A Monogenic Local Gabor Binary Pattern for Facial Expression Recognition

Abubakar M. Ashir, Alaa Eleyan

Mechatronic Engineering Department, Mevlana University Konya, Turkey
{aashir, aeleyan}@mevlana.edu.tr

Abstract

The paper implements a monogenic-Local Binary Pattern (mono-LBP) algorithm on a local Gabor Pattern (LGP). The algorithm is applied at different scales of the Gabor kernel with different normalization schemes. Results from the two best performing normalization algorithms with mono-LBP are fused at score level to obtain an improved performance. Moreover, performance comparison is done with other variants of LGP algorithm and also the effects of various normalization techniques are investigated. Experimental results on JAFFE facial expression database show that the new technique has the best average performance of 93% compared to its counterparts using distance metrics as a classifier.

Keywords— Facial Expression Recognition, Local Gabor Patterns, monogenic Local Binary Patterns, Local Binary pattern.

1. Introduction

Facial Expression Recognition (FER) has been one of the areas which have drawn a lot of interests and attentions of the researchers in the field of computer vision and pattern recognition. This may not be unconnected to the need for human-machine interaction (HMI), surveillance systems, robotics applications and many others. Quiet a handful number of feature extraction and classifier algorithms have been proposed and implemented for this task (FER). Gabor kernel has being one of the most robust feature extraction algorithm and widely exploited in FER and face recognition due to its ability to approximate receptive fields of simple cells in the primary visual cortex of human eyes [1] , multi-resolution approach and direction selectivity.

Following successful implementation of Gabor kernels in iris recognition by Doughman [2] and coupled with the success of local binary pattern (LBP) algorithm, a number of variants of Gabor algorithms emerged over time. These Gabor variants are sometimes referred to Local Gabor Patterns (LGP). LGP algorithms exploited various Gabor feature channels such as magnitude, phase, imaginary and real channels. For instance, [3] proposes Local Gabor Binary Patterns (LGBP) which encodes Gabor magnitude with LBP operator at different resolution and orientations to form the feature vector. The proposed LGBP was reported to have improved performance for face recognition [4]. Proposed Local Gabor Phase pattern (LGPP) variants and applied it for face recognition. LGPP essentially encodes both real and imaginary parts of the Gabor features using Douglas method and then the result is further encoded using what is called Local XOR Pattern (LXP). In search for robustness and improved performance, other LGP were proposed such as Histogram of Gabor Phase Pattern (HGPP), Local Gabor Phase difference Pattern (LGPDP) and a host of others which are quite relevant to specific problems. In general, these LGP algorithms come with additional cost of computation, extensive memory usage and in most cases feature vector dimensionality reduction becomes necessary.

In our work, we borrowed a rotation invariant monogenic LBP proposed in [4] for texture classification. Instead of encoding the Gabor magnitude channels with LBP as is the case in (LGBP), we encoded these channels with monogenic LBP which, within the context of this work, we referred to mono-LGBP. Furthermore, the results are computed at different resolution (scales) of the Gabor kernel under different normalization algorithms. At each scale, results of the proposed method with the best two performing normalization technique are fused at the score level to obtain the overall performance of the method.

The paper is divided into five sections. Section I covers the introduction while section II briefly discusses Gabor kernel, LBP and M-LBP. Section III introduces the normalization schemes used and section IV presents the experimental results. Section V summarizes the findings.

2. LGP Feature Extraction Operators

A brief literature background about the Gabor kernels and some of the operators used for LBP algorithms are discussed below.

2.1. Gabor Wavelet Transform

Gabor filter is basically a modulation of a Gaussian function with a sinusoidal plane wave. Therefore the result of convolution of Gabor kernel, $\psi_{\mu,v}(z)$ with an image, *I(z)* is represented as $G_{\mu,v}(z)$ in Eqn. 1.

$$G_{\mu,v}(z) = I(z) * \psi_{\mu,v}(z), \tag{1}$$

Here, z = (x, y) which is the 2D pixel's index along x and y plane and operator '* ' is the 2D convolution operator. μ and v are the orientation and the scales of the kernel, respectively. The kernel is defined as:

$$\psi_{\mu,v}(z) = \frac{\|k_{\mu,v}\|}{\sigma^2} e^{\left(-\frac{\|k_{\mu,v}\|^2 \|z\|^2}{2\sigma^2}\right)} \left[e^{-ik_{\mu,v}z} - e^{-\sigma^2/2}\right], \tag{2}$$

where $\|.\|$ is the norm operator and σ is the standard deviation of the distribution. The vector $k_{\mu,v}$ is defined as:

$$k_{\mu,v} = k_v e^{-i\phi_\mu} \tag{3}$$

where $k_v = k_{max}/f^v$ and $\phi_\mu = \pi\mu/8$; k_{max} is the maximum frequency, ϕ_μ is the kernel's orientation and f is the spacing between the kernels in the frequency domain [4]-[6].

2.2. Local Binary Pattern

Due to its relative simplicity, LBP has been applied successfully in many applications. The algorithm uses 3×3 windows of neighborhood pixels in the image to determine the new value of a pixel being considered [7]-[9]. Consider Fig.1, initially the algorithms probes all the 8-neighbood pixels around pixel *I(z),* any pixel greater than *I(z)* is assigned a binary bit value 1 while those whose values are less than or equals to *I(z) are* assigned bit value 0. 8-bit code is generated and is converted to decimal as the new value for *I(z).* The operation is applied to all the pixels in the image.

100	240	30
20	I(z)=120	185
70	100	200

→

0	1	0
0		1
0	0	1

Fig. 1. LBP operator

The LBP code for pixel *I(z)* can be computed by arranging the results of the operation starting from top-left corner clockwise is '01011000' which is equivalent to 152 in decimal.

2.3. Local Binary XOR Operator

LXP is very similar to LBP except that it applies XOR to 3×3 pixels neighborhood to decide the new value of a pixel. Due to the fact that it applies XOR operator the pixels values must be converted to zeros and ones before being applied [10,11]. For instance results from an image convolved with Gabor kernel may be formatted to logical by deciding that any value greater than zero is assigned a logical zeros while those with zeros and below are assigned logical ones. Fig. 2 shows how LXP is applied to the logically formatted image. For the new value of *I(z)* is to be determined, all the 8-neighborhood pixels are XOR-ed with *I(z)* and the resulting 8 bit codes are converted to decimal.

1	1	0
0	I(z)=1	0
0	1	0

→

0	0	1
1		1
1	0	1

Fig. 2. XLP operator

For instance, in the figure above the new LXP code for pixel *I(z)* starting from top-left corner clockwise is '00111011' which is equivalent to 115 in decimal.

2.4. Monogenic Local Binary Pattern

The motivation for this algorithm comes from the monogenic signal theory. It combines the local phase information, the local surface type information, and the traditional LBP to improve the performance of LBP in texture classification [4]. Based on this theory, three features are combined together to form monogenic 3-D texton feature vector to determine monogenic LBP. These features are; phase, φ_c, *r*otation *i*nvariant *u*niform pattern LBP, LBP^{riu2} and the monogenic curvature tensor S_c based on higher order *Riesz transforms*. Eqn. 4-6 describe these features. For more details refer to [4].

$$LBP_{P,R}^{riu2} = \begin{cases} \sum_{p=0}^{p} s(g_p - f_c) & if\ U(LBP_{p,r}) \leq 2, \\ P+1, & otherwise \end{cases}, \tag{4}$$

where;

$$U(LBP_{p,r}) = |s(g_{p-1} - g_c) - s(g_o - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(g_{p-1} - g_c)| \tag{5}$$

Superscript "riu2" means the use of rotation invariant "uniform" patterns that have U value of at most 2; s is the sign function; g_c corresponds to the gray value of the center pixel of the local neighborhood and g_p (p= 0,…, P-1) correspond to the gray values of P equally spaced pixels on a circle of radius R.

Phase φ_c, is defined as ;

$$\varphi_c = \varphi \Big/ (\pi/M) , \tag{6}$$

where M=5.

The last parameter S_c is defined by:

$$S_c = \begin{cases} 0, & \det(T_e) \leq 0 \\ 1, & else \end{cases}, \tag{7}$$

where $\det(T_e)$ is the determinant of the monogenic curvature tensor.

3. Normalization Schemes

Normalization techniques are quite often being used without much regards to the effect they can have on the general statistical distribution of the vectors to be normalized [8]-[10]. For instance, in fusion of the score levels of a various classifiers, a normalization scheme can be deploy to bring the scores within the same range. But in a vector sense, the normalization algorithm is more of a vector transform from one vector space to another. Hence the choice for a compatible normalizer becomes important as this may distort the vectors there by improving or decreasing the class seperability between two distinct class vectors [8,10]. Due to this fact, we investigated some of the most common normalization techniques to show how they affect class vectors distribution. Four normalization techniques are examined in this paper.

3.1. Z-Score Normalization

It is one of the most common normalization schemes. It uses the arithmetic mean and standard deviation of the vector. Z-score has a record of good performance on a set of data with Gaussian distribution. However, it is not robust due to the fact that it depends on the mean and standard deviation of the data which are both sensitive to outliers [13]. For a data point S_k, Z-score computes the new normalized value S'_k, using Eqn. 8.

$$S_k' = \frac{S_k - \mu}{\sigma} \quad , \tag{8}$$

where μ and σ are the mean and standard deviation of the distribution respectively.

3.2. MIN-MAX Normalization

Is one of the simplest of all the normalization techniques. This operator shifts the data sets within an interval [0, 1]. It can easily be seen that this technique is also not robust because presence of outliers in the distribution may affect the contribution of the majority datasets. Eqn. 9 defines min-max operator.

$$S_k' = \frac{S_k - min}{max - min}, \tag{9}$$

where *max* is the maximum data value in the distribution and *min* is the minimum data value of the distribution.

3.3. Median-MAD Normalization

The median and median absolute deviated as abbreviated (Median-MAD), are less sensitive to outliers and points at the extreme ends of the distribution. Therefore, this technique is robust. However, for distributions other than Gaussian, median and MAD are poor estimates of the location and the scales parameters [13]-[15]. Therefore, the scheme does not preserve the original distribution and does not transform the datasets into a common numerical range [13]. The equation below defines the median-MAD operation.

$$S_k' = \frac{S_k - median}{MAD}, \tag{10}$$

where *median* is the median of the distribution and MAD is the median of the absolute deviation from the median defined as $MAD = median(|S_k - median|)$.

3.4. Tangent-hyperbolic (Tanh) Normalization

Tanh normalization has been successfully used in many normalization schemes [14]. The tanh estimator is robust and very efficient. It is defined as;

$$S_k' = \frac{1}{2}\left\{\tanh\left(0.01\left(\frac{S_{k} - \mu_{GH}}{\sigma_{GH}}\right)\right) + 1\right\} \quad , \tag{11}$$

where μ_{GH} and σ_{GH} are the mean and standard deviation estimates, respectively.

Quite a number of normalization schemes do exist, for example Decimal Scaling normalization which is useful for data in logarithmic scales and Euclidean normalization. The ability of particular normalization algorithm to capture statistical distribution of a dataset will make it worthwhile.

4. Simulation Results

The proposed algorithm is implemented using a JAFFE database [5]. The database contains of 213 samples images of seven basic facial expressions (i.e. Neutral, Happy, Sad, Surprise, Anger Disgust and Fear) collected from 10 different subjects. To make it automatic 210 samples were used. 66.6% of the data were used for training and the remaining for testing.

The algorithm extracted the Gabor features from each samples at different orientation (i.e. **θ**=8) and from scale 1 through 3. At each orientation and scale monogenic LBP approach is used to extract features for the proposed mono-LGBP algorithm. Moreover, each mono-LGBP feature vector is normalized with four different normalization techniques as explained in section 3. Euclidean distance classifier is used to compute the distance metrics. The results from the two best performing normalization algorithm (Z-score and Tanh) with mono-LGBP are fused together using simple sum rule. In the same way three other LGP algorithms (Gabor-mag, LGBP and LGPP) were implemented to compare the results with the proposed method. Fig.3 shows training samples from JAFFE database of four different subjects with 7 basic facial expressions (e.g. Neutral, Happy, Sad, Surprise, Anger Disgust and Fear) from left to right accordingly. Tables 1-3 present simulation results using one to three of the Gabor kernel scales.

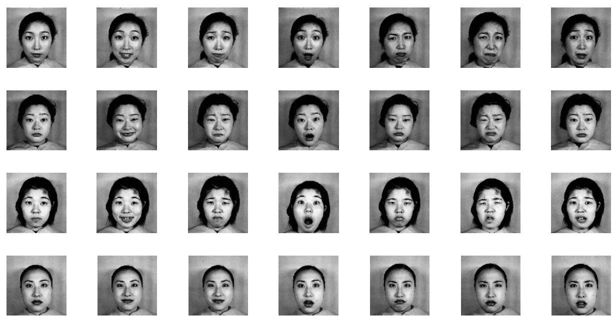

Fig. 3. Examples of images from the JAFFE Database

It is worth noting that the proposed mono-LGBP algorithm performance increases with the increase of Gabor scales. The fused results from Z-score and Tanh normalization algorithms gives a better performance. This is because in mono-LGBP, each of the two normalization schemes has been able to uniquely recognize some poses which are not being recognize by the other. Hence, the fusion of these results will lead to improving performance. The same can not be said for the other LGP. For example, Gabor-magnitude (Gabor-mag) has the best result with all the Z-score, tanh and min-max normalizations but unfortunately, they all pointed at the same recognition classes. Therefore fusing their results does not improve the performance. The same for LGPP and LGBP.

Fig. 4 shows graphical comparative of recognition performances of the proposed algorithm and the other algorithms.

Table 1. Experimental Results at one scale

Feature Extractors	**Normalization Algorithms**				
	Non	Z-score	min-max	M-MAD	tanh
G-Mag	88.6	90.0	88.6	90.0	90.0
LGPP	67.1	68.6	67.1	67.1	70.0
LGBP	34.3	40.0	34.3	40.0	50.0
mono-LGBP	64.3	64.3	57.1	57.1	67.1
mono- GBP	**Z-Score+ tanh = 77.14**				

Table 2. Experimental Results at two scales

Feature Extractors	**Normalization Algorithms**				
	Non	Z-score	min-max	M-MAD	tanh
G-Mag	90	91.4	91.4	90.0	91.4
LGPP	77.1	78.6	77.1	77.1	78.6
LGBP	50.0	50.0	40.0	50.0	77.1
mono-LGBP	80.0	80.0	74.3	71.4	80.0
mono-LGBP	**Z-Score+ tanh = 84.29**				

Table 3. Experimental Results at three scales

Feature Extractors	**Normalization Algorithms**				
	Non	Z-score	min-max	M-MAD	tanh
G-Mag	90	91.4	91.4	90	91.4
LGPP	77.1	82.1	80.1	77.1	85.7
LGBP	71.4	71.4	60.0	71.4	80.1
LGBP-mono	87.1	91.4	82.9	90	91.4
Mono-LGBP	**Z-Score + tanh = 92.83**				

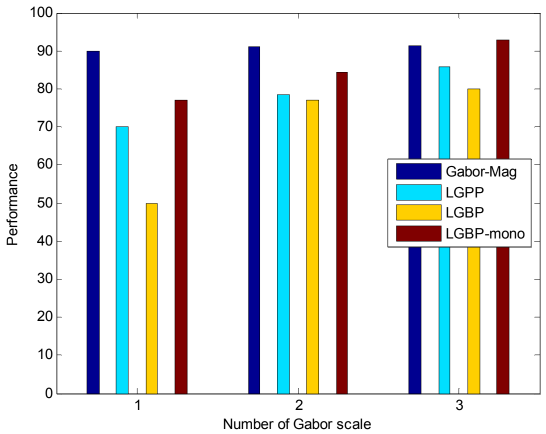

Fig. 4. Performance comparison betwwrn different algorithms.

5. Conclusions

A new approach for FER was proposed and implemented. The performance of the approach was compared with the existing LGP algorithms. The new approach was able to achieve better performance approximately 93% which is comparable to the best known results of FER on JAFFE using simple Euclidean classifier. The normalization scheme has also showed that it can influence a great deal of performance of a feature extraction algorithm if properly selected.

6. References

[1] C. Wei-Lun, D.Chaoa, Jian-Jiun and K. Author, " Facial Expression Recognition based on Improved Local Binary Pattern and Class-regularized Locality Preserving Projection ", *International Journal of Signal Processing*, Vol 117, pp. 1-10 Dec. 2015.

[2] J. Doughman, "High confidence recognition of Persons by Iris Patterns", *IEEE International Carnahan Conference on Security Technology*, pp. 254-263, 2001.

[3] J. Yanxia and R. Bo "Face Recognition using Local Gabor Phase Characteristics", *IEEE international conference on Intelligence and Software Engineering*, pp. 1-4, 2010.

[4] L. Zhang, L. Zhang, Z. Guo, and D. Zhang, "Monogenic-LBP: a New Approach for Rotation Invariant Texture Classification", *IEEE 17th International Conference on Image Processing (ICIP).* pp. 2677 - 2680, 2010.

[5] J.L. Michael, A. Shigeru, K. Miyuki and G. Jiro, "Coding Facial Expressions with Gabor Wavelets", *3rd IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 200-205, 1998.

[6] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active Shape Models, Their Training and Application," *Computer Vision and Image Understanding*, Vol. 61, No. 1, pp. 38-59, 1995.

[7] T. Ahonen, A. Hadid, and M. Pietikainen, " Face description with local binary patterns " *IEEE International Symposium on Computer, Consumer and Control (IS3C), 2014* , pp. 520 - 524, 2014.

[8] T. Ahonen, A. Hadid, and M. Pietikainen, " Face description with local binary patterns " *IEEE International Symposium on Computer, Consumer and Control (IS3C), 2014* , pp. 520 - 524, 2014.

[9] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multi-resolution Gray Scale and Rotation Invariant Texture Analysis with Local Binary Patterns", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 7, pp. 971-987, 2010.

[10] C. Liu and H. Wechsler, "Independent Component Analysis of Gabor Features for Face Recognition", *IEEE Transactions on Neural Networks,* Vol. 14, No. 4, pp. 919-928, 2003.

[11] A. Garg and V. Choudhary. "Facial Expression Recognition Using Principal Component Analysis'', *International Journal of scientific Research Engineering and Technology*, Vol.1, No. 4, pp. 039-042, 2012.

[12] M. Turk and A. Pentland, "Eigen faces for Recognition", *Journal of Cognitive Neuro-science*, Vol. 3, No. 1, pp. 71-86, 1991.

[13] S. Ribaric, I. Fratric, "Experimental Evaluation of Matching-Score Normalization Techniques on Different Multimodal Biometric Systems", *IEEE Mediterranean Electrotechnical Conference, MELECON* , pp.498-501, 2006.

[14] N. Nandakumar ,K. J, Anil and A.A Arun, "Score Normalization in Multimodal Biometric Systems" , *The Journal of Pattern Recognition Society, Elservier,* Vol. 38, pp. 2270-2285, Dec. 2005

[15] M. Sigdel, S. Dinc, M.S. Sigdel, M.L. Pusey, and R.S. Aygun "Evaluation of Normalization and PCA on the Performance of Classifiers for Protein Crystallization Images" *IEEE conference on SOUTHEASTCON*, pp. 1-6, 2014.