

# Body Movements for Affective Expression: A Survey of Automatic Recognition and Generation

Michelle Karg<sup>1</sup>, Ali-Akbar Samadani<sup>1</sup>, Rob Gorbet<sup>2</sup>, Kolja Kühnlenz<sup>3</sup>, Jesse Hoey<sup>4</sup>, Dana Kulić<sup>1</sup>

**Abstract**—Body movements communicate affective expressions and, in recent years, computational models have been developed to recognize affective expressions from body movements or to generate movements for virtual agents or robots which convey affective expressions. This survey summarizes the state of the art on automatic recognition and generation of such movements. For both automatic recognition and generation, important aspects such as the movements analyzed, the affective state representation used, and the use of notation systems is discussed. The survey concludes with an outline of open problems and directions for future work.

**Index Terms**—Movement Analysis, Recognition of Affective Expressions, Generation of Affective Expressions.

## 1 INTRODUCTION

AFFECTIVE computing aims to enhance human-computer interaction (HCI) and human-robot interaction (HRI) through affective communication to create a more intuitive, engaging, and entertaining interaction. During the interaction, affective states can be expressed and recognized through facial expressions, speech, body movements, and physiological parameters. Automatic recognition of human affective expressions and generation of expressive behavior for virtual avatars and robots are key challenges in this research area. Several surveys address detection of affective states in general [1], [2], [3], [4], [5], through facial or/and audio expressions [6], [7] and generation of affective expressions [4], [8]. As facial expressions and speech dominate during face-to-face interaction, these are the modalities that have been predominantly studied in communication of nonverbal behavior, psychology, and computer science to date [9], [10], [11], [2]. Yet, there exists evidence from communication of nonverbal behavior and psychology research that body movements also convey affective expressions, e.g.,

[12], [13], [14], [15]. Considering body movement as a modality for affective computing is particularly suitable in situations where the affective state is estimated from a distance [16], to retrieve expressions which are less susceptible to social editing [17], and to communicate affective states which are easier conveyed through movement [18].

A recent survey [19] reviews the literature on affect recognition from body posture and movement, and discusses the main challenges in affect recognition from body posture and movement, including inter-individual differences, impact of culture and multi-modal recognition, and the challenges in collecting appropriate datasets and ground truth labeling. Computational models have been developed for both *automatic recognition* and *generation* of affect-expressive movements<sup>1</sup>. A large body of work has emerged in recent years developing these computational models; this survey is intended to synthesize the findings of these studies, and to identify key contributions and open research questions. Two significant characteristics of these computational models are (1) the representation of movements in physical space and time, and (2) the representation of affect.

To provide a comprehensive overview of affect-expressive movements studied to date in HCI/HRI, we introduce a suitable movement categorization and summarize works studying similar movements. We discuss the use of movement notation systems in automatic recognition and generation of affect-expressive movements. Movement notation systems, commonly used in the dance community, can provide a systematic approach for the choice of movement descriptors and facilitate knowledge transfer between communication, psychology, and computer science. This elaboration on the

<sup>1</sup>Department of Electrical and Computer Engineering, University of Waterloo, Canada.

<sup>2</sup>Centre for Knowledge Integration, University of Waterloo, Canada.

<sup>3</sup>Electrical Engineering and Information Technology, Technische Universität München, Germany.

<sup>4</sup>David R. Cheriton School of Computer Science, University of Waterloo, Canada.

The authors acknowledge the support of the Toronto Rehabilitation Institute, which receives funding under the Provincial Rehabilitation Research Program from the Ministry of Health and Long-Term Care in Ontario. The views expressed do not necessarily reflect those of the Ministry. This work was also supported by the Natural Sciences and Engineering Research Council of Canada.

The authors wish to thank Prof. Gentiane Venture and Dr. Ritta Baddoura for their additional suggestions on the emotion literature. The authors also wish to thank the anonymous reviewers for the thorough and insightful review comments. Their suggestions significantly improved the quality and readability of the paper.

1. In this work, we introduce the term *affect-expressive movement* to mean that subset of expressive movements whose purpose is to convey affect.

movements studied to date and the use of movement notation systems for both automatic recognition and generation provides complementary information to the previous survey discussing the importance of postural and dynamic features for automatic recognition [19].

A categorical or dimensional approach can be used for representing affective states. For both automatic recognition and generation studies, we analyze the set of considered affective states and their representation. We report on the common results regarding the expressiveness of affective states and whether these results can be linked to psychological studies. This detailed analysis facilitates the selection of a suitable representation for affective states [19].

The outline of the article is as follows: Section 2 briefly summarizes relevant background information on the definition, expression, and representation of affective states in HCI/HRI. Section 3 introduces a movement categorization for summarizing the state of the art in recognition and generation of affect-expressive movements, and discusses movement notation systems and their utility for affect-expressive movement analysis. Selected studies on human perception of affect-expressive movements are discussed in section 4. Studies on automatic recognition of affect-expressive movements are summarized in section 5. Section 6 provides an overview of generative models to synthesize affect-expressive movements for virtual agents or robots. Current achievements are summarized in section 7 and open questions are discussed. Concluding remarks are drawn in section 8.

## 2 AFFECT: DEFINITION, EXPRESSION, AND REPRESENTATION IN HCI/HRI

*Affective* phenomena (broadly termed affect [20]) include emotions, feelings, moods, attitudes, temperament, affective dispositions, and interpersonal stances [1], [21], [22], [23]. A categorization of these terms based on event focus, intrinsic and transactional appraisal, synchronization, rapidity of change, behavioral impact, intensity, and duration is provided in [22], [24], [23], e.g. emotions change rapidly, are short-term and intense, whereas *mood* covers a longer time-span and changes more slowly. Among the affective phenomena, emotions are those most widely studied in HCI and HRI [21, p.4]. Scherer defines *emotion* as “an episode of interrelated, synchronized changes in the states of all or most of the five organismic subsystems in response to the evaluation of an external or internal stimulus event as relevant to major concerns of the organism” [22]. The five subsystems are: cognitive (for evaluation), physiological (for regulation), motivational (for preparation of an action), subjective feeling, and motor *expression* (for communication) [22]. Theories of emotion expression are often based on facial expressions; fewer studies have been conducted that investigate the extent to which existing theories predict and explain bodily expressions of emotion [25], [26], [27], [28]. Effective emotional communication requires

both the ability to send or encode one’s own emotion in an appropriate and comprehensible manner and the ability to receive or decode the emotions of others [21]. For virtual agents and robots, these two skills refer to automatic generation and recognition of emotional expressions, respectively. These two modules can be integrated into higher-level computational models which cover the generation of appropriate emotions, and relations to cognitive and motivational subsystems [29].

The terms affect and emotion have been used inconsistently, often interchangeably, in the field of HCI/HRI, and both terms can be found in studies on automatic recognition and generation of affect-expressive movements. Here, we are concerned with body movements that can convey both short-term emotions and long-term moods, e.g., depression influences the kinematics of walking [30]. When covering both long-term and short-term affective phenomena, we use the broader term *affective states* in this survey. Within this survey, we use the term emotion when studies explicitly address emotions.

Automatic recognition of affective states is based on observing expressions. Humans can control their expressions to a certain extent, e.g., using display rules to achieve a social goal [11, p.72]. This gives rise to the possibility of a difference between the internal experience and external expression. This possible discrepancy is most widely studied for the communication of emotions. The communication of emotions can be either spontaneous or strategic [31], [11], [14]. Spontaneous communication is involuntary and the content is non-propositional<sup>2</sup>, whereas strategic or symbolic communication is goal-oriented and the content is propositional [31], [11]. Considering automatic recognition, it is important to note that an observed expression may not necessarily be the observable manifestation of an internal state, but rather, displayed to achieve a social goal. Within the scope of this survey, which focuses on HCI and HRI studies, this difference is not further explored as most studies on automatic recognition and generation assume a correspondence between the expressed and internally felt state.

Evidence from psychology indicates that affective states are expressed in body movements [33], [34], [35], [36], [15]. During daily human-human interaction, humans pay attention to expressive body and activity cues almost as often as to expressive facial, indirect verbal, and context cues [33]. A recent study highlights the utility of bodily cues for the discrimination between intense positive and negative emotions [18]. Gelder *et al.* suggest that bodily cues are particularly suitable for communication over larger distances, whereas facial expressions are more suitable for a fine-grained analysis of affective expressions [35]. Studies comparing bodily expression versus facial expressions are summarized in [19], and

2. One can apply an analysis “true or false?” to a proposition. [31, p. 7 referring to [32]]

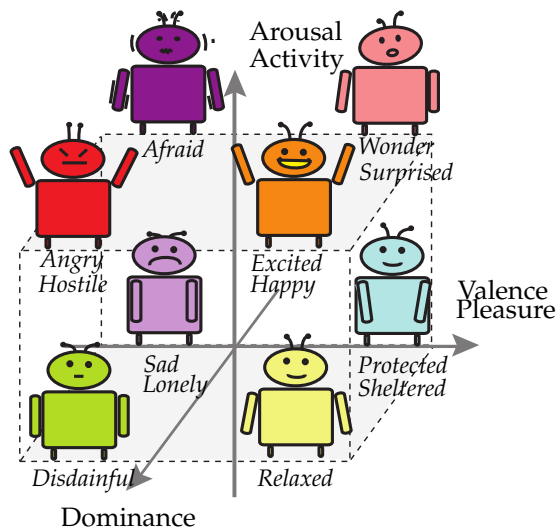


Fig. 1. The PAD model

Kleinsmith *et al.* conclude that bodily expressions are an important modality for non-verbal communication in HCI/HRI. Even though the expression and recognition of affective states from gestures and body motion is relatively unexplored in comparison to studies on facial expression and physiology [37], these modalities are advantageous 1) for perception from a distance, because bodily expressions are more easily visible from a distance than subtle changes in the face [16], [35], [38], 2) to analyze types of expressions which are less susceptible to social editing, because people are often less aware of their bodily than their facial expressions [39], [17], and 3) to convey affective states which are more easily expressed through movement, e.g., intense positive or negative emotions [18].

Affective states can be represented using a set of distinct categorical labels, or a dimensional model. *Categorical labels* describe affective states based on their linguistic use in daily life. Different sets of categorical labels can be chosen depending on the study. Most frequently, happiness, sadness, fear, and anger are included, a subset of the basic emotions [40]. A *basic emotion* is defined by a set of neural correlates in the brain, a specific set of bodily expressions, and a motivational component for action tendencies [40]. A popular set of basic emotions contains anger, happiness, sadness, surprise, disgust, and fear [41], [42].

An affective state is represented in a *dimensional model* as a point on a continuum spanned by a set of independent dimensions. A popular example is a circumplex model, where similar affective states are arranged to lie adjacent to each other on a circle, and dissimilar affective states are arranged to lie opposite each other [43]. A common model applied in affective computing is the PAD-model, with the dimensions *pleasure* (or *valence*), *arousal*, and *dominance* [44]. Arousal corresponds to the level of activation, mental alertness, and physical activity. Dominance represents the amount of control over

others and the surroundings versus feeling controlled by external circumstances. Categorical labels can be mapped to the continuous PAD space, e.g., happiness, amusement, and contentment are related to high pleasure, whereas anger, fear, and sadness are related to low pleasure [45], [46], [47], [48], as qualitatively illustrated in Fig. 1<sup>3</sup>. A dimensional representation may relate more to the underlying physiological changes [26] and Barrett suggests that categorical labels are a result of subjective categorization of emotions using conceptual knowledge, e.g., similar to color perception [49].

Most studies on affect recognition to date use categorical labels, while fewer studies have applied a dimensional representation [2]. Challenges for dimensional affect recognition include unbalanced datasets, differences in the inter-observer agreement on the dimensions, and handling of categories which are not covered by the dimensions or overlap with other categories in a dimensional approach [2]. Recent works in HCI/HRI have investigated the use of a dimensional representation and further exploration of the utility of dimensional representations in HCI/HRI is advised [7], [4].

### 3 AFFECT-EXPRESSIVE MOVEMENTS

Affective states can be expressed through body movements in various ways, e.g., by whole-body gestures, arm gestures, or modulation of functional movements. To refer to this set of movements and given the discussion in section 2, we introduce the term *affect-expressive movements*, because expressions and not internal experiences are analyzed and among these only expressions that are related to affect. In this section, we introduce a categorization of movements for summarizing the current state of the art in automatic recognition and generation of affect-expressive movements. Next, a representative review of movement notation systems is presented. We focus on movement notation systems that have been applied to study affective expressiveness and discuss a set of criteria to evaluate the suitability of the notation systems for computational analysis of affect-expressive movements.

#### 3.1 Categorization of Affect-Expressive Movements

The current state of the art on automatic generation and recognition of affect-expressive movements analyzes a large variety of body movements. An affective state can be expressed via modulation of a single movement or by selecting from a library of different movements. For instance, anger can be communicated via modulation of a movement, e.g., increasing the walking speed, or through selecting a specific movement type, e.g., making a fist. Studies discussed in this survey analyze affective expressiveness either through modulation of a single movement, use of different movement types, or a combination of both. Considering the definition of a gesture

<sup>3</sup> Categorical terms for each octant in the PAD space in Fig. 1 are selected based on pleasure, arousal, and dominance ratings in [48].

as “a movement of part of the body, especially a hand or the head, to express an idea or meaning” [50], gestures generally combine type selection with movement modulation. We consider that postures are snapshots of movements, and not separate entities.

We introduce the following categorization for the large variety of movements studied to date:

- *Communicative movements* cover a broad range of movements, which are performed in daily life and may convey affective states, e.g., gestures [51], [52], [38], [53]. A communicative movement can select from a library of movement types to express an affective state and may be accompanied by a modulation level.
- The objective of *functional movements* is to perform a task unrelated to the expression of affect, e.g., walking or knocking [54], [55], [56]. In contrast to the former category, affective states can only be expressed through modulation of a functional movement and expressiveness is secondary to function.
- *Artistic movements*, such as choreographed [57] and non-choreographed dancing [58], can display exaggerated expressions and can consist of movement types which do not occur during daily life. Artistic movements vary in terms of movement type to express an affective state and may be accompanied by a modulation level. For choreographed artistic movements, different affective states are expressed only by modulation, because the movement type is specified by the choreography, e.g., choreographed dancing [57].
- *Abstract movements* are used neither to explicitly accomplish a task nor to communicate a meaning (e.g., lifting the arms [59]). Expressivity of a movement can be analyzed independently of a possible symbolic meaning of the movement itself and movement types can be selected which do not necessarily occur during daily life.

Affective states can also be conveyed by a lack of movement (e.g., freezing in terror). However, this type of response is rarely addressed in HCI/HRI studies to date, and is therefore omitted from the rest of the paper.

### 3.2 Movement Notation Systems

In computational human movement analysis, movements are commonly represented in terms of joint angle trajectories (e.g., [60]), or derived discrete or time-series features (e.g., maximum velocity [59]). Movement analysis based on these kinematic features is usually computationally expensive due to the high-dimensional nature of the movement trajectories and more importantly, does not necessarily capture expressive qualities which are critical for affect-expressive movement analysis. Movement notation systems provide an efficient tool for systematic and compact representation of movements that capture both structural characteristics and

expressive qualities of the movements. As FACS<sup>4</sup> pushed forward the research on computational analysis of facial expressions, movement notation systems could also advance the computational analysis of affect-expressive movements by providing an objective and systematic movement representation.

A good movement notation system achieves the following criteria proposed by Guest [62]: universality (capability of coding all forms of movement), comprehensiveness (covering every aspect of a movement), movement analysis (anatomically and physiologically sound movement coding), versatility in movement description (truthful representation of intention and expressivity in addition to structure), flexibility in application, logicity (consistent presentation of similar actions using logically related symbols and codes), visuality (readable visual presentation), legibility (distinctive and discrete coding symbols or categories), and practicability (ease of use and integrability with modern technologies). Expanding on the criteria proposed by Guest, in order to limit coders’ bias, the notation systems should not need any contextual assumption or perceptual inference for notating an observed movement. A comparison between notation systems can be found in [63], [64]. However, a more recent categorization by Burgoon *et al.* enables a more systematic comparison between the notation systems and their utilities [14]. Burgoon *et al.* divide movement notation systems into *functional*<sup>5</sup> and *structural* approaches [14]. We focus on notation systems that are primarily designed for coding bodily movements and do not discuss notation systems that use bodily movements only as a peripheral indicators of affective expressions (e.g., SPAFF [65]).

*Functional approaches* describe the communicative function of a displayed movement using verbal labels. The Ekman and Friesen formulation of kinesic behaviours into five categories (emblems, illustrators, affective displays, regulators, and manipulators) is an example of a functional notation system [9].

*Structural approaches* are primarily concerned with the question of what bodily movements look like and provide detailed notation of posture and movement dynamics [14]. As a result, they provide sufficient structural and expressive details for movement replication, and are more appropriate for computational affect-expressive movement analysis.

Inspired by linguistic notation systems, Birdwhistell proposed a structural movement notation system that parallels phonemic transcription in linguistics [66]. Birdwhistell referred to non-verbal communicative body movements as *kinesics* [66], and introduced *kine* (the smallest perceivable body motion e.g., raising eye

4. Face action coding system (FACS) provides a comprehensive set of action units that can be used to objectively describe any type of facial movement [61].

5. In Section 3.1, functional movements are discussed, which should not be confused with functional movement notation systems discussed here.

brows), *kineme* (a group of movements with a same social meaning e.g., one nod, two nods, three nods), and *kinemorphs* (a combination of kinemes forming a gesture) followed by *kinemorphic* classes and complex kinemorphic constructs, which are analogous to sentences and paragraphs in linguistics. Birdwhistell used motion qualifiers and action modifiers that define: 1) the degree of muscular tension involved in executing a movement, 2) duration of the movement, and 3) the range of the movement. The *kinemograph* is introduced as a tool for notating individual kines and their direction at different body sections. The Birdwhistell system is capable of micro analysis of body movements as its kines capture barely perceivable body motion ranging from 1/50 seconds to 3 seconds in duration [66]. Birdwhistell emphasizes the importance of context for inferring the meaning of an observed movement. A Birdwhistell-inspired annotation was used to extract semantic areas in emoticons<sup>6</sup> for automatic recognition of their expressed emotions [67].

The Laban system is another prominent example of a structural movement notation system, which was developed for writing and analyzing both the structure and expressivity of movements in dance choreography [68], [69]. The Laban notation system has four major components: Body, Space, Effort, and Shape. Body indicates the active body parts, and the sequence of their involvement in a movement. Space defines where in space a movement is happening, and the directions of the body and body parts. Laban Effort and Shape components provide a comprehensive set of descriptors for a qualitative characterization of a movement [70]. Effort describes the inner attitude toward the use of energy along four bipolar components: Space, Weight, Time, and Flow, with their extremes being Indirect/Direct, Light/Strong, Sustained/Sudden, and Free/Bound, respectively. Shape consists of Shape Flow, Directional, and Shaping/Carving, all of which describe dynamic changes in the movement form [70]. *Labanotation* is the Laban notation tool that makes use of a set of abstract geometric symbols to notate an observed movement. Computational Laban analysis has been carried out for movement recognition (e.g., [71]), and generation (e.g., [72]), and to relate Laban components to low-level movement features e.g., velocity and acceleration [73], [74], [75] and different affective expressions [76].

Delsarte [77] classifies emotion as a form of expression in gestures and divides the body into zones within which mental, moral, and vital components are defined. He identifies nine laws that contribute to the meaning of a movement: altitude, force, motion (expansion, contraction), sequence, direction, form, velocity, reaction, and extension. The Delsarte system has been used for automatic generation of affect-expressive full-body [78] and hand and arm [79] movements. In [79], participants'

perception of a set of Delsarte-generated hand and arm movements displayed on an animated agent was shown to be consistent with the Delsarte model prediction.

Recently, Dael *et al.* proposed BAP (body action and posture), a structural notation system for a systematic description of temporal and spatial characteristics of bodily expression of emotions that, analogous to FACS, introduces 141 behavioural categories for coding action, posture, and function of an observed body movement [25]. BAP segments body movements into localized units in time and describes them at three levels: anatomical (articulation of different body parts), form (direction and orientation of the movements), and functional (behavioural classes categorized in kinesics emblems, illustrators, manipulators). BAP anatomical and form variables are Boolean (0 for absence and 1 for presence), while functional variables are ordinal (1 for very subtle and 5 for very pronounced). BAP was developed using the GEMEP corpus of emotion portrayals. Since the movements are captured from the knees upwards in GEMEP, the current version of BAP does not code whole body postures and leg movements. BAP also does not code dynamic movement characteristics such as velocity, acceleration, and energy. BAP reliability has been demonstrated by assessing intercoder agreement (two coders) on occurrence, temporal precision, and segmentation of posture and action units [25]. To the best of our knowledge, there is a single report on the application of BAP for computational analysis of affect-expressive movements, in which BAP behavioural categories are employed for recognition of 12 affective states encoded in 120 movements demonstrated by 10 actors [37]. Recently, AutoBAP has been proposed for automatic annotation of posture and action units based on BAP anatomical and form (and not functional) coding guidelines [80].

For automated affect-expressive movement recognition and generation, there is a need for consistent and quantitative description of movements, leading to a preference for structural notation systems that provide a fixed number of distinct movement descriptors (legibility criterion). These descriptors should be mappable to measurable movement characteristics e.g., joint position, velocity, or they should be readily quantifiable. Furthermore, such a mapping (or quantification) should be reversible for generation studies in which low-level movement trajectories are needed to produce a desired affect-expressive movement. The reversible map enables modifying low-level motion trajectories for a given set of specified quantitative descriptors characterizing a desired affective expression [72].

Despite their proven suitability for movement coding, except for BAP, the structural notation systems do not explicitly provide quantitative measures, which is perhaps the main barrier to their application in computational movement analysis. In addition, the extensive attention to microanalysis (e.g., Birdwhistell system [66]), and the need for special training (e.g., Laban system) ham-

6. An emoticon is a string of symbols used in text communication to express users' emotions [67].

per their adoption in affective computing. Furthermore, some notation systems require the coder to infer the meaning or function of an observed movement e.g., Del-sarte [77]. However, the correspondence between movements and affective expressions is not transcultural and transcontextual and there may be idiosyncratic, gender-specific, or age-specific differences in affect-expressive movements [81]. Such movement/affective expression discrepancies result in a drawback for the notation systems that code the meaning or function of an observed movement. In addition, the amount and intensity of an affect-expressive movement is important for computational analysis; hence, the preference for structural notation systems that code such information. A table summarizing the characteristics of the prominent structural notation systems is provided in the supplementary material.

## 4 HUMAN PERCEPTION OF AFFECT-EXPRESSIVE MOVEMENTS

Various psychological studies indicate that humans are not only capable of recognizing the intended action [82], but also gender [83], identity [84], and affective state [82], [13], [85], [28] from body movements. In general, the perception of affective state is multimodal, e.g., sound and body movements both influence the affective interpretation of music [86].

Many researchers have tested the human perception of affective state through dance and body movements, using a variety of stimuli including both full-light (FL) and point-light (PL) videos [87], [88], [85], [89], and reported above chance recognition rates. Research on the capacity of humans to recognize affective states helps to develop a better understanding of human perception of affect-expressive movements, and can offer insight into whether movements of full-body or isolated body parts are capable of communicating affective expressions. Furthermore, perceptual studies can inform computational analysis by identifying the movement features most salient to affective perception in human observers.

In this section, an overview of the affect-expressive movement perception literature is presented to motivate and support the computational analysis of affect-expressive movements covered in the following sections.

### 4.1 Full Body Movements and Gait

Boone *et al.* [87] report the following six movement cues used by participants for the perception of affect-expressive dance movements: changes in tempo (anger), directional changes in face and torso (anger), frequency of arms up (happiness), duration of arms away from torso (happiness), muscle tension (fear), the duration of time leaning forward (sadness). Camurri *et al.* suggest that the duration of the movement, quantity of the movement (the amount of observed movement relative to the velocity and movement energy represented), and

contraction index (measured as the amount of body contraction/expansion) play key roles in the perception of affect from dance movement [88].

Studies from psychology indicate that affective states can be expressed during walking and recognized above chance level by human observers (e.g., videos [90], [91], animation [54], [92]). These studies show that affective states modulate gait parameters and kinematics such as walking speed, shoulder and elbow range of motion, and head orientation [90], [30], [91], [92]. For instance, sadness and depression are characterized by reduced walking speed, arm swing, and vertical head movements as well as slumped postures and larger lateral body sway [30]. In gait-based analysis using the PAD model, human observers recognize differences in arousal better than differences in pleasure [54].

### 4.2 Upper Body Movements

There are fewer works on affect perception from individual body parts in isolation from the rest of the body. There are situations where only individual body parts are observable due to occlusion of the rest of the body (e.g., the head and hands movements in a video call). Furthermore, there is interest to display affect-expressive movements on embodiments that, due to kinematic constraints, are incapable of full-body movements and rather display only isolated limb movements (e.g., [93], [94]). Therefore, it is important to explore expression and perception of affect through individual body parts.

Ekman and Friesen suggest that head orientation is an indicator of gross affective state (i.e., positive vs. negative) as well as intensity of emotion [10]. Busso *et al.* conducted a user study using the PAD model to evaluate the perception of affect from head motion during affective speech [95]. They report that head motion corresponding to different affective states is characterized by distinct motion activation, range, and velocity.

There is evidence that hand and arm movements are most significant for distinguishing between affective states [12], [96], [36]. Different affective states conveyed with hand and arm movements are recognized above chance level (FL videos of hand and arm movements [97], [98], [99], [100], PL animation of affective drinking and knocking movements [28], animated anthropomorphic and non-anthropomorphic hand models displaying abstract movements [101], [102]). Velocity, acceleration, and finger motion range are frequently reported as important hand and arm movement features for distinguishing different affective states [100], [28]. For instance, in [100], happy movements were characterized by indirect arm trajectories, angry movements were forceful and fast, whereas sad movements were slow and weak. Perceived arousal was found to be correlated with velocity, acceleration, and jerk (rate of change of acceleration) of the arm movement [28]. Affective state is also recognized above chance level during sign language communication, even for observers who do not understand the language being signed [98], [99].

### 4.3 Gender Differences in Perception

The effect of gender on the perception of bodily expression of affect is largely unexplored. Differences in the perception of affect-expressive movements may exist due to the gender of the demonstrator and/or observer. In general, reports on gender differences in the perception of affective state mainly focus on facial expressions. There are various and sometimes contradicting findings on the abilities of men and women in decoding facial expressions [103]. Women perceive affective states through facial expressions more accurately than men [104], [105], [106]. Men are found to be superior in recognizing angry facial expressions, e.g., [107], [108], [109], whereas women are found to be better at perceiving happy facial expressions, e.g., [110], [111] and sad facial expressions, e.g., [109]. However, in a recent study on decoding affect-expressive movements, male observers outperformed female observers in recognizing happy movements, whereas the female observers were better at recognizing angry and neutral knocking movements [112]. In another study, no significant gender differences in the perception of affective hand and arm movements were observed [97].

Other studies investigate the role of the demonstrator's gender in the perception of affect-expressive movements. In a user study, participants tended to apply social stereotypes to infer the gender of a point-light display throwing a ball with different emotions. Angry movements were judged to be demonstrated by men and sad movements were more likely to be attributed to women [113]. Similarly, the perception of fearful gait is facilitated if the walker is female [114], due to kinematic similarities between fearful gait and natural female gait.

These sometimes conflicting findings illustrate the important role that gender might play in the perception and demonstration of affect-expressive movements, and emphasize the importance of considering gender in studying affect-expressive movements. For computational affect-expressive movement analysis, to remove (or control) the potential role of gender, databases should contain a balanced number of male and female demonstrators, and the reliability of the databases should be evaluated with both male and female observers.

### 4.4 Embodiment

The embodiment of an artificial agent expressing an affective state can be physical (robotic agent) or virtual (animated agent). Physical and virtual embodiments can be further subdivided into anthropomorphic (human-like kinematics and appearance e.g., physical anthropomorphic [115], virtual anthropomorphic [116]) and non-anthropomorphic (non-human-like kinematics and appearance e.g., physical non-anthropomorphic [117], virtual non-anthropomorphic [102]).

It is well known that humans can perceive affective states from non-anthropomorphic demonstrators. For example, humans can perceive life-like affective states

from the movement of abstract geometrical shapes [118]. However, there are conflicting reports on the role of embodiment in perception. In some studies, the perception of affect-expressive movements was not influenced by non-anthropomorphic appearance [101] and kinematics [117] of demonstrators, whereas other studies have shown that non-anthropomorphic kinematics and appearance of display embodiment may influence the affect-expressive movement perception [119], [120], [102]. For instance, in a recent study [102], the non-anthropomorphic appearance was found to significantly affect participants' ratings of affect-expressive movements demonstrated on animated human-like and frond-like embodiments (see Fig. 2).

Movement features salient to the perception of affect from animated geometric agents include absolute velocity, relative angle, relative velocity, relative heading, relative vorticity, absolute vorticity and relative distance [121]. The arousal component of perceived affect-expressive movements displayed on embodiments with non-anthropomorphic kinematics [117], [122] and appearance [102] is found to be correlated with the velocity and acceleration of the movements. In [122], the valence component of perceived affect-expressive movements, displayed by an interactive device with non-anthropomorphic kinematics, is related to the smoothness of the movements.

In a recent study, the impact of the embodiment on the participants' perception of affect-expressive movements varied between male and female participants [123]. For instance, male participants correctly recognized arousal and valence characteristics of angry movements regardless of the embodiment, whereas female participants associated less arousal and less-negative valence to the non-anthropomorphic embodiment displaying the angry movements.

There are also conflicting reports on differences between perceiving affective states from physical embodiments (e.g., robot) and virtual embodiments depicted in video or animation. Some studies report no significant difference [124], [125], whereas significant differences between the perception of affect-expressive movements from videos of an actor and his animated replicates are reported in [116]. Therefore, the embodiment of affective displays may influence their interaction with human users and merits further exploration to identify the role of display embodiment in the perception of demonstrated social, behavioural, and affective cues.

## 5 AUTOMATIC RECOGNITION OF AFFECT-EXPRESSIVE MOVEMENTS

Although affective phenomena have been studied for over a century and a half by psychologists [126], [127], it was only in the 1990s that research into systems capable of automatic recognition of affective states attracted the interest of engineers and computer scientists [128]. Early research was conducted on facial expressions

and identified as challenges, among others, the subconscious nature of the perception of affective expressions as well as contextual, interpersonal, and intercultural differences in expressing and perceiving affect [7], [6]. Automatic recognition of facial expressions has been investigated for both person-dependent and person-independent recognition<sup>7</sup> of acted facial expressions of basic emotions. Most recent works move towards the recognition of spontaneous expressions, representing affective states as dimensions rather than distinct categories, and multi-modal recognition based on facial expressions, linguistic and non-linguistic components of speech, and body movements [7], [6], [2], [4]. While recognition from body movements is not as mature as facial expression recognition, most of the challenges described above for automatic recognition of facial expressions have been investigated in individual case studies for body movements.

In the following, we summarize the type of sensors recording the movements, the consideration of acted, elicited, or natural expressions, and the components of an automatic recognition system for affect-expressive movements. Then, we provide an overview on the current achievements on analyzing upper-body and whole-body movements.

### 5.1 Sensors for Recording Affect-Expressive Movements

The following approaches have been applied to record expressive motions:

- Computer vision [38], [129], [130], [100], [59], [58], [52], [131], [132], [133], [134],
- Motion capture [53], [56], [54], [135], [136], [137], [55], [51], [138],
- Pressure sensors [139], [140].

Optical motion capture data provides high resolution data which facilitates recognition. However, optical motion tracking is limited to methodological studies, while wireless inertial motion tracking [53], [138], data gloves, bend sensors, pressure sensors, or computer vision algorithms are used to record movements in a naturalistic environment. Most studies that approach real-world scenarios are based on computer vision and track regions of the hands, the head, the shoulders, the neck, the upper body, or the complete body [38], [129], [59], [141]. A framework for extracting a minimum representation of affective states from the position and dynamics of the head and hands from video data is provided in [132]. To circumvent the difficulties that accompany computer vision algorithms (e.g., different lighting, background conditions, viewing angle, and occlusions), other approaches suitable for HCI include inertial motion tracking [53], [138] or, for seated postures, instrumenting the

<sup>7</sup> *Person-dependent* recognition refers to the case where both training and test data come from the same individual, and *person-independent* recognition refers to the case where the training and test data come from different individuals.

seat and the back of the chair with pressure sensors [139], [140]. These sensors are especially suitable to insure privacy by avoiding the use of video data from which the user's identity might be easily retrieved.

### 5.2 Acted, Elicited, or Natural Expressions

Automatic recognition of affective states from facial expression is moving towards the use of databases that consist of natural expressions [7], [6], [2], [4]. A similar trend is observed for affective body movements. Early studies are based on acted or elicited affective expressions. Recent work has considered the recording of natural expressions during HCI [53], [142], [139]. Still, the number of such studies is small in comparison to works on facial expression.

Most studies rely on data from acted or elicited affective expressions. Dancers demonstrate the movements in [100], [58], [57] and non-professional actors in [51], [59], [131]. Affective expressions are either elicited by a human-computer conversation [129], or by imagining a scenario. The scenario is either a short story [38], [143], [56] or one of a person's own memories [54], [55]. A small number of studies investigate natural expressions during HCI. The scenarios are playing a computer game [53], playing a video game [142], and using an automatic tutoring system [139]. Kleinsmith *et al.* discuss advantages and disadvantages of recording acted and elicited expressions and conclude that works on acted and elicited expressions provide useful insights, but that more studies are necessary which consider natural expressions for real-world HCI/HRI applications [19].

Furthermore, the datasets differ regarding annotation. Ground truth is considered to be the affective expression which actors are asked to act, the intended elicited expression, e.g., [56], [54], or the affective state which is most frequently labeled by observers, e.g., [144], [142]. This aspect is discussed extensively in [19], and Kleinsmith *et al.* introduce a novel approach for reliably estimating the affective expression for a small number of annotators [142]. When moving towards natural expressions, a reliable estimate for the ground truth is necessary and recognition algorithms need to be developed which can handle label uncertainty and multiple labels for an expression.

A table listing the available affect-expressive movement databases is provided in the supplementary material.

### 5.3 Components of an Automatic Recognition System for Affect-Expressive Movements

Developing an automatic recognition system for affect-expressive movements consists of the following four steps (e.g., [57], [132], [56], [129]):

- 1) Estimation of motion trajectories from sensor data,
- 2) Temporal segmentation (based on time windows or movement primitives),



- 3) Construction of a feature set and dimensionality reduction (when necessary),
- 4) Detection of affective expression.

Steps 1 and 2 address human movement analysis in general. Step 1 reconstructs a description of the movements, e.g., joint angle trajectories, from sensor data. For this purpose, various sensors can be used, see section 5.1; the motion reconstruction approach used depends on the sensor data. Step 2 partitions the continuous time series data into short-duration segments for subsequent analysis. Many studies investigating affect-expressive movements use pre-segmented data, which are either manually segmented or only individual movement segments are recorded. Only a small number of studies include automatic segmentation, e.g., based on the quality of movement [57] or motion energy [56]. For training and evaluation of the classifier, these segments need to be labeled.

A large variety of features can be constructed from movement data. Most approaches transform the time series data describing the motions to a set of time-independent variables using descriptive measures. In [60], functional dimensionality reduction is used to extract relevant features directly from time-series data. In general, the following three approaches are commonly used for constructing the feature space: 1) A set of hand-selected features is created describing human movement and when necessary this set is reduced to relevant features by dimensionality reduction, e.g., [38], [129], [51], [55], [54], [52], [53]. This approach is not grounded in psychological theories about affective expressions and is particularly suitable when the sensor data cannot easily be related to a kinematic or shape-based model of human motion, e.g. for pressure sensors integrated in a seat [139]. 2) In the second approach, features are selected based on findings from perceptual studies in psychology [54]. 3) The third approach utilizes a notation system to create a set of high-level descriptors as features [100], [59], [58], [141], [138], [57]. These three approaches are not mutually exclusive and can be combined to find a comprehensive feature set, e.g., in [57]. Across the three approaches, movement speed is commonly selected as a feature in most studies.

Step 4 maps low-level motion features or high-level descriptors to affective states. Classifiers such as Support Vector Machine, Naive Bayes, Nearest Neighbor, or Multi-Layer Perceptron are used for categorical labels. When a dimensional representation is used, each dimension is either categorized into low, neutral, and high and a classifier is trained, or regression techniques are applied, e.g., Support Vector Regression. Each technique can be applied for either person-independent or person-dependent recognition. Reported recognition rates from several studies are summarized in the supplementary material.

To date, only a small number of studies have made use of movement notation systems [131], [100], [59], [141], [138]. The lack of use of the notation systems can be

explained by 1) a lack of quantification of the notation systems, 2) a lack of awareness of the notation systems, which are most commonly used in the dance community and psychology, in engineering and computer science, and 3) a lack of systematic and quantitative mapping between the components of the notation system and affective states. The majority of the approaches map kinematic motion features directly to affective states or dimensions. This approach provides reliable recognition rates when only movement data is available, comparable recognition rates to facial expressions or speech in multi-modal scenarios, and improves overall accuracy in multi-modal systems when combined with other modalities [38], [129], [141], [145].

An additional challenge when developing systems that automatically analyze affective expressions is that expressions differ between individuals in both intensity and quality. Factors contributing to these differences are, among others, differences in early child development, personality, life history of a person, gender, and culture [146], [147], [14], [15]. Inter-individual differences in the expressiveness and movement performance lead to noticeable decreases in accuracy for person-independent recognition [56], [59], [54], [51]. This difference is particularly large for functional movements where expressiveness is secondary to function [54], [56]. For a group of known users during HCI/HRI, the recognition accuracy can be improved when the algorithm is adapted and trained for each user. An approach to circumvent person-specific training would be to cluster affect-expressive movements in sets of expression styles. Persons who express affective states in a similar style are grouped into a cluster and a form of style-specific recognition is developed which is between person-dependent and person-independent recognition. Kleinsmith *et al.* discuss the dependence of affective expressions on the culture [19]. As nonverbal behavior differs between cultures [14], [15], this aspect needs to be considered when recording databases, annotating databases, developing algorithms, and transferring a system to a different cultural area.

## 5.4 Upper Body Movements

*Abstract* and *functional movements* have been investigated in studies using optical motion capture. Pre-defined abstract movements are studied to analyze recognition of affective states based only on modulation of the specified movement, e.g., arm-lowering motions [100], both raising and lowering the arm [59], and closing and opening the hand [102]. The results of these studies [100], [59], [102] support the hypothesis that information about affective states can be retrieved from the way a motion is performed and recognition is not only limited to the detection of expressive gestures. This result is relevant for functional movements, where expressiveness is only possible through trajectory modulation, e.g. knocking [56].

Studies that approach naturalistic scenarios are mostly based on computer vision. Estimation of head pose and

gesture types using computer vision is covered in the surveys [148] and [149]. To generalize to a wide range of movements, classifiers for recognition of affective states are trained on a variety of *communicative* movements, which include gestures, head motions, torso movements, or combinations of them. These movements can vary in type selection and style. Considering a set of four affective states, Glowinski *et al.* suggest that the discriminative features may be features describing movement style rather than specific gesture types [132]. This facilitates recognition, because learning a complete set of possible gesture types that might occur during an interaction can be avoided. It has not yet been investigated whether this finding holds also for discriminating between a larger set of affective states.

Recognition of affective states from observing only head movements is investigated in [150], [151], [152], [153]. Using only a small set of possible movement types (e.g., nodding, shaking, tilting, no movement, leaning forward and backward) and the way they are performed (modulation), results in person-dependent recognition rates above chance level. Person-independent recognition rates above chance level are reported when combining head movements with arm gestures or observing only arm gestures and torso movements, e.g., [53], [38]. By observing only torso movement during HCI, e.g. by instrumenting the seat and the back of the chair in front of the computer with pressure sensors, it is possible to classify the two-class case whether a specific affective state is expressed or not, but the recognition rate for distinguishing between several affective states is low [139]. Hence, the most promising approach to achieve robust person-independent recognition is to create a feature set that captures head movements, torso motion, and arm gestures.

Several studies investigate multi-modal systems which combine the recognition of affective states from body movements, facial expressions, and speech [38], [129], [52], [141], [131], [153]. For example, [38] uses both body movements and facial expression to recognize affect, and finds that recognition accuracy is improved over either modality alone. The results of these studies indicate that gestures and upper body movements are suitable modalities for multi-modal systems to recognize affective states and improve recognition accuracy of the overall system. Furthermore, it might be easier for a vision-based system to model and recognize affective states from a combination of gestures and head motions than from atomic movements in the face [38].

## 5.5 Gait, Dance & Whole-Body Gestures

Many of the studies on recognition of affective states from whole-body movements use optical motion capture providing high accuracy of the reconstructed joint angles [55], [135], [136], [54], [137], [51], [60]. Naturalistic scenarios are approached in [142], [57], [58], [131] using computer-vision or body-worn sensors for motion

recording. The state of the art can be subdivided into works considering *functional* movements such as walking, *artistic* movements such as choreographed and non-choreographed dancing, and those analyzing *communicative* movements such as whole-body gestures.

For *walking*, the primary task is locomotion; affective expression through variations in the walking style is secondary. Gait databases are characterized by highly dimensional, temporally dependent, and highly variable data vectors. Efficient dimensionality reduction is essential, achieved, e.g., by an adapted blind source separation algorithm which considers joint-specific time delays [137]. Recognition rates for person-dependent training are higher than person-independent training [136], [54], [55]. Therefore, a two-stage approach is proposed in [54], which consists of first identifying the person and then applying person-dependent recognition of affective expressions. Conversely, identifying a person by his/her walking style is affected by expressiveness; accuracy for identification decreases by about 10% when the affective state differs between the training and test set [54].

The recognition of emotions from *dance* movements is studied in [57], [58]. Basic emotions are recognized in both studies without the need to relate a set of specific expressive gestures to each emotion. This is necessary to allow the dancers to perform any motion and not restrict the dancers to a set of motions.

Gestures communicate information about affective states and/or emphasize words during verbal interaction (i.e., kinesic regulators). This makes them especially suitable for recognizing affective states. A set of acted, expressive *whole-body gestures* are considered in [51] and [144]. Even though it can be expected that the expression of affective states by gestures has common characteristics among subjects, inter-individual variations are sufficiently large to influence recognition accuracy [51]. Kessous *et al.* [131] use the EyesWeb software [154] to extract features from whole-body gestures and accuracy is larger than for features extracted from facial expressions or speech in their study. The highest recognition rate is achieved by feature-level fusion of the three modalities (gesture, facial expression, and speech) [131]. Dimensionality reduction of high-dimensional whole-body gestures is directly applied to the time series data using both supervised and unsupervised techniques in [60].

Work on *static postures* considers snapshots of movements as the input signal and excludes dynamic information. The discriminability of affective postures is studied in, e.g., [144], [142]. Feature extraction and classification do not process temporal information. These studies show that, in addition to the dynamic characteristics of movement, postural configurations contain information about the affective state. In most studies, postural cues are not explicitly separated from dynamic cues. Hence, recognition of affective expressions from body movements is usually based on a feature set combining postural and dynamic movement information.

## 6 GENERATION OF AFFECT-EXPRESSIVE MOVEMENTS

This section reviews the research on the generation of affect-expressive movements for virtual agents for HCI and robots for HRI. It is believed that the character of the virtual agent or robot becomes more believable if emotional body language stylizes the movements [116]. Both facial and bodily expressions have been investigated for human-like, zoomorphic, or cartoon-like virtual and physical embodiments [155], [156], [157], [158], [152]. Affect-expressive movements are particularly suitable for simplified humanoids without sophisticated hardware for displaying facial expressions. For virtual agents, bodily and facial expressions have been combined in a few studies, providing an enhancement of the character's expressiveness [155]. The generation of affective expressions for robotic heads and virtual faces can utilize the FACS [155], [156], [157], [159]. Such a standardization facilitates the development of automatic generation of affect-expressive movements as it provides engineers and computer scientists with a common psychologically-accepted framework and enables a systematic comparison between results of different studies. Unfortunately, such a coding system has not been developed and evaluated for the generation of expressive body movements with respect to generalization across a variety of different movements and embodiments.

Automatic generation of affect-expressive movements includes the following steps:

- 1) Selection of desired affective state
- 2) Movement type selection
- 3) Movement modulation
- 4) Trajectory generation (when necessary)
- 5) Robotics: Motor commands

The movement type is either selected based on the desired affective state (step 1) or to accomplish a functional task. Modulation of the movement type (step 3) adds affective expressiveness to a functional or abstract movement and enhances expressiveness for communicative or artistic movements. When step 2 and 3 provide key poses, step 4 generates smooth trajectories for each joint, which link the joint angle configurations from the start, the intermediate, to the end frames. These trajectories are used directly to control the joints of a virtual agent. An additional step 5 is required for real robots to translate these trajectories into motor commands. The desired movement type and its modulation (step 2 and 3) can be either designed by artists, dancers, photographers, and other experts, e.g., [160], [161], or automatically generated by a computational model. The latter approach is preferable in affective computing because it facilitates replication and generalization for different situations and different embodiments. This section focuses on summarizing the approaches which provide explicit *computational models* for relating desired affective states to movement features.

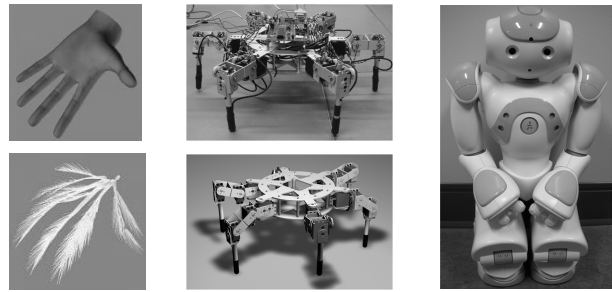


Fig. 2. Generation of affect-expressive movements is studied for a wide variety of embodiments, e.g., for an artificial hand or a frond-like structure [102], a real or animated hexapod [162], and a humanoid [163].

### 6.1 State of the Art

The animation industry has extensive experience developing believable character movement. Animators at Walt Disney studios have proposed a set of 12 design principles to create believable characters [164], [165], [166]. Out of these 12 design principles, four are associated with the expression of affective states and can be relevant for enhancing the believability of a virtual agent or robot [164], [165], [166], [167], [168]. Experienced animators recommend that only one clearly defined affective state is expressed at a time. Expressing an affective state can be a secondary action, when it is started before or after the primary action, e.g., displaying happiness while drinking. These two principles consider that the attention of the observer may be drawn to only one action at a time and a simultaneous secondary action may be easily missed by the observer. Third, appropriate timing, such as fast or slow movement, influences the expression. And fourth, exaggeration can be utilized to enhance the visibility of an expression. The former two principles influence the *selection* of the affect-expressive movement; timing and exaggeration influence the *generation* of the specific motion path for the selected expression. In the following, different approaches are summarized for a computational model which generates affect-expressive movements. Models for generation have been applied to a wide variety of embodiments, see Fig. 2.

#### 6.1.1 Virtual Agents

Studying bodily expressiveness with animations has the advantage that the full range of possible joint angles can be investigated, whereas hardware limitations can decrease the range of possible movements for physical embodiments such as robots. Using animations also facilitates the evaluation of a computational model. A model can be easily tested regarding its generalizability to different embodiments and to different movements. Perceptual user studies are commonly used to evaluate the computational model [102], [169], [170]. An alternative is to compare the generated motion trajectories to recorded affect-expressive movement trajectories [171]. The computational model relies on a relationship be-

tween movement features and affective states. This relationship can be designed manually from motion capture templates [171], [172], or extracted by machine learning [102].

For example, the motion of an animated figure can be described by a Fourier series for each joint angle over time. Unuma *et al.* propose a model which continuously transitions between two behaviors, e.g., neutral and sadness [172]. The transition is controlled by linear interpolation between the models of the two behaviors. A template for each affective state and each movement needs to be recorded with optical motion tracking to obtain the Fourier series. Amaya *et al.* also rely on optical motion capture for affect-expressive movements [171]. They derive an abstract description for expressiveness based on the speed and spatial amplitude of the movement. Other approaches that rely on recording sample movements with motion capture and interpolating between these templates describe a movement using radial basis functions [173], a Hidden Markov Model [174], or nonlinear inverse optimization [175]. Chi *et al.* extend the number of movement descriptors by including the Effort and Shape components of Laban [72]. The Effort and Shape components can be varied to achieve expressiveness for virtual characters. The intention of [72] is to provide animators with a tool to design affective expressions by high-level descriptors. However, a computational model between the descriptors and affective states is not integrated. A behavior markup language is proposed in [176], [177], which formulates an XML format for behavior generation for virtual agents. Each behavior is encoded by the developer and can have attributes such as duration and amount.

Deriving high-level descriptors for affective expressions facilitates the transfer to other movements, but requires that two relations are defined: 1) the relationship between the movement features and the descriptors and 2) the relationship between the high-level descriptors and affective states. The studies [169], [170] circumvent the need for the second relationship by mirroring high-level descriptors of human affective expressions during HCI. They observe the human's gestures with video analysis and use contraction index, velocity, acceleration, and fluidity as descriptors for human movements. These descriptors are mapped to spatial extension, temporal extension, power, and fluidity of the virtual agent. This approach is evaluated in a perceptual user study for a repetitive hand motion (called beating in the original text) [169], [178], and choosing different gestures depending on the expressed affective state in [170].

Machine learning can be utilized to automatically derive a set of descriptors for affect-expressive movements. One possible approach is proposed in [102], where descriptors represent the main modes of variation in a collection of affect-expressive movements, obtained using functional PCA. These descriptors were then used to generate affect-expressive movements, which were perceived as intended by observers in a user study [102].

### 6.1.2 Robotics

The concept of *proximity zones* in HRI follows the theory of proxemics for human-human interaction in psychology [36], [8]. The distance between human and robot is divided into intimate (0m–0.46m), personal (0.46m–1.22m), and social (1.22m–3.66m) proximity zones. Body movements are especially suitable for conveying affective states during HRI in the *personal* and *social* proximity zone [8]. Several approaches have been undertaken to automatically generate expressive motions for robots. They can be subdivided into studies which utilize Laban to describe affect-expressive movements and studies which model kinematic parameters directly.

A relationship between motion parameters and affective state is directly derived in [162], [179], [163]. Beck *et al.* obtain a key pose for each affective state, arrange these in the valence and arousal space, and generate new expressions by blending between the joint angles of the key poses [163]. In [162], the gait of a hexapod robot is adapted to display affective states by variations of the gait parameters step length, height, and time. Values for each parameter corresponding to low or high pleasure, arousal, and dominance are retrieved from a human gait database and are mapped automatically to the kinematics of the hexapod. To find a general approach which modifies an arbitrary motion so that affective nuances are expressed, Nakagawa *et al.* propose to divide an arbitrary motion into the velocity and expansiveness of the motion and a basic posture [179]. Velocity and expansiveness correlate with arousal and the basic posture relates to the expressed level of valence, with a contracted posture for low valence and an open posture for high valence. Evaluation for the movements pointing and waving shows that the type of motion has an effect on the expressiveness, and that the observers recognized the intended affective nuances except for the combination of high valence and low arousal. The approaches [162], [163], [179] derive a direct relation between motion parameters and affective states.

Laban provides a scheme to describe motions, but is not in itself a computational model for relating affective states to bodily expression. Several research groups have proposed the use of Laban to derive a computational model for the generation of affect-expressive movements. To derive a computational model, first the relationship between an affective expression and its Laban-based movement descriptors needs to be established. This can be achieved by manually designing a set of possible relations [94], taking a relation between affective state and movement from psychological studies, e.g., [180] who refer to [88], or evaluating the relation for a specific robot in a user study [181], [158], [182]. In these studies, affect-expressive movements are modeled either using only the Effort component [180] or both Effort and Shape components [181], [158], [94]. Secondly, the mapping between the Laban components and the joint angles and velocities needs to be specified for the

kinematics of the robot. Reporting this mapping, e.g., in [181], [158], facilitates the transfer of a published approach for a specific robot to new studies with similar kinematics.

## 6.2 Influence of Embodiment on Expressiveness

Virtual agents and robots can differ in their embodiment. Perceptual user studies have shown that differences in the embodiment can lead to differences in the perception of affective states, and that the embodiment-specific impact on perception limits the range of affective states that can be conveyed through movement (see section 4.4). The results from perceptual user studies indicate that a general computational model for generation of affective expressions needs to consider the influence of embodiment on expressiveness. Individual case studies have addressed this influence for a virtual versus a physical robot [162], different appearances [94], and anthropomorphic versus non-human like appearance [102]. Each study evaluated a single aspect in a user study and the interplay between aspects has not been investigated.

Takahashi *et al.* investigate the perception of affective states for a teddy bear robot and a black-suited teddy bear robot performing the same expressive motion [182]. Their results show that the perception of anger and disgust depend on the appearance (regular versus black-suited) of the robot whereas the perception of joy, sadness, surprise, and fear did not depend on the appearance. Differences in human perception are also reported for generated affect-expressive movements displayed on a virtual hand with human-like or frond-like appearances (Fig. 2), e.g., sad movements displayed on the frond-like hand were perceived as happy movements [102]. However, Saerbeck *et al.* find no difference in expressiveness between an iCat and a Roomba robot when the motion characteristics acceleration and curvature are altered to convey affective expressions [117]. These studies show that differences in the kinematics and the appearance can lead to differences in the perception of affect-expressive movements.

## 7 DISCUSSION

In the following, we summarize the key issues for automatic recognition and generation of affect-expressive movements. We elaborate what movements have been investigated, whether affective states are represented as categorical labels or dimensions, and what movement features and computational models are used. We conclude by outlining future research directions.

### 7.1 What Movements Have Been Investigated?

The perception of affective states from bodily motions has been studied in psychology for a variety of affect-expressive movements: functional movements, e.g., [91], [90], [28], [54], [183], [92], communicative movements, e.g., [98], [99], artistic movements, e.g., [87], [88], [85],

[184], and abstract movements, e.g., [89], [13], [102], [100]. Perceptual studies mostly focus on analyzing the full body and less attention has been directed to affective expression through individual body parts in isolation from the rest of the body [102], [95].

Motivated by these results from psychological studies, several machine learning algorithms have been developed to recognize affect-expressive movements. Algorithms learn either to detect an affective state for one specific movement or for various movements. In the latter case, the training data contains several movements for the same affective state and the approach attempts to generalize across different movements. Several studies analyze communicative movements, e.g., upper-body gestures [51], [52], [38], [53], and demonstrate that affective states can be recognized from gestures. A small number of studies analyze the recognition of affective states from artistic movements [58], [57], and abstract movements [59], [100]. The results of [57], [59], [100] indicate that affective states can be retrieved not only from different movement types for each affective state but also from the way a movement is performed. Functional movements such as walking and knocking have been investigated in [54], [55], [56]. As the primary task is walking or knocking, affective expression is secondary and it is displayed mainly by modulation of the functional movement. These variations can be retrieved from high-resolution motion-capture data.

The generation of affect-expressive movements is often confined to a small set of movement types as exemplars for each affective state. Abstract movements, e.g., hand movements [102], and functional movements, e.g., drinking, knocking, kicking [171], clapping [169], have been studied for computer animations. Implementations on robots mainly focus on upper-body movements, e.g., [179], [158], [94]. Exceptions are studies on the functional movements gait [162] and planar locomotion [180]. Evaluation of the generated affect-expressive movements show a strong relationship between movement type and style for expressiveness. Modulation of the same movement can lead to different affective expressions [162], [102]. Some affective states may be more easily conveyed by a set of specific movements, e.g., happiness by waving. This is supported by the results of [179], [181], [158], [169], [178] for arm movements. On the other hand, some movements may be unsuitable for conveying certain affective states, e.g., anger in a study-specific conversational gesture [169]. Mancini *et al.* hypothesize that the movement type can signal a particular affective state so strongly that a different affective state can be difficult to communicate [169]. Possible explanations for this interaction are that some movements are only performed during certain affective states in everyday life, the meaning of the movement type dominates for some movements, and some movements can not be modified to express certain affective states. More detailed research on the role of movement type and style on the expressiveness would facilitate the generalization of algorithms

for automatic recognition and generation across different movements.

## 7.2 Affective Dimensions or Categorical Labels ?

Most perceptual studies on affect-expressive movements use a set of categories representing the affective states. Fewer studies investigate perception of affective dimensions through body movements. Similarly, the majority of the proposed approaches for automatic recognition and generation of affect-expressive movements refer to a set of categorical labels and only a few utilize a dimensional approach.

When using categorical labels, a subset of the basic emotions is predominantly studied, particularly happiness, sadness, and anger for both automatic recognition and generation. This is in accordance with results from perceptual user studies in psychology, which show that the basic emotions anger, sadness, and happiness are conveyed through body movements [90], [13]. Often, the discrimination between anger and sadness, which differ mainly in arousal, is more accurate than the discrimination of states which differ mainly in valence (recognition: [56], [54], [59], generation: [182], [180]). An exception is [51], where joy and anger are most easily distinguishable. Only a few studies investigate fear [51], [100], [132]. Sawada *et al.* [100] analyze the complete set of Ekman's basic emotions and find that surprise, disgust, and fear are difficult to express by simple arm movements. The observation that some affective states are less easily conveyed by body movement or movements of individual body parts is also supported by perceptual user studies in psychology, e.g., [89], [92], [91]. These findings highlight the importance of both appropriate movement and body part selection to facilitate communication. For some affective states, other contextual cues or modalities (e.g., facial expression) might also be needed.

A smaller number of studies utilize a dimensional approach for representation (automatic recognition: [144], [54], [129], [142], [138], [132], [150], [151], generation: [162], [179], [163]). They typically all use the dimensions of valence and arousal. Each dimension is either categorized [144], [54] or a regression model is applied to model continuous values for valence and arousal [129], [138], [150]. Better discrimination of arousal compared to valence is reported for automatic recognition in [54], [138]. The few studies on generative models using a dimensional approach report similar results: differences in valence are harder to communicate by modifying the functional movement walking than differences in arousal [162], and the combination of high valence and low arousal may not be clearly represented in affect-expressive movements [179], [163].

## 7.3 What Movement Features and Computational Models are Used?

The set of possible features which can be calculated from movement data is large; calculated feature sets

differ largely among the studies. The feature sets can be derived heuristically, by dimensionality reduction techniques, or from findings of psychological studies. Most studies include speed as a feature. The contribution of speed to the perception of affective states is confirmed by perceptual studies in psychology, e.g., [92], [91], [88]. Furthermore, a feature or multiple features describing the spatial extent of a movement is included in the majority of the studies.

The computational models can be divided into *direct models* which directly relate movement features to affective states, and *two-stage models* which first model the relation between descriptors and affective states, and secondly model the relation between movement features and descriptors, e.g., fluidity. Several studies utilize a movement notation system for designing a two-stage model (Laban: e.g., [59], [58], [138], [57], [158], [182], BAP: [37]). Considering the generation of affect-expressive movements, the use of a movement notation system facilitates generalization across different movements and embodiments. The recently proposed BAP system provides a more detailed categorization of body movements than the Laban system, and its applicability to affect-expressive movements has been shown in [37].

Experimental evaluation of a computational model for generation of affect-expressive movements through user studies is essential. The design of such user studies varies widely across the works on generative models. Recommendations on designing and conducting human studies in HRI are summarized in [185].

## 7.4 Directions for Future Work

To enhance generalization across individual movements and move towards integration into interactive scenarios, the following future directions are suggested:

### 7.4.1 A Notation System for Affect-Expressive Movements

Movement features and descriptors vary largely across studies on automatic recognition and generation. A psychologically accepted and validated notation system that relates affective states to bodily expression would provide a common basis facilitating interdisciplinary research. Developing computational models on such a common framework enables comparison between different studies, provides a foundation for the choice of movement features and descriptors, and facilitates the fusion with other modalities to convey affective states. From the computational perspective, it is desired that such a framework describes which movement features represent specific affective states, what influence cultural and inter-individual differences have on bodily expression, and how the embodiment influences expressiveness. The first approaches towards this direction have been undertaken, utilizing the Laban or BAP movement notation system.

However, there are some limitations to the application of the current notation systems for computational affect-expressive movement analysis. Except for BAP, the existing notation systems do not provide a validated set of behavioral action units with defined boundaries and phase relationships such as those of FACS. The adoption of a notation system such as BAP for systematic coding of nuances in bodily expression of emotion could drive forward the research in computational analysis of affect-expressive movements. To the best of our knowledge, there is only one study on validating the reliability of BAP, and that study is conducted by BAP's authors [25]. More experiments are needed to verify BAP intercoder consistency in coding bodily movements either as a whole or movements of individual body parts. Furthermore, to enable the application of the movement notation systems in computational movement analysis, measurable physical correlates (e.g., velocity) of the notation descriptors (e.g., Laban Effort components [75]) need to be identified and used for quantifying the descriptors.

#### 7.4.2 Representation of Affective States: Choice of States and Dimensions

Automatic recognition and generation have used either a dimensional or categorical representation of affective states. Choosing an appropriate representation also includes the consideration of how detailed the representation should be and which affective states should be covered [1]. To date, most studies include the categories of anger, happiness, and sadness, or the dimensions of pleasure and arousal.

Findings from psychology indicate that body movements are not limited to conveying only basic emotions [97] and future studies may consider including affective states beyond basic emotions. This is also motivated by individual studies in HCI/HRI investigating additional affective states including mood and feelings (recognition: frustration, elation, panic, amusement, relief, despair, interest, irritation, pride, boredom, confusion, delight, flow, frustration, innocence, and laughter types [59], [139], [130], [131], [132], [186]; generation: pride, disgust, surprise, fear, relaxation, and nervousness [94], [182], [158], [180], [163]). It is further of interest which affective phenomena are reliably conveyed by body movements. Affect-expressive movements may also provide the utility to detect affective states which are less clearly conveyed through other modalities or which can only be reliably recognized by observing more than a single modality, e.g., pride, which human observers can only decode reliably using both facial and bodily cues [27]. Taking a multi-modal approach opens the opportunity to consider a broader range of affective states which have relevance for HCI/HRI.

Considering a dimensional representation, it may be beneficial to include additional dimensions beyond pleasure and arousal, and future work may address determining the appropriate choice of dimensions that are most relevant for HCI/HRI applications. Individual

studies on automatic recognition and generation indicate that also other dimensions may be expressed and perceived from body movements (automatic recognition: dominance/power, potency, avoidance, expectation, intensity and interest [54], [144], [150], [140], [53]; generation: dominance[162]). Adding the dimensions potency and expectation (unpredictability) is motivated by the studies [187], [188], which emphasize that the two dimensions pleasure and arousal are not sufficient to describe a large variety of emotions. Furthermore, Cowie et al. recommends that engagement is an essential dimension for HCI/HRI, but it is rarely studied in psychology [24, p.17] or used, e.g., as evaluation parameter in [152].

#### 7.4.3 Generalization for Different Embodiments

A computational model for the generation of affect-expressive movements has not been proposed which generalizes across different kinematics and appearances. Developing such a model remains a future research area. Its evaluation can be performed for either physical robots or virtual agents. A user study on affect-expressive movements comparing the perception of affective walking styles for a physical robot and its virtual animation and similar user studies on facial expressions show that the expressiveness of a robot and of its animated display differ only slightly [162], [124], [125]. This enables an evaluation of computational models for a large variety of kinematics and appearances in virtual environments, where animations can be more easily modified than the hardware design of a real robot. When the kinematics/dynamics and the appearance of the animation and a physical robot are the same, results on the performance of the computational model in terms of expressiveness of the animated robot can be transferred to the physical robot and design guidelines for developing the hardware of the real, affect-expressive robot can be provided.

#### 7.4.4 The Benefit of Integrating Context Knowledge

Context knowledge summarizes information about the environment, interaction partners (e.g., information about gender, personality traits, and culture), current activities, and interactions. The current approaches for automatic recognition and generation of affect-expressive movements largely do not take context knowledge into account. A future direction is the development of adaptive systems to select appropriate movements through interaction with the user. A robot may learn to choose from a set of affect-expressive movements these motions which most likely match the expectations of the current interaction partner. Furthermore, analyzing the context when a movement is performed may lead to more robust recognition of affective states. When the context is considered, actions can indicate affective states, such as freezing during a conversation. Taking the context into account enhances automatic recognition and generation of affect-expressive movements by interpreting whether taken actions express an affective state, reasoning how

probable the expressed affective state is in the current scenario, minimizing the risk that movements are misinterpreted as affective expressions even though they are only functional, and adapting the behavior of a virtual agent or robot to the expectations of a user.

Multi-modal recognition provides a methodology to detect affective states conveyed through movements in the context of other modalities, e.g., [38], [129], [52], [141], [131], [153]. Combining several modalities, such as conveying affective states through facial expressions, speech, and bodily expressions provides more reliable estimates for automatic recognition and enhances believability and congruency for generation. This also provides a methodology to include affective expressions which are difficult to express by a single modality, e.g., pride [27].

#### 7.4.5 Integration into Interactive Scenarios

The long term goal of research on affective expressions is to integrate the automatic recognition and generation of affect-expressive movements into HCI/HRI scenarios. Key challenges include dealing with expressions that are natural and variable in duration and intensity, and of integrating these two modules into higher-level cognitive and interaction models.

Acted expressions may differ from natural expressions occurring during HCI/HRI interaction [2]. The transition from acted or elicited expressions to natural expressions can be guided by results of psychological studies on affective expressions in daily life and during human-human interaction, e.g., only a small number of expressions are perceived as intense, most are intermediate, and some are not emotional in a natural scenario [24]. This study indicates that machine learning algorithms need to consider different intensities of affective expressions. This has been done by qualifying categorical labels with a level of intensity (e.g., strong happy, weak happy) [188], and, for the dimensional representation, moving towards regression techniques [129] or by adding intensity as an additional dimension [188]. Furthermore, affect-expressive movements can vary in duration, and onsets and offsets of a movement can be indistinct. Considering automatic recognition, this influences the length of the segmentation windows: the expression of very brief emotions can be missed with large windows; but short windows may result in analyzing incomplete movements. Here, analyzing individual movements can be avoided by generalization across different movements, e.g., by general motion descriptors, [132], [38], [131]. The use of a movement notations system can provide guidance for choosing such motion descriptors, e.g., the Effort and Shape components of Laban.

When moving towards HCI/HRI scenarios, it is often required to integrate automatic recognition of affective states and/or generation of affective expressions into higher-level cognitive models. Cognitive models can include, e.g., perception of context cues, motivational processes, emotion generation processes, and action selection processes. When integrating generative models

into a cognitive system, the interface protocol should specify parameters for intensity, duration, and repetition of an affective expression [176], [177]. Furthermore, generative models can be shaped by personality traits, e.g., angry and happy characters have preferences for different expressions [152].

Furthermore, the interaction between several individuals or between a virtual agent/robot and individuals can be modeled. Here, psychological theories on human-human nonverbal interaction can provide valuable guidance [14], [189], [190], [191]. This is closely related to social signal processing in HCI/HRI [192], [193]. Most of the reviewed approaches detect the affective state of one individual. Extensions to the interaction and relation of several subjects or analysis of group behavior should be considered. An example considering affect-expressive movements is the study by Varni *et al.*, who propose a real-time algorithm to analyze the synchronization of affective behavior, empathy, and the emergence of functional roles in a social group during music performances and music listening [194].

## 8 CONCLUDING REMARKS

To date, essential achievements have been made to enable automatic recognition and generation of affect-expressive movements in HCI and HRI for a limited set of specific movements, primarily in laboratory settings. For both recognition and generation, communicative, functional, artistic, and abstract movements have been studied. However, only a small subset of all possible movements, i.e., gestures, walking, waving, and head movements, have been explored in detail. In these studies, affective states are represented either by categorical labels or a dimensional approach. A few studies demonstrate that affective states beyond basic emotions and beyond changes in pleasure and arousal can be communicated through affect-expressive movements. When modeling the relationship of affective states and movement features, movement notation systems (Laban, Delsarte, or BAP) can provide guidance for choosing motion descriptors. The use of a common movement notation system for affect-expressive movements would facilitate the transfer of knowledge from psychological studies to their applications in HCI/HRI and the generalization across movement types and embodiments. Finally, the development of agents capable of recognizing and generating affect-expressive movements during natural, interactive scenarios, integrating context knowledge and adapting to individual users remains a key challenge for future work.

## REFERENCES

- [1] A. Calvo and S. D'Mello, "Affect detection: an interdisciplinary review of models, methods, and their applications," *IEEE Trans. Affect. Comp.*, vol. 1, no. 1, pp. 18–37, 2010.
- [2] H. Gunes and M. Pantic, "Automatic, dimensional and continuous emotion recognition," *Int. Journal of Synthetic Emotions*, vol. 1, no. 1, pp. 69 – 99, 2010.



- [3] L. Devillers, L. Vidrascu, and L. Lamel, "Challenges in real-life emotion annotation and machine learning based detection," *NEURAL NETWORKS*, vol. 18, pp. 407–422, 2005.
- [4] H. Gunes, B. Schuller, M. Pantic, and R. Cowie, "Emotion representation, analysis and synthesis in continuous space: A survey," in *IEEE Int. Conf. FG, Workshop Emotion Synthesis, Representation, and Analysis in Continuous Space*, (Santa Barbara, CA, USA), 2011.
- [5] H. Gunes, M. Piccardi, and M. Pantic, "From the lab to the real world: affect recognition using multiple cues and modalities," in *Affective computing: focus on emotion expression, synthesis, and recognition* (J. Or, ed.), pp. 185–218, Vienna, Austria: InTech Education and Publishing, 2008.
- [6] M. Pantic and L. J. M. Rothkrantz, "Automatic analysis of faial expressions: The safe of the art," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, pp. 1424–1445, December 2000.
- [7] Z. Zeng, M. Pantic, G. Roisman, and T. Huang, "A survey of affect recognition methods: Audio, visual, and spontaneous expressions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, pp. 39–58, Jan. 2009.
- [8] C. Bethel and R. Murphy, "Survey of non-facial/non-verbal affective expressions for appearance-constrained robots," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 38, pp. 83–92, Jan. 2008.
- [9] P. Ekman and W. Friesen, "The repertoire of nonverbal behavior: Categories, origins, usage, and coding," *SEMIOTICA*, vol. 1, no. 1, pp. 49–98, 1969.
- [10] P. Ekman and W. Friesen, "Head and body cues in the judgement of emotion: A reformulation," *PERCEPT MOTOR SKILL*, vol. 24, pp. 711–724, 1967.
- [11] S. Planalp, *Communicating emotion: Social, moral, and cultural processes*. Cambridge University Press, 1999.
- [12] H. Wallbott, "Bodily expressions of emotion," *EUR J SOC PSYCHOL*, vol. 28, pp. 879–896, 1998.
- [13] M. Coulson, "Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence," *J NON-VERBAL BEHAV*, vol. 28, no. 2, pp. 117–139, 2004.
- [14] J. Burgoon, L. Guerrero, and K. Floyd, *Nonverbal communication*. Allyn and Bacon, 2010.
- [15] D. Matsumoto, M. Frank, and H. Hwang, *Nonverbal Communication: Science and Applications*. Sage Publications, 2013.
- [16] R. D. Walk and K. L. Walters, "Perception of the smile and other emotions of the body and face at different distances," *B PSYCHONOMIC SOC*, vol. 26, p. 510, 1988.
- [17] P. Ekman and W. Friesen, "Detecting deception from the body or face," *J PERS SOC PSYCHOL*, vol. 29, pp. 288–298, 1974.
- [18] H. Aviezer, Y. Trope, and A. Todorov, "Body cues, not facial expressions, discriminate between intense positive and negative emotions," *SCIENCE*, vol. 338, no. 6111, pp. 1225–1229, 2012.
- [19] A. Kleinsmith and N. Bianchi-Berthouze, "Affective body expression perception and recognition: A survey," *IEEE Trans. Affect. Comp.*, vol. 4, pp. 15–38, Jan.-Mar. 2013.
- [20] A. Ortony, *Affective Information Processing*, ch. 2: Affect and Emotions in Intelligent Agents: Why and How?, pp. 11–21. Springer, 2009.
- [21] K. Scherer, *Blueprint for affective computing: A sourcebook*, ch. Emotion and emotional competence: conceptual and theoretical issues for modelling agents, pp. 3–20. Oxford University Press, 2010.
- [22] K. Scherer, "What are emotions? and how can they be measured?," *SOC SC INFORM*, vol. 44, no. 4, pp. 695–729, 2005.
- [23] C. Derbaix and M. T. Pham, "Affective reactions to consumption situations: A pilot investigation," *J ECON PSYCHOL*, vol. 12, no. 2, pp. 325–355, 1991.
- [24] R. Cowie, N. Sussman, and A. Ben-Zeev, "Emotion: Concepts and definitions," *Emotion-Oriented Systems*, pp. 9–30, 2011.
- [25] N. Dael, M. Mortillaro, and K. Scherer, "The body action and posture coding system (BAP): Development and reliability," *J NONVERBAL BEHAV*, pp. 1–25, 2012.
- [26] I. B. Mauss and M. D. Robinson, "Measures of emotion: A review," *COGNITION EMOTION*, vol. 23, no. 2, pp. 209–237, 2009.
- [27] J. L. Tracy and R. W. Robins, "Show your pride: evidence for a discrete emotion expression," *PSYCHOL SCI*, vol. 15, no. 3, pp. 194–197, 2004.
- [28] F. Pollick, H. Paterson, A. Bruderlin, and A. Sanford, "Perceiving affect from arm movement," *COGNITION*, vol. 82, pp. B51–B61, Dec. 2001.
- [29] S. Marsella, J. Gratch, and P. Petta, *Blueprint for affective computing: A sourcebook*, ch. Computational models of emotion, pp. 21–41. Oxford University Press, 2010.
- [30] J. Michalak, N. Troje, J. Fischer, P. Vollmar, T. Heidenreich, and D. Schulte, "Embodiment of sadness and depression - gait patterns associated with dysphoric mood," *PSYCHOSOM MED*, vol. 71, pp. 580–587, 2009.
- [31] R. Buck, *The communication of emotion*. Guilford Press New York, 1984.
- [32] B. Russel, *The principles of mathematics*. London: Allen & Unwin, 1903.
- [33] S. Planalp, "Varieties of cues to emotion in naturally occurring situations," *COGNITION EMOTION*, vol. 10, no. 2, pp. 137–154, 1996.
- [34] B. de Gelder and B. de Gelder, "Why bodies? twelve reasons for including bodily expressions in affective neuroscience," *PHILOS T ROY SOC B*, vol. 364, no. 1535, pp. 3475–3484, 2009.
- [35] B. de Gelder, J. Van den Stock, H. K. Meeren, C. Sinke, M. E. Kret, and M. Tamietto, "Standing up for the body. Recent progress in uncovering the networks involved in the perception of bodies and bodily expressions," *NEUROSCI BIOBEHAV R*, vol. 34, no. 4, pp. 513–527, 2010.
- [36] M. Argyle, *Bodily communication*. Methuen, 1988.
- [37] N. Dael, M. Mortillaro, and K. Scherer, "Emotion expression in body action and posture," *EMOTION*, 2011.
- [38] H. Gunes and M. Piccardi, "Automatic temporal segment detection and affect recognition from face and body display," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 39, pp. 64–84, Feb. 2009.
- [39] P. Ekman, "Darwin, deception, and facial expression," *Annals of the New York Academy of Sciences*, vol. 1000, no. 1, pp. 205–221, 2003.
- [40] J. L. Tracy and D. Randles, "Four models of basic emotions: A review of Ekman and Cordaro, Izard, Levenson, and Panksepp and Watt," *Emotion Review*, vol. 3, no. 4, pp. 397–405, 2011.
- [41] P. Ekman and W. V. Friesen, "A new pan-cultural facial expression of emotion," *MOTIV EMOTION*, vol. 10, pp. 159–168, Jun. 1986.
- [42] P. Ekman, "Are there basic emotions?," *PSYCHOL REV*, vol. 99, no. 3, pp. 550–553, 1992.
- [43] R. Plutchik and H. Conte, eds., *Circumplex models of personality and emotions*. APA, 1997.
- [44] A. Mehrabian, "Pleasure-arousal-dominance: a general framework for describing and measuring individual differences in temperament," *CURR PSYCHOL*, vol. 14, no. 4, pp. 261–292, 1996.
- [45] L. F. Barrett and J. A. Russell, "The structure of current affect: Controversies and emerging consensus," *CURR DIR PSYCHOL SCI*, vol. 8, no. 1, pp. 10–14, 1999.
- [46] J. Mikels, B. Fredrickson, G. Larkin, C. Lindberg, S. Magold, and P. Reuter-Lorenz, "Emotional category data on images from the international affective picture system," *BEHAV RES METHODS*, vol. 37(4), pp. 626–630, 2005.
- [47] W. Havlena and M. B. Holbrook, "The varieties of consumption experience: Comparing two typologies of emotion in consumer behavior," *J CONSUM RES*, vol. 13, pp. 394–404, 1986.
- [48] J. Russell and A. Mehrabian, "Evidence for a three-factor theory of emotions," *J RES PERS*, vol. 11, pp. 273–294, 1977.
- [49] L. F. Barrett, "Solving the emotion paradox: Categorization and the experience of emotion," *PERS SOC PSYCHOL REV*, vol. 10, no. 1, pp. 20–46, 2006.
- [50] *Oxford Dictionaries.*, ch. "Gesture". Oxford University Press, April 2010.
- [51] A. Kapur, N. Virji-Babul, G. Tzanetakis, and P. Driessen, "Gesture-based affective computing on motion capture data," in *Int. Conf. ACII*, pp. 1–7, 2005.
- [52] S. G. C. Shan and P. W. McOwan, "Beyond facial expressions: Learning human emotion from body gestures," in *Brit. Mach. Vis. Conf.*, 2007.
- [53] P. R. D. Silva, M. Osano, and A. Marasinghe, "Towards recognizing emotion with affective dimensions through body gestures," in *Int. Conf. FG*, 2006.
- [54] M. Karg, K. Kühnlenz, and M. Buss, "Recognition of affect based on gait patterns," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 40, pp. 1050–1061, Aug 2010.
- [55] D. Janssen, W. Schöllhorn, J. Lubienetzki, K. Fölling, H. Kokenge, and K. Davids, "Recognition of emotions in gait patterns by

- means of artificial neural nets," *J NONVERBAL BEHAV*, vol. 32, pp. 79–92, June 2008.
- [56] D. Bernhardt and P. Robinson, "Detecting affect from non-stylised body motions," in *Int. Conf. ACII*, pp. 59–70, 2007.
- [57] A. Camurri, B. Mazzarino, and G. Volpe, "Expressive interfaces," *Cognition, Technology & Work*, vol. 6, pp. 15–22, 2004.
- [58] H. Park, J.-I. Park, U.-M. Kim, and W. Woo, "Emotion recognition from dance image sequences using contour approximation," in *Structural, Syntactic, and Statistical Pattern Recognition, LNCS 3138*, pp. 547–555, Springer, 2004.
- [59] G. Castellano, S. Villabla, and A. Camurri, "Recognising human emotions from body movement and gesture dynamics," in *Int. Conf. ACII*, pp. 71–82, 2007.
- [60] A. Samadani, A. Ghodsi, and D. Kulić, "Discriminative functional analysis of human movements," *PATTERN RECOGN LETT*, vol. 34, no. 15, pp. 1829 – 1839, 2013.
- [61] P. Ekman and W. Friesen, *The Facial action coding system*. Consulting Psychologists Press, Stanford Univ., Palo Alto, 1977.
- [62] A. Guest, *Dance notation: The process of recording movement on paper*. Dance Horizons New York, 1984.
- [63] K. White, *The bodily action research of Ray L. Birdwhistell and Paul Ekman: Implications for oral interpretation theory*. PhD thesis, Univ. Arizona., 1983.
- [64] R. H. Rozensky and L. F. Honor, "Notation systems for coding nonverbal behavior: A review," *J BEHAV ASSESS*, vol. 4, no. 2, pp. 119–132, 1982.
- [65] J. A. Coan and J. M. Gottman, "The specific affect coding system (spaff)," *Handbook of emotion elicitation and assessment*, pp. 267–285, 2007.
- [66] R. Birdwhistell, *Kinesics and context: Essays on body motion communication*. Univ of Pennsylvania Pr, 1970.
- [67] M. Ptaszynski, J. Maciejewski, P. Dybala, R. Rzepka, and K. Araki, "CAO: A fully automatic emoticon analysis system based on theory of kinesics," *IEEE Trans. Affect. Comp.*, vol. 1, no. 1, pp. 46–59, 2010.
- [68] R. Laban and F. Lawrence, *Effort*. Macdonald and Evans, 1947.
- [69] R. Laban and L. Ullmann, *The mastery of movement*. ERIC, 1971.
- [70] I. Bartenieff, *Effort-Shape analysis of movement: The unity of expression and function*. Albert Einstein College of Medicine, Yeshiva University, 1965.
- [71] J. Rett, *Robot-human interface using laban movement analysis inside a bayesian framework*. PhD thesis, Univ. Coimbra, January 2009.
- [72] D. Chi, M. Costa, L. Zhao, and N. Badler, "The EMOTE model for effort and shape," in *Conf. Computer graphics and Interactive Techniques*, pp. 173–182, ACM Press/Addison-Wesley Publishing Co., 2000.
- [73] T. Nakata, T. Mori, and T. Sato, "Quantitative analysis of impression of robot bodily expression based on laban movement theory," *JOURNAL-ROBOTICS SOCIETY OF JAPAN*, vol. 19, no. 2, pp. 104–111, 2001.
- [74] M. Kapadia, I. Chiang, T. Thomas, N. I. Badler, and J. Kider, "Efficient Motion Retrieval in Large Databases," in *Proceedings of the symposium on Interactive 3D graphics and games, I3D, ACM*, 2013.
- [75] A. Samadani, S. Burton, R. Gorbet, and D. Kulić, "Laban effort and shape analysis of affective hand and arm movements," in *Int. Conf. ACII*, pp. 343 – 348, 2013.
- [76] T. Lourens, R. Van Berkel, and E. Barakova, "Communicating emotions and mental states to robots in a real time parallel framework using laban movement analysis," *ROBOT AUTON SYST*, vol. 58, no. 12, pp. 1256–1265, 2010.
- [77] T. Shawn, *Every little movement: a book about François Delsarte*. Dance Horizons, 1968.
- [78] M. Nixon, P. Pasquier, and M. El-Nasr, "Delsartemap: Applying delsarte's aesthetic system to virtual agents," in *Intelligent Virtual Agents* (J. Allbeck, N. Badler, T. Bickmore, C. Pelachaud, and A. Safonova, eds.), vol. 6356 of *Lecture Notes in Computer Science*, pp. 139–145, Springer Berlin Heidelberg, 2010.
- [79] S. Marsella, S. Carnicke, J. Gratch, A. Okhmatovskaia, and A. Rizzo, "An exploration of delsarte's structural acting system," in *Int. Conf. on IVA, LNCS 4133*, pp. 80–92, Springer, 2006.
- [80] E. Velloso, A. Bulling, and H. Gellersen, "AutoBAP: Automatic coding of body action and posture units from wearable sensors," in *Int. Conf. ACII*, pp. 135 – 140, 2013.
- [81] J. A. Harrigan, "Proxemics, kinesics, and gaze," *The new handbook of methods in nonverbal behavior research*, pp. 137–198, 2005.
- [82] S. Brownlow, A. Dixon, C. Egbert, and R. Radcliffe, "Perception of movement and dancer characteristics from point light displays of dance," *PSYCHOL REC*, vol. 47, pp. 411–421, 1997.
- [83] L. Kozlowski and J. Cutting, "Recognizing the sex of a walker from a dynamic point-light display," *PERCEPT PSYCHOPHYS*, vol. 21, pp. 575–580, 1977.
- [84] J. Cutting and L. Kozlowski, "Recognizing friends by their walk: Gait perception without familiarity cues," *B PSYCHONOMIC SOC*, vol. 9, pp. 353–356, 1977.
- [85] W. Dittrich, T. Troscianko, and S. L. ans D. Morgan, "Perception of emotion from dynamic point-light displays represented in dance," *PERCEPTION*, vol. 25, pp. 727–738, 1996.
- [86] W. Thompson, P. Graham, and F. Russo, "Seeing music performance: Visual influences on perception and experience," *SEMIOTICA*, vol. 2005, no. 156, pp. 203–227, 2005.
- [87] R. Boone and J. Cunningham, "Children's decoding of emotion in expressive body movement: The development of cue attunement," *DEV PSYCHOL*, vol. 34, no. 5, pp. 1007–1016, 1998.
- [88] A. Camurri, I. Lagerlöf, and G. Volpe, "Recognizing emotion from dance movement: comparison of spectator recognition and automated techniques," *INT J HUM-COMPUT ST*, vol. 59, no. 1, pp. 213–225, 2003.
- [89] A. Atkinson, W. Dittrich, A. Gemmell, and A. Young, "Emotion perception from dynamic and static body expressions in point-light and full-light displays," *PERCEPTION*, vol. 33, pp. 717–746, 2004.
- [90] J. Montepare, S. Goldstein, and A. Clausen, "The identification of emotions from gait information," *J NONVERBAL BEHAV*, vol. 11, pp. 33–42, 1987.
- [91] E. Crane and M. Gross, "Motion capture and emotion: Affect detection in whole body movement," in *Int. Conf. ACII*, pp. 95–101, 2007.
- [92] C. Roether, L. Omlor, A. Christensen, and M. Giese, "Critical features for the perception of emotion from gait," *J VISION*, vol. 9, no. 6, pp. 1–32, 2009.
- [93] P. Beesley, "Kinetic architectures and geotextile installations," *Riverside Architectural Press*, 2010.
- [94] K. Takahashi, M. Hosokawa, and M. Hashimoto, "Remarks on designing of emotional movement for simple communication robot," in *IEEE ICIT*, pp. 585–590, 2010.
- [95] C. Busso, Z. Deng, M. Grimm, U. Neumann, and S. Narayanan, "Rigid head motion in expressive speech animation: Analysis and synthesis," *IEEE Audio, Speech, Language Process.*, vol. 15, no. 3, pp. 1075–1086, 2007.
- [96] J. Fast, *Body language*. Pocket, 1988.
- [97] L. Carmichael, S. Roberts, and N. Wessell, "A study of the judgment of manual expression as presented in still and motion pictures," *J SOC PSYCHOL*, vol. 8, no. 1, pp. 115–142, 1937.
- [98] J. Reilly, M.L.McIntire, and H. Seago, "Affective prosody in american sign language," *Sign Language Studies*, vol. 75, pp. 113–28, 1992.
- [99] J. Hietanen, J. Leppänen, and U. Lehtonen, "Perception of emotions in the hand movement quality of finnish sign language," *J NONVERBAL BEHAV*, vol. 28, no. 1, pp. 53–64, 2004.
- [100] M. Sawada, K. Suda, and M. Ishii, "Expression of emotions in dance: Relation between arm movement characteristics and emotion," *PERCEPT MOTOR SKILL*, vol. 97, no. 3, pp. 697–708, 2003.
- [101] A. Samadani, B. DeHart, K. Robinson, D. Kulić, E. Kubica, and R. Gorbet, "A study of human performance in recognizing expressive hand movements," in *IEEE Int. Symp. RO-MAN*, pp. 93–100, 2011.
- [102] A. Samadani, E. Kubica, R. Gorbet, and D. Kulić, "Perception and generation of affective hand movements," *Submitted to Int. Journal of Social Robotics*, vol. 5, no. 1, pp. 35–51, 2013.
- [103] J. Hall, "Gender effects in decoding nonverbal cues," *PSYCHOL BULL*, vol. 85, no. 4, pp. 845–857, 1978.
- [104] J. Hall and D. Matsumoto, "Gender differences in judgments of multiple emotions from facial expressions," *EMOTION*, vol. 4, no. 2, pp. 201–206, 2004.
- [105] G. Kirouac and F. Dore, "Accuracy of the judgment of facial expression of emotions as a function of sex and level of education," *J NONVERBAL BEHAV*, vol. 9, no. 1, pp. 3–7, 1985.
- [106] B. Montagne, R. Kessels, E. Frigerio, E. De Haan, and D. Perrett, "Sex differences in the perception of affective facial expressions: Do men really lack emotional sensitivity?," *COGN PROCESS*, vol. 6, no. 2, pp. 136–141, 2005.

- [107] H. Wagner, C. MacDonald, and A. Manstead, "Communication of individual emotions by spontaneous facial expressions," *J PERS SOC PSYCHO*, vol. 50, no. 4, pp. 737–743, 1986.
- [108] N. Rotter and G. Rotter, "Sex differences in the encoding and decoding of negative facial emotions," *J NONVERBAL BEHAV*, vol. 12, no. 2, pp. 139–148, 1988.
- [109] M. Mandal and S. Palchoudhury, "Perceptual skill in decoding facial affect," *PERCEPT MOTOR SKILL*, vol. 60, no. 1, pp. 96–98, 1985.
- [110] L. Brody, "Gender, emotional expression, and parent-child boundaries," *Emotion: Interdisciplinary perspectives*, pp. 139–170, 1996.
- [111] U. Donges, A. Kersting, and T. Suslow, "Women's greater ability to perceive happy facial emotion automatically: Gender differences in affective priming," *PLoS ONE*, vol. 7, p. e41745, 07 2012.
- [112] A. Sokolov, S. Krüger, P. Enck, I. Krägeloh-Mann, and M. Pavlova, "Gender affects body language reading," *Frontiers in psychology*, vol. 2, 2011.
- [113] K. Johnson, L. McKay, and F. Pollick, "He throws like a girl (but only when he's sad): Emotion affects sex-decoding of biological motion displays," *COGNITION*, vol. 119, pp. 265–280, 2011.
- [114] S. Halovic and C. Kroos, "Facilitating the perception of anger and fear in male and female walkers," in *AISB, Symp. Mental States, Emotions and their Embodiment*, pp. 3–7, 2009.
- [115] Aldebaran-Robotics company. URL: [www.aldebaran-robotics.com](http://www.aldebaran-robotics.com).
- [116] A. Beck, B. Stevens, K. A. Bard, and L. Cañamero, "Emotional body language displayed by artificial agents," *ACM Trans. Interact. Intell. Syst.*, vol. 2, pp. 1–29, Mar. 2012.
- [117] M. Saerbeck and C. Bartneck, "Perception of affect elicited by robot motion," in *ACM/IEEE Int. Conf. HRI*, pp. 53–60, 2010.
- [118] F. Heider and M. Simmel, "An experimental study of apparent behavior," *AM J PSYCHOL*, vol. 57, no. 2, pp. 243–259, 1944.
- [119] J. Hodgins, J. O'Brien, and J. Tumblin, "Perception of human motion with different geometric models," *AM J PSYCHOL*, vol. 4, no. 4, pp. 307–316, 1998.
- [120] R. McDonnell, S. Jörg, J. McHugh, F. Newell, and C. O'Sullivan, "Evaluating the emotional content of human motions on real and virtual characters," in *Symp. APGV*, pp. 67–74, ACM, 2008.
- [121] P. Blythe, P. Todd, and G. Miller, "How motion reveals intention: Categorizing social interactions," pp. 257–285, 1999.
- [122] J. Lee, J. Park, and T. Nam, "Emotional interaction through physical movement," in *Int. Conf. HCI: Intelligent Multimodal Interaction Environments*, pp. 401–410, Springer, 2007.
- [123] A. Samadani, R. Gorbet, and D. Kulić, "Gender differences in the perception of affective movements," in *Accepted for publication in Int. Workshop on Human Behaviour Understanding*, vol. LNCS vol.7559, 2012.
- [124] C. Bartneck, J. Reichenbach, and A. Van Breemen, "In your face, robot! the influence of a character's embodiment on how users perceive its emotional expressions," in *Design and Emotion*, pp. 32–51, 2004.
- [125] A. Powers, S. Kiesler, S. Fussell, and C. Torrey, "Comparing a computer agent with a humanoid robot," in *ACM/IEEE Int. Conf. HRI*, pp. 145–152, 2007.
- [126] C. Darwin, *The expression of emotions in man and animals*. Chicago University Press, 1872.
- [127] W. James, "What is an emotion?," *MIND*, vol. 9, pp. 188–205, 1884.
- [128] R. Picard, *Affective Computing*. MIT Press, 1997.
- [129] M. Nicolaou, H. Gunes, and M. Pantic, "Continuous prediction of spontaneous affect from multiple cues and modalities in valence-arousal space," *IEEE Trans. Affect. Comp.*, vol. 2, pp. 92–105, apr.-jun. 2011.
- [130] T. Meservy, M. Jensen, J. Kruse, J. Burgoon, J.F. Nunamaker, D.P. Twitchell, G. Tschepnakis, and D. Metaxas, "Deception detection through automatic, unobtrusive analysis of nonverbal behavior," *IEEE Intell. Syst.*, vol. 20, pp. 36–43, Sep. 2005.
- [131] L. Kessous, G. Castellano, and G. Caridakis, "Multimodal emotion recognition in speech-based interaction using facial expression, body gesture and acoustic analysis," *J Multimodal User Interfaces*, vol. 3, no. 1, pp. 33–48, 2010.
- [132] D. Glowinski, N. Dael, A. Camurri, G. Volpe, M. Mortillaro, and K. Scherer, "Toward a minimal representation of affective gestures," *IEEE Trans. Affect. Comp.*, vol. 2, no. 2, pp. 106–118, 2011.
- [133] J. F. Grafsgaard, K. E. Boyer, E. N. Wiebe, and J. C. Lester, "Analyzing posture and affect in task-oriented tutoring," in *Int. Florida Art. Intel. Res. Soc. Conf.*, pp. 438–443, 2012.
- [134] D. Glowinski and M. Mancini, "Towards real-time affect detection based on sample entropy analysis of expressive gesture," in *Affective Computing and Intelligent Interaction*, pp. 527–537, Springer, 2011.
- [135] M. Karg, R. Jenke, K. Kühnlenz, and M. Buss, "A two-fold PCA-approach for inter-individual recognition of emotions in natural walking," in *Int. Conf. MLDM*, 2009.
- [136] M. Karg, R. Jenke, W. Seiberl, K. Kühnlenz, A. Schwirtz, and M. Buss, "A comparison of PCA, KPCA and LDA for feature extraction to recognize affect in gait patterns," in *Int. Conf. ACII*, pp. 195–200, 2009.
- [137] L. Omlor and M. Giese, "Extraction of spatio-temporal primitives of emotional body expressions," *NEUROCOMPUTING*, vol. 70, no. 10–12, pp. 1938–1942, 2007.
- [138] D. Amelynck, M. Grachten, L. V. Noorden, and M. Leman, "Toward e-motion-based music retrieval a study of affective gesture recognition," *IEEE Trans. Affect. Comp.*, vol. 3, no. 2, pp. 250–259, 2012.
- [139] S. D'Mello and A. Graesser, "Automatic detection of learner's affect from gross body language," *APPL ARTIF INTELL*, vol. 23, no. 2, pp. 123–150, 2009.
- [140] S. Mota and R. W. Picard, "Automated posture analysis for detecting learner's interest level," in *Conf. CVPRW*, vol. 5, pp. 49–49, IEEE, 2003.
- [141] K. Karpouzis, G. Caridakis, L. Kessous, N. Amir, A. Raouzaoui, L. Malatesta, and S. Kollias, "Modeling naturalistic affective states via facial, vocal, and bodily expressions recognition," in *Artificial Intelligence for Human Computing, LNCS 4451*, pp. 91–112, Springer, 2007.
- [142] A. Kleinsmith, N. Bianchi-Berthouze, and A. Steed, "Automatic recognition of non-acted affective postures," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 41, no. 4, pp. 1027–1038, 2011.
- [143] Y. Ma, H. Paterson, and F. Pollick, "A motion capture library for the study of identity, gender and emotion perception from biological motion," *BEHAV RES METHODS*, vol. 38, no. 1, pp. 134–141, 2006.
- [144] A. Kleinsmith and N. Bianchi-Berthouze, "Recognizing affective dimensions from body posture," in *Int. Conf. ACII*, pp. 48–58, Springer, 2007.
- [145] A. Lim and H. G. Okuno, "Using speech data to recognize emotion in human gait," in *Human Behavior Understanding*, pp. 52–64, Springer, 2012.
- [146] D. Keltner and P. Ekman, *Encyclopedia of Psychology*, vol. 3, ch. Emotion, pp. 162–166. London: Oxford University Press, 2000.
- [147] J. Hall, J. Carter, and T. Horgan, *Gender and Emotion*, ch. 5: Gender Differences in nonverbal communication of emotion, pp. 97–117. Cambridge University Press, 2000.
- [148] E. Murphy-Chutorian and M. M. Trivedi, "Head pose estimation in computer vision: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 4, pp. 607–626, 2009.
- [149] S. Mitra and T. Acharya, "Gesture recognition: A survey," *IEEE Trans. Syst., Man, Cybern. C, Appl., Reviews*, vol. 37, no. 3, pp. 311–324, 2007.
- [150] H. Gunes and M. Pantic, "Dimensional emotion prediction from spontaneous head gestures for interaction with sensitive artificial listeners," in *Int. Conf. on IVA*, pp. 371–377, Springer, 2010.
- [151] H. Monkaresi, R. A. Calvo, and M. S. Hussain, "Automatic natural expression recognition using head movement and skin color features," in *Int. Working Conf. on Advanced Visual Interfaces*, pp. 657–660, ACM, 2012.
- [152] M. Schroder, E. Bevacqua, R. Cowie, F. Eyben, H. Gunes, D. Heylen, M. ter Maat, G. McKeown, S. Pammi, M. Pantic, C. Pelachaud, B. Schuller, E. de Sevin, M. Valstar, and M. Wollmer, "Building autonomous sensitive artificial listeners," *IEEE Trans. Affect. Comp.*, vol. 3, pp. 165–183, apr.-jun. 2012.
- [153] S. Scherer, M. Glodek, G. Layher, M. Schels, M. Schmidt, T. Brosch, S. Tschechne, F. Schwenker, H. Neumann, and G. Palm, "A generic framework for the inference of user states in human computer interaction," *J Multimodal User Interfaces*, vol. 6, no. 3–4, pp. 117–141, 2012.
- [154] A. Camurri, B. Mazzarino, and G. Volpe, "Analysis of expressive gesture: The eyesweb expressive gesture processing library,"

- in *Gesture-based Communication in Human-Computer Interaction, LNAI 2915*, pp. 460–467, Springer, 2004.
- [155] S. Hyniewski, R. Niewiadomski, M. Mancini, and C. Pelachaud, *Blueprint for affective computing*, ch. 5.1: Expression of affects in embodied conversational agents, pp. 213–221. Oxford University Press, 2010.
- [156] C. Breazeal, *Designing Sociable Robots*. MIT Press, 2002.
- [157] K. Kühnlenz, S. Sosnowski, and M. Buss, *Human-Robot Interaction*, ch. 12: Evaluating Emotion Expressing Robots in Affective Space, pp. 235–245. I-Tech Education and Publishing, 2007.
- [158] M. Masuda and S. Kato, “Motion rendering system for emotion expression of human form robots based on laban movement analysis,” in *IEEE Int. Symp. RO-MAN*, pp. 344–349, 2010.
- [159] Y. Zhang, Q. Ji, Z. Zhu, and B. Yi, “Dynamic facial expression analysis and synthesis with MPEG-4 facial animation parameters,” *EEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 10, pp. 1383–1396, 2008.
- [160] M. Zecca, Y. Mizoguchi, K. Endo, F. Iida, Y. Kawabata, N. Endo, K. Itoh, and A. Takanishi, “Whole body emotion expression for KOBAN humanoid robot - preliminary experiments with different emotional patterns,” in *IEEE Int. Symp. RO-MAN*, pp. 381–386, 2009.
- [161] A. Beck, L. Canamero, and K. Bard, “Towards an affect space for robots to display emotional body language,” in *IEEE Int. Symp. RO-MAN*, pp. 491–496, 2010.
- [162] M. Karg, M. Schwimbeck, K. Kühnlenz, and M. Buss, “Towards mapping emotive gait patterns from human to robot,” in *IEEE Int. Symp. RO-MAN*, pp. 258–263, sept. 2010.
- [163] A. Beck, A. Hiolle, A. Mazel, and L. Cañamero, “Interpretation of emotional body language displayed by robots,” in *Int. Workshop AFFINE*, pp. 37–42, 2010.
- [164] J. Lasseter, “Principles of traditional animation applied to 3D computer animation,” in *ACM Int. Conf. SIGGRAPH*, pp. 35–44, 1987.
- [165] F. Thomas and O. Johnston, *Disney Animation: The Illusion of Life*. Abbeville Press, New York, 1981.
- [166] I. V. Kerlow, *The art of 3D: computer animation and effects*. John Wiley & Sons, 2004.
- [167] J. Bates, “The role of emotion in believable agents,” *COMMUN ACM*, vol. 37, pp. 122–125, Jul. 1994.
- [168] B. Blumberg, *Biologically inspired intelligent robots*, ch. 3: Biometric Animated Creatures, pp. 47–71. Spie Press, 2003.
- [169] M. Mancini, G. Castellano, C. Peters, and P. W. McOwan, “Evaluating the communication of emotion via expressive gesture copying behaviour in an embodied humanoid agent,” in *Int. Conf. ACII*, pp. 215–224, 2011.
- [170] G. Castellano and M. Mancini, “Analysis of emotional gestures for the generation of expressive copying behaviour in an embodied agent,” in *Gesture-Based Human-Computer Interaction and Simulation, LNCS 5085*, pp. 193–198, Springer, 2009.
- [171] K. Amaya, A. Bruderlin, and T. Calvert, “Emotion from motion,” in *Conf. on Graphics Interface '96, GI '96*, pp. 222–229, 1996.
- [172] M. Unuma, K. Anjyo, and R. Takeuchi, “Fourier principles for emotion-based human figure animation,” in *ACM Int. Conf. SIGGRAPH*, pp. 91–96, ACM, 1995.
- [173] C. Rose, M. F. Cohen, and B. Bodenheimer, “Verbs and adverbs: Multidimensional motion interpolation,” *IEEE Comput. Graph. Appl.*, vol. 18, no. 5, pp. 32–40, 1998.
- [174] M. Brand and A. Hertzmann, “Style machines,” in *Conf. SIGGRAPH*, pp. 183–192, ACM, 2000.
- [175] C. K. Liu, A. Hertzmann, and Z. Popović, “Learning physics-based motion style with nonlinear inverse optimization,” *ACM T GRAPHIC*, vol. 24, no. 3, pp. 1071–1081, 2005.
- [176] S. Kopp, B. Krenn, S. Marsella, A. N. Marshall, C. Pelachaud, H. Pirker, K. R. Thorisson, and H. Vilhjalmsón, “Towards a common framework for multimodal generation: The behavior markup language,” in *Int. Conf. on IVA*, pp. 21–23, 2006.
- [177] M. Schröder, H. Pirker, M. Lamolle, F. Burkhardt, C. Peter, and E. Zovato, “Representing emotions and related states in technological systems,” in *Emotion-Oriented Systems*, pp. 369–387, Springer, 2011.
- [178] G. Castellano, M. Mancini, C. Peters, and P. McOwan, “Expressive copying behavior for social agents: A perceptual analysis,” *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 42, no. 3, pp. 776–783, 2012.
- [179] K. Nakagawa, K. Shinozawa, H. Ishiguro, T. Akimoto, and N. Hagita, “Motion modification method to control affective nuances for robots,” in *Int. Conf. IROS*, pp. 5003–5008, 2009.
- [180] E. Barakova and T. Lourens, “Expressing and interpreting emotional movements in social games with robots,” *PERS UBIQUIT COMPUT*, vol. 14, pp. 457–467, 2010.
- [181] M. Masuda, S. Kato, and H. Itoh, “Emotion detection from body motion of human form robot based on laban movement analysis,” in *Int. Conf. PRIMA, LNAI 5925*, pp. 322–334, 2009.
- [182] T. Matsumaru, “Discrimination of emotion from movement and addition of emotion in movement to improve human-coexistence robot’s personal affinity,” in *IEEE Int. Symp. RO-MAN*, pp. 387–394, 2009.
- [183] C. Roether, L. Omlor, and M. Giese, “Lateral asymmetry of bodily emotion expression,” *CURR BIOL*, vol. 18, no. 9, pp. 329–330, 2008.
- [184] I. Lagerlöf and M. Djerf, “Children’s understanding of emotion in dance,” *EUR J DEV PSYCHOL*, vol. 6, no. 4, pp. 409–431, 2009.
- [185] C. L. Bethel and R. R. Murphy, “Review of human studies methods in hri and recommendations,” *International Journal of Social Robotics*, vol. 2, no. 4, pp. 347–359, 2010.
- [186] H. Griffin, M. Aung, B. Romera-Paredes, C. McLoughlin, G. McKeown, W. Curran, and N. Bianchi-Berthouze, “Laughter type recognition from whole body motion,” in *Int. Conf. ACII*, pp. 349–355, 2013.
- [187] J. R. Fontaine, K. Scherer, and E. R. P. Ellsw, “The world of emotion is not two-dimensional,” *PSYCHOL SCI*, vol. 18, pp. 1050–1057, 2007.
- [188] G. McKeown, M. F. Valstar, R. Cowie, and M. Pantic, “The semaine corpus of emotionally coloured character interactions,” in *Proc. IEEE ICME*, pp. 1079–1084, 2010.
- [189] J. Cappella and J. Greene, “A discrepancy-arousal explanation of mutual influence in expressive behavior for adult-adult and infant-adult dyadic interaction,” *COMMUN MONOGR*, vol. 49, pp. 89–114, 1982.
- [190] K. Floyd, *Communicating affection: Interpersonal behavior and social context*. Cambridge University Press, 2006.
- [191] L. K. Guerrero and K. Floyd, *Nonverbal communication in close relationships*. Lawrence Erlbaum Associates, Publishers, 2006.
- [192] A. Pentland, “Social signal processing,” *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 108–111, 2007.
- [193] A. Vinciarelli, M. Pantic, D. Heylen, C. Pelachaud, I. Poggi, F. D’Errico, and M. Schröder, “Bridging the gap between social animal and unsocial machine: A survey of social signal processing,” *IEEE Trans. Affect. Comp.*, vol. 3, no. 1, pp. 69–87, 2012.
- [194] G. Varni, G. Volpe, and A. Camurri, “A system for real-time multimodal analysis of nonverbal affective social interaction in user-centric media,” *IEEE Trans. on Multimedia*, vol. 12, no. 6, pp. 576–590, 2010.

**Michelle Karg** is a post-doctoral fellow at the Electrical & Computer Engineering Department at the University of Waterloo. Her research interests are machine learning for human movement analysis applied to affective computing, rehabilitation, and clinical movement analysis.

**Ali-Akbar Samadani** is a Ph.D Candidate at the University of Waterloo’s Electrical and Computer Engineering Department. His research interests are affective computing, human movement recognition and generation, and human-machine interaction.

**Rob Gorbet** is an Associate Professor in the University of Waterloo’s Centre for Knowledge Integration. His research interests are in the design and creation of interactive artistic installations, identification and display of affect, and interdisciplinary collaboration.

**Kolja Kühnlenz** is a Senior Researcher with the Institute of Automatic Control Engineering, Technische Universität München. His research interests include robotics, robot vision, attention, visual servoing, social robotics, and emotions.

**Dr. Jesse Hoey** is an associate professor in the Cheriton School of Computer Science at the University of Waterloo. He is also an adjunct scientist at the Toronto Rehabilitation Institute where he is co-leader of the AI and Robotics Research Team. Dr. Hoey’s research is on artificial intelligence and health informatics.

**Dana Kulić** is an Assistant Professor with the Department of Electrical and Computer Engineering, University of Waterloo, Canada. Her research interests include robot learning, humanoid robots and human motion analysis.