



Local Directional Threshold based Binary Patterns for Facial Expression Recognition and Analysis

V Uma Maheswari^{1*}, Vara Prasad², S Viswanadha Raju³

¹Department of Computer Science and Engineering

Vardhaman College of Engineering, Hyderabad 501 218, India.

²BMS College of Engineering, Bangalore 566 019, 501 218, India.

³JNTUH College of Engineering Jagityal 505501, India.

*Corresponding author E-mail: umasridhar11@gmail.com

Abstract

In this paper, proposing a novel method to retrieve the edge and texture information from facial images named local directional standard matrix (LDSM) and local dynamic threshold based binary pattern (LDTBP). LBP and LTP operators are used for texture extraction of an image by finding difference between center and surrounding pixels but they failed to detect edges and large intensity variations. Thus addressed such problems in proposed method firstly, calculated the LDSM matrix with standard deviation of horizontal and vertical pixels of each pixel. Therefore, values are encoded based on the dynamic threshold which is calculated from median of LDSM values of each pixel called LDTBP. In experiments used LFW facial expression dataset so used SVM classifier to classify the images and retrieved relevant images then measured in terms of average precision and average recall.

Keywords: LDSM (Local Directional Standard Matrix), LDTBP (Local Dynamic Threshold based Binary Pattern), SVM (Support Vector Machine) Classifier, edge detection, facial expression recognition.

1. Introduction

Due to volatile expansion of internet usage and digital equipment like cameras, smart phones, tabs, surveillance etc. generating abundant data frequently so it makes system difficult to manage the data to the administrator as it is. It demands a novel and dire system which is retrieving features from the existing data such as color, shape and texture etc. in addition image classification is gaining importance day by day due to its tremendous demand in various areas such as medical, entertainment, crime analysis, face recognition, facial expression analysis, gender identification, military, object recognition or object tracking, government applications etc. in classification, feature vector generation plays an imperative role can be calculated from the either low-level features, high-level features or combinational features are used to retrieve the features from an image. Therefore, texture feature extraction is the branch of texture analysis is the prominent step in feature extraction. Though different descriptors are proposed like LBP, LTP, LTrp, LDTP, LGP etc[8][9][17][7]. and also proved to be successful in various applications but fails in some applications due to generate the generalized feature descriptor for image classification. The most extensive and comprehensive survey about content based image retrieval current techniques are mentioned in [1 – 4].

2. Related work

Zhen et. al [6] has proposed combinational features such as range, space and direction for face recognition and applied convolutional multidirectional and multiscale gabor filter on decomposed facial images. Jun et. al [7] has used local gradient patterns for face

recognition and has been proved as insensitive to global intensity values. Timo Ojala et. al [8] had introduced a Local binary patterns which is silver lining method in texture feature extraction and LBP is familiar for face recognition and expression analysis. Murala et. al [9] introduced local tetra patterns for texture feature extraction in CBIR based on the horizontal and vertical directions of nth order derivatives and Murala has also proposed local extrema patterns as extension of local derivative patterns based on directions towards pixel intensity values [10]. Vipparthi et. al [11] has proposed a color directional quinary patterns by dividing the image pixels into RGB channels then applied Directional binary code in four directions and differences are substituted with four quantized values. Varaprasad et. al [12] introduced detection of potholes for independent vehicle in an unstructured environment based on road images taken from sensors. Viswanadha Raju et. al [13] proposed amalgamation of Color and texture features used for feature extraction from an image in content based images retrieval (CBIR) system that extended to retrieve sub images. Dubey et. al [14] has proposed the multichannel decoded LBPs to reduce the dimensions by concatenating the channels separately. Mostafijur et. al [15] proposed discriminative ternary census transform histogram (DTCTH) for image representation and classification based on the dynamic threshold that is calculated from square root of center pixel. Majunder et. al [16] has proposed new automatic expression system (AFERS) using deep learning techniques with combination of geometric feature feature, LBP, fusion of both then classified using SOM classifier. Ryu et. al [17] has introduced local directional ternary patterns (LDTP) for facial expression recognition and analysis. Zia et. al [18] proposed local directional rank histogram pattern (LDRHP) by calculating the eight histograms of eight local directional strengths of an given image.



Gray level co-occurrence matrix (GLCM), Gabor filters, and morphology profiles are used for texture and shape feature extraction from the neighboring pixels[19].GLCM is used for melanoma image segmentation [20].

3. Local patterns

3.1. Local binary patterns

Local binary pattern operator is introduced by Timo Ojala [8]. It is been used as a texture feature descriptor which is invented for texture classification. LBP is calculated by differentiating the center pixel and neighbouring 8 pixels values then encoded with binary values based on comparison using Equ (1).

$$LBP_{P,R}(I) = \sum_{i=0}^{p-1} s(g_i - g_c) \times 2^i \quad (1)$$

$$\text{Where, } s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & \text{else} \end{cases}$$

'P' and 'R' are the surrounding pixel values and radius of the pixels from center pixel, 'I' is an given image, ' g_i ' and ' g_c ' are the i th neighboring pixel and center pixel respectively. 'x' is the difference of the center and neighbouring pixel gray values.

3.2. Local ternary patterns

Tan and Triggs et. al [5] introduced local ternary patterns with three values instead of binary values in LBP then divided into upper and lower patterns with positive and negative substitution by Eqn (2).

$$LTP(I) = \sum_{i=0}^p u(x) \times 3^i \quad (2)$$

$$\text{Where, } u(x) = \begin{cases} 1, & x \geq g_c + \delta \\ 0, & |x - g_c| < \delta \\ -1, & x \leq g_c - \delta \end{cases}$$

'x' and ' g_c ' are same as mentioned in LBP ' δ ' is the user defined threshold.

4. Gray Level Co Occurrence Matrix (GLCM)

GLCM matrix is statistical matrix method of analysing textural features calculating from the spatial relationship between the neighboring and referenced pixel. Thus, matrix is formed with the values sum of the frequency of occurrences in rows with that referenced 'y' pixel in element(x, y) in the given image. Here, the amount of different gray level values in an image is considered as the size of the GLCM matrix.

5. Proposed method

LTP, CDLQP [5][11] are encoded based on static threshold value and LGP, DTCTH [7][15] etc. are used dynamic threshold which is calculated from image pixels are motivated to propose this new feature for facial expression analysis. Proposed patterns are calculated based on LDSM matrix using Eqn (3), which is calculated from standard deviation of 4 directional pixels of each pixel and place the value in the center to construct LDSM matrix as shown in Fig.1.

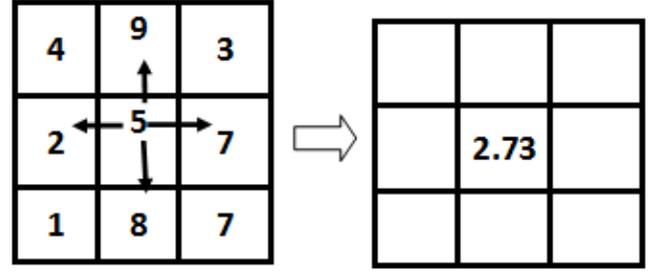


Fig.1 calculation of standard deviation of 4 directional pixels.

and place into center to form the LDSM matrix binary patten is generated based on dynamic threshold value is named local dynamic threshold based binary patterns (LDTBP) calculation is illustrated in Fig. 2. If the LDSM matrix value is greater than or equal to threshold value encoded with '1', otherwise encoded with '0'.

$$LDSM = \sum_{i=0}^{p-1} f'(x) \times 2^i \quad (3)$$

Where ' $f'(x)$ ', is the function which is used to calculate the LDSM matrix from Equ (3)

$$f'(x) = \sqrt{\frac{\sum_{i=0}^{p-1} (g_{\theta} - \bar{g}_{\theta})^2}{p}} \Big|_{\theta=0^{\circ}, 90^{\circ}, 180^{\circ}, 270^{\circ}} \quad (3)$$

where, ' g_{θ} ' is the neighboring horizontal and vertical in directions $0^{\circ}, 90^{\circ}, 180^{\circ}, 270^{\circ}$ pixels and ' \bar{g}_{θ} ' is the mean of neighboring $0^{\circ}, 90^{\circ}, 180^{\circ}, 270^{\circ}$ pixels, 'p' is the number of pixels around given each pixel in directions $0^{\circ}, 90^{\circ}, 180^{\circ}, 270^{\circ}$.

$$f''(x) = \begin{cases} 1, & f'(x) \geq T \\ 0, & \text{Otherwise} \end{cases} \quad (4)$$

Where, ' T ', is a dynamic threshold which is calculated as follows:

$$T = \text{median}\{s(x) : x = 1, 2, \dots, 8\} \quad (5)$$

here 'x' is the value of LDSM matrix

5.1. Algorithm:

Proposed algorithm steps are illustrated in Fig. 1.

1. Upload the RGB image and convert into grayscale.
2. Construct the LDSM matrix using Eqn(3), illustrated in Fig.2.
3. Encode the matrix values based on dynamic threshold calculated from Eqns (5) and generate the pattern from Eqn (4).
4. Construct the histogram for LDTBP pattern and generate the feature vector.
5. Classify the images using SVM classifier using proposed feature vector.
6. Apply similarity measure algorithms using Equ (6) & (7).
7. Retrieve the expected images based on top matches using Equ (10).

Table.1 shows the retrieval images expression wise compared to other existing methods.



Fig. 2. Framework for the proposed method ie. Local Directional Threshold based Binary Patterns (LDTBP).

5.2. Advantages over other methods:

- As shown in Fig.2, edges are detected using LDSM matrix efficiently than variance values and no need to apply any filters like GLBP, GLTP to preprocess the images.
- Applied only 4 directions instead of calculating 8 generally. Thus, requires less computational cost compared to others.
- Calculated LDTBP patterns to retrieve the local texture of an image, so gives the more feature.

6. SVM Classifier

SVM is the supervised learning technique and it is proved as better classifier in many applications due to its efficiency. SVM is

3 X 3 Query image

4	9	3
2	5	7
1	8	7

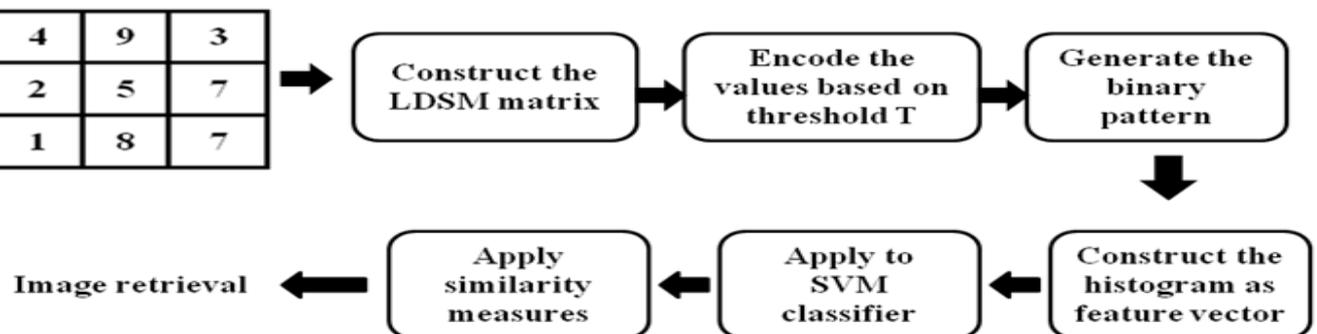


Fig.3 Framework for the proposed method ie.. Local Directional Threshold based Binary Patterns (LDTBP).

LFW dataset:

used in detecting the facial emotion detection to classify the different emotional expressions. It provides the hypothesis function employed to predict the output for the further input [21]. Kotsia and Pitas used SVM for expression classification associated with Gabor wavelet. SVM productivity is always the label utilized to classify the grid information of six essential facial expressions under examination. It resolves only one optimization problem but multi class SVM's are used to construct the six basic facial expressional rules [22]. SVM can be used for linear classification by using two classes and consequent class labels as '1' and '-1' [23]. SVM can also enhance to solve regression problems as support vector regression.

LFW facial database consists around 2338 cropped images of different subjects and emotions of different people in various illuminations as shown in Fig.3 which are frequently used in experiments to detect faces and facial expressions. In our experiment we used SVM classifier to classify the expressions based on the feature vector. The margin and kernel parameter are been set over the cross validation on the training dataset. Results are calculated using Eqn(11) & (12) in terms of ARP and ARR are illustrated in Fig.4 & Fig.5.



Fig.4 Sample images of LFW face dataset of different people.

7. Similarity Measurements

Image feature descriptor is acquired from the feature extraction of an image is known as feature vector [33][41]. Subsequently, every image is represented by the feature vector in the database. The main perspective of this is to get the top matching images to a given query image. This makes the system to retrieve the top matched images by reducing the gap between the given query image and the database images. This can be done by four mentioned distance similarity measurements.

Manhattan (or) L1 Distance:

$$D_s(Q_i, T_i) = \sum_i |f_i(Q_i) - f_i(T_i)| \quad (6)$$

Euclidean or L2 Distance:

$$D_s(Q_i, T_i) = \sqrt{(f_i(Q_i) - f_i(T_i))^2} \quad (7)$$

Where, ' Q_i ' and ' T_i ' are the query image and reference image, ' $f_i(Q_i)$ ' and ' $f_i(T_i)$ ' are feature of the query and reference image

Precision and Recall are the accuracy measurement terms described for a given image query I_q

$$\text{Precision}(I_q, n) = \frac{1}{n} \sum_{i=1}^{|\text{DB}|} \left| \left(\Delta(\lambda(I_i), \lambda(I_q)) \right) \middle| \text{Rank}(I_i, I_q) \leq n \right| \quad (8)$$

$$\text{Recall}(I_q, n) = \frac{1}{N_G} \sum_{i=1}^{|\text{DB}|} \left| \left(\Delta(\lambda(I_i), \lambda(I_q)) \right) \middle| \text{Rank}(I_i, I_q) \leq n \right| \quad (9)$$

' N_G ' is the images relevant to query in databases, 'n' says the number of top images, ' $\lambda(I_i)$ ' is the category of image x. 'Rank()' function provides the rank of a particular image with respect to query image ' I_q ' amongst database images and ' $|\text{DB}|$ ' represents the total images in a database.

$$\Delta(\lambda(I_i), \lambda(I_q)) = \begin{cases} 1 & \lambda(I_i) = \lambda(I_q) \\ 0 & \lambda(I_i) \neq \lambda(I_q) \end{cases} \quad (10)$$

ARP and ARR are defined as follows:

$$\text{ARP} = \frac{1}{|\text{DB}|} \sum_{i=1}^{|\text{DB}|} \text{Precision}(I_i, n) \quad (11)$$

$$\text{ARR} = \frac{1}{|\text{DB}|} \sum_{i=1}^{|\text{DB}|} \text{Recall}(I_i, n) \quad (12)$$

Table 1. Confusion matrix of top ten images retrieved form dataset

	Happy	sad	Disgust	Surprise	Angry	Neutral
Happy	7	0	2	1	0	0
sad	0	8	1	0	1	0
Disgust	0	1	7	0	1	1
Surprise	1	0	0	9	0	0
Angry	0	0	0	0	9	1
Neutral	0	2	0	0	0	8

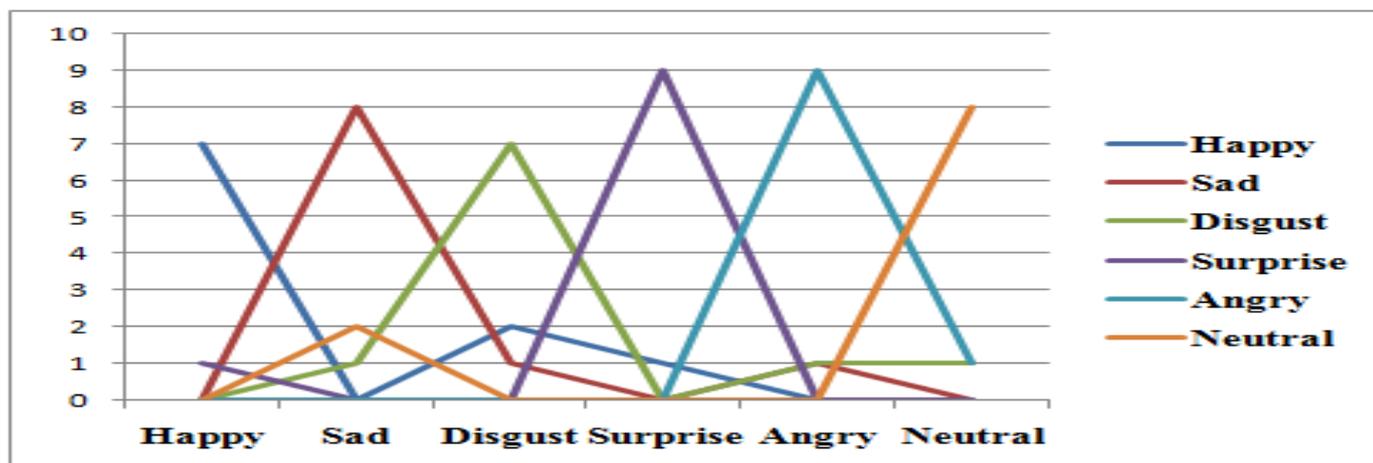


Fig.5 Illustration of expression recognition among top ten images on LFW DB.

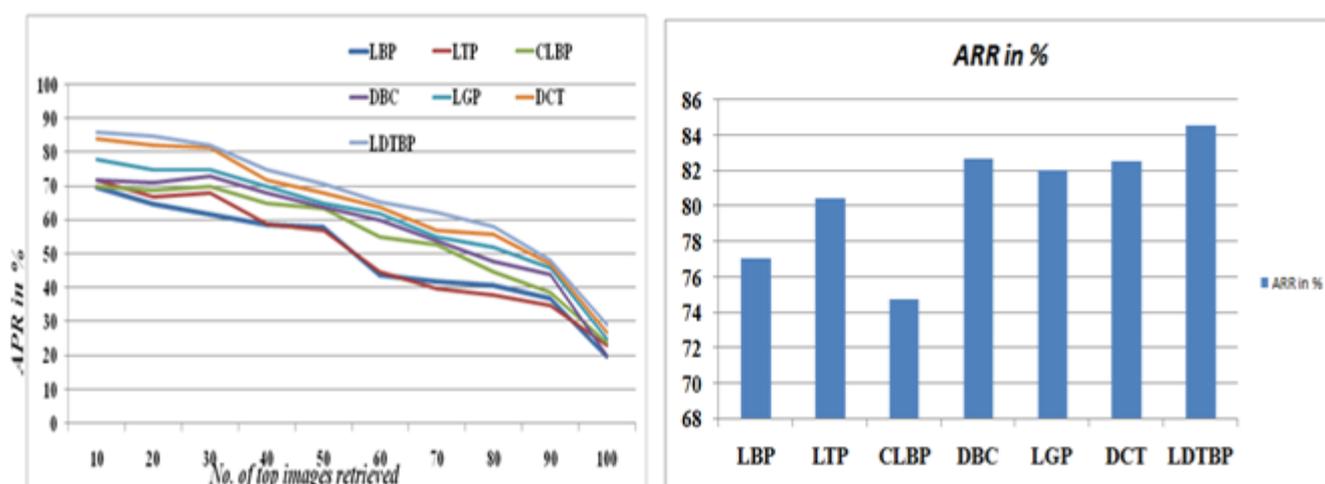


Fig.6 Comparison of existing methods in terms of ARP and ARR % on LFW DB

8. Conclusion

The proposed method works along with directions and threshold based binary substitution using LDSM matrix and LDTBP patterns. Thus, it proved as better method to detect the edges and extract the local texture features which helps to identify the expected image. Proposed method used statistical values to detect the edges instead of filters and dynamic threshold helps to retrieve the feature correctly from an any image rather than using static threshold to specific image. Furthermore, we used four directions only to reduce the computational cost, can also use diagonal pixels also to improve accuracy in edge detection.

References

- [1] M. Kokare, B. N. Chatterji, and P. K. Biswas, "A survey on current content based image retrieval methods," *IETE J. Res.*, 48,(3)&(4), 2002, pp. 261–271.
- [2] Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma, "A survey of content-based image retrieval with high-level semantics," *Pattern Recogn.*, 40,(1), Jan. 2007, pp. 262–282.
- [3] Y. Rui and T. S. Huang, "Image retrieval: Current techniques, promising directions and open issues," *J. Visual Commun. Image Represent.* 10,(1), Mar. 1999, pp. 39–62.
- [4] T. Ojala, M. Pietikainen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions," *Pattern Recogn.* 29,(1), Jan. 1996, pp. 51–59.
- [5] Xiaoyang Tan and Bill Triggs, "Enhanced Local Texture Feature Sets for Face Recognition under Difficult Lighting Conditions", *IEEE Transactions on Image Processing*, 19, (6), 2010, pp. 1635–1650.
- [6] Z. Lei, S. Liao, M. Pietikainen, and S. Z. Li, "Face recognition by exploring information jointly in space, scale and orientation," *IEEE Trans. Image Process.*, 20, (1), Jan. 2011 pp. 247–256.
- [7] Bongjin Jun, Daijin Kim, "Robust face detection using local gradient patterns and evidence accumulation", *Pattern Recognition*, 45, 2012, pp. 3304–3316.
- [8] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns", *IEEE Trans. Pattern Anal. Mach. Intell.*, 24,(7), Jul. 2002, pp. 971–987.
- [9] Subrahmanyam M, Maheshwari R. P, Balasubramanian R, "Local Maximum Edge Binary Patterns: A New Descriptor for Image Retrieval and Object Tracking", *Signal Processing*, 2012, pp. 1467–1479.
- [10] S. Murala, R. P. Maheshwari, and R. Balasubramanian, "Local tetra patterns: a new feature descriptor for content-based image retrieval," *IEEE Transactions on Image Processing*, 21, 5, 2012, pp. 2874–2886.
- [11] Subrahmanyam Murala · R. P. Maheshwari · R. Balasubramanian, "Directional local extrema patterns: a new descriptor for content based image retrieval", *International Journal of Multimedia Information Retrieval*, 1, (3), Mar. 2012, pp. 191 – 203.
- [12] Santhosh Kumar Vipparthi, S. K. Sagar, "Color Directional Local Quinary Patterns for Content Based Indexing and Retrieval", *Human-centric Computing and Information Sciences*, May. 2014, pp. 1 – 13.
- [13] Golla Varaprasad, Sachin Bharadwaj Sunda Murthy, "Detection of potholes in autonomous vehicle", *IET Intelligent Transport Systems*, 2013, pp. 543 – 549.
- [14] S. Viswanadha Raju, J. Sreedhar, "Query Processing for Content Based Image Retrieval", *International Journal of Soft Computing and Engineering*, 1,(5), November 2011, pp. 122 – 131.
- [15] Shiv Ram Dubey, Satish Kumar Singh, Rajat Kumar Singh, " Multichannel Decoded Local Binary Patterns for Content-Based Image

- Retrieval", *IEEE Transactions on Image Processing*, 25, (9), Sep. 2016, pp. 4018 – 4032.
- [16] Md. Mostafijur Rahman*, Shanto Rahman, "DTCTH: a discriminative local pattern descriptor for image classification", *EURASIP Journal on Image and Video Processing*, 30, 2017.
- [17] Anima Majumder, Laxmidhar Behera, "Automatic Facial Expression Recognition System Using Deep Network-Based Data Fusion", 48, (1), 2018, pp. 103 – 114.
- [18] Byungyong Ryu; Adín Ramírez Rivera, "Local Directional Ternary Pattern for Facial Expression Recognition", *IEEE Transactions on Image Processing*, 26, (12), 2017, pp. 6006 – 6018.
- [19] Md. Zia Uddin; Weria Khaksar; Jim Torresen, "Facial Expression Recognition Using Salient Features and Convolutional Neural Network", *IEEE Access*, Vol.5, 2017, pp. 26146 – 26161.
- [20] Maryam Imani, Hassan Ghassemian, "GLCM, Gabor, and morphology profiles fusion for hyperspectral image classification", *Electrical Engineering (ICEE), 2016 24th Proc. Int. Conf. IEEE, Shiraz, Iran*, 10-12 May 2016, pp. 1-5.
- [21] Mahamed Hassan, Mohammed Hossny, "Skin lesion segmentation using Gray Level Co-occurrence Matrix", *Systems, Man, and Cybernetics (SMC), 2016 IEEE International Conference on, 9-12 Oct. 2016, Budapest, Hungary*, pp. 820-825.
- [22] Anvita Bajpai, Kunal Chadha, "Real-time Facial Emotion Detection using Support Vector Machines", *International Journal of Advanced Computer Science and Applications*, 2010, 1, (2), pp. 37 – 40.
- [23] Hongzan Jiao, Yanfei Zhong, "An Unsupervised Spectral Matching Classifier Based on Artificial DNA Computing for Hyperspectral Remote Sensing Imagery", *IEEE Transactions on Geoscience and Remote Sensing*, 2014, 52, (8), pp. 4524 – 4538.
- [24] Irene Kotsia and Ioannis Pitas, Senior Member, "Facial Expression Recognition in Image Sequences Using Geometric Deformation Features and Support Vector Machines", *IEEE Transactions on Image Processing*, 2007, 16, (1), pp. 172 – 187.