Robust Face Authentication Based on Dynamic Quality-weighted Comparison of Visible and Thermal-to-visible images to Visible Enrollments

Khawla Mallat[§], Naser Damer[⋆†], Fadi Boutros[⋆†], Jean-Luc Dugelay[§]
[§]Digital Security Department, EURECOM, Sophia Antipolis, France
[⋆]Fraunhofer Institute for Computer Graphics Research IGD, Darmstadt, Germany
[†]Mathematical and Applied Visual Computing, TU Darmstadt, Darmstadt, Germany
Email: khawla.mallat@eurecom.fr

Abstract—We introduce, in this paper, a new scheme of score level fusion for face authentication from visible and thermal face data. This proposed scheme provides a fast and straightforward integration into existing face recognition systems and does not require recollection of enrollment data in thermal spectrum. In addition to be used as a possible countermeasure against spoofing, this paper investigates the potential role of thermal spectrum in improving face recognition performances when employed under adversarial acquisition conditions. We consider a context where individuals have been enrolled solely in visible spectrum, and their identity will be verified using 2 sets of probes: visible and thermal. We show that the optimal way to proceed is to synthesis a visible image from the thermal face in order to create a synthetic-visible probe; and then to fuse scores resulting from comparisons between visible gallery with both visible probe and synthetic-visible probe. The thermal-to-visible face synthesis is performed using a Cascaded Refinement Network (CRN) and face features were extracted and matched using LightCNN and Local Binary Patterns (LBP). The fusion procedure is performed based on several quality measures computed on both visible and thermal-to-visible generated probes and compared to the visible gallery images.

Index Terms—Score fusion, thermal-to-visible face synthesis, quality assessment

I. Introduction

The growing necessity for digital and physical security has spreadingly led to the deployment of biometric systems. Particularly, face recognition has received a lot attention these last decades for its wide range of applications, from law enforcement and security systems to increasing everyday life safety. This is mainly motivated by the fact that face recognition is considered as a fast, passive and non intrusive systems compared to other biometric traits. Nevertheless, it is important to admit that face recognition technology is still not fully reliable, giving that it is still encountering challenges due to poor data quality, missing information and diverse threats. Taking into account that most face recognition systems are based on visible spectrum, variable or low illumination conditions have been proved to be one of the major challenges. Moreover, face recognition technology is also threatened by presentation attacks that endeavour to spoof the system. Some prompt actions have been taken such as requiring an eye blink, smile or other visual reaction to prove the liveness of the user, yet this can be easily tricked using video replay attacks. More advanced solutions, developed in order to detect face presentation attack, are available in [1]. Although, face recognition systems are extensively implemented for border and access control and surveillance systems. Thereby, it is necessary to seek solutions of presentation attack detection that are cost effective and easy to integrate with existing face recognition systems. Thermal imagery is considered a natural spoofing countermeasure, as the heat emitted by the human face provide evidence of the user's liveness. Thermal imagery technology has drastically advanced during the last couple of decades, and thermal cameras have evolved to become affordable and user friendly. Even though thermal imaging solutions are significantly advancing, they still suffer from poor performances due to low image resolution, lack of color, and poor texture and geometric information. Several studies of face recognition have thoroughly focused on bringing visible and thermal spectra together to benefit from the advantages of each. Particularly, cross-spectrum face recognition, aims in our case of study to identify a person imaged in thermal spectrum from a gallery containing face images acquired in the visible spectrum. We draw focus on image synthesis strategy for cross-spectrum face recognition [2]–[4], consisting in generating visible-like images from thermal captures. These generated visible-like images will be then matched against a gallery of visible faces. Opting for this strategy is essential to enable its integration in the existing face recognition systems as well as manual face verification by human examiners. It has been revealed that synthesis based cross-spectrum face recognition outperformed visible spectrum based face recognition systems when engaged in poorly lit environments. Although, generated images are still far from optimal compared to visible images when confronting other sorts of variations, such as occlusion or head pose variations. In an attempt to achieve a more

robust and accurate face recognition system, we suggest to fuse scores obtained while matching visible face probes with visible face gallery and the scores obtained by matching thermal-to-visible generated images from thermal face against the same visible face gallery. Based on the intuition that quality can be indicative of the utility of a face sample, we propose score fusion scheme weighted by quality similarity scores.

Conventional multi-modal fusion systems usually perform the authentication process separately on each modality considering gallery and probe samples from the same modality. However, the novelty of this work lies in considering solely one set of gallery images acquired in visible spectrum. This is motivated by the need of a fast and convenient integration of thermal imagery in existing face recognition systems, as it has been proven effective for nighttime face recognition. Quality-weighted score fusion of visible and thermal spectrum will ensure a robust face authentication operational at daytime and nighttime.

The remainder of the paper is organised as follows. Section II presents the related work in thermal-to-visible face synthesis and quality based fusion. Section III describes the proposed method of quality-weighted comparison of visible and thermal-to-visible images to visible enrollments. Section IV reports the experimental setup, followed by results and discussion. Conclusions are presented in Section V.

II. RELATED WORK

Interest in cross spectrum face recognition based on thermal-to-visible face synthesis has grown over the last several years. It enables a straightforward integration in the existing face recognition systems trained on data acquired in visible spectrum. Moreover, it offers an interpretable representation allowing human examiners to verify face matches across different spectra more efficiently.

Li et al. [4] is considered one of the first to investigate thermal-to-visible face synthesis. The proposed approach consists in a learning framework that benefits from the local linearity in the spatial domain of the image as well as in the image manifolds. Organizing the image patches and improving the generated visible face were performed using Markov random fields. More recently, the rise of deep learning have significantly boosted the state-of-the-art in numerous domains of computer vision. It offers an opportunity to approach the problems which were hardly solvable with conventional machine learning. Therefore, several works were based on Generative Adversarial Neural network (GAN) to synthesize visible faces from thermal images. Wang et al. [2] used a conditional GAN to synthesize faces from thermal images. The authors proposed to integrate facial landmark loss in the CycleGAN model [5] that portrays face identity preserving features. Another work, inspired from pix2pix model [6], is presented by Zhang et al. [3]. This work uses the same generator as pix2pix network coupled with multitask discriminator that performs closed-set face recognition in order to preserve the face identity information from the thermal inputs.

Our own recent work [7] presented a solution using Cascaded Refinement Networks (CRN) [12] to generate high-quality color visible image. The proposed network is trained using contextual loss function [13], enabling it to be inherently scale and rotation invariant. This work has reported higher recognition performance [7] and better visual quality [8] compared to models proposed in [3], [6]. Underlining the motivation of face synthesis from thermal-to-visible spectrum, it has been proved that face recognition performance reported on the synthesized images is significantly higher than the performance reported on visible spectrum when operated in poorly lit environments, as it was improved by 37.5%. However, under standard illumination conditions, face recognition based solely on visible spectrum outperforms by far systems based on generated visible-like face images, even under other challenging operating scenarios such as occlusion or head pose variation.

Several fusion and modality selection solutions were proposed, in setting multimodal biometric systems, based on quality assessment of the biometric sample. Good quality image usually yields to a robust matching performance. Fierrez-Aguilar et al. [9] introduced one of the earliest works of biometric quality fusion at score level, integrating quality information into a Bayesian statistical model for multimodal biometric classification. Using a unimodal biometric system, Vatsa et al. [10] proposed fusing RGB channel based on quality scores to improve the performance of iris recognition. Zhou et al. [11] presented quality based eye recognition by segmenting the eye into iris ans scelra and performing classification on the selected region as reported by its quality.

III. QUALITY-WEIGHTED SCORE FUSION

In this section, we describe in details the proposed fusion solution. First, we present thermal-to-visible face synthesis model used in this paper. This model will provide an estimation of the visible information based on the thermal input when it is initially missing in the visible spectrum. Then, we define the two face recognition systems used to compare face samples and obtain their matching scores. Subsequently, we list the quality assessment metrics considered in this paper. Finally, we describe the proposed quality-weighted fusion scheme.

A. Thermal-to-visible face synthesis

Thermal-to-visible face synthesis model, presented in [7], is based on Cascaded Refinement Networks (CRN) [12]. Choosing CRN as the basic block for our image synthesis model was motivated by the fact that it considers multi-scale information and requires training a limited number of parameters resulting in high resolution image generation. CRN is a convolutional neural network that consists of inter-connected refinement modules. Each module consists of only three layers, input, intermediate, and output layer. The first module considers the lowest resolution space (4x4 in our case). This resolution is duplicated in the successor modules until the last module (128x128 in our case), matching the target image resolution.

Fig. 1: Samples of visible images generated from thermal captures. Each group of three images have the source thermal image (TH) on the left, the generated visible image (G$_{VIS}$) in the middle and the ground-truth visible image (VIS) on the right for comparison.

To control the training of our CRN network, we used the contextual loss function (CL) [13]. This loss function fulfills our needs, as it is robust to not well aligned data and neglects outliers at the pixel level (in comparison to pixel level loss [6], [14]). Gramm loss [15] can satisfy the two aforementioned conditions, however, unlike CL, it does not constrain the content of the generated image since it describes the image globally.

The CL function can be computed between the source (thermal) and the generated images, and between the target (ground-truth) and the generated images. The source-generated loss aims at saving the details of the source image such as detailed boundaries. The target-generated loss maintains the properties of the target image in the generated image, e.g. target image style. In our case, as will be presented in the section IV, the source and target training image pairs are of identical faces. Therefore, the target-generated loss also maintains the detailed properties of the face in the source image.

Training of thermal-to-visible face synthesis model was carried out using numerous facial variations taken in controlled illumination conditions. For more details, one can refer to [7].

The images in Fig. 1 are samples of the generated visible images from thermal faces acquired in poorly lit environment. For each sample, we show the source image *(TH)*, the generated visible image *(G$_{VIS}$)*, and the ground-truth visible image *(VIS)*, in this order from left to right. We can observe that the thermal-to-visible face synthesis model has succeeded to generate a faithful estimation of the visible face. In other words, this step is essential to provide the missing visible information due to lack of illumination.

B. Face feature extraction and matching

We present here the face comparison systems used to obtain the matching scores on which the fusion will be applied. We selected the state-of-the-art system based on deep learning embeddings and a second system based on handcrafted features.

LightCNN [22] is a new implementation of CNN for face recognition designed to have fewer trainable parameters and to handle noisy labels. This network introduces a new concept of maxout activation in each convolutional layer, called Max-Feature-Map, for feature filter selection. This network has achieved better performance than CNNs while reducing computational costs and storage space. When evaluated on the LFW database, LightCNN achieved face recognition accuracy of 99.33%. We used the learned network with 29-layers to obtain embeddings of 256-dimension from face images. Embeddings extracted from gallery and probe templates are compared using cosine similarity.

Local Binary Pattern (LBP) was originally introduced by Ojala et al. [23] for texture analysis, but later on it was thouroughly explored in numerous applications. Particularly, it has shown its efficiency for face analysis. LBP represent a binary pattern that describes the local neighborhood of each pixel of the face image. The obtained LBP features are then concatenated to create a single histogram feature vector of 256-dimensions. Histograms extracted from gallery and probe image samples are compared using Chi-square dissimilarity measure.

C. Quality assessment metrics

Most often, quality of face samples reflects its convenience in providing a correct and accurate identification with a high

matching score. High quality samples often deliver highly informative features, yet low quality samples suffer heavily from noisy data and missing information. Therefore, selecting quality assessment metrics is very critical in boosting or lowering recognition performance.

We present, in this paper, a number of quality metrics in order to study the impact of each on face recognition performance.

- **Lightening symmetry [16]:** it quantifies the symmetry between sub-regions of an image and can be measured as the difference between the histogram of intensity in each half sub-region.
- **Brightness [17]:** is given by the average value of the image intensity histogram.
- **Contrast [17], [18]:** can be defined as the scale difference between maximum and minimum intensity values in an image.
- **Exposure [19]:** indicates the amount of light in the image and can be measured using image statistical measures.
- **Global Contrast Factor (GCF) [20]:** is the weighted sum of local contrast for various resolutions of the image.
- **Blur [21]:** is based on the fact that sharp images have thin edges and blurry images have wider edges, blur is expressed as the edge width.
- **Sharpness [18]:** is defined as the sum of gradients at every pixel intensity.

D. Proposed fusion scheme

We illustrate, in Fig. 2, the proposed asymmetric approach of quality weighted fusion at score level. Let Q_{VIS}, $Q_{G_{VIS}}$ and $Q_{Gallery}$ denote the quality measures of the visible image probe, the quality of the thermal-to-visible generated image probe, and the quality of visible gallery image, respectively. During authentication, we calculate the quality similarity scores of the original visible image and the thermal-to-visible generated image by determining their similarity to $Q_{Gallery}$, as follow:

$$QS_i = e^{\frac{Q_{Gallery}-Q_i}{Q_{Gallery}}}, where\, i \in \{VIS, G_{VIS}\} \quad (1)$$

Once the quality scores are obtained, they are normalized using min-max normalization. Then, we compute the weight to be assigned to each entity, as $w_i = \frac{QS_i}{QS_{VIS}+QS_{G_{VIS}}}$, $i \in \{VIS, G_{VIS}\}$. The closer Q_i is to $Q_{Gallery}$, the higher the weight will be assigned to i. Next, the face matching scores, denoted by S_i, are computed. S_{VIS} are obtained by comparing the visible image probe to the visible gallery set. $S_{G_{VIS}}$ are calculated by performing face comparison between the generated visible-like image and the visible gallery set. The obtained matching scores are then normalized. The overall fused score is computed using the weighted exponential sum rule, as follow:

$$S_{fused} = \sum_i w_i e^{S_i}, where\, i \in \{VIS, G_{VIS}\} \quad (2)$$

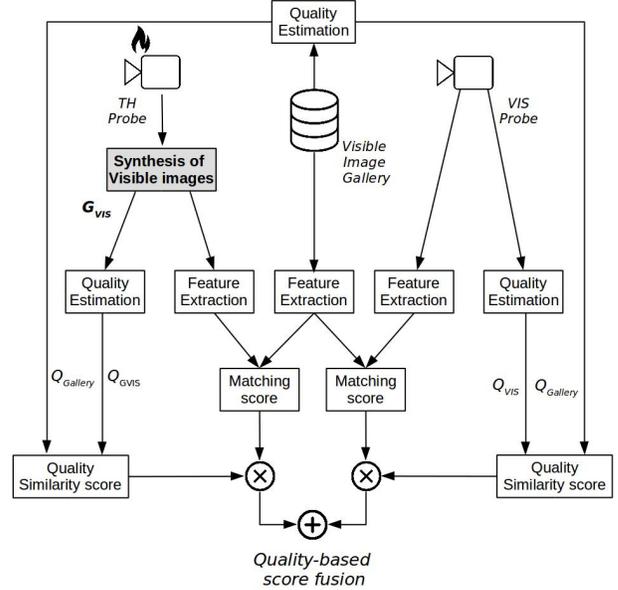$$w_i = \frac{QS_i}{QS_{VIS} + QS_{G_{VIS}}}, where\, i \in \{VIS, G_{VIS}\} \quad (3)$$

Fig. 2: Framework of the proposed quality-based score fusion scheme, where *VIS*, *TH* and *G_VIS* denote the visible image, the thermal image and the generated visible like image from the thermal capture, respectively.

Simply put, the quality weight will play a role in determining whether the visible sample is reliable enough to provide an accurate authentication. The quality of visible samples deteriorates mostly due to lack to illumination. Thereupon, the proposed fusion scheme will favor the generated visible-like sample as it is estimated from thermal inputs that are immune to illumination variations. The proposed method is summarized in the algorithm 1.

IV. EXPERIMENTS AND RESULTS

In this section, we present the database used for training the model of thermal-to-visible face synthesis and ultimately for performing face authentication based on quality weighted fusion. Then, we detail the evaluation protocol used to assess the proposed fusion approach. Finally, we present the obtained results followed by an analysis of the impact of different quality assessment metrics on face authentication performance.

A. Database

We used the VIS-TH face database [24] for the development and the evaluation of our solution. The database is publicly available[1] and contains face images in both visible spectrum with pixel resolution of 1920×1080 and thermal spectrum of pixel resolution 160×120 with a spectral response range of 7.5 - 13.5μm. Unlike the few existing databases of visible and thermal face, this database is acquired simultaneously using the dual sensor camera Flir DUO R [25] considering

---

[1]Visible And Thermal Paired Face Database: http://vis-th.eurecom.fr/

Algorithm 1: Quality-weighted score fusion

Input *Probe Samples:* set of samples acquired simultaneously in visible and thermal spectrum under various facial variations.
    Gallery Samples: set of neutral face samples acquired solely in visible spectrum.

for $p \in$ *Probe Samples* **do**
  $VIS \leftarrow$ Read Visible Image (p)
  $TH \leftarrow$ Read Thermal Image (p)
  $G_{VIS} \leftarrow$ Thermal-to-Visible face synthesis (TH) as per Sec.III-A
  $Q_{VIS} \leftarrow$ Quality Estimation(VIS)
  $Q_{G_{VIS}} \leftarrow$ Quality Estimation (G_{VIS})
  for $g \in$ *Gallery Samples* **do**
    $Gallery \leftarrow$ Read Visible Image (g)
    $Q_{Gallery} \leftarrow$ Quality Estimation$(Gallery)$
    $QS_{VIS}(p,g) \leftarrow$ Quality Similarity Score $(Q_{VIS}, Q_{Gallery})$ as per Eq.1
    $QS_{G_{VIS}}(p,g) \leftarrow$ Quality Similarity Score $(Q_{G_{VIS}}, Q_{Gallery})$ as per Eq.1
    $S_{VIS}(p,g) \leftarrow$ Matching Score $(VIS, Gallery)$ as per Sec.III-B
    $S_{G_{VIS}}(p,g) \leftarrow$ Matching Score$(G_{VIS}, Gallery)$ as per Sec.III-B
  end
end
Min-Max normalization of QS_{VIS}, $QS_{G_{VIS}}$, S_{VIS} and $S_{G_{VIS}}$
Compute weights w_{VIS} and $w_{G_{VIS}}$ as per Eq.3
$S_{fused} \leftarrow$ Quality-weighted score fusion $(w_{VIS}, S_{VIS}, w_{G_{VIS}} S_{G_{VIS}})$ as per Eq.2
return *the overall fused score* S_{fused}

---

a wide range of facial variations. The database contains in total 2100 images collected from 50 subjects of different ages, gender, and ethnicities. For the evaluation, we have considered 5 subsets of the database split per facial variation, as illustrated in Fig. 3.
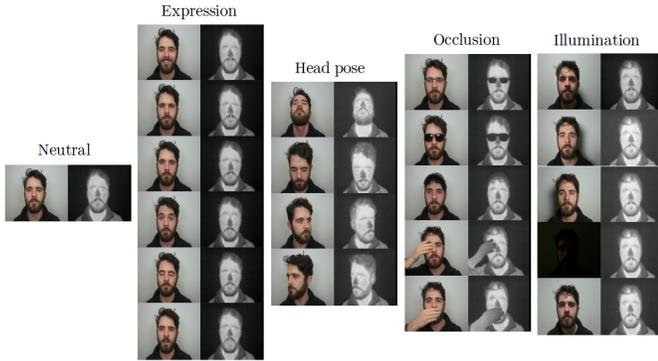

Fig. 3: Illustration of visible and thermal images for various facial variations.

B. Experimental protocol and results

A detailed protocol of training and evaluating thermal-to-visible face synthesis is described in [7].

For face authentication experiment, we consider as gallery set the neutral face image acquired in visible spectrum and as probe set all the remaining face variations from visible and thermal-to-visible generated face images. We use LBP and LightCNN for face features extraction. Feature vectors from gallery set and probe sets are compared to obtain the matching scores of the two entities. In parallel, quality measures are computed using 7 different quality assessment metrics and quality similarity scores are then deduced. We perform quality weighted fusion at score level as described

above. The performance of our proposed fusion approach is compared to the performance of fusing scores obtained from matching visible probes and thermal probes against a common visible gallery set.

We present, in table I, rank-1 identification of LightCNN and LBP systems under expression, head pose, occlusion and illumination variations. In this table, we report firstly the identification performance of each of the following setups separately: matching visible probe, original thermal probe and thermal-to-visible synthesized faces against visible gallery. We observe that face identification using the generated visible-like images yields to better performance than when using thermal images, which proves the efficiency of thermal-to-visible face synthesis in reducing the gap between visible and thermal spectra. Although, the generated visible-like images are still far from optimal and that is marked for all the facial variations. Concretely, this discrepancy is mostly significant when the facial variations are more prominent, as it is the case for different head poses and occlusions, and to a less degree facial expression and illumination variations.

To highlight the main motivation of thermal spectrum usage in face authentication, we display, in Fig. 4, the receiver operating characteristic (ROC) curve of the three setups aforementioned for face images that were acquired in total darkness. We can clearly observe that the setup based on thermal-to-visible synthesized images provides significantly higher performance compared to the setup based on visible images. This affirms the efficacy of thermal imagery in most of the challenging scenarios such as poorly lit environments. Also, we note that the setup based on thermal-to-visible synthesized images outperforms the thermal based setup, which proves the efficiency of thermal-to-visible face synthesis in reducing spectral gap between visible and thermal data.

Furthermore, we can evidently perceive that face authentication using deep learning embeddings (LightCNN) outperforms
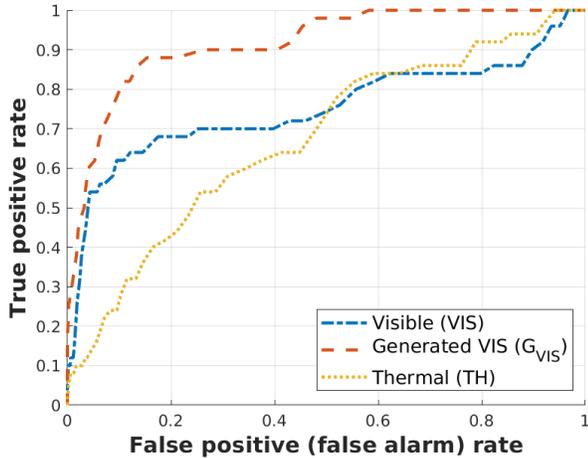
Fig. 4: ROC curves in dark environment using LightCNN system.

hand-crafted features (LBP) which confirms the assertions presented in [26].

To assess the impact of each quality metric used in this paper, we report rank-1 identification of quality weighted fusion of visible images and generated visible-like images (denoted as *(VIS, G_{VIS})* in table I) for each quality metric, where Q^1, Q^2, Q^3, Q^4, Q^5, Q^6 and Q^7 denote lightning symmetry, brightness, contrast, exposure, GCF, blur and sharpness, respectively. Q^{avg} refers to using the average quality score of the 7 quality metrics. Furthermore, we consider quality weighted score fusion of visible face images and original thermal images (denoted as *(VIS,TH)* in table I) as a baseline. We note that the described fusion scheme using the thermal-to-visible face synthesis outperforms considerably the plain fusion of visible and thermal images. This divergence in performance certifies the proficiency of thermal-to-visible face synthesis in bringing the two spectra closer together.

Regarding the impact of each quality metric, we can determine that the proposed quality weighted score fusion shows nearly similar performances for all the quality assessment metrics. Nevertheless, we note that weighing the matching score to be fused with the average quality score results in a better performance in most cases.

To get a deeper understanding of the performance of our proposed fusion scheme, we plot the ROC curves, in Fig. 5. We compute the ROC curve over all the facial variations contained in the database, so as to demonstrate the efficacy of our proposed approach in a wide range of operative scenarios. The plot confirms our previous observation, as we can see that all the considered quality assessment metrics impact the performance of the fused system similarly. Conclusively, we observe that the proposed fusion based approach in this paper outperforms face authentication operating solely on visible data. It is fair to admit that the difference of performance is not considerably large, that is due to the distribution of the variations within the database, as it contains only one sample

that highlights the thermal imagery usage.

V. CONCLUSION

Integrating thermal imagery in face recognition systems does not only confront spoofing attacks, but it also tackles the poor illumination challenge for visible spectrum. Therefore, we have proposed, in this paper, a new scheme of score level fusion for robust face authentication from visible and thermal face data that enables straightforward integration in the existing face recognition systems. The proposed system operates according to the following protocol in face recognition: individuals had been enrolled solely in visible spectrum (i.e. gallery) but can be afterwards controlled by dual visible and thermal acquisition (i.e. probe). Considering that the gap between the visible and thermal spectra is important, it was necessary to include a step where we generated visible-like images from thermal inputs. This solution benefits from the quality measures of the visible gallery and probe faces to assign weights for visible and thermal samples. The results report an interesting improvement in face recognition performance compared to when using solely visible samples. In addition, results have proved the efficiency of thermal-to-visible face synthesis in providing more accurate face authentication system.

REFERENCES

[1] R. Ramachandra, B. Busch, "Presentation Attack Detection Methods for Face Recognition Systems: A Comprehensive Survey", ACM Comput. Surv. New York, NY, USA, vol. 50, pages 8:1–8:37, April 2017.

[2] Z. Wang, Z. Chen, and F. Wu, "Thermal to visible facial image translation using generative adversarial networks", IEEE Signal Processing Letters, 25(8):1161–1165, Aug 2018

[3] T. Zhang, A. Wiliem, S. Yang, and B. Lovell. TV-GAN: Generative adversarial network based thermal to visible face recognition. In 2018 International Conference on Biometrics (ICB), pages 174–181, Feb 2018.

[4]] J. Li, P. Hao, C. Zhang, and M. Dou. Hallucinating face from thermal infrared images. In Proceedings of the International Conference on Image Processing, ICIP 2008, San Diego, California, USA, pages 465–468, October 12-15, 2008.

[5] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. "Unpaired imageto-image translation using cycle-consistent adversarial networks". 2017 IEEE International Conference on Computer Vision (ICCV) 2017.

[6] P. Isola, J. 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 2017, pages 5967–5976. IEEE Computer Society, Honolulu, HI, USA, July 21-26, 2017.

[7] K. Mallat, N. Damer, F. Boutros, A. Kuijper, J. Dugelay, "Cross-spectrum thermal to visible face recognition based on cascaded image synthesis", 2019 International Conference on Biometrics (ICB), June 4-7, 2019.

[8] N. Damer, F. Boutros, K. Mallat, F. Kirchbuchner, J. Dugelay, A.Kuijper, "Cascaded Generation of High-quality Color Visible Face Images from Thermal Captures", submitted to 27th European Signal Processing Conference, EUSIPCO 2019, Coruna, Spain, Sept 2019.

[9] J. Fierrez-Aguilar, J. Ortega-Garcia, J. Gonzalez-Rodriguez, J. Bigun, Discriminative multimodal biometric authentication based on quality measures, Pattern Recognition, 2005.

[10] M. Vatsa, R. Singh, A. Noore, Integrating image quality in 2-SVM biometric match score fusion, International Journal of Neural Systems 17 (05), pages 343–351, 2007.

[11] Z. Zhou, E. Y. Du, C. Belcher, N. L. Thomas, E. J. Delp, Quality fusion based multimodal eye recognition, in: IEEE International Conference on Systems, Man, and Cybernetics, pages 1297–1302, 2012.
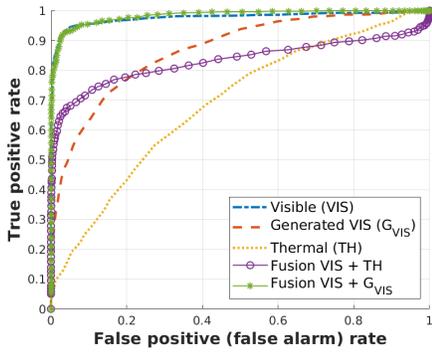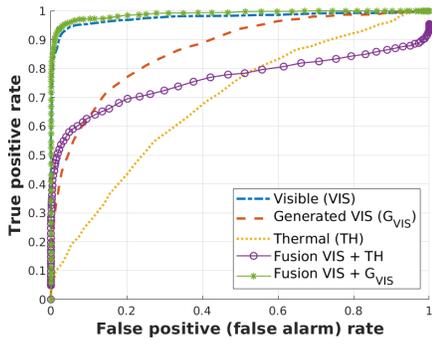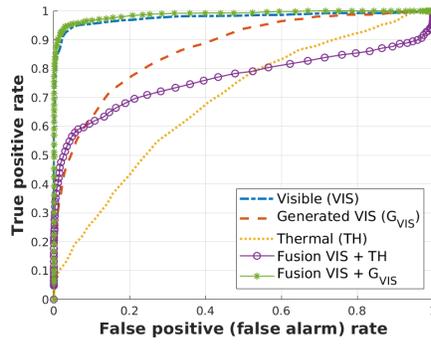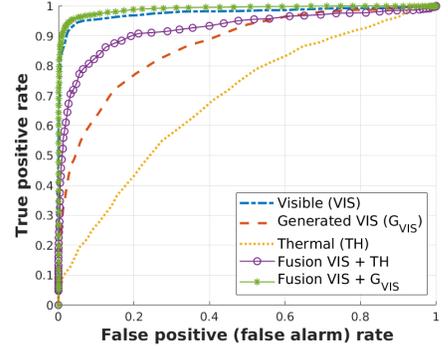
(a) Lightning symmetry
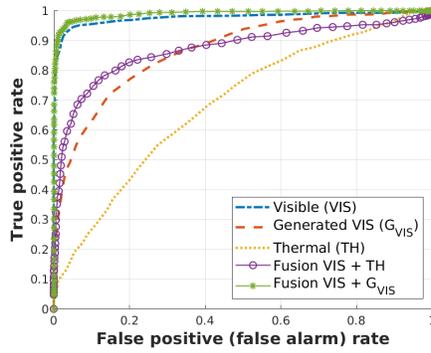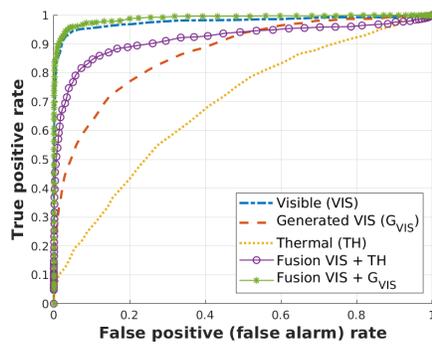
(b) Brightness
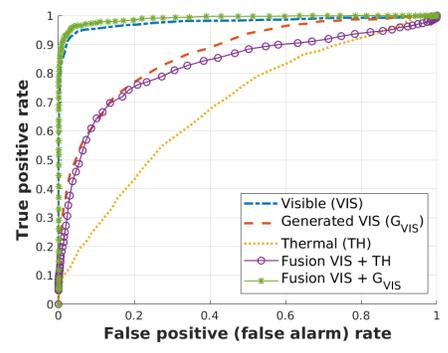
(c) Contrast
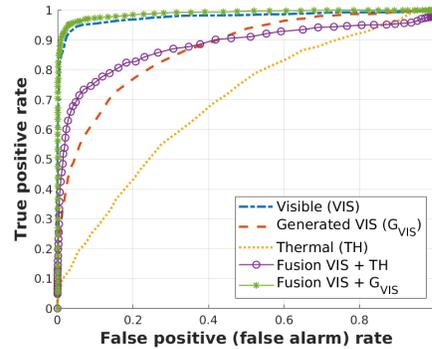
(d) Global Contrast Factor (GCF)

(e) Exposure

(f) Blur

(g) Sharpness

(h) Average

Fig. 5: ROC curve deduced over all the facial variations in VIS-TH database [24] using LightCNN

TABLE I: Rank-1 identification across multiple facial variations.

		LightCNN					LBP				
		VIS	TH	G_{VIS}	(VIS, TH)	(VIS, G_{VIS})	VIS	TH	G_{VIS}	(VIS, TH)	(VIS, G_{VIS})
EXPRESSION	Q^1	0.996	0.240	0.700	0.580	0.986	0.973	0.033	0.593	0.370	0.883
	Q^2				0.933	0.993				0.486	0.873
	Q^3				0.883	0.993				0.626	0.890
	Q^4				0.490	0.99				0.583	0.906
	Q^5				0.513	0.993				0.520	0.930
	Q^6				0.917	0.987				0.383	0.916
	Q^7				0.797	0.990				0.500	0.880
	Q^{avg}				0.797	0.990				0.740	0.936
HEAD POSE	Q^1	0.81	0.16	0.353	0.587	0.813	0.833	0.027	0.347	0.167	0.607
	Q^2				0.647	0.840				0.273	0.633
	Q^3				0.613	0.827				0.38	0.653
	Q^4				0.393	0.853				0.286	0.686
	Q^5				0.446	0.846				0.293	0.686
	Q^6				0.653	0.820				0.160	0.660
	Q^7				0.586	0.820				0.240	0.626
	Q^{avg}				0.626	0.826				0.333	0.660
OCCLUSION	Q^1	0.988	0.144	0.468	0.600	0.932	0.880	0.060	0.440	0.228	0.632
	Q^2				0.808	0.972				0.244	0.684
	Q^3				0.756	0.976				0.404	0.708
	Q^4				0.508	0.964				0.316	0.748
	Q^5				0.556	0.940				0.300	0.736
	Q^6				0.856	0.972				0.164	0.716
	Q^7				0.792	0.928				0.188	0.680
	Q^{avg}				0.748	0.980				0.372	0.740
ILLUMINATION	Q^1	0.872	0.176	0.648	0.772	0.792	0.600	0.048	0.448	0.080	0.432
	Q^2				0.832	0.880				0.136	0.604
	Q^3				0.848	0.896				0.352	0.668
	Q^4				0.644	0.896				0.276	0.444
	Q^5				0.656	0.896				0.312	0.520
	Q^6				0.796	0.904				0.236	0.516
	Q^7				0.768	0.896				0.192	0.536
	Q^{avg}				0.816	0.896				0.272	0.604

[12] Q. Chen and V. Koltun. "Photographic image synthesis with cascaded refinement networks" In IEEE International Conference on Computer Vision, ICCV 2017, pages 1520–1529. IEEE Computer Society, Venice, Italy, October 22-29, 2017.

[13] R. Mechrez, I. Talmi, and L. Zelnik-Manor. The contextual loss for image transformation with non-aligned data. 2018 - 15th European Conference In Computer Vision - ECCV, Proceedings, volume 11218 of Lecture Notes in Computer Science, pages 800–815. Springer,Munich, Germany, September 8-14, 2018.

[14] L. Xu, J. S. J. Ren, C. Liu, and J. Jia. Deep convolutional neural network for image deconvolution. In Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, pages 1790–1798, December 8-13 2014, Montreal, Quebec, Canada.

[15] L. A. Gatys, A. S. Ecker, and M. Bethge. Image style transfer using convolutional neural networks. In 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, pages 2414–2423. IEEE Computer Society, Las Vegas, NV, USA, June 27-30, 2016.

[16] W. Pankaj, R. Kiran, R. Raghavendra and B. Christoph, "Assessing face image quality for smartphone based face recognition system", 5th International Workshop on Biometrics and Forensics (IWBF), pages 1–6, Coventry, UK, April2017.

[17] F. T. Zohra, A. D. Gavrilov, O. Z. Duran and M. Gavrilova," A linear regression model for estimating facial image quality", 16th International Conference on Cognitive Informatics and Cognitive Computing (ICCI* CC), pages 130–138, 2017.

[18] X. Gao, S. Z. Li, R. Liu, and P. Zhang, Standardization of face image sample quality, in International Conference on Biometrics. Springer, pp. 242–251, 2007

[19] M. V. Shirvaikar, An optimal measure for camera focus and exposure, in System Theory, 2004. Proceedings of the Thirty-Sixth Southeastern Symposium on. IEEE, pages 472–475, 2004.

[20] K. Matkovic, L. Neumann, A. Neumann, T. Psik, and W. Purgathofer, Global contrast factor-a new approach to image contrast. Computational Aesthetics, vol. 2005, pages 159–168, 2005.

[21] N. D. Narvekar and L. J. Karam, A no-reference image blur metric based on the cumulative probability of blur detection (cpbd), IEEE Transactions on Image Processing, vol. 20, no. 9, pages 2678–2683, 2011.

[22] X. Wu, R. He, Z. Sun, and T. Tan. A light CNN for deep face representation with noisy labels. IEEE Transactions on Information Forensics and Security, 13(11) pages 2884–2896, 2018.

[23] Ojala, T., Pietikinen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. Pattern Recognit. 29(1), 5159, 1996.

[24] K. Mallat and J. Dugelay, "A benchmark database of visible and thermal paired face images across multiple variations", International Conference of the Biometrics Special Interest Group (BIOSIG), pages1–5, Darmastadt, Germany, 26-28 Sept 2018.

[25] Flir systems: http://www.flir.com/

[26] S. Setiowati, Zulfanahri, E. L. Franita, and I. Ardiyanto, A review of optimization method in face recognition: Comparison deep learning and non-deep learning methods, in 2017 9th International Conference on Information Technology and Electrical Engineering (ICITEE), pp. 1–6, Oct 2017.