Heterogeneous Face Matching: NIR images to VIS Images

Sandhya R.Waddhavane
M.E Student
Department of Computer Engineering
KKWIEER, Nashik, India.
Savitribai Phule Pune University,Pune

S.M.Kamalapur, PhD
Associate Professor
Department of Computer Engineering
KKWIEER, Nashik, India.
Savitribai Phule Pune University,Pune

ABSTRACT
Heterogeneous (cross spectral) face matching is very important in many of the security applications, especially at night time face recognition where query images are near infrared face images and gallery images are generally visible light images. At night time environment near infrared cameras are used for imaging and images can be captured at various distances as object is not at fix position. As distance increases the quality of face image get degrade and it becomes difficult to match the query near infrared face image with the gallery images. The aim of proposed work is to implement an efficient face matching technique that resolves the problem of cross distance together with cross spectral face matching. Learning based image restoration is an approach to deal with this problem. In this the face images at long distances are restored first and then restored face images are matched with VIS (database) images. The proposed work improves the face matching performance by normalizing near infrared and visible light face images using Difference of Gaussian filter and extracted HOG features for heterogeneous face matching.

Keywords
Acquisition system, Cross spectral, Cross distance, Heterogeneous face, Image restoration, Near infrared (NIR) face image, Visible light (VIS) face image

1. INTRODUCTION
Now a day, heterogeneous face matching becomes an important and essential. Because in many of the applications such as E-passport, video surveillance, criminal investigation, photo based identification, and law enforcement; enrollment (probe) face images and gallery face images are of different modalities. The face images are captured using different acquisition system such as infrared images are captured using infrared devices while VIS images are captured using visible light cameras. As images are captured using different acquisition system, image formation characteristics differs such as quality of camera, sensors used in camera. All such kinds of face images are called heterogeneous face images. The heterogeneous (NIR-VIS) face images are shown in figure 1.

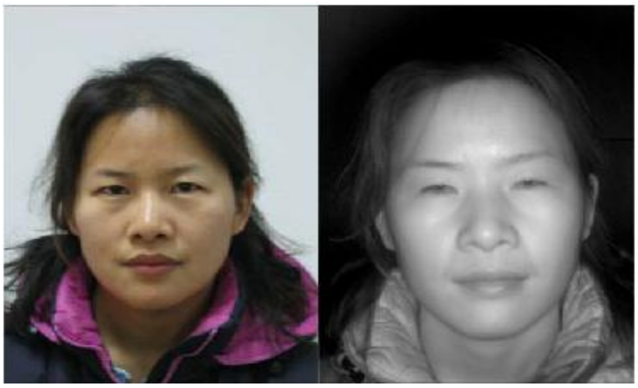

Fig 1. Example of VIS-NIR face image

The visible light cameras are able to capture face images at long distance but it is expensive, only works in day time and also not suitable in secret operation. The use of near infrared devices is an option for nighttime scenario because it has some advantages. [8] like i) Near infrared face images are generally not affected by surrounding temperature ii) Near infrared light is invisible to human iii) Near infrared illumination can easily penetrate glasses and iv) NIR illuminators are cheap as compared to other sensors. NIR-VIS face matching is required in surveillance system, as at night time environment NIR illuminator is used in cameras for imaging and database contains VIS images so it becomes heterogeneous matching.

Most of the heterogeneous face matching techniques available for matching face images which are captured at same distance. If the query NIR face images are captured at various and long distances, the quality of the face image and ultimately performance of the system get degrade. Learning based image restoration is an approach to improve the quality of degraded face image. In this work, trying to restore the long distanced NIR face image and then match restored NIR image with VIS face images contain in database. For better performance here normalizing NIR and VIS face images using Difference of Guassian (DoG) filter and HOG (Histogram Oriented Gradient) descriptor for feature extraction.

The remainder of paper is organized as follows: Section 2 discussed the related work done and its shortcoming. An overview of the proposed scheme is given in section 3. Section

4 discuss the experimental results of the system. Finally, conclude the paper.

2. LITERATURE SURVEY
All the existing heterogeneous face matching work is grouped into following three different categories: (1) Projection based (2) Invariant feature based (3) Synthesis based.

In first category, images from both modalities are projected to a common space. To learn more distinctive features of NIR images and VIS images, Canonical Correlation Analysis (CCA) is used by Yi et al. [3]. For better classification, CCA learning is performed between features instead of performing between images. The main drawback of the algorithms proposed in [3] is over-fitting. To overcome it Liao et al. [13] proposed a new method. In this approach, images are normalized using Difference of Guassian (DoG) filter and then Multi-scale Block LBP (MB-LBP) features are extracted. To improve the efficiency and reduce the time complexity, graph embedding and spectral regression based techniques combined with regularization techniques are used in [16]

In second category, used invariant features of both the modalities that are robust in illumination variation conditions. To reduce the visual appearance difference between VIS and NIR face images Goswami et al. [9] used some effective preprocessing techniques includes Difference of Gaussian (DoG) filtering, gamma correction, and contrast equalization. Then Local Binary Pattern (LBP) descriptor is used to extract feature combine with linear discriminant analysis (LDA).To address heterogeneous face matching problem Huang et al. [7] extract three different modality-invariant features namely, sparse coefficients (SC), least square coefficients (LSC), and quantized distance vector (QDV).

To maximize the correlation between encoded NIR and VIS face images Gong and Zheng [4] proposed a new feature descriptor. As it reduces within class variation, offers better performance than Local B inary Pattern (LBP) and Multiscale Local Binary Pattern (MLBP). For better correlation Lei et al. [3] proposed learning-based face descriptor also designed a new hyperplane based encoding method to encode infrared and optical face images. To avoid the over-fitting, face matching is done in two steps based on local feature based discriminant analysis (LFDA).

In third category, synthesize one modality face image based on other modality and the synthesized image can then be used directly for homogeneous face matching. Based on analysis-by-synthesis framework Wang et al. [15] proposed face mapping method, called face analogy. Further improved synthesis of VIS image using locally-linear embedding (LLE) by Chen et al. [11], which is based on cross-domain dictionary learning. VIS images can be synthesized patch-by-patch by finding the best matching patch for each patch of the input NIR image. Xiong et al [6] presented a new synthesis based method for VIS-NIR face mapping in which, a 3D model is used to perform pose rectification. Most recently, the image restoration method based on locally-linear embedding (LLE) is proposed by Kang et.al [2].This method solved the problem of cross-distance as well as cross-spectral face matching.

3. PROPOSED WORK

This section describes overall proposed system of the NIR-VIS face matching.

3.1 Process block Diagram

The overall system architecture of the system is shown in the figure 2. The system input is various distanced NIR image and the output will be matching VIS images that would be various distanced or same distanced. Function of every block of system is given below:

3.1.1 Preprocessing:

Due to different image formation characteristics, both NIR and VIS differs in appearance. Additionally, NIR images captured at various distances hence there is quality degradation. To tackle with these problems photometric and geometric normalization is applied on both types of images. All the face images are rotated by some angle, cropped and scaled to the same size. Here used Difference of Gaussian (DoG) filter to decrease illumination variations, to remove noise and

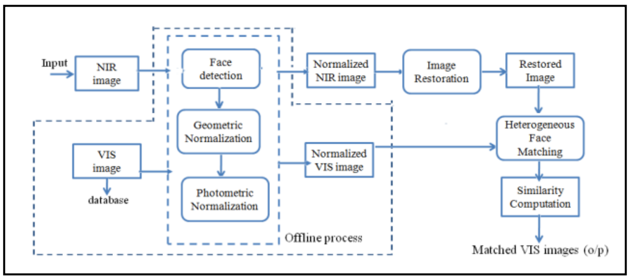

Fig 2. System Block diagram

high frequency details. It is given as:

$$D(\boldsymbol{x_1}, \boldsymbol{y_1} \mid \sigma_0, \sigma_1) = G(\boldsymbol{x_1}, \boldsymbol{y_1}, \sigma_0) - G(\boldsymbol{x_1}, \boldsymbol{y_1}, \sigma_1))* I(\boldsymbol{x_1}, \boldsymbol{y_1})$$

Where

$$G(x_1, y_1, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x_1^2+y_1^2)/2\sigma^2}$$

σ_0 and σ_1 are the width of two Gaussian filters ($\sigma_0 < \sigma_1$)

3.1.2 Image Restoration:

For the quality improvement of the degraded NIR images, a learning based image restoration method is used. It is based on manifold learning theory, this paper consist of two manifolds one for low quality NIR face images (M^L) and other for high quality NIR face images (M^H). It is basically carried out in two steps:

1) Dictionary building of patches
2) Image restoration

3.1.2.1 Dictionary building:

Given training data consist of pairs of high quality and low quality NIR face images.

i) For each image pair first randomly sample n pairs of corresponding patches.
ii) The set of high and low quality patches constitute two dictionaries of patches respectively.
iii) These dictionaries will be used for learning mapping between high quality and low quality patches.
iv) Both high and low quality patches are group into K different clusters which reduces the computational cost.

3.1.2.2 *Image restoration*

The input to the system is low quality NIR face image. In this first NIR image is sampled into patches similar to the training data. Restoration is carried out as follow:

- Let $\{P_i^L$, j=1,2,......N$\}$ are low quality patches. So for each patch P_i^L , find its closest cluster ($C_{\hat{k}}^L$) by comparing with all cluster means and it is given as

$$\hat{k} = \arg min \,||\mu_k^L - P_j^L||_{L^2}$$

Where,

μ_k^L is mean of cluster

- Compute reconstruction weights $\{\omega_t$, t=1,2,.....T$\}$

by using its neighbouring patches. Here set T=5. An embedding of P_j^L in low quality manifold (M^L) given as

$$\varepsilon_j = ||p_j^L - \sum_{p_t^L \in N_j^L} \omega_t \; P_t^L||_L^2$$

Where,

N_j^L is set of neighbouring patches in M^L

- The restored patch is computed from the corresponding T neighbouring patches in M^H and it is given as:

$$P_j^H = \sum_{P_t^H \in N_j^H} \omega_{t\; P_t^H}$$

Where,

N_i^H is set of neighbouring patches in M^H correspond to N_j^L

- The modified restoration is given as

$$P_j^H = \frac{1}{2}\left(\mu_{\mathrm{k}}^H\right) + \frac{1}{2}\left(\sum_{P_t^H \in N_j^H} \omega_t \; P_t^H\right)$$

- After restoration of all patches, to get the final restored face images; the overlapping patches are averaged.

3.1.3 *Heterogeneous Face Matching*

After restoration, the restored NIR face image matched with normalized VIS face image. Histogram Oriented Gradient (HOG) descriptors are used extract the features. The images of size 200×250 divided into 32×32 overlapping block, results in total 154 blocks. Each block consist of 2×2 cells with size 8×8 and using 8 orientation bins results into 4928 dimensions. The similarity between VIS and NIR face image is computed using Euclidean distance..

4. EXPERIMENTAL SETUP

All Experimentation is performed on Intel core i3 Processor and 3GB RAM. The operating system is windows 7(32 bit) with visual studio 10.

4.1 Dataset

The proposed NIR-VIS face matching system is evaluated using the Long Distance Heterogeneous Face (LDHF) dataset [8]. This dataset is challenging because it contains both NIR and VIS images of 100 subjects at various distances (1m, 60m, 100m, and 150m).This dataset contains indoor as well as outdoor images. From this we have applied restoration for 60m, 100m, and 150m NIR images and gallery data includes 400 VIS images.

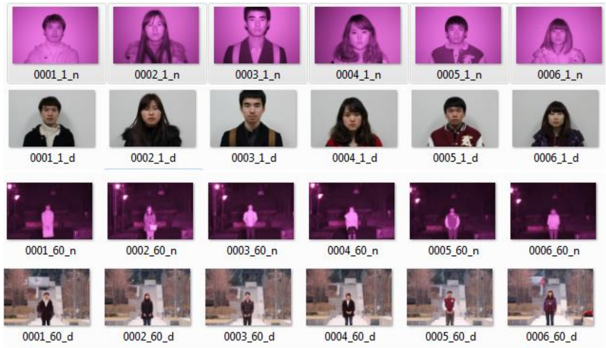

Fig 3. Images from LDHF dataset (1m NIR-first row,1m VIS-second row, 60m NIR-third row,60m VIS-fourth row)

4.2 Performance measure

The results of face matching are computed using identification rate and it is given as:

Identification rate: It is ratio of correctly identified faces to the total no of tested face images.

4.3 Results

1)Face detection:

As images are captured at various distance there is extra noise, so need to detect face from image for face recognition.First step is to take any NIR image and then detect face from this image Figure 4 demonstrates the face detection.

2)Normalization:

To improve the face matching performance and accuracy photometric normalization is applied on the both NIR and VIS face images.Here we have used difference of guassian filtering to reduce the illumination variations and noise. Figure 8 demonstrates the face normalization using DoG filter.

3)Restoration:

If the NIR images are captured at long distance then we need to restore that NIR face image for better accuracy. We have restored face images which are captured at 60m,100m,and 150m standoff distance.Using restoration process query images are matched to correct gallery VIS images. I have used 80 persons for dictionary building and 20 persons for restoration out of 100 persons. Figure 6 shows some examples of 150m and 100m NIR face image restoration and their corresponding true matched VIS images.

4)Feature extraction:

A fter restoration the NIR image is ready for feature extraction. Here, Histogram Oriented Gradient (HOG) features are extracted from both restored NIR face image and VIS images and then computed similarity between these two using eucliden distance.Figure 7 demonstrates the feature extraction,similarity computation using euclidean distance and matching VIS images for given query NIR image

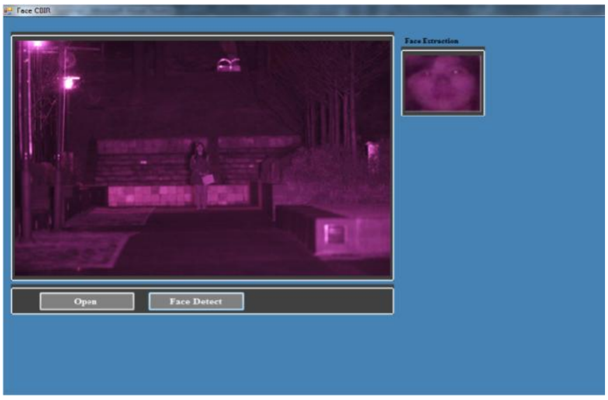

Fig 4.Face detection

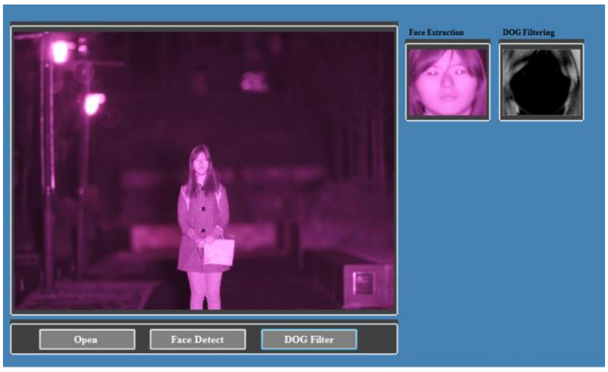

Fig 5. DoG filtering

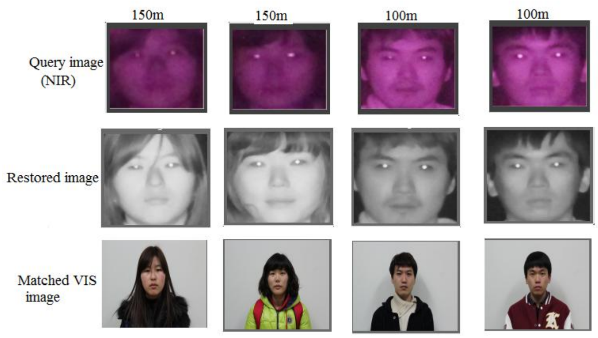

Fig 6. Examples of restoration and matching result (first row-input NIR face image,second row-restored NIR image ,third row-matched VIS image)

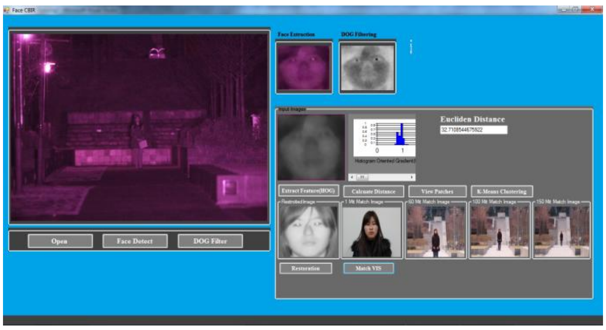

Fig 7. Feature extraction, similarity computation and matching VIS images

4.4 Performance Analysis

The Performance of a system can be analyzed using the CMC curve and is calculated using identification rate for various standoff distances. We have tested system for 60m, 100m, and 150m standoff NIR images with various distanced VIS (database) images. Fig 8, Fig 9 and Fig 10 shows the CMC curve of face identification using 60m, 100m and 150 m NIR images respectively with various ranks. Accuracy of proposed approach is greater than existing one.

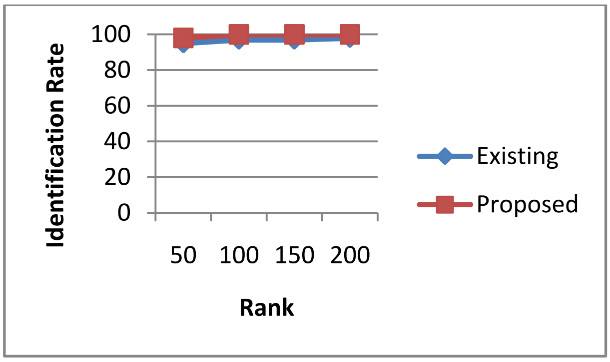

Fig 8. CMC curve of 60m NIR to VIS face identification

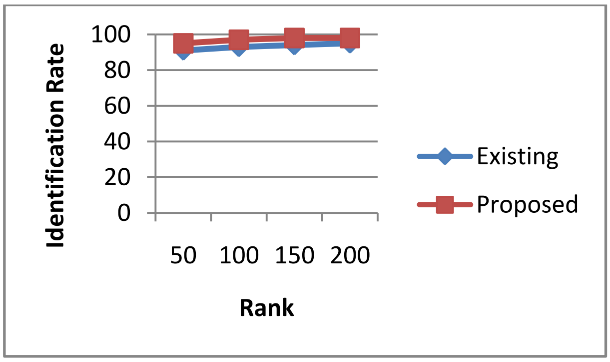

Fig 9. CMC curve of 100m NIR to VIS face identification

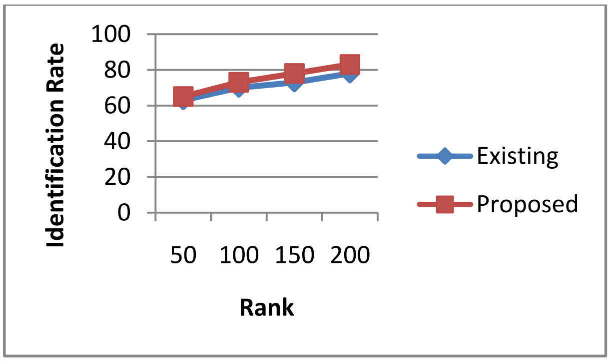

Fig 10. CMC curve of 150m NIR to VIS face identification

5. SUMMARY AND CONCLUSION

Heterogeneous face images are of different modality so there is great appearance difference between visible light and near infrared light images. Different VIS-NIR face matching studies have addressed bridging the cross-modal gap with a variety of synthesis, projection and feature-based approaches. This paper presented heterogeneous face matching system, which is used for matching near infrared face image to visible light face image. In this, used learning based image restoration method for restoring the long distanced NIR images. The performance of system is improved by using Difference-of-Gaussian filter to reduce illumination variation and HOG descriptor for feature extraction.

6. REFERENCES

[1] Sandhya R.Wadhavane and Prof. Dr. S. M. Kamalapur, “Heterogeneous Face Matching: NIR images to VIS images”, cPGCON 2016, Fifth Post Graduate Conference of Computer Engineering, March 2016

[2] D. Kang, et al., *Night time face recognition at a large standoff: cross-distance and cross-spectral matching*, Pattern Recognition(2014), \\http://dx.doi.org/10.1016/j.patcog.2014.06.004

[3] Zhifeng Li, Senior Member, IEEE, Dihong Gong, Yu Qiao, Senior Member, IEEE, and zacheng Tao, Senior Member*,” Common Feature Discriminant Analysis for Matching Infrared Face Images to Optical Face Images*,” June 2014

[4] Dihong Gong and Jiangyu Zheng. 2013. “*A Maximum Correlation Feature Descriptor for Heterogeneous Face Recognition”*. In Asian Conference on Pattern Recognition (ACPR).135–139

[5] B. Klare and A. K. Jain,“ *Heterogeneous face recognition using kernel prototype similarities*,” *IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 6, pp. 1410–1422, Jun. 2013*

[6] Xiong Pengfei, Lei Huang, and Changping Liu. 2012. “*A method for heterogeneous face image synthesis*”, In The IAPR International Conference on Biometrics (ICB). 1–6.

[7] Likun Huang, Jiwen Lu, and Yap-Peng Tan. 2012. “*Learning modality-invariant features for heterogeneous face recognition*” ,In International Conference on Pattern Recognition(ICPR). 1683–1686

[8] H. Maeng, S. Liao, D. Kang, S.-W. Lee, A .K. Jain, “*Nighttime face recognition at long distance: cross-distance and cross-spectral matching*,” in: Proceedings of ACCV,2012,pp.708–721

[9] D. Goswami, Chi-Ho Chan, D. Windridge, and J. Kittler. 2011. “*Evaluation of face recognition system in heterogeneous environments (visible v_s NIR)*”, In ICCV. 2160–2167.

[10] H. Maeng, H.C. Choi, U. Park, S. W. Lee, A. K. Jain, *NFRAD: near-infrared face recognition at a distance*, in: of IJCB,2011,pp.1–7.

[11] Jie Ch1en, Dong Yi, Jimei Yang, Guoying Zhao, S.Z. Li, and M. Pietikainen. 2009. “*Learning mappings for face synthesis from near infrared to visual light images*”, In CVPR .156–163.

[12] B. Klare and A. Jain, “*Heterogeneous Face Recognition: Matching NIR to Visible Light Images*,” *Proc. Int’l Conf. Pattern Recognition,2010.*

[13] S. Liao, D. Yi, Z. Lei, R. Qin, and S. Li. “*Heterogeneous face recognition from local structures of normalized appearance*”, In *Proc.3rd ICB*, 2009

[14] Z. Lei and S. Li. “*Coupled Spectral Regression for matching heterogeneous faces*”, In Computer Vision and Pattern Recognition, 2009. CVPR*2009.IEEE Conference on, pages 1123–1128, June 2009.2*

[15]R.Wang, J.Yang, D. Yi, and S. Z. Li. “*An analysis-by synthesis method for heterogeneous face biometrics”*, In Proceedings of *IAPR/IEEE International Conference on Biometrics,2009. 2*

[16]Z. Lei and S. Li. “*Coupled Spectral Regression for matching heterogeneous faces*”, In Computer Vision and Pattern Recognition, 2009. CVPR *2009.IEEE Conference on, pages 1123–1128, June 2009.2*

[17] D. Yi, R. Liu, R.-F. Chu, Z. Lei, and S. Z. Li, “*Face matching between near infrared and visible light images,*” in *Proc. Int. Conf. Adv. Biometrics*, 2007, pp. 523–530

[18] Lin, D., Tang, X.: “*Inter-modality face recognition”*, In: Leonardis, A. Bischof, H. Pinz, A. (eds.) *ECCV 2006. LNCS, vol. 3954, pp. 13–26. Springer, Heidelberg (2006)*