An Enhanced Local Texture Pattern for Effective Face Feature Description

Md. Moniruzzaman, Shaikh Jeeshan Kabeer, and Muhammad Mahbub Alam*

***Abstract*—Feature representation technique for facial images is of paramount importance in fields ranging from biometrics and HCI to digital media, gaming and security. Researchers are focusing on new ways to perform automatic recognition of facial expression primarily because the human face is capable of expressing an amazing range of feelings and emotions. Human expression detection is challenging because of variation of images in intensity, pose, lighting, occlusions etc. In this paper, a new micro pattern based feature representation technique is proposed for facial images, referred to as Angular Response Pattern (ARP). The proposed method applies directional Sobel masks on each pixel in the original image and generates edge responses. Pairs of the response values are used to calculate the angle information and based on this angle values ARP code is generated for each pixel. The image is divided into a number of sum-images and a histogram is generated for each sub image. Each bin of the histogram is updated with the magnitude information of the corresponding pixel. The feature vector is generated by concatenating all the histogram. The final histogram is then further smoothed by applying a square root operation. The proposed ARP method has been applied on publicly available benchmarked image databases for performance evaluation. Performance analysis with state-of-the-art approaches shows that the proposed method outperforms them all. Thus our approach exhibits robustness against illumination variations and random noise. This approach can further be applied towards applications such as person identification, age estimation, gender classification.**

***Keywords*—Local texture pattern, expression recognition, feature descriptor, angular response pattern, SVM.**

* Corresponding author.

Md. Moniruzzaman and Muhammad Mahbub Alam are with the CSE department of Islamic University of Technology (IUT), Gazipur, Bangladesh. e-mail: milton@iut-dhaka.edu and mma@iut-dhaka.edu.

Shaikh Jeeshan Kabeer was with CSE department of IUT and currently doing his masters at University of Calgary.

Manuscript received October 05, 2015; revised December 12, 2015.

I. Introduction

The availability of increased computing abilities have propelled different research directions of image processing to flourish, leading to greater and wider areas of applications. The advent of these new applications in the domain requires better and efficient methods of image processing. One of the key components of most image processing based recognition system is the feature representation techniques. Feature representation techniques define the efficiency of the system as well as the resilience of the system against various challenges. Among the different directions of image processing, human facial expression analysis has received quite a lot of attention in recent times due to the inherent significance in fields ranging from biometrics and HCI to digital media, gaming and security [1]–[3]. Researchers are focusing on new ways to perform automatic recognition of facial expression primarily because the human face is capable of expressing an amazing range of feelings and emotions [4]. It is a subconscious process which occupies a major portion of our conveyed messages and at the same time gives way to out likes, dislikes while also providing vital feedback cues [5]. Thus the proper identification of facial expression can go a long way towards achieving intelligent and effective communication with computers.

Human expression detection is challenging because facial expression vary with each individual. At the same time real world scenarios introduces variations in intensity, pose, lighting, occlusions make this task more challenging [6]. An effective feature representation technique is essential which will help not only towards proper expression detection but also help in other areas such face detection and gender classification [7], [8]. Over the years researchers have applied a wide variety of feature representation for facial images. All feature representation techniques perform some data transformation for an enhanced image representation. A popular approach is to use holistic information, where features are generated based on the entire intensity value distribution. One such example is [9] which employs Principal Component Analysis (PCA) PCA that transforms intensity values to an orthogonal eigenspace, where the dimensionality is reduced with respect to the first and second moments. Independent Component Analysis (ICA) is higher order generalization of PCA employing higher moments, [10] is an application of ICA in this domain. [11] uses Linear Discriminant Analysis (LDA), which performs intensity value transformation using between class and within class variance.

However many researchers use local feature representation instead of holistic ones. One popular set of approaches is based on the transformation of intensity values to gradient and orientation data. SIFT [12] is a very poplar algorithm which falls under this category. SIFT computes histograms of locally oriented gradients around certain points and represent the information in 128D space. [13] introduced Histogram of oriented objects (HOG) which works by partitioning an image into smaller segments, calculating the HOG for each segment and finally generating the features by concatenating the HOG for the smaller blocks. Another popular approach is Gabor wavelets [14] which captures frequency and orientation information from facial images using frequency decomposition, which are then used for feature construction.

Of the many local feature representation techniques micro pattern based techniques is the most popular choice due to the inherent simplicity and relative easiness of implementation. Micro pattern based features are found by utilizing the surrounding pixels information with respect to the center pixel [15]. Local Binary Pattern (LBP) [16] is by far the most popular micro pattern based feature encoding scheme which uses two levels encoding, but LBP is sensitive to noise and cannot effectively distinguish regions with similar intensity values. Another very similar approach is Local Ternary Pattern (LTP) [7] which introduces thresholding function and uses three level encoding. Although LTP is more robust compared it still suffers from sensitivity towards similar intensity values along the threshold boundary and the usage of three level of encoding increases the feature size by a great margin. Local Directional Pattern (LDP) [8] is another approach which introduces direction information into the encoding process. Although better then LBP and LTP, LDP encodes a limited number of directions into the encoding process and the lack of a threshold function means that LDP suffers from some sensitivity towards similar response values of the neighbors. Directional Ternary Pattern (DTP) [6] is a three level encoding mechanism that also uses a threshold function which enables it to perform better compared to the other micro pattern techniques mentioned earlier. However DTP still suffers from the sensitivity issues of similar response values and the direction information that is encoded gives a rough estimation.

To overcome the limitations of the processes mentioned above this paper proposes a new facial feature representation method, referred to as Angular Response Pattern (ARP), which has been developed considering the weaknesses of existing feature representation techniques. The proposed approach overcomes the shortcomings of the existing methods by encoding direction of transition values and magnitude information into the feature encoding process. The proposed approach has been applied on a publicly available image database and performance analysis shows better results compared to existing feature encoding techniques under different experimental setup.

This paper is broken down into the following chapters. Chapter II provides a literature review which covers the related micro-pattern based facial feature representation techniques. Chapter III describes the proposed algorithm in detail focusing on the encoding and representation schemes. Chapter IV presents experimental analysis which highlights the performance of ARP in comparison to other popular and related approaches. The final segment is chapter V that concludes the paper with a summary of the work and a direction on the possible improvements of our proposed approach.

II. Literature Review

In this section some popular micro pattern based facial feature representation methods are reviewed. One of the early local pattern based feature description technique is Local Binary Pattern (LBP) [16] which operate by defining a local neighborhood around each center pixel. The neighbors of each pixel are encoded as 0 or 1 depending on the intensity value of the center pixel; thus the center pixel value acts as threshold. The resulting binary bit pattern around each pixel is assigned to the center pixel which in turn is converted to an equivalent value and added to the histogram bin which is the LBP representation of the image. An extension to the LBP method called Uniform LBP (ULBP), works on repeating patterns that appear throughout the image with little bitwise changes. Based on the analysis of patterns, [17] found that at most two bitwise transitions are capable of representing significant local features. Thus these patterns are identified and included in the feature representation process. The basic problem with LBP based techniques is that the binary pattern of each pixel is defined by the comparative magnitude of the center pixel. A small variation in pixel intensity values would cause the calculation for the center pixel to falsely represent some edge information in the smooth region.

Another popular micro pattern facial feature description technique is Local Ternary Pattern (LTP) which is capable of overcoming the drawbacks of LBP. LTP [7] is similar to LBP except for the fact that is uses three level encoding to capture more intensity variation information. LTP also uses a threshold value t which is used in the encoding process. Intensity values of neighbor pixels which are within the threshold bound $\pm t$ are encoded as 0. For intensities above $+t$ or below $-t$ the values are quantized as $+1$ and -1 respectively. Thus a binary bit pattern is generated for every pixel in the original image. Compared to LBP, LTP is more robust to small variations in intensity values due to the introduction of a threshold value . Although the introduction of a three level encoding scheme allows more intensity variation to be captured it also increases the feature size which in turn also increases classification complexity. At the same time sensitivity to intensity values along the threshold boundary still remain.

Another approach is Local Directional Pattern (LDP) [8] which calculates edge responses using Kirsch masks. Based on the response values are quantized to binary levels 0 and 1. Based on the highest k magnitude of edge responses, values are encoded as 1 in k locations and 0 for the remaining directions. LDP is capable of capturing variations of intensity in at most three directions. Based on the relative position of the binary codes, binary bit pattern is generated for each image pixel, which is later used for generating histograms of the image. The eight directional Kirsch masks generate high response or similar responses in respective directions based on whether the image texture contains edges or smooth regions. Although the directions of intensity variations are encoded, only direction transitions may be captured. Moreover the absence of a threshold value along with the usage of binary encoding means that similar intensity values such as in smooth regions, with minor variations in intensity may represent a large variation in the binary coded image.

Directional Ternary Pattern (DTP) [6] overcomes the problems of LDP. Similar to LDP, DTP works on the edge response values rather than the grey intensities, since response values provide more stable [8] and effective feature description information. DTP applies Robinson mask on the original image and then calculates the average μ of the edge responses. If responses are within the boundary $\pm t$, where t is the defined threshold, then the encoding is 0 in the corresponding neighbor position. If the response is above $\mu + t$ or below $\mu - t$ then the corresponding value is encoded as $+1$ or -1 respectively. The use of three level encoding mechanism allows DTP to capture more response variations. To reduce the size of the feature representation, DTP code in split into two parts. In addition the introduction of a threshold value allows DTP to be more robust to noise and small variations in intensity compared to LDP, LBP and other similar approaches. Although the performance of DTP is better compared to the earlier mentioned approaches, it is still sensitive to similar types of response values in the neighborhood positions and the use of three level encoding increases the feature size.

All the approaches described above are sensitive to response values in similar range. For those which use a threshold function the sensitivity arises for values which are close to the threshold boundary. Secondly for approaches like LDP and DTP the encoding process can capture the general direction of changes in the response values, however that information is not directly encoded into the feature to provide any real sense in direction. To overcome these challenges, in this paper a new feature representation technique called Angular Response Pattern (ARP) is introduced. The proposed approach overcomes the problems identified above by introducing angle information into the feature pattern, which is calculated using response values. Using pairs of response values, the angle of the resultant response is evaluated and the information is used in the pattern representation. A histogram representation is then created which is updated using the magnitude of the resultant response values. In the final step a square root function is applied on the histogram data to perform smoothing which helps to eliminate the effects of noise. The introduction of angle and magnitude information in the feature representation ensures better performance compared to other related approaches mentioned above, as shown in section IV. Moreover the encoding process is made such that the resultant histogram bin size is kept low; as a result the feature representation using ARP is smaller compared to any of the earlier mentioned methods above and also other feature representation techniques. The comparative smaller size of representation will mean that applications such as classification can be done much more effectively and efficiently. The following chapter contains a detailed account of the proposed approach.

III. Proposed Approach

The proposed approach overcomes the problems identified above by introducing angle information into the feature pattern, which is calculated using response values. Using pairs of response values, the angle of the resultant response is evaluated and the information is used in the pattern representation. As earlier work have shown better performance by including angle information in the feature encoding process, we applied a similar technique which was applied on the edge responses rather than grey intensity values with the motivation that the performance would be improved. Our motivation bore fruit as performance analysis shows that our proposed approach achieved better classification accuracy compared to other related approaches. A histogram representation of the image is then created which is updated using the magnitude of the resultant response values. In the final step a square root function is applied on the histogram data to perform smoothing which helps to eliminate the effects of noise. The introduction of angle and magnitude information in the feature representation ensures better performance compared to other related approaches mentioned before. Moreover, the encoding process is made in a way to keep the resultant histogram bin size low; and as a result the feature representation using ARP is smaller compared to any of the earlier mentioned methods above and also other feature representation techniques. The comparative smaller size means that classification can be done much more effectively and efficiently.

A. Feature Representation

For each pixel in the original image a neighborhood is defined containing P neighbors with radius R. The radius R defines the neighborhood over which the calculations will be done for encoding. For $R = 1$ the pixels which are adjacent to the center pixel in eight directions will be considered, for $R = 2$ pixels which are 1 pixel apart from the center pixel C are considered. In this manner changing the value of R will change the neighborhood under consideration. Using the grey intensity values, a response matrix is calculated using Sobel Mask. Sobel masks have mirrored values with opposite sign, thus encoding only four response values will be sufficient to represent feature vector. It helps reducing feature size of ARP. As shown in Figure 1, the masks are rotated 45 degrees clockwise to produce different orientations which will generate eight different responses for the center pixel when applied on the original image.

$$\underset{N}{\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}} \quad \underset{E}{\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}} \quad \underset{S}{\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}} \quad \underset{W}{\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}}$$

$$\underset{NW}{\begin{bmatrix} -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}} \quad \underset{SW}{\begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{bmatrix}} \quad \underset{SE}{\begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -2 \end{bmatrix}} \quad \underset{NE}{\begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix}}$$

Fig. 1: Eight different orientation of Sobel mask for calculating eight directional edge responses.

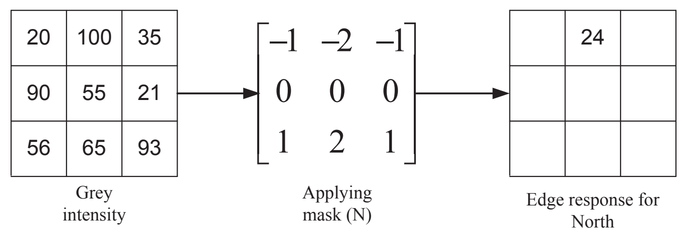


Fig. 2: Edge response calculation in East direction.

The value of the response for each center pixel is defined by the original grey level intensity. High intensity variation within neighborhood correspond to higher response values and vice versa, thus a transition from a relative high to low or the reciprocal indicates the presence of an edge. Figure 2 shows the calculation of two edge responses in North and East direction, respectively.

The response values obtained after applying the Sobel mask is denoted as r_i $(i = 0, 1, \ldots, 7)$ and are used for further processing. Here r_0, r_1, r_2 and r_3 represents response values of NW, N, NE and E

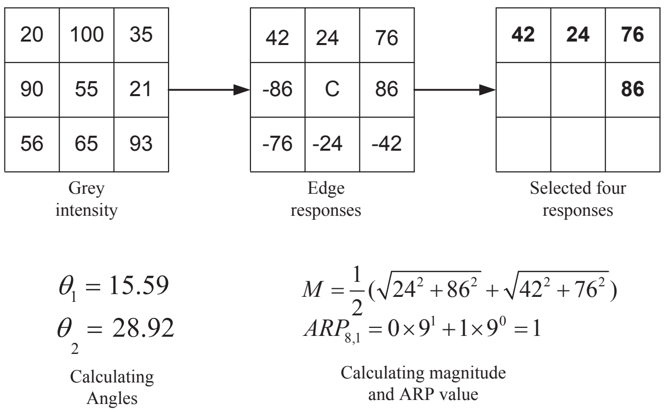


Fig. 3: Sample ARP value calculation.

respectively and r_4, r_5, r_6 and r_7 represents response values of SE, S, SW and W respectively. As we can observe that, the values r_0, r_1, r_2, r_3 and r_4, r_5, r_6, r_7 are mirror opposites. Thus, this approach only uses one set of these values for further calculations, i.e., r_0, r_1, r_2, r_3 . Using r_0 and r_2 an angle $A1_{P,R}$ will be evaluated using Eq. 1, which is the angle of the resultant response between r_0 and r_2. Similarly, r_1 and r_3 will be used to calculate the angle of the resultant response between r_1 and r_3, using Eq. 2.

$$A1_{P,R} = \left\lfloor \frac{S}{180} tan^{-1}\Big(\frac{r_0}{r_2}\Big) \right\rfloor \quad (1)$$

$$A2_{P,R} = \left\lfloor \frac{S}{180} tan^{-1}\Big(\frac{r_1}{r_3}\Big) \right\rfloor \quad (2)$$

The range of the angles between the pairs of responses mentioned in the equations is between 0 and 180 degrees, which is divided into S segments. Considering an $M \times N$ image, $A1_{P,R}$ and $A2_{P,R}$ will be calculated for each pixel. Using these values ARP pattern will be generated for each pixel according to Eq. 3 where $i \in M$ and $j \in N$ and S is the total number of segments.

$$ARP_{P,R}(i,j) = A1_{P,R}(i,j) \times S + A2_{P,R}(i,j) \quad (3)$$

Figure 3 shows an example of ARP code calculation for a pixel where $P = 8$ and $R = 1$. At first eight responses are calculated, then using four of the response values two angles and magnitudes are computed. Finally, by combining the values corresponding to both of the angle values one ARP code is generated.

The ARP patterns generated for each pixel (i, j) is used to represent the original image as an ARP image in which angle information are encoded. We also have constructed an array containing magnitude values obtained from the edge responses, which will be used while generating feature vector. Feature vector size of ARP is 81 which is less than many other well known feature descriptor. This reduced feature size makes ARP more feasible for practical implementation. Table 1 shows the feature size of various feature descriptor.

TABLE 1: Feature size of various feature descriptor.

Descriptor	Feature Size
LBP	256
LTP	6561
DTP	6561
cDTP	81
Sobel LBP	512
ARP	81

B. Feature Vector

After the original cropped image has been converted to an ARP image, it is partitioned into 3×3, 5×5 and 10×10 smaller sub images. For each ARP sub images histograms are generated using Eq. 4 and Eq. 5 where H_k represents the $k-th$ bin of the histogram.

$$H_k = \sum_{i=1}^{M}\sum_{j=1}^{N} f\Big(ARP_{P,R}(i,j), k\Big) \qquad [k \in K] \quad (4)$$

$$f\Big(ARP_{P,R}(i,j), x\Big) = \begin{cases} M_{P,R}(i,j), & ARP_{P,R}(i,j) = x \\ 0, & \text{Otherwise} \end{cases} \quad (5)$$

While generating the histogram, the $k-th$ bin is updated with the magnitude value of the pixel which is calculated using Eq. 5 where $M_{P,R}(i,j)$ is the magnitude of the resultant response values of image pixel that is calculated using the following equation:

$$M_{P,R} = \left\| \frac{1}{2}\Big(\sqrt{r_1^2 + r_3^2} + \sqrt{r_0^2 + r_2^2}\right\| \quad (6)$$

The histograms obtained for each of the ARP sub-images are merged together to generate the complete histogram of the original image. Figure 4 summarizes the feature vector generation where the ARP image is partitioned into 10×10 sub images. The complete histogram generated from ARP encoding now represents

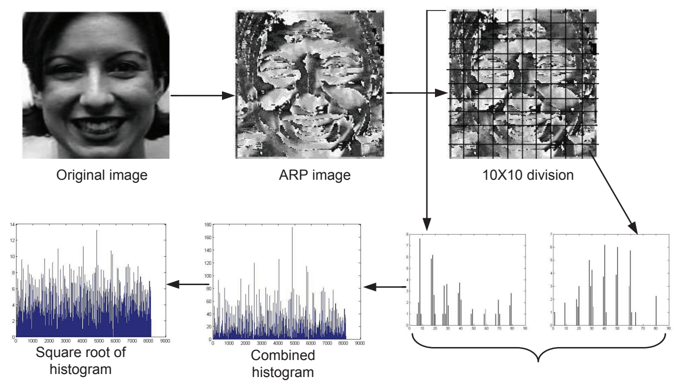


Fig. 4: Summary of ARP Feature vector generation.

Fig. 5: Sample Images from Cohn-Kanade Image Database.

the image with respect to the features, i.e. angle and magnitude of the resultant response values.

C. Histogram Smoothing

The final step of the proposed approach performs a smoothing operation on the histogram to reduce the contribution of smooth regions in the feature representation and also to alleviate the effects of noise in the original image. The smoothing is done by applying a square root function on which was calculated earlier. The final smoothed bin values are stored in HF_k, which is calculated as Eq. 7.

$$HF_k = \sqrt{H_k} \tag{7}$$

The smoothed histogram is then used to evaluate the performance of ARP feature representation procedure by evaluating the features using different popular classifiers.

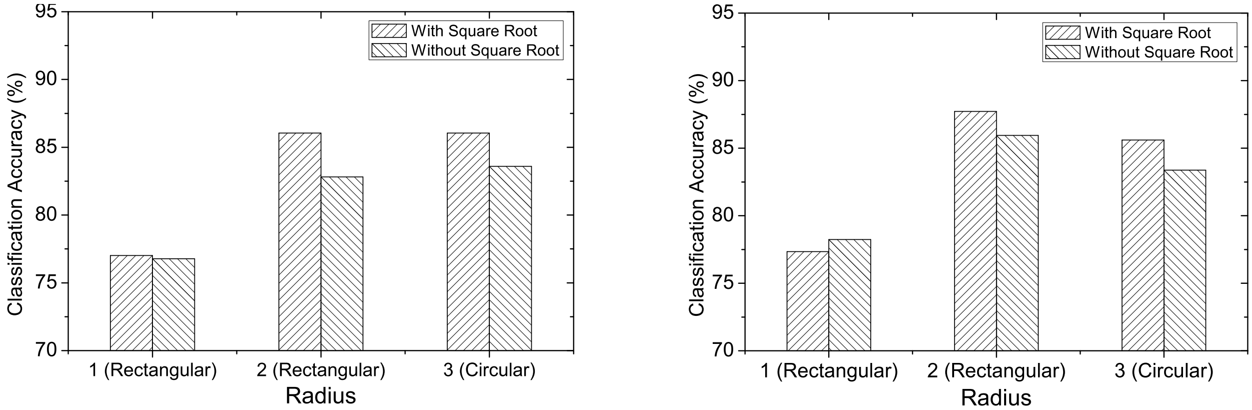
IV. Experimental Results and Performance Analysis

We used Matlab for generating feature vector and Weka Classification tool was used to evaluate the encoding process against different classifiers. The proposed ARP system was applied on Cohn-Kanade (CK) facial image database. The database consists of images of 97 persons. The database contains sequence of images, where each sequence starts from normal expression to the peak of a particular expression. Among all of the images 896 are is selected for this study consisting of six different expressions i.e. happy, sad, anger, disgust, surprise and fear. For further experiment 104 neutral expression images are added to check performance of our proposed method against other methods. The images were of different variations with respect to the orientation, illumination, pose and age factors. Using the two eye positions as ground truth, the images were cropped to a reduced dimension of 150×150 pixels. Figure 5 shows some sample images from this database.

We run our test by adding different types of Gaussian white noises with sample images. We also used different commonly used classifier (Adaboost, Linear SVM, Bagging, SMO, Random forest) for evaluating performance of our method.

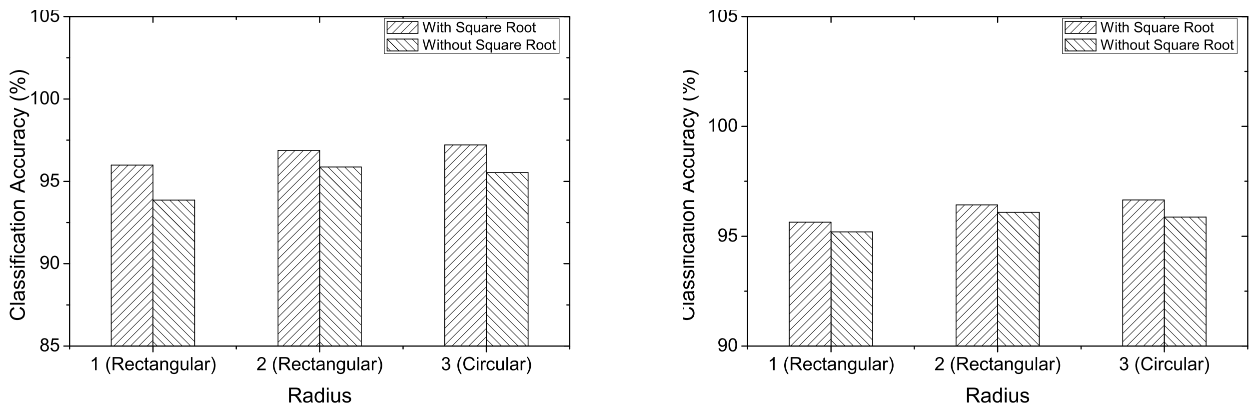
Figures 6 through 8 shows the performance of the proposed ARP encoding procedure on different experimental setup. We have divided original image into 3×3, 5×5 and 10×10 sub images. For each of the case we have tested performance of ARP for different angle segments values of $S = 9$ and $S = 12$. ARP was run on different neighborhood configurations of $R = 1(R1)$, rectangular $R = 2(R2)$ and circular $R = 2(R2')$. For circular $R2'$ the corner pixel values are calculated using the average of the corner pixels of $R1$ and $R2$ neighborhood definitions. The classification accuracy was generated using SVM classifier of Weka with ten-fold cross validation. The figures show the performance of ARP with (With SQ) and without (Without SQ) square root smoothing respectively. As denoted by the figures in almost all the cases the performance is better with smoothing and thus reinforcing our claim that smoothing alleviates noise and improves performance.

Figure 6 shows performance for a 3×3 sub-partitioned image. We have tested it for both $S = 9$



(a) $s = 9$ (b) $s = 12$

Fig. 6: Performance of ARP for dividing original image into 9 sub images.



(a) $s = 9$ (b) $s = 12$

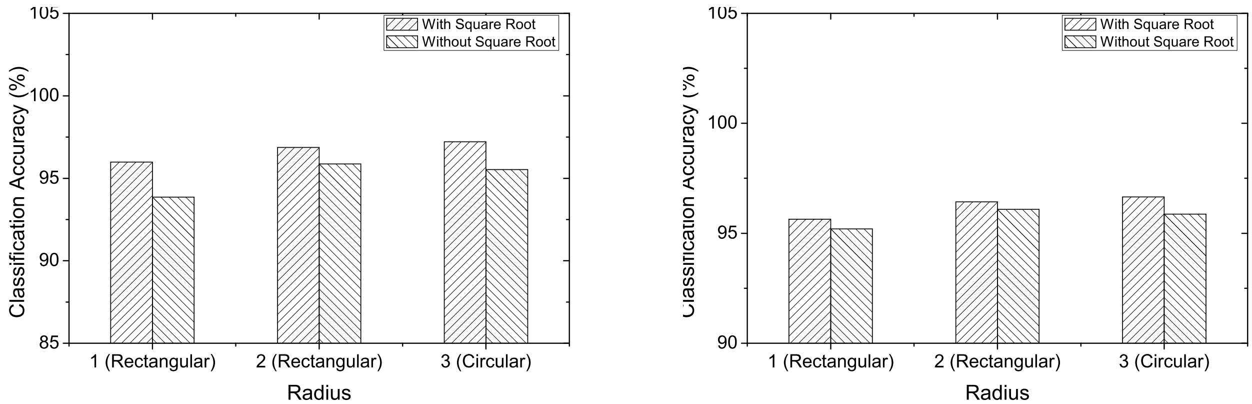
Fig. 7: Performance of ARP for dividing original image into 25 sub images.

and $S = 12$ for three different radius values. For most of the cases ARP shows better performance when histogram smoothing is done by performing square root. Here best accuracy (87.72%) is obtained for $S = 9$ and $R = 2$(Rectangular) with square root of histogram.

Figure 7 shows performance for a 5×5 sub-partitioned image. Performance for three different radius values are calculated for both $S = 9$ and $S = 12$. In this case ARP also produces better result with histogram smoothing. Here best accuracy (97.21%) is obtained for $S = 9$ and $R = 2$(Circular) with square root of histogram.

Figure 8 shows performance for a 10×10 sub-partitioned image. Here we also find better results for histogram smoothing and dividing angle into 9 equal parts. Best accuracy is obtained here for $S = 9$ and $R = 2$(Rectangular) with square root of histogram which is (97.99%). Dividing the image in more sub division results better result because it encodes more local information. But if the total number of sub division is less, then a large portion of the image is grouped together and considered as a single unit, thus losing some local information. However selecting best radius depends on the resolution of the image. For our test the the resolution of the image was 150×150 and radius 2 shows the best performance.

For highlighting the robustness of the proposed approach and at the same time show that ARP is not classifier dependent, the proposed approach has been evaluated with different classifiers as shown by Table 2. ARP was run with $R = 2$ for both $S = 9$ and $S = 12$ while considering the effects of both including and excluding the magnitude information from the histogram. All the results are consistently better when magnitudes are considered and thus this result backs



(a) $s = 9$ (b) $s = 12$

Fig. 8: Performance of ARP for dividing original image into 100 sub images.

TABLE 2: Performance of ARP with various classifier.

Classifier	$S = 9$		$S = 12$	
	X	Y	X	Y
SVM	95.64	97.99	95.09	97.54
Adaboost	69.62	77.00	71.87	76.63
SMO	86.94	92.07	88.50	91.41
Bagging	66.41	70.65	64.62	69.53
Random Forest	72.43	77.68	64.17	71.20

X = Without magnitude Y = With magnitude

TABLE 3: Confusion matrix for ARP.

Expression	F	Sd	D	A	H	Sr
Fear	98.65	0	0	0	0.675	0.675
Sad	0	94.81	0	5.19	0	0
Disgust	0	2.22	97.78	0	0	0
Angry	2.14	0	0	95.72	0	2.14
Happy	0	0	0	0	100	0
Surprise	0	0	0	0	0	100

F = Fear, Sd = Sad, D = Disgust, A = Anger, H = Happy and Sr = Surprise

our claim that including magnitude information into the histogram improves the classification accuracy.

Table 3 shows the confusion matrix for ARP. This test is conducted for 6 expression data set. Here $S = 9$ and $R = 2$. Original is divided into 10×10 sub images. From the result we observe that, our descriptor generates discriminative feature vector happy and surprise. For both of these cases accuracy was 100%. For few images sad expression is misclassified as angry. Fear expression is misclassified ad happy and surprise.

TABLE 4: Performance comparison ARP for 8 neighbors in radius 1.

Feature descriptor	3×3	5×5	10×10
cDTP [6]	75.67	95.98	96.65
Sobel LBP [18]	79.80	90.51	93.30
LBPV [19]	69.42	85.38	88.95
LGP [2]	81.25	93.30	81.25
ARP(Proposed)	77.01	95.98	97.09

For comparative analysis the proposed ARP approach has been compared with other popular feature representation techniques which have been applied on the same image database. For classification SVM classifier with linear kernel from Weka was used with ten-fold cross validation.

Table 4 shows the performance of ARP in comparison to cDTP, Sobel LBP, LBPV and LGP for R1 neighborhood definitions. ARP was executed for with square root smoothing and angle was divided into 9 equal parts. The experiment was run for different sizes of image sub-partitions and it is observed that for most of the cases ARP performs better compared to other approaches.

From the result presented in the table 5, it can be observed that ARP shows better performance compared to other approaces with classification accuracy of 97.99% for 10×10 sub-partitions of original images.

We have further investigated performance of our approach against other approaches by adding normal image along with six expression images. The result shows that ARP performs better for this seven expression datasets as well. Table 6 summarizes the result.

TABLE 5: Performance comparison ARP for 8 neighbors in radius 2.

Feature descriptor	3×3	5×5	10×10
cDTP [6]	81.36	96.09	97.32
Sobel LBP [18]	84.49	92.52	95.53
LBPV [19]	75.78	86.38	88.84
LGP [2]	90.29	96.31	95.21
ARP(Proposed)	86.05	96.87	97.99

TABLE 6: Performance analysis of ARP for seven expressions.

Feature descriptor	3×3	5×5	10×10
cDTP [6]	90.9	96.2	96.2
Sobel LBP [18]	84.0	91.3	94.8
LBPV [19]	70.8	84.9	87.7
LGP [2]	79.5	92.5	94.6
ARP(Proposed)	86.6	96.5	96.9

Real life images often come with random noises. We have taken 896 images with six expression from CK database and added Gaussian white noises with them and compared the results of ARP with other methods.

Table 7, table 8 and tabel 9 shows comparative performance analysis of ARP with LGP and cDTP with presence of Gaussian noise on original images. We have compared performance of ARP for three different scenarios. We divided original image into 9, 25 and 100 sub images and tested the performance of ARP for all of the cases. We have added Gaussian white noise of different variances with original images. Table 7 shows the performance after adding noise with mean 0 anv variance 0.005. Table 8 shows performance for same images sets with noise variance 0.003 and Table 9 shows performance for same images sets with noise variance 0.001. From the results it is obvious that ARP is more resistant to noise than other feature descriptors.

TABLE 7: Performance analysis of ARP with presence of Gaussian noise ($V = 0.005$).

Feature descriptor	3×3	5×5	10×10
cDTP [6]	57.03	72.01	77.01
LGP [2]	64.06	82.59	89.95
ARP	64.73	88.39	91.85

TABLE 8: Performance analysis of ARP with presence of Gaussian noise ($V = 0.003$).

Feature descriptor	3×3	5×5	10×10
cDTP [6]	75.88	89.51	91.74
LGP [2]	63.01	78.91	79.91
ARP	70.31	90.95	93.53

TABLE 9: Performance analysis of ARP with presence of Gaussian noise ($V = 0.001$).

Feature descriptor	3×3	5×5	10×10
cDTP [6]	69.87	86.27	86.83
LGP [2]	83.03	94.08	95.98
ARP	77.12	95.13	96.42

V. Conclusion

This paper proposes a local feature based image description technique that uses the angle and magnitude information of response values. Results shows that the proposed approach, titled ARP performs better compared to the existing and popular local pattern based feature representation techniques, achieving the highest accuracy on the image database. The inclusion of angle information into the pattern representation enables ARP to capture transitions in response values along with the directions in which the transitions take place. As transition indicates significant features such as edges in almost all kinds of images, ARP is capable of representing these changes more efficiently and thus outperforms other methods mentioned earlier. ARP can be further extended for other image classification tasks such as gender classification among others. For further improvement added mechanisms can be incorporated with ARP which can explicitly identify and discard smooth regions, is expected to make the proposed approach a more robust and efficient local texture based image representation technique.

References

[1] Z. Zhang and Z. Zhang, “Feature-based facial expression recognition: Sensitivity analysis and experiments with a multilayer perceptron,” *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 13, pp. 893–911, 1999.

[2] B. Jun, I. Choi, and D. Kim, “Local transform features and hybridization for accurate face and human detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1423–1436, 2013.

[3] A. Butalia, M. Ingle, and P. Kulkarni., “Facial expression recognition for security,” *IJMER*, vol. 2, no. 4, pp. 1449–1453, Aug. 2012.

[4] C. Darwin, *The expression of the emotions in man and animals / by Charles Darwin.* New York ;D. Appleton and Co.,, 1916.

[5] A. Mehrabin, *Communication without words.* Psychology Today, 1968.

[6] F. Ahmed, "Effective facial feature representation based on directional micro-patterns," *M.Sc. Thesis, Islamic University of Technology, Bangladesh*, 2012.

[7] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *Trans. Img. Proc.*, vol. 19, no. 6, pp. 1635–1650, Jun. 2010.

[8] T. Jabid, M. H. Kabir, and O. Chae, "Robust facial expression recognition based on local directional pattern," *ETRI Journal*, vol. 32, no. 5, pp. 784–794, 2010.

[9] M. Turk and A. Pentland, "Face recognition using eigenfaces," in *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR '91., IEEE Computer Society Conference on*, Jun 1991, pp. 586–591.

[10] P. Comon, "Independent component analysis, a new concept?" *Signal Process.*, vol. 36, no. 3, pp. 287–314, Apr. 1994.

[11] P. N. Belhumeur, J. a. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.

[12] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[13] M. Dahmane and J. Meunier, "Emotion recognition using dynamic grid-based hog features." in *FG*. IEEE, 2011, pp. 884–888.

[14] J. Ou, X.-B. Bai, Y. Pei, L. Ma, and W. Liu, "Automatic facial expression recognition using gabor filter and expression analysis," in *Computer Modeling and Simulation, 2010. ICCMS '10. Second International Conference on*, vol. 2, Jan 2010, pp. 215–218.

[15] M. Pietikäinen, G. Zhao, A. Hadid, and T. Ahonen, *Computer Vision Using Local Binary Patterns*, ser. Computational Imaging and Vision. Springer, 2011, no. 40.

[16] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," *Image Vision Comput.*, vol. 27, no. 6, pp. 803–816, May 2009.

[17] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.

[18] S. Zhao, Y. Gao, and B. Zhang, "Sobel-lbp," in *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, Oct 2008, pp. 2144–2147.

[19] Z. Guo, L. Zhang, and D. Zhang, "Rotation invariant texture classification using lbp variance (lbpv) with global matching," *Pattern Recogn.*, vol. 43, no. 3, pp. 706–719, Mar. 2010.

Md. Moniruzzaman received his B.Sc degree in Computer Science and Information Technology and M.Sc degree in Computer Science and Engineering from Islamic University of Technology (IUT) , Bangladesh in 2009 and 2015, respectively. He has been working as a faculty member in the Computer Science and Engineering (CSE) department in Islamic University of Technology (IUT), Dhaka, Bangladesh. Currently, he is working as a Lecturer in the department.

Shaikh Jeeshan Kabeer received his B.Sc degree in Computer Science and Information Technology from Islamic University of Technology (IUT), Bangladesh in 2011. He worked as a faculty member in the Computer Science and Engineering (CSE) department in Islamic University of Technology (IUT), Dhaka, Bangladesh. Currently, he is working toward his masters degree at University of Calgary.

Muhammad Mahbub Alam received his B.S. degree in Applied Physics and Electronics and M.S. degree in Computer Science from the University of Dhaka, Bangladesh in 1998 and 2000, respectively. He received his PhD from Kyung Hee University, South Korea in February 2009. Since 2002, he has been working as a faculty member in the Computer Science and Engineering (CSE) department in Islamic University of Technology (IUT), Dhaka, Bangladesh. Currently, he is working as a Professor in the department. His research interests include wireless networks, and performance analysis of networking systems.