

Towards automatic analysis of gestures and body expressions in depression

Marwa Mahmoud
University of Cambridge
Computer Laboratory
Cambridge, UK
marwa.mahmoud@cl.cam.ac.uk

Peter Robinson
University of Cambridge
Computer Laboratory
Cambridge, UK
peter.robinson@cl.cam.ac.uk

ABSTRACT

Depression is a common mental disorder and one of the main causes of disease burden worldwide. Several studies in depression address the relation between non-verbal cues and different levels of depression. Manual coding of non-verbal cues is the common practice for running such studies, which is time consuming and non-objective. Recent research has looked into automatic detection of cues associated with depression. However, most of the work has focussed on facial cues such as facial expressions, gaze and head pose. Few studies have looked into multimodal features for analysis of depression, mainly focusing on facial movements, head movements and vocal prosody. Body gestures are an understudied modality in that field. We propose to investigate assessment of depression using automatic detection of non-verbal signals of body gestures. Moreover, we propose the use of multimodal fusion of features to incorporate body as well as face and head for better inference of depression level. Automatic analysis of such body cues can serve as a tool for experimental psychologists. Also, it can assist physicians in diagnosing by providing quantitative measures after or during face to face sessions or telemedicine sessions or even in systems like a virtual coach.

Keywords

Gestures; Body expressions; Assessment of depression; Mental health

1. INTRODUCTION

Nonverbal communication plays a central role in how humans communicate and empathize with each other. The ability to read nonverbal cues is essential to understanding, analyzing, and predicting the actions and intentions of others. Nonverbal cues include facial expressions, hand gestures, body posture and tone of voice. These nonverbal cues may indicate expression of emotions and mental states or even some medical conditions such as pain, depression

and anxiety. As technology becomes more widespread and personalised, fields like ‘affective computing’ and ‘socially aware computing’ have emerged to investigate the possibility of building machines and intelligent virtual agents that can sense non-verbal signals and respond to natural human behaviour.

2. NON-VERBAL CUES IN DEPRESSION

Depression is a common mental disorder and one of the leading causes of disease burden worldwide [15]. Moreover, psychological disorders, such as depression and anxiety, are common conditions that are frequently experienced by people with dementia and their carers. Several studies in depression address the relation between non-verbal cues and different levels of depression. Manual coding of non-verbal cues is the common practice for running such studies, which is time consuming and non-objective. Recent research has looked into automatic detection of cues associated with psychological disorders, like depression [18]. Most of this work has focussed on facial cues such as facial expressions, gaze and head pose [1] or on physiology measurements [17, 16]. Few studies have looked into multimodal features, especially for analysing depression, mainly focusing on facial movements, head movements and vocal prosody [6, 9, 20].

Body gestures are an understudied modality in that field. Joshi et al [10] looked into relative body parts movements in assessment of depression, but they did not consider any specific hand gestures or body expressions. Several studies in affective computing show that body expressions and gestures are significant visual cues that can complement facial expressions [5] and they can be utilised in automatic detection of human internal states [3, 2].

3. PROPOSED APPROACH

We propose to investigate assessment of depression using automatic detection of non-verbal signals of gestures and body expressions as novel cues that can help in automatic assessment of psychological distress. Moreover, we propose the use of multimodal fusion of features to incorporate body expressions as well as the face and head features for better inference of depression level. For example, depressed patients often exhibit more ‘self-touches’ or self-adaptors [18]. Fidgeting - which includes gestures such as tapping or rhythmically shaking hands or feet - is often seen and reported in both anxiety and depression [8]. Depressed patients also often engage in “self-adaptors” [7], such as rhythmically touching, hugging or stroking parts of the body or self-grooming, such as repeatedly stroking their

hair [8]. Scherer et al. [18] studied indicators of psychological distress, including Depression, Anxiety and PTSD. They demonstrated that subjects with psychological conditions exhibit on average longer self-touches and fidget with both hands (e.g. rubbing, stroking) and legs (e.g. tapping, shaking).

To that end, our proposed approach is to build on our previous work on automatic detection of self-touches, especially on the face area [14], and the work on automatic detection of multimodal descriptors of rhythmic body movements associated with psychological distress [13]. Proposed features include: self-touch, hand shape, hand action, fidgeting behaviour and overall body/joints movements. We propose using spatiotemporal features - such as Space-time interest points [11, 12, 19, 14] and Histograms of Oriented gradients [4, 21] - that proved to be successful in automatic classification of challenging behaviour descriptors.

4. CONCLUSION

We have presented briefly our motivation and proposed approach in investigating automatic analysis of gestures and body expressions in assessment of depression. Body expressions are an understudies modality in this emerging field. This might be due to the many challenges involved in this area of research, such as availability of datasets, annotations, building the appropriate gesture taxonomy and evaluation. Recent studies have looked into automatic detection of facial expressions and head gestures in assessment of psychological disorders. However, we argue that analysis of gesture and body expressions is an additional modality that is worth exploring.

5. REFERENCES

- [1] S. Alghowinem, R. Goecke, J. F. Cohn, M. Wagner, G. Parker, and M. Breakspear. Cross-cultural detection of depression from nonverbal behaviour. In *IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, volume 1, pages 1–8. IEEE, 2015.
- [2] D. Bernhardt and P. Robinson. Detecting affect from non-stylised body motions. In *Affective Computing and Intelligent Interaction*, pages 59–70. Springer, 2007.
- [3] G. Castellano, S. D. Villalba, and A. Camurri. Recognising human emotions from body movement and gesture dynamics. In *Affective computing and intelligent interaction*, pages 71–82. Springer, 2007.
- [4] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition*, volume 1, pages 886–893. IEEE, 2005.
- [5] B. de Gelder. Why bodies? twelve reasons for including bodily expressions in affective neuroscience. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1535):3475–3484, 2009.
- [6] H. Dibeklioğlu, Z. Hammal, Y. Yang, and J. F. Cohn. Multimodal detection of depression in clinical interviews. In *ACM International Conference on Multimodal Interaction*, pages 307–310. ACM, 2015.
- [7] P. Ekman and W. V. Friesen. The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, 1(1):49–98, 1969.
- [8] L. A. Fairbanks, M. T. McGuire, and C. J. Harris. Nonverbal interaction of patients and therapists during psychiatric interviews. *Journal of abnormal psychology*, 91(2):109, 1982.
- [9] J. M. Girard and J. F. Cohn. Automated audiovisual depression analysis. *Current opinion in psychology*, 4:75–79, 2015.
- [10] J. Joshi, A. Dhall, R. Goecke, and J. F. Cohn. Relative body parts movement for automatic depression analysis. In *Affective Computing and Intelligent Interaction (ACII)*, pages 492–497. IEEE, 2013.
- [11] I. Laptev. On space-time interest points. *International Journal of Computer Vision*, 64(2-3):107–123, 2005.
- [12] I. Laptev, M. Marszałek, C. Schmid, and B. Rozenfeld. Learning realistic human actions from movies. In *Computer Vision and Pattern Recognition*, pages 1–8. IEEE, 2008.
- [13] M. Mahmoud, L.-P. Morency, and P. Robinson. Automatic multimodal descriptors of rhythmic body movement. In *ACM International Conference on Multimodal Interaction*, pages 429–436. ACM, 2013.
- [14] M. M. Mahmoud, T. Baltrušaitis, and P. Robinson. Automatic detection of naturalistic hand-over-face gesture descriptors. In *ACM International Conference on Multimodal Interaction*, pages 319–326. ACM, 2014.
- [15] C. D. Mathers and D. Loncar. Projections of global mortality and burden of disease from 2002 to 2030. *Plos med*, 3(11):e442, 2006.
- [16] D. J. McDuff, J. Hernandez, S. Gontarek, and R. W. Picard. Cogcam: Contact-free measurement of cognitive stress during computer tasks with a digital camera. In *Computer and Human Interaction Conference (CHI)*. ACM, 2016.
- [17] R. W. Picard, S. Fedor, and Y. Ayzenberg. Multiple arousal theory and daily-life electrodermal activity asymmetry. *Emotion Review*, 8(1):62–75, 2015.
- [18] S. Scherer, G. Stratou, G. Lucas, M. Mahmoud, J. Boberg, J. Gratch, L.-P. Morency, et al. Automatic audiovisual behavior descriptors for psychological disorder analysis. *Image and Vision Computing*, 32(10):648–658, 2014.
- [19] Y. Song, L.-P. Morency, and R. Davis. Learning a sparse codebook of facial and body microexpressions for emotion recognition. In *ACM International conference on multimodal interaction*, pages 237–244. ACM, 2013.
- [20] M. Valstar, B. Schuller, K. Smith, T. Almaev, F. Eyben, J. Krajewski, R. Cowie, and M. Pantic. AVEC 2014: 3d dimensional affect and depression recognition challenge. In *the 4th International Workshop on Audio/Visual Emotion Challenge*, pages 3–10. ACM, 2014.
- [21] X. Zhu and D. Ramanan. Face detection, pose estimation, and landmark localization in the wild. In *Computer Vision and Pattern Recognition*, pages 2879–2886. IEEE, 2012.