

Recognition of Emotions in Gait Patterns by Means of Artificial Neural Nets

Daniel Janssen · Wolfgang I. Schöllhorn · Jessica Lubienetzki ·
Karina Fölling · Henrike Kokenge · Keith Davids

Published online: 11 January 2008
© Springer Science+Business Media, LLC 2008

Abstract This paper describes an application of emotion recognition in human gait by means of kinetic and kinematic data using artificial neural nets. Two experiments were undertaken, one attempting to identify participants' emotional states from gait patterns, and the second analyzing effects on gait patterns of listening to music while walking. In the first experiment gait was analyzed as participants attempted to simulate four distinct emotional states (normal, happy, sad, angry). In the second experiment, participants were asked to listen to different types of music (excitatory, calming, no music) before and during gait analysis. Derived data were fed into different types of artificial neural nets. Results showed not only a clear distinction between individuals, but also revealed clear indications of emotion recognition in nets.

Keywords Emotion · Gait · Music · Neural network · Pattern recognition

Introduction

The question “How is it going?” is often posed when a person is interested in another individual's mood state. A fascinating aspect of this question is the vocabulary which is used. The word *going* (from *gait* as a mode of locomotion) is used to ascertain information on feelings and emotional states or on health in general. Thus, in common linguistic usage, there seems to be a connection between movement and inner feelings, which can be

D. Janssen (✉) · W. I. Schöllhorn
Training and Movement Science, University of Mainz, Albert Schweitzer Strasse 22,
55099 Mainz, Germany
e-mail: djanssen@uni-mainz.de

J. Lubienetzki · K. Fölling · H. Kokenge
Training Science, University of Muenster, Leonardo Campus 15, 48149 Münster, Germany

K. Davids
School of Human Movement Studies, Queensland University of Technology, Brisbane, Australia

confirmed by interpreting the etymology of the word *emotion* (from Latin *e*: *out of* and *movere*: *move*).

Whether or not this connection exists and whether bodily expression is affected by emotional states has been the subject of several previous studies. According to some previous research, humans are able to recognize emotions from a person's posture (e.g., Coulson 2004), facial expression (e.g., Ekman and Friesen 1978), or body movement (e.g., Clarke et al. 2005; De Meijer 1989; Dittrich et al. 1996; Montepare et al. 1987; Montepare et al. 1999; Walk and Homan 1984, Wallbott and Scherer 1986). While these studies primarily focused on the human ability of perception and recognition of emotions in movements, computer-assisted research has tended to focus on recognizing emotions from speech patterns (e.g., Nicholson et al. 2000; Nwe et al. 2003; Park et al. 2005), from facial expressions (e.g., Ichimura et al. 2001; Ioannou et al. 2005; Kaiser and Wehrle 1992), or in both speech and facial expressions (e.g., Fragopanagos and Taylor 2005). More recently in the field of human–machine communication, the recognition and application of emotional elements has received increasing interest for the improvement of machine-based speech interpretation, lie detection, intelligent tutoring applications, and for the enhancement of interactive and realistic computer games (Fragopanagos and Taylor 2005).

However, research on acquiring emotional information by observing body movements by means of automated computer assistance is rare. For example, Camurri et al. (2003) and Sawada et al. (2003) studied computerized emotion recognition techniques in analyzing dance movements. In an ongoing program of work, a comparison between that technique and the subjective judgements of spectators was drawn. The computer-based technique was found to achieve 71.4% of the spectators' level of performance (Camurri et al. 2004). In other work, Gunes and Piccardi (2005) developed a multi-modal method to study computer-based recognition of emotional information from both facial expressions and upper body gestures. In their study, participants were filmed while interpreting different emotional states. Manually selected video frames were extracted for the computer-based emotion recognition. This multi-modal approach (facial expression and upper body gestures) led to an increase in emotion recognition rates when contrasted with the use of only a single modality.

However, not only movement characteristics, postures and facial expressions seem to be affected by emotions. Further support for the emotional impact on bodily processes are provided by data from studies by Coombes et al. (2005, 2006), who were able to identify changes in movement coordination and force production processes, respectively, after visual presentation of emotional stimuli to participants. Similar findings based on changes in hand movement patterns while listening to different types of music, were provided by Camurri et al. (2006). Their results suggested that even music is able to influence movement characteristics.

However, when an individual is observing another individual, expressed emotional content is not the only information which can be received. It has been found that humans are also able to recognize the individuality of others from facial information, even in a crowd (e.g., Bichot and Desimone 2006), or from a person's individual walking style (e.g., Cutting and Kozlowski 1977). In these contexts, dynamic data seem to provide more reliable information to support recognition processes than static data (e.g., Schöllhorn et al. 2002). Interestingly, most research on the analysis of individual gait patterns has been conducted in clinical settings or on biometric identification processes (e.g., Benabdelkader et al. 2004). For example, Schöllhorn et al. (2002) demonstrated that participants ($n = 13$ females walking in dress shoes with different heel heights) could be identified with recognition rates of up to 100%, from analysis of only 200 ms of their gait pattern, using

kinetic and kinematic data obtained with artificial neural networks in the recognition process. Kinetic data (3D ground reaction forces during gait) were derived with a force plate while kinematic data (3D angles and angular velocities of ankle and knee) were computed using a four-camera motion analysis system.

Due to methodological issues, there has been a strong emphasis on investigating the recognition of individuals and emotional influences, independently of each other. Research on simultaneous recognition of individuals and their emotions has tended to be neglected. For this reason, our experiments studied whether it was possible to recognize individuals by their gait patterns, and on refined level to distinguish emotions that were directly or indirectly evoked by imagination or by musical influence. Accordingly, we sought to examine three main questions: (1) Can neural networks recognize individuals by their kinetic and kinematic gait patterns? (2) Can neural networks distinguish different emotional states simulated within the individual gait patterns? and (3) How do participants' gait patterns change when listening to music? More specifically, does gait pattern provide any information about characteristics of the type of music that a participant might be listening to?

An overview of the study's organization is given in Fig. 1. Two experiments were undertaken: In the first experiment participants were asked to imagine several emotional states while walking. Kinetic data were derived and prepared for the recognition of individual gait patterns with a multilayer perceptron (MLP; Rumelhart et al. 1986), representative of supervised learning, and for emotion recognition with a self-organizing map (SOM; Kohonen 1982), representative of unsupervised learning. In the second experiment kinetic and kinematic data were recorded when participants were walking while listening to different types of music. In this second experiment, kinetic data were treated as in Experiment 1, and kinematic data were fed into a MLP for recognition of individual gait patterns and into a custom-made SOM for emotion recognition. Details on neural networks used for data analysis are provided in the next section. The experiments are described in more detail in subsequent sections of this article.

Neural Networks

In gait analysis, previous research has primarily focused on; (i) the human capacity to perceive and recognize individual, gender- or age-specific gait patterns; (ii) the emotional influence on perception and recognition processes; or (iii) the recognition of intentionally disguised movement patterns. This body of work has tended to adopt more subjective

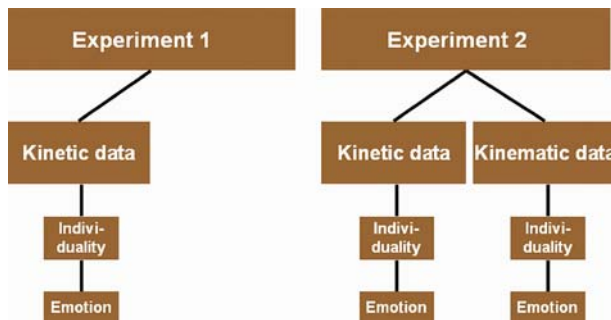


Fig. 1 Schematic overview of the two experiments with included types of variables and aimed recognition area

methods of movement recognition (for an overview see Richardson and Johnston 2005). Over the past 15 years, a considerable amount of research on gait analysis has been conducted using quantitative and more objective computer-based methods such as artificial neural nets (ANNs). Due to their nonlinear recognition and classification methods, ANNs are more suitable for tasks such as pattern or speech recognition in comparison to linear methods (a review in the context of clinical biomechanics is provided by Schöllhorn 2004). ANNs in general can be viewed as heavily simplified models of parts of the human brain. Several different types of ANN models exist, including models with supervised, unsupervised, and reinforcement learning methods. However, the architecture of ANNs mostly consists of a mathematical graph and, depending on the topology of the network, it is either cyclic or acyclic. In addition, shortcut connections within the architecture may exist as well. The graph's vertices are substituted by neurons, which receive input signals (modeled by means of numerical values) from their dendrites (connected presynaptic neurons) over the edges (weighted connections). The connections between neurons are weighted with different strengths (also numerical values) in order to simulate different synaptic weights in accordance with the neurobiological model of origin.

Within the field of ANNs the most popular nets, *inter alia*, are the supervised multilayer perceptrons (MLP) and the unsupervised self-organizing maps (SOM). From a statistical point of view, these two approaches can be seen as hypothesis verifying (supervised) and hypotheses generating (unsupervised) approaches. For the use of a MLP, the data need to be divided into training and test data. Training data are used to train the net, whereas test data are used for the application. A validation data set may exist as well, but is not considered in our experiments. During the training process the elements of the network (more precisely the weights and biases) are modified, so that for each input pattern the given network output is made more similar to the desired output. For example, if a gait pattern belongs to person A, but the network allocates it to person B, the inner units are changed, so that after some training the network correctly associates this specific pattern with person A. If this occurs for all available inputs over several hundreds or thousands of trials, the network learns to recognize specific patterns well, while still being able to generalize and associate novel patterns with the appropriate individuals. The information is not stored in a database-like unit, but is implicitly distributed among the neurons and connection values.

When using a SOM for classification or data mining, the category that a pattern belongs to, need not be known *a priori*. The network is able to classify, separate, and distinguish high-dimensional input patterns by their similarities. This approach is well-suited for studying data where there is a lack of knowledge of membership of classes or categories, since the net is able to find clusters (classes) of similar input-patterns within the data set automatically and to map them onto groups of similar or neighbored neurons. A mostly two-dimensional graphical output space is used to illustrate the mapping of the data onto the net.

Further types of neural networks, including combinations of supervised and unsupervised approaches can be found in Haykin (1998).

Experiment 1

Method

For studying the emotional states simulated by participants (walking normally, with joy, with sadness, and with anger) kinetic data were recorded from 22 healthy male and female

participants using a 40×60 cm Kistler force platform 9821B with a frequency of 1,000 Hz. The experimental setup is illustrated in Fig. 2. Participants had to walk a distance of approximately 7 m at a self-determined walking speed registered by two pairs of double light-barriers after a short period (about two and a half minutes) of internalization of the emotion to be simulated by imagination. Participants were asked to feel sad, angry, or happy, depending on the assigned emotion. Simulations of these emotions were aided by encouraging participants to remember a particular occasion when they felt each specific emotion. The order of emotions to be simulated in walking was randomly pre-assigned. With their 3rd to 5th foot contact, participants were required to hit the force platform with the right foot without incurring any unnatural step lengths or movements. Participants had to perform several trials until data on gait from three consecutive error-free trials were collected per participant for each emotion. Error-free in this context meant that participants were asked to avoid looking at the platform, and data were only recorded if they hit the platform centrally with their 3rd to 5th stride. The ground reaction forces in x -, y -, and z -dimensions were acquired by means of commercially available software (Dasy Lab 6.00.03). In order to remove, or at least to minimize the influence of speed and body weight on the recognition process, all measurements were normalized by amplitude and time. The vertical ground reaction force was divided by the participant's weight, and the horizontal forces were scaled into a unique interval. Since data were recorded at 1,000 Hz for each trial, approximately 600 data points were available for analysis per dimension. These points were time-normalized to 100 data points for the z -, 50 for the x -, and 50 for the y -dimensions, leading to an optimal ratio of computational effort and precision for use with the neural nets. Subsequently, a synthetic model gait pattern was built by calculating the mean of all available ground reaction forces from all participants in x -, y -, and z -dimensions. This reference gait pattern was then subtracted from all single gait patterns in order to extract the individual deviations from the model pattern. This process allowed us to extract solely the individual characteristics of each participant's gait patterns for further analysis. All calculations were based on the software package Matlab R2006a.

In order to analyze individuality in gait patterns, the data set, consisting of all available gait patterns from the 22 participants, was split into training and test data at a ratio of 2:1 and thereafter presented to a supervised MLP with 200-111-22 (input-hidden-output) neurons (one output neuron per participant). Recognition rate was computed by expressing the number of correct recognitions (correct identifications of gait patterns that belonged to a particular participant) as a percentage. Recognition rate was additionally averaged using cross-validation (Schöllhorn et al., in press) in order to obtain precise rates of identification (see Appendix for a more precise description of the nets' architectures).

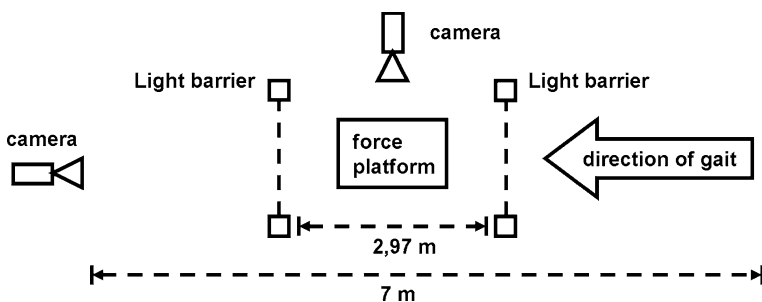


Fig. 2 Schematic depiction of the data collection apparatus

For intra-individual emotion recognition during gait, an unsupervised SOM with 5×3 neurons was chosen. All available gait patterns from a single participant were in each case fed into the network. From gait analysis, it is known that dynamic data provides more information than static data (Richardson and Johnston 2005). In addition, Schöllhorn et al. (2002) underlined that time-continuous data sets are better suited for achieving individual information than time discrete data sets. Hence, similar to Schöllhorn et al. (2002), two approaches were applied: (i) using time discrete parameters (minima, maxima, positions of minima, positions of maxima, integral, length, and mean of the curves in x -, y -, and z -dimensions); and (ii) using time continuous data (whole time courses) as inputs for the SOM. The SOM classified and clustered the gait patterns according to their similarities. Because recognition rates are normally not provided by a SOM, a customized algorithm was developed in order to ascertain the emotion-distinction quality of the net. For each gait pattern the demonstrated emotion was documented. With this knowledge a kind of retrospective performance analysis was conducted, analyzing whether gait patterns with the same demonstrated emotions formed clusters, which implied that these patterns should be more similar (see Appendix).

Results

A person recognition rate of 95.3% was achieved with the MLP trained with 176 gait patterns from all 22 participants and tested with 88 unknown gait patterns of these individuals. Hence, 95.3% of the gait patterns were allocated to the correct individuals.

Intra-individual emotion recognition (identifying the four emotional states of an individual's gait patterns correctly) was realizable with up to 100% accuracy for some participants. Figure 3 provides an exemplary spread of a participant's gait patterns over the output space of the SOM (5×3 neurons). In all parts of the figure the SOM output is shown. Parts (b–e) of the figure illustrate which neurons classified the participant's gait patterns from the simulated emotional state. The bigger a filled hexagon appears, the more gait patterns of the same emotion are classified onto this neuron. It can be seen, that gaits of the same emotions are predominantly classified into the same region, and these gaits are more strongly distinguished from the gaits associated with other emotions in most cases. This observation is supported from data shown in part (a) of the figure (the unified distance matrix), where the real similarities between classified gait patterns (represented by activated neurons) can be interpreted from the brightness of the background. The cluster borders are illustrated in black. Neighbored gaits inside this region are more dissimilar, because the real distances between them are greater (borders can be thought of as hills in a 3D landscape). By contrast, white planes (valleys in this sense) indicate that gaits in these areas have strong similarity. The emotions *anger*, *joy*, and *sadness* are distributed over the map with maximum distances to each other, the control condition *normal* is closest to the emotional state *sadness*.

Average emotion recognition rates were also calculated. For the time discrete parameters an average recognition rate of 80.8% ($SD = 11.5$; maximum = 100.0%) was achieved from analysis of data from all participants pooled. By including the whole time courses of action, when considering time continuous data, performance could be increased to 83.7% ($SD = 12.5$; maximum = 100.0%) and thus both methods revealed findings significantly above the level of chance (25%). However, no statistical difference was reported between both emotion recognition rates using the Mann–Whitney Test ($Z = -.82$, $p = .41$).

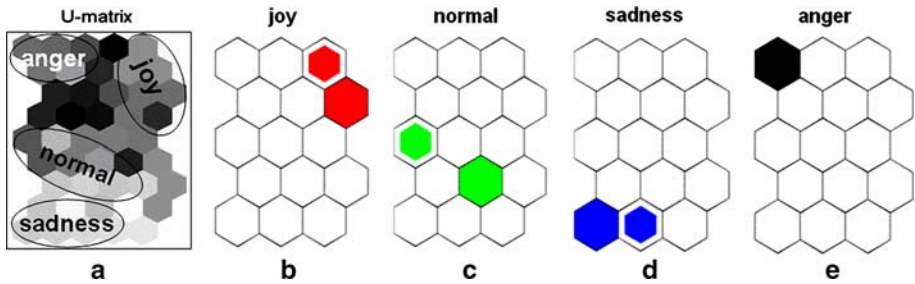


Fig. 3 Unified distance matrix (a) and activated neurons ordered by simulated emotions (b–e) of a participant's kinetic gait patterns classified with the SOM. The filled hexagons (b–e) show onto which neuron gait patterns for the given emotional state are classified. The bigger a hexagon appears the more gaits are classified onto this neuron. Neighborhood can be interpreted as similarity. The distance matrix (a) shows the real distances between the data vectors. Black planes indicate borders, white planes donate clusters. Within a cluster similarity is higher than between neighbored neurons of a border-region. Emotion-regions are given as well

Experiment 2

Method

Using the same data processing as in Experiment 1, with respect to observing changes in participants' gait patterns when listening to music defined as excitatory and calming, and including a no-music condition, kinetic and kinematic data were recorded from 16 healthy participants different from those in Experiment 1. As excitatory music a track from a German techno group (*Scooter—Back to the heavyweight jam*; 130 beats per minute [bpm]) was chosen, whereas for calming music a track in a world music style (*Bjørnstad/Darling/Christensen/Rypdal—The sea*; free tempo) was selected. The latter track had a free tempo of about 80 bpm, with slowly changing melodies and harmonies. Due to its smoothness it was designated as calming. Differences between these tracks were used to intuitively define music as excitatory or calming. The experimental setup was similar to the first experiment (see Fig. 2). However, since Experiment 1 showed the possibility of recognizing emotions by means of ground reaction forces, kinematic data were also recorded in Experiment 2, in order to scrutinize whether or not visible effects in the gestalt of the movement pattern were observable. Therefore, participants were filmed by two synchronized orthogonally positioned cameras with a frequency of 25 Hz. For kinematic analysis, markers were attached on the top of the manubrium sterni, the left and right acromion, the epicondylus lateralis, the processus steyloideus ulnae, the left and right spina iliaca anterior superior, the trochanter major, the lateral end of the femur (knee), the left and right patella, the articulatio tibo fibulare talare, the calcaneus, the phalanx distalis, and the hallux. Through this, 3D angles and angular velocities of arm, hip, knee, and ankle could be digitized.

Two and a half minutes before and during the walking procedure, participants listened to the randomly pre-assigned music type using headphones. The last double-step before touching the force platform with the right foot was chosen for kinematic data acquisition, beginning and ending with the toes of the right foot leaving the floor. The walking procedure was repeated until data from three error-free gait trials per music type were logged (for criteria see Experiment 1). The derived kinetic data were processed as described in Experiment 1. The processing of the kinematic data was as follows: With the aid of Adobe

Premiere 6.0, the single video sequences were cut to the length of a double-step (see above) before 3D angles and angular velocities of arm, hip, knee, and ankle were manually digitized with the commercially available software Simi Motion 5.0 (SIMI Reality Motion Systems). The generated files were mathematically scaled to the same time-length. Due to the sampling rate of 25 Hz, the smallest number of sampled data vectors was 21, so that all other files were mathematically normalized to this global minimum of 21 discrete values per angle and angular velocity. A further normalization was not necessary as angles and angular velocities were recorded using the same scale for all participants.

To examine person recognition rates, kinetic and kinematic data from gait patterns were fed to a 200-108-16-MLP and a 168-92-16-MLP, respectively. The configuration of the MLPs was the same as described above and the architectures only differed in the number of available or selected data points (200 as in Experiment 1 for the kinetic data and 168 for the kinematic data: $[4 \text{ angles} + 4 \text{ angular velocities}] \cdot 21 \text{ data points}$). Intra-individual emotion recognition on the basis of the kinetic data was managed using the same 5×3 SOM structure as described in the first experiment. However, as the input dimension of the kinematic data (four angles and four angular velocities) was considerably higher, a self-implemented network called 2SOM (due to its structure that consisted of two series-connected SOMs) was chosen for classification. Within the 2SOM, the first SOM (SOM A) had the task of reducing the data dimension, whereas the second SOM (SOM B) took care of the classification (see Fig. 4). This procedure is based on the original work of Bauer and Schöllhorn (1997) and Barton et al. (2006). Further information on these procedures is provided in the Appendix. For both net types, again, time-discrete as well as time-continuous parameters were chosen for input (inside the 2SOM, this differentiation was implemented after the dimension reduction of SOM A).

Results

The average person recognition rate with the MLP was 99.3%. In each case the net was trained with 96 kinetic gait patterns of 16 different participants and tested with another unknown 48 gait patterns. The MLP that was trained and tested with 96 and 48 kinematic gait patterns, respectively, achieved a person recognition rate of 96.9%, a performance rate that was slightly lower than the MLP trained with the kinetic data.

Intra-individual emotion recognition for the participants listening to calming, excitatory, or no music revealed recognition rates of up to 100% for the kinetic and kinematic data. Figure 5 illustrates a representative spread of an individual participant’s kinematic gait

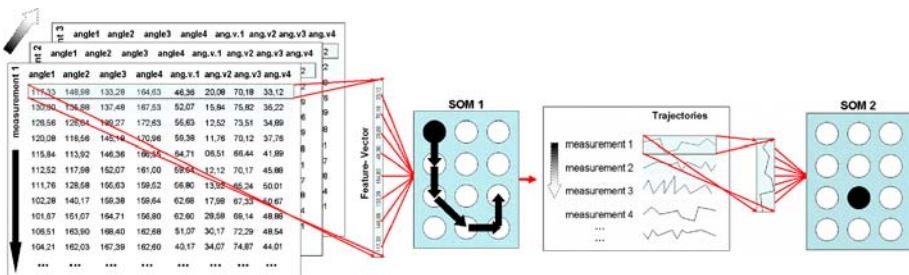


Fig. 4 Architecture of the 2SOM. Trajectories of activated neurons on SOM A are used as input vectors for SOM B

patterns ordered by music types and the unified distance matrix. The gait patterns with *calming* and *no* music in this case seem to be more similar, whereas the pattern with *excitatory* music defines a single cluster. The average recognition rates for the kinetic gait data from all participants pooled was 77.8% ($SD = 9.1$; maximum = 100.0%) and 82.6% ($SD = 9.9$; maximum = 100.0%) for the time discrete and time continuous data, respectively, regarding all participants. On average, the recognition rates for the kinematic data were 73.0% ($SD = 11.5$; maximum = 88.9%) and 79.2% ($SD = 13.4$; maximum = 100.0%) for the time-discrete and time-continuous data, respectively, a performance rate that was slightly lower than obtained with the kinetic data. As in the first experiment, the Mann–Whitney-Test delivered no statistical differences between the two approaches ($Z = -1.49$, $p = .14$ for the kinetic data; $Z = -.40$, $p = .16$ for the kinematic data). In the case of the second experiment, chance level was about 33.3% due to the three types of music. Hence, all rates were significantly above this level.

General Results

Apart from specific results, there were some outcomes common to both experiments. To begin with, the strong individuality of the gait patterns in both experiments warrants attention. Even when combining the kinetic data of the two experiments, person recognition was achieved at a 98.5% success rate for all participants ($n = 38$). Kinematic data are believed to be more convenient for gait pattern recognition since they provide more individual information (Schöllhorn 2004). In our experiments, the recognition rates for kinematic data were slightly lower than for kinetic data. However, the high level of individuality in the gait patterns was not surprising, since in previous work Schöllhorn et al. (2002) were able to distinguish individuals with recognition rates of up to 100%. This finding might explain why inter-individual emotion recognition was not realizable in the present study. The individuality of the gait patterns was just too dominant and gaits were mainly classified by each individual, regardless of simulated emotion.

On a more subtle level, a kind of *finer structure* within the participants' individual gait patterns was discovered, as emotion recognition was achievable with recognition rates of up to 100% (see Figs. 3 and 5). A more detailed analysis of the spreads of data in

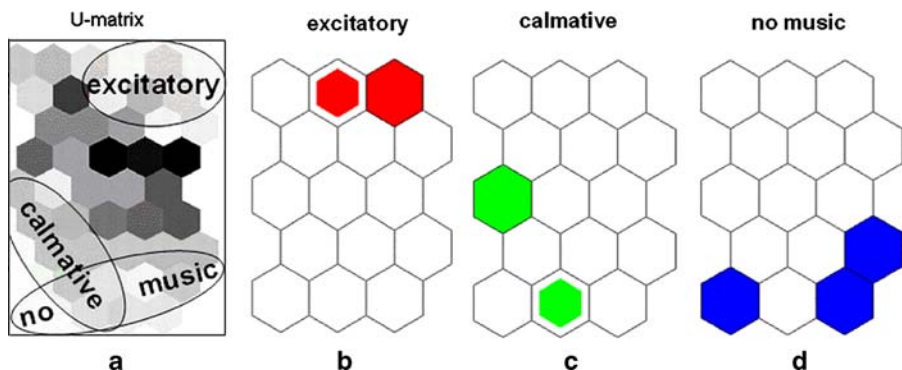


Fig. 5 Unified distance matrix (a) and activated neurons ordered by music types (b–d) of a participant's kinematic gait patterns classified with the 2SOM. The filled hexagons (b–d) indicate the activated neurons for the given music type. The distance matrix (a) shows the real distances between the data vectors. Black planes indicate borders. White planes donate clusters. Emotion-regions are given as well

Experiment 1 showed a clearly visible tendency. In nearly all cases, the gait patterns with the highest degree of differences in arousal levels (Bradley et al. 2001) were clustered well away from each other. Arousal, in this instance, can be viewed as a tendency from low (e.g., *sadness*) to high (e.g., *joy*). In most cases, this trend was also confirmed for the gait patterns in the second experiment. An overview of the recognition rates and results is given in Tables 1 and 2.

Discussion

In the two experiments reported in this paper, kinetic and kinematic gait patterns served as input for three different neural networks (in different configurations) with the task of recognizing individuality and emotional states from individual gait patterns. In the case of person identification, recognition rates of up to 99.3% were achieved by the nets. Inter-individual emotion recognition levels, by means of the same data analysis approach, were not as high and remained around chance level. Hence, the results suggested that intra-individual emotion recognition performance was successfully accomplished using self-organizing maps. Kinetic as well as kinematic gait patterns delivered recognition rates of up to 100%. Even the kinematic data of participants, listening to music while walking, delivered sufficient information for emotion recognition (based on the assumption that emotions in this case were influenced by music). Taken together, the results of both experiments showed: (i) the potential of a more objective emotion recognition approach in human gait with artificial neural nets in a diagnostic way (Experiment 1); and (ii) that by using this diagnosis tool, it was possible to observe changes in participants' gait patterns induced by music while walking (Experiment 2).

Although intra-individual emotion distinction and recognition was successful in most cases, for some participants emotion recognition rates were lower. Thus, it was not possible to identify the type of music listened to by analyzing the gait pattern for all participants. Hence, two consequences can be drawn: First, the approach has to be improved to get better results. Second, emotion expression (more precisely: effects on the gestalt of the movement induced by music) is highly individual and individually pronounced with different levels of magnitudes. Potential improvements in sports or exercise performance by listening to music, for example, may only be promising for a few athletes since not all athletes will benefit from it. In this context, the effects of music on sports performance are controversial (Tenenbaum et al. 2004) and have been reported for physiological processes such as changes in hormone concentrations (e.g., Yamamoto et al. 2003), heart rate (e.g., Guzzetta 1989) or sports performance (e.g., Becker et al. 1994; Ferguson et al. 1994). In the latter studies, participants' performance was increased after listening to specific kinds of music. These changes in performance may, as in this study, be provoked by changes in the characteristic or gestalt of the movement.

Table 1 Person recognition rates with the MLP

	Kinetic data (%)	Kinematic data (%)
Experiment 1	95.3	–
Experiment 2	99.3	96.9
All available data	98.5	–

– Not recorded

Table 2 Recognition rates of intra-individual emotion recognition with the SOM and 2SOM

	Kinetic data (%)	Kinematic data (%)
Experiment 1		
Time discrete	81.3	–
Time continuous	83.7	–
Experiment 2		
Time discrete	77.8	73.0
Time continuous	82.6	79.2

Note. In Experiment 1 chance level was about 25%, in Experiment 2 chance level was about 33%
– Not recorded

Whether or not the relative distances of the observed emotional states (c.f. Figs. 3, 5) provide information about emotional traits needs further research. However, if the *normal* gait pattern is identified as very similar to the simulated *sad* gait pattern, two different implications can be drawn. First, participants were not able to distinguish between *normal* and *sad* simulations of gait at the motor control level. Second, *normal* gait patterns may reflect a *sad* or rather depressive trait of participants. In conclusion, the results suggest further research would be useful, particularly for music therapies or therapies in general where exploring the emotional processes of patients through non-verbal behavioral indices is desired. Self-organizing maps provide a graphical, interpretable output that allows retracing the development and changes in intra-individual gait patterns. In this vein, for example, the development of a “propelling” to “pulling” gait (e.g., Sloman et al. 1982; Sloman et al. 1987) in depressive patients could be monitored and tracked over a specific period of time.

Finally, a further application of the research in this paper may consist of enhancing existing emotion recognition system technology (e.g., emotion recognition in speech or in facial expression). The approach of recognizing emotions through movements may improve the reliability and validity of existing computerized expert systems, leading to developments in security and perhaps clinical usages. For example, since the individual occurrence of emotional expressions differed from participant to participant in the current study, traditional clinical practices with orientation towards general ‘person-independent’ models may not seem to support patients optimally. However, as shown in this paper, when the focus of the therapeutic program is on individual progression, better clinical support may be guaranteed.

Acknowledgments We wish to thank Larry Katz and Veronica Everton-Williams for their useful comments.

Appendix

The Nets

The processing was implemented using Matlab R2006a, the Neural Network Toolbox V5.0 and the SOM Toolbox (Vesanto et al. 2000). Before the data were fed into the networks, a further normalization to the interval $[-1$ to $1]$ was completed in order to prepare the data for the nets. For the kinetic data, the MLP consisted of three layers with 200 neurons (50 x -, 50 y -, and 100 z -data-points) in the first layer, about $(n+c)/2$ neurons in the hidden layer, where n is the number of input neurons and c is the number of desired classes and

also the number of output neurons. The MLPs for the kinematic data were built analogically except for the first layer which contained 168 neurons ($21 \text{ data points} \cdot 8 \text{ angles and angular velocities}$). As an activation function, the tangens hyperbolicus was chosen in all layers. The nets were initialized with the Nguyen–Widrow function (Nguyen and Widrow 1990) and trained with the scaled conjugate gradient algorithm (Møller 1993). Training lengths were set to 500 epochs in general and 600 epochs for the classification with all data from 38 participants respectively. Recognition rates were calculated counting the misclassifications and expressing them as a percentage using cross-validation.

The SOM algorithms were used in two different ways. The normal SOM was used to classify the data as usual; the 2SOM implied two connected SOMs (see Fig. 4). Within this architecture, the first SOM (SOM A) served as data reduction; the second SOM (SOM B) took note of classification tasks. To achieve this, the data vectors were presented as feature-vectors. One single vector then included the values of all angles and angular velocities for a particular point in time, what Bauer and Schöllhorn (1997) call a more ‘coordination-oriented’ approach. In this way, the process information of the movement was represented by the trajectory built by the successively activated neurons in a two-dimensional space. The trajectory of activated neurons was converted to a new data vector by using the x - and y -coordinates from each neuron as new data points. Finally, the second SOM classified the actual measurement. All used SOMs were two-dimensional maps with a hexagonal lattice and a rectangular shape, as recommended by Kohonen et al. (1995). A comparatively small lattice of 5×3 neurons was preferred for the data. As a training method batchtrain was chosen. The learning rate was initially set to 0.5 for the rough training phase and then reduced to 0.05 for the fine tuning phase. Training length was set to 10 epochs in rough training and 40 epochs in fine tuning. Further training was not necessary and yielded no better results. The calculation of recognition rates was achieved as follows: During repeated presentation of training data, the distances from each classified gait pattern to the ‘emotion clusters’ (not the clustering on the SOM, but the collection of gait patterns from one simulated emotion) of all emotions were calculated. If the distance to the emotion-cluster, to which the gait pattern belonged to, was not the shortest, the classification was counted as a misclassification.

References

- Barton, G., Lees, A., Lisboa, P., & Attfield, S. (2006). Visualisation of gait data with Kohonen self-organising neural maps. *Gait & Posture*, *24*, 46–53.
- Bauer, H. U., & Schöllhorn, W. (1997). Self-organizing maps for the analysis of complex movement patterns. *Neural Processing Letters*, *5*, 193–199.
- Becker, N., Brett, S., Chambliss, C., Crows, K., Haring, P., Marsh, C., et al. (1994). Mellow and frenetic antecedent music during athletic performance of children, adults, and seniors. *Perceptual and Motor Skills*, *79*, 1043–1046.
- Benabdelkader, C., Cutler, R. G., & Davis, L. S. (2004). Gait recognition using image self-similarity. *Eurasip Journal on Applied Signal Processing*, *2004*, 572–585.
- Bichot, N. P., & Desimone, R. (2006). Finding a face in the crowd: Parallel and serial neural mechanisms of visual selection. In S. Martinez-Conde, S. L. Macknik, L. M. Martinez, J.-M. Alonso, & P. U. Tse (Eds.), *Progress in brain research; visual perception – fundamentals of awareness: Multi-sensory integration and high-order perception* (pp. 147–156). Elsevier.
- Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion*, *1*, 276–298.
- Camurri, A., Castellano, G., Ricchetti, M., & Volpe, G. (2006). Subject interfaces: Measuring bodily activation during an emotional experience of music. *Gesture in Human-Computer Interaction and Simulation*, *3881*, 268–279.

- Camurri, A., Mazzarino, B., & Volpe, G. (2004). Expressive interfaces. *Cognition, Technology & Work*, 6, 15–22.
- Camurri, A., Lagerlof, I., & Volpe, G. (2003). Recognizing emotion from dance movement: Comparison of spectator recognition and automated techniques. *International Journal of Human-Computer Studies*, 59, 213–225.
- Clarke, T. J., Bradshaw, M. F., Field, D. T., Hampson, S. E., & Rose, D. (2005). The perception of emotion from body movement in point-light displays of interpersonal dialogue. *Perception*, 34, 1171–1180.
- Coombes, S. A., Cauraugh, J. H., & Janelle, C. M. (2006). Emotion and movement: Activation of defensive circuitry alters the magnitude of a sustained muscle contraction. *Neuroscience Letters*, 396, 192–196.
- Coombes, S. A., Janelle, C. M., & Duley, A. R. (2005). Emotion and motor control: Movement attributes following affective picture processing. *Journal of Motor Behavior*, 37, 425–436.
- Coulson, M. (2004). Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of Nonverbal Behavior*, 28, 117–139.
- Cutting, J. E., & Kozlowski, L. T. (1977). Recognizing friends by their walk – gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, 9, 353–356.
- De Meijer, M. (1989). The contribution of general features of body movement to the attribution of emotions. *Journal of Nonverbal Behavior*, 13, 247–268.
- Dittrich, W. H., Troscianko, T., Lea, S. E., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception*, 25, 727–738.
- Ekman, P., & Friesen, W. (1978). *The facial action coding system*. Palo Alto: Consulting Psychologists Press.
- Ferguson, A. R., Carbonneau, M. R., & Chambliss, C. (1994). Effects of positive and negative music on performance of a karate drill. *Perceptual and Motor Skills*, 78, 1217–1218.
- Fragopanagos, N., & Taylor, J. G. (2005). Emotion recognition in human–computer interaction. *Neural Networks*, 18, 389–405.
- Gunes, H., & Piccardi, M. (2005). Fusing face and body display for bi-modal emotion recognition: Single frame analysis and multi-frame post integration. *Affective Computing and Intelligent Interaction, Proceedings*, 3784, 102–111.
- Guzzetta, C. E. (1989). Effects of relaxation and music-therapy on patients in a coronary-care unit with presumptive acute myocardial-infarction. *Heart & Lung*, 18, 609–616.
- Haykin, S. (1998). *Neural networks: A comprehensive foundation*. Prentice Hall PTR.
- Ichimura, T., Ishida, H., Terauchi, M., Takahama, T., & Isomichi, Y. (2001). Extraction of emotion from facial expression by parallel sand glass type neural networks. In *Proceedings of KES 2001, 5th International Conference on Knowledge Based Intelligent Information Engineering Systems and Allied Technology, Osaka, Japan, 6–8 Sept. 2001* (Vol. 2, pp. 988–992). Amsterdam: IOS Press.
- Ioannou, S. V., Raouzaoui, A. T., Tzouvaras, V. A., Mailis, T. P., Karpouzis, K. C., & Kollias, S. D. (2005). Emotion recognition through facial expression analysis based on a neurofuzzy network. *Neural Networks*, 18, 423–435.
- Kaiser, S., & Wehrle, T. (1992). Automated coding of facial behavior in human–computer interactions with FACS. *Journal of Nonverbal Behavior*, 16, 67–84.
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43, 59–69.
- Kohonen, T., Hynninen, J., Kangas, J., & Laaksonen, J. (1995). *The self-organizing map program package, version 3.1*. SOM Programming Team of the Helsinki University of Technology, Laboratory of Computer and Information Science, Espoo.
- Møller, M. F. (1993). A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks*, 6, 525–533.
- Montepare, J., Koff, E., Zaitchik, D., & Albert, M. (1999). The use of body movements and gestures as cues to emotions in younger and older adults. *Journal of Nonverbal Behavior*, 23, 133–152.
- Montepare, J. M., Goldstein, S. B., & Clausen, A. (1987). The identification of emotions from gait information. *Journal of Nonverbal Behavior*, 11, 33–42.
- Nguyen, D., & Widrow, B. (1990). Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. *Proceedings of the International Joint Conference on Neural Networks*, 3, 21–26.
- Nicholson, J., Takahashi, K., & Nakatsu, R. (2000). Emotion recognition in speech using neural networks. *Neural Computing & Applications*, 9, 290–296.
- Nwe, T. L., Foo, S. W., & De Silva, L. C. (2003). Speech emotion recognition using Hidden Markov models. *Speech Communication*, 41, 603–623.
- Park, C. H., Byun, K. S., & Sim, K. B. (2005). The implementation of the emotion recognition from speech and facial expression system. *Lecture Notes in Computer Science*, 3611, 85–88.

- Richardson, M. J., & Johnston, L. (2005). Person recognition from dynamic events: The kinematic specification of individual identity in walking style. *Journal of Nonverbal Behavior*, *29*, 25–44.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*, 533–536.
- Sawada, M., Suda, K., & Ishii, M. (2003). Expression of emotions in dance: Relation between arm movement characteristics and emotion. *Perceptual and Motor Skills*, *97*, 697–708.
- Schöllhorn, W. I. (2004). Applications of artificial neural nets in clinical biomechanics. *Clinical Biomechanics*, *19*, 876–898.
- Schöllhorn, W. I., Jäger, J. M., & Janssen, D. (2008). Artificial neural network models of sports motions. In Y.B. Hong & R. Bartlett (Eds.), *Routledge handbook of biomechanics and human movement science*. Routledge: London.
- Schöllhorn, W. I., Nigg, B. M., Stefanyshyn, D. J., & Liu, W. (2002). Identification of individual walking patterns using time discrete and time continuous data sets. *Gait & Posture*, *15*, 180–186.
- Slovan, L., Berridge, M., Homatidis, S., Hunter, D., & Duck, T. (1982). Gait patterns of depressed patients and normal subjects. *American Journal of Psychiatry*, *139*, 94–97.
- Slovan, L., Pierrynowski, M., Berridge, M., Tupling, S., & Flowers, J. (1987). Mood, depressive illness and gait patterns. *Canadian Journal of Psychiatry*, *32*, 190–193.
- Tenenbaum, G., Lidor, R., Lavyan, N., Morrow, K., Tonnel, S., Gershgoren, A., et al. (2004). The effect of music type on running perseverance and coping with effort sensations. *Psychology of Sport and Exercise*, *5*, 89–109.
- Vesanto, J., Himberg, J., Alhoniemi, E., & Parhankangas, J. (2000). *SOM Toolbox for Matlab 5. Report A57*. Helsinki University of Technology, Neural Networks Research Centre, Espoo.
- Walk, R. D., & Homan, C. P. (1984). Emotion and dance in dynamic light displays. *Bulletin of the Psychonomic Society*, *22*, 437–440.
- Wallbott, H. G., & Scherer, K. R. (1986). Cues and channels in emotion recognition. *Journal of Personality and Social Psychology*, *51*, 690–699.
- Yamamoto, T., Ohkuwa, T., Itoh, H., Kitoh, M., Terasawa, J., Tsuda, T., et al. (2003). Effects of pre-exercise listening to slow and fast rhythm music on supramaximal cycle performance and selected metabolic variables. *Archives of Physiology and Biochemistry*, *111*, 211–214.