Chapter 30
Ascertaining Human Identity in Night Environments

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

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 · N.A. Schmid
Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV 26506, USA
e-mail: Thirimachos.Bourlai@mail.wvu.edu

N. Kalka
e-mail: Nathan.Kalka@mail.wvu.edu

D. Cao
e-mail: dcao@mix.wvu.edu

B. Decann
e-mail: bdecann@mix.wvu.edu

Z. Jafri
e-mail: zjafri@mix.wvu.edu

F. Nicolo
e-mail: fnicolo@mix.wvu.edu

C. Whitelam
e-mail: cwhitela@mix.wvu.edu

J. Zuo
e-mail: jzuo@mix.wvu.edu

D. Adjeroh
e-mail: Donald.Adjeroh@mail.wvu.edu

B. Cukic
e-mail: Bojan.Cukic@mail.wvu.edu

J. Dawson
e-mail: Jeremy.Dawson@mail.wvu.edu

L. Hornak
e-mail: Lawrence.Hornak@mail.wvu.edu

A. Ross
e-mail: Arun.Ross@mail.wvu.edu

N.A. Schmid
e-mail: Natalia.Schmid@mail.wvu.edu

B. Bhanu et al. (eds.), *Distributed Video Sensor Networks*,
DOI 10.1007/978-0-85729-127-1_30,

Abstract Understanding patterns of human activity from the fusion of multimodal sensor surveillance sources is an important capability. Most related research emphasizes improvement in the performance of biometric systems in controlled conditions characterized by suitable lighting and favorable acquisition distances. However, the need for monitoring humans in night environments is of equal if not greater importance. This chapter will present techniques for the extraction, processing and matching of biometrics under adverse night conditions in the presence of either natural or artificial illumination. Our work includes capture, analysis and evaluation of a broad range of electromagnetic bands suitable for night-time image acquisition, including visible light, near infrared (IR), extended near IR and thermal IR. We develop algorithms for human detection and tracking from night-time imagery at ranges between 5 and 200 meters. Identification algorithms include face, iris, and gait recognition, supplemented by soft biometric features. Our preliminary research indicates the challenges in performing human identification in night-time environments.

Keywords Visible/infrared spectrum · Face/iris/gait recognition · Soft biometrics · Performance evaluation

1 Introduction

Biometrics plays a pivotal role in human identification. Determination of human activity patterns in groups relies on a combination of local and global subject differentiation. Local differentiation can be obtained through soft biometrics whereby each subject is differentiated within the group and tracked. Global subject identification uses strong biometrics where the actual identity of individuals is ascertained from comparisons with biometric watch lists or other biometric databases. Within the last two decades, we notice improvement in the performance of biometric systems in outdoor daytime conditions at various acquisition distances. Subject presentations support understanding of human network activity in daylight. However, it is of equal if not greater importance to redirect the focus of such research to night environments. The extraction, processing and matching of biometrics under adverse night conditions in the presence of either available natural or artificial illumination is an open area of study.

In order to motivate on the scope of the study, we begin by surveying the spectral ranges of potential interest for use in biometric recognition in night environments under either active or passive illumination. Table 1 summarizes the electromagnetic bands of interest, their wavelength range, detection devices and illumination sources. The visible, near IR, extended IR and thermal IR bands have the best intersection of night vision and illumination ranges of operational interest. Further, the technology is available for our study.

In this chapter we provide a broad analysis and evaluation of techniques used for ascertaining the identity in night environments. The contributions of this work

Table 1 Imaging ranges of interest for night environments

Electromagnetic Band	Wavelength Range	Detection Source	Illumination Source	Viability
Ultraviolet	<450 nm	GaN Detectors, Schottky Devices	GaN LEDs, discharge sources, Lightning, Explosions	Low
Visible	450–750 nm	High sensitivity Si CCDs	Ambient light, Broadband sources	High High
Near IR	750–1100 nm	Si CCDs	AlGaAs LEDs, LDs	High
SWIR	1100–2500 nm	InGaAs Focal Plane	InGaAs LEDs, LDs	High
Thermal IR	7–14 μm	BST (Bolometric), CdTe Devices	Subject reflected ambient thermal	High
Submillimeter, Millimeter	0.1–3 mm	High-f electronics (SiGe), Optoelectronics	Gyrotron, far IR or free electron lasers	Low
Microwave	centimeters	High-f diodes, array	Magnetron, klyston	Low

are four-fold. *First* (see *Sect. 2*—Nicolo, Zuo, Schmid), the problem of Color-NIR cross-spectral iris matching will be broadly studied. In this section we will show some preliminary results of a nonlinear adaptive model to predict the value of the NIR channel from a visible range iris image. *Second* (see *Sect. 3*—Bourlai, Kalka, Jafri, Whitelam, Cukic, Dawson, Hornak, Ross), we will discuss the problem of face verification across the short wave infrared spectrum (SWIR). We will elaborate on the biometric image acquisition process, the sensors and the data collection setup used. We will also demonstrate the possibility for SWIR inter-wavelength matching. *Third* (see *Sect. 4*—Decann, Ross), we will describe our efforts on the extraction of gait curves, i.e., a spatio-temporal representation of silhouette shape dynamics, for human gait identification. *Fourth* (see *Sect. 5*—Cao, Adjeroh), we will describe our preliminary work towards categorizing and grouping people in surveillance video using metrological features (e.g. human body shape) and other soft biometrics (e.g. gait, age, etc.). The chapter will conclude in *Sect. 6* where we will discuss our current results, and provide some future research directions.

2 Color-NIR Cross-Spectral Iris Matching

Iris images captured in near infrared (NIR) band are a traditional input to iris recognition systems [3]. In recent years, however, biometricians turned their attention to iris images acquired in the visible band of electromagnetic spectrum [10]. This trend is supported by a variety of factors: (1) optical cameras in visible range are cheap and characterized by a very high resolution; (2) in terms of the application, most of security and surveillance camera systems (installed for monitoring various human

activities) are visible range cameras. Apart from monitoring activities these cameras may capture face or iris images, which further may be used to authenticate a suspicious or violent individual. Thus iris images acquired by the cameras in visible spectrum are query images, that have to be compared against enrolled database of irises, which is traditionally formed from the iris images in the NIR range. Although some publications are available on the topic (for example, publications involving UBIRIS dataset and WVU multispectral iris dataset), the investigation of matching NIR and visible range iris images has been very limited. Generally, the problem of cross-spectral iris matching remains to be solved.

In this section, we, propose a nonlinear adaptive model to predict the value of the NIR channel from a visible range iris image. The predicted value of the NIR channel will be compared with real NIR iris images by using the classical log-Gabor filter-based algorithm [8]. The results of the performance evaluation are promising and outperform NIR vs. IR cross match comparison.

2.1 Multispectral Iris Dataset and Data Used in Simulations

Multispectral iris data collected at WVU (see [2] for detailed description of the dataset) are used to demonstrate the performance of the proposed method. MS3100 camera manufactured by Geospatial Systems, Inc. was employed for data requisition. The MS3100 camera is a 3-chip multispectral digital camera with 1.4M+ pixels per sensor. The multispectral iris dataset collected at WVU involves data of 35 users (68 classes, 232 images). Each iris snapshot outputs the data acquired by three different spectral channels, which are acquired and saved simultaneously. They are: NIR, R and G/B channels. They are perfectly aligned and synchronized. In our cross match test, 25 users (50 classes, 150 images) were used as the testing data, and 10 images from the remaining users were used as the training data to design a prediction model.

2.2 Proposed Predictive Model

Due to a variety of pigmentation color in irises, all irises can be broadly partitioned in three groups: blue, brown, and green. Each color mentioned above generates its own predictive model. Thus, the models for the blue irises, brown irises and green irises have to be trained separately. It has been empirically confirmed that a prediction model trained on a blue iris performs well only on blue irises.

The predictive model is trained as follows. Training in this case implied estimation of a nonlinear mapping between input and output parameters. For each pixel in a snapshot (composed of three channels R, G/B, and NIR) selected for training, the input and output parameters of the mapping with the pair (value of R channel, value of G/B channel) as an input and a value of NIR channel as an output are recorded

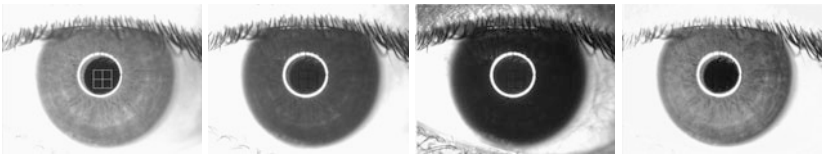

Fig. 1 (**a**) The original NIR channel; (**b**) the R channel; (**c**) the G/B channel; and (**d**) the predicted NIR channel using (**b**) and (**c**) as the input to adaptive mapping

and passed to the training tool, such as a Neural Network (NN). The type of NN and the number of layers in a NN are selected adaptively based on the availability of training data. Once the mapping is estimated from training data, it will be further used to predict NIR image of a testing iris images from R and G/B channels.

An example showing an iris snapshot (its original three channels: R, B/G, and NIR) and a generated NIR channel are shown in Fig. 1. From this example, it is clear that the predicted NIR channel visually is very similar to the NIR channel and substantially differs from the R and G/B channels. Note, however, that the predicted NIR channel is slightly over-smoothed compared to the original NIR channel, and this may reflect on recognition performance of the predicted NIR channel.

2.3 Recognition Performance

To verify the recognition performance all testing iris images were segmented using the robust iris segmentation algorithm developed by Zuo and Schmid [18]. The data were further encoded using an elaborated version of the Masek's iris encoding algorithm (see [17] for more details). Three similarity matrices ware formed and analyzed: (1) true NIR imagery vs. true NIR imagery (viewed as the best achieved performance); (2) true NIR vs. predicted NIR channel illustrate the promise of the proposed predictive model); and (3) true NIR vs. R channel (viewed as the baseline comparison).

The ROC curves for these experiments can be found in Fig. 2. The horizontal axis is FAR (False Acceptance Rate) and the vertical axis is GAR (Genuine Acceptance Rate). The average genuine scores, the average impostor scores, d-prime and the Equal Error Rate (EER) are summarized in Table 2.

The achieved performance so far is approximately in the middle of the baseline and the upper bound. The upper bound is not reachable because not all information in NIR channel is predictable. However, in our future work, more sophisticated prediction will be developed to release its full potential.

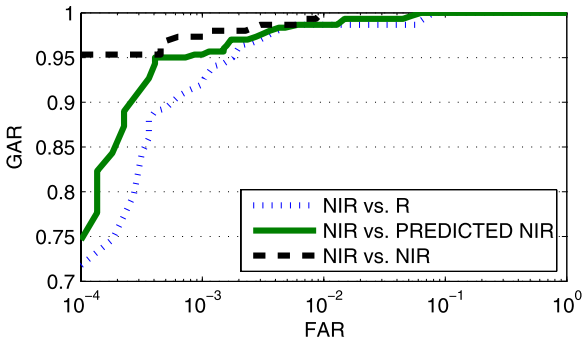

Fig. 2 The performance improvement due to the proposed predictive model

Table 2 Performance of the three cross comparisons

	Mean of Genuine Scores	Mean of Impostor Scores	d'	EER
NIR vs. NIR	0.0671	0.4093	7.9266	0.0069
NIR vs. Predicted NIR	0.0967	0.3907	6.8603	0.0131
NIR vs. R	0.1300	0.4118	5.5570	0.0137

3 Short Wave Infrared Face Verification

Recognition of faces using different imaging modalities in the visible and infrared spectrum has become an area of growing interest [12, 16]. Short wave infrared (SWIR) imagery represents a viable alternative to visible imaging in the search for a robust and practical identification system. Since SWIR 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.

There are several benefits in using SWIR spectrum for face recognition: (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 night-time environment. For example, in Fig. 3(a) we can compare a face image (1) in the visible spectrum (controlled vs. uncontrolled conditions), and (2) at different SWIR wavelengths captured either at a day or night 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 SWIR images, e.g., Fig. 3(b) illustrates the skin texture of a subject in the visible and SWIR spectrum respectively.

In this work we study the problem of face verification across the SWIR spectrum. We utilize several face recognition techniques using the West Virginia University Multispectral database (WVUM). The database is composed of 30 subjects, and is collected using visible, multispectral and SWIR cameras.

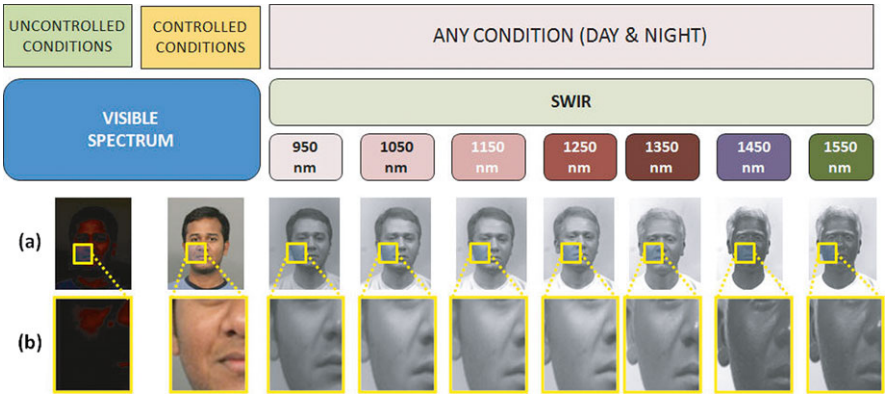

Fig. 3 Face images of a subject: (**a**) in the visible spectrum captured at controlled or uncontrolled conditions; (**b**) at different SWIR wavelengths. The advantages of SWIR over the visible spectrum are: (1) in harsh environmental conditions characterized by unfavorable lighting and pronounced shadows (such as a night environment) SWIR images are not affected; (2) Facial features—such as texture—cannot be observed in the visible spectrum under uncontrolled conditions. However, they are observable in SWIR images

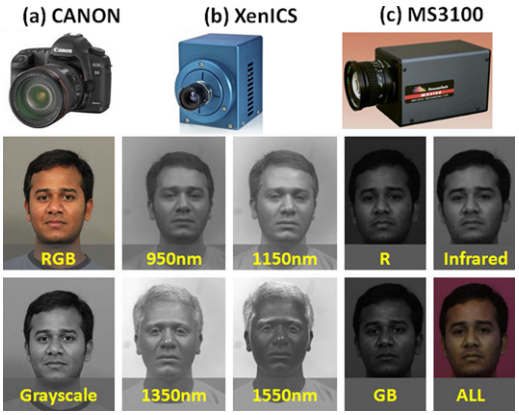

Fig. 4 (**a**) *Canon EOS 5D Mark II* ultra-high resolution digital camera. (**b**) MS3100 multispectral camera. (**c**) XenICS SWIR camera

3.1 SWIR Data Collection

Three different types of cameras have been used in this work, i.e. a Canon EOS 5D Mark II, a DuncanTech MS3100, and a XenICS Xeva InGaAs. These cameras are described below and they cover the visible, NIR, and a significant range of the SWIR spectrum:

Canon EOS 5D Mark II: This digital SLR camera (www.canon.com) has a 21.1-megapixel full-frame CMOS sensor with DIGIC 4 Image Processor, and a vast ISO Range of 100–6400. It has also *Auto Lighting Optimizer* and *Peripheral Illumination Correction* that enhance its performance. In this work, the Canon (see Fig. 4(a)) is used to obtain standard RGB, ultra-high resolution frontal pose face images.

XenIC: A XenIC XEVA-818 camera served as the imager for this study (see Fig. 4(b)). The camera has an Indium Gallium Arsenide (InGaAs) 320 × 256 Focal Plane Array (FPA) with 30 μm 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 InfraRed (SWIR) band) across which the InGaAs FPA has largely uniform quantum efficiency. Response falls rapidly at wavelengths lower than 950 nm and near 1700 nm. In the work of this section it is used for the visible-IR, and IR-IR (intra-wavelength) face recognition comparative studies.

DuncanTech MS3100: This camera (see Fig. 4(c)) incorporates three CCD and three band-pass prisms behind the lens to simultaneously capture four different wavelength bands (no explicit image registration or alignment across multiple channels is necessary). The IR and Red (R) sensors of the multispectral camera are two separate Sony ICX205AL sensors whose spectral response ranges from 400 nm to 1000 nm. The Green (G) and the Blue (B) channels are recorded on the same Sony RGBICX205 RGB sensor, with a blue response from 400 nm to 550 nm, and a green response from 400 nm to 650 nm. Note that the IR and Red sensor outputs an image of size 1392 × 1040. The G and the B images are recorded on a RGB Bayer pattern sensor and are, therefore, one-third the resolution of the other images. Then G/B images are extracted and scaled (linear interpolation) to have the same resolution as the IR and R images.

These cameras were used to acquire the *WVU Multispectral* database that is composed of three datasets, i.e. visible, multispectral, and SWIR. The purpose of acquiring such a database was to capture face images under different spectra in order to benefit from the information variability provided by different modalities, and study the recognition performance of different techniques. The live face capture setups we used are illustrated in Fig. 5(a) and (b). The database was collected over two sessions spanning one month. In total, 30 subjects participated, 22 male and eight female.

Each session of the *Visible* dataset has 15 frontal face images for each subject, resulting in 450 images in total. In each session of the *Multispectral* dataset, we acquired face images at three different poses, i.e. full frontal, and left/right at $+/-67.5$ degrees. For each pose and with a single shot we captured four images that are associated with the Green/Blue, Red, IR, and RGB/Color-IR (combined) component respectively (360 images in total). 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, viz. a set of hardware/lenses that were placed in front of the camera before each shot. The wavelength range covered starts from 900 nm, and goes up to 1700 nm, using filters that are 100 nm wide and centered at 950, 1050, . . . , 1650 nm. Thus, nine samples per pose (no filter, and 950–1650 nm/100 nm) have been acquired, resulting in 810 images in total (27 samples/session$\times$30 subjects).

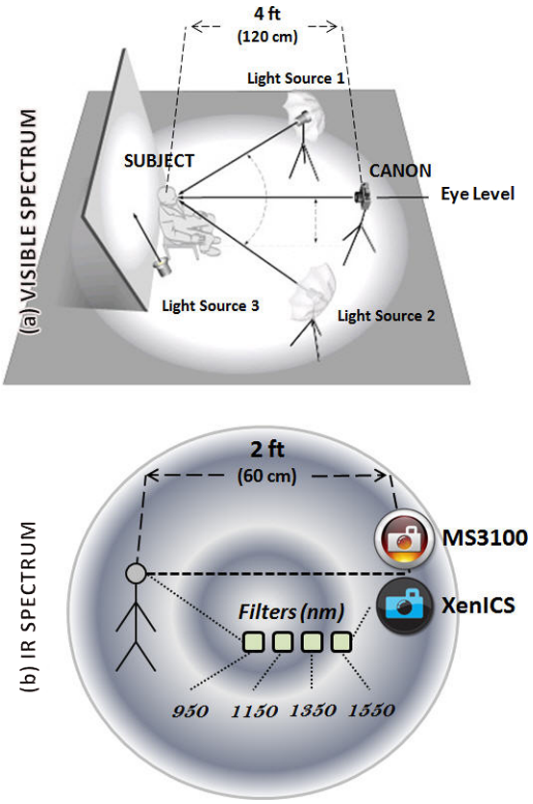

Fig. 5 Image capture setups. (**a**) Visible Spectrum: the live subject-capture setup using the Canon digital camera (taken from US Department of US State Department, Bureau of Consular Affairs). More details have been added by the authors. (**b**) Visible-IR Spectrum: live subject-capture setup using the XenICS and MS3100 cameras

3.2 Methodology

Preprocessing and Normalization We employed the Viola & Jones *face detection* algorithm to the original data. It is used to localize the spatial extent of the face and determine its boundary. In the next step, a *geometric normalization* scheme is applied to images acquired after detection. The normalization scheme compensates for slight perturbations in the frontal pose, and it is composed of eye detection and affine transformation [15]. Eye detection is based on a template matching algorithm where the eye coordinates are obtained. Then the canonical faces are constructed by applying affine transformation. Eventually, all faces are warped to the same dimension of 300×300. Then, *photometric normalization* is employed by a simple contrast and brightness adjustment.

Matching Experiments We employed two face recognition methods, i.e. *Linear Discriminant Analysis* (LDA) and *Principal Components Analysis* (PCA) with k-nearest neighbor (k-NN). PCA was used only in the IR-IR fusion study where we have only one sample per subject for each session.

By using the visible and IR datasets of *WVU Multispectral* database, three different types of face verification experiments have been performed: (1) Visible-Visible,

(2) Visible-IR, and (3) IR-IR. The *first* experimental scenario is a high quality vs. high quality study for the purpose of establishing a baseline for comparison. In the *second* experiment we compare the infrared subset of the MS3100 camera, and the SWIR subset of the XenICS camera, with their high resolution counterparts acquired in the visible spectrum. The goal here is to investigate the possibility of matching SWIR images to visible images. In the *third* experiment we analyze intra-wavelength matching.

3.3 Results

SWIR Face Verification Table 3 presents the results for all experiments in terms of Equal Error Rate (EER). The first and second experiment utilized ten images for training, while only one image was available for training in experiment three. In regards to testing, experiment one utilized three images, while only one image was available for testing in experiments two and three.

For experiment one (Visible-Visible), LDA achieves an EER of 11.78% which fairly consistent with the literature. In experiment two, we compare visible to SWIR wavelengths, starting at 950 nm and ranging to 1550 nm in 200 nm intervals. The overall trend for experiment two illustrates that as wavelength increases, performance degrades. This degradation in performance is more specific at high SWIR wavelengths. Experiment three investigates intra-wavelength matching. Performance is consistently better than the inter-wavelength matching but follows a similar trend, i.e., as wavelength increases performance degrades.

The main question here is why for both inter- and intra-wavelength matching the appearance based face recognition algorithms employed (i.e., LDA and PCA) demonstrate degrading performance as wavelength increases. As illustrated in Fig. 3, at longer wavelengths (i.e., above 1350 nm) human faces—with respect to intensity and texture—are vastly different when imaged under shorter wavelengths.

Table 3 Equal Error Rate (EER%) for all Verification experiments utilizing LDA (all inter-wavelength scores) and PCA (IR intra-wavelength scores). Nofilter corresponds to images taken in absence of IR cut-filters with the XenIC camera

		Session Two						
		Visible	MS3100	Nofilter	950 nm	1150 nm	1350 nm	1550 nm
Session One	Visible	11.78	15.60	38.30	29.00	31.74	45.63	48.60
	MS3100		13.33					
	Nofilter			19.22				
	950 nm				13.68			
	1150 nm					19.96		
	1350 nm						20.15	
	1550 nm							23.23

This difference is specifically noted in skin pigmentation, viz., the longer the wavelength, the darker the pigmentation. Similarly, facial hair at longer wavelengths becomes bleach white (see Figs. 3, 4). To help compensate for these differences we have employed the contrast limited adaptive histogram equalization (CLAHE) method [11] that performed better that other photometric normalization techniques (e.g., histogram equalization) [1]. In this work a reduction in EER was observed when using *Identity Tools G8*[1] for both inter- and intra-wavelength matching. In the same work we also concluded that match score fusion, an alternative to image normalization, can help reduce error rates at longer wavelengths.

4 Gait Curves for Human Recognition in a Night-Time Environment

The solution presented in this chapter concerns the extraction of gait curves for human gait identification. Gait curves are, in essence, a spatio-temporal representation of silhouette shape dynamics and were inspired from previous work by Wang et al. [14]. Thus, this approach is representative of the model-free variety. The primary advantage of a model-free approach is simplicity, as features are strictly derived from silhouette shape dynamics. However, silhouette distortions such as clothing changes are consistently an area of concern as they can impact the overall shape dynamics. Evaluation of the data was performed on the CASIA Night Gait Dataset.[2]

4.1 Methodology

Preprocessing and Silhouette Extraction Thermal imagery differs significantly from that of the visible spectrum. Lower contrasts and resolutions are common, and pixel variation between the subject and background may be small. This is especially the case for the CASIA Night Gait Dataset. Thus, prior to attempting silhouette extraction, the intensity values of the k^{th} input image I_k are adjusted such that 1% of the pixel values are saturated at the extrema of 0 and 1 respectively, resulting in image I_{adj}. Background subtraction is then used to create a difference image in time. Background noise and additional anomalies are removed through a series of threshold filters and morphological operations. The resulting binary image is denoted by $S = \{s_{i,j}\}$ where $i = 1, 2, \ldots, n_h$, $j = 1, 2, \ldots, n_w$, n_h is the height of the image and n_w is the width of the image.

$$I_{\text{diff}} = \text{abs}(I_{\text{adj}} - I_0), \tag{1}$$

$$S = \begin{cases} 1, & I_{\text{diff}} > 0, \\ 0, & I_{\text{diff}} \leq 0. \end{cases} \tag{2}$$

[1] http://www.l1id.com/.

[2] Website: http://www.cbsr.ia.ac.cn/english/Gait%20Databases.asp.

Spatio-temporal Feature Extraction As a silhouette traverses across the viewing plane, key markers can be collected at each frame. In this scheme, the markers of interest lie in identification of the coronal plane of a silhouette. Treating the coordinates of S as a matrix, let $S_{\min}$ represent the minimum row entry whose value is one. This entry represents the top of the head, and is fairly consistent across gait sequences. The column that this entry occurs in is designated as the identifier for the coronal plane, which divides a silhouette into halves.

Following calculation of the coronal plane coordinates the left and rightmost pixel locations of the outermost silhouette contour are obtained for each row in the silhouette. That is, for the pth row where $p \in [S_{\min}, S_{\max}]$ these pixels are denoted as g_p^{left} and g_p^{right}, respectively. Subtraction of the column position of the coronal to these point sets yields a space normalized contour, denoted as the gait curve, G_k, for the k^{th} frame in the sequence. Thus, the evolution of the gait curve across several frames can be regarded as a spatio-temporal feature for shape based analysis.

The next step regards representation of a set of gait curves as a discriminatory feature. An arbitrary function, $F(G_1, G_2, \ldots, G_N)$ whose input is a set of gait curves, can manipulate this information to produce an output encompassing the shape dynamics of a particular video sequence. The Procrustes meanshape [7, 13] is a mathematically elegant measure of representing and evaluating shape set similarity and is chosen to represent $F(\mathbf{G})$. This measure is particularly attractive because subsequent distance comparisons are invariant to translation, rotation and scale. In order to use this measure, the dimension for all G_k's must be normalized such that they are equal. The set is also vectorized by conversion from spatial coordinates to the complex plane. Subtraction of the vector mean aligns each gait curve at the origin.

$$\mathbf{z}_k = \text{Re}(G_k) + j\text{Im}(G_k); \tag{3}$$

$$\bar{\mathbf{z}} = \sum_{i=1}^{k} \frac{\mathbf{z}_i}{k}; \tag{4}$$

$$\mathbf{u}_k = \mathbf{z}_k - \bar{\mathbf{z}}; \tag{5}$$

$$\mathbf{u} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N]; \tag{6}$$

$$S_u = \sum_{j=1}^{N} \frac{(\mathbf{u}_j \mathbf{u}_j^T)}{(\mathbf{u}_j^T \mathbf{u}_j)}. \tag{7}$$

Following the series of K frames, a series of K vectorized gait curves are created. Extracting the first eigenvector of scatter matrix S_u results in the averaged gait curve, denoted as $\bar{\mathbf{G}}$. Examples for three different subjects are found in Fig. 6. The procrustes distance between any two shape representations, $(\bar{\mathbf{G}}_1, \bar{\mathbf{G}}_2)$, is then

$$\mathbf{d}(\bar{\mathbf{G}}_1, \bar{\mathbf{G}}_2) = 1 - \frac{|\bar{\mathbf{G}}_1^T \bar{\mathbf{G}}_2|^2}{\|\bar{\mathbf{G}}_1\|^2 \|\bar{\mathbf{G}}_2\|^2}, \tag{8}$$

where, the smaller the resulting value, the similar the shapes.

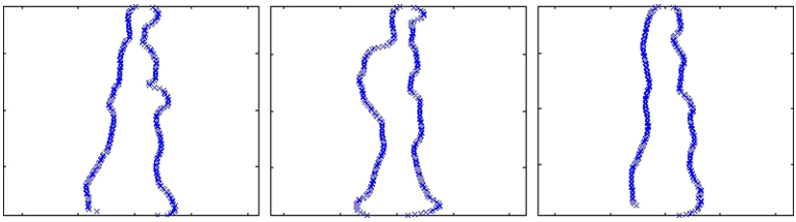

Fig. 6 Primary eigenvector of scatter matrix S_u for three different gait curves

Table 4 Experimental results on CASIA night gait dataset

EXP	Type	Probe Size	Gallery Size	Match Score (Rank = 1)	Match Score (Rank = 3)	Match Score (Rank = 5)
A	Normal vs. Normal	612	612	91.00%	96.01%	97.70%
B	Normal vs. Bag	612	306	25.50%	41.20%	49.30%
C	Normal vs. Slow	612	306	65.40%	78.40%	80.70%
D	Normal vs. Fast	612	306	69.90%	80.40%	85.30%
E	Slow vs. Slow	306	306	85.00%	92.20%	94.10%
F	Fast vs. Fast	306	306	79.10%	85.90%	88.90%
G	Bag vs. Bag	306	306	81.00%	86.90%	89.20%
H	All Sequences	1530	1530	88.40%	93.30%	95.20%

4.2 Experiments and Results

Classification of Human Gait The objective of this experiment is to illustrate the performance of gait curve matching as a function of walk type. A probe (test) set is compared against a gallery (training) set. Since the CASIA dataset offers scenarios where subjects walk at normal pace, fast pace, slow pace, and with a bag, there are four distinct test sets. A fifth set is used to represent the combination of all data. Evaluation by comparing differing sets of pace aids is illustrating the effects of walking speed and silhouette distortion. Also, for the experiments that evaluate similar or all paces, the matching probe sequence is left out of the gallery. These results are summarized in Table 4.

4.3 Discussion

As previously stated, the experiments reflect the effectiveness of the algorithm based on variability in walking type. From Table 4, the results show that speed has a profound impact on classification. The classification rates for rank-1 is significantly more accurate for similar speeds than those that differ. Noting these results, the functionality of this algorithm can be concluded to be optimal when the gallery is

diverse enough to include multiple paces. This can be verified directly by observing the results of experiment H, as well as through noting that experiments E, F and G performed better than their C, D, and B counterparts despite only having one training sample to match against. In more recent work, the variability caused due to backpacks has been significantly mitigated resulting in improved performance for Experiment B [4].

5 Soft Biometrics—Body Measurement

For problems such as identification under night-time environments or identification at a distance, most of the traditional hard biometrics such as fingerprints, face, and iris may not be readily available. Further, under such challenging situations, the video or image sequence from a single modality will typically be corrupted by various sources of noise, such as atmospheric noise, weather, and uneven ambient illumination. In such environments, extraction of traditional primary biometrics (such as facial features and iris-based features) from one imaging source could be very challenging. One approach to the problem is the use of multispectral signals from certain biometric traits such as face and iris to improve our ability to identify people under such night-time environments. A major part of this chapter has been devoted to this approach.

A complementary approach could be to exploit potential secondary or soft biometric traits [6] that could be automatically extracted from such typically poor quality video or images. Thus, metrological features (such as human body shape, anthropometric measurements, and geometrical features that can be extracted from these measurements) and other soft biometrics (such as gait, age, gender, weight, skin texture, etc.), provide a first approach to initial categorization and grouping of people in such videos. This could be used for a quick appraisal of the individual, and perhaps eliminate unlikely candidates. Depending on the application, the soft biometric features could also provide us with an easier and cheaper method to handle subjects at a relatively far distance.

Key challenges in the use of static soft biometrics such as body size and body geometry in automated human identification or human grouping include:

(1) Automated extraction of the measurements from an image or from a video sequence;
(2) Handling of missing body parts and errors in the automated extraction;
(3) Human modeling and prediction using the extracted body measurements;
(4) Generating robust and discriminative features from the extracted measurements.

An important step in addressing the above problems is to identify possible similarities between body part dimensions, or measurements, for instance, based on the pair-wise correlation between them. Towards this end, we performed a brief experiment using the CAESAR anthropometric dataset. The CAESAR (Civilian American and European Surface Anthropometry Resource) dataset (http://store.sae.org/caesar/) contains anthropometric measurements on 2,400 U.S.

& Canadian and 2,000 European civilians (both men and women), ages 18–65. The dataset contains both 3D surface points and 1D measurements, including 40 traditional measurements, manually acquired using a measuring tape and a caliper. Our work on soft biometrics is based on 1D measurements obtained from the 2,400 US and Canadian civilians.

We used 45 human body measurements or attributes (including gender and weight) from the dataset, and computed the pair-wise correlation coefficient between them. Using the correlation coefficients, we generate the correlation graph, which captures a form of similarity relationships between the sizes and dimensions of different human body parts. A node in the correlation graph corresponds to a body measurement or attribute, while an edge between two nodes implies that the corresponding measurements represented by the nodes are correlated (up to a certain threshold on the correlation coefficient). Figure 7 shows the correlation graph between some of the measurements in the CAESAR dataset. As can be observed from the figure, the measurements generally fall into two basic groups: the 2D groups containing mainly circumferences of certain body parts, and the 1D groups that contain mainly one-dimensional measurements, such as lengths and heights. It is surprising that the simple correlation measure is able to partition the measurements into various groups with a physical meaning. An interesting on-going work is the characterization of the correlation graph using methods of information network analysis, such as centrality and betweenness measures [5, 9]. We have successfully

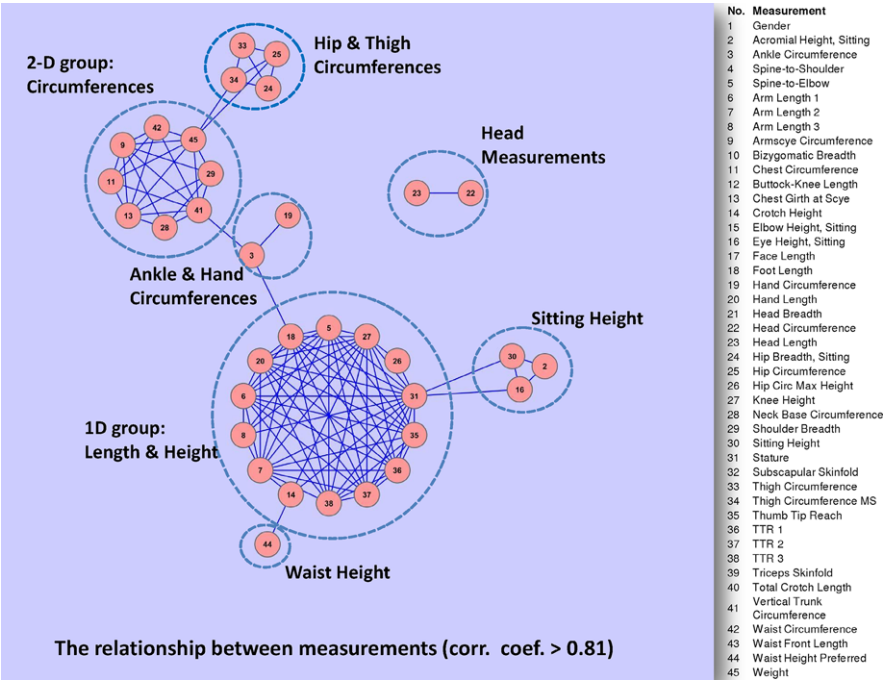

Fig. 7 Grouping of human body measurements based on their correlation graph

used the measurements to predict certain soft biometrics, such as weight and gender, to a high accuracy. We are currently investigating how the relationships between the measurements, for instance, as captured in the figure, can be used to address the above key challenges.

6 Summary

In this chapter we presented a preliminary study on the problem of ascertaining identity in night-time environments, and highlighted a new set of challenges that current technology does not fully address. In particular, we showed how different biometric modalities operate on different sub-bands of the visible and IR spectra. In Sect. 2 we looked at the challenges associated with cross-spectral iris matching, and demonstrated the feasibility of matching irides imaged under the visible spectrum to irides imaged under the traditional NIR spectrum. In Sect. 3 we studied the performance of multispectral face recognition utilizing only various bands from the SWIR spectra. Our results indicate that as spectral wavelength increases recognition performance degrades. It is often the case that hard biometrics such as iris or face may not be readily available in a night-time environment as a result of image noise or subject stand-off distance. Thus, in Sect. 4 we presented a gait analysis as a supplementary modality for identity verification under such operating environments. Our results indicate that subject velocity has a profound impact on identification rate. Finally, in Sect. 5 we described the utilization of complementary information such as soft biometrics.

Acknowledgements This work was funded in part by grants from the Office of Naval Research, and the National Science Foundation's Center for Identification Technology Research (CITeR). The authors are grateful to all colleagues and students, staff and faculty at West Virginia University for their support and assistance with the data collection process.

References

1. Bourlai, T., Kalka, N.D., Ross, A., Cukic, B., Hornak, L.: Cross-spectral face verification in the short wave infrared band. In: International Conference on Pattern Recognition (2010)
2. Boyce, C., Ross, A., Monaco, M., Hornak, L., Li, X.: Multispectral iris analysis: A preliminary study. In: Computer Vision and Pattern Recognition Workshop on Biometrics (CVPRW), pp. 51–59 (2006)
3. Daugman, J.: Iris recognition. Am. Sci. **89**, 326–333 (2001)
4. Decann, B., Ross, A.: Gait curves for human recognition, backpack detection, and silhouette correction in a nighttime environment. In: Proc. of SPIE Conference on Biometric Technology for Human Identification VII, Orlando, USA (2010)
5. Girvan, M., Newman, M.E.: Community structure in social and biological networks. Proc. Natl. Acad. Sci. **99**, 7821–7826 (2002)
6. Jain, A.K., Dass, S.C., Nandakumar, K.: Soft biometric traits for personal recognition systems. In: Proceedings of International Conference on Biometric Authentication, pp. 731–738 (2004)

7. Kent, J.: New directions in shape analysis. In: Informatics and Mathematical Modeling, pp. 115–127. Wiley, Chichester (1992)
8. Masek, L.: Recognition of human iris patterns for biometric identification. Bachelor's Dissertation, The School of Computer Science and Software Engineering, The University of Western Australia (2003)
9. Newman, M., Barabasi, A.-L., Watts, D.J.: Structure and Dynamics of Networks. Princeton University Press, Princeton (2006)
10. Proenca, H., Filipe, S., Santos, R., Oliveira, J., Alexandre, L.A.: The UBIRIS.v2: A database of visible wavelength images captured on-the-move and at-a-distance. IEEE Trans. Pattern Anal. Mach. Intell. **99** (2009) (PrePrints)
11. Reza, A.: Realization of contrast limited adaptive histogram equalization (clahe) for real-time image enhancement. VLSI Signal Process. **38**, 35–44 (2004)
12. Socolinsky, D., Selinger, A., Neuheisel, J.: Face recognition with visible and thermal imagery. Comput. Vis. Image Underst. **91**, 72–114 (2003)
13. Stegmann, M., Gomez, D.: A brief introduction to statistical shape analysis. Inf. Math. Model. (2002)
14. Wang, L., Ning, H., Hu, W., Tan, T.: Gait recognition based on procrustes shape analysis. In: ICIP, pp. 433–436 (2002)
15. Whitelam, C., Jafri, Z., Bourlai, T.: Multispectral eye detection: A preliminary study. In: International Conference on Pattern Recognition (2010)
16. Yoshitomi, Y., Miyaura, T., Tomita, S., Kimura, S.: Face identification using thermal image processing. In: WRHC, pp. 374–379 (1997)
17. Zuo, J., Schmid, N.A.: On a local ordinal binary extension to gabor wavelet-based encoding for improved iris recognition. In: Proceedings of 2007 Biometric Consortium Conference (BCC'07) (2007)
18. Zuo, J., Kalka, N.D., Schmid, N.A.: A robust iris segmentation procedure for unconstrained subject presentation. In: Proceedings of 2006 Biometric Consortium Conference (BCC'06) (2006)