A Novel Facial Expression Recognition Method Using Fast BEMD Based Edge Detection

James Z. Zhang[†], Zijing Qin[††], Peter C. Tay[†], and Robert D. Adams[†]

Department of Engineering and Technology
Kimmel School, Western Carolina University
Cullowhee, NC 28723 U.S.A.
Email: [†]zhang, ptay, radams{@email.wcu.edu}, [††]ginger227227@gmail.com

***Abstract*— Traditionally, human facial expressions are recognized using standard images. These methods of recognition require subjective expertise and high computational costs. This article presents a novel facial expression recognition method using edge maps generated from standard images that can significantly improve computational efficiency. The edge maps are generated from a novel Bi-dimensional Empirical Mode Decomposition (BEMD) based edge detection method. In this paper, the BEMD edge detection algorithm is discussed, the facial expression decision metrics are developed, and detection results of facial expression databases are presented. The success rates of recognition suggest that the proposed method be a potentially efficient method for human facial expression detection and other similar recognition applications.**

***Index Terms*—** facial expressions, edge detection, BEMD

1. INTRODUCTION

Human facial expression recognition has been an active research area for many years, mainly in the fields of psychology and cognitive sciences. Basic emotions can be categorized with six basic facial expressions: "Happy," "Sad," "Fear," "Surprise," "Angry," and "Disgust" [1-5]. The aim of the research focuses on the studies of identifying human behaviors and intents. A number of methods have been developed for facial recognition including Gabor wavelets [6] [7], linear discriminant analysis [8], local feature analysis [9], and independent component analysis [10]. These methods are based on the analysis of original human facial expression images which is labor intensive and of high computational cost. Because human expertise are needed in decision-making and interpretation processes, real-time determination of a facial expression is still far from reality. Hence, an automatic analysis method with reduced computational complexity is needed to detect nonverbal behaviors conveyed by facial expressions.

Edge detectors may be a natural candidate for such a method. As one of the primary means for image processing, edge detection of an image significantly reduces the amount of data to be processed and filters out information that may be regarded as less relevant, while preserving the important structural properties of an image. This makes the subsequent task of interpreting the information contents in the original image substantially simplified [11]. There are many ways to perform edge detection. The gradient method detects the edges by looking for the maxima and minima in the first derivatives of the image. Most predominant edge detectors such as Sobel [12], Prewitt [13], and Canny [14] are based on gradient measures. Other methods for edge detection include the Marr operator [15], Nalwa-Binford, the Sarkar-Boyer, and 2D-wavelet edge detectors [16] [17] [18].

Recently, a BEMD-based edge detector was presented [19]. This detector extracts the edge map of an image from its first Bi-dimensional Intrinsic Mode Function (BIMF). Our research applies this detector to facial expression recognition and the results are presented in this article.

2. BEMD DETECTOR AND DETECTION CRITERIA

2.1 BEMD Edge Detector

BEMD is a generalization of the EMD method [20]. BEMD algorithm extracts Bi-dimensional Intrinsic Mode Functions (BIMF) that are intrinsic to the original image through a sifting procedure. The BIMFs of a signal obtained by EMD are expected to have the following properties:

1. The number of extrema and the number of zero-crossings must either be equal or differ at most by one.
2. At any point, the mean value of the envelope defined by the local maxima and the envelope by the local minima is zero.

978-1-4244-9721-8/10/$26.00 ©2010 IEEE

Denote I as the original image, a BIMF as F, and the residue as R. In the decomposition process, the i^{th} BIMF F_i is obtained from its source image S_i, where S_i is a residue image obtained as $S_i = S_{i-1} - F_{i-1}$ and $S_i=I$. It requires one or more iterations to obtain F_i, where the Intermediate Temporary State of BIMF (ITS-BIMF) in j^{th} iteration can be denoted as F_{Tj}. The sifting procedure of the BEMD process can be summarized as follows:

1. Set i=1. Take I and set S_i=I.
2. Set j=1. Set $F_{Tj}=S_i$.
3. Obtain the local maxima and local minima maps of F_{Tj}.
4. Form the upper envelope and lower envelope of F_{Tj} by 2D-interpolation of the maxima and minima points.
5. Find the mean envelope: $M_{Ej} = \dfrac{U_{Ej} + L_{Ej}}{2}$
6. Calculate F_{Tj+1} : $F_{Tj+1} = F_{Tj} - M_{Ej}$
7. Check whether F_{Tj+1} follows the BIMF properties.
8. If F_{Tj+1} meets the criteria given in step 7, take $F_{Tj} = F_{Tj+1}$. Set i=i+1, $S_i = S_{i-1}$. And proceed to the next step (Finding the BIMF's). If not, set j=j+1, repeat step 3-8 until the stopping criteria is fulfilled.

A fast algorithm [21] was employed in our edge detector. BEMD decomposes an image *"I"* into *"K"* BIMFs (F_k being the k^{th} BIMF) and a residue term *"R"*. The original image can be reconstructed by:

$$I = \sum_{k=1}^{K} F_k + R \qquad (1)$$

The stopping criteria for BIMF extraction include many parameters but the most important one is the standard deviation defined in (2). This parameter affects the quality of the first BIMF, therefore the quality of the edge map.

$$SD = \sum_{x=1}^{M} \sum_{y=1}^{N} \frac{|F_{k+1}(x,y) - F_k(x,y)|^2}{|F_k(x,y)|^2} \qquad (2)$$

Because the first BIMF contains highest frequency contents in the signal, edge information of an image is contained in this BIMF. Therefore, for edge detection, extraction of the first BIMF is sufficient. The edge map can be obtained with a properly chosen threshold level, and morphological operations are used to further enhance the edge map [20].

2.2 Facial Expression Recognition Criteria

Facial expressions are generally determined by extraction of certain facial features. Based on these features, expressions are divided into two main categories: Category I (C-I) includes "Happy," "Sad," and "Fear," and Category II (C-II) includes "Surprise", "Angry", and "Disgust." Some predominant features include mouth width D1, mouth depth D2, and the distance between eyebrow and lower edge of eyelid D4, as shown in figure 1, respectively [22].

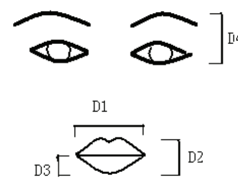


Fig. 1 Facial skeleton and characteristic distances

In this research, we found that the distance between one corner of the mouth and the lower edge of the lips, D3, to be an important parameter to distinguish "Happy" and "Sad" expressions. Therefore, we included D3 as the 4th metric.

The procedure of developing facial expression recognition criteria is summarized as follows:

1. A reference dataset is chosen. The dataset consists of 7 images: A "Neutral" face and 6 typical facial expressions.
2. D1-D4 are measured on edge map of the "Neutral" face, these measurements are used as references.
3. D1-D3 of each expression are measured and D4 is only needed for the expressions of "Angry," "Surprise" and "Disgust". These data are then converted as percentage changes from corresponding "Neutral" measurements.

The above criteria lead to the following facial expression recognition procedures (detailed in the next section):

1. D1 is used to determine if the expression is in C-I or C-II.
2. For C-I expressions, D3 is applied to distinguish the facial expressions in this category: "Happy," "Sad," and "Fear."
3. For C-II expressions, D2 and D4 are used to recognize the facial expressions "Angry," "Surprise," and "Disgust."

3. EXPERIMENTATION AND RESULTS

3.1 Determination of Recognition Criteria

The reference dataset chosen for developing recognition criteria is shown in figure 2.

Fig. 2 Reference Dataset for Recognition Criteria

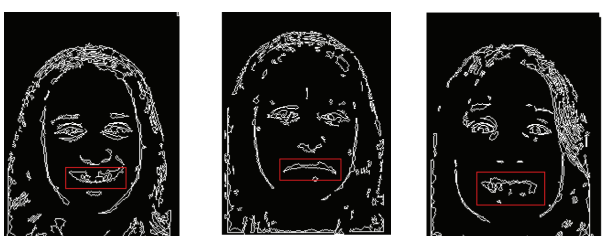
The dataset consists of 7 images representing (from top left to bottom right) "Neutral", "Happy", "Sad", "Fear", "Surprise", "Disgust", and "Angry", respectively.

After extracting the first BIMF of the "Neutral" image and applying edge detection and morphological operations, D1 through D4 were measured on the resulting edge map, as shown in figure 3. The measurements are in number of pixels and are used as references for other facial expressions.

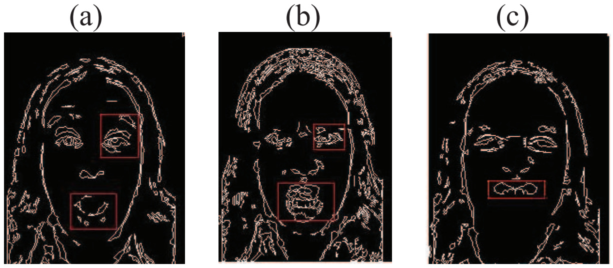
(a) (b) (c)

Fig.3 Neutral facial expression: (a) reconstructed image from BIMFs, (b) the 1st BIMF, (c) extracted edge image

D1-D3 were then measured on the edge maps for the other six expressions as shown in figure 4. D4 is only needed for "Surprise and "Disgust" as shown in figure 4(d) and 4(e).



(a) (b) (c)



(d) (e) (f)

Fig.4 Extracted facial edge images of six emotions: (a) Happy, (b) Sad, (c) Fear, (d) Surprise, (e) Disgust, (f) Angry

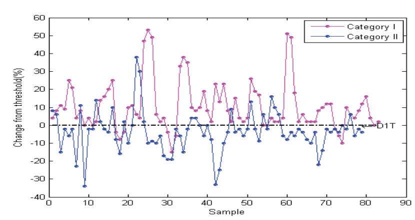
Table 1 shows the decision metrics for facial expressions. In each cell, the value shows the number of pixels measured for each metric, and the percentages are the percent-change of each metric from its corresponding "Neutral" expression. It can be observed from the table that an increase in D1 puts expressions in C-I and a decrease in C-II. In C-I, a significant increase in D3 identifies "Happy," a significant decrease for "Sad," and the rest are "Fear." In C-II, a decrease in D4 recognizes "Angry," and a large change of both D2 and D4 identifies "Surprise" and an insignificant change in D2 and D4 is the expression "Disgust."

Table 1 Criteria for Expression Recognition

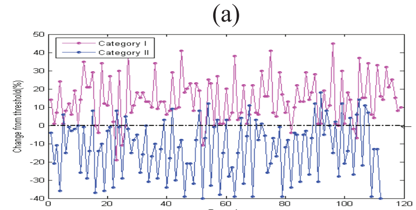
	Happy	Sad	Fear	**Neutral**	Surprise	Angry	Disgust
D1	61 (20%)	60 (18%)	61 (20%)	**51 (0%)**	43 (-16%)	45 (-12%)	46 (-10%)
D2	Not Used for Detection			**24 (0%)**	40 (67%)	36 (50%)	29 (21%)
D3	25 (79%)	11 (-21%)	21 (50%)	**14 (0%)**	Not Used for Detection		
D4	Not Used for Detection			**29 (0%)**	44 (52%)	24 (-17%)	30 (3%)

3.2 Recognition Using Facial Expression Databases

The above described algorithm and criteria were used to determine facial expressions with images from JAFFE, Japanese Female Facial Expressions [7], and TFEID (Taiwanese Facial Expression Image Database) [23]. These databases were chosen to demonstrate the detection accuracy as well as the effects of cultural differences on facial expressions. JAFFE contains 213 images of 7 expressions (6 typical expressions + 1 neutral) posed by 10 Japanese models. Around 233 "rimmed" gray scale images from TFEID were used for this research. In this database, 20 male models and 20 female models were instructed to pose 8 expressions (6 typical expressions+1 neutral+1 contempt). Because "Contempt" is not included in the JAFFE database, we omitted this expression in our research so that better comparisons could be conducted. For each group, we first tested categorical detection rates, and then performed expression detection for each specific facial expression.



(a)



(b)

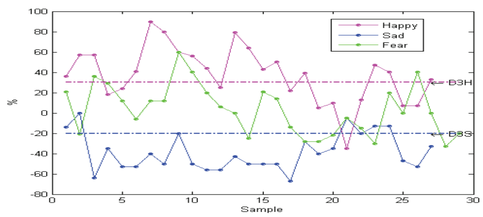
Fig. 5 Categorical detection using D1 (a) JAFFE (b) TFEID

Figure 5 shows categorical detection results using D1. D1T is the threshold of D1 determined by "Neutral" expression. An increase of D1 from D1T decides the expressions in C-I, otherwise, the expression belongs to C-II. The success rates were 89% and 78% for JAFFE, 93% and 79% for TFEID respectively, as shown in table 2.

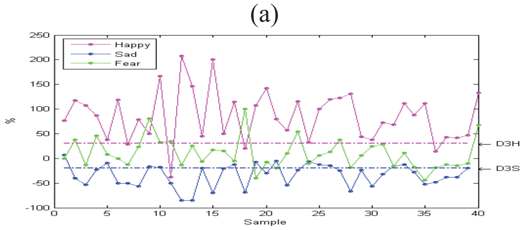
Table 2 Categorical Detection Rates

Database	Category I	Category II
JAFFE	89%	78%
TFEID	93%	79%

Figure 6 shows the results for distinguishing "Happy," "Sad" and "Fear" using D3. D3H and D3S refer to D3 thresholds for "Happy" and "Sad," respectively. Any D3 values fall between D3H and D3S are classified as "Fear."



(a)



(b)

Fig. 6 C-I detection using D3 (a) JAFFE (b) TFEID

In our experiment, it was found that facial expressions of certain emotions of the tested databases are different from the reference set. This is especially true for C-II expressions. Table 3 shows the percentage changes in D2 and D4 on the subjects in databases under test. As can be observed, D2 for "Angry" has a significant negative change from "Neutral." Hence, we used D2 to identify "Angry" first. A combination of D2 and D4 are then used to separate "Surprise" and "Disgust" as discussed in the previous section. Once this adjustment was made, the adjusted metrics were applied in detecting expressions for all subjects in both databases.

Table 3 Changes of D2 and D4 in JAFFE

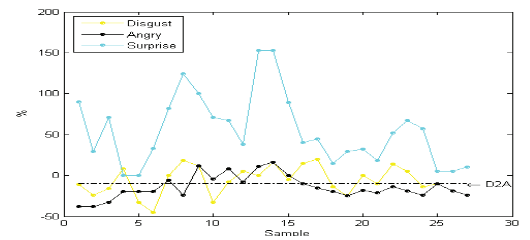
	Neutral	Angry	Disgust	Surprise
D2	**0%**	-14%	-4%	55%
D4	**0%**	-16%	-9%	18%

This variance could be due to the differences between the ways of expressing emotions of different cultural groups. As reported in [24], "…cultural differences occurred on multiple emotion scales for each expression…" Asian groups tend to be more reserved in expressing their feelings through their facial expressions in C-II category than their American counterparts. Instead of opening their mouth when expressing emotions like "Angry" and "Disgust," the Asian subjects tend to close their lips tighter when express category II emotions.

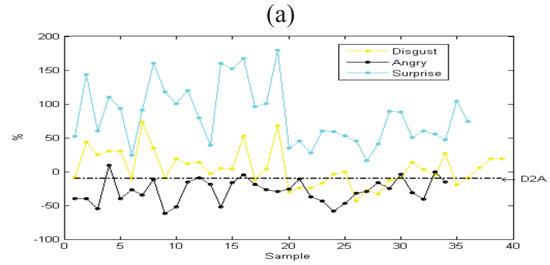
(a) (b) (c) (d)

Fig. 7 Japanese Female Facial Expressions: (a) Neutral, (b) Fear, (c) Disgust (d) Angry

Figure 7 shows an example of original images from the JAFFE database. It can be observed that there are very little differences among measured metrics among these facial expressions, especially in D2 and D4, which are crucial for separating "Disgust" and "Angry."



(a)



(b)

Fig.8 Detection of "Angry" using D2 (a) JAFFE (b) TFEID

Figure 8 shows the results for facial expressions detection in C-II. D2A is the threshold to recognize "Angry" from "Surprise" and "Disgust." As demonstrated from figure 8, "Angry" has the most significant change from "Neutral," which shows a negative change in D2. A -10% threshold was set to recognize "Angry" in both JAFEE and TFEID.

Finally, "Surprise" and "Disgust" were separated using D4. As shown in table 3, "Surprise" shows a decrease in D4 for "surprise" but an increase in D4 for "Disgust." A 5% threshold was set to distinguish these two groups, where a change greater than 5% identifies "Surprise" while a change less than 5% indicates "Disgust."

Table 4 Success Detection Rates

Database	Category I			Category II		
	Happy	Sad	Fear	Disgust	Angry	Surprise
JAFFE	63%	81%	67%	68%	70%	93%
TFEID	90%	76%	73%	51%	82%	97%

Table 4 summarizes the successful detection rates using our method. While detection rates are relatively high for both databases in all categories, one can easily observe that detection of "Disgust" is relative low for both databases.

Despite cultural differences in expressions, further research needs to be done to improve the recognition rates.

4. CONCLUSIONS

In this paper, a computationally efficient method for facial expression recognition was developed. This method uses a fast BEMD algorithm to extract the first BIMF that contains high frequency contents (edge information) of an image. An edge map was created from this BIMF and used for expression recognition. A set of recognition criteria was developed through previous literature and our experiments. A new metric was added to increase recognition success rates. To show the robustness of this method, criteria were established by expressions of an American model, and databases from two different (yet related) cultural groups. Same metrics were applied to the images from both databases. Overall results are encouraging, at the same time, it is suggested by the results that further research needs to be done to further improve the detection rates. One potential solution may be to add yet another metric to strengthen the recognition rates.

5. ACKNOWLEDGMENT

The authors would like to thank Dr. Linda Camras for providing reference dataset used in this research.

5. REFERENCES

[1] G. Donato, M. S. Bartlett, J. C. Hager, P. Ekman, and T. J. Sejnowski, "Classifying Facial Actions." *IEEE Trans. on Pattern Analysis and Machine Intelligence,* vol. 21, no. 10, pp.974-989, 1999.

[2] J. Bassili, "Emotion Recognition: The Role of Facial Movement and Relative Importance of Upper and Lower Areas of Faces." *J. Personality and Social Psychology,* vol. 37, pp. 817-823, 1979.

[3] P. Ekman and E.L. Rosenberg, *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*. New York: Oxford University Press, 1997.

[4] W. Zhao, R. Chellappa, P.J. Philips, and A. Rosenfeld, "Face Recognition: A Literature Survey." *ACM Computing Surveys,* vol.35, no. 4, pp. 399-459, 2003.

[5] M. Pantic and A. Vinciarelli, "Implicit Human Centered Tagging." *IEEE Signal Processing Magazine*, vol. 26, iss. 6, Nov. 2009.

[6] J.G. Daugman, "Complete Discrete 2D Gabor Transform by Neural Networks for Image Analysis and Compression." *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 36, pp. 1,169-1,179, 1988.

[7] M. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba, "Coding Facial Expressions with Gabor Wavelets." *Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition,* pp. 200-205, April 14-14, 1998, Nara, Japan.

[8] M. Lades and J. VorbruÈggen, J. Buhmann, J. Lange, W. Konen, C. Malsburg, and R. WuÈrtz, "Distortion Invariant Object Recognition in the Dynamic Link Architecture." *IEEE Trans. on Computers*, vol. 42, no. 3, pp. 300-311, Mar. 1993.

[9] P. N. Belhumeur, J. P. Hespanha, and D.J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection." *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, July 1997.

[10] M. S. Bartlett and T. J. Sejnowski, "Viewpoint Invariant Face Recognition Using Independent Component Analysis and Attractor Networks." *Advances in Neural Information Processing Systems,* M. Mozer, M. Jordan, and T. Petsche, eds., vol. 9, pp. 817-823, Cambridge, Mass., 1997.

[11] L. Liang, Z. Ping, "An Edge Detection Algorithm of Image Based On Empirical Mode Decomposition." *Second International Symposium on Intelligent Information Technology Application, 2008*, pp.128-132

[12] I. Sobel, *Camera Models and Machine Perception*. PhD thesis. Stanford University, CA, 1970.

[13] J. Prewitt, "Object Enhancement and Extraction." *Picture Process Psychopict*, 1970, pp. 75-149.

[14] J. Canny, "A computational approach to edge detection." *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 8, pp. 679-698, 1986.

[15] D. C. Marr and E. Hildreth, "Theory of Edge Detection." *Proc Roy Soc London*, vol. B275, pp. 187-217, 1980.

[16] V. S. Nalwa and T. O. Binford, "On Detecting Edges." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, pp. 699-714, Nov. 1986.

[17] S. Sarkar and K.L. Boyer, "Optimal Infinite Impulse Response Zero Crossing Based Edge Detectors." *Computer Vision, Graphics, and Image Processing: Image Understanding,* vol. 54, pp. 224-243, Sept. 1991.

[18] A. Walczak, L. Puzio, "Adaptive Edge Detection Method for Images." *Opto-Electronics Review*, 16 No.1 , pp 60–67, 2008.

[19] J. Z. Zhang, Z. Qin, "Edge Detection Using Fast Bidimensional Empirical Mode Decomposition and Mathematical Morphology." *Proceedings of the IEEE Southeast Conference*, pp. 139 – 142, March, 2010.

[20] N.E.Huang, Z.Shen, S.R.Long, et al. "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis." *Proc. Roy. Soc., London*. A, vol.454, pp. 903-995, 1998.

[21] C. Damerval, S. Meignen, and C. Perrier, "A Fast Algorithm for Bidimenstional EMD." *IEEE Signal Processing Leters*, vol. 12, no.10, pp 701-704, 2005.

[22] K. Ghanem, A. Caplier, and M. Kholladi, "Estimation Of Anger, Sadness And Fear Expressions Intensity Based on the Belief Theory." *VISAPP 2008: Proceedings of the Third International Conference on Computer Vision Theory and Applications,* pp.452-460, 2008.

[23] L. F. Chen and Y. S. Yen, "Taiwanese Facial Expression Image Database" [http://bml.ym.edu.tw/~download/html]. Brain Mapping Laboratory, Institute of Brain Science, National Yang-Ming University, Taipei, Taiwan. 2007.

[24] N. Yrizarry, D. Matsumoto, and C. Wilson-Cohn, "American-Japanese Differences in Multiscalar Intensity Ratings of Universal Facial Expressions of Emotion." *Motivation and Emotion,* vol. 22, no.4, pp. 315-327, 1998.