BULLETIN OF THE POLISH ACADEMY OF SCIENCES
TECHNICAL SCIENCES, Vol. 70(6), 2022, Article number: e143554
DOI: 10.24425/bpasts.2022.143554

CONTROL AND INFORMATICS

Hand-drawn face sketch recognition using rank-level fusion of image quality assessment metrics

Sami MAHFOUD[1]*, **Abdelhamid DAAMOUCHE**[1], **Messaoud BENGHERABI**[2], and **Abdenour HADID**[3]

[1] University M'Hamed Bougara of Boumerdes, Institute of Electrical and Electronic Engineering, Laboratory of Signals and Systems, Boumerdes, 35000, Algeria
[2] Center for Development of Advanced Technologies, P.O. Box 17 Baba-Hassen 16303, Algiers, Algeria
[3] Sorbonne University Abu Dhabi, Sorbonne Center for Artificial Intelligence, Abu Dhabi, UAE

Abstract. Face Sketch Recognition (FSR) presents a severe challenge to conventional recognition paradigms developed basically to match face photos. This challenge is mainly due to the large texture discrepancy between face sketches, characterized by shape exaggeration, and face photos. In this paper, we propose a training-free synthesized face sketch recognition method based on the rank-level fusion of multiple Image Quality Assessment (IQA) metrics. The advantages of IQA metrics as a recognition engine are combined with the rank-level fusion to boost the final recognition accuracy. By integrating multiple IQA metrics into the face sketch recognition framework, the proposed method simultaneously performs face-sketch matching application and evaluates the performance of face sketch synthesis methods. To test the performance of the recognition framework, five synthesized face sketch methods are used to generate sketches from face photos. We use the Borda count approach to fuse four IQA metrics, namely, structured similarity index metric, feature similarity index metric, visual information fidelity and gradient magnitude similarity deviation at the rank-level. Experimental results and comparison with the state-of-the-art methods illustrate the competitiveness of the proposed synthesized face sketch recognition framework.

Key words: face sketch recognition; synthesized face sketch; rank-level fusion; IQA metrics.

1. INTRODUCTION

Face Sketch Synthesis (FSS) plays a central role in both digital social entertainment and criminal investigations for law-enforcement [1, 2]. For instance, in public safety applications, it is common that a face photo of a suspect, captured by a surveillance camera, is in low resolution with pose/light variations, occlusion or even worse, it cannot be available due to its fraudulent concealment [3]. In such a practical scenario, the sketch drawn by forensic experts from video surveillance or according to the eyewitness descriptions is a valuable tool for reducing and deducting potential suspects from a large-scale mugshot database. However, directly matching the sketch drawn by forensic experts to the mugshot database using algorithms developed for face recognition performs poorly due to the great discrepancy in their texture and imaging modes [2]. Therefore, face sketch synthesis procedure is often used to reduce the large texture mismatch between mugshot photos and sketch images. An intelligent sketch-based face recognition system relies on automatic face sketch synthesis from photographs. The mugshot photos are transformed into sketches and then the probe sketches drawn by forensic experts are used to identify the suspect from synthesized sketch gallery [4].

*e-mail: s.mahfoud@univ-boumerdes.dz

Manuscript submitted 2022-05-09, revised 2022-08-09, initially accepted for publication 2022-10-22, published in December 2022.

FSS techniques can be grouped into two main categories, data-driven and model-driven [5]. The data-driven methods synthesize a face sketch from similar training sketch patches through a linear combination. The data-driven methods are divided into two classes. The first class synthesizes each sketch patch independently, e.g., the Locally Linear Embedding (LLE) [6] and the Spatial Sketch Denoising (SSD) [7]. The second class considers neighbouring constraints and mainly refers to probabilistic graphical model-based methods such as the Markov Random Field (MRF) [4] and the Markov Weight Field (MWF) [8]. The model-driven methods refer to models that learn a mathematical function offline from the training photo-sketch pairs, which map a photo (patch) to a sketch (patch). For instance, Wang and al. [9] utilized a linear regressor to learn the mapping from the training photo-sketch pairs, which synthesize a sketch at a breakneck speed.

On the other hand, deep learning has recently attracted increasing attention, and has been widely applied to close topics such as image super-resolution, image style transfer, image fusion and image classification [10–13]. For the problem in hand, Zhang *et al.* [14] developed a model based on a Convolutional Neural Network (CNN) to learn the end-to-end photo-sketch mapping. Recently, inspired by the significant success of Generative Adversarial Networks (GAN) [15] in image-to-image translation problems [16], various GAN-based FSS methods [2, 17–19] have achieved compelling progress. While the data-driven or exemplar methods suffer from detail loss, the model-driven methods are characterized by micro-details preservation at the cost of requiring more training data [5].

When face sketch synthesis is used for law enforcement, IQA metrics and face sketch recognition accuracy are usually considered as criteria to evaluate the face sketch synthesis performance [20]. IQA metric delivers a measure to estimate the quality of synthesized face sketch regarding the image distortions [21, 22]. Face photo-sketch recognition accuracy is used as an indirect way to assess the performance of face sketch synthesis methods [23]. Xiao and Gao [24] proposed an objective IQA, designed according to the consistency of face sketch-photo recognition, where they confirmed the premise that IQA metric should predict the matching performance and vice versa. The most recent FSS techniques target structure-consistence, identity preservation and realistic texture demonstrating high recognition rate and good visual quality [2, 17, 18, 25, 26]. Another promising direction of research in face sketch recognition was introduced in [27]. In this work, the authors introduce the issue of fusion at different levels of multiple stylistic sketches for suspect identification.

Inspired by the previous work and motivated by the progress made in full-reference image quality assessment achieved by the fusion of single quality metrics as demonstrated in [26, 28, 29], we propose in this paper a simple yet efficient face sketch recognition system based on the fusion of multiple full-reference image quality metrics. To validate the proposed approach and as a proof of concept, the training-free system proposed in [23] is taken as a baseline. The training-free feature is simply achieved by visual quality metrics through matching ground truth and synthesized sketches. The same recognition pipeline is used to evaluate the recognition performance of the recently proposed Bayesian Face Sketch Synthesis [30]. Our contribution in this paper is threefold. First, to the best of our knowledge, this is the first work that introduces the combination of multiple quality metrics for face sketch recognition task. Second, by embedding the fusion of multiple IQA metrics into the face recognition framework, the proposed approach simultaneously evaluates the performance of face sketch synthesis methods and conduct the face recognition application. Third, the proposed recognition strategy is benchmarked and compared favourably to the state-of-the-art methods, achieving a significant gain in performance. As it is expensive to collect large-scale photo sketches for training matchers, our proposed method is training-free, overcoming this issue.

The rest of this paper is organized as follows: the next section is dedicated to the overview of the Face Sketch Recognition framework using single Image Quality Assessment Metrics and the description of the four IQA metrics. In Section 3, we will present the proposed fusion strategy. Section 4 provides the experimental evaluation and results. Finally, the conclusion and perspectives are presented in Section 5.

2. REVISITING IMAGE QUALITY ASSESSMENT METRICS FOR FACE SKETCH RECOGNITION

Motivated by the strong correlation between face sketch recognition accuracy and IQA metrics, a training-free framework for sketch recognition is proposed in [23]. The rationale behind the IQA-based methods for sketch recognition is to compute the IQA scores between a probe sketch, which is taken as the reference image (ground truth), and each synthesized sketch in the gallery database, taken as the distorted image. IQA scores, obtained by the full-reference IQA metrics, are arranged in descending/ascending order according to the IQA metric utilized to form the list of matching identities. The synthesized face sketch which obtains the best IQA score, in the nearest neighbor sense, is identified as the suspect.

In this training-free recognition framework, four data-driven sketch synthesis methods are applied to transform mugshot photos into sketches, namely, LLE , SSD , MRF and MWF methods. To conduct the matching operation, four common full-reference image quality assessment metrics are explored: the Structured Similarity Index Metric (SSIM) [21], the Visual Information Fidelity (VIF) [31], the Feature Similarity Index Metric (FSIM) [32] and the Gradient Magnitude Similarity Deviation (GMSD) [22]. On the one hand, the four selected IQA metrics are widely adopted in the literature and have been shown to provide good performance in many previous works [28, 29, 33]. On the other hand, any existing IQA metric can be easily embedded into our proposed framework to further enhance the performance of the face sketch matching. It is also important to note that all these IQA metrics are designed for gray-scale images.

Hereafter, I_r and I_t denote the probe sketch image drawn by the experts and its corresponding synthesized sketch, respectively.

2.1. Structural Similarity Index Metric

The SSIM [21] algorithm performs similarity measurement in three steps: luminance comparison l, contrast comparison c, and structure comparison s. Their respective formulas are provided in (1), (2) and (3):

$$l = \frac{2\mu_r\mu_t + T_1}{\mu_r^2 + \mu_t^2 + T_1}, \tag{1}$$

$$c = \frac{2\sigma_r\sigma_t + T_2}{\sigma_r^2 + \sigma_t^2 + T_2}, \tag{2}$$

$$s = \frac{\sigma_{r,t} + T_3}{\sigma_r\sigma_t + T_3}, \tag{3}$$

where μ_r, μ_t represent the mean intensities of probe sketch image and its synthesized sketch, respectively. σ_r, σ_t are the variances of the probe sketch image and its synthesized sketch respectively, and $\sigma_{r,t}$ is the covariance. T_1, T_2, and T_3 are positive stabilizing constants chosen to prevent the denominator from becoming too small. The overall similarity is a function of combination:

$$SSIM = [l(I_r,I_t)]^{\alpha}\,[c(I_r,I_t)]^{\beta}\,[s(I_r,I_t)]^{\gamma}, \tag{4}$$

where α, β, and γ are positive constants chosen to indicate the relative importance of each component.

2.2. Visual Information Fidelity

The VIF [31] algorithm models natural images in the wavelet domain using Gaussian scale mixtures (GSMs). It quantifies how much information is preserved in the distorted image from the reference image. VIF measure consists of three components: source model, distortion model, and human visual system (HVS) model. The VIF quality measure is calculated as follows:

$$VIF = \frac{\sum\limits_{j \in subbands} I_t(C^j; F^j \mid z^j)}{\sum\limits_{j \in subbands} I_r(C^j; E^j \mid z^j)}, \tag{5}$$

where j is the subband index, and $I_t(C^j;F^j \mid z^j)$ and $I_r(C^j;E^j \mid z^j)$ are the corresponding mutual information of the j-th subband in the synthesized sketch image I_t and its reference sketch image I_r, respectively.

2.3. Feature Similarity Index Metric

The FSIM [32] assumes that HVS understands an image mainly according to its low-level features. Therefore, it computes the quality estimates based on the phase congruency PC as the primary feature, and incorporates the gradient magnitude GM as a complementary feature. The FSIM similarity between the I_r and I_t can be represented as:

$$FSIM = \frac{\sum\limits_{x \in \Omega} S_{PC}.S_G.\max(PC_r(x), PC_t(x))}{\sum\limits_{x \in \Omega} \max(PC_r(x), PC_t(x))}, \tag{6}$$

where $PC_r(x)$ and $PC_t(x)$ are PC maps computed for I_r and I_t respectively at location x, Ω is the image spatial domain. S_{PC} and S_G are defined as follows:

$$S_{PC} = \frac{2PC_r(x).PC_t(x) + T1}{PC_r^2(x).PC_t^2(x) + T1}, \tag{7}$$

$$S_G = \frac{2GM_r(x).GM_t(x) + T2}{GM_r^2(x).GM_t^2(x) + T2}, \tag{8}$$

where $GM_r(x)$ and $GM_t(x)$ are GM maps for I_r and I_t at location x, respectively. $T1$ and $T2$ are positive constants utilized to increase the stabilities of S_{PC} and S_G, respectively.

2.4. Gradient Magnitude Similarity Deviation

The GMSD [22] focuses on computational efficiency of the quality prediction, by simply computing pixel-wise Gradient Magnitude Similarity (GMS) followed by applying standard deviation pooling to the GMS map as Gradient Magnitude Similarity Mean (GMSM). The GMSD similarity between the I_r and I_t is computed as follows:

$$GMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(GMS(i) - \frac{1}{N}\sum_{i=1}^{N} GMS(i)\right)^2}, \tag{9}$$

where the gradient magnitude similarity $GMS(i)$ between I_r and I_t at location i is computed as follows:

$$GMS(i) = \frac{2GM_r(i)GM_t(i) + c}{GM_r^2(i)GM_t^2(i) + c}, \tag{10}$$

where $GM_r(i)$ and $GM_t(i)$ are the GM maps for I_r and I_t at location i, respectively. c is a positive constant that supplies numerical stability.

3. FUSION OF IQA METRICS FOR SYNTHESIZED FACE SKETCH RECOGNITION

Although numerous IQA Metrics have been proposed during the last years, there is no single quality measure that significantly beats others [29]. Different fusion-based methods are assessed to exploit the diversity and complementarity of different metrics in the context of image quality assessment [33, 34]. Considering the IQA measures presented in Section 2, our proposed face sketch recognition framework is based on their rank-level fusion.

Rank-level fusion has been discussed in [35,36] including the highest rank method, the Borda count method and the weighted Borda count method. The first two methods do not employ any statistical information about the matcher performance in the combination process. In other words, these two approaches do not require a training phase. In contrast to the weighted Borda count method, which requires a training phase to calculate weights for different matchers. Moreover, the advantage of the Borda count method over the highest rank one is its ability to account for the variability in ranks due to the use of a large number of classifiers while the highest rank method suffers from its inherited tie problem. Therefore, we choose Borda count for fusion in our work.

The rank-level combination using Borda count is the most commonly used method for unsupervised rank-level fusion. In Borda count method, the fused rank is calculated as the sum of the ranks of individual matchers:

$$R_i = \sum_{i=1}^{c} r_{i,j}, \tag{11}$$

where $r_{i,j}$ denotes the rank of j^{th} sample using i^{th} matcher, c refers to the number of used matchers.

The Borda count method using, equation (11), assumes that the IQA matchers are statistically independent and that all of them perform well [35]. However, this assumption may not always held as the IQA metrics used in different matchers are extracted from the same sketch probe. This makes Borda count method highly vulnerable to the effect of weak classifiers. To enhance the performance of Borda count method, the Nanson function [37], known also as the Borda elimination, is used. One of the Nanson function properties is to eliminate the weakest rank, i.e.,

$$\max r_{i,j} = 0. \tag{12}$$

In this implementation, the weakest rank is therefore firstly eliminated and then the regular Borda count is computed on the remaining ranks equation (11).

The block diagram of the proposed framework is depicted in Fig. 1. Firstly, all mugshot photos are synthesized into sketches by face sketch algorithms, e.g., LLE [6]. Subsequently, given

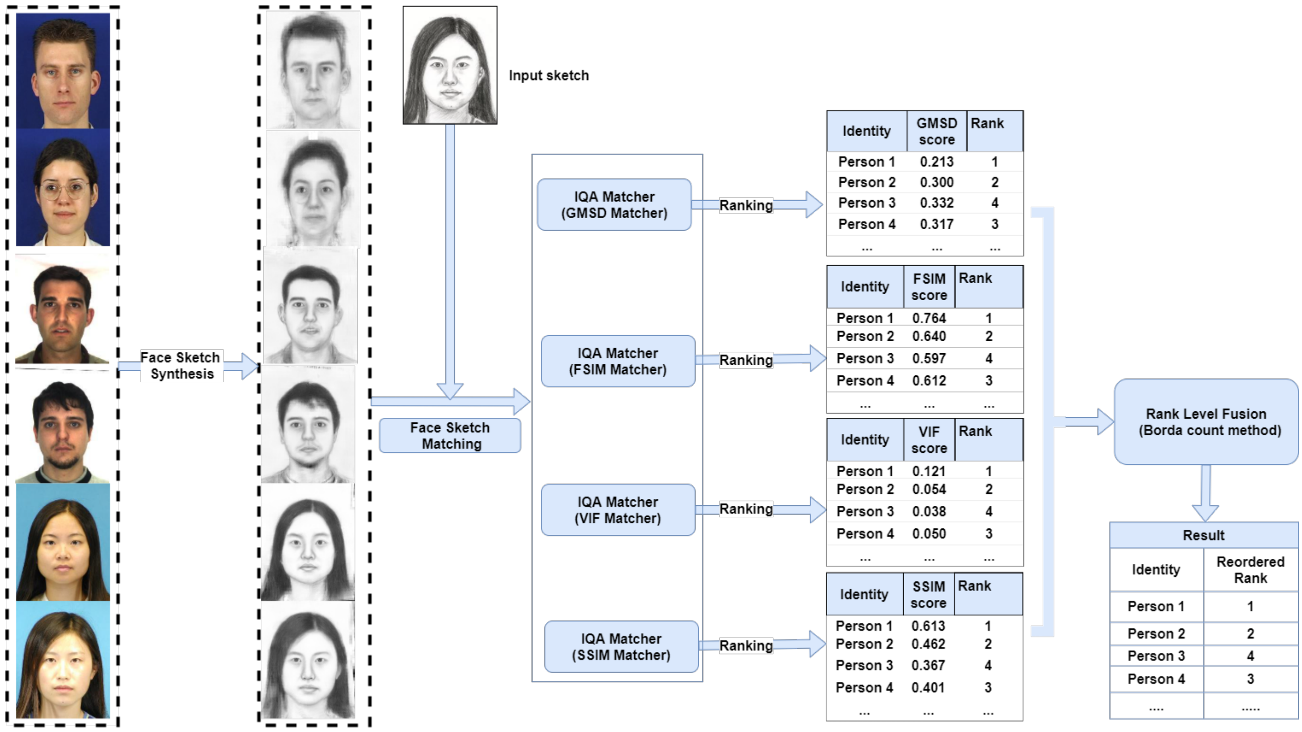

Fig. 1. Block diagram of the proposed method for face sketch recognition. Four IQA-based match scores are computed and fused using Borda count method

a probe sketch, taken as the reference image and each synthesized sketch in gallery, taken as the distorted image, four matching sores are computed by the considered IQA full-reference metrics. The previous step produces a ranking list for each IQA matcher. Finally, the ranked lists of the different individual IQA matchers are fused using Borda count method.

The proposed framework offers two major advantages; it improves the performance of face sketch recognition based on a single IQA metric and it has the desirable property of being training free, as it is expensive to collect photo sketches at large scale for matchers learning.

4. EXPERIMENTAL RESULTS

In our experiments, we adopted the full-reference IQA baseline presented in Section 2. We followed the same experimental protocol and we used the same benchmark. Face sketches used in this benchmark are from the Chinese University of Hong Kong (CUHK) face sketch database (CUFS). It is composed of three sub-datasets, i.e., the Chinese University of Hong Kong (CUHK) student database [38] which includes 188 faces, the AR database [39] with 123 faces, and 295 faces from the XM2VTS database [40]. For each face, there is a face sketch drawn by an artist based on the corresponding photo.

An illustrative example of the synthesized face sketch by the four exemplar-based methods considered in this work is shown in Fig. 2. To conduct training-free recognition, 338 synthesized face sketches (100 CUHK + 43 AR + 195 XM2VTS) were used as the probe set, and the corresponding ground-truth sketches drawn by the artist were taken as the gallery set (it is also the reference image for IQA). Table 1 shows the recognition rate (rank-1 accuracy) of each synthesis method using both single IQA and our framework dubbed IQA-fusion. From Table 1, it can be seen that the proposed recognition framework achieves a substantial gain in performance compared to single IQA matchers for all the four synthesis techniques. This gain is more pronounced for the LLE with an absolute gain of 6.8% w.r.t to the best IQA matcher achieved by FSIM. An important observation

Fig. 2. Synthesized face sketch examples by four data-driven methods on three datasets. The first column is the input photo and the second column is the corresponding sketch drawn by the artist. The third to the last column are the results of LLE, SSD, MRF and MWF methods. These three face photos are from the CUHK Student, AR, and XM2VTS dataset respectively.

to highlight is that the MWF synthesis technique which yields the best recognition accuracy for all the IQA metrics beneficiates also from the IQA-fusion with an absolute gain of 4.73% w.r.t to the best IQA matcher (VIF). This gain in performance can be justified by the diversity and complementarity of single IQA matchers.

This gain in performance is also manifested by the Cumulative Match Characteristic (CMC) curves shown in Fig. 3. The outperformance of the IQA-fusion approach especially among first several ranks is apparent.

Table 1
The recognition accuracy of the single IQA-based methods and the IQA-fusion framework. The best recognition rates are in bold

Method	SSIM	VIF	FSIM	GMSD	Ours
LLE (%)	75.15	82.54	82.84	81.95	**88.17**
SSD (%)	78.99	84.02	74.85	75.44	**85.21**
MRF (%)	78.11	74.26	83.43	77.22	**85.50**
MWF (%)	82.54	84.91	84.02	84.62	**89.64**

It is worth noting that the proposed fusion framework is tested on four classical exemplar-based synthesis echniques. To further demonstrate the effectiveness of our proposal, we compare the CMC using the proposed IQA-fusion strategy applied on State Of The Art (SOTA) sketch synthesis technique based on Generative Adversarial Learning (GAN) for generating a structure-consistent and realistic texture sketch proposed in [17], with two face sketch recognition systems also based on SOTA FSS techniques while adopting the same testing protocol used in our study. The first system is based on the Bayesian sketch synthesis proposed recently in [30], the matching is based on the SSIM and VIF matchers. And the second system is utilized the GAN for detail-preserving proposed in [19], and the recognition is based on the pretrained ResNet-50 model (trained on VGGface2 database) [41] for face embedding and the Euclidean distance to measure the feature similarity.

Table 2 shows the CMC from rank-1 to rank-10 of the proposed IQA-fusion applied on LLE, MWF and GAN-based FSS

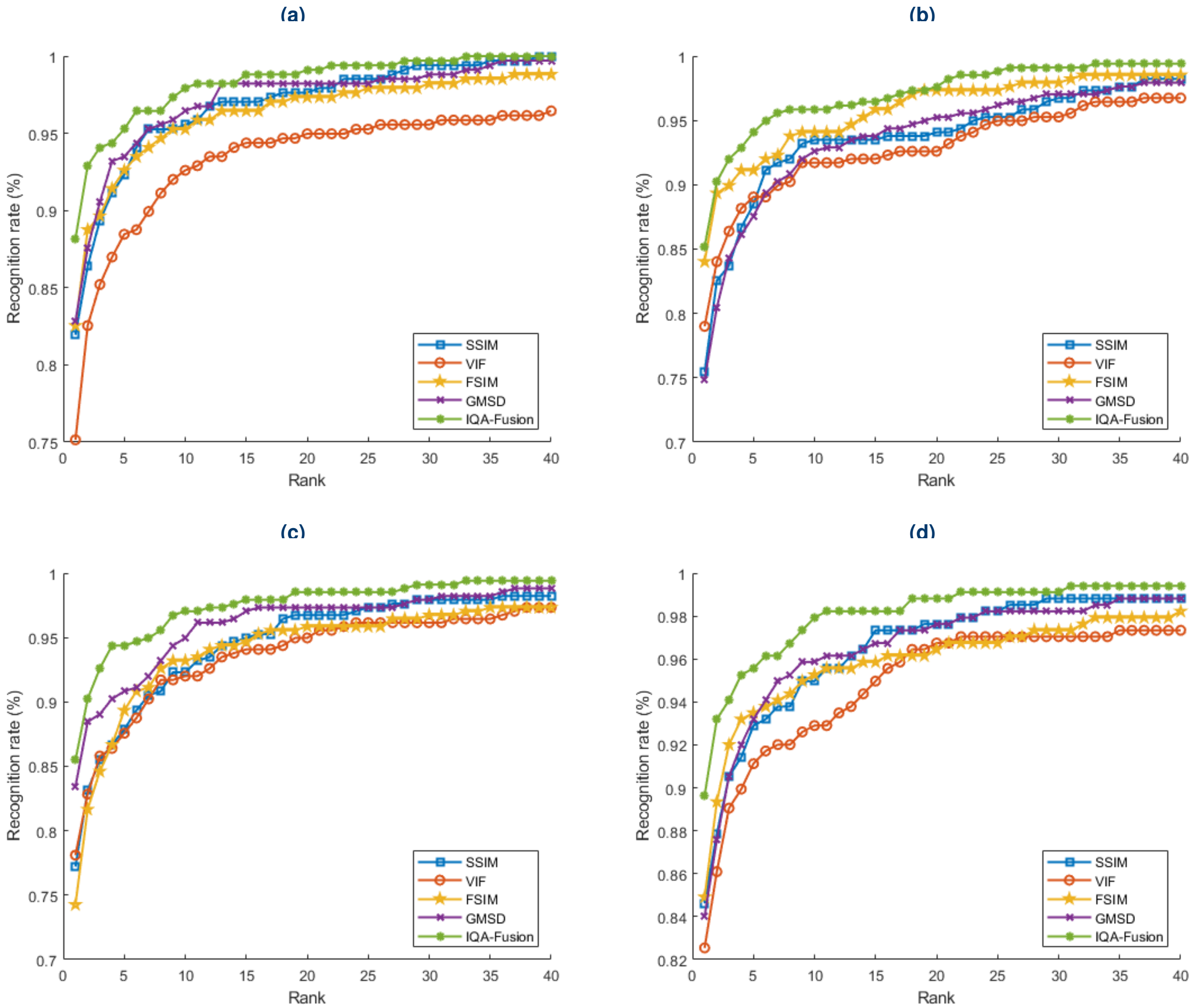

Fig. 3. CMC curves using single and the proposed fusion of IQA-based methods on face sketches generated by (a) LLE, (b) SSD, (c) MRF and (d) MWF

Table 2

CMC using the proposed IQA-fusion on face sketch synthesized by LLE, MWF and GAN-based method respectively, the Bayesian face sketch synthesis using SSIM and VIF matchers and finally GAN-based method using ResNet-50 trained on VGGFace2. The best recognition rates are in bold

Rank	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
LLE(%) (IQA-Fusion (Ours))	89.64	93.79	94.67	94.97	94.67	95.27	95.86	96.15	96.75	97.63
MWF(%) (IQA-Fusion (Ours))	89.64	92.01	93.79	94.97	95.86	96.75	97.34	97.34	97.34	98.22
Bayesian(%) (SSIM) [30]	90.53	93.79	94.67	95.56	95.86	97.34	97.34	97.63	97.63	97.63
Bayesian (%) (VIF) [30]	89.35	93.20	94.08	96.45	97.63	**98.52**	**98.52**	**98.82**	98.82	98.82
GAN(%) (ResNet-50) [19]	83.4	/	/	/	96.2	/	/	/	/	97.9
GAN(%) (IQA-Fusion (Ours))	**94.97**	**96.75**	**97.93**	**97.93**	**98.52**	**98.52**	**98.52**	98.52	**99.11**	**99.11**

methods, respectively and the two compared sketch recognition systems. From Table 2, it can be seen that the proposed IQA-fusion framework makes the classical sketch generation technique, represented by LLE and MWF, competitive with the advanced Bayesian method using a single IQA for matching and compares favourably with the recognition framework using state-of-the-art architecture dedicated to face recognition and applied on GAN-based FSS technique. Whereas our proposed method using GAN-based FSS outperforms all the other approaches by a high margin. The last observation confirms that despite the research effort to make the synthesized face more vivid and realistic, it still suffers from shape exaggeration. In contrast, the state-of-the-art deep learning models are generally characterised by texture bias, especially when considering the last layers for recognition task [42].

Besides the face sketch recognition application, the proposed IQA-fusion framework could also be utilized to compare the performance of different face sketch synthesis techniques. In Table 2, considering, the CMC of the three FSS techniques to which the proposed IQA-Fusion framework was applied, it can be seen that the model-driven methods represented by GAN-based FSS perform better than the data-driven methods represented by LLE and MWF, respectively. Furthermore, embedding customized IQA metrics, particularly those designed for synthesized face sketches, such as [43], may improve the framework proposed even better. Actually, the proposed fusion framework could be generalized to other heterogeneous face matching applications [44] such as matching visible light to near-infrared (VIS-NIR) face images and high-resolution to low-resolution (HR-LR) face images.

5. CONCLUSIONS

In this paper, we proposed a training-free face sketch recognition method based on the rank-level fusion of image quality assessment metrics. The Borda count method is employed to fuse four full reference assessment metrics for implementing the proposed framework, i.e., the SSIM, VIF, FSIM, and GMSD methods. We utilized five face sketch synthesis techniques to transform the photos in the database into sketches, including four Data-driven methods and one Model-driven method. Experimental results illustrate that the proposed pipeline outperforms the single IQA-based matchers. We also compared the proposed fusion IQA metrics-based recognition framework with two state-of-the-art sketch recognition frameworks based on two recently proposed face sketch synthesis techniques, where the obtained recognition accuracies confirm the effectiveness of our proposal. Our work constitutes a preliminary study on the topic of the fusion of multiple quality metrics for face sketch recognition. The application of the proposed fusion framework on recent face sketch synthesis techniques and the investigation of the current IQA metrics is an immediate perspective of this work. In addition, this paper highlights the need to consider the fusion and combination of IQA metrics as a loss function during the process of face sketch synthesis.

To support the principle of reproducible research, we make our code publicly available through Github repository named IQA-Fusion.

DECLARATION OF COMPETING INTERESTS

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ACKNOWLEDGEMENTS

This work is financially supported by the DGRSDT (General Directorate of Scientific Research and Technological Development) and CDTA (Center for Development of Advanced Technologies) Algeria, as part of the CDTAs triennial research protocol 2019–2021 (Project code: N004-1/BIOSMC/DTELE COM/CDTA/PT 19-21).

CODE, DATA, AND MATERIALS AVAILABILITY

The code and data to generate the results, tables and figures in this work are available in a Github repository at https://github.com/SamiMahfoud/IQA-Fusion.

REFERENCES

[1] N. Balayesu and H.K. Kalluri, "An extensive survey on traditional and deep learning-based face sketch synthesis models," *Int. J. Inf. Technol.*, vol. 12, no. 3, pp. 995–1004, Nov 2020, doi: 10.1007/s41870-019-00386-8.

[2] Y. Fang, W. Deng, J. Du, and J. Hu, "Identity-aware cyclegan for face photo-sketch synthesis and recognition," *Pattern Recognit.*, vol. 102, p. 107249, Jun 2020, doi: 10.1016/j.patcog.2020.107249.

[3] J.C. Klontz and A.K. Jain, "A case study of automated face recognition: The boston marathon bombings suspects," *Computer*, vol. 46, no. 11, pp. 91–94, Nov 2013, doi: 10.1109/MC.2013.377.

[4] X. Wang and X. Tang, "Face photo-sketch synthesis and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 11, pp. 1955–1967, Nov 2009, doi: 10.1109/TPAMI.2008.222.

[5] P. Li, B. Sheng, and C.L.P. Chen, "Face sketch synthesis using regularized broad learning system," *IEEE Trans. Neural Networks Learn. Syst.*, pp. 1–15, 2021, doi: 10.1109/TNNLS.2021.3070463.

[6] Q. Liu, X. Tang, H. Jin, H. Lu, and S. Ma, "A nonlinear approach for face sketch synthesis and recognition," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1. IEEE, 2005, pp. 1005–1010, doi: 10.1109/CVPR.2005.39.

[7] Y. Song, L. Bao, Q. Yang, and M.-H. Yang, "Real-time exemplar-based face sketch synthesis," in *Computer Vision – ECCV 2014*. Springer International Publishing, 2014, pp. 800–813, doi: 10.1007/978-3-319-10599-451.

[8] H. Zhou, Z. Kuang, and K.K. Wong, "Markov weight fields for face sketch synthesis," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, Jun 2012, pp. 1091–1097, doi: 10.1109/cvpr.2012.6247788.

[9] N. Wang, M. Zhu, J. Li, B. Song, and Z. Li, "Data-driven vs. model-driven: Fast face sketch synthesis," *Neurocomputing*, vol. 257, pp. 214–221, Sep 2017, doi: 10.1016/j.neucom.2016.07.071.

[10] A. Cichocki, T. Poggio, S. Osowski, and V. Lempitsky, "Deep learning: Theory and practice," *Bull. Pol. Acad. Sci. Tech. Sci.*, vol. 66, no. 6, pp. 757–759, 2018, doi: 10.24425/bpas.2018.125923.

[11] X. Yang, Y. Zhang, and D. Zhou, "Deep networks for image super-resolution using hierarchical features," *Bull. Pol. Acad. Sci. Tech. Sci.*, vol. 70, no. 1, p. e139616, 2022, doi: 10.24425/bpasts.2021.139616.

[12] K. Hawari and I. Ismail, "The automatic focus segmentation of multi-focus image fusion," *Bull. Pol. Acad. Sci. Tech. Sci.*, vol. 70, no. 1, p. e140352, 2022, doi: 10.24425/bpasts.2022.140352.

[13] M. Grochowski, A. Kwasigroch, and A. Mikołajczyk, "Selected technical issues of deep neural networks for image classification purposes," *Bull. Pol. Acad. Sci. Tech. Sci.*, vol. 67, no. 2, pp. 363–376, 2019, doi: 10.24425/bpas.2019.128485.

[14] L. Zhang, L. Lin, X. Wu, S. Ding, and L. Zhang, "End-to-end photo-sketch generation via fully convolutional representation learning," in *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval*. ACM, jun 2015, pp. 627–634, doi: 10.1145/2671188.2749321.

[15] I. Goodfellow *et al.*, "Generative adversarial nets," *Adv. Neural Inf. Process. Syst.*, vol. 27, 2014, doi: 10.1145/3422622.

[16] P. Isola, J.-Y. Zhu, T. Zhou, and A.A. Efros, "Image-to-image translation with conditional adversarial networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Jul 2017, pp. 1125–1134, doi: 10.1109/cvpr.2017.632.

[17] X. Li, F. Gao, and F. Huang, "High-quality face sketch synthesis via geometric normalization and regularization," in *2021 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, Jul 2021, pp. 1–6, doi: 10.1109/ICME51207.2021.9428348.

[18] J. Yu, X. Xu, F. Gao, S. Shi, M. Wang, D. Tao, and Q. Huang, "Toward realistic face photo–sketch synthesis via composition-aided GANs," *IEEE Trans. Cybern.*, vol. 51, no. 9, pp. 4350–4362, Sep 2021, doi: 10.1109/tcyb.2020.2972944.

[19] W. Wan, Y. Yang, and H.J. Lee, "Generative adversarial learning for detail-preserving face sketch synthesis," *Neurocomputing*, vol. 438, pp. 107–121, May 2021, doi: 10.1016/ j.neucom.2021.01.050.

[20] N. Wang, X. Gao, J. Li, B. Song, and Z. Li, "Evaluation on synthesized face sketches," *Neurocomputing*, vol. 214, pp. 991–1000, nov 2016, doi: 10.1016/j.neucom.2016.06.070.

[21] 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, doi: 10.1109/tip.2003.819861.

[22] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 684–695, Feb 2014, doi: 10.1109/tip.2013.2293423.

[23] N. Wang, J. Li, L. Sun, B. Song, and X. Gao, "Training-free synthesized face sketch recognition using image quality assessment metrics," *arXiv preprint arXiv:1603.07823*, 2016, ArXiv:1603.07823.

[24] B. Xiao and X. Gao, "Visual quality assessment of the synthesized sketch," in *2013 Ninth International Conference on Natural Computation (ICNC)*. IEEE, Jul 2013, pp. 317–321, doi: 10.1109/icnc.2013.6817993.

[25] Y. Lin, K. Fu, S. Ling, J. Wang, and P. Cheng, "Toward identity preserving face synthesis between sketches and photos using deep feature injection," *IEEE Trans. Ind. Inf.*, vol. 18, no. 1, pp. 327–336, jan 2022, doi: 10.1109/tii.2021.3074989.

[26] Y. Peng, J. Xu, Z. Luo, W. Zhou, and Z. Chen, "Multi-metric fusion network for image quality assessment," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE, Jun 2021, pp. 1857–1860, doi: 10.1109/cvprw 53098.2021.00205.

[27] C. Peng, X. Gao, N. Wang, and J. Li, "Face recognition from multiple stylistic sketches: Scenarios, datasets, and evaluation," *Pattern Recognit.*, vol. 84, pp. 262–272, Dec 2018, doi: 10.1016/j.patcog.2018.07.014.

[28] M. Oszust, "Decision fusion for image quality assessment using an optimization approach," *IEEE Signal Process. Lett.*, vol. 23, no. 1, pp. 65–69, Jan 2016, doi: 10.1109/LSP.2015.2500819.

[29] K. Ding, K. Ma, S. Wang, and E. P. Simoncelli, "Comparison of full-reference image quality models for optimization of image processing systems," *Int. J. Comput. Vision*, vol. 129, no. 4, pp. 1258–1281, Jan 2021, doi: 10.1007/s11263-020-01419-7.

[30] N. Wang, X. Gao, L. Sun, and J. Li, "Bayesian face sketch synthesis," *IEEE Trans. Image Process.*, vol. 26, no. 3, pp. 1264–1274, Mar 2017, doi: 10.1109/TIP.2017.2651375.

[31] H. Sheikh and A. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, Feb 2006, doi: 10.1109/tip.2005.859378.

[32] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug 2011, doi: 10.1109/tip.2011.2109730.

[33] K. Okarma, P. Lech, and V.V. Lukin, "Combined full-reference image quality metrics for objective assessment of multiply distorted images," *Electronics*, vol. 10, no. 18, p. 2256, sep 2021, doi: 10.3390/electronics10182256.

[34] S. Athar and Z. Wang, "A comprehensive performance evaluation of image quality assessment algorithms," *IEEE Access*, vol. 7, pp. 140 030–140 070, 2019, doi: 10.1109/ACCESS.2019.2943319.

[35] A. Kumar and S. Shekhar, "Personal identification using multibiometrics rank-level fusion," *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, vol. 41, no. 5, pp. 743–752, Sep 2011, doi: 10.1109/TSMCC.2010.2089516.

[36] A. Abaza and A. Ross, "Quality based rank-level fusion in multibiometric systems," in *2009 IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*. IEEE, Sep 2009, pp. 1–6, doi: 10.1109/BTAS.2009.5339081.

[37] E.M.S. Niou, "A note on nanson's rule," *Public Choice*, vol. 54, no. 2, pp. 191–193, 1987, doi: 10.1007/BF00123006.

[38] X. Tang and X. Wang, "Face photo recognition using sketch," in *Proceedings. International Conference on Image Processing*, vol. 1. IEEE, 2002, pp. I–I, doi: 10.1109/ICIP.2002.1038008.

[39] A.M. Martinez, "The ar face database," *CVC Technical Report24*, 1998. [Online]. Available: https://www2.ece.ohio-state.edu/~aleix/ARdatabase.html

[40] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre, "Xm2vtsdb: The extended m2vts database," in *Second International Conference on Audio and Video-Based Biometric Person Authentication*, vol. 964, 1999, pp. 965–966. [Online]. Available: https://www.semanticscholar.org/paper/b62628ac06bbac998a3ab825324a41a11bc3a988

[41] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman, "VGGFace2: A dataset for recognising faces across pose and age," in *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*. IEEE, May 2018, pp. 67–74, doi: 10.1109/FG.2018.00020.

[42] K.L. Hermann, T. Chen, and S. Kornblith, "The origins and prevalence of texture bias in convolutional neural networks," *arXiv preprint arXiv:1911.09071*, 2019, arxiv.org/abs/1911.09071v3.

[43] D.-P. Fan, S. Zhang, Y.-H. Wu, Y. Liu, M.-M. Cheng, B. Ren, P. Rosin, and R. Ji, "Scoot: A perceptual metric for facial sketches," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE, oct 2019, pp. 5612–5622, doi: 10.1109/iccv.2019.00571.

[44] M. Cho, T. Kim, I.-J. Kim, K. Lee, and S. Lee, "Relational deep feature learning for heterogeneous face recognition," *IEEE Trans. Inf. Forensics Secur.*, vol. 16, pp. 376–388, 2021, doi: 10.1109/TIFS.2020.3013186.