

Short Communication

Fan-Beam Projection-Based Feature Extraction for Facial Expression Recognition

A. Sherly Alphonse *

Ponjesly College of Engineering, Nagercoil, India

* Correspondence: sherls82@gmail.com

Received: 10 January 2022; **Accepted:** 18 January 2022; **Published:** 19 January 2022

Abstract: This paper presents a novel method of feature extraction using Fan beam projection-based data. The Fan-beam projection covers the image completely and hence gathers all the important information. Even though the image quality is distorted, this type of feature extraction method helps to gather all the important information as there is a huge volume of projection data. Also, the use of multiple detectors speeds up the entire process. All the projections of the image together form a sinogram image which is unique for each facial expression image. Hence, the sinogram image is divided into grids and the histogram formation results in a feature vector for each image. The classification of these feature vectors using Radial Basis Function-based Extreme learning Machine (RBF-ELM) results in high classification accuracy for all the datasets.

Keywords: feature; expression; fan-beam; emotion; classification; image

1. Introduction

Feature extraction is the process through which the essential and non-redundant data are extracted from a raw image that increases the between-class distance and decreases the within-class distance. There are lots of feature extraction techniques in literature like Local Binary Pattern (LBP), Local Directional Pattern (LDP), Maximum Response-based Directional Texture Pattern (MRDTP) etc. LBP sometimes can't represent the structure effectively in its feature vector as it leaves the center pixel information. LDP and MRDTP filter the image using Kirsch masks and the prominent edges are considered for the calculation. This may leave out some of the needed information. When these feature vectors are used for the classification of facial expressions, there is a reduction of accuracy. This paper proposes a novel method of extracting features called fan-beam projection-based feature extraction. This method considers the projection data taken from an image at different angles. The source is rotated at different angles of certain intervals. The distance between the source and the object's origin point is initialized so that it covers the entire image. Thus, the projection data obtained using fan-beams are a matrix whose dimensions are determined by the number of sensors used and a number of angles used for the rotation of the fan-beams. This projection data

captures unique information from the raw image. So, it is converted to feature vector by observing the micro-patterns in the projection data. These feature vectors can group the images from the same class very effectively.

2. Materials and Methods

The human eyes can very well recognize the patterns in the data. But if the dimension of the data is more than three it becomes more difficult to recognize it using a human eye. So, feature extraction techniques are used to recognize patterns and to perform classification using a classifier. This paper proposes a novel feature extraction technique that focuses on fan-beam transforms and its projection data. After cropping the portion of the image from the available raw images in the datasets, the images are converted to gray scale images. Then, the feature extraction techniques can be applied for extracting the feature vectors. The fan-beam projections are the variations of Radon transforms. Using fan-beam transformation functions [1], the projections of an image can be taken at specific directions. In this proposed work, the features are extracted from the projection data obtained using Fan-beam geometry. Here, the distance from the source of fan-beam to the rotation center is chosen. should be chosen so that the source of fan-beam

is outside the image rotated at all angles as in Figure 1. The distance of the diagonal image is calculated as

$$L = a^2 + b^2 \quad (1)$$

where, a and b are the height and width of the image. Then, D is chosen to be greater than. D is the distance between the fan-beam vertex and the center of the object used for rotation. This rotation center is the center pixel which is given as

$$\text{floor}((\text{size}(\text{image}) + 1) / 2) \quad (2)$$

The values used for rotation angles are obtained incrementally in degrees. The beams diverging from the source here are like fan. The source emits the beams and the detector receives the image as in Figure 1. Angle β represents the position of the source and the coordinates σ, β are in the ranges as $\pi / 2 \leq \sigma \leq \pi / 2$ and $0 \leq \beta \leq \pi / 2$. The projection rays are represented using σ and β as in Figure 2.

The coordinates of the fan beams are V and θ as in (3) and (4).

$$V = D \sin \sigma \quad (3)$$

$$\theta = \sigma + \beta \quad (4)$$

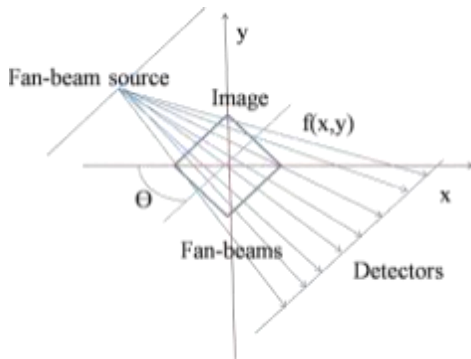


Figure 1. Fan beam geometry.

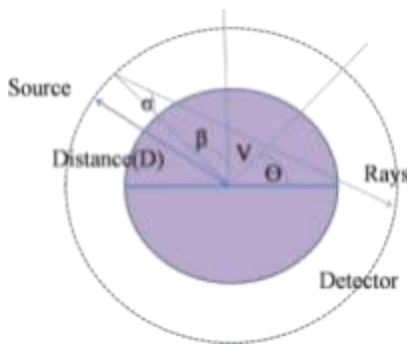


Figure 2. Fan beam projection at rotation angle θ .

The projection data obtained at one rotation angle are represented as one column of the sinogram. Thus, the

fan-beam projection data or sinogram is computed from the image I. This method of collecting sinograms by projections using fan-beams is faster and also represent a large amount of information regarding each image. This sinogram has a unique pattern for each image as in Figure 3.

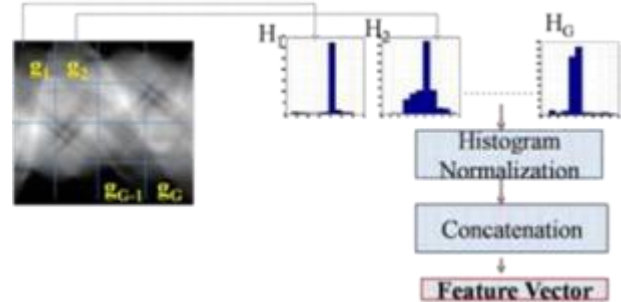


Figure 3. Fan-beam projection-based feature vector for a JAFFE image.

Then the sinogram for each image is divided into grids and the histograms for each grid are collected to form a feature vector f_{vi} as in (5).

$$f_{vi} = \langle H_1, H_2, \dots, H_G \rangle \quad (5)$$

Where i the feature vector of the image i , and G is the total number of grids. For classifying the feature vectors to seven different categories of emotions (Anger, Disgust, Happy, Fear, Sadness, Surprise, and Neutral) four different classifiers have been used and their performance variations with respect to time and accuracy have been discussed in the experimental results section.

3. Experimental Results and Discussion

In the experiments five different datasets like Japanese Female Facial Expressions (JAFFE) [2], Multimedia Understanding Group (MUG) [3], Man-Machine Interaction (MMI) [4,5], Static Facial Expressions in the Wild (SFEW) [6,7], Oulu, the Machine Vision Group, Chinese Academy of Science and Institute of Automation(Oulu-CASIA) [8] etc. that are collected under different constraints and environments are taken for the analysis of the performance of the proposed method. In Table1, the total number of images used from each of the dataset and the accuracies obtained are given. Ten-fold cross-validation is used for the classification purpose. Grid search algorithm is used for fixing the parameter values. In Table 1, (Radial Basis Function-based Extreme Learning Machine) RBF-ELM is used for classification.

Table 1. Comparison of the classification accuracy obtained using some of the existing feature descriptors and the proposed one.

Feature Descriptor \ Datasets	JAFFE	MUG	MMI	SFEW	Oulu-CASIA
	213 Images	567 Images	1050 Images	1394 Images	3500 Images
LBP	90.2	91.2	80.0	28.0	83.4
LDP [10]	92.1	92.3	80.3	30.9	83.5
LTP [11]	93.2	95.6	79.3	30.2	84.3
MRDTP [12]	94.3	95.6	82.4	34.4	86.3
MRDNP [12]	92.8	94.7	82.3	34.1	86.0
E-Gabor [13]	92.8	92.4	81.4	34.9	85.9
MDP [14]	96.2	98.4	83.0	35.4	87.5
Proposed	96.6	98.5	83.1	35.6	88

In Table 2, (One Vs. One Support Vector Machine) OVO-SVM achieves good accuracy by consuming less time for multi-classification compared to (One Vs. All Support Vector Machine) OVA-SVM. In (Deep Neural Network) DNN, activation function is used. DNN takes a huge amount of time for training and classification. But, RBF-ELM attains good accuracy in less time as in Table 2.

Table 2. Comparison of the classification accuracy obtained using different classifiers for JAFFE and the time used.

Classifiers	Accuracy (%)	Time Taken (Seconds)
OVA-SVM [15]	96.6	1.212
OVO-SVM [15]	96.8	1.114
RBF-ELM [16]	96.6	0.415
DNN [17]	96.5	19.014

In Table 3, for an image of size 202 the sinogram size is 71 because 71 diverging beams are used to cover that image completely and 360 rotation angles are used. Finally, after dividing that sinogram into grids of sizes 10 and obtaining the histograms using 10 bins the feature vectors of length 840 is obtained. Similarly, for images of various resolutions accordingly the number of beams needed to cover the image varies. Therefore, the feature vector size and hence the accuracy also varies.

Table 3. Comparison of the classification accuracy, dimension of the Sinograms and feature vector sizes obtained using the proposed approach for JAFFE images of different imagesizes.

Imagesize	Sinogram Size	Feature Vector Size	Accuracy (%)
202×202	71×360	840	96.6
302×302	119×360	1320	97.1
102×102	35×360	360	92.3

4. Conclusions

In this paper, a novel, simple and effective feature extraction algorithm based on Fan-beam projection is proposed, that is based on unique sinograms produced as a result of the projection of facial expression images. This method is faster, simpler and also achieves high classification accuracy when compared to the existing methods. The performances have been tested on five datasets. Also, four different types of classifiers have also been used and the best classifier is indicated.

References

1. Kak, A.C., Slaney, M., 1988. Principles of Computerized Tomographic Imaging. IEEE Press, NY. pp. 92- 93.
2. Lyons, M., Akamatsu, S., Kamachi, M., Gyoba, J., April 1998. Coding facial expressions with gabor wavelets. Third IEEE International Conference on Automatic Face and Gesture Recognition. Washington, DC, USA. pp. 200-205.
3. Aifanti, N., Papachristou, C., Delopoulos, A., April 2010. The MUG facial expression database. 11th Int. Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), Italy. pp. 1-4.
4. Pantic, M., Valstar, M., Rademaker, R., Maat, L., July 2005. Web-based database for facial expression analysis. In Multimedia and Expo, IEEE International Conference, Amsterdam. pp. 5.
5. Valstar, M., Pantic, M., 2010. Induced disgust, happiness and surprise: an addition to the MMI facial expression database. In Proc. 3rd Intern. Workshop on EMOTION (satellite of LREC): Corpora for Research on Emotion and Affect. pp. 65.
6. Dhalla, A., Goecke, R., Lucey, S., Gedeon, T., 2012. Collecting large, richly annotated facial expression databases from movies. IEEE MultiMedia. 19, 34-41.
7. Dhalla, A., Goecke, R., Joshi, J., Sikka, K., Gedeon, T., 2014. Emotion recognition in the wild challenge 2014: Baseline, Data and Protocol. ACM ICMI.
8. Zhao, G., Huang, X., Taini, M., Li, S.Z., Pietikainen, M., 2011. Facial expression recognition from near-infrared videos. Image and Vision Computing. 29(9), 607- 19.
9. Ojala, T., Pietikainen, M., Maenpaa, T., 2002. Multiresolution gray-scale and rotation invariant texture

- classification with local binary patterns. *IEEE Transactions on pattern analysis and machine intelligence*. 24(7), 971-87.
10. Jabid, T., Kabir, M.H., Chae, O., 2010. Robust facial expression recognition based on local directional pattern. *ETRI journal*. 32(5), 784-94.
 11. Tan, X., Triggs, B., 2010. Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE transactions on image processing*. 19(6), 1635-50.
 12. Alphonse, A.S., Dharma, D., 2018. Novel directional patterns and a Generalized Supervised Dimension Reduction System (GSDRS) for facial emotion recognition. *Multimedia Tools and Applications*. 77(8), 9455-88.
 13. Alphonse, A.S., Dharma, D., 2017. Enhanced Gabor (E-Gabor), Hypersphere-based normalization and Pearson General Kernel-based discriminant analysis for dimension reduction and classification of facial emotions. *Expert Systems with Applications*. 90, 127-45.
 14. Alphonse, A.S., Dharma, D., 2017. A novel Monogenic Directional Pattern (MDP) and pseudo-Voigt kernel for facilitating the identification of facial emotions. *Journal of Visual Communication and Image Representation*. 49, 459-70.
 15. Huang, G.B., Zhou, H., Ding, X., Zhang, R., 2012. Extreme learning machine for regression and multi-class classification. *IEEE Transactions on Systems, Man, and Cybernetics*. 42(2), 513-529.
 16. Chang, C.C., Lin, C.J., 2011. LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technology*. 2(3), 27.
 17. Keyvanrad, M.A., Homayounpour, M., Aug. 2014. A brief survey on deep belief networks and introducing a new object oriented toolbox (DeeBNet). *arXiv:1408.3264 [cs]*.



Copyright © 2022 by the author(s). Published by UK Scientific Publishing Limited. This is an open access article under the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).