An investigation of the minimal visual cues required to recognize emotions from human upper-body movements

Donald Glowinski, Antonio Camurri, Paolo Coletta
InfoMus Lab – Casa Paganini
University of Genoa
Piazza Santa Maria in Passione, 34 - 16123 – Genoa
00390103532201
Donald.Glowinski@unige.it

Fabrizio Bracco, Carlo Chiorri
Disa-Department of Anthropological Sciences/psychology unit- University of Genoa
Corso Andrea Podestà, 2 – 16128 - Genoa
003920953709
Fabrizio.Bracco@unige.it
Carlo.Chiorri@unige.it

Anthony Atkinson
Department of Psychology, Durham University,
Science Laboratories, South Road, Durham, DH1 3LE, UK
a.p.atkinson@durham.ac.uk

ABSTRACT
This paper illustrates a research project on the visual cues that account for the ability to recognize emotions from upper-body movements. We describe an experiment that evaluated whether human observers could discriminate between high and low arousal emotions by means of impoverished displays representing the motion and shape information of the head and the hands. We investigated the effects of figure-ground segregation and inversion of the displays on (i) the recognition of emotion, (ii) the perception of biological motion and (iii) the animacy experience.

Categories and Subject Descriptors
H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems – *animations, evaluation/methodology*.

General Terms
Experimentation, Human Factors, Design

Keywords
Point Light Display (PLD), animacy, emotional portrayals

1. INTRODUCTION
Current research using impoverished visual displays (e.g. point-light displays) indicates that people can judge the emotional implication of an action from both form and motion information conveyed by human body movement (for a review see [1]). Santos et al. [2] found that the perception of biological motion *per se* might be influenced by the structural similarities between the point light display (PLD) and the human body. However, the inference of an emotional content is not completely dependent upon a human-like, organism-shaped structure in motion. Heider and Simmel [3] consistently showed participant's strong tendency to attribute intentions, thoughts and desires to moving two-dimensional shapes, thus resulting in an animacy experience. Most PLD studies focused on facial expression [5] on specific body limbs [6]or on the full body [7,8,9]. It would be rather new to observe whether a minimal point-light configuration that refers to the upper body could convey sufficient information for the discrimination of emotional states.

The upper body can be considered a privileged source of expressive information: Wallbott [10] relieved on that different types of hand and arm postures and movements are specific for some emotions (e.g. arms stretched sideways for terror). He also found that external (lateral or frontal) hand or arm movements are most frequent during 'active' emotions like hot anger, cold anger, elated joy and terror. In a related field of communication, sign language production, reveal that small detailed motions are performed in and around the face and upper body region where the receiver (looking at the signer's face) can naturally observe gestures in high acuity [11].

However this PLD-based approach may not be able to distinguish whether the perception of an emotional content is better supported by changes in shape of a human-like form or change in shape of a more abstract-geometrical form (animacy experience). As a first tentative step, Glowinski et al. [4] recently implemented a gesture-based automatic analysis procedure on a subset of emotion portrayals that distinguished between low and high arousal emotions. Successful discrimination of high v. low arousal emotion was achieved on the basis of two expressive motor cues, "energy" and "bounding-triangle perimeter", that accounted for, respectively, the movement activity and the spatial occupation of arms with respect to the body of actors portraying an emotion. However, it is still unclear whether human observers could distinguish between low and high arousal emotions on the basis of the two representations that have been used by the computer to achieve the analysis (i.e. the movement of the 3 points corresponding to the head and the two hands, and the changes in shape of the "bounding triangle" that relates these three points). Our study aims to determine whether human observers can use these displays to perform the same emotion discriminations. A first issue is to investigate whether the 3-dots light display or a simple geometrical form related to upper-body movements are sufficient for the discrimination of emotional expressions. A second issue is to investigate whether the recognition of a human form is a pre-requisite for successful emotion labelling. A third and final issue is to understand whether the attribution of an intention (related to an animacy experience) is a preliminary stage for inferring an emotion.

Such research should lead to concrete applications for affective computing, in particular for novel networked environment (e.g.

palm, cell phone) where specific visual constraints and limited computational resources should be considered for the communication of an expressive content.

2. PROBLEM OVERVIEW

PLD studies show that emotional expressions can be retrieved from kinematic patterns (transformational invariants), and geometric relevant patterns (structural invariants) [11,12,13,14]. Usually, 13 dots corresponding to body joints and the head are displayed to characterize the human body structure and produce a biological motion. Considering the relevance of the upper-body part for non-verbal communication, we tested a “reduced” PLD with three dots corresponding to head and hands.

However, such a minimal spatial configuration can create perceptual ambiguities. Without any contextual information, the three points can be considered as three interactive agents or as the three parts of a unique entity (referring here to the actor). The first situation has been extensively studied as test-case for assessing animacy experience: given such cognitively challenging stimuli, humans have tendency to anthropomorphize, and to give attribution to these simple moving geometric forms. In a recent study, Santos et al.[2] show that the animacy experience is reinforced in scenarios where at least two objects are interacting. In such situation, the observer is placed outside the frame of reference and evaluates object relations within the system. Our tendency for anthropomorphism is acknowledged for observable scenes that do not necessarily include oneself as an agent in the interaction [15,16]. In a situation of human-human expressive communication however, such feeling of self-involvement would be likely to occur and PLD-based studies previously mentioned still revealed our great sensitivity to motion and form’s cues to identify human activities and infer their mental states.

As our study deals with the emotion recognition based on a reduced – potentially ambiguous- three points displays animations, it follows from this brief overview that we should consider both processes respectively related to biological motion identification and to animacy experience not separately, but with respect to their possible interactions. To consider how these two processes can be related one to the other, we adopted a novel approach developed by McAleer et al. [17,18] that ensured a continuity between the production of display for animacy and for the PLD animations for they are both based on the same real human motions.

3. METHOD

3.1 Results from automated video analysis

Our experiment is based on a recent study that determined two expressive motor cues related to the kinematics of the upper part of the body: “energy” and “bounding-triangle perimeter”. These motor cues account for, respectively, the *movement activity* and the *expansiveness* of an actor portraying emotions and enabled a computer model to distinguish between low and high arousal emotions (elated or angry vs. relief or sad). Glowinski et al. [19] developed a color skin-tracking algorithm to extract the three blobs corresponding to the three main body extremities (hands and head) from video of professional actors portraying emotions (see Figure 1). Following a layered approach, Glowinski et al.’s analysis started from this low-level physical specification (the 3-D coordinates of head and hands and their related velocity) to determine the significant static and dynamic cues. Forty portrayals were selected from the GEMEP (Geneva Multimodal Emotion Portrayals) corpus, which consists of more than 7000 audio-video emotion portrayals, representing 18 emotions portrayed by 10 actors [20]. These selected portrayals cover the basic combinations of the major affective dimensions, arousal and valence.

3.2 Perceptual test

A rating study was created to test the perceptual validity of the two motor cues related to upper-body movements. The stimuli consisted of (i) 3 dots corresponding to the head and the two hands of actors, (ii) empty (transparent) triangles representing the area bounded by the three dots, and (iii) full (opaque) triangles with dots. Since we are studying the perception of emotion and animacy-related issues with a minimal display configuration (i.e., three moving points), it could be unclear whether, in the three dots condition, observers consider the stimulus as three independent objects or as three related parts of the same entity. This perceptual ambiguity was addressed by adding two displays (transparent and full triangle with dots) that were supposed to progressively reinforce the gestalt-based figure-ground segregation process. By linking the three points with edges (empty triangle), we enhance the bounding process that is commonly considered a central process for object perception (gestalt closure principle). By filling in the triangle (white on uniform black background), we created an opacity which is considered as a main object property [21].

A last relevant issue concerns the relation between emotion recognition and biological motion identification. The visual analysis of human motion is constrained by the hierarchical structure of the human body. The upside-down presentation of a PLD impairs biological motion perception as well as emotion recognition accuracy [22,23,24,25,26]. To assess the relevance of biological coherence for emotion recognition, we presented the displays in four tilted versions (0°, 90°, 180°, 270°), thus maintaining kinematic properties, but varying the human structural properties in order to prevent the study participants from recognizing human movement. In our design, the rotations of the image with respect to the horizontal line (90°, 270°) were added because in the case of three points, the upside-down presentation can potentially be considered as a normal posture: two points high above another one can naturally refer to someone keeping her/his arms outspread.

Participants were evaluated on their ability to recognize these stimuli as portraying (i) human or object-like movement and (ii) physical or intentional movement, and as (iii) expressing emotional content (see Figure. 1). We hypothesized that emotion recognition accuracy would gain from higher ratings related to (i) human-like movement and (ii) intentional action.

Emotion recognition is a complex perceptual and cognitive task that could be difficult to accomplish with such a poorly informative display. Recognition accuracy is the result of a progressive accumulation of information that may grow at an implicit level and that elicits a response whenever it overcomes a threshold. Threshold definition is typical of subliminal perception research [27] and is based on two main methods in collecting dependent variables: free response or alternative-forced-choice task (AFC). It is assumed that in free response it is possible to get a subjective threshold of conscious access to information, since the observer manages to give the correct answer without cues. The amount of information needed to reach the subjective threshold is higher than that affording the reach of an objective threshold, which is defined by the choice among some alternatives that are used as cues by the observer. The identification of the correct response thanks to the cue given by one of the alternatives needs less information and, therefore, the objective threshold is lower than the subjective threshold.

According to this rationale, we tested two groups where participants were submitted two rating methods. First group participants were proposed a 5-AFC with only the four emotions as response options (elation, anger, relief, sadness). In this situation, observers were forced to choose among the four options which correspond to the whole set of possible emotions. The restriction of the response options to these 4 labels may help the observers access information acquired when viewing the portrayals., On the other hand, we designed for the second group participants a 6-AFC with the possibility to choose an "other emotion" label if they thought the emotion represented was different from the four proposed ones. In this view, this method was similar to a free response task, . n In both studies observers also saw a neutral action as a control condition (a functional gesture without affective connotation).

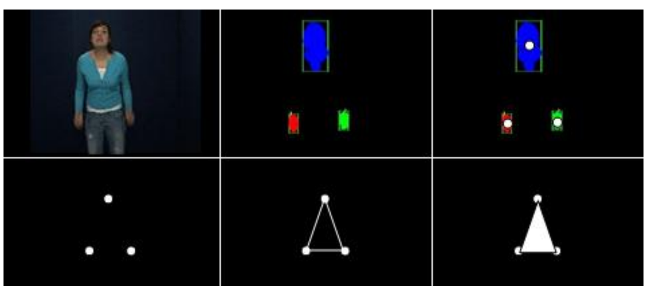

Figure. 1. Above, from left to right: (i) original image from a selected portrayal of the GEMEP data base, (ii) three extracted blobs based on color skin-tracking algorithms (iii) dots corresponding to the blob's centre of gravity. Below, from left to right, the three displays used as stimuli for the perceptual experiment: (i) three dots, (ii) empty (transparent) triangle with dots, (iii) full triangle with dots

3.2.1 Design

The study was implemented with a 3 × 4 × 5 × 4 × 2 mixed factorial design. Within-subject independent variables were Display (3: full triangle, empty triangle and points), Biological coherence (4: points or triangle representing original head and hands position and their respective 90°, 180°, 270° tilted versions), Emotions (5: anger, sadness, relief, joy and neutral), and Actor (4 different actors performing emotional portrayals). Group (2: absence vs presence of the "none of the above emotions" in emotion recognition options) was a between-subject factor. As a control condition for the Biological coherence factor we tilted the original clip along the four axes. As a control condition for the Emotion factor we added a neutral (i.e., non-emotional movement) performed by each actor and shown in the several display and coherence conditions. The purpose of this solution was to limit subject acquiescence to emotional displays.

3.2.2 Material

Point-light animations and empty or full triangles representing body movements of 4 actors performing actions with 4 emotions (anger, joy, relief, sadness) and with non-emotional gestures as a control condition. The selection of the 4 actors followed two kinds of consideration: avoiding gender bias (2 women and 2 men) and ensuring a constant visibility of the three points (head and hands) during the portrayal (sequence with no overlapping and no occlusion). The control condition was not originally presented in the GEMEP core-set. We recorded neutral, common human movement, i.e. movement that are related to the execution of daily life actions without any peculiar expressive content (e.g. pouring water in a glass). Five actors were chosen, whose body properties (size, build) approximated the original actors from the selected emotional portrayals. The recordings, using a dv cam standing in frontal view, were then started without any additional intervention from the experimenter. Recordings and psychophysical experiments were performed with informed consent of participants. All further data processing was done using EYESWEB XMI [28], an open platform for multimodal analysis,(body parts tracking) and MATLAB 7.0 (filtering, motor cues computation) following the protocol previously applied for the video material of the GEMEP core-set [20]: an application based on skin-color detection was used again to automatically track actors' head and hands and extract their trajectories. A preliminary pilot experiment showed that the emotional rating of such portrayals was lower than the experimental ones.

The three types of 2-D animations (PLD, transparent and opaque triangles) were created using Matlab. The three points (which also correspond to the three vertexs of the triangle) were determined by the head and hands positions previously extracted. The animation's size was thus actor-specific. The video processing utility Virtual Dub was then used for image compression, resizing to 340x340 resolution, and for 90°,180° and 270° rotations of the animations. The final square format of the screen was chosen to maintain a unique frame of reference that was independent of the display orientation. The 340x340 resolution was selected as it corresponds to the average standard of mobile devices (screen pixel resolution usually start from a 160x160 and reaches up to 640x480).

Stimuli were displayed and participants' responses were recorded using the EyesWeb-Mobius on a 15-inch monitor screen (60 Hz frame rate; 1440 × 900 pixel resolution). Within the Mobius graphical interface (see figure 2), stimuli were presented on a uniform black background display with a frame size of 7 cm in height and width viewed from a distance of about 50 cm, which subtended approximately 11,3 degrees of visual angle. White color was chosen for displaying the figure following previous examples of PLD. The Mobius platform, developed by InfoMus Lab, is an open source tool to create graphical user interface. It uses EyesWeb XMI as the underlying real-time engine to process multimodal data and let the researcher focus on the user front-end design. First applications have been developed in the biomedical research field (http://biomobius.trilcentre.org/). The Mobius platform is composed of a designer and a runtime components. The first one is used by the researcher to design the visual layout; the second component runs the experiment displaying data to the user and recording his reactions; it can also run on mobile devices (phones, PDAs, etc).

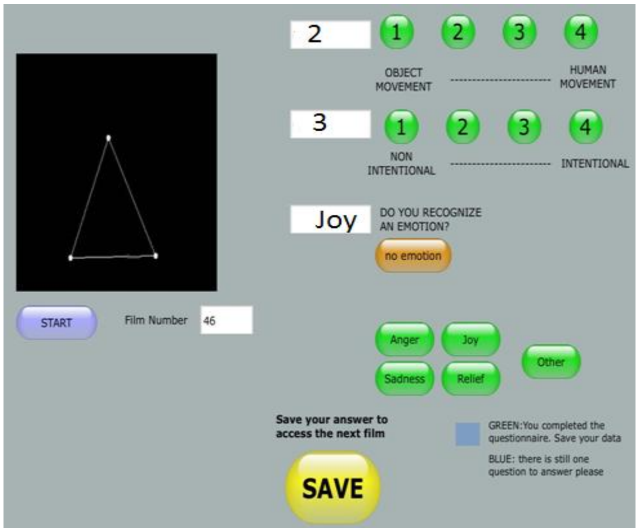

Figure. 2. Display of the EyesWeb-Mobius interface used for the stimulus presentation and for the real-time recording of participants' responses.

3.2.3 Participants
Altogether 24 individuals were recorded. Seventeen were assigned to the first group (ten females, mean age 31 years 4 months, range 24-40 years) and seven were assigned to the second group (five females, mean age 29 years 8 months, range 26-40 years). They all had normal or corrected-to normal vision.

3.2.4 Procedure
Before being presented the set of training stimuli, participants were instructed to (i) rate on a four-point semantic differential scale whether the stimuli were representing human or object-like movements (later named Q1), (ii) indicate whether stimuli were produced by passive compliance to physical laws or generated by an internal motivation (later named Q2) and (iii) identify the emotional content of the same stimuli, distinguishing them from non-emotional gestures. During the test session, participants were shown 240 (3 displays × 4 biological coherence positions × 5 emotions × 4 actors) animations counterbalanced for each factor. Each animation had a mean (± s.d.) duration of 2.75 ± 1.7 s (anger), 2 ± 0.81 s (joy), 2.5 ± 0.57 s (relief), 1.25 ± 0.5 s (relief) and 4.75 ± 2.36 s (neutral). The clip was repeated three times, after which participants had to perform the abovementioned three tasks.

3.2.5 Statistical Analysis
A modified path analysis model [29,30] was built to examine the direct and indirect effects of the independent variables (IVs: Group, Actor, Emotion, Orientation, Display) on human and animacy ratings and on emotion recognition performance. Four hierarchical log-linear submodels were tested: (i) IVs effects on Q1, (ii) IVs effects on Q2, (iii) the association of Q1 and Q2, and (iv) IVs, Q1 and Q2 effects on emotion recognition performance. In each case, a forward selection procedure was used to find the most parsimonious submodel. For each submodel, the likelihood ratio chi-square test statistic L^2 with its corresponding degrees of freedom was computed. These separate test statistics and their degrees of freedom were then summed up to obtain an overall test statistic. Standardized values of log-linear parameters λ of the model were (i) subsequently used to evaluate the significant ($p < .05$) deviations of observed frequencies from those expected under the assumption of no-association between variables and (ii) consequently test the hypotheses of this study.

4. RESULTS AND DISCUSSION
The forward selection procedure led to choose submodels with no higher order interaction effect than three-way interactions (Model (i) $L^2(4761) = 4713.99$, $p = .68$; (ii) $L^2(4995) = 4972.32$, $p = .59$; (iii) saturated model could not be rejected; (iv) $L^2(2890) = 2871.26$, $p = .59$; overall $L^2(12646) = 12557.57$, $p = .71$). The inspection of standardized λ values in the corresponding subtables revealed that high ratings of human motion and animacy were predictive of a better performance in emotion recognition, that is, participants could distinguish between neutral and emotional portrayals.

This study has demonstrated that emotional recognition accuracy depends both on the identification of a biological motion and on the animacy experience. Our results provide partial support for the so called theory of mind that claims that mental state are attributed to others to understand their behaviour. Our results show that either a moving body-related form, organism-shaped structure or simple geometric form can trigger an emotion recognition process as it creates the conditions for mental states attribution.

Our study points out that a minimal 3 points display animation corresponding to head and hands could be sufficient to give rise to the perception of a biological motion. To our knowledge, these results relieved on for the first time that this reduce PLD, when modeling main extremities of the upper-body part, can convey sufficiently kinematical and form-related information about the human body. This reveals an interesting trade-off concerning the sensitivity to human motion [31]: the more informative is the represented body area, the less could be the point light number.

Low human ratings (observer mainly perceived object movement) associated with high animacy ratings led to a good discrimination between high vs low arousal emotions. These results are consistent with previous studies on animacy experience where a moving 2d geometric shape can be observed as an intentional agent. In cases where the stimulus were considered as an object (low human ratings) with high intentional ratings, it could further be attributed a particular emotional states; the observer being able to distinguish between low and high arousal emotions. This arousal emotion recognition accuracy increased for high ratings of human movement.

A main effect of group was found for emotion recognition accuracy. Performances were better when participants performed a five alternative-forced-choice task (first group). When the “other emotion” label (second group) was available, participants were more likely to misclassified stimuli. This result suggested that participants in group one outperformed those in group two in emotion recognition. The presence of only four labels representing all the possible emotions to be selected, may facilitate the recognition task, since the subject may perceive an emotion and is forced to choose among the proposed alternatives. On the other hand, when “other emotion” alternative answer is available, participants may tend to choose between the four specific emotional labels only when they have enough confidence in the correctedness of their perception.

Screening of effects of display and orientation on human ratings revealed that higher human movement ratings were associated with PLD. This result is consistent with a large body of evidence about PLD, even if the number of light points was quite smaller (i.e., 3) than the 13-point light display that is usually employed to represent the full body structure. Interestingly enough, we could not find the same pattern of results for the transparent and the opaque triangle, whose perceptual appearance should be more consistent with the human body (e.g., figure-ground segregation). Nevertheless higher animacy ratings occurred when full triangle was presented, and this is consistent with findings about animacy experience showing that the attribution of intention can be retrieved even in objects, which are characterized by low human-like features.

Normal and 180° tilted stimuli were associated with higher human ratings. The 90° and 270° tilted stimuli were associated with lower human ratings (high object ratings). These results contradict the robust effect of inversion on the biological motion identification from PLD [26]. When 3-point light displays are shown upside-down we still observed a high human ratings, since the orientation did not seem to impair the perception of human structural properties. It appears that, at least for 3-point displays, the human vulnerability to inversion effect occurs when the display is tilted by 90° and 270°.

5. CONCLUSION

The present study only partially replicated previous findings that two expressive motor cues related to the kinematics of the upper body part could distinguish between high and low arousal emotions. The perceptual experiment we carried out on the same selected emotional portrayals unveiled the importance of minimal visual cues for emotion recognition performance relying on existent literature on PLD. We reviewed and refined original solutions to test minimal display configuration consisting of three points, related the most informative body parts (i.e., head and the two hands), and developed a new solution for assessing animacy experience. Since techniques described in literature to generate animacy displays do not allow for distinguishing between object and human motion (parametric variation of simple motion pattern), we grounded our animacy display on real actor performance. Through this procedure, we could reaffirm the role played by human motion perception and animacy in emotion recognition. We further confirmed the effectiveness of the PLD when compared to the simple 2d moving geometric shapes in enhancing the perception of human movement. The 3-point light display was vulnerable to the inversion effect in case of horizontal, rather than vertical, orientations of the stimulus. This is a new and notable feature of this kind of display, and it may be argued that other relevant properties may be found in further investigations.

We also pointed out a possible methodological flaw of classical methods of assessing emotion recognition accuracy. The availability of the “other emotion” response category may have led more conservative subjects to avoid a specific emotion labelling even when they perceived an emotional content. On the other hand, the absence of such option drove them to rely on information also present at a less explicit level.

Lastly, this study can be considered as an instance of successful implementation of Mobius and EyesWeb to run experiments and record data in real-time. Given its modular architecture, the proposed setup could be easily extended to run similar experiments on mobile devices.

6. AKNOWLEDGEMENTS

The This work has been partially supported by the EU-IST Project SAME (Sound And Music for Everyone Everyday Everywhere Every way). We greatly thank Yulia Munoz and Chiara Noera for their precious help.

7. REFERENCES

[1] Blake, R., & Shiffrar, M. 2007. Perception of human motion. Annual Review of Psychology, 58, 47–73.

[2] Santos, N.S., David, N., Bente, G., Vogeley, K. 2008 Parametric induction of animacy experience. Conscious Cognition. Volume 17, Issue 2, Pages 409-556 (June 2008).

[3] Heider, F., & Simmel, M. 1944. An experimental study of apparent behavior. American Journal of Psychology, 57, 243–249.

[4] Hill H, Jonno Y, Johnston A. 2003. Comparing solid-body with point-light animations. Perception 32:561–66.

[5] Pollick, F. E., Paterson, H., Bruderlin, A., & Sanford, A. J. 2001. Perceiving affect from arm movement. Cognition, 82. B51-B61.

[6] Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. 2004. Emotion perception from dynamic and static body expressions in point-light and full-light displays. Perception, 33, 717–746.

[7] De Meijer, M. 1989. The contribution of general features of body movement to the attribution of emotions. Journal of Nonverbal Behavior 13, 247–268

[8] Dittrich WH, Troscianko T, Lea SEG, Morgan D. 1996. Perception of emotion from dynamic point-light displays represented in dance. Perception 25:727–38

[9] Wallbott, H. G. 1998. Bodily expression of emotion. European Journal of Social Psychology, 28, 879-896.

[10] Siple, P. 1978. Visual constraints for sign language communication. Sign Language Studies, 19, 95–110.

[11] Dittrich WH. 1993. Action categories and the perception of biological motion. Perception 22:15–22

[12] Johansson, G. 1973. Visual perception of biological motion and a model for its analysis. Percept. Psychophys. 14:195–204

[13] Thornton IM, Pinto J, Shiffrar M. 1998. The visual perception of human locomotion. Cogn. Neuropsychol. 15:535–52

[14] Verfaillie K. 2000. Perceiving human locomotion: priming effects in direction discrimination. Brain Cogn. 44:192–213

[15] Abell, F., Happe, F., & Frith, U. 2000. Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. Cognitive Development, 15, 1–16.

[16] Scholl, B. J., & Tremoulet, P. D. 2000. Perceptual causality and animacy. Trends in Cognitive Sciences, 4(8), 299–309.

[17] McAleer P., Mazzarino B., Volpe G., Camurri A., Smith K., Paterson H.M. & Pollick F.E., 2004 Perceiving Animacy and Arousal in Transformed Displays of Human Interaction Journal of Vision 4 (8) pp 230.

[18] McAleer, P. & Pollick, F.E. 2008 Understanding Intention from Minimal Displays of Human Activity Behavior Research Methods (in press)

[19] Banziger, T., & Scherer, K.R. 2007. Using actors portrayals to systematically study multimodal emotion expression: the GEMEP corpus. In A. Paiva, R. Prada & R.W. Picard (Eds.), Lecture Notes in Computer Science: vol. 4738. Affective Computing and Intelligent Interaction, Second International Conference, ACII 2007, Lisbon, Portugal, September 12-14, 2007, Proceedings (pp. 59 - 70). Berlin: Springer Verlag.

[20] Glowinski, D., Dael, N., Volpe, G., Camurri, A., Scherer, K.. 2008. Technique for automatic emotion recognition by body gesture analysis, IEEE Workshop on CVPR for Human Communicative Behavior Analysis, June 24-26.

[21] Palmer, S. E., & Rock, I. 1994 Rethinking perceptual organization: The role of uniform connectedness. Psychonomic Bulletin and Review, 1 (1), 29-55.

[22] Bertenthal & Pinto, 1994 Bertenthal BI, Pinto J. 1993. Complementary processes in the perception and production of human movement. In Dynamic Approaches to Development: Vol. 2 Approaches, ed. LB Smith, E Thelen, pp. 209–39. Cambridge, MA: MIT Press

[23] Pavlova M, M, Sokolov A. 2000. Orientation specificity in biological motion perception. Percept. Psychophys. 62:889–98

[24] Shipley, T. F. 2003. The effect of object and event orientation on perception of biological motion. Psychological Science, 14, 377–380.

[25] Troje NF. 2003. Reference frames for orientation anisotropies in face recognition and biological-motion perception. Perception 32:201–10

[26] Atkinson, A.P., Tunstall, M.L., & Dittrich, W.H. 2007. Evidence for distinct contributions of form and motion information to the recognition of emotions from body gestures. Cognition, 104, 59-72.

[27] Cheesman J, . Merikle, p.m 1986. Distinguishing conscious from unconscious perceptual processes. Canadian Journal of Psychology, 40, 343-367.

[28] Camurri, A., De Poli, G., Leman, M., and Volpe, G. 2005. Toward Communicating Expressiveness and Affect in Multimodal Interactive Systemsfor Performing Art and Cultural Applications, IEEE Multimedia, 12,1 (Jan.2005), 43-53.

[29] Goodman, L. A. 1973. The analysis of multidimensional contingency tables when some variables are posterior to others: a modified path analysis approach. Biometrika, 60(1), 179–192.

[30] Hagenaars, J A. 1998. Categorical Causal Modeling. Sociological Methods and Research, 26(4), 436–486.

[31] Neri P, Morrone MC, Burr D. 1998. Seeing biological motion. Nature 395:894–96