IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 24, 2020, accepted March 22, 2020, date of publication March 27, 2020, date of current version April 15, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2983837

Diagonal Symmetric Pattern-Based Illumination Invariant Measure for Severe Illumination Variation Face Recognition

CHANGHUI HU[1,2,3], FEI WU[1], JIAN YU[1], XIAOYUAN JING[1], XIAOBO LU[2], AND PAN LIU[3]

[1]College of Automation and College of Artificial Intelligence, Nanjing University of Posts and Telecommunications, Nanjing 210023, China
[2]School of Automation, Southeast University, Nanjing 210096, China
[3]School of Transportation, Southeast University, Nanjing 210096, China

Corresponding author: Changhui Hu (hchnjupt@126.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61802203 and Grant 61702280, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20180761, Grant BK20170900, and Grant BK20191377, in part by the China Postdoctoral Science Foundation under Grant 2019M651653, in part by the Jiangsu Planned Projects for the Postdoctoral Research Funds under Grant 2019K124, in part by the National Postdoctoral Program for Innovative Talents under Grant BX20180146, in part by the Natural Science Foundation of the Jiangsu Higher Education Institutions of China under Grant 18KJB520034, and in part by the Nanjing University of Posts and Telecommunications Science Foundation under Grant NY218119.

ABSTRACT The center symmetric pattern (CSP) was widely used in the local binary pattern based facial feature, whereas never used to develop the illumination invariant measure in the literature. This paper proposes a novel diagonal symmetric pattern (DSP) to develop the illumination invariant measure for severe illumination variation face recognition. Firstly, the subtraction of two diagonal symmetric pixels is defined as the DSP unit in the face local region, which may be positive or negative. The DSP model is obtained by combining the positive and negative DSP units in the even $\times$ even block region. Then, the DSP model can be used to generate several DSP images based on the 2×2 block or the 4×4 block by controlling the proportions of positive and negative DSP units, which results in the DSP2 image or the DSP4 image. The single DSP2 or DSP4 image with the arctangent function can develop the DSP2-face or the DSP4-face. Multi DSP2 or DSP4 images employ the extended sparse representation classification (ESRC) as the classifier that can form the DSP2 images based classification (DSP2C) or the DSP4 images based classification (DSP4C). Further, the DSP model is integrated with the pre-trained deep learning (PDL) model to construct the DSP-PDL model. Finally, the experimental results on the Extended Yale B, CMU PIE, AR, and VGGFace2 face databases indicate that the proposed methods are efficient to tackle severe illumination variations.

INDEX TERMS Severe illumination variations, diagonal symmetric pattern, center symmetric pattern, single sample face recognition.

I. INTRODUCTION

Face recognition has been a hot research topic for decades due to its wide application prospects [1]–[3]. The illumination variation problem is inevitable in face recognition, even for the deep learning features [4]–[5]. Severe illumination variations are considered as tough issues for the face images in the outdoor environment, such as the driver face images in the intelligent transportation systems [6]. Hence, it is still significance to address illumination variations in face recognition, especially for severe illumination variations. As numerous approaches have been proposed to tackle severe illumination variations, some significant works are selected to review in this paper.

The illumination holding based approach and the illumination eliminating based approach are two categories of methods to address illumination variations in face recognition. The illumination holding based approach [7] aims to recover the normal lighting of the illumination variation face image. The illumination eliminating based approaches [8]–[12] remove the illumination and obtain the illumination independent features from the illumination variation face image. As the illumination recovery could cause the face discriminant information distortion in the illumination holding based

The associate editor coordinating the review of this manuscript and approving it for publication was Andrea F. F. Abate.

approach, the illumination eliminating based approach is more efficient to tackle severe illumination variations. In fact, most of illumination eliminating based approaches were developed based on the lambertian reflectance model [13]. The face reflectance [8], the face high-frequency facial features [5], [9], [10], and the face illumination invariant measures [6], [10]–[12] are very efficient to tackle severe illumination variations.

The face reflectance [8] employed the lambertian reflectance model [13] to estimate the reflectance and the illumination from the illumination contaminated face image simultaneously. The logarithmic total variation (LTV) model [9] extended the lambertian reflectance model to extract the small scale facial structures (i.e. the high-frequency facial features) and the large scale facial structures of the illumination contaminated face image. The high-frequency single value decomposition face (HFSVD-face) [10] firstly used the frequency interpretation of the single value decomposition algorithm to extract the high-frequency facial features of the illumination contaminated face image, Recently, HFSVD-face was extended to the orthogonal triangular with column pivoting (QRCP) decomposition algorithm, which resulted in that the QRCP decomposition was first used to construct the QRCP-face and the normalized QRCP-face [5].

The illumination invariant measures [6], [10]–[12] constructed the reflectance based pattern by eliminating the illumination of the face image. The Weber-face [11] proposed a simple reflectance based pattern that the differences of the center pixel and its neighboring pixels were divided by the center pixel in the 3×3 block region, which can eliminate the illumination of the face image under the lambertian reflectance model [13], since the face illumination invariant measure assumes that illumination intensities of neighborhood pixels are approximately equal in the face local region. Then, the pixel domain based Weber-face was extended to the logarithm domain, and several illumination invariant measures were proposed such as [6], [10], and [12], since the illumination invariant measure of the logarithm domain was proved to have better tolerance to illumination variations than that of the pixel domain in mathematics [10].

Nowadays, the deep learning features [14]–[17] are the best for face recognition, which require massive available face images to train. VGG [14] was trained by 2.6M internet face images (2622 persons and 1000 images per person). The matching/non-matching face patches of 200M internet face images were employed to train Facenet [15]. CASIA-WebFace face images (10575 persons and 0.49M images) were used to train Cosface [16]. Arcface [17] was trained by 85742 persons and 5.8M internet face images. As large scale face images for training the deep learning model are collected via internet, the deep learning features performed very well on internet face images. However, the internet face images are not with severe illumination variations, thus the deep learning features performed unsatisfactorily under severe illumination variations [5], [6].

A. MOTIVATION

The illumination variation problem is one of the tough issues for the face images in the outdoor environment. Zhang *et al.* [4] claimed that the current deep learning based face recognition required a huge amount of labeled face data, which were unable to cover the infinite illumination variations that can occur in real-life applications. Hence, illumination processing approaches continue to be crucially important for further widening the application field of face recognition. Hu *et al.* [5] indicated that the performance of the current deep learning feature exhibited unsatisfactorily under severe illumination variations.

In many practical applications, such as the driver face recognition in the intelligent transportation systems [6], it is very difficult to collect enough image pairs with severe illumination variations to train an illumination robust deep learning model or fine-tune a well-trained deep learning network. Severe illumination variations are still challenging for the commonly used deep learning features [14]–[17], since the sate-of-the-art deep learning models were trained by massive light internet face images without considering severe illumination variations.

The illumination invariant measures [6], [10]–[12] were efficient to tackle severe illumination variations, which were developed via the subtraction of the center pixel and its neighboring pixel in the face local region. In the literature, the CSP model that employed the subtraction of two center symmetric pixels in the face local region has never been used to construct the illumination invariant measure. In fact, the subtraction of the center pixel and its neighboring pixel, and the subtraction of two center symmetric pixels are quite different facial features.

Our previous work [6] first developed the positive and negative illumination invariant units to construct the generalized illumination robust (GIR) model, that made the illumination invariant measure independent of the weights associated with multi face local regions.

Inspired by the CSP model and the GIR model, we are motivated to develop a novel diagonal symmetric pattern based illumination invariant measure to tackle severe illumination variation face recognition.

B. CONTRIBUTION

In this paper, a novel local model named diagonal symmetric pattern (DSP) is proposed to develop the illumination invariant measure for severe illumination variations. Compared to the previous works such as [7]–[12], the new contributions are:

(1) To the best of our knowledge, the center symmetric pattern (CSP) has never been used to develop the illumination invariant measure in the literature. This paper proposes a novel pixel-wise DSP model, which employs the subtraction of two diagonal symmetric pixels in the face local region to construct the illumination invariant measure. The proposed DSP model is

diagonal symmetry, whereas horizontal asymmetry and vertical asymmetry.

(2) Different from the traditional illumination invariant measure that employed the odd × odd block region such as the 3×3 block or the 5×5 block to have a center pixel, the proposed DSP model uses the even × even block region such as the 2 × 2 block or the 4 × 4 block, and the center pixel is unnecessary. To the best of our knowledge, the 2 × 2 block has never been used in the LBP based facial feature or the illumination invariant measure in the literature.

(3) The proposed DSP model employs the relationships of the diagonal symmetric pixels, and the diagonal symmetric is the only relationship of pairs of pixels in the DSP model. The proposed DSP model can fully use all pixels in the even × even block region, and none pixel is ignored.

(4) The illumination invariant measure is developed by the DSP model based on the 2 × 2 block region or the 4 × 4 block region, which can generate multi DSP2 or DSP4 images. Further, the DSP model is integrated with the pre-trained deep learning model to tackle severe illumination variation face recognition.

C. ORGANIZATION

The remainder of this paper is organized as follows. Section II describes the related works. Section III elaborates the illumination invariant measure based on the diagonal symmetric pattern. Section IV presents the classification algorithm. Section V gives the experiments, and Section VI concludes this paper.

II. RELATED WORKS

The local binary pattern (LBP) based approach was an efficient hand-crafted facial descriptor, and robust to various facial variations. The center symmetric local binary pattern (CSLBP) [18] employed the center symmetric pairs of 8 pixels in the local block region with the size of 3 × 3, whereas the center pixel was not used. The symmetric pixel pairs of the CSLBP were diagonal symmetry, horizontal symmetry or vertical symmetry. Recently, the center symmetric quadruple pattern (CSQP) [19] extended the CSP to the quadruple space, which divided the 4 × 4 block into 4 sub-blocks with the size of 2 × 2, and pixels of diagonally opposite sub-blocks were compared to generate a binary pattern of 8 bits. The attractive-and-repulsive center symmetric local binary patterns (ACS-LBP and RCS-LBP) [20] employed the four triplets corresponding to the vertical and horizontal directions, and the two diagonal directions by establishing relations between the center pixel and the center symmetric pairs of pixels in the block with the size of 3×3. The cascaded asymmetric local pattern (CALP) [21] utilized three pairs of pixels of the vertical direction and three pairs of pixels of the horizontal direction in the block with the size of 3 × 3. The fully center symmetric dual cross pattern (FCSDCP) [22] used not only the center symmetric pairs of pixels of the four directions (i.e. the vertical and horizontal directions, and the two diagonal directions) but also the non-center-symmetric pixels in the block with the size of 5 × 5.

The multiscale logarithm difference edgemaps (MSLDE) [12] extracted facial features from multi edges of six local block regions, whose sizes were from 3 × 3 to 13 × 13. The local near neighbor face (LNN-face) [10] was attained from five local block regions, whose sizes were from 3 × 3 to 11 × 11. In [10] and [12], different weights were assigned to different local edges or blocks, whereas the edge based generalized illumination robust face (EGIR-face) and the block based generalized illumination robust face (EGIR-face) [6] equally treated different edges and blocks, and removed the weights associated with multi edges and blocks. Recently, the reflectance ratio and histogram equalization (RRHE) based illumination normalization framework [23] conducted histogram equalization on the ratio of each pixel intensity to its local mean pixel intensity, where the local mean pixel intensity was calculated by the block with size of 3 × 3, 5 × 5, 7 × 7, 9 × 9 or 11 × 11.

Nowadays, the illumination invariant measure mainly utilized relationships between the center pixel and the pixels around the center in the face local region. The CSP based approaches developed many efficient relationships of the center symmetric pixels which were widely used in the LBP based facial features, whereas seldom employed in the illumination invariant measure.

III. ILLUMINATION INVARIANT MEASURE BASED ON DIAGONAL SYMMETRIC PATERN

A. PROBLEM FORMULATION

Based on the commonly used assumption that illumination intensities of neighborhood pixels are approximately equal in the face local region, the illumination invariant measure aims to eliminate the illumination of the face image with severe illumination variations. Currently, the illumination invariant measure depends on relationships between the center pixel and its neighbors, which is limited to use the odd × odd block region. In order to break through this limitation, we propose the DSP model, which takes the relationships of the diagonal symmetric pixels in the even × even block region to construct the illumination invariant measure.

B. THE DIAGONAL SYMMETRIC PATTERN

The CSP model employs the center symmetric pixels or sub-blocks to develop the facial feature in the face local region. The center of the CSP model can be a center pixel in the odd × odd face local region such as [18], [20]–[23], or the intersection of the horizontal and vertical axes that divide the even × even face local block region into 4 sub-blocks such as [19]. In fact, the current illumination invariant measure always employed the odd × odd face local region with a center pixel, and never used the even × even face local region whose center is the intersection of the horizontal and vertical axes.

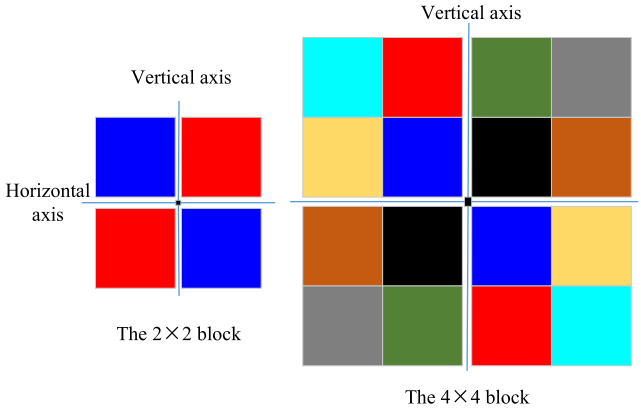


FIGURE 1. The proposed pixel-wise DSP model.

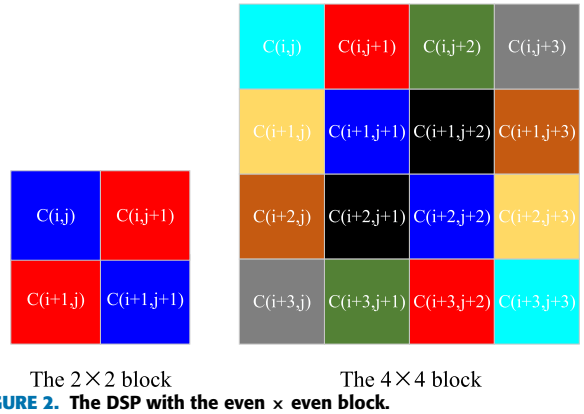


FIGURE 2. The DSP with the even × even block.

In this paper, we propose the pixel-wise DSP model, and its symmetric center is the intersection of the horizontal and vertical axes, which can divide an even × even block region into 4 pixels or sub-blocks. Pairs of pixels of the DSP model are strictly diagonal symmetry about the intersection of the horizontal and vertical axes.

Fig.1 shows the proposed pixel-wise DSP model. The 2×2 block and the 4×4 block incorporate 4 pixel-blocks and 16 pixel-blocks respectively. A pixel-block represents one pixel in the face local region, and two pixel-blocks with the same color such as two blue blocks, are diagonal symmetric pixels. The subtraction of the two diagonal symmetric pixels is defined as the DSP unit in this paper.

It is appropriate to indicate that the proposed DSP model strictly follows pixel-wise diagonal symmetry, whereas the CSQP [19] was sub-block-wise diagonal symmetry, and pairs of pixels of the CSQP are neither diagonal symmetry nor center symmetry. The CSLBP [18] was pixel-wise center symmetry, and pairs of pixels of the CSLBP were diagonal symmetry, horizontal symmetry or vertical symmetry. Pairs of pixels of the proposed DSP model are diagonal symmetry, whereas horizontal asymmetry and vertical asymmetry.

C. ILLUMINATION INVARIANT MEASURE UNDER THE DSP WITH THE EVEN × EVEN BLOCK

Suppose m $\geq$ n, the logarithm image C is with m rows and n columns. Fig.2 shows the DSP with the even × even block, which describes a 2 × 2 block and a 4 × 4 block of the logarithm image C, where (i,j) denotes the location of the image point of the i-th row and the j-th column. C(i,j) denotes the image point intensity at the location (i,j). From the lambertian reflectance model [13], the logarithm image can be presented as C(i,j) = ln(R(i,j)) + ln(L(i,j)), where R and L are the reflectance and the illumination respectively.

According to Section III-B, the subtraction of the two diagonal symmetric pixels is defined as the DSP unit. Due to the assumption that illumination intensities are approximately equal in the face local region, we can obtain the DSP unit in the even × even block as

$$\begin{aligned} &I\left(i+d,j+p\right)\\ &=C\left(i+d,j+p\right)-C\left(i+N-d-1,j+N-p-1\right)\\ &=lnR\left(i+d,j+p\right)-lnR\left(i+N-d-1,j+N-p-1\right). \end{aligned} \tag{1}$$

where N = 2 is for the 2×2 block, and N = 4 is for the 4×4 block. p is the column variate and $p = 0, \cdots, N-1$. d is the row variate and $d = 0, \cdots, N/2-1$. In (1), $I\left(i+d,j+p\right)$ is the logarithm reflectance subtraction of the DSP pixel pair in the even × even block. It is easy to know that $I\left(i+d,j+p\right)$ is independent of the illumination. Hence, the DSP unit is illumination invariant.

Accordingly, the illumination invariant measure under the DSP with the N × N block (IIM-DSP) can be presented as

$$\text{IIM-DSP}\left(i,j\right)=\sum_{d=0}^{\frac{N}{2}-1}\sum_{p=0}^{N-1}I\left(i+d,j+p\right). \tag{2}$$

where N is even.

In point view of numerical sign, the subtraction of two diagonal symmetric pixels may be positive or negative in (1). Hence, we can obtain the positive and negative DSP units from the N × N block. As $I\left(i+d,j+p\right)=0$ contributes nothing to the illumination invariant measure, we divide $I\left(i+d,j+p\right)$ ($p=0,\cdots,N-1$ and $d=0,\cdots,\mathrm{N}/2-1$) into positive DSP units and negative DSP units in the N × N block, where $I_{+}\left(i+d,j+p\right)>0$ and $I_{-}\left(i+d,j+p\right)<0$ denote the positive DSP unit and the negative DSP unit respectively. Equation (2) can be re-written as

$$\begin{aligned} \text{IIM-DSP}\left(i,j\right)&=\sum_{d=0}^{\frac{N}{2}-1}\sum_{p=0}^{N-1}I\left(i+d,j+p\right)\\ &=\sum_{d=0}^{\frac{N}{2}-1}\sum_{p=0}^{N-1}I_{+}\left(i+d,j+p\right)\\ &\quad+\sum_{d=0}^{\frac{N}{2}-1}\sum_{p=0}^{N-1}I_{-}\left(i+d,j+p\right). \end{aligned} \tag{3}$$

Equation (3) is developed by combining the positive and negative DSP units in the even × even block. It is possible to generate multi IIM-DSPs by different proportions of positive and negative DSP units, which results in the DSP image as below

$$\begin{aligned} \text{DSP}\left(i,j\right)&=\alpha\sum_{d=0}^{\frac{N}{2}-1}\sum_{p=0}^{N-1}I_{+}\left(i+d,j+p\right)\\ &\quad+(2-\alpha)\sum_{d=0}^{\frac{N}{2}-1}\sum_{p=0}^{N-1}I_{-}\left(i+d,j+p\right). \end{aligned} \tag{4}$$

where α is the weight to control the balance of the positive and negative DSP units. When $\alpha = 1$, the DSP image in (4) is equal to the IIM-DSP in (3). The DSP face (DSP-face) is obtained by the DSP image with the arctangent function, which is presented as

$$\begin{aligned}&\text{DSP-face}(i,j)\\&= arctan\left(4\left(\alpha\sum_{d=0}^{\frac{N}{2}-1}\sum_{p=0}^{N-1} I_{+}(i+d,j+p)\right.\right.\\&\left.\left.+(2-\alpha)\sum_{d=0}^{\frac{N}{2}-1}\sum_{p=0}^{N-1} I_{-}(i+d,j+p)\right)\right). \quad (5)\end{aligned}$$

where the parameter 4 is the same as [12] recommended.

Some DSP2 images and DSP2-faces are shown in Fig.3. Some DSP4 images and DSP4-faces are shown in Fig.4, where Fig.3 and Fig.4 share the same original images. It can be seen that the DSP images vary from dark to bright with increase in value of α. Compared with [6], [10] and [12], the DSP image and the DSP-face are quite different from previous illumination invariant measures.

From Fig.3 and Fig.4, some detail features of the DSP2 image and the DSP4 image are different. As the 4×4 block covers larger face area than the 2×2 block, the DSP4 image can extract large scale features of the face image, whereas the DSP2 image extracts small scale features of the face image. The same conclusion can also be done for the DSP2-face and the DSP4-face.

Accordingly, the proposed DSP model can easily extend to the 6×6 block. Based on our test, the performance of DSP6-face lags behind the performances of DSP2-face and DSP4-face. The reason can be explained as that the 6×6 block is larger than the 4×4 block and the 2×2 block, the pixel pairs of the edge of the 6×6 block are with large distances, which may violate the assumption that illumination intensities of neighborhood pixels are approximately equal in the face local region.

IV. THE CLASSIFICATION ALGORITHM

A. SINGLE DSP IMAGE CLASSIFICATION

From [6], [10], [12] and [24], the illumination invariant measure with the saturation function is more efficient than the illumination invariant measure without the saturation function for the single measure image recognition under the template matching classification method such as the nearest neighbor classifier, since the high-frequency interference seriously impacts the recognition of the illumination invariant measure under the template matching classification method, whereas the high-frequency interference can be well tackled by the saturation function.

Hence, the illumination invariant measure with the saturation function (i.e. DSP2-face and DSP4-face) are employed to address the single DSP image recognition under the nearest neighbor classifier, and the parameter $\alpha = 0.4$ in (5) is adopted, which is the same as [6] recommended for severe illumination variations.

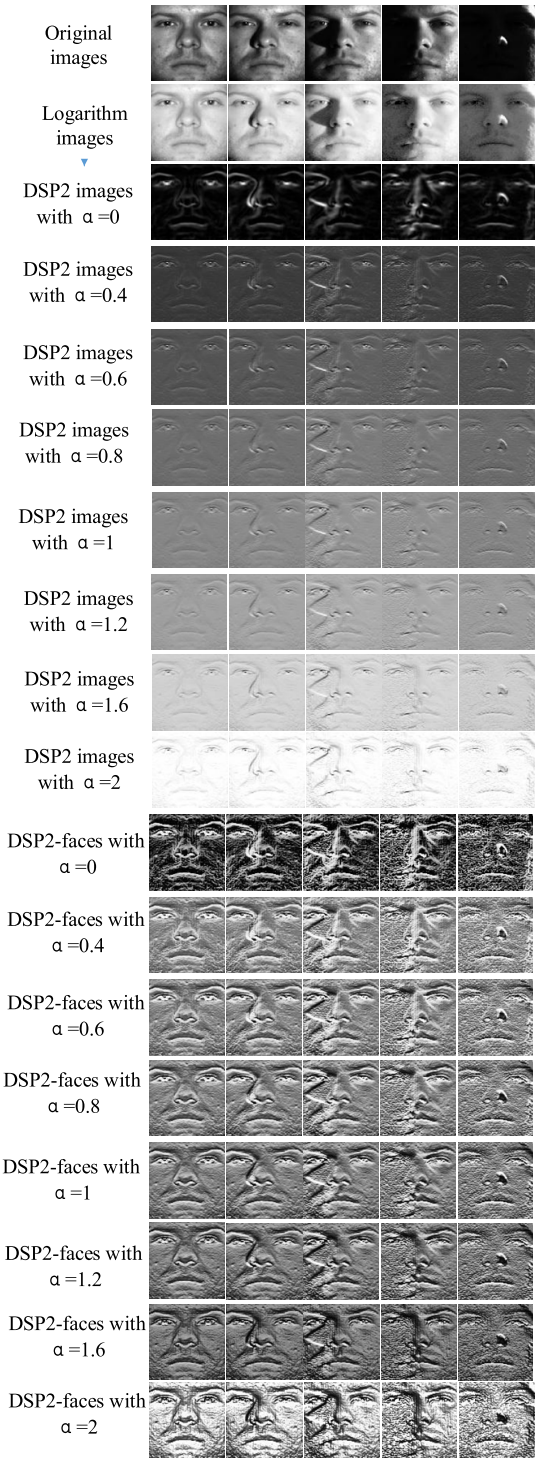


FIGURE 3. Some DSP2 images and DSP2-faces with different parameters.

B. MULTI DSP IMAGES CLASSIFICATION

Although the template matching classification method is sensitive to noise such as the high-frequency interference, the sparse representation classification (SRC) [25] is robust to noise. In many practical applications, such as the driver face recognition in the intelligent transportation systems [6],

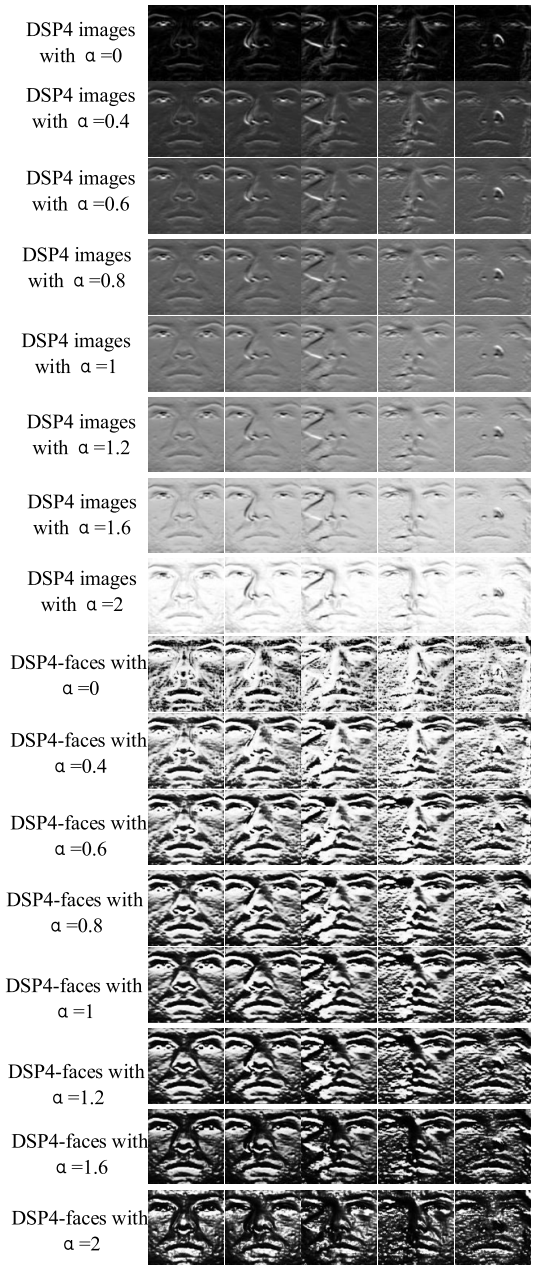


FIGURE 4. Some DSP4 images and DSP4-faces with different parameters.

severe illumination variations and single sample problem coexist. Similar with [6, Formulas (8) and (9)], the extended sparse representation classification (ESRC) [26] under multi DSP images (i.e. DSP2 images or DSP4 images) can be employed to tackle severe illumination variation face recognition with single sample problem.

For ESRC with multi DSP images, the DSP model is used to generate multi training DSP images of the single training sample by different parameter α. Multi training DSP images contain more intra class variations of the single training sample as shown in Fig.3 and Fig.4, which can improve the representation ability of ESRC. In this paper, we select three DSP images with $\alpha = 0.4$, 1, and 1.6 to form multi training DSP images of each single training sample, which is the same as recommended by [6]. Accordingly, the DSP image of the test image is generated by $\alpha = 1$. As multi images of the generic subject can produce enough face intra class variation information, it is unnecessary to generate multi DSP images for the generic image. The DSP image of each generic image is generated by $\alpha = 1$.

In this paper, ESRC with multi DSP2 images is termed as the DSP2 images based classification (DSP2C). ESRC with multi DSP4 images is termed as the DSP4 images based classification (DSP4C). The Homotopy method [27] is used to solve the L1-minimization problem in DSP2C and DSP4C.

C. MULTI DSP IMAGES AND THE PRE-TRAINED DEEP LEARNING MODEL BASED CLASSIFICATION

Similar with [6, Formulas (10) and (11)], the proposed DSP model can be integrated with the pre-trained deep learning model. ESRC [26] can be used to classify the state-of-the-art deep learning feature. The representation residual of DSP2C or DSP4C can be integrated with the representation residual of ESRC of the deep learning feature to conduct classification, which is termed as multi DSP images and the pre-trained deep learning model based classification (DSP-PDL).

In this paper, the pre-trained deep learning models VGG [14] and ArcFace [17] are adopted. Multi DSP2 images and VGG (or ArcFace) based classification is briefly termed as DSP2-VGG (or DSP2-ArcFace), and multi DSP4 images and VGG (or ArcFace) based classification is briefly termed as DSP4-VGG (or DSP4-ArcFace).

V. EXPERIMENTS

A. FACE DATABASES

This paper proposes the DSP model to tackle severe illumination variations. The performances of the proposed methods are validated on the Extend Yale B [28], CMU PIE [29], AR [30] and VGGFace2 test [31] face databases.

The Extended Yale B face database [28] contains grayscale images of 28 persons. 64 frontal face images of each person are divided into subsets 1-5 with illumination variations from slight to severe. Subsets 1-5 consist of 7,12,12,14 and 19 images per person respectively. As the deep learning feature requires the color face image, three RGB channels use the same grayscale image for the experiments on Extended Yale B.

The CMU PIE [29] face database incorporates color images of 68 persons. 21 images of each person from each of C27 (frontal camera), C29 (horizontal 22.5° camera) and C09 (above camera) in CMU PIE illum set are selected. CMU PIE face images are with slight/moderate/severe illumination variations. From [29], pose variation of C29 is larger than that of C09.

The AR database [30] consists of color images of 126 persons in two sessions. 100 persons (50 males and 50 females)

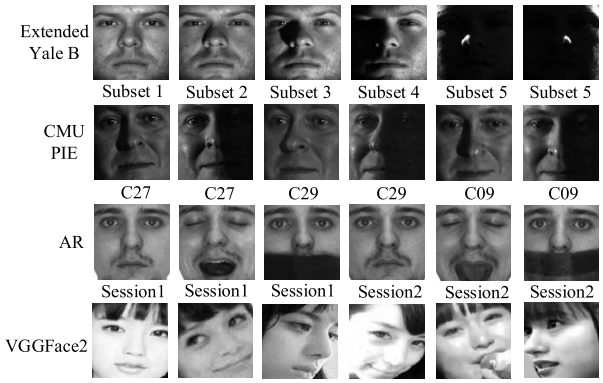


FIGURE 5. Some images from Extended Yale B, CMU PIE, AR and VGGFace2 face databases.

in session 1 and session 2 are selected, and 10 images of each person are selected, which include variations of expression (neutral, smile, anger and scream), illumination (left light, right light and all side lights) and scarf.

The VGGFace2 test database [31] incorporates color images of 500 persons, which are with large variations in pose, age, illumination, ethnicity and profession. The VGGFace2 test database is a large scale test set with more than 160,000 images.

Extended Yale B, CMU PIE and AR are with small size in comparison with the large scale face database VGGFace2, whereas the three face databases are with benchmark illumination variations, which are commonly used to validate the illumination invariant measures.

In this paper, all cropped face images and the experimental setting are the same as [6]. The recognition rates of Tables 1, 2 3 and 4 are from [6, Tables 3, 4, 5 and 7] except for the proposed methods and RRHE [23]. Fig.5 shows some cropped images from our employed face databases.

B. COMPARED METHODS

1) THE PROPOSED METHOD

DSP2-face, DSP4-face, DSP2C, DSP4C, DSP2-VGG/ ArcFace and DSP4-VGG/ArcFace. Three DSP images (i.e. $\alpha = 0.4$, 1, and 1.6) are generated for DSP2C, DSP4C, DSP2-VGG/ArcFace and DSP4-VGG/ArcFace.

2) THE HIGH-FREQUENCY FACIAL FEATURE AND THE LOCAL BINARY PATTERN DESCRIPTOR

HFSVD-face [10] and CSQP [19]. The parameters are the same as [10] and [19] recommended.

3) THE ILLUMINATION INVARIANT MEASURE

Weber-face [11], MSLDE [12], LNN-face [10], EGIR-face [6], BGIR-face [6], MSLDE+ESRC, LNN-face+ESRC, RRHE [23], EGIRC [6] and BGIRC [6]. LNN-face+ESRC represents that LNN-faces of the face images are classified by ESRC, and the same interpretation can also be done for MSLDE+ESRC. EGIRC and BGIRC [6] employed three GIR images of each single training sample.

4) THE PRE-TRAINED DEEP LEARNING MODEL

VGG [14] and ArcFace [17], VGG/ArcFace+ESRC, EGIR-VGG/ArcFace and BGIR-VGG/ArcFace [6]. The 4096D VGG feature and the 512D ArcFace feature are used. VGG/ArcFace+ESRC has the same interpretation as LNN-face+ESRC. EGIR-VGG/ArcFace represents the integration of the EGIRC and the ESRC based VGG/ArcFace, and the same interpretation can also be done for BGIR-VGG/ArcFace.

5) ORIGINAL AND LOG

Original and LOG represent the pixel image without any processing and the logarithm image, which are directly used as facial features for recognition.

6) THE SOURCE CODE LOCATION

The code of Weber-face [11] was downloaded at http://luks.fe.uni-lj.si/sl/osebje/vit-omir/face_tools/INFace/index.html. The code of VGG [14] was downloaded at http://www.robots.ox.ac.uk/_vgg/soft- ware/vgg_face/. The code of ArcFace [17] was downloaded at https://github.com/deepinsight/insightface, and the third party pre-trained model model-r100-ii was adopted and downloaded at https://pan.baidu.com/s/1wuRT f2YIsKt76-TxFufsRNA, which was trained by MS1MV2 (85742 persons and 5.8M images). The code of Homotopy [27] was downloaded at http://www.eecs.berkeley.edu/_yang/soft-ware/l1benchmark/, where the error tolerance $\varepsilon = 0$ is used.

The compared methods (Original, LOG, HFSVD-face, Weber-face, CSQP, MSLDE, LNN-face, RRHE, EGIR-face, BGIR-face, VGG, ArcFace, DSP2-face and DSP4-face) employ the nearest neighbor (NN) classifier with Euclidean distance for the classification. HFSVD-face, Weber-face, MSLDE, LNN-face, RRHE, EGIR-face, BGIR-face, DSP2-face and DSP4-face are termed as the shallow illumination invariant approaches in this paper.

C. EXPERIMENT RESULTS

Tables 1, 2 and 3 list average recognition rates of the compared methods on Extended Yale B, CMU PIE and AR datasets. Table 4 lists recognition rates of some compared methods on VGGFace2 test database.

1) EXTENDED YALE B

The Extended Yale B face database is with extremely challenging illumination variations. Face images in subsets 1-3 are with slight/moderate illumination variations, where subsets 1-2 face images are with slight illumination variations, and subset 3 face images are with small scale cast shadows. Face images in subsets 4-5 are with severe illumination variations, where subset 4 face images are with moderate scale cast shadows, and subset 5 face images are with large scale cast shadows (or severe holistic illumination variations).

From Table 1, we can conclude some important results as below.

TABLE 1. The average recognition rates (%) of the compared methods on the Extended Yale B face database.

Method	Subsets 1-3	Subset4	Subset5	Subsets 4-5	Total
Original	48.52	19.63	15.79	14.96	20.21
LOG	49.58	32.55	32.51	26.41	22.39
HFSVD-face [10]	94.92	83.02	97.97	90.18	86.49
Weber-face [11]	87.07	58.66	92.52	77.66	74.21
CSQP [16]	83.13	59.04	87.67	74.55	65.53
MSLDE [12]	81.30	53.35	81.45	66.79	60.27
LNN-face [10]	84.83	61.59	92.02	77.98	70.32
RRHE [23]	91.89	**91.85**	97.12	94.50	91.76
EGIR-face [6]	77.20	61.69	88.12	74.54	66.74
BGIR-face [6]	77.99	70.15	93.27	82.17	72.75
DSP2-face	88.05	74.40	96.17	87.53	84.75
DSP4-face	84.01	62.97	94.84	80.41	77.01
VGG [17]	86.31	47.14	27.67	30.90	45.32
ArcFace [20]	85.56	53.28	30.93	35.49	49.71
MSLDE+ESRC	90.19	66.41	92.17	80.38	75.60
LNN+ESRC	92.55	76.08	97.11	88.36	82.70
VGG+ESRC	94.19	61.90	40.60	43.58	57.75
ArcFace+ESRC	91.35	58.55	36.16	41.78	55.95
EGIRC [6]	95.88	75.62	96.31	86.84	83.59
BGIRC [6]	96.31	78.53	97.30	89.27	86.69
DSP2C	90.58	90.79	**97.74**	**94.90**	**92.13**
DSP4C	92.90	78.31	97.30	89.55	88.66
EGIR-VGG	98.33	81.79	84.30	80.69	82.28
BGIR-VGG	**98.42**	82.45	82.84	80.19	83.53
EGIR-ArcFace	97.92	79.49	83.13	79.19	78.24
BGIR- ArcFace	97.95	79.19	81.73	78.30	78.97
DSP2-VGG	97.63	82.40	80.48	78.27	86.22
DSP4-VGG	97.52	81.58	88.52	83.13	87.49
DSP2-ArcFace	96.32	77.78	79.93	75.42	80.75
DSP4-ArcFace	97.17	79.71	88.95	82.21	84.53

1) DSP2-face and DSP4-face outperform EGIR-face, BGIR-face, MSLDE and LNN-face on all Extended Yale B datasets, except that DSP4-face lags behind BGIR-face on subset 4 and subsets 4-5. Although moderate scale cast shadows of subset 4 images are not as severe as large scale cast shadows of subset 5 images, moderate scale cast shadows incorporate more edges of cast shadows than large scale cast shadows as shown in Fig.5. Edges of cast shadows of face images violate the assumption of the illumination invariant measure that illumination intensities are approximately equal in the face local region.
2) DSP2C and DSP4C perform better than DSP2-face and DSP4-face. There may be two main reasons, one is that multi DSP images contain more intra class variation information than the single DSP image, and the other one is that ESRC is more robust than NN under illumination variations. DSP2C outperforms DSP4C, EGIRC and BGIRC under severe illumination variations such as on all Extended Yale B datasets except on subsets 1-3, where DSP2C lags behind DSP4C, EGIRC and BGIRC. Moreover, DSP2C performs much better than DSP4C, EGIRC and BGIRC on subset 4. The reason may be that DSP2C uses the smaller 2×2 block region in comparison with DSP4C, EGIRC and BGIRC, which means that less noises are introduced into the illumination invariant measure.
3) VGG/ArcFace was trained by large scale light internet face images, without considering severe illumination variations, which performs well on subsets 1-3, but unsatisfactorily under severe illumination variations such as on subsets 4-5. Despite DSP2-VGG/ArcFace and DSP4-VGG/ArcFace are not able to attain the highest recognition rates on all datasets, DSP2-VGG/ArcFace and DSP4-VGG/ArcFace achieve very high recognition rates on all Extended Yale B datasets, and outperform VGG/ArcFace and VGG/ArcFace+ESRC. Hence, the DSP-PDL model can have the advantages of both the DSP model and the pre-trained deep learning model to tackle face recognition.

TABLE 2. The average recognition rates (%) of the compared methods on the CMU PIE face database.

Method	C27	C29	C09	C27+C29	C27+C09
Original	30.31	30.17	27.52	20.88	19.97
LOG	31.19	30.06	27.04	20.18	19.66
HFSVD-face [10]	94.50	87.30	91.71	52.82	51.21
Weber-face [11]	89.17	84.00	89.17	49.46	46.42
CSQP [16]	86.36	82.46	83.21	51.97	49.81
MSLDE [12]	81.01	77.57	80.04	46.89	48.41
LNN-face [10]	89.26	84.67	88.29	50.29	51.32
RRHE [23]	99.42	97.55	98.98	53.86	50.74
EGIR-face [6]	82.12	83.50	83.33	47.75	47.66
BGIR-face [6]	89.30	89.25	89.72	50.06	49.26
DSP2-face	99.37	**99.16**	99.09	55.88	50.40
DSP4-face	98.02	97.46	97.76	56.14	51.97
VGG [17]	87.33	76.91	86.67	79.78	83.69
ArcFace [20]	91.90	78.02	97.51	79.57	86.62
MSLDE+ESRC	91.68	88.46	90.46	57.05	58.35
LNN+ESRC	95.09	91.85	94.70	57.08	59.38
VGG+ESRC	95.73	89.02	94.90	91.70	94.00
ArcFace+ESRC	94.89	81.40	97.85	83.48	89.32
EGIR-VGG	98.88	95.48	98.52	93.95	94.35
BGIR-VGG	99.08	95.91	98.88	94.40	95.06
EGIR-ArcFace	98.40	93.38	99.07	88.65	89.17
BGIR-ArcFace	98.66	93.92	99.37	88.92	89.94
DSP2-VGG	**99.49**	97.69	99.48	**94.87**	**95.90**
DSP4-VGG	99.48	98.13	99.36	94.70	94.71
DSP2-ArcFace	99.36	96.70	**99.87**	88.90	90.74
DSP4-ArcFace	99.45	97.60	99.79	89.48	89.11

2) CMU PIE

Some CMU PIE face images are bright (i.e. slight illumination variations), and other face images are with partial dark (i.e. moderate/severe illumination variations). Illumination variations of CMU PIE are not as extreme as those

TABLE 3. The average recognition rates (%) of the compared methods on the AR face database.

Method	AR1	AR2	AR1+AR2
Original	14.11	13.21	13.82
LOG	19.17	18.08	17.14
HFSVD-face [10]	63.76	58.49	53.28
Weber-face [11]	49.15	47.43	42.17
CSQP [16]	50.19	47.67	43.81
MSLDE [12]	45.30	43.21	38.99
LNN-face [10]	50.53	48.76	44.23
RRHE [23]	56.78	51.29	47.60
EGIR-face [6]	44.41	42.11	37.02
BGIR-face [6]	44.35	42.21	37.08
VGG [17]	60.58	59.71	58.61
ArcFace [20]	41.41	40.37	39.10
DSP2-face	47.77	44.66	39.94
DSP4-face	41.98	42.04	36.77

of Extended Yale B. From Fig.5, images in each of C27, C29 and C09 are with the same pose (i.e. frontal, 22.5° profile and downward respectively), whereas images in each of C27+C29 and C27+C09 incorporate two face poses (i.e. frontal pose and non-frontal pose).

From Table 2 , DSP2-face and DSP4-face achieve very high recognition rates on C27, C29 and C09, and perform much better than EGIR-face, BGIR-face, MSLDE and LNN-face on all CMU PIE datasets. However, DSP2-face and DSP4-face lag behind VGG/ArcFace on C27+C29 and C27+C09. It can be seen that DSP2-face and DSP4-face are very robust to severe illumination variations under fixed pose, whereas sensitive to pose variations. Although DSP2-face and DSP4-face outperform EGIR-face, BGIR-face, MSLDE and LNN-face under pose variations, the shallow illumination invariant approaches perform unsatisfactorily to pose variations.

However, DSP2-VGG/ArcFace and DSP4-VGG/ArcFace perform very well on all CMU PIE datasets, which illustrates that the DSP-PDL model can achieve satisfactorily results under both severe illumination variations and pose variations. Hence, the DSP-PDL model is robust to both illumination variations and pose variations. As DSP2-face and DSP4-face are superior to EGIR-face and BGIR-face under severe illumination variations, DSP2-VGG/ArcFace and DSP4-VGG/ArcFace are slightly better than EGIR-VGG/ArcFace and BGIE-VGG/ArcFace on C27, C29 and C09, whereas they achieve similar performances on C27+C29 and C27+C09, since VGG/ArcFace is the dominant feature under pose variations.

3) AR

AR face images are with frontal pose, slight illumination and moderate/severe expression variations as well as scarf occlusion. Illumination variations of AR are not as severe as those of Extended Yale B and CMU PIE.

From Table 3, DSP2-face achieves higher recognition rates than EGIR-face, BGIR-face and MSLDE on all AR datasets, whereas lags behind LNN-face. DSP4-face slightly lags behind other illumination invariant measures. As the illumination variation problem is not very tough for the AR face images, the performance of the DSP model is limited.

From Tables 1 2 and 3, the recent RRHE [23] with the 3×3 block obtains high recognition rates compared with other illumination invariant measures, whereas RRHE is still sensitive to pose variations.

TABLE 4. The recognition rates (%) of some compared methods on the VGGFace2 test set.

MSLDE	LNN-face	EGIR-face	BGIR-face
3.53	3.07	2.93	2.55
CSQP	VGG	ArcFace	ArcFace+ESRC
3.46	28.80	34.84	35.67
EGIRC	BGIRC	EGIRC-VGG	BGIRC-VGG
3.54	3.20	41.98	44.19
DSP2-face	DSP4-face	DSP2-VGG	DSP4-VGG
1.56	3.04	47.27	43.76

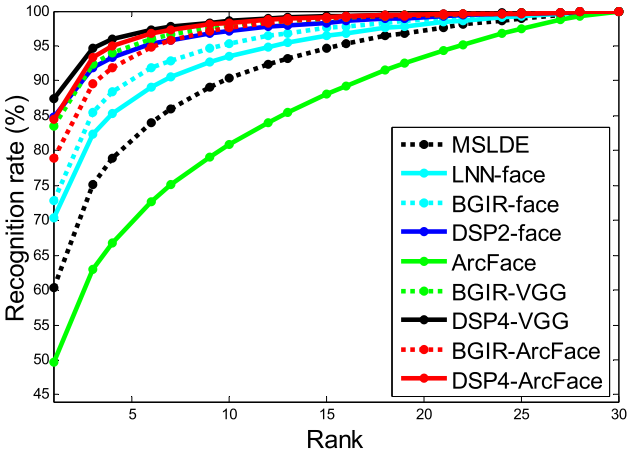


FIGURE 6. CMC of some compared methods on Extended Yale B database.

4) VGGFACE2

VGGFace2 images are composed of large scale bright internet face images with large pose/expression variations, and illumination variations of VGGFace2 are not as severe as those of Extended Yale B and CMU PIE.

From Tables 4, the shallow illumination invariant approaches achieve very low recognition rates compared with the deep learning feature VGG/AracFace. DSP4-face is superior to EGIR-face and BGIR-face, whereas DSP2-face lags behind EGIR-face and BGIR-face. DSP2-VGG and DSP4-VGG achieve high recognition rates, since VGG is the dominant feature of the DSP-PDL model on the VGGFace2 test set.

D. CMC CURVES OF SOME COMPARED METHODS

Fig.6, Fig7, Fig.8 and Fig.9 show the cumulative match characteristic (CMC) curves of some compared methods on Extended Yale B, CMU PIE and VGGFace2 respectively. These CMC curves follow the same experiment protocols as the corresponding datasets in Tables 1, 2 and 4. Recognition

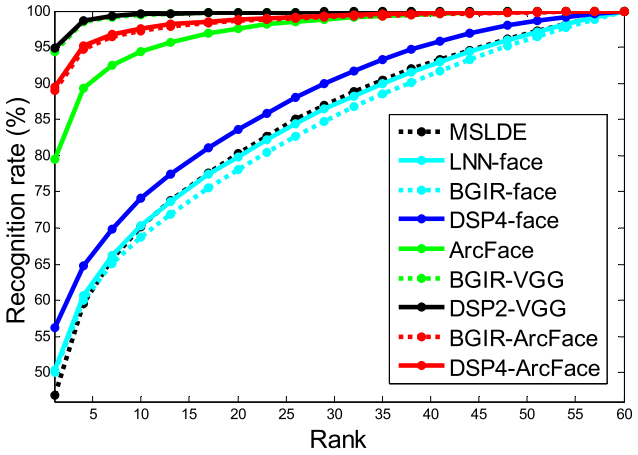


FIGURE 7. CMC of some compared methods on C27+C29 of CMU PIE database.

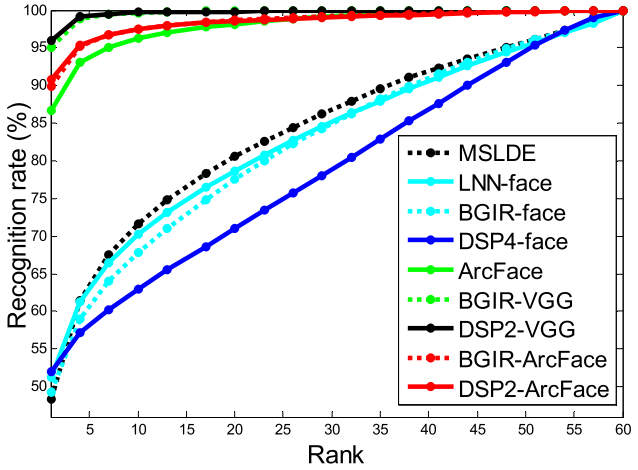


FIGURE 8. CMC of some compared methods on C27+C09 of CMU PIE database.

rates of rank=1 in Fig.6, Fig.7, Fig.8 and Fig.9 are equal to recognition rates of corresponding datasets in Tables 1, 2 and 4 respectively. It can be seen that the proposed methods show consistent improvement in recognition rate with increasing ranks.

E. THE PROPOSED METHODS

In comparison with the data-driven based deep learning methods VGG [14] and ArcFace [17] that required a training processing, the proposed DSP2-face and DSP4-face are the model-driven based illumination invariant measures, which do not depend on large scale face images training, such as previous illumination invariant measures EGIR-face [6], BGIR-face [6], RRHE [23], LNN-face [10] and MSLDE [12]. Moreover, DSP2-face or DSP4-face in (5) employs only one parameter α.

From Tables 1, 2, 3 and 4, on the one hand, DSP2-face is superior to DSP4-face on Extended Yale B datasets, C27, C29 and C09 of CMU PIE, AR datasets. These datasets are with frontal or fixed face pose and illumination variations, which illustrates that DSP2-face outperforms DSP4-face under frontal or fixed face pose with illumination variations.

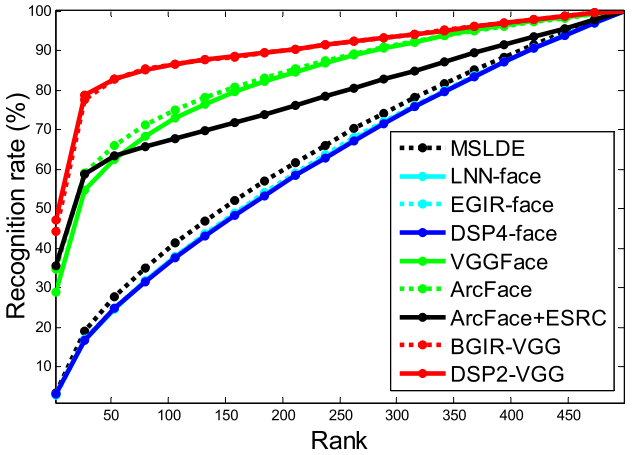


FIGURE 9. CMC of some compared methods on VGGFace2 test set.

On the other hand, DSP2-face lags behind DSP4-face on C27+C29 and C27+C09 of CMU PIE, and VGGFaces2 test set. These datasets are with pose variations, which indicates that DSP4-face outperforms DSP2-face under both pose and illumination variations.

From experimental results of Extended Yale B and CMU PIE datasets, DSP2-face and DSP4-face achieve higher recognition rates than EGIR-face and BGIR-face except on subsets 1-3 of Extended Yale B, which illustrates that DSP2-face and DSP4-face are superior to EGIR-face and BGIR-face under severe illumination variations. However, DSP2-face and DSP4-face lags behind EGIR-face and BGIR-face under slight illumination variations, which can be seen from the experimental results of Extended Yale B subsets 1-3 and VGGFace2 test set.

DSP2-VGG/ArcFace and DSP4-VGG/ArcFace achieve very high recognition rates on all datasets of Extended Yale B, CMU PIE and VGGFace2 test, except on Extended Yale B subset 5 with extremely severe illumination variations, since the pre-trained deep learning model is restricted to frontal face images with severe illumination variations, whereas this is insufficient to deny that DSP2-VGG/ArcFace and DSP4-VGG/ArcFace are the best approaches to tackle severe illumination variations. Hence, the DSP-PDL model is able to have the advantages of both the DSP model (i.e. robust to severe illumination variations) and the pre-trained deep learning model (i.e. robust to slight illumination variations and large pose variations) to tackle face recognition.

F. COMPUTATIONAL COMPLEXITY

Similar to LNN-face [10], DSP2-face or DSP4-face consists of four steps: 1) image logarithm transformation, 2) Gaussian smooth, 3) DSP2 image or DSP4 image generation, and 4) arctangent calculation. For the image with size of $m \times n$, the computational complexity of DSP2 image or DSP4 image generation is $O(mn)$. Both the logarithm transformation and the arctangent calculation are factorized to the polynomial sum with limited accuracy, and the computation complexity of each of them is $O(mn)$. The computation complexity of

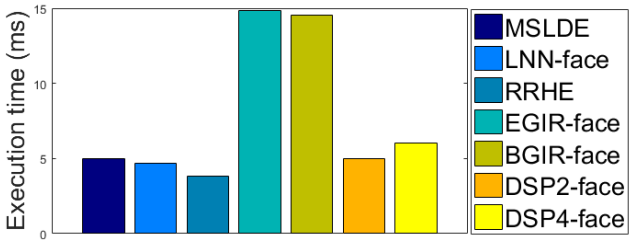


FIGURE 10. Average execution time of some compared methods.

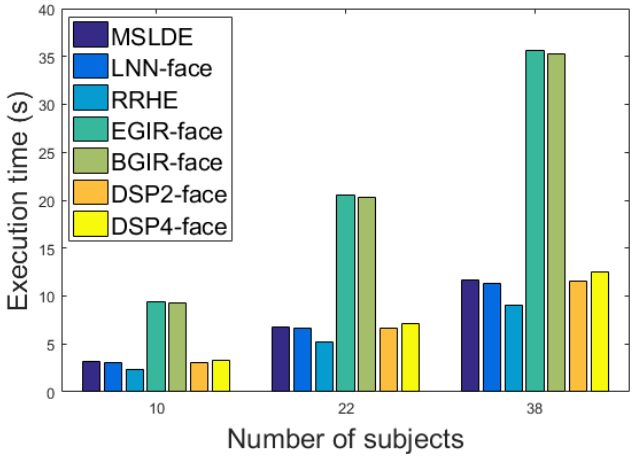


FIGURE 11. Execution time of some compared methods under different number of subjects.

Gaussian smooth is $O(mn)$ from [10]. Thus, the computational complexity of DSP2-face or DSP4-face is $O(mn)$.

All 2432 images of Extended Yale B are selected to test. MSLDE [12], LNN-face [10], RRHE [23], EGIR-face [6] and BGIR-face [6] extract facial features of 2432 images with the size of 50×50, due to that these methods use images with the size of 50×50 in above experiments. The tests are conducted on the PC with Core (TM) i7-9700F 3.00GHz processor and 8.00GB RAM using MATLAB R2016a.

Fig.10 lists average execution time (millisecond) of some compared methods, where the average execution times of MSLDE, LNN-face, RRHE, EGIR-face, BGIR-face, DSP2-face and DSP4-face are 4.98ms, 4.68ms, 3.84ms, 14.82ms, 14.57ms, 4.97ms and 6.02ms respectively. As the max local regions of EGIR-face and BGIR-face are 6×6 and 5×5 respectively, the former consumes more time than the latter. The same conclusion can also be done for MSLDE and LNN-face. MSLDE and LNN-face did not require to calculate positive and negative illumination invariant measures, which results in that MSLDE and LNN-face consume less time than EGIR-face and BGIR-face. Although DSP2-face and DSP4-face have to calculate positive and negative illumination invariant measures, they employ the 2×2 block and the 4×4 block respectively. Thus DSP2-face and DSP4-face consumes less time than EGIR-face and BGIR-face. RRHE with the 3×3 block is the fastest algorithm in Fig.10, since RRHE does not conduct logarithm transformation, Gaussian smoothing and arctangent calculation. Fig.11 lists execution time (second) of some compared methods under different number of subjects. It can be seen from Fig.11 that execution time increases linearly with increase in number of subjects.

VI. CONCLUSION

This paper proposes the DSP model to address severe illumination variation face recognition. DSP2-face and DSP4-face achieve higher recognition rates compared with previous illumination invariant approaches EGIR-face, BGIR-face, LNN-face and MSLDE under severe illumination variations. DSP2C and DSP4C are efficient to severe illumination variations, due to the fact that multi DSP images cover more discriminative information of the face image. Further, the proposed DSP model is integrated with the pre-trained deep learning model to have the advantages of both the DSP model and the pre-trained deep learning model.

Although the proposed DSP model can efficiently tackle severe illumination variations, the authors realized that further work must be done to improve the proposed model under pose variations and well illumination.

REFERENCES

[1] C. Zou, K. I. Kou, L. Dong, X. Zheng, and Y. Y. Tang, "From grayscale to color: Quaternion linear regression for color face recognition," *IEEE Access*, vol. 7, pp. 154131–154140, Oct. 2019.

[2] Y. Chen, Y. Zhao, Y. He, F. Xu, W. Jia, J. Lian, and Y. Zheng, "Face identification with top-push constrained generalized low-rank approximation of matrices," *IEEE Access*, vol. 7, pp. 160998–161007, Oct. 2019.

[3] M. Hu, C. Yang, Y. Zheng, X. Wang, L. He, and F. Ren, "Facial expression recognition based on fusion features of center-symmetric local signal magnitude pattern," *IEEE Access*, vol. 7, pp. 118435–118445, Jul. 2019.

[4] W. Zhang, X. Zhao, J.-M. Morvan, and L. Chen, "Improving shadow suppression for illumination robust face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 3, pp. 611–624, Mar. 2019.

[5] C.-H. Hu, X.-B. Lu, P. Liu, X.-Y. Jing, and D. Yue, "Single sample face recognition under varying illumination via QRCP decomposition," *IEEE Trans. Image Process.*, vol. 28, no. 5, pp. 2624–2638, May 2019.

[6] C.-H. Hu, Y. Zhang, F. Wu, X.-B. Lu, P. Liu, and X.-Y. Jing, "Toward driver face recognition in the intelligent traffic monitoring systems," *IEEE Trans. Intell. Transp. Syst.*, early access, Oct. 14, 2019, doi: 10.1109/TITS.2019.2945923.

[7] J.-W. Wang, N. T. Le, J.-S. Lee, and C.-C. Wang, "Color face image enhancement using adaptive singular value decomposition in Fourier domain for face recognition," *Pattern Recognit.*, vol. 57, pp. 31–49, Sep. 2016.

[8] X. Fu, D. Zeng, Y. Huang, X.-P. Zhang, and X. Ding, "A weighted variational model for simultaneous reflectance and illumination estimation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2782–2790.

[9] T. Chen, W. Yin, X. Sean Zhou, D. Comaniciu, and T. S. Huang, "Total variation models for variable lighting face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 9, pp. 1519–1524, Sep. 2006.

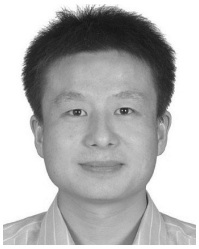
[10] C. Hu, X. Lu, M. Ye, and W. Zeng, "Singular value decomposition and local near neighbors for face recognition under varying illumination," *Pattern Recognit.*, vol. 64, pp. 60–83, Apr. 2017.

[11] B. Wang, W. Li, W. Yang, and Q. Liao, "Illumination normalization based on Weber's law with application to face recognition," *IEEE Signal Process. Lett.*, vol. 18, no. 8, pp. 462–465, Aug. 2011.

[12] Z.-R. Lai, D.-Q. Dai, C.-X. Ren, and K.-K. Huang, "Multiscale logarithm difference edgemaps for face recognition against varying lighting conditions," *IEEE Trans. Image Process.*, vol. 24, no. 6, pp. 1735–1747, Jun. 2015.

[13] B. K. P. Horn, *Robot Vision*. Cambridge, MA, USA: MIT Press, 1997.

[14] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in *Proc. Brit. Mach. Vis. Conf.*, Sep. 2015, pp. 1–12.
[15] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 815–823.
[16] H. Wang, Y. Wang, Z. Zhou, X. Ji, D. Gong, J. Zhou, Z. Li, and W. Liu, "CosFace: Large margin cosine loss for deep face recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 5265–5274.
[17] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "ArcFace: Additive angular margin loss for deep face recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 4690–4699.
[18] M. Heikkilä, M. Pietikäinen, and C. Schmid, "Description of interest regions with local binary patterns," *Pattern Recognit.*, vol. 42, no. 3, pp. 425–436, Mar. 2009.
[19] S. Chakraborty, S. K. Singh, and P. Chakraborty, "Centre symmetric quadruple pattern: A novel descriptor for facial image recognition and retrieval," *Pattern Recognit. Lett.*, vol. 115, pp. 50–58, Nov. 2018.
[20] Y. El Merabet, Y. Ruichek, and A. El Idrissi, "Attractive-and-repulsive center-symmetric local binary patterns for texture classification," *Eng. Appl. Artif. Intell.*, vol. 78, pp. 158–172, Feb. 2019.
[21] S. Chakraborty, S. K. Singh, and P. Chakraborty, "Cascaded asymmetric local pattern: A novel descriptor for unconstrained facial image recognition and retrieval," *Multimedia Tools Appl.*, vol. 78, no. 17, pp. 25143–25162, Sep. 2019.
[22] T. Tuncer and S. Dogan, "Improving facial image recognition based neutrosophy and DWT using fully center symmetric dual cross pattern," *Int. J. Image, Graph. Signal Process.*, vol. 11, no. 6, pp. 35–44, Jun. 2019.
[23] J. Yadav, N. Rajpal, and R. Mehta, "An improved illumination normalization and robust feature extraction technique for face recognition under varying illuminations," *Arabian J. Sci. Eng.*, vol. 44, no. 11, pp. 9067–9086, Nov. 2019.
[24] T. Zhang, Y. Yan Tang, B. Fang, Z. Shang, and X. Liu, "Face recognition under varying illumination using gradientfaces," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2599–2606, Nov. 2009.
[25] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.
[26] W. Deng, J. Hu, and J. Guo, "Extended SRC: Undersampled face recognition via intraclass variant dictionary," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 9, pp. 1864–1870, Sep. 2012.
[27] D. L. Donoho and Y. Tsaig, "Fast solution of ℓ_1-norm minimization problems when the solution may be sparse," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 4789–4812, Nov. 2008.
[28] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 643–660, Jun. 2001.
[29] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression database," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 12, pp. 1615–1618, Dec. 2003.
[30] A. M. Martinez and R. Benavente, "The AR face database," CVC, New Delhi, India, Tech. Rep. #24, Jun. 1998.
[31] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman, "VGGFace2: A dataset for recognising faces across pose and age," in *Proc. 13th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, May 2018, pp. 67–74.



CHANGHUI HU received the Ph.D. degree from the School of Automation, Southeast University, Nanjing, China, in 2017. He is currently a Lecturer with the School of Automation, Nanjing University of Posts and Telecommunications. He also holds a postdoctoral position with the School of Transportation, Southeast University. His research interests include image processing and pattern recognition. He has received many research awards, such as the Best Doctoral Dissertation Award from the China Intelligent Transportation Systems Association, and the Prize from the Science and Technology Award of Jiangsu Province.

FEI WU received the Ph.D. degree in computer science from the Nanjing University of Posts and Telecommunications, China, in 2016. He is currently with the College of Automation, Nanjing University of Posts and Telecommunications. He has authored over 30 scientific articles. His research interests include pattern recognition, artificial intelligence, and computer vision.

JIAN YU received the B.E. degree from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2017, where he is currently pursuing the master's degree with the College of Automation and College of Artificial Intelligence. His research interests include pattern recognition, machine learning, and large-scale image analysis.

XIAOYUAN JING received the Ph.D. degree in pattern recognition and intelligent system from the Nanjing University of Science and Technology, in 1998. He was a Professor with the Department of Computer, Shenzhen Research Student School, Harbin Institute of Technology, in 2005. He is currently a Professor with the College of Automation and College of Artificial Intelligence, Nanjing University of Posts and Telecommunications, and with the School of Computer, Wuhan University, China. He has published over 100 scientific articles on the international journals and conferences such as TPAMI, TIP, TIFS, TSMCB, TMM, TCSVT, TCB, CVPR, AAAI, IJCAI, ACMMM, and so on.

XIAOBO LU received the Ph.D. degree from the Nanjing University of Aeronautics and Astronautics. He did his Postdoctoral Research with the Chien-Shiung Wu Laboratory, Southeast University, from 1998 to 2000. He is currently a Professor with the School of Automation and the Deputy Director with the Detection Technology and Automation Research Institute, Southeast University. He is the coauthor of *An Introduction to the Intelligent Transportation Systems* (Beijing, China Communications, 2008). His research interests include image processing, signal processing, pattern recognition, and computer vision. He has received many research awards, such as the First Prize in Natural Science Award from the Ministry of Education of China and the prize from the Science and Technology Award of Jiangsu Province.

PAN LIU received the Ph.D. degree in civil engineering from the University of South Florida, Tampa, FL, USA, in 2006. He is currently a Professor with the School of Transportation, Southeast University, Nanjing, China. His research interests include traffic operations and safety and intelligent transportation systems. He was a recipient of the Distinguished Young Scientist Foundation of NSFC, in 2013, and the Outstanding Young Scientist Foundation of NSFC, in 2019.