

# CAN NO-REFERENCE IMAGE QUALITY METRICS ASSESS VISIBLE WAVELENGTH IRIS SAMPLE QUALITY?

Xinwei Liu<sup>1,2,\*</sup>, Marius Pedersen<sup>2</sup>, Christophe Charrier<sup>1</sup>, and Patrick Bours<sup>2</sup>

<sup>1</sup>Normandie Univ, UNICAEN, ENSICAEN, CNRS, GREYC, 14000 Caen, France

<sup>2</sup>NTNU - Norwegian University of Science and Technology, Gjøvik, Norway

## ABSTRACT

The overall performance of iris recognition systems is affected by the quality of acquired iris sample images. Due to the development of imaging technologies, visible wavelength iris recognition gained a lot of attention in the past few years. However, iris sample quality of unconstrained imaging conditions is a more challenging issue compared to the traditional near infrared iris biometrics. Therefore, measuring the quality of such iris images is essential in order to have good quality samples for iris recognition. In this paper, we investigate whether general purpose no-reference image quality metrics can assess visible wavelength iris sample quality.

**Index Terms**— Quality assessment, visible wavelength, iris biometrics, no-reference, image quality metric

## 1. INTRODUCTION

The iris is one of the most commonly used modalities for biometric recognition and it is a mature technology used in many government and civilian applications (e.g. border control in the United Kingdom and United Arab Emirates). However, most of the existing systems rely on heavy imaging constraints captured in a stop-and-stare interface, at close distances and using near infrared (700-900 nm) wavelengths with sufficient quality. In recent years, thanks to the development of imaging technologies, there are more and more iris recognition systems that operate in the visible wavelength (VW) and in less constrained environments [1–7]. The VW iris imaging systems lead to acquire degraded iris samples due to less constrained environments that makes the sample quality assessment a major issue. Existing iris sample quality assessment methods [8, 9] cannot handle the specificity of VW iris data and there are quite a few approaches developed for such iris images [10]. Since VW iris images are much closer to ordinary color images than near infrared iris images, it is interesting to investigate whether general purpose no-reference image quality metrics (IQMs) can assess VW iris samples due to the lack of methods specific for VW iris image

quality assessment. Towards this goal, we propose to 1) conduct the iris sample quality assessment by using no-reference IQMs, and 2) evaluate the performance of no-reference IQMs based on the performance of iris recognition system in order to answer ‘Can no-reference IQMs assess visible wavelength iris sample quality?’.

In this paper, we introduce the related works of this research field in Section 2. In Section 3, the experimental setup is presented. Then the experimental results will be illustrated in the next section. Finally, the conclusions and future works are presented in Section 5.

## 2. RELATED WORK

With some minor exceptions, the large majority of the existing iris recognition methods (both near infrared and VW iris) follow the statistical pattern recognition model and can be separated into four stages: segmentation, normalization, feature extraction and comparison [11]. The human iris is an internal organ, naturally protected, visible from the exterior and enables data acquisition. VW iris recognition systems will inevitably constitute a tradeoff between data acquisition constraints and recognition accuracy. This research topic receives growing interest and several recent publications focus on this aspect. The ‘Iris-on-the-move’ project [1] is an example of an image acquisition system to make the recognition process less intrusive for subjects. Fancourt *et al.* [2] presented that it is possible to acquire sufficiently high-quality images at a distance of up to 10 meters. Smith *et al.* [3] investigated that the iris structure can be captured in the VW spectra, discovering the possibility of using multispectral data to improve recognition performance. Park and Kim [4] captured out of focus VW iris images. Raja *et al.* [5] first developed a new recognition scheme and adapt it to smartphone based VW iris images, and then proposed an improved VW iris imaging solution for increasing the performance of VW iris recognition system [6]. Boyce *et al.* [7] studied the image acquisition wavelength of revealed components of the iris, and identified the important role of iris pigmentation.

In order to ensure high performance of biometric system based on VW iris, the challenge is to assess the iris quality during the acquisition process [12]. However, only a few pub-

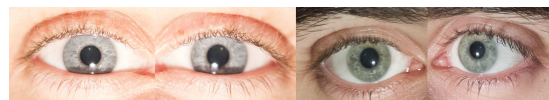
\*This research is supported by the Conseil Regional Basse-Normandie Grant 14P02048 and Research Council of Norway through project number 221073: HyPerCept - Color and Quality in Higher Dimensions.

lications investigated VW iris quality and most of these works are from Proença *et al.* [10, 11]. Due to the limited number of approaches for VW iris sample quality assessment, meanwhile, VW iris images are much closer to ordinary color images than near infrared iris images, it is interesting to investigate whether no-reference IQMs can assess VW iris samples. No-reference IQMs provide quality assessment of an image without the need of any reference image or its features and they are mostly for natural images. The properties of no-reference IQMs make it possible to use them to assess VW iris images. There are many no-reference IQMs, for more details we refer to [13].

### 3. EXPERIMENTAL SETUP

The iris recognition system used in this paper is OSIRIS (Open Source for IRIS) version 4.1 [14]. The OSIRIS reference system is an open source iris recognition system developed in the framework of the BioSecure project [14]. The VW iris database used in this paper is part of the **GC<sup>2</sup> multi-modality biometric database**. This database is on the final stage of the development by the authors of this paper and will be available soon to the research field. The iris database contains 50 subjects, both left and right eyes are acquired by two different cameras: a Canon D700 camera with Canon EF 100mm f/2.8L Macro Lens (18 Megapixels) and a Google Nexus 5 smartphone embedded camera (8 Megapixels). 15 samples are taken for each camera per eye. In order to avoid that segmentation errors corrupt the results for OSIRIS system, 31 and 25 subjects are selected from reflex camera and smartphone, which were accurately segmented based on visual inspection. Totally, 1680 iris samples (930 from reflex camera and 750 from smartphone) are used in this paper. Only small natural distortions appear in the images. Some examples from the database are shown in Figure 1.

We select 13 no-reference IQMs which are commonly used in the research field and their Matlab code are publicly available: BIQI [15], BLIINDS2 [16], BRISQUE [17], ILNIQU2 [18], JNBM [19], SSEQ [20], CONTRAST [21], DCTSP [22], PWN [23], AQI and AQIP [24], SSH [25], and SH [26]. BIQI is a two-step framework based on natural scene statistics. BLIINDS2 is a distortion-agnostic metric based on a natural scene statistics model of discrete cosine transform coefficients. BRISQUE is a natural scene statistic-based distortion-generic model that operates in the spatial domain. ILNIQU2 integrates the features of natural image statistics derived from multiple cues and learns a multivariate Gaussian model of image patches from a collection of pristine natural images. JNBM is a perceptual-based image sharpness/blurriness metric. SSEQ is based on spatial and spectral entropies. CONTRAST is a method for contrast distorted images based on the principle of natural scene statistics. DCTSP is a algorithm for blurred images and it is based on block-based discrete cosine transform statistics and



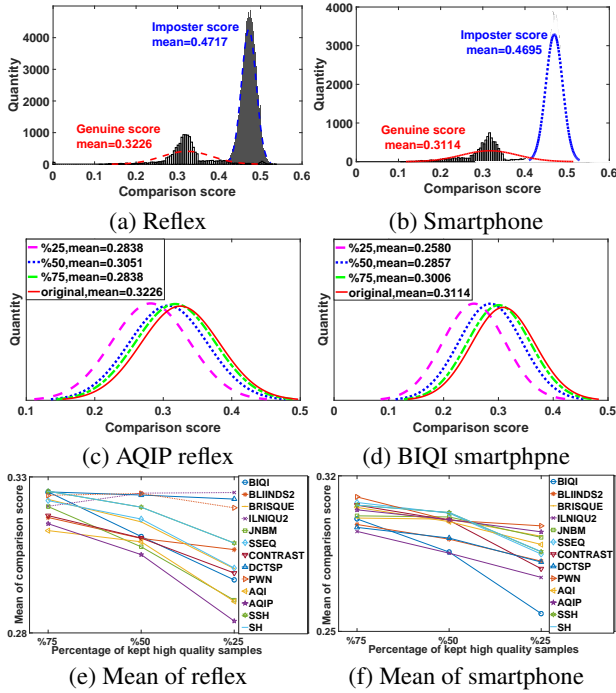
**Fig. 1.** Iris samples from the database. The first two samples are from reflex camera and the last two are from smartphone.

linear prediction method. PWN uses a perceptually weighted local noise into a probability summation model. AQI and AQIP are based on measuring the variance of the expected entropy of a given image upon a set of predefined directions. SSH measures image sharpness. SH is a definition of a sharpness index that is closely related to the notion of global phase coherence. Some of these no-reference IQMs are designed for distortion-specific use: CONTRAST for measuring image contrast; JNBM, SH, and SSH for measuring image sharpness; DCTSP for measuring image sharpness and block-based artifacts; and PWN for measuring image noise. Some of them are designed for non-specific use: AQI, AQIP, BIQI, BLIINDS2, BRISQUE, ILNIQU2, and SSEQ. All assess image quality in grayscale channel. The conversion from color image to grayscale image is by the Matlab function `'rgb2gray'`.

## 4. EXPERIMENTAL RESULTS

### 4.1. Distribution and mean of comparison scores

In order to evaluate the performance of the IQMs, we first plot the original comparison score for both reflex camera and smartphone in Figure 2 (a) and (b). The x axis represents the score and the y axis represents the quantity of the comparison. The histograms of the comparison score are obtained from the genuine (comparison between samples from the same subject) and imposter (comparison between samples from different subjects) comparisons for all image samples. The line plots correspond to the fitted normal distributions and the mean of the score is given as well. In general, high quality biometric samples could generate relatively good genuine comparison scores (in our case, the score closer to zero the more similar two iris samples), which are well separated from impostor comparison scores [27]. This can be observed from Figure 2 (a) and (b). Then we show two examples of the histogram of the genuine comparison scores when omitting low quality samples by using selected no-reference IQMs in Figure 2 (c) and (d). Grother *et al.* [27] proposed to use quality-bin based approaches to evaluate the image quality assessment methods. They believed if a certain percentage of low quality samples are excluded from the dataset, the comparison score would decrease (in our case) and the Equal Error Rate (EER) (when False Match Rate (FMR) and False None Match Rate (FNMR) are equal) would decrease as well. Because the scale of the quality score for each IQM is different and the linearity of the quality score is unknown, we omit the percentile low quality samples and keep 75%, 50%, and 25%



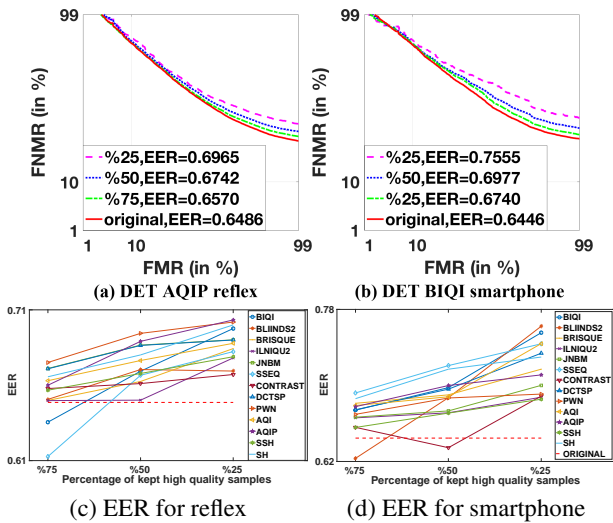
**Fig. 2.** Comparison scores and their mean values with and without omitting low quality samples.

of high quality samples for each of the IQMs. We conduct an inspection for all omitted low quality samples in order to ensure that at least there is one sample per subject. In the two plots, the red continuous line represents the original comparison score; the green '-' line represent the comparison scores when we keep the 75% highest quality iris samples; the blue colon line represents the comparison scores when we keep the 50% highest quality iris samples; and the magenta dot line represents the comparison scores when we keep the 25% highest quality iris samples. We only illustrate the best IQM for reflex camera and smartphone, and for the rest of IQMs we only plot the mean values with omitting low quality samples in Figure 2 (e) and (f). From Figure 2 (c) and (d) we can see that, there is only slightly left shift for genuine score. In Figure 2 (e) and (f), all mean values of comparison score become lower when omitting more low quality samples for smartphone, however, the mean values become higher at 50% (equals to right shift in previous plots) for reflex camera when using ILNIQU2 and PWN to omit low quality samples. The rest of the no-reference IQMs have similar performance with these two IQMs. The above findings indicate that the selected no-reference IQMs fail to assess the iris image quality well according to the performance of OSIRIS system. AQIP and BIQI have lowest mean of comparison score for reflex camera and smartphone, respectively, while ILNIQU2 and PWN have the highest mean score compared to the others.

## 4.2. DET curves and EER

An IQM is useful if it can at least give an ordered indication of an eventual performance. Rank-ordered Detection Error Trade-Off (DET) characteristics curve is one of the most commonly used and widely understood method used to evaluate the performance of quality assessment approaches. The DET used here plots FNMR versus FMR. As mentioned before, we also obtain EER as an indicator to examine the performance of IQMs. Figures 3 (a) and (b) illustrate the two examples of DET curves with EER for data with and without omitting low quality samples for reflex camera by using AQIP, and for smartphone by using BIQI. Since all DET curves for both reflex camera and smartphone are very similar we only show two examples here. If a DET curve is closer to the left-bottom point, it means that this set of data lead to a higher iris recognition performance. Meanwhile, the lower EER values the better system performance. From Figure 3 (a) and (b) we can see that, the DET curves with omitting low quality samples (magenta, blue, and green lines, same as Figure 2) are on the right-top side or very close to DET curve without omitting low quality samples (the red continuous line). Figure 3 (c) and (d) represent the EER values from all IQMs for reflex camera and smartphone, respectively. From Figure 3 (c) and (d) we can observe that, most of the EER values with omitting low quality samples are higher than EER values without omitting low quality samples (the red dash line represents the EER calculated from original comparison score). This means that there is no system performance improvement after omitting low quality samples. However, the EER values from BIQI and SSEQ when keeping 75% of the high quality score are lower than original EER value for reflex camera, and EER values from CONTRAST when keeping 50% and BLINDS2 when keeping 75% of the high quality score are lower than original EER value for smartphone.

Additionally, we calculate EER values for both cameras by omitting lowest quality iris sample one by one until only one highest quality iris sample remains. Figure 4 (a) and (b) illustrate the change of EER values for both cameras. The x axis represents the number of omitted low quality samples. The y axis represents the EER value. We slightly shift the plot to the right-top side for each IQM in order to show the results more clear. As introduced before, when we omit low quality iris samples, the EER value will decrease if the IQMs can predict iris sample quality. From both plots in Figure 4 we can see that, there is no obvious decrease for EER values (from the left to the right) until only a few high quality samples left. However, for most of the IQMs, we can find an increase or fluctuation in the EER values at the end of the line. Another interesting finding is that, most of the lines have a big drop in the end due to only high iris samples are remains, except the second last line (BLINDS2) for both cameras and the fourth last line (ILNIQU2) for smartphone. This means that high quality iris samples given by BLINDS2 and ILNIQU2 cannot



**Fig. 3.** Examples of DET curves with EER for original comparison score with and without omitting low quality samples.

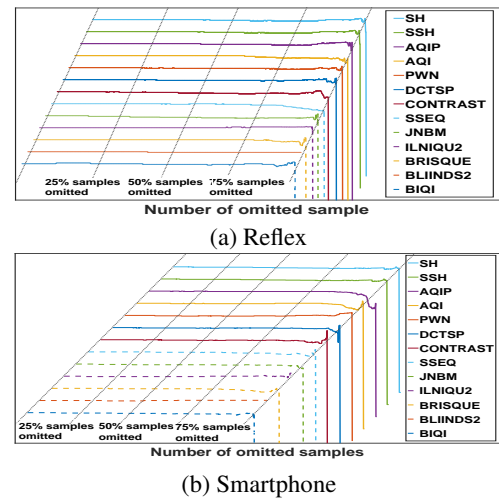
enhance the performance of the OSIRIS system compared to low quality samples.

### 4.3. Computational complexity

We also notice that some of the no-reference IQMs are designed for real images so they are time consuming, which is not suitable for real time iris recognition. The IQMs whose computation speed is less than 1 second: BIQI, JNB, SSEQ, DCTSP, PWN, SSH, and SH; whose computation speed is between 1 and 5 seconds: BRISQUE, ILNIQ2, CONTRAST, AQI, and AQIP; and whose computation speed is more than 5 seconds: BLINDS2.

## 5. CONCLUSIONS AND FUTURE WORKS

In this paper, we evaluate the performance of no-reference IQMs on VW iris images. From the experimental results we can conclude that, the selected 13 no-reference IQMs cannot assess well the quality of VW iris samples from selected  $GC^2$  database. By looking at EER values, only BIQI and SSEQ increase the performance when keeping 75% of the high quality samples for reflex camera; only CONTRAST and PWN increase the performance when keeping 50% and 75% of the high quality samples for smartphone, respectively. Since all these IQMs are designed for natural images but VW iris images contain rich texture information with unique structure components. Therefore, VW iris images are different from natural images, so it can be the reason for the low performance. Another possible reason when these IQMs fail to assess iris sample quality can be that, the quality attributes applied in these IQMs do not appear in some of the iris images and these attributes are not sensitive for OSIRIS iris recog-



**Fig. 4.** EER values with omitting low quality iris samples one by one until the best quality sample left. The x axis represents the number of omitted low quality samples. The y axis represents the EER value. We slightly shift the plot to the right-top side for each IQM in order to show the results more clear.

nition system. The contribution of this work can be used for the development of quality assessment methods for VW iris, moreover, for contactless biometric modalities. The experimental results indicate that biometric modality specific IQMs might be required. It has been shown that red channel has more discriminatory information than the other channels. It is interesting to investigate if the experimental results will be different if only red channel is used. Other VW iris databases and iris recognition systems could also be used to evaluate the performance of selected metrics.

## 6. REFERENCES

- [1] J. R. Matey, D. Ackerman, J. Bergen, and M. Tinker, "Iris recognition in less constrained environments," in *Advances in Biometrics*, pp. 107–131. Springer, 2008.
- [2] C. Fancourt, L. Bogoni, K. Hanna, Y. Guo, R. Wildes, N. Takahashi, and U. Jain, "Iris recognition at a distance," in *International Conference on Audio-and Video-Based Biometric Person Authentication, NY, USA, July*. Springer, 2005, pp. 1–13.
- [3] K. N. Smith, V. P. Pauca, A. Ross, T. Torgersen, and M. C. King, "Extended evaluation of simulated wavefront coding technology in iris recognition," in *Biometrics: Theory, Applications, and Systems, 2007. First IEEE International Conference on, Washington DC, Sept.* IEEE, 2007, pp. 1–7.
- [4] K. R. Park and J. Kim, "A real-time focusing algorithm

- for iris recognition camera,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 35, pp. 441–444, 2005.
- [5] K. B. Raja, R. Raghavendra, V. K. Vemuri, and C. Busch, “Smartphone based visible iris recognition using deep sparse filtering,” *Pattern Recognition Letters*, vol. 57, pp. 33–42, 2015.
  - [6] K. B. Raja, R. Raghavendra, and C. Busch, “Visible iris imaging: A novel imaging solution for improved iris recognition,” in *2015 IEEE International Conference on Imaging Systems and Techniques (IST)*, Macau, China, Sept. IEEE, 2015, pp. 1–6.
  - [7] C. Boyce, A. Ross, M. Monaco, L. Hornak, and X. Li, “Multispectral iris analysis: A preliminary study,” in *2006 Conf. on Computer Vision and Pattern Recognition Workshop*, NY, USA, June. IEEE, 2006, pp. 51–51.
  - [8] N. D. Kalka, J. Zuo, N. A. Schmid, and B. Cukic, “Image quality assessment for iris biometric,” in *Defense and Security Symposium, Orlando, USA, Apr.* Int. Society for Optics and Photonics, 2006, p. 62020D.
  - [9] Y. Chen, S. C. Dass, and A. K. Jain, “Localized iris image quality using 2-d wavelets,” in *International conference on biometrics*. Springer, 2006, pp. 373–381.
  - [10] H. Proença, “Quality assessment of degraded iris images acquired in the visible wavelength,” *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 1, pp. 82–95, 2011.
  - [11] H. Proença, S. Filipe, R. Santos, J. Oliveira, and L. A. Alexandre, “The ubiris. v2: A database of visible wavelength iris images captured on-the-move and at-a-distance,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, pp. 1529–1535, 2010.
  - [12] X. Liu, M. Pedersen, and C. Charrier, “Image-based attributes of multi-modality image quality for contactless biometric samples,” in *3rd International Conference on Signal Processing and Integrated Networks (SPIN)*, New Delhi, India, Feb., 2016.
  - [13] V. Kamble and K.M. Bhurchandi, “No-reference image quality assessment algorithms: A survey,” *Optik-International Journal for Light and Electron Optics*, vol. 126, no. 11, pp. 1090–1097, 2015.
  - [14] G. Sutra, B. Dorizzi, S. Garcia-Salicetti, and N. Othman, “A biometric reference system for iris, osiris version 4.1,” *Telecom Sud Paris, France, Tech. Rep*, 2012.
  - [15] A. K. Moorthy and A. C. Bovik, “A two-step framework for constructing blind image quality indices,” *IEEE Signal Processing Letters*, vol. 17, pp. 513–516, 2010.
  - [16] M. A. Saad, A. C. Bovik, and C. Charrier, “Dct statistics model-based blind image quality assessment,” in *18th IEEE International Conference on Image Processing, Brussels, Belgium, Sept.*, 2011, pp. 3093–3096.
  - [17] A. Mittal, A. K. Moorthy, and A. C. Bovik, “No-reference image quality assessment in the spatial domain,” *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695–4708, 2012.
  - [18] L. Zhang, L. Zhang, and A. C. Bovik, “A feature-enriched completely blind image quality evaluator,” *IEEE Transactions on Image Processing*, vol. 24, no. 8, pp. 2579–2591, 2015.
  - [19] R. Ferzli and L. J. Karam, “A no-reference objective image sharpness metric based on the notion of just noticeable blur (jnb),” *IEEE transactions on image processing*, vol. 18, no. 4, pp. 717–728, 2009.
  - [20] L. Liu, B. Liu, H. Huang, and A. C. Bovik, “No-reference image quality assessment based on spatial and spectral entropies,” *Signal Processing: Image Communication*, vol. 29, no. 8, pp. 856–863, 2014.
  - [21] Y. Fang, K. Ma, Z. Wang, W. Lin, Z. Fang, and G. Zhai, “No-reference quality assessment of contrast-distorted images based on natural scene statistics,” *IEEE Signal Processing Letters*, vol. 22, no. 7, pp. 838–842, 2015.
  - [22] Y. Han, X. Xu, and Y. Cai, “Novel no-reference image blur metric based on block-based discrete cosine transform statistics,” *Optical Engineering*, vol. 49, pp. 050501–050501, 2010.
  - [23] T. Zhu and L. Karam, “A no-reference objective image quality metric based on perceptually weighted local noise,” *EURASIP Journal on Image and Video Processing*, vol. 2014, no. 1, pp. 1–8, 2014.
  - [24] S. Gabarda and G. Cristóbal, “Blind image quality assessment through anisotropy,” *Journal of Optical Society of America*, vol. 24, no. 12, pp. B42–B51, 2007.
  - [25] A. Leclaire and L. Moisan, “No-reference image quality assessment and blind deblurring with sharpness metrics exploiting fourier phase information,” *Journal of Mathematical Imaging and Vision*, pp. 145–172, 2015.
  - [26] G. Blanchet and L. Moisan, “An explicit sharpness index related to global phase coherence,” in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing*, Tokyo, Japan, Mar., 2012, pp. 1065–1068.
  - [27] Patrick Grother and Elham Tabassi, “Performance of biometric quality measures,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 29, no. 4, pp. 531–543, 2007.