N. Siddartha* et al.
(IJITR) INTERNATIONAL JOURNAL OF INNOVATIVE TECHNOLOGY AND RESEARCH
Volume No.4, Issue No.5, August – September 2016, 3844–3847.

Automatic Face Recognition Based On Learning to Rank for Image Quality Assessment

Mr. N. SIDDARTHA
Student,M.Tech,
Department of ECE,
Sri Chaitanya Institute of Technological Sciences
Karimnagar India.

Mr. K.MAHESH
Assistant professor
Department of ECE,
Sri Chaitanya Institute of Technological Sciences
Karimnagar India.

Abstract: **Automatic face recognition technology has attracted a great amount of attention from both academia and industry in the recent trends. It is usually possible for practical recognition systems to capture multiple face images from each subject. Selecting face images with high quality for recognition is a promising stratagem for improving the system performance. We propose a simple and flexible framework for face image quality assessment, in which multiple feature fusion and learning to rank are used. The proposed method is simple and can adapt to different recognition methods. To demonstrate the overall effectiveness of the proposed method, we use heuristic criteria for data selection in our experiments.**

Keywords: **Face Quality; Automatic Face Recognition; Learning To Rank**

I. INTRODUCTION

Many researchers have been proposed for identifying or verifying personal identities based on face images [1]–[3]. However, the effectiveness of automatic face recognition is challenged by variations in illumination, pose, occlusion and expression in the captured face images [4] largely because the face image acquisition process is non-contact in nature. Such problems become even more serious in real applications with uncooperative users and uncontrolled environments. Human face is believed to be an ideal biometric feature for personal authentication because it is universal, discriminative, non-intrusive, and easy to obtain.

During the past two decades, automatic face recognition technology has attracted a great amount of attention from both academia and industry. Although many approaches have been proposed for improving the robustness of face recognition against different kinds of face image quality degradation [5]–[7], it is still widely understood that most face recognition methods achieve better performance on high quality face images [8]. Take face verification vendor tests for example. In the Multiple Biometrics Evaluation (MBE) organized by NIST in year 2010, on a face database consists of high quality visa photo images, the lowest error rate reported was 0.3% (False Rejection Rate at False Acceptance Rate = 0.001) [9]. However, on the LFW database [10] made up of wild face images collected from the web, the latest reported result indicates a corresponding error rate of no less than 18% [11], which is nearly two orders of magnitude worse than that in MBE.

In many practical video based face recognition systems, it is actually possible to acquire multiple face images from the target subjects. Selecting high quality face images for recognition can not only improve the system robustness and suppress false alarms, but also reduce the overall computation load considering that face feature extraction is usually complex. Berrani and Garcia were among the first to study this problem and proposed to use robust PCA for removing low quality face images as outliers [12]. This method, however, cannot be applied in situations like video surveillance in which low quality face images dominate.

A more straightforward approach to solve this problem is face image quality assessment, of which most existing methods are based on the analysis of specific facial properties. Yang *et al.* adopted a tree structure for pose estimation and used the results for evaluating face image quality [13]. Gao *et al.* proposed to use the degree of facial asymmetry to quantify the face quality degradation caused by non-frontal illumination and poses [14]. Sellahewa *et al.* directly used the universal image quality index [15][16] for measuring the face image quality in terms of illumination distortion in comparison to a specified reference face image [17].Wong *et al.* proposed a patch based probabilistic model for quality assessment trained on reference face images with frontal poses, uniform illumination and neutral expressions [18]. However, the effectiveness of these methods are limited by the applicability of the artificially defined facial properties and empirically selected reference face images. To solve this problem, we propose a simple and flexible framework for face image quality assessment, in which multiple feature fusion and learning to rank are used.

II. SYSTEM DESIGN MODEL

A big difference between ordinary data privacy and video data privacy is the amorphous nature of the latter, and the difficulty in processing it automatically to extract useful information. A video clip can convey negligible amounts of information or may contain very detailed and specific information. Privacy is hard to define, even for explicit textual information such as name, address and social security number fields in a database, knowledge of which can be used for identity theft, fraud and the mining of copious information about the individual from other databases.

III. FACE NORMALIZATION

Ideally, only image pixels inside the human face should be used for assessing face quality. This can be realized, for example, by locating contour landmarks and generating a specific mask for each face in the image. However, this can be time consuming and may cause difficulties in subsequent feature extraction due to shape irregularity.

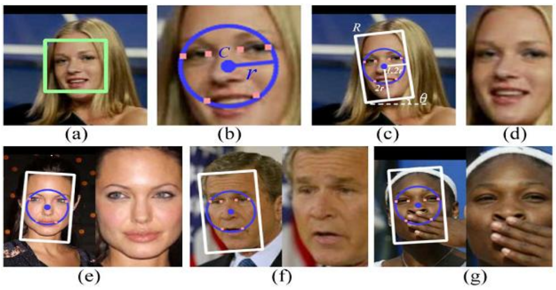

Fig. 1. Face normalization using smallest enclosing circle.

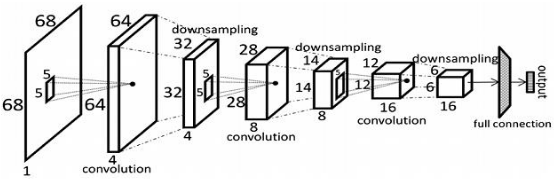

Fig. 2. Structure of the CNN for landmark location.

On the other hand, most face detectors [19] simply output square bounding boxes which obviously deviate from human face shapes and may contain a considerable amount of non-face information. In addition, in-plane rotation of faces should not be treated as quality degradation since most face recognition systems are able to handle it properly [4]. Based on all these considerations, we propose the face normalization process shown in Fig. 1. Fig. 1(a) shows the face detection [19] result on an image from the LFW database. The detected face area is then resized to pixels and passed to a CNN (Convolutional Neural Network) for landmark location [20].We use the eye/mouth corners because these landmarks are clearly defined and covermost face regions. Fig. 2 shows the structure of the CNN which contains three convolution layers and three downsampling layers.

Altogether 164 convolution kernels of size are used and the output of the network is the vector form of the landmark coordinates. We randomly select 10000 images from LFW for training and the remaining images are used for testing. The average landmark location error is 1.4 pixels on the test set. Fig. 1(b) shows the located landmarks. To normalize the face area and eliminate in-plane rotation, we first calculate the center and radius of the smallest circle that enclose all the landmarks using the linear time algorithm proposed in [21].We then place a rectangle of size centered at as is shown in Fig. 1(c). Obviously, all the landmarks are guaranteed to be enclosed by . Suppose the coordinates of the four eye corners are , the orientation of the rectangle can be determined by equation (1), in which and are the mean values of the horizontal and vertical coordinates respectively. Thus, the shorter side of is parallel to the line that best fits the four eye corners. The normalized face area in Fig. 1(d) can thus be achieved by rotating the rectangular area inside around by angle,

$$\theta = \arctan\left(\left(\sum_{i=1}^{4} x_i y_i - 4\bar{x}\bar{y}\right)\right) / \left(\sum_{i=1}^{4} {x_i}^2 - 4\bar{x}^2\right)\Bigg) \quad (1)$$

More face normalization results on LFW images are shown in Fig. 1(e), (f), (g). It can be observed that the normalized faces are compact and guaranteed to contain main facial parts. The normalized faces are then used as inputs to the face quality assessment process to be introduced in the next section. The proposed normalization method is somewhat robust to inaccuracy in landmark location. Nevertheless, in case that multiple landmarks are significantly incorrectly located simultaneously, the normalization result may deteriorate and lead to a low face quality assessing result. This problem, however, can be tolerated in our work considering that such a situation, for most cases, does indicate very low face quality.

IV. FACE QUALITY ASSESSMENT

It is in general difficult to explicitly define and quantify the quality of a face image. There have been mainly two approaches for solving this problem in previous research. The first one is to empirically use certain facial properties, such as the resolution, pose angle, or illumination parameters, to quantify face image quality [13][14]. The second one is to compare a face image to selected 'standard' faces and use their discrepancies for measuring face quality [17][18]. Both approaches are inflexible and lack of applicability since neither of them has taken into account the possible differences among face recognition methods. For a face recognition algorithm good at solving the occlusion problem [7], Fig. 1(g) is probably more preferable than Fig 1(f). On the contrary, for a recognition method in which poses can be properly handled [6], Fig. 1(f) should be considered of higher quality. Also, face image quality should be considered in a relative manner. For most recognition methods, Fig. 1(d) is better than Fig. 1(f) but worse than Fig 1(e) in terms of face quality. Based on the above considerations, we propose a simple and flexible face quality assessment approach based on learning to rank [22]. Suppose a face recognition method is tested on two different face databases and ; and the recognition performance on is better than it is on .

This indicates that for this specific recognition method, face images in have higher quality than those in . We note this as . Let and be two images selected from and respectively; and let be the function that transform a face image to a feature vector. Define a linear form quality assessment function ,and our goal is to find the value of rank weight that satisfies as many constraints in equation (2) as possible. Also, images in the same face database should be considered of similar face quality. This can be expressed by the equality constraints in equations (3) and (4). Considering the ranking nature of this formulation, we name the value of as the RQS (Rank based Quality Score) of .I

$$w^T f(I_i) > w^T f(I_j); \quad \forall I_i \in A, \forall I_j \in B$$

$$w^T f(I_i) = w^T f(I_j); \quad \forall I_i \in A, \forall I_j \in A$$
$$w^T f(I_i) = w^T f(I_j); \quad \forall I_i \in B, \forall I_j \in B$$

The above problem formulation is identical to that in [23] and thus can be transformed into a convex max-margin formulation shown in equation (5) by introducing non-negative slack variables. , and are constants balancing the degree of slackness allowed by the corresponding constraints. The primal optimization problem defined by equation (5) can be efficiently solved using Newton's method. The proposed formulation can be extended to multiple databases and features. For multiple feature fusion, we use a two level learning stratagem.

V. SIMULATION RESULTS

To demonstrate the overall effectiveness of the proposed method, we use heuristic criteria for data selection in our experiments. Three sets of data and are prepared. Consists of face images selected from face databases collected in controlled environments, such as FERET, FRGC and a Chinese ID card photo database in our laboratory. Consists of face images selected from two real world face databases: LFW [10] and AFLW. Consists of non-face natural images in which the face detector [19] generates false positive detection results. Each dataset contains 6000 images and among which 5000 are used for training and 1000 are used for testing.

VI. CONCLUSION

We propose to apply the proposed method to a specific face recognition system; we suggest selecting the training datasets accordingly for achieving better performance. To formulate the face image quality assessing problem in a relative manner and use learning to rank for solving it. For practical systems, the proposed RQS value can be used for improving face detection robustness, controlling the face quality in registration and selecting high quality images for recognition. It is also possible to use RQS for evaluating the confidence of different face images in multi-instance face recognition

VII. REFERENCES

[1] M. A. Turk and A. P. Pentland, "Eigenfaces for recognition," *J. Cognitive Neurosci.*, vol. 3, no. 1, pp. 71–86, 1991.

[2] P. N. Belhumeur, J. P. Hespanha, and D. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 2005.

[3] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, Dec. 2006.

[4] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfield, "Face recognition: A literature survey," *ACM Comput. Surv.*, vol. 35, no. 4, pp. 399–458, 2003.

[5] L. Wiskott, J.-M. Fellous, N. Kuiger, and C. Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 775–779, Jul. 1997.

[6] V. Blanz and T. Vetter, "Face recognition based on fitting a 3D morphable model," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 9, pp. 1063–1074, Sep. 2003.

[7] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[8] P. Grother and E. Tabassi, “Performance of biometric quality measures,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 4, pp. 531–543, Apr. 2007.

[9] P. J. Grother, G. W. Quinn, and P. J. Pillips, “Report on the evaluation of 2D still-image face recognition algorithms,” NIST Interagency Report 7709, 2010.

[10] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, “Labeled faces in the wild: A database for studying face recognition in unconstrained environments,” Univ. of Mass., Tech. Rep. 07-49, 2007.

[11] H. Fan, Z. Cao, Y. Jiang, Q. Yin, and C. Doudou, “Learning deep face representation,” *arXiv:1403.2802*, 2014.

[12] S. A. Berrani and C. Garcia, “Enhancing face recognition from video sequences using robust statistics,” in *Proc. IEEE Int. Conf. Video Signal Based Surveill.*, 2005, pp. 324–329.

[13] Z.Yang,H.Ai, B.Wu, S. Lao, and L. Cai, “Face pose estimation, and its application in video shot selection,” in *Proc. Int. Conf. Pattern Recog.*, 2004, pp. 322–325.

[14] X. Gao, S. Z. Li, R. Liu, and P. Zhang, “Standardization of face image sample quality,” in *Proc. Int. Conf. Biometrics*, 2007, p. 242251.

[15] Z. Wang and A. C. Bovik, “A universal image quality index,” *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.

[16] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[17] H. Sellahewa and S. A. Jassim, “Image-quality-based adaptive face recognition,” *IEEE Trans. Instrum. Meas.*, vol. 59, no. 4, pp. 805–813, Apr. 2010.

[18] Y. Wong, S. Chen, S. Mau, C. Sanderson, and B. C. Lovell, “Patchbased probabilistic image quality assessment for face selection, and improved video-based face recognition,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recog. Worshop*, 2011, pp. 74–81.

[19] P. Viola and M. J. Jones, “Robust real-time face detection,” in *Proc. Int. J. Comput. Vis.*, 2004, vol. 57, no. 2, pp. 137–154.

[20] Y. Sun, X. Wang, and X. Tang, “Deep convolutional network cascade for facial point detection,” in *Proc. IEEE Conf. Comput. Vis. and Pattern Recog.*, 2013, pp. 3476–3483.

[21] E. Welzl, “Smallest enclosing disks (balls and ellipsoids),” *Lecture Notes in Computer Science 555*, pp. 359–370, 1991.

[22] T. Joachims, “Optimizing search engines using clickthrough data,” in *Proc. ACM Conf. Knowledge Discovery and Data Mining*, 2002.

[23] D. Parikh and K. Grauman, “Relative attributes,” in *Proc. Int. Conf. Comput. Vis.*, 2011.

[24] O. Chapelle, “Training a support vector machine in the primal,” *Neural Comput.*, vol. 19, no. 5, pp. 1155–1178, 2007.

