Emotion Recognition for Exergames using Laban Movement Analysis

Haris Zacharatos* University of Cyprus

Christos Gatzoulis† Bahrain Polytechnic

Yiorgos Chrysanthou‡ University of Cyprus

Andreas Aristidou§ University of Cyprus

Abstract

Exergames do not have the capacity to detect whether the players are really enjoying the game-play. The games are not intelligent enough to detect significant emotional states and adapt accordingly in order to offer a better user experience for the players. We propose a set of body motion features, based on the Effort component of Laban Movement Analysis (LMA), that are used to provide sets of classifiers for emotion recognition in a game scenario for four emotional states:concentration, meditation, excitement and frustration. Experimental results show that, the system is capable of successfully recognizing the four different emotional states at a very high rate.

CR Categories: I.2.10 [Computing Methodologies]: Artificial Intelligence—Vision and Scene Understanding I.4.7 [Computing Methodologies]: Image Processing and Computer Vision—Feature Measurement J.0 [Computer Applications]: General—;

Keywords: Emotion recognition, Laban Movement Analysis, Exergames

1 Introduction

New advances in non-intrusive user interfaces that use natural human gestures as input have resulted in high popularity of a new game genre called Exergaming. Exergames go beyond the passive gameplay activity that traditional controllers such as gamepads, keyboard and mouse offer, and require game players to become physically active. Through this, exergames are often used to promote a healthy lifestyle for both casual gamers that use such interfaces at home but also for special categories of users who need to advance their physical activity in order to improve specific health conditions [GERLING et al. 2010] [PAPASTERGIOU 2009]. Further to this, exergames provide a novel and livelier game experience that can also augment the fun factor, however research in this area is still in the early stages [KIILI and MERILAMPI 2010] [ISBISTER et al. 2011] [ISBISTER et al. 2012].

A major issue of the available exergames is that they do not have the capacity to detect whether the players are really enjoying the game-playing. The games are not intelligent enough to detect significant emotional states and adapt according to them in order to offer a better user experience for the players. While facial and audio information have been used successfully to detect emotions on users of desktop applications [GOLDMAN and SRIPADA 2005] [BUSSO et al. 2004] [FASELA and LUETTIN 2003] [VERVERIDIS and KOTROPOULOS 2006], exergame players express their emotions using their bodies as these modalities are more active and energetic during exergaming. Existing research that attempts to recognize emotions using human motion data does not achieve sufficient recognition rates, and is based on training the system with low level feature data that is very vague (such as rotation of a given joint on a given axes etc) and is selected without firm justification from movement analysis theories. Some recent approaches in robotics do achieve good quality recognition [MASUDA et al. 2009] [MASUDA et al. 2010], however their task is more simplified since robots perform mechanic and predetermined movements while expressive human movement is more complex and non-deterministic. Therefore it is not clear how applicable these methods are to real game playing situations.

This paper is a step towards overcoming the above limitations, by providing a novel method that achieves high recognition rates using real human motion data, captured during genuine game playing. It presents an emotion recognition model that makes use of human motion data dynamics derived from the widely accepted and applied movement analysis theories of Laban [LABAN 1974]. The features that are used to describe the emotional state vector are derived from the theories of Laban on Effort movement qualities. Four different game-playing related emotional states (excitement, frustration, meditation and concentration) are studied and training features extracted so that they can distinguish either single emotions or subsets of the above mentioned emotion set. As shown in the results, the system is capable of successfully recognizing the four different emotional states at a very high rate.

2 Related Work

Recently there is an increased pursuit in the field of Affective Computing, for automatic recognition of bodily expressions. Most early automatic recognition systems relied on corpora that had been acted. More recent studies are using non-acted data with body posture and movements, validating the results by using human observers. Different approaches have been used in order to get higher recognition rates for some basic emotions. Castelano [CASTELLANO et al. 2007], uses movement qualities such as amplitude, speed and fluidity of movement to infer emotions. Similar to this, Savva [SAVVA et al. 2012] tried to recognize emotions from animation by using low level features such as angular velocity, acceleration for the body's Arm, Hand and Right Forearm and body directionality for spine and head. In this approach, the recognition rate is average for individual emotions (happy 58%, concentrate 36%) and higher when categorization is been done for high and low intensity of emotions 67% and 70% respectively.

Many studies have been done for motion analysis by using Laban theory's such as [CHEN et al. 2011], [OKAJIMA et al. 2012] and [SANTOS et al. 2009]. [CAMURRI et al. 2004] examined emotion in dance. The results ranged between 31% and 46% for recognizing four emotions, far below from observer recognition rate of 56%. Lourens [LOURENS et al. 2010] extracted low level features from video and used Labanotation experts to classify the video clips to four emotional states manually. Another study used Laban features like whole-body movement, inclination of the body and area, to extract four emotions, pleasure, anger, sadness and relaxation from a robot that has limited ways of movement [MASUDA et al. 2010].

*e-mail:haris@cellock.com

†e-mail:christos.gatzoulis@polytechnic.bh

‡e-mail:yiorgos@cs.ucy.ac.cy

§a.m.aristidou@gmail.com

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Motion in Games 2013, November 7 – 9, 2013, Dublin, Ireland.

Copyright © ACM 978-1-4503-2546-2/13/11 $15.00

Although they used observers to classify the robot movements to emotions, they have not used automatic recognition techniques for classification. They used empirical estimation of correlation between Laban features and emotional set. In our approach we use human motion capture data to extract some of the Laban features for the body's extremity parts such as arms, legs and head and then perform automatic recognition techniques on those features to classify our emotional set: excitement, frustration, concentration and meditation. For our experiment, 'meditation'is the mental state during which users ignore the environment and focus on themselves. For example, a breathing moment, stretching and any other movement which draws the user's attention to his own body.

3 Laban Movement Analysis

Laban Movement Analysis (LMA) is a theory for observing, describing, notating and interpreting human motion. It was originally developed by dance artist and theorist Rudolf Laban in the early 20th century. The method focuses on the relationships between internal state, intention and attention and their effects on all human motions. One of the strong points of LMA is the ability to describe expressive content of movements, which makes it excellent for emotion and behavior analysis. Many researchers have been trying to create a computational form of LMA for motion analysis [BADLER et al. 1993] [ZHAO and BADLER 2005] [ZHAO 2002]. Nakata [NAKATA et al. 2002], reproduced expressive movements in a robot that could be interpreted as emotions by a human observer.

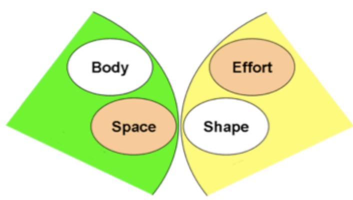

Figure 1: *Four major Components of Laban Movement analysis. Adopted by ZHAO 2002*

Theory divides LMA in four components shown in Figure 1. In this work we will focus on the Effort component that deals with the expressiveness and describes the dynamic qualities of the movement and the inner attitude towards using energy. By selecting a set of suitable features from the trajectories described by hands, foot and head, the effort component can be used as one descriptor for expressive movements. Laban sees Effort as the inner impulse-a movement sensation, a thought, a feeling or emotion- from which movement originates; it constitutes the interface between mental and physical components of movement. The inner impulse is expressed by way of Motion Factors. Every human movement including thought has potential to engage the four motion factors: Space, Weight, Time and Flow. Table 1 shows the motion factors, the underlying cognitive process associated with and the bipolar quality between two extremes of Effort component.

Motion Factor	*Cognitive process*	*Extremes*
Space	Attention-Thinking	Indirect-Direct
Weight	Intention-Sensing	Light-Strong
Time	Decision-Intuiting	Sustained-Sudden
Flow	Progression-Feeling	Free-Bound

Table 1: *Effort motion factors*

Space Motion Factor

As observed by Maletic [MALETIC 2005], motion factors have correlation with cognitive processes. The emphasis on attitudes toward Space can be associated with the cognitive capacities of orienting, attending and organizing. It addresses the quality of active attention to the surroundings. The two Extremes are Direct (Concentrate, Focused, pinpointing, narrowing down) and Indirect (multi-focused, with all-round attention).

Weight Motion Factor

The predominance of Weight qualities may indicate sensing or sensibility for assuming light or firm Intention towards an action. It senses the physical mass and its relationship with gravity. The two Extremes are Light(Accepting or Adjusting to gravity, delicate, lesser muscular tension) and Strong (resisting the pull of gravity, firm, forcefull).

Time Motion Factor

A great frequency of Time qualities may indicate an intuitive readiness for Decision making. Its mastery gives a calm or alert approach to thought or movement actions. The two Extremes are Sustained (Calm, slow tempo of movement) and Sudden (Excited, immediate, unexpected).

Flow Motion Factor

The emphasis on Flow can be associated with the emergence of feelings that, depending on the interaction with self or others, free or bind the continuity of movement and give either a controlled and carefull or exuberant and outgoing Progression. The two Extremes are Free (Accepting the continuity of movement, go with the flow) and Bound (Resisting the flux of movement, controlled, restrained).

4 Methodology

4.1 Data Collection and Processing

Thirteen players (ten male and three female) were asked to play sports games for 30 minutes each on the Xbox integrated with the Microsoft Kinect [MICROSOFT]. The motion data was collected using a PhaseSpace Impulse X2 motion tracking system with 8 cameras. A camera was also used to record all the sessions on video, in order to aid at a later stage the annotation of emotional states. The data annotation was done in a two-step process. First motion clips (of size no longer than 2 seconds) that potentially exhibit one of the 4 investigated emotions were extracted manually. Special care was given not to include frames at the beginning and end of the clip that are not significantly expressive in order to reduce noise in the learning process. A total of 309 clips were extracted. In the second step, four different observers through a multiple-choice questionnaire annotated each of the extracted clips, resulting on an agreement on 197 clips that became the ground truth for our system.

4.2 Feature Analysis

In the current implementation, the Space and Time motion factors of the Effort component were implemented. According to Laban [LABAN 1974], the Space motion factor represents the persons attention to the surroundings. It is related to attention and thinking. Indirect Space is multi-focused with all-around attention, while direct is focused with a tendency to align joints and bend. Laban states that concentrated behavior has direct space quality. Through observation of motion data it is easy to see that excitement and frustration are not focused movements, while meditation is a state of focusing on ones whole body rather on a single point. In the current study, Space motion factor is used to try to recognize concentrate emotional states from the other three emotional states. Space motion factor is implemented similar to Masuda [MASUDA et al. 2009], but taking into consideration the above theories of Laban.

With experimentation we have used and discarded multiple features like Quaternion Velocity and Acceleration, Torsion, Corner Curvature, Angular Displacement, Angular Velocity and Acceleration, Swivel Angles, Sternum Height [ZHAO 2002].

One of the features finally used is the percentage of narrowing down P_{ND} in the clip, which is calculated as the difference of the initial Y position of the head minus the average head Y position of the clip, divided by the initial Y position. Through observation, it is easy to see that in concentration clips, the player tends to bend resulting in significantly lower head positions throughout the clip frames.

$$P_{ND} = (Y_{InitialHead} - \bar{Y})/Y_{InitialHead} \tag{1}$$

Further to the above feature, to highlight a prospective focus of the movement to a given point (direct behaviour) the face direction $\overrightarrow{F}$ and the unit movement vectors of the four extremity points of the skeleton are used.

$$S = \{\overrightarrow{L}_{hand}, \overrightarrow{R}_{hand}, \overrightarrow{L}_{foot}, \overrightarrow{R}_{foot}\}$$

The dot product of the face vector with each of the four extremity movement vectors are calculated at each frame of the clip.

$$\forall x \in S, F \cdot x \tag{2}$$

Their signs are tested to see if the angle of each pair of vectors (hands or foot) is above or below 90° (resulting in indirect or direct movement). For each extremity point, the average of the dot product values of all the direct frames and the indirect frames are calculated as two separate features.

At the end eight features are calculated for all four extremity points. Together with the P_{ND} feature, they form the Space feature vector, as seen in Table 2 .

Feature	*Description*
P_{ND}	Percentage of narrow down
$DotL_{hand}Direct$	$(\overrightarrow{F} \cdot \overrightarrow{L}_{hand})$ for direct frames
$DotR_{hand}Direct$	$(\overrightarrow{F} \cdot \overrightarrow{R}_{hand})$ for direct frames
$DotL_{foot}Direct$	$(\overrightarrow{F} \cdot \overrightarrow{L}_{foot})$ for direct frames
$DotR_{foot}Direct$	$(\overrightarrow{F} \cdot \overrightarrow{R}_{foot})$ for direct frames
$DotL_{hand}inDirect$	$(\overrightarrow{F} \cdot \overrightarrow{L}_{hand})$ for indirect frames
$DotR_{hand}inDirect$	$(\overrightarrow{F} \cdot \overrightarrow{R}_{hand})$ for indirect frames
$DotL_{foot}inDirect$	$(\overrightarrow{F} \cdot \overrightarrow{L}_{foot})$ for indirect frames
$DotR_{foot}inDirect$	$(\overrightarrow{F} \cdot \overrightarrow{R}_{foot})$ for indirect frames

Table 2: *The Space feature vector*

The Time component represents the speed of the movement. According to Laban, it has to do with decision and intuition. Sustained movements are calm, with slow tempo, while sudden movements, are immediate, excited, unexpected and with fast tempo. Labans theory about Time and emotions correlates:

(a) {meditation,concentration} $\in$ Sustained

(b) {frustration,excitement} $\in$ Sudden

Time is implemented using the positional velocity(υ), acceleration(α) and jerk(j) (acceleration derivative) for the extremities of the body, both hands and foot. In Figures 2,3 and 4 we show the average velocity, acceleration and jerk of each extremity joint across all the clips. It shows that for the left foot the variation is small and thus does not contribute much and can be omitted from the feature set. The final feature set for the Time component comprises nine features, positional velocity, acceleration and jerk for left hand, right hand and right foot respectively as seen in Table 3.

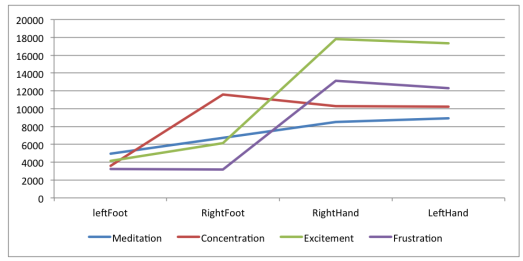

Figure 2: *Average Velocity*

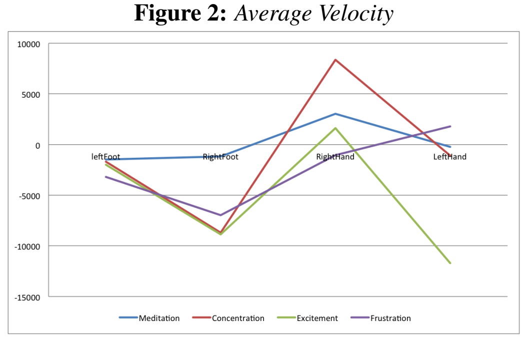

Figure 3: *Average Acceleration*

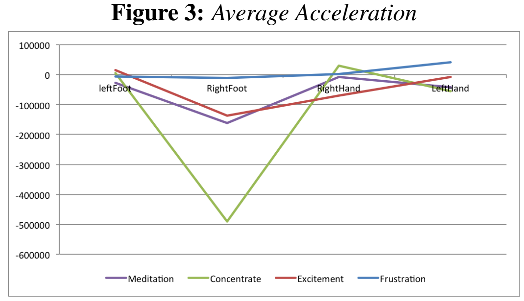

Figure 4: *Average Jerk*

Feature	*Description*
$L_{hand}V$	Velocity(υ) for right hand
$R_{hand}V$	Velocity(υ) for left hand
$R_{foot}V$	Velocity(υ) for right foot
$L_{hand}A$	Acceleration(α) for right hand
$R_{hand}A$	Acceleration(α) for left hand
$R_{foot}A$	Acceleration(α) for right foot
$L_{hand}J$	Jerk(j) for right hand
$R_{hand}J$	Jerk(j) for left hand
$R_{foot}J$	Jerk(j) for right foot

Table 3: *The Time feature vector*

4.3 Machine Learning Approach

To assess the validity of the selected feature sets for the Space and Time motion factors, and to measure the success in recognition of the targeted emotional states, the following tests were conducted:

Test1: Annotate all non-concentrate clips as one category and test to see if Space factor can distinguish between concentrate and non-concentrate clips. This can be used during automatic measurements of concentration on exergames in which acute cognitive benefits such as temporal improvements in concentration are being evaluated [GAO and MANDRYK 2012].

Test2: Annotate all excitement and frustration clips as the one category and all meditation and concentration clips as another and attempt to see how well the Time factor can recognize between the two categories. This can be used in a scenario where the valence of the emotion state of the user is required to be measured.

Test3: Attempt to recognize all four emotions against all others using a combined feature set.

We have used WEKA [HALL et al. 2009] to distinguished all 4 emotions against all others using a combined feature set. The whole data set was divided to 10 folds and each fold was used once as a testing set, while the rest acted as training sets. All the three tests have computed by Multi Layer Perceptron Classification algorithm. The results presented in this paper are the averages of the 10 trials.

5 Results

For Test 1 showed an overall 92,38% recognition rate for the binary set of concentrate emotion or other. As seen in Table 4, 36 from the 44 clips were recognized as concentrate, and 146 from 153 clips as other.

Concentration	*Other*	
$\mathbf{36}(\mathbf{82}\%)$	8	Concentration
7	$\mathbf{146}(\mathbf{95}\%)$	Other

Table 4: *Concentrate or not classification using the Space factor*

For Test 2 we have define a binary set of emotional states, when the clip is sustained or sudden. It showed a 91,87% recognition rate for the binary set of emotion, with 86 out of 96 clips were recognized as Concentrate or Meditation and 95 out of 101 clips recognized as Excitement or Frustration. The confusion matrix can be seen on Table 5.

Concentrate-Meditation	*Excitement-Frustration*	
$\mathbf{86}(\mathbf{90}\%)$	10	Con.-Med.
6	$\mathbf{95}(\mathbf{94}\%)$	Exc.-Fru.

Table 5: *Concentrate-Meditation vs Excitement-Frustration classification using the Time factor*

For Test 3, this time with all the four emotional states available we combined space features and time features in one set, with overall classification of 85.27%, with Kappa statistic 0.8031. The Confusion matrix can be seen on Table 6.

Meditation	*Concentrate*	*Excitement*	*Frustration*	
$\mathbf{45}(\mathbf{87}\%)$	2	0	5	Med.
ti ka 5	$\mathbf{39}(\mathbf{89}\%)$	0	0	Con.
1	1	$\mathbf{39}(\mathbf{83}\%)$	6	Exc.
3	1	5	$\mathbf{45}(\mathbf{83}\%)$	Fru.

Table 6: *All four emotions classification using the combined Space and Time feature set*

6 Conclusion

From the results achieved we can conclude that Laban Movement Analysis is a valid and promising approach for emotion recognition from body movements due to the abstract level of Labans technique. Specifically we have shown that two of Effort's component motion factors, Time and Space can result to high emotion recognition rates. The implementation of the rest of the Laban motion factors and components is one of our current goals and part of our future work. It is anticipated that this will further improve our recognition rates. This is very important as emotion recognition systems must be very accurate before they can be used within games to adapt game behavior, as any emotion recognition mistakes can have the opposite effect on the player's experience. It would also be interesting to integrate the method to an automatic emotion recognition system capable to be used by Exergames.

Acknowledgements

This project is co-financed by the European Regional Development Fund and the Republic of Cyprus through the Research Promotion Foundation (DIDAKTOR/0311/73)

References

BADLER, N., WEBBER, B., AND PHILLIPS, C. 1993. Simulating humans. *Computer Graphics, Animation and Control, Oxford Univ Press.*

BUSSO, C., DENG, Z., YILDIRIM, S., AND BULUT, M. 2004. Analysis of emotion recognition using facial expressions, speech and multimodal information. *6th international conference on Multimodal interfaces (ICMI '04).ACM, New York, NY, USA,*, 205–211.

CAMURRI, A., MAZZARINO, B., RICCHETTI, M., TIMMERS, R., AND VOLPE, G. 2004. Multimodal analysis of expressive gesture in music and dance performances. *Gesture-Based Communication Human-Computer Interactions*, 59–70.

CASTELLANO, G., VILLALBA, S., AND CAMURRI, A. 2007. Recognising human emotions from body movement and gesture dynamics. *Affective Computing and Intelligence Interaction*, LNCS 4738, 71–82.

CHEN, J., LIN, W., TSAI, K., AND DAI, S. 2011. Analysis and evaluation of human movement based on laban movement analysis. *Tamkang Journal of Science and Engineering 13*, 3, 255–264.

FASELA, B., AND LUETTIN, J. 2003. Automatic facial expression analysis: a survey. 259–275.

GAO, Y., AND MANDRYK, R. 2012. The acute cognitive benefits of casual exergame play. *In CHI '12: Proceedings of the 30th international conference on Human factors in computing systems., Austin, Texas, USA.*, 1863–1872.

GERLING, K., SCHILD, J., AND MASUCH, M. 2010. Exergame design for elderly users: the case study of silverbalance.

7th International Conference on Advances in Computer Entertainment Technology (ACE '10). ACM, New York, NY, USA,, 66–69.

GOLDMAN, A., AND SRIPADA, C. 2005. Simulationist models of face-based emotion recognition. *Cognition 94*, 3, 193–213.

HALL, M., FRANK, E., AND HOLMES, G. 2009. The weka data mining software an update. *SIGKDD Explorations 11*, 1.

ISBISTER, K., RAO, R., SCHWEKENDIEK, U., HAYWARD, E., AND LIDASAN, J. 2011. Is more movement better?: a controlled comparison of movement-based games. *6th International Conference on Foundations of Digital Games ACM, New York, NY, USA*, 331–333.

ISBISTER, K., KARLESKY, M., FRYE, J., AND RAO, R. 2012. Scoop!: using movement to reduce math anxiety and affect confidence. *the International Conference on the Foundations of Digital Games ACM, New York, NY, USA*, 228–230.

KIILI, K., AND MERILAMPI, S. 2010. Developing engaging exergames with simple motion detection. *14th International Academic MindTrek Conference: Envisioning Future Media Environments ACM, New York, NY, USA*, 103–110.

LABAN, R. 1974. *The Language of Movement: A Guidebook to Choreutics*. Plays, Boston, Mass, USA.

LOURENS, T., BERKEL, R., AND BARAKOVA, E. 2010. Communicating emotions and mental states to robots in a real time parallel framework using laban movement analysis. *Robot. Autonomus Systems 58*, 12, 1256–1265.

MALETIC, V. 2005. Dance dynamics: Effort and phrasing.

MASUDA, M., KATO, S., AND ITOH, H. 2009. Emotion detection from body motion of human form robot based on laban movement analysis. *Principles of Practice in Multi-Agent Systems, 12th International Conference, Nagoya, Japan*.

MASUDA, M., KATO, S., AND ITOH, H. 2010. Laban-based motion rendering for emotional expression of human form robots. *PKAW, LNAI*, 6232, 49–60.

MICROSOFT. Kinect controller for xbox360: http://www.xbox.com/en-us/kinect.

NAKATA, T., MORI, T., AND SATO, T. 2002. Analysis of impression of robot bodily expression. *Journal of Robotics and Mechatronics*, 14, 27–36.

OKAJIMA, S., WAKAYAMA, Y., AND OKADA, Y. 2012. Human motion retrieval system based on lma features using interactive evolutionary computation method. *Innovations in Intelligent Machines SCI*, 376, 117–130.

PAPASTERGIOU, M. 2009. Exploring the potential of computer and video games for health and physical education: A literature review. *Computers and Education*, 53, 603–622.

SANTOS, L., PRADO, J., AND DIAS, J. 2009. Human robot interaction studies on laban human movement analysis and dynamic background segmentation. *IEEE international conference on robots and systems*.

SAVVA, N., SCARINZI, A., AND BIANCHI-BERTHOUZE, N. 2012. Continuous recognition of player's affective body expression as dynamic quality of aesthetic experience. *IEEE Transactions on Computational Intelligence and AI in Games 4*, 3.

VERVERIDIS, D., AND KOTROPOULOS, C. 2006. Emotional speech recognition: Resources, features, and methods. 1162–1181.

ZHAO, L., AND BADLER, N. 2005. Acquiring and validating motion qualities from live limb gestures. *Graphical Models 67*, 1, 1–16.

ZHAO, L. 2002. *Synthesis and acquisition of laban movement analysis qualitive parameters for communicative gestures*. PhD thesis, University of Pennsylvania.