Cross-spectral Face Recognition in Heterogeneous Environments: A Case Study on Matching Visible to Short-wave Infrared Imagery

Nathan D. Kalka   Thirimachos Bourlai   Bojan Cukic   Lawrence Hornak

{Nathan.Kalka,Thirimachos.Bourlai,Bojan.Cukic,Lawrence.Hornak}@mail.wvu.edu

West Virginia University, LDCSEE, 6201, Morgantown, WV, 26505, U.S.A

Abstract

In this paper we study the problem of cross spectral face recognition in heterogeneous environments. Specifically we investigate the advantages and limitations of matching short wave infrared (SWIR) face images to visible images under controlled or uncontrolled conditions. The contributions of this work are three-fold. First, three different databases are considered, which represent three different data collection conditions, i.e., images acquired in fully controlled (indoors), semi-controlled (indoors at standoff distances $\geq$ 50m), and uncontrolled (outdoor operational conditions) environments. Second, we demonstrate the possibility of SWIR cross-spectral matching under controlled and challenging scenarios. Third, we illustrate how photometric normalization and our proposed cross-photometric score level fusion rule can be utilized to improve cross-spectral matching performance across all scenarios. We utilized both commercial and academic (texture-based) face matchers and performed a set of experiments indicating that SWIR images can be matched to visible images with encouraging results. Our experiments also indicate that the level of improvement in recognition performance is scenario dependent.

1. Introduction

Most face recognition (FR) systems are based on images captured in the visible range of the electromagnetic spectrum (380-750 *nm*). However, in harsh environmental conditions characterized by unfavorable lighting and pronounced shadows (such as a nighttime environment [2]), human recognition based only on visible spectral images may not be feasible [14, 8]. Thus, recognition of faces in the infrared spectrum has become an area of growing interest [16, 10, 15].

The infrared (IR) spectrum can be further divided into multiple spectral bands. The boundaries between these bands can vary depending on the scientific field involved
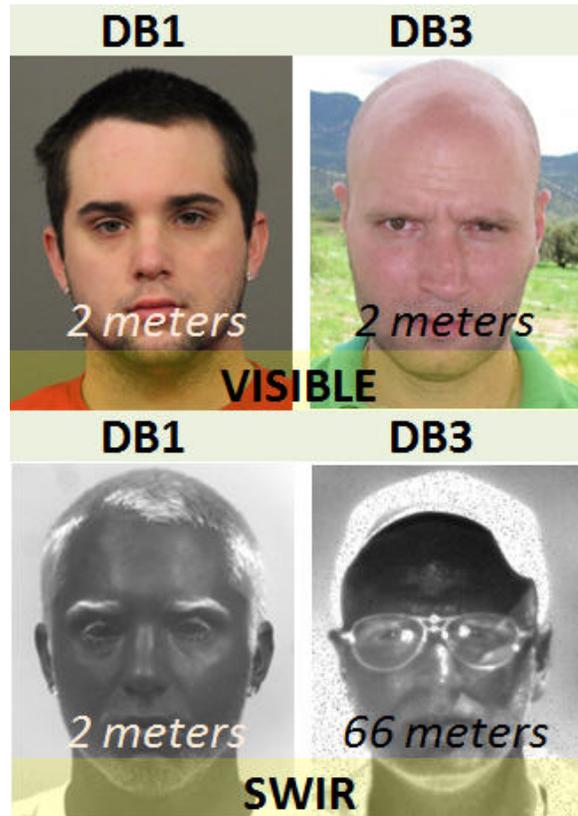

Figure 1. Illustration of visible and SWIR imagery present in databases DB1 and DB3 respectively.

(e.g., optical radiation, astrophysics, or sensor technology [5]). The spectral bands used in this work are the visible (used for baseline and cross-spectral experiments) and the Short-Wave Infrared (used for cross-spectral experiments only); the SWIR band is a part of the reflected IR (active) band that ranges, in our experiments, from 0.9-1.9μm. SWIR has a longer wavelength range than NIR and is more tolerant to low levels of obscurants like fog and smoke. Differences in appearance between images sensed in the visible and the active IR band are due to the properties of the object

1

being imaged. Regions in the SWIR band require an external light source. However, the advantage is that a SWIR imaging system can take advantage of sunlight, moonlight, or starlight, and can remain unobtrusive and covert since the reflected IR light is invisible to the human eye.

In the visible spectrum, human faces from different ethnic groups can exhibit different reflectance characteristics. The problem of finding invariants related to skin color in the visible spectrum could be addressed in the IR spectrum by using a calibrated IR sensor. In addition, since IR and visible imagery capture intrinsically different characteristics of the observed faces, intuitively, a better face description could be found by utilizing the complimentary information present across the two spectra. The main benefits in using SWIR spectrum for face recognition are the following: (a) the external source of illumination is invisible to the human eye making it suitable for covert applications; (b) it can be useful in a nighttime environment; (c) SWIR imagery can be combined with visible-light imagery to generate a more complete image of the human face; and (d) facial features that are not observed in the visible spectrum may be observable in the SWIR spectrum. In addition, its proximity to the visible spectrum makes it particularly relevant for use in biometric-related face applications.

While previous FR studies have mainly concentrated in visible and NIR imagery, FR in SWIR spectrum, specifically at 1550 *nm*, has received limited attention in the literature. The main problem is that prior work focused in SWIR FR, researchers concentrated on applying FR algorithms on face images acquired under fully controlled conditions [1]. However, in uncontrolled scenarios (long range recognition, operational conditions), there is a need for efficient Intelligence and Surveillance Reconnaissance (ISR) interoperability, i.e., operational teams (e.g., armed forces) are required to effectively manage, access and use ISR to improve command and control, and enhance information sharing and situational understanding to improve the effectiveness of operations while minimizing collateral damage in a complex environment.

The aforementioned challenging scenarios motivated us, and therefore in this paper we study the problem of cross spectral face recognition in heterogeneous environments. Specifically, we investigate the advantages and limitations of matching short wave infrared (at 1550 *nm*) probe face images to visible (gallery) images acquired under variable scenarios: visible images were collected under controlled and semi-controlled conditions (full frontal faces, facial expressions, indoors and outdoors, short range, fixed standoff distance to 7 feet or 2 meters), while SWIR images were captured under (i) fully controlled indoor conditions; (ii) semi-controlled conditions (full frontal faces, indoors, long ranges, i.e., up to 348 feet or 106 meters); and (iii) uncontrolled conditions (variable poses, face expressions, occlu-

sion, outdoors, night and day, variable range, i.e., up to 1312 feet or 400 meters). Three different matching/encoding algorithms were utilized, namely, Local Binary Patterns (LBP) [11], Local Ternary Patterns (LTP) [11], and a commercial face matcher (Identity Tools G8) provided by L1 Systems[1].

Our experimental results indicate that our proposed methodology (i.e., using a cross-photometric score level fusion scheme) performs better than baseline (single matchers before photometric normalization) cross-spectral FR performance, in the most challenging (uncontrolled) scenario described above, by 25% using G8 in combination with the proposed fusion rule. Similar performance improvement was noted in the dataset acquired under semi-controlled conditions, i.e., our fusion rule performed better than the baseline by 10% at 50 *m*, and by 15% at 106 *m*.

1.1. Contributions

The contributions of this work are three-fold. First, three different databases are considered, which encompass three different data collection scenarios (controlled, semi-controlled and uncontrolled). Second, a set of experiments is performed in order to demonstrate the feasibility of cross-spectral matching under controlled and challenging scenarios. Third, we illustrate how (a) the usage of independent or combined photometric normalization techniques, and (b) cross-photometric score level fusion can be utilized to improve cross-spectral matching performance across all scenarios. In this method we first applied independent or combined photometric normalization techniques to both the gallery and probe images. Then, the best scores, in terms of recognition performance, were fused at the score level. Cross-spectral matching is useful in practical scenarios: in law enforcement, for example, the mug shots are mainly acquired in the visible spectrum (as described in the ANSI/NIST-ITL 1-2000 standard) while a probe image may be acquired in the greater IR spectrum. The purpose of using cross-photometric score level fusion was because we are dealing with a heterogeneous problem (where gallery and probe sets have face images acquired in different spectral bands) while at the same time the probe images were acquired under variable scenarios. Thus, we argue that cross-photometric score level fusion can bridge the spatial representation of independent gallery (visible) and probe (SWIR) images so that we can achieve higher matching performance.

1.2. Paper Organization

The rest of this paper is organized as follows. Section 2 describes the visible and SWIR imagery used in this work. Section 3 provides a summary of the pre-processing techniques used, face recognition algorithms employed, the pro-

---

[1]www.l1id.com

| Database | Scenario | Standoff Distance | | Cameras | | Sessions | Subjects | Imgs/Subject | |
		Variable/Fixed	Meters	Visible	SWIR			Visible	SWIR
DB1	CI	F	2	5D Mark II	XenIC	2	50	2-14	27
DB2	SCI	F	50 & 106	Powershot	Goodrich	1-2	50	6	4-6
DB3	UO	V	60 - 400	Powershot[2]	Goodrich	2-5	16	1 - 6	1-8

Table 1. Database and scenario statistics - DB1 was collected in a controlled indoor (CI) environment at a fixed stand-off distance, with minor variability in regards to the acquired imagery. DB2 was collected in a semi-controlled indoor (SCI) environment with moderate variability in terms of distance. DB3 was collected in an uncontrolled outdoor (UO) environment during the day and night (SWIR only). SWIR image acquisition was opportunistic; no restrictions were placed on the subjects, resulting in significant variability with respect to pose, expression, illumination, stand-off distance, and occlusion (sunglasses and headwear).

posed cross-photometric score level fusion rule, and experiments conducted. Section 4 presents the results while conclusions are drawn in Section 5.

2. Experimental Setup

Three unique facial image databases (DB1, DB2, and DB3) were considered to facilitate the proposed study in heterogeneous environments. Each database consists of a set of visible gallery images and SWIR probe images, which were acquired in different environmental conditions and under different scenarios. The following subsections describe each database in more detail as well as the equipment used for acquiring imagery in the different spectra.

2.1. Equipment

1. **Canon EOS 5D Mark II**: This digital SLR camera has a 21.1-megapixel full-frame CMOS sensor with DIGIC 4 Image Processor and a vast ISO range of 100-6400. It also has Auto Lighting Optimizer and Peripheral Illumination Correction that enhances its capability. In this work, the Mark II is used to obtain standard RGB, ultra-high resolution frontal pose face images in the visible spectrum.

2. **Canon PowerShot SX110**: The SX110 digital SLR has a 9-megapixel CCD sensor with an ISO range from 80-1600.

3. **XenICS Xeva-818**: This camera has an Indium Gallium Arsenide (InGaAs) 320×256 Focal Plane Array (FPA) with 30m pixel pitch, 98% pixel operability and three stage thermoelectric cooling. The XEVA-818 has a relatively uniform spectral response from 950 - 1700 *nm* wavelength (SWIR band) across which the InGaAs FPA has largely uniform quantum efficiency. Response falls rapidly at wavelengths lower than 950 *nm* and greater than 1700 *nm*.

4. **Goodrich SU640**: The SU640 is an Indium Gallium Arsenide (InGaAs) video camera featuring high-sensitivity and wide dynamic range. This model has a 640×512 FPA with 25m pixel pitch, and $> 99\%$

pixel operability. The spectral sensitivity of the SU640 ranges uniformly from 700 - 1700 *nm* wavelength. Similar to the XenICS, response falls rapidly at wavelengths lower than 700 *nm* (as opposed to 950 *nm* for XenICS) and greater than 1700 *nm*.

2.2. Databases

The following is a short description of each database utilized in our experiments. Additional information may be found in table 1.

1. **DB1** - Collected in a controlled indoor (CI) environment, comprised of 50 subjects over two sessions. High quality visible imagery was captured with a Canon 5D Mark II with an average inter-ocular distance of 563.08 pixels (min=347, max=947) and standard deviation of 144.36 pixels. SWIR images were collected using a Xeva-818 using broadband tungsten illumination with an average inter-ocular distance of 63.16 pixels (min=47, max=93) and a standard deviation of 10.12 pixels. Images were collected at three different poses, i.e., full frontal and left/right at +/-67.5 degrees. For each pose, images were obtained with and without employing a band pass filter. The wavelength of the bandpass filters starts from 950 *nm* and goes up to 1650 *nm* in steps of 100 *nm*. In this paper we only utilized the SWIR imagery captured at 1550 *nm*.

2. **DB2** - Collected in a semi-controlled indoor (SCI) environment, composed of 50 subjects over two sessions. Visible imagery was captured utilizing a Canon PowerShot SX110 with an average inter-ocular distance of 119.59 pixels (min=52, max=139) and a standard deviation of 7.97 pixels. SWIR imagery was collected at 1550 *nm* with a Goodrich SU640 at a stand-off distance of 50m and 106m, respectively, utilizing proprietary optics and laser illumination. Inter-ocular distance, averaged across both distances, was 61.79 pixels

---

[2]DB3 visible imagery was acquired with multiple cameras including an Canon PowerShot.

(min=53, max=75) with a standard deviation of 3.78 pixels.

3. **DB3** - Collected in an uncontrolled outdoor (UO) environment during day and night (SWIR only). This dataset is composed of 16 subjects over multiple sessions. Visible images were acquired with multiple cameras including a Canon PowerShot SX110. The average inter-ocular distance for the visible imagery was 113.07 pixels (min=26, max=317) with a standard deviation of 59.90 pixels. SWIR imagery was collected at 1550 *nm* with a Goodrich SU640 at variable stand-off distances, ranging from 60 to 400 meters. Inter-ocular distance averaged around 110.22 pixels (min=47, max=146) with a standard deviation of 17.22 pixels. It is also important to note that the SWIR imagery was collected opportunistically. That is, subjects were uncooperative, and no constraints were in place to minimize expression, pose, stand-off distance, and occlusion (sunglasses and headgear).

3. Methodological Approach

In this paper we experiment with both commercial and academic FR algorithms. While pre-processing methods utilized by the commercial software are not known, the proposed research software [11] employed the following pre-processing routines.

3.1. Photometric Normalization

Cross spectral face recognition is a challenging problem because the physical interaction of different electromagnetic waves (i.e., visible vs. SWIR) with materials (in our case facial skin) will be different, resulting in different reflection, transmission and scattering properties. As such, texture, contrast, etc. is different when dealing with visible and SWIR face images. Photometric normalization algorithms have traditionally been employed to compensate for changes in illumination, such as ambient variation or strong shadows [9]. Similarly, in this work we employ photometric normalization for the purpose of facilitating cross spectral matching. More specifically, we employed contrast limited adaptive histogram equalization (CLAHE) [7], single scale retinex [3] (SSRlog) with logarithmic transformation, single scale retinex (SSRatan) with arc-tangent transformation, SSRlog followed by CLAHE, and SSRatan follow by CLAHE (see Fig. 2 for sample face images normalized by the aforementioned techniques). These algorithms are described as follows:

1. **CLAHE** - Operates on local regions (8x8 for our experiments) in the image and applies histogram equalization to each region. Mathematically, it is described

as:

$$f(n) = \frac{N-1}{M} \times \sum_{k=0}^{n} h(k), \qquad (1)$$

where M and N are the number of pixels and gray level bins in each sub-region, and h is the histogram of each sub-region. To increase contrast without amplifying noise, CLAHE redistributes each histogram such that the height falls below the clip limit threshold (.01 in our experiments). More specifically, gray level counts above the clip limit are uniformly redistributed among the gray levels below the clip threshold. Each sub-region is then subsequently combined using bilinear interpolation.

2. **SSR** - The image is decomposed into illumination $L(x, y)$ (the amount of light falling on the targeted object), and reflectance $R(x, y)$ (the amount of light reflected from the surface of the targeted object). Illumination is estimated as a low-pass version of the original image, while reflectance is found by dividing the original image by the estimated illumination followed by a non-linear transformation such as the logarithm or arc-tangent. Mathematically this is described as:

$$I(x, y) = L(x, y) \times R(x, y), \qquad (2)$$
$$L(x, y) = I(x, y) * G_\sigma(x, y), \qquad (3)$$

where G_σ is a Gaussian of scale σ, and $*$ denotes convolution. Finally, the reflectance is estimated as:

$$R(x, y) = \kappa \left(\frac{I(x, y)}{L(x, y)} \right), \qquad (4)$$

where κ is a non-linear transformation such as the logarithm or arc-tangent function.

3. **SSR Combinations** - A common problem with SSR is that certain regions of the image can become oversaturated or "washed out", which can have a negative impact on texture based approaches to face recognition. Furthermore, "halo" artifacts may be introduced depending on the scene and scale value chosen for the Gaussian smoothing function. Modifications to SSR have been introduced in the literature to help alleviate this problem, such as multi-scale retinex [6], but at the cost of increased processing time. In this work, we also experiment with the combination of SSR followed by CLAHE normalization to help compensate for the aforementioned degradations.

3.2. Face Detection and Geometric Normalization

First, the Viola & Jones *face detection* [12] algorithm was applied to the visible and SWIR images from databases DB1, DB2, and DB3 (visible only). It was used to localize
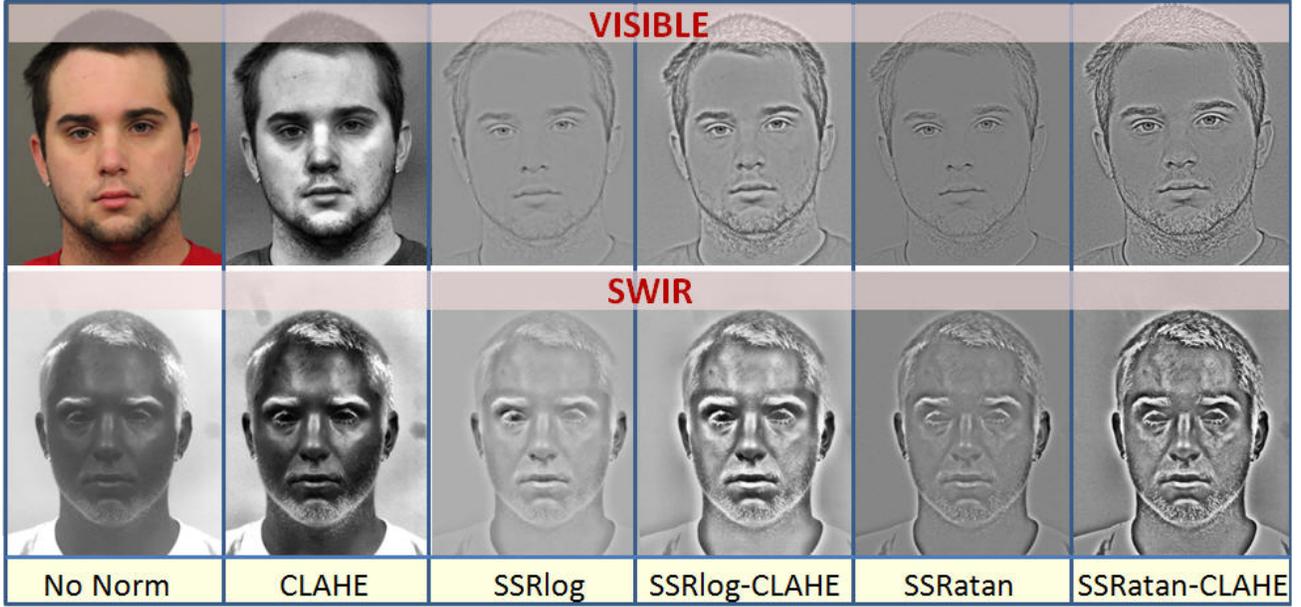
Figure 2. Illustration of photometric normalization for visible and SWIR (1550 *nm*) imagery. Note that for columns 3 and 5, regions of the face become over-saturated, obscuring local texture. The application of SSR followed by CLAHE, columns 4 and 6, reduces over-saturation, thereby increasing the contrast of local textures in the face region.

the spatial extent of the face and determine its boundary.

A *geometric normalization* scheme was applied to images acquired after face detection. The normalization scheme compensated for slight perturbations in the frontal pose, and consisted of eye detection and affine transformation. Automated eye detection was performed using a template matching algorithm where the coordinates of the eye were automatically obtained [13]. Traditional face and eye detection techniques did not work when evaluating images from DB3 acquired in the SWIR band. Thus, eyes centers were located by manual annotation. After the eye centers were found, the canonical faces were automatically constructed by applying an affine transformation. Finally, all faces were canonicalized to the same dimension of 150×130.

3.3. Face Recognition Methods

Both commercial and research software was employed to perform the face recognition experiments: (1) Commercial software such as Identity Tools G8 provided by L1 Systems; (2) standard texture based face recognition methods such as LBP and LTP [11]. The LBP operator was introduced as a texture descriptor. Patterns in an image are computed by thresholding 3×3 neighborhoods based on the value of the center pixel. Then, the resulting binary pattern is converted to a decimal value. The local neighborhood is defined as a set of sampling points evenly spaced on a circle. The LBP operator used in our experiments is described as $LBP_{P,R}^{u^2}$,

where P refers to the number of sampling points placed on a circle with radius R. The symbol u^2 represents the uniform pattern, which accounts for the most frequently occurring pattern in our experiments. The pattern is important because it is capable of characterizing local regions that contain edges and corners. The binary pattern for pixels, lying in a circle f_p, $p = 0, 1, \ldots, P - 1$ with the center pixel f_c, is mathematically computed as follows:

$$S(f_p - f_c) = \left\{ \begin{array}{ll} 1 & \text{if } f_p - f_c \geq 0; \\ 0 & \text{if } f_p - f_c \leq 0. \end{array} \right\} \quad (5)$$

Following this a binomial weight 2^P is assigned to each sign $S(f_p - f_c)$ to compute the LBP code,

$$LBP_{P,R} = \sum_{p=0}^{P-1} S(f_p - f_c) 2^P. \quad (6)$$

LBP is invariant to monotonic gray-level transformations. However, one disadvantage is that LBP tends to be sensitive to noise in homogeneous image regions since the binary code is computed by thresholding the center of the pixel region.

Consequently, LTP [11] has been introduced to overcome such a limitation, where the quantization is performed as follows:

$$S(f_p - f_c) = \left\{ \begin{array}{ll} 1 & \text{if } f_p - f_c \geq t; \\ 0 & \text{if } |f_p - f_c| \leq t. \\ -1 & \text{if } f_p - f_c \leq -t. \end{array} \right\} \quad (7)$$
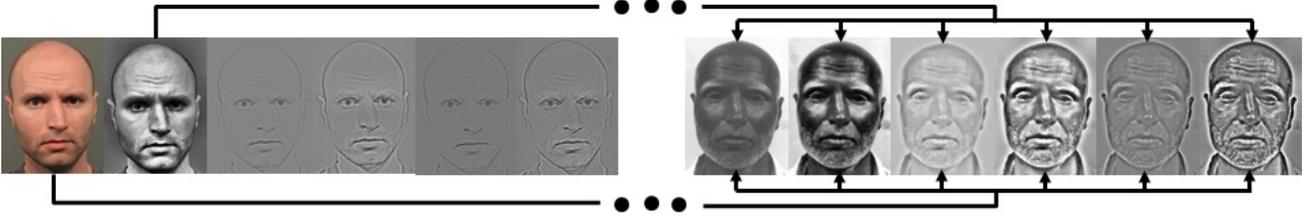
Figure 3. Cross-photometric score level fusion illustration. Gallery and probe images are heterogeneously matched across different photometrically normalized imagery, which can provide additional information for match score fusion as opposed to intra-photometric matching.

The output of this operator is a 3-valued pattern, as opposed to a binary pattern. Furthermore, the threshold t, can be adjusted to produce different patterns. The user-specific threshold also makes the LTP code more resistant to noise.

3.4. Cross-Photometric Score Level Fusion

In [4], Mendex-Vazquez et al. demonstrated that when operating in the visible spectrum, cross-photometric score level fusion can be utilized to improve matching performance under variable lightning conditions. In this work, we propose a cross-photometric match score fusion rule applied to images acquired in the visible and SWIR spectra, before and after photometric normalization. Our work is different not only in the application (i.e., facilitation of cross-spectral matching; challenging scenarios), but also in the selection and combination of photometric normalization techniques. Intuitively, intra-photometric matching is beneficial when variability is limited to a single degradation (i.e., illumination) or degradations which are consistent across the observed data.

In this paper we demonstrate that, when facial images are collected with multiple sources of variability present (i.e., illumination, blur, low contrast, etc.), cross-photometric match score fusion provides better performance by taking advantage of the information present across all photometrically normalized imagery. This is facilitated by photometrically normalizing both gallery and probe images by the aforementioned techniques described in subsection 3.1, resulting in six ($n = 6$) gallery and six probe images per comparison (this process is illustrated in Fig. 3). Matching is then performed cross-photometrically between the 36 combinations. This is mathematically characterized as:

$$\overline{S}_{ij} = \bigwedge_{i=1}^{n} \bigwedge_{j=1}^{n} m\left(Gt_i, Pt_j\right), \qquad (8)$$

where Gt and Pt are the gallery and probe templates respectively. Matching function $m()$ corresponds to the matching algorithms described in subsection 3.3, while $\bigwedge$ simply indicates iteration. Then, the max or min fusion rule is applied on the vector of resulting match scores, $\overline{S}_{ij}$, for similarity (G8) or distance (LBP/LTP) respectively.

3.5. Cross-Spectral Matching

Utilizing the datasets illustrated in Table 1, three different scenarios, described in subsection 2.1, were explored: controlled indoor (CI), semi-controlled indoor (SCI), and uncontrolled outdoor (UO). For each scenario the following set of experiments were conducted:

1. *Visible to visible (baseline)*
2. *Visible to SWIR (before photometric normalization)*
3. *Visible to SWIR (after photometric normalization)*
4. *Visible to SWIR (Proposed cross-photometric fusion)*

In the first experiment, matching is performed on images acquired in the visible spectrum for the purpose of establishing a baseline for comparison. The second experiment establishes a baseline for cross-spectral matching prior to photometric normalization. Specifically, gallery images acquired in the visible spectrum are matched against imagery captured in the SWIR spectrum. Similarly, the third experiment evaluates the performance of cross-spectral matching after the application of the photometric normalization algorithms described in subsection 3.1. Finally, the fourth experiment applies the proposed cross-photometric fusion rule.

The identification performance of the system is evaluated through the cumulative match characteristic (CMC) curve. The CMC curve measures the $1 : m$ identification system performance, and judges the ranking capability of the identification system.

4. Results and Discussion

For all matching experiments only one visible gallery image and one SWIR image (at 1550 *nm*) is utilized per subject. CMC Performance for experiments 1-4 across all databases is illustrated in Fig. 4. The proposed fusion rule is presented only for the best performing matcher (performance for all matchers is listed in Table 2).

The experimental results indicate that, although cross-spectral matching is a very challenging problem, when gallery (visible) face images are compared against SWIR face images, and all images were acquired under fully con-
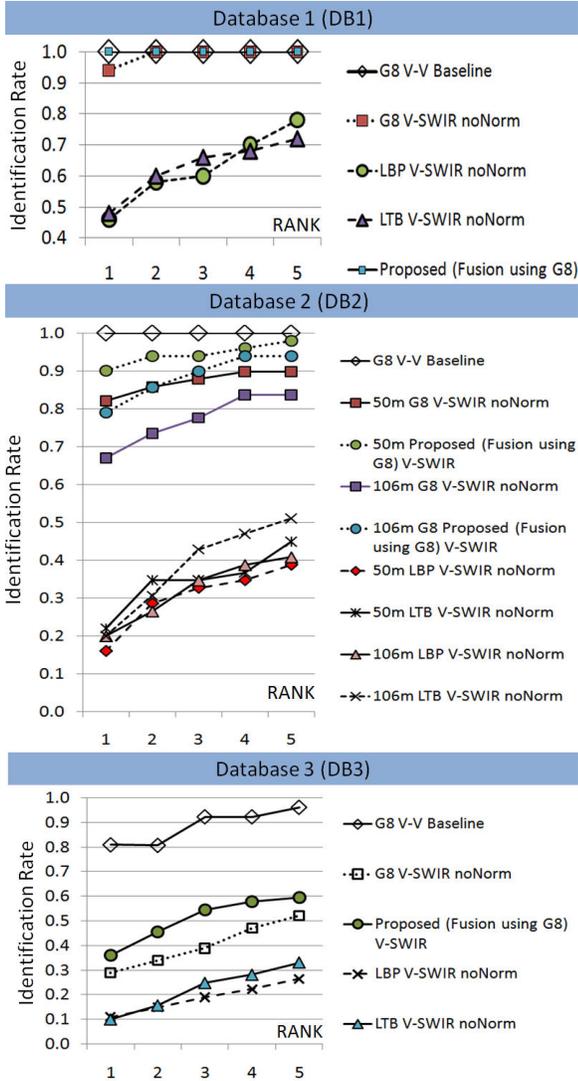
Figure 4. CMC curves comparing the performance of baseline (visible to visible) and baseline (visible to SWIR) to the proposed fusion rule with G8.

	DB1	
	v-v	v-s (w/o Norm, w Norm, Proposed)
G8	1	(0.94, 0.96-CLAHE, **1**)
LBP	0.96	(0.46, **0.56**-CLAHE, **0.56**)
LTP	0.96	(0.48, 0.62-CLAHE, **0.64**)
	DB2-50m	
	v-v	v-s (w/o Norm, w Norm, Proposed)
G8	1	(0.82, 0.88-SSRlogCLAHE, **0.90**)
LBP	1	(0.16, **0.41**-SSRatanCLAHE, 0.29)
LTP	1	(0.22, **0.39**-SSRatanCLAHE, 0.24)
	DB2-106m	
	v-v	v-s (w/o Norm, w Norm, Proposed)
G8	1	(0.67, 0.76-SSRlogCLAHE, **0.80**)
LBP	1	(0.20, **0.22**-SSRatanCLAHE, **0.22**)
LTP	1	(0.20, **0.29**-SSRlog, 0.18)
	DB3	
	v-v	v-s (w/o Norm, w Norm, Proposed)
G8	0.81	(0.29, **0.37**-SSRlog, **0.37**)
LBP	0.69	(0.11, 0.14-SSRatan, **0.15**)
LTP	0.69	(0.10, **0.13**-SSRlog, 0.11)

Table 2. Rank 1 identification rates when utilizing G8, LBP, and LTP for cross-spectral matching. Experiment 1 is indicated by v-v while experiments 3-4 are indicated as v-s (w/o Norm, w Norm, proposed) respectively. Note that for experiment 3, v-s (w Norm), only performance for the single best photometric normalization technique is provided which is indicated within the table for each matcher.

In the most challenging (DB3 - uncontrolled) scenario, all FR matchers do not perform well. One of the problems is that in the baseline experiment of the uncontrolled scenario, the number of subjects used is quite modest (16). Another problem is that different subjects were acquired under different conditions (the gallery images were acquired under variable distances and illumination conditions; also subjects were not cooperative when acquiring the probe imagery). The advantage of using G8, in combination with the proposed fusion rule when experimenting with database, DB3, is that it performs better than baseline (single matchers before photometric normalization) cross-spectral system performance by approximately 25%.

5. Conclusions

In this paper our focus was on investigating the problem of cross spectral face recognition in heterogeneous environments. Specifically we investigate the advantages and limitations of matching SWIR face images against visible images under variable conditions. In terms of pre-processing, different photometric normalization techniques were used.

trolled conditions (DB1 database), the identification rate can be very high (100% at rank-1). Interestingly, the identification rates at rank-1 in cross-spectral matching experiments is comparable to the baseline identification rates, i.e., when performing intra-spectral (baseline scenario - when using both gallery and probe images that were all collected in the visible spectrum) matching experiments.

When using our proposed fusion-based approach, in semi-controlled conditions (DB2 database) at a stand-off distance of 50 m, the identification rate (90%) is reasonably comparable to the baseline rate. However, when the stand-off distance is more than doubled (106 m), the identification rate at rank-1 drops another 11%, resulting in 79% accuracy.

In addition, two research and one commercial matcher were employed to perform a set of baseline experiments. For the purpose of this work, three different databases are considered representing three different data collection conditions, i.e., images acquired in fully controlled indoors, semi-controlled indoors, and uncontrolled outdoor environments. Our results indicate that, across all datasets used, the application of photometric normalization improves recognition performance. While different face matchers were applied in all three datasets, the best performance results (when compared to the cross spectral baseline scenario, i.e., visible to SWIR matching before normalization) were acquired when using our proposed cross-photometric fusion rule in conjunction with G8 across all experimental scenarios (based on the aforementioned databases).

Our future plans are to develop an improved normalization scheme that will bridge the spatial representation of face images acquired under visible and SWIR bands. This is expected to result in improved matching performance.

Acknowledgments

References

[1] M. A. Akhloufi, A. Bendada, and J. C. Batsale. Multispectral face recognition using non linear dimensionality reduction. In *Proceedings of SPIE, Visual Information Processing XVIII*, volume 7341, 2009. 2

[2] T. Bourlai, N. Kalka, D. Cao, B. Decann, Z. Jafri, F. Nicolo, C. Whitelam, J. Zuo, D. Adjeroh, B. Cukic, J. Dawson, L. Hornak, A. Ross, and N. A. Schmid. Ascertaining human identity in night environments. In *Distributed Video Sensor Networks*, pages 451–467. Springer, 2011. 1

[3] D. Jobson, Z. Rahman, and G. Woodell. Properties and performance of a center/surround retinex. *Image Processing, IEEE Transactions on*, 6(3):451 –462, mar 1997. 4

[4] H. Méndez-Vázquez, J. Kittler, C.-H. Chan, and E. García-Reyes. On combining local dct with preprocessing sequence for face recognition under varying lighting conditions. In *Proceedings of the 15th Iberoamerican congress conference on Progress in pattern recognition, image analysis, computer vision, and applications*, CIARP'10, pages 410–417, Berlin, Heidelberg, 2010. Springer-Verlag. 6

[5] J. L. Miller. *Principles Of Infrared Technology: A Practical Guide to the State of the Art*. Springer, 1994. 1

[6] Z. Rahman, D. Jobson, and G. Woodell. Multi-scale retinex for color image enhancement. In *Image Processing, 1996. Proceedings., International Conference on*, volume 3, pages 1003–1006, September 1996. 4

[7] A. Reza. Realization of Contrast Limited Adaptive Histogram Equalization (CLAHE) for real-time image enhancement. *VLSI Signal Processing*, 38:35–44, 2004. 4

[8] A. Selinger and D. A. Socolinksy. Face recognition in the dark. In *CPRW*, pages 129–134, June 2004. 1

[9] J. Short, J. Kittler, and K. Messer. A comparison of photometric normalisation algorithms for face verification. In *Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on*, pages 254 – 259, may 2004. 4

[10] D. Socolinsky, A. Selinger, and J.Neuheisel. Face recognition with visible and thermal imagery. *CVIU*, 91:72–114, 2003. 1

[11] X. Tan and B. Triggs. Enhanced local texture feature sets for face recognition under difficult lighting conditions. *Trans. Img. Proc.*, 19:1635–1650, June 2010. 2, 4, 5

[12] P. Viola and M. Jones. Robust real-time face detection. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 2, page 747, 2001. 4

[13] C. Whitelam, Z. Jafri, and T. Bourlai. Multispectral eye detection: A preliminary study. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 209 –212, August 2010. 5

[14] J. Wilder, P. Phillips, C. Jiang, and S. Wiener. Comparison of visible and infra-red imagery for face recognition. In *Automatic Face and Gesture Recognition, 1996., Proceedings of the Second International Conference on*, pages 182 –187, October 1996. 1

[15] D. Yi, S. Liao, Z. Lei, J. Sang, and S. Z. Li. Partial face matching between near infrared and visual images in mbgc portal challenge. In *Proceedings of the Third International Conference on Advances in Biometrics*, ICB '09, pages 733–742, Berlin, Heidelberg, 2009. Springer-Verlag. 1

[16] Y. Yoshitomi, T. Miyaura, S. Tomita, and S. Kimura. Face identification using thermal image processing. *WRHC*, pages 374–379, 1997. 1