# Multiscale Representation for Partial Face Recognition Under Near Infrared Illumination

Lingxiao He, Haiqing Li, Qi Zhang, Zhenan Sun and Zhaofeng He
Center for Research on Intelligent Perception and Computing,
Institute of Automation, Chinese Academy of Sciences, Beijing, P.R. China, 100190
{lingxiao.he, hqli, qi.zhang, znsun}@nlpr.ia.ac.cn, hezhf@irisking.com

## Abstract

*Near infrared (NIR) partial face images acquired in iris recognition in less constrained environments contain plentiful identity information which have not been fully exploited. In this paper, a NIR partial face recognition (PFR) algorithm is designed according to the characteristics of NIR partial face images acquired in iris recognition systems. In the preprocessing stage, the eye corners are utilized for alignment because the eye regions are usually visible. In the feature representation stage, highly compact and discriminatory features are extracted based on Multiscale Double Supervision Convolutional Neural Network (MDSCNN) from several multiscale patches which are cropped from the aligned images. Weights of each multiscale patch are learned to improve the performance of PFR. Finally, the dissimilarity between two partial face images is calculated as the weighted $\ell_2$ distance between corresponding patches. A new NIR partial face (NIR-PF) database is constructed for extensive evaluation, which includes 5300 images acquired from 276 subjects when they walk towards an iris imaging device. Experimental results on the NIR-PF and CASIA-IrisV4-Distance databases demonstrate the effectiveness and efficiency of the proposed algorithm.*

## 1. Introduction

The accuracy of iris recognition drops dramatically in unconstrained scenarios because of poor image quality. Many complex imaging apparatus have been designed to capture high quality iris images in less constrained environments, but their capture volume and standoff distance are still limited [3][11]. It is inevitable capture defocused and motion-blurred NIR images when users walk towards an iris imaging device. As shown in Fig. 1(a) and (b), these images are unsuitable for iris recognition. In an iris recognition system at a distance, the high resolution partial faces caused by limited field of view contain plentiful identity information. However, current iris recognition at a distance or on the move systems have not taken advantage of partial faces. Therefore, it is necessary to develop a practical NIR PFR system to recognize PF as shown in Fig. 1(c).

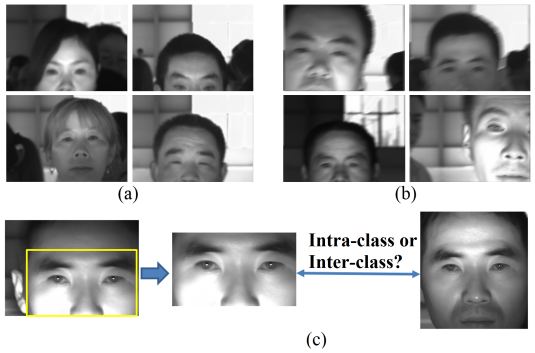


Figure 1. (a) Example of some defocus blurred images. (b) Example of some motion blurred images. (c) Example of partial face recognition.

Unlike partial face recognition (PFR) under visible illumination which has been comprehensively investigated [9]. PFR under NIR illumination has not been received attention until the Multiple Biometrics Grand Challenge (MBGC) in 2008 and 2009 [23][25]. The PFR algorithms for visible-light images can be directly transferred to NIR images but not obtain optimized performance in practical system. NIR partial face images in iris recognition systems present several characteristics compared to visible-light images: 1) Eye regions are usually visible in NIR images in an iris recognition system and not occluded by sunglasses, which can be used for face alignment. 2) As a byproduct of iris imaging devices, the resolution of faces in NIR images is much higher than that in visible-light images.

In this paper, we exploit the characteristics of NIR partial face images to design a fast and accurate NIR PFR algorithm. Firstly, the proposed algorithm aligns partial faces according to the inner and outer corner of each eye. Secondly, highly compact and discriminatory features are extracted based on Multiscale Double Supervision Convolu-

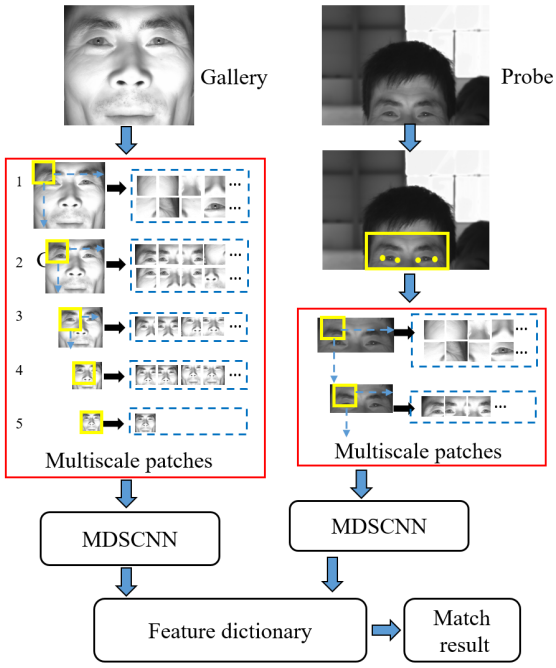


Figure 2. The proposed NIR partial face recognition algorithm. We extract multiscale patches features using MDSCNN for both gallery and probe face images. And then, we implement the patch-to-patch matching to search the minimum matching subject.

tional Neural Network (MDSCNN) from several multiscale patches which are cropped from the aligned images. The multiscale representation of partial faces contains local texture information in high resolution images and global structure information in coarse scale. Thirdly, weights of each multiscale patch is learned to improve the performance of PFR. Finally, the dissimilarity between two partial face images is calculated as the weighted $\ell_2$ distance between corresponding patches searched around by eyes centers. The flowchart of the algorithm is shown in Fig. 2.

The paper investigates the possibility of PFR in a NIR iris recognition system. The major contributions of our work are three-fold:

- NIR PF images are first exploited to assist iris recognition systems at a distance and a novel algorithm based on CNNs and triplet loss is proposed with promising results.
- We explore the significance of the portions of a face for recognition. And we arrive at a conclusion that periocular portion is much more important to face recognition than the other portions of a face.
- We developed a novel NIR-PF dataset to address the problem of PFR. The database will be released to the public to promote the research on NIR iris and face recognition in less constrained environments.

## 2. Related work

Face recognition is a hot topic in computer vision in recently years, which has achieved encouraging results on some public databases. However, the PFR has not been fully researched in the literature and it is confronted with some difficulties, such as partial face detection, partial face feature presentation.

Partial face detection is of great importance in partial face system. All current face detection algorithms based on deep learning [8][24] have achieved excellent results, but they would fail to work if the probe image is an arbitrary face patch. Likewise, partial face alignment also plays an important role for improving identification accuracy. However, face key-points are unavailable for aligning partial face because of incomplete face.

Partial face recognition is attractive for researching. In order to solve partial face recognition, some pioneer works have been proposed. Some approaches [5] [13] [16] use only certain patches of holistic faces, such as the left or right face, periocular, mouth, nose etc. [2] [10] [12] [14] divide the aligned face image into several patches and fuse the matching results by patch-to-patch matching. Nevertheless, it is difficult to detect and align facial portions in practice scenarios when partial face randomly appears.

Hu *et al.* [6] develop a partial face recognition approach based on local feature representation, where the similarity between each probe patch and gallery face is computed by the instance-to-class distance with the sparse constraint. Weng *et al.* [20] propose a Metric Learned Extended Robust Point Matching (MLERPM) algorithm to discriminatively match local features sets of a pair of gallery and probe samples. In their work, affine parameters need to be learned. However, some local features around each key-point are extracted, the computational cost will be extensive when there are many key-points. Besides, the performance of these algorithms seriously degrade when some of the facial portions are invisible or cannot be detected or aligned.

Liao *et al.* [9] propose an alignment-free face representation algorithm for partial face recognition, where multiple key-points descriptors are extracted for facial features representation, sparse representation based classification (SRC) [22] is used for classification. Their algorithm achieves good performance. However, the computational cost of their algorithm is extensive with the increase of register identities, which may limit its practical applications.

Although these proposed algorithms are capable of addressing the problem of partial face recognition, there are still many limitations in PFR: 1) Single local feature representation or single global feature representation is not very robust for PFR. 2) Partial face detection has not been solved in the state-of-the-art algorithms. Hence, we propose an al-

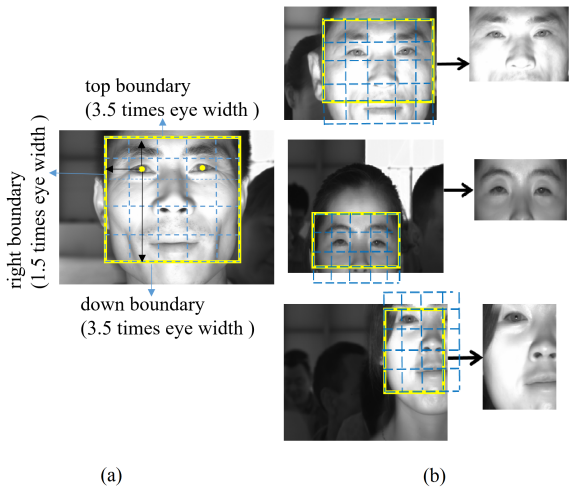


Figure 3. (a) The top, down, right, left boundary in a figure. (b) Some results of partial face detection.

gorithm based on MDSCNN that is able to extract distinguishing features, which helps to present robust features. Unlike visible-light face images, one or two eyes are visible in an iris recognition system at a distance under near infrared illumination. Therefore, eyes can be used to detect partial face.

## 3. The proposed method

### 3.1. Multiscale patches extraction

In visible-light face recognition system in unconstrained scenarios, eyes are easily occluded. Partial face detection is hard problem for face recognition system on this occasion. However, in NIR iris acquisition system at a distance, most images with one eye or two full eyes can be successfully detected. Therefore, patches in upper part are much more important than that in lower part. We utilize the characteristic of iris recognition system at a distance to detect partial face. If there is no full eye in a frame, we will detect the next frame. We use 9,000 positive samples and one million negative samples to train the eye detector with Haar features. Then, the left and right eye are combined to train a corner detection model using the algorithm in [15]. Consequently, the partial face is aligned by rotating and normalizing to the same size using the eyes corners.

In our proposed algorithm, the inner and outer eye corners are taken as eye key-points. And we use these eye key-points to attain a partial face bounding box. Therefore, a partial face is cropped into patches in order to attain bounding box of a partial face. The width and height of each patch are equal to the width of eyes, because the width of eye stay stable in our lifetime. Therefore, we design a rule to decide a partial face as shown in Fig. 3(a). 1) The distance between the left eye center and the left boundary is 1.5 times the width of the left eye. So do the right boundary. 2) The distance between eye center and the top boundary is 1.5 times the eye width. 3) The distance between eye center and the down boundary is 3.5 times the eye width. Fig. 3(b) show some partial face detection results. Eye corners can be used to detect PF and available multiscale patches.

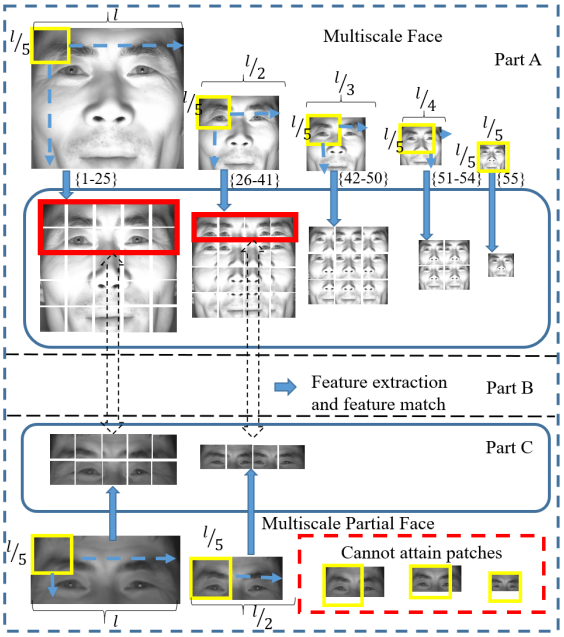


Figure 4. Multiscale representation for partial faces. Part A shows multiscale patches of a holistic face. Part B represents feature extraction and feature matching. Part C represents multiscale patches of a partial face.

If the size of the whole face is $l \times l$, a five scales face image pyramid is built by down-sampling as shown in Fig. 4-Part A. The sizes of a face image pyramid from the bottom scale to top scale are $l \times l$, $l/2 \times l/2$, $l/3 \times l/3$, $l/4 \times l/4$, $l/5 \times l/5$ respectively. A sliding window of $l/5 \times l/5$ with strides of $l/5 \times l/5$, $l/10 \times l/10$, $l/15 \times l/15$, $l/20 \times l/20$, $l/25 \times l/25$ is used to attain the patches. Therefore, a holistic face image pyramid from the bottom scale to top scale have 25, 16, 9, 4, 1 patches respectively. The 55 patches contains different scales patches, which helps to extract local features and structured features and is also beneficial to present features robustly. As shown in the Fig. 4 - Part C, the same patches extraction algorithm is used for a partial face. But patches cannot be attained in the top 3-scale because the height of the partial face is smaller than the width of the sliding window.

### 3.2. Deep multiscale representation

CNN is capable of extracting highly compact features, which can learn discriminatory features from raw pixel values of multiscale face patches. Fig. 5 shows the detail structure of our designed MDSCNN, which contains four convolutional layers followed by max-pooling as in [7] and one full-connected layer (highly compact features). Three cor-

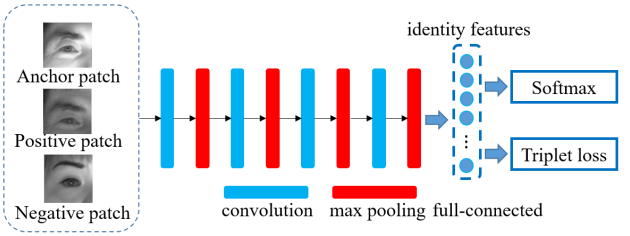


Figure 5. The detail structure of MDSCNN.

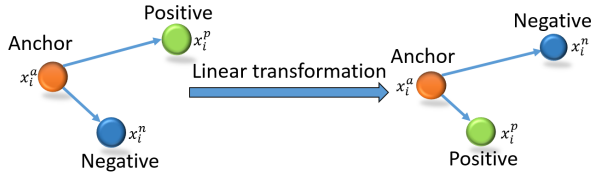


Figure 6. The Triplet Loss minimizes the distance between an anchor and a positive, both of which have the same identity, and maximizes the distance between the anchor and a negative of a different identity.

responding face patches are used as the input of CNN. The first convolutional layer filters the input three face patches with 48 kernels of size $5 \times 5$ at a stride of 1 pixel. The second convolutional layer takes as input (after max-pooling) the output of the first convolutional layer and filters it with 48 kernels of size $5 \times 5 \times 48$. The third convolution layer has 48 kernels of size $5 \times 5 \times 48$ connected to the outputs (after max-pooling) of the second convolutional layer. The fourth convolution layer has 48 kernels of size $5 \times 5 \times 48$. The convolutional operation is expressed as

$$y^{j(r)} = \max(0, b^{j(r)} + \textstyle\sum_i k^{ij(r)} * x^{i(r)}) \quad (1)$$

where $x^i$ and $y^j$ are the $i$-th input map and the $j$-th output respectively. $k^{ij}$ is the convolutional kernel between the $i$-th input map and the $j$-th output map. $b^j$ is the bias of the $j$-th output map. We use ReLU nonlinearity ($y = \max(0, x)$), which is able to extract rotation and scale invariant features. The full-connected layer has 128 neurons with highly compact features. The output layer have double supervisory information. One of supervisory information is *softmax loss*, which can classify different subjects. And the other supervisory information is *triplet loss* [17], which helps to improve recognition accuracy.

Softmax loss can be regard as a classifier which can identify different subjects. The softmax loss output probability which predict the probability distribution over $c$ different identities by computing the empirical *softmax log-loss*:

$$J_1(\theta) = -\frac{1}{m}\left[\sum_{i=1}^{m}\sum_{c_t=1}^{N} sign\{y^{(i)} = c\} \log \frac{e^{\theta_c^T x^{(i)}}}{\sum_{l=1}^{k} e_{\theta_l}^T x^{(i)}}\right] \quad (2)$$

where $sign\{\cdot\}$ is the indicator function, $sign\{y^{(i)} = c\} = 1$ when $y^{(i)} = c$, otherwise $sign\{y^{(i)} = c\} = 0$. $x \in \mathbb{R}^D$ is the output of the final full-connected layer.

**Triplet loss.** In order to achieve high performance in partial face recognition, we strive for an embedding $f(x)$, from a patch $x$ into a feature space $\mathbb{R}^D$. In the feature space $\mathbb{R}^D$, the squared distance between patches from the same identity is small, whereas the squared distance between patches from different identities is large. The motivation is that all patches of faces from the same identity are projected onto one single point in the embedding space $\mathbb{R}^d$. The embedding is represented by $f(x) \in \mathbb{R}^D$. It embeds a patch $x$ into a $D$-dimensional Euclidean space, and we constrain this embedding to live on the $D$-dimensional hypersphere. This loss is motivated in [18] in the context of nearest-neighbor classification. As shown in Fig. 6, in Euclidean space the aim of triplet loss is to ensure a patch $x_i^a$ (anchor) of a specific person closer to all other patches $x_i^p$ (positive) than to any patches $x_i^n$ (negative) of other persons [19]. The *triplet loss* $J_2(\theta)$ is minimized through Euclidean distance as follows:

$$J_2(\theta) = \frac{1}{m}\sum_{i}^{m}[\|f(x_i^a) - f(x_i^p)\|_2^2 - \|f(x_i^a) - f(x_i^n)\|_2^2 + \alpha] \quad (3)$$

where $f(x) = Wx$, $W$ is a transformation matrix, $x$ is patch feature vector.

The loss function of the MDSCNN is

$$J(\theta) = J_1(\theta) + J_2(\theta) \quad (4)$$

$J(\theta)$ is a non-convex function, gradient descent is susceptible to local optimal. So the key step is to compute the partial derivatives of $Eq.$ 4. We make use of Stochastic Gradient Descent (SGD) [21] to optimize the parameters of the network.

We train 55 different MDSCNNs, one for each multiscale patch. The features of each multiscale patch are denoted as $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{55}$. Let $\mathbf{X}_c = (\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{55})$, where each descriptor is an $M$-dimensional vector ($M$ is the feature length). If enrollment dictionary contains $c$ subjects, a gallery dictionary for the $c$ subjects is built as $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_c\}$.

### 3.3. Weights learning

Different portions of face have different weights for face recognition. Therefore, it is necessary to learn weights of each multiscale patch to improve the performance of PFR.

Let us consider three face images $\mathbf{X}^a$, $\mathbf{X}^p$ and $\mathbf{X}^n$ with 55 multiscale patches

$$\begin{aligned} \mathbf{X}^a &= (x_1^a, x_2^a, ..., x_{55}^a) \\ \mathbf{X}^n &= (x_1^n, x_2^n, ..., x_{55}^n) \\ \mathbf{X}^p &= (x_1^p, x_2^p, ..., x_{55}^p) \end{aligned} \quad (5)$$

where $\mathbf{X}^a$, $\mathbf{X}^p$ come from the same identity and $\mathbf{X}^a$, $\mathbf{X}^n$ come from different identities. Each multiscale patch has different weights $\alpha_i$. Features of the corresponding multiscale patches are calculated by $\ell_2$ distance and we fuse the matching results.

We aim to attain $\alpha_i$ to solve multiscale features matching minimization problem that is to ensure the distance of the same identity closer and the distance of different identities farther. Therefore, the loss function can be defined as

$$D = arg\min \sum_{i=1}^{55} \alpha_i (d_i^{(ap)} - d_i^{(an)} + \beta) \quad s.t \ \sum_{i=1}^{55} \alpha_i = 1, \alpha_1 \geq 0 \tag{6}$$

where $d_i^{(ap)} = \|\mathbf{X}_i^a - \mathbf{X}_i^p\|_2$ , $d_i^{(an)} = \|\mathbf{X}_i^a - \mathbf{X}_i^n\|_2$, the algorithm of Lagrange multipliers (named after Joseph Louis Lagrange [4]) is a strategy for finding the local minima of a function subject to equality constraints.

After learning the weights of each multiscale patch, the dissimilarity between two partial face images is calculated as the weighted $\ell_2$ distance between corresponding patches to search the subject of minimum matching distance.

## 4. Experiments

To verify the effectiveness of our partial face recognition algorithm, we provide experimental results on two databases: NIR-PF, CASIA-IrisV4-Distance.

### 4.1. Database

**NIR-PF database.** NIR-PF database is a new partial face database, which includes 276 subjects with multi partial face images. Each subject contains a sequence of face images (16-20 face images), half of which can cover full face. In the database, Partial face images are captured under NIR camera while people moving that exhibits different arbitrary part of a face. Face images are also different in distance and view, including variations in scales, lighting conditions, focus. Fig. 7 shows some examples of the NIR-PF database, the size of each face image is 3840×2748.

**CASIA-IrisV4-Distance database.** [1] The CASIA-IrisV4-Distance database contains both iris and face patterns as shown in Fig. 8, which includes 152 subjects with 2567 images in total. Each image from the CASIA-Iris-Distance contains two eyes and partial frontal face regions with a high resolution (2352×1728) and is captured at about 3 meters away by near-infrared cameras. CASIA-IrisV4-Distance is a challenging and multi-modal database including variations of lighting conditions, glasses, expression, etc. It can be used to face recognition and iris recognition.

### 4.2. Experiments on the NIR-PF database

This experiment mainly investigates the performance of MDSCNN. The training set and testing set of the two databases are shown in Table 1. In our training, triplet loss is used in our CNNs. Therefore, big scale triplet samples can be attained to avoid overfitting.

Figure 7. Examples of NIR-PF database. Partial face in an image is randomly, which contains left face, right face, or upper face, ect.

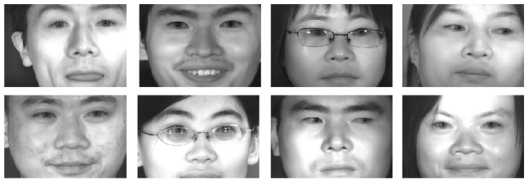

Figure 8. Examples of the CASIA-IrisV4-Distance database.

Table 1. Settings of the training set and testing set of the two databases.

| Database | | NIR-PF | CASIA-IrisV4-Distance |
|---|---|---|---|
| Training | subjects | 171 | 60 |
| | images | 3200 | 900 |
| Testing | subjects | 105 | 92 |
| | images | 2100 | 1380 |

**Comparison algorithm.** The SIFT-SRC [9], MLERPM [20] implemented by ourselves are used for comparison.

**Evaluation Protocol.** We use the Receiver Operating Characteristic (ROC) curves to evaluate the performance of these algorithms.

Fig. 9 shows the ROC curves of different algorithms. The results show that the proposed MDSCNN perform better than other algorithms. MDSCNN could extract local and global features, which improve the robustness of features. However, MLERPM, SIFT-SRC only adopt local features around key-points. Besides, weighted patch matching can improve the recognition performance.

Fig. 10 shows the weights of each multiscale patch. Red columns represent multiscale patches around the periocular portion, which suggests that weights of periocular region at each scale are bigger than other portions of a face. Therefore, periocular portion is much more important to face recognition than other portions of a face. The order of importance for recognition is as follows: periocular> nose> mouth. In addition, the weights increase with the in-

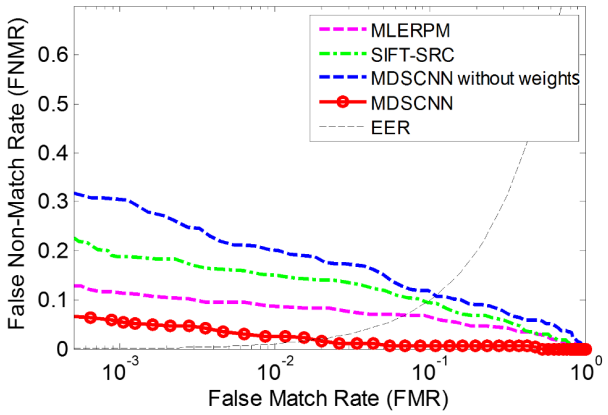


Figure 9. ROC curves on NIR-PF database.

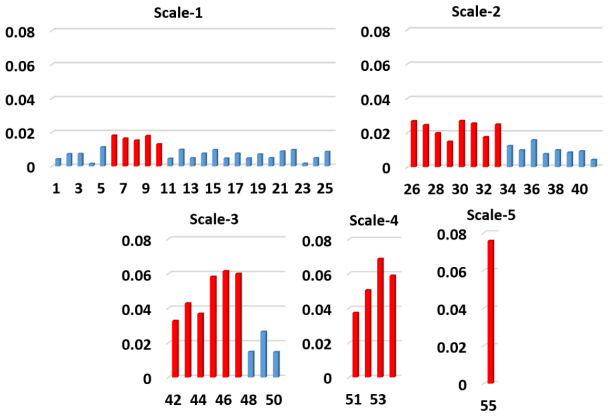


Figure 10. Weights of mulitiscale patches.

creasing of scale because patches of big scale contain more information.

### 4.3. Experiments on the CASIA-IrisV4-Distance database

To demonstrate robustness of MDSCNN, we perform another experiment on the CASIA-IrisV4-Distance database. Face images in the CASIA-IrisV4-Distance database are captured under near infrared illumination, including variations in lighting conditions, glasses, expression, which can verify performance of MDSCNN at the best. All gallery face images are first cropped to 300×300 pixels according to eyes detection in the same manner to achieve partial face recognition. Next, a patch at random position of a random size $h \times w$ is cropped to represent a partial face for each cropped probe image, where both $h$ hand $w$ and are uniformly distributed in [60,240]. According to $h$ hand $w$, we can get different sizes of probe face patches, such as 20% deletion, 40% deletion, 60% deletion. Fig. 11 shows some instances of such randomly cropped face patches. We use 20% deletion face, 40% deletion face, 60% deletion face to do experiment respectively.

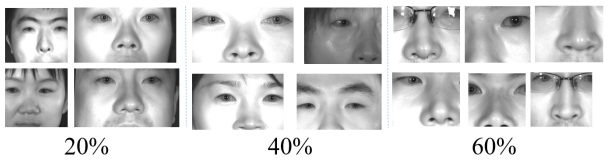


Figure 11. Examples of partial face from CASIA-IrisV4-Distance. We randomly cut off part of face to make up partial face database that contains three different sizes probe face of 20% deletion, 40% deletion and 60% deletion.

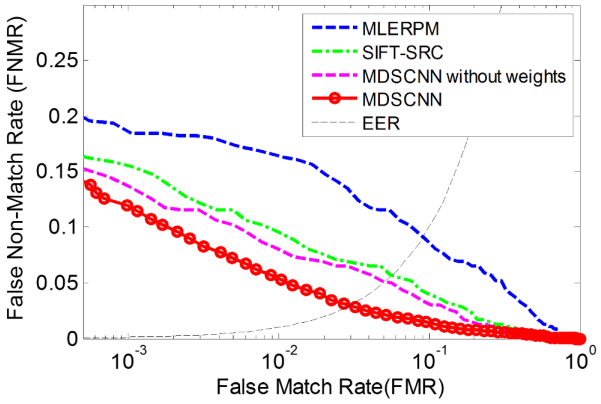


Figure 12. ROC curves on CASIA-IrisV4-Distance.

The SIFT-SRC, MLERPM are used for comparison. Fig. 12 shows the experimental results (20%, 40% 60% cropping faces are all tested). A conclusion can be drawn from Fig. 12 that the proposed MDSCNN achieves a better performance than SIFT-SRC and MLERPM. The experimental results suggest that MDSCNN is capable of extracting local features and coarse features. Multiscale feature presentation contributes to improving the accuracy of PFR.

## 5. Conclusions

In this paper, we propose a novel partial face recognition algorithm for NIR images captured in iris recognition systems. Several characteristics of NIR partial face images have been considered in both the preprocessing and feature representation procedures. The proposed algorithm aligns partial faces according to the inner and outer corner of each and cuts partial faces into multiscale patches, which contributes to local and global feature extraction. Secondly, highly compact and discriminatory features are extracted based on the MDSCNN from several multiscale patches which are cropped from the aligned images. Thirdly, the weight of each multiscale patch is learned to improve the performance of PFR. Finally, the dissimilarity between two partial face images is calculated as the weighted $\ell_2$ distance between corresponding patches. Extensive experimental results on the NIR-PF database and the CASIA-IrisV4-Distance database have demonstrated that the multiscale double supervision feature representation based on

CNN of partial faces is highly discriminative and efficient.

We will integrate the proposed algorithm into a multi-modal identity recognition system which fuses NIR face and iris information. The integration strategy of iris and partial face will further improve the performance of iris recognition system at a distance.

## 6. Acknowledgements

This work is supported by the National Natural Science Foundation of China (Grant No.61403389) , the Beijing Nova Programme (Grant No. Z141101001814090), the Beijing Talents Fund (Grant No. 2015000021223ZK30) and the Beijing Municipal Science and Technology Project-Grant No. Z141100003714131, Z161100000216144).

## References

[1] Chinese Academy of Sciences Insitute of Automation (CASIA). CASIA Iris Image database, http://biometrics.idealtest.org/. 2010. 5

[2] T. Ahonen, A. Hadid, and M. Pietikainen. Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12):2037–2041, 2006. 2

[3] F. Bashir, P. Casaverde, D. Usher, and M. Friedman. Eagle-eyes: a system for iris recognition at a distance. In *IEEE Conference on Technologies for Homeland Security*, pages 426–431, 2008. 1

[4] D. P. Bertsekas. Constrained optimization and lagrange multiplier methods. 1996. 5

[5] S. Gutta, V. Philomin, and M. Trajkovic. An investigation into the use of partial-faces for face recognition. In *Proceedings of International Conference on Automatic Face and Gesture Recognition*, pages 33–38, 2002. 2

[6] J. Hu, J. Lu, and Y.-P. Tan. Robust partial face recognition using instance-to-class distance. In *Visual Communications and Image Processing*, pages 1–6, 2013. 2

[7] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012. 3

[8] H. Li, Z. Lin, X. Shen, J. Brandt, and G. Hua. A convolutional neural network cascade for face detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5325–5334, 2015. 2

[9] S. Liao, A. K. Jain, and S. Z. Li. Partial face recognition: Alignment-free approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(5):1193–1205, 2013. 1, 2, 5

[10] A. M. Martínez. Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(6):748–763, 2002. 2

[11] J. R. Matey, O. Naroditsky, K. Hanna, R. Kolczynski, D. J. LoIacono, S. Mangru, M. Tinker, T. M. Zappia, and W. Y. Zhao. Iris on the move: Acquisition of images for iris recognition in less constrained environments. *Proceedings of the IEEE*, 94(11):1936–1947, 2006. 1

[12] R. Min, A. Hadid, and J.-L. Dugelay. Improving the recognition of faces occluded by facial accessories. In *IEEE International Conference on Automatic Face & Gesture Recognition and Workshops*, pages 442–447, 2011. 2

[13] H. Neo, C. Teo, and A. B. Teoh. Development of partial face recognition framework. In *International Conference on Computer Graphics, Imaging and Visualization*, pages 142–146, 2010. 2

[14] K. Pan, S. Liao, Z. Zhang, S. Z. Li, and P. Zhang. Part-based face recognition using near infrared images. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–6, 2007. 2

[15] S. Ren, X. Cao, Y. Wei, and J. Sun. Face alignment at 3000 fps via regressing local binary features. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1685–1692, 2013. 3

[16] M. Savvides, R. Abiantun, J. Heo, S. Park, C. Xie, and B. Vijayakumar. Partial & holistic face recognition on frgc-ii data using support vector machine. In *International Conference on Computer Vision and Pattern Recognition Workshop*, pages 48–48, 2006. 2

[17] F. Schroff, D. Kalenichenko, and J. Philbin. Facenet: A unified embedding for face recognition and clustering. *arXiv preprint arXiv:1503.03832*, 2015. 4

[18] J. Wang, Y. Song, T. Leung, C. Rosenberg, J. Wang, J. Philbin, B. Chen, and Y. Wu. Learning fine-grained image similarity with deep ranking. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1386–1393, 2014. 4

[19] K. Q. Weinberger and L. K. Saul. Distance metric learning for large margin nearest neighbor classification. *The Journal of Machine Learning Research*, 10:207–244, 2009. 4

[20] R. Weng, J. Lu, J. Hu, G. Yang, and Y.-P. Tan. Robust feature set matching for partial face recognition. In *IEEE International Conference on Computer Vision*, pages 601–608. 2, 5

[21] D. R. Wilson and T. R. Martinez. *Neural Networks*, 16(10):1429–1451, 2003. 4

[22] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma. Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(2):210–227, 2009. 2

[23] J. Yang, S. Liao, and S. Z. Li. Automatic partial face alignment in nir video sequences. In *Advances in Biometrics*, pages 249–258. Springer, 2009. 1

[24] S. Yang, P. Luo, C.-C. Loy, and X. Tang. From facial parts responses to face detection: A deep learning approach. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3676–3684, 2015. 2

[25] D. Yi, S. Liao, Z. Lei, J. Sang, and S. Z. Li. Partial face matching between near infrared and visual images in mbgc portal challenge. In *Advances in Biometrics*, pages 733–742. Springer, 2009. 1