Towards a Multimodal Repository of Expressive Movement Qualities in Dance

Stefano Piana, Paolo Coletta, Simone Ghisio, Radoslaw Niewiadomski, Maurizio Mancini,
Roberto Sagoleo, Gualtiero Volpe and Antonio Camurri
InfoMus Lab,
University of Genoa, DIBRIS
Viale Francesco Causa 13
Genoa, Italy
{paolo.coletta|maurizio.mancini|roberto.sagoleo|gualtiero.volpe|antonio.camurri}@unige.it,
stefano.piana@dist.unige.it, simoneghisio@gmail.com, radoslaw.niewiadomski@dibris.unige.it

ABSTRACT

In this paper, we present a new multimodal repository for the analysis of expressive movement qualities in dance. First, we discuss guidelines and methodology that we applied to create this repository. Next, the technical setup of recordings and the platform for capturing the synchronized audio-visual, physiological, and motion capture data are presented. The initial content of the repository consists of about 90 minutes of short dance performances movement sequences, and improvisations performed by four dancers, displaying three expressive qualities: Fluidity, Impulsivity, and Rigidity.

Author Keywords

Dance; Multimodal repository, Expressive movement; Automated analysis of movement qualities.

ACM Classification Keywords

H.1.2: Information Systems: User/Machine Systems: Human factors; H.5.m. Information Systems: User/Machine Systems: Miscellaneous.

INTRODUCTION

Expressive quality of movement is one of the most important aspects in dance performance. Recently, computational models and analysis techniques of movement qualities became a matter of interest for computer scientists [4, 19]. Several algorithms were developed to automatically detect and analyze different qualities of a dancer's movement (see e.g., [13] for a recent review). But research on movement qualities is often fragmentary and unsystematic. One of the

major difficulties is the lack of multimodal datasets that could be used by researchers for systematic analysis of a large spectrum of movement qualities. Different algorithms for single expressive qualities are developed and tested on datasets created *ad-hoc*, with the participation of different dancers and using different and non-standardized technical setups. This makes very difficult a comparative analysis of movements displaying different qualities. The existing repositories often do not focus on single movement qualities, but rather on more general aspects of dance activity, e.g., its cultural heritage or training. Consequently, they are not appropriate for studying single movement qualities. These repositories may also lack a precise definition of the recorded qualities. Finally, the multimodal nature of dance performance was often neglected in existing state of the art research.

The main goal of this work is the definition of the requirements and the creation of an initial version of a new multimodal repository for the analysis of movement qualities. Each fragment in the repository refers to a specific movement quality and consists of a series of synchronized multimodal data: motion capture data, audio, multiple videocameras, IMU sensors, and other biometric data. A main use of the repository is to study, develop, and compare software modules for the real-time analysis of expressive qualities. Our specific objective is to study expressive quality in dance performance in the framework of EU ICT H2020 DANCE Project.

In this paper, we present the initial recordings we carried out toward the creation of such a multimodal repository. We motivate the need for such a repository (Section 3) and we provide the guidelines for the recordings development process (Section 4). Then, we present the technical setup supporting such development process. Data synchronization of multimodal data is assured by the technological platform based on EyesWeb XMI (Section 5). For the initial recordings presented in the paper, professional dancers were invited and asked to perform short dance performances

focusing on specific movement qualities: for example, a number of sessions focused on Impulsivity and Fluidity.

STATE OF THE ART

Several multimodal datasets of the full body movements exist (e.g., [11, 14, 17], but only a few of them are explicitly dedicated to dance movement. Essid et al. [6] developed a freely available multimodal corpus of couples (student/teacher) dancing predefined salsa choreographies. The corpus consists of 1 hour and 40 minutes of data including synchronized audio rigs, multiple videos and RGB-D cameras, as well as IMU sensors.

Repositories of video-only dance movements exist. Alaoui et al. [1] recorded professional dancers performing four full-body movement types from Emio Greco's vocabulary: Breathing, Jumping, Expanding, Reducing. Then, they extracted a set of 6 movement qualities to drive a mass-spring system: Verticality, Extension, Leg opening, Shifting of weight, Periodicity, Quantity of Motion. By showing subjects the animation of the Mass-spring system they demonstrated that movement qualities only can successfully communicate the 4 target movement qualities.

Camurri at al. [4, 5] recorded professional dancers performing non-propositional movements with emotional content. They carried out subjective (i.e., human) and objective (i.e., automated) evaluation of the communicated emotional state and found out that subjects ratings can be explained by a small set of low-level features.

A few studies collected data from motion capture. Kitsikidis et al. [10] exploited multiple Kinect depth sensors to compute the body joints position of a dancer performing traditional Greek dances. They propose a fuzzy logic model to detect the quality of the dancer's performance, considering the evaluation of an expert of traditional dances as baseline. Truong et al. [18] use SVM algorithms for recognition of 3D gesture descriptors based on the Laban notation. For this purpose, they use publicly available Cambridge Gestural Performance Database [7] that contains Kinect data of several basic iconic and metaphoric gestures. Similar approach was taken also by Ran et al. [16], who built a dataset of 550 segments, and developed machine learning algorithms to detect Laban Movement Analysis (LMA) qualities.

MOTIVATION

Unfortunately, in the existing dance datasets described above, only few and sometimes unreliable sensors are available (e.g., video only, Kinect only, see for example [1]) while in other cases expressivity of movement is not addressed at all (e.g., [10]). Finally, dance sequences are usually not freely available, along with a detailed description of the recording setup, details on data streams, synchronization information, and so on. To the best of our knowledge, a large multimodal dataset with such characteristics, aiming at gathering a wide variety of expressive qualities in the dance performance, is still missing.

The aim of this work is to build a new repository being a collection of recordings of expressive dance movements by focusing on the following characteristics:

- the repository contains multimodal data of various dancers (with different dance backgrounds),
- the repository contains movements explicitly referring to and aimed at expressive qualities,
- the repository supports the development of scalable automated analysis software modules.

Multimodality

Although our research focuses on the expressive quality of full-body movement, other modalities include complementary information useful to detect expressive qualities. For example, respiration is of paramount importance to explain expressive features, and is strongly related to physical activity: previous works analyze respiration in walking and running rhythm [3, 8], rowing [2]. Respiration is also used to detect emotions [9, 12]. We hypothesize that different respiration patterns can be observed for movements with different expressive qualities. Our multimodal recordings include the respiration and vocal utterance by using a wireless microphone close to the nose.

Reuse

Existing datasets are often created *ad-hoc* (e.g., [1, 10]), to evaluate specific algorithms. Our repository is designed to study various expressive qualities from different perspectives and with different dancers. Dancers with different background participate in recordings.

Scalability

While using the MoCap system provides us the precise data on the movement, it is difficult to use such systems in ecological scenarios (e.g., during the live performance, improvisation with the public, or outdoor public spaces). In the perspective of a scalable platform, it might be useful to develop techniques to estimate the quality of movement from different modalities: when the MoCap data is not available (e.g., in a live dance performance), other modalities (e.g., RGB-D data or acceleration captured by wearable devices) may be exploited. For this reason, the dataset contains redundant data provided by devices less precise than MoCap but, at the same time, less cumbersome like, e.g., IMU (Inertial Measurement Unit) sensors and RGB-D cameras.

GUIDELINES AND DESIGN STRATEGIES

In the framework of the H2020 DANCE Project[1], we collect feedback from blind and non-blind persons, dancers and choreographers, and we perform feasibility studies and proof-of-concepts in scientific experiments and public events. The set of guidelines and strategies used to create the first version of the dance repository is described in the following sections.

Strategy 1: Expressivity as modulation of basic (carrier) movements

Very simple, basic movements, not evoking signs or language correlates, not stereotypical, can be used as carriers of different expressive qualities. These short, "neutral" movements can be *modulated* by different qualities (e.g., a simple "infinite" movement can be performed in a fluid, relaxed or in a jerky, rigid way).

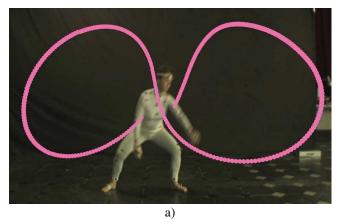Examples of simple movements used as "carriers" of movement qualities are the following:

- **Movement 1**: SineWave & StraightBack air drawing. Draw on air (with a finger, a hand or whole body, depending on the Space scale) one period of a horizontal sinusoidal movement followed by a straight line to return to the start position. It can be a single movement or repetitive.
- **Movement 2**: "Infinite shape air drawing", or "left-right loop". Draw on air the symbol of "infinite". Repetitive movement.
- **Movement 3**: Ellipsoid. Draw on air ellipsoid movements, with different Pathway, Direction, Inclination.

These three types of movements have the following characteristics:

- they are very simple and basic human movements, not requiring any particular skill,
- they can be performed rhythmically, periodically;
- they can be performed with a number of expressive qualities (sudden, fluid, rigid, contracted/expanded...)
- they can be scaled up and down in the space scale;
- they can be "building bricks" for a large set of more complex movements: e.g. the movements of an orchestra conductor, the ancillary movements accompanying speech;
- there is not a fixed starting position, each movement can be performed starting from different initial postures, initial states, originating variations of the movement in terms of difficulty, equilibrium/balance, and of the above list of variables.

Figure 1 presents an example of the execution of a simple movement, the "infinite" symbol, with different movement

---

qualities: Fluidity Vs. Rigidity (the movement trajectory is a projection of 3D MoCap data on 2D image).
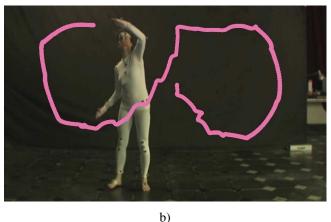

a)

b)

Figure 1. Execution of a simple movement, the "infinite" symbol, with different movement qualities: a) fluid; b) rigid.

Strategy 2: Expressivity evoked by textual images and metaphors

To induce specific movement qualities, the dancer is instructed with *textual images* and metaphors while doing the assigned basic movement. This approach is often exploited by choreographers to instruct dancers.

For each quality, multiple scenarios were proposed. For instance, in the case of the expressive quality called *Impulsivity*, dancers are asked:
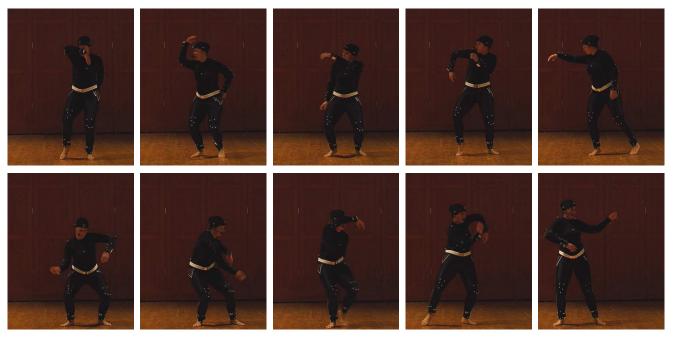
Figure 2. An improvised sequence of fluid movements.

"to move with a high fluidity and to imagine to be unexpectedly touched by a hot stick, or pushed by someone".

In the case of the expressive quality called *Fluidity* dancers are asked:

"to move as a fish in the water, or to reach the periphery of your personal space while imagining energy flowing along your body in a wave-like manner from the ground to the periphery; for example, imagine to throw a stone using your entire body by exploiting the ground as the origin of the energy flow and alternating your right and left arm".

Figure 3: An example of Strategy 3: while a blindfolded dancer is performing a fluid "infinite" movement (Movement 2), another person standing behind her touches her arm in unexpected moments with a soft stick.

For the expressive quality *Rigidity* dancers are asked:

"to imagine to have a pain in one part of the body (e.g., in a back)".

Additionally, in the case of certain bipolar qualities, e.g., *Rigidity*, the dancers are asked to repeat some of the above scenarios using both ends of the quality scale (e.g., high Rigidity Vs. low Rigidity).

The study of an expressive feature requires a sufficient number of different motion segments representing that feature in different ways, strategies and by different dancers. This means, in this case, the need to design more than one metaphorical description to evoke a given feature.

Strategy 3: Expressivity induced by external events
The dancer can be induced to express a feature by an external event. Let us consider the case of *Impulsivity*, i.e., a sudden movement with no preparation (no antagonist muscles prepare the movement, see [15]). To induce this feature, we adopt a strategy that implies unexpected physical contact:

1. the dancer is blind-folded;
2. the dancer is explained that, when she perceives some touch on her body, she has to imagine that it is a hot stick she has to avoid;
3. while the dancer is performing, for example, a fluid basic movement (e.g., the "infinite" movement described above), another person at a distance around the dancer touches her body at random times with a stick (see Figure 3).

Strategy 4: Improvisation

At different stages of the recording session (e.g., the beginning and the end), dancers are asked to perform an improvised choreography containing movements that, in the opinion of the dancer, better express the expressive quality under study. This is usually also part of a preparatory "warm-up" phase, as well as a "recovery" phase after the repetitions of the same basic movements on a specific quality performed following one of the strategies mentioned above.

By improvising a short dance performance, dancers are free to experiment with the expressive quality they have been asked to perform. Furthermore, the comparative analysis of improvised performances carried out at different stages of the recording session (e.g., at the beginning Vs at the end) can be used to highlight differences in the movement quality induced by the experience gathered by the dancer during the recording session. Figure 2 shows several frames of an improvised dance: the dancer aims to express Fluidity.

RECORDINGS

Data

Following the guidelines described in the previous section, we recorded a set of short performances of dancers asked to exhibit full body movements with a requested expressive quality.

At the beginning of each session, dancers were given definitions of the expressive quality by means of the *textual images* proposed in the previous section. For each expressive quality, following instructions were provided to the dancers:

1. to perform an improvised choreography containing movements that, in the opinion of the dancer, express convincingly the quality;
2. to perform several repetitions of predefined sequences movements by focusing on the expressive quality;
3. to repeat, at the end of session, the improvised choreography as in 1.

Four female dancers were recorded so far. The dancers have different dance backgrounds (classic dance, pop, contemporary dance), and different levels of professional experience.

In the first set of recording sessions, four dancers performed several repetitions of the following expressive features: *Fluidity*, *Rigidity*, and *Impulsivity*. Each trial was repeated at least two times and all trials, including the improvised choreographies, had a duration of 1.5 to 2 minutes. The total recordings duration is about 90 minutes. Table 1 presents the number of recorded trials for each quality. Trials are segmented into smaller chunks each corresponding to a uniform, coherent sequence of movements, e.g., each Impulsivity trial contains from 5 to 19 segments (mean = 9.8, std = 4.66) of the duration from 3 to 12 seconds.

The recordings were annotated by experts in dance and experts in expressive movement analysis. Apart of Fluidity, Impulsivity, and Rigidity, other qualities were observed in the recordings, including Symmetry, Lightness, Stillness, Slowness, Balance (see [19] for detailed definitions and overview of different movement qualities).

	Number of trials	Total duration
Fluidity	23	47 min 37 sec
Impulsivity	11	12 min 31 sec
Rigidity	18	29 min 29 sec
Total	52	89 min 37 sec

Table 1. List of recordings in the first version of the repository. Each trial has duration of 1.5 to 2 minutes.

Technical setup: the multimodal recording platform

The DANCE software platform, based on the EyesWeb XMI research platform, allows for synchronized recording, playback, and analysis of a multimodal stream of data.

The overall architecture for multimodal recordings is shown in Figure 4. The performer movements are captured by a Motion Capture system (MoCap recorder), in our case a Qualisys with 13 cameras. The performer wears a headset microphone, which is used to record breathing/respiration sound. She also wears five IMU units, two on the wrists, two on the ankles, and one on the coccyx. Finally, two broadcast quality video cameras, one from the front and one from the side, record HD video signals. All these data are synchronized. The synchronization mechanism works as follows. On the MoCap recorder computer, EyesWeb generates a reference clock used by all other recorders. The generated reference clock is sent to all the other connected devices in a format compatible with each specific device. As an example, the Qualisys Motion Capture system receives such clock encoded in an audio stream, in SMPTE format. Also, the two broadcast video-cameras and the Audio recorder use the same SMPTE clock encoded as an audio signal. The IMU recorder receives the reference clock via network, through the OSC protocol.

To guarantee synchronization the platform keeps track, for every recorded frame or sample, of the timestamp indicating the time at which the data sample was recorded. As a matter of facts, not all streams can be hardware-synchronized (e.g., with a genlock signal), thus, a software synchronization is performed by keeping track of these timestamps in separate files, and using such information when playing back the data. IMU sensors or Kinect are examples of devices, which are synchronized in this way.
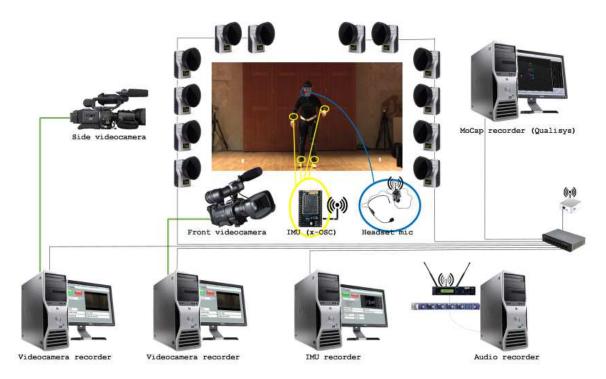
Figure 4. The DANCE multimodal recording platform. Several streams of data acquired by sensors on different modalities (video, motion capture, audio, respiration) are recorded in a synchronized way.

Data streams

We recorded data streams from the following devices:

- one Qualisys motion capture system, tracking a number of on-body markers (two configurations are available, one using rigid bodies and one with a higher, redundant number of markers), at 100 frames per second; resulting data consists of the 3D positions of the markers, as well as 3D rotations of rigid bodies if available;
- one wireless microphone (Mono, 48 kHz) placed close to the dancer's nose, recording the sound of breathing/respiration;
- 5 IMU sensors placed on the dancer's body limbs; the data is captured at 50 frames per second; each frame consists of 9 values: (x, y, z) of accelerometer, gyroscope, and magnetometer;
- 2 video cameras (1280x720, at 50fps);
- one Kinect2 RGB-D sensor, collecting video, depth map and user's silhouette (512x448, at 30fps).

Standards

Our system complies with a set of standards that makes the integration of existing (if any, see Section 2) or future multimodal recordings easy. The 3D motion capture data as well as the 3D IMU data is saved in the standard tab-separated values (tsv) format. Audio and video recordings use standard codec compressions. Data is synchronized using SMPTE timecodes - another standard that is widely used in multimedia content production.
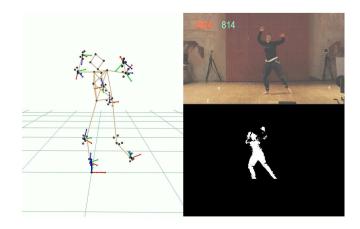

Figure 5. Synchronized multiple data streams captured by the DANCE multimodal recording platform: stick-figure visualization of 3D data, camera view, user's silhouette.

Examples

Figure 5 shows an example of the recorded synchronized data: 1) a stick-figure visualization of the all markers on the

left (including the markers of rigid bodies), 2) the corresponding video frame in the top-right corner, 3) the dancer's silhouette captured by the Kinect 2 sensor on the bottom-right corner.
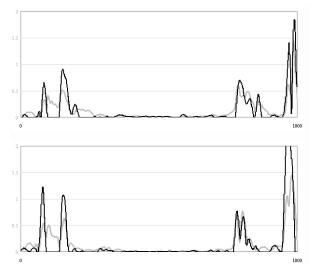

Figure 6. Synchronized data of left (above) and right (bottom) ankle from x-OSC accelerometer (black) and MoCap (grey) on 1000 consecutive frames (10 seconds).

Figures 6 - 8 show examples of the signal synchronization. In the Figure 6 one can see the acceleration computed from markers placed on the right and left ankle (grey color) and the acceleration absolute value of x-OSC IMU correspond sensors placed near to that marker (black color).
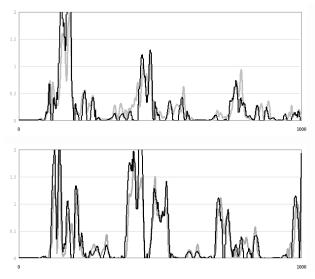

Figure 7. Synchronized data of right (above) and left (below) wrist from x-OSC accelerometer (black) and MoCap marker (grey) on 1000 consecutive frames (10 seconds).

Figure 7 shows IMU and x-OSC accelerations of the right and left wrist. One can see that corresponding signals (MoCap, and x-OSC) are synchronized.

Figure 8 shows the audio volume of respiration (black) and full body kinetic energy on computed 1000 consecutive frames (10 seconds) of MoCap data of an Impulsivity trial.
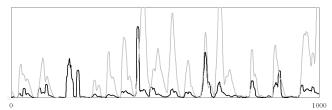

Figure 8. Synchronized audio volume of respiration (black) and full-body kinetic energy on 1000 consecutive frames (10 seconds).

CONCLUSION

In this paper, the initial version of the DANCE multimodal repository for analysis of expressive movement qualities in dance performances , a set of recording guidelines, and the technical setup and platform are presented. To create the repository we are using the multimodal platform for capturing synchronized audio-video streams, IMU sensors, and MoCap data of dance segments.

The main goal of the DANCE Project is to study the expressive quality of the movement from the perspective of the observer. The repository is currently under intensive growth, and dance students as well famous choreographers are currently collaborating to its development: this will be described in a paper in preparation. Further, we are working also on perceptual studies, to evaluate and validate whether observers (both expert and non-expert) perceive the expressive qualities as we planned during the recordings.

The data of our repository will be made publicly available on the DANCE Project website in the near future (see dance.dibris.unige.it for more details).

REFERENCES
1. Sarah Alaoui, Frederic Bevilacqua, and Christian Jacquemin. 2015. Interactive Visuals as Metaphors for

Dance Movement Qualities. *ACM Transactions on Interactive Intelligent Systems* Vol. 5, Issue 3.

2. Anthony H. Bateman, Alison H. McGregor, Anthony M.J. Bull, Peter M.M. Cashman, Robert C. Schroter. 2006. Assessment of the timing of respiration during rowing and its relationship to spinal kinematics. *Biol Sport*, 23: 353-365.

3. Paolo Bernasconi, Jana Kohl. 1993. Analysis of co-ordination between breathing and exercise rhythms in man. *J. Physiol* 471: 693–706.

4. Antonio Camurri, Barbara Mazzarino, Matteo Ricchetti, Renee. Timmers, Gualtiero Volpe. 2004. Multimodal analysis of expressive gesture in music and dance performances, in *Human-Computer Interaction*, Springer Verlag, v. LNAI 2915, 20-39.

5. Antonio Camurri, Ingrid Lagerlöf, Gualtiero Volpe. 2003. Recognizing emotion from dance movement: comparison of spectator recognition and automated techniques, *International Journal of Human-Computer Studies*, 59, 1–2: 213-225.

6. Slim Essid, Xinyu Lin, Marc Gowing, Georgios Kordelas, Anil Aksay, Philip Kelly, Thomas Fillon, Qianni Zhang, Alfred Dielmann, Vlado Kitanovski, Robin Tournemenne, Aymeric Masurelle, Ebroul Izquierdo, Noel E O'Connor, Petros Daras, Gaël Richard. 2013. A multi-modal dance corpus for research into interaction between humans in virtual environments. *Journal of Multimodal User Interfaces* 7:157–170.

7. Simon Fothergill, Helena Mentis, Pushmeet Kohli, Sebastian Nowozin. 2012. Instructing people for training gestural interactive systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '12). ACM, New York, NY, USA, 1737-1746. DOI=http://dx.doi.org/10.1145/2207676.2208303

8. Charles P. Hoffmann CP, Gerald Torregrosa G, Benoit Bardy. 2012. Sound Stabilizes Locomotor-Respiratory Coupling and Reduces Energy Cost. PLoS ONE 7(9): e45206

9. Kim Jonghwa; Elisabeth Andre. 2008. Emotion recognition based on physiological changes in music listening. In *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30, 12: 2067-2083.

10. Alexandros Kitsikidis, Kosmas Dimitropoulos, Erdal Yilmaz, Stella Douka, Nikos Grammalidis. 2014. Multi-sensor technology and fuzzy logic for dancer's motion analysis and performance evaluation within a 3D virtual environment. In *Universal Access in Human-Computer Interaction. Design and Development Methods for Universal Access,* 379-390. Springer International Publishing.

11. Ksenia Kolykhalova, Antonio Camurri, Gualtiero Volpe, Marcello Sanguineti, Enrico Puppo, Radoslaw Niewiadomski. 2015. A Multimodal Dataset for the Analysis of Movement Qualities in Karate Martial Art. In Proceedings *of INTETAIN 2015*, Turin, Italy.

12. Myagmarbayar Nergui, Lin Pai-hui, Gota Nagamatsu, Mikio Waki, Mihoko Otake. 2013. Automatic detection of laughter using respiratory sensor data with smile degree. In *Proceedings of the International Conference on Advances in Computing, Communications and Informatics*, 8 - 11.

13. Radoslaw Niewiadomski, Maurizio Mancini, Stefano Piana. 2013. Human and virtual agent expressive gesture quality analysis and synthesis, Rojc, M., Campbell, N. (eds.) Coverbal Synchrony in Human-Machine Interaction, CRC Press, Science Publishers (USA), 269-292.

14. Radoslaw Niewiadomski, Maurizio Mancini, Giovanna Varni, Gualiero Volpe, Antonio Camurri. 2016. Automated Laughter Detection From Full-Body Movements. In *Human-Machine Systems, IEEE Transactions on*, 46, 1: 113-123.

15. Radoslaw Niewiadomski, Maurizio Mancini, Gualtiero Volpe, and Antonio Camurri. 2015. Automated Detection of Impulsive Movements in HCI. In *Proceedings of the 11th Biannual Conference on Italian SIGCHI Chapter* (CHItaly 2015). ACM, New York, NY, USA, 166-169. DOI=http://dx.doi.org/10.1145/2808435.2808466

16. Bernstein Ran, Shafir Tal, Tsachor Rachelle, Studd Karen, Schuster Assaf. 2015. Multitask learning for Laban movement analysis. In *Proceedings of the 2nd International Workshop on Movement and Computing* (MOCO '15). ACM, New York, NY, USA, 37-44. DOI=http://dx.doi.org/10.1145/2790994.2791009.

17. Leonid Sigal, Alexandru O. Balan Michael J. Black. 2010. HumanEva: Synchronized Video and Motion Capture Dataset and Baseline Algorithm for Evaluation of Articulated Human Motion *International Journal of Computer Vision*, 87, 1-2: 4-27.

18. Arthur Truong, Hugo Boujut, Titus Zaharia.. 2016. Laban descriptors for gesture recognition and emotional analysis. *The Visual Computer*, 32, 1: 83-98.

19. Gualtiero Volpe, Antonio Camurri, A., Stefano Piana, Maurizio Mancini, Radoslaw Niewiadomski, Nicola Ferrari, Corrado Canepa, The Dancer in the Eye: Towards a Multi-Layered Computational Framework of Qualities in Movement, accepted to 3rd International Symposium on Movement and Computing, MOCO 2016.