Facial Expression Recognition Using Weighted Distance Transform

Syed Muhammad Rafi[1] Shahzad Nasim[2]

Sheikh Muhammad Munaf[3] Mohsin Khan[4]

Abstract

Facial emotions of humans transfer non-verbal signals which have a dynamic role in interactive communication. Human-machine interface evolves according to facial expression recognition because both have a significant relationship. Psychology, ethical science, and robotics are necessary applications of facial expression recognition. A lot of work has been done already on feature extraction, face detection, and the famous techniques used for expression recognition. Weighted distance is the basic method of this research. It is used for recognition of all basic human emotions, such as anger, happiness, disgust, fear, neutral, sadness, and surprise. For the extraction of weighted distance paths, a fast-marching algorithm is used, and the seed point is taken on the nose tip of the human face. Diverse number of paths have also been taken, and they have had an effect on facial expression recognition. Intensity variation is the main motivation to use Weighted Distance Transform. Because the facial intensity variations or facial curvatures of most human beings are different, the accuracy of final evaluation may be increased and achieved in a respective manner by applying it. JAFFE (Japanese Female Facial Expression) database is used, and it is composed of 213 facial images of 10 Japanese female models with all seven basic emotions. The dimensions of JAFFE database are 256x256, and all the images are frontal position view. Twenty points are labelled for the calculation of feature vectors. Different mathematical measures are calculated as a feature vector of this geometric representation. Diverse seed locations are also being taken during research. Total four seed locations have been taken, and dissimilar number of points have also been applied for achieving better grades in the final evaluation. In classification, KNN is used and it illustrates reasonable results. In the end, validation is done with famous techniques of facial expression recognition.

Keywords: Weighted Distance, Human Behavior, Psychological Aspects, Euclidean, Fast Marching, HCI, Robotics, KNN

1. Introduction

Human faces are entities of great status in our daily lives. They present us with the

[1,3,5]*Department of Software Engineering, Faculty of Engineering, Science, Technology and Management, Ziauddin University, Karachi*

[2]*Department of Management Science and Technology, Begum Nusrat Bhutto Women University, Sukkur*

personality of the soul we are looking at, and convey information on attractiveness, age, and other traits. Neuroscience, social psychology, and cognitive science are the main applications of human emotions. They play a prominent role in human cognition [1]. These emotions of human expression could also reflect a vital role in one to one communication [2]. By using facial expression, sign languages encode the part of grammar [3]. Darwin proposed the basic rules of expression, and grouped various kinds of expressions into similar categories like hatred-anger, low spirit-dejection etc. He also stated that human expression facial fit in with human evolution [4]. Following figure-1 showed the idea of Darwin's thinking in the 19th century. It illustrated the behavior of a person with electrical equipment against the normal kind of joke.

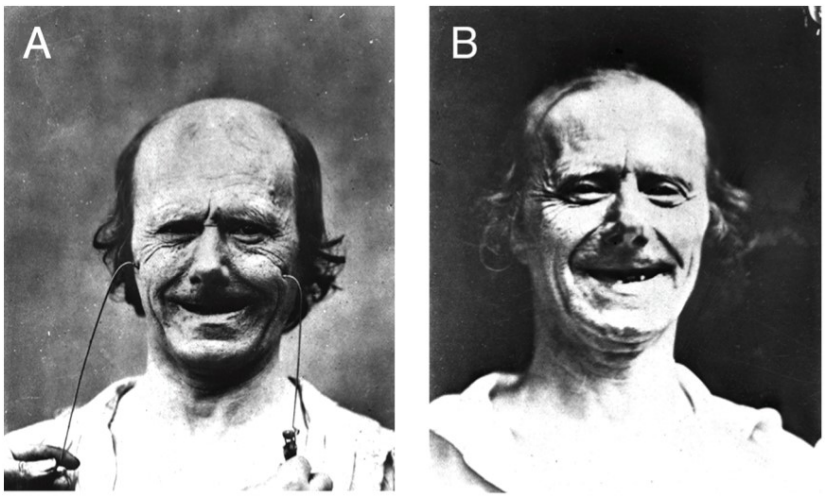

Figure 1: (A) Smile produced when zygomatic major muscles were electrically stimulated (B) Smile generated when subject was told a joke

Facial expression of emotion is much involved in the advancement of many scientific areas. It is because computer vision and machine learning researchers are concerned with developing computational representation of the perception of human face emotion to assist studies in the above sciences [5]. Computerized models of human emotions are very necessary in human computer interaction (HCI) systems, and also the key development of artificial intelligence [6]. Facial expressions are an explicit means by which people accommodate to their social ecology [7]. Face recognition has a vital value of security issues but facial expression recognition always has above hand over face recognition. Psychological research [8], match to distinct universal emotions classified six facial expressions: fear, anger, happiness, disgust, sadness, and surprise as seen in following figure-2.

Figure 2: Different Human Emotions

As discussed above, facial expression has been studied by psychologists, clinical practitioners, and also actors and artists who are interested to read about facial expressions to enhance their capability.

The famous book of Le Brun "The Perfect Imitation of Gemini Facial Expression" [9] was the key major book to achieve the best artistic ability in the 18th century. However, over the last quarter century, with the advances in the field of computer vision, animators, computer graphics, and robotics, computer scientists started delivering great interest to explore facial expressions. Shape characterization on microscopic images is also a reasonable approach in medical diseases evaluation [10]. Robotics is also the main factor of facial expression recognition, especially in humanoid robots. Many scientists believe that robots are the final destination of the facial expression recognition system. As the robots begin to interact more and more with humans, they need to develop extra and sharp intelligence in terms of understanding human moods and emotions. This is the basis of human computer interaction (HCI) community to build computers close to humans. Robots and affect sensitive HCI have also opened a new domain to use expression recognition systems in Animations, Telecommunication, Video Games, Automobile Safety, and Education-related Software etc. [11].

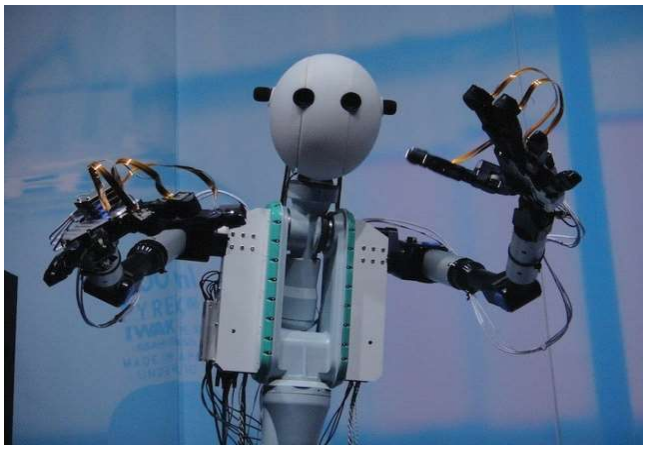

Figure 3: Telesar V robot can see, feel, and hear

For facial expression recognition, there is the available wide range of databases on spontaneous and posed emotions. In this research, JAFFE [12] database is used which

consists of 213 images of 10 Japanese female models. This database covers all major emotions, including anger, disgust, happiness, fear, sadness, surprise, and neutral. These images are captured in the Psychology Department at Kyushu University.

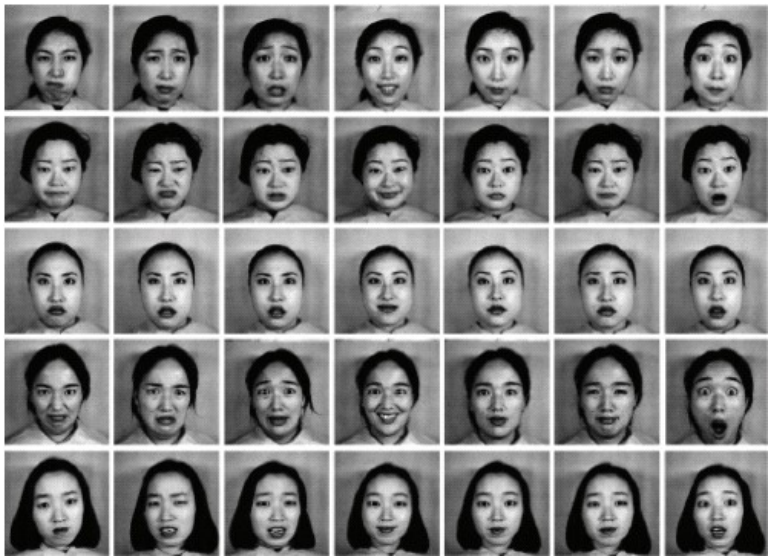

Figure 4: Anger, Disgust, Fear, Happy, Neutral, Sad, Surprise (JAFFE DATABSE)

Weighted distance is widely used for 2D images, surface segmentation, meshes, and feature extraction in different research articles [13, 14, 15]. For the extraction of facial curvatures of 2-D frontal face images, weighted distance is used because it retains intensity curvatures which works in the spatial-intensity domain. It considers the local intensity variations and also uses spatial distance between neighboring pixels. The weighted distance includes redundant information related to facial intensity curvatures. For extracting exact features, multiple paths are drawn using weighted distance transform. For the calculation of weighted distance paths, the starting point is required, and it is called the seed point. Paths are extracted by taking different seed points on the faces.

A wide variety of information is included in these paths. That's why it is very essential to parameterize the algorithm. For this purpose, different feature vectors are extracted. The dimensionality of these feature vectors is also reduced for achieving actual information. After parameterization, these feature vectors are used for the final calculation of recognition rates. For this purpose, we used the Euclidean distance for classification and in the end, comparisons are also done for the validation of method. Later in this research, a literature review and methodology is defined, then the result is shown in a respective manner to comprehensively cover a whole area of research and terminate it with references.

2. Literature Review

Facial expression recognition has attracted researchers because of its diversity. Even face recognition [16] also have respective involvement but facial expressions evaluated in many

aspects. Both Face recognition and facial expression recognition have lots of similarity in concurrence to approach. Many people from other fields, including psychologists, neurologists, and computer scientists are involved in it. Action Unit Detection in particular and emotion recognition in general has been studied broadly in the last few decades. It is not possible to review whole field comprehensively here. That's why only relevant works appear and focusing on related methods. There are two major approaches for human emotion recognition that cover most of the area of facial research. They are geometric based approach and appearance based approaches.

2.1. Geometric based Approaches

Geometric based approaches use location, distance, angle, and other relations between the face components. In this approach, it is important to discover the exact location of the face components [17]. Most of the researches on above approach are mostly about Facial Action Coding Units (AUs)[18]. AUs were mostly based on facial muscle movements. Zheng and Ji [19] also used above approach using Dynamic Bayesian Networks (DBNs). They detected 26 face features by marking around the areas of eyes, nose and mouth. The work of Kotsia and Pitas [20] also shows some significant effect. They used candid grid nodes to the facial landmarks to build a facial wire frame model for human emotion recognition and for classification purpose used Support Vector Machine (SVM). Valster et al. [21, 22] declared that geometric based systems are better than appearance based approaches. They used fiducial point on the face to extract geometrical features.

2.2. Appearance based Aproaches

Appearance based approaches use the texture or color arrangement of whole or some part of the image. In another way, this approach is mostly holistic. The local features of appearance based are much easier to calculate. Ahonen et al. [23] proposed a Local Binary pattern (LBP) method for still images. LBP was proposed by Ojala et al. [24], used texture analysis and achieved better results. Gabor filter is also very famous holistic and appearance based approaches. Gabor filters are time and memory intensive [25] for facial representations. In holistic, whole image is given as an input. Edwards et al. [26] used principal component analysis (PCA) to create Active Appearance Model (AAM). They constructed a multivariate multiple regression for modelling the relationship between the AAM displacement and the image differently. It also matched the AAM to input image in recognition phase. Images in holistic are constrained to be normalized and properly aligned. Holistic approaches also perform better in face recognition techniques [27].

2.3. Hybrid based Approaches

Holistic and local features are merged into the Hybrid approaches. For the representation

of face Yoneyama et al. [34] used a hybrid approach. For a normalized facial image, Yoneyama et al. [34] fit an 8x10 quadratic grid and in the every 8x10 regions. Geometric and Appearance are also combined to make Hybrid based approaches [28]. Weighted distance is also used but only for 3D images. It was never used before for 2D images. As described above, intensity variations are different of every two persons. That's why this algorithm got better results than previous approaches. Following figure-5 shows the implementation steps of the proposed method.

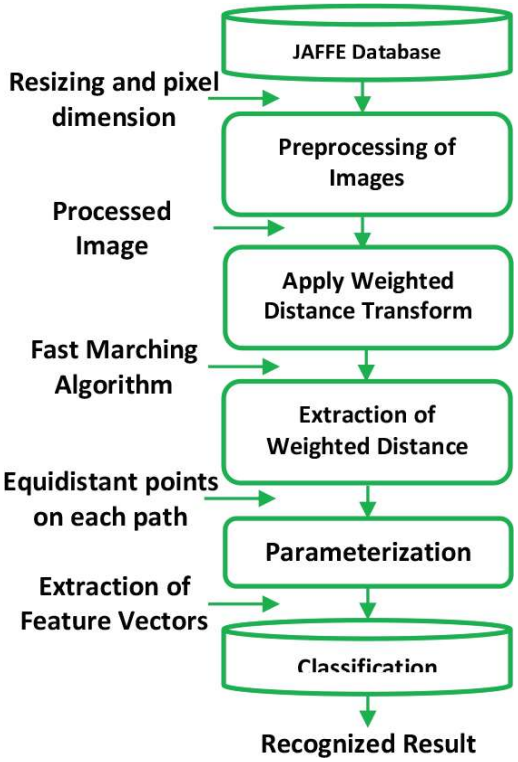

Fig 5: Implementation Steps

3. Proposed Method

The main theme of this research is facial expression recognition. Weighted distance is a very powerful technique and fast marching algorithm is used as described above for the extraction of geodesic paths. The method covered all seven major emotions that are described according to behavioral neurosciences: angry, happy, sad, fear, surprise, disgust, and neutral. The intensity variation of facial curvatures is always different of every two persons. Weighted distance only focuses on intensity variations of facial curvature and that is the requirement of this algorithm. Because of this intensity variation in human faces, we can capture more and more information which is very helpful to recognize images accurately.

3.1. Pre-processing of Image Database

The 2D images of human faces retain the intensity of pixels. Facial intensity curvatures

could be extracted from the 2D image through intensity variations of face. The resizing of the given particular image is necessary if it is not similar to the stored images. But JAFFE database that is used for the algorithm is already contained 256x256 pixels that are the requirement of our algorithm. The calculation of the weighted distance transform requires a point and this seed point is located at the nose tip on frontal faces. Algorithm took 200 images out of total 213 images of JAFFE database. 20 images of 10 persons are characterized to recognize the facial expressions. For all individual emotions, data sets are organized in all expressions separately. In this step, happy, angry, sad, surprise, fear, disgust, neutral are arranged in 3 to 4 images of every emotion along static image in each directory of facial emotions images. All images are set to 256x256 pixels, which is already the original dimension of this database.

3.2. Apply and Extract Weighted

Distance Transform

The weighted distance transform is an effort to extort intensity curvatures from a 2-dimenasional image. So the method depends on facial intensity curvatures. A fast marching algorithm is used for the extraction of paths. This fast marching algorithm [29, 30] was first introduced by James A. Sethian for solving boundary value problems which mostly related to closed curves.

$$F(a)|T(a)|=1$$

The starting point was taken from the nose tip of the frontal face because it is the requirement of weighted distance transform. This starting point is actually the seed location of database images and this work must be done manually. The algorithm took four dissimilar locations on the face to examine the facial intensity curvatures contained in geodesic distance. The weighted distance or geodesic distance calculated from every location of seed point can be seen in fig 6.

Figure 6: Seed points on different locations

As the distance transform resulted in redundant intensity curvatures, there is a need to extract more weighted distance paths in the face. It is a further attempt that extracts

additional discriminating intensity curvatures. For calculating paths, the algorithm needs ending points in addition to their corresponding seed points. We take distinct loci of end points at the elliptical edge of the face to cover the whole area of the face. Firstly, we have taken 10 paths but for obtaining sufficient information and improving the result we took 20, 30, 40, 50, 60, 70, 80, 90 and 100 paths on each individual emotions of the human face. Weighted distance and calculation of paths on the face are also extracted by using the same fast marching algorithm as seen in fig 7.

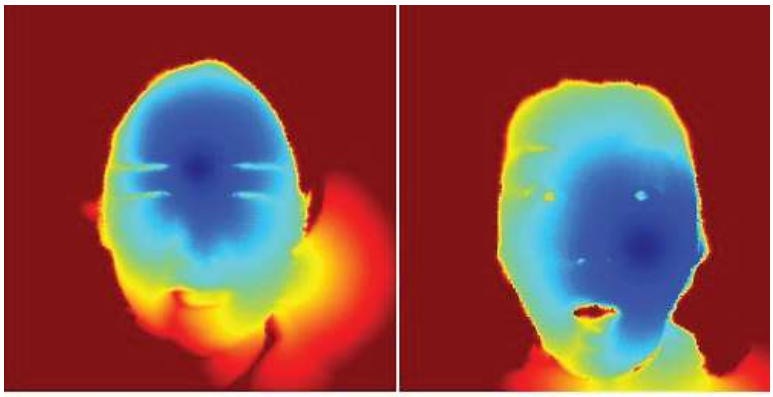

Figure 7: Weighted Distances of above seed points

3.3. Parameterization

For comparison, weighted distance paths needed to be parameterized. Selective curvatures of a path retain the main information regarding facial intensity curvatures, the parametric illustration of a path should effectively reflect the discriminating curvatures in order to enable facial expression recognition. Firstly, the algorithm is parameterized by taking two equidistant points on each path. As the number of points (taken on a path) has a direct relation to the contents of curvatures reflected in parametric representation, that's why more points are taken on paths can be seen in fig 9.

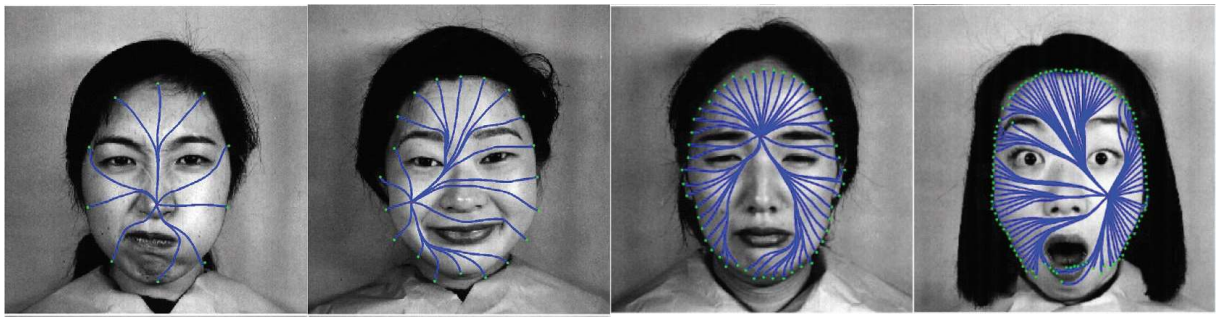

Figure 8: Paths with different seed locations

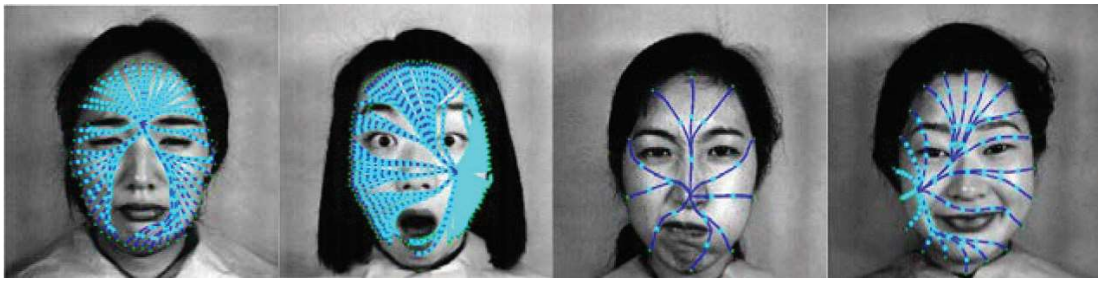

Figure 9: Equidistant points on each path

The algorithm obtained different kinds of information for every point taken from the parametric representation of a path. Coordinates, displacement from seed point, distance along the path between two successive points, and displacement between equivalent points of contiguous paths are calculated after the parameterization process. This collected information is utilized for constructing a feature space of each image.

3.4. Classification

Above feature vectors are used for classification. For this KNN (K-nearest neighbor) is used, in which the classification is done by the matching of neighboring pixels. Euclidean distance is actually following the KNN approach. Following equation represents the Euclidean distance between two points. The result showed the classification in terms of different parameters.

$$d(p,q) = \sqrt{(p_1 - p_{2)}^{\ 2} + (q_1 - q_{2)}^{\ 2}} \quad (Eq\ 2)$$

4. Results

The method is evaluated on the basis of different parameters. The evaluation is performed using four images for every individual emotion (anger, happiness, disgust, fear, surprise, sad, and neutral) of a person, where two images are taken for training and the remaining two for testing. The intensity curvatures are represented by the facial curvatures of the path retained in weighted distance. By increasing the number of paths, the representation of intensity curvatures are also improved. For this purpose, different numbers of paths are drawn on the face and algorithm took 10 to 100 paths. The representation of paths is also based on the number of equidistant points and in this research 2 to 20 points on each path are selected for showing its parametric representation. The result showed that if the number of points increases, the results became much better. The location of seed points also affected the weighted distance and four different locations for seed points have been taken. Each seed point generated different patterns for each location, as can be seen in figure 8. The effects of different feature vectors that were extracted through weighted distance transform also have significant impact on the algorithm. As discussed above, the feature space of the images are constructed through these feature vectors.

Table.1 showed the classification accuracy of all basic human emotions taken from JAFEE database, and the graph (fig 10) also presented the accuracy proposed method against the LAP (local arc pattern).

Table.1 showed the classification accuracy of all basic human emotions taken from JAFEE database, and the graph (fig 10) also presented the accuracy proposed method against the LAP (local arc pattern).

Table 1: Accuracy of Every individual

Emotions	Proposed Method	Local Arc Pattern (LAP)
Angry	97	100
Disgust	95	93
Fear	87	84
Happy	98	96
Neutral	99	100
Sad	98	87
Surprise	96	96

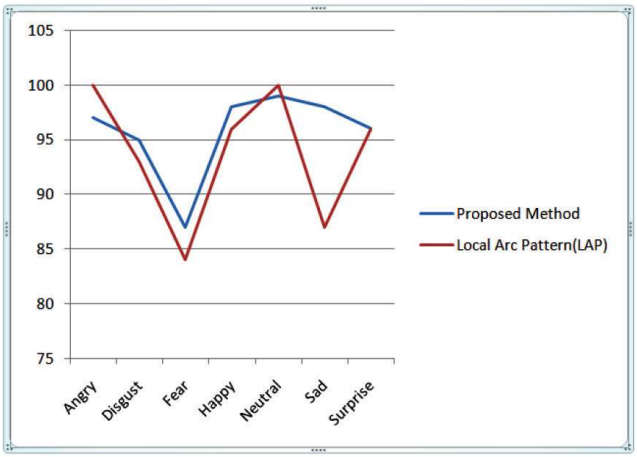

Figure 10: Comparison against LAP

5. Comparison

The Table 2 reflected the comparative evaluation where the better performance of the proposed method in terms of recognition rate can be seen. Local Arc Pattern (LAP) of Shahid et al [31], Gabor filter of Guo and Dyer [32] and M.R. Mahmood et al. [33] are compared with the proposed method.

Table 2: Comparison with LAP and Gabor Filter

Researches	Methods	Classifier	%
Proposed	Weighted Distance	Euclidean	95.71
Shahidulet al. (2013)	LAP	SVM	94.41
Guo and Dyer (2003)	Gabor Filter	Linear programming	91.00
M. R. Mahmoodet al. (2021)	Chi Square	RF/KNN	94.2

6. Conclusion and Future Work

This research showed the facial intensity curvatures could take part in a vital role in facial expression recognition. The extraction of intensity variations of human faces by using weighted distance transform showed capable results. The determination of exact loci of end points of paths created the major problem if using other databases but here it showed better results. Comparison against LAP and Gabor Filter showed comparable results which suggested that local features extraction played important role in facial expression recognition. Robustness and accuracy also increased when taking more reliable database and also some more work on facial face detection methods.

This work based on static or posed images using weighted distance transform while in future it can also be done on spontaneous images, which is actual requirement of psychological aspects. Also for face detection automatic ways are also used to achieve better result. Even if semi-automatic way is used in this research for face detection but more automatic may get better results.

References

[1] Damasio A. R., Descartes error: Emotion, reason, and the huma brain, 1stedn, G. P. Putnams Sons, (1995).

[2] Schmidt K. L. and Cohn J. F., Human facial expressions as adaptations: Evolutionary questions in facial expression research, Amer. J. of phy. anthro., 116(33), 3-24, (2001).

[3] Wilbur R. B., Nonmanuals, semantic operators, domain marking, and the solution to two outstanding puzzles in asl, Sign Lang. & Ling., 14(1), 148-178, (2011).

[4] Darwin C., The expression of the emotions in man and animals, 3rdedn, Oxford Univ. Press, (1998).

[5] Martinez A. M., Matching expression variant faces, Vision Research, 43(9), 1047-1060, (2003).

[6] Pentland A., Looking at people: Sensing for ubiquitous and wearable computing, Pattern. Analy. and Mach. Intelli., IEEE Trans., 22(1), 107–119, (2000).

[7] Martinez A., Du S., A model of the perception of facial expressions of emotion by humans: Research overview and perspectives, The J. of Mach. Lear. Resear., 98888, 1589-1608, (2012).

[8] Andersen E. N., Unconscious processing of emotional content in hybrid faces, Ph.D. thesis, Inst. of Psycho. Univ. of Oslo, 2011.

[9] Montagu J., The Expression of the Passions, final edn, Yale Univ. Press, (1994).

[10] A. Kokou M., and Antoine V., Shape characterization on phase microscopy images using a dispersion indicator: Application to amoeba cells, Res. J. of Comp. and Info. Tech. Sci., 1(5), 8–12, (2013).

[11] Bettadapura V., Face expression recognition and analysis: the state of the art, arXiv preprint arXiv,1203.6722, (2012).

[12] JAFFE database, http://www.kasrl.org/jaffe.html, (Last accessed 30/11/2020).

[13] Peyre G. and Cohen L., Surface segmentation using geodesic centroidal tesselation, Proc. 2nd Intr. Symp. on 3D Data Proces., Visuali. and Trans., (2004).

[14] Toivanen P. J., New geodosic distance transforms for gray-scale images, Pattern Recog. Letters, 17(5), 437-450, (1996).

[15] Ahsraf M., Sarim M. and Shaikh A. B., Raffat S. K., Siddiq M., Face recognition using weighted distance transform, Res. J. of Rec. Sci., (2013).

[16] Sadeghi R., A comparative face recognition algorithm for dark places, Res. J. of Rec. Sci., 2(9), 92–94, (2013).

[17] Shan, C., Gong, S., &McOwan, P. W., Robust facial expression recognition using local binary patterns,ICIP 2005, IEEE Intr. Conf. on Image Proces.,2, II-370, (2005).

[18] Ekman P. and Friesen W. V., Facial action coding system: A Technique for the Measurement of Facial Movement, 1stedn,Consult. Psycho. Press, Palo Alto, CA., USA, (1978).

[19] Zhang, Y. and Ji Q., Active and dynamic information fusion for facial expression understanding from image sequences, Patter. Analys. and Mach.Intelli., IEEE Trans., 27(5), 699-714, (2005).

[20] Kotsia I. and Pitas I., Facial expression recognition in image sequences using geometric deformation features and support vector machines, ImageProces., IEEE Trans., 16(1), 172-187, (2007).

[21] Valstar M. F., Patras I. and Pantic M., Facial action unit detection using probabilistic actively learned support vector machines on tracked facial point data,Comp. Visi. and Patter. Recog. Work., CVPR Work., IEEE Comp. Socie., Conf., 76-84, (2005).

[22] Valstar M. and Pantic M., Fully automatic facial action unit detection and temporal analysis,Comp. Visi. and Patter. Recog. Work., CVPR Work., IEEE Comp. Socie., Conf., 149-156, (2006).

[23] Ahonen T., Hadid A. and Pietikainen M., Face description with local binary patterns: Application to face recognition, Patter. Analys. and Mach. Intelli., IEEE Tran., 28(12), 2037-2041, (2006).

[24] Ojala T., Pietikäinen M. and Harwood D., A comparative study of texture measures with classification based on featured distributions, Pattern Recog., 29(1), 51-59, (1996).

[25] Bartlett M. S., Littlewort G., Frank M., Lainscsek C., Fasel I. andMovellan J., Recognizing facial expression: machine learning and application to spontaneous behavior,Comp.Visi. and Pattern Recog., CVPR 2005, IEEE Comp.Socie. Conf.,2, 568-573, (2005).

[26] Cootes T. F., Edwards G. J. and Taylor C. J, Active appearance models,Comp.Visi., ECCV,Springer Berlin, 484-498, (1998).

[27] Muhammad Sharif S. M. M. R., Shah J. H., Sub-holistic hidden markov model for face recognition, Res. J. of Rec. Sci., 2(5), 10–14, (2013).

[28] MendozaD. S., Masip D., Baró X., andLapedriza À., Emotion Detection Using Hybrid Structural and Appearance Descriptors,Model. Deci. for Artifi.Intelli., 105-116, (2013).

[29] Sethian J. A., Fast marching methods, SIAM review, 41(2), 199–235, (1999).

[30] Sethian J. A., A fast marching level set method for monotonically advancing fronts, Proc. of the National Academy of Sciences, 93(4), 1591–1595, (1996).

[31] Islam, M. S., and Auwatanamongkol S., Facial Expression Recognition Using Local Arc Pattern, Asian J. of Inf. Tech., 12(4), 126-130, (2013).

[32] Guo G., and Dyer C. R., Simultaneous feature selection and classifier training via linear programming: A case study for face expression recognition.Comp.Visi. and Patter.Recog., Proc. 2003 IEEE Comp.Socie. Conf., 1, I-346, (2003).

[33] M.R. Mahmood, M.B. Abdulrazzaq, S.R. Zeebaree, A.K. Ibrahim, R.R. Zebari, and H.I. Dino, "Classification techniques' performance evalutation for facial expression recognition." Indonesian Journal of Electrical Engineering and Computer Science, vol. 21 no.2, pp.176~1184. 2021.