RRAR: A NOVEL REDUCED-REFERENCE IQA ALGORITHM FOR FACIAL IMAGES

Jiazhen ZHU, Yuchun FANG, Pengjun JI, Moad-EL ABDL, Wang DAI

School of Computer Engineering and Science, Shanghai University, Shanghai, China

ABSTRACT

Image Quality Assessment (IQA) aims at automatically predicting the perceptual quality of targets with low computation complexity and high precision. However, it is usually very hard to combine all these merits into one algorithm. In this paper, we propose simple yet efficient facial image quality assessment algorithm---Reduced-Reference Automatic Ranking (RRAR) for face recognition. The RRAR contains a quality control stage and quality ranking stage based on modified structural similarity---Reduced-Reference of SSIM as the reduced reference IQA module. Experimental results show that the proposed algorithm increases the precision of face recognition with low memory consumption and computation complexity and works exceptionally well with face images captured under uncontrolled environment.

Index Terms— Image Quality Assessment, Structural Similarity, Face Recognition

1. INTRODUCTION

Face image quality assessment is a challenging and necessary task in the area of face image recognition because of wide variety of distortions during acquisition, processing, compression, storage, transmission and reproduction of facial images. So far, the variations of illumination and pose are major factors leading to poor quality face image that will prominently deteriorate the performance of face recognition system.

Among the existing IQA methods, Full-Reference IQA (FRIQA) could automatically evaluate quality of images with predefined references. However, it is intractable for automatic face recognition system since it is impossible to define a reference for each subject. For face recognition, Reduced-Reference IQA (RRIQA) and Non-Reference IQA (NRIQA) are two reliable methods. Usually, NRIQA method is designed only for handling one type of IQA. Shiekh et al. use and quantify the dependencies in the natural scene statistics to assess the images [1]. In [2], Zhu et al. defined a new metric H, which quantify the amount of blur and noise, by decomposing the gradient matrix of images. Normally, NRIQA is of high complexity due to lack of reference. Hence, we attempt to utilize a RRIQA model, a hot topic in QA domain. Lv et al. used the Fast Johnson-Lindenstrauss transform of hashing features to assess image quality [3]. Li and Wang evaluate image by the extracted statistical features from a divisive normalization transformation [4]. Abdelkaher et al. decomposed images to intrinsic mode function [5]. The normal way of RRIQA is to extract features from image for IQA.

In online face recognition system, face IQA is helpful for prompt the accuracy. In [6], Yang et al. directly calculated luminance point of high-energy to assess images' precision in the spectrum. Rein-Lien et al. proposed to use an overall quality score in the automatic face recognition [7]. Mahoor et al. assess the quality of facial images based on the eigenface technique [8]. However, the IQA algorithm must be fast as a module of the whole recognition system.

In this paper, we propose a fast and simple rank type RRIQA algorithm that could help to filter the low-quality and select the good-quality facial images. We firstly modify the Structural Similarity (SSIM) [9] (a fast FRIQA metric) into a RRIQA method by constructing the automatically updated average face as reference. It is fused with a Support Vector Machine （SVM）based pose classifier to eliminate the bad quality image and utilized for quality ranking. Experiments show that our algorithm is fast and effective.

The remainder of this paper is organized as follows. In Section 2, we introduce the framework of RRAR. The modified SSIM for reduced-reference IQA is explained in Section 3 in detail. Fusion of modified SSIM and SVM is described in Section 4. The Section 5 analyzes the experimental results and the conclusions are presented in Section 6.

2. FRAMEWORK OF RRAR

The proposed RRAR mainly contains two parts: a quality control stage and a quality ranking stage and its framework is shown in Fig. 1.

Quality control stage aims at eliminating the facial-part images with dramatic pose variation in depth, detected non-face or partial face, and low quality images due to variation of lighting condition. The quality control is realized by performing a decision level fusion of modified SSIM and a trained SVM pose classifier. We obtain the reasonable thresholds learned with the sample sets for both of them. The modified SSIM could eliminate the non-face or partial-face by using the average face image as reference since it is an effective structure checking metric. For the same reason, it could also help to avoid facial images with uneven lighting around face area. For facial images of depth-rotation, we train a SVM pose classifier offline. The undesirable face images of very poor quality will largely be rejected in this stage. The quality ranking stage uses the output of the modified SSIM as scores to rank the face images and select "better" ones for registration or login.

In this framework, with the face recognition system working on, the average face image could be updated automatically with the accumulation of face images in the database. One of its benefit is that only a small amount of images is good enough to be used as reference, which save the memory consumption and computation complexity.

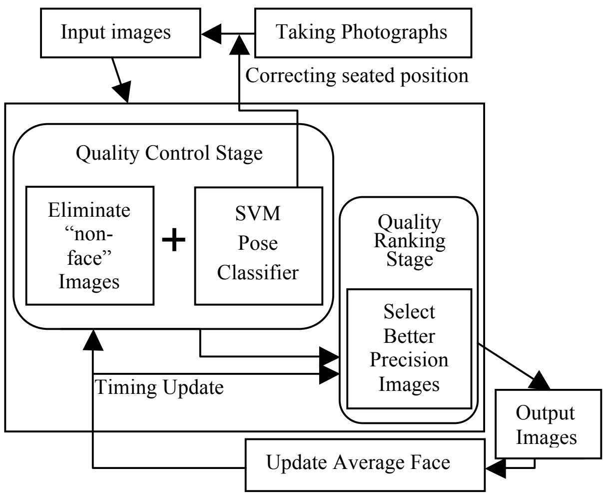


Fig.1.Framework of RRAR

3. MODIFIED SSIM

Among various popular FRIQA metrics such as Mean Square Error (MSE), Peak-Signal-to-Noise-Ratio (PSNR) and Visual Information Fidelity (VIF) [11], we pick SSIM as the face IQA metric out of the following consideration. The MSE and PSNR are simple but not coherent with human visual system, the VIF is proposed out of the traits of human visual system but computationally complex. The SSIM is proposed based on the hypothesis that the human vision system is highly adapted to structural information [10, 12]. Face appearance is just one of such object, especially the frontal face images broadly used in face recognition, no matter the ethnicity, age, gender and other demographics of the subject, are of strong structure due to relatively fixed position of facial features.

The formula of SSIM is defined as

$$SSIM(x,y) = [l(x,y)]^{\alpha}[c(x,y)]^{\beta}[s(x,y)]^{\gamma} \qquad (1)$$

Where *l(x, y)* is the mean value estimates the light term, *c(x, y)* is the standard deviation is the contrast term and *s(x, y)* is the structural term, the exponential coefficient is a control of contribution of each part.

For using the usual SSIM, a reference is necessary to each image, while it is impossible to provide a reference to any subject or face image. To solve this problem, we propose to construct the average face, which firstly is built among the existing images in the database, as reference and update it with the environment of the face recognition system. The so-called "average face image" is an aggregate of many human faces as illustrated in Fig. 2. All the images of pixel value are added at each same pixel point. Then, that value is divided by the number of images. Though average face is not real, it looks very much like a real by human visual system.

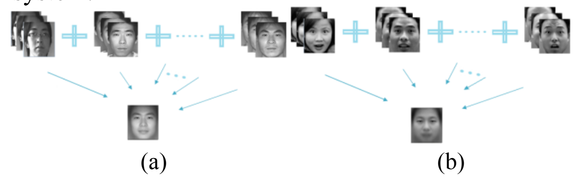


Fig.2. the process of creating an average image ((a) average image gained in face images database (b) average image gained in CASPEAL database)

Structurally, all face images should be close to the average face with SSIM metric, but the non-face, partial-face and lighting-uneven face are all different from the average face. The advantages of taking average face image as reference not only lie in low-storage required reference, but also high efficiency when combined with the SSIM. This simple modified SSIM contributes in both stages of RRAR.

4. FUSION OF MODIFIED SSIM AND SVM

Examples of "depth-rotation" of face images are shown in the Fig. 3. Since part of the face is occluded due to such pose variation, the recognition performance drops drastically with most of the available system. To tract this problem, we train a machine learning model with SVM to classify face images of depth rotation and frontal faces.

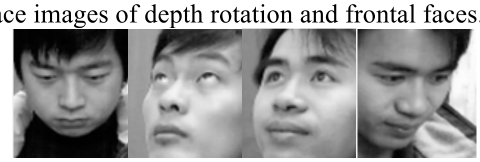
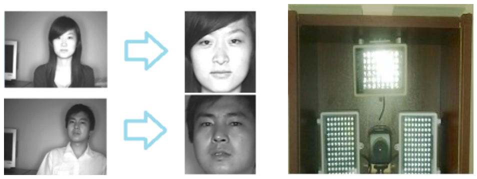


Fig.3. face images with obvious pose variations in depth

SVM is primarily addressed to solve the binary classification problem through minimizing the structural risk. Given the set of training samples $\{x_i, y_i\}$, $i = 1,\ldots,l$, where $x_i \in R^N$ and its label $y_i \in \{-1,1\}$. The solution to linearly separable problem is $f(x) = \sum_{i \in SV} \alpha_i y_i x_i \cdot x + b$ (α_i is the coefficient, SV means the label of support vectors) through solving a quadratic problem [11]. The generalization of SVM to nonlinear problems can be easily realized through replacement of the inner product term with a kernel function that satisfies the Mercer's condition.

5. EXPERIMENTAL RESULTS AND ANALYSIS

We tested RRAR on the two image databases---a subset of CAS-Peal [13], which is the largest domain about obverse face image in the whole CAP-Peal database, have 376 subjects. (5 images per person) and a self collected face database from 50 subjects (10 images per person) collected with the equipment shown in Fig.4.(b) in uncontrolled environment.



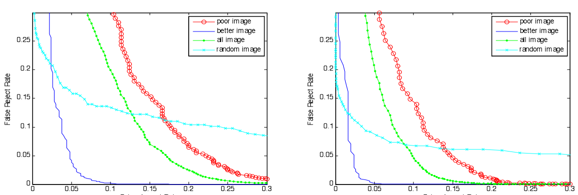
(a) (b)

Fig.4. (a) the part of face database (b) images of collecting equipments.

5.1. Feature Extraction

We adopt the bench mark face recognition feature （Local Gabor Binary Pattern）LGBP and two types of multi-directional (Rotation Invariant Uniform Local Binary Pattern) RIU-LBP features mentioned in our previous work in [14] to evaluate the proposed RRAR and denote them as LBP1, LBP2.

Figure 5 shows that, whatever we choose the features and whatever the dimensions of feature are, it does not change that the order of four different lines from inside to outside. That means the effect of modified SSIM is steady when there are changes in the aspect of feature or its dimensions. In our system, we choose LBP1 with 560 dimensions in order to increase the speed.

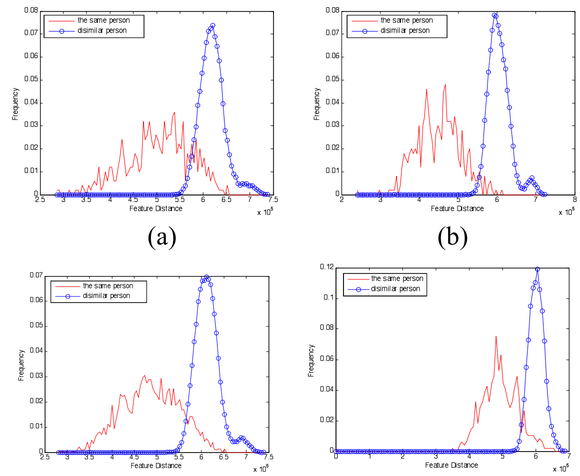


(a) (b)

Fig.5. comparison among LBP1 and LGBP with face database ((a) LBP1 ROC curve (b) LGBP ROC curve)

5.2. The Effect of Modified SSIM

We first calculate the SSIM values between the average face and all images in the face database. Ranking by the value of modified SSIM, we pick half poor images and half better images for each subject and denote them as big group 1 and big group 2 respectively. And then we calculate LBP1, LBP2 and LGBP features in these image subsets. In each group, we calculate the pair-wise distance between images both from the same subject and from different subjects. Then the histogram is constructed to estimate the distribution of distances from pairs of images from the same subject and those from different subjects as shown in Fig. 6(a)-(d) respective for big group 1, big group 2, all images and randomly picked image subsets. The crossing areas in the figures are the number of misclassified samples. The bigger the area, the lower the correct recognition rate will be and vice versa. The crossing area of better images is the smallest in the four histograms. We also plot the Receiving Operating Characteristic (ROC) curves in Figure 7. It is obvious in both figures that the set of better images picked with modified SSIM is of the minimum error rate and the performance on this set is dramatically better than those on the other image sets.



(a) (b) (c) (d)

Fig.6. Comparison among poor images, better images, all images and random images with face images database ((a) histogram of poor images (b) histogram of better images (c) histogram of all images (d) histogram of random images)

Moreover, by comparison between Figure 5(a) and Figure 7, the curves for the better images is obvious different from those of the poor images. Hence, adopting the modified SSIM as quality ranking metrics could improve the performance of face recognition. Moreover, such improvement is more distinctive in Figure 5(a) (obtained with the 500 images captured in uncontrolled environment) than in Figure 7 (obtained in CASPeal). Hence, the modified SSIM is more suitable for practical environment.

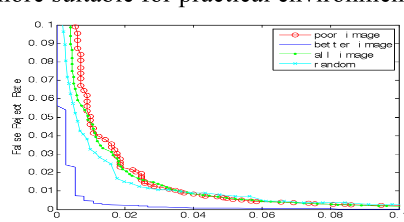


Fig.7. LBP1 ROC curve with CASPeal database

5.3. Estimate Thresholds of Quality Control Stage

The thresholds stated in Section 4 are obtained with the following algorithms:

Step 1: Preparing training sets of positive (good quality face images) and negative samples (non-face images).

Step 2: Obtain modified SSIM metric values on the training sets.

Step 3: Estimate the distribution of value of modified SSIM values respectively for positive and negative samples, see the example shown in Fig. 8, by which the threshold is set at the Minimum Error Rate (MRR) point 0.4.

Similarly, we pick the threshold for SVM classifier, see the example shown in Fig. 9 with threshold 0 as minimum error rate point. Part of the False Accept Rate (FAR) and False Reject Rate (FRR) is listed for comparison in Table 1 obtained with various value of thresholds.

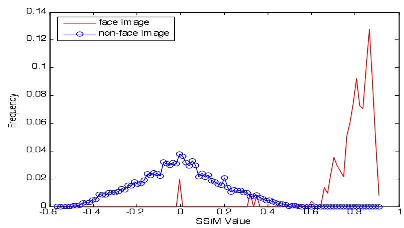


Fig.8. face and non-face images of histogram

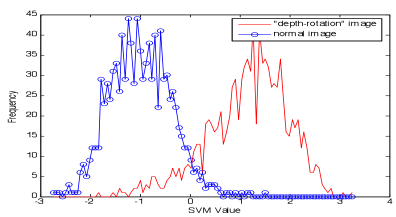


Fig.9. "depth-rotation" and normal images of histogram

Table 1. Performance of Threshold

non-face threshold	SVM Pose threshold	FAR (%)	FRR(%)
0.35	-0.5	9	3
	0	14	6
	0.5	20	11
0.4	-0.5	4	10
	0	5	3
	0.5	12	5
0.45	-0.5	4	19
	0	8	14
	0.5	13	8

6. CONCLUSIONS

In this paper, we propose a new face IQA algorithm RRAR to improve the performance of face recognition. The RRAR contains two parts: a quality control stage and a quality ranking stage. Both parts utilize a modified SSIM proposed by us as a reduced-reference QA module. Through building up the average face image as reference, the normal SSIM could be easily applied for face image QA. The quality control step of RRAR fuses modified SSIM and SVM to filter the poor-quality face images, the former specially serves to eliminate bad-illumination-condition images, the latter serves to eliminate face images with obvious pose variations in depth. The quality ranking stage use the output of the modified SSIM as scores to rank the face images, in which the average face could be automatically updated. The experiments on two face database show that the proposed RRAR could prominently improve the performance of face recognition through quality assessment of low memory consumption and computation complexity, especially in the situation of practical uncontrolled environment.

7. ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China (60605012), the Natural Science Foundation of Shanghai (08ZR1408200), the Open Project Program of the National Laboratory of Pattern Recognition of China, the Innovation Fund for the Undergraduate Students of Shanghai University (CXGJ09-10) and the Shanghai Leading Academic Discipline Project (J50103).

8. REFERENCES

[1] H.R. Sheikh, A.C. Bovik and L. Cormack, "No-reference quality assessment using natural scene statistics: JPEG2000," *IEEE Trans. Image Processing*, vol. 14, pp. 1918-1927, Nov. 2005.

[2] X. Zhu and P. Milanfar, "A no-reference sharpness metric sensitive to blur and noise," *Quality of Multimedia Experience,* pp. 64-69, Jul. 2009.

[3] X. Lv and Z.J. Wang, "Reduced-reference image quality assessment based on perceptual image hashing," In *Proc. IEEE Int. Conf Image Processing,* pp. 4361-4364, 2009.

[4] Q. Li and Z. Wang, "Reduced-reference image quality assessment using divisive normalization-based image representation," *IEEE Journal of Selected Topics in Signal Processing,* vol. 3, pp. 202-211, 2009.

[5] A.A. Abdelouahad and M. El Hassouni, "A reduced reference approach based on bidimensional empirical mode decomposition for image quality assessment", *I/V Communication and Mobile Network*, pp. 1-4, 2010.

[6] F. Yang, J. Su and J. Dai, "Fast quality assessment of face images for face recognition," Control Conference, pp. 531-535, Jul. 2008.

[7] R.-L.V. Hsu, J. Shah, and C. Jersey, "Quality Assessment of Facial Images," Biometric Consortium Conf., pp. 1-6, Sept. 2006

[8] M. Abdel-Mottaleb and M.H. Mahoor, "Application notes-Algorithms for Assessing the Quality of Facial Images," Computational Intelligence Magazine, IEEE, vol. 2, pp. 10-17, May. 2007

[9] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: From error measurement to structural similarity," *IEEE Trans. IP*, vol. 13, pp. 600-612 Apr. 2004.

[10] Z. Wang, and A.C. Bovik, "Mean squared error: Love it or leave it? A new look at Signal Fidelity Measures," *Signal Processing Magazine,* IEEE, vol. 26, pp. 98-117, 2009.

[11] H.R. Sheikh and A.C. Bovik, "Image information and visual quality," *IEEE Trans. IP,* vol. 15, pp. 430–444, Feb. 2006.

[12] Z. Wang, A.C. Bovik, and L. Lu, "Why is image quality assessment so difficult?," In *Proc. IEEE Int. Conf Acoustics, Speech, and Signal Processing,* vol. 4, pp. 3313-3316, Aug. 2002.

[13] W. Gao, B. Cao, S. Shan, et al, "The CAS-PEAL Large-Scale Chinese Face Database and Baseline Evaluations," *IEEE Trans. Systems, Man and Cybernetics, Part A: Systems and Humans,* vol. 38, pp. 149-161, Jan. 2008.

[14] Y Fang, J Luo, and C Lou, "Fusion of multi-directional rotation invariant uniform LBP features for face recognition", *Intelligent Information Technology Application,* vol. 2, pp. 332-335, Nov. 2009.