

The Perception of Bodily Expressions of Emotion and the Implications for Computing

Winand H. Dittrich[¶] and Anthony P. Atkinson[‡]

[¶]*University of Hertfordshire;*

[‡]*Durham University*

UK

1. Introduction

Increasingly rapid technological developments in robotics, human-computer interaction (HCI), and the computer and video games industry have led to the development of quite sophisticated, intelligent robots and realistic, computer-generated characters. These robots and characters vary in their human-like qualities, from visual and behavioural replicas of humans, to androids, aliens or other beings that have some human characteristics, to creatures that have little in common with humans. The most common human-like qualities evidenced by these robots and characters are intelligence, looks, movements and language. Yet until very recently they were often not endowed with much in the way of emotions, and when they were they typically appeared emotionally underdeveloped, stilted, or false. For example, in James Cameron's film *Terminator 2*, Arnold Schwarzenegger plays an emotionless but intelligent humanoid robot. The film character explains that it is equipped with a microchip based on connectionist neural-net architecture. He comes to understand the patterning of human emotional behaviour but, of course, he does not have feelings or emotions. This was also true of other animated or computer-generated characters in films. Even if the characters gave rise to highly emotional video game or film scenes, the users or audience were in agreement that the characters were emotionally impoverished or did not have emotions at all. In large part this was and still is a technological issue: advances made in endowing machines and computer-generated characters with realistic emotions have not kept pace with the rapid advances in endowing such machines and characters with realistic intelligence, looks, movements and language. Nonetheless, given that these robots and characters are designed to interact with humans, and that emotions are so central to successful social interactions, effective implementation of emotions in these robots and characters is vital.

The latest developments in HCI include an area that is concerned with implementing emotions with mostly human-like characteristics into HCI technology. About 10 years ago, even a completely new field of computer science emerged that is based on such an attempt, namely "affective computing" (Picard, 1997). Since then, various attempts have been made either to implement emotions into machines or have machines that express emotional reactions. A close examination of these examples reveals two ways of implementing emotions and emotional reactions in HCI technology. First, emotions are seen as

constituting a state variable affecting the range or diversity of responses available to the machine or interface (Suzuki et al., 1998; Arkin et al., 2003) without changing the responses themselves. Second, emotions are linked to overt responses shaping directly the response characteristics of the machine or interface (e.g., Breazeal, 2003; Fujita, 2001), i.e. changing the response characteristics noticeably. This variation in how to implement emotions in machines seems to mirror the ambiguities in the understanding of human emotions.

Before the scientific study of emotions had disappeared almost completely in the wake of behaviourism, emotions were often introduced as the undesired side effect of behaviour. "I was so enraged" one says, "that I couldn't think straight". Emotions and cognitions were considered strange bedfellows and strictly separated for decades (for reviews and discussion, see e.g., Prinz, 2004; Oatley, 2004; Mandler, 2007). This concept of emotion as distinct from and irrelevant to cognition, in other words rather a disturbing factor in information processing, seems no longer tenable. Emotions play an important part in cognition and in the adaptive significance of ongoing behavioural decisions (e.g., Schachter & Singer, 1962; Damasio, 1994). We argue that the psychology of emotions, especially that which characterizes the relation between specific emotional states and the related facial, bodily, and vocal patterns, may provide a useful theoretical framework for model-driven applications in affective computing.

For this purpose a functional approach to emotion research seems to be the most promising (e.g., Frijda, 1986; Izard, 1991; Lazarus, 1991; Plutchik, 1980; Scherer, 1984). Schneider and Dittrich (1989) proposed a functional system-analytic approach to studying emotions. Following this approach, emotion can be described as a phylogenetically evolved mechanism for behavioural adaptation. Adaptation can take place both with respect to preparation for appropriate responses to environmental challenges and to providing chances for evaluation and communication of intentions in order to optimize response characteristics (see Schneider & Dittrich, 1989). In line with this approach, we shall highlight two essential functions of emotions with respect to computing and summarize research on the perception of emotional expressions. The two functional processes that are essential for emotional processing consist of (a) the intra-individual regulation of cognition and actions and (b) the inter-individual regulation in communication.

In human psychology, emotion research has concentrated on the study of, on the one hand, basic and unique human emotions as shown in facial, vocal and body expressions and their role in communication, and on the other hand, on affective states and internal bodily reactions and their role in cognition or behavioural organization. With respect to the latter research programme, Arbib and Fellous (2004) point out that the neural and chemical basis of animal function differs greatly from the mechanics and electronics of current machines, such that implementing emotions in such machines will require functional theories of emotion that abstract away from the details of the biological substrates. Our concern in this chapter is primarily with the former research programme and its implications for affective computing. Nonetheless, it is important to note that in animals these two aspects of emotions (expression/communication and behavioural organization) co-evolved and remain intimately linked, and as we shall point out below, implementing one might be aided by implementing the other.

The recognition of emotions in communicative social situations relies strongly on visual cues. Following Darwin, Paul Ekman and colleagues (e.g., Ekman et al., 1972; Ekman, 1992) proposed that there are only a limited number of basic and unique facial expressions. They

characterized six emotional facial expressions on the basis of a unique subset of facial muscle movements: disgust, surprise, fear, joy, anger, and sadness. These emotions were found in infants and blind people, who had no opportunity to imitate them. Also, the appearance, range and interpretation of these emotions are supposed to be similar in Papua New Guinea or Hatfield, England. These facial expressions of basic emotions have been systematically characterized in terms of individual facial muscle movements using the Facial Action Coding System, or FACS (Ekman & Friesen, 1978).

As exemplified by the above work, the study of emotion perception has relied primarily on visually presented static facial expressions. While faces are a vitally important – arguably the most important – source of cues in non-verbal communication in humans, they are not the only source of such cues. In the visual modality, postures and movement of the body and its parts also make a substantial contribution to non-verbal communication, including the communication of emotions; yet relatively little work has been devoted to the study of the perception of emotional body expressions. A central aim of this chapter is to review the current state of knowledge about what cues the human visual system uses in the perception of body postures and movements, with particular focus on emotion perception, and about the functional organization of the underlying neural systems.

An important reason why robots and animated characters can appear emotionally impoverished is that they rarely express emotions in realistic ways. The attribution of emotions by an observer or interactant depends a lot on the character's emotional expressions – not just what she does but how she expresses her (real, simulated, or intended) inner feelings. Only if someone shows emotional reactions are we likely to attribute emotional states to this person. Indeed, humans even tend to interpret very simple cues as indications of emotional reactions. For example, in animated scenarios a simple change of colour might give rise to emotional interpretations in the viewer or user (e.g. Miwa et al., 2001). Furthermore, when viewing a moving abstract stimulus, people tend to attribute social meaning and purpose to the movement (e.g., Heider & Simmel, 1944; Dittrich & Lea, 1994) (reviews: Dittrich, 1999; Scholl & Tremoulet, 2000).

Research on visual recognition has been dominated by studies about the recognition of objects. In these studies, the roles of form and motion cues have been investigated and various models of object recognition rely on the dissociation of the processing of form and motion cues. In sections 2 and 3 we will outline major lines of research in human psychology and neuropsychology that have applied this approach to the study of form and motion cues in the perception of social objects, and specifically bodies, and in the perception of emotional expressions from postures and movements of the body. Then in section 4 we shall place this work in the wider context of affective computing.

2. The perception of bodily form and motion

2.1 The visual cues that humans use to perceive bodies and their motion

The *form* of the human body could be represented in several different ways, demarcating points on a configural-processing continuum, from part-based to holistic processing. Thus bodies could be represented in terms of individual body parts or features, the relative positions of those parts (i.e., first-order spatial relations), the structural hierarchy of body parts (i.e., first-order configuration plus information about the relative position of features with respect to the whole body), or in terms of whole body posture templates (Reed et al., 2006). Unlike faces, the relative positions of body parts change as people move, which

suggests the need for a relatively fine-grained structural description of the spatial relationships among body parts.

A series of experiments by Reed and colleagues (2006) suggests that the recognition of body postures depends on the processing of the structural hierarchy of body parts. This study drew on the well-known inversion effect in face recognition, that turning faces upside down impairs the ability to recognize their identity more than inverting nonface objects impairs the recognition of their identity. It is generally considered that face inversion disrupts configural processing, specifically the coding of second-order relational information, that is, the metric distances amongst features (e.g., Diamond & Carey, 1986; Rhodes et al., 1993; Maurer et al., 2002). In Reed et al.'s (2006) study, participants had to judge whether two sequentially presented images were the same or different. Each pair of images was presented either upright or inverted. Performance was significantly impaired for inverted compared to upright whole-body postures but not houses, replicating their earlier finding (Reed et al., 2003). The matching of isolated body parts (arms, legs, heads) was unaffected by inversion, indicating that, as with isolated facial features, individual body parts do not evoke configural processing. Disrupting first-order spatial relations, by rearranging the body parts around the trunk (by e.g., putting the arms in the leg and head positions), abolished the inversion effect, indicating that such first-order configural cues do not contribute to body posture recognition. Presenting half-body postures that were divided along the vertical midline (i.e., left or right halves), which preserves the structural hierarchy of body parts but disrupts holistic template matching, did not abolish the body inversion effect. In contrast, presenting half-body postures that were divided along the horizontal midline (the waist), which preserves salient parts (e.g., both arms) but disrupts structural hierarchy information, did not produce an inversion effect. Thus the particular form of configural processing critical to body posture recognition, as indexed by the presence of an inversion effect, appears to be the structural hierarchy of body parts, that is, the positions of body parts relative to themselves and to the whole body.

There are three main classes of information pertaining to the *movements* of human bodies: the changes of structural or form information over time (including motion-mediated structural information), kinematics (e.g., velocity, acceleration, displacement) and dynamics (motion specified in terms of mass and force). Considerable attention has been given to the role of kinematics in specifying cues for action and person perception (e.g., Westhoff & Troje, 2007). Typically, these studies employ *point-light* or *patch-light* displays of human or other biological motion, in which static form information is minimal or absent but motion information (kinematics and dynamics) and motion-mediated structural information are preserved (Johansson, 1973). Point-light displays of body movements provide a sufficient basis for observers to discriminate biological motion from other types of motion, and to make accurate judgements about the people making the movements, including sex from gait (e.g., Barclay et al., 1978), identity from gait (Richardson & Johnston, 2005) or actions (Loula et al., 2005), the weight of boxes from the lifting movement (Runeson & Frykholm, 1981), and complex individual or social actions from whole-body movements (Dittrich, 1993). Some of this evidence shows equivalent or near equivalent performance with point-light compared to full-light (or solid-body) displays, in which the whole body is visible (e.g., Runeson & Frykholm, 1981), which suggests that static form cues are rather less important than motion cues and may often be unnecessary for successful judgements about people and their actions based on their visible behaviour. Evidence for the relative importance of

kinematic cues comes from studies that measure the effects on recognition of changes in certain kinematic or structural dimensions of point-light stimuli. For example, accuracy in judging the sex of point-light walkers was influenced more by “body sway” than by the ratio of shoulder to hip width, in Mather and Murdoch’s (1994) study, and was greater when point-light walkers were normalized with respect to their size (thus providing only motion information) than when they were normalized with respect to their motion information (thus providing only size cues), in Troje’s (2002) study.

It has been argued that the ability to discriminate at least simple biological movements in point-light displays may be based on relatively low-level or mid-level visual processing that does not involve the reconstruction of the form of body parts or of the whole body, either from static form or motion-mediated structural cues (e.g., Casile & Giese, 2005; Mather et al., 1992). Nevertheless, neuropsychological and neurophysiological evidence demonstrates that form information can indeed subserve biological motion perception from point-light displays (e.g., Hirai & Hiraki, 2006; McLeod et al., 1996; Peelen et al., 2006; Vaina et al., 2002). The processing of changes in the form of the body over time may be particularly important (e.g., Beintema & Lappe, 2002), especially in the context of more sophisticated tasks, such as recognizing emotional states or complex actions (Casile & Giese, 2005; Giese & Poggio, 2003). This conclusion gains some support from inversion effects in biological motion perception. The spontaneous identification of point-light motion displays as biological motion is impaired when they are shown upside down (Bertenthal & Pinto, 1994; Pavlova & Sokolov, 2000; Shipley, 2003; Troje, 2003), even given prior knowledge about display orientation (Pavlova & Sokolov, 2003). Moreover, neural activation characteristic of upright biological motion displays is attenuated or absent when such displays are inverted (Grossman & Blake, 2001; Pavlova et al., 2004). Inversion of point-light displays also disrupts the ability to distinguish the identity of the actors from their actions (Loula et al., 2005), and sex judgements based on gait tend to be reversed (Barclay et al., 1978). While it is likely that inversion of biological motion disrupts the processing of dynamic cues related to movement within the earth’s gravitational field (Barclay et al., 1978; Bertenthal et al., 1987; Pavlova & Sokolov, 2000; Shipley, 2003), there is also some evidence to suggest that inversion of whole-body movements impairs the processing of configural information (Lu et al., 2005; Pinto & Shiffrar, 1999).

2.2 The human brain contains regions specialized for processing bodily form and motion

The form of the human (or primate) body is a category of visual object for which there appears to be both selectivity and functional specialization in higher-level visual cortices. By *selectivity* we mean the extent to which a mechanism is activated by or operates over a particular stimulus class, such as faces or bodies, as compared to other stimulus classes. By *functional specialization* (or function for short) we mean a mechanism’s specificity for performing a particular process. Evidence for body-selective visual mechanisms comes from studies of both humans and non-human primates (reviewed by Peelen & Downing, 2007). In humans, the evidence points to two distinct regions, dubbed the extrastriate body area (EBA), located in lateral occipitotemporal cortex (Downing et al., 2001), and the fusiform body area (FBA), located in fusiform gyrus (Peelen & Downing, 2005; Schwarzlose et al., 2005). The EBA and FBA respond selectively to human bodies and body parts compared with objects, faces, and other control stimuli, despite considerable anatomical overlap

between the FBA and the face-selective fusiform face area (FFA) (Peelen & Downing, 2005; Schwarzlose et al., 2005; Peelen et al., 2006) and between the EBA, motion processing area V5/MT, and object-form-selective lateral occipital complex (Peelen et al., 2006; Downing et al., 2007).

With respect to functional specialization, the EBA represents the static structure of viewed bodies (Downing et al., 2006; Peelen et al., 2006; Michels et al., 2005), although these representations appear to be at the level of individual body parts rather than at the level of whole-body configuration (Taylor et al., 2007; Urgesi et al., 2007a). As discussed above, configural cues in body perception include the relative positions of body parts and the positions of those parts with respect to the whole body (Reed et al., 2006), and there is evidence indicating that the processing of one or other or both of these configural cues is more a function of the FBA than of the EBA (Taylor et al., 2007).

Another region implicated as having a critical role in processing configural body cues is left ventral premotor cortex (Urgesi et al., 2007a). This region of inferior frontal cortex is known for its role in both the planning of motor actions (Johnson & Grafton, 2003) and in the visual discrimination of such actions (Grafton et al., 1996; Pobric & Hamilton, 2006; Urgesi et al., 2007b), which has led to the suggestion that it forms part of a system for simulating the observed action to allow it to be understood (Gallese et al., 2004; Rizzolatti & Craighero, 2004).

The EBA appears to constitute a critical early stage in the perception of other people (Chan et al., 2004), rather than a later processing stage via, for example, top-down effects related to imaginary gestures and movement (de Gelder, 2006). Evidence in support of this claim comes from recent studies using either intracranial recordings or transcranial magnetic stimulation (TMS). Pourtois et al. (2007) recorded highly body-selective visual evoked potentials over the EBA of a patient that started approximately 190ms and peaked 260ms after stimulus onset. Consistent with this finding are reports of selectively impaired perception of body form following application of TMS over EBA at 150–250 ms (Urgesi et al., 2004; Urgesi et al., 2007a) and at 150–350ms (Urgesi et al., 2007b) post-stimulus onset. Despite this evidence, however, it is entirely possible that, in addition to its role in the early visual processing of body form, the EBA also plays a role in later processing stages of person perception. Little is yet known about the timing of the FBA and ventral premotor cortex involvement in body and person perception, although given that they preferentially represent configural over body-part cues it is likely that their initial involvement occurs subsequent to that of the EBA. Nonetheless, as Taylor et al. (2007) comment, a strictly serial model is probably too simplistic, given the widespread bi-directional connectivity in visual cortex.

Our brains contain systems specialized for processing the *movements* of bodies and their parts (including faces), in addition to those systems specialized for processing bodily facial form. (Although as we shall soon see, the computations performed by these biological motion-processing systems may well draw on form information.) Important early evidence came from neuropsychological lesion studies, which demonstrated spared ability to discriminate biological motion stimuli despite severe impairments in discriminating other types of motion (Vaina et al., 1990; McLeod et al., 1996). However, not all aspects of biological motion perception are normal in such 'motion blind' patients. For example, McLeod et al. (1996) report a case of a subject who was able to describe accurately a variety of actions from whole-body movements represented in point-light displays, but was unable

to report in which direction the figure was facing, or whether it was approaching or retreating from her. Furthermore, this same patient was severely impaired at identifying natural speech from point-light or fully illuminated facial movements, despite being unimpaired in recognizing speech-patterns from face photographs (Campbell et al., 1997).

In patients with relatively spared biological motion perception despite deficits in perceiving other sorts of motion, the lesions are restricted to ventral and middle occipito-temporal cortices, sparing superior temporal and parietal areas. Electrophysiological and neuroimaging studies confirm a particularly important role for superior temporal cortex in the perception of body and facial movement (Puce & Perrett, 2003; Allison et al., 2000). Single-cell recording studies in monkeys revealed neurons in superior temporal sulcus (STS) and superior temporal gyrus (STG), especially in the anterior portion of superior temporal polysensory area (STPa), selective for various types of face, limb and whole body motion (e.g., Jellema et al., 2000; Oram & Perrett, 1994; Perrett et al., 1985). Functional imaging studies in humans show that whole-body movements as represented in point-light displays elicit activation in pSTS compared to a variety of non-biological movements (e.g., Bonda et al., 1996; Grossman et al., 2000; Grossman & Blake, 2002; Pelphrey et al., 2003; Peuskens et al., 2005; Vaina et al., 2001). Regions of posterior and middle STS and surrounding superior and middle temporal gyri are also selectively activated by movements of the face or other body parts, as represented in fully-illuminated displays, compared to static images of the same body parts (Wheaton et al., 2004) and to non-biological motion (Puce et al., 1998). Disruption of the activity of right pSTS using TMS has confirmed a critical role for this region in perceiving body movement (Grossman et al., 2005). More recently, a lesion-overlap study with 60 brain-damaged subjects showed that impairments in the ability to discriminate whole-body from non-biological motion in point-light displays were most reliably associated with lesions in posterior temporal and ventral premotor cortices, which corresponded with the regions whose activity in neurologically intact subjects was selective for the same point-light whole-body movements (Saygin, 2007). The critical involvement of ventral premotor cortex in this study confirms earlier studies showing selectivity in this region for point-light whole-body movements (Saygin et al., 2004; Pelphrey et al., 2003).

There are also reports of selectivity to biological motion, in the form of whole-body movements, in the posterior inferior temporal sulcus/middle temporal gyrus (Grossman & Blake, 2002; Michels et al., 2005; Peuskens et al., 2005; Saygin et al., 2004), which might reflect activation of body-selective neurons in the EBA or motion-selective neurons in the overlapping V5/MT. There is even a report of selectivity to whole-body movements in the face-selective FFA (Grossman & Blake, 2002), which might reflect activation of body-selective or face-selective neurons, or of both body- and face-selective neurons. These last two issues have been resolved by a recent study: biological (whole-body) motion selectivity in occipitotemporal cortex was correlated on a voxel-by-voxel basis to body selectivity (i.e., EBA and FBA activation) but not to face selectivity (i.e., FFA activation) or to non-biological motion selectivity (i.e., V5/MT activation) (Peelen et al., 2006).

Neuroimaging studies in humans have also revealed distinct regions of STS selective for the movements of different body parts. While face, hand, mouth, and leg movements activate substantially overlapping regions of right pSTS (Thompson et al., 2007; Wheaton et al., 2004; Pelphrey et al., 2005), movements of the face (Thompson et al., 2007; Wheaton et al., 2004) and mouth (Pelphrey et al., 2005) are also associated with activity along the mid-posterior STS, as are leg movements (Wheaton et al., 2004). Moreover, whereas both facial speech

(principally mouth) and visually similar but linguistically meaningless facial movements activate right pSTS, speech and non-speech facial movements also elicit dissociable patterns of temporal cortex activation, with speech movements activating traditional language processing areas in both hemispheres, including auditory cortex (Campbell et al., 2001; Calvert et al., 1997). In addition to activating pSTS, hand motion is associated with activity in inferior right pSTS and inferior parietal lobule (Thompson et al., 2007), extending into middle occipital and lingual gyri (Pelphrey et al., 2005), whereas eye movements are associated with activity in more superior and posterior portions of the right pSTS (Pelphrey et al., 2005) and elicit stronger responses in these pSTS regions for mutual than for averted gaze (Pelphrey et al., 2004). Other areas, including ventral premotor and intraparietal cortex, also show differential selectivity to the motion of different body parts, in a somatotopic manner (Wheaton et al., 2004; Buccino et al., 2001).

Selectivity to whole-body movements relative to non-biological motion is evident as early as 80-100ms post-stimulus onset over the left parieto-occipital region (Pavlova et al., 2006; Pavlova et al., 2004), regardless of whether the participant is attending to the stimuli (Pavlova et al., 2006). Subsequent selectivity for whole-body motion is evident at several different stages, each associated with different brain regions, including fusiform and superior temporal cortices, but typically only when the stimuli are attended (2006; Pavlova et al., 2004; 2007; Hirai et al., 2005; Jokisch et al., 2005; Hirai & Hiraki, 2006).

The distribution of responses in STS and surrounding cortex to the motion of different body parts suggests a functional organization in which distinct but overlapping patches of cortex extract body-part specific representations of biological motion, with a posterior region of STS, especially in the right hemisphere, encoding a higher-level representation of biological motion that is not dependent on the particular body part generating that motion. Consistent with the first part of this hypothesis is the considerable evidence for an important role for areas of STS in the integration of motion and form information, especially that related to social perception (e.g., Beauchamp, 2005; Oram & Perrett, 1996; Puce et al., 2003; Vaina et al., 2001). With respect to the second part of this hypothesis, there is some debate over whether pSTS analyzes local image motion and higher-level optic flow or some more global motion of the whole figure (Lange & Lappe, 2006; Giese & Poggio, 2003; Thompson et al., 2005; Beintema & Lappe, 2002). While this issue has yet to be fully resolved, recent evidence is building up in favour of the latter proposal. Two computational models of biological motion perception (Giese & Poggio, 2003; Lange & Lappe, 2006) propose that a ventral form pathway derives 'snapshots' that represent the various static postures comprising a movement sequence. Neuroimaging evidence indicates these snapshots are derived by the EBA and FBA (Peelen et al., 2006). In one model (Giese & Poggio, 2003), these snapshots are summated and temporally smoothed in ventral visual areas on the basis of local image motion information derived in separate areas, including pSTS. The other model (Lange & Lappe, 2006) proposes that more superior cortical areas, especially pSTS, temporally integrate sequences of intact body configurations, a suggestion also supported by neuroimaging evidence (Thompson et al., 2005; Peuskens et al., 2005). Nonetheless, it is possible that pSTS both analyzes local image motion, at an early stage, and, at later stage, the more global motion information related to changes in body and body part configurations over time, subsequent to the analysis of individual configurations of body form in the EBA and FBA and facial form in the face-selective areas of occipital and fusiform cortex and perhaps also STS.

2.3 Putting it all together: How we perceive and understand bodily actions

We have been emphasizing the use of structural form and motion information in the perception and identification of bodies (as distinct from other visual objects), body postures and movements. The evidence reviewed above points to the following account. The human visual system contains mechanisms specialized for processing the form of the human body, with distinct mechanisms encoding the form of body parts and configural relations between those parts. The ability to recognize (at least non-emotional) body postures relies on configural cues towards the template end of a continuum that extends from part-based to holistic processing, specifically, on cues specifying the structural hierarchy of body parts. The human visual system also contains mechanisms specialized for processing the motion of human bodies and body parts, as distinct from other forms of motion. Some mechanisms encode local image motion and higher-level optic flow, and a full model of how these various routines and processes are integrated and work together is captured in the Interactive Encoding Model of e-motion perception (Dittrich, 1999). He argues that it seems unlikely that one or both types of these mechanisms will be found to be specific to the processing of biological motion. Other mechanisms, which have been suggested to be specific to biological motion (see reviews below), encode the translations of body form over time, integrating form information captured in snapshots of the moving body with the local motion and optic flow information. Again, Dittrich (1999) argues that neither of these general mechanisms (snapshot capturing or optic flow) seem specific to or even necessary for perceiving biological motion. Instead, as proposed in the Interactive Encoding Model, the processing of biological motion depends on a particular way of neuronal coding in the brain by motion integrators (instead of motion detectors) and three cognitive routines depending on the motion information available. The first routine is strictly associated with the analysis of the structural components of human motion to reconstruct 3D body-related emotion information out of the 2D motion trajectory. The second routine is part of the working memory system and allows the application of cognitive constraints relating to human emotions and their motion trajectories for the 3D reconstruction. The third routine relies on visual semantics related to emotion categories as stored in long-term memory.

Our review here has been necessarily brief. For readers interested in finding out more about the intricacies of human body and biological motion perception, in addition to Dittrich's (1999) Interactive Encoding model, we suggest the following reviews: de Gelder (2006), Giese and Poggio (2003), Puce and Perrett (2003), Blake and Shiffrar (2007), and Peelen and Downing (2007).

Is such an image-processing account, or an extension of it, sufficient for explaining our ability to perceive and understand bodily *actions*? The majority of human postures and movements are not aimless but are directed towards some purpose or goal and thus reflect that person's intentions and may also or instead reflect their emotional and other internal states. Moreover, humans are not passive observers but like the people whose postures and movements they are observing, have intentions and emotions and act in a purposive, goal-directed manner. Research and theory over the past decade or so suggests that one (and perhaps the main or even the only) route to understanding others' actions depends on the observer's own action capabilities (Rizzolatti & Craighero, 2004; Gallese et al., 2004). One productive source of evidence for this view is the body of findings showing neural mechanisms with dual functions in action perception and action production. For example, so-called mirror neurons in the premotor cortex of monkeys were found to respond not only

when the monkey prepares to perform an action itself, but also when the monkey observes the same visually presented action performed by someone else (e.g., Rizzolatti et al., 1996). Various supportive findings suggesting the existence of a mirror neuron system, if not actual mirror neurons, have also been obtained in humans. Observing another's actions results in desynchronization in motor cortex as measured with magnetoencephalography (Hari et al., 1998), and lowers the threshold for producing motor responses when transcranial magnetic stimulation is used to activate motor cortex (Strafella & Paus, 2000). Imitating another's actions via observation activates premotor cortex in functional imaging studies (Iacoboni et al., 1999); moreover, such activation is somatotopic with respect to the body part that is observed to perform the action, even in the absence of any overt action on the part of the subject (Buccino et al., 2001).

These and numerous other findings indicating the existence of a human mirror neuron system suggest a simulation account of action understanding, according to which observing another perform an action triggers in the observer an offline simulation of the viewed action. Work on motor control indicates ways in which such simulations may be computationally instantiated, in the form of forward models, inverse models, or both (Miall, 2003; Grush, 2004; Wolpert et al., 2003). Forward models use copies of the motor commands to map the current sensory states and motor commands to the future sensory and motor states that will result once the current motor commands have been executed. Inverse models perform the opposite transformations, by mapping sensory representations associated with the intended action to the motor commands to execute the action. One suggestion, for example, is that proposed by Wolpert et al. (2003) and also comprehensively envisaged in the Interactive Encoding Model (Dittrich, 1999): When observing another's action, the observer's brain generates a set of motor commands that would be produced given the observed movements and the current state of the observed person. Rather than driving the observer's own motor behaviour, these motor commands are used to predict the sensory and motor consequences of the observed action, which are then compared with the observed new state of the actor.

3. Bodily form and motion cues in emotion perception

There is compelling evidence that the kinematics of body and body-part movements are at least sufficient, and may often be important, in furnishing cues for the perception of emotional expressions. For example, using point-light knocking and drinking arm movements as stimuli, Pollick et al. (2001) found that judgements of anger and happiness were more likely when the movements were fast and jerky, and that judgements of sadness were more closely associated with slow and smooth movements. And Sawada et al. (2003) reported that arm movements made with the intention of expressing joy, sadness, or anger varied in their velocity, acceleration, and displacement, and that differences in these factors predicted the ability of observers to distinguish between the three types of emotional expression. Nonetheless, there is also evidence that form-related cues in moving bodies and faces, in addition to kinematics, contribute to emotion perception. Bassili (1978) reported greater emotion classification accuracy for full-light compared to point-light facial movements, except for happy expressions. Dittrich (1991) found equivalent emotion recognition performance for point-light face stimuli in which the dots demarcated key facial structures (e.g., eyes, mouth) and those in which the dots were positioned randomly on the face. This result contrasts with Hill et al.'s (2003) finding that sex judgements from facial

movements were more accurate with spatially normalized than pseudo-random dot placement, and thus highlights the relationship between form and motion information in specifying cues for emotion perception. In our own work, we have reported a reduction in emotion recognition performance with point-light (Dittrich et al., 1996) and patch-light (Atkinson et al., 2004) compared to full-light displays of body movements.

Building on this earlier work, we have demonstrated robust effects of stimulus inversion and motion reversal on the classification of basic emotions from patch-light and full-light movie clips of bodily expressions (Atkinson et al., 2007). Inverting the 3-second long movies significantly impaired emotion recognition accuracy, but did so more in the patch-light than in the full-light displays, indicating that inversion disrupts the processing of form cues more than it does the processing of kinematic and dynamic cues. Playing the movies backwards also significantly impaired emotion recognition accuracy, but this effect was only marginally greater for the patch-light than for the full-light displays, providing qualified support for the importance of the sequencing of changes in form to judgements of emotions from body gestures. While we cannot be certain that our stimulus manipulations completely eliminated all cues other than kinematics, even when in combination, the substantial reduction in emotion classification performance, especially for the inverted, reversed patch-light displays, attests to the importance of form cues in emotion perception; conversely, the fact that emotion classification performance was still substantially above chance, even in the inverted, reversed patch-light displays, attests to the importance of kinematics in providing cues for emotion perception. While it is likely that inversion of biological motion disrupts the processing of dynamic cues related to movement within the earth's gravitational field, if that were *all* that inversion impaired, then we should not have seen a greater effect of orientation for the patch-light compared to full-light stimuli.

The results of this study provide partial support for Dittrich et al.'s (1996) modified version of Walk and Homan's (1984) "alarm hypothesis", insofar as the identification of fearful and disgusted body movements was disproportionately impaired by inversion, suggesting a more important role for static form cues in the recognition of these emotions compared to the other emotions. On this reasoning, however, one would also expect a similar effect for anger, the identification of which was not disproportionately impaired by inversion. Consistent with the idea that an important diagnostic feature of fearful body movements is that they often involve cowering or retreating, which when reversed would appear as advancements, fear recognition was also disproportionately impaired by motion reversal.

What specific form-related cues are utilized in emotion perception from body expressions? One suggestion is that the overall shape of particular body postures, such as their angularity or roundedness, informs emotion judgements (Aronoff et al., 1992). The inversion effects that we found (Atkinson et al., 2007) highlight the importance of relational or configural cues, adding weight to previous claims that configural information plays an important role in subserving emotion perception from body expressions (Dittrich et al., 1996; Stekelenburg & de Gelder, 2004). In contrast, the effects of motion reversal tentatively suggest a possible role for spatiotemporal cues (changes in form over time) in emotion recognition. Given the conventional and sometimes symbolic (Buck, 1984) nature of our actors' movements (see Atkinson et al., 2004 for details), we speculate that configurations of static form and their changes over time are more closely associated with representations of *what* people do with their bodies than with how they move them, the latter being specified mostly by kinematics (see also Giese & Poggio, 2003).

In emphasizing the roles of configural form cues and kinematics in the recognition of emotions from body movements, we do not wish to deny the possible importance of simulation accounts of emotion recognition. As with action understanding in general (discussed in Section 2.3), purely image-processing accounts may not be sufficient for emotional expression understanding, or at any rate may not detail the only means by which we can understand others' emotional expressions. One way in which we might be able to recognize the emotional state of another is via our perception of an emotional response within ourselves (Adolphs, 2002; Atkinson & Adolphs, 2005; Gallese et al., 2004; Atkinson, 2007; Heberlein & Adolphs, 2007; Goldman & Sripada, 2005). One version of this idea is that a visual representation of another's expression leads us to experience what that person is feeling (i.e., emotional contagion), which allows us then to infer that person's emotional state. That is, the grounds for inferring the viewed person's emotional state is knowledge from the 'inside'; experiencing the emotion for oneself (even in an attenuated or unconscious form) is an important, perhaps necessary, step to accurate judgements about the other's emotion. A different but conceivably compatible idea is that coming to know what another is feeling involves simulating the viewed emotional state via the generation of a somatosensory image of the associated body state (Adolphs, 2002), or simulating the motor programs for producing the viewed expression (Carr et al., 2003; Gallese et al., 2004; Leslie et al., 2004).

4. Implications for computer vision and AI

The implementation of an emotion model into computer vision and AI has widespread consequences for future technologies and research approaches. For example, virtual reality scenarios recently have tried to include emotions in various ways. We will briefly mention the implications of the psychology of emotion expression and recognition for affective computing in such fields as AI, HCI, robotics and telecommunications. The use of emotions for future technologies will be discussed and some strengths and weaknesses of the application of emotional behaviours will be addressed.

In the previous sections we reviewed evidence for and argued that the ability to identify bodily expressions of emotion relies on specific types of both visual form and motion cues, and that the relative reliance on these different types of cue can vary across emotions. Several suggestions for affective computing can be drawn from this work, especially when viewed in the light of the Interactive Encoding Model proposed by Dittrich (1999). He argues for the integration of low-level processing of structural motion routines with more conceptually-driven semantic processing. These integrative processes are strictly resource-dependent but amenable to learning through continuous updating in working memory (see Figure 1).

A prediction of the Interactive Encoding model is that there is a strategic transition from input-driven to conceptually-driven processing as the performer develops more elaborate cognitive processes. The level of stimulus encoding therefore seems to be variable, depending on the amount and type of information available to the observer. Such an information flow characterized by multiple interactions and continuous updating is linked together through what are termed 'motion integrators'. The idea is that 'motion integrators' operating at a perceptual and cognitive level of visual processing and requiring attention are integral to skilled pattern recognition in a domain such as emotion recognition and can be directly translated to affective computing (see Dittrich, 1999). This model is, on the one

hand, brain-inspired, as recent models of artificial computing are, but on the other hand, guided by the functional process model of emotion processing. Emotions are seen as grounded in the overall interactive nature of input-output characteristics of the artificial device (where between input and output one has to infer some intervening variables, such as homeostatic variables of the system state, or emotional variables), and as playing a regulatory role, both internally and externally.

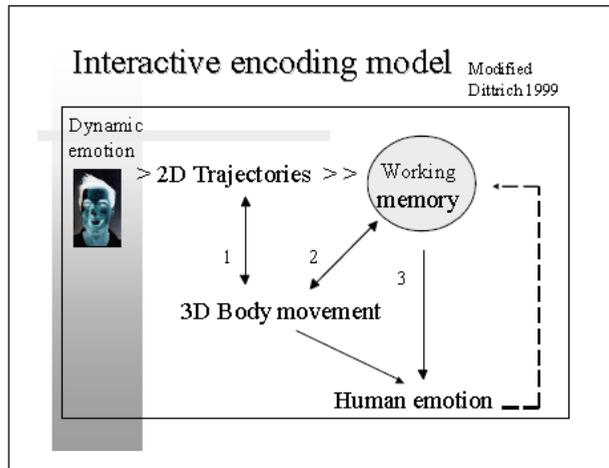


Fig. 1. Sketch of the Interactive Encoding Model for E-Motion Perception. (Note: ‘Dynamic emotions’ include body as well as facial expressions, and are usually dynamic events.)

In this sense, a developmental or ethological approach to implementing emotions in artifacts (“understanding by building”), favoured by many robot designers at present (e.g., Breazeal, 2002; Pfeifer, 2002; Brooks & Stein, 1994), might, however, not necessarily be regarded as the only or most efficient way to equip artifacts with emotional capabilities. As Brooks (1991; 2002) quite rightly pointed out, the notion of embodiment seems one possible avenue of how to achieve the necessary grounding of emotions. However, for implementing and linking biological input-output characteristics, several avenues and theoretical models from the behavioural and brains sciences become available. For example, we argue that the role of emotions is not restricted to expressive display characteristics; rather, emotions have a much wider impact on behavioural regulation in multi-layered and multi-agent artifacts. They form part of more general structures or are instrumental for a wide range of adaptive functions in an organism (Schneider & Dittrich, 1989). Analogously, the roles of emotions in artifacts may reach from a simple indication of its state (e.g., energy needs, spatial arrangements) to the highest levels of mental regulation (e.g., consciousness, imagination or turn-taking, and intentionality). Following this line of argument, we consider the role of emotions as ideally suited to fit perfectly Brooks’s (1991) demand for structures that form part of the control systems in a way that fits exactly into existing structures, and are both totally distributed and also local in their operations. In this sense, we argue that the role of emotions and emotion recognition is not only crucial but the key for any artificial device or robot to pass the ‘Total Turing Test’ as specified by Harnad (2000). Therefore, an Emotional Turing Test has to form part of any Total Turing Test. The Emotional Turing Test will satisfy

Harnad's (2000) useful criterion of some kind of embodiment for a Turing test in which there is no screen between the candidate and the interrogator. Why and how could it be?

In a useful overview article "Multiagent systems: milestones and new horizons", which summarizes the immense progress made by a second generation of multiagent architectures, Sen (1997) quite rightly sees the field at a critical juncture. Nevertheless, it is revealing that in the whole article the term 'emotion' is not mentioned once. Similarly, when Schaal (1999) asked the question "Is imitation learning the route to humanoid robots?" not once is emotion addressed. This situation has dramatically changed over the last 10 years, not the least through Picard's (1997) influential book "Affective Computing". One of the early researchers to implement social features and emotional signals in robots is Cynthia Breazeal (for overviews, see Breazeal, 2002; 2003) with KISMET. But see also the pioneering work of Takeuchi and Nagao (1992), who developed a computer interface with synthetic facial displays. In these approaches the importance of emotional displays is strongly recognized and seen as crucial for the successful development of robots. For example, KISMET's skills include the ability to direct the robot's focus to establish shared reference ('joint viewing'), the ability to display readable expressions to the human, the ability to recognize expressive movements such as praise and rejection, the ability to take turns during social learning and the ability to regulate interaction to provide sensible learning situations ('emotion-driven learning'). Often, as in the case of KISMET, the implementation of expressive movements and emotion recognition is based on (a) facial expressions, (b) gestures or body movement in the visual modality, and (c) affective speech or affective sound variations in the auditory modality.

Generally, in affective computing, emotions can mirror the two regulatory roles emotions have, namely intra-individual and inter-individual. Emotions can be used, on the one hand, in the form of emotional visual expressions for social interactions and communication and, on the other hand, in order to organise and prioritise behaviours. The first aspect is more concerned with the relationship between people whereas the second aspect refers to the internal regulation of behaviours through selection, coordination or shifting of output priorities. The latter roles seem closely linked to some of the most evolved mental activities in humans such as consciousness or intentionality and would constitute the core of any "autonomous agent" (e.g., Dean, 1998; Kozma & Fukuda, 2006). Nevertheless, one needs to acknowledge that sometimes emotions are of no direct use, as Arbib (2005) quite rightly pointed out. Queuing for the bus seems a useful activity but showing anger about the length of the queue might be of no use when it is the last bus to make the journey into town. In other words, to display emotions or not seems closely linked to the debate on "autonomous agents".

Occasionally, in addition to changes in the form of the body and in its movements, changes in body colour are associated with emotional responses in humans; for example, a reddening of the face in anger or increased pallor in fear (Drummond & Quah, 2001; Drummond, 1997). Similarly, colours could be used to generate some kind of emotional display in virtual reality scenarios and computing. However, as the association between colours and emotional interpretations is quite vague and often subjective, if not spurious, such attempts seem less fruitful in the long run. Colours might be used as a common frame of reference as with traffic lights, but emotional significance should not be assumed as self-evident; rather, it needs to be defined for each instance as, for example, a red light should be taken unambiguously as a stop signal. It does not seem fruitful to rely on peoples' feelings

or interpretations of colour as emotional signals. Various attempts to use colour as emotional signals in sports environments testify to the difficulties of relying on emotional colour information. A controversial study by Frank and Gilovich (1988), for example, reported that the wearing of black uniforms in professional sports leads to increased aggression and more penalties. Movement from a non-black to a black uniform team results in an increase in penalties. When watching sport events on a colour monitor where teams wear black and non-black uniforms, increases in perceived rule violations for the referees and the actions of the players themselves are reported by observers. A more recent study by Hill and Barton (2005), which seems not to be as methodologically problematic as the one by Frank and Gilovich (1988), found, across a range of competitive sports, a tendency for individuals or teams who wore red to have a higher probability of winning. While the underlying principles for such an advantage remain unclear, their uncovering will have the potential for greatly stimulating affective computing as well as psychology. Yet, whereas colour is without doubt a very strong stimulus, its significance for emotion processing is not as clear-cut as necessary to suggest in a straightforward way its use in affective computing.

One implication from findings on how we perceive visual facial expressions can be seen in the development of automatic systems to detect and categorize facial expressions. In order to do this the coding of facial behaviours has to be automated. A prime candidate for such attempts, Ekman and Friesen's (1978) Facial Action Coding System (FACS), has been used because of the precise description of facial movements and the standardized catalogue of corresponding emotion labels. The development of such automated systems is not without its difficulties, however, primarily because of the wide range of physiognomic differences between individuals, which affects the variability of visual display characteristics such as eyebrow or mouth movements. Kaiser and Wehrle (1992) tried to solve such problems by adopting a combined approach using FACS as an expert system teaching a connectionist network fuzzy rules to measure facial movements automatically. The combination of fuzzy rules together with an artificial neural network has proven more successful in overcoming the variability problem than traditional methods such as expert systems or connectionist systems on their own.

The use of visual signals as emotional indicators is particularly relevant for telecommunications, and highlights two opposing principles. On the one side, despite the ever-increasing storage capacities of computer chips one tenet in telecommunication is nevertheless the reduction of the information flow between sender and receiver to save processing costs and minimise errors. On the other side, emotional signals are more often than not characterized by subtle changes in form or motion. Now the dilemma for the telecommunication architect seems obvious. First, there is the question of how much information is necessary to portray the emotional state in addition to the identity. Second, there is the question of how to build a system that is able to pick up reciprocal behaviour. Such new systems would have to go far beyond the FACS model or the concept of a small set of basic emotions (see Ortony & Turner, 1990). But which psychological emotion theory should be embraced instead?

One option is the appraisal model of emotions. Quite interestingly, the majority of the computer models of emotions, if they refer expressly to psychological theories, are based on the so-called appraisal theories. The fascination with these theories most likely stems from

the fact that they can be converted into programme code in a more or less direct way. Consider, for example, Ortony, Clore and Collins's view (1988), which assumes that emotions develop as a consequence of certain cognitions and interpretations. This view concentrates exclusively on the cognitive elicitors of emotions: events, agents and objects. Their central assumption is that emotions represent value-based reactions to our perceptions of the world. One can be pleased about the consequences of an event or not; one can endorse or reject the actions of an agent or one can like or not like aspects of an object. Furthermore, events can have consequences for others or for oneself, and an acting agent can be another or oneself. The consequences of an event for another can be desirable or undesirable; the consequences for oneself can lead to relevant or irrelevant expectations. Relevant expectations for oneself finally can be differentiated again according to whether they actually occur or not. The authors further define a set of global and local intensity variables, which operate over all three emotion categories. Although no formalization of their model is explicitly provided, every emotion can be described using a formal notation, which makes their model so attractive to the computing community. A problem with the model is that not all stimuli can be directly evaluated because, quite often, there are strong pre-wired stimulus-response relationships for which explicit appraisal procedures would not apply or stimuli that are processed subliminally in emotion processing without further appraisal. Thus whereas the phenomenon of implicit emotion processing is common in humans, it seems a major unsolved challenge for full implementation in affective computing.

Another elaborate appraisal model has been suggested by Scherer (1984; Ellsworth & Scherer, 2003), in which five functionally defined subsystems at various levels of consciousness are involved with emotional processes. An information-processing subsystem evaluates the stimulus through perception, memory, forecast and evaluation of available information. A supporting subsystem adjusts the internal condition through control of neuroendocrine, somatic and autonomous states. A leading subsystem plans, prepares actions and selects between competitive motives. An acting subsystem controls motor expression and visible behaviour. Finally, a monitor subsystem controls the attention that is assigned to the present states and passes the resulting feedback on to the other subsystems. Emotion-related appraisals are implemented by the information-processing subsystem, which in Scherer's model are called stimulus evaluation checks, and which lead to changes in the other subsystems. Each emotion can be clearly characterized by a mix of the stimulus evaluation checks and subchecks.

A very interesting appraisal model has been suggested in the form of the "communicative theory of emotions" (Oatley & Johnson-Laird, 1996). This model proposes a hierarchy of simultaneously operative processing modules, which work asynchronously on different tasks. These instances are coordinated by a central control or operating system. This control system contains a model of the entire system. The functioning of the whole system depends on communication between modules. According to Oatley and Johnson-Laird (1996) there are two kinds of communication between modules: propositional or symbolic, through which actual information about the environment is conveyed, and non-propositional, whose primary role is related to the emotions. The task of the non-propositional form of communication is less to convey information and more to shift the entire system of modules

into a state of increased attention, the so-called emotion mode. This function seems not unlike the global interrupt programs on computers.

The appraisal models of emotion outlined above are just three of the most prominent of a wide range of cognitive approaches (see also e.g., Frijda, 1986, 1993; Roseman et al., 1996; Lerner & Keltner, 2000; Levine et al., 2001). This variety of models, all linked to one or another kind of appraisal process, clearly demonstrates a dilemma for the computing community. Which of the many models should be adopted as the most promising in affective computing? Again, it seems as if the hope for general answers to such a question is futile. Instead, computer scientists concentrate more on the criteria of applicability for their solutions. In this sense, questions of implementation and availability of programming tools have had priority in the past. In the future, there are challenges ahead to look beyond valence models of emotion in psychology and equip artifacts with a range of different emotional processes that work together in the end. To highlight the challenges ahead and hint at possible solutions for the most complicated area of emotional processing linked to consciousness and intentionality, we focus on a concrete example below.

At the other end of the spectrum when considering the regulatory contributions of emotional features in multi-agent systems, the concepts of meaning and intentionality have to be taken more seriously (see Dittrich, 1999). It is argued that progress is made only if designers move away from the gestalt or beyond a valence view when combining multi-feature input characteristics and fully embrace a more holistic view based on the overall input-output and intervening variable states, including emotional states. For a holistic design approach the question of emotional regulation between multiple agents and, finally, intentionality seems of utmost importance. However, implementing intentional features in artifacts is one of the greatest challenges affective computing faces, as it requires the combination of emotional and cognitive states. Intentionality seems to have two aspects to it, namely target-directed and goal-directed. Both functions fall together sometimes, for example, when a pack of wolves target and hunt a sheep herd to catch a prey. Such a rather narrative scenario has been simplified, simulated and studied by Dittrich and Lea (1994), in order to investigate directly the perception of intentional motion. A visual display was developed in which multiple objects moved randomly except one (see Figure 2). This one object always moved towards one of the other randomly moving objects. This is typical for predator/prey scenarios. Three separate factors were identified, all of which contributed to the perception of intentional motion in an object by an observer: (1) the object moved in a direct fashion; (2) the object moved faster than other objects; and (3) the goal towards which the object was moving was visible. The fact that identification of intentional motion depended on three separate factors suggests that the perception of intentionality is a rather complex concept depending on the integration of different dynamic features of multiple objects. Therefore, one of the challenges for the architecture of artificial devices is how such 'motion integrators' can be designed in multi-layered systems. The successful implementation of such e-motion features in artifacts would not only be seen as a major step in realizing the tenets of embodiment but as central to understand the workings of the animal and human brains at a psychological level. The notion of intelligence would be demystified in a major way. Only then would interaction with robots be able to constitute significant relationships in people's lives. Real-life scenarios could be entered into or re-enacted and virtual reality become a reality in life.



Fig. 2. The wolf-sheep scenario: The thin lines denote the movement trace of random moving objects. The thick line is the movement trace of the objects imbued with intention (for more detailed explanation, see text).

5. Conclusion

We suggest two recommendations for affective computing. First, affective computing researchers should heed the findings of psychology and neuroscience. We grant that machines that recognize emotional expressions or that behave emotionally need not achieve these feats in exactly the way humans and their brains do. Nonetheless, a better understanding of human perception and human emotions can aid affective computing by narrowing down the search space of possible solutions and by providing a rich set of clues for alternative solutions within that space. We have tried to give a flavour of how emotion perception, as well as emotions more generally, are implemented in humans. In summary: Emotions have roles in both communication and behavioural organization. The communication of emotions is multimodal. With respect to the perception of bodily expressed emotions, humans rely on a mixture of visual form and motion cues, particularly configural form and motion as well as kinematic cues. The relative reliance on these different types of cue can vary across emotions. Evidence reviewed here suggests the functional architecture of human emotion perception is both modular and interactive. In particular, visual emotion perception involves, *inter alia*, subsystems specialized for processing the form and motion of others' bodies, and other, modality-general subsystems that evaluate and signal the value of stimuli and events to the observer's own welfare and interests. At least in some cases, these latter processes subsequently modulate the activity of the visual mechanisms. Evidence also suggests that existence of other subsystems that simulate observed expressions or actions and/or that simulate the changes in internal body state associated with the viewed expression or action. Attempts to emulate human emotion recognition in machines will require detailed knowledge not only of how all these different subsystems operate but also of how they interact, so it is heartening to note that this is now a focus of research in cognitive neuroscience. With respect to the role of emotions in behavioural organization, research on human emotions indicates the importance of basic

level regulatory functions as well as high-order mental activities related to consciousness and intentionality, which also in turn influence perception and communication.

Our second recommendation is that, on the basis of the findings from psychology and neuroscience, affective computing researchers might want to give serious consideration to the question of whether giving a machine the capacity to experience emotions might be one way in which it could achieve reliable and efficient emotion recognition. Indeed, it might also or instead be beneficial to consider the flipside of this question, that is, whether installing an ability to recognize emotions in others might be an important or even necessary step in building a machine that experiences genuine emotions.

Affective computing could be a powerful tool to produce a new generation of artificial devices. However, we suggest that a new approach to affective computing based on the above recommendations may support the design of emotional agents as well as the implementation of emotional reactions into different types of machinery. It is possible that the features of human emotions that we have outlined above are not directly applicable to AI or computer science models of emotions. Even so, we suggest that the role of emotions in regulating the internal functioning and external behaviour of multi-layered and multi-agent systems needs to form the core of affective computing attempts to simulate the overall distribution and at the same time local display of emotional characteristics.

6. References

- Adolphs, R. (2002). Recognizing emotion from facial expressions: Psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews*, 1(1), 21-62.
- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences*, 4(7), 267-278.
- Arbib, M. A. (2005). Beware the passionate robot. In J.-M. Fellous & M. A. Arbib (Eds.), *Who needs emotions? The brain meets the robot*, Oxford University Press, New York.
- Arbib, M. A., & Fellous, J. M. (2004). Emotions: from brain to robot. *Trends in Cognitive Sciences*, 8(12), 554-561.
- Arkin, R. C., Fujita, M., Takagi, T., & Hasegawa, R. (2003). An ethological and emotional basis for human-robot interaction. *Robotics and Autonomous Systems*, 42(3-4), 191-201.
- Aronoff, J., Woike, B. A., & Hyman, L. M. (1992). Which are the stimuli in facial displays of anger and happiness? Configurational bases of emotion recognition. *Journal of Personality and Social Psychology*, 62(6), 1050-1066.
- Atkinson, A. P. (2007). Face processing and empathy. In T. F. D. Farrow & P. W. R. Woodruff (Eds.), *Empathy in mental illness* (pp. 360-385), Cambridge University Press, Cambridge.
- Atkinson, A. P., & Adolphs, A. (2005). Visual emotion perception: Mechanisms and processes. In L. F. Barrett, P. M. Niedenthal & P. Winkielman (Eds.), *Emotion and consciousness* (pp. 150-182), Guilford Press, New York.
- 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(5), 717-746.

- 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(1), 59-72.
- Barclay, C. D., Cutting, J. E., & Kozlowski, L. T. (1978). Temporal and spatial factors in gait perception that influence gender recognition. *Perception and Psychophysics*, 23(2), 145-152.
- Bassili, J. N. (1978). Facial motion in the perception of faces and of emotional expression. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 373-379.
- Beauchamp, M. S. (2005). See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex. *Current Opinion in Neurobiology*, 15(2), 145-153.
- Beintema, J. A., & Lappe, M. (2002). Perception of biological motion without local image motion. *Proceedings of the National Academy of Sciences*, 99(8), 5661-5663.
- Bertenthal, B. I., & Pinto, J. (1994). Global processing of biological motions. *Psychological Science*, 5(4), 221-225.
- Bertenthal, B. I., Proffitt, D. R., & Kramer, S. J. (1987). Perception of biomechanical motions by infants: implementation of various processing constraints. *Journal of Experimental Psychology: Human Perception & Performance*, 13(4), 577-585.
- Blake, R., & Shiffrar, M. (2007). Perception of human motion. *Annual Review of Psychology*, 58, 47-73.
- Bonda, E., Petrides, M., Ostry, D., & Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *Journal of Neuroscience*, 16(11), 3737-3744.
- Breazeal, C. (2003). Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies*, 59(1-2), 119-155.
- Breazeal, C. L. (2002). *Designing sociable robots*, MIT Press, Cambridge, MA.
- Brooks, R. (2002). *Robot: The future of flesh and machines*, Allen Lane/Penguin, London.
- Brooks, R. A. (1991). *Intelligence without reason*. Paper presented at the 12th International Joint Conference on Artificial Intelligence, Sydney, Australia.
- Brooks, R. A., & Stein, L. A. (1994). Building brains for bodies. *Autonomous Robots*, 1(1), 7-25.
- Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience*, 13, 400-404.
- Buck, R. (1984). *The communication of emotion*, Guilford Press, New York.
- Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C. R., McGuire, P. K., et al. (1997). Activation of auditory cortex during silent lipreading. *Science*, 276(5312), 593-596.
- Campbell, R., MacSweeney, M., Surguladze, S., Calvert, G., McGuire, P., Suckling, J., et al. (2001). Cortical substrates for the perception of face actions: an fMRI study of the specificity of activation for seen speech and for meaningless lower-face acts (gurning). *Cognitive Brain Research*, 12(2), 233-243.
- Campbell, R., Zihl, J., Massaro, D., Munhall, K., & Cohen, M. M. (1997). Speechreading in the akinetopsic patient, L.M. *Brain*, 120, 1793-1803.

- Carr, L., Iacoboni, M., Dubeau, M. C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences*, 100(9), 5497-5502.
- Casile, A., & Giese, M. A. (2005). Critical features for the recognition of biological motion. *Journal of Vision*, 5(4), 348-360.
- Chan, A. W., Peelen, M. V., & Downing, P. E. (2004). The effect of viewpoint on body representation in the extrastriate body area. *Neuroreport*, 15(15), 2407-2410.
- Damasio, A. R. (1994). *Descartes' error: Emotion, reason and the human brain*, Picador, London.
- de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, 7(3), 242-249.
- Dean, J. (1998). Animats and what they can tell us. *Trends in Cognitive Sciences*, 2(2), 60-67.
- Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, 115, 107-117.
- Dittrich, W. (1991). Das Erkennen von Emotionen aus Ausdrucksbewegungen des Gesichts. *Psychologische Beiträge*, 33, 366-377.
- Dittrich, W. H. (1993). Action categories and the perception of biological motion. *Perception*, 22(1), 15-22.
- Dittrich, W. H. (1999). Seeing biological motion – Is there a role for cognitive strategies? In A. Braffort, R. Gherbi, S. Gibet, J. Richardson & D. Teil (Eds.), *Gesture-based communication in human-computer interaction* (pp. 3-22), Springer, Berlin.
- Dittrich, W. H., & Lea, S. E. (1994). Visual perception of intentional motion. *Perception*, 23(3), 253-268.
- Dittrich, W. H., Troscianko, T., Lea, S., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception*, 25(6), 727-738.
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, 293(5539), 2470-2473.
- Downing, P. E., Peelen, M. V., Wiggett, A. J., & Tew, B. D. (2006). The role of the extrastriate body area in action perception. *Social Neuroscience*, 1(1), 52-62.
- Downing, P. E., Wiggett, A. J., & Peelen, M. V. (2007). Functional magnetic resonance imaging investigation of overlapping lateral occipitotemporal activations using multi-voxel pattern analysis. *Journal of Neuroscience*, 27(1), 226-233.
- Drummond, P. D. (1997). Correlates of facial flushing and pallor in anger-provoking situations. *Personality and Individual Differences*, 23(4), 575-582.
- Drummond, P. D., & Quah, S. H. (2001). The effect of expressing anger on cardiovascular reactivity and facial blood flow in Chinese and Caucasians. *Psychophysiology*, 38(2), 190-196.
- Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, 6(3-4), 169-200.
- Ekman, P., & Friesen, W. V. (1978). *The facial action coding system*, Consulting Psychologist's Press, Palo Alto, CA.
- Ekman, P., Friesen, W. V., & Ellsworth, P. (1972). *Emotion in the human face*, Pergamon Press, New York.
- Ellsworth, P. C., & Scherer, K. R. (2003). Appraisal processes in emotion. In R. J. Davidson, K. R. Scherer & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 572-595), Oxford University Press, New York.

- Frank, M. G., & Gilovich, T. (1988). The dark side of self-perception and social perception: Black uniforms and aggression in professional sports. *Journal of Personality and Social Psychology*, 54(1), 74-85.
- Frijda, N. H. (1986). *The emotions*, Cambridge University Press, Cambridge.
- Frijda, N. H. (1993). The place of appraisal in emotion. *Cognition & Emotion*, 7(3-4), 357-387.
- Fujita, M. (2001). AIBO: Toward the era of digital creatures. *The International Journal of Robotics Research*, 20(10), 781-794.
- Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, 8(9), 396-403.
- Giese, M. A., & Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. *Nature Reviews Neuroscience*, 4(3), 179-192.
- Goldman, A. I., & Sripada, C. S. (2005). Simulationist models of face-based emotion recognition. *Cognition*, 94(3), 193-213.
- Grafton, S. T., Arbib, M. A., Fadiga, L., & Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography. 2. Observation compared with imagination. *Experimental Brain Research*, 112(1), 103-111.
- Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., et al. (2000). Brain areas involved in perception of biological motion. *Journal of Cognitive Neuroscience*, 12(5), 711-720.
- Grossman, E. D., Battelli, L., & Pascual-Leone, A. (2005). Repetitive TMS over posterior STS disrupts perception of biological motion. *Vision Research*, 45(22), 2847-2853.
- Grossman, E. D., & Blake, R. (2001). Brain activity evoked by inverted and imagined biological motion. *Vision Research*, 41(10-11), 1475-1482.
- Grossman, E. D., & Blake, R. (2002). Brain areas active during visual perception of biological motion. *Neuron*, 35(6), 1167-1175.
- Grush, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *Behavioral and Brain Sciences*, 27(3), 377-396; discussion 396-442.
- Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., & Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: a neuromagnetic study. *Proceedings of the National Academy of Sciences*, 95(25), 15061-15065.
- Harnad, S. (2000). Minds, machines and Turing: The indistinguishability of indistinguishables. *Journal of Logic, Language, and Information*, 9(4), 425-445.
- Heberlein, A. S., & Adolphs, R. (2007). Neurobiology of emotion recognition: Current evidence for shared substrates. In E. Harmon-Jones & P. Winkielman (Eds.), *Social neuroscience: Integrating biological and psychological explanations of social behavior*, Guilford Press, New York.
- Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology*, 57, 243-259.
- Hill, H., Jinno, Y., & Johnston, A. (2003). Comparing solid-body with point-light animations. *Perception*, 32(5), 561-566.
- Hill, R. A., & Barton, R. A. (2005). Red enhances human performance in contests. *Nature*, 435(7040), 293-293.

- Hirai, M., & Hiraki, K. (2006). The relative importance of spatial versus temporal structure in the perception of biological motion: An event-related potential study. *Cognition*, 99(1), B15-B29.
- Hirai, M., Senju, A., Fukushima, H., & Hiraki, K. (2005). Active processing of biological motion perception: an ERP study. *Cognitive Brain Research*, 23(2-3), 387-396.
- Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, 286, 2526-2528.
- Izard, C. E. (1991). *The psychology of emotions*, Kluwer Academic Amsterdam.
- Jellema, T., Baker, C. I., Wicker, B., & Perrett, D. I. (2000). Neural representation for the perception of the intentionality of actions. *Brain and Cognition*, 44(2), 280-302.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, 14(2), 201-211.
- Johnson, S. H., & Grafton, S. T. (2003). From 'acting on' to 'acting with': the functional anatomy of object-oriented action schemata. *Progress in Brain Research*, 142, 127-139.
- Jokisch, D., Daum, I., Suchan, B., & Troje, N. F. (2005). Structural encoding and recognition of biological motion: evidence from event-related potentials and source analysis. *Behavioural Brain Research*, 157(2), 195-204.
- Kaiser, S., & Wehrle, T. (1992). Automated coding of facial behavior in human-computer interactions with FACS. *Journal of Nonverbal Behavior*, 16(2), 67-84.
- Kozma, R., & Fukuda, T. (2006). Intentional dynamic systems: Fundamental concepts and applications. *International Journal of Intelligent Systems*, 21(9), 875-879.
- Lange, J., & Lappe, M. (2006). A model of biological motion perception from configural form cues. *Journal of Neuroscience*, 26(11), 2894-2906.
- Lazarus, R. S. (1991). *Emotion and adaptation*, Oxford University Press, New York.
- Lerner, J. S., & Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgement and choice. *Cognition & Emotion*, 14(4), 473-493.
- Leslie, K. R., Johnson-Frey, S. H., & Grafton, S. T. (2004). Functional imaging of face and hand imitation: towards a motor theory of empathy. *Neuroimage*, 21(2), 601-607.
- Levine, L. J., Prohaska, V., Burgess, S. L., Rice, J. A., & Laulhere, T. M. (2001). Remembering past emotions: The role of current appraisals. *Cognition & Emotion*, 15(4), 393-417.
- Loula, F., Prasad, S., Harber, K., & Shiffrar, M. (2005). Recognizing people from their movement. *Journal of Experimental Psychology: Human Perception and Performance*, 31(1), 210-220.
- Lu, H., Yuille, A., & Liu, Z. (2005). Configural processing in biological motion detection: Human versus ideal observers. *Journal of Vision*, 5(8), 23-23.
- Mandler, G. (2007). *A history of modern experimental psychology: From James and Wundt to cognitive science*, MIT Press/ Bradford Books, Cambridge, MA.
- Mather, G., & Murdoch, L. (1994). Gender discrimination in biological motion displays based on dynamic cues. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, 258(1353), 273-279.
- Mather, G., Radford, K., & West, S. (1992). Low-level visual processing of biological motion. *Proceedings of the Royal Society of London Series B: Biological Sciences*, 249(1325), 149-155.

- Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, 6(6), 255-260.
- McLeod, P., Dittrich, W., Driver, J., Perrett, D., & Zihl, J. (1996). Preserved and impaired detection of structure from motion by a "motion-blind" patient. *Visual Cognition*, 3(4), 363-391.
- Miall, R. C. (2003). Connecting mirror neurons and forward models. *Neuroreport*, 14(17), 2135-2137.
- Michels, L., Lappe, M., & Vaina, L. M. (2005). Visual areas involved in the perception of human movement from dynamic form analysis. *Neuroreport*, 16(10), 1037-1041.
- Oatley, K. (2004). *Emotions: A brief history*, Blackwell, Oxford.
- Oatley, K., & Johnson-Laird, P. N. (1996). The communicative theory of emotions: Empirical tests, mental models, and implications for social interaction. In L. L. Martin & A. Tesser (Eds.), *Striving and feeling: Interactions among goals, affect, and self-regulation* (pp. 363-393), Erlbaum, Hillsdale, NJ.
- Oram, M. W., & Perrett, D. I. (1994). Responses of anterior superior temporal polysensory (STPa) neurons to "biological motion" stimuli. *Journal of Cognitive Neuroscience*, 6(2), 99-116.
- Oram, M. W., & Perrett, D. I. (1996). Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. *Journal of Neurophysiology*, 76(1), 109-129.
- Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions*, Cambridge University Press, Cambridge.
- Ortony, A., & Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review*, 97(3), 315-331.
- Pavlova, M., Birbaumer, N., & Sokolov, A. (2006). Attentional modulation of cortical neuromagnetic gamma response to biological movement. *Cerebral Cortex*, 16(3), 321-327.
- Pavlova, M., Lutzenberger, W., Sokolov, A., & Birbaumer, N. (2004). Dissociable cortical processing of recognizable and non-recognizable biological movement: Analysing gamma MEG activity. *Cerebral Cortex*, 14(2), 181-188.
- Pavlova, M., Lutzenberger, W., Sokolov, A. N., Birbaumer, N., & Krageloh-Mann, I. (2007). Oscillatory MEG response to human locomotion is modulated by periventricular lesions. *Neuroimage*, 35(3), 1256-1263.
- Pavlova, M., & Sokolov, A. (2000). Orientation specificity in biological motion perception. *Perception and Psychophysics*, 62(5), 889-899.
- Pavlova, M., & Sokolov, A. (2003). Prior knowledge about display inversion in biological motion perception. *Perception*, 32(8), 937-946.
- Peelen, M. V., & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *Journal of Neurophysiology*, 93(1), 603-608.
- Peelen, M. V., & Downing, P. E. (2007). The neural basis of visual body perception. *Nature Reviews Neuroscience*, 8(8), 636-648.
- Peelen, M. V., Wiggett, A. J., & Downing, P. E. (2006). Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion. *Neuron*, 49(6), 815-822.

- Pelphrey, K. A., Mitchell, T. V., McKeown, M. J., Goldstein, J., Allison, T., & McCarthy, G. (2003). Brain activity evoked by the perception of human walking: Controlling for meaningful coherent motion. *Journal of Neuroscience*, 23(17), 6819-6825.
- Pelphrey, K. A., Morris, J. P., Michelich, C. R., Allison, T., & McCarthy, G. (2005). Functional anatomy of biological motion perception in posterior temporal cortex: An fMRI study of eye, mouth and hand movements. *Cerebral Cortex*, 15(12), 1866-1876.
- Pelphrey, K. A., Viola, R. J., & McCarthy, G. (2004). When strangers pass: Processing of mutual and averted social gaze in the superior temporal sulcus. *Psychological Science*, 15(9), 598-603.
- Perrett, D. I., Smith, P. A., Mistlin, A. J., Chitty, A. J., Head, A. S., Potter, D. D., et al. (1985). Visual analysis of body movements by neurones in the temporal cortex of the macaque monkey: a preliminary report. *Behavioural Brain Research*, 16(2-3), 153-170.
- Peuskens, H., Vanrie, J., Verfaillie, K., & Orban, G. A. (2005). Specificity of regions processing biological motion. *European Journal of Neuroscience*, 21(10), 2864-2875.
- Pfeifer, R. (2002). Robots as cognitive tools. *International Journal of Cognition and Technology*, 1, 125-143.
- Picard, R. W. (1997). *Affective computing*, MIT Press, Cambridge, MA.
- Pinto, J., & Shiffrar, M. (1999). Subconfigurations of the human form in the perception of biological motion displays. *Acta Psychologica*, 102(2-3), 293-318.
- Plutchik, R. (1980). *Emotion: A psychoevolutionary synthesis*, Harper and Row, New York.
- Pobric, G., & Hamilton, A. F. (2006). Action understanding requires the left inferior frontal cortex. *Current Biology*, 16(5), 524-529.
- Pollick, F. E., Paterson, H. M., Bruderlin, A., & Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition*, 82(2), B51-61.
- Pourtois, G., Peelen, M. V., Spinelli, L., Seeck, M., & Vuilleumier, P. (2007). Direct intracranial recording of body-selective responses in human extrastriate visual cortex. *Neuropsychologia*, 45(11), 2621-2625.
- Prinz, J. J. (2004). *Gut reactions: A perceptual theory of emotion*, Oxford University Press, New York.
- Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience*, 18(6), 2188-2199.
- Puce, A., & Perrett, D. (2003). Electrophysiology and brain imaging of biological motion. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, 358(1431), 435-445.
- Puce, A., Syngeniotis, A., Thompson, J. C., Abbott, D. F., Wheaton, K. J., & Castiello, U. (2003). The human temporal lobe integrates facial form and motion: evidence from fMRI and ERP studies. *Neuroimage*, 19(3), 861-869.
- Reed, C. L., Stone, V. E., Bozova, S., & Tanaka, J. (2003). The body-inversion effect. *Psychological Science*, 14(4), 302-308.
- Reed, C. L., Stone, V. E., Grubb, J. D., & McGoldrick, J. E. (2006). Turning configural processing upside down: Part and whole body postures. *Journal of Experimental Psychology: Human Perception and Performance*, 32(1), 73-87.

- Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition*, 47(1), 25-57.
- Richardson, M. J., & Johnston, L. (2005). Person recognition from dynamic events: The kinematic specification of individual identity in walking style. *Journal of Nonverbal Behavior*, 29(1), 25-44.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169-192.
- Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3, 131-141.
- Roseman, I. J., Antoniou, A. A., & Jose, P. E. (1996). Appraisal determinants of emotions: Constructing a more accurate and comprehensive theory. *Cognition & Emotion*, 10(3), 241-277.
- Runeson, S., & Frykholm, G. (1981). Visual perception of lifted weight. *Journal of Experimental Psychology: Human Perception and Performance*, 7(4), 733-740.
- Sawada, M., Suda, K., & Ishii, M. (2003). Expression of emotions in dance: relation between arm movement characteristics and emotion. *Perceptual and Motor Skills*, 97(3 Pt 1), 697-708.
- Saygin, A. P. (2007). Superior temporal and premotor brain areas necessary for biological motion perception. *Brain*, 130(9), 2452-2461.
- Saygin, A. P., Wilson, S. M., Hagler, D. J., Bates, E., & Sereno, M. I. (2004). Point-light biological motion perception activates human premotor cortex. *Journal of Neuroscience*, 24(27), 6181-6188.
- Schaal, S. (1999). Is imitation learning the route to humanoid robots? *Trends in Cognitive Sciences*, 3(6), 233-242.
- Schachter, S., & Singer, J. E. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, 69(5), 379-399.
- Scherer, K. R. (1984). On the nature and function of emotion: A component process approach. In K. R. Scherer & P. Ekman (Eds.), *Approaches to emotion*, Lawrence Erlbaum, Hillsdale, NJ.
- Schneider, K., & Dittrich, W. (1989). Functions and evolution of emotions (Germ.). In K. Scherer (Ed.) *Enzyklopaedie der Psychologie*, Bd. C/IV/3 (pp.41-115). Goettingen: Hogrefe.
- Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, 4(8), 299-309.
- Schwarzlose, R. F., Baker, C. I., & Kanwisher, N. (2005). Separate face and body selectivity on the fusiform gyrus. *Journal of Neuroscience*, 25(47), 11055-11059.
- Sen, S. (1997). Multiagent systems: milestones and new horizons. *Trends in Cognitive Sciences*, 1(9), 334-340.
- Shipley, T. F. (2003). The effect of object and event orientation on perception of biological motion. *Psychological Science*, 14(4), 377-380.
- Stekelenburg, J. J., & de Gelder, B. (2004). The neural correlates of perceiving human bodies: an ERP study on the body-inversion effect. *Neuroreport*, 15(5), 777-780.

- Strafella, A. P., & Paus, T. (2000). Modulation of cortical excitability during action observation: a transcranial magnetic stimulation study. *Neuroreport*, 11(10), 2289-2292.
- Suzuki, K., Camurri, A., Ferrentino, P., & Hashimoto, S. (1998). *Intelligent agent system for human-robot interaction through artificial emotion*. Paper presented at the IEEE International Conference on Systems, Man, and Cybernetics, San Diego, CA, USA.
- Takeuchi, A., & Nagao, K. (1992). *Communicative facial displays as a new conversational modality*. Tokyo: Sony Computer Science Laboratory.
- Taylor, J. C., Wiggett, A. J., & Downing, P. E. (2007). fMRI analysis of body and body part representations in the extrastriate and fusiform body areas. *Journal of Neurophysiology*, 98 (3), 1626-1633.
- Thompson, J. C., Clarke, M., Stewart, T., & Puce, A. (2005). Configural processing of biological motion in human superior temporal sulcus. *Journal of Neuroscience*, 25(39), 9059-9066.
- Thompson, J. C., Hardee, J. E., Panayiotou, A., Crewther, D., & Puce, A. (2007). Common and distinct brain activation to viewing dynamic sequences of face and hand movements. *Neuroimage*, 37(3), 966-973.
- Troje, N. F. (2002). Decomposing biological motion: a framework for analysis and synthesis of human gait patterns. *Journal of Vision*, 2(5), 371-387.
- Troje, N. F. (2003). Reference frames for orientation anisotropies in face recognition and biological-motion perception. *Perception*, 32(2), 201-210.
- Urgesi, C., Berlucchi, G., & Aglioti, S. M. (2004). Magnetic stimulation of extrastriate body area impairs visual processing of nonfacial body parts. *Current Biology*, 14(23), 2130-2134.
- Urgesi, C., Calvo-Merino, B., Haggard, P., & Aglioti, S. M. (2007a). Transcranial magnetic stimulation reveals two cortical pathways for visual body processing. *Journal of Neuroscience*, 27(30), 8023-8030.
- Urgesi, C., Candidi, M., Ionta, S., & Aglioti, S. M. (2007b). Representation of body identity and body actions in extrastriate body area and ventral premotor cortex. *Nature Neuroscience*, 10(1), 30-31.
- Vaina, L. M., Cowey, A., LeMay, M., Bienfang, D. C., & Kikinis, R. (2002). Visual deficits in a patient with 'kaleidoscopic disintegration of the visual world'. *European Journal of Neurology*, 9(5), 463-477.
- Vaina, L. M., Lemay, M., Bienfang, D. C., Choi, A. Y., & Nakayama, K. (1990). Intact "biological motion" and "structure from motion" perception in a patient with impaired motion mechanisms: A case study. *Visual Neuroscience*, 5(4), 353-369.
- Vaina, L. M., Solomon, J., Chowdhury, S., Sinha, P., & Belliveau, J. W. (2001). Functional neuroanatomy of biological motion perception in humans. *Proceedings of the National Academy of Sciences*, 98(20), 11656-11661.
- Walk, R. D., & Homan, C. P. (1984). Emotion and dance in dynamic light displays. *Bulletin of the Psychonomic Society*, 22, 437-440.
- Westhoff, C., & Troje, N. F. (2007). Kinematic cues for person identification from biological motion. *Perception & Psychophysics*, 69, 241-253.

- Wheaton, K. J., Thompson, J. C., Syngeniotis, A., Abbott, D. F., & Puce, A. (2004). Viewing the motion of human body parts activates different regions of premotor, temporal, and parietal cortex. *Neuroimage*, 22(1), 277-288.
- Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London, B: Biological Sciences*, 358(1431), 593-602.



Affective Computing

Edited by Jimmy Or

ISBN 978-3-902613-23-3

Hard cover, 284 pages

Publisher I-Tech Education and Publishing

Published online 01, May, 2008

Published in print edition May, 2008

This book provides an overview of state of the art research in Affective Computing. It presents new ideas, original results and practical experiences in this increasingly important research field. The book consists of 23 chapters categorized into four sections. Since one of the most important means of human communication is facial expression, the first section of this book (Chapters 1 to 7) presents a research on synthesis and recognition of facial expressions. Given that we not only use the face but also body movements to express ourselves, in the second section (Chapters 8 to 11) we present a research on perception and generation of emotional expressions by using full-body motions. The third section of the book (Chapters 12 to 16) presents computational models on emotion, as well as findings from neuroscience research. In the last section of the book (Chapters 17 to 22) we present applications related to affective computing.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Winand H. Dittrich and Anthony P. Atkinson (2008). The Perception of Bodily Expressions of Emotion and the Implications for Computing, Affective Computing, Jimmy Or (Ed.), ISBN: 978-3-902613-23-3, InTech, Available from:

http://www.intechopen.com/books/affective_computing/the_perception_of_bodily_expressions_of_emotion_and_the_implications_for_computing

INTECH

open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2008 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.