ACII 2015 Tutorial  
September 21, 2015  
Xi’an (China)

A Research Platform for Synchronised Individual/Group Affective/Social Signal Recording and Analysis

M. Mancini, R. Niewiadomski, G. Volpe, A. Camurri  
Casa Paganini – InfoMus Research Centre  
DIBRIS, University of Genoa, Italy  
www.casapaganini.org
Summary

• The Casa Paganini–InfoMus Research Centre
• Conceptual Framework
• Automated analysis of multimodal features of non-verbal behaviour
  – Individual: expressive gesture, emotion
  – Group: synchronisation, entrainment, leadership
• The EyesWeb XMI Software Platform
  – Non-Verbal Social Signals Software Library
The monumental building of Santa Maria delle Grazie La Nuova
Research

• Cross-fertilization between research in science and technology and humanistic and artistic research.

• Art for ICT: Artistic and humanistic theories as source of inspiration for scientific-technological research.

• ICT for Art: Research results from science and technology as a source of inspiration for art languages and artistic projects.
Research

Real-time analysis of expressive gesture and non verbal social signals  *FP7 FET SIEMPRE*

Interactive sonification
Sensory substitution

*H2020 ICT DANCE*

(Socio-mobile) active music listening  *FP7 ICT SAME*
Research

**FP7 ICT ASC INCLUSION**
Therapy and rehabilitation: interactive serious games to support autistic children to learn to recognize and express emotions

Interactive software for music education (**FP7 ICT MIROR, H2020 ICT TELMI**) serious games, edutainment, active embodied experience of cultural audiovisual content

“Viaggiatori di Sguardo” permanent installation, Palazzo Ducale, Genoa
Interactive sonification to support chronic pain (with UCL, Nadia Berthouze)

Interactive systems for rehabilitation of children (with Gaslini Children Hospital) (Intetain 2015)

Rehabilitation exercises for Parkinson disease (ICT CARE HERE, EU CA CAPSIL)

Motion Composer (Wechsler et al)

Camurri et al 2003 “Application of multimedia techniques in the physical rehabilitation of Parkinson's patients”, Journal of Visualization and Computer Animation, 14(5)
## Research grounded on cross-fertilisation of ICT and art

<table>
<thead>
<tr>
<th>From artistic project…</th>
<th>…to S&amp;T research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Music Theatre Opera “Outis” Luciano Berio Teatro Alla Scala di Milano (1996)</td>
<td>Invisible interfaces for on-stage interaction and synchronisation with audio</td>
</tr>
<tr>
<td>Music Theatre Opera “Un Avatar del Diavolo”, Roberto Doati, La Biennale Venezia (2005)</td>
<td>Tangible acoustic interfaces: give the sense of touch to everyday objects</td>
</tr>
<tr>
<td>Museum “Enrico Caruso”, permanent interactive installation “Sala della Musica”, Firenze (2011-)</td>
<td>Non-verbal behaviour analysis for individual and group interaction with cultural heritage content</td>
</tr>
<tr>
<td>EU FET11 Closing Performance: TanGO Touching Music” (6 May 2011)</td>
<td>Performance built upon scientific results of the European ICT FET SIEMPRE Project.</td>
</tr>
<tr>
<td>Study of music joint performance: string quartets, orchestra sections with conductor, audience response</td>
<td>S&amp;T research in EU ICT FET SIEMPRE Project</td>
</tr>
</tbody>
</table>
Expressive Gesture

Example: Real-time measure and sonification of the space between the two dancers

Automated Analysis of Emotion from Expressive Gesture features

“Mappe per Affetti Erranti”, Festival della Scienza 2007

Each dancer embodies a human voice (bass, tenor, contralto, soprano); Each voice sings with the emotion expressed by the body gesture of the corresponding dancer. Example: Hesitant -> Whispering voice.

(1min ca.) dancers express different emotions: singing voices incoherent (2:30min ca.) dancers converge to Joy: all singing voices joyful and synchronized

• G.Volpe, A.Camurri (2011) “A system for embodied social active listening to sound and music content” ACM Journal on Computing and Cultural Heritage, 4(1)
Automated Analysis of Emotion from Expressive Gesture features

Luciano Berio
Music theatre opera “Cronaca del Luogo”
Salzburg Festival 1999
(video)

Expressive Gesture and Music

- Singing voice
  Original

A “microscope” on expressive gesture cues
(4 times longer)

(Rolf Inge Godoy, Oslo University)
Conceptual framework

Layered model for multimodal expressive gesture

- physical signals
- low-level features
- mid-level features; maps and shapes
- concepts, structures

Camurri et al 2005, IEEE Multimedia J
Conceptual framework

concepts, structures

mid-level features; maps and shapes

low-level features

physical signals

Real-time (ms)
Local

0,5 – 3s
Predictive models

Camurri et al 2005, IEEE Multimedia J
Conceptual framework

- physical signals
- low-level features
- mid-level features; maps and shapes
- concepts, structures

Example

Fluidity: smoothness of body joints + “wave-like” coordination

Smoothness one joint

Joints Positions, Velocities

Camurri et al 2005, IEEE Multimedia J
Conceptual framework

self

- concepts, structures
  - mid-level features; maps and shapes
    - low-level features
      - physical signals

other

- concepts, structures
  - mid-level features; maps and shapes
    - low-level features
      - physical signals

Synchronization (of Low Level and Expressive Features)
Research inspired by the arts and humanistic theories: Laban’s Effort Theory, Schaeffer’s Morphology, Gesture in Visual Arts

<table>
<thead>
<tr>
<th>Axes</th>
<th>Indulging Effort</th>
<th>Fighting Effort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space</td>
<td>Flexible</td>
<td>Direct</td>
</tr>
<tr>
<td>Time</td>
<td>Sustained</td>
<td>Quick</td>
</tr>
<tr>
<td>Weight</td>
<td>Light</td>
<td>Strong</td>
</tr>
<tr>
<td>Flow</td>
<td>Free</td>
<td>Bound</td>
</tr>
</tbody>
</table>

Example of Mid-Level Features: Laban Theory of Effort
EU-ICT Project MIROR

- Embodied and reflexive applications, to support children in exploring rhythm, melody, and harmony by means of their own body.

- Interaction paradigm: full-body mimicking a character

- Mapping of Laban’s movement qualities to elements of the musical structure.

EU-ICT Project ASC-INCLUSION

Serious games for teaching autistic children to recognize and express emotions by non-verbal full-body expressive gesture automated analysis of emotions

Example of High-Level Features: Emotions

S. Piana et al 2014 “Real time automated recognition of emotions from body gesture”, IDGEI 2014
Automated analysis of laughter from full-body movement

EU-ICT-FET Project ILHAIRE


Real-time multimodal analysis of non-verbal affective and social signals in ecological environments

**Music as ideal test-bed:**
- Non-verbal communication of emotion
- Involves social interaction between the musicians in an ensemble and with the audience
- Common, shared goal (no “cheating”)
- Can “speak” at the same time (good to study synchronization)

**Focus:**
- Focus on processes/dynamics rather than on states
- Predictive models for higher-level cues
Case studies

- Violin duo, classical western music, 2007-2008
- String quartet, classical western music, 2009-2013 (EU-FET Project SIEMPRE)
- Orchestra section, classical western music, 2010-2013 (EU-FET Project SIEMPRE)
- Duo, Hindustani music, 2014-2015
Violin duo

Dataset: Multimodal recordings performed during Premio Paganini 2006 International Violin Competition
Recordings and dataset

• Participants: four violin players - two pairs
• Material: a canon at unison from Musical Offering by J.S. Bach
• Conditions: player asked to act four emotions: Anger, Joy, Sadness, and Pleasure plus a deadpan condition, with and without visual feedback, repeated three times
• Recordings:
  – 2 b/w video-cameras: 720 x 576, 25 Hz
  – Height: 5-meters, taking the head of the performers
  – EyesWeb XMI application for synchronised recordings
Head motion tracking

Centre of Mass trajectory and speed extracted
Analysis

• 60 video recordings
• No time alignment of the two performances (canon): only the common part of the performance was taken into account

![Diagram showing time alignment]

• Each player modeled as a component of a complex system:
  – State vector: \((x, y, v_x, v_y)\) of head’s CoM
Phase synchronisation (PS)

- “Adjustment of rhythms of oscillating objects due to their weak interaction” [Pikovsky et al., 2001]
- More specifically: locking of phase/frequencies & unlocking of amplitude

- Idea: PS as low-level social signal toward an indirect measure of empathy
Recurrence and Recurrence Plots

- **Recurrence**: fundamental features of many dynamical systems
- **Recurrence Plots** [Eckmann, 1987]: time-time visualization of recurrences
- **Recurrence matrix**:

  \[ R_{ij} = \Theta \left( \varepsilon_i - \| \vec{x}_i - \vec{x}_j \| \right) \quad i, j = 1...N \]

  \[ R_{ij} = \begin{cases} 
  0 &: x_i \approx x_j \\
  1 &: x_i \approx x_j 
  \end{cases} \]
Recurrence and Recurrence Plots

Figure from Marwan et al., (2007). Recurrence plots for the analysis of complex systems. *Physics Reports*, 438, 237-329
Recurrence Quantification Analysis (RQA)

- **Small-scale patterns:**
  - single dots
  - vertical lines
  - diagonal lines

- **RQA:** quantification of small-scale patterns

  E.g., Recurrence Rate (RR):

  \[
  RR(\varepsilon) = \frac{1}{N^2} \sum_{i,j=0}^{N} R_{i,j}(\varepsilon)
  \]
Acoustic AND visual channels: 10 findings of PS
Acoustic channel only: 14 findings of PS

On average: $|CPRa| = 0.57$  $|CPRav| = 0.38$

No significant difference among emotions

BUT

Indication of a significant role of positive emotions.

String Quartet

“Self-Managed Team”: all musicians contribute equally to the task [Gilboa & Tal-Shmotkin, 2010, 2012]

(Picture: Quartetto di Cremona during an experiment at Casa Paganini-InfoMus)
Which multimodal cues explain the difference between playing alone or in ensemble?
Research questions

- Which multimodal cues explain who is the leader?
- Which multimodal cues explain the difference between a performance *engaging* an audience and a simply *correct* - but «cold» - performance?
The EU ICT FET Project
SIEMPRE

• When people perform a task as part of a joint action, their behaviour is not the same as it would be if they were alone: it is adapted to facilitate shared understanding (or sometimes to prevent it).

• Joint performance of music offers a test-bed for ecologically valid investigations of the way non-verbal behaviour facilitates joint action.

• SIEMPRE scenarios: music ensembles (quartets), orchestra section with conductor, audience.
Solo
Ensemble

Quartetto di Cremona
Solo or in Ensemble ? (1)
Solo or in Ensemble? (2)
Solo vs. Ensemble: measures

- Subjective centre of the quartet: the “ear”.
- Distance of heads from ear.
- Sample Entropy as measure of complexity.
Solo vs. Ensemble: results

- Significant effect of Solo vs. Ensemble condition.

\[ F_{1,135} = 119.984, \ p < .001 \]

Entrainment and leadership in an orchestra with conductor

Orchestra of Music Conservatory of Genoa, conductors Pietro Borgonovo and Sera Tokay (Italian Institute of Technology and Casa Paganini - InfoMus)
Entrainment

- **Affective entrainment**: formation of interpersonal bonds. Related to the pleasure in moving the body to music and being in time with others.

- **Temporal entrainment**: automatic movements that occurs when listening to musicians play (covert activation of motor areas of the brain), observed in complex rhythmic timing and exchange between partners or ensemble members in music or dance.

*(Phillips-Silver & Keller 2012)*
Soft Entrainment \((Yoshida 2002)\)

- Boerner’s model of orchestra conduction, focusing on entrainment within and between sections.
- Soft-entrainment: alternation of high and low entrainment \((Yoshida, 2002)\).
Orchestra: measures

- Recordings of Motion Capture and audio of two string sections and conductor. Dataset from (D’Ausilio et al., 2012).
- Focus on the z-component of bow trajectory of string sections (taken as predominant direction of bow movement).
Orchestra: some results

- Effectiveness of Leadership allowed ranking the performances of the two conductors in the dataset along ten different takes.
- The ranking confirms a previous study by (D’Ausilio et al PlosOne 2012) on the same dataset using another technique (Granger Causality).

Does synchrony follow inter-cultural shared patterns in music performances? Does the “soft entrainment” model apply also to non Western music culture?

Hindustani performances

In collaboration with M. Clayton (Durham University) and P. Keller (University of Western Sydney)
Hindustani performances

- **Objective**: investigate whether synchrony follows intercultural shared patterns
- **Dataset**: videos from different music performances from different cultures. Preliminary results from a considerable archive of digital audiovisual recordings of Hindustani performances, collected at Durham University
Hindustani performances

- **Performance**: analysis of a duo
  - Murad Ali (MA): highly-regarded sarangi (bowed lute)
  - Gurdain Rayatt (GR): tabla accompanist

- **Synchrony in Hindustani performances**:
  - A soloist will either return to a refrain just before beat one (sam) in the tala, and/or will conclude an improvisation at this point
  - In between these points it may be regulated by a soloist’s hand gestures (used to direct the drummer to adjust the tempo), or affirmed by gestures such as synchronised head nods on the sam (beat one)
System architecture (based on EyesWeb)
Extracted movement features

- **Quantity of motion:**
  - amount of detected movement (integration of pixel-wise differences between consecutive frames) [Mazzarino et al., 2003]

- **Y coordinate of head’s CoM:**
  - head’s barycentre coordinate on the vertical axis. It was extracted to analyse joint head movements (nods) that typically characterise the end of musical (tala) cycles

- **Head X displacement:**
  - the overall translation and rotation components of the head movements, computed from optical flow. Used as approximation of gaze direction
Preliminary results

• The final part of each cycle is the one displaying the highest synchronisation between the two musicians, confirming Yoshida’s findings on soft-entrainment.

• Future work: more sophisticated analysis of synchronisation of relevant events, exploiting Event Synchronisation (and our Multi-class and Multi-set extensions).

Technological outputs: EyesWeb

- Open software platform supporting fine-grain synchronised recordings of multimodal (audio, video, MoCap, biometric) data, performing pre-processing and analysis of multimodal signals in real-time
- Modular, flexible and adaptable
- Widely employed for developing real-time dance, music, and multimedia apps. Adopted by universities, industry, artists, cultural institutions
- Adopted in many EU projects
- Windows and Android (mobiles)
- Supports wide range of sensing and actuating devices
- Free download
Technological outputs
The SIEMPRE Platform for synchronised multimodal recordings

Multiple audio sources
Multiple video sources
Biometric sensors
Motion Capture
Low-cost sensors (e.g., Kinect2, Leap)
Sensors on Android mobiles
Our sensor based on Ethernet HD camera (50fps) + Asus Xtion (MotionComposer)
Technological outputs

Recordings can be previewed immediately, without any data processing or data conversion.
Technological outputs

Libraries for real-time analysis of body movement: motion features (e.g., kinematics, amount of movement, impulsivity, directness, fluidity, and so on) can be computed, stored on file, and viewed in real-time or off-line.
Technological outputs

EyesWeb Library for Real-Time Social Signals Processing: Synchronisation (Recurrence Quantification Analysis, Event Synchronisation), Leadership (e.g., chronemic leadership, analysis based on Graph Theory), …
Technological outputs: apps

Systems for: social active experience of music, interactive dance, experience of cultural heritage, rehabilitation, education
Understanding behaviour of groups of users: applications

- Active experience of music (and audiovisual content): users are enabled to interactively operate on content by creating, distributing, retrieving, modifying, and molding it in real-time.
Applications: iDJ

iDJ: A concept of application for embodied cooperation as a paradigm for formulating social queries.

Prototype presented at EU ICT 2013 Exhibition, Vilnius, Podium performance.

Applications: iDJ

- **Input:** full-body motion features
e.g., Motion Index, Contraction Index, Fluidity Index,…

- **Analysis of synchronisation**
i.e., whether the users are dancing in a tight-knit way

- **Analysis of dominance**
i.e., identification of the dominant user
Applications: Sync4All

Collaboration among users enables active listening to a music piece.

Each user rhythmically and freely moves her mobile phone trying to synchronise with the other users.

Synchronisation is measured as Phase Synchronisation of gestures (RQA, Cross-Recurrence Plots).

Applications: Sync4All
Main objective: to investigate techniques for sensory substitution in blind people, to enable perception of and participation in non-verbal, artistic whole-body experiences

Current work: sonification of motion features extracted from sample dance performances
DANCE – EU H2020 ICT Project

• “To see, close your eyes”

• Investigating how affective and relational qualities of body movement can be expressed, represented, and analyzed by the auditory channel.

• Objectives: to understand the meaning of “closing the eyes”,
  – the perception of expressiveness and entrainment in dance,
  – the participation to the emotion conveyed by a sequence of movements in space,
  – the understanding of the non-verbal language of bodies that communicate,
  – imagining and questioning concrete ways to listen to a choreography, feel a ballet.
DANCE – Dancing in the Dark

Collaboration with choreographers and dance groups
- To create archives of multimodal data on individual and group movement qualities
- To support and experiment project results in artistic productions
On-going projects: martial arts

- Karate: two «kata» performed by experts and students (all «black belt»)
- Hypothesis: measurable differences in intra-personal synchronisation of limbs can predict observers’ ratings of performance quality
- Event Synchronisation as measure of intra-personal synchronisation
Ongoing project on martial arts

- Karate: two «kata» performed by experts and students (all «black belt»)
- Hypothesis: measurable differences in intra-personal synchronization of body parts.
- Use Event Synchronization to measure intra-personal synchronization
Future research directions

• Aesthetic appreciation is one of the most intangible aspects of higher cognition
• Exploring the rules governing aesthetic experience has a great potential for future ICT:
• Measuring aesthetic experience in new applications: augmented reality, mobiles, embodied social media
More info and contacts

- Videos: www.youtube.com/InfoMusLab
- Papers: ftp.infomus.org/pub/Staff/AntonioCamurri
- www.casapagananini.org
- ricerca@unige.it

Research team
Antonio Camurri, Gualtiero Volpe, Corrado Canepa, Paolo Coletta, Nicola Ferrari, Simone Ghisio, Maurizio Mancini, Stefano Piana, Alberto Massari, Radoslaw Niewiadomski, Paolo Alborno, Ksenia Kolykhalova, Damiano Malafronte