A Cascade Face Recognition Approach with Illumination Compensation Using Local Features

Guilherme Felippe Plichoski, Leyza Dorini[1], Heitor Silvério Lopes[1], Chidambaram Chidambaram
Departamento de Sistemas de Informação – São Bento do Sul
Departamento de Informática – Curitiba[1]
Universidade do Estado de Santa Catarina – UDESC
Universidade Tecnológica Federal do Paraná – UTFPR[1]
Brasil
gui.sbs@gmail.com, leyza@utfpr.edu.br, hslopes@utfpr.edu.br, chidambaram@udesc.br

***Abstract*— Face Recognition (FR) has made significant progress in last two decades arriving at high rates of recognition. However, in FR systems, in addition to the factors such as pose, focus and expression, uncontrolled lighting conditions remains as one of the major challenges. Hence, in this work, we tried to investigate this issue using the supervised cascade approach that combines two different features extraction techniques. In the cascade approach, image features are extracted using the Local Binary Patterns (LBP) and the Edge Histogram Descriptor (EHD). These are applied in face images obtained under two different illumination conditions. Aiming to compensate the illumination variation, the images are pre-processed using the Phase Congruency (PC) and the Histogram Equalization (HE) techniques. This approach is evaluated using mainly the local features which are extracted from sub-images. The results obtained show that the cascade approach can improve effectively the recognition rates. The experiment with LBP as a first stage produced the best rates of recognition. In addition to this, among the results obtained from different combination of techniques, the images pre-processed with HE achieved 95.65% of recognition.**

Keywords—Face Recognition; Cascade Approach; Illumination Variation; Phase Congruency; Edge Histogram Descriptor

I. INTRODUCTION

Facial recognition (FR) still a subject of wide research due to persistent challenges that should be overcome for practical face recognition systems. FR can be applied in security and, surveillance areas, biometric identification, human-computer interaction and computer animation. Recognition process can be performed through many ways. However, in most cases, the basic steps consist in features extraction and comparison of images. Although FR has advanced in the recent years, challenges that require specific investigation such as facial expressions, occlusion, scale and illumination variation are difficult to overcome. In real-world applications, the lighting conditions cannot be controlled. Additionally the physics of illumination setup is difficult to model and recover [1]. Hence, it becomes one of the major challenges in the FR process, for example, the surveillance cameras that capture images usually from a crowed public places, such as an airport or a mall. In order to compensate the illumination and to extract the adequate features, the images can go through a pre-processing step, which may contribute to achieve better results. Many researchers have tested different image filters, feature extraction and comparison methods trying to compensate this influence [1]. Based on this context, we have attempted to compensate the variation illumination using images obtained under two different lighting conditions.

Several attempts have been made applying image filters and normalization techniques to compensate the illumination. Arandjelovic and Cipolla [1] proposed a framework to match face sets or video sequences based on a set of simple image processing filters, for example, high-pass filter, directional derivatives and Laplacian-of-Gaussian filters. Likewise, the authors [8] applied homomorphic filter to images in spatial domain to reduce the effects of illumination. In addition to this work, the authors Fan and Zhang [2] also worked with homomorphic filter to compensate illumination variation by improving the contrast through Histogram Equalization (HE). Though there are many works discussing about the illumination variation, we still believe strongly that developing new approaches in different directions is effectively a contribution to FR research works.

Under different illumination conditions, image normalization becomes a central task in FR. Illumination compensation becomes one of the important tasks before feature extraction and matching to improve the FR rate. The main issue discussed in this work can be approached through three steps: pre-processing and normalization, face representation and invariant feature extraction. One of the image pre-processing and normalization techniques is HE, which is stable under different lighting conditions [10]. HE is a method used to enhance contrast in images by extending the intensity range. It implies in mapping the pixels intensity distribution to a wider distribution uniformly [20]. Another technique that we have attempted is PC, which provides a significant measure of visual features of an image, and it is characterized by its invariance to brightness and contrast [9]. Hence, it provides an absolute measure of the significance of feature points [11]. Within the scope of cascade approach, the way the both techniques, HE and PC, are tested is also a contribution of the present work.

To construct robust FR algorithms, in addition to compensation of illumination, extraction of invariant features through different face representation methods becomes one of the main tasks. In this context, LBP and EHD are selected to extract the features from pre-processed face images. LBP is a texture operator that labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number. Its robustness to gray-scale changes makes LBP an operator that compensates some of illumination variation [3]. On the other side, EHD is a descriptor that characterizes image texture. It is useful even when the underlying region is not homogeneous in texture properties [21]. Additionally it reduces the dimensionality of the features representation.

One of the main contributions of this work is the cascade approach to conduct the FR process. It consists of two stages and in each stage, one of the following methods will be used: Local Binary Patterns (LBP) and Edge Histogram Descriptor (EHD). According the literature survey, we can realize that no individual method can obtain high rate of recognition for all types of face images. In other words, construction of robust FR approaches becomes difficult using a single method. In this sense, one alternative way is to develop FR using different feature extraction methods, which may have capacity to extract different features from images obtained under varying conditions. The cascade approach was initially discussed in the classical work developed by Viola and Jones [12] through a machine learning approach for visual object detection. In this work they applied a method by combining increasingly more complex classifiers in a cascade. The same author also published another paper [13] using the same technique for real-time applications. More recently, Li and colleagues [7] proposed in their work a cascade architecture built on conventional neural networks with powerful discriminative capability for face detection. Likewise, Zhang and Guo [19] presented another work using the cascade approach for both still and video-based face detection.

In the pre-processing step, in addition to RGB images, other two classes of images will be created by applying the PC and the HE independently, before the feature extraction stage of cascade approach. Though the techniques used in this work are already explored in different applications, to the best of our knowledge, no similar work can be found in the literature.

In many applications involving images or videos, the region of image from where the features are extracted becomes important. Generally extracted features can be treated as global or local. Global features describe the image as a whole meanwhile local features represent a small part of the image. Likewise, global features have the potential of generalizing an entire image, which may provide some guidelines for class discrimination purposes [14]. Similarly, local features can somehow improve the robustness of recognition systems. For example, the authors [22] combines the traits of global features in first stage and the local features in second stage to construct a robust FR system. In EHD, the technique itself extracts features in different forms [21]. In the case of LBP, it is common extracting features globally.

This paper is organized as follows: in Section II we explain how LBP works and how it has been used in FR process. Likewise, in Section III, EHD is addressed. In the subsequent Sections IV and V, we discuss about PC and HE respectively. The methodology for recognition process applied in this work is exposed in Section VI. The conclusions and future works are discussed in Section VII.

II. Local Binary Patterns

LBP is a descriptor that reduces efficiently local structures in images. In last years, it has been used in several fields of study like image processing and computer vision. In recent years, it has been exploited in many applications, for example, face image analysis. It is worth pointing out that the LBP is tolerance regarding monotonic illumination changes and simple to compute [3]. In fact, the LBP original operator label the pixels of an image with decimal numbers called LBP codes. To do this, each pixel is compared with its eight neighbors in a neighborhood 3x3 subtracting by the value of center pixel. The negative results are replaced with 0 and positive results, with 1. The labeling is done by concatenating the binary digits clockwise, starting from the upper left pixel. This process is shown in "Fig. 1" [21].

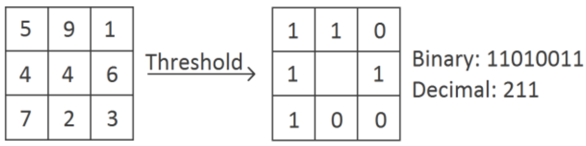

Fig. 1. Basic LBP operator process.

LBP features are based on the texture of an image. Approaches of face analysis based on LBP has been widely studied and become one of most popular applications in recent years. Following the basic and original procedure, several other variations of LBP has been created to improve the performance in different applications. Among these works, we can mention a LBP variation called multi-resolution uniform local Gabor binary patterns (MULGBP) [8] with homomorphic filtering in spatial domain for histogram extraction. These histograms are combined in order to obtain the average and then to be compared with a database of histograms. The results show that the approach applied to FR is robust in images with illumination variation. Yang and Chen [4] proposed another work that used LBP operator. To improve the performance of FR, they proposed a framework that is focused on two aspects: image pre-processing and sub-space representation. In pre-processing step the goal was to eliminate the effects of illumination. Hence they proposed a novel method based on invariant rotation LBP and gradient direction. The experiments show that the results are equal to state-of-art methodologies.

III. Edge Histogram Descriptor

Edge Histogram Descriptor (EHD) technique captures the spatial distribution of the edges in images [21]. This provides a set of standard tools that describes an image. Edge histograms for an image are constructed by combining sub-images into three different types of histograms: Local (L-EHD), Semi-Global (SG-EHD) and Global (G-EHD) edge histograms. It consists in dividing an image in 4x4 sub-images, as shown in “Fig. 2”, and for each sub-image, EHD is extracted. More generally, according to the MPEG-7 Color and Texture Core Experiments [21], the L-EHD represents the spatial distribution of different types of edges. The other two descriptors, G-EHD and SG-EHD are just obtained combining of histogram of local structures [17].

In MPEG-7 standard, edges are grouped into five categories vertical, horizontal, 45 diagonal, 135 diagonal and non-directional edge, as shown in “Fig. 3”. To extract these five edges, each sub-image is further divided into non-overlapping square blocks of 2 x 2 pixels. In each of these blocks, five edge-oriented detectors (filters) are applied to compute the edge strength, as shown in “Fig. 4”. Thus, the edge strength of each image block is calculated to determine the type of edge. In other words, the image is divided into image-blocks from which mean values are obtained by convolving with five filter coefficients to obtain edge magnitudes. If the maximum value for edge magnitudes is greater than a threshold value, it is assigned an edge type.

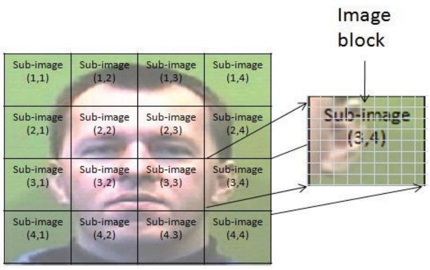

Fig. 2. 4x4 sub-images for EHD local features extraction.

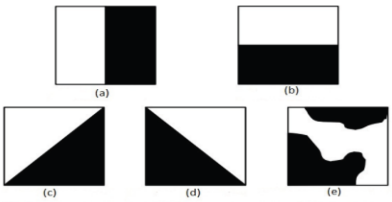

Fig. 3. (a) vertical (b) horizontal (c) 45-diagonal (d) 135-diagonal (e) non-directional edge.

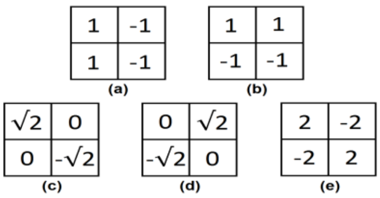

Fig. 4. Five filters for edge detection, (a) vertical, (b) horizontal, (c) 45-diagonal, (d) 135-diagonal and (e) non-directional edge.

The G-EHD represents the edge distribution for whole image space which is obtained by combining edge descriptors of all sub-images. For SG-EHD, four connected sub-images are grouped as shown in “Fig. 5”. Consequently, 13 different clusters can be obtained and for each one, five different edge distribution will be formed.

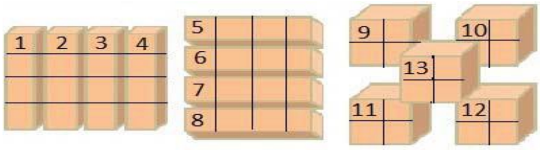

Fig. 5. Semi-Global (SG-EHD) sub-images.

The extraction of facial features using EHD is effectively applied in many works. For example, in FR process proposed by [17], PCA (Principal Component Analysis) feature extraction method based on a semi-supervised technique using EHD was discussed. In this work, the authors mentioned that EHD are not only useful in FR but also reduces considerably the dimensionality of the features of the image and computational complexity. On the other side, the authors [18] proposed another work using EHD. They performed eye detection in facial images compensating illumination effects and obtained results with high accuracy and low computational complexity.

IV. Phase Congruency

Phase Congruency (PC) technique provides a significant measure of visual features of an image. It is performed based on [9] Local Energy Model. In papers related to FR, PC is not common, although this technique is characterized by its invariant capacity to brightness and contrast. The Local Energy Model is based on the concept that the salient image features are represented by points in an image whose fourier components are maximally in phase. The article written by [15] proposes a strategy of feature selection with PC to enhance recognition rate in facial images affected by illumination, partial occlusion and expressions variation. To construct a robust FR invariant to illumination, the classification technique is carried out on the phase congruency maps of face images. The results present a significant increase in algorithm classification performance. PC technique was also used by [5]. The authors proposed a FR method based on the combination of Gabor wavelets and PC which is used to detect salient features in an efficient way. At first, the authors obtained image features from local frequency using Gabor filters, and then, PC technique. In this work, using ORL database, they obtained 98% of recognition rate.

V. Recognition process

FR still a subject of widely research due to persistent challenges that need to be overcome for practical face recognition systems. Nowadays, according to type of application, face images are generally obtained from different places which in fact lead to the problem of uncontrolled lighting conditions. As discussed in the previous sections, it is not an easy task recognizing all type of images using only one technique. Hence, we propose the cascade approach focusing on the illumination variation issue. In this section, the methodology is briefly described. The way the recognition process is conducted becomes one of the main contributions of the present work.

The supervised cascade approach is illustrated in “Fig. 6”. It involves three main steps: preprocessing, face-representation, feature extraction and recognition. At pre-processing stage, PC and HE techniques are independently applied to compensate the illumination in all images. In the next stage, one of the face representation techniques is applied followed by feature extraction. During the FR process, one face image (query image) will be searched in the base images. This process is performed comparing the query image with base images one by one which results in a face image with maximum similarity.

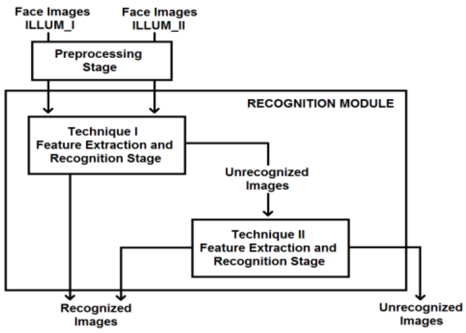

Fig. 6. Supervised Cascade Recognition Process

As discussed in the previous Section IV, the feature extraction technique can be LBP or EHD at each stage. Using EHD, only local features are extracted meanwhile, from LBP, both local and global features are extracted for the comparison purposes. Similar to EHD, the LBP using local features (L-LBP) is a variation of LBP in which features are extracted from face image that is divided into 4 x 4 sub-images. During the recognition step, a match with maximum similarity is selected as the best face image among all comparisons.

To evaluate the proposed approach and to check whether the classified image is same as the query image, the final decision of recognition is done in a supervised way by comparing the face codes. According to the proposed approach, only unrecognized images obtained from the first stage are sent to the second stage. The matching procedure is based on Euclidean distance measure. The classification process in this approach is done using the K-NN (k-nearest neighbor) algorithm in which k is equal to1.

VI. EXPERIMENTS AND RESULTS

FR research has advanced in the last years, but still a subject of research due to persistent challenges that need to be overcome such as illumination variation. Our work attempts to compensate this variation with techniques PC and HE in the pre-processing step and extracting features using LBP and EHD for comparison and matching. The experiments were conducted in a PC-Intel Core i5 with 6 GB memory and under platform Linux.

The sample of pre-processed images are shown in “Fig. 7”. In addition to the original RGB image, two more categories are obtained using PC (RGB-PC) and HE (RGB-HE).

Fig. 7. (a) RGB, (b) RGB-PC, (c) RGB-HE.

To discuss the focus of this work, we used [6] the database, which contains 138 face images of different individuals. This database was initially constructed under two different lighting conditions: controlled illumination (ILLUM_I) and uncontrolled illumination (ILLUM_II). In ILLUM_I, face images were captured under specific lighting system mounted for this purpose meanwhile the images of ILLUM_II were obtained under natural lighting conditions of the room in which the acquisition process was conducted. After image acquisition, face normalization for size, orientation and illumination becomes a critical issue in FR systems. Hence, following the international rules, all images were normalized and standardized in geometric coordinates through the eye detection process [14], except for illumination. The samples of cropped face images used in this work are shown in “Fig. 7”. The size of the images is 550 x 550 pixels and the base contains 138 face images of individuals.

Fig. 7. Examples of face images (first row – ILLUM-I, second row – ILLUM-II).

Another important step is the definition of edge strength as a threshold when EHD is used as a face representation technique. Hence, the threshold setting for experiments were done in order to determine the optimal values so that they can achieve high recognition rates. The main parameter tuned was edge strength threshold. An image block contributes to the histogram descriptor only if the maximum strength value of any edge type is greater than the threshold value. The value that yields the best recognition rates was determined based on a course to fine-tuning considering descriptors composed by features extracted from different types of images. From the fine-tuning analysis, the overall best recognition rate was obtained using the edge strength is about 3 [14]. Therefore, this value is maintained as a threshold in our experiments.

The whole set of experiments were conducted following the cascade approach as described in the Section V. Though the EHD technique proposes three type of features, in all experiments, we have attempted to work with global (EHD) as well as local features (L-EHD) that are extracted from the sub-images. Similarly, in the case of LBP, the both features, L-LBP and LBP, were extracted from the pre-processed face images. The results are shown in "Table 1". The first column shows the combination of the methods that are tested in the cascade approach meanwhile the other columns present the recognition rates under different types of images. For comparison purposes, the "Fig. 9 to Fig. 12" demonstrates the performance of each combination of cascade approach individually.

According to the experiments carried out in this work and the results shown in the "Table 1", among the proposed methods, LBP + L-EHD and L-LBP + L-EHD provide better results almost in all types of images. Among the three types of images, compensation through HE results in a high rate. RGB color space provides very low recognition. It can be interpreted as the direct influence of pre-processing step in the recognition process, more specifically, the capacity of HE to normalize the illumination variation created by different lighting conditions in ILLUM_II images.

TABLE I. FACE RECOGNITION RATES (%)

METHOD	TYPE OF IMAGES: RGB-HE	RGB-PC	RGB
LBP + L-EHD	95.65	71.73	63.76
L-EHD + LBP	90.57	68.84	62.31
L-LBP + L-EHD	94.92	71.01	65.21
L-EHD + L-LBP	90.57	68.84	63.76

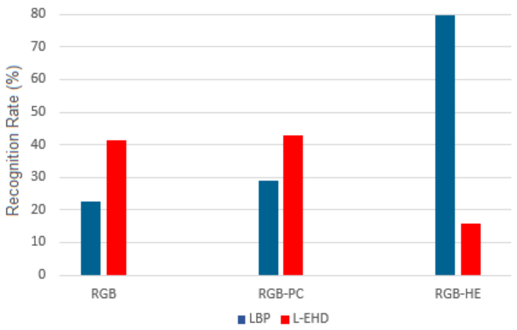

Fig. 9. Recognition rate – Cascade approach LBP + L-EHD.

The performance of each feature extraction technique can be seen in Figures 9 to 12. As shown in "Fig. 9 and Fig. 10", using images RGB and RGB-PC, L-EHD performs much better than LBP. However, this is not the same when we analyze the images RGB-HE. In this category, it is worth pointing out that the LBP (80%) has more discriminative power to recognize the images than L-EHD. From the cascade approach, it can be observed the discriminative power a face representation technique.

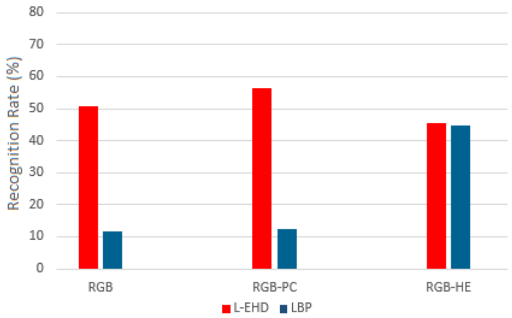

Fig. 10. Recognition rate – Cascade approach L-EHD + LBP.

For the comparison purposes, we have also done the experiments with LBP using the same concept of local features as L-EHD. The results are shown in "Fig. 11 and Fig. 12". In this category, we can observe that no significant improvement can be obtained using L-LBP approach in comparison with LBP. The performance is similar to the LBP (global).

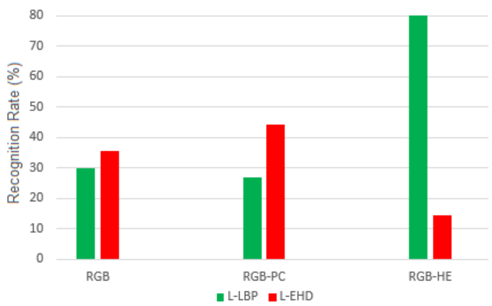

Fig. 11. Recognition rate – Cascade approach L-LBP + L-EHD.

The performance of the experiments demonstrates the limited discriminative power of edges using EHD in FR. Focusing on the cascade approach, the L-EHD, as a first stage, has the same performance with LBP and L-LBP (shown in "Fig. 10 and Fig. 12"). However, the recognition rate increases slightly in RGB-PC regarding RGB. On the other side, using RGB and RGB-PC, the performance of L-EHD is much better than the category RGB-HE.

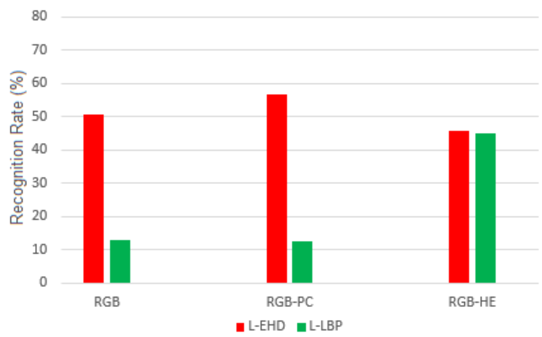

Fig. 12. Recognition rate – Cascade approach L-EHD + L-LBP

VII. Conclusion

In this paper, we proposed a supervised FR process based on cascade approach, using LBP and EHD feature extraction techniques, by pre-processing the face images with HE and PC techniques. We can conclude that the cascade approach technique provides a significant increase in recognition rate in all cases in combination with the pre-processing step. As discussed in the previous section, LBP + L-EHD combination provides the best results in all category of images. RGB color space alone is not adequate to be used in FR process when we work with different lighting conditions. Pre-processing the images certainly improves the recognition rates which can be seen in the results of this work. The RGB-PC provides a slight increase in the FR rates which stays around 8%. Among the three categories of images evaluated, RGB images pre-processed with HE technique yields the best rate achieving 95.65% using the database [6]. Another important observation can be drawn from the results is the discriminative power of local and global features in FR process. It is clear that the LBP with local features as well as global features achieve almost the same recognition rates. The L-EHD, as a first of stage of cascade approach, different to the LBP, performs equal in all categories of images. In other words, the L-EHD can recognize approximately 50% of face images independent of the category. Furthermore, since only one method is not alone able to recognize the images acquired under different conditions, it can be said that the cascade approach becomes useful to improve the recognition rates. In future works, other databases will be tested to evaluate the proposed approach. In the present work, only the basic LBP proposed by the authors was implemented. Another direction for future will be to test the improved LBP versions and other illumination compensation techniques found in the literature.

References

[1] O. Arandjelovic and R. Cipolla, “A methodoloy for rapid illumination-invariant face recognition using image processing filters”, Computer Vision and Image Understanding, Vol.113, 2009, pp.159-171.

[2] C.N. Fan and F. Zhang, "Homomorphic filtering based illumination normalization method for face recognition," Pattern Recognition Letters 32.10, 2011, pp. 1468-1479.

[3] D. Huang, C. Shan, M. Ardabilian, Y. Wang and L. Chen, “Local Binary Patterns and Its Application to Facial Image Analysis: A Survey,” IEEE Transactions on Systems, vol. 41, Issue 6, 2011, pp. 765 – 781.

[4] D. Yang and W. Chen, “An illumination insensitive framework using robust illumination normalization and spectral regression kernel discriminant analysis for face recognition,” Signal Processing, Communications and Computing (ICSPCC), IEEE International Conference, 2015, pp.1-5.

[5] E. Bezalel and U. Efron, "Efficient face recognition method using a combined phase congruency/Gabor wavelet technique," International Society for Optics and Photonics, Optics & Photonics, pp. 59081K-59081K, 2005.

[6] F. das Prodossimo, C. Chidambaram and H. S. Lopes, “Proposta de uma metodologia para construção de um banco de imagens faciais normalizadas,” In: VIII Workshop de Visão Computacional. [S.l.: s.n.], 2012.

[7] H. Li, Z. Lin, X. Shen, J. Brandt and G. Hua, "A convolutional neural network cascade for face detection," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 5325-5334.

[8] H. Shahamat and A.A. Pouyan, “Face recognition under large illumination variations using homomorphic filtering in spatial domain,” J. Visual Communication and Image Representation 25, 2014, pp. 970-977.

[9] K. Souza, B.C.G. da Silva, R.J. Ferrari and J. Carvalho-Tavares, “Automatic detection of leukocytes from intravital video microscopy using the phase congruency technique,” Medical & Biological Engineering & Computing, vol. 53, Issue 2, 2015, pp 151-163.

[10] X.Xie, J. Lai and W. Zheng, “ Extraction of Illumination invariant facial features from a single image using nonsubsampled contourlet transform”, Pattern Recognition, vol. 43, 2010, pp. 4177-4189.

[11] P. Kovesi, “Image features from phase congruency,” Journal of computer vision research, vol. 1, no. 3, 1999, pp. 2–26.

[12] P. Viola and M.J. Jones, “Rapid object detection using a boosted cascade of simple features,” The proceedings of the IEEE Computer Society Conference, Vol. 1, 2001, pp. I-511.

[13] P. Viola and M.J. Jones, "Robust real-time face detection," International journal of computer vision 57.2, 2004, pp. 137-154.

[14] C. Chidambaram, “A Contribution for Single and Multiple Faces Recognition using Feature-Based Approaches,” Ph.D. Thesis, Federal University of Technology-Paraná. 2013.

[15] S. Gundimada and V.K. Asari, “A Novel Neighborhood Defined Feature Selection on Phase Congruency Images for Recognition of Faces with Extreme Variations,” International Journal of Computer, Electrical, Automation, Control and Information Engineering vol. 1, no. 7, 2007, pp. 25-31.

[16] S. Gundimada and V.K. Asari, “Facial Recognition Using Multisensor Images Based on Localized Kernel Eigen Spaces,” IEEE Transactions on Image Processing, vol. 18 Issue 6, 2009, pp. 1314 – 1325.

[17] S. Rahman, S.M. Naim, A.A. Farooq and M.M. Islam, “Performance of PCA Based Semi-supervised Learning in Face Recognition Using MPEG-7 Edge Histogram Descriptor,” Journal of Multimedia, Vol 6, No 5, 2011, pp. 404-415.

[18] C. Jung, T. Sun and L. Jiao, “Eye detection under varying illumination using the retinex theory,” Neurocomputing 113, 2013, pp. 130-137.

[19] Ping Zhang and Xi Guo, “A cascade face recognition system using hybrid feature extraction”, Digital Signal Processing, vol. 22, 2012, pp. 987–993

[20] R. Gonzalez and R.Woods, Digital Image Processing, 3rd ed., Prentice Hall, 2009.

[21] B.Manjunath, J-R, Ohm, V.V, Vasudevan, and A. Yamada, “Color and Texture Descriptors”, IEEE Transactions on Circuits And Systems for Video Technology, Vol. 11, No. 6, 2001, pp.703-715.

[22] C. Singh, E. Walia, N. Mittal, “Robust two-stage face recognition approach using global and local features”. The Visual Computer, v. 28, No. 11, 2012, pp. 1085–1098.