the Restrepo's order parameter *r* (see above). However, in order to take into account possible negative couplings, the absolute value of the coupling strength was used in the computation, i.e.,

$$d_i = \sum_{j=1}^{N} |C_{ij}|$$

The Synchronisation Index *SI* was then computed for each bar of each single music phrase. To this aim, the time-average of the Restrepo's order parameters r_1 and r_2 for the two orchestra sections mentioned in D4.1 was computed on time windows corresponding to each bar the music phrase is composed by.
Moreover, at the beginning of each phrase, except the first one, the time-average of the order parameters was computed on a time window centred on the starting sample of the phrase and having a width of one bar, i.e., half bar before the beginning of the phrase and half after it. This was done for taking into account possible soft entrainment patterns where the first peak falls very close to the beginning of the phrase. For the first phrase only the half bar after the beginning was considered.
Similarly, at the end of each phrase, except the last one, the time-average of the order parameters was computed on a time window centred on the ending sample of the phrase and having a width of one bar. This was done for taking into account possible soft entrainment patterns where the second peak falls very close to the end of the phrase. For the last phrase only the half bar before the end was considered. Figure 7 shows the course of *SI* in a trial.
In order to compute the effectiveness of leadership (*EoL*), the number of times entrainment and soft entrainment occur in each music piece has to be computed. To this aim, the threshold *t* for entrainment was chosen equal to 0.4. As for soft entrainment, the following rule was applied:

1 Soft entrainment occurs in a music phrase (i.e., soft_ent = 1), if the values of *SI* in the phrase follow the typical soft-entrainment pattern (first peak, valley, second peak).
2 Soft entrainment does not occur (i.e., soft_ent = 0), if *SI* does not show such a pattern.
3 Soft entrainment partially occurs (i.e., soft_ent = 0.5), if either *SI* follows the typical pattern, but with a temporal shift of around a bar, or the music phrase contains two bars only and the value of *SI* for the first bar is lower that the value for the second bar.

Figure 8 depicts an example of how the soft entrainment component was computed.

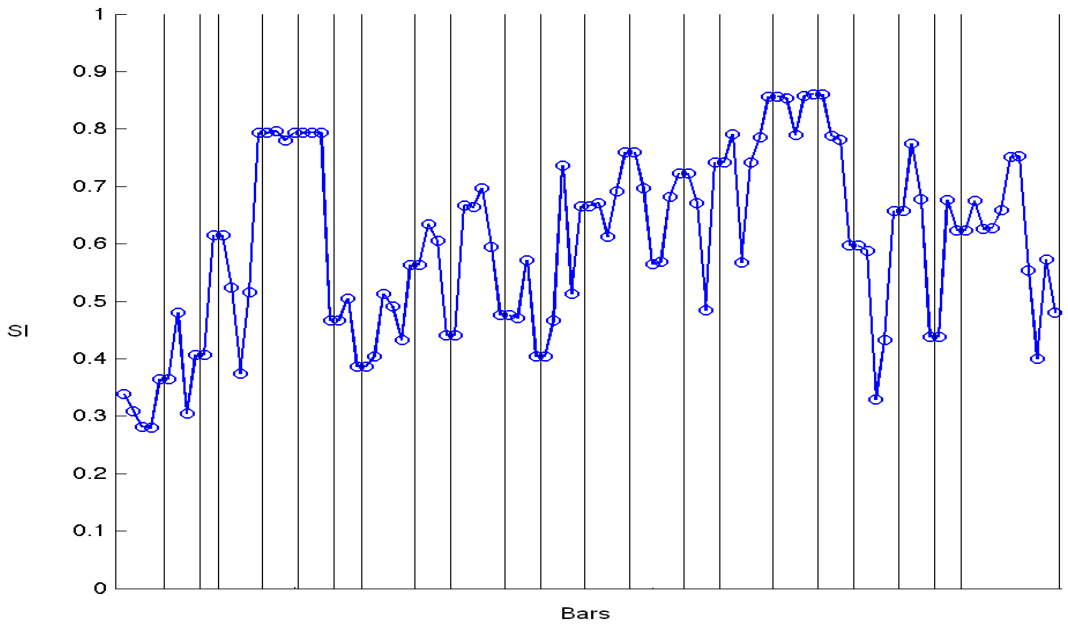


Figure 7 The course of SI in a trial

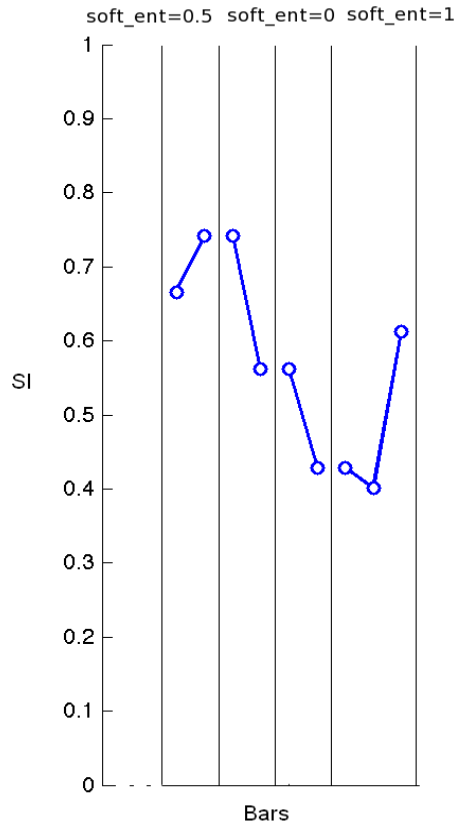


Figure 8 Example of computation of soft entrainment in a short fragment of a trial

Finally, the *EoL* metric was computed choosing the weights α and β equal to 0.5.

Table 2 summarises the results showing the overall value of *EoL* and its components (ent is the level of entrainment, soft_ent is the soft entrainment) for each excerpt and for each trial. At the present, a questionnaire-based evaluation with expert musicians (e.g. professional players, music teachers) is being carried out to assess these results.

Table 1 EoL results

Piece	Trial	Conductor	EoL	ent	soft_ent
P1	T1	new	0.67	22	7.5
P1	T2	new	0.45	14	6
P1	T1	own	0.52	20	3
P1	T2	own	0.48	18	3
P2	T1	new	0.38	6	3
P2	T2	new	0.77	11	4.5
P2	T3	new	0.69	9	3
P2	T1	own	0.60	12	2.5
P2	T2	own	0.46	9	2
P2	T3	own	0.62	12	3

References

S.Boerner, D. Krause, and D. Gebert, “Leadership and cooperationin orchestras,” Human Resource Development International, vol. 7, no. 4, pp. 465–474, 2004.

T. Yoshida, S. Takeda, and S. Yamamoto, “The application of entrainment to musical ensembles,” in II International Conference on Music and Artificial Intelligence (ICMAI), Edinburgh, Scotland, 2002.

J. A. Acebron, L. L. Bonilla, C. J. P´erez Vicente, F. Ritort, and R. Spigler, “The kuramoto model: A simple paradigm for synchronization phenomena,” Reviews of Modern Physics, vol. 77, no. 1, pp. 137–185, 2005.

H. Daido, “Onset of cooperative entrainment in limit-cycle oscillators with uniform all-to-all interactions: bifurcation of the order function,” Physica D: Nonlinear Phenomena, vol. 91, no. 1-2, pp. 24 – 66, 1996.

A. Arenas, A. Daz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, “Synchronization in complex networks,” Physics Reports, vol. 469, no. 3, pp. 93–153, 2008.

W.-X. Wang, L. Huang, Y.-C. Lai, and G. Chen, “Onset of synchronization in weighted scale-free networks,” Chaos: An Interdisciplinary Journal of Nonlinear Science, vol. 19, no. 1, p. 013134, 2009.

D. Cumin and C. Unsworth, “Generalising the kuramoto model for the study of neuronal synchronisation in the brain,” Physica D: Nonlinear Phenomena, vol. 226, no. 2, pp. 181 – 196,

2007.

S. Strogatz, “From kuramoto to crawford: exploring the onset of synchronization in populations of coupled oscillators,” Physica D: Nonlinear Phenomena, vol. 143, no. 1-4, pp. 1–20, 2000.

J. G. Restrepo, E. Ott, and B. R. Hunt, “Synchronization in large directed networks of coupled phase oscillators,” Chaos, vol. 16, 2005.

3.3 Evaluating leadership in music ensemble performance via measures of network centrality and game theory

In the last year of the project, UNIGE will investigate further techniques to measure the degree of leadership of musicians in the orchestra experiments by exploiting tools and methods from graph theory and game theory.

The starting point of the proposed feasibility study is the representation of the orchestra as a network, where the nodes are the conductor and the musicians. The weighted connections between pairs of nodes provide a measure of the "degree of similarity" and the "degree of influence" between each pair of nodes. In the simplest model that we plan to consider, such connections and their weights are fixed and a-priori given. In more complex and more realistic models, they are allowed to be time-varying and estimated, e.g., from the available physical measurements. In general, a-priori and a-posteriori knowledge can be combined to identify the weights. Once the orchestra has been represented as a network, the degree of leadership of its nodes can be measured by applying tools from the theory of social networks and game theory. Finally, the models can be validated by experts providing a ground-truth knowledge of the degree of leadership or via analysis of the music score.

Measures of leadership based on the theory of social networks

These measures of leadership are based on the idea that there is a strong connection between centrality of a node in a social network and the relevance of its role in the network itself (Hanneman and Riddle, 2005). Indeed, a more "central" node has more opportunities to interact and is less dependent on the other nodes. So, *a measure of centrality of a node in a network can be interpreted also as a measure of its degree of leadership*. The following are possible ways to measure centrality in the simple model of a an undirected and unweighted network (Hanneman and Riddle, 2005).

Degree centrality. This measure is based on the intuitive fact that nodes with more direct links have more opportunities to interact. *The most central node is the one with the highest degree*, where the degree of a node is the number of its neighbors in the network. This measure is particularly simple to compute, since it requires only a *local knowledge* of the network. However, locality is at the same time a limitation, since the global structure of the network is not taken into account.

Closeness centrality. It is based in the idea that nodes close to more nodes have more opportunities to interact. This is quantified via the *lengths of shortest paths* connecting them. For an undirected and unweighted network, the length of a shortest path between two generic nodes A and B is the minimum number of edges that are required to reach B when starting from A. Closeness centrality of a node is then defined as the inverse of the sum of the lengths of the shortest paths between itself and all the other nodes in the network. *The most central node is the one that minimizes the sum of the lengths of such shortest paths*. With respect to degree centrality, this measure of centrality takes into account the structure of the *whole network*, at the expense of a larger computational burden required to compute the lengths of the shortest paths. Moreover, it can applied only to the connected components of the network.

Betweenness centrality. According to this measure of centrality, *the most central node is the one with the capability of being an intermediary for the largest number of pairs of other nodes*. This can be

quantified as follows. Suppose, for simplicity, that there exists only one shortest path between any given pair of nodes. Then, for each node i, let SPi be the number of the shortest paths between any (unfixed) pair of other nodes, going through i (the definition can be extended to the case in which there exist possibly more than one shortest path between any pair of nodes). The most central node is *the one that maximizes the* SP_i. Betweenness centrality can be obtained as a by-product of an algorithm computing closeness centrality. A drawback is that a large proportion of nodes in a network generally does not lie on any shortest path, therefore they receive the same score 0.

Eigenvalue-based centrality. This is indeed a family of measures of centrality. They take into account the structure of the whole network by representing it via its *adjacency matrix* and finding suitable *eigenvectors/eigenvalues*. In such a way, tools from spectral graph theory (Chung, 1997) come into play. The adjacency matrix for an undirected graph is defined as follows: in entry (i,j) it contains 1 if the two nodes i and j are directly connected by an edge, otherwise 0. An example of eigenvalue-based centrality measure is *Katz centrality*, which is defined in terms of a vector of "prestiges", whose dimension is equal to the number of nodes in the network. The "prestige" of a node is recursively defined as the sum of the "prestiges" of its neighbors divided by their respective degrees. It can be proven that the vector of "prestiges" is a unit eigenvector of the (normalized) adjacency matrix. Google's ranking algorithm, PageRank, is based on these ideas.

More general measures of centrality have been recently proposed in (Opsahl et al., 2010). In general, there is no measure of centrality that is always superior to the others. Indeed, the performances depends on the particular context (Hanneman and Riddle, 2005), though in some cases it is possible to justify a-priori the suitability of a particular measure to a specific context (Opsahl et al., 2010).

As to the choice of the model, we plan to use at first undirected and unweighted networks; in a refined analysis, the identification of the weights/directions of the links has to be addressed. Indeed, a more realistic model of the orchestra experiments is a network with weighted and directed connections. Some measures of centrality were extended in the literature to weighted/directed networks. In our investigations, we plan to exploit such extensions and/or to develop new ones, better-suited to the orchestra context. For instance, in the definition of degree centrality, the degree of a node can be replaced by the sum of the weights of its edges in the case of an undirected weighted network. For a directed unweighted network it can be replaced by its in-degree or out-degree (the in-degree of a node is the number of arcs entering the node and its out-degree is the number of arcs exiting it).

Measures of leadership based on cooperative game theory

Another possible way to measure the degree of leadership of nodes is offered by *cooperative game theory*.

This approach takes the hint from a problem investigated in cooperative games, namely, finding a "fair" allocation among the players of the "social utility" of a group of cooperating players. Such an approach has been used, e.g., to identify influential users in social networks (Naramanam and Narahari, 2010) and to rank genes in gene expression networks (Moretti et al., 2010). The setting to which these ideas apply is represented by *cooperative games with transferable utility*, or *TU games*, defined as follows.

TU games. Given a finite set N of players, called *grand coalition*, one assigns a *utility* v(S) to each subset S of N, called *subcoalition*. In general, such a utility function has to satisfy some constraints. For instance, for the empty subcoalition S=∅ one has $v(\emptyset) = 0$ and for two subcoalitions S_1 and S_2 with empty intersection, $v(S_1 \cup S_2) \geq v(S_1) + v(S_2)$. The latter constraint can be interpreted as an

incentive for the players to form larger subcoalitions, hence to form the grand coalition. Moreover, *there is the possibility to transfer utility from each player to each other one*: in such games one assumes the existence of a common currency that is valued equally by all the players (hence the name of games with transferable utility, or TU games). *The problem consists in finding an allocation* $x_1, x_2, \ldots, x_{|N|}$ *among the players of the utility v(N) of the grand coalition, such that* $x_1 + x_2 + \cdots + x_{|N|} = v(N)$*, and some other requirements are satisfied*. Among these requirements, it is usually imposed that the allocation is "*fair*" (for a rigorous definition of the concept of fairness, see Tijs, 2003). Loosely speaking, this means that if player i contributes to the utility of the grand coalition more than player j, it should be rewarded with a value x_i larger than x_j. So, a player i with a very large value x_i in a fair allocation of v(N) may be interpreted as a leader, and the vector x with components $x_1, x_2, \ldots, x_{|N|}$ may be used to rank players according to the importance of their roles in the grand coalition.

The following are two possibilities of defining the vector x in such a way that suitable constraints are satisfied (in both cases, a fairness constraint is imposed).

Banzhaf power index:

$$x_i = \frac{\sum_{S \subseteq N} (v(S) - v(S \setminus \{i\}))}{2^{|N|} - 1}.$$

Shapley value:

$$x_i = \frac{\sum_{S \subseteq N} \left((v(S) - v(S \setminus \{i\})) \frac{(|S| - 1)!(|N| - |S|)!}{|N|!} \right)}{2^{|N|} - 1}.$$

In both indices, the summation is actually performed only on the subcoalitions S containing the player i. Power indices can be approximately computed via Monte-Carlo evaluation (Naramanam and Narahari, 2010), and for certain utility functions v(S) the evaluation, e.g., of the Shapley value can be significantly simplified (Aadithya et al., 2010).

TU games on networks. TU games can be defined on networks (Aadithya et al., 2010). In such games, the players are, e.g., the nodes of the network (conductor and musicians, in our context) and the subcoalitions can be identified with suitable subnetworks. Since these are games on networks, in order to define the value v(S) of a subcoalition S the topology of the network has to be taken into account. Some possible ways of defining specification of v(S) are provided in (Grofman and Owen , 1982), (Naramanam and Narahari, 2010) and (Aadithya et al., 2010). In particular, for the cases considered in (Aadithya et al., 2010), the Shapley value can be computed efficiently.

In the context of the orchestra experiments, the reason for which TU games on networks are of interest is that *the group of musicians may be modeled as a network, and the musicians as players in a cooperative game*.

The following are some points that have to be addressed in the proposed framework.

In a first step, known centrality measures/power indices will be used for the orchestra experiments. Different indices may be used together and suitably combined to estimate leadership. In a refined analysis, new centrality measures/power indices, well-suited to the context, may be developed. One

may be interested in using measures suitable to evaluate *leadership of single musicians* or others suitable to evaluate how much *the orchestra as a whole is "centralized".*

For each experimental setup, one has to *identify the network topology* (each configuration of the orchestra has to be modeled via different topologies - nodes and connections) and model the interaction/mutual influence between each pair of individuals via weights on the links. Each single node (musician/conductor) may be a set of features, associated to various quantities measured during the experiments. The weight and direction of the connection between each pair of nodes depend on the topology of the network in each experimental setup. E.g., the connection between two musicians or a musician and the conductor may be stronger when they can look at each other. The weights may be chosen via experiments or signal processing. *Audio influence* and *visual influence* between the conductor and the musicians and among the musicians have to be taken into account; so, two kinds of weights may be used. Moreover, in general the weights are *time-varying*: for a given network topology, the weight of the connection between two nodes may change, depending, e.g., on some measure of similarity between the various features associated with the two nodes.

For each experimental setup, various centrality measures/power indices can be compared. An open point is the choice of different centrality measures/power indices for audio and visual interaction, versus the use of a unique centrality measure/power index. A comparison can be done among the various indices and the models can be *validated* by the analysis of the performance by experts and/or processing the experimental measurements.

References

K. V. Aadithya, B. Ravindran, T. P. Michalak, and N. R. Jennings, "Efficient Computation of the Shapley Value for Centrality in Networks", in Proc. 6th Workshop on Internet and Network Economics (WINE 2010), Lecture Notes in Computer Science, vol. 6484, pp. 1-13, Springer, Berlin Heidelberg, 2010.

F. R. K. Chung, Spectral Graph Theory. AMS, 1997.

B. Grofman and G. Owen, "A Game Theoretic Approach to Measuring Degree of Centrality in Social Networks", Social Networks, vol. 2, pp. 213-224, 1982.

R. A. Hanneman and M. Riddle, "Introduction to social network methods". Riverside, CA: University of California, Riverside, 2005 (online book).

S. Moretti, V. Fragnelli, F. Patrone, and S. Bonassi, " Using Coalitional Games on Biological Networks to Measure Centrality and Power of Genes", Bionformatic, vol. 26, pp. 2721-2730, 2010.

R. Naramanam and Y. Narahari, "A Shapley Valued-Based Approach to Discover Influential Nodes in Social Networks", IEEE Transactions on Automation Science and Engineering, vol. 99, pp. 1-18, 2010.

T. Opsahl, F. Agneessens, and J. Skvoretz, "Node Centrality in Weighted Networks: Generalizing Degree and Shortest Paths", Social Networks, vol. 32, pp. 245-251, 2010.

S. Tijs, Introduction to Game Theory. Hindustan Book Agency, New Delhi, India, 2003.

4. TECHNIQUES FOR ANALYSIS OF AUDIENCE

Audience response to musical performance is one of the key parts of SIEMPRE. To study it, we need methodologies that extract features related to entrainment and emotional contagion in listeners (audience member) as they relate to musical content and specific performance situations. Some of these techniques have been developed in the context of work that deals directly with audiences, others in the context of studies with individual listeners, which provide the large quantities of controlled data that are needed to develop measures.

4.1 Autonomic response to music listening

Techniques used in to measure autonomic response to music listening are mainly consolidated techniques for data analysis, such Principal Component Analysis, Factor Analysis, correlation. A novel approach consists of applying Granger causality to the lens model and dynamic emotional judgments.
The proposition of the Lens model is a good way to better understand the attribution of emotional characteristics to the music. With the new method of dynamic judgments (cf deliverable D1.1 and D2.2), we will use Granger causality method with two levels of prediction: the acoustic parameters and the elements presents in the musical structure (figure 9).
In figure 10, each X represents a set of links. For instance, X1 is decomposable into pitch or fundamental frequency or the note D. Similarly, X2 represents the rhythm, decomposable into quaver or quarter note and so on.

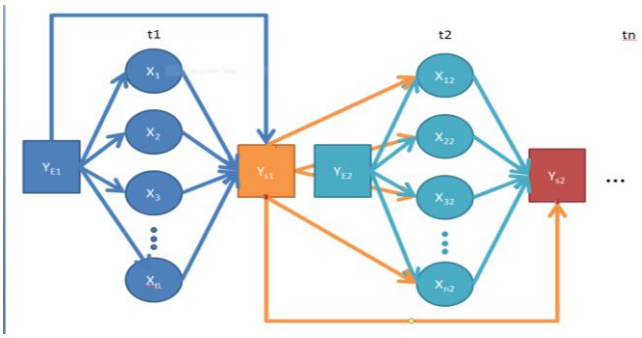


Figure 9: Towards a dynamic version of the Lens model applied to the attribution of emotional characteristics to the music.

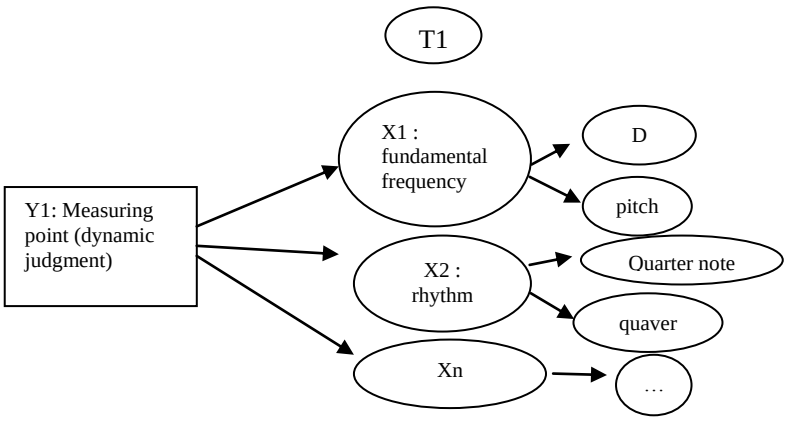


Figure 10: Example of analysis at Time 1.

The attribution of emotional characteristics to music may be the result of the interaction of these sets of elements (i.e. eighth notes with the note E may be the attribution of emotional characteristics of GEMS dimensions as joy or wonder).

We are also currently collaborating with Olivier Lartillot (developer of the MIR toolbox) and the Geneva University of Music to refine the selection of relevant elements and the tools to better understand the emotional response of the audience to the music.

4.2 Measuring audience temperature variations

In this section we describe a novel method to measure (inter-subject) correlated temperature variations within the audience. The goal of the method is two-fold: 1) to capture temperature changes on the faces (of an audience) that simultaneously occur in a given interval time; 2) to be robust to artifact due to movement. The following is a step-by-step description of the proposed method:

1. Randomly select a frame from the thermographic video and then manually define the Regions of Interests (ROIs), e.g., ellypses that circumscribe the faces (see Fig. 11)
2. For each pixel within a ROI and whose temperature is compatible with human physiology compute temperature changes (ΔT) between two consecutive frames. Only ΔTs that are compatible with human physiology are accepted (see Fig 12).
3. For each ROI, sum all the ΔTs. Compute the mean of each of those sums (mRΔT)
4. At each time step *t* compute:
    - the Integral of mRΔT within an observation window *W* (ImRΔT).. ImRΔT is a measure of the overall temperature change in a given interval time
    - the Pearson's correlation (i.e., mean of all the pairwise correlations) between RΔTs within the *W* observation window (CmRΔT). CmRΔT is a measure of intra-subjects synchronicity in a given interval time.
5. Compute the "Synchronous Temperature Change" (i.e., inter-subject correlated temperature variations):

$$STC = ImR\Delta T * CmR\Delta T$$

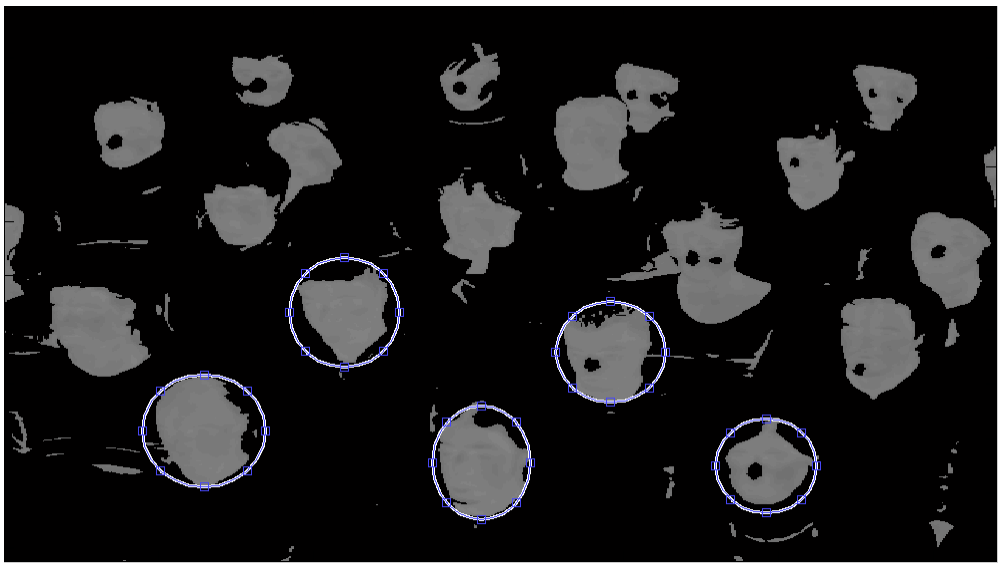


Figure 11 Step 1, manual identification of the Regions of Interest

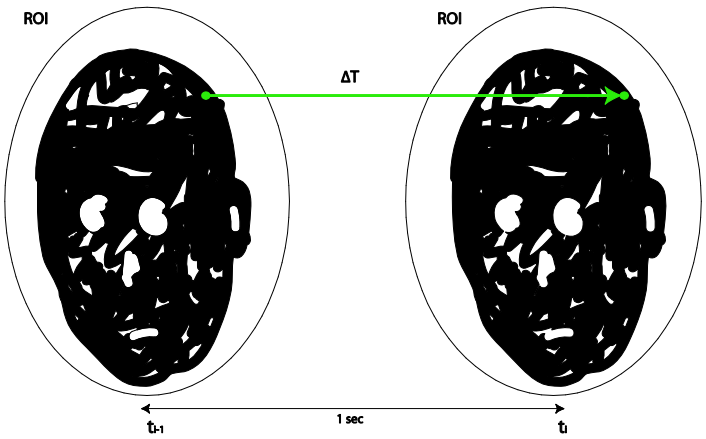


Figure 12 Step 2. Black pixels are pixels whose temperature and ΔT are compatible with human physiology.

Two other complementary approaches are under development to track temperature changes in a small audience (N=15 to 20).
Description of the processing steps: 1° Extraction of data from the thermocamera in MATLAB, 2° Manual registration of the thermal images (alignement and normalization of images across trials and participants, see Figure 13, 14, and 15), 3° Spatial downsampling, 4° Baseline correction based on pre-stimulus period, 5° Computation of the results based on the experimental design using anatomical ROIs or Independent Component Analysis (ICA), see Figure 16 and 17. Statistical analysis are based on permutations (N=200) across experimental conditions (shuffling observations and building up a distribution of differences obtained by hazard).

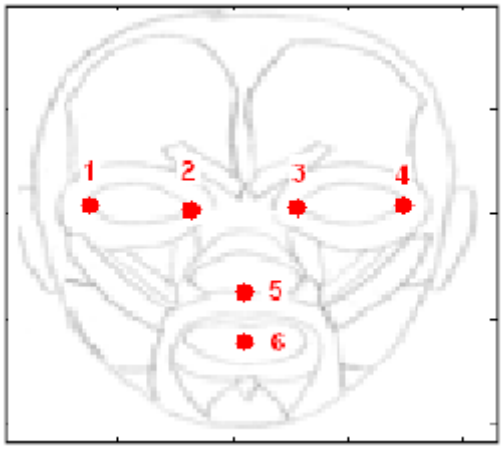


Figure 13: Landmarks on the face for normalization of face participants in the NeuroTherma toolbox (developed by S. Jarlier, Swiss Center for Affective Sciences).

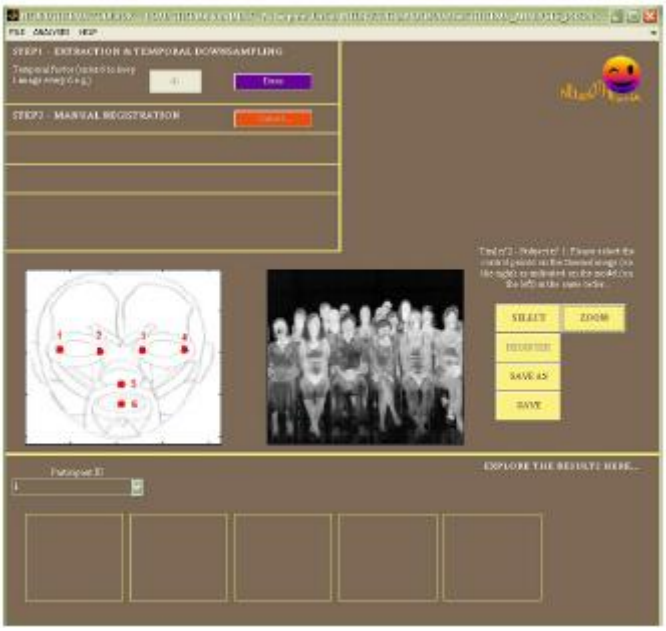


Figure 14: Matlab toolbox for thermographic analysis for small audience.

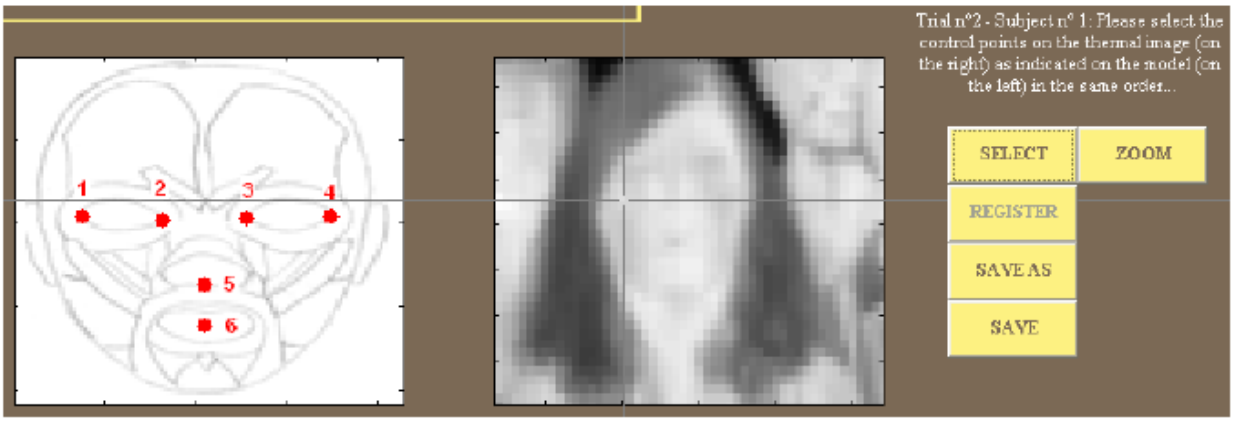


Figure 15: Landmarks manually defined for each face.

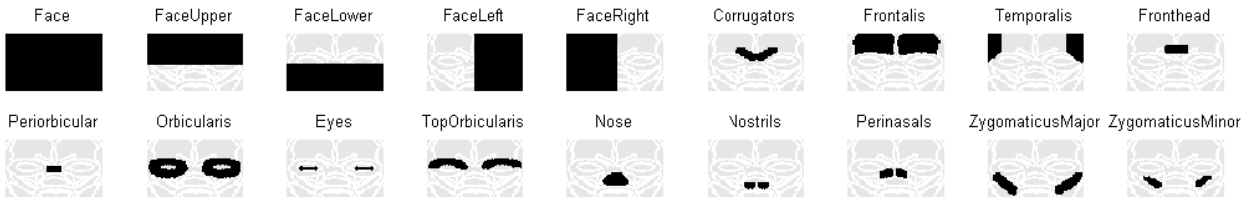


Figure 16: Definition of anatomical regions of Interest (ROI).

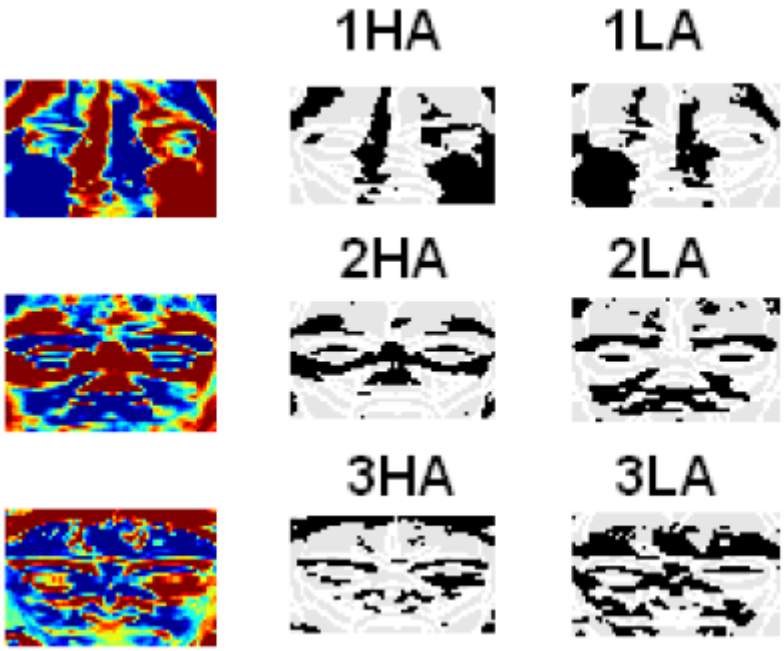


Figure 17: Examples of three components (HA=high amplitude for increase of temperature, and LH= low amplitude for decrease of temperature) based on Independent Components Analysis (ICA).

5. CONCLUSIONS

There is no clear line between techniques that are genuinely new and those that are adaptations of existing techniques to deal with data that has some unusual features. The focus in this deliverable has been on developments that are clearly in the first category. There is much more work in SIEMPRE that is innovative to some extent, but that we have treated for the purpose of this deliverable as belonging to the second.

In the long run, one of the outcomes of SIEMPRE will be a repository of techniques that are suited to processing the kind of data that we have studied. In that context, it will make sense to be inclusive, and to ensure that the community has access to as wide a range of analyses as possible. In this context of this deliverable, the aim has simply been to make clear that the repository will include some genuinely interesting innovations

6. PUBLICATIONS

D'Ausilio, A., Badino, L., li, Yi., Tokay, S., Craighero, L., Canto, R., Aloimonos, Y., Fadiga, L. Communication in orchestra playing as measured with Granger Causality. Intetain 2011 (Intelligent Technologies for Interactive Entertainment), Genova, 25-27 May. Also in: A. Camurri, C. Costa, and G. Volpe (Eds.): INTETAIN 2011, LNICST 78, pp. 273–275, 2012.

D'Ausilio A., Tokay S., Aloimonos Y., Li Y., Craighero L., Canto R., Badino L., Fadiga L. (2012) Coordinated action in orchestra. PlosONE. In Press.

D.Glowinski, A.Camurri, M.Mancini, R.Cowie (*in preparation*) "Playing alone or in ensemble: can movement regularity explain the difference?".

D. Glowinski, M. Mancini, N. Rukavishnikova, V. Khomenko, A. Camurri, (2011) Analysis of Dominance in Small Music Ensemble. Proceedings of the AFFINE satellite workshop of the ACM ICMI 2011 Conference, Alicante, Spain.

Jaimovich, J., Ortiz, M., 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 2012 Conference on New Interfaces for Musical Expression (NIME 2012), Ann Arbor, Michigan. New Interfaces for Musical Expression. Ann Arbor, Michigan, p. (In Press).

Jaimovich, J., Coghlan, N. & Knapp, R.B., 2012. Emotion in Motion: A Study of Music and Affective Response. In Proceedings of the 9th International Symposium on Computer Music Modeling and Retrieval (CMMR) Music and Emotions. Symposium on Computer Music Modeling and Retrieval. London, England, p. (In Press).

Marchini, M., Papiotis, P., Maestre, E. (2012) Timing synchronization in string quartet performance: a preliminary study. International Workshop on Computer Music Modeling and Retrieval (CMMR12), London, UK (2012)

Panagiotis Papiotis, Esteban Maestre, Marco Marchini, and Alfonso Peréz : Synchronization of intonation adjustments in violin duets: towards an objective evaluation of musical interaction. In Proceedings of the 14th International Conference on Digital Audio Effects (DAFx-11), Paris, France (2011)

Panagiotis Papiotis, M. Marchini, E. Maestre, and A. Perez. Measuring ensemble synchrony through violin performance parameters: a preliminary progress report. In Proceedings of the 4th International ICST Conference on Intelligent Technologies for Interactive Entertainment (INTETAIN 2011), Genoa, Italy (2011)

Panagiotis Papiotis, Marco Marchini, Alfonso Peréz, and Esteban Maestre. Measuring music ensemble synchrony in a string quartet through analysis of intonation adjustments. International Journal of Arts and Technology (IJART), submitted (2012)

Panagiotis Papiotis, Marco Marchini, and Esteban Maestre. Computational analysis of solo versus ensemble performance in string quartets: Dynamics and Intonation. In Proceedings of the 12th International Conference of Music Perception and Cognition (ICMPC12), Thessaloniki, Greece (2012).