Body Movements for Affective Expression: A Survey of Automatic Recognition and Generation - SUPPLEMENTARY MATERIAL -

Michelle Karg[1], Ali-Akbar Samadani[1], Rob Gorbet[2], Kolja Kühnlenz[3], Jesse Hoey[4], Dana Kulić[1]

1 MOVEMENT NOTATION SYSTEMS

Table 1 summarizes the characteristics of the prominent structural notation systems in terms of the evaluation criteria discussed in Section 3.2 of the survey and their application in computational recognition and generation of affect-expressive movements to date.

	Birdwhistell	Laban	Delsarte	BAP
Universality	✓	✓	✓	-
Movement analysis (anatomically and physiologically-based)	✓	✓	-	✓
Versatility	✓	✓	✓	✓
Flexibility in application	✓	✓	✓	✓
Visuality[a]	✓	✓		✓
Logicality	✓	✓		✓
Legibility	✓	✓	✓	✓
Practicability[a]	✓	✓	✓	✓
Coder's perceptual inference is not needed	-	✓	-	-
Codes the meaning/function	✓	-	✓	✓
Codes intensity	✓	✓	✓	-
Detection and recognition applications	[1]	[2], [3], [4] [5], [6]	-	[7] [8]
Generation applications	-	[9], [10], [11] [12], [13], [14]	[15] [16]	-

TABLE 1: Prominent structural notation systems

a) The application of Birdwhistell, Laban, and Delsarte notation systems should be conducted in consultation with experts in the corresponding systems.

[1]*Department of Electrical and Computer Engineering, University of Waterloo, Canada.*
[2]*Centre for Knowledge Integration, University of Waterloo, Canada.*
[3]*Electrical Engineering and Information Technology, Technische Universität München, Germany.*
[4]*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.

2 Databases

Body movement recordings in the form of video or marker-based motion capture databases are used by researchers working on developing automatic recognition and generation sytems for affect-expressive movements. Affect-expressive movements are either *acted*, *elicited* (e.g., self-induction by recalling an emotional experience), or *natural*, where the latter is preferable but often not accessible in experiments. Affect-expressive movements demonstrated by professional actors might be more expressive as compared to naive demonstrators [17]. Although no significant differences between recorded motions of lay-actors (two years of amateur acting experience) and novices was found in [18], the lay-actors reported less inhibition during the recording of affect-expressive movements than did the novices. Table 2 lists the publicly available affect-expressive movement databases.

Dataset	Demonstrators	Movements	Category	Affective states	Labeling	Samples	Data
Atkinson *et al.* [19],[20]	10 actors	Acted full body movements	Communicative	Anger, sadness, happiness, disgust, fear, surprise	Intended emotions	180	FL & PL[a] videos
Cam3D [21]	7 non-actors	Elicited hand gestures and facial expressions	Communicative	Cognitive states: thinking, concentrating, unsure, confused and triumphant. Affective states: anger, surprise, frustration, boredom, neutral	81 observers	108	Video
EIMOCAP [22]	10 actors	Scenario-based and elicited hand gestures and facial expressions	Communicative	Anger, happiness, sadness, surprise, frustration, disgust, fear, excitement	Demonstrators' self-assessment[b], 8 observers	10039	Motion capture
FABO [23]	23 non-actors	Natural scenario-based upper body movements and facial expressions	Communicative	Anger, surprise, fear, happiness, anxiety, disgust, boredom, sadness, uncertainty, neutral	Intended emotions, demonstrators' self-assessment, 6 observers	1900	Video
GEMEP core set [24]	10 actors	Acted scenario-based full body movements	Communicative	Admiration, Amusement, Extreme displeasure, Tenderness, Disgust, Despair, Pride, Shame, Anxiety, Interest, Irritation, Joy, Contempt, Panic, Pleasure, Relief, Surprise, Sadness	Intended emotions, 57 observers	145	Video
Pollick *et al.* [25]	30 non-actors	Scenario-based (scripted) walking, knocking, lifting, throwing	Functional	Anger, happiness, neutral, sadness	Intended emotions	4080	Motion capture
UCLIC [26]	13 non-actors	Acted full body movements	Communicative	Anger, fear, happiness, sadness	Intended emotions	183	Motion capture
UCLIC [27]	11 non-actors	Non-acted full body movements	Communicative	Unlabled	-	36	Motion capture

TABLE 2: Available affect-expressive movement databases.
a) FL: Full-light, PL: Point-light. b) Demonstrators' self assessment is available for the elicited affective states.

3 Performance of Automatic Recognition

Studies on the automatic recognition of affective states differ in the type of sensors recording the movement, affect-expressive movements, affective categories investigated, the type of affective expression (acted, elicited, or natural), feature selection, and classification. Hence, results of different studies often cannot be directly compared. Two general conclusions can be drawn across the different studies:

- Considering an individual study, the recognition rate is higher for person-dependent than for person-independent training of a classifier. But, differences in the design of the study, e.g., type of sensor, affect-expressive movement, or affective categories, can influence reported recognition rates more than person-dependent or person-independent training, see Tab. 3.
- When the number of affective categories increases, it may become more difficult to distinguish between the categories. Only a small number of studies have investigated more than 4 affective categories, see Tab. 3.

When a dimensional representation of affective states is used, each dimension can be discretized enabling the use of a classification approach. Alternatively, a regression technique can be applied to each dimension providing a continuous measure. Several performance measures can be applied such as recognition rate, correlation index (R), mean squared error (MSE), or root mean squared error (RMSE). Besides the differences between the studies

Study	Year	Movement	Affective States	Feature Extraction	Classifier	Recog. Rate
Amelynck *et al.*[28]	2012	Arm motion	Pleasure, arousal	Correlation Index	Linear regression	Pleasure: $R = .37$ Arousal: $R = .73$
Bernhardt *et al.* [29]	2007	Knocking [25]	Neutral, happy, angry, sad	5 selected features	SVM	50%, 81%*
Camurri *et al.*[6]	2004	Choreographed full body dance	Anger, fear, grief, joy	Amount and quality of movement	Decision tree	40%
Castellano *et al.* [30]	2007	Raising and lowering arms	Angry, joyful, pleasant, sad; 8 emotions: and desperate, interested, irritated, prideful	Correlation based FS, wrapper FS	DTW, NN, J48, HNB	4 emotions: 63%, 66%*, 8 emotions: 41%, 47%*
D'Mello and Graesser [31]	2009	Sitting posture	Boredom, confusion, delight, flow, frustration	112 selected features	Bayesian, SVM, kNN, trees, ...	39%
De Silva *et al.* [32]	2006	Gestures	Sad, frustrated, joyful, happy	Key-Points	HMM	79%
Gunes *et al.* [33]	2009	Gesture (upper body)	12 affective states	PCA	BayesNet, SVM, Random forest, Adaboost, HMM	77%
Gunes & Pantic [34]	2010	Head Gestures	1)Pleasure, 2)Arousal, 3)Expectation, 4)Intensity, 5)Power	Head angle	HMM, SVR	1)MSE=.15*, .06*, .10* 2)MSE=.13*, .07*, .06* 3)MSE=.13*, .11*, .09* 4)MSE=.10*, .05*, .09* 5)MSE=.14*, .12*, .12*
Janssen *et al.* [35]	2008	Gait	Neutral, happy, sad, angry Calm, excited, neutral	200 features 368 features	ANN ANN	84%* 79%*
Karg *et al.* [36]	2010	Gait	Neutral, happy, sad, angry	PCA,LDA	SVM, NN, NB	69%, 95%*
Karg *et al.* [36]	2010	Gait	Valence, arousal, dominance	PCA, LDA	NN	Valence: 88%*, Arousal: 97%*, Dominance: 96%*
Kapur *et al.* [37]	2005	Gestures	Sad, joy, anger, fear	8 selected features	Logistic regression, SVM, J48, NB, ANN	85%, 93%*
Kessous *et al.* [38]	2010	Gestures & Body Movement	Anger, despair, interest, pleasure, sadness, irritation, joy, pride	Software EyesWeb[39]	Bayesian classifier	67%
Nicolaou *et al.*[40]	2011	Shoulder	Valence, arousal (continuous)	5 points on shoulder and torso	SVR, BLSTM-NN	RMSE for valence: 0.21* RMSE for arousal: 0.29*
Park *et al.* [5]	2004	Non-choreographed full body dance	Happy, surprised, angry, sad	SVD	MLP	73%
Samadani *et al* [41]	2013	Hand Movement	Angry, sad, happy	Functional SPCA	NN	97%*
Samadani *et al* [41]	2013	Full-body movement	Angry, sad, happy, fearful	Functional SPCA	NN	54%*
Savva *et al.* [42]	2012	Full-body movement	Frustration, anger, happiness, concentration, surprise, sadness, boredom, relief	Dynamic features	Recurrent Neural Network	55%, 61%*
Sawada *et al.* [43]	2003	Arm movement	Joy, sadness, anger	3 selected features	Discriminant analysis	78%
Shan *et al.* [44]	2007	Arm gestures	7 emotions	PCA	SVM, NN	73%*

TABLE 3: Recognition of affective expressions from body movements (* indicates person-dependent rates).

summarized in the paragraph above, studies using a dimensional representation differ further in the quantization of each dimension and the evaluation criterion, so that results are not directly comparable.

Different techniques can be used for feature extraction, feature selection (FS) and classification. Common dimensionality reduction methods are principal component analysis (PCA), functional supervised PCA (SPCA) linear discriminant analysis (LDA), and singular value decomposition (SVD). For classification, nearest neighbor (NN), dynamic time warping (DTW), decision trees (J48), Naive Bayes (NB), Hidden Naive Bayes (HNB), support vector machines (SVM), artificial neural networks (ANN), and multi-layer perceptrons (MLP) can be used. For regression, only a small number of techniques are used such as regression, support vector regression (SVR), and bidirectional Long Short-Term Memory neural networks (BLSTM-NN).

REFERENCES

[1] 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.

[2] J. Rett, *Robot-human interface using laban movement analysis inside a bayesian framework.* PhD thesis, Univ. Coimbra, January 2009.

[3] 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.

[4] 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.

[5] 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.

[6] A. Camurri, B. Mazzarino, and G. Volpe, "Expressive interfaces," *Cognition, Technology & Work*, vol. 6, pp. 15–22, 2004.

[7] N. Dael, M. Mortillaro, and K. Scherer, "Emotion expression in body action and posture.," *EMOTION*, 2011.

[8] 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.

[9] 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.

[10] 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.

[11] K. Takahashi, M. Hosokawa, and M. Hashimoto, "Remarks on designing of emotional movement for simple communication robot," in *IEEE ICIT*, pp. 585–590, 2010.

[12] E. Barakova and T. Lourens, "Expressing and interpreting emotional movements in social games with robots," *PERS UBIQUIT COMPUT*, vol. 14, pp. 457–467, 2010.

[13] 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.

[14] 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.

[15] 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.

[16] M. Nixon, P. Pasquier, and M. El-Nasr, "Delsartmap: 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.

[17] 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.

[18] 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.

[19] 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.

[20] A. Atkinson, Q. Vuong, and H. Smithson, "Modulation of the face-and body-selective visual regions by the motion and emotion of point-light face and body stimuli," *NEUROIMAGE*, vol. 59, no. 2, pp. 1700–1712, 2011.

[21] M. Mahmoud, T. Baltrušaitis, P. Robinson, and L. Riek, "3D corpus of spontaneous complex mental states," in *Int. Conf. ACII*, pp. 205–214, Springer, 2011.

[22] C. Busso, M. Bulut, C. Lee, A. Kazemzadeh, E. Mower, S. Kim, J. Chang, S. Lee, and S. Narayanan, "IEMOCAP: Interactive emotional dyadic motion capture database," *LANG RESOUR EVAL*, vol. 42, no. 4, pp. 335–359, 2008.

[23] H. Gunes and M. Piccardi, "A bimodal face and body gesture database for automatic analysis of human nonverbal affective behavior," in *ICPR*, vol. 1, pp. 1148–1153, 2006.

[24] T. Bänziger, M. Mortillaro, and K. R. Scherer, "Introducing the Geneva multimodal expression corpus for experimental research on emotion perception.," *Emotion*, vol. 12, no. 5, pp. 1161–1179, 2011.

[25] 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.

[26] A. Kleinsmith, P. De Silva, and N. Bianchi-Berthouze, "Cross-cultural differences in recognizing affect from body posture," *INTERACT COMPUT*, vol. 18, no. 6, pp. 1371–1389, 2006.

[27] 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.

[28] 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.

[29] D. Bernhardt and P. Robinson, "Detecting affect from non-stylised body motions," in *Int. Conf. ACII*, pp. 59–70, 2007.

[30] G. Castellano, S. Villabla, and A. Camurri, "Recognising human emotions from body movement and gesture dynamics," in *Int. Conf. ACII*, pp. 71–82, 2007.

[31] 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.

[32] P. R. D. Silva, M. Osano, and A. Marasinghe, "Towards recognizing emotion with affective dimensions through body gestures," in *Int. Conf. FG*, 2006.

[33] 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.
[34] 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.
[35] 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.
[36] 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.
[37] 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.
[38] 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.
[39] 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.
[40] 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.
[41] A. Samadani, A. Ghodsi, and D. Kulić, "Discriminative functional analysis of human movements," *PATTERN RECOGN LETT*, vol. 34, no. 15, pp. 1829 – 1839, 2013.
[42] A. S. N. Savva and N. Bianchi-Berthouze, "Continuous recognition of player's affective body expression as dynamic quality of aesthetic experience," *IEEE Trans. Comp. Intell. AI in Games*, vol. 4, pp. 199–212, Sep. 2012.
[43] 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.
[44] S. G. C. Shan and P. W. McOwan, "Beyond facial expressions: Learning human emotion from body gestures," in *Brit. Mach. Vis. Conf.*, 2007.