ISSN 2229-5518

A Survey on the Available Algorithms for Recognizing Depression Disorders through Gesture Recognition

Prof.Shamla Mantri, Sonal Gadiya, Sunny Bangale, Ramolee Chaudhari, Sonakshi Nathak

Abstract-- Depression is one of the most common mental health disorders with strong adverse effects on personal and social functioning. The absence of any objective diagnostic aid for depression leads to a range of biases in the diagnosis and ongoing monitoring. Psychologists use various visual cues to quantitate depression such as facial expressions, eye contact and head movements. This study throws light upon the contribution of gestures and facial points for mood analysis. A framework based on space-time interest points is proposed for the analysis of facial movements. The experiments are performed on real-world clinical data where video clips of patients and their answers to questionnaires are recorded during interactive sessions. The diagnosis is done and appropriate action is taken according to severity and scale of the depression in the patient.

Index Terms-Classifier Training, Face Registration, Feature Extraction, Feature Reduction, Gesture Recognition, Image Processing, Mood Detection.

1.Introduction

A state of low mood is a hindrance to normal behavior and functioning of a person.Person may lose interest in activities that once were pleasurable and may contemplate, attempt, or commit suicide. WHO Global Burden of Disease (GBD) report by Mathers proposed unhappiness as the leading cause of disability world-wide (20% of the population). People from all ages suffer from depression, which is also a major cause for suicide and is increasing at an alarming rate. One major reason for such a high suicidal rate is the absence of timely diagnosis of depression in early stages. Effective and deterministic diagnostic technology can provide these objective means and assist physicians in both initial depression diagnosis and ongoing monitoring. With the help of advancement in computer vision technology, it is possible to build systems, which can estimate neurological problems such as depression, and stress. In this paper for detecting depression in person ,we do the analysis on facial expression. Emotions are not only responsible for discernment functions such as rational decision making, perception and learning, but are also important for interpersonal communication. Regardless of its severity and high pre, there currently exist no laboratory-based measures of illness expression, course and recovery. This compromises optimal patient care, constituted of the burden of disability. As healthcare charge increases globally, the provision of effective health surveillance systems and diagnostic aids is highly important. With the advancement in affective sensing and machine learning, computer support diagnosis can and will play a important role in providing an unbiased assessment.

In this paper we survey the algorithms and techniques that are going to be implemented for our system. For example, for classifier here are two algorithm FACS and AAM. Based on their features and efficiency we will decide which algorithm is better so that we can get the maximized and accurate result for our system.

The process in which our System is going to work is as follows:

- Face Detection.
- Feature Extraction.
- Feature Reduction.
- Feature Classifier.
- Analysis and Final Output.

2. Previous Work

Early work demonstrated that untrained observers were able to identify severity of depressed individuals and depression from non-verbal behavior [10], [11]. These findings triggered a long history of research and development for investigating the facial expressions of patients having depression. Overall, this work indicates that depression is characterized by the weakening of smiles produced by the jugal major muscle (e.g., [12], [13], [14], [15]). Studies have also found major differences in terms of other facial expressions of the subjects. But, these results were largely individual differences in personality traits [16]. Any between-group differences could have been due to either stable individual differences or current depression of an individual. The studies that displayed this issue by following depressed subjects over time have yielded very interesting results. For instance, the patterns of facial expression in the duration of an intake interview showed signs of a later clinical improvement, and patterns of facial expression changed over the course of treatment [17], [13], [14], [19].

Additionally, the majority of previous work explored an extremely limited and a meager sample of facial expressions. However, perfect interpretation of the functionality of a facial expression requires comprehensive and a detailed description of its structure. For instance, the most prominent finding in depression so far has been the weakening of jugal major activity that led many researchers to conclude that depression is marked by weakening of expressed positive expression. However, jugal major contraction has been linked to all of the following, in addition to positive expression, when paired with the activation of other facial muscles: politeness, pain and embarrassment.

Many previous studies examined the facial expressions of depressed participants in contexts of low sociality like viewing static images or film clips alone. Moreover, the research has found that the frequency of certain facial expressions can be very low in such contexts and conditions .

For the further exploration of this issue, a number of studies on depression [19], [20], [21], [22] compared smiles with and without the orbicularis oculi muscle contraction. The the orbicularis oculi muscles are well known as the Duchenne. marker. Because the results of these studies were ambiguous, more research is needed on the occurrence of different types of smiles when an individual is facing depression. Similarly, no studies have totally examined the occurrence of negative smiles in depression. These smiles include the contraction of facial muscles that are related to negative emotions. The negative smiles were more common in currently depressed participants than in those without in any current depression.

Finally, a majority of the previous work used manual facial expression analysis or visual coding. The training required for this and the coding itself is considerably time consuming even though humans can become highly reliable at coding even complex facial expressions. In an effort to remove this time burden, a great deal of research has been devoted to the development of automated systems for facial expression analysis. Even though initial applications to clinical science are beginning to be developed, very few automatic systems have been trained or tested on populations facing depression.

3. Methods

3.1 Working

3.1.1 Face Registration

In the first stage of the automated FACS system (Fig 1), we detect facial regions and track facial landmarks defined on the contours of eyebrows, eyes, nose, lips, etc. in videos. For each frame of the video, an approximate region of a face is detected by a face detector. Within the face region, we search the exact location of facial landmarks using a deformable face model such as Active Shape Model (ASM) for several reasons. ASM is the simplest and fastest method among different models, which fit our need to track a large number of frames in our videos. In order to register two facial images such as the reference image and the sensed image, a set of points called landmark points is utilized. Facial landmark points indicate the location of the important facial components like eye corners, nose tip, jaw line. In order to automatically interpret a face image, an efficient method for tracking the facial landmark points is required.

ISSN 2229-5518

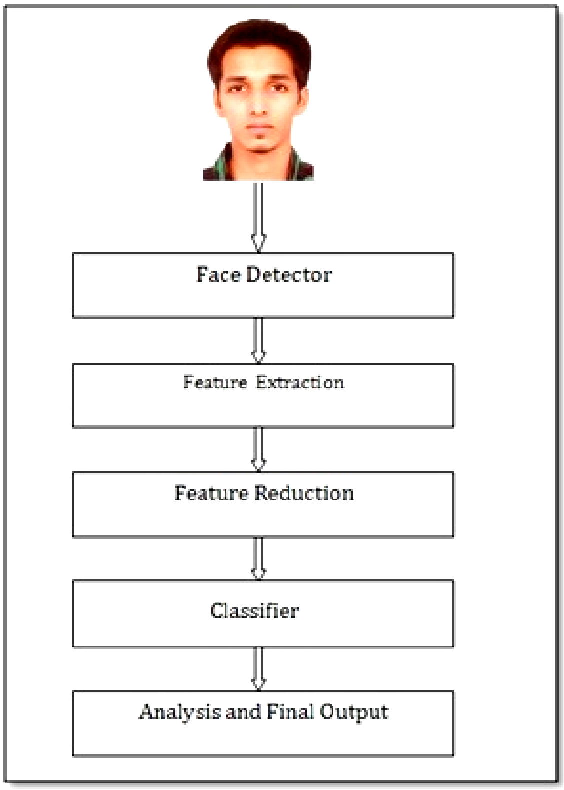

Fig 1: Steps for automatic coding

3.1.2 Feature Extraction

Face recognition's core problem is to extract information from images. This feature extraction process can be defined as the procedure of extracting relevant information from a face image. This information must be valuable to the later step of identifying the subject with an acceptable error rate. The feature extraction process must be efficient in terms of computing time and memory usage.

Feature extraction involves several steps – Feature Extraction, Feature Selection (Fig.2) [1]. These steps may overlap, and dimensionality reduction could be seen as a consequence of the feature extraction and selection algorithms.(Table 1) [1] Ultimately, the number of features must be carefully chosen. Too less or redundant features can lead to a loss of accuracy of the recognition system. We can make a distinction between feature extraction and feature selection. Both terms are usually used interchangeably. Nevertheless, it is recommendable to make a distinction.

A feature extraction algorithm extracts features from the data. It creates those new features based on transformations or combinations of the original data. In other words, it transforms or combines the data in order to select a proper subspace in the original feature space.

On the other hand, a feature selection algorithm selects the best subset of the input feature set. It discards non-relevant features. Feature selection is often performed after feature extraction. So, features are extracted from the face images, then a optimum subset of these features is selected.

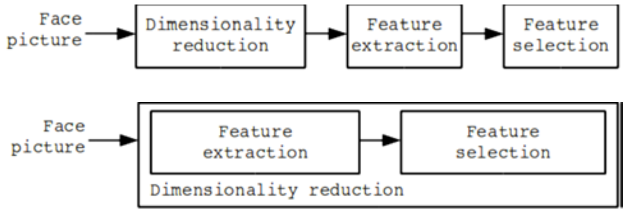

Fig 2: Feature extraction

TABLE 1
Feature extraction methods

Active Shape Models (ASM)	Statistical method searches boundaries
Active Appearance Models (AAM)	Evolution of ASM uses shape and texture
Facial Action Coding System (FACS)	Represents expression with a set of Action units (AU)

The Facial Action Coding System (FACS) [1] that defines the human face by a number of Action Units (AUs) (Table 2) [1] and represents the facial expressions by different combinations of these AUs was proposed. Since the classification into AUs is based on facial anatomy, practically all expressions can be represented by this coding scheme. Hence, FACS is by far the most widely used method for facial expression recognition. However, one of the inherent difficulties with the FACS coding scheme is that it requires a highly trained human expert to manually score each frame of a video.

TABLE 2
Action Units

AU	Name	Incidence
1	Inner Brow Raiser	0.292
2	Outer Brow Raiser	0.196

ISSN 2229-5518

4	Brow Lowerer	0.322
5	Upper Lip Raiser	0.172
6	Cheek Raiser	0.206
7	Lip Tightener	0.201
9	Nose Wrinkler	0.125
11	Nasolabial Deepener	0.056
12	Lip Corner Puller	0.187
15	Lip Corner Depressor	0.150
17	Lower Lip Depressor	0.041
20	Lip Stretcher	0.130
23	Lip Tightener	0.100
24	Lip Pressor	0.096
25	Lips Part	0.484
26	Jaw Drop	0.164
27	Mouth Stretch	0.137

Active appearance models (AAM) [1] simultaneously describe the shape and texture variation of faces Herein, shape refers to the relative positions of feature points (such as eye corners, mouth corners, nose tip, etc.) in the face, whereas texture refers to the shape-normalized visual appearance of the face (for instance, eye color, skin color, malls, etc.). Active appearance models thus consist of two sub models:
(1) a shape model that models the location of facial feature points and
(2) a texture model that models the shape normalized facial texture.

Shape model:- To train the shape model of an active appearance model, a data set of face images is required in which facial feature points—for instance, mouth corners, eye corners, and nose tip—are manually annotated. The feature points are required to be relatively dense, in such a way that a triangulation constructed on the feature points approximately captures the geometry of the face, i.e., in such a way that the imaginary triangles between the feature points correspond to roughly planar surfaces of the face. The manual annotation of a collection of face images is time-consuming, but it only needs to be done once for a fixed collection of faces. If later on, we encounter a new face image, we can automatically determine the facial feature point locations by fitting the active appearance model on the new face image using the procedure described in Fitting.
Texture model:- To model the facial texture we use the feature point annotations to construct a data set of face images in which all feature points have exactly the same location. This is achieved by warping each face image onto the base shape m using the feature point annotations as control points. We can use the resulting shape-normalized face images to construct a texture model that describes features such as eye color, skin color, lip color, malls, etc.

3.1.3 Feature Reduction

In many pattern classification applications, the number of features is extremely large. This size makes the analysis and classification of the data a very complex task. In many pattern classification applications, the number of features is extremely large and this size makes the analysis and classification of the data a very complex task. In order to extract the most important and efficient features of the data, several linear techniques, such as Principle Component Analysis (PCA) and Independent Component Analysis (ICA), and nonlinear techniques, such as Kernel PCA and Manifold Learning, have been proposed.

The origin of PCA is from Eigen Vectors. PCA belongs to linear transforms based on the statistical techniques. This method provides a powerful tool for data analysis and pattern recognition which is often used in signal and image processing as a technique for data dimension reduction or their de-correlation as well. Its advantage is that it can be used as a compression method of data without any loss of information. Principle Component Analysis (PCA) is a mathematical procedure that uses Linear Transformations to map data from high dimensional space to low dimensional space.
With minimal effort PCA provides a roadmap for how to reduce a complex data set to a lower dimension to reveal the sometimes hidden, simplified structures that often underlie it. Importance of PCA is manifested by its use in so many different fields of science and life.

PCA is a way of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences. PCA can also be used to compress the data, by reducing the number

ISSN 2229-5518

of dimensions, without much loss of information. Therefore it is important to keep in mind that PCA best serves to represent data in simpler, reduced form.

The PCA is a statistical data analysis method that transforms the initial set of variables into an assorted set of linear combinations, known as the principal components (PC), with specific properties with respect to variances. This condenses the dimensionality of the system while maintaining information on the variable connections. PCA uses a signal based representation criterion where the purpose of feature extraction is to represent the samples accurately in a lower dimensional space whereas the alternate technique, Linear Discriminant Analysis (LDA) deploys a classification based approach.

Linear Discriminant Analysis easily handles the case where the within-class frequencies are unequal and their performances have been examined on randomly generated test data. This method maximizes the ratio of between-class variance to the within-class variance in any particular data set thereby guaranteeing maximal separability.

Although it might sound intuitive that LDA is superior to PCA for a multi-class classification task where the class labels are known, this might not always be the case. For example, comparisons between classification accuracies for image recognition after using PCA or LDA show that PCA tends to outperform LDA if the number of samples per class is relatively. In practice, it is also not uncommon to use both LDA and PCA in combination: E.g., PCA for dimensionality reduction followed by an LDA.

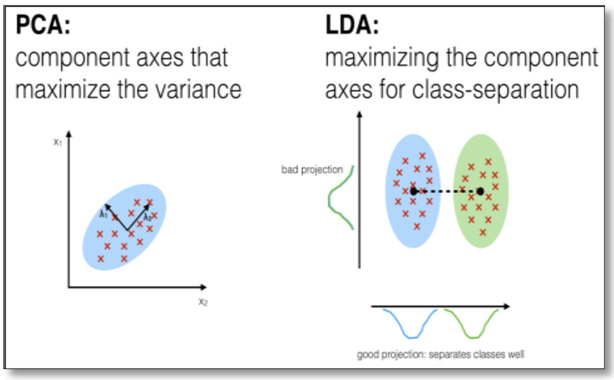

Fig 3: PCA and LDA component axes

Linear Discriminant Analysis (LDA) is a procedure used to overcome dimensionality reduction. It is used mainly in the Small Sample Size (SSS) problem. This issue can in a large set of data. A drawback of the original LDA is that at least one non-singular scatter matrix is required for computation. When there are insufficient numbers of samples in the data set, this condition fails, i.e. the data dimension generally exceeds the number of available data values.

Advantages of Discriminant Analysis:
1. Multiple dependent variables
2. Reduced error rates
3. Easier interpretation of Between-group

Differences: each discriminant function measures something unique and different.(Table 3) [2]

TABLE 3
Comparison of PCA and LDA

Principal Components Analysis	**Linear Discriminant Analysis**
The prime difference between LDA and PCA is that PCA does more of feature classification.	The prime difference between LDA and PCA is that LDA does data classification.
The shape and location of the original data sets changes when transformed to a different space.	LDA doesn't change the location but only tries to provide more class separability and draw a decision region between the given classes.
PCA calculates best discriminating components without knowledge about groups	LDA calculates the best discriminating components about groups which are defined by the client (user).

3.1.4 Classifier Training

In machine learning, feature classification aims to assign each input value to one of a given set of classes. A well-known classification technique is the Support Vector Machine (SVM). A well-known classification technique is the Support Vector Machine (SVM). The SVM classifier applies the kernel trick which uses dot product, to keep computational loads reasonable.

ISSN 2229-5518

There are three concepts that are key in building a classifier - similarity, probability and decision boundaries.

Similarity: This approach is intuitive and simple. Patterns that are similar should belong to the same class. This approach has been used in the face recognition algorithms implemented later. The idea is to establish a metric that defines similarity and a representation of the same-class samples. For example, the metric can be the euclidean distance. The representation of a class can be the mean vector of all the patterns belonging to this class.

Probability- Some classifiers are build based on a probabilistic approach. Bayes decision rule is often used. The rule can be modified to take into account different factors that could lead to miss-classification. Bayesian decision rules can givean optimal classifier, and the Bayes error can be the best criterion to evaluate features. Therefore, a posteriori probability functions can be optimal.

Decision boundaries- This approach can become equivalent to a Bayesian classifier. It depends on the chosen metric. The main idea behind this approach is to minimize a criterion (a measurement of error) between the candidate pattern and the testing patterns. One example is the Fisher's Linear Discriminant. Often FLD and LDA are used interchangeably. It's closely related to PCA.

SVM is a method widely used is the support vector classifier. It is a two class classifier, although it has been expanded to be multiclass. The optimization criterion is the width of the margin between the classes, which is the distance between the hyper plane and the support vectors. These support vectors define the classification function. Support Vector Machines (SVM) are originally two-class classifiers. There are two main strategies

1. On-vs-all approach. A SVM per class is trained. Each one separates a single class from the others.

2. Pairwise approach. Each SVM separates two classes. A bottom-up decision tree can be used, each tree node representing a SVM . The coming face's class will appear on top of the tree.

There can be different training sets, collected in different conditions and representing different features. Each training set could be well suited for a certain classifier. Those classifiers could be combined. .One single training set can show different results when using different classifiers. A combination of classifiers can be used to achieve the best results. .Some classifiers differ on their performance depending on certain initializations. Instead of choosing one classifier, we can combine some of them.

An alternate to svm is adaboost using gabor wavelengths. The appearance features that have been successfully employed for emotion recognition are local binary pattern (LBP) operato, histogram of orientation gradients (HOG), local Gabor binary patterns (LGBP), local directional pattern (LDP), non-negative matrix factorization (NMF) based texture feature, Gabor filter based texture information [16], principle component analysis (PCA), linear discriminant analysis (LDA), *etc.*

Among the appearance-based techniques, the theory of NMF has recently led to a number of promising works. In, an analysis of the effect of partial occlusion on facial expression recognition is performed, using a method based on Gabor wavelets texture information extraction, supervised image decomposition method based on discriminant NMF (DNMF), and shape-based method. A technique called graph-preserving sparse NMF (GSNMF) was introduced by Zhi *et al*. The GSNMF is an occlusion-robust dimensionality reduction technique, which transforms high-dimensionality facial expression images into a locality-preserving subspace, with sparse representation the best result was obtained by selecting a subset of Gabor filters using AdaBoost (Table 4) [2] and then training SVM (Table 5) [2] on the output of the filters selected by the AdaBoost.

TABLE 4 Confusion matrix for facial expression recognition in percentages, using multi-class AdaBoost with 125 feature vectors.

	Anger	Disgust	Fear	Happiness	Sadness	Surprise
Anger	95	5	0	0	0	
Disgust	0	95	1.67	3.33	0	0
Fear	0	0	92	8	0	0
Happiness	0	0	3.08	96.92	0	0
Sadness	6.67	0	0	0	93.33	0
Surprise	0	0	0	1.25	0	98.75

TABLE 5: Confusion matrix for facial expression recognition in percentages, using SVM with boosted features (100 AdaBoost selected features).

	Anger	Disgust	Fear	Happiness	Sadness	Surprise
Anger	92.5	2.5	0	0	5	0
Disgust	0	96.67	1.67	0	1.67	0
Fear	0	0	92	8	0	0
Happiness	0	3.08	1.54	95.38	0	0
Sadness	6.67	0	3.33	0	87.67	3.33
Surprise	0	0	2.5	0	2.5	95

4. Conclusion

We developed a system for automatic action unit recognition. Our system uses conditional random fields to predict action unit states from features extracted using active appearance models. The performance of our system is promising, and it can be used in real time. Accuracy in detecting depression was 88% for manual FACS and 79% for AAM.

From the above study we can conclude that PCA is importantly use for feature classification and LDA for data classification. But our main aim is on reducing the features which are of our interest. Large data can be computed using PCA but LDA fails to do the same. Its drawback is it atleast required one non-singular matrix for computation. As our project aims at Feature reduction Technique PCA is best algorithm for it. PCA can handle small as well as large scale data.Hence it is suggested that PCA can be best suited for our project. PCA can be described as an "unsupervised" algorithm, since it "ignores" class labels and its goal is to find the directions (the so-called principal components) that maximize the variance in a dataset. In contrast to PCA, LDA is "supervised" and computes the directions ("linear discriminants") that will represent the axes that maximize the separation between multiple classes.

Two methods for facial expression recognition, by using either multi-class AdaBoost with DTW, or by using SVM on the boosted features, are proposed in this part. The geometric features are extracted from the sequences of facial expression images, based on tracking results of facial landmarks. The proposed facial expression recognition system is fully automatic, in which the landmark initialization and tracking is based on the EBGM method. Each extracted geometric feature vector is used to build a single weak classifier, which is based on the similarity between the input feature vector and the prototypical feature vector of each facial expression. A multi-class AdaBoost algorithm is used, to select the subset of discriminative feature vectors. A recognition accuracy of 95.17% using feature selective multi-class AdaBoost, and 97.35% using SVM on boosted features is achieved. Feature selection using AdaBoost and expression classification using SVM gives the best recognition accuracy. The discriminative feature vectors for each facial expression are also determined, and the result is analyzed, depending on the face region from which the facial landmark tracking results are contributed, to build the feature vectors

References

[1] J. F. Cohn, T. S. Kreuz, I. Matthews, Y. Yang, M. H. Nguyen, M. T. Padilla, F. Zhou, and F. De la Torre, "Detecting Depression from Facial Actions and Vocal Prosody," in Proc. Affective Computing and Intelligent Interaction (ACII2009), 2009, pp. 1–7.

[2] Deepak Ghimire and Joonwhoan Lee, "Geometric Feature-Based Facial Expression Recognition in Image Sequences Using Multi-Class AdaBoost and Support Vector Machines"

[3] Vikram Solunke, Pratik Kudle, Abhijit Bhise, Adil Naik, Prof. J. R. Prasad, " A Comparison between Feature Extraction Techniques for Face Recognition".

[4] Dr. Sanjay Kumar , Er. Ankur Chauhan, "A Survey on Image Feature Selection Techniques".

[5]Tuan Anh, Sander Mägi, "Principal Component Analysis", Final Paper in Financial Pricing NATIONALCHENGKUNGUNIVERSITY, 6/17/2009

[6] S. Balakrishnama, A. Ganapathiraju, ,9571 "LINEAR DISCRIMINANT ANALYSIS - A BRIEF TUTORIAL"

[7] Ajeet Singh, BK Singh and Manish Verma, 2012 "Comparison of HGPP, PCA, LDA, ICA and SVM"

[8] Onsen TOYGAR and Adnan ACAN, "Face Recognition using PCA, LDA and ICA Approaches on Colored Images "

[9] Suganthy, M. and P. Ramamoorthy, 2012 "Principal Component Analysis Based Feature Extraction, Morphological Edge Detection and Localization for Fast Iris Recognition "

[10] Waxer P. 1974.Nonverbal cues for depression. Journal of Abnormal Psychology.

[11] Waxer P. 1976.Nonverbal cues for depth of depression: Set versus no set. Journal of Consulting and Clinical Psychology.

[12] Chentsova-Dutton YE, Tsai JL, Gotlib IH. 2010.Further evidence for the cultural norm hypothesis:positive emotion in depressed and control european american and asian american women. Cultural Diversity and Ethnic Minority Psychology.

[13] Gaebel W, Wolwer W. 1967. Facial expressivity in the course of schizophrenia and depression. European Archives of Psychiatry and Clinical Neuroscience.

[14] Jones IH, Pansa M. 1979. Some nonverbal aspects of depression and schizophrenia occurring during the interview. Journal of Nervous and Mental Disease.

[15] Tremeau F, Malaspina D, Duval F, Correa H, Hager-Budny M, Coin-Bariou L, Macher JP, Gorman JM. 2005. Facial expressiveness in patients with schizophrenia compared to depressed patients and non-patient comparison subjects. American Journal of Psychiatry.

[16] Kotov R, Gamez W, Schmidt F, Watson D. 2010. Linking big personality traits to anxiety, depressive, and substance use disorders: A meta-analysis. Psychological Bulletin.

[17] Fossi L, Faravelli C, Paoli M. 1984.The ethological approach to the assessment of depressive disorders. Journal of Nervous and Mental Disease; Journal of Nervous and Mental Disease.

[18] Sakamoto S, Nameta K, Kawasaki T, Yamashita K, Shimizu A. 1997. Polygraphic evaluation[of laughing and smiling in schizophrenic and depressive patients. Perceptual and Motor Skills.

[19] Carney RM, Hong BA, O'Connell MF, Amado H. 1981.Facial electromyography as a predictor of treatment outcome in depression. The British Journal of Psychiatry.

[20] Gaebel W, Wolwer W. 1967. Facial expressivity in the course of schizophrenia and depression. European Archives of Psychiatry and Clinical Neuroscience.

[21] Renneberg B, Heyn K, Gebhard R, Bachmann S. 2005. Facial expression of emotions in borderline personality disorder and depression.

[22] Tsai JL, Pole N, Levenson RW, Munoz RF. 2003. The effects of depression on the emotional responses of spanish-speaking latinas. Cultural Diversity and Ethnic Minority Psychology.