

A MONOGENIC LOCAL GABOR BINARY PATTERN FOR FACIAL EXPRESSION RECOGNITION

¹Alaa ELEYAN, ²Abubakar M. ASHIR

¹Avrasya University, Electrical & Electronics Engineering Department, Trabzon, TURKEY

²Selçuk University, Electrical & Electronics Engineering Department, Konya, TURKEY

¹aeleyan@avrasya.edu.tr, ²ashir4real@yahoo.com

(Geliş/Received: 10.01.2017; Kabul/Accepted in Revised Form: 19.04.2017)

ABSTRACT: The paper implements a monogenic-Local Binary Pattern (mono-LBP) algorithm on Local Gabor Pattern (LGP). The proposed approach initially features from the samples using LGP at different scales and orientation. The extracted LGP features are further enhanced by decomposing it into three monogenic LBP channels before being recombined to generate the final feature vector. Different Normalization schemes are applied to the final feature vector. Two best performing normalization algorithms with mono-LBP are fused at score level to obtain an improved performance using K-Nearest Neighbor classifier with L_{1-norm} as a distance metrics. Moreover, performance comparison is done with other variants of LGP algorithm and also the effects of various normalization techniques are investigated. Experimental results from JAFFE and TFEID facial expression databases show that the new technique has improved performance compared to its counterparts.

Key Words: Facial expression recognition, Local gabor patterns, Monogenic local binary patterns.

Yüz İfade Tanınması İçin Bir Monojenik Yerel Gabor İkili Desen

ÖZ: Bu makale, yerel Gabor Desen (LGP) üzerinde monojenik-Yerel İkili Desen (mono-LBP) algoritmasını uygular. Önerilen algoritma Gabor çekirdeğinin farklı ölçeklerinde ve farklı normalizasyon şemaları ile uygulanır. Mono-LBP ile en iyi performans gösteren normalleştirme algoritmalarından elde edilen sonuçlar, geliştirilmiş bir performans elde etmek için skor düzeyinde birleştirilmiştir. Üstelik, performans karşılaştırması diğer LGP algoritmasının türevleri ile yapılmıştır ve ayrıca çeşitli normalleştirme tekniklerinin etkileri araştırılmaktadır. JAFFE yüz ifadesi veritabanında yapılan deneysel sonuçlarına göre, önerilen yaklaşım bir sınıflandırıcı olarak mesafe metrikini kullanarak mevcut algoritmalarla kıyasla en iyi ortalama performansa sahip olduğunu göstermektedir.

Anahtar Kelimeler: Yüz ifade tanıma, Yerel gabor desen, Monojenik yerel ikili desen.

INTRODUCTION

Facial Expression Recognition (FER) has recently been one of leading field drawing 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 (Chao *et al.*, 2015). Quite a handful number of feature extraction and classifier algorithms have been proposed and implemented in this field (FER). Gabor kernel has been 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, multi-resolution approach and direction selectivity (Chao *et al.*, 2015).

Following successful implementation of Gabor kernels in iris recognition by 2001 (Doughman, 2001) and coupled with the success of local binary pattern (LBP) algorithm, several variants of Gabor algorithms emerged over times. These Gabor variants are sometimes referred to as Local Gabor Patterns (LGP). LGP algorithms exploit various Gabor feature channels such as magnitude, phase, imaginary and real channels. For instance, (Yanxia and Bo, 2010) proposed 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. In (Zhang *et al.*, 2010), the authors 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 (LGPD) 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.

A rotation invariant monogenic LBP which was proposed for texture classification in (Zhang *et al.*, 2010) is used in this work. 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, is 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 and normalization schemes. Section III describes the proposed approach and section IV presents the experimental results. Section V summarizes the findings.

FEATURE EXTRACTION AND NORMALIZATION SCHEMES

A brief literature background on the feature extraction operators and normalization schemes deployed in the course of this work are discussed below.

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, $\psi_{\theta,v}(z)$ with an image, $I(z)$ is represented as $G_{\theta,v}(z)$ in Eqn. 1.

$$G_{\theta,v}(z) = I(z) * \psi_{\theta,v}(z), \quad (1)$$

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

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

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

$$k_{\theta,v} = k_v e^{-i\phi_\theta} \quad (3)$$

where $k_v = k_{max}/f^v$ and $\phi_\theta = \pi\theta/8$; k_{max} is the maximum frequency, ϕ_θ is the kernel's orientation and f is the spacing between the kernels in the frequency domain (Eleyan *et al.*, 2008; Liu and Wechsler, 2003; Lyon *et al.*, 1998; Cootes *et al.*, 1995).

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 (Ahonen *et al.*, 2006; Ojala *et al.*, 2010, Tran *et al.*, 2014). Consider Figure 1, initially, the algorithms probes the 8-neighborhood pixels around pixel z . Any pixel greater than z is assigned a binary bit value 1 otherwise, assigned bit value 0. An 8-bit code is generated and is converted to decimal and recorded as the new value for z . The operation is applied to all the pixels in the image.

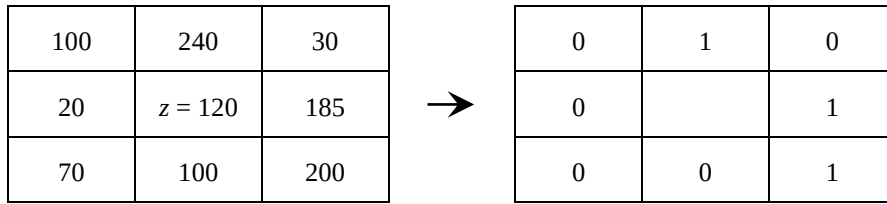


Figure 1. LBP operator

The LBP code for pixel 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.

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 (Zhang *et al.*, 2007). 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. Figure 2 shows how LXP is applied to the logically formatted image. For the new value of z is to be determined, all the 8-neighborhood pixels are XOR-ed with z and the resulting 8 bit codes are converted to decimal.

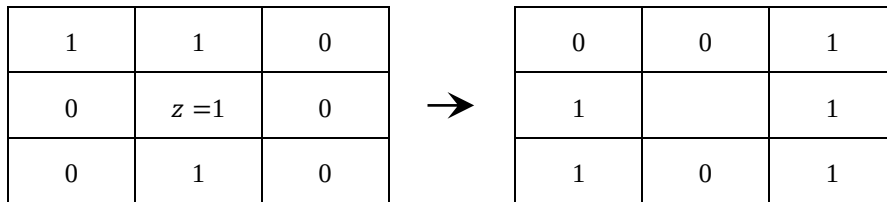


Figure 2. LXP operator

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

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 (Zhang *et al.*, 2010). 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 , rotation invariant uniform pattern LBP, LBPriu2 and the monogenic curvature tensor S_c based on higher order Riesz transforms. Eqn. 4-6 describe these features. For more details refer to (Zhang *et al.*, 2010).

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

where;

$$U(LBP_{p,r}) = |s(g_{p-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(g_{p-1} - g_c)| \quad (5)$$

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 ($p = 0, \dots, P - 1$) correspond to the gray values of P equally spaced pixels on a circle of radius R .

Phase φ_c , is defined as ;

$$\varphi_c = \varphi / (\pi/M) \quad (6)$$

where $M = 5$.

The last parameter S_c is defined by:

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

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

Normalization Operators

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 (Ribaric and Fratric, 2006; Nandakumar *et al.*, 2005). For instance, in fusion of the score levels of various classifiers, a normalization scheme can be deployed 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 separability between two distinct class vectors. 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.

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 (Ribaric and Fratric, 2006). For a data point S_k , Z-score computes the new normalized value S'_k , using Eqn. 8.

$$S'_k = \frac{S_k - \mu}{\sigma} \quad , \quad (8)$$

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

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.

$$S'_k = \frac{S_k - \min}{\max - \min} \quad (9)$$

where $\max$ is the maximum data value of the distribution, and $\min$ is the minimum data value of the distribution.

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 (Ribaric and Fratric, 2006; Sigdel *et al.*, 2014, Nandakumar *et al.*, 2005). Therefore, the scheme does not preserve the original distribution and does not transform the datasets into a common numerical range (Ribaric and Fratric, 2006). The equation below defines the median-MAD operation.

$$S'_k = \frac{S_k - \text{median}}{\text{MAD}} \quad (10)$$

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

Tangent-hyperbolic (Tanh) Normalization

Tanh normalization has been successfully used in many normalization schemes (Nandakumar *et al.*, 2005). The tanh estimator is robust and very efficient. It is defined as;

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

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.

PROPOSED APPROACH

The proposed approach has both encompasses the critical stages of a facial expression algorithms. Initially During feature extraction, the proposed approach extract Gabor features from each sample using different orientation (i.e. $\theta = 8$) and scale (1 to 3) of the Gabor filter. For each Local Gabor features extracted at a specified orientation and scale, monogenic LBP is further applied to the extracted Gabor Features using equations (4, 6 and 7). The three monogenic LBP features ($LBP_{P,R}^{riu2}$, φ_c , and S_c) are combined together to form a single monogenic 3-D texton feature vector. The monogenic 3-D texton feature vector is adopted as the final feature vector known as the mono-LGBP algorithm in the context of this work.

Moreover, as a way of exploiting performance, a normalization scheme is applied at the feature level of the proposed approach. Hence different normalization approaches are applied to the mono-LGBP feature vector before classification. Four different normalization techniques are investigated as explained in chapter 3.

In the classification stage, KNN has being used with l_2 -norm as the distance measure. Each Euclidean representation of mono-LGBP feature with different normalization algorithm is classified separately using KNN. Based on the performance of the normalization representation of the mono-LGBP feature, two of the best normalization schemes are fused at score level of the classifier using a simple *sum rule* to obtain a better performance. Figure depicts the flowchart of the proposed approach.

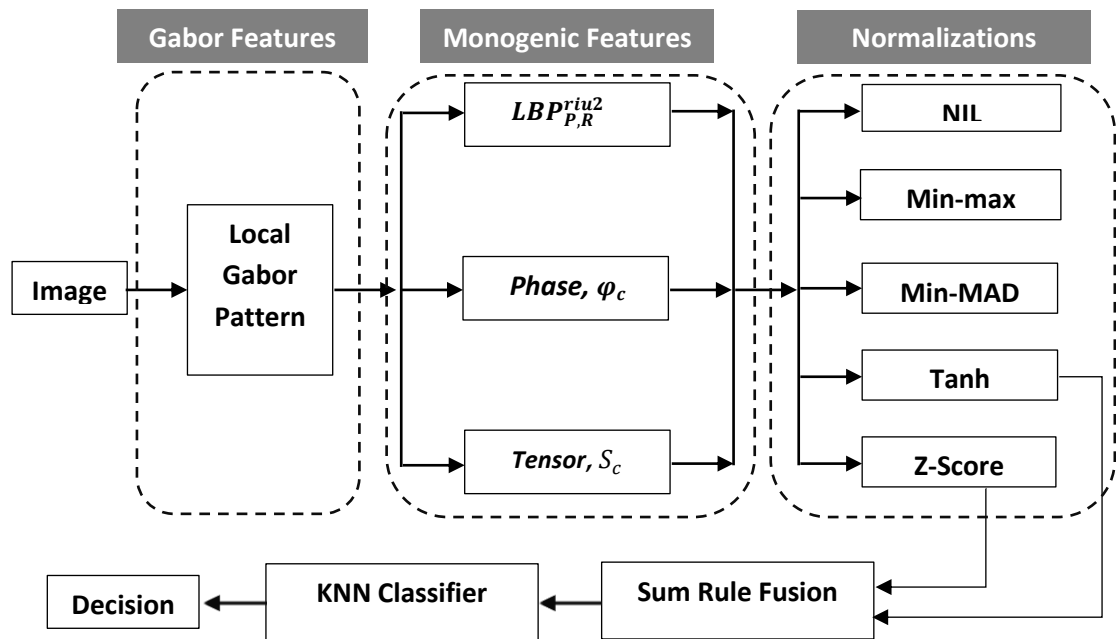


Figure 3. Proposed approach Flowchart

EXPERIMENTAL RESULTS

The proposed algorithm is implemented using two different facial expression databases which include Japanese Female Facial Expression JAFFE (Lyon *et. al* 1998) and Taiwanese Facial Expression Image Database TFEID (Chen and Yen, 2007). JAFFE database contains a total of 213 samples images of seven basic facial expressions (i.e. Neutral, Happy, Sad, Surprise, Anger Disgust and Fear) collected from 10 different subjects. The number of samples per expression in each subject ranges from 2 to 4. While the current public TFEID database consist of facial images from 20 male models each acquired

with a frontal view between 0° to 45° . It constitutes 8 expressions with *contempt* as the eighth expression in addition to the seven basic expressions found in JAFFE.

During training all the sample images were grouped into emotions classes (i.e. 7 for JAFFE and 8 for TFEID) irrespective of subjects to which they belong to (i.e. person-independent). Using Leave-One-Pose-Out (LOPO) procedures, one sample is drawn from each class for training while the remaining samples are used for training. This process is repeated and rotated until each sample is uniquely use as a training set. The overall performance is given as the average performance of the entire number of times the training is repeated.

Table 1 to 6 display the experimental results from the proposed approach in comparison to its counterparts. Results from different normalizations schemes and three other variants of LGP algorithms (Gabor-magnitude features, LGBP and LGPP) were implemented to compare the results with the proposed approach for the two database used. The same experimental procedures were adopted in throughout the experiments. Similarly, results from the fusion of the proposed mono-LGBP were also included to show the leverages of the fusion techniques over the non-fusion approach. Figure 4 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, respectively.

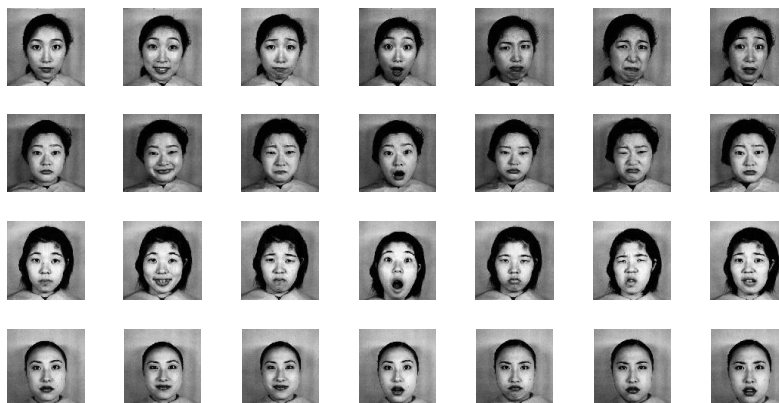


Figure 4. Examples of images from the JAFFE Database

Results Discussion

It is worth noting that the proposed mono-LGBP algorithm performance increases with the increase of Gabor scales (see Tables 1-6). 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 cannot 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.

Table 1. Experimental Results at one scale ($\theta = 8$, $v = 1$) using JAFFE Database.

Feature Extractors	non	Z-score	min-max	M-MAD	tanh
LGP-Magnitude	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
Fused mono-LGBP	Z-Score+tanh = 77.14				

Table 2. Experimental Results at two scales ($\theta = 8$, $v = 2$) using JAFFE Database.

Feature Extractors	non	Z-score	min-max	M-MAD	tanh
LGP-Magnitude	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
Fused mono-LGBP	Z-Score+tanh = 84.29				

Table 3. Experimental Results at three scales ($\theta = 8$, $v = 3$) using JAFFE Database.

Feature Extractors	non	Z-score	min-max	M-MAD	tanh
LGP-Magnitude	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
mono-LGBP	87.1	91.4	82.9	90	91.4
Fused mono-LGBP	Z-Score+tanh = 92.83				

Table 4. Experimental Results at one scale ($\theta = 8$, $v = 1$) using TFEID Database.

Feature Extractors	non	Z-score	min-max	M-MAD	tanh
LGP-Magnitude	87.6	92.1	90.4	93.0	93.3
LGPP	70.1	73.6	69.1	70.1	75.5
LGBP	50.3	59.0	49.3	52.0	65.0
mono-LGBP	64.3	68.3	59.6	60.8	73.4
Fused mono-LGBP	Z-Score+tanh = 82.15				

Table 5. Experimental Results at two scales ($\theta = 8$, $v = 2$) using TFEID Database.

Feature Extractors	non	Z-score	min-max	M-MAD	tanh
LGP-Magnitude	90.6	94.4	92.6	91.0	94.4
LGPP	80.3	84.6	80.9	81.1	83.9
LGBP	60.0	60.9	50.2	60.4	87.1
mono-LGBP	90.0	92.0	82.3	90.4	96.0
Fused mono-LGBP	Z-Score+tanh = 96.29				

Table 6. Experimental Results at three scales ($\theta = 8$, $v = 3$) using TFEID Database.

Feature Extractors	non	Z-score	min-max	M-MAD	tanh
LGP-Magnitude	89.2	90.5	91.4	93.3	92.4
LGPP	77.1	82.1	80.1	77.1	85.7
LGBP	71.4	71.4	60.0	71.4	80.1
mono-LGBP	91.1	96.4	92.9	94.5	97.4
Fused mono-LGBP	Z-Score+tanh = 97.9				

CONCLUSION

A new approach for facial expression recognition was proposed and implemented. The performance of the proposed approach was compared with the existing LGP algorithms using different normalization schemes. The new approach was able to achieve better performance approximately 92.8% with JAFFE

database and 97.9% with TFEID database. The results are comparable to the best-known results of facial expression recognition on JAFFE database and TFEID in the literature using KNN as a classifier. The normalization schemes further indicate that a great deal of performance can be realized with a proper application of normalization algorithm to extracted feature vectors. The results also confirmed the effectiveness of the fusion technique deployed in the proposed approach.

REFERENCES

- Ahonen T., Hadid A., Pietikainen M., 2006, "Face Description with Local Binary Patterns: Application to Face Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28(12), pp. 2037-2041.
- Chao, W. L., Ding, J. J., Liu, J. Z., 2015, "Facial Expression Recognition based on Improved Local Binary Pattern and Class-regularized Locality Preserving Projection", *Signal Processing*, Vol. 117, pp. 1-10.
- Chen, L. F., Yen, Y. S., 2007, "Taiwanese Facial Expression Image Database", *Brain Mapping Laboratory, Institute of Brain Science, National Yang-Ming University*, Taipei, Taiwan.
- Cootes, T. F., Taylor, C. J., Cooper, D. H., Graham, J., 1995, "Active Shape Models, Their Training and Application," *Computer Vision and Image Understanding*, Vol. 61(1), pp. 38-59.
- Doughman, J., 2001, "High Confidence Recognition of Persons by Iris Patterns", *IEEE International Carnahan Conference on Security Technology*, London, England, pp. 254-263, 16-19 October 2001.
- Eleyan, A., Demirel, H., Özkaramanli, H., 2008, "Complex Wavelet Transform-Based Face Recognition", *EURASIP Journal on Advances in Signal Processing*, Vol. 2008, pp. 1-13.
- Liu, C., Wechsler, H., 2003, "Independent Component Analysis of Gabor Features for Face Recognition", *IEEE Transactions on Neural Networks*, Vol. 14(4), pp. 919-928.
- Lyons, M., Akamatsu, S., Kamachi, M. and Gyoba, J., 1998, "Coding Facial Expressions with Gabor Wavelets", *3rd IEEE International Conference on Automatic Face and Gesture Recognition (AFGR)*, Nara, Japan, pp. 200-205, 14-16 April 1998.
- Nandakumar, K., Jain, A., Ross, A., 2005, "Score Normalization in Multimodal Biometric Systems", *The Journal of Pattern Recognition Society*, Elsevier, Vol. 38, pp. 2270-2285.
- Ojala, T., Pietikainen, M., Maenpaa, T., 2010, "Multi-resolution Gray Scale and Rotation Invariant Texture Analysis with Local Binary Patterns", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24(7), pp. 971-987.
- Ribaric S., Fratric I., 2006, "Experimental Evaluation of Matching-Score Normalization Techniques on Different Multimodal Biometric Systems", *IEEE Mediterranean Electrotechnical Conference, MELECON*, Malaga, Spain, pp. 498-501, 16-19 May 2006.
- Sigdel, M., Dinc, S., Sigdel, M. S., Pusey, M. L., Aygun, R. S., 2014, "Evaluation of Normalization and PCA on the Performance of Classifiers for Protein Crystallization Images" *IEEE Conference on SOUTHEASTCON*, Lexington, KY, USA, pp. 1-6, 13-16 March 2014.
- Tran C. K., Lee T. F., Chang L., Chao P. J., 2014, "Face Description with Local Binary Patterns and Local Ternary Patterns: Improving Face Recognition Performance Using Similarity Feature-Based Selection and Classification Algorithm", *International Symposium on Computer, Consumer and Control*, Taichung, Taiwan, pp. 520-524. 10-12 June 2014.
- Yanxia, J. Bo, R., 2010, "Face Recognition using Local Gabor Phase Characteristics", *IEEE International Conference on Intelligence and Software Engineering*, Wuhan, China, pp. 1-4, 10-12 December 2010.
- Zhang, B., Shan, S., Chen, X., Gao, W., 2007, "Histogram of Gabor Phase Patterns (HGPP): A Novel Object Representation Approach for Face Recognition", *IEEE Transaction on Image Processing*, Vol. 16(1), pp. 57-68.
- Zhang L., Zhang L., Guo Z., Zhang D., 2010, "Monogenic-LBP: A New Approach for Rotation Invariant Texture Classification", *17th IEEE International Conference on Image Processing (ICIP)*, Hong Kong, China, pp. 2677 – 2680, 26-29 September 2010.