

# Étude de l'évaluation de la qualité d'images d'iris pour un système de reconnaissance biométrique

Xinwei LIU<sup>1,2</sup>, Marius PEDERSEN<sup>1</sup>, Christophe CHARRIER<sup>2</sup>, Patrick BOURS<sup>1</sup>

<sup>1</sup>NTNU - Norwegian University of Science and Technology, Gjøvik, Norway

<sup>2</sup>Normandie Univ, UNICAEN, ENSICAEN, CNRS, GREYC, 14000 Caen, France

xinwei.liu@unicaen.fr, marius.pedersen@hig.no  
christophe.charrier@unicaen.fr, patrick.bours@ntnu.no

**Résumé** – La performance d'un système de reconnaissance biométrique d'iris est conditionnée par la qualité des images obtenues par le capteur. Habituellement, l'acquisition de telles données s'effectue dans des conditions contrôlées et dans le domaine du proche infra rouge, ce qui garantit une qualité minimale au détriment d'un système d'acquisition spécifique et onéreux. Grâce à l'évolution des techniques d'imagerie, il est possible d'acquérir des images d'iris dans le spectre visible réduisant ainsi le coût du système biométrique. Dès lors se pose le problème de l'évaluation de la qualité des images d'iris obtenues dans le spectre visible. Dans cet article, dix méthodes d'évaluation de la qualité des images sans-référence est réalisée afin de déterminer si elles sont performantes lorsque l'on évalue la qualité des images d'iris. Les résultats mettent en évidence un comportement varié selon la métrique et le système d'acquisition.

**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, contactless 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 contactless iris sample quality, in order to ensure a high performance level of any iris-based biometric system.

## 1 Introduction

Parmi les différentes modalités utilisées pour la reconnaissance et l'identification d'un individu, l'analyse de l'iris est une approche de plus en plus répandue. Elle est notamment utilisée par plusieurs gouvernements et applications civiles, telles que le contrôle aux frontières au Royaume-Uni et aux Émirats Arabes Unis. Néanmoins, la plupart des systèmes existants déploient des procédés spécifiques pour l'acquisition des images, fonctionnant le plus souvent en champs proche et nécessitant l'utilisation d'une gamme de spectre de longueur d'onde en proche infrarouge (700-900 nm). La qualité des images ainsi acquises permet de garantir un niveau de performance du système complet, le plus souvent au détriment du coût financier des systèmes d'acquisition qui sont très souvent onéreux. Grâce aux récents développements des technologies d'acquisition des images, il est maintenant possible d'acquérir des images d'iris dans le domaine du spectre visible à l'aide de capteurs peu onéreux (appareil photos numériques, caméras embarquées dans des téléphones) et ainsi développer des systèmes biométriques basés sur l'analyse de l'iris opérant de facto dans des environnements moins contraints [1, 2]

De tels systèmes d'acquisition sans contact acquièrent dès lors des images d'iris dans un environnement non contrôlé qui

peuvent présenter ou non des dégradations. Il convient alors de pouvoir évaluer la qualité des images d'iris afin de garantir un seuil de performance du système de reconnaissance ou d'identification biométrique.

Cependant, les méthodes actuelles d'évaluation de la qualité des images d'iris [3, 4] ne sont pas adaptées pour les images acquises dans le domaine visible (puisque spécifiques au spectre du proche infra rouge). De plus, peu de méthodes d'évaluation de la qualité des images d'iris dans le domaine visible sont disponibles [5].

Les images ainsi acquises pouvant être considérées comme des images couleur naturelles, il convient de quantifier et analyser les performances des métriques d'évaluation de la qualité des images existantes avant d'envisager le développement d'une mesure de qualité spécifique. Étant donné que l'on cherche à quantifier la qualité d'une image fournie en sortie d'un capteur, il n'est pas possible de procéder à une comparaison avec une image dite de référence. Nous nous intéresserons alors aux méthodes d'évaluation de la qualité des images sans-référence.

Dans cet article, nous présentons une comparaison des méthodes d'évaluation de la qualité des images, réputées les plus efficaces au sens de la corrélation avec le jugement humain, pour des images d'iris. Dans notre cas, le critère retenu sera la mesure de performance du système biométrique utilisé.

## 2 Qualité des images d’iris

La grande majorité des systèmes de reconnaissance biométrique basée sur l’iris (en proche infra-rouge ou dans le spectre visible) s’appuient sur des modèles de reconnaissance de forme. Les méthodes suivent le processus suivant qui est divisé en quatre étapes : 1) un processus de segmentation suivi 2) d’un processus de normalisation auquel s’ajoute 3) un processus d’extraction de caractéristiques pour terminer par 4) un algorithme de comparaison.

Les particularités physiologiques de l’iris induisent que les systèmes d’acquisition des images d’iris se font sans contact entre l’oeil et le capteur. Les systèmes biométriques associés sont calibrés en trouvant le bon compromis entre contraintes d’acquisition et précision de la reconnaissance. Dans la littérature, plusieurs travaux [6, 7] ont étudié l’impact de la qualité de l’acquisition dans le proche infra-rouge sur les performances du système biométrique. Cependant, hormis les travaux de Pronça *et al.* [5], il existe peu de travaux portant sur l’évaluation de la qualité des images d’iris dans le domaine visible.

## 3 Protocole d’évaluation

### 3.1 Qualité des images sans-référence

Il existe de nombreuses techniques d’évaluation de la qualité des images dès lors que l’on ne possède pas d’images de référence à laquelle il est possible de se comparer. Parmi toutes les techniques existantes, 10 algorithmes ont été étudiés en raison de leur fort taux de corrélation avec le jugement humain et de la disponibilité du code. Le tableau 1 présente les métriques retenues. Trois catégories de métriques de qualité sont retenues : 1) les mesures basées sur un apprentissage des statistiques de scènes naturelles (détails dans [8]), 2) les méthodes basées sur la mesure de l’entropie locale dans l’image et 3) des schémas algorithmiques basés sur des types de dégradation spécifique.

TABLE 1 – Algorithmes d’évaluation de la qualité des images sans référence testés

Métriques	commentaires
BIQI [8]	Apprentissage (SVM)
BRISQUE [8]	Apprentissage (modèle probabiliste)
BLLINDS2 [8]	Apprentissage (modèle probabiliste)
ILNIQUE [8]	Apprentissage (modèle probabiliste)
SSEQ [8]	Mesure des entropies spatiales et spectrales
AQIP [9]	Approche basée sur la mesure d’entropie
CONTRAST [10]	Mesure du contraste des images
PWN [11]	Approche basée sur la mesure du bruit local
JNBM [11]	Mesure basée sur le flou et la netteté
DCTPS [12]	Mesure du flou

### 3.2 Système biométrique et bases d’images

Le système biométrique de reconnaissance basé sur l’iris utilisé est OSIRIS (Open Source IRIS) qui a été développé dans le cadre du projet BioSecure [13].

La base de données des images d’iris a été spécifiquement développée dans le cadre de ces travaux. Les deux iris (gauche et droit) des sujets ont été photographiés en utilisant deux systèmes d’acquisition différents : un appareil photo Canon D700 de 18 mégapixels avec un macro objectif Canon EF 100mm f/2.8L ainsi qu’un appareil photo embarqué dans un téléphone Google Nexus 5 de 8 mégapixels. 15 échantillons par œil et par appareil photos ont été acquis. 31 sujets ont participé à la prise de photos avec l’appareil photo Canon D700, tandis que 25 sujets ont participé à la campagne d’acquisition avec le téléphone intelligent Google Nexus 5. Finalement, 1680 échantillons d’iris ont ainsi été collectés (930 à partir de l’appareil photo Canon et 750 à partir du téléphone). La figure 1 présente quelques exemples d’images d’iris obtenues.

### 3.3 Critères de performance

Afin de pouvoir quantifier l’impact de la qualité des échantillons biométriques sur la performance du système, le taux d’égale erreur (*Equal Error Rate*) EER est utilisé. IL s’agit d’un point de performance permettant de quantifier la performance des systèmes biométriques. L’EER est obtenue à partir de la courbe ROC lorsque la valeur du taux de fausse acceptation (*False Rate Acceptation*) FAR est égale à la valeur du taux de faux rejets FRR (*False Reject Rate*). Ces deux derniers taux dépendent d’un seul de décision qui est fixé en fonction du niveau de sécurité que l’on accorde au système biométrique. Plus le seuil sera bas et plus le taux de fausses acceptation sera élevé ; et inversement pour un seuil élevé. La courbe DET (*Detection Error Trade-off*), qui est une variante de la courbe ROC, permet d’évaluer les performances des algorithmes d’évaluation de la qualité de images. Dans le cas des courbes DET, les axes représentent les taux d’erreur du système et le taux de fausse non correspondance (*False Non Match Rate*) FNMR.

Dans cette étude, les courbes DET ainsi que les points de fonctionnement EER sont calculés en retirant un pourcentage des images dont les scores de qualité prédits par chaque métrique comme étant les plus faibles. Ainsi, quatre situations sont générées : la première correspond à la situation où la totalité des images est conservée, la deuxième (resp. troisième et quatrième) correspond à la situation où les images pour lesquelles les scores de qualités les situent dans les 75% (resp. 50% et 25%) meilleures sont conservées.

## 4 Résultats

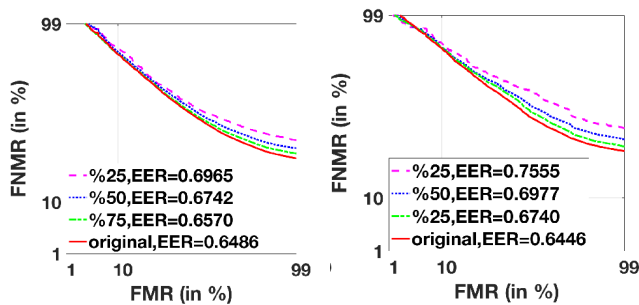
Les figures 2 (a) et (b) présentent deux exemples de courbe DET obtenues pour deux métriques de qualité testées, AQIP et BIQI, respectivement pour les images acquises en utilisant l’appareil photo Canon et le téléphone Google Nexus 5. Ces

TABLE 2 – Moyenned scores de qualité de chaque méthode d'évaluation pour les images capturées avec un téléphone Google Nexus 5 et un appareil photo numérique Canon D700

	BIQI	BRISQUE	BLLINDS2	ILNIQUE	SSEQ	AQIP	CONTRAST	PWN	JNBM	DCTPS
Téléphone	69.1	41.8	69.9	42.5	26.7	0.0014	3.3	0.163	12.7	3.21
Reflex	60.47	38.59	62.6	25.35	18.44	0.002	2.67	0.21	24.43	3.34



FIGURE 1 – Exemple d'iris photographiés. Les deux premières images ont été acquises avec l'appareil photo numérique Canon D700 et les deux dernières avec le téléphone Google Nexus 5.

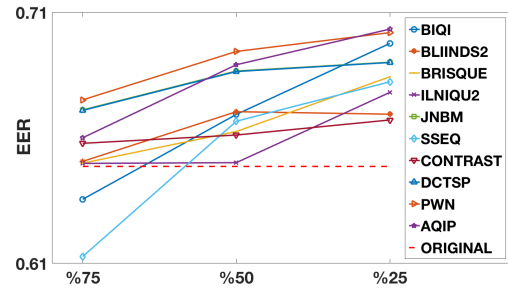


(a) Courbe DET pour la métrique (b) Courbe DET pour la métrique AQIP sur les images obtenues avec BIQI sur les images obtenues avec l'appareil photo

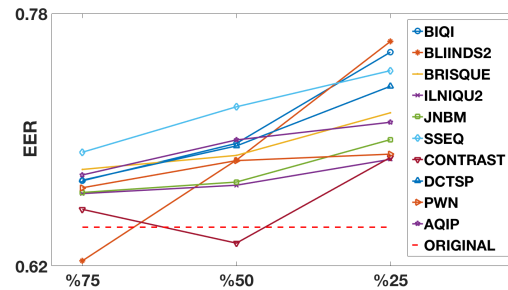
FIGURE 2 – Evolution des courbes DET en fonction du pourcentage des images non prises en compte pour les images obtenues avec (a) l'appareil photo et (b) le téléphone.

deux exemples sont représentatifs des courbes obtenues quelle que soit la métrique de qualité étudiée. On observe que lorsque l'on retire des images qui sont jugées de mauvaise qualité par la métrique considérée, les courbes DET ont tendance à avoir un point de fonctionnement (EER) plus élevé. Ceci peut-être interprété de plusieurs manières. Tout d'abord les métriques, qui ont été développées pour des images naturelles, ne sont pas nécessairement de bons indicateurs de la qualité des images d'iris pour un système de reconnaissance biométrique basé sur cette modalité. Deuxièmement, le système OSIRIS utilisé est suffisamment robuste pour accepter en entrée des images d'iris dégradées, dans une certaine limite.

Les figures 3 (a) et (b) présentent les valeurs EER obtenues par chaque méthode d'évaluation de la qualité des images obtenues par l'utilisation de l'appareil photo Canon D700 et du téléphone Google Nexus 5. On remarque que quel que soit l'appareil utilisé pour acquérir les images et quelle que soit la métrique de qualité appliquée, plus on retire d'images lors de la phase d'enrôlement, plus la valeur du point de fonctionnement EER augmente, ce qui est cohérent avec les constatations effectuées précédemment. On observe néanmoins une disparité



(a) EER obtenus à partir des images de l'appareil photo



(b) EER obtenus à partir des images du téléphone

FIGURE 3 – Évolution des valeurs EER en fonction du pourcentage des images non prises en compte pour les images obtenues avec (a) l'appareil photo et (b) le téléphone.

de fonctionnement des métriques.

Si l'on considère les images obtenues à partir d'un appareil photo (Figure 3(a)), nous pouvons noter que la métrique SSEQ permet d'avoir la valeur EER la plus faible lorsque l'on conserve 75% des images de meilleure qualité. La métrique permettant d'avoir la deuxième valeur EER la plus faible est la méthode BIQI. Finalement, on remarque que l'ensemble des méthodes d'évaluation de la qualité des images par apprentissage (BIQI, BRISQUE, BLIINDS2 et ILNIQUE) ont les valeurs d'EER les meilleures après SSEQ. Les métriques exploitant des types de dégradations spécifiques sans apprentissage ont les valeurs d'EER les plus élevées. Lorsque l'on passe à 50% d'image de meilleure qualité, on observe une dégradation des résultats. La méthode CONTRAST est celle qui présente la plus grande stabilité de la valeur EER quel que soit le pourcentage d'images conservées. ILNIQUE est la méthode qui a une meilleure stabilité en terme de valeur de l'EER que l'on conserve 75% ou 50% des échantillons de meilleure qualité.

À partir de la figure 3(b), on note que la méthode d'évaluation de la qualité qui a une valeur EER la plus faible lorsque l'on conserve 75% des images de meilleure qualité est BLIINDS2,

alors que pour 25% des images conservées, la valeur EER obtenue est la plus grande. La métrique CONTRAST permet d'avoir globalement une valeur EER la plus faible et ce quel que soit le pourcentage d'images conservées. Inversement, alors la méthode SSEQ donnait de bonnes performances sur les images acquises avec un appareil photographique, cette méthode fournit des valeurs d'EER parmi les plus élevées.

À la lecture conjointe des deux figures 3(a) et (b), plusieurs constatations peuvent être formulées :

1. Les remarques formulées pour les images obtenues à partir d'un appareil photo reflex ne peuvent pas être directement transposées pour les résultats obtenus sur des images capturées à l'aide d'un téléphone. En effet, nous observons un comportement erratique des métriques testées. Ceci peut-être dû au fait que le capteur numérique entre les deux appareils ne présentent les mêmes caractéristiques et donc influent directement sur la qualité des images acquises.
2. Les métriques qui sont basées sur une phase d'apprentissage (telles que BIQI, BRISQUE, BLINDS2 et ILNIQE) ont été entraînées sur des images naturelles et non des images d'iris. Or une image d'iris, même si elle peut être considérée comme une image naturelle dans une certaine mesure, est une image présentant des singularités spécifiques non apprises. Ce qui pourrait expliquer les performances modérées de ces métriques.
3. Chacune des métriques s'appuie sur une combinaison d'attributs qui ont été conçus pour des images naturelles. Or ces derniers peuvent ne pas être tous adaptés pour des images d'iris et de facto non pertinents pour le système de reconnaissance biométrique OSIRIS. Ceci semble être confirmé par le tableau 2 dans lequel on peut observer que les notes moyennes obtenues situent les images dans une gamme de qualité moyenne alors que les images n'ont subies aucune dégradation et sont réputées être de bonne qualité.

## 5 Conclusion

Diverses techniques d'évaluation de la qualité des images développées pour noter la qualité des images naturelles ont été testées sur des images d'iris. L'objectif est de vérifier dans quelles mesures elles sont adaptées pour prédire les performances d'un système biométrique de reconnaissance basé sur la modalité de l'iris. Les résultats obtenus montrent que les métriques étudiées ne sont pas adaptées pour filtrer les images d'iris tout en garantissant une qualité minimum en entrée d'un système de reconnaissance biométrique.

Les travaux futurs porteront sur la définition d'une méthode d'évaluation de la qualité spécifique aux images d'iris dans le spectre visible. Il convient de noter que même si les attributs des métriques testées ne sont pas tous adaptés, certains peuvent l'être pour la mesure de la qualité des images d'iris.

## Références

- [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] 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.
- [3] 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.*, p. 62020D, Int. Society for Optics and Photonics, 2006.
- [4] Y. Chen, S. C. Dass, and A. K. Jain, "Localized iris image quality using 2-d wavelets," in *International conference on biometrics*, pp. 373–381, Springer, 2006.
- [5] 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.
- [6] 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.*, pp. 1–7, IEEE, 2007.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] 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.
- [11] 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.
- [12] 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.
- [13] 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.