Affective Computing For Empathic Behaviour Change

Mathias Fuchs
University of Bern
Master of computer science
Matr. Nr.: 09-923-764
Email: fuchsmat@students.unibe.ch

***Abstract*—Humans strive to build machines that can interact with humans in a humanoid way. This is why it is crucial for a computer to be able to understand in which emotional state the user is in. To achieve such a feat there are different approaches. Within the research area of affective computing, a large part of the studies focusses on facial expressions and changes in speech. These expressions are good to recognize the emotional state of a human during social interaction, however they may not be suitable in other situations for example recognizing emotions from a greater distance [1]. In this paper I give an overview over the body language recognition approaches done today and propose a model which analyses emotions based on the way a human subject walks.**

I. Introduction

For a long time people were convinced that human behaviour is "all nurture and no nature" [2]. However already Darwin [3] suggested that along with the facial expression, the human body movements and the gestures also represent the state of mind and the corresponding emotions of humans. We know today that body language pays a very important contribution to understand the affective state of a person [4], [5]. Surprisingly only 7% of human communication are made of words and 55% are made up of non-verbal communication [2], [6], [7]. The idea of this paper was to focus on micro expressions in body language. However there are no bodily micro expressions as in facial micro expressions. A micro expression is a "very fast facial movement lasting less than one-fifth of a second" [2]. Body language in comparison can be subconscious however it can be consciously changed more easily, than a facial micro expression. This is the reason that I focus on the emotion detection via body language in this paper.

There are various uses for emotion detection by body language. Some of those are detecting the *affective state* of a person, *lie detection*, the degree of *accessibility* towards another person etc. Different indicators for interpreting body language can be *body position and distance* [5], [7], *body movement* [1], [6] and *hand form* [8], [9]. This list is not necessarily concluding, but those are the parts that I focussed on in this work.

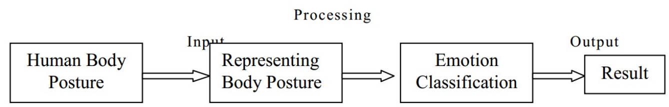

Fig. 1. The phases of a bodily emotion detection system [6]

In the field of affective computing however, it is not only a challenge to interpret the body language, but also to detect the position of the human posture and get useful data out of it. This leads to the two main challenges: *1. Detection of the posture* and *2. How to interpret the posture representation* (see Figure 1 [6]).

In the following work I present papers that provide approaches for both, the detection of the human posture as well as for interpreting the found results. In the following sections first some general things about body language are explained, then I present various papers with different approaches. Finally I propose a theoretical model of how to recognize human emotion by gait.

II. Emotions through body language

As already stated, 55% of our communication consists of non-verbal cues, like body language. The expression of emotions has been studied extensively [10]–[13]. According to Eckman [2] there are 6 basic emotions:

- Anger
- Disgust
- Fear
- Happiness
- Sadness
- Surprise

We can find clear signs of all those emotions in our faces (see Figure 2[1]), because those signs are involuntary micro expressions. Recently more research has been done in the field of detecting emotions through body language, like body movement and body pose [14]–[16]. In Figure 3 we can see a representation of the body pose for the 6 basic emotions. In Figure 4 we can see a possible way on how to interpret certain bodily signs based on the body position and body movement.

The form of the hand can also give information about the emotional state of a person. For example open palms could

[1]https://hubpages.com/health/Facial-Expressions-Emotions-and-Feelings

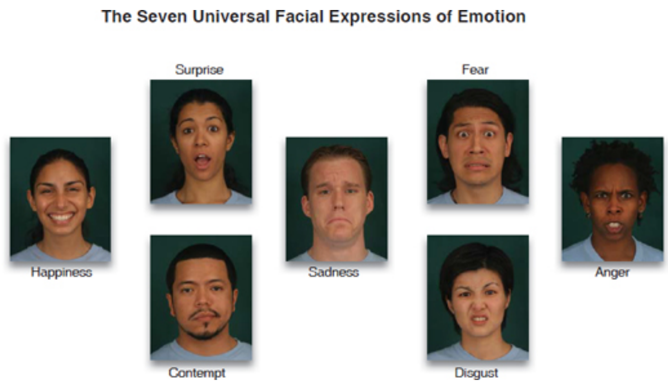

Fig. 2. Facial expression of the six basic emotions (and contempt)

Fig. 3. Bodily signs of the six basic emotions [19]

mean pleasure/ openness, closed hands towards the chest could mean a sense of pride and clenched fists could mean anger [17], [18].

In the following sections I summarize a few approaches that extract and/or interpret body language in different ways.

III. Emotion detection through body position

A. Lie Detection based on Facial Micro Expression, body Language and Speech Analysis

A very interesting approach was done by Barathi [20].

1) Body pose extraction: To extract the body poses from videos the Limb Action Model Converter [21] has been used. This converter uses Microsoft Kinect as a base. The Limb Action Model extracts 10 limbs from a body posture: "Spine to center shoulder, center shoulder to head, left/right shoulder to left/right elbow, left/right elbow to left/right wrist, left/right hip to left/right knee and left/right knee to left/right ankle" [21]. In Figure 5 we can see how the posture is represented after the extraction.

Emotion	Body Posture
Anger	Head backward, no chest backward, no abdominal twist, arms raised forwards and upwards, shoulders lifted.
Joy	Head backward, no chest forward, arms raised above shoulder and straight at the elbow, shoulders lifted.
Sadness	Head forward, chest forward, no abdominal twist, arms at the side of the trunk, collapsed posture.
Surprise	Head backward, chest backward, abdominal twist, arms raised with straight forearms.
Pride	Head backward or lightly tilt, expanded posture, hands on the hips or raised above the head.
Fear	Head backward, no abdominal twist, arms are raised forwards, shoulders forwards.
Disgust	Shoulders forwards, head downwards.
Boredom	Collapsed posture, head backwards not facing the interlocutor.

Fig. 4. A table of bodily signs for different emotions [6]

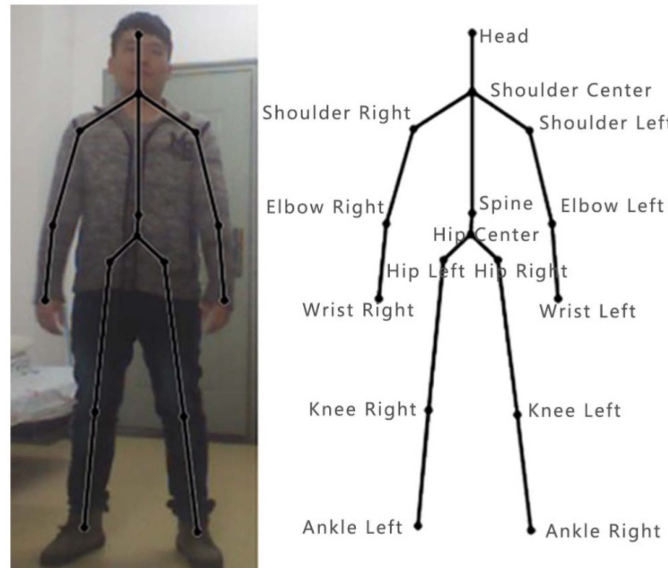

Fig. 5. Skeleton joints extracted with the Limb Angle Model [21]

2) Interpretation: As the paper focusses on lie detection, they defined the following signs for lying [20], [22], [23]:

- Increasing hand to face/mouth gestures
- Nose touching: Because of an adrenaline rush, the capillaries open up, which causes the nose to itch
- Place the hand close to or over the mouth
- Small gestures like lip biting, hands rubbing, fidgeting
- Clenched fist, crossed arms

According to those specifications they trained their system with images of liars that were exhibiting those typical body language cues for lying and pictures of people who were not lying. All the images are subjected to the Limb Action Model Converter. Finally they clustered the converted pictures with k-means. Sadly they did not provide an evaluation for the body language part of their method.

B. EDBL - Algorithm for Detection and Analysis of Emotion Using Body Language

1) Body pose extraction: EDBL relies on a pose estimation which is using postelets for humnan parsing [24]. Different postelets are extracted from images. A linear SVM classifier is trained for detecting the presence of each postelet. In the end we get a complete model of the human. See Figure 6 for a graphical representation of the process.

Based on the postelet representation a line graph (see Figure 7 of the extracted pose is created and used for the interpretation of the body posture.

2) Interpretation: To figure out the emotional state of a person, the position of the shoulders is interpreted. There is

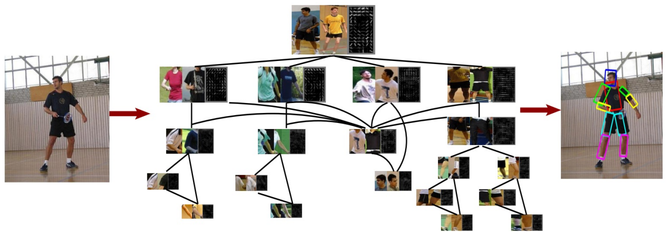

Fig. 6. Graphical illustration of the postelet model [24]

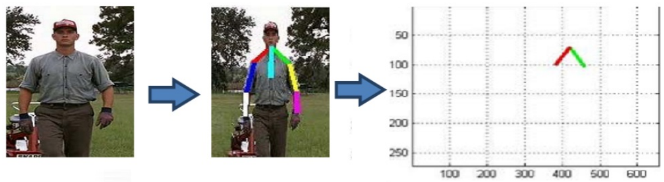

Fig. 7. EDBL stick pose representantion [18]

differentiated between three different types of poses: 1. normal/ calm, 2. confused, amazed or in doubt and 3. depressed/ not interested. Figure 8 shows the three different types. First the slope of the normal shoulder position θ is calculated. An initialisation pose is needed. Then it is compared to other images of the same person and if the slope is bigger than the slope of θ then it is a sign of doubt, confusion or being amazed. If the slope is smaller than the slope of θ then it is a sign of a depressed, not-interested or lazy emotional state.

IV. Affect Detection from Body Language during Social HRI

1) Body pose extraction: This approach of McColl et al [5] also uses Microsoft Kinect as a base to extract the body pose. It creates an ellipsoid model as seen in Figure 9. To do this, the Kinect sensor first performs a human body extraction. Then the extracted body poses are observed to see if a static pose is displayed. Once a static pose is identified the segmented body parts are fitted with ellipsoids [5]. The static body poses are further explained in subsubsection IV-2.

Fig. 8. Meaning of shoulder positions. (a) shows a person in normal or calm position, (b) shows a person in confused or amazed state and (c) shows a person in a depressed or not interested pose. [18]

Fig. 9. An ellipsoid model od the human body built on the basis of Microsoft Kinect. [18]

Trunk Orientation	Accessibility Level	Arm Orientation	Finer-Scaling
Upper/Lower trunk: T/N or N/T combined with upright or forward leans, T/T with all possible leans	IV	T	12
		N	11
		A	10
Upper/Lower trunk: T/N or N/T except positions that involve upright or forward leans	III	T	9
		N	8
		A	7
Upper/Lower trunk: N/N, A/N, N/A, T/A, A/T with all possible leans	II	T	6
		N	5
		A	4
Upper/Lower trunk: A/A with all possible leans	I	T	3
		N	2
		A	1

Fig. 10. Different accessibility levels defined to determine the level of accessibility of the person against the system [5]

2) Interpretation: To figure out the emotional state of the subject, static body poses are used. The poses are based on the Nonverbal Interaction States Analysis of the Davis Nonverbal States Scale [25]. Body angle, trunk lean and arm position are evaluated. The resulting metrics are the following:

- **Three different body angles:** *Toward(T)*: 0-3 angle from the robot, *Neutral(N)*: 3-15 angle from the robot, *Away(A)*: ¿15 from the robot
- **Trunk lean:***Upright*: The shoulders are over the hips, *Forward/ backward lean*: The shoulders are closer/ farther away than the hips in relation to the robot, *Right/ left lean*: The right/ left shoulder is tilted past the right/ left hip.
- **Arm positions:** *T*: The arms are closer to the robot than the upper trunk, *A*: The arms are farther from the robot than the trunk, *N*: else

According to those metrics 4 accessibility levels [5] have been defined (see Figure 10). Level I-IV, where I is least accessible and IV is the most accessible state. The arm orientation is used for a finer scaling of the accessibility levels.

This approach classified acessibility with an accuracy of 88% [5].

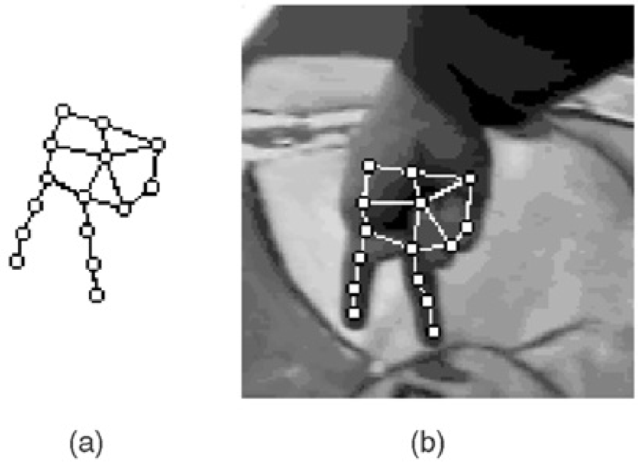

Fig. 11. Picture of an elastic graph of a hand, and a new hand that matches it [9]

V. Emotion detection through hand position

A. A system for person-independent hand posture recognition against complex backgrounds

1) Hand pose extraction: Most approaches in this field, require the hand to be in front of a static background or the hand to be the only skin coloured item in the picture, in order for it to be recognized [26]–[29]. The approach of Triesch and Malsburg [9], proposes a solution using elastic graph matching (EGM) with multiple feature types. EGM is "a neurally inspired object recognition architecture" [30]. In EGM the views of objects are visualized as a labeled 2-d graph. [9]. The nodes of the graph contain a local image description and the edges are labelled with a distance vector which represents the distance between the nodes.

2) Interpretation: Various graphs are created from training images and then images of new hands are matched to the most similar existing image. In **??** we can see an example of a graph, which is matched to a new hand. Like this we can easily recognize if a hand is clenched to a fist (e.g. anger).

The reported accuracy of this approach is 92.9% accuracy on simple backgrounds and 85.8% accuracy on complex background.

VI. Emotion detection through gait

VII. Emotion detection through gait

VIII. Conclusion

IX. Future work

References

[1] M. Karg, "Pattern recognition algorithms for gait analysis with application to affective computing," Ph.D. dissertation, Technische Universität München, 2012.
[2] P. Eckman, "Emotions revealed," *St. Martins Griffin, New York*, 2003.
[3] C. Darwin, *The expression of the emotions in man and animals, 2nd ed.*, London, 1980.
[4] P. Ekman and W. V. Friesen, "Head and body cues in the judgment of emotion: A reformulation," *Perceptual and motor skills*, vol. 24, no. 3 PT 1, pp. 711–724, 1967.
[5] D. McColl and G. Nejat, "Affect detection from body language during social hri," in *RO-MAN, 2012 IEEE*. IEEE, 2012, pp. 1013–1018.
[6] S. Singh, N. Sethi, and V. Sharma, "Significance of bodily movement for detection and analysis of emotions: A."
[7] A. Mehrabian, "Significance of posture and position in the communication of attitude and status relationships." *Psychological Bulletin*, vol. 71, no. 5, p. 359, 1969.
[8] C.-C. Wang and K.-C. Wang, "Hand posture recognition using adaboost with sift for human robot interaction," in *Recent progress in robotics: viable robotic service to human*. Springer, 2007, pp. 317–329.
[9] J. Triesch and C. Von Der Malsburg, "A system for person-independent hand posture recognition against complex backgrounds," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 12, pp. 1449–1453, 2001.
[10] S. S. Tomkins, *Affect, imagery, consciousness*. Springer Verlag, 1962.
[11] P. EKMANN, "Universal facial expressions in emotion," *Studia Psychologica*, vol. 15, no. 2, p. 140, 1973.
[12] P. Ekman, "Facial expression and emotion." *American psychologist*, vol. 48, no. 4, p. 384, 1993.
[13] N. H. Frijda, *The emotions*. Cambridge University Press, 1986.
[14] B. De Gelder, "Towards the neurobiology of emotional body language," *Nature Reviews Neuroscience*, vol. 7, no. 3, p. 242, 2006.
[15] J. Grezes, S. Pichon, and B. De Gelder, "Perceiving fear in dynamic body expressions," *Neuroimage*, vol. 35, no. 2, pp. 959–967, 2007.
[16] H. K. Meeren, C. C. van Heijnsbergen, and B. de Gelder, "Rapid perceptual integration of facial expression and emotional body language," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 45, pp. 16 518–16 523, 2005.
[17] R. Beale and C. Peter, *Affect and emotion in human-computer interaction*. Springer, 2008.
[18] S. Singh, V. Sharma, K. Jain, and R. Bhall, "Edbl-algorithm for detection and analysis of emotion using body language," in *Next Generation Computing Technologies (NGCT), 2015 1st International Conference on*. IEEE, 2015, pp. 820–823.
[19] K. Schindler, L. Van Gool, and B. de Gelder, "Recognizing emotions expressed by body pose: A biologically inspired neural model," *Neural networks*, vol. 21, no. 9, pp. 1238–1246, 2008.
[20] S. Barathi. C, "Lie detection based on facial micro expression, body language and speech analysis," vol. V5, 02 2016.
[21] J. Du and D. Chen, "3d action recognition based on limb angle model," in *Information Science and Technology (ICIST), 2014 4th IEEE International Conference on*. IEEE, 2014, pp. 304–307.
[22] C. Kinsey Goman, "12 ways to spot a liar at work," *Forbes.com*, 2012.
[23] M. Soorjoo, *The Black Book of Lie Detection*, 2009.
[24] Y. Wang, D. Tran, and Z. Liao, "Learning hierarchical poselets for human parsing," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 1705–1712.
[25] M. Davis and D. Hadiks, "Non-verbal aspects of therapist attunement," *Journal of Clinical Psychology*, vol. 50, no. 3, pp. 393—405, 1994.
[26] Y. Cui, D. L. Swets, and J. J. Weng, "Learning-based hand sign recognition using shoslif-m," in *Computer Vision, 1995. Proceedings., Fifth International Conference on*. IEEE, 1995, pp. 631–636.
[27] E. Hunter, "Posture estimation in reduced-model gesture input systems," in *Proc. Int'l Workshop on Automatic Face and Gesture Recognition*, 1995, pp. 290–295.
[28] M. Kohler, *Technical details and ergonomical aspects of gesture recognition applied in intelligent home environments*. Citeseer, 1997.
[29] B.-W. Min, H.-S. Yoon, J. Soh, T. Ohashi, and T. Ejima, "Visual recognition of static/dynamic gesture: Gesture-driven editing system," *Journal of Visual Languages & Computing*, vol. 10, no. 3, pp. 291–309, 1999.
[30] M. Lades, J. C. Vorbruggen, J. Buhmann, J. Lange, C. Von Der Malsburg, R. P. Wurtz, and W. Konen, "Distortion invariant object recognition in the dynamic link architecture," *IEEE Transactions on computers*, vol. 42, no. 3, pp. 300–311, 1993.