Journal of Theoretical and Applied Information Technology
10th September 2013. Vol. 55 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

NORMALIZED CROSS-CORRELATION BASED FILTERING PROCESS FOR FACE LOCALIZATION

GHASSAN MARWAN ABDULFATTAH[1], MOHAMMAD NAZIR AHMAD[2]
[1]Faculty of Computing, 81310,Universiti Teknologi Malaysia
Skudai, Johor.
[2]Faculty of Computing, 81310,Universiti Teknologi Malaysia
Skudai, Johor.
E-mail: [1]g_h_assan@hotmail.com, [2]drmnazir09@gmail.com

ABSTRACT

In the real-time face recognition systems, locating faces in the input image should be fast and accurate. `Therefore, the need to introduce a face localization method with less complexity rises. Recently, Template-matching Approach has been used by different researchers to find the region of interest (face location) with different positions, lighting and cluttered backgrounds. However, some matching errors always occur with regard to the appearance of parts of the image as face location because of the variations of image pixels values. In this paper, faster and reliable pre-processing method is proposed to reduce the effects of the noise on image pixels values by applying a filtering process. The result showed an important and significant improvement in localization accuracy or precision compared to using the traditional NCC alone. Since this method was implemented, there has been an eleven percent (11%) increase in accuracy compared to the traditional NCC (80% localization accuracy). In this study, Yale University's database is used to evaluate the proposed method.

Keywords: *Image Processing, Face Localization, Normalized Cross Correlation*

1. INTRODUCTION

Face recognition has caught the attention of many researchers in the last few decades. Therefore, numerous research efforts have been made on a wide range of facial biometrics [17] to introduce reliable features for recognition. Before extracting these features, there is a need to determine the region of interest (ROI) which contains only the face because some parts of the input images are not needed for the recognition process. In the real-time face recognition systems, locating faces in the input image should be fast and accurate. Thus the need to introduce face localization method with less complexity rises. Essentially, face localization is considered the first step whenever it is a special case of face detection. In the face localization problem, there is an already existing face in the input image and the goal of which is to determine the location of the face. However, face localization from the input image is considered to be a challenging task due to variations in scale, pose, occlusion, illumination, facial expressions and clutter backgrounds. One of the important surveys on face recognition and some detection techniques can be found in [1]. Recently, an outstanding survey about the relevant literature of faces detection and localization was written by Yang et al. [2] the survey covered the problems of detection and clearly highlighted the research gap. Moreover, it introduced some solutions for these problems. Another critical survey can be found in [12].

In this paper, simple and fast technique is proposed to enhance the quality of the input image by using a smoothing filter which will reduce the noise and the effect of the image variations on the matching process and which will increase the accuracy of the localization. A normalized cross correlation (NCC) will be used to compute the similarity matching between the face template and rectangular blocks of the input image to locate the face.

The rest of this paper is organized as follows: in the following section we explain the related and previous work and highlight the research gaps in the localization problem. In Section 3, we introduce the proposed method. In Section 4, experimental results and comparisons with (the) recent methods will be presented; the last section is allocated for the conclusion and the recommendation for further studies.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2. RELATED WORK

With regard to the literature, face detection (localization) methods are classified into four main categories as follows: Knowledge-based method [3], Feature invariant method [4-13], Template matching method [5-15] and Appearance-based method [6-14]. Template matching approaches have been widely used for face localization because of its accuracy and appropriateness for real-time systems [15]. In such approaches, the first step is to create a template based on some human face criteria stored in the system database. Later, correlation between the stored template and the input image parts are computed to locate the faces. The advantages of these methods are that they are resistant to noise, simple to implement and do not take a long time to locate the candidate's faces from the input images [7]. However, it is not sufficient for locating faces in images with clutter background and illumination because some parts of the image appear to be face location regarding the variations of the image. One of the simplest methods of the template matching approach is to gather set faces samples to obtain an average face and then save them in the database as a template. Later, for the localization process, the template is called up and passed through the input image and the image block, the highest similarity correlation score is supposed to be the correct face location. From this method, the process can be called the filter match method where the input image constantly convolves or interacts with a flipped version of the average face as a filter. Statistically, filter matching assumes additive white Gaussian noise (AWGN) which is very bad for image variations such as clutter background, illumination and expressions [7]. To reduce the effect of high face variation problems, the Eigenface approach is adopted to enhance matched filter performance [8] which makes linear combination for Eigenfaces of the average face it also assumes that each face should be closed to this linear combination. Nevertheless, the Eigenface approach has problems which are reflected in the variations of the face and in the noise as well [9]. Due to this problem, there is always some localization error where non-face blocks may give high matching similarity to the linear combination of the average face and its Eigenfaces are more than the face blocks. Therefore, the Eigenfaces method can give a good detection rate when the noise is white noised clutter. Meng et al. [7], proposed a new method to localize the human faces using linear discriminants from gray scale images. To minimize the Bayesian error, they developed an optimal discriminant template by modeling faces and non-faces as Gaussian distribution. In addition, they compared their results with the matched filter and the Eigenfaces methods and it was 92.7% using the University of Michigan's face database. Some researchers used another method to locate the faces in the input images by using face anthropometric templates such as eye, mouth and nose. Campadelli et al [16] introduced a good survey on face locating using anthropometric template based on the eye as *a* reference.

One of the widely used methods to compute the correlation between two images is similarity measurements such as Normalized Cross Correlation (NCC) [10-11]. However, NCC still does not give good matching accuracy because the variation in illumination and clutter background sometimes produces miss-matching face localization. This problem can be eliminated by using a filtering process which changes the properties of the image and increases the similarity process which will reflect on the localization accuracy. Huaibin et al [20] proposed a new method based on anisotropic filter, for denoising and smoothing images which can also be used for edge detection, but it is still not that much effective for images with illumination and clutter background variations.

3. PROPOSED METHOD

As the first step in the proposed method, the template face will be created by gathering a set of faces from the users of the system together as in figure1. Two steps will follow, preprocessing and matching.

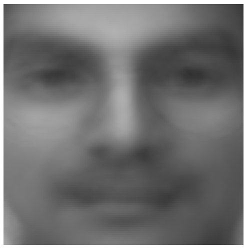

Figure 1: Template face

3.1. Preprocessing Step

Actually it is a low pass filter or, more precisely, it will be like a smoothing filter. Therefore, there is a need to understand the concept of this filter and

Journal of Theoretical and Applied Information Technology
10th September 2013. Vol. 55 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

describe how the mechanism works in order to use it *in* this step.

Generally, image filtering is done either for enhancing the image and detecting the edges through high frequencies or smoothing the image through the low frequencies. Image can be filtered in two domains namely; frequency domain and spatial domain. The operation of filtering starts by transforming the image to the frequency domain, then *multiplying* the transformed image with the frequency filter. The result will be transformed again to the spatial domain. Spatial domain operation can be defined as convolving the input image A with filter B, and it can be formulated by using the following equation:

$$C(i, j) = A(i, j)\Theta B(i, j) \qquad (1)$$

Practically, there is a difference between the mathematical operation and the multiplication result in the implementation, but the result remains the same in the frequency space. Therefore, there is a need to use a discrete or finite Kernel this kernel is to be passed over the input image and multiply the kernel elements with image pixels. For example, if there is a square Kernel with size L×L, the output image can be calculated by the following equation:

$$C(i, j) = \sum_{m=-\frac{M}{2}}^{\frac{M}{2}} \sum_{n=-\frac{M}{2}}^{\frac{M}{2}} A(m, n)B(i-m, j-n) \qquad (2)$$

Where M is the dimension of the filter matrix B.

A numbers of Kernels were used previously to determine the operations characteristics of the input image based on the form and the size of these kernels. One of the most important Kernels is mean filtering. it is used for smoothing which is the same in this case. It's a simple method for smoothing the image and easy to implement. Moreover, it is used to decrease the variation between the values of the image pixels and its neighbor pixels. It is very useful in reducing the noise.

To implement this kind of kernel, replace the values of the image pixels by the mean of neighbor's pixels. Thus, the process will reduce the noise and eliminate the variance between the values of the image pixels. The following matrix illustrates the mean Kernel.

$1/n^2$	$1/n^2$	$1/n^2$
$1/n^2$	$1/n^2$	$1/n^2$
$1/n^2$	$1/n^2$	$1/n^2$

Figure 2: n-Mean Kernel

In this case, the problem is not a noise one rather an illumination one. Therefore, there is a need to modify the smoothing filter reasonably to be used in face localization. This can be done by adding each pixel in the input image to the neighbor pixel, and this process increases the brightness of the image and reduces the difference between the image parts, but in the cost of the image dimensions, if we have the following image matrix with size 4×4:

5	1	4	3
2	0	6	5
1	1	2	8
9	6	3	7

Figure 3: Image matrix with size 4x4

The output matrix will be as follows:

8	11	18
4	9	21
17	12	20

Figure 4: Output matrix

Now, the increase in image pixels values of the output image compared to the original image and the new image becomes less by 1 is clear. To see this operation in real image, Figure 5 provides that:

A B

Figure 5: A is the original image and B is the filtered image

Journal of Theoretical and Applied Information Technology
10th September 2013. Vol. 55 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3.2. Matching Process

In this step the original and filtered images will be divided into sub-images (window) with the same size of the template face, and then NCC will be used to calculate the correlation between the template face image and the sub-images in the filtered image using the following equation:

$$c = \underline{\mu}^{\mathrm{T}} \frac{\underline{x} - d(\underline{x})}{\left\| \underline{x} - d(\underline{x}) \right\|} = (\underline{\mu} - d(\underline{\mu}))^{\mathrm{T}} \frac{\underline{x}}{\left\| \underline{x} - d(\underline{x}) \right\|} \quad (3)$$

Where, C gives the maximum probability of a face, $\underline{\mu}$ is a face template, $d\,\underline{\mu}$ its DC component, and $\underline{x}$ is any window of the filtered image. But, in this case, another equation proposed by [18] will be used to calculate the correlation coefficients which are simpler *easier* to implement. It's as follows:

$$R = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{\sum_m \sum_n (A_{mn} - \bar{A})^2 * \sum_m \sum_n (B_{mn} - \bar{B})^2}} \quad (4)$$

Where A is the template face and B is the sub-image window.

Based on the face location in the filtered image, the corresponding face location in the original image will be determined now by using block index. Figure 6 shows face localization using the proposed method as *an* example image from the Yale Database [19].

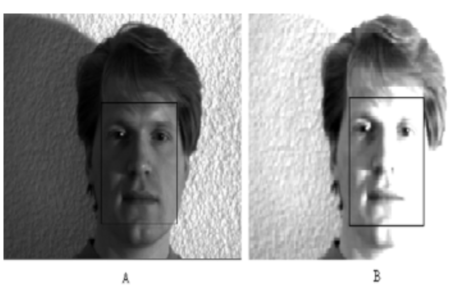

Figure 6: A the corresponding face location to the face location in B

4. RESULT AND DISCUSSION

The Yale database established by Yale University [19] was used to evaluate the proposed method. 11 images of 15 subjects with a total of 165 images were tested during the experiments. The images of each subject were in different cases such as: center-light, with/without glasses, happy, sad, left-light, with/without glasses, normal, right-light, sad, sleepy, surprised, winking and illuminated. Few examples of these images are shown in Figure 7:

Figure 7: Samples from Yale database

At the beginning, the template image was prepared and saved in the system database and the face template size used is *was* 100×120. Then the N-mean Kernel was used to increase the brightness of the input image by increasing the image pixels values which looks like the denoising process. After that, a dynamic window was created after the pre-processing step with a size similar to the template image and passed through the input image. For the matching process, NC was used to calculate the maximum correlation between the template image and the dynamic window in the input image, and then a recorded matrix was constructed to record correlations values. According to the maximum correlation in the recorded matrix, the position of the dynamic window is cropped: Table 1 presents the localization accuracy using a number of kernels. From the Table, we can notice the decrease in the localization accuracy when the dimension of the mean kernel is increased and this is due to the increase in brightness of the input image. Therefore, 2-mean kernel appears to be the optimum kernel compared to each other. Figure 8 shows some of the correct and incorrect faces locations on the Yale dataset using the proposed method.

Table 1: Result of using different N-mean kernels with NC

Kernel	2-mean	3-mean	4-mean	5-mean	6-mean
Result	91%	86%	73%	68%	56%

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

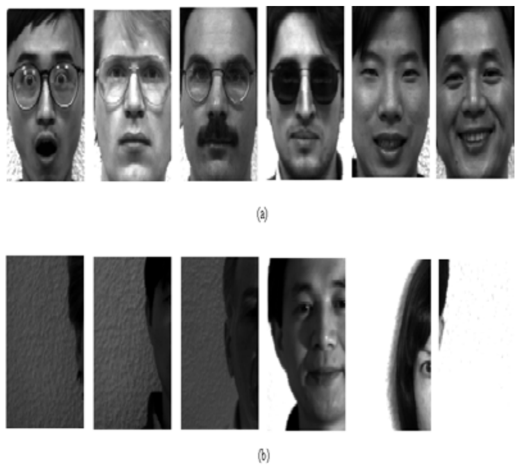

Figure 2: Samples of face localization: correct location in the top (a) and incorrect location in the bottom (b)

Table 2 shows the comparison on face localization in terms of result and time using the proposed method with the proposed methods using NC alone, method in [3], [7], [11], [15] and [20] on Yale dataset.

Table 2: Comparison between the proposed method and NC, LDA methods on Yale dataset

Method	Result (%)	Time (s)
NC	80	5
Rectangular Knowledge Rule and Face Structure	90	17
LDA	92.7	15
Normalized Cross Correlation With Adaptive Multilevel Winner Update	91	10
Particle Swarm Optimization (PSO)	97.4	30
Anisotropic filter + NNC	89	8
2-mean kernel + NC	91	8

In the above Table, the proposed method was compared to six of the existing methods on Yale database under the same conditions. From the table, we can notice that the proposed method gives good localization result in terms of time because it is less complex compared to the PSO method [11] and LDA method [7] as real time systems need a shorter time. On the other hand, the NCC method showed less localization time compared to the proposed method but in the cost of reliability. Moreover, the proposed method showed superiority compared to the methods in [11] and [20] where the problem was illumination although there were filtering methods. To sum up, face localization using the pre-process stage is preferred rather than using NCC only for illumination problems compared to other filtering methods.

5. CONCLUSION

In this paper a robust and simple technique is proposed to improve the performance of NCC as a similarity measurement for the template matching approach. This technique is based on increasing the pixels values of the input image by using n-means kernel. It will reduce the effects of illumination problems. The result showed the significant improvement in NCC performance by using the proposed technique. Further work will focus on developing another similarity measurement equation to increase the accuracy and reduce the complexity.

REFERENCES:

[1]. R. Chellappa, C. L. Wilson and S. Sirohey, "Human and machine recognition of faces: a survey," Proceedings of the IEEE, vol. 83, pp. 705-741, 1995.

[2]. Ming-Hsuan Yang, D. J. Kriegman and N. Ahuja, "Detecting faces in images: a survey," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 24, pp. 34-58, 2002.

[3]. Zhao Fei and Qiao Qiang, "Face detection based on rectangular knowledge rule and face structure," in Information Science and Engineering (ICISE), 2009 1st International Conference on, 2009, pp. 1235-1239.

[4]. S. Jeng, H. Y. M. Liao, C. C. Han, M. Y. Chern and Y. T. Liu, "Facial feature detection using geometrical face model: An efficient approach," Pattern Recognit, vol. 31, pp. 273-282, 3, 1998.

[5]. A. K. Jain, Y. Zhong and M. Dubuisson-Jolly, "Deformable template models: A review," Signal Process, vol. 71, pp. 109-129, 12/15, 1998.

[6]. M. -. Yang, N. Ahuja and D. Kriegman, "Face recognition using kernel eigenfaces," in Image Processing, 2000. Proceedings. 2000 International Conference on, 2000, pp. 37-40 vol.1.

[7]. Lingmin Meng and T. Q. Nguyen, "Frontal face localization using linear discriminant," in Signals, Systems, and Computers, 1999. Conference Record of the Thirty-Third

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Asilomar Conference on, 1999, pp. 745-749 vol.1.

[8]. A. Pentland, B. Moghaddam and T. Starner, "View-based and modular eigenspaces for face recognition," in Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94, 1994 IEEE Computer Society Conference on, 1994, pp. 84-91.

[9]. Lingmin.Meng and Truong.Q.Nguyen, "Frontal face detection using multi-segment and wavelet," in 1999.

[10]. D. Tsai and C. Lin, "Fast normalized cross correlation for defect detection," Pattern Recog. Lett., vol. 24, pp. 2625-2631, 11, 2003.

[11]. Shou-Der Wei and Shang-Hong Lai, "Fast Template Matching Based on Normalized Cross Correlation With Adaptive Multilevel Winner Update," Image Processing, IEEE Transactions on, vol. 17, pp. 2227-2235, 2008.

[12]. J. Ddaugman, "Face detection: a survey," Journal of computer vision and image understanding, vol. 83, no.3, pp. 236-274, 2001.

[13]. Jing-Wein Wang, "Face Localization in Clutter Background," in Intelligent Hiding and Multimedia signal processing, 2008 International Conference on, 2008, pp. 1235-1239.

[14]. S. M. Hassan, Wei-Yun, E. K. Teoh," Fast Face Detection and Localization from Multi-Views using Statistical Approach" in Information Science and Engineering (ICISE), 2011 International Conference on, 2011, pp. 1235-1239.

[15]. A. Perez, M. Carlos, I. Juan, A. Pablo, M. Claudio," Face and Iris Localization using Templates Designed by Particle Swarm Optimization" Pattern Recognition Letters, vol. 31, pp. 857-868, 2010.

[16]. P. Campadelli, R. Lanzarotti," Eye Localization: A Survey," The Fundamentals of Verbal and Nonverbal Communication and their Biometrical Issues, Vol.18, pp. 234-245, 2007.

[17]. W. Zhao, R. Chellappa, A. Rosenfeld and P. J. Phillips, "Face Recognition: A Literature Survey," vol. 35, pp. 399-458, December, 2003.

[18]. R. c. Gonzalez, R. E. Woods,"Digital Image Processing," Second ed, prentice Hall, 2001.

[19]. Yale Face Database, http://cvc.yale.edu/projects/yalefaces/yalefaces.html vol. 2010.

[20]. H. Wang, Y. Wang, W. Ren," Image Denoising Using Anisotropic Second and Fourth Order Diffusions Based on Gradient Vector Convolution," Computer Science and Information Systems, Vol. 9, pp. 1493-1512, December, 2012.