International Journal of Interdisciplinary Innovative Research & Development (IJIIRD)
ISSN: 2456-236X
Vol. 03 Issue 01 | 2018

Facial Emotion Recognition Database: A Brief Review

Vishal D. Bharate[1], Devendra S. Chaudhari [2]
[1] Assistant Professor, Electronics & Telecommunication Department, Sinhgad Academy of Engineering, Pune, Maharashtra, India
[2] Professor, Electronics & Telecommunication Department, Government College of Engineering, Jalgaon, Maharashtra, India

ABSTRACT

Human-computer interaction has gained much importance in day to day life of human being. Many developments in the era of human-computer interaction gave rise for finding the solutions in the domain of daily needs of human beings. Emotion recognition is one of the critical areas where many numbers of researchers are investigating new technologies and optimizing the algorithms for ease of the system. The human emotions are contributed in facial expressions are about 55% compared to all other ways. The database required for enhancing the accuracy and predicting the exact class of human emotion from facial expressions requires a sufficient and qualitative database. This paper reviews all the human emotion databases for facial expression recognition available. The various datasets are listed, and the detailed summary is presented in this research paper. The availability of the database, along with the web address, is also listed in the paper.

Keyword: - Facial expression recognition; Human Emotion Recognition; facial action coding system; facial action unit

1. INTRODUCTION

Over the past decades, facial expression recognition got much more importance. In various applications its being used extensively. The analysis and correct recognition of human emotions is the prime motto of the researchers. The various applications need the exact recognition of emotion. Applications like surveillance, robotics, patient monitoring are emerging to give services to end-users without any hurdle. To enhance the accuracy of the human emotions for providing services to these applications system requires a robust and meaningful database.

The objective of this review paper is to have a ready reference for the collection of the datasets available publicly. To enhances the research in the field of emotion recognition using facial expressions using this publicly available dataset is the primary task. In day to day life, Interpersonal communication is carried out using facial expressions. Facial expressions are the primary source of conveying the state of mind that is emotion. In animations and affective computing, many numbers of researchers are using facial expressions for human emotion recognition.

The traditional facial expression recognition uses three steps, namely, 1. Face detection, 2. Feature extraction and 3. Classification of the expression into the different emotion class. Several face detection algorithms are readily available in the literature. Majorly different action units (AU) are decided once the face is detected. The face is categorized into various parts such as the nose, chick, chin, eye, eyebrow etc. The AUs and landmarks are fixed on the face to recognize emotion correctly.

From these AUs and landmarks, various temporal and spatial features are extracted. Finally, the pre-trained classifier may be supervised, or unsupervised is applied for the detection of emotion to fall in the correct emotion class. The collection of various datasets is readily available for face detection. However, there is no prominent collection in case of facial expression recognition is available. Many of the datasets are unequal and are modified according to their use in the different applications. Thus there is a limitation in a thorough overview and review of available datasets for human emotion recognition using facial expressions. There is diversity in the databases available. This diversity includes the number of participants, the method used for capturing the image, background, the mood of participants during recording etc. This diversity in the available datasets creates challenges while applying features on the dataset.

The major challenge in the dataset is that it is not robust. For different datasets, the age limit is restricted. Another challenge in the dataset is lightning conditions. Frontal images are only considered for preparation of datasets. Head pose, Occlusions, Difference in the facial aspects of human face structure, which varies from region to region are another set of challenges. Due to all these numbers of challenges, one feature may give the best result, but the same feature may fail to give the optimum performance on another dataset.

Some datasets have been prepared by considering the controlled environment, while some have considered a less controlled environment. This gives rise to background challenges in the processing of the images available in the datasets. Many of the researchers have used cropped faces from available datasets while some of them use the same images for recognition of human emotions.

2. INTRODUCTION TO FACIAL EXPRESSION RECOGNITION DATABASE

In the field of Facial Expression Recognition Database, many of datasets used for relative comparison and extensive experimentation. Usually, human emotions from facial expressions use either 2D static images or 2D videos. Handling of significant pose variations and delicate facial behaviours makes 2D base analysis difficult. Finely structured variations in the natural, spontaneous facial expressions make an analysis of 3D facial expressions relatively easy. The subsequent section gives a brief introduction to all available 2D and 3D video sequence datasets.

Compound Emotion (CE)[2]: CE contains 5060 images corresponding to 22 categories of primary and compound emotions for its 230 human subjects (130 females and 100 males, mean age of 23). Most ethnicities and races are included, including Caucasian, Asian, African, and Hispanic. Facial occlusions are minimized, with no glasses or facial hair. Male subjects were asked to shave their faces as cleanly as possible, and all participants were also asked to uncover their forehead to fully show their eyebrows. The photographs are color images taken using a Canon IXUS with a pixel resolution of 3000 × 4000.

The Extended Cohn-Kanade Dataset (CK+)[1]: CK+ contains 593 video sequences on both posed and non-posed (spontaneous) emotions, along with additional types of metadata. The age range of its 123 subjects is from 18 to 30 years, most of whom are female. Image sequences may be analyzed for both action units and prototypic emotions. It provides protocols and baseline results for facial feature tracking, AUs, and emotion recognition. The images have pixel resolutions of 640 × 480 and 640 × 490 with 8-bit precision for grey-scale values.

Japanese Female Facial Expressions (JAFFE)[13]: The JAFFE database contains 213 images of seven facial emotions (six basic facial emotions and one neutral) posed by ten different female Japanese models. Each image was rated based on six emotional adjectives using 60 Japanese subjects. The original size of each facial image is 256 pixels × 256 pixels. All images are stored in .tiff format. It is extensively used in several research papers. Denver Intensity of Spontaneous Facial Action Database (DISFA)[3]: DISFA consists of 130,000 stereo video frames at high resolution (1024 × 768) of 27 adult subjects (12 females and 15 males) with different ethnicities. The intensities of the AUs (0–5 scale) for all video frames were manually scored using two human experts in FACS. The database also includes 66 facial landmark points for each image in the database. The original size of each facial image is 1024 pixels × 768 pixels.

The Karolinska Directed Emotional Face (KDEF)[9]: This database contains 4900 images of human emotional facial expressions. The database consists of 70 individuals, each displaying seven different emotional expressions photographed from five different angles. The original size of each facial image is 562 pixels × 762 pixels.

Binghamton University 3D Facial Expression (BU-3DFE)[4]: Because 2D still images of faces are commonly used in FER, Yin et al. at Binghamton University proposed a database of annotated 3D facial expressions, namely, BU-3DFE 3D. It was designed for research on 3D human faces and facial expressions, and the development of a general understanding of human behaviour. It contains a total of 100 subjects, 56 females and 44 males, displaying six emotions. There are 25 3D facial emotion models per subject in the database, and a set of 83 manually annotated facial landmarks associated with each model. The original size of each facial image is 1040 pixels × 1329 pixels.

MMI[6]: MMI consists of over 2900 video sequences and high-resolution still images of 75 subjects. It is fully annotated for the presence of AUs in the video sequences (event coding), and partially coded at the frame-level, indicating for each frame whether an AU is in a neutral, onset, apex, or offset phase. It contains a total of 238 video sequences on 28 subjects, both males and females. The original size of each facial image is 720 pixels × 576 pixels.

Binghamton-Pittsburgh 3D Dynamic Spontaneous (BP4D-Spontaneous)[8]: BP4D-spontaneous is a 3D video database that includes a diverse group of 41 young adults (23 women, 18 men) with spontaneous facial expressions. The subjects were 18–29 years in age. Eleven are Asian, six are African-American, four are Hispanic, and 20 are Euro-Americans. The facial features were tracked in the 2D and 3D domains using both person-specific and generic approaches. The database promotes the exploration of 3D spatiotemporal features during subtle facial expressions for a better understanding of the relation between pose and motion dynamics in facial AUs, as well as a deeper understanding of naturally occurring facial actions. The original size of each facial image is 1040 pixels × 1329 pixels.

Extended Yale B face (B+)[5]: This database consists of a set of 16,128 facial images taken under a single light source, and contains 28 distinct subjects for 576 viewing conditions, including nine poses for each of 64 illumination conditions. The original size of each facial image is 320 pixels × 243 pixels.

Table 1 shows a summary of these publicly available databases. It is necessary to create the emotions on the face of participants for which emotion has to be captured. Based on this, facial emotions are classified into posed, spontaneous and induced. Definitely mood of the participant at the time of capturing the emotion may hamper the accuracy of the dataset.

2.1 Posed

Posed emotion is nothing but an assumption of the feeling of the human being to be transformed on the face of a human being. It is captured by giving training to the participants. Most of the databases consist of posed facial expressions. Examples are CK, Chen-Huang and Banse-Scherer. It is a simple technique to have the facial expression on the face of a human being using posed expressions. In an earlier time, all databases were created utilizing this technique. They may be exaggerated or missing the slight details in that expression. Due to these technical difficulties when these datasets are used they give minimum accuracy in the classified results. The professional actors are asked to create such kind of database to overcome these problems.

Spontaneous

Real-life scenarios are being recorded to form spontaneous emotion datasets. When humans don't know the recording is going on, he or she can express emotion in natural form. However, to prepare these kinds of datasets, it is sturdy and requires much time to prepare it. In this context, to prepare datasets most of the challenging issues discussed so far occur. There is a possibility of occurrence of several errors while capturing the dataset. The labeling for these kinds of datasets is done manually. Also, it may happen that the same person may have changed the facial look in between may create the problem.

Also, there exists another kind of datasets derived from the films, YouTube videos and television serials. These videos are fitted to achieve the mean of posed and spontaneous facial expressions. Such type of datasets containing posed as well as sometimes spontaneous reactions of the participant. Another issue to create such type of datasets is that always frontal for angled images cannot be captured for all the participants.

Induced

The most natural form of emotions of an individual is induced form. The audiovisual information is presented in front of the participant. The videos are recorded to achieve the desired goal of capturing the emotion while this audiovisual clip is played in front of the participant. Several datasets are prepared to employ this technique. There is much improvement in the preparation of dataset in terms of capturing natural emotions. Fake or overemphasized expressions can be overcome by utilizing this technique.

Table 1. A summary of publicly available databases of Facial Expressions.

Name of Database	Database Description	Web Link of Database
CE [2]	5060 images corresponding to 22 categories of primary and compound emotions 230 human subjects (130 females and 100 males, mean age 23) Includes most ethnicities and races The image resolution of 3000 × 4000	http://cbcsl.ece.ohio-state.edu/ dbform_compound.html
CK+ [1]	593 video sequences on both posed and non-posed (spontaneous) emotions 123 subjects from 18 to 30 years in age Provides protocols and baseline results for facial feature tracking, action units, and emotion recognition Image resolutions of 640 × 480, and 640 × 490	http://www.consortium.ri.cmu. edu/ckagree/

JAFFE [13]	213 images of seven facial emotions Ten different female Japanese models Six emotion adjectives by 60 Japanese subjects The image resolution of 256 × 256	http://www.kasrl.org/jaffe_info.html
DISFA [3]	130,000 stereo video frames at high resolution 27 adult subjects (12 females and 15 males) 66 facial landmark points for each image The image resolution of 1024 × 768	http://www.engr.du.edu/mmahoor/DISFA.htm
KDEF [9]	4900 images of human facial expressions of emotion 70 individuals, seven different emotional expressions with 5 different angles The image resolution of 562 × 762	http://www.emotionlab.se/resources/kdef
BU-3DFE [4]	3D human faces and facial emotions 100 subjects in the database, 56 females and 44 males, with about six emotions 25 3D facial emotion models per subject The image resolution of 1040 × 1329	http://www.cs.binghamton.edu/~lijun/Research/3DFE/3DFE_ Analysis.html
MMI[6]	Over 2900 video sequences and high-resolution still images of 75 subjects 238 video sequences on 28 subjects, male and female. The image resolution of 720 × 576	https://mmifacedb.eu/

3. CONCLUSIONS

This paper gives a brief introduction to human emotion recognition using facial expressions. For the beginners, in the field of facial expression recognition, this paper gives the ready reference for the datasets available. The paper has elaborated challenges in the datasets, information related to datasets and availability of the datasets. To make research robust in the field of human-computer interaction broad experimentation is required. To make a robust feature, it has to be tested on various datasets along with different supervised and unsupervised classifiers. Different applications based on facial expression recognition can be optimized and designed efficiently with the rapid growth in facial expression recognition. Patient monitoring, depression, robotic applications etc can be made more and more impactful with the help of facial expression recognition.

5. REFERENCES

[1] Lucey, P.; Cohn, J.F.; Kanade, T.; Saragih, J.; Ambadar, Z.; Matthews, I. The extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression. In Proceedings of the IEEE Conference on Computer Vision and Pattern RecognitionWorkshops, San Francisco, CA, USA, 13–18 June 2010; pp. 94–101.
[2] Tao, S.Y.; Martinez, A.M. Compound facial expressions of emotion. Natl. Acad. Sci. 2014, 111, E1454–E1462.
[3] Mavadati, S.M.; Mahoor, M.H.; Bartlett, K.; Trinh, P.; Cohn, J. DISFA: A spontaneous facial action intensity database. IEEE Trans. Affect. Comput. 2013, 4, 151–160.
[4] Yin, L.; Wei, X.; Sun, Y.; Wang, J.; Rosato, M.J. A 3D Facial Expression database for facial behaviour research. In Proceedings of the International Conference on Automatic Face and Gesture Recognition, Southampton, UK, 10–12 April 2006; pp. 211–216.
[5] B+. Available online: https://computervisiononline.com/dataset/1105138686
[6] MMI. Available online: https://mmifacedb.eu/
[7] Yan, W.J.; Li, X.; Wang, S.J.; Zhao, G.; Liu, Y.J.; Chen, Y.H.; Fu, X. CASME II: An improved spontaneous micro-expression database and the baseline evaluation. PLoS ONE **2014**, 9, e86041.
[8] Zhang, X.; Yin, L.; Cohn, J.; Canavan, S.; Reale, M.; Horowitz, A.; Liu, P.; Girard, J. BP4D-Spontaneous: A high resolution spontaneous 3D dynamic facial expression database. Image Vis. Comput. **2014**, 32, 692–706.
[9] KDEF. Available online: http://www.emotionlab.se/resources/kdef.
[10] Die große MPI Gesichtsausdruckdatenbank. Available online: https://www.b-tu.de/en/graphic-systems/databases/the-large-mpi-facial-expression-database.
[11] Kaulard, K.; Cunningham, D.W.; Bülthoff, H.H.; Wallraven, C. The MPI facial expression database—A validated database of emotional and conversational facial expressions. PLoS ONE 2012, 7, e32321.
[12] Lyons, M.J.; Akamatsu, S.; Kamachi, M.; Gyoba, J. Coding facial expressions with Gabor wave. In Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition, Nara, Japan, 14–16 April 1998; pp. 200–205.
[13] M. Kamachi, M. Lyons and J. Gyoba, "The Japanese female facial expression (JAFEE) database," International Conference on Automatic Face and Gesture Recognition, 1997.

ISSN: 2456-236X

Vol. 03 Issue 01 | 2018

6. BIOGRAPHIES

	**Vishal D. Bharate** obtained BE, from Shivaji University, Kolhapur and M.Tech. from Dr. Babasaheb Ambedkar University, Lonere. He is pursuing Ph. D. from Government College of Engineering, Amravati. He has been engaged in teaching, research for about 10 years. Presently he is working as Assistant Professor, Department of Electronics and Telecommunication Engineering at Sinhgad Academy of Engineering, Pune, Maharashtra, India. He has published research papers and presented papers in international conferences at Chennai, Kanyakumari. He has published 1 Journal and 4 Conference papers. His present research and teaching interests are in the field of Digital Image Processing, Digital Signal Processing, Computer Vision and Machine Learning.
	**Devendra S. Chaudhari** obtained BE, ME, from Marathwada University, Aurangabad and PhD from Indian Institute of Technology, Bombay, Powai, Mumbai. He has been engaged in teaching, research for about 30 years. He has worked as Head Electronics and Telecommunication, Instrumentation, Electrical, Research and incharge Principal at Government Engineering Colleges. Presently he is working as Head, Department of Electronics and Telecommunication Engineering at Government College of Engineering, Jalgaon. Dr. Chaudhari published research papers and presented papers in international conferences abroad at Seattle, USA and Austria, Europe. He is fellow of IE, IETE and life member of ISTE, BMESI and member of IEEE (2007). He is the recipient of Best Engineering College Teacher Award of ISTE, New Delhi, Gold Medal Award of IETE, New Delhi, Engineering Achievement Award of IE (I), Nashik. He has organized various Continuing Education Programmes and delivered Expert Lectures on research at different places. He has also worked as ISTE Visiting Professor and visiting faculty member at Asian Institute of Technology, Bangkok, Thailand. His present research and teaching interests are in the field of Biomedical Engineering, Digital Signal Processing and Analogue Integrated Circuits.