Towards Affective Agent Action: Modelling Expressive ECA Gestures

Björn Hartmann
Stanford University Computer Science Dept.
Stanford, CA 94305, USA
bjoern @cs.stanford.edu

Maurizio Mancini
IUT Montreuil, University of Paris-8
Montreuil, France
m.mancini @iut.univ-paris8.fr

Catherine Pelachaud
IUT Montreuil, University of Paris-8
Montreuil, France
c.pelachaud @iut.univ-paris8.fr

ABSTRACT

To enable individualized displays of personality and emotion in Embodied Conversational Agents (ECAs), a generic agent architecture is augmented to generate variable idiosyncratic gesturing behaviors. A set of dimensions of expressivity that characterize individual variability is proposed along with a mapping of the identified dimensions onto low-level animation parameters. Gesture synthesis is modified at multiple planning stages. Semantic information about the structure and communicative function of the behaviors is taken into account to guide modifications. The implementation is tested in two evaluation studies with large groups of non-expert users.

Keywords

Affective systems; Agent-based interfaces; Procedural animation

INTRODUCTION

Embodied Conversational Agents (ECAs) are a powerful user interface paradigm, aiming to transfer the inherent richness of human-human interaction to human-computer interaction. ECAs are virtual embodied representations of humans that communicate multimodally with the user through voice, facial expression, gaze, gesture, and body movement. Effectiveness of an agent is dependent on her ability to suspend the user's disbelief during an interaction. To increase believability and life-likeness of an agent, she has to express emotion and exhibit personality in a consistent manner [22]. Human individuals differ not only in their reasoning, their set of beliefs, goals, and their emotional states, but also in their way of expressing such information through the execution of specific behaviors. To replicate this variability in the actions of agents, we move away from a generic action model and simulate individualized agents that portray idiosyncratic behaviors. Specifically, in this paper we describe and evaluate a new method that augments an existing gesture synthesis architecture to achieve parameterized, *expressive* actions – actions that can reflect internal and external influences and a personal communicative style.

Until now most ECA systems have concentrated solely on defining computational models of behavior selection – *which* behaviors to choose for a given communicative act. We now investigate the qualitative aspects of coverbal behavior – *how* people differ in their ways of performing behaviors that accompany acts of speech. High-level agent functions such as emotion, personality, culture, role and gender can modify actions in complex and competing ways. Since the nature of these influences is not well understood, we restrict our attention to generating phenomenologically accurate behaviors without claiming to correctly represent internal processes (cf. Nass et al. [27]).

We make two contribution in this paper. First, we develop an intermediate level of behavior parametrization – a set of *dimensions of expressivity* based on a survey of psychology literature. We regard expressivity parameters as useful enabling tools to mediate between holistic, qualitative communicative goals and low-level animation parameters (see Fig. 1). Second, we describe an implementation that modifies hand-arm gesture according to the expressivity parameters in a *semantically sensitive* fashion: information about the structure of a gesture, its communicative function and its relation to neighboring gestures within a longer utterance is taken into account to guide expressivity modifications.

BACKGROUND: HUMAN EXPRESSIVITY

Our approach to capturing bodily expressivity is driven by a perceptual standpoint – how expressivity is perceived by oth-

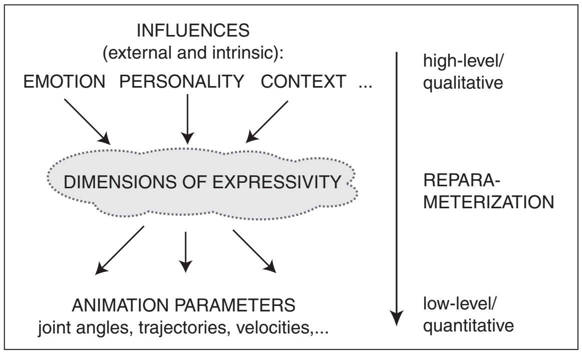

Figure 1: Tri-level mapping of high level agent functions via dimension of expressivity to low-level animation parameters.

ers. We do not attempt to model internals such as muscle activation patterns that underly these outwardly visible signals. Researchers in social psychology have investigated how various influences affect perceived bodily behaviors. We review some of the most pertinent studies below.

Wallbot and Scherer [36] had judges encode their impressions of behavior along the following five categories: slow/fast, small/expansive, weak/energetic, small movement activity/large movement activity, and unpleasant/pleasant. In a later study, Wallbot [35] found that besides static pose configurations, three dynamic dimensions could be reliably identified by observers: amount of movement activity, expansiveness/spatial extension, and movement dynamics/energy/power. Gallaher [12] found four significant dimensions of variability in personal encoding style: expressiveness - energetic communication; animation - energy in acts not directly related to communication; expansiveness - use of space, elbow position; and coordination - smoothness, fluidity. Ball and Breese [2] outline Collier's finding on correlations between temporal and spatial tendencies in gesture/posture and personality/emotion – movement frequency and speed were related to emotional arousal, as was the size of overall body outline. A summary of Pollick's research [31] points out limits of dissecting movement features and ascribing discrete values to them: the degree and manner in which this style is dependent on spatial and temporal encoding is not trivial and varies between different movements. The importance of considering human posture as a communication channel has been stressed as well [26, 14] but we do not yet integrate these findings in our own work.

While we base our method on the studies above, related work has used methods from the field of human movement observation within the dance community. The most prominent system of notation is Laban movement analysis (kinetography) [20]. Laban uses five dimensions of classification: Body, Space, Shape, Effort, and Relationship. Each dimension is further subdivided into a set of parameters, e.g., four each for the Effort and Shape dimensions. Spatial extension (Space) is captured as well as movement information (Shape) and intention (Effort).Other major approaches to movement notation include the Benesh system [3] and the Eshkol-Wachman system [11]. The former is used mostly for ballet and is better suited to represent acted, planned forms of movements, not spontaneous nonverbal behavior. The latter lacks the expressive power to describe execution manner and subtle temporal features – precisely what we are interested in.

STATE OF THE ART

We now turn to a review of existing research in Embodied Conversational Agents and expressive behaviors of virtual agents in particular. Research that has particularly informed our current approach is highlighted.

Expressive ECAs

An overview of recent ECA implementations can be found in [7, 32]. Cassell et al. presented an early implementation of multimodal communicative agents in [6]. Gesture production as a subsystem of agent behavior generation has been a focus of investigation in recent years as well. We can distinguish between systems addressing the problem of gesture selection (*which* gestures to execute) and gesture animation (*how* to execute them).

Selection algorithms have mostly been concerned with semantic aspects of gesturing, often using McNeill's method of classification [25] into *beats* – rhythmic emphases; *iconics* – gestures that refer to concrete objects or properties in the physical world; *metaphorics* – gestures for abstract concepts; *emblems* – culturally specific symbols; and *deictics* – pointing movements. Cassell et al. select suitable nonverbal behaviors to accompany user-supplied text based on a linguistic analysis [8]. Noot and Ruttkay approach intersubject variability in GESTYLE [29] by choosing between atomic behaviors based on 'style dictionaries.'

Gesture animation is concerned with realistic movement generation of an agent's upper limbs based on computer graphics and computer animation techniques such as hierarchical modelling of skeletal joint chains and inverse kinematics solvers [21]. Animation systems often introduce a custom representation language to describe sequential and parallel components of gestures [19, 15]. Fiume and Neff have presented a convincing simulation of muscle tension to add expressivity to pre-existing actions [28]. Their system also controls animation of the spine, but their dynamics simulation is computationally too expensive for our near-realtime requirement.

EMOTE

EMOTE by Chi et al. [9] is most closely related to our work as it also introduces an intermediate level of parametrization to obtain expressive gestures. EMOTE implements Laban principles of Effort and Shape to render gesture performances more expressive by varying kinematic and spatial spline properties of existing keyframed actions. It is being attempted to tie the EMOTE system to OCC and OCEAN models of emotion and personality [1]. We do not seek to surpass the fine-grained control of EMOTE's parameter set. However, EMOTE can only modify – not generate – gestures. By implementing EMOTE as a generic filter on pre-existing behaviors, a wide variety of motions can be manipulated. But important information about the semantic structure of the behaviors can no longer be used to guide adaptations. EMOTE thus applies transformation at the level of limb articulation without taking the type of gesture that that limb configuration represents into account. To overcome this problem, EMOTE allows phrasing control of its parameters; however, establishing the correspondence between keyframes and underlying movement structure and meaning must be established manually.

On the other hand, in our approach, we rely on psychology literature to obtain a set of expressivity parameters. We uphold the link between gesture modification and the Cassell/McNeill approach to gesture classification. We thus incorporate expressivity calculation into the *synthesis* of gesturing behavior and can use the high-level semantics available during gesture planning to guide and constrain adaptations. In addition, we can not only modify gestures at a low animation spline level, but can also control higher levels such as repetition of gesture phases or suppression of entire gestures. A final point of divergence is related to the evaluation methodology. Evaluation of EMOTE was carried out by a small number of trained observers. We use a large number of untrained subjects.

Figure 2: Our GRETA agent.

EXPRESSIVE AGENT ARCHITECTURE

GRETA, our multimodal agent (see Fig. 2), interprets utterance text marked up with communicative functions [10] to generate synchronized speech, face, gaze and gesture animations. Text-to-speech conversion is accomplished through Festival [5]. The timing output from speech synthesis is used to set coincidence constraints for separate engines modelling face/gaze and gesture animation. Both engines produce animation data in MPEG4-compliant FAP/BAP format [33], which in turn drive a facial and skeletal body model in OpenGL. A detailed description of the architecture can be found in [30, 15, 4].

To enable the thus-far generic, deterministic agent for expressivity control, we augmented the existing architecture (see Fig. 3). This paper is concerned only with the implementation of the bottom part of the diagram – the animation engine that instantiates behaviors according to given expressivity settings. Another research group is currently working on constructing the *Expressivity Specification Module* that supplies the text markup used as input for our animation system [24]. We therefore mention the involved left-hand processes only briefly: a global agent definition structure was added which determines individual preferences in the expressive dimensions as well as a priority queue indicating a general preference of modality use (face, gaze, hand and body gesture). These static agent definition variables seek to capture the net outcome on behavior a particular agent personality may have without having to model the personality structure itself. The Expressivity Specification Module also adds local tag level information carrying expressivity values for each communicative function tag.

Set of Expressivity Attributes

Based on the aggregate evidence of the surveyed studies in section , we propose to capture expressivity with a set of six attributes which we describe below in qualitative terms. As part of an individualized agent's definition, personal default values for the expressivity attributes are defined. These values can be supplanted by including expressivity information in communicative function tags within our markup language.

- *Overall Activation*: quantity of movement during a conversational turn (e.g., passive/static or animated/engaged).
- *Spatial Extent*: amplitude of movements (e.g., amount of space taken up by body)
- *Temporal Extent*: duration of movements (e.g., quick versus sustained actions)
- *Fluidity*: smoothness and continuity of overall movement (e.g., smooth/graceful versus sudden/jerky)
- *Power*: dynamic properties of the movement (e.g., weak/relaxed versus strong/tense)
- *Repetition*: tendency to rhythmic repeats of specific movements.

Each of the attributes is float-valued and defined over the interval $[-1, 1]$, where the zero point corresponds to the actions our generic agent without expressivity control would perform. Overall activation is singly float-valued and ranges from 0 to 1, where 0 corresponds to a complete absence of nonverbal behavior. Overall activation, Fluidity and Power

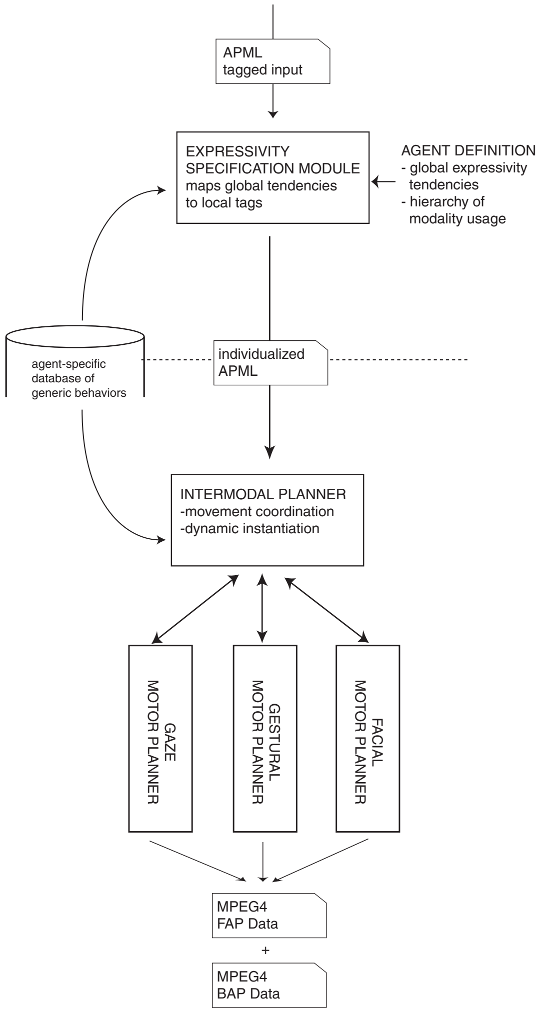

Figure 3: Agent architecture outline.

act on the entire agent animation calculated for a conversational turn, while the other parameters generate only local changes specific to one communicative act. Though not represented in our literature review, the concept of repetition was added from its appearance in an annotated gesture corpus [23].

MAPPING EXPRESSIVITY INTO GESTURE ANIMATION PARAMETERS

Given a particular type of action and a set of values in the expressive space, how can we modify non-verbal behavior production to communicate the appropriate expressive content? Based on an analysis of the communicative functions contained in the marked up input text, the gesture system chooses a matching prototype gesture. Subsequently, we dynamically instantiate the gesture according to a set of expressivity attributes and synchrony constraints with speech. To do so, we need a suitable representation of gestures, which is outlined first. Next, our implementation for gesture instantiation and modification is described. It is our primary aim to preserve the semantic value of each gesture during the expressivity modifications. Effective strategies have to adjust behavior on multiple levels - from abstract planning (whether to search for a gesture for a given text at all), to gesture phase-level modifications (whether or not to repeat a stroke), down to adjusting velocity profiles of key pose transitions.

In the following, let the variables oac, spc, tmp, flt, pwr and rep stand for the *overall activation*, *spatial extent*, *temporal extent*, *fluidity*, *power* and *repetition* values we are trying to express.

Augmented gesture specification language

In the past, we devised an abstract keyframe based scheme for gesture synthesis [15]. The gesture specification language is a sequence of key poses of the action with regard to wrist location, palm orientation and hand shape. Sets of key poses are grouped into the gesture phases defined by McNeill [25]. Our specification language was augmented by attributed defining which features of a gesture carry its semantic meaning and are thus invariable, and which features can be modulated to add expressivity. Conservation of salient features is especially important for *iconic* gestures, which refer to concrete objects or processes in the real world; and it is essential for *emblems*, which depend on exact reproduction. We hypothesize that *metaphorics* and *beat* gestures can be modified to a greater degree. Furthermore, an explicit description of the temporal aspect of each gesture was added. Where previously, kinematics were implied through the timing of the key frames, timing is now calculated using motion functions and the specification language only expresses qualitative velocities for frame-to-frame transitions such as *speed=(slow—normal—fast)*.

Overall Activation

A filtering is applied at the level of the intermodal planner, which assigns gesture prototypes to input text mark up tags. Each input tag carries a summary weight attribute that captures how important stressing the tag's content through nonverbal signals is. Communicative functions tags for which this activation attribute does not surpass a given agent's overall activation threshold are not matched against the behavior database and thus no nonverbal behavior is generated at all. A similar principle of activity filtering was presented and implemented by Cassell et al. in [8].

Spatial Extent

The space in front of the agent that is used for gesturing is represented as a set of sectors following McNeill's diagram [25]. We expand or condense this entire space through scaling. Wrist positions in our gesture language are defined in terms of these sectors. Represented by their center coordinates, the location of the sectors can be scaled asymmetrically using a simple scaling matrix on homogenous coordinates. For meaningful scaling, we establish sector center coordinates $\vec{p_i}$ relative to the agent's solar plexus. Then the modified sector centers are given by:

$$\vec{p_i'} = \begin{bmatrix} I & \vec{spc} \\ 0 & 1 \end{bmatrix} \cdot \vec{p_i}$$

with:

$$\vec{spc} = \begin{pmatrix} 1.0 + spc \cdot spc_{agent_{horiz}} \\ 1.0 + spc \cdot spc_{agent_{vert}} \\ 1.0 + spc \cdot spc_{agent_{front}} \end{pmatrix}$$

$spc_{agent_{horiz}}$, $spc_{agent_{vert}}$, and $spc_{agent_{front}}$ are individual scaling factors in the horizontal, vertical and frontal directions that can define individualized patterns of space use. To find the location of articulation for a gesture, we compute a point in the dynamically resized gesture quadrant set in the gesture definition file and then calculate joint angles needed to reach that target with the IKAN inverse kinematics package [34]. Note that this technique is conceptually similar to EMOTE's kinematic reach space. While inverse kinematics are computationally expensive, they provide the only way of addressing arm movement in terms of goal positions. In a complex articulated joint chain such as a human arm, changing forward kinematics (i.e., joint angles) directly yields non-linear and unpredictable results.

Adjusting the elbow *swivel angle* (Tolani [34]) also directly changes the space taken up by the agent - extended elbows enlarge the body's silhouette. Since we are using inverse kinematics to position the wrist, we can control each arm's swivel angle θ for every key position:

$$\theta' = \left\{ \begin{array}{ll} \min(\theta \cdot (1.0 + 0.5 \cdot spc), \pi/2) & spc \geq 0 \\ \max(\theta \cdot (1.0 + 0.5 \cdot spc), 0) & spc < 0 \end{array} \right.$$

Temporal Extent

Starting from the synchronicity constraint on the end of the gesture stroke to coincide with the stressed affiliate in speech (McNeill [25]), we can calculate preceding and proceeding frame times from invariant laws of human arm movement described in [13]. The gesture specification language includes a high-level notion of how quickly the gesture phases should be performed. During the planning phase, the actual distance travelled by the wrist joint in space is approximated by linear segments through key points. The duration to complete each segment can be derived from a simplification of Fitt's law as

$$T = a + b \cdot \log_2(\|\vec{x_n} - \vec{x_{n+1}}\| + 1)$$

The value of the velocity coefficient b has been established as 10^{-1} for average speed movements by Kopp. Using this value as a starting point, the speed of a gesture segment can be adjusted as follows:

$$b = (1 + 0.2 \cdot tmp) \cdot 10^{-1}$$

Since we still have information about which part of the movement corresponds to which gesture phase, we can for example selectively amplify the stress of the gesture by increasing only the speed of the stroke to accentuate the gesture. We note that Fitt's law only applies to linear movements. Approximating the arm trajectory with piecewise linear segments is a first draft. In the future we will have to model additional duration functions (e.g., the two-thirds power law [13]) for curved movements.

Fluidity

This concept seeks to capture the smoothness of single gestures as well as the continuity between movements (the inter-gestural rest phases). We achieve low-level kinematic control through varying the continuity parameter of Kochanek-Bartels splines [18] for the kinematic interpolants. Once again, this idea is close to EMOTE timing and fluidity control. In our implementation, we set the continuity parameter $cont$ of the spline of the position interpolation spline for the wrist end-effector of each arm to equal the fluidity setting: $cont = flt$.

On a higher planning level, larger fluidity increases the minimum timing threshold for retracting arms to a neutral position on the sides of the torso in between two gestures. During below-threshold pauses, arms are not retracted; instead two neighboring gestures are directly connected by interpolating between the retraction position of a previous gesture and the preparation position of the following gesture.

Power

To visualize the amount of energy and tension invested into a movement, we once again look at the dynamic properties of gestures. Powerful movements are expected to have higher acceleration and deceleration magnitudes. This behavior is modelled with the tension and bias parameters of the kinematic TCB-spline: $tension = pwr$ and $bias = pwr$.

We also hypothesize that tense, powerful performances will be characterized by different hand shapes. If the configuration of the hand is not indicated as fixed in the gesture specification, high power settings will contract the hand towards a fist shape.

Finally, power also influences the retraction magnitude for repetitive gestures (see below). The parameters controls how far the arm travels from the stroke end back to the next stroke start for each repetition: $retraction_{mag} = 0.6 + 0.3 \cdot pwr$.

We are aware of the possibly conflicting definitions of Power and Fluidity and have not yet fully addressed this problem. Composition of two functions that modify the same underlying variables carries with it the risk of losing discrete meanings.

Repetition

We have previously introduced the technique of stroke expansion [15] to capture coarticulation/superposition of beats onto other gestures. Stroke expansion repeats the meaning-carrying movement of a gesture so that successive stroke ends fall onto the stressed parts of speech following the original gesture affiliate. It is possible to control the extent of repetition by selectively increasing the 'horizon' or lookahead distance that the stroke repetition algorithm analyzes beyond the original enclosing rheme tag of an utterance. In the original implementation, proceeding *Hstar* pitch accents tags were looked for within the *rheme* clause that introduces the new information linked to the original gesture. The rheme clause has a known duration t_{rheme}. For $rep = 0$, we keep the default algorithm. For $rep = 0$, no expansion is attempted. For $(-1 < rep < 0)$, we only look into the span from the beginning of the rheme section to $\frac{1}{1-rep} \cdot t_{rheme}$ for stroke expansion matches. For $(0 < rep \leq 1)$, we push the lookahead horizon back to $(x \cdot (1 + rep)) \cdot t_{rheme}$, where $x > 1$. A good value for x remains to be established.

EXAMPLE

We show how setting parameters can generate a qualitatively different animation. Our system represents only a building block towards realizing affective action - exactly how motion quality is changed by the emotional state of an actor is still an open question in experimental psychology. Wallbott [] had progressed the most before his untimely death. For now, we use qualitative labels that are neutral with respect to emotion and personality, such as "abrupt." "Abrupt" action is realized by modifying "neutral" action with the following parameters (derived from intuition with added trial-and-error fine tuning): $oac = 0.6$, $spc = 0$, $tmp = 1$, $flt = -1$, $pwr = 1$, $rep = -1$.

EVALUATION

We can and must distinguish two different levels of evaluation of our system, corresponding to the two different mappings outlined in Figure 1. First, we can test whether a given implementation is effective in realizing the behavioral changes implied by each parameter of expressivity. Second, we can test how setting multiple parameters to reflect an expressive intent affect overall believability of our agent. However, the second type of evaluation has to rely on the underlying implementation - thus we first conducted a test of individual parameters. Specifically, we evaluated the following hypothesis: *The chosen implementation for mapping single dimensions of expressivity onto animation parameters can be recognized and correctly attributed by users.* 52 subjects were asked to identify a single dimension and direction of change in forced-choice comparisons between pairs of animations.

In the first test, in 41.3% of cases the correct parameter was identified. The dimensions $spatialextent$ and $temporalextent$ had the highest recognition rates - 72% and 73% of users correctly attributed a change in expression to the respective dimension. FINISH

The second test with 54 subjects was conducted as a preference ranking task with four options per trial to test the following hypothesis: *Combining parameters in such a way that they reflect a given communicative intent will result in more believable overall impression of the agent.* FINISH

INSERT HERE: summary of interpretation of results. Details of the evaluation studies can be found in [16].

FUTURE WORK

We have presented an approach to introduce meaningful variability into gesture synthesis for ECAs. As our evaluation results indicate, a subset of expressivity parameters and their combinations is well recognized by users. However, some parameter implementations need to be refined.

Furthermore, we still have to find a systematic way to account for interaction effects between parameters. Since the parameters are not all independent of each other, complex combinations can result in incoherent animations after all expressivity modifications are applied. We currently lack good negotiation strategies for such cases.

We also acknowledge that our current model only introduces a limited amount of variability into behavior execution that cannot capture the wide range of human behavior adaptations. To allow for the possibility of choosing between altogether different representations of a given communicative function, integration between gesture selection and gesture modification is necessary.

The unidirectional top-down processing from abstract goals to concrete animation parameters is a simplification of actual human behavior. Bottom-up theories also exist which posit that behavior causes change in emotional state [17]. In between these extreme points, feedback mechanisms may guide adaptation during performance within the human motor system.

Much work is needed on intermodal synchrony and arbitration when expressivity computations are expanded beyond gestures to include face and gaze control. At the present moment, modalities are synchronized only through their relationship to speech time constraints. The engines themselves calculate behaviors independently. It also remains to be resolved whether our chosen expressivity mappings exhibit uniform meaning for high level functions across modalities, e.g., if a gesture with increased spatial usage carries a similar emotional message as a facial display with increased amplitude.

ACKNOWLEDGMENTS

Part of this research is supported by the EU FP6 Network of Excellence "HUMAINE", IST contract 507422.

REFERENCES

1. J. Allbeck and N. Badler. Toward representing agent behaviors modified by personality and emotion. In *Embodied Conversational Agents at AAMAS'02*. ACM Press, 2002.

2. G. Ball and J. Breese. Emotion and personality in a conversational agent. In S. P. J. Cassell, J. Sullivan and

E. Churchill, editors, *Embodied Conversational Characters*. MITpress, Cambridge, MA, 2000.

3. R. Benesh and J. Benesh. *Reading Dance, The Birth of Choreology*. Souvenir Press (E and A) Ltd., 1977.
4. E. Bevacqua and C. Pelachaud. Modelling an italian talking head. In *Auditory-Visual Speech Processing AVSP'03*, Annecy, France, 2003.
5. A. Black, P. Taylor, R. Caley, and R. Clark. Festival. http://www.cstr.ed.ac.uk/projects/festival/.
6. J. Cassell, C. Pelachaud, N. Badler, M. Steedman, B. Achorn, T. Becket, B. Douville, S. Prevost, and M. Stone. Animated conversation: rule-based generation of facial expression, gesture & spoken intonation for multiple conversational agents. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, pages 413–420. ACM Press, 1994.
7. J. Cassell, J. Sullivan, S. Prevost, and E. Churchill, editors. *Embodied Conversational Agents*. MIT Press, Cambridge, MA, 2000.
8. J. Cassell, H. H. Vilhjálmsson, and T. Bickmore. Beat: the behavior expression animation toolkit. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 477–486. ACM Press, 2001.
9. D. Chi, M. Costa, L. Zhao, and N. Badler. The EMOTE model for effort and shape. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 173–182. ACM Press/Addison-Wesley Publishing Co., 2000.
10. B. DeCarolis, C. Pelachaud, I. Poggi, and M. Steedman. APML, a mark-up language for believable behavior generation. In H. Prendinger and M. Ishizuka, editors, *Life-Like Characters*, Cognitive Technologies, pages –. Springer, 2004.
11. N. Eshkol and A. Wachman. *Movement Notation*. Widenfeld and Nicolson, London, 1958.
12. P. E. Gallaher. Individual differences in nonverbal behavior: Dimensions of style. *Journal of Personality and Social Psychology*, 63(1):133–145, 1992.
13. S. Gibet, J.-F. Kamp, and F. Poirier. Gesture analysis: Invariant laws in movement. In A. Camurri and G. Volpe, editors, *Gesture-Based Communication in Human-Computer Interaction - GW 2003*, number 2915 in LNAI, pages 1–9. Springer, 2004.
14. J. A. Harrigan. Listener's body movements and speaking turns. *Communication Research*, 12(2):233–250, April 1985.
15. B. Hartmann, M. Mancini, and C. Pelachaud. Formational parameters and adaptive prototype instantiation for mpeg-4 compliant gesture synthesis. In *Proceedings of the Computer Animation 2002*, page 111. IEEE Computer Society, 2002.
16. B. Hartmann, M. Mancini, and C. Pelachaud. Design and evaluation of expressive gesture synthesis for embodied conversational agents. *Submitted for publication*, 2005.
17. W. James. *The Principles of Psychology*. Henry Holt, New York, 1890.
18. D. H. U. Kochanek and R. H. Bartels. Interpolating splines with local tension, continuity, and bias control. In H. Christiansen, editor, *Computer Graphics (SIGGRAPH '84 Proceedings)*, volume 18, pages 33–41, 1984.
19. S. Kopp and I. Wachsmuth. A knowledge-based approach for lifelike gesture animation. In W. Horn, editor, *ECAI 2000 Proceedings of the 14th European Conference on Artificial Intelligence*. IOS Press, 2000.
20. R. Laban and F. Lawrence. *Effort: Economy in body movement*. Plays, Inc, Boston, 1974.
21. T. Lebourque and S. Gibet. High level specification and control of communication gestures: The gessyca system. In *Proceedings of the Computer Animation 1999*, page 24. IEEE Computer Society, 1999.
22. A. B. Loyall and J. Bates. Personality-rich believable agents that use language. In W. L. Johnson and B. Hayes-Roth, editors, *Proceedings of the First International Conference on Autonomous Agents (Agents'97)*, pages 106–113, Marina del Rey, CA, USA, 1997. ACM Press.
23. C. Martell, P. Howard, C. Osborn, L. Britt, and K. Myers. FORM2 kinematic gesture corpus. Video recording and annotation, 2003.
24. V. Maya, M. Lamolle, and C. Pelachaud. Influences on embodied conversational agent's expressivity: Towards an individualization of the ecas. In *Proceedings of AISB*, Leeds, UK, 2004.
25. D. McNeill. *Hand and Mind - What gestures reveal about thought*. The University of Chicago Press, Chicago, IL, 1992.
26. A. Mehrabian. Significance of posture and position in the communication of attitude and status relationships. *Psychological Bulletin*, 71(5), 1969.
27. C. Nass, K. Isbister, and E. Lee. Truth is beauty: Researching embodied conversational agents. In S. P. J. Cassell, J. Sullivan and E. Churchill, editors, *Embodied Conversational Agents*. MITpress, Cambridge, MA, 2000.
28. M. Neff and E. Fiume. Modeling tension and relaxation for computer animation. In *Proceedings of the 2002 ACM SIGGRAPH/Eurographics symposium on Computer animation*, pages 81–88. ACM Press, 2002.
29. H. Noot and Z. Ruttkay. Gesture in style. In A. Camurri and G. Volpe, editors, *Gesture-Based Communication in*

Human-Computer Interaction - GW 2003, number 2915 in LNAI, page 324. Springer, 2004.

30. C. Pelachaud, V. Carofiglio, B. D. Carolis, and F. de Rosis. Embodied contextual agent in information delivering application. In *First International Joint Conference on Autonomous Agents & Multi-Agent Systems (AAMAS)*, Bologna, Italy, July 2002.
31. F. E. Pollick. The features people use to recognize human movement style. In A. Camurri and G. Volpe, editors, *Gesture-Based Communication in Human-Computer Interaction - GW 2003*, number 2915 in LNAI, pages 10–19. Springer, 2004.
32. H. Prendinger and M. Ishizuka, editors. *Life-Like Characters*. Cognitive Technologies. Springer, 2004.
33. G. Taubin. SNHC verification model 7.0 - needs to be updated. Technical report, MPEG-4, 1998.
34. D. Tolani. *Inverse Kinematics Methods for Human Modeling and Simulation*. PhD thesis, University of Pennsylvania, 1998.
35. H. G. Wallbott. Bodily expression of emotion. *European Journal of Social Psychology*, 28:879–896, 1998.
36. H. G. Wallbott and K. R. Scherer. Cues and channels in emotion recognition. *Journal of Personality and Social Psychology*, 51(4):690–699, 1986.