Cross-spectral Face Verification in the Short Wave Infrared (SWIR) Band

Thirimachos Bourlai   Nathan Kalka   Arun Ross   Bojan Cukic   Lawrence Hornak

West Virginia University

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

Abstract—The problem of face verification *across* the short wave infrared spectrum (SWIR) is studied in order to illustrate the advantages and limitations of SWIR face verification. The contributions of this work are two-fold. First, a database of 50 subjects is assembled and used to illustrate the challenges associated with the problem. Second, a set of experiments is performed in order to demonstrate the possibility of SWIR cross-spectral matching. Experiments also show that images captured under different SWIR wavelengths can be matched to visible images with promising results. The role of multispectral fusion in improving recognition performance in SWIR images is finally illustrated. To the best of our knowledge, this is the first time *cross-spectral SWIR face recognition* is being investigated in the open literature.

Keywords-Short Wave Infrared; Face Verification; Cross-spectral Matching; Fusion.

I. INTRODUCTION

Most face recognition (FR) systems are based on images captured in the visible range of the electromagnetic spectrum (380-750nm). However, in harsh environmental conditions characterized by unfavorable lighting and pronounced shadows (such as a nighttime environment [1]), human recognition based only on visible spectral images may not be feasible [2]. Thus, recognition of faces in the infrared spectrum has become an area of growing interest [3], [4].

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 (e.g., optical radiation, astrophysics, or sensor technology [5]). The IR bands discussed in this work are based on the response of various detectors (see Figure 1).

The IR spectrum comprises of the reflected IR (active) and the thermal IR (passive) bands. The active band (0.7-$2.5\mu m$) is divided into the NIR (near infrared) and the SWIR (shortwave infrared) spectrum. The 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 being imaged. Unlike thermal IR imagery that can be acquired without any external illumination in day or night environments, regions in SWIR band require an external light source. For example, in order to acquire reflective imagery near the visible spectrum with low light levels, image intensification in the NIR is needed through the use of a micro-channel plate (MCP) [6]. At the same time, an *active* IR system can take advantage of sunlight, moonlight, or starlight, and can remain unobtrusive and covert since the reflected IR light is for the most part 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. While previous studies have shown that the NIR spectrum yielded promising results [7]–[9] and that fusion of thermal and visible face images improved face recognition [10], the SWIR spectrum has received limited attention in the literature (see research gap in Figure 1).

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 spectra may be observable in SWIR images. In addition, its proximity to the visible spectra makes it particularly relevant for use in biometric face applications.

A. Goals and Contributions

Most FR systems acquire images of the face in the visible and NIR range of the electromagnetic spectrum. To the best of our knowledge, this paper represents the first attempt in the literature to investigate the problem of face verification in the SWIR spectrum. Thus, multiple spectral bands within the SWIR range are considered. Further, the possibility of performing cross-spectral facial matching is studied. Cross-spectral matching is also useful in practical scenarios: in law enforcement, for example, the mugshots are acquired only 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.

In this work we investigate face recognition on the West Virginia University Multi-spectral (WVUM) database consisting of visible and SWIR face images of 50 subjects. Four different experiments have been performed. The first
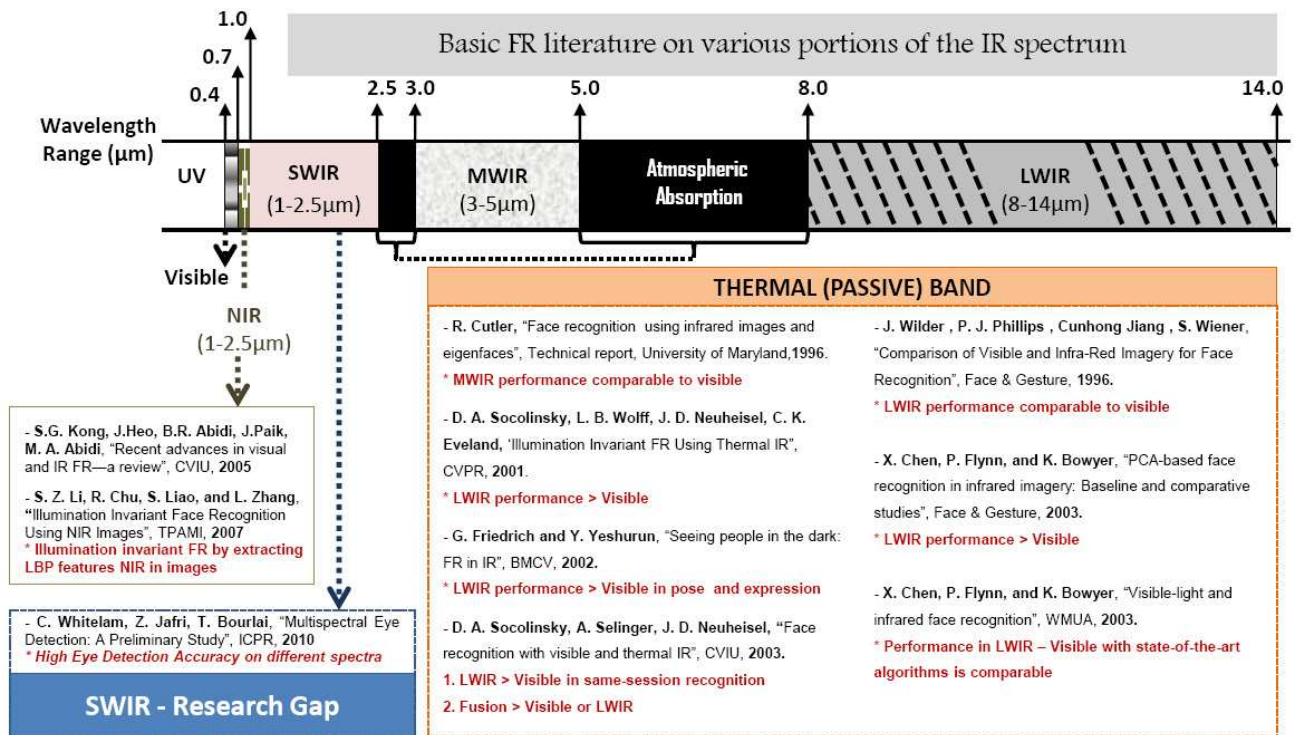
Figure 1: Literature map and SWIR research gap.

is the matching of high quality face images obtained in the visible spectrum for the purpose of establishing a baseline. The second experiment compares facial images captured in the SWIR spectrum with those in the visible spectrum, and the third compares SWIR images obtained at different wavelengths. The fourth experiment involves fusion of multiple algorithms for improving the performance of SWIR face verification.

B. Paper Organization

The rest of this paper is organized as follows. Section II describes the multi-spectral imagery used in this work. Section III provides a summary of the face recognition algorithms employed and the experiments conducted. Section IV presents the results while conclusions are drawn in section V.

II. MULTISPECTRAL IMAGERY

Two different types of cameras have been used in this work, viz., a Canon EOS 5D Mark II, and a XenICS Xeva InGaAs. These cameras are described below and they cover the visible spectrum and a significant portion of the SWIR spectrum.

- **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, Canon (Fig. 2(a)) is used to obtain standard RGB, ultra-high resolution frontal pose face images in the **visible** spectrum.

- **XenICS**: A XenICS XEVA-818 camera served as the imager for obtaining **SWIR** images (see Fig. 2(b)). The camera has an Indium Gallium Arsenide (InGaAs) 320×256 Focal Plane Array (FPA) with $30\mu m$ pixel pitch, 98% pixel operability and three stage thermoelectric cooling. The XEVA-818 has a relatively uniform spectral response from 950 - 1700 nm wavelength (the Short Wavelength Infra-Red (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.

The aforementioned cameras were employed for data collection in order to assemble the WVU Multispectral Face database.

The live face capture configuration we used is illustrated in Fig. 3 (a) and (b). The database was assembled across two sessions spanning over a time period of 1 month. In the beginning of the first session, the subjects were briefed about the data collection process after which they signed a consent document. In total, 50 subjects (30 male + 20 female) participated in this experiment. Each session of the visible spectrum dataset has 25 frontal face images of each subject, resulting in a total of 1,250 images.
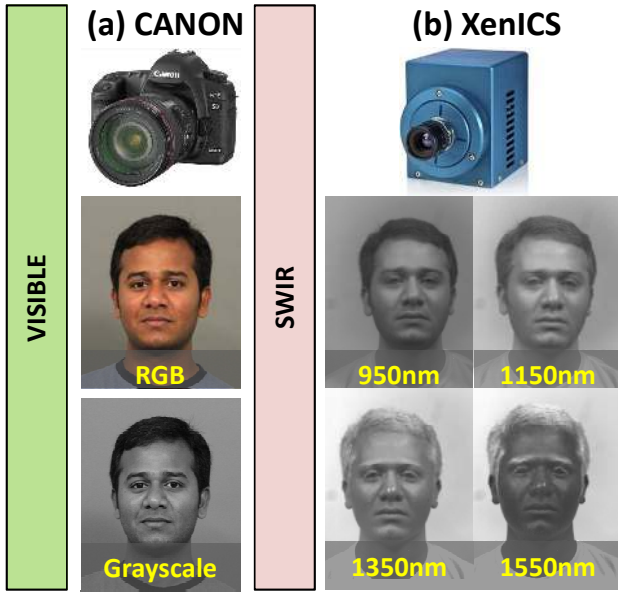
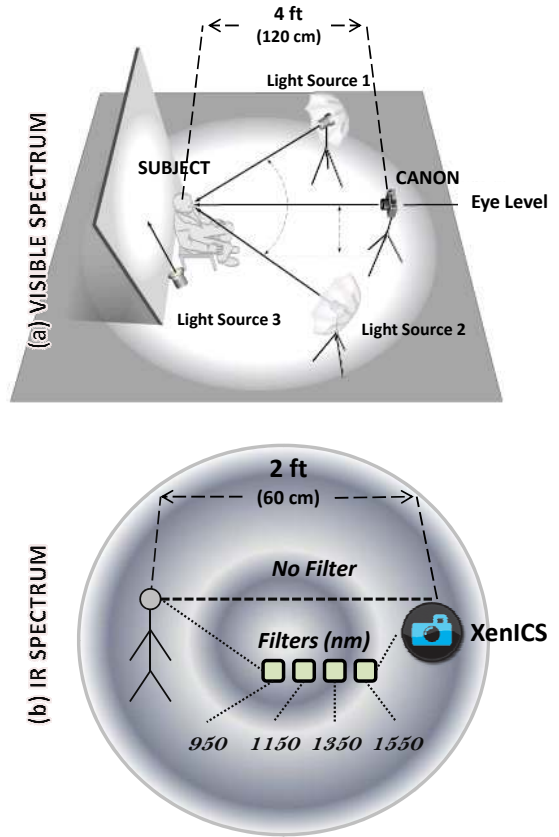Figure 2: (a) *Canon EOS 5D Mark II*. (b) XenICS SWIR.

Figure 3: Image capture setup: (a) Visible Spectrum: The live subject-capture setup using the Canon digital camera (taken from US Department of US State Department, Bureau of Consular Affairs). More details have been added by the authors. (b) SWIR Spectrum: subject-capture using the XenICS camera.

For the SWIR dataset, we acquired face images at three different poses, i.e., full frontal and left/right at +/-67.5 degrees. For each pose, we obtained images with and without employing a band pass filter. The wavelength of the band-pass filters starts from 950nm and goes up to 1650nm in steps of 100nm. Thus, nine samples per pose (no filter and 950nm-1650nm in steps of 100nm) were acquired resulting in a total of 1,350 images (27 samples/session×50 subjects).

III. METHODOLOGY AND EVALUATION

In our experiments we used both commercial and academic software. While the pre-processing routines employed by the commercial software are not known, the academic software employed the following pre-processing method.

A. Pre-Processing Method

First, the Viola & Jones *face detection* algorithm was applied to the original data. It was used to localize the spatial extent of the face and determine its boundary.

In the next step, a *geometric normalization* scheme was applied to images acquired after detection. The normalization scheme compensated for slight perturbations in the frontal pose, and consisted of eye detection and affine transformation [11]. Eye detection was based on a template matching algorithm where the eye coordinates were obtained. Then the canonical faces were constructed by applying an affine transformation. Eventually, all faces were warped to the same dimension of 300×300.

For *photometric normalization* and in order to facilitate cross-spectral SWIR matching, we employed the contrast limited adaptive histogram equalization (CLAHE) technique

[12]. CLAHE operates on local regions (8x8 for our experiments) in the image and applies histogram equalization, illustrated in equation (1), to each sub-region.

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

Here, M and N are the number of pixels and grey 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, grey level counts beyond the clip limit are uniformly redistributed among the grey levels below the clip threshold. Finally each sub-region is combined using bilinear interpolation. CLAHE normalized imagery is illustrated in Fig 4.

B. Face Recognition Methods & Fusion

In the next step, three face recognition methods were used: Principal Components Analysis (PCA) with k-nearest

		Session Two					
		Visible	Nofilter	950nm	1150nm	1350nm	1550nm
Session One	Visible	(0.00 n/a)	(0.99 2.10)	(1.00 0.44)	(0.59 1.19)	(1.47 2.37)	(5.60 5.00)
	Nofilter		(2.02 0.00)	(0.00 0.00)	(0.00 0.00)	(2.00 2.00)	(6.02 4.00)
	950nm			(0.00 0.00)	(0.00 0.00)	(2.00 2.00)	(8.00 4.00)
	1150nm				(0.00 0.00)	(2.00 2.00)	(7.97 3.97)
	1350nm					(0.14 0.00)	(2.00 2.00)
	1550nm						(8.02 3.97)

Table I: EER for all verification experiments utilizing G8 before and after CLAHE normalization.
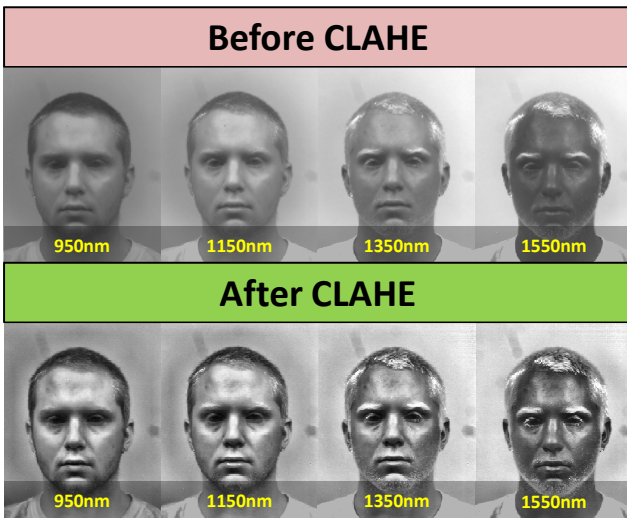

Figure 4: CLAHE normalization.

neighbor (k-NN), Verilook[1], and Identity Tools G8® (G8). PCA and Verilook were used only in the SWIR versus SWIR fusion study where we have only one sample per subject per pose per spectral band in each session. In addition, two fusion algorithms were evaluated: the likelihood ratio based fusion rule and the sum of scores fusion method with min-max normalization.

C. Experiments

By using the visible and SWIR images in the assembled database, four different types of face verification experiments were performed:

1) Visible vs Visible
2) Visible vs SWIR
3) SWIR vs SWIR
4) Fusion

The *first* experimental scenario matches high quality images in the visible spectrum for the purpose of establishing a baseline for comparison. In the *second* experiment we compare the SWIR image against the Visible spectra images. In the *third* experiment we analyze cross-spectral matching

[1]http://www.neurotechnology.com

within the SWIR band. Our *fourth* experiment utilizes fusion algorithms on images corresponding to a specific SWIR spectral band whose performance was experimentally observed to be inferior.

D. Performance Evaluation

To evaluate the performance of the system, the Equal Error Rate (EER-%) criterion is employed (other measures such as d-prime or the FRR at a pre-defined FAR may be used). This is determined from the Receiver Operating Characteristic (ROC) Curve computed on the second session. The lower the EER, the better the system performance.

IV. EXPERIMENTAL RESULTS

Table I presents the results for the first three experiments in terms of EER. Each entry in the table has two values representing the EER for G8 before and after CLAHE normalization, respectively. Experiments 1 and 2 utilized ten images for training while only one image was available for training in experiment 3. With regards to testing, experiment 1 utilized three images while only one image was available for testing in experiments 2 and 3.

For experiment 1, G8 performed exceptionally well, achieving perfect separation between genuine and imposter distributions. In experiment 2, we compare visible and SWIR images ranging from $950nm$ to $1550nm$ in $200nm$ intervals. G8 consistently provides good performance across all wavelengths. The overall trend for experiment two illustrates that as wavelength increases, performance degrades. However, even when testing images at $1550nm$, G8 achieved an EER of *5.6%* which is a very encouraging result. Experiment 3 investigated cross-spectral matching across SWIR. Performance is consistently good across all wavelengths with the exception of $1550nm$, which proved to be the most challenging case. Table I also shows that CLAHE had a positive impact on matching cross-spectral images. Improvement in performance is specifically observed at $1550nm$.

The final experiment involved fusion at $1550nm$, since it proved to be the most difficult wavelength for matching. Here we fuse the match scores for PCA, Verilook, and G8 using the simple sum (SS) and likelihood ratio (LR) fusion rules and the results are presented in Fig. 5. Both fusion algorithms perform better than G8 at $1550nm$ resulting
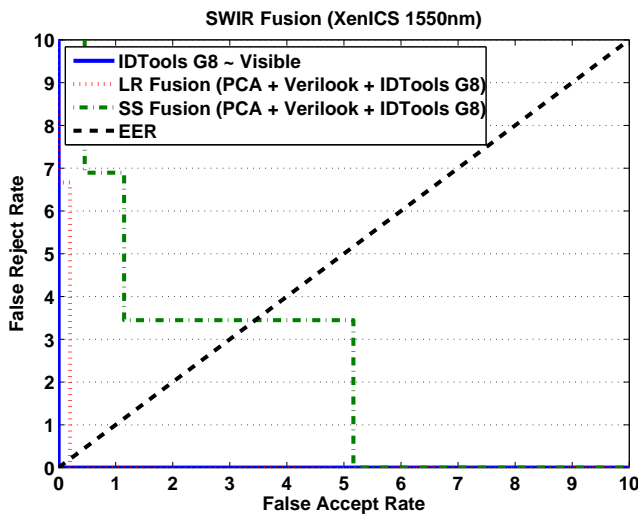
Figure 5: Fusion across recognition algorithm for SWIR 1550nm. Also shown is the ROC for Identity Tools G8 on visible spectrum data for comparison.

in an EER of 0.10% and 3.45% for LR and SS fusion, respectively.

V. DISCUSSION AND CONCLUSION

We have presented a study on the problem of face verification *across* the short wave infrared spectrum (SWIR). The experiments were performed using the WVU Multispectral database. Experimental results show that, unlike long wave infrared or thermal, images captured under different SWIR wavelengths can be matched to visible images with promising results (see table I). This is especially important when intra-wavelength matching is extremely challenging (i.e., nighttime vs. day time environments).

Related to the photometric normalization employed, CLAHE improved the performance of both commercial algorithms for SWIR cross-spectral matching. This improvement was primarily seen when comparing lower to higher wavelengths. The drawback of CLAHE was that it did not always improve performance when matching visible to SWIR images. Finally, we illustrated that multispectral fusion can significantly improve recognition performance in SWIR matching scenarios.

ACKNOWLEDGMENT

REFERENCES

[1] 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*. Princeton Univeristy Press, 2010.

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

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

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

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

[6] S. Kong, J. Heo, B. Abidi, J. Paik, and M. Abidi. Recent advances in visual and infrared face recognition: A review. *CVIU*, 97(1):103–135, 2005.

[7] H. Chang, A. Koschan, B. Abidi, and M. Abidi. Physics-based fusion of multispectral data for improved face recognition. *ICPR*, 3:1083–1086, 2006.

[8] H. Chang, Y. Yao, A. Koschan, B. Abidi, and M. Abidi. Spectral range selection for face recognition under various illuminations. In *ICIP*, pages 2756–2759, October 2008.

[9] Z. Pan, G. Healey, M. Prasad, and B. Tromberg. Face recognition in hyperspectral images. *TPAMI*, 25:1552–1560, 2003.

[10] S. Singh, A. Gyaourova, G. Bebis, and I. Pavlidis. Fusion of infrared and visible images for face recognition. *ECCV*, 2004.

[11] C. Whitelam, Z. Jafri, and T. Bourlai. Multispectral eye detection : A preliminary study. In *ICPR*, 2010.

[12] A. Reza. Realization of contrast limited adaptive histogram equalization (clahe) for real-time image enhancement. *VLSI Signal Processing*, 38:35–44, 2004.