Emilya: Emotional body expression in daily actions database

Nesrine Fourati, Catherine Pelachaud

Institut Mines-Telecom, CNRS LTCI
37/39 rue Dareau 75014 Paris
nesrine.fourati@telecom-paristech.fr, catherine.pelachaud@telecom-paristech.fr

Abstract

The studies of bodily expression of emotion have been so far mostly focused on body movement patterns associated with emotional expression. Recently, there is an increasing interest on the expression of emotion in daily actions, called also non-emblematic movements (such as walking or knocking at the door). Previous studies were based on database limited to a small range of movement tasks or emotional states. In this paper, we describe our new database of emotional body expression in daily actions, where 11 actors express 8 emotions in 7 actions. We use motion capture technology to record body movements, but we recorded as well synchronized audio-visual data to enlarge the use of the database for different research purposes. We investigate also the matching between the expressed emotions and the perceived ones through a perceptive study. The first results of this study are discussed in this paper.
Keywords: Movement task, Emotion, Motion capture

1. Introduction

Since the last two decades, the study of emotion expression in movement received a lot of interest through a wide range of approaches: social-psychological experimental studies (Dael et al., 2011; Meijer, 1989; Wallbott, 1998), automatic recognition of affects from body movements (Gong et al., 2010; Bernhardt and Robinson, 2007) or computational models of multimodal behavior for the design of affective embodied conversational agents (ECAs).

Darwin was one of the first researchers to define specific body movement pattern to each emotion. Lately, further studies (Dael et al., 2011; Tracy and Robins, 2008; Wallbott, 1998; Coulson, 2004) have been conducted to associate distinct patterns of movement and postural behaviors with some emotions using Body Action/Posture Units and/or communicative gestures (such as arms crossed in front of chest for Pride and self-manipulators for Shame (Wallbott, 1998)).

Bodily expression of emotions can also be signaled and described by the way a person is doing an action, usually defined as non-emblematic movement (Gross et al., 2010) or non-stylised body motion (Bernhardt and Robinson, 2007; Gong et al., 2010). In this paper, we interchangeably use the terms actions and movement tasks to refer to them. Previous approaches mainly focused on one single movement task such as walking (Montepare et al., 1987) or knocking at the door (Pollick et al., 2001; Gross et al., 2010).

The quantitative change of movement for the analysis and the synthesis of expressive and emotional body movements is as important as the qualitative description of body movement change (Gross et al., 2010). While the qualitative description of body movements refers to the description of (High-Level) emotion-related movement characteristics, the quantitative (Low-Level) description is required to build computational models for emotional and expressive movement behavior. Three-dimensional coordinate data provide an accurate description of body movement and can be readily provided by motion capture techniques. However there are still few databases that recorded bodily emotional behaviors using motion capture technology (Ma et al., 2006; Kleinsmith and Berthouze, 2007; Tilmanne and Dutoit, 2011). Furthermore, only few studies considered the validation of the emotional behaviors recorded in the database by studying the matching between the expressed (or the intented) emotion and the perceived one (Bänziger et al., 2006; Gross et al., 2010).

In this paper, we describe the collection of our new Emotional body expression in daily actions database (Emilya) database, a new repository of expressive and emotional body movements. The matching between the expressed and perceived emotion is also investigated through a perception study. The database can be used for the qualitative and quantitative characterization of emotion in different non-emblematic body movements.

2. Related work on expressive and emotional body movements corpora

The study of the expression of emotions and affects in body movement received recently an increasing amount of interest. Traditionally, the studies of bodily emotional behaviors have widely relied on database of emotion expression where actors portrayed some emotions.

Several scholars focus on the communication of emotion through body action, posture units and communicative gestures (Wallbott, 1998; Dael et al., 2011). In these works, the actor is usually asked to freely express an emotion without any constraints. In the work of Wallbott (Wallbott, 1998), 12 professional actors were asked to portray 14 emotions following a scenario-based induction approach. Similarly in GEMEP (Geneva Multimodal emotion portrayals) database (Bänziger et al., 2006), 10 professional actors portrayed 15 affective states under the direction of a professional stage director (Bänziger et al., 2006). Some researches highlight the importance of studying emotion expression during an interaction. IEMOCAP (interactive emotional dyadic motion capture) database is a recent corpus that was created for the analysis of speech and gestures in expressive human communication (Busso et al., 2008).

Other researches focused on the expression of emotion in daily actions. In non-emblematic body motion, the emotion is communicated implicitly while the subject is performing some movement tasks (e.g. walking). These studies investigate the effect of the emotion on movement qualities. Most existing databases do not encompass large variety of emotional states and movement tasks. So far, scholars tend to focus mainly on a subset of the six basic emotions (Ma et al., 2006) (Hicheur et al., 2013) (Kleinsmith and Berthouze, 2007). Similarly only a limited range of motions, often reduced to one movement task (e.g. walking), is studied. Several works were based on a database that contain emotional behaviors of a single movement task such as walking (Hicheur et al., 2013) (Roether et al., 2009) (Montepare et al., 1987) or knocking at the door (Gross et al., 2010) (Pollick et al., 2001). Furthermore, different body movement coding schema were used to describe the qualitative movement characteristics from one movement task to another.

Our aim is to elaborate a multi-level description of emotional body movements. We propose a new database of emotion body expression in daily actions where we consider a larger set of emotions and movement tasks.

One of the most agreed upon issues in the collection of emotional expression databases is the naturalness and spontaneity of emotional behavior. While naturalistic data involve many important information such as physiological arousal which gives rise to changes in muscle tone or reactivity, using acted data has also potential advantages like ensuring the structure of the data and the quality of recording (Cowie et al., 2011). Due to the difficulties related to the collection of spontaneous and natural emotional expressions, emotion induction methods are often used. Several approaches were proposed in the literature to induce affective states on demand (Cowie et al., 2011). Those approaches aim to achieve an acceptable compromise between data naturalness and the advantage of acting data. Scenario based approach is one of the most used approach for data induction (Dael et al., 2011; Wallbott, 1998; Montepare et al., 1987). In scenario based approaches, the actor is asked to read a written copy of a scenario and he is told to imagine how he would feel in the corresponding situation.

We can distinguish mainly two types of recording equipment; audio-visual recording (Bänziger et al., 2006) and 3D motion capture recording (Ma et al., 2006). Audio-visual recording provides 2D digital video that contains the visual content of body movement and/or audio information. Although the analysis of digital video can be used for producing computational models of multimodal behavior (Camurri et al., 2004), 3D motion capture of body movement provides more accurate information about the 3D posture and movement of some specific body joints of the subject, allowing one to produce more accurate computational models (Roether et al., 2009). In the collection of our database, we record both digital videos and 3D motion capture data providing a rich multimodal dataset of emotional and expressive body movement.

3. Emotional expression recording

In this section, we describe our database.

Actors: The actors were eleven (6 females and 5 males) graduate students. The mean age was 26 ranging from 23 to 28. They were motivated to participate to the construction of our database and they gave informed consent that their motion capture data as well as their video could be used and published for research purpose. Although most of them have received theater courses since a long time, a professional acting director was hired to give them 7 training sessions regarding the use of body movements to express emotional states. The acting director was also invited for one recording session (using only one scenario to elicit each emotion) and he was paid for his services. Each training session lasted three hours in which the acting director tried to keep the actors at ease as much as possible. A principal purpose of those sessions was to make aware to the actors how to use their body to express affects through actions.

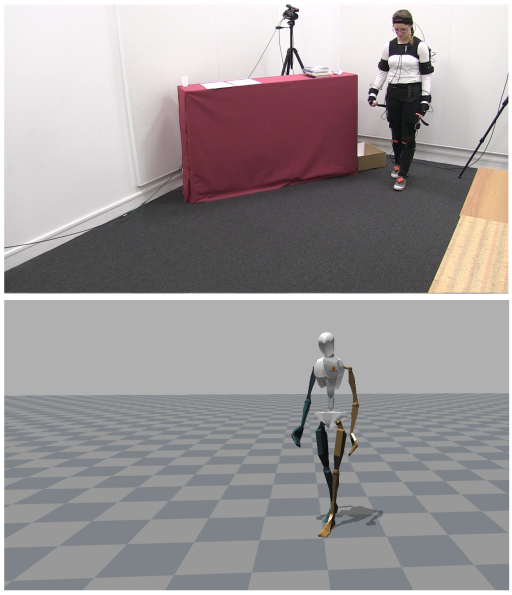

Figure 1: The extraction of the same posture from one video sequences (related to the second camera) and 3D motion capture data

Emotions: The emotions considered in our study are Joy, Anger, Panic Fear, Anxiety, Sadness, Shame, Pride and Neutral. Those emotions were selected to cover the arousal and valence dimensions. It has been shown in previous works that the expression of those emotional states can last a period of time (Dael et al., 2011), which makes their expression through body movement more or less extensible unlike reactive emotions such as surprise.

A scenario-based approach was adopted for data induction as it provides a good compromise between its reliability to induce affects on demand and its simplicity from the perspective of the actor. Each scenario includes the description of a situation, which is assumed to elicit a given emotional state (Bänziger et al., 2006). For each motion sequence recording, the actor was asked to read the scenario which is written on a paper, to imagine that he/she is living this situation and to perform the proposed movement tasks.

Four different scenarios were proposed for each emotional state but only three of them were used. The fourth one was used to replace a scenario when the actor felt unable to imagine the specific situation. Most of the scenarios that we used were adapted from those used in the study of Dael et al. (Dael et al., 2011) and Scherer et al. (Scherer et al., 1991). The other scenarios were validated through an informal perceptive study where ten participants had to attribute an emotion chosen from a closed list to each scenario. Two scenarios used respectively for Anger and Sadness emotions are provided here as examples:

”I hoped to sleep late on Sunday morning, but my neighbor started very noisy work in his house at 7am. I felt so angry that I decided to go and scold him.”

”I got a call to tell me that my favorite aunt suddenly died.”

During the recording sessions, the order of emotions, scenarios as well as actions was randomized from one actor to another.

Movement tasks: We consider a wide range of daily actions that involve the whole body as well as upper body and arms in particular. The movement tasks are walking, sitting down, knocking at the door, lifting and throwing objects (a piece of paper) with one hand, and moving objects (books) on a table with two hands. For the walking actions, we asked the actors to walk back and forth along the long side of the room. Walking was divided into a simple walk and a walk with an object in hand to capture two types of arms behavior during walking action. Some of those actions were already used in past studies and considered as relevant to discriminate between different styles of the same movement type (Tilmanne and Dutoit, 2011; Pollick et al., 2001; Gross et al., 2010). We asked the actors to perform each action four times in a row to capture a large set of data. A continuous sequence consisting of the series of all the actions with just one trial per action was also recorded. Continuous sequences were recorded only for two scenarios for each emotion.

Figure 2: Still frames depicting three viewpoints (corresponding to three cameras) relevant for the visualization of the Moving books action

4. Body movement recording

In our database, we record both audio-visual and motion capture data. We use specific hardware to perform the synchronization between those recording types. The motion capture took place in a professional audio-visual recording studio in our institute, Telecom ParisTech in France. For each sequence of body movements, we recorded 3D motion capture file as well as video file.

3D motion capture: We used the inertial motion capture system Xsens (Xsens, 2005 2014) to record the 3D motion data of the whole body. The MVN Mounting straps were used. 17 sensors were used to capture the movements of 23 body segments. The orientation and the position information are obtained for each body joint.

Video acquisition: Besides the 3D motion capture data obtained for each sequence of movement, we recorded four MXF (Material eXchange Format) video files of full HD resolution (1280*720) from four cameras placed in the four corners of the studio. Two cameras were dedicated to capture a general view of the room while the other two cameras were placed carefully to capture the face and the upper body when the actor performed actions involving mainly upper body. Canon XF105 cameras were used for this purpose. We recorded audio even though we did not explicitly ask actors to express emotions through the voice.

The synchronization of video and 3D motion capture data: The video files were synchronized with the motion capture files through some specific hardware. The Rosendahl nanosyncs HD, a professional video and audio sync reference generator, was used to generate a common Time Code (TC). The generated TC is read from the cameras through the outlet Genlock/TC. The Alpermann card PLC PCIe is used to read the TC in the computer. Using the MVN time code plug-in of the MVN Studio software (Xsens software), we were able to read the TC from the Alpermann card. Thus, for each sequence of movement, the corresponding motion capture file contains the same TC as the video files recorded from the cameras. The TC integrated in motion capture and video files is useful to extract the same sequence of movement in both audio-visual and motion capture data (See Figure 1).

5. Data post-processing and quantification

As explained in section 3., the actors were asked to express emotions in different movement tasks, where each movement task is repeated four times. While the segmentation of the motion sequence into different movement tasks was performed manually, the segmentation of the same movement task into single trials was performed automatically for most of the actions.

Data post-processing: Three types of segmentation process were used in order to separate multiple repetitions of the same action; manual (for Lift, Throw and Sit down actions), semi-automatic (for Knock at the door, Move books actions) and automatic (for Walk action). A detailed description of the beginning and the end of each movement task was provided as a reference for manual segmentation. Although we always give the instructions to move to a neutral pose (called N−Pose) between each successive action repetitions, actors sometimes forget to perform this step. Therefore, a manual segmentation was performed for the actions of Knocking at the door and Moving books when the actor did not return to N−Pose between successive trials. The automatic segmentation of these actions (in which the actor returned to N−Pose) was performed automatically based on kinematic measures related to the position of the hands along the vertical axis. The automatic segmentation of the walking action was based on the work conducted on the analysis of walking and turning task (Fourati and Pelachaud, 2013). Therefore, the separation between walking and turning tasks was based on the detection of the Turn Interval Time in which the relationship between hips and shoulders is the most linear.
Since video and motion capture data were synchronized, we can find the start and the end of the video sequences that correspond to motion capture data using the TC information. Consequently, the process of video files segmentation (MXF files) was based on the results of motion capture files segmentation using MediaInfo JAVA library, VirualDub tool and the start and end TC of the segmented motion capture files. The videos resulted from the video segmentation process were compressed and converted to AVI format (Audio Video Interleave).

Data quantification: Each actor was asked to express 8 emotional states, each described through 3 scenarios (except for Neutral which was described through 2 scenarios only). Thus, at the end of each recording session, we obtained several files each depicts the expression of one emotion successively in 7 movement tasks, each repeated four times. As a result, we obtained 23 motion captures files (7 emotions * 3 scenarios + Neutral emotion * 2 scenarios) and 92 (24 motion sequence * 4 Cameras) video files (corresponding to the four cameras).
We recorded also the expression of emotion successively in 7 actions each one performed once. We call them continuous sequences. The segmentation process led to the duplication of the motion sequence according to the different actions and the individual repetition of each action.
So in total, we obtained around:

176 continuous motion capture sequences = 11 actors * 8 emotions * 2 scenarios,

1771 motion capture sequences of one movement task repeated 4 times = 11 actors * 23 motion sequences * 7 actions,

7084 motion capture sequences of one movement task trial = 11 actors * 23 motion sequences * 7 actions * 4 repetitions,

The number of video files is equal to the number of motion captures files multiplied by the number of cameras. For example, only two viewpoints (corresponding to two cameras) allow visualizing the action Sitting Down as the chair is placed in a particular corner, while three viewpoints (corresponding to three cameras) recorded Lifting, Throwing and Moving objects actions (See Figure 2). For instance, we obtained 704 videos of continuous motion sequences where each set of four videos corresponds to the same motion showed from four different viewpoints.
The duration of motion sequences depends on the movement task and the expressed emotion. Motion sequences of one movement task repetitions last around 30 $\pm$ 20 seconds while motion sequences of one movement task trial last around 5.5 $\pm$ 3 seconds.

6. Expressed emotions evaluation

A perception study was conducted to evaluate the emotional body expression acted out by the actors. The content of stimuli used to judge the emotional content can be based on real or on virtual actors (where body movements are reproduced via a software agent). In fact, some features such as facial expression, fingers motion and respiration are included in the visual content of emotion expression by real actors but not visualized with virtual actors. Such features can influence the perception of emotions expressed by real actors. Since the main goal of our study is to characterize emotional body movement, we conducted this study using movies where body movements are reproduced on a virtual character instead of the video of the real actors.

Stimuli creation: Computer avatars have been used in various body movement perception studies (Hicheur et al., 2013; Roether et al., 2009). Different body models were proposed: Point-Light display (Pollick et al., 2001), skeleton-based body model (Griffin et al., 2013; Kleinsmith et al., 2011), and virtual animated character involving gender feature (Hicheur et al., 2013). A skeleton-based body model is the most appropriate body model in our perception study: First because we do not aim to include extra features (such as culture and gender) in emotions perception study; second because it has been shown that the perception accuracy of emotions was reliably greater for Full-Light display (e.g. 3D skeleton-based model) than Point-Light display of body movement (Atkinson et al., 2004). The default 3D Studio MAX biped model proposed in the character studio feature of 3D Studio MAX software was used in our perception study.
Among the emotional behaviors that we aim to evaluate in our study, some body behaviors involve movements that imply the whole body displacement in the space. We

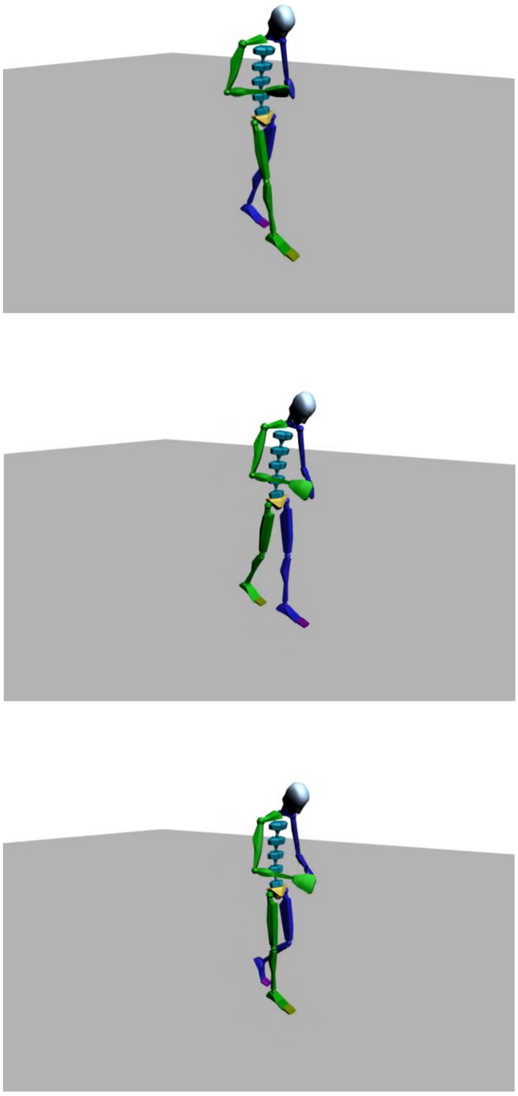

Figure 3: Snapshots of one created stimuli (Sad walking) using the 3D Studio MAX biped model; the virtual camera motion is synchronized with the avatar displacement in the space while keeping a stable distance between the avatar and the camera along the animation.

parametrize the motion of the virtual camera to follow the avatar motion in stimuli creation (Fourati et al., 2014). The virtual camera motion follows a non uniform linear style and it is synchronized with the motion of the avatar (Fourati et al., 2014). The automatic creation of stimuli was implemented using MAXScript scripting language of Autodesk 3DS Max software (See Fig. 3).

Database selection and distribution: Our database consists in more than 7000 motion capture files depicting emotional behaviors sequences. The evaluation of the expressed emotions in all the motion sequences is a cumbersome task. However, we can assume that choosing randomly a single trial from the four repetitions for each action and a single scenario from the three scenarios proposed to elicit each emotional state will be enough to cover more or less the content of the database. This selection process lead to 664 motion sequences.

Since the total number of videos under the database selection (664 movies) is a huge number to assign to each observer, we define a set of movies for each observer. This set of movies has to represent a significant difference for each feature that may affect the content of the emotional behavior. We distinct three main features: actor, movement task and emotion. We define one set of movies as 16 videos containing a variety of 2 actors, 4 emotional states and 2 movement tasks. 42 set of movies, each composed of 16 videos, were used to cover the database selection.

6.1. Protocol

In our evaluation study, we asked participants to rate both the emotion perception and some body expressive features. In this paper report results from the emotion perception task. Participants were asked to visualize a video where the emotional body expression is reproduced through a virtual agent, and to judge the perceived emotion. We use the Likert-type scale to evaluate emotion perception. The participants had to rate the perception of each emotion (8 emotions) on a five-level Likert item, where the levels are ”Strongly disagree”, ”Disagree”, ”Undecided”, ”Agree” and ”Strongly Agree”. The statement used for all the Likert items starts with ”The actors expresses ”. The movement task performed by the virtual actor was indicated for each video.

In order to evaluate the perception of emotions, we addressed the following hypothesis: The perception of an expressed emotion is significantly different and higher for videos expressing this emotion than from the videos expressing other emotions.

6.2. Amazon Mechanical Turk and demographics of participants

The popular crowd-sourcing website Amazon Mechanical Turk (AMT) was used to collect the results of emotion perception as it provides an easy access to a large, stable and diverse subject pool. TurkGate tools (Grouping and Access Tools for External surveys) (Darlow and Goldin, 2013) were used to better control the use of Mechanical Turk with external HITS.

1008 participant took part in our study (56.01% of females and 43.98% of males). Each video was evaluated 24 times (42 groups of videos * 24 annotation per video = 1008 participants). Participants were 36.01 years old on average with 15 and 76 as respectively the minimum and maximum age, and 33 as the median age. The percentage of participants who spent the majority of their life in U.S was 98.21%.

6.3. Analysis and Results

One-way Anova was conducted to evaluate the difference of rating the perception of each emotion between the different expressed emotions. We performed the multiple comparison Tukey-Kramer test to conclude which pairs of means are significantly different. Since we aim to study whether the perception of an expressed emotion is significantly different for videos expressing this emotion from the

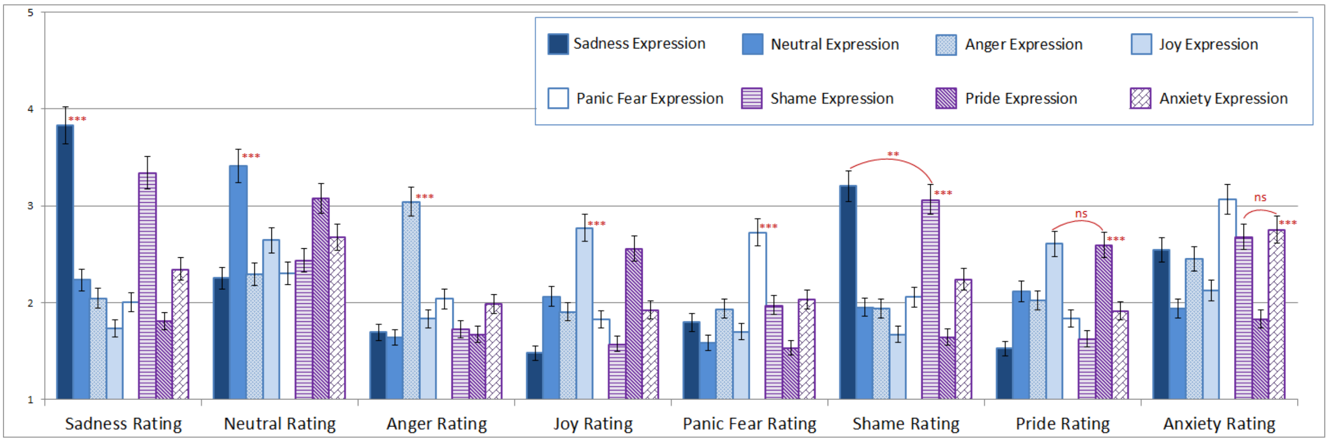

Figure 4: The mean rating of each emotion perception from all the videos of the virtual character. '***', '**' and 'ns' stands respectively for significant difference with p<0.001, a significant difference with p<0.01, non-significant difference. The error bars indicate 95% confidence interval. The mean rating is graduated from 1 to 5 which stands respectively for the following agreement levels: "Strongly disagree", "Disagree", "Undecided", "Agree", "Strongly Agree".

videos expressing other emotions, the Tukey-Kramer test was used for this purpose.

Statistical results: The one-way Anova test has shown that the emotion expressed by actor has a main effect on the rating of each emotion (See Table 1).
The mean rating of each emotion perception in all the videos are shown in Figure 4. The Tukey-Kramer test has shown that the perception of Sadness was significantly different and higher for video expressing Sadness from the video expressing other emotions (p<0.001). The same result was found for the perception of Neutral, Anger, Joy and Panic Fear. A significant difference was found between the perception of Shame in videos expressing Shame and videos expressing other emotions (p<0.01), however the perception of Shame in videos expressing Sadness received the highest mean rating (mean=3.20) (See Figure 4). No significant difference was found between the perception of Pride in videos expressing Joy and videos expressing Pride (p<0.05). No significant difference was found also between the perception of Anxiety in videos expressing Anxiety and videos expressing Shame (p<0.05).

Rated Emotion	Anova results
Sadness	(F=237.77, p<0.001)
Anger	(F=394.57, p<0.001)
Panic Fear	(F=263.77, p<0.001)
Joy	(F=345.30, p<0.001)
Neutral	(F=206.004, p<0.001)
Shame	(F=566.83, p<0.001)
Anxiety	(F= 237.77, p<0.001)

Table 1: The results of One-Way Anova showing the significant difference between the perception of each emotion in all the videos

Discussion: As shown in Figure 4, the mean ratings of emotion perception were above the "Undecided" agreement level for the correct perception of Sadness, Anger and Neural. However, the mean ratings were below the "Undecided" agreement level for Joy and Panic Fear perception. That means that participants mostly disagree or they are undecided about the perception of Joy and Panic Fear in videos showing the expression of the same emotion(respectively Joy and Panic Fear expressions), but they disagree much more with the perception of Joy and of Panic Fear in the other conditions. This result suggests that the perception of Joy and of Panic Fear was more difficult than the perception of Sadness, Anger and Neutral across all the actors and movement tasks. This discrepancy in the results could come either from the actor or from the movement task. Further analyses measuring the effect of the actor and of the action have to be performed.
Despite Shame perception in videos depicting Shame expression received significant result (p<0.01), participants showed high level of confusion with Sadness as they perceived Shame in videos expressing Sadness with high mean rating. Participants perceived Anxiety similarly in videos expressing Anxiety and videos expressing Shame, but above all they perceived Anxiety in videos expressing Panic Fear. Such a confusion could be explained as the emotions of Anxiety and Panic Fear belong to the same family of emotions. As the arousal level of Anxiety is lower than the arousal level of Panic Fear, the perception of Anxiety in videos expressing Panic Fear could be easier than the perception of Anxiety in videos expressing Anxiety.
The perception of Pride was similar (no significant difference) in videos expression Joy and in videos expression Pride. The mean rating of Pride perception in videos expressing Pride and videos expressing Joy were lower than the "Undecided" agreement level. Overall, this result suggests that participants mostly disagree or they are undecided about the perception of Pride in videos expressing positive emotions, but they disagree much more with the perception of Pride in videos expressing negative emotion.
Overall, the perception of the emotions considered as "Basic Emotions" in videos depicting the expression of the same emotions received the highest and the most signifi-

cant rating while Shame, Anxiety and Pride were confused at some level with similar emotions; Shame with Sadness, Anxiety with Panic Fear and Shame and finally Pride with Joy.

The perception of negative emotions, in particular Anger and Sadness, and neutral emotion received higher rating compared to the perception of positive emotions (Joy and Pride). Overall we obtained significant results for the perception of most of the emotions. The low mean rating of the perception of some emotion can be due to the effect of movement task or actor expression on the recognition of emotion. A deeper analysis to study the effect of movement task and actor on the emotion perception would be helpful to better understand the results. Providing the results of emotion perception for each actor and movement task would allow for a better description of the database.

7. Conclusion

In this paper, we introduce a new database of emotional body expression in daily actions. This database constitutes a rich repository of emotional expression in body movements. Eleven actors expressed 8 emotions while performing 7 movement tasks. Each movement task was repeated four times to capture a wide range of data. Our database includes synchronized audio-visual and motion capture recording. While the video files include facial and bodily emotional expression visualized from different camera viewpoints, motion capture files include three-dimensional data of the whole body movement.

We investigate the matching between the emotions that we asked the actors to express (through a scenario-based approach) and the emotion that can be perceived by third-party observers through a perception study. This study was done on Mechanical Turk crowd-sourcing website. MAXScript scripting language of Autodesk 3DS Max software was used to automatically create the stimuli of the perception study. The statistical analysis has shown that the emotions Anger, Sadness, Neutral, Panic Fear and Joy were significantly better perceived in the videos showing the expression of the same emotions than from the videos showing the expression of other emotions. Shame perception received also significant mean rating, but the expression of Shame was confused with the expression of Sadness. Confusion was also found among the expression of Pride and Joy as well as the expression of Anxiety and Shame.

In our future work, we aim to study of the effect of actors and movement tasks on the perception of emotion and the evaluation of emotion perception for each actor and each movement task. After having conducted these perception analyses, we will offer this database to the research community.

8. Acknowledgements

A special thanks to all the actors, the acting director Mr Robert Bennett, Mr Giovanni Nolli from Xsens support, Mr Gilbert Pappalia for his help to resolve several technical issues, Mrs Janina Breidaka who helped us to organize and segment a large part of data, and finally Sabrina Campano for her advises. This research is partially supported by the National project Anipev1 (FUI Acteur Virtuel) and the FUI Play-Serious project.

9. References

Anthony P Atkinson, Winand H Dittrich, Andrew J Gemmell, and Andrew W Young. 2004. Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception*, 33(6):717–746.

Tanja Bänziger, Hannes Pirker, and Klaus Scherer. 2006. GEMEP-GEneva Multimodal Emotion Portrayals: A corpus for the study of multimodal emotional expressions. *Proceedings of LREC*.

Daniel Bernhardt and Peter Robinson. 2007. Detecting affect from non-stylised body motions. *Affective Computing and Intelligent Interaction*, 4738:59–70.

Carlos Busso, Murtaza Bulut, Chi-Chun Lee, Abe Kazemzadeh, Emily Mower, Samuel Kim, Jeannette N. Chang, Sungbok Lee, and Shrikanth S. Narayanan. 2008. IEMOCAP: interactive emotional dyadic motion capture database. *Language Resources and Evaluation*, 42(4):335–359, November.

Antonio Camurri, Barbara Mazzarino, and Gualtiero Volpe. 2004. Analysis of Expressive Gesture: The EyesWeb Expressive Gesture Processing Library. *In Gesture-based Communication in Human-Computer Interaction, LNAI 2915 (Springer Verlag)*, pages 460–467.

Mark Coulson. 2004. Attributing emotion to static body postures: recognition accuracy, confusions, and viewpoint dependence. *Journal of Nonverbal Behavior*, 28(2):117–139.

Roddy Cowie, Ellen Douglas-Cowie, Margaret McRorie, Ian Sneddon, Laurence Devillers, and Amir Noam. 2011. Issues in Data Collection. In Paolo Petta, Catherine Pelachaud, and Roddy Cowie, editors, *Emotion-Oriented Systems: The HUMAINE Handbook*, pages 197–212.

Nele Dael, Marcello Mortillaro, and Klaus R Scherer. 2011. Emotion expression in body action and posture. *Emotion*, 12(5):1085–1101.

Adam Darlow and Gideon Goldin. 2013. Turk-Gate (Version 0.4.0) [Software]. Available from http://gideongoldin.github.com/TurkGate/.

Nesrine Fourati and Catherine Pelachaud. 2013. Head, Shoulders and Hips Behaviors during Turning. *Worshop on Human Behavior Understanding, HBU 2013, in conjunction with ACM Multimedia 2013*, 8212:223–234.

Nesrine Fourati, Jing Huang, and Catherine Pelachaud. 2014. Dynamic stimuli visualization for experimental studies of body language. *MMC 2014, the 10th Workshop on Multimodal Corpora, held at LREC 2014, 27 May, Reykjavik, Iceland.*

Liyu Gong, Tianjiang Wang, Chengshuo Wang, Fang Liu, Fuqiang Zhang, and Xiaoyuan Yu. 2010. Recognizing affect from non-stylized body motion using shape of Gaussian descriptors. *Proceedings of the 2010 ACM Symposium on Applied Computing - SAC '10*, page 1203.

Harry J. Griffin, Min S.H. Aung, Bernardino Romera-Paredes, Ciaran McLoughlin, Gary McKeown, William Curran, and Nadia Bianchi-Berthouze. 2013. Laughter

Type Recognition from Whole Body Motion. *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, pages 349–355, September.

M Melissa Gross, Elizabeth A Crane, and Barbara L Fredrickson. 2010. Methodology for Assessing Bodily Expression of Emotion. *Journal of Nonverbal Behavior*, 34(4):223–248.

Halim Hicheur, Hideki Kadone, Julie Grèzes, and Alain Berthoz. 2013. Perception of emotional gaits using avatar animation of real and artificially synthesized gaits. *Humaine Association Conference on Affective Computing and Intelligent Interaction*.

A Kleinsmith and N Berthouze. 2007. Recognizing Affective Dimensions from Body Posture. *Affective Computing and Intelligent Interaction*, LNCS-4738:1–12.

Andrea Kleinsmith, Nadia Bianchi-Berthouze, and Anthony Steed. 2011. Automatic Recognition of Non-Acted Affective Postures. *IEEE Transactions on Systems Man and Cybernetics Part B Cybernetics*, 41(4):1027–1038.

Yingliang Ma, Helena M Paterson, and Frank E Pollick. 2006. A motion capture library for the study of identity, gender, and emotion perception from biological motion. *Behavior research methods*, 38(1):134–41, February.

Marco Meijer. 1989. The contribution of general features of body movement to the attribution of emotions. *Journal of Nonverbal Behavior*, 13(4):247–268.

Joann M Montepare, Sabra B Goldstein, and Annmarie Clausen. 1987. The identification of emotions from gait information. *Journal of Nonverbal Behavior*, 11(1):33–42.

Frank E Pollick, H M Paterson, A Bruderlin, and A J Sanford. 2001. Perceiving affect from arm movement. *Cognition*, 82(2):B51–B61.

Claire L Roether, Lars Omlor, Andrea Christensen, and Martin A Giese. 2009. Critical features for the perception of emotion from gait. *Journal of Vision*, 9(6):1–32.

Klaus R Scherer, Rainer Banse, Harald G Wallbott, and Thomas Goldbeck. 1991. Vocal cues in emotion encoding and decoding. *Motivation and Emotion*, 15(2):123–148.

Joëlle Tilmanne and Thierry Dutoit. 2011. Continuous control of style through linear interpolation in hidden markov model based stylistic walk synthesis. In *Cyberworlds (CW), 2011 International Conference on*, pages 232–236, Oct.

Jessica L Tracy and Richard W Robins. 2008. The nonverbal expression of pride: evidence for cross-cultural recognition. *Journal of Personality and Social Psychology*, 94(3):516–530.

Harald G Wallbott. 1998. Bodily expression of emotion. *European Journal of Social Psychology*, 28(6):879–896.

Xsens. 2005-2014. Xsens Technologies, http://www.xsens.com.