The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM)
ISSN: 2602-3199

The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM)

Volume 1, Pages 316-330

ICONTES2017: International Conference on Technology, Engineering and Science

AN EMOTION ANALYSIS ALGORITHM AND IMPLEMENTATION TO NAO HUMANOID ROBOT

Onder Tutsoy
Adana Science and Technology University

Fatma Gongor
Adana Science and Technology University

Duygun Erol Barkana
Yeditepe University

Hatice Kose
Istanbul Technical University

Abstract: Humanoid robots are extensively becoming an essential part of the social life. It is crucial for humanoid robots to understand the emotions of the people for efficient human-robot interaction. Even though a great number of facial emotion analysis algorithms have been developed and a number of them have been implemented to humanoid robots, there are still gaps in improving accuracy, computational burden and speed of these algorithms. This paper proposes a 4-stage emotion analysis algorithm and then presents its application to NAO humanoid robot. Initially, the robot detects the face using Viola-Jones algorithm. Later, important facial distance measurements are taken with geometric based facial distance measurement technique. Then, facial action coding system technique is used to detect movements of the measured facial points. Finally, measured facial movements are evaluated to understand instant emotional properties of the person. Although this algorithm can be implemented to all humanoid robots, in this research, it has been specifically applied to NAO humanoid robot. The reliability of the emotion analysis is verified by analyzing each terminal decision made based on the facial distance measurements. In addition, the accuracy, computational burden and speed of the algorithm are assessed to show the effectiveness of the algorithm.

Keywords: Emotion analysis, facial action coding system, geometric based facial distance measurement technique, facial expression, Viola-Jones algorithm

Introductıon

Humanoid robots (HR) are expected to perform various advanced tasks such as making decision about content of a communication based on emotions of a human. If an HR thinks that a person is sad, angry or happy; then, it can change the communication accordingly or it can leave the person alone for him/her to relax. Therefore, designing autonomous HR with efficient emotion analysis capability leads to more sophisticated social robots design.

A great number of facial emotion analysis algorithms have been developed in the literature. Gratch and Marella's stated importance of understanding emotions in education, and focused on computational approaches to make facial emotion analysis [25]. Silvia et al. proposed a facial emotion detection algorithm, and examined it in noisy dynamic environments [21]. This algorithm detects the facial muscle movements with a facial action coding system (FACS) and then classifies these movements to reach a decision about emotion of a human [22]. Ballihi et al. developed a further emotion analysis algorithm to only detect negative and positive facial expressions [24]. This algorithm takes into account not only the facial muscle movements, but also the posture of the upper body in emotion analysis.

- Selection and peer-review under responsibility of the Organizing Committee of the conference

*Corresponding author: Onder Tutsoy- E-mail: otutsoy@adanabtu.edu.tr

A number of facial emotion analysis algorithms have been implemented to various HR in the literature. Recently, Breazeal used Arousal_Valence-Stance (AVS) algorithm to obtain facial expressions and implemented it to Kismet HR, designed by MIT [27]. However, any information about the facial emotion recognition accuracy of the algorithm is not available. Barros et al. considered a different approach by applying convolutional neural networks to classify facial emotions [23]. This algorithm was implemented to iCub HR and its learning or generalization ability of the emotional states was assessed. They examined the algorithm with only two experiments and reported that the facial emotion recognition accuracy was 74.2%.

Zhang et al. used Facial Expression Database (FED) covering approximately 2000 emotional facial images collected from over 200 subjects to train feedforward neural network [31]. This algorithm was applied to NAO HR where the facial emotion recognition accuracy was 71.3%. These results indicate that there is still a room to improve accuracy of the facial emotion recognition algorithms implemented to HR. It is important to note that together with the accuracy of these algorithms, their computational burden and also speed should be analysed.

In this paper, proposed facial emotion analysis algorithm has four components: face detection algorithm, facial distance measurement technique, facial action coding system, and classifying emotional properties of a human during human-robot interaction. Various algorithms have been proposed in the literature for face detection. Initial form of Viola and Jones algorithm detects only frontal faces [1-2], whereas its succeeding algorithms are able to detect faces from various angles and profiles. Other face detection algorithms such as rotation invariant neural network based 3D face detection algorithm, which has been experimented on wider range of face images than Viola-Jones algorithm, has previously been proposed [3]. Even though this algorithm shared a number of commonalities with Viola-Jones algorithm, the face detection speed and accuracy have been lower. Another, face detection algorithm have been developed by Rowley-Baluja-Kanade and Schneiderman-Kanade, which uses an invariant scale detector and resizes the scanning window by preserving the image resolution [3-4]. It was stated that Viola-Jones algorithm required less detection time compared to this algorithm even though the amount of scanning time has been same [3]. The key reason behind faster Viola-Jones algorithm is the implementation of a boosting algorithm called Ada-Boost algorithm, which selects the best features, and then classifies these features [5]. Thus, in this paper Viola-Jones algorithm is selected for face detection.

When face is detected; then, the second step for the emotion analysis is the specification of the facial distances by either using appearance based or geometric based approaches. In this paper, a geometric based approach is preferred to measure facial distances; particularly, relative sizes and positions of the important facial parts. Geometric based facial distance measurement approach has filters, such as Canny filter, to detect the eyes or mouth region, and also has transform methods, such as Hough transform, to detect the nose region [13]. Geometric based facial distance measurement technique is computationally expensive, but it is more robust to variation in face position, scale, size and head orientation [14].

As the distances of each facial component have been measured, the next stage of the emotion analysis mechanism is to track the facial muscle movements. In this paper, Facial Action Coding System (FACS) technique is preferred because of its effective and realistic facial muscles tracking ability. There exist more techniques such as Arousal-Valence-Stance (AVS) to detect emotions [11]. AVS technique is promising but FACS provides accurate clarity in facial expression analysis as it detects each muscle movement on the face. FACS technique considers 64 Action Units (AUs) to express the state of facial expressions [10].

In the last stage of the facial character analysis, emotional states of the human is recognized by classifying facial the measurements and facial muscle movements. Islam and Loo modified a two-stage fuzzy reasoning model for facial emotion classification [28]. In the first stage of the fuzzy reasoning, the facial distance measurements and facial muscle movements are fuzzified and then linked to action units (AUs). In the second stage of the fuzzy reasoning, these AUs are fuzzified and then transferred to the emotion space to classify basic emotions such as surprise, sadness, fear, anger, and happiness. They evaluated the algorithm on the extended Cohn-Kanade (CK) database by using 150 image sequences and reported an average 87% accuracy for facial emotion recognition. Kharat and Dudul considered support vector machines (SVM) for basic emotional expressions classification [30]. They tested the algorithm in simulation environment on available data rather than in dynamic environment and claimed 94% accuracy for emotion recognition. Tsalakanido and Malassiotis used rule based classification to recognize emotions [29]. They applied both the geometric and appearance based approaches to obtain facial measurements from 2D and 3D images and conducted experiments with 52 participants in dynamic environmentto assess accuracy where it was around 70%. Thus, it is clear that the accuracy of the facial emotion recognition algorithms drops significantly if real time experiments are conducted in noisy dynamic environments.

In the rest of the paper, Section 2 presents the proposed emotion analysis algorithm and Section 3 briefly reviews Viola-Jones face detection algorithm. A geometric based facial distance measurement technique to measure each determined facial point distances is presented in Section 4. Section 5 presents FACS technique to determine, and track the facial muscle movements. Section 6 presents emotion extraction process of the person. The implementation of the proposed emotion analysis mechanism to the NAO robot is given in Section 7. Section 8 analyses the simulation and experimental results. Finally, conclusion and possible future work are given in Section 9.

Proposed And Experımentally Implemented Emotıon Analysıs Algorıthm

The proposed emotion analysis algorithm has five key stages as shown in Figure 1. In the first stage, images are taken from NAO humanoid robot's camera, and then at the second stage, face is detected from the taken image by using Viola-Jones algorithm. At the third stage, distance measurements are taken from the detected face with Geometric based facial distance measurement technique. Then in the fourth stage, movements of the measured facial points are detected using FACS. Finally, at the fifth stage, measured facial movements are classified to understand instant emotional properties of the person.

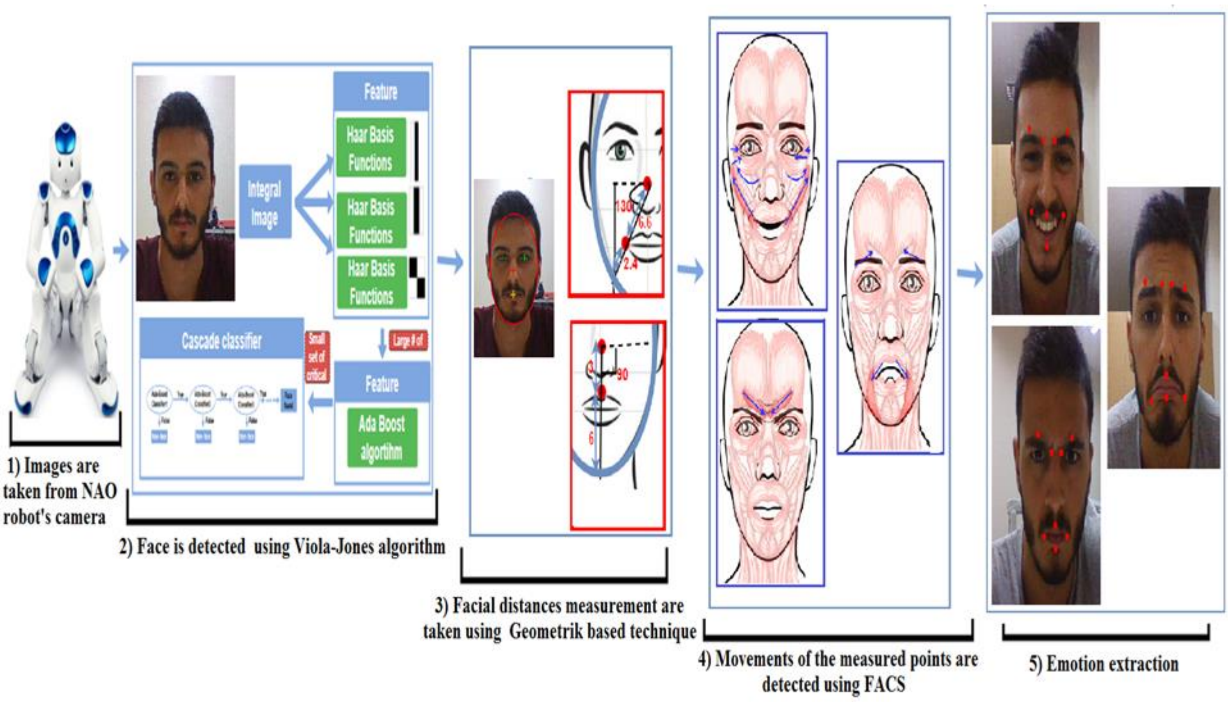

Figure 1: Proposed emotion analysis architecture

In the rest of the paper, these stages of the algorithm are reviewed, implemented to NAO HR and the results are analyzed.

Vıola-Jones Algorithm

This section presents a brief review of Viola-Jones algorithm considered for face detection. Please see the recent research paper for more detailed explanation of the algorithm [15]. Viola-Jones algorithm initially divides the taken image into rectangles to speed up the feature extraction process of face detection. The number of the rectangles varies, but generally, two, three or four rectangles are considered to extract features as shown in Figure 2.

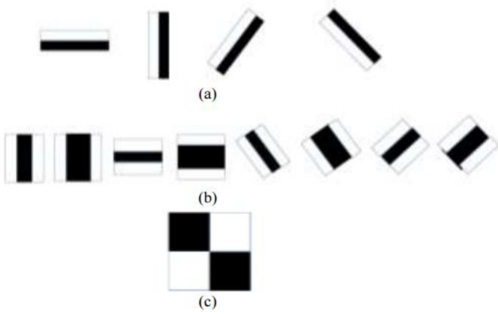

Figure 2: Rectangle based feature collection With a) Two rectangles, b) Three rectangles, And c) Four rectangles

The next step in Viola-Jones algorithm is to specify colour of the images, which is determined based on the intensity of the rectangle region. Figure 3 shows two and three rectangles for eyes region and their colours.

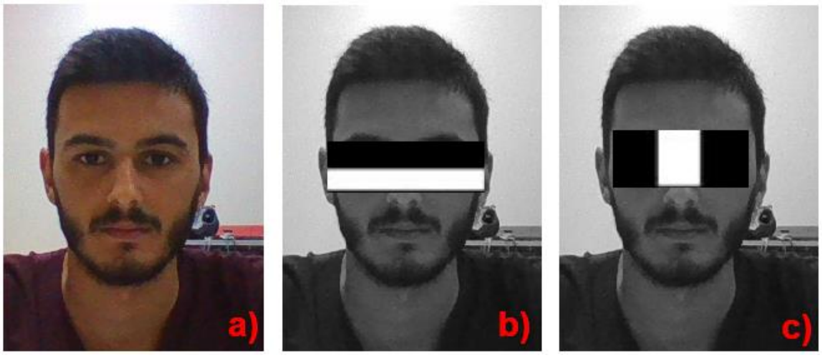

Figure 3: (a) Taken image, (b) Two rectangles for the eyes region (c) Three rectangles for the eyes region

After extracting features with rectangles, the next step is to determine sum of features in each rectangle and then the overall image. Sum of each rectangle can be represented as:

$$s\left(x_i, y_j\right) = s\left(x_{i-1}, y_j\right) + i\left(x_i, y_j\right), \text{ for each j} \quad (1)$$

where i is the indices of x and y axes and $s\left(x_i, y_j\right)$ is the current sum of features, $s\left(x_{i-1}, y_j\right)$ is the previous total sum, where $s\left(x_0, y_j\right) = 0$, and $i\left(x_i, y_j\right)$ is the each partitioned area of the rectangle where y_j is constant. Finally, the total area of the image can be found.

$$ii(x_f, y_j) = ii(x_f, y_{j-1}) + s(x_f, y_j) \quad (2)$$

where $ii(x_f, y_j)$ is the successive sum of the features in the rectangles with the final (terminal) value x_f and y_j is the j^{th} rectangle, $ii(x_f, y_{j-1})$ previous iterative sum, and $s(x_f, y_j)$ is the total area of each rectangle.

Then, the next step is to normalize the sum of the rectangles and find a temporary labelling error. This error is used to create weak and strong candidates based on a threshold. After summing up features within rectangles having different intensities, Ada-Boost algorithm uses these summed features to label them as weak or strong candidates considering the desired object. Ada-Boost algorithm requires a threshold value to distinguish weak and strong candidates. Later, labelled weak and strong candidates are compared with the testing samples specified for each facial component. As candidates pass from one stage, then they are transferred to next stage and finally the face is detected.

Geometrıc Based Facial Distance Measurement Technique

The detected face with Viola-Jones algorithm is assigned to a coordinate system to measure facial distances where the center of face and its parts such as eye, nose, and mouth are specified as shown in Figure 4.

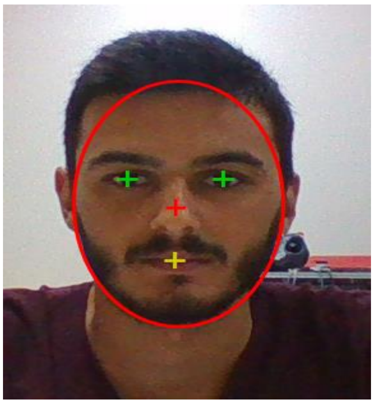

Figure 4. Elliptical representation of Face; Center of face, eye, nose and mouth detection

When the centres of each facial parts are found, then eight facial points are positioned to determine the movements of these facial parts. The distances between eyes is used to identify the centre point of the face, followed by the nose and mouth (Figure 5).

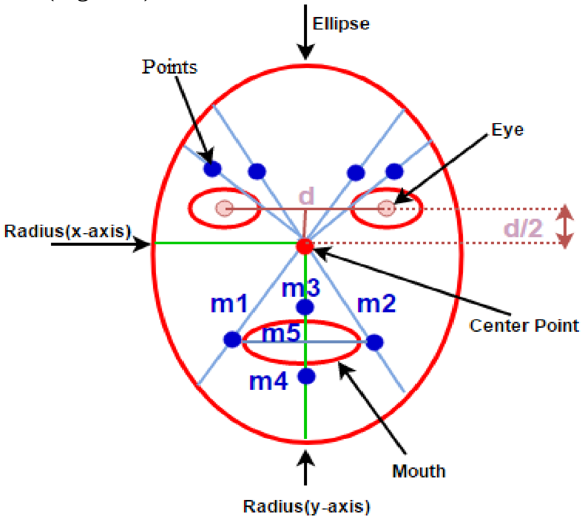

Figure 5. Placement of facial points in geometric model

Figure 5 shows the placement of eight facial points (blue) on a person's face from the centre point. The system places a centre point after the computation of the distance between two eyes. Hence, half distance from the eye to mouth is used to position the centre point. Then, an ellipse is located around the person's face with reference to the centre point. The radius of the ellipse is determined based on the facial features previously extracted by using Viola-Jones algorithm. Next, a vertical line is drawn from the centre point to the intersection point between the ellipse and x-axis of the centre point of the face.

Eight points are placed on the upper and lower face with reference to the centre of the person's face; four points each are placed on the upper face, and other four ones are placed on the lower face (Figure 6). These facial points' positions are used to calculate each facial point's distances to each other to extract the facial expression of the person.

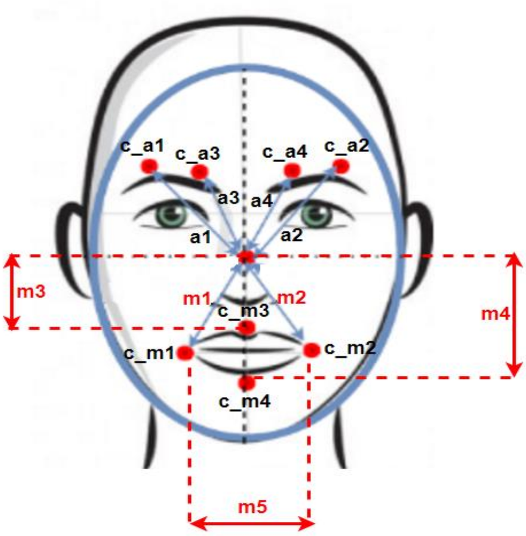

Figure 6. Total of eight facial point positions with their specific names

Placement of eight points is done at a certain angle, and distance from the centre point. The first and second points of the upper face (c_a1 and c_a2) are placed at 45 degree angle from the x-axis on the left and right sides of face, respectively. Later, the radius of the ellipse at the 45 degrees angle is calculated. It is determined that a mean point distance ratio 6.5:9 with ellipse radius at 45 degrees angle provided the best position for c_a1 and c_a2 points on the upper face (Figure 7)

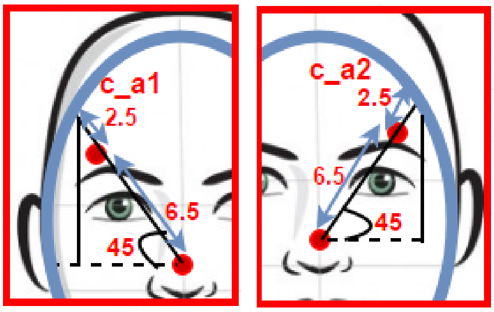

Figure 7.Point placement: The positions of two upper face points from centre point with the respective angles and distances

Same method is also used for lower face. A complete description of the ratio calculation, and placement of points on the lower face of each person is shown in Figure 9.

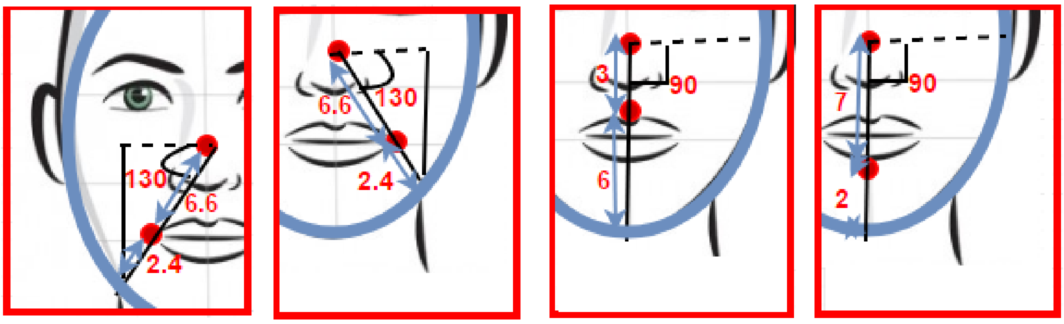

Figure 8. Points placement: The positions of lower face markers from center point with the respective angles and distance

The distances between the center point and other facial points are referred as facial distance measurements and used to extract emotional states of a person. Each distance measurements are calculated using Pythagorean Theorem and every point is assigned by their own x-y coordinates. For example in Figure 9, in the left mouth column, line m1 is the hypotenuse of a right triangle, wherein the line parallel to the x- axis is dx (the difference

between x-coordinates of the centre point x_c and c_m1 (x_{m1})); and line parallel to the y-axis is d_y (the difference between y-coordinates of the centre point y_c and c_m1 (y_{m1})).
The formula to compute m1 is given by equation (3);

$$H^2 = \left(X_c - X_{m1}\right)^2 + \left(Y_c - Y_{m1}\right)^2$$
(3)

where H is Hypotenuse value.

Therefore, the formula for m1 computation is given as in equation (4).

$$F_d(m1) = \sqrt{\left(X_c - X_{m1}\right)^2 + \left(Y_c - Y_{m1}\right)^2}$$
(4)

where F_d represents the distance measurement of mouth.
In a similar way, the distance of each point from the centre point is calculated using equation (4). The coordinates of each point are calculated using trigonometry formulas.

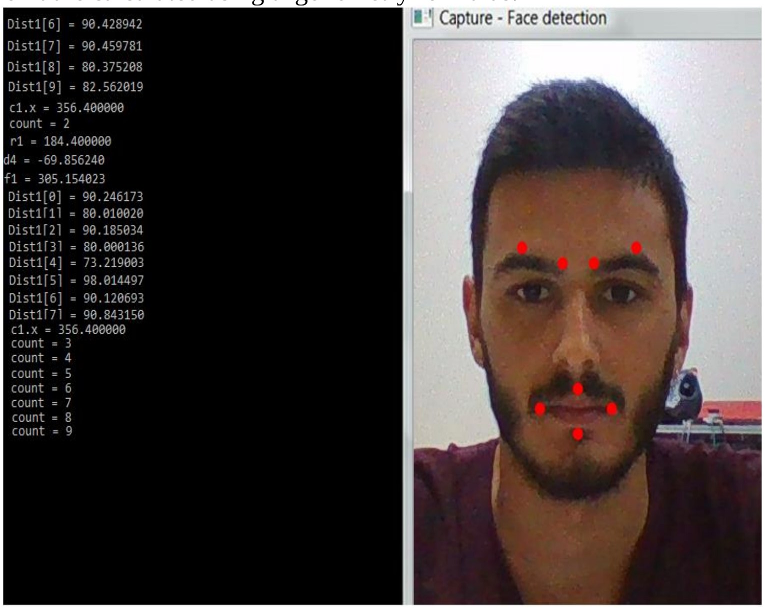

Figure 9. Facial point placement representation and each point's position: (Left: The coordinate and distance calculation, Right: Point placements at specific position

Facial Action Coding System (Facs)

When the facial distance measurements are computed with Geometric based technique, then the Facial Action Coding System (FACS) technique is used for recognizing and labelling facial expressions to describe the movement of determined facial points (muscles of the face). The FACS takes every part of the face, and breaks it down into possible movements. Each movement is given an Action Unit (AU) number. For instance, an inner brow raise has own AU of 1, a cheek raiser has an AU of 6, and a lip corner puller results in an AU of 12. There are nearly 46 basic AUs and nearly 100 AUs in total. Facial expressions are often made up of combinations of these AUs. For example, a code combination of AU6 and AU12 together represent a sincere and involuntary smile. Table 1 describes a relevant subset of AUs needed for simulation of the facial expressions for the six basic emotions.

Table 1. Set of aus needed for basic emotions

Basic expressions	Involved Aus
Suprise	AU 1, 2, 5, 15, 16, 20, 26
Fear	AU 1, 2, 4, 5, 15, 20, 26
Disgust	AU 2, 4, 9, 15, 17
Anger	AU 2, 4, 7, 9, 10, 20, 26
Happiness	AU 1, 6, 12, 14
Sadness	AU 1, 4, 15, 13

Since the measurement analysis only shows the changes in the geometric distance measurements of facial parts, the effect of AUs is required to be mapped to the observable changes in these facial measurements to analyze facial expressions. Table 2 describes a mapping between AUs and the movement of facial points. For example, AU 6, 12, and 14 all are involved in stretching c_m1 and c_m2; AU 4 and 9 pull down the inner eyebrow points c_a3 and c_a4 down.

Table 2. Mapping of AUs to measured facial points

Aus		Facial points	
AU no:	ACTION	ID	ACTION
1	up	c_a3,c_a4	up
2	up	c_a1,c_a2	up
4	down	c_a3,c_a4	down
5	up	c_m3,c_m4	up
6	up	c_m1,c_m2	strecth
7	tight	c_m1,c_m2	tight
9	wrinkle	c_a3,c_a4	down
12	pull	c_m1,c_m2	strecth
14	dimple	c_m1,c_m2	strecth
20	strecth	c_m1,c_m2	strecth
23	tight	c_m1,c_m2	tight

Emotion Extraction

In the last stage of this work, after labelling all facial points (muscles) movements using FACS, emotion interpretation is performed. In this sense, emotions are defined as strong, rush and relatively unstable mental processes which are followed by some events. The simplest way to divide emotions is to categorize them as negative, positive or neutral. In a set of negative emotions are situated such as sadness, anger or fear. The second set (positive emotions) contains emotions such as happiness and positive surprise. The last one (neutral category) contains, for example, bliss. In [12] six basic emotions and these emotions have been recognized over entire population which are anger, sadness, happiness, surprise, fear, and disgust (Figure 11). A set of facial features, which characterized an expression of each basic emotions has also been described as follows in [12].

1. **Sadness:** Inner corner of the eyebrows are raised, eyelids are loose and lip corners are pulled down.
2. **Happiness:** Muscle around the eyes is tightened, cheeks are raised and lip corners are raised diagonally.
3. **Fear**: Eyebrows are pulled up and together, upper eyelids are pulled up and mouth is stretched.
4. **Surprise:** Entire eyebrows are pulled up, eyelids are also pulled up and mouth is widely open.
5. **Anger:** Eyebrows are pulled down, upper lids are pulled up, lower lids are pulled up and lips may be tightened.
6. **Disgust:** Eyebrows are pulled down, nose is wrinkled and the upper lip is pulled up.

Figure 10. Facial expression of six basic emotions a) Sadness, b) Happiness, c) Fear, d) Surprise, e) Anger, f) Disgust

This paper only concentrates on happiness, sadness, and angriness.

The key point in happiness is that only outer lips are raised towards the ears as shown in Figure 12.

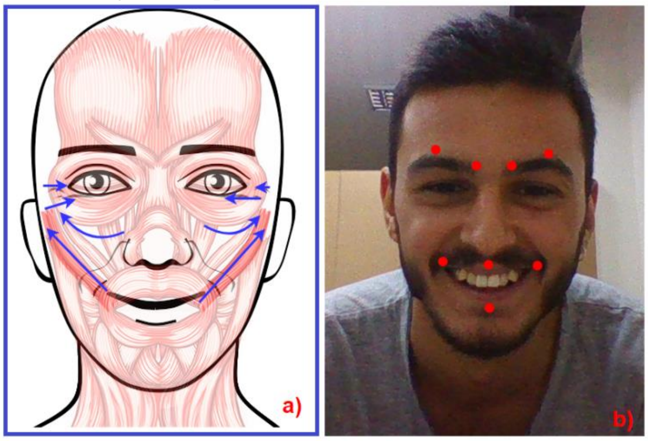

Figure 11. Happiness analysis; a) Muscular movement representation, b) Point movement representation

The key is to detect angriness is how the inner eyebrows come together and down simultaneously. This can also be accompanied by on open mouth, widened eyes, or flared nostrils. However, this paper only focuses on the movements of the eyebrows and mouth (Figure 13).

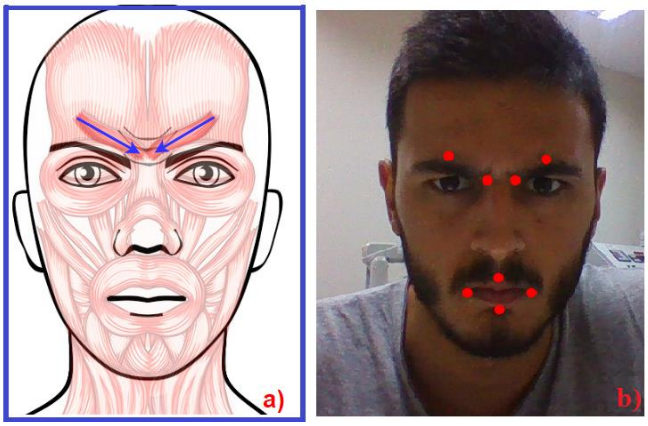

Figure 12. Angriness analysis, a) Muscular movement representation, b) Point movement representation

Finally, to analyze sadness, there are two key features which are the raising of the inner eyebrows and the pulling down of the outer lips (Figure 14).

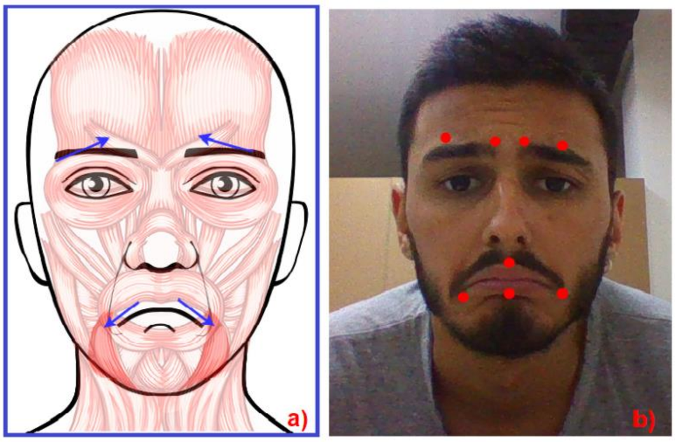

Figure 13. Sadness analysis, a) Muscular movement representation, b) Point movement representation

Implementation of the Emotion Analysis Algorithm to Nao Robot

This section briefly introduces the NAO humanoid robot and the implementation of the proposed emotion analysis algorithm.

NAO Humanoid Robot: An Overview

Nao humanoid robot is an autonomous, programmable bipedal robot developed by Aldebaran Robotics [16]. It is 58-cm tall and it is able to move, recognize, hear and even talk to human beings. Nao robot is a platform of Two-Legged Standard League [17].

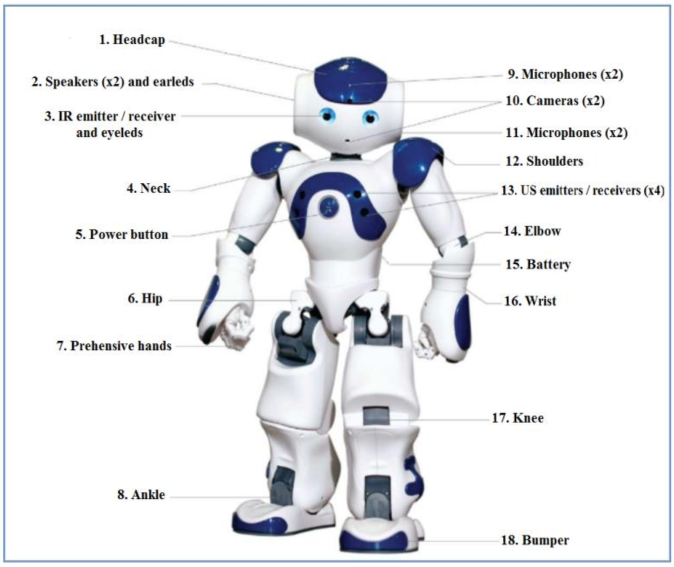

Figure 14. Nao robot

NAO requires images through two identical video cameras located in the forehead and inside the mouth to understand the emotion suing facial expressions. Both of the cameras are providing resolution up to 1280×960 at 30 frames per second.

Each NAO camera provides a vertical vision range of 47.64 degrees as well as a horizontal vision range of 60.97 degrees. Figure 16 illustrates the vision range of two cameras of NAO.

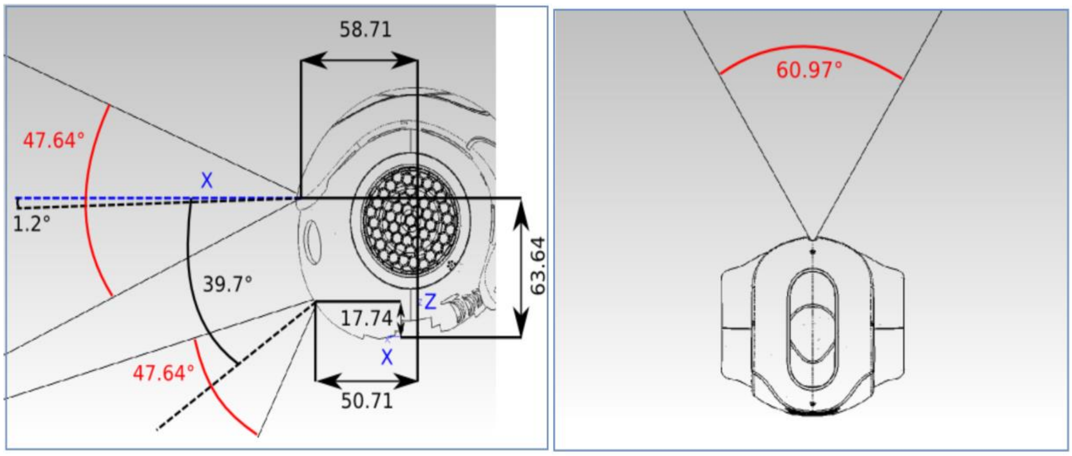

Figure 15. Vision and horizontal ranges of nao robot

Choregraphe software [18] will be used to select cameras so that NAO can detect the human face. In the next step, OpenCV [19] as well as Python [20] language have been used to further process, analyze, and understand faces.

Simulation and Experimental Results

This section evaluates the emotion analysis algorithm both in simulation and experimental environments. Initially, experimental setting is introduced and then the results are assessed.

Simulation and Experimental Settings

Facial expression analysis is performed to extract one person's three emotional states, which are happiness, sadness, and angriness in five trials. Initially, Table 1 shows the various facial distance ratios obtained with Geometric based technique to analyse facial point's movements.

Table 3: Geometric based facial distance measurement results for 1 person

Facial point placement and distanc e measurements for one person in 5 trials								
Points	Left Eye_c1		Left Eye_c2		Right Eye_c1		Right Eye_c2	
Trials	Distance	Angle in	Distance	Angle in	Distance	Angle in	Distance	Angle in
1	0,64	45,45	0,63	64,03	0,46	44,11	0,46	64,06
2	0,63	45,03	0,62	63,37	0,45	45,14	0,45	65,03
3	0,64	44,76	0,63	64,01	0,45	44,33	0,45	64,94
4	0,63	44,92	0,63	63,95	0,46	44,06	0,46	65,3
5	0,64	45,32	0,62	64,05	0,45	45,04	0,46	64,88
Average	0,636	45,096	0,626	63,882	0,454	44,536	0,456	64,842
Std Dev	0,00489898	0,254448423	0,00489898	0,258178233	0,00489898	0,462454322	0,00489898	0,416576524
Points	Left Mouth		Right Mouth		Upper Mouth		Lower Mouth	
Trials	Distance Ratio	Angle in degree	Distance Ratio	Angle in degree	Distance Ratio	Angle in degree	Distance Ratio	Angle in degree
1	0,62	129,01	0,62	130,01	0,22	90,99	0,21	90,43
2	0,61	130,23	0,61	129,21	0,21	89,67	0,21	90,04
3	0,62	131,02	0,62	131,03	0,22	91,03	0,21	90,09
4	0,61	129,4	0,62	129,48	0,21	89,45	0,21	91,17
5	0,62	129,98	0,61	130,07	0,22	90,08	0,22	91,32
Average	0,616	129,928	0,616	129,96	0,216	90,244	0,212	90,61
Std Dev	0,00489898	0,693956771	0,00489898	0,624883989	0,00489898	0,657437449	0,004	0,537661604

To perform emotion analysis, corresponding measurements determined with Geometric based technique for the eyes and mouth are initially obtained. In terms of the measurements; for the eyes, distance ratio of Left Eye_c1 (left corner distance to the centre of eye), Left Eye_c2 (left corner distance to the centre of eye), Right Eye_c1 (right corner distance to the centre of eye) and Right Eye_c2 (right corner distance to the centre of eye) are used to find the centre of eyes and distances to each other. Similarly, for the mouth measurements, left mouth (left corner distance to the centre of mouth), right mouth (right corner distance to the centre of mouth), upper mouth (top corner distance to the centre of mouth) and lower mouth (bottom corner distance to the centre of mouth) distance ratios are used to find the size of mouth (width-height). Similarly, distance ratios of each facial parts from the centre point to the left eye, right eye, mouth, above and below mouth are approximately 0.63, 0.45, 0.61, and 0.21, respectively.

After having the facial distance measurements, to determine facial expressions, specifically created rules have been defined based on [12]. For each facial expression a list of associated facial muscle changes is determined and translated in changes of facial measurement values. In this term, a subset of 7 important action units (AU1, AU2, AU4, AU7, AU12, AU15, AU27 [12]) and three facial expressions (happiness, sadness, angriness) are presented in Tables 3, respectively. In order to classify the facial muscle changes, these changes are transformed into a set of parameters, which describe the increase or decrease values in the facial measurement D_i with respect to the corresponding value in a neutral expression R_i and inc(D_i)/dec(D_i) denotes that the value of D_i has increased/decreased as shown in Table 3.

Table 4. Rules for recognizing facial action units

AU1	Raise the inner eyebrow part if inc(D1)>10 and inc(D2)>10 or inc(D3)>30 then AU1= true
AU2	Raise the outer eyebrow part if inc(D4)>12 and inc(D5)>12 then AU2=true
AU4	Lower the eyebrows if (dec (D6)<10 or dec(D7)<10 and dec(D8)<15) then AU4=true
AU12	Pull lip corners upwards if inc(D9)>6 and dec(D10)>6 then AU12=true
AU15	Press lip corners downwards if (D11)=0 and inc(D12)>8 and not dec(D13)>15 then AU15=true
AU27	Strech the mouth and pull the lower jaws downward if (D11)>= 1 cm and inc(D14)>80 then AU27=true
AU7	Tight lip corners if (dec (D12)<5 then AU4=true

Labels	Corresponding distance measurements
D1	Distance between c_a3 and face center
D2	Distance between c_a4 and face center
D3	Forehead wrinkling
D4	Distance between c_a1 and face center
D5	Distance between c_a2 and face center
D6	Distance between c_a3 and c_a4
D7	Distances of c_m1 to c_m3+c_m3 to c_m2
D8	Distances of c_m1 to c_m4+c_m4 to c_m2
D9	Mouth opening
D10	Distance between c_m1 and face center
D11	Distance between c_m2 and face center
D12	Distance between face center and D8

Results

Having done the facial distance measurements, FACS assesses these measurements to reach a decision about the emotional state of the person from their facial muscle movement measurements (AUs).

Figure 17 presents the detection rates for different AU events. The mean detection rate of 90.01% is achieved.

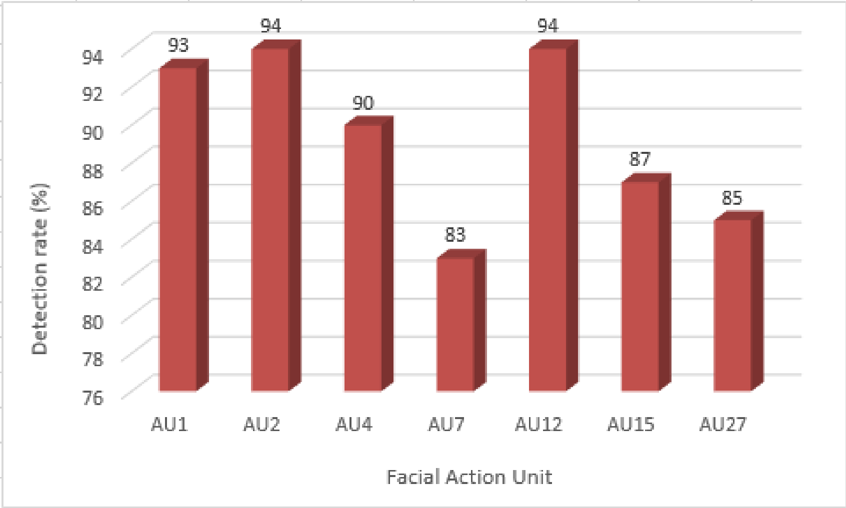

Figure 16. Facial action unit detection rates

The evaluation results are presented as a confusion matrix in Table 5. The element (i,j) of this table represents the percentage of sequence frames describing expression i, which are assigned emotion label j. The average facial expression rate is 91% for each emotional states. The highest misclassification error is reported for angriness, which 1 out of 4 times is classified as neutral. This can be attributed to the fact that most trials expressed angriness by slightly tightening lip corners and it is directly associated with relative low rate observed for AU7 and AU23.

Table 5. Facial expression rates (%)

True/classified	Neutral	Hapiness	Sadness	Angriness
Neutral	96,52	1,06	1,87	1,36
Happiness	5,02	90,88	4,73	2,71
Sadness	2,53	5,76	90,07	0,01
Angriness	13,25	0,43	0,43	88,99

Based on the facial distance measurements and facial muscle movements, three simple statistical features (mean, root mean square, and variance) are obtained to evaluate each emotional state of a person.

Table 6. Statistical results for each emotional states

Emotional states	Mean	Root mean square error	Variance
Happiness	90,99	0,004	0,109
Sadness	90,01	0,025	0,199
Angriness	89,03	0,005	0,113

Now, algorithms for face detection, facial point distance measurements, corresponding facial movement and physiognomy based emotion analysis are transferred to the Nao's interface through Choreographe as shown in Figure 18. In this application, all steps are represented with boxes, each box has own functions and these functions are connected to each other. In this paper, it is only concentrated on three emotional states which are happiness, sadness and angriness as show in Figure 18.

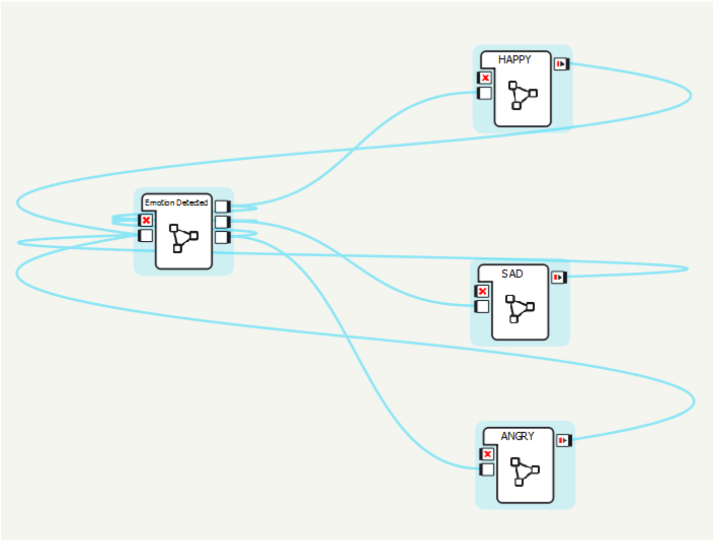

Figure 17. Emotion analysis application in choreographe

Conclusion and Further Research

In this paper, facial emotional analysis algorithm is reviewed and applied to the NAO humanoid robot. Initially, Viola-Jones algorithm is used for face detection, then, important facial distance measurements are obtained with Geometric based facial distance measurement technique. Then, facial muscle movements are obtained using Facial Action Coding System (FACS) technique. Finally, obtained facial movements are evaluated with physiognomy science to reveal emotional properties of the person. Even though, the proposed algorithm can be implemented to all humanoid robots, in this research, it has been specifically applied to NAO humanoid robot. This research will be extended and applied to various sophisticated HRI cases in near future.

References

Viola, P., & Jones, M. J. (2004). Robust real-time face detection. International journal of computer vision, 57(2), 137-154.

Jensen, O. H. (2008). Implementing the Viola-Jones face detection algorithm (Master's thesis, Technical University of Denmark, DTU, DK-2800 Kgs. Lyngby, Denmark).

Schneiderman, H. & Kanade, T. (2000). A Statistical Method for 3D Object Detection Applied to Faces and Cars.. CVPR (p./pp. 1746-1759), : IEEE Computer Society. ISBN: 0-7695-0662-3

Henry A. Rowley, Shumeet Baluja, and Takeo Kanade,"Neural network-based face detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.20, no. 1, pp. 23–38, 1998.

T. Kanade, "Picture Processing System by Computer Complex and Recognition of Human Faces," Kyoto University, Japan, PhD. Thesis 1973.

Schapire, R. (n.d.). Explaining AdaBoost. 1st ed. [ebook] Princeton. Available at: https://www.cs.princeton.edu/~schapire/papers/explaining-adaboost.pdf [Accessed 30 Nov. 2014].

R. Brunelli and T. Poggio, "Face recognition: features versus templates," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.15, pp.1042- 1052, 1993.

Brunelli, Roberto, and Tomaso Poggio. "Face recognition: Features versus templates." IEEE transactions on pattern analysis and machine intelligence15.10 (1993): 1042-1052.

Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. Environmental psychology and nonverbal behavior, 1(1), 56-75.

Ekman, P. (1993). Facial expression and emotion. American psychologist, 48(4), 384.

Loutfi, A., Widmark, J., Wikstrom, E., & Wide, P. (2003, July). Social agent: Expressions driven by an electronic nose. In Virtual Environments, Human-Computer Interfaces and Measurement Systems, 2003. VECIMS'03. 2003 IEEE International Symposium on (pp. 95-100). IEEE.

P. Ekman and W. Friesen. Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto, 1978.

Asteriadis, S., Nikolaidis, N., & Pitas, I. (2010). A review of facial feature detection algorithms. Advances in Face Image Analysis: Techniques and Technologies: Techniques and Technologies, 42.

Dhall, S., Sethi, P. (2014).Geometric and Appearance feature analysis for facial expression recognition. International Journal of Advanced Engineering Technology.

Gongor, F., Tutsoy, O. (2017). NAO Humanoid Robot Makes Facial Character Analysis. International Journal of Social Robotics.

https://www.ald.softbankrobotics.com/en

Chown, E., & Lagoudakis, M. G. (2014, July). The standard platform league. In *Robot Soccer World Cup* (pp. 636-648). Springer, Cham.

"Choregraphe User Guide." Aldebaran Robotics. Web. Aug 2012.

http://opencv.org/

"Python NAOqi API ." Aldebaran Robotics. Web. Aug. 2012.

Sebanz, N., Knoblich, G., & Prinz, W. (2005), How two share a task: Corepresenting stimulus-response mappings. Journal of Experimental Psychology: Human Perception and Performance, 31, 1234–1246.

Michael, J. (2011). Shared emotions and joint action. Review of Philosophy and Psychology, 2(2), 355–373.

Barros, P., Weber, C., & Wermter, S. (2015, November). Emotional expression recognition with a cross-channel convolutional neural network for human-robot interaction. In Humanoid Robots (Humanoids), 2015 IEEE-RAS 15th International Conference on (pp. 582-587). IEEE.

L. Ballihi, A. Lablack, B. Amor, I. Bilasco, and M. Daoudi, "Positive/negative emotion detection from RGB-D upper body images," in Face and Facial Expression Recognition from Real World Videos, ser. Lecture Notes in Computer Science, Q. Ji, T. B. Moeslund, G. Hua, and K. Nasrollahi, Eds. Springer International Publishing, 2015, vol. 8912, pp. 109–120.

Marsella, S., Gratch, J., and Petta, P, "Computational Models of Emotion." In Scherer, K.R., BA.nziger, T., & Roesch, E. (Eds.) A blueprint for an affectively competent agent: Cross-fertilization between Emotion Psychology, Affective Neuroscience, and Affective Computing. Oxford: Oxford University Press, 2010.

Le, Q.A., Hanoune, S. and Pelachaud, C. "Design and Implementation of an expressive gesture model for a humaoid robot." 11th IEEE-RAS International Conference of Humanoid Robots (Humanoids 20122), 2011.

Breazeal, C., "Emotion and sociable humanoid robot". Int. J. of Human Computer Studies, vol. 59, pp.119-155, 2003.

Islam, M. N., & Loo, C. K. (2014, November). Geometric feature-based facial emotion recognition using two-stage fuzzy reasoning model. In *International Conference on Neural Information Processing* (pp. 344-351). Springer, Cham.

Tsalakanidou, F., & Malassiotis, S. (2010). Real-time 2D+ 3D facial action and expression recognition. *Pattern Recognition*, *43*(5), 1763-1775.

Kharat, G. U., & Dudul, S. V. (2008). Human emotion recognition system using optimally designed SVM with different facial feature extraction techniques. *WSEAS Transactions on Computers*, *7*(6), 650-659.

Zhang, L., Jiang, M., Farid, D., & Hossain, M. A. (2013). Intelligent facial emotion recognition and semantic-based topic detection for a humanoid robot. *Expert Systems with Applications*, *40*(13), 5160-5168.