A Two-fold PCA-Approach for Inter-Individual Recognition of Emotions in Natural Walking

Michelle Karg, Robert Jenke, Kolja Kühnlenz, and Martin Buss

Technische Universität München, München, Germany,
mkarg@tum.de, robert.jenke@mytum.de, koku@tum.de, mb@tum.de,
WWW home page: http://www.lsr.ei.tum.de

Abstract. This paper describes recognition of emotions of an unkown person during natural walking. As gait data is redundant, high dimensional and variable, effective feature extraction is essential. A combination of two consecutive Principal Component Analyses (PCA) and Fourier Transformation is used for data reduction. Naive Bayes, 1-Nearest Neighbor and Support Vector Machine (SVM) are compared for classification. Best accuracy is achieved for Naive Bayes with 72% for the four emotions sad, neutral, happy and angry within 1.65 seconds recording time. Slight increase in recognition accuracy is achieved, if the first PCA is replaced by Kernel PCA for SVM classification. The emotions angry and sad, both with high contrast in arousal, are better recognizable than neutral and happy. It is concluded that gait is capable to reveal the emotional state of a walker above chance level.

1 Introduction

Affective computing and its benefit is investigated within the last decade. This involves recognition of human emotions and designing appropriate reactions for a computer or a robot to enhance interaction [1].

Emotion recognition is usually based on the analysis of facial expressions, acoustic as well as linguistic features in speech or postures [2]. It is expected that combination of different modalities increases accuracy and robustness of current recognition systems. Human observers are highly sensitive to inconsistency in emotional expressions of different modalities [3]. The credibility of an expressed emotion strongly depends on consistence of different modalities. For this reason this work investigates if human motions are capable to reveal information about the emotional state of a person. As daily and ordinary motion we choose natural walking.

First, this paper gives a compact overview of current studies which investigate recognition of emotions in human motions in section 2. Our analysis uses 3D optical tracking recordings of Ma et al. [4]. The data set, feature extraction and classification is described in section 3. This study focuses on inter-individual recognition of emotions. Feature extraction is based on a twofold Principal Component Analysis (PCA) and Fourier Transform. For comparison, the second PCA is omitted and accuracy is determined for this case. Also, the nonlinear extension

Kernel PCA is investigated as an option to PCA. Results regarding classification accuracy are presented in section 4. The paper ends with a discussion of currently achieved results in recognition of emotions in gait patterns in section 5 and finally draws the conclusion that emotion recognition above chance level is possible. Particularly, recognition of emotional states which differ in arousal, like anger and sad, give better results.

2 State of The Art

Due to Lazarus, emotion is defined as the combination of physiological disturbance, action tendencies, which are not necessary acted out, and subjective experience [5]. Recognition of emotions in body motion is mostly related to discrete categorization of emotion. For facial expression, the discrete categorizations happy, sad, angry, surprise, disgust and fear are commonly used. In recognition of emotions in body motion usually a subset of these emotions is investigated, as emotions like disgust are hard to express an thus to recognize in posture or walking.

In 1996, Dittrich et al. investigated the expression of emotions in dance [6]. Point-light videos were recorded from trained dancers, who conveyed the emotions fear, anger, grief, joy, surprise and disgust. Human observers classified the videos to 88% correct. Also, Pollick et al. showed that the affective states afraid, angry, happy, neutral and sad can be recognized during knocking [7] by human observers with an error rate of 29%. Bernhard et al. applied SVM classification to the same data set and reached an average sensitivity of 81% [8]. Statistical features in this study are maximum distance of hand from body, average hand speed, acceleration and jerk. Castellano et al. followed a different approach. They investigated how emotions affect a defined motion. The motion was swinging arms, which is a nondescript motion for expression of emotions. The average accuracy for four emotions (anger, joy, pleasure, sadness) was 68% [9].

Crane et al. analyzed if human observers can identify emotional states in natural walking. The average recognition rate was 72.38%. Further on, emotions significantly affect step frequency, head orientation and shoulder as well as elbow range of motion [10]. Anymore, human perception of emotions in walking is affected by gender. Female kinematics are similar to kinematics of fearful gait and male kinematics facilitate the perception of anger, as shown by Halovic and Kroos [11]. Janssen et al. achieved intra-individual recognition rates of 99,3% for recognition of emotions in walking patterns using artificial neural nets, whereas inter-individual recognition levels remain around chance [12].
Several approaches have been explored in the field of gait data analysis in recent years. Among these are multivariate statistics, fractal dynamics, fuzzy systems, neural networks and time-frequency analysis [13, 14].

3 Methods

This study analyzes inter-individual recognition of emotions in natural walking. As gait data sets are redundant, high dimensional and variable, efficient feature extraction is crucial for further classification.

3.1 Gait Data Set

F. Pollick together with his team at the University of Glasgow, Scotland, developed a motion capture library to study perception of identity, gender and emotion from biological motion [4]. They recorded 30 non-professional actors while they performed walking, knocking, lifting and throwing actions with a Falcon Analog optical motion capture system (60 Hz). The emotions sad, angry, happy and neutral were chosen because they represent emotional states which may last for an extended period of time. A story for each emotion was told to the participants, in which they had to imagine themselves in a specific situation. Then the participants walked in a triangle for 25 seconds. The position over time of 33 passive markers, which are attached to the skin of the subjects, is available in .csm format. The data can be downloaded from paco.psy.gla.ac.uk/data.php and is used for this study.

Sections of straight walking are extracted from the data sets, each with a length of 1.65 sec. The position of 33 markers of 29 participants within this time window is used for feature extraction and classification.

Using the triangle walk complicates recognition of emotions and causes falsification by extracting unfavourable features. Thus, sections of straight walking are extracted from the data sets, each with a length of 1.65 sec. The algorithm finding these sections examines the marker close to the bury centre. Start and end points of straight movement of this marker indicates straight walking. Since different lengths of straight walking for each recording are returned they need to be cut down to the shortest afterwards, i.e. a frame length of 99 time steps in this case. The position of 33 markers of 29 participants within this time window is used for feature extraction and classification.

3.2 Feature Extraction

One frame contains the 3-dimensional position of 33 markers over time, which affords a 99 dimensional vector $\mathbf{x}(t)$. A clipped recording of one gait i consists of 99 frames, so that the dimension of the input data $\mathbf{X}_i$ is $R^{99\times 99}$. The complete data set contains 116 gaits. Due to the curse of dimensionality and redundancy, efficient dimension reduction is required before applying different classifiers. On on hand, statistical features based on joint angles and joint centers calculated by kinematic transformation of $\mathbf{x}(t)$ can be used for classification. On the other hand, dimension reduction techniques can be directly applied to the marker-based presentation of a gait. This study investigates the performance of the latter one using variations of PCA. In this case, no prior expert knowledge is required and the task of the dimension reduction technique is to select features comparable to expert knowledge.

Structural and dynamical cues are extracted by two PCAs and Fourier Transformation (FT). The procedure is leant on Eigenpostures and Eigenwalkers as proposed by Troje for gender recognition in walking [15]. In addition, PCA is replaced by Kernel PCA (KPCA) and the accuracy is compared to the PCA-FT-PCA approach.

First, PCA is applied to the data matrix $\mathbf{X}_i$ of a single gait i. This results in a mean posture vector $\mathbf{p}_{mean,i}$ and 99 Eigenvectors, each of them is 99 dimensional. The first four Eigenvectors, named Eigenpostures $\mathbf{p}_{k,i}$, are used for further data analysis according to Troje [15]. The coefficients $c_{k,i}(t)$ describe the temporal development of each Eigenposture $\mathbf{p}_{k,i}$. A single gait $\mathbf{p}_i(t)$ at a specific time step is approximated by

$$\mathbf{p}_i(t) = \mathbf{p}_{mean,i} + \sum_{k=1}^{4} c_{k,i}(t)\mathbf{p}_{k,i} \qquad \text{with} \;\; \mathbf{p}_i(t),\; \mathbf{p}_{mean,i},\; \mathbf{p}_{k,i} \in R^{99\times 1} \; . \quad (1)$$

The Eigenpostures are specific for each walker and each emotional expression. Dynamic information of one individual gait is represented by the changes over time of the coefficients $c_{k,i}(t)$, whose main frequency and phase shift is extracted with FT. The leakage effect is reduced using a Hamming filter. The phase shift of the first Eigenposture $\phi_{1,i}$ is always set to zero and the others $\phi_{2,i}, \phi_{3,i}, \phi_{4,i}$ are adapted. A single gait $\mathbf{p}_i(t)$ is modelled as follows

$$\mathbf{p}_i(t) = \mathbf{p}_{mean,i} + \mathbf{p}_{1,i}\sin\left(2\pi f_{1,i}t\right) + \sum_{k=2}^{4} \mathbf{p}_{k,i}\sin\left(2\pi f_{k,i}t + \phi_{k,i}\right) \; . \quad (2)$$

All parameters $\mathbf{p}_{mean,i}$, $\mathbf{p}_{k,i}$, $f_{k,i}$ and $\phi_{k,i}$ are time-independent. Composition of the mean posture vector $\mathbf{p}_{mean,i}$, the four Eigenpostures $\mathbf{p}_{k,i}$, the four frequencies $f_{k,i}$ and three phase shifts $\phi_{k,i}$ results in the vector $\mathbf{w}_i$, which describes a single walk. Its dimension is $5 * 99 + 4 + 3 = 502$. Reduced versions $\mathbf{w}_{dyn,i}$, $\mathbf{w}_{EP1,i}$, $\mathbf{w}_{EP2,i}$, $\mathbf{w}_{EP3,i}$ and $\mathbf{w}_{EP4,i}$ of $\mathbf{w}_i$ exclude the mean posture or contain only single Eigenpostures. A second PCA transformation is applied to the set of $29*4 = 116$ walks $\mathbf{w}_i$, including all emotional walks of all participants. This results in up to 99 Eigenvectors, called Eigenwalkers $\mathbf{w}_{EW,j}$. Instead of classification of a 502 dimensional vector $\mathbf{w}_i$, only the coefficients d_{ij} of the second PCA build the feature vector for classification. A single walk $\mathbf{w}_i$ is described by the mean $\mathbf{w}_{mean}$ of all walks and individual coefficients d_{ij} for the Eigenwalkers $\mathbf{w}_{EW,j}$

$$\mathbf{w}_i = \mathbf{w}_{mean} + \sum_{j=1}^{99} d_{ij}\mathbf{w}_{EW,j} \; . \quad (3)$$

Figure 1 summarizes applied feature extraction. The following variations of the algorithm are investigated. Either the first PCA, the second PCA or both are replaced with KPCA. Polynomial kernels with different degrees and a radial

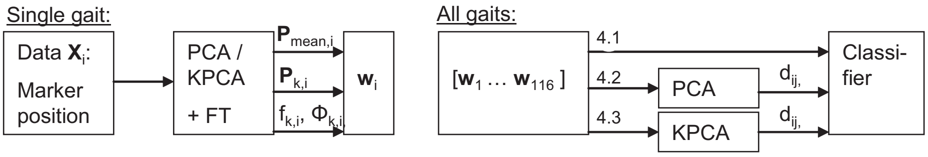

Fig. 1. First, PCA and FT is applied to the data of a single gait. Then a second PCA calculates the Eigenwalkers of all gaits which contain most variance. In all stages PCA can be replaced by KPCA for comparison.

bais function kernel were tested for KPCA. In each case the dimension of the Eigenvectors increases. Furthermore, the second PCA is skipped and a Suppport Vector Machine is directly applied to vector $\mathbf{w}_i$ of all walks, to evaluate if the last PCA stage is actually necessary. The subsections 4.1, 4.2 and 4.3 give results for each case.

3.3 Classification

Naive Bayes, 1-Nearest Neighbor using Euclidean distance and L1 soft-margin Support Vector Machine (SVM) with a radial basis function as kernel [16] ($c = 1.0, \gamma = 0.01$) were used for classification. As the number of data sets is small, the recognition rate is calculated using leave-one-out cross validation.

4 Results and Discussion

Recognition rates are calculated for the combinations PCA, KPCA, PCA-FT-PCA, KPCA-FT-PCA, PCA-FT-KPCA and KPCA-FT-KPCA for feature extraction. The approaches are compared regarding accuracy. For the best result, the confusion matrix is presented additionally. Expression of emotions varies within subjects, ground truth of labelled data is not assured in affective computing and also gait patterns are highly individual, so that recognition rates of 100% are not expected. In general, recognition rates range between 60% and 95% for person-independent recognition of emotions in speech and facial expressions [17].

4.1 Single PCA or KPCA

It is expected, that applying a second PCA the data set containing all Applying PCA to the data set of each gait, train the classifier with the gaits of 28 persons and recognize the four emotional states of the excluded person's gait patterns lead to recognition rates around chance level. Table 1 lists the recognition rates for applying SVM to a dynamic description $\mathbf{w}_{dyn}$ excluding the mean posture and reduced descriptions of a walk containing only single Eigenpostures (EP). PCA can be replaced by KPCA, then dimension of the Eigenpostures is 116.

Best result is achieved for KPCA with a polynomial kernel and degree $d = 3$. In comparison to the other emotions, the emotional state sad is better recognizable. Recognition rate for this case is 55%. Table 2 shows the corresponding confusion matrix. Applying SVM to data sets containing complete description $\mathbf{w}$ of all gaits is not suitable, as in all cases recognition rate is 25% which is equal to chance level.

Table 1. Recognition rate for SVM applied to Eigenpostures, frequencies and phase shifts.

Feature Extraction	**1st EP**	**2nd EP**	**3rd EP**	**4th EP**	$\mathbf{w}_{dyn}$
PCA	34	29	27	19	26
KPCA (polynomial kernel, d=1)	24	29	26	34	27
KPCA (polynomial kernel, d=2)	20	33	30	34	16
KPCA (polynomial kernel, d=3)	35	**55**	42	30	45
KPCA (polynomial kernel, d=4)	25	27	31	26	35
KPCA (radial basis function)	25	28	24	34	24

Table 2. Confusion Matrix for KPCA (polynomial kernel, d=3) with SVM.

	angry	**happy**	**neutral**	**sad**	acc(%)
angry	15	9	2	3	52
happy	6	14	5	4	48
neutral	4	5	15	5	52
sad	1	4	4	20	69

4.2 Twofold PCA

First, the combination of PCA-FT-PCA is used for feature extraction. Figure 2 shows the accuracy depending on the number of Eigenwalkers $\mathbf{w}_{EW,j}$. Eigenwalkers are sorted according to descending variance. A Naive Bayes classifier is trained with the coefficients $d_{i,j}$ of the related Eigenwalkers. Best recognition rate is achieved for applying the second PCA only to the first Eigenposture. A maximum of 65% is reached for 39 Eigenwalkers. Accuracy for angry, happy, neutral and sad gaits is 69%, 45%, 62% and 83% respectively. Increasing the number of Eigenwalkers decreases accuracy, which is traced back to redundant features and a higher ratio between number of features to number of training sets. Table 3 shows the corresponding confusion matrix. The emotional states sad and angry are better distinguishable.

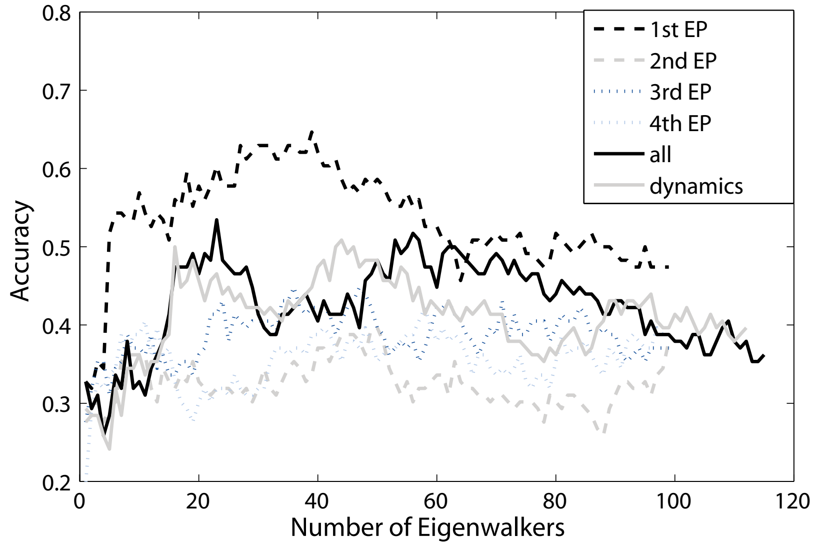

Fig. 2. Highest accuracy is achieved if the second PCA is applied to the first Eigenposture of all gaits. Furthermore, accuracy varies due to redundant features for Naive Bayes.

In comparison to SVM and 1NN, Naive Bayes reaches best recognition rate. However, 1NN and SVM are less sensible to redundant features and accuracy is almost stable over number of Eigenwalkers. Plot 3 shows accuracy over number of Eigenwalkers exemplary for applying PCA to the first Eigenposture. 1NN has lowest accuracy, also for using other Eigenpostures or the complete vector **w** for calculation of the Eigenwalkers. For SVM, accuracy reaches 50% for using 20 Eigenwalkers, calculated based on complete vectors **w**.

So far, the order of the Eigenwalkers represents the variance in gait patterns. As redundant features are crucial for Naive Bayes, backward elimination can give the optimal coefficients $d_{i,j}$ for emotion recognition in walking. Recognition rates above 70% are reached, if the number of Eigenwalkers is between 11-52.

Table 3. Confusion matrix for Naive Bayes (Accuracy: 65%) using the 1st Eigenposture.

	angry	**happy**	**neutral**	**sad**	acc(%)
angry	20	4	3	2	69
happy	6	13	7	3	45
neutral	3	6	18	2	62
sad	2	2	1	24	83

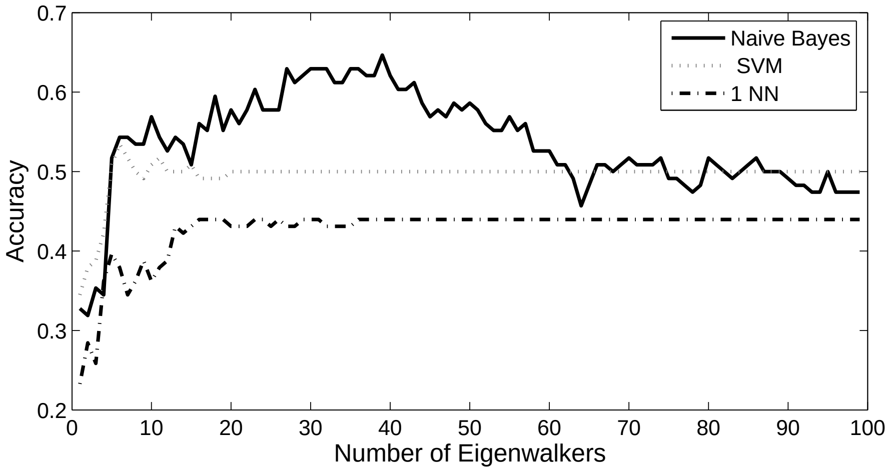

Fig. 3. Naive Bayes reaches highest accuracy in comparison to SVM and 1NN using the first Eigenposture for calculation of the Eigenwalkers.

Backward elimination returns best accuracy of 72% for a feature vector containing the coefficients of 20 Eigenwalkers. The Eigenwalkers are 5, 6, 10, 13, 14, 15, 18, 23, 27, 30, 31, 47, 52, 53, 70, 80, 83, 84, 90 and 98. Table 4 shows the confusion matrix for this case. Accuracy for each emotion is above 69%.

Table 4. Confusion Matrix for Naive Bayes with Best Eigenwalkers (Accuracy: 72%).

	angry	**happy**	**neutral**	**sad**	acc(%)
angry	20	6	1	2	69
happy	4	22	2	1	76
neutral	3	5	20	1	69
sad	2	3	2	22	76

4.3 Substitution of PCA with KPCA

As accuracy depends strongly on the number of Eigenwalkers for Naive Bayes, SVM is chosen for comparison between PCA and KPCA. Recognition rate varies less with number of features. Slight improvement of recognition rate for SVM classification is achieved, if the first PCA is replaced by KPCA with a quadratic polynomial kernel. In this case accuracy is 52% and KPCA is applied to the second Eigenposture. Replacing only the second PCA with KPCA, either radial basis function or a polynomial kernel, does not improve recognition rate. Best kernels for all combinations of two subsequent KPCAs are two quadratic polynomial kernels. In this case accuracy reaches 52%, if the first KPCA is applied to the second Eigenposture. In comparison to two consecutive PCAs, the number of required Eigenwalkers decreases until recognition rate reaches a stable limit, though no significant increase in recognition rate is observed. The emotional states sad and angry are easier to recognize than happy and neutral for the combinations which reach 52% accuracy.

5 Discussion

Recognition of emotions in gait patterns is a challenge from data mining point of view, as data sets are high-dimensional comparing to the number of samples and redundant. Also, recognition of emotions in gait patterns is not trivial for human observers. An inter-individual recognition rate of 72% is achieved for distinguishing four emotions in gait patterns by using a twofold PCA for feature extraction, finding the most relevant features by backward elimination and applying Naive Bayes. In comparison to recognition rates of emotions in speech or facial expressions, recognition of emotions in gait patterns with this approach is within the range of published work. Yet gait is not a primary indicator for emotions, so a deficit is expected comparing to recognition based on facial expressions or speech. Due to low number of individuals, inter-individual recognition rate for four emotions in walking remains around chance level in [12]. High number of individuals is required to compensate the variance in walking among individuals. It is concluded that gait is capable to give cues about ones emotional state. However, further improvement is required. This includes intra-individual recognition, markerless tracking and further improvement of feature extraction as well as classification.

Within this work a simple twofold PCA approach for feature extraction shows good performance regarding accuracy. Removing the second PCA, leads to lower recognition rates. Also, a nonlinear extension of PCA, KPCA, has been compared to the performance of PCA. Accuracy does not significantly improve, but less features are required for same recognition rate. In general the emotional states sad and angry are better recognizable. The emotions sad and angry are also better recognizable in knocking [8]. This rises the question, if these emotions are generally easier to express and thus to recognize in motions.

6 Conclusion and Future Work

This work uses a combination of PCA-FT-PCA for feature extraction to recognize four emotional states in walking. Applying backward elimination leads to a selection of features, which reach an accuracy of 72% for person independent recognition. The emotions angry and sad are in general better recognizable in gait patterns. Furthermore, future identification of the recognition rate of human observers for this data set allows comparison to performance of automatic recognition. In addition, the PCA-FT-PCA approach can be compared to predefining gait parameters, like elbow range of motion, and classification based on these parameters.

ACKNOWLEDGMENTS
This work is supported within the DFG excellence initiative research cluster "Cognition for Technical Systems - CoTeSys", see also www.cotesys.org.

References

1. R. Picard, *Affective Computing*. Cambridge: MIT Press, 2002.
2. C. Peter and R. Beale, *Affect and Emotion in Human-Computer Interaction*. Springer, 2008.
3. Y. Bar-Cohen and C. Breazeal, *Biologically inspired intelligent robots*. Spie Press, 2003.
4. Y. Ma, H. Paterson, and F. Pollick, "A motion capture library for the study of identity, gender and emotion perception from biological motion," *Behaviour Research Methods*, vol. 38, pp. 134–141, 2006.
5. R. Lazarus, *Emotion & Adaptation*. Oxford University Press, 1991.
6. W. Dittrich, T. Trosciancko, and S. L. ans D. Morgan, "Perception of emotion from dynamic point-light displays represented in dance," *Perception*, vol. 25(6), pp. 727–738, 1996.
7. F. Pollick, H. Paterson, A. Bruderlin, and A. Sanford, "Perceiving affect from arm movement," *Cognition*, vol. 82, pp. B51–B61, 2001.
8. D. Bernhardt and P. Robinson, "Detecting affect from non-stylised body motions," in *Affective Computing and Intelligent Interaction*, 2007.
9. G. Castellano, S. Villabla, and A. Camurri, "Recognising human emotions from body movement and gesture dynamics," in *Affective Computing and Intelligent Interaction*, 2007.
10. E. Crane and M.Gross, "Motion capture and emotion: Affect detection in whole body movement," in *Affective Computing and Intelligent Interaction*, 2007.
11. S. Halovic and C. Kroos, "Facilitating the perception of anger and fear in male and female walkers," in *AISB, Symposium on Mental States, Emotions and their Embodiment*, 2009.
12. 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," *Nonverbal Behaviour*, vol. 32, pp. 79–92, 2008.
13. T. Chau, "A review of analytical techniques for gait data: Part 1: Fuzzy, statistical and fractal methods," *Gait and Posture*, vol. 13, pp. 49–66, 2001.
14. T. Chau, "A review of analytical techniques for gait data: Part 2: Neural network and wavelet methods," *Gait and Posture*, vol. 13, pp. 102–120, 2001.
15. N. Troje, "Decomposing biological motion: A framework for analysis and synthesis of human gait patterns," *Journal of Vision*, vol. 2, pp. 371–387, 2002.
16. C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines," 2001. Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm.

17. Z. Zeng, M. Pantic, G. Roisman, and T. Huang, "A survey of affect recognition methods: Audio, visual, and spontaneous expressions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 1, pp. 39–58, 2009.