NFRAD: Near-Infrared Face Recognition at a Distance

Hyunju Maeng[a], Hyun-Cheol Choi[a], Unsang Park[b], Seong-Whan Lee[a] and Anil K. Jain[a,b]

[a]Dept. of Brain and Cognitive Eng. Korea Univ., Seoul, Korea
[b]Dept. of Comp. Sci. & Eng. Michigan State Univ., E. Lansing, MI, USA 48824

{hjmaeng, hcchoi}@korea.ac.kr, parkunsa@cse.msu.edu, swlee@image.korea.ac.kr , jain@cse.msu.edu

Abstract

Face recognition at a distance is gaining wide attention in order to augment the surveillance systems with face recognition capability. However, face recognition at a distance in nighttime has not yet received adequate attention considering the increased security threats at nighttime. We introduce a new face image database, called Near-Infrared Face Recognition at a Distance Database (NFRAD-DB). Images in NFRAD-DB are collected at a distance of up to 60 meters with 50 different subjects using a near-infrared camera, a telescope, and near-infrared illuminator. We provide face recognition performance using FaceVACS, DoG-SIFT, and DoG-MLBP representations. The face recognition test consisted of NIR images of these 50 subjects at 60 meters as probe and visible images at 1 meter with additional mug shot images of 10,000 subjects as gallery. Rank-1 identification accuracy of 28 percent was achieved from the proposed method compared to 18 percent rank-1 accuracy of a state of the art face recognition system, FaceVACS. These recognition results are encouraging given this challenging matching problem due to the illumination pattern and insufficient brightness in NFRAD images.

1. Introduction

Face recognition is one of the most popular research areas in biometrics for its wide applications in access control, government ID cards, forensics, and surveillance systems [12]. Due to the non-contact and remote sensing capabilities of the face biometric, it is regarded as the only viable biometric trait to be used in surveillance systems for subject identification. Integrating face recognition capability in surveillance systems will significantly improve the security of the monitoring area.

Among the many challenges in face related applications, face recognition at a distance is regarded as rather unique and the most imminent problem for subject identification in surveillance applications. Typically, the large intra-class variations due to pose, lighting, and expression
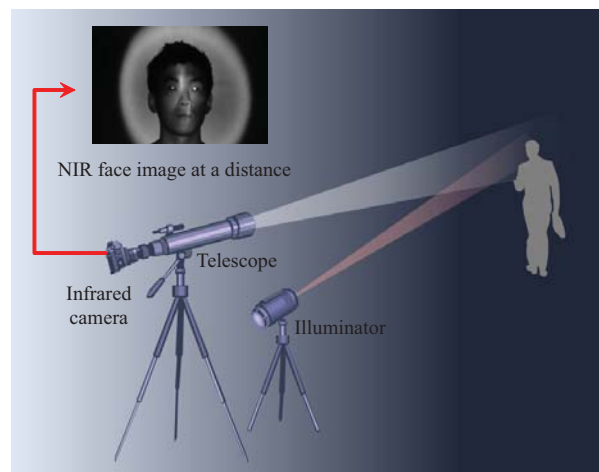

Figure 1. Schematic of the Near-infrared Face Recognition at a Distance (NFRAD).

variations make face recognition a challenging problem in all the applications using face biometric. Even though all the aforementioned problems exist in surveillance scenarios, the low resolution problem due to the large standoff between the camera and subject is considered the most fundamental problem in face recognition at a distance (i.e., in surveillance environments). With progress in face recognition techniques over the past decades, the performance of face recognition has significantly improved, given frontal pose and normal lighting conditions. However, with low resolution images, the face recognition performance is still poor even if the face image appears as frontal under normal lighting conditions. Therefore, it is essential to address the low resolution problem first, and then take advantage of the advanced face recognition techniques that have been developed for face image under reasonable conditions (i.e., near-frontal pose and uniform lighting conditions).

Considering the increased security threats in nighttime, face recognition techniques in surveillance systems need to be coupled with the capability of face recognition at nighttime. Use of a search light would be one option to enable

1

face recognition at night, but it will violate the requirement of covert surveillance. Therefore, the use of infrared light is considered as a promising method of face recognition at night in surveillance applications. The advantage of the use of near-infrared light for face recognition in surveillance at nighttime has the following advantages [10]: (1) near-infrared illuminator is not visible to the human eye and keeps the surveillance operation covert, (2) NIR images captured are not affected by ambient temperature compared to thermal images, and (3) the additional cost involved in using the near-infrared illuminator is relatively low. Therefore, face recognition in near-infrared spectrum has been investigated actively by many research groups. However, near-infrared face recognition has been mostly studied at relatively close distances [5]; the use of near-infrared light source at a long distance has not yet received much attention. Fig. 1 shows the schematic of the proposed NFRAD system using a near-infrared camera, a telescope, and a near-infrared illuminator.

Table 1. Summary of image databases for FRAD

Database	IRIS-LDHM [14]	FRAD [10]
No. of subjects	Indoor: 55; Outdoor: 48	Outdoor: 97
Distance	Indoor: 10 to 20 meters Outdoor: 50 to 300 meters	15 to 80 meters
Resolution	640×480	4752×3168
Format	Still image (BMP) and video sequence (AVI)	Still image

There are a number of closeup face image databases publicly available for performance evaluation of face recognition systems. AR [7], BANCA [2], FERET [9] and FRGC [3] are some of the representative face image databases collected at a close distance and with constant zoom. A few research groups have collected long distance face image databases for face recognition at a distance (FRAD) [13]. Yao et al. [14] collected a face video database, IRIS-LDHM, at a long distance with high magnifications, both in indoor and outdoor settings. Rara et al. [10] collected FRAD database in both indoor and outdoor environments. They also proposed a FRAD system as far as 1,000 meters in challenging imaging conditions. Table 1 summarizes two representative face image databases collected at a long distance. While these FRAD databases include face images at a distance, they do not include near-infrared long distance face images which are captured at night. To address the problem of face recognition at a distance in nighttime environment, we have collected a FRAD database in nighttime using near-infrared spectrum, called

NFRAD-DB. To the best of our knowledge, this is the first attempt to build a face image database for the study of near-infrared face recognition at a distance. The rest of this paper is organized as follows. Section 2 explains the sensor configuration, section 3 describes the data acquisition setup, and section 4 describes the face recognition experiments using the collected database, and section 5 concludes this paper.

2. Sensor Configuration

For collecting FRAD database, additional devices should be appended to the basic image acquisition system. Yao et al. [14] set up an imaging system for outdoor sequence collection using a telescope and DSLR camera. To acquire long distance images, the telescope is used with an afocal connection and eyepieces. Rara et al. [10] built their own passive stereo acquisition setup to collect the FRAD database using two CCD cameras mounted on pan-tilt units on two separate tripods.

To collect the NFRAD-DB, we also built our own imaging system. NFRAD imaging system was built with the Sky-Watcher SK MAK180 telescope (focal length of 2700 mm) coupled with a Nikon D80 DSLR camera for high resolution image capture. Nikon DSLR is connected to the eyepiece of the telescope. To connect these two devices, we used a DSLR camera-eyepiece connector ring. Although there are many cameras and camcorders which have a night vision functionality (by using their integrated compact near-infrared illuminators), these illuminators are not sufficiently strong to be used at a long distance. Moreover, most of the commercial cameras and even cameras with night vision features have built in IR (and UV) filters in the sensor. We removed the UV filter from the DSLR to prevent it from blocking NIR rays from the illuminator and used this modified DSLR for data collection.

The near-infrared light is also essential for collecting nighttime images, but the light should be sufficiently strong to illuminate subjects at a long distance. Fig. 2 shows a subject's face images with and without the presence of near-infrared light. For acquiring nighttime face images at vari-

(a)	(b)

Figure 2. Nighttime images acquired using a near-infrared illuminator: (a) illuminating the face of the subject and (b) pointing the near-infrared light next to the subject. This shows that the face is visible only by the near-infrared illuminator not by any other ambient lighting source.
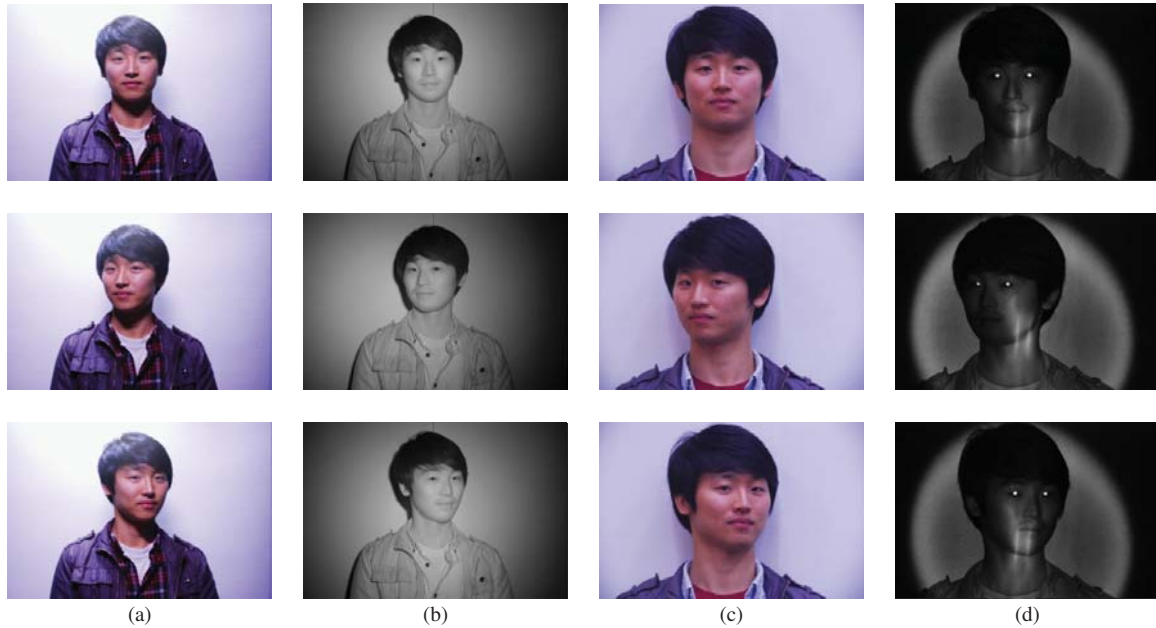
Figure 4. Example images in NFRAD-DB: the data for each subject consists of frontal, left, and right facial images captured at (a) 1 meter during the daytime, (b) 1 meter during the nighttime, (c) 60 meters during the daytime, and (d) 60 meters during the nighttime.
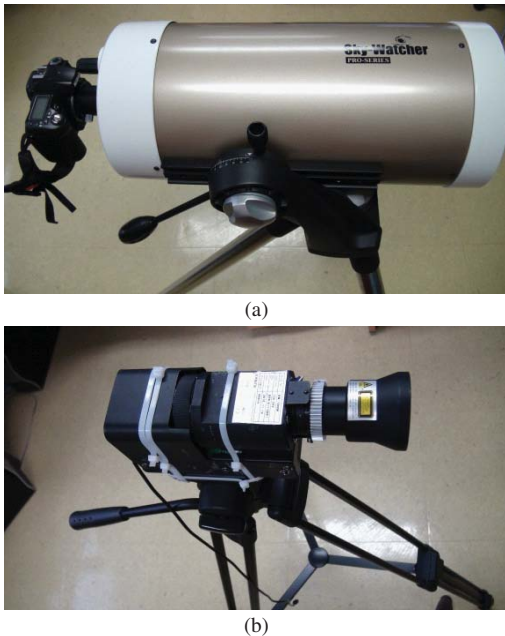

Figure 3. Sensors used in the NFRAD-DB collection: (a) Sky-Watcher telescope connected to a Nikon camera and (b) near-infrared illuminator.

ous distances, two near-infrared illuminators are used. One is LANICS illuminator which can emit near-infrared rays up to a distance of about 1,000 meters, so it is sufficient to put light on a face at a long distance; this is placed next to the telescope to illuminate in the same direction as the telescope. The range of wavelengths of the near-infrared ray is from 810 nm to 960 nm (peak at 868 nm). Another near-infrared illuminator is 24 IR LED built in a CCTV (GUARDCAM, SS81 IR24) which is appropriate for short distances. Fig. 3 shows the imaging system used in the NFRAD database collection.

3. Data Acquisition

The NFRAD database contains visible images and NIR images at a long distance (60 meters)[1] and a short distance (1 meter). We captured frontal, left, and right facial images of 50 subjects under the following conditions: (1) 1 meter-day, (2) 1 meter-night, (3) 60 meter-day, and (4) 60 meter-night. All the images were collected in an indoor environment on the Korea University campus; the subjects were university students. Table 2 describes the database specification. The total size of the database is 1.25 GB.

Table 2. Summary of the NFRAD-DB

No. of subjects	50
Distance	1 meter and 60 meters
Lighting condition	Fluorescent light (Day), Infrared light (Night)
Resolution	3872×2592
Format	Still image (JPEG)

---

[1]This distance was limited by the size of the room that we had access to on the Korea University Campus.

3.1. Short-distance Database

Short-distance (1 meter) database includes 6 still images per subject and contains frontal face, left face, and right face during the day (visible light) and at night (near-infrared light). Nighttime images are collected at 1 meter standoff in a dark room. Day images are collected under a fluorescent light by using the adapted DSLR camera with SIGMA 28-80 mm F2.8-4 DS zoom lens. The night images are collected (with no visible light) using the adapted DSLR camera without the UVIR filter and a near-infrared illuminator of 24 IR LED which is appropriate for short distance capture. Figs. 4(a) and 4(b) show examples of short-distance images. Short-distance face database has 50 subjects (74 percent males and 26 percent females). The image resolution is 3872×2592 pixels and all the images are stored in the JPEG format.

3.2. Long-distance Database

For long-distance database collection during the day (visible light) and at night (near-infrared light), the proposed NFRAD imaging system is used. Long-distance database also includes 6 still images per subject at 60 meters and contains frontal, left, and right facial images during the day (visible light) and at night (near-infrared light). Still images are also collected in indoor environments, with no visible light. Day images are collected under fluorescent light
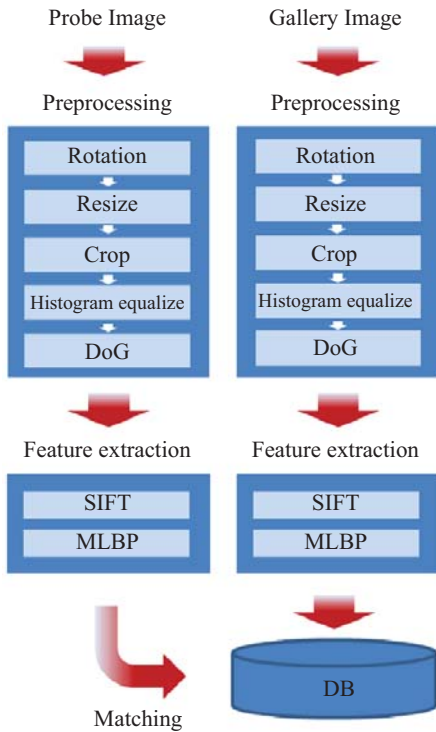

Figure 5. Schematic of the overall face recognition process.

by Sky Watcher telescope and the modified DSLR camera, and night images are collected using near-infrared light provided by an LANICS illuminator which is appropriate for long distance. Fig. 4(c) and 4(d) show example images in the long-distance database. Long-distance database has 50 subjects which are the same as those in the short-distance database. The image resolution is 3872×2592 pixels and all the images are stored in the JPEG format.

4. Face Recognition

Long-distance night images in NFRAD-DB have unique illumination patterns (Fig. 4(d)). The pattern is due to the shape of the illuminator's light source array, and it poses difficulty in face recognition. To compensate for this, we applied preprocessing to the captured face images to reduce the effect of the illumination pattern and enhance the image quality. For representing a face, we extracted features using SIFT and MLBP local descriptors.

In order to evaluate the performance of face recognition on NFRAD-DB, we tested various methods of face recognition: (1) a COTS matcher, FaceVACS [1], (2) DoG-SIFT, and (3) DoG-MLBP. The face recognition experiments used short distance (1 meter) day images (50 images of 50 subjects) as gallery. Additional 10,000 images of 10,000 subjects from the MORPH database [11] were also added to the gallery to increase the complexity of the face identification task. Even though the images in MORPH are different from the gallery in terms of pose, overall face size, ethnicity, etc., it is the only large public domain face image database available. We used 50 frontal NIR images at a short-distance and 50 frontal visible and NIR facial images at a long-distance of 50 subjects as probe images. Fig. 5 shows the overall schematic of the face recognition process.
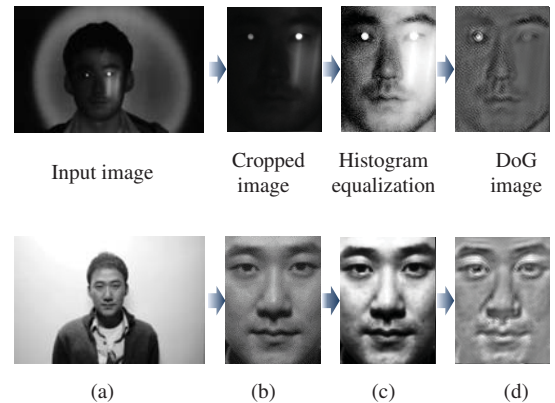

Figure 6. Schematic of the preprocessing for two different face images: (a) input image, (b) cropped image using eye position, (c) histogram equalized image, and (d) DoG image (filter size=10, σ=0.5). Images in the first and second row correspond to long-distance night (probe) and short-distance day image (gallery), respectively.

4.1. Preprocessing

Preprocessing involves face detection, segmentation, histogram equalization, and Difference of Gaussian (DoG) filtering. The face is detected based on eye locations in the image, and then the image is rotated according to the left and right eye locations which are aligned with the horizontal direction. Based on the distance between the eyes, the images are resized and cropped to a size of 90×120 pixels. With normalized images, histogram equalization is applied to spread out the most frequent intensity values. In particular, this method is useful for images with background and foreground that are either both bright or both dark. Therefore, we can mitigate the presence of illumination pattern in long-distance night images. We also performed DoG filtering on the input face images (Fig. 6). We generated various DoG filtered images by changing the window size and standard deviation (σ =0.5~0.75) of the Gaussian kernel.

4.2. Feature extraction

4.2.1 SIFT

Scale Invariant Feature Transform (SIFT) [6] is one of the most popular local feature representation methods. It has been successfully used in feature vector construction from face images. [4] It consists of two stages: (1) key point extraction, and (2) descriptor calculation in a local neighborhood at each key point. In our algorithm, we extract 42 key points from the normalized face image by using a grid of landmarks within a preset radius. The interval of grid is 12 pixels, and the SIFT descriptors are constructed in 24×24 windows for calculating the gradient magnitude and orientation of the SIFT descriptor. The gradient magnitude at a pixel is calculated according to its orientation and summarized into an 8-bin orientation histogram. Fig. 7 shows the schematic of the SIFT descriptor construction for feature extraction. Through this process, we obtain a 128-dimensional feature vector for each key point. The final feature vector is obtained by concatenating the feature vector at each key point and has a dimensionality of 5,376. Euclidean distance is used to calculate the similarity between the probe and gallery images both using the SIFT and MLBP descriptors.

4.2.2 MLBP

Multi-scale Local Binary Pattern (MLBP) [8] is a popular method for extracting feature vectors from face images. MLBP descriptor is the extended version of the LBP descriptor by using multiple radii and neighborhood size. The process of constructing MLBP descriptor is similar to that of SIFT descriptor in using sub-windows to calculate the local histograms. In MLBP feature extraction, we used two different radii, $R = 1, 2$, and three different neighborhood

sizes, $P = 4, 8$, and 12. Fig. 8 shows the parameter sets that were used in constructing MLBP descriptors.

4.3. Experiment results

In this section, we describe face recognition results for three different subsets of the NFRAD-DB: (1) 1 meter-night images, (2) 60 meter-day images, and (3) 60 meter-night images. We first compare rank-1 recognition rates on these three sub-databases using a commercial state of the art matcher, FaceVACS. We used 10,050 gallery images, including 10,000 additional images from the MORPH database, and used 50 probe images in each experiment. All gallery and probe images are in frontal facial pose. The face recognition performances of FaceVACS on 1 meter-night images and 60 meter-day images are perfect (100% accuracy). However, its performance on 60 meter-night probe images shows extremely low accuracy (0%) due to the illumination pattern which makes it difficult to detect the eyes as well as recognize the faces.

In order to focus on the recognition capability, we labeled the eye positions in all the probe images (60 meter-night images) manually. The recognition experiment with manually detected eyes was repeated with 10,050 gallery images. Note that FaceVACS was also provided manually detected eye locations.

Besides FaceVACS, we used two different methods of face recognition on 60 meter-night images: (1) DoG-SIFT and (2) DoG-MLBP. Fig. 9 shows the face identification results with all the three different face recognition meth-
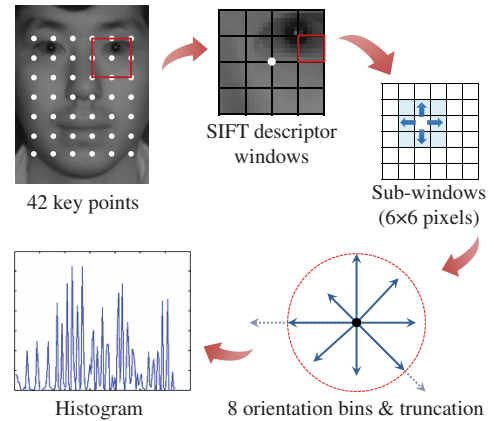

Figure 7. Schematic of the SIFT descriptor construction.
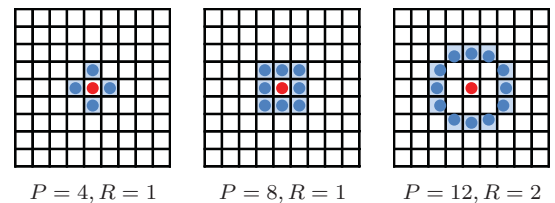

$P = 4, R = 1$ $P = 8, R = 1$ $P = 12, R = 2$

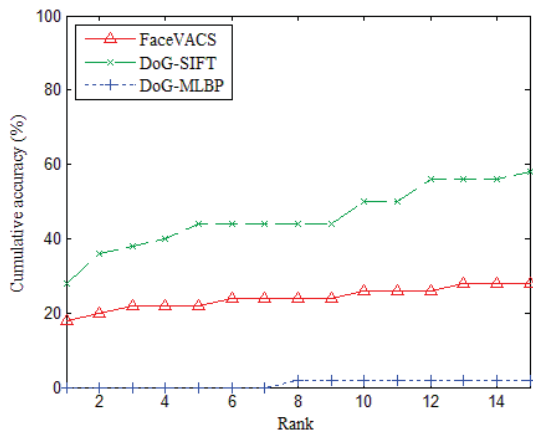Figure 8. Schematic of the MLBP descriptor construction.

Figure 9. CMC curves showing the face recognition performance of 50 night face images (probe) at a distance (60 meters) with 10,050 closeup visible gallery images.
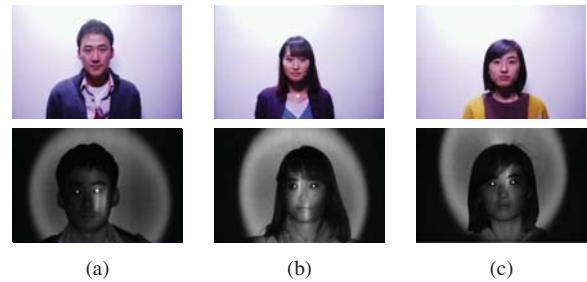

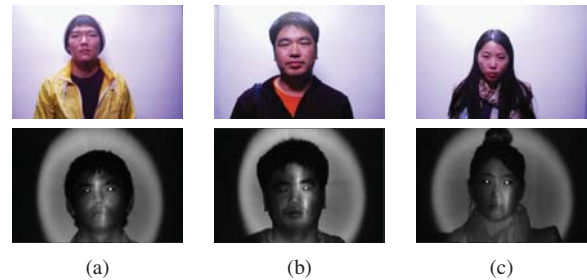Figure 10. Example gallery (top row) and probe (bottom row) images successfully matched at rank-1 using DoG-SIFT.

Figure 11. Example gallery (top row) and probe (bottom row) images matched at rank- (a) 5956, (b) 2869, and (c) 1435 using DoG-SIFT.

ods. Note that while FaceVACS shows only an 18% rank-1 accuracy, DoG-SIFT shows a higher rank-1 accuracy of 28%. On the contrary, DoG-MLBP performs miserably (0%) even with manually detected eye locations. Given a 6×6 sub-window, SIFT extracts 8 features corresponding to 8-orientation bins, while MLBP extracts 74 uniform LBP features using 4 and 8 neighborhood with R=1. The resulting 74-dimensional MLBP features are more sensitive to image translation and rotation compared to the 8-dimensional SIFT features.

4.4. Conclusions and Future Work

We have collected a unique face database, called NFRAD-DB, that includes normal visible still images as well as NIR images at a long distance. We have evaluated a state of the art face recognition engine, FaceVACS and two additional face recognition algorithms (DoG-SIFT and DoG-MLBP) on this database. For the 60 m night image database, the results of FaceVACS showed an 18% rank-1 accuracy using 10,050 gallery images. Matching results using the DoG-SIFT showed a higher rank-1 accuracy than that of FaceVACS. Although there are some challenges related to illumination pattern problem in NIR images at a long distance, this database can be used to evaluate face recognition for surveillance systems operating in nighttime.

We are in the process of collecting additional images to increase the size of the database to include up to 100 subjects. Our goal is to collect face images as far as 100 meters under visible light as well as with near-infrared illuminator. We also plan to reduce the challenge of near-infrared illumination problem by using a better quality illuminator which has a uniform brightness across the face without forming a noisy pattern. Additional research topics include automatic

eye detection under various illumination conditions as well as invariant feature extraction methods to mitigate the image distortion effects observed at long distance.

References

[1] FaceVACS Software Developer Kit, http://www.cognitec-systems.de/.

[2] E. Bailly-Bailliere, S. Bengio, F. Bimbot, M. Hamouz, J. Kittler, J. Mariethoz, J. Matas, K. Messer, V. Popovici, F. Poree, B. Ruiz, and J. Thiran. The BANCA database and evaluation protocol. In *Proc. Audio and Video-Based Biometric Person Authentication*, volume 2688, pages 625–638. 2003.

[3] J. R. Beveridge, G. H. Givens, P. J. Phillips, and B. A. Draper. Factors that influence algorithm performance in the face recognition grand challenge. *Computer Vision and Image Understanding*, 113(6):750–762, 2009.

[4] B. Klare and A. K. Jain. Heterogeneous face recognition using kernel prototype similarities. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2011 (pending).

[5] S. Z. Li, C. RuFeng, L. ShengCai, and Z. Lun. Illumination invariant face recognition using near-infrared images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 29(4):627–639, 2007.

[6] D. G. Lowe. Distinctive image features from scale invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[7] A. Martínez and R. Benavente. The AR face database. Technical Report 24, Computer Vision Center, 1998.

[8] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24(7):971–987, 2002.

[9] P. Phillips, H. Wechsler, J. Huang, and P. Rauss. The FERET database and evaluation procedure for face-recognition algorithms. *Image and Vision Computing*, 16(5):295–306, 1998.

[10] H. Rara, S. Elhabian, A. Ali, T. Gault, M. Miller, T. Starr, and A. Farag. A framework for long distance face recognition using dense- and sparse-stereo reconstruction. In *Proc. Int. Symposium on Advances in Visual Computing: Part I*, pages 774–783, 2009.

[11] K. Ricanek and T. Tesafaye. MORPH: A longitudinal image database of normal adult age-progression. In *Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition*, pages 341–345, 2006.

[12] S. Z. Li and A. K. Jain (Eds.). *Handbook of Face Recognition, Second Edition*. Springer-Verlag, 2011.

[13] M. Tistarelli, S. Z. Li, and R. Chellappa (Eds.). *Handbook of Remote Biometrics: for Surveillance and Security*. Springer, 2009.

[14] Y. Yao, B. R. Abidi, N. D. Kalka, N. A. Schmid, and M. A. Abidi. Improving long range and high magnification face recognition: Database acquisition, evaluation, and enhancement. *Computer Vision and Image Understanding*, 111(2):111–125, 2008.