Edge Detection for Facial Expression Recognition

Jesús García-Ramírez, Ivan Olmos-Pineda, J. Arturo Olvera-López, Manuel Martín Ortíz

Faculty of Computer Science, Benemérita Universidad Autónoma de Puebla, Av. San Claudio y 14 sur. Puebla, Pue. C.P. 72570, México
gr_jesus@outlook.com,{iolmos, aolvera, mmartin}@cs.buap.mx

Abstract. Nowadays, Facial Expression Recognition is an active research subarea of computer vision because of its applications in different human activities. In this paper a method for detecting edge information in human face images is introduced, where Robinson Edge technique is applied. In addition, a thresholding threshold values are proposed to reduce the noise in the images with the aim to improve results. Based on experiments of the first stage, it is possible to detect facial expression in eyes, eyebrows, mouth, and forehead with good accuracy.

Keywords. Facial Expression Recognition, Image pre-processing, Edge Detection.

1 Introduction

Computer vision is a research area which has several applications, in most of the digital image processing is applied for detecting relevant image information. An example of computer vision application is the Facial Expression Recognition (FER) that can be applied to detect mental disorders, detect whether a person is lying, detection of emotions, among others. Face images are used in FER systems, and it is common to find Regions of Interest (ROI's) because processing the whole image could be computationally expensive. To extract the ROI's information, there exist several methods like those based on image textures, others are based on image thresholding or locate image points. The main ROI's that are extract in face image are those containing eyebrows, eyes, nose and mouth.

The process of FER system starts with the image pre-processing using filters like smoothing, border detection, transformations into a color space different of RGB, among others. The image pre-processing stage processes the face images for the feature extraction stage in which the relevant information is extracted. Commonly the pre-processing stage has as main objective to find the border information in face images, but in most of the times it is difficult to do, because transitions between the background and the face are very soft, for that reason it is important to find the edge information and denoise the images as much as possible.

The face image processing is commonly to applied over RGB (Red, Green and Blue) images; however, the edge detection algorithms are applied to gray scale images. Sev-

eral edge detection algorithms have been developed, like Otsu and Isodata, these algorithms use the image histogram to find a threshold for binarize images, methods like Laplace and Sobel apply a convolutional operation to every pixel in the image. Robinson edge detection is a filter based on find maximum transitions in different directions, depending of the convolutional matrix applied over images.

The edge detection in face images is a non trivial task because images have soft transition in the edges, for dealing with this problem a thresholding process based on Robinson edge detection is proposed in this paper to find relevant information in face images for the feature extraction stage, the earlier results are presented too. The images used in this work were taken from MMI Facial Expression Database [12].

This paper is organized as follows: Section 2 introduces the problem to solve, Section 3 presents the related works of FER, Section 4 shows the proposed thresholding method for edge detection, finally in Section 5 the conclusions and future work are presented.

2 Facial Expression Pre-processing

The main objective of a Computer Vision System (CVS) is to process images as human does, to simulate a human eye a CVS uses an image capture device. To analyze the image, a computer system uses a learner to classify or detect what is happening in the images, in this case the facial expression from a captured human image.

The interactions between humans have the characteristic that emotions can be expressed in a conversation, on the other hand, computers cannot neither express nor show emotions, for this reason the natural interaction between a computer system and a human is computationally difficult.

The process that several researchers follows to develop FER system is described in this paragraph. The input of the system is an image or multiple images (video), then a pre-processing stage is applied, image filters, border detection algorithms, threshold methods are applied to prepare the images for the next stage that is the feature extraction, in this stage, the images are encoded to numeric values, which are descriptive about the image content. Finally the system uses a model previously constructed (by a learner or classifier) to classify person's facial expression in the input image.

The pre-processing stage consist in applying image filters, this process could be computationally expensive because every pixel in the image is taken into account to extract the necessary information. In most of the cases, it is not enough applying only one image filter, but as many filters are applied over the image more runtime is required.

In Fig. 1 it is shown the general approach commonly followed to obtain edges from face images: first a gray scale operator is applied to input images, then filters about smoothing, threshold, border detection, among others are applied to find edge information, finally a binarization is commonly used for preserving in either black or white the edge information.

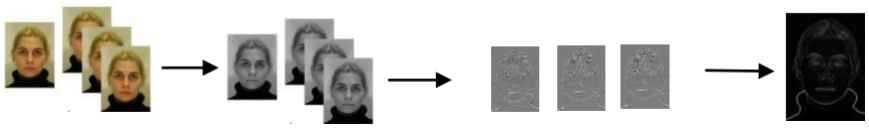


Fig. 1. Process used for pre-processing face images: a) input as image or a set of images (video), b) a gray scale transform is applied c) different images filters are applied, d) a binary image as output.

3 Related Works

FER is a research area developed since 1970 when Paul Ekman presented the six universal facial expressions: happiness, anger, sadness, disgust, surprise, and fear. This is an important contribution because most of the approaches find these emotions and others extend them with the neutral expression [1]. Then Ekman presents the Facial Action Coding System (FACS), this codification system specifies 9 AUs in the upper face, and 18 in the lower face. In addition, there are 14 head positions and movements, 9 eye position and movements, 5 miscellaneous actions units, 9 action descriptors, 9 gross behavior and 5 visible codes [2].

The first stage of a FER system consist in pre-process the images, in this process the system prepares the images for the feature extraction, before applying any image filter a transformation to another color space distinct to RGB color model like YUV or HSV is carried out. Torres et al. report a FER system with transformation to other color spaces; it is shown that they can provide different information by the RGB color space as is shown in Fig. 2 [3].

In [4] a comparison among four edge detectors: Robert, Sobel, Laplace and Canny is reported, having this last the best performance. Bourel et al. [5] locate 12 points in face images related to the eyebrows, nose, mouth and eyes, then distances between some of them are found and they are stored for the feature selection stage. Deepak Ghimire et al. locate points in face image to segment the face into regions and find a Local Binary Pattern (LBP) histogram for each region [6]. Other approaches combine models 3D and LBP to find the ROIs in images [7].

The feature extraction stage consists in registering the information of pre-processed images; there exist three methods to register the information: the whole face as a full entity, only the ROI's of the face and face edge. An example for register the whole face is to split the face image and use Gabor filters [8], for local registration a histogram of a 3D image is saved [9,10]. Other example of local registration is to save a region LBP histogram [6]. Cohen Sebe et al. register information about Motion Units [11].

The final stage in facial expression recognition process is the classification, it uses the feature extraction information to find a model, those models are used to predict or classify, the model extraction is based on a dataset i.e. the training set, when the dataset is categorized it is called supervised learning, otherwise it is called unsupervised learning. N. Sebe et al. present a comparison among four classifiers: Naïve Bayes, Bayesian

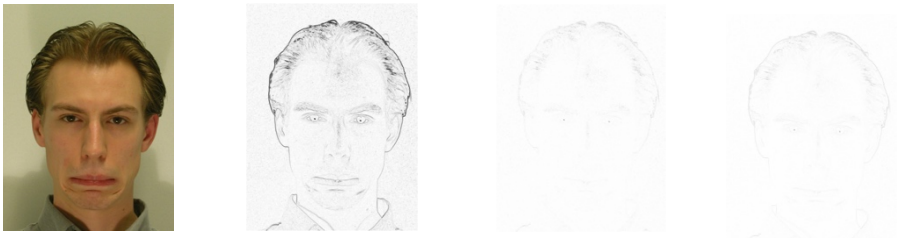
Network, Decision Trees, and Nearest Neighbors, among these, the best performance was obtained by Decision Tree MC4 [11]. Two approaches that implements a 3D model and a classification with Hidden Markov Models (HMM) are presented in [10, 13] but they only find three expressions (Happy, Sad, Surprise), other approaches that implement HMM for classify are presented in [14, 15], Neural Networks are used in [8], and also Support Vector Machines in [6].

a) b) c) d)

Fig. 2. a) Image in RGB color model, b) channel H, c)channel S, d) channel V.

4 Edge detector based on Robinson

In this section the proposed methodology for detecting borders in a face image is presented. Several edge detectors can be applied to face images, but there are cases in which edge detectors do not obtain a good performance because the image has soft transitions in edges as it can be seen in Fig. 3. In this figure it is shown the result of apply Sobel, Laplace and Gradient edge detector filters, as a visual comparing these border detectors have not a good performance in face images.



a) b) c) d)

Fig. 3. Result image of apply edge detectors, a) RGB image, b) Sobel, c) Laplace, d) Gradient

Other Edge detectors can be applied like Robinson algorithm; it is based on detect maximum transitions in different directions in an image. This edge detector uses a convolutional matrix of dimension 3 by 3 over an image, there are many variants of the convolutional matrix, the main difference is that this matrix is rotated. In our method the initial convolutional matrix is rotated 90 degrees as a result four convolutional matrixes are applied to faces images and the border is found according to the maximum among four different orientations (90,180,270 and 360 degrees). The result images are presented in the Fig. 4: Fig. 4a shows RGB images that are the input of the pre-processing stage; at Fig. 4b it is shown the result of applying the Robinson filter with a convolutional matrix rotated 90 degrees to images from Fig. 4a.

The next step is to denoise the result images after Robinson filter is applied, and then a threshold is found. The function for find the new values in the image is described in expression (1).

$$I'(x,y) = \begin{cases} 0 \ if \ \ 0 < I(x,y) \leq \gamma_1 \\ 128 \ if \ \ \gamma_1 < I(x,y) \leq \gamma_2 \\ 255 \ if \ \ \gamma_2 < I(x,y) \leq 255 \end{cases} \quad (1)$$

Where $I(x, y)$ is an intensity pixel value in (x, y) position, the main objective of this process is to binarize the image and get only the pixels with information about borders, the best values in this case of $\gamma_1 = 105$ and $\gamma_2 = 160$, the noise of the Robinson edge detector appears in the image results of this process (see Fig. 4c). This noise can be reduced with a smooth filter, the filter that is applied is the median filter that takes the median value of a pixel in a 3 by 3 region and it is assigned to the center pixel. The result after applying median filter discards both salt and pepper noise located in isolated regions.

Fig. 4d shows the results after a median filter is applied, it can be seen that the images show the information about the edges of the ROI's in face images and the wrinkles, this information is relevant when micro-expressions are studied.

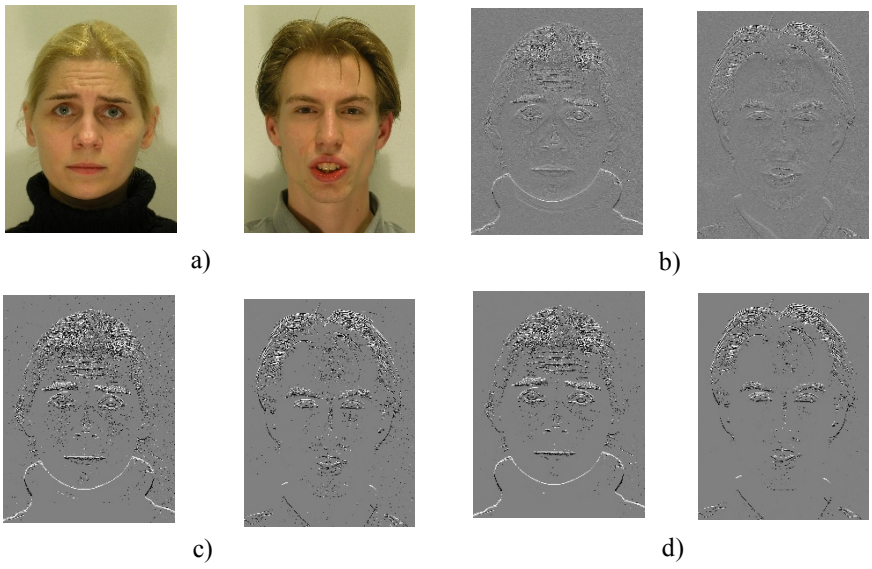


Fig. 4. Result images in the pre-processing stage a) images in RGB color space, b) result of apply Robinson edge detector, c) result of apply thresholding method, d) result of apply median filter.

In Fig. 5 it is shown the algorithm of our approach, this algorithm has as input an RGB color model image *I*, then the Robinson edge detector with a convolutional matrix rotated 90 degrees is applied to I (I_R). For denoise the salt and pepper noise in I_R a median filter is applied (I_M), finally a binarization process is applied to I_M following the expression (1), this is the result image and is the output of the algorithm.

Binarizing Method based on Robinson Edge Detector(BMRE)
Input: I (Image in RGB color model), Output: R (Result image in gray scale)
01. $I_{GS} \leftarrow$ Transform I to gray scale
02. $I_R \leftarrow$ Apply Robinson edge detector rotated 90 degrees to I_{GS}
03. $I_M \leftarrow$ Apply Median filter to $I_R$for discard salt an pepper noise
04. $R \leftarrow$ Binarize the image with the rule (1)

Fig. 5. BMRE algorithm for binarizing a face image in RGB color model.

In Fig. 6b it is shown a pseudo 3D intensity model of the image and figure 6b depicts the same model for the thresholded image. Based on Fig. 6b, it is clear that the high intensity values show the border transition in image, and the low intensity values shows noise in the image. From the pseudo 3D information, it can be noticed that after applying our method for edge detection the main delimiting points defining eyes, mouth and eyebrows are preserved which is useful to extract features about the expression in the image.

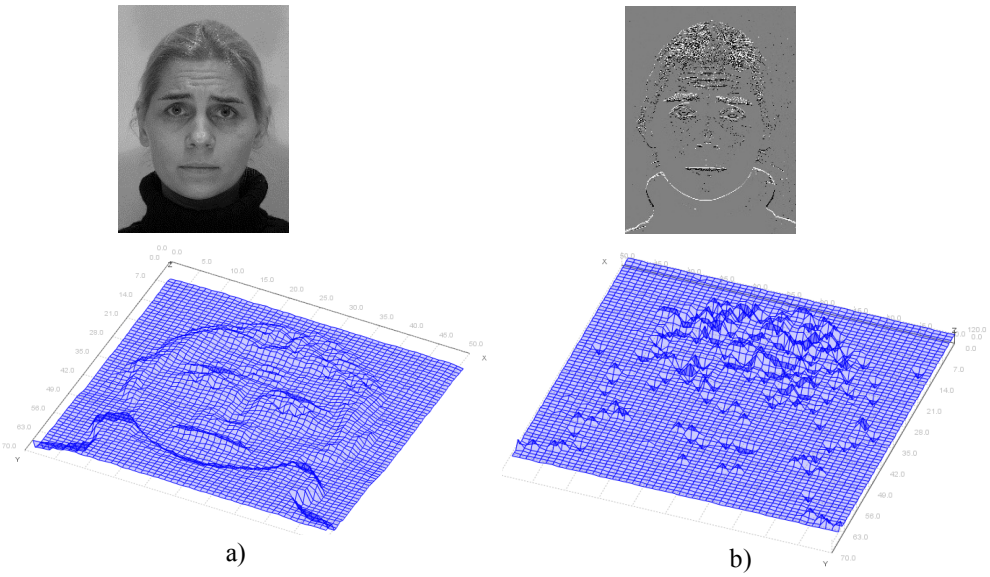


Fig. 6. Pseudo 3D intensity model that shows intensity of pixels a) shows the image in gray scale image b) shows the thresholded obtained by our method.

In Fig. 7 it is shown the detail of the result images of the proposed method. In Fig. 7a-b shows the result images after applying our method, in Fig. 7c-d it is shown the left eye, note that in the four images the edge information is visualized, the eyebrows are important to get the expression that the person shows. In Fig. 7e-f the mouth information is shown, this region shows relevant information about the feeling of the person in face images. Finally in the Fig. 7g-h information of the upper face are shown, note the expressions marks in the front face, in Fig 7h there are not marks information but hair information is marked.

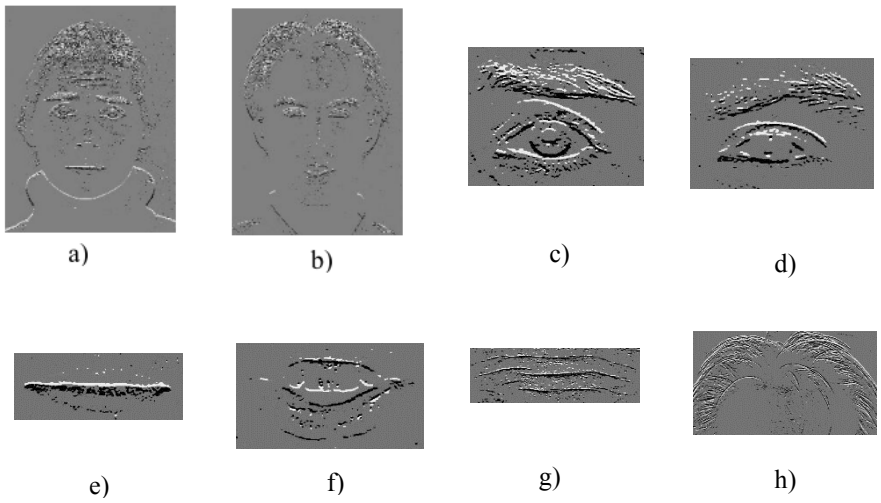


Fig. 7. Detail of the face information in result images a) and b); eye and eyebrow edge detail c) and d); mouth information, e) and f); upper face mark lines g) and h).

5 Conclusions

The image pre-processing stage is the first stage of the systems of FER systems, as it can be seen in this paper pre-process stage consist in apply image filters to find information of ROI's from face images, the first results of our approach are presented. We propose a method to find the ROIs' in face images with an approach that applies Robinson edge detector to get the edge information, then a noise reduction filter is applied.

Finally, a pseudo 3D gray intensity model is analyzed for finding more information about the edge, mainly the transitions between intensities are better noticed in the pseudo 3D model.

For future work, we will continue analyzing images into the frequencies space, then we will work on the second stage of the feature extraction and classification. In addition, we will construct a facial expression database from both visible and thermal (infrared) ranges, which will be used for testing our future FER approach.

Acknowledgements

The first author of this work thanks to CONACyT for supported this work by the Mastering Scholarship 701191.

References

1. Ekman P.: Strong evidence for universals in facial expressions: a reply to Russell's mistaken critique, Psychological Bulletin, 115(2), 268–287 (1994)
2. Cohn J., Ambadar Z., Ekman P.: Observer-based measurement of facial expression with the Facial Action Coding System. In: J. A. Coan & J. J. B. Allen (Eds.), The handbook of emotion elicitation and assessment. New York, pp. 203-221 (2007)
3. Torres L., Reutter J., Lorente L: The importance of the color information in face recognition. In: Proceedings of the International Conference on Image Processing, (ICIP 99), Vol. 3, pp. 627–631 (1999)
4. Xiaoming C., Wushan C.: Facial expression recognition based on edge detection, International Journal of Computer Science & Engineering Survey, 6(2) 1– 9 (2015)
5. Bourel F., Chibelushi C., Low A.: Robust Facial Expression Recognition Using a State-Based Model of Spatially-Localized Facial Dynamics. In: Proceedings of the Fifth IEEE International Conference Automatic Face and Gesture Recognition, pp. 106–111 (2002)
6. Ghimire D., Jeon S, Lee J., Park S.: Facial Expression Recognition based on Local Region Specific Features and Support Vector Machines, Multimedia Tools and Applications, 1–19 (2016)
7. Essa I., Pentland A.: Coding Analysis, Interpretation, and recognition of facial expression, IEEE Transaction on Pattern Analysis and Machine Intelligemce, 19(7), pp. 757–763 (1997)
8. Wenfei G., Cheng X., Venkatesh Y., Dong H., Han L.: Facial expression recognition using radial encoding of local Gabor Features and classifier synthesis, Pattern Recognition 45(1), pp. 80–91 (2012)
9. Tang H., Huang T.: 3D facial expression recognition based on automatically selected features. In: Proceedings of the Computer Vision and Pattern Recognition (CVPR 2008), pp. 1–8 (2008)
10. Sandbach G., Zafeiriou S., Pantic M.: Local normal binary patterns for 3D facial action unit detection. In: proceedings of the International Conference of Image Processing (ICIP 2012), pp. 1813–1816 (2012)
11. Cohen I., Sebe N., Gozman F, Cirelo M., Huang T.: Learning bayesian network classifiers for facial expression recognition both labeled and unlabeled data. In: Proceedings of Computer Vision and Pattern Recognition (CVPR 2003), pp. I–595–I–601 (2003)
12. Pantic M., Valstar M., Rademarker R., Maat L.: Web-based database for facial expression Analysis. In: Proceedings of International Conference on Multimedia and Expo, pp. 5-10 (2005)
13. Le V., Tang H., and Huang T.: Expression recognition from 3D dynamic faces using robust spatio-temporal shape features. In: Proceedings of Face and Gesture Recognition (FG 2011), pp. 414–421 (2011)
14. Sariyanidi E., Gunes H., Cavallaro A.: Automatic Analysis of Facial Affect: A survey of Registration, Representation, and Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 37(6), 1113–1133 (2015)
15. Corneau C., Oliu M., Cohn J., Escalera S.: Survey on RGB, 3D, Thermal, and Multimodal Approaches for Facial Expression Recognition: History, Trends, and Affect-related Applicants. IEEE Transactionon Pattern Analysis and Machine Intelligence 99, 1–20 (2015)