Non-Uniform Patch Based Face Recognition via 2D-DWT

Zheng-Hai Huang[a,c], Wen-Juan Li[a,*], Jin Shang[a,b], Jun Wang[a,b], Ting Zhang[a,c]

[a] *Center for Applied Mathematics of Tianjin University, Tianjin 300072, P.R. China*
[b] *Center for Combinatorics, LPMC-TJKLC, Nankai University, Tianjin 300071, P.R. China*
[c] *Department of Mathematics, School of Science, Tianjin University, Tianjin 300072, P.R. China*

Abstract

In this paper, we propose a method for face recognition by using the two-dimensional discrete wavelet transform (2D-DWT) and a new patch strategy. Based on the average image of all training samples, by using integral projection technique for two top-level's high-frequency sub-bands of 2D-DWT, we propose a non-uniform patch strategy for the top-level's low-frequency sub-band. This patch strategy is more suitable to reflect the structure feature of face image; and it is better for retaining the integrity of local information. By applying the obtained patch strategy to all samples, we obtain patches of training samples and testing samples; and then, give the final decision by using the nearest neighbor classifier and the majority voting. Experiments are run on the AR, FERET, Extended Yale B and LFW face databases. The obtained numerical results show that the new face recognition method outperforms the traditional 2D-DWT method and some state-of-the-art patch based methods.

Keywords: Face recognition, two-dimensional discrete wavelet transform, integral projection, non-uniform patch

*Corresponding author. Email address: liwenjuan@tju.edu.cn (W.J. Li). Tel.: 86-22-27403615, fax: 86-22-27403615.

1. Introduction

The two-dimensional discrete wavelet transform (2D-DWT) is a popular tool in image processing and computer vision due to its nice features of space-frequency localization and multiresolutions [1, 2, 3, 4, 5, 6, 7, 8]. Through 2D-DWT, a face image is transformed into four parts: a low frequency sub-band and three high frequency sub-bands. Generally, the k-level 2D-DWT ($k > 1$) means that the low frequency sub-band obtained in the $(k-1)$-level 2D-DWT is further decomposed into a low frequency sub-band and three high frequency sub-bands. Since the low frequency sub-band plays a dominant role in four sub-bands for approximation of the original image, it is usually solely used for face recognition [9]. However, as the high frequency sub-bands also contain some important information, the method in [10] introduces a joint of pixel-level and feature-level fusion at the top-level's wavelet sub-bands. In this paper, we try to show another use of the high frequency sub-bands.

In recent years, the strategy of patch has been adopted in many face recognition methods. A modular eigenspace description technique is used to incorporate salient features such as the eyes, nose and mouth, in an eigenfeature layer in 1994 [11], which is the first patch based method. In the subsequent earlier methods, the whole image is partitioned into pathes for feature extraction [11, 12, 13, 14, 15]. These methods show better performance when confronted with some problem such as illumination and the singular in linear discriminant analysis (LDA), to some extent, which draw the attention upon an image from global to local. By combining patch technique with some other methods, many effective face recognition methods have been proposed recently in the literature [16, 17, 18, 19, 20, 21, 22, 23, 24]. From the above papers, it is not difficult to find that there are two basic kinds of patches: one is overlapped patches as in [16, 17]; and another is non-overlapped patches as in [22, 23]. Among these papers, [16] investigates the face recognition problem via the overlapping energy histogram of the DCT coefficients, and it is an earlier paper which adopts overlapped patch technique for face recognition; in [17], the method of collaborative representation based classification with patch, denoted by PCRC, is proposed to classify the query sample; and in [22, 23], each face image is divided into smaller regions of the same size and the two-dimensional principal component analysis is used in each region. Generally, the overlapped patches is more robust. In all of these methods, images are partitioned uniformly; and hence, some salient features (like eyes and nose) may be set apart. A natural question is *how can face images be*

partitioned according to its own geometrical structure automatically? We will give part of the answer to this question in this paper.

In order to obtain better patches of a face image, we will use integral projection technique. It is known that integral projection is a method for extraction of facial features, which is first proposed by Kanade [25]. It is well known that integral projections are used to detect salient features and locate boundaries of local regions; and they are usually implemented on frontal and standard faces [25, 26, 27]. In this paper, we use the idea of integral projection to propose a new patch strategy which is suitable to many different conditions including different illumination, expressions and some small changes in pose.

In addition, in patch based methods, an important issue is how can the classification outputs of all patches be combined to classify the query sample. There are many methods to do it, such as majority voting [24, 28, 29], probabilistic model [30], linear weighted combination [15], and kernel plurality [31]. The method in [24] describes majority voting clearly, where each face image is represented as a spatial arrangement of image patches, and a smooth non-linear mapping for the corresponding patches are sought using Volterra kernels; then each patch casts a vote towards image classification after been classified independently; finally the class with the maximum votes is chosen. Similarly, in this paper, we will use the idea of majority voting.

Motivated by the methods mentioned above, we propose a new method for face recognition, which is based on 2D-DWT and a non-uniform patch strategy. We denote it by NUPDWT. In NUPDWT, we first propose a patch strategy, whose process is divided into the following five steps. In Step 1, we give the average image of all training samples. In Step 2, 2D-DWT is performed on the average image to gain the top-level's four sub-bands: LL, LH, HL and HH (Note: for a k-level 2D-DWT, the top-level means the k-th level). In Step 3, the vertical integral projection and the horizontal integral projection are implemented on the sub-bands LH and HL, respectively; and then, from the two resulting projection vectors, we search for all of the extreme points. In Step 4, by using the obtained extreme points, we give non-overlapped patches of the top-level's low-frequency sub-band. In Step 5, based on the obtained non-overlapped patches, we further give overlapped patches of the top-level's low-frequency sub-band. Furthermore, by using the above patch rule to obtain all patches of the top-level's low-frequency sub-band of all samples, we classify all patches of each test sample by the nearest neighbor classifier (NN), and give the final decision according to the majority voting.

The rest of this paper is organized as follows. In Section 2, we give a brief introduction of 2D-DWT and the integral projection technique. Then the proposed method is discussed in Section 3. The numerical experimental results on the AR, FERET, Extended Yale B and LFW face image databases are reported in Section 4. Finally, conclusions are given in Section 5.

2. Technologies of 2D-DWT and Integral Projection

2.1. 2D-DWT

2D-DWT is a very popular and commonly used transform for image processing. It is based on one-dimensional wavelet transform and discrete wavelet transform (DWT) [1]. Firstly, a two-dimensional scaling function, $\Phi(x, y)$, and three two-dimensional wavelet functions, $\Psi^H(x, y)$, $\Psi^V(x, y)$, and $\Psi^D(x, y)$, are required [5]. They are usually showed as the products of two one-dimensional functions,

$$\begin{aligned}
&\Phi(x, y) = \phi(x)\phi(y),\\
&\Psi^H(x, y) = \psi(x)\phi(y),\\
&\Psi^V(x, y) = \phi(x)\psi(y),\\
&\Psi^D(x, y) = \psi(x)\psi(y),
\end{aligned}$$

where $\phi(\cdot)$ is a one-dimensional scaling function and $\psi(\cdot)$ is a one-dimensional wavelet function. These wavelets measure functional variations (intensity variations for images) along different directions [9]: Ψ^H measures variations along columns (for example, horizontal edges), Ψ^V measures variations along rows (like vertical edges), and Ψ^D measures variations along diagonals.

The two-dimensional discrete scaled and translated basis functions are defined as

$$\begin{aligned}
&\Phi_{j,m,n}(x, y) = 2^{j/2}\Phi(2^j x - m, 2^j y - n),\\
&\Psi^i_{j,m,n}(x, y) = 2^{j/2}\Psi^i(2^j x - m, 2^j y - n), i \in \{H, V, D\},
\end{aligned}$$

where j is a scale and m, n are the translation quantities. The corresponding transform of image $f(x, y)$ of size $M \times N$ is expressed as

$$\begin{aligned}
&W_\Phi(j_0, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)\Phi_{j_0,m,n}(x, y),\\
&W^i_\Psi(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)\Psi^i_{j,m,n}(x, y), i = \{H, V, D\},
\end{aligned}$$

where j_0 is an arbitrary starting scale and normally designed as 0. This transformation is the so-called two-dimensional discrete wavelet (2D-DWT).

2D-DWT is implemented as a set of filter banks, comprising of a cascaded scheme of high-pass and low-pass filters. The final result obtained is four non-overlapped multi-resolution sub-bands: LL, HL, LH and HH. The HL, LH and HH are all the high frequency components of the image. Further decomposition can be conducted on the LL sub-band, just as Figure 1 shows. Each sub-band can be thought of a smaller version of the image representing different properties. Take a simple image for example, just as Figure 2 shows: The band LL is a coarser approximation to the original image; the bands LH and HL record the changes of the image along horizontal and vertical directions, respectively; and the HH band shows the changes of the image along diagonal direction.

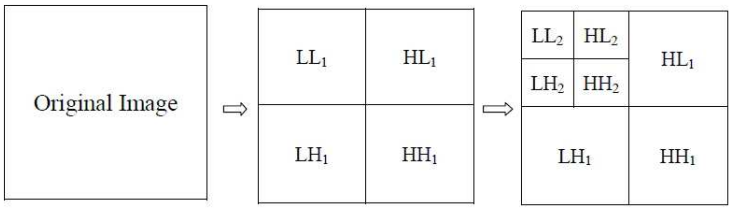

Figure 1: The demonstration of 2D-DWT

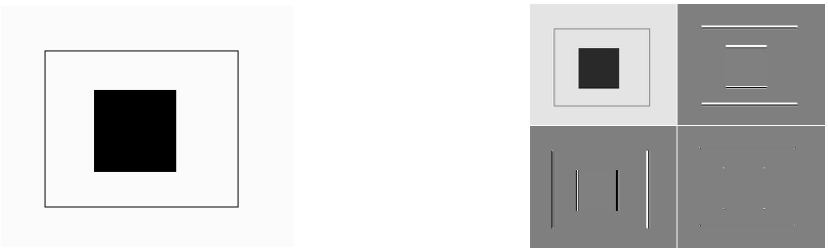

Figure 2: The original image and its sub-bands after 1 level 2D-DWT

2.2. Technology of Integral Projection

Integral projection is a method for extraction of facial features, which was first proposed by Kanade [25]. It is a useful technology for reflecting the

changes along one fixed direction by analyzing the accumulative changes. For a given image $f(x, y)$, we use $H(y)$ and $V(x)$ to represent the horizontal and vertical projections, respectively. Then, over the image area $\Omega = [x_1, y_1] \times [x_2, y_2]$, $H(y)$ and $V(x)$ are defined as:

$$H(y) = \sum_{x=x_1}^{x_2} f(x, y), \quad \forall y \in [y_1, y_2];$$
$$V(x) = \sum_{y=y_1}^{y_2} f(x, y), \quad \forall x \in [x_1, x_2].$$

3. The Proposed Method - NUPDWT

On the one hand, the sub-bands LH and HL after 2D-DWT can exhibit the texture's changes along horizontal and vertical directions, respectively, just as Figure 2 shows. On the other hand, integral projection is a good technology to make the texture changes statistically significant along horizontal and vertical directions. Thus it is logical to adopt integral projection on LH and HL to analyze the texture's changes. When the corresponding direction's integral projection is respectively applied to the sub-bands LH and HL, the changes will be more intuitive. Then the sub-band LL, which has the same size with LH and HL, can be partitioned based on the intuitive changes. Furthermore, compared with the original image, LL eliminates some noise and reduces the dimension which are both beneficial for recognition. We can illustrate the idea on face image in Figure 3, where the bottom part is the vertical integral projection of the sub-band LH; and the upper part is the corresponding partition which is performed on LL along horizontal direction. From this figure, we are amazed to find that Region 1 locates the boundary of nose along horizontal direction, and meanwhile, Regions 2 and 3 locate the boundary of eyes along horizontal direction.

According to the basic idea introduced above, we can find that our patch strategy is able to automatically partition the face images into non-uniform patches, essentially because of the special geometrical structure and physical meaning of the 2D-DWT's sub-bands. Based on this idea, we propose the method - NUPDWT.

There are two parts in this section. The first one describes the strategy of non-uniform patch; and in the second one, the part of classification is described and the whole steps of the method are listed.

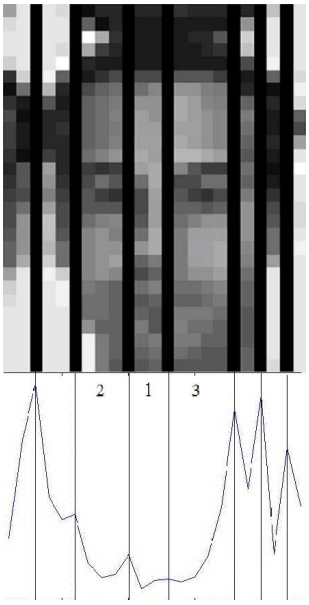

Figure 3: An example for vertical integral projection and the corresponding partition

3.1. The Strategy of Non-Uniform Patch

Suppose that a set of training images with size of $m' \times n'$ from C different subjects is given and the i-th subject contains N_i training samples, $i \in \{1, 2, \ldots, C\}$. Let $\mathcal{X}^i = (X_1^i, X_2^i, \ldots, X_{N_i}^i)$, where the matrix $X_j^i \in R^{m' \times n'}$ represents the j-th training image of the i-th subject, and $j \in \{1, 2, \ldots, N_i\}$. The average face of training images can be calculated by

$$\bar{X} = \frac{1}{N} \sum_{i=1}^{C} \sum_{j=1}^{N_i} X_j^i, \tag{1}$$

where $N = \sum_{i=1}^{C} N_i$ is the total number of training samples. $\bar{X}$ can give a general view of the current training database, so its patch strategy is suitable to most images. Thus we can directly apply the patch strategy of $\bar{X}$ to all images.

After disposing $\bar{X}$ with the 2D-DWT, we get the top-level's sub-bands LL, HL and LH with the size of $m \times n$. Then take the absolute value of LH and HL because what we need are the absolute variations of LH and HL in fact. We can observe the texture variations along horizontal direction after performing a vertical integral projection upon LH. Suppose the matrix of

LH is $A = (a_{ij}) \in R^{m\times n}$, the integral projection can be calculated by

$$v_j = \sum_{i=1}^{m} a_{ij}, \quad j \in \{1, 2, \ldots, n\}. \tag{2}$$

Next, we find all of v_j's extreme points, that is to say, we can obtain the set of extreme points by

$$I_x = \left\{ \hat{j} \in \{1, 2, \ldots, n\} \mid v_{\hat{j}-1} \leq v_{\hat{j}} \ \& \ v_{\hat{j}} \geq v_{\hat{j}+1} \right\}.$$

These extreme points reflect the saltation of texture along horizontal direction, and the regions between two extremes are relatively stable. So, the extreme points can mark the locally boundaries along horizontal direction just as Figure 3 shows.

Similarly, we get the horizontal integral projection of HL and the set of extreme points, denoted by I_y. Then, we give the partition of LL according to sets I_x and I_y just as the first graph in Figure 4. The second graph in Figure 4 shows the partition of an image according to the method of PCRC [17]. From Figure 4, it is easy to see that our method shows its advantage on retaining the integrity of local information.

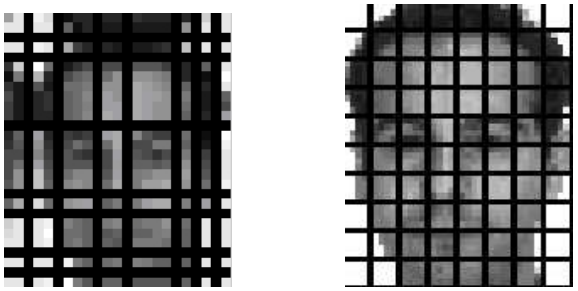

Figure 4: Examples for the patch strategies of NUPDWT and PCRC

The above patch strategy is given based on the average image $\bar{X}$. Though different images may have different tiny changes compared with the average image, for convenience, we apply the above strategy to all samples. In addition, in order to enhance stability and retain the integrity of local information further, we consider the overlapped patch. In our method, the overlapped patches combine four small blocks, that is, they contain both two small blocks along two directions, just as shown in the first graph of Figure

5: Patch $A_1A_3C_3C_1$ is the first overlapped patch, and Patch $B_1B_3D_3D_1$ is the second overlapped patch, where their overlapped part is $B_1B_3C_3C_1$. In the method of PCRC [17], the size of patch was designed as 10×10 (overlap was 5 pixels), which is shown in the second graph of Figure 5.

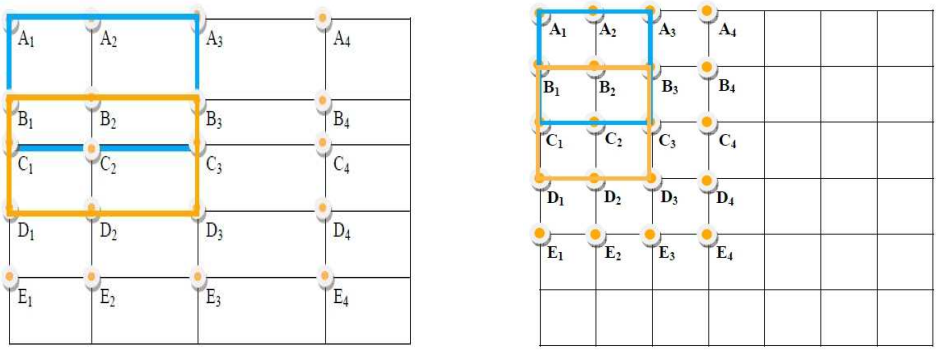

Figure 5: The overlapped patch methods of NUPDWT and PCRC

As the information on face is quite maldistribution, some important areas will be divided if the face image is partitioned uniformly. So our patches are reasonable with different sizes and better for retaining the integrity of local information when compared with the traditional partition schemes, e.g., uniformly partition, just as Figure 4 shows. Furthermore, different from manual partition or partition after finding some fixed points, our strategy is automatical through analyzing the image's geometrical structures, i.e., analyzing the corresponding direction's integral projection upon the sub-bands HL and LH to find the breakpoints and get the smooth regions.

3.2. Classification and Method Description

As we know, in recognition process, classifier plays a vital role. There are many classifiers such as NN [32], sparse representation-based classification (SRC)[19], support vector machine (SVM) [33] and so on.

In this paper, we apply the strategy of patch given in above subsection to all samples to obtain the non-uniform patches, then classify them with NN and give the final decision according to the majority voting.

The whole steps of the method are listed in the following.

- Preparatory Phase:

    1. Get the average image of all training samples as (1).

2. Decompose the average image with 2D-DWT to get four top-level's sub-bands: LL, HL, LH and HH. Take the absolute value of HL and LH.
3. Perform a vertical integral projection on LH as (2), then find the set I_x of extreme points.
4. Perform a horizontal integral projection on HL to get the set I_y of extreme points.

- Training Phase:

1. Read in all training samples, and decompose them with the same 2D-DWT.
2. Partition LL of every training sample according to sets I_x and I_y. Extract the non-uniform patches from each sample and save them.

- Test Phase:

1. Read in and decompose a query sample with the same 2D-DWT, then partition its LL according to sets I_x and I_y.
2. Extract the non-uniform patches and classify all of them by NN.
3. Give the final decision according to the majority voting.

4. Numerical Experiments

We use the AR, FERET, Extended Yale B and LFW face databases to test the performance of the proposed method.

The comparison methods contain four kinds, the first one is some state-of-the-art patch based methods, including patch based NN (PNN) [31], PCRC [17], patch based SRC (PSRC) [19], modular image principal component analysis (MIMPCA) [22, 23]; the second is majority voting based method Volterra [24]; the third is 2D-DWT based method TWSBF [10] which is combined with PCA here and the traditional 2D-DWT method (TDWT) [9] whose classifier is SVM with the parameters of $C = 128$ and $\gamma = 0.0078425$; and the last one, named UPDWT, is designed by ourselves in order to verify the superiority of non-uniform patch strategy used in this paper. In this method, integer projection technique is not used and the patch strategy is designed as follows: after we obtain the top-level's low-frequency sub-bands,

an overlapped uniform patch method is used for the top-level's low-frequency sub-bands, where overlapped method is the same as the one of NUPDWT.

Our experiments are made on a PC platform with 64-bit win8 operating system, Intel Core i5-3550S CPU, and 8G memory.

4.1. Parameter setting

Different from other patch based methods in which the size of patch are designed artificially and fixed, the size of patch in our method is unfixed. But the two parameters: the level of 2D-DWT and the combined number of patches, together influence the size of patch and eventually influence the number of final votes. Obviously, after partition, the size of the patch will get bigger but the number of final votes will be less as the combined number of patches increases. The level of 2D-DWT is directly related to the whole size of LL and the number of the extreme points, so the choice of the level is very important to the size of patch. Take the practice on the AR database for example. We randomly select 13 images of each person for training, and the remaining images are used for testing. Experiments of the level of 2D-DWT and the combined number of patches (which are marked as l and cp in the following) are shown in Figure 6, where l changes from 1 to 4, and cp has four choices as 1×1, 2×2, 3×3, 4×4. The highest recognition rate is obtained when $l = 1$ and $cp = 2 \times 2$, which shows that the patches have suitable sizes and stable property in this case. And these two parameters in other databases are all chosen through this way.

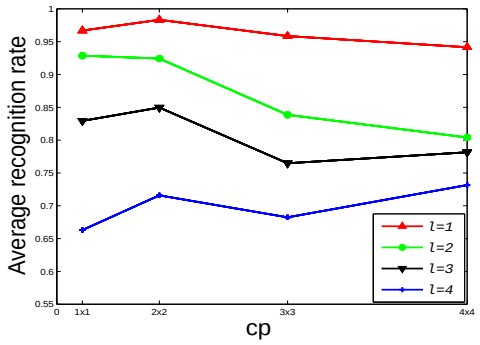

Figure 6: The parameters on the AR database

According to the original settings in references and the experiments above, some important parameters are exhibited in Table 1, where, l, s, o represent

for the level of 2D-DWT, the size of patch and the size of overlap, respectively.

Table 1: The setting of parameters in all methods on different databases

Method	AR	$FERET$	$ExtendedYaleB$	LFW
l in TDWT	2	4	3	2
l in TWSBF	2	4	3	2
l in UPDWT	1	1	1	1
l in NUPDWT	1	1	1	1
s in PNN	10×10	10×10	10×10	10×10
s in PSRC	20×20	20×20	20×20	20×20
s in Volterra	8×8	8×8	8×8	8×8
s in MIMPCA	10×10	20×20	24×28	16×16
s in PCRC	16×16	26×26	60×60	26×26
o in PCRC	8×8	13×13	30×30	13×13
s in UPDWT	8×8	8×8	28×28	8×8
o in UPDWT	4×4	4×4	14×14	4×4

Besides, the penalty parameter λ in PCRC is respectively set as 0.001, 0.1, 0.001, 0.001 in these database. For MIMPCA, the final dimensions of each patches after PCA are all set as 5, and the final dimension of PCA in TDWT is 20. For Volterra, the Volterra kernel is set as 3 and the approximation order is 2 on the databases except Extended Yale B where it is 1.

In our experiments, the chosen 2D-DWT is the 4-th order Daubechies ('db4') wavelet. For comparison, the levels of 'db4' in methods NUPDWT and UPDWT are designed as the same. However, the method TDWT and TWSBF are equipped with the levels to gain as good performance as possible in different databases. The experiments are all repeated 20 times.

4.2. Experiments on the AR Database

The AR database contains over 4,000 color face images of 126 people (70 men and 56 women), including frontal views of faces with different facial expressions, lighting conditions and occlusions. Images are of 768 by 576 pixels and of 24 bits of depth. The images of most persons were taken in two sessions (separated by two weeks). Each section contains 13 color images and 120 individuals (65 men and 55 women) participated in both sessions. In our experiments, AR database which we use contains 1,560 gray level images of 60 individuals (each person has 26 different images) and each image is

manually cropped and resized to 50×40 pixels. Figure 7 shows 26 sample images of one person in the AR database.

Figure 7: Images of the AR database

We compare NUPDWT with TDWT, TWSBF, PNN, PSRC, PCRC, MIMPCA, Volterra and UPDWT on the AR database.

We randomly select $7, 9, 11, 13, 15$ images of each person for training, respectively, and in every case, the remaining images are used for testing. The average recognition rates of TDWT, TWSBF, PNN, PSRC, PCRC, MIMPCA, Volterra and NUPDWT are shown in Table 2.

Table 2: The average recognition rates on the AR database

Method	7	9	11	13	15
TDWT	81.09%	82.97%	88.53%	94.33%	93.34%
TWSBF	88.35%	92.99%	96.56%	97.08%	97.90%
PNN	85.29%	87.74%	91.49%	94.16%	95.67%
PSRC	77.33%	82.74%	84.07%	85.11%	88.09%
PCRC	88.85%	91.79%	93.20%	94.58%	94.99%
MIMPCA	81.09%	85.93%	87.74%	90.24%	94.49%
Volterra	84.97%	88.80%	90.73%	92.96%	94.39%
UPDWT	85.51%	91.89%	94.12%	94.99%	95.98%
NUPDWT	**92.67%**	**96.41%**	**97.88%**	**98.21%**	**98.34%**

In addition, we extract the faces with scarf and faces with sunglasses as two small sub-databases. In each sub-database, there are 60 individuals with 6 different images each person, and each image is also manually cropped and resized to 50×40 pixels. We randomly choose 3 images for training, and the remaining images are used for testing. The average recognition rates are shown in Table 3.

Table 3: The face recognition rates on the AR's sub-databases

Scenarios	Sunglasses	Scarf
TDWT	80.56%	83.03%
TWSBF	83.50%	83.50%
PNN	70.14%	69.11%
PSRC	88.67%	85.44%
PCRC	91.92%	90.83%
MMPCA	85.06%	86.47%
Volterra	80.56%	83.03%
UPDWT	86.69%	87.08%
NUPDWT	**92.16%**	**94.33%**

From Table 2 and Table 3, it can be easily seen that the proposed NUPDWT achieves obviously higher recognition rates on the AR face database in all circumstances, no matter whether there are occluded images in the training and testing set or not, which proves that our method performs well for the images with occlusion.

4.3. Experiments on the FERET Database

The FERET database is one of the most useful databases for face recognition. The images are taken under different poses and illumination conditions. The FERET database in our experiment contains 1,200 face images of 200 people and the sizes of all images are tailored into 80×80. Figure 8 shows 18 sample images of 3 persons in the FERET database.

Figure 8: Images of the FERET database

We randomly choose $1, 2, 3, 4, 5$ images of each person for training, respectively, and in every case, the remaining images for testing. The average

recognition rates of each methods are shown in Table 4, and it can be easily seen that the recognition rates of NUPDWT are higher than the other methods on the database of FERET. In addition, we can see our method can also be used to solve problem with small changes of pose.

Table 4: The average recognition rates on the FERET database

Method	1	2	3	4	5
TDWT	30.61%	40.47%	49.82%	59.75%	63.20%
TWSBF	27.04%	37.58%	47.78%	51.68%	55.55%
PNN	34.13%	44.60%	53.32%	58.66%	60.50%
PSRC	11.79%	21.30%	25.10%	35.84%	42.30%
PCRC	13.61%	22.12%	24.13%	25.68%	30.32%
MMPCA	43.82%	60.90%	72.79%	80.95%	82.98%
Volterra	27.16%	33.04%	47.26%	56.25%	63.20%
UPDWT	34.78%	55.32%	67.61%	69.44%	82.20%
NUPDWT	**44.12%**	**65.14%**	**76.82%**	**85.91%**	**90.63%**

4.4. Experiments on the Extended Yale B Database

The Extended Yale B database contains images of 38 individuals, and each person has 64 different images with the size of 192×168 pixels under 9 poses and 64 illumination conditions. In our experiments we choose the former 40 images per person. The examples of this database are shown in Figure 9.

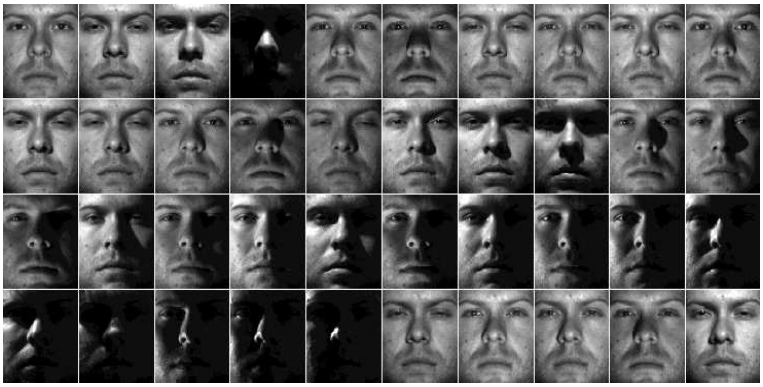

Figure 9: Images of the Extended Yale B database

We compare NUPDWT with TDWT, TWSBF, PNN, PSRC, PCRC, MIMPCA, Volterra and UPDWT on the Extended Yale B database.

We randomly choose 20 images per person for training, and the remaining samples for testing. The average recognition rates are shown in Table 5. RR in table represents "recognition rate".

Table 5: Recognition rates on the Extended Yale B database

Method	$AverageRR$	$MinRR$	$MaxRR$
TDWT	89.78%	81.05%	96.32%
TWSBF	40.12%	33.03%	44.74%
PNN	95.11%	92.76%	96.84%
PSRC	94.18%	88.68%	97.36%
PCRC	97.43%	90.13%	**100%**
MIMPCA	75.13%	70.79%	80.92%
Volterra	71.65%	62.24%	76.05%
UPDWT	93.33%	88.68%	96.58%
NUPDWT	**98.74%**	**97.10%**	99.87%

From Table 5, it can be easily seen that on the database of Extended Yale B, the average recognition rate of NUPDWT are higher than the other methods and the difference between the $MinRR$ and $MaxRR$ is the smallest in NUPDWT, which indicates that the proposed method is more robust than other methods.

What's more, we can confirm from these experimental results that the use of non-uniformly patch strategy alleviates the influence of illumination variations even when such noise is presented in both training and testing samples.

4.5. Experiments on the LFW Database

The LFW database contains images of 5749 different individuals in unconstrained environment [34]. The database here is a sub-database which gathers the subjects including no less than ten samples. It contains 158 individuals and each person has 10 different images with the size of 64×64 pixels. The examples of this database are shown in Figure 10.

We compare NUPDWT with TDWT, TWSBF, PNN, PSRC, PCRC, MIMPCA, Volterra and UPDWT on the LFW database.

Figure 10: Images of the LFW database

We randomly choose $1, 3, 5, 7, 9$ images per person for training, respectively, and in every case, the remaining images for testing. The average recognition rates are shown in Table 6.

Table 6: The average recognition rates on the LFW database

Method	1	3	5	7	9
TDWT	4.37%	7.19%	11.39%	17.23%	18.96%
TWSBF	3.54%	5.48%	7.06%	8.37%	9.72%
PNN	4.67%	8.42%	10.70%	12.51%	14.37%
PSRC	5.28%	8.64%	9.87%	11.16%	12.91%
PCRC	**6.26%**	10.91%	13.70%	15.36%	16.32%
MMPCA	4.12%	6.36%	7.75%	9.11%	10.47%
Volterra	2.91%	4.60%	5.49%	6.30%	6.00%
UPDWT	5.86%	10.96%	14.22%	17.34%	18.95%
NUPDWT	5.94%	**11.17%**	**14.42%**	**17.34%**	**19.02%**

From Table 6, one can see when training number is 1, NUPDWT has the second higher recognition rate only to PCRC. As the training number increases, NUPDWT and UPDWT perform much better than other methods, while the recognition performance is improved by NUPDWT to a larger extent. However, due to the unconstrained environment such as great changes on expression especially on pose, the recognition effectiveness of all the comparative methods on the LFW database are not ideal.

4.6. Experiments Analysis

It can be seen from the experiments implemented above that the recognition rates of our method are always higher than the methods based on 2D-DWT(such as TDWT and TWSBF) due to the consideration of the subbands' properties and the using of the patch strategy; when compared with

some patch-based methods, such as PNN, PCRC, PSRC, and MIMPCA, the advantage of the non-uniform patch strategy in our method can be exhibited; though our method and Volterra adopt the same voting strategy, our method outperforms Volterra because the patches in our method are non-uniform so that the integrity of some important local information can be maintained.

5. Conclusions

Considering the physics meaning of 2D-DWT's sub-bands, applying the technology of integral projection on the sub-bands, and employing the idea of overlapped patches, we proposed an overlapped non-uniform patch strategy in this paper. By using NN and majority voting, we accomplished the classification. Thus the new method was proposed. It contained the advantages of taking better consideration of the sub-bands' properties and retaining the integrity of local information. We implemented the new method as well as TDWT, TWSBF, PNN, PSRC, PCRC, MIMPCA, Volterra and UPDWT on the AR, FERET, Extended Yale B and LFW face databases, and the numerical results showed that the new face recognition method outperformed other methods, which revealed the superiority of the proposed overlapped non-uniform patch strategy.

The strategy of patch in our method was proposed based on the average image of training samples and complemented on the top-level's sub-band LL of all samples. If every sample can be further partitioned according to its own geometry structure, then the obtained patches will be better for retaining the integrity of local information, which maybe lead to more effective method for face recognition. In addition, the proposed method performed well with frontal or near-frontal facial images, but it can't effectively deal with the images under great changes on pose, as 2D-DWT cannot solve pose problem effectively. Face recognition under complex conditions, especially the different poses is a challenging and valuable issue. These are our study topics in the future.

Acknowledgments

This work was partially supported by the National Natural Science Foundations of China (Grants 11171252 and 11431002).

References

[1] I. Daubechies, Ten lectures on wavelets, CBMS-NSF Series in Applied Mathematics, Capital City Press, Montpelier, Vermont, 1992.

[2] C.L. Fan, X.T. Chen, N.D. Jin, Research of face recognition based on wavelet transform and principal component analysis, in: 2012 Eighth International Conference on Natural Computation (ICNC), 2012, pp. 575-578.

[3] K.C. Kwak, W. Pedrycz, Face recognition using fuzzy integral and wavelet decomposition method, IEEE Transactions on Systems Man and Cybernetics Part B:Cybernetics, 34(4)(2004) 1666-1675.

[4] D. Li, W. Pedrycz, N.J. Pizzi, Fuzzy wavelet packet based feature extraction method and its application to biomedical signal classification, IEEE Transactions on Biomedical Engineering, 52(6)(2005) 1132-1139.

[5] R.C. Gonzalez, R.E. Woods, Digital Image Processing, Third Edition, Publishing House of Electronics Industry, BeiJing, 2012, pp. 508-543.

[6] J.T. Chien, C.C. Wu, Discriminant waveletfaces and nearest feature classifiers for face recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(12)(2002) 1644-1649.

[7] P. Nicholl, A. Amira, DWT/PCA face recognition using automatic coefficient selection, in: 2008 4th IEEE International Symposium on Electronic Design, Test and Applications (DELTA '08), Hong Kong, 2008, pp. 390-393.

[8] E. Gumus, N. Kilic, A. Sertbas, O.N. Ucan, Evaluation of face recognition techniques using PCA, wavelets and SVM, Expert Systems with Applications, 37(2010) 6404-6408.

[9] B. Li, Y.H. Liu, When eigenfaces are combined with wavelets, Knowledge-Based Systems, 15(2002) 343-347.

[10] Z.H. Huang, W.J. Li, J. Wang, T. Zhang, Face recognition based on pixel-level and feature-level fusion of the top-level's wavelet sub-bands, Information Fusion, 22(2015) 95-104.

[11] A. Pentland, B. Moghaddam, T. Starner, View-based and modular eigenspaces for face recognition, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '94), 1994, pp. 84-91.

[12] R. Gottmukkal, V.K. Asari, An improved face recognition technique based on modular pca approach, Pattern Recognition Letters, 24(2004) 429-436.

[13] T. Ahonen, A. Hadid, M. Pietikäinen, Face recognition with local binary patterns, in: Proceedings of the Eighth European Conference Computer Vision (ECCV), 2004, pp. 469-481.

[14] S. Chen, J. Liu, Z. Zhou, Making FLDA applicable to face recognition with one sample per person. Pattern Recognition, 37(7)(2004) 1553-1555.

[15] X. Tan, S. Chen, Z. Zhou, F. Zhang, Recognizing partially occluded, expression variant faces from single training image per person with SOM and soft k-NN ensemble, IEEE Transactions on Neural Networks, 16(4)(2005) 875-886.

[16] R. Tjahyadi, W. Liu, S. An, S. Venkatesh, Face recognition via the overlapping energy histogram, in: Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI), 2007, pp. 2891-2896.

[17] P. Zhu, L. Zhang, Q. Hu, C.K. Shiu, Multi-scale patch based collaborative representation for face recognition with margin distribution optimization, in: Proceedings of European Conference on Computer Vision (ECCV), 2012, pp. 822-835.

[18] B. Wang, W.F. Li, Q.M. Liao, Face recognition based on nonsubsampled contourlet transform and block-based kernel Fisher linear discriminant, in: 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2012, pp. 1533-1536.

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

[20] R. Atta, M. Ghanbari, An efficient face recognition system based on embedded DCT pyramid, IEEE Transactions on Consumer Electronics, 58(4)(2012) 1285-1293.

[21] L. Yu, Z. He, Q. Cao, Gabor texture representation method for face recognition using the Gamma and generalized Gaussian models, Image and Vision Computing, 28(1)(2010) 177-187.

[22] J.F. Pereira, G.D.C. Cavalcanti, T.I. Ren, Modular Image Principal Component Analysis for Face Recognition, in: Proceedings of International Joint Conference on Neural Networks (IJCNN), 2009, pp. 2481-2486.

[23] G.D.C. Cavalcanti, T.I. Ren, J.F. Pereira, Weighted modular image principal component analysis for face recognition, Expert Systems with Applications, 40(2013) 4971-4977.

[24] R. Kumar, A. Banerjee, B. Vemuri, Volterrafaces: Discriminant analysis using volterra kernels, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2009, pp. 150-155.

[25] T. Kanade, Picture processing by computer complex and recognition of human faces, Technical Report, Kyoto University, Department of Information Science, 1973.

[26] R. Brunelli, T. Poggio, Face recognition: features versus templates, IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(10)(1993) 1042-1052.

[27] C. Garcia, G. Zikos, G. Tziritas, Wavelet packet analysis for face recognition, Image and Vision Computing, 18(2000) 289-297.

[28] X.J. Wei, C.T. Li, Fixation and saccade based face recognition from single image per person with various occlusions and expressions. in: 2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2013, pp. 70-75.

[29] Y. Chen, T.T. Do, T.D. Tran, Robust face recognition using locally adaptive sparse representation, in: 17th IEEE International Conference on Image Processing (ICIP), 2010, pp. 1657-1660.

[30] D. Lin, X. Tang, Recognize high resolution faces: From macrocosm to microcosm, in: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2006, pp. 1355-1362.

[31] R. Kumar, A. Banerjee, B.C. Vemuri, H. Pfister, Maximizing all margins: Pushing face recognition with kernel plurality, in: 2011 IEEE International Conference on Computer Vision (ICCV), 2011, pp. 2375-2382.

[32] R. Duda, P. Hart, D. Stork, Pattern Classification, 2nd ed, John Wiley Sons, 2001, pp. 174-188.

[33] V.N. Vapnik, Statistical Learning Theory, Wiley-Interscience, New York, 1998.

[34] G.B. Huang, M. Mattar, T. Berg, E. Learned-Miller, Labeled faces in the wild: a database for studying face recognition in unconstrained environments, Technical Report, University of Massachusetts, Amherst, 2007.