

# MATCHING HETEROGENEOUS PERIOCCULAR REGIONS: SHORT AND LONG STANDOFF DISTANCES

*Zhicheng X. Cao and Natalia A. Schmid*

West Virginia University, Lane Department of CSEE, Morgantown, WV 26506-6109

## ABSTRACT

Partial face recognition has been a problem of interest for more than a decade. Most of previous publications on partial face recognition assume intra spectral matching. Matching Short Wave Infrared (SWIR), Middle Wave Infrared (MWIR) or Near Infrared (NIR) images of partial face to a gallery of color images is a much more challenging task. The photometric properties of images in these four spectral bands are highly distinct. Because of the limited space - and also sufficient interest to this biometric - in this paper we present results of cross spectral matching applied to periocular regions. Equipped with a well developed automatic recognition algorithm for heterogeneous face, we demonstrate that the algorithm can be tuned and applied to periocular regions for a positive cross spectral matching of SWIR, MWIR and NIR periocular regions to visible periocular regions at short (1.5 m) and long (50 and 106 m) standoff distances. Our numerical analysis demonstrates the results of the matching. To the best of our knowledge, the performance evaluation presented in this paper is the first of its kind.

**Index Terms**— heterogeneous periocular recognition, heterogeneous partial face recognition, intra spectral matching, cross spectral matching, short wave infrared, middle wave infrared, near infrared

## 1. INTRODUCTION

Periocular recognition has been an active research area in the past five years [1, 2, 3]. Periocular regions can be categorized as a part of face, perhaps most visible in the presence of facial occlusions. On the other hand, it can add to reliable iris recognition (in visible and NIR spectra), especially when the quality of iris images is low. Although many research challenges such as unconstrained subject's presentation, uneven illumination, and partial occlusions have been previously addressed in the literature [1, 4, 5, 6], many challenges remain.

---

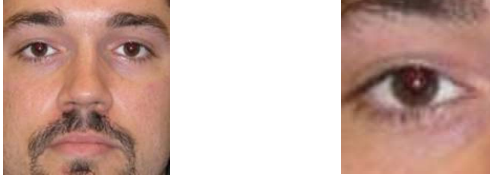
The authors would like to thank Brian Lemoff of West Virginia High Technology Consortium Foundation for providing the Pre-TINDERS and TINDERS datasets employed in the described experiments. The authors are also grateful to Scott McCallum, System Analyst from Pinellas County Sheriff's Office, for providing the PCSO face dataset.

Furthermore, as new practical applications evolve, new challenges offered by the applications arise and hence a need for development of new algorithms to mitigate them. Surveillance at night or in harsh environments has appealed to new imaging modalities - NIR, SWIR, MWIR and Long Wave Infrared (LWIR) - and evolved into new application.

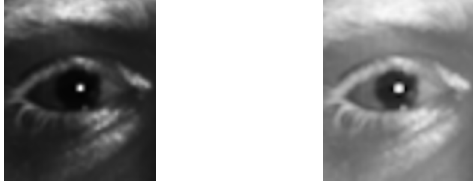
Intra spectral (within the same spectral modality) matching of periocular regions have been previously explored (see for example, [1, 4, 5, 6, 7]). However, results of heterogeneous matching of periocular regions have not been previously reported, and this motivates the experiments described in our paper. If addressed, heterogeneous matching of periocular regions will provide a baseline for the development of new improved algorithms for heterogeneous face and periocular recognition at night or in challenging environments. In addition to cross spectral matching we explore the effect of varying standoff distances on the recognition performance.

Few publications on cross spectral face recognition have appeared in the literature. Most of them were focused on algorithms for and analysis of matching NIR, SWIR, MWIR and LWIR face images to a gallery of visible face images [8, 9, 10, 11, 12, 13]. Some publications assumed short standoff distances, while others explored the case of varying standoff distances [14, 11, 8]. Popular algorithms such as Local Binary Pattern (LBP), Scale-Invariant Feature Transform (SIFT), Histogram of Oriented Gradients (HOG) [15], Gabor Ordinal Measures (GOM) [16] and their variants have been used for feature extraction and matching in the past [17]. In this work, we use a method based on Gabor LBP and Generalized LBP combined with Gabor Weber descriptors. The algorithm was originally developed in our lab for cross spectral face recognition [8, 12]. We demonstrate that the algorithm can be adapted to successfully match heterogeneous periocular regions at short (1.5 m) and long (50 and 106 m) standoff distances. Its performance is compared to that of the original LBP and the GOM-based algorithm. The latter is one of the newly introduced state-of-the-art algorithms for face recognition.

The remainder of the paper is organized as follows. Sec. 2 explains the adopted automatic matching algorithm and describes all modifications applied to it. Sec. 3 presents experimental set up and the results of numerical analysis. A brief summary is provided in Sec. 4.



(a) aligned and cropped face (b) aligned and cropped eye  
**Fig. 1.** Sample images of cropped face and periocular region.



(a) original eye in SWIR (b) log-transformed eye

**Fig. 2.** Illustration of log-transformation.

## 2. MATCHING ALGORITHM

We adopt a Gabor filter-based approach at the initial encoding stage, followed by an encoding scheme that involves three operators to extract robust features across different spectral bands. These three operators are designed to encode both magnitude and phase of filtered images. Details of this scheme can be found in [8]. Here we will provide only a brief description of the encoding method and highlight modifications.

### 2.1. Alignment and Cropping

Preprocessing steps involve image normalization, alignment, cropping and a simple intensity normalization. A canonical transformation is applied to all original face images in datasets resulting in images of size  $120 \times 112$ . The images are aligned by using the position of the eyes and nose. These anchor points are manually selected. Geometric transformations - rotation, scaling and translation are applied to each image such that the eyes and nose tip are in fixed positions (see Fig. 1 (a)). The aligned face images are further cropped and rescaled to generate periocular region (for the right eye) of size  $88 \times 72$  for usage in the following experiments (see Fig. 1 (b)). To do this, we crop the right eye from the face image with the pupil as center. Color images are converted to grayscale images. SWIR, MWIR and NIR images are pre-processed using a simple nonlinear transformation given by  $\log(1 + X)$  where  $X$  is the input image, as shown in Fig. 2.

### 2.2. Filtering and Encoding

Each normalized image in any spectrum is processed with a bank of Gabor filters at 2 different scales and 8 orientations generating 16 responses in total. The Gabor filter is given by:

$$G(z, \theta, s) = \frac{\|\mathcal{K}(\theta, s)\|}{\sigma^2} \exp \left[ \frac{\|\mathcal{K}(\theta, s)\|^2 \|z\|^2}{2\sigma^2} \right] \times \left[ \exp \{i\mathcal{K}(\theta, s)z\} - \exp \left\{ -\frac{\sigma^2}{2} \right\} \right], \quad (1)$$

where  $\sigma^2$  is the variance of the Gaussian kernel,  $\mathcal{K}(\theta, s)$  is the wave vector and  $s$  is a scale. The magnitude and phase of the wave vector determine the scale and orientation of the oscillatory term and  $z = (x, y)$ . A normalized and pre-processed face image  $I(z)$  is convolved with the Gabor filter  $G(z, \theta, s)$  at orientation  $\theta$  and scale  $s$  resulting in the filtered image  $Y(z, \theta, s)$ .

Both magnitude and phase responses are used to effectively encode heterogeneous data block by block. The magnitude response is encoded using two distinct operators: SWLD [18] and uniform LBP [19]. For encoding the phase response we adopt a uniform generalized LBP operator (GLBP). All operators consider the relationship among 12 neighbors at both radii  $r = 1$  and  $r = 2$ .

The SWLD operator is defined as:

$$SWLD_{l,r,12}(x) = \mathcal{Q}_l \left\{ \tan^{-1} \left[ \sum_{i=0}^{11} \left( \frac{x_i - x}{x} \right) \right] \right\}, \quad (2)$$

where  $x_i$  are the neighbors of  $x$  at radius  $r$  and  $\mathcal{Q}_l$  is a uniform quantizer with  $l$  quantization levels.

The uniform LBP operator is described as

$$LBP_{r,12}^u(x) = \mathcal{U} \left\{ \sum_{i=0}^{11} \mathcal{I}\{x_i - x\} 2^i \right\}, \quad (3)$$

where  $\mathcal{U}$  is the uniform pattern mapping and  $x_i$  are the neighbors of a value  $x$  at radius  $r$  and  $\mathcal{I}(\cdot)$  is the unit step function.

The uniform GLBP operator is a generalization of the encoding method introduced in [20]. It is defined as

$$GLBP_{r,12,t}^u(x) = \mathcal{U} \left\{ \sum_{i=0}^{11} \mathcal{T}_t\{x_i - x\} 2^i \right\}, \quad (4)$$

where  $\mathcal{T}_t$  is a thresholding operator based on threshold  $t$ .

For each of the 16 Gabor filters, there are 3 local operators working at 2 different radii, which yield 96 encoded responses in total for each input image.

### 2.3. Matching

Each encoded response is divided into 99 non-overlapping  $8 \times 8$  square blocks. Then 135-bin histograms of each block are normalized and concatenated to be treated as a probability mass function, resulting in a template of length  $135 \times 99 = 13,365$  features. When a distance between two images is evaluated, it is expressed as a sum of distances for all feature vector pairs.

Sum of two I-divergence distances [21] is used as the main metric to compare the feature vectors of images captured at different wavelengths. For two images  $A$  and  $B$  with the feature vectors  $H_A$  and  $H_B$ , respectively, the symmetric I-divergence distance is defined as:

$$D_{KL}(A, B) = \sum_{k=1}^K (H_A(k) - H_B(k)) \log \frac{H_A(k)}{H_B(k)}, \quad (5)$$

where  $K$  is the length of the feature vectors  $H_A$  or  $H_B$ .

### 3. NUMERICAL ANALYSIS

#### 3.1. Datasets

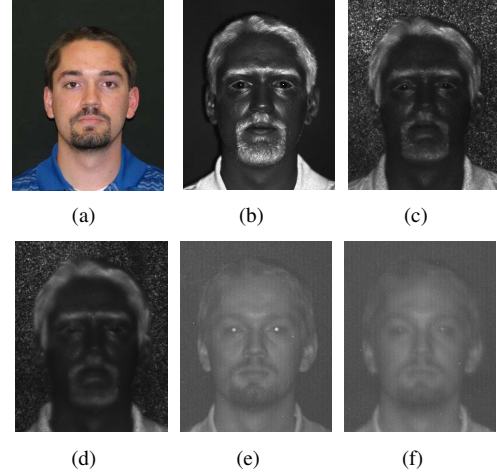
In our experiments we use three datasets. The first two datasets, Pre-TINDERS (Tactical Imager for Night/Day Extended-Range Surveillance) and TINDERS are collected by the Advanced Technologies Group, West Virginia High Tech Consortium (WVHTC) Foundation. The third dataset is collected by Pinellas County Sheriff's Office (PCSO).

Pre-TINDERS dataset is composed of 48 frontal face classes (total of 384 images) at two wavelengths - visible and 1550 nm SWIR. Images are acquired at a short standoff distance of 1.5 m in a single session. Four images per class are available in each spectral band - two images have neutral expression and two images show the talking person. A 1550 nm light source is used to illuminate the face in the SWIR spectral band. The original resolutions of the acquired images (see Fig. 3) are  $640 \times 512$  (png format) for SWIR images and  $1600 \times 1200$  (jpg format) for color images.

TINDERS dataset is composed of 48 frontal face classes each represented by visible, NIR (980 nm) at two standoff distances (50 and 106 m), and SWIR at two standoff distances (50 and 106 m) images. At each distance and spectrum, four or five images per class are available - two or three have neutral expression and two have talking expression. A total of 478 images with the resolution  $640 \times 512$  (png format) are available in SWIR band. A total of 238 images with the resolution  $640 \times 512$  (png format) are available in the NIR band. The visible (color) images with the resolution  $640 \times 480$  (jpg format) are collected at a short distance and in two sessions (3 images per session), and all of them have neutral expression, resulting in a total of 288 images. Sample images from the Pre-TINDERS and TINDERS datasets are shown in Figure 3.

The PCSO dataset is composed of color and MWIR images of 1000 subjects. Images are acquired at a short standoff distance of 1.5 m. Each class is represented by two MWIR images collected in two different sessions and one color image. The resolution for all images in both visible and MWIR spectral bands are  $620 \times 480$  in jpg format.

It is important to note that although the original resolution of images in the Pre-TINDERS, TINDERS, and PCSO datasets is varied, we crop and normalize them to be of the same size for each experiment described below. This is done to ensure a fair comparison. Prior to matching, heterogeneous



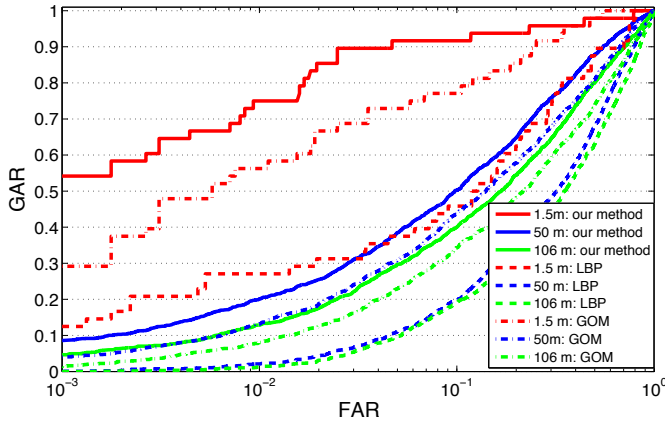
**Fig. 3.** Sample images: (a) visible, (b) SWIR 1.5 m, (c) SWIR 50 m, (d) SWIR 106 m, (e) NIR 50 m, (f) NIR 106 m.

periocular regions are extracted from the heterogeneous images. The right eye region is normalized to the size  $88 \times 72$ .

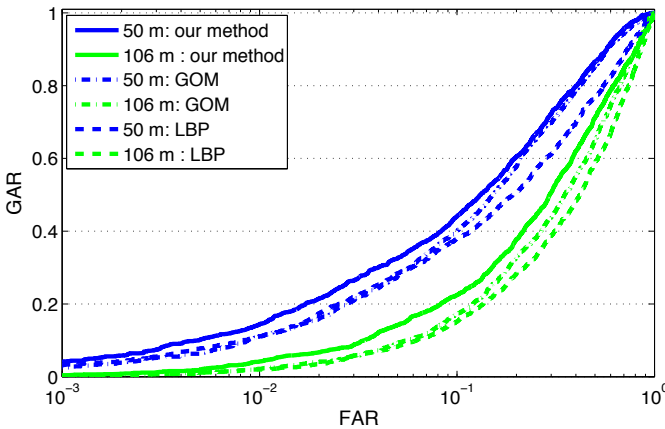
#### 3.2. Matching SWIR and visible

Our first experiment involves matching SWIR periocular regions to visible periocular regions. The heterogeneous images are encoded using three algorithms: (1) Gabor followed by Weber, LBP and GLBP (our algorithm), (2) the original LBP and (3) the GOM-based algorithm [16]. The performance of the LBP is used as a benchmark. The results of matching are displayed as Receiver Operating Characteristic (ROC) curves in Figure 4. Visible light images form the gallery. All SWIR images are involved as the test images. The results are shown for right eye only. The curves are parameterized by three standoff distances. Note that for the case of short standoff distance (pre-TINDERS dataset) the value of Genuine Accept Rate (GAR) is 0.75 at False Accept Rate (FAR) set to 0.01 for our algorithm compared to the GAR of 0.28 and 0.56 at the same FAR for LBP and GOM, respectively. This comparison clearly demonstrates the advantage of using our algorithm over both LBP and GOM for short standoff distances. However, as the standoff distance increases to 50 m and then to 106 m, the difference in matching performance of the three algorithms diminishes too. Performance of all algorithms drops significantly.

Note that the useful information that helps matching heterogeneous images is contained in the gradients of intensity and in their relative distribution, that is, their relative location and their density. Due to long standoff distances the overall quality of SWIR images in TINDERS dataset is reduced. This, in turn, affects the quality of informative gradients and their distribution in SWIR images compared to visible images or SWIR images in Pre-TINDERS dataset. As demonstrated, these low quality gradients can be detected equally well (or



**Fig. 4.** Matching SWIR vs. visible. Short and long standoff distances.



**Fig. 5.** Matching NIR vs. visible. Long standoff distance.

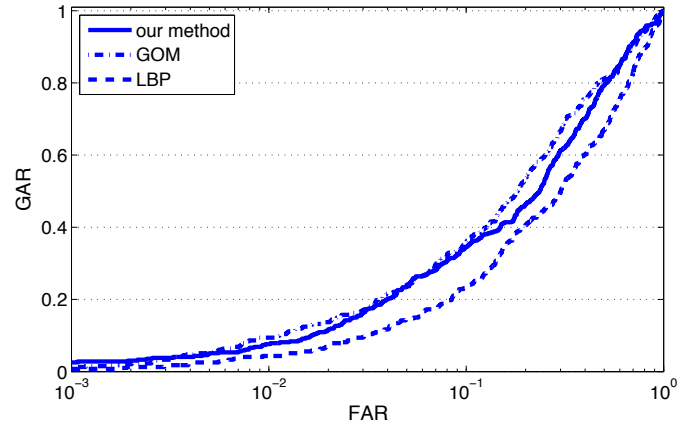
equally poorly) by LBP, GOM and our algorithm.

### 3.3. Matching NIR and visible

The results of matching NIR periocular regions to visible periocular regions are shown in Figure 5. Two standoff distances are tested: 50 m and 106 m. At 50 m, our algorithm reaches GAR value of 0.44 at FAR set to 0.1 while LBP and GOM reach GAR values of 0.37 and 0.40 at the same value of FAR, respectively. When the standoff distance increases to 106 m, our algorithm, LBP and GOM produce GAR values of 0.22, 0.15 and 0.17, respectively, at FAR equal to 0.1. Again, this experiment demonstrates that a small region of face becomes a weak identifier for the case of heterogeneous face matching. Apart from this, NIR images at large standoff distances are extremely noisy due to insufficient illumination (see Fig. 3). This requires taking the quality of the images into account.

### 3.4. Matching MWIR and visible

In our last experiment, we match MWIR periocular regions to visible periocular regions. The results shown in Figure 6 are for 200 periocular classes (right eye only). Color images



**Fig. 6.** Matching MWIR vs. visible. Short standoff distance.

constitute the gallery while MWIR images (two per class) are test images. It is clear that all the LBP, GOM and our algorithms are not designed to deal with such a large spectral gap. MWIR images display distribution of heat in a subject's face and periocular regions, while visible images are formed as patterns characterizing the reflectivity properties of subject's skin. Again, the only common information for heterogeneous matching is in gradient lines and their relative distribution and density. MWIR and visible periocular regions do not contain much of this type of information in common. Nevertheless, we achieved GAR of 0.35, 0.24 and 0.36 at FAR set to 0.1 by applying our algorithm, LBP and GOM respectively.

## 4. SUMMARY

This paper presented the analysis of matching performance of heterogeneous periocular regions. The encoding and matching algorithm have been previously developed in our lab for the purpose of heterogeneous face recognition and adapted to work with smaller heterogeneous regions such as periocular regions. We presented the results of matching SWIR, NIR and MWIR periocular regions to visible periocular regions. Both short (1.5 m) and long (50 m and 106 m) standoff distances were considered. Our algorithm substantially outperformed both LBP and GOM when they are applied to heterogeneous periocular regions collected at short standoff distances (matching SWIR to visible regions). As the standoff distance increases, the matching performance of the heterogeneous periocular regions drops. This drop is attributed to a relatively low quality of imagery at long standoff distances (SWIR vs. visible and NIR vs. visible). When matching MWIR periocular regions to visible regions at a short standoff distance, our algorithm displayed relatively low performance. In this case it is due to the different nature of MWIR (measures the heat of a body) and visible (measures reflected light) imagery.

The results of performance evaluation presented in this work can be used as a baseline for the development of new algorithms for matching heterogeneous periocular regions.

## 5. REFERENCES

- [1] U. Park, A. Ross, and A. K. Jain, "Periocular biometrics in the visible spectrum: a feasibility study," in *Proc. of IEEE 3rd Int. Conf. on Biometrics: Theory, Applications and Systems*, 2009, pp. 1–6.
- [2] D. L. Woodard, S. J. Pundlik, J. R. Lyle, and P. E. Miller, "Periocular region appearance cues for biometric identification," in *Proc. of IEEE Computer Society Conf. on Computer Vision and Pattern Recognition Workshops*, 2010, pp. 162–169.
- [3] S. Bharadwaj, H. S. Bhatt, M. Vatsa, and R. Singh, "Periocular biometrics: When iris recognition fails," in *Proc. of IEEE Int. Conf. on Biometrics: Theory, Applications and Systems*, 2010, pp. 27–29.
- [4] K. P. Hollingsworth, S. S. Darnell, P. E. Miller, D. L. Woodard, K. W. Bowyer, and P. J. Flynn, "Human and machine performance on periocular biometrics under near-infrared light and visible light," *IEEE Trans. on Inf. Forensics and Security*, vol. 7, no. 2, pp. 588–601, 2012.
- [5] U. Park, R. R. Jillela, A. Ross, and A. K. Jain, "Periocular biometrics in the visible spectrum," *IEEE Trans. on Inf. Forensics and Security*, vol. 6, no. 1, pp. 96–106, 2011.
- [6] C. W. Tan and A. Kumar, "Human identification from at-a-distance images by simultaneously exploiting iris and periocular feature," in *Proc. of Int. Conf. on Pattern Recognition*, 2012, pp. 553–556.
- [7] S. Liao, A. K. Jain, and S. Z. Li, "Partial face recognition: Alignment-free approach," *IEEE Trans. on Pattern Analysis and Mach. Int.*, vol. 35, no. 5, pp. 1193–1205, 2013.
- [8] F. Nicolo and N. A. Schmid, "Long range cross-spectral face recognition: Matching swir against visible light images," *IEEE Trans. on Inf. Forensics and Security*, vol. 7, no. 6, pp. 1717–1726, 2012.
- [9] S. Liao, D. Yi, Z. Lei, R. Qin, and S. Z. Li, "Heterogeneous face recognition from local structures of normalized appearance," in *Proc. of IAPR/IEEE Int. Conf. on Biometrics*, 2009, pp. 209–218.
- [10] B. Klare and A. K. Jain, "Heterogeneous face recognition: Matching nir to visible light images," in *Proc. of Int. Conf. on Pattern Recognition*, 2010, pp. 1513–1516.
- [11] N. D. Kalka, T. Bourlai, B. Cukic, and L. Hornak, "Crossspectral face recognition in heterogeneous environments: A case study on matching visible to short-wave infrared imagery," in *Proc. of Int. Joint Conf. on Biometrics*, 2011, pp. 1–8.
- [12] J. Zuo, F. Nicolo, N. A. Schmid, and S. Boothapati, "Encoding, matching and score normalization for cross spectral face recognition: Matching swir versus visible data," in *Proc. of IEEE 5th Int. Conf. on Biometrics: Theory, Applications and Systems*, 2012, pp. 203 – 208.
- [13] B. F. Klare and A. K. Jain, "Heterogeneous face recognition using kernel prototype similarities," *IEEE Trans. on Pattern Analysis and Mach. Int.*, vol. 35, no. 6, pp. 1410–1422, 2013.
- [14] T. Bourlai, N. Kalka, A. Ross, B. Cukic, and L. Hornak, "Cross-spectral face verification in the short wave infrared (swir) band," in *Proc. of Int. Conf. on Pattern Recognition*, 2010, pp. 1343 – 1347.
- [15] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, 2005, pp. 886 – 893.
- [16] Z. Chai, Z. Sun, H. Mendez-Vazquez, R. He, and T. Tan, "Gabor ordinal measures for face recognition," *IEEE Trans. on Inf. Forensics and Security*, vol. 9, no. 1, pp. 14–26, 2014.
- [17] J. Xu, M. Cha, J. L. Heyman, S. Venugopalan, R. Abiantun, and M. Savvides, "Robust local binary pattern feature sets for periocular biometric identification," in *Proc. of IEEE Int. Conf. on Biometrics: Theory, Applications and Systems*, 2010, pp. 1–8.
- [18] J. Chen, S. Shan, C. He, G. Zhao, M. Pietikainen, X. Chen, and W. Gao, "Wld: A robust local image descriptor," *IEEE Trans. on Pattern Analysis and Mach. Int.*, vol. 32, no. 9, pp. 1705–1720, 2010.
- [19] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. on Inf. Forensics and Security*, vol. 24, no. 7, pp. 971–987, 2002.
- [20] Y. Guo and Z. Xu, "Local gabor phase difference pattern for face recognition," in *Proc. of Int. Conf. on Pattern Recognition*, 2008, pp. 1–4.
- [21] S. Kullback and R. A. Leibler, "On information and sufficiency," *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, 1951.