Head Pose and Movement Analysis as an Indicator of Depression

Sharifa Alghowinem[1,5], Roland Goecke[2,1], Michael Wagner[2,1], Gordon Parker[3], Michael Breakspear[4,3]

[1]*Australian National University, Canberra, Australia,*
[2]*University of Canberra, Canberra, Australia,*
[3]*University of New South Wales, Sydney, Australia,*
[4]*Queensland Institute of Medical Research, Brisbane, Australia,*
[5]*Ministry of Higher Education: Kingdom of Saudi Arabia,*

`sharifa.alghowinem@anu.edu.au, roland.goecke@ieee.org, michael.wagner@canberra.edu.au,`
`g.parker@blackdog.org.au, mjbreaks@gmail.com`

Abstract—**Depression is a common and disabling mental health disorder, which impacts not only on the sufferer but also their families, friends and the economy overall. Our ultimate aim is to develop an automatic objective affective sensing system that supports clinicians in their diagnosis and monitoring of clinical depression. Here, we analyse the performance of head pose and movement features extracted from face videos using a 3D face model projected on a 2D Active Appearance Model (AAM). In a binary classification task (depressed vs. non-depressed), we modelled low-level and statistical functional features for an SVM classifier using real-world clinically validated data. Although the head pose and movement would be used as a complementary cue in detecting depression in practice, their recognition rate was impressive on its own, giving 71.2% on average, which illustrates that head pose and movement hold effective cues in diagnosing depression. When expressing positive and negative emotions, recognising depression using positive emotions was more accurate than using negative emotions. We conclude that positive emotions are expressed less in depressed subjects at all times, and that negative emotions have less discriminatory power than positive emotions in detecting depression. Analysing the functional features statistically illustrates several behaviour patterns for depressed subjects: (1) slower head movements, (2) less change of head position, (3) longer duration of looking to the right, (4) longer duration of looking down, which may indicate fatigue and eye contact avoidance. We conclude that head movements are significantly different between depressed patients and healthy subjects, and could be used as a complementary cue.**

Index Terms: depression, head pose, mood classification

I. INTRODUCTION

Clinical depression, unlike mood fluctuation, is a common mental disorder that lasts longer and causes disability and reduced functionality. A recent World Health Organization (WHO, 2012) survey estimated that 350 million people worldwide are affected by depression. With increased severity level, it might lead to suicide. Moreover, it causes more than two-thirds of suicides each year [1]. The suicide risk is more than 30 times higher among depressed patients than that of the population without these disorders [2]. Although treatment of depression disorders has proven to be effective in most cases [3], misdiagnosing depressed patients is a common barrier [4]. Moreover, the assessment methods of diagnosing depression rely almost exclusively on patient-reported or clinical judgments of symptom severity [5], risking a range of subjective biases. Based on the WHO, the barriers to effective diagnosis of depression include a lack of resources and trained health care providers. Our goal here is to investigate the general characteristics of depression, which we hope will lead to an objective affective sensing system that assists clinicians in their diagnosis and monitoring of clinical depression. Ultimately, we hope to assist patients with depression to monitor the progress of their illness.

In this study, we investigate whether head pose and movement hold discriminative power for detecting depression. We extract head pose and movement features from face videos using a 3D face model projected onto person-specific Active Appearance Models (AAM), trained with 46 points. We examine the performance of these features using Gaussian Mixture Models (GMM) and Support Vector Machines (SVM) for classification on a real-world, clinically validated dataset of 30 patients with severe depression and 30 healthy control subjects. We also analyse the statistical significance of those features between the two groups. In this paper, we investigate the differences in the eye movement while expressing positive and negative emotions, as well as differences in males and females from both groups.

II. BACKGROUND

Psychological research on depressives' head movements have received far less attention compared with the studies on gestures and facial expressions. Simple behaviours such as head movement could reflect cues about mood, emotions, personality, or cognitive processing [6]. An ethological study on depressed patients behaviour noticed that behavioural elements are pronounced more in the head and hand regions compared with other body regions [7]. A study on the social behaviour of depressed patients found that they show significantly less head nodding than healthy controls [8]. Another study found that depressed subjects are more likely to position their head downward than healthy ones [9]. Studying eye contact, [8] found that depressed patients engage in less eye contact with others than non-depressed

persons, illustrated by gazing away and a downward head pose. A study investigating depressed patients' involvement in a conversation, showed a low involvement reflected by less head nodding, less head movements, less looking and gesturing during speech [10].

Automatic head pose estimation using computer vision techniques has been surveyed in [11]. Such methods include template matching, where the head image is matched to the nearest group of images that approximate the head pose. Such a method requires a huge variety of head pose images. Deformable models are another method for head pose estimation, including the AAM, where specific facial points are labelled and trained to create a 2D model, then the head pose is estimated using the direction of the first principal component of the principal components analysis [11]. Another method of head pose estimation is determining the geometry between local features, such as the eyes, mouth, and nose tip. Another technique approximated the head pose using the shape of the head by a 3D cylinder [12]. Knowing that not all methods could estimate all three dimensions of the head pose (pitch, yaw and roll), [13] estimated 3D head pose by combining AAM with Pose from Orthography and Scaling with ITerations (POSIT) [14]. POSIT is an algorithm that detects and matches at least 4 points in a 2D image to a 3D object, then finds the geometry of the 3D object (orientation and translation) [14]. Once the estimated orientation is calculated, the pose of the object could be extracted. POSIT is a useful alternative to popular pose algorithms because it does not require an initial pose estimation, and because it is easy to implement as well as faster to run [14]. Inspired by [13], in this paper, we also estimate the head pose using POSIT algorithm with 2D points of face AAM projected into 3D face model.

To the best of our knowledge, the head pose and movement analysis in this paper is the first attempt for automatic depression detection. We analyse the differences between depressed patients and healthy controls in head pose and movement behaviour in a recorded interview. Beside analysing the differences in head pose and movements in general, we specifically investigate these differences while expressing positive and negative emotions. In addition, we investigate the potential effect of gender-dependent classification of depression. After extracting features from each frame in the video recordings, we use a hybrid classifier based on a GMM for each subject and a SVM for classification. We also analyse the head pose and movement features statistically in terms of functionals to identify how head pose and movement differ between both groups.

III. METHOD

A. Real-World Clinically Validated Data

The Black Dog Institute, which is a clinical research facility in Sydney, Australia offering specialist expertise in depression and bipolar disorder, collects data from patients

Table I
DURATION OF THE INTERVIEWS (IN MIN) OF DEPRESSED AND CONTROL SUBJECTS

Questions	Both genders		Male Only		Female Only	
	Depressed	Control	Depressed	Control	Depressed	Control
All 8 questions	309.2	199.8	126.6	112.0	182.5	87.8
"Good News"	34.7	21.3	14.2	9.8	20.5	11.5
"Bad News"	37.6	28.7	16.5	17.9	21.0	10.8

who had been diagnosed with depression and from healthy controls, who have no history of mental illness. To date, data from over 40 depressed subjects plus over 40 controls (both females and males) has been collected.

In this paper, a gender-balanced subset of 30 depressed subjects and 30 controls was analysed. We acknowledge that the amount of data used here is relatively small, but this is a common problem in similar studies [15], [16]. As we continue to collect more data, future papers will be able to report on a larger dataset. The audio-video experimental paradigm contains several parts, including an interview with the subjects. The interview was conducted by asking specific open questions (in 8 question groups) to describe events that had aroused significant emotions. In this paper, the interview part with all 8 question groups is used for analysing spontaneous head pose and movements. We also compare head pose and movements when expressing positive and negative emotions by analysing two related questions from the interview: "Can you recall some recent *good news* you had and how did that make you feel?" and "Can you recall news of *bad or negative nature* and how did you feel about it?" For simplicity, these two questions will be referred to as "Good News" and "Bad News", resp. We assume that these questions elicit the emotions, even though the answers were not validated for certain emotions. The interview was manually labelled to separate each question, where the total duration was 509min for all 60 subjects used in this study. The duration of the interviews of depressed and control subjects for all 8 questions, the "Good News" question and the "Bad News" question used in this paper is shown in Table I.

B. Building AAM

On average, 30 images were automatically selected (almost every 250 frames, 25 fps) per subject having different head position variation. Those images were annotated using 68 points (see Figure 1 (a)). These annotated points and images were used to build subject-specific face AAM, using linear parameters to update the model in an iterative framework as a discriminative fitting method [17]. The points of the trained model were initialised in a semi-automated manner, i.e. face detection [18] was performed in the first frame, then a fitting function was called. If the fitting function was not accurate enough, the steps of changing the model location and the calls for the fitting function
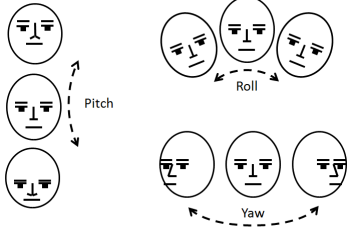
Figure 2. Head rotation angles: Yaw, roll and pitch

were manually repeated before the tracking was started for the entire video. The trained model fits on and tracks the subject's face for the entire interview, producing the position of the 68 landmarks in each frame.

C. Head Pose estimation and the 3D Face Model Fitting

To obtain the 3D pose of the subject's head, we project a 3D face model onto the 2D AAM. For the 3D model, we use a 58-points 3D face statistical anthropometric model [13] (see Figure 1 (b)). Since our 2D AAM uses 68 points, and the 3D model used 58 points, we chose only the 46 points that correspond in position for both models (see Figure 1 (c)). The resulting 46 points of the 3D model are projected on the acquired tracked 46 points of the 2D AAM to estimate the head pose. The head pose is estimated using the POSIT algorithm, which first approximates the pose using the Pose from Orthography and Scaling (POS) algorithm [14]. POS assumes the image was obtained by a scaled orthographic projection. POSIT adds multiple iterations to POS, that is, the rotation and scale matrices are re-estimated until no improvement to the pose projection is detected.

D. Feature Extraction

Once the head pose is estimated, three angles are extracted from the rotation matrix: yaw, roll, and pitch (see Figure 2). Defining the rotation matrix R as:

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (1)$$

rotation angles could be extracted as follows:

$$Yaw = \tan^{-1}(r_{21}/r_{11}) \quad (2)$$
$$Roll = \tan^{-1}(r_{31}/\sqrt{r_{32}^2 + r_{33}^2}) \quad (3)$$
$$Pitch = \tan^{-1}(r_{32}/r_{33}) \quad (4)$$

These three angles are extracted for each frame (25 fps), then their velocity and acceleration are extracted to give in total 9 features per frame. Outlier frames, which are caused by incorrect fitting of the AAM or failing to converge the 3D model into, are detected using Grubbs' test for outliers [19], those detected outlier frames are skipped.

Also a total of 100 statistical features ("functionals") where extracted, which are:

- Maximum, minimum, range, mean, variance, and standard deviation for all 9 low-level features mentioned earlier. (6×9)
- Maximum, minimum, range and average duration of: looking left, right, up and down, tilting clockwise and anticlockwise. (4×6)
- Looking direction duration rate, and looking direction to looking direction to the opposite side duration rate for all directions mentioned above. (2×6)
- Change looking direction rate for for all directions mentioned above. (1×6)
- Total number of changes of looking direction for yaw, roll, pitch, and all directions. (1×4)

The above duration features are detected when the feature in question is higher than a threshold. The threshold is the average of the feature in question plus the standard deviation of that feature for each subject.

E. Classification and Evaluation

Classification methods and techniques could be divided in two categories: generative models and discriminative models. Generative models, such as GMM, learn to cover the subspace that belongs to one class. Discriminative models, such as SVM, learn boundaries between two classes. Different classifiers have been used in emotion recognition. In this paper, we use a hybrid classifier that combines GMM with SVM for the frame-by-frame features. For the GMM, we use 7 mixtures empirically chosen and fixed to ensure consistency in the comparison, with a diagonal covariance matrix. For the hybrid classifier, we create a GMM model for each subject, then feed thse models including mean, variance and weights to the SVM classifier. The use of GMMs served as dimensionality reduction, as well as to get the same number of features to be fed to the SVM regardless of the subject's interview duration [20]. For the functional features, we use an SVM for classification.

The head rotation and movement patterns are classified in a binary subject-independent scenario (i.e. depressed/non-depressed). A manual gender separation was used in our study to test the overall accuracy for each gender group. To mitigate the effect of the limited amount of data, a leave-one-subject-out cross-validation was used, without any overlap between training and testing data for all classification tasks. To evaluate the performance of the system, several statistical measures could be evaluated, such as recall or precision. In this paper, the average recall (AR) is computed.

To analyse the significant differences of the functionals, a two-tailed T-test for two samples is computed assuming unequal variances and $p = 0.05$. The state of the T-test is calculated to identify the direction of effect. Moreover, as a feature selection method, the statistical features that passed the T-test were selected for the SVM classification in order to reduce dimensionality and maximise the recognition rate.

(a) 2D AAM face model using 68 landmarks   (b) 3D face model using 58 landmarks   (c) The 3D Model projected onto the 2D Model Using Only Matched 46 Points
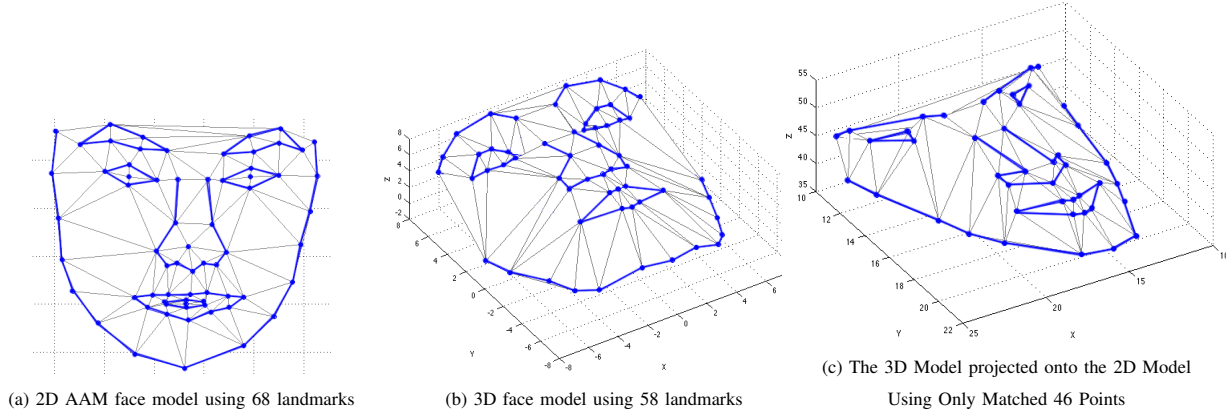
Figure 1.   3D to 2D AAM projection for estimating head pose

Table II
AVERAGE RECALL IN % FOR THE SVM CLASSIFICATION RESULTS OF DIFFERENT FEATURES

Question	All Questions			"Good News" Question			"Bad News" Question		
Features per subject	Gender Independent	Males	Females	Gender Independent	Males	Females	Gender Independent	Males	Females
Low-level features	68.8	70.1	64.0	60.3	67.0	83.5	59.3	83.5	76.8
Statistical features (passed T-test)	76.8	75.6	82.6	73.3	77.8	63.4	65.2	67.9	67.0
Average	72.8	72.9	73.3	66.8	72.4	73.5	62.3	75.7	71.9

IV. RESULTS

A. Classification Results

Automatic classifications of depression from both frame-by-frame features and functionals of head pose and movement is performed and the results are shown in Table II. Although the head pose and movement would be used as complementary cue in detecting depression in practice, their recognition rates are impressive on their own, giving 71.2% on average, which illustrates that they hold effective cues in diagnosing depression.

Comparing low-level and functional features, low-level features give 68.8% AR using the hybrid classifier, while functionals give 76.8% AR, which are both above chance level. In general, the functional features recognition rate is significantly higher than the low-level features results. This result implies that the selection of features that passed the T-test is a useful method for reducing dimensionality and feature selection. There are a few exceptions where low-level modelling performed better than functionals, such as modelling the "good news" question for females and modelling the "bad news" question for both gender-dependent. This signifies that the feature selection should be tuned for a specific T-test for both genders separately [21].

We also investigate the influence of gender-dependent classification. As shown in Table II, the comparison between the recognition results in males and females on average are not significant. A study about gender differences in depression behaviour pointed out that depressed women are more likely to be detected than depressed men, which might be caused by women amplifying their moods, while men are more likely to engage in distracting behaviours that dampen their mood when depressed [22]. This finding was supported by a previous study on speech cues to detect depression, where a significantly higher depression detection rate has been found in depressed women than in depressed men [23]. Although our head pose and movement classification results do not support the conclusions of the previous studies of gender differences [22], [23], it might indicate a physical abnormality rather than a behavioural one in head movement.

While acknowledging the potential impact of the huge reduction of training data from using all 8 questions to using only one question, we investigate the differences in expressing positive and negative emotions between depressed and control subjects. This is done by evaluating the "Good News" and "Bad News" questions from the interview. For the "Good News" question (positive emotion), recognising depression is almost as accurate as when using all 8 questions. Getting good recognition rates from such a small dataset indicates the clearly noticeable differences in expressing positive emotions in depressives and controls. On the other hand, analysing the "Bad News" question (negative emotion), gives worse recognition rates than using all 8 questions or the positive question. This indicates that both groups express negative emotions in the same or a similar manner. Linking this finding with the previous one, we conclude that positive emotions are expressed less in depressed subjects at all times, that negative emotions dominate in

Question	All Questions		"Good News" Question		"Bad News" Question	
Feature	Direction	P val.	Direction	P val.	Direction	P val.
Max. Pitch velocity	C>D	0.01	NS	NS	NS	NS
Max. Pitch acceleration	C>D	0.02	NS	NS	NS	NS
Rng. Pitch acceleration	C>D	0.03	NS	NS	NS	NS
Max. Dur. of Looking right	D>C	0.02	NS	NS	NS	NS
Rng. Dur. of Looking right	D>C	0.05	NS	NS	NS	NS
Avg. count of looking left	C>D	0.01	NS	NS	NS	NS
Avg. count of tilting ACW	C>D	0.06	NS	NS	NS	NS
Rng. Dur. of Looking Down	NS	NS	D>C	0.02	NS	NS
Avg. Dur. of Looking Down	D>C	0.05	D>C	0.01	NS	NS
Avg. Dur. of steady Yaw mov.	D>C	0.03	NS	NS	D>C	0.04
Avg. Dur. of continues Yaw mov.	C>D	0.02	NS	NS	NS	NS
Avg. Dur. of steady Roll mov.	NS	NS	NS	NS	D>C	0.03
Avg. Dur. of continues Roll mov.	C>D	0.06	NS	NS	NS	NS

NS: Not Significant, D: Depressed, C: Control, Dur.: duration, Rng.: Range, ACW: anticlockwise

depressives [24], [25] and, hence, negative emotions has less discriminatory power than positive emotions in detecting depression.

Given the huge reduction in the sample size, both gender-dependent and emotion-dependent investigations give relatively good recognition rate. This results could be explained by the thin-slicing theory, that is, when using a different smaller part of the interview, the performance is similar if not better than using the whole interview. This is based on the psychological thin slicing theory saying that using a brief observation or thin slice of behaviour can be used to predict psychological outcome at levels above that expected by chance [26].

B. Statistical Analysis

Table III shows only the significant T-test (of any of All, "Good News", and "Bad News" questions) results of analysing the functionals extracted from different parts of the interviews. Velocity and acceleration of pitch and yaw were faster in controls, which display that depressed patients move their head more slowly. Slower movements for depressives is expected as an indication of fatigue, which is a common symptom of depression [27]. Given that the interviewer approximately stands in a centred position in front of the interviewee, maximum and range duration of looking to the right was longer in depressives, which might be an indicator of avoiding eye contact with the interviewer [8]. The average number of times healthy controls move their head left to right and roll their head clockwise to anticlockwise, is higher than for depressed patients. This finding indicates that depressives' head movements are significantly less than healthy controls, which is inline with [10], [9]. We also find that the average duration of looking down is longer in depressives, which could be an indicator of avoiding eye contact with the interviewer [8], [9]. Also, steady head movements (or pause) left to right, on average is longer in duration with depressed patients. At the same time, continuous head movements left

to right and roll clockwise to anticlockwise, average duration is longer in healthy controls. These findings are another indication of having less and slow overall head movement in depressives. We conclude that head movements in general are significantly different between depressed patients and healthy subjects due to psychomotor retardation.

We also investigate the differences while expressing positive and negative emotions. When expressing positive emotions, only the average of looking down and the range duration are significantly longer in depressives. This result shows that even during positive emotions, depressives continue to look down to avoid eye contact. Steady head movement was not significant while expressing positive emotions, which might imply that depressives move their head more and faster during positive emotions. Only steady head movements left to right and roll clockwise to anticlockwise were still significant while expressing negative emotions. While being significant during overall interview and positive emotions, the looking down duration was not significant when expressing negative emotions. These results might suggest that both depressives and healthy controls express negative emotions in a similar pattern of head movements.

V. CONCLUSIONS AND FUTURE WORK

Depression is a serious psychological disorder that affects mood, thoughts, and the ability to function in everyday life. Moreover, the assessment methods of diagnosing depression rely almost exclusively on patient-reported or clinical judgments of symptom severity, risking a range of subjective biases. We have presented work aiming at an objective diagnostic aid supporting clinicians in their diagnosis of depression. In this paper, we examine the head pose and movement patterns in depressives compared to healthy subjects using interviews from real-world clinically validated data. We extracted head pose and movement features from face videos using a 3D face model projected onto user-specific face AAM. In a binary classification task (depressed vs. non-depressed), we modelled low-level and statistical functional features for classification. Although the head pose and movement would be considered as complementary cues in detecting depression in practice, their recognition rates were impressive on their own, giving 71.2% on average, which illustrates that these features hold effective cues in diagnosing depression. Investigating gender-dependent classification influence did not yield significantly differences, which may indicate a physical abnormality such as psychomotor retardation rather than behavioural differences. On the other hand, differences in expressing positive and negative emotions between depressed and control subjects were significant. Recognising depression using positive emotions was higher than using negative emotions, concluding that positive emotions are expressed less in depressed subjects at all times, and that negative emotions have less discriminatory power than positive emotions in detecting depression.

The thin-slicing theory was observed in this work from a reduced sample size in both gender-dependent and emotion-dependent investigations, which gave relatively good recognition rate. Statistical analyses on head pose and movement behavioural patterns found that depressives had slower head movements, which may indicate fatigue. We also found the duration of looking to the right was longer in depressives, which might be an indicator of avoiding eye contact with the interviewer. We also found that overall change of head position in depressives was significantly less than healthy controls. We also found that the average duration of looking down was longer in depressives, which could be an indicator of avoiding eye contact with the interviewer. Based on these findings, we conclude that head movements in general are significantly different between depressed and healthy subjects, and could be used for detecting depression.

Fusing speech features, facial expression, body posture, eye movement with this current research will be a next step towards multi-modal system. A known limitation is the fairly small number of (depressed and control) subjects. As data collection is ongoing, we anticipate to report on a larger dataset in the future.

VI. ACKNOWLEDGEMENT

REFERENCES

[1] U. D. of Health and H. Services, "Healthy people 2010: Understanding and improving health," *Health San Francisco*, vol. 2nd, no. 46, p. 62 p., 2000.

[2] S. B. Guze and E. Robins, "Suicide and primary affective disorders." *British Journal of Psychiatry*, pp. 437–438, 1970.

[3] L. G. Kiloh, G. Andrews, and M. Neilson, "The long-term outcome of depressive illness." *The British Journal of Psychiatry*, vol. 153, no. DEC, pp. 752–757, 1988.

[4] N. Niedermaier, E. Bohrer, K. Schulte, P. Schlattmann, and I. Heuser, "Misdiagnosed patients with bipolar disorder: co-morbidities, treatment patterns, and direct treatment costs." *The Journal of clinical psychiatry*, vol. 66, no. 11, pp. 1619–1623, 2004.

[5] J. C. Mundt, P. J. Snyder, M. S. Cannizzaro, K. Chappie, and D. S. Geralts, "Voice acoustic measures of depression severity and treatment response collected via interactive voice response (ivr) technology," *Journal of neurolinguistics*, vol. 20, no. 1, pp. 50–64, 2007.

[6] D. Heylen, "Head gestures, gaze and the principles of conversational structure," *International Journal of Humanoid Robotics*, vol. 3, no. 03, pp. 241–267, 2006.

[7] J. Pedersen *et al.*, "An ethological description of depression," *Acta psychiatrica scandinavica*, vol. 78, no. 3, pp. 320–330, 1988.

[8] L. Fossi, C. Faravelli, and M. Paoli, "The ethological approach to the assessment of depressive disorders," *The Journal of nervous and mental disease*, vol. 172, no. 6, pp. 332–341, 1984.

[9] P. Waxer *et al.*, "Nonverbal cues for depression." *Journal of Abnormal Psychology*, vol. 83, no. 3, p. 319, 1974.

[10] W. W. Hale III, J. H. Jansen, A. L. Bouhuys, J. A. Jenner, and R. H. van den Hoofdakker, "Non-verbal behavioral interactions of depressed patients with partners and strangers: The role of behavioral social support and involvement in depression persistence," *Journal of affective disorders*, vol. 44, no. 2-3, pp. 111–122, 1997.

[11] E. Murphy-Chutorian and M. M. Trivedi, "Head pose estimation in computer vision: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 4, pp. 607–626, 2009.

[12] K. Seo, "Face pose estimation system by combining hybrid ica-svm learning and re-registration," in *Asian Conference on Computer Vision ACCV 2004*, 2004, pp. 27–30.

[13] P. Martins and J. Batista, "Monocular head pose estimation," *Image Analysis and Recognition*, pp. 357–368, 2008.

[14] D. F. Dementhon and L. S. Davis, "Model-based object pose in 25 lines of code," *International Journal of Computer Vision*, vol. 15, no. 1, pp. 123–141, 1995.

[15] A. Ozdas, R. Shiavi, S. Silverman, M. Silverman, and D. Wilkes, "Analysis of fundamental frequency for near term suicidal risk assessment," *IEEE Conf. Systems, Man, Cybernetics*, pp. 1853–1858, 2000.

[16] E. Moore, M. Clements, J. Peifer, and L. Weisser, "Critical analysis of the impact of glottal features in the classification of clinical depression in speech." *IEEE Trans. on Bio-medical Eng.*, vol. 55, no. 1, pp. 96–107, Jan. 2008.

[17] J. Saragih and R. Goecke, "Iterative Error Bound Minimisation for AAM Alignment," in *ICPR2006*, vol. 2, Aug. 2006, pp. 1192–1195.

[18] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR 2001*, vol. 1. IEEE, 2001, pp. I–511 – I–518.

[19] F. Grubbs, "Procedures for detecting outlying observations in samples," *Technometrics*, vol. 11, no. 1, pp. 1–21, 1969.

[20] S. Alghowinem *et al.*, "A Comparative Study of Different Classifiers for Detecting Depression from Spontaneous Speech," in *ICASSP*. IEEE, May 2013.

[21] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *The Journal of Machine Learning Research*, vol. 3, pp. 1157–1182, 2003.

[22] S. Nolen-Hoeksema, "Sex differences in unipolar depression: Evidence and theory." *Psychol*, no. 101, pp. 259–282, 1987.

[23] S. Alghowinem *et al.*, "From Joyous to Clinically Depressed: Mood Detection Using Spontaneous Speech," in *Proc. FLAIRS-25*, 2012.

[24] P. Ekman, "Moods Emotions And Traits," in *P. Ekman & R. Davidson (Eds.) The Nature of Emotion: Fundamental Questions*. New York: Oxford University Press, 1994, pp. 15–19.

[25] P. Ekman and A. J. Fridlund, "Assesment Of Facial Behvior In Affective Disorders," in *Depression and Expressive Behavior. x*. Hillsdale, N.J: Lawrence Erlbaum, 1987, pp. 37–56.

[26] N. Ambady and R. Rosenthal, "Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis." *Psychological Bulletin*, vol. 111, no. 2, p. 256, 1992.

[27] M. Prendergast, *Understanding Depression*. VIC Australia: Penguin Group, 2006.