Detail-Enhance Face Illumination Normalization Based on LDCT-Wavelet

Songbai Pu, Huachuang Wang
Institute of Optics and Electronics Chinese Academy of Sciences, Sichuan Province, China
pusongbai1@126.com

Abstract - As a kind of self-carry and unique biological characteristic, face has been widely used in the authentication field. Although in recent years face recognition has achieved rapid progress, there are still some technical fortresses difficult to overcome. Illumination is one of the major challenges, a large number of researches have showed that the accuracy of face recognition is highly dependent on illumination variations. In this paper a novel DCT-Wavelet approach is proposed, through rescaling the low-frequency coefficients based on entropy optimized by rand-shake and wavelet de-noising by maximizing detail energy with strategy adaptive differential evolution (SaDE), image can normalize illumination and enhance detail at the same time. Experiments showed this approach has outstanding advantage over other LDCT methods and the error rate of face recognition can be decreased approximately 3%.

Index Terms - LDCT, Entropy, Rand-Shake, Wavelet, Detail-Energy SaDE

1. Introduction

Face is one of the most important distinguish labels. As a weak invasive biometric, face has been widely used in areas of building access control, identification, and surveillance [1].

Differences caused by illumination variations have proven to be much more significant than differences between individuals [2], what’s more, as the preprocessing step of face recognition, the result of illumination normalization would decide subsequent processing. Various approaches have been proposed to solve the problem; the mainstream methods can be divided into three categories: face and illumination modelling, illumination invariant feature extraction and normalization [3].

Face and illumination modelling often need complex computation and a large number of reference images which can not satisfy the real-time requirement, illumination invariant feature would lose global information of image which may be useful for latter recognition. Based on upper factors, illumination normalization has become a research hot spot. The discrete cosine transformation (DCT) proposed in is one of the most representative normalization methods [4].

The original image is transformed into log domain and then operates DCT; the low frequency coefficients usually contain the illumination part. Traditional DCT approach just zero part of the low frequency coefficients which would weaken contour of face and increase the high-frequency noise relatively, both of the two factors will reduce the accuracy of face recognition. Virendra P. Vishwakarma proposed a rescaling DCT coefficients approach to keep low-frequency face contour while normalizing illumination [5], in his approach the low-frequency coefficients are divided by a constant (>1) instead of being zeroed, in addition he also suggests to increase the DC coefficient by 10% to enhance contrast, but he did not propose any principle to decide the constant and increase factor. Perez and Castillo proposed a method applied Genetic Algorithms to search appropriate weights to re-scale low-frequency DCT coefficients [6], although the rescale weights are reasonable and adaptive, the computation complexity is horrible. Xiaouqin Zhao designed two lighting condition (LC) coefficients lies in horizontal and vertical direction to decide rescale factor of DCT coefficients [7], this method uses the spatial information of face-organ (e.g. eyes and nose) to estimate the illumination proportion in DCT low-frequency coefficients, thus it needs to detect face-organ accurately at first.

In this paper, entropy is used as the criterion to decide how to rescale low-frequency DCT coefficients, larger the entropy is, more reasonable the rescaled coefficients are. To accelerate calculation, we use rand-shake to approximate the optimal value. Considering high-frequency noise may pollute entropy, we use wavelet to de-noise before normalizing illumination and to compensate the loss of facial contour caused by LDCT on low-frequency, we increase the coefficients of high-frequency wavelet decomposed layers before wavelet reconstitution, to increase the accuracy strategy adaptive differential evolution (SaDE) is used for continuous optimizing. Experiments show that this method can normalize illumination and enhance detail at the same time.

2. Discrete Cosine Transform

A. Discrete Cosine Transform in Log Domain

Logarithm transform is often used in image enhancement to expand dark pixels and simplifying the multiplication and division operation [2]. In a simple situation, the image gray level \( f(x, y) \) can be assumed to be proportional to the reflectance \( r(x, y) \) and the illumination \( l(x, y) \) [8].

\[
f(x, y) = r(x, y) \times l(x, y)
\]

Taking logarithm transform on (1), we have

\[
\log f(x, y) = \log r(x, y) \times \log l(x, y)
\]

This work is partially supported by The Chinese Academy of Sciences Innovation Fund # A08K001.
Thus the relation between reflectance $r(x, y)$ and illumination $l(x, y)$ could be changed from multiplication to addition, $\log f(x, y)$ would be closer to $\log r(x, y)$. The illumination variation is weakened somewhat.

As illumination mainly located in low frequency, traditional approaches usually zero a number of low frequency LDCT coefficients in zig-zag order to normalize illumination. The algorithm flow shows as follows:

Step1: transform image into log domain.
Step2: calculate LDCT coefficients of log-image.
Step3: zero low-frequency coefficients in zig-zag order.
Step4: inverse LDCT coefficients back to log-image.
Step5: inverse log-image to reconstruction image.

B. Rescale low-frequency DCT coefficients

Based on a simple assumption that the illumination has a multiplication relation with original image, just as (1), thus we have

$$C(u,v) = a(u)a(v)\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} \log f(x,y) \times \cos \left( \frac{2\pi(2x+1)v}{2N} \right) \cos \left( \frac{2\pi(2y+1)u}{2M} \right) $$
$$= a(u)a(v)\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} \log f(x,y) \times \cos \left( \frac{2\pi(2x+1)v}{2N} \right) \cos \left( \frac{2\pi(2y+1)u}{2M} \right) $$
$$= a(u)a(v)\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} \log f(x,y) \times \cos \left( \frac{2\pi(2x+1)v}{2N} \right) \cos \left( \frac{2\pi(2y+1)u}{2M} \right) $$
$$= a(u)a(v)\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} \log f(x,y) \times \cos \left( \frac{2\pi(2x+1)v}{2N} \right) \cos \left( \frac{2\pi(2y+1)u}{2M} \right) $$

(3)

And then

$$\log f(x,y) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} a(u)a(v)C_u(u,v)$$

$$\times \cos \left( \frac{2\pi(2x+1)v}{2M} \right) \cos \left( \frac{2\pi(2y+1)u}{2N} \right)$$

$$= \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} a(u)a(v)[C_u(u,v) - C_u(u,v)]$$

$$\times \cos \left( \frac{2\pi(2x+1)v}{2M} \right) \cos \left( \frac{2\pi(2y+1)u}{2N} \right)$$

(4)

As we can see from (4) the original log-image $\log f(x,y)$ could be reconstructed from reflected image LDCT coefficient $C_u(x,v)$ which equals original image LDCT coefficients $C(x,y)$ minus illumination LDCT $C_u(x,v)$, low-frequency information contains not only illumination but also face contour, that is $C_u(x,y)$ in low-frequency should not be zero, otherwise the result of inverse LDCT would loss contour information of face. Many approaches use a constant to down rescale the low-frequency LDCT coefficients instead of making them zero, it need to satisfy the constraint $C(x,y) - C_u(x,y) = C(x,y) = a$ in log domain, where $a$ is a constant, but as we can conclude from (3), if the illumination size is irregular, the $C_u(x,y)$ would differ from each other a lot in low-frequency, that means the upper constraint can not be satisfied, so the down scale factor should be decided independently and designed as a constant.

3. Entropy based LDCT with Rand-Shake Optimization

Entropy represents the uncertainty of the information system; it has been widely applied to image segmentation [9], image enhancement [10], image fusion [11], image compression and threshold selecting [12], [13]. Entropy often used as indicators of the amount of information, the purpose of illumination normalization is to increase the useful information, so entropy could be used as the criterion to judge which normalization result is best.

To decide the scale factor in original domain or subtrahend in log domain for every LDCT coefficient in continuous domain insufferable. Here we introduce a rand-shake optimization method to simplify this process. The algorithm flow shows as follows:

Step1: transform image into log domain.
Step2: calculate LDCT coefficients of log-image.
Step3: zero low-frequency DCT coefficients in zig-zag order.
Step4: inverse LDCT coefficients back to log-image.
Step5: inverse log-image to reconstruction image.

Step1: transform image into log domain.
Step2: calculate LDCT coefficients of log-image.
Step3: zero low-frequency DCT coefficients in zig-zag order.
Step4: inverse LDCT coefficients back to log-image.
Step5: inverse log-image to reconstruction image.

Step3: zero $m \ (m = 1... \frac{M \times N}{2})$ low frequency coefficients

$$C(0,0), C(0,1), C(1,0)... C(x,y)$$

, where total number is $m$ in zig-zag order, and inverse it back to log-image, then calculate the entropy $E_m \ (m = 1... \frac{M \times N}{2})$ of each image, at last select the value $m$ with the biggest entropy as Fig.1 shows.

Step4: if $n$ is selected at last step, make a random number $R(x,y)$ between 0.3 to 0.8 for each $C(x,y)$, the total number of $R(x,y)$ is $n$ and change $C(x,y)$ into $C(x,y) \times (1 - R(x,y))$, then inverse LDCT coefficients back to log-image and calculate the entropy, recode all the $R(x,y)$ when the entropy is the biggest compared to others in history.

Step5: loop step 4 until the biggest entropy increases less than $10^{-2}$ or iteration times has exceed 500.

In Fig.2, the number from C(0,0) to C(x,y) in zig-zag order is $n$ which is decided by step 3 as Fig.1 shows, and $P(x,y) = C(x,y) \times R(x,y)$, $R(x,y)$ is the random number produced by step 3.
4. Wavelet Transform

In upper LDCT process, we assumed that the result of illumination normalization is better when its reconstructed image has larger entropy, but in fact this assumption often is impracticable because of noise. Noise can destroy the homogeneity of image, so the uncertainty of the image information will increase, and then the entropy will increase too.

To avoid the disturbance of noise when using entropy to normalize illumination, we use wavelet to de-noise at first. As wavelet has both spatial and frequency resolution, face detail could be kept while removing noise, and to enhance the detail accurately, we use strategy adaptive differential evolution (SaDE) to decide reconstruction scale factor.

A. Mallat decomposition

As the character of multi-resolution, image can be decomposed into pyramid structure by wavelet, Mallat decomposition is one of the most common methods [14], the decomposition process shows as follows:

\[
\begin{align*}
C(i,0) &= \sum_{l,n} h_{2i} g_{2n} c_{l,n} \\
D^{1,1}_{i,n} &= \sum_{l,n} h_{2i} g_{2n} d^{1,1}_{l,n} \\
D^{1,2}_{i,n} &= \sum_{l,n} g_{2i} h_{2n} c_{l,n} \\
D^{1,3}_{i,n} &= \sum_{l,n} g_{2i} h_{2n} d^{1,3}_{l,n}
\end{align*}
\]

Where \(c_{l,n}\) stands for the \(j\) level low frequency coefficient, and \(d^{1,1}_{l,n}\) stands for the \(j+1\) level low-frequency coefficient, \(d^{1,2}_{i,n}\) to \(d^{1,3}_{i,n}\) stands for \(j+1\) level high frequency coefficient. Fig.3 shows the structure of three-layer composition:

The reconstruction process is the inverse of decomposition:

\[
\begin{align*}
c_{l,n} &= \sum_{j=0}^{j=3} \sum_{l,n} h_{2j} g_{2n} c_{l,n} \\
D^{1,1}_{i,n} &= \sum_{j=0}^{j=3} \sum_{l,n} h_{2j} g_{2n} d^{1,1}_{l,n} \\
D^{1,2}_{i,n} &= \sum_{j=0}^{j=3} \sum_{l,n} g_{2j} h_{2n} c_{l,n} \\
D^{1,3}_{i,n} &= \sum_{j=0}^{j=3} \sum_{l,n} g_{2j} h_{2n} d^{1,3}_{l,n}
\end{align*}
\]

B. Wavelet Denoising

As we know noise mainly concentrates in high-frequency part, and the frequency range is small, detail lies in high-frequency too, but unlike noise, the frequency range of detail is large [15], that means the noise may only stay in \(HH1, HL1, LH1\) layers as Fig.3 shows, and the detail may stay in \(HH1, HL1, LH1, HL2, LH2, HH2, HH3, HL3, LH3\) layers. Fig.4 and Fig.5 show that the detail information of noised image is concealed in noise in \(HL1, LH1\) compared to original image and the detail information is more significant than noise in \(HL2, LH2, HL3, LH3\) which means noise becomes weak and detail keeps in low-frequency just as our argument. Wavelet can process every individual pixel in every frequency, thus it can process noise and detail independently to keep detail while de-noising.

Base on the upper argument, we consider the pixel on original image which has large valve in high-frequency Mallat layer (e.g. \(HH1, HL1, LH1\)) and small value in lower high-frequency Mallat layer (e.g. \(HL2, LH2, HH2, HH3, HL3, LH3\)) to be the noise pixel, then we make the high-frequency Mallat coefficient(locates \(HH1, HL1, LH1, HL2, LH2, HH2, HH3, HL3, LH3\)) of noise pixels zero, nothing is done to detail pixels, thus the detail is kept while the noise is removed.
for each Mallat decomposition layer

\[ u_y = x_y + K \times (x_{n,j} - x_j) + F \times (x_{n,j} - x_{n,j}) \]  

\[ (11) \]

where \( u_y \) stands for the value of dimension \( j \) for individual \( i \), \( F \) stands for the scale factor, \( CR \) stands for a random between 0 to 1, \( j_{rand} \) stands for a random between 1 to the number of dimensions \( n \) for every individual. Each type of mutation has own advantage for different data set. Many researches suggest mixing them together to increase adaptability for heterogeneous data set.

A. Detail-Energy Function

Differential evolution needs an energy function, in LDCT process, we use entropy to judge weather the result of inverse LDCT is good, but here, entropy is no longer a good criterion as energy function. The reason is explained as follows

Firstly, to maximize entropy, the rescale factor of high frequency layer would be much larger than low frequency layer, this would weaken contour information a lot as it will be concealed in the highly enlarged detail information. Research has figured out that the large scale information in low-frequency is also very important to accuracy of face recognition [17].

Secondly, as wavelet reconstruction is very sensitive to wavelet coefficient, arbitrarily rescale coefficients would lead to plaque effect in reconstruction image as Fig.6 shows, so there should be a strict constraint for rescale factor, but entropy can not supply a constraint like this while optimizing.

According to the two reasons, it is necessary to find a new energy function for differential evolution. The constraint of rescale factor to avoid plaque effect is hard to solve, but we could judge weather the rescale factor is reasonable from the reconstruction image, if there is plaque effect in reconstruction image, the smooth extent of full image will be smaller than reconstruction image that does not rescale coefficient, and to enhance detail of face we use a criterion named edge-intensity as a part of energy function. The smooth value \( f_{smooth} \) is calculated as follows [18]

**Step1:** calculate \( k(x,y) = I(x,y) - \frac{1}{8} \sum I(u,v) \) for each pixel, where \( U \) are the eight pixels around.

**Step2:** calculate the probability density \( p(x,y) \) of \( k(x,y) \)

**Step3:** calculate the smooth value of full image

\[ f_{smooth} = \int p \times k^2 dk \]  

\[ (12) \]
The edge-intensity value \( f_{\text{smooth}} \) is designed as follows

Firstly, get edges of full image by Canny operator with parameter \( \theta, 3 \), then calculate the number of pixels \( S \) which lie in a continuous edge with a size larger than 20 pixels, and finally calculate the edge-intensity value by \( f_{\text{edge}} = \frac{S}{N} \), where \( N \) is the number of all pixels.

As the plaque often is comprised of small grating, the edge size can not exceed 20 pixels, so the plaque effect will not disturb the calculation of real facial detail.

The energy function is defined as

\[
f = f_{\text{smooth}} + f_{\text{edge}} \quad (13)
\]

B. Strategy-Adaptive Differential Evolution Equations

The select of mutation type and parameters e.g. \( CR \) and population size \( (NP) \) is very important [19]. Stom and Price suggested that \( NP \) should be five to ten times of individual dimension\( (D) \), \( CR \) should be approximate 0.5, and scale factor should be 0.4 to 1, and other researchers also proposed many different parameter select strategies, e.g. \( NP=\{3D,8D\} , F=0.6, CR=\{0.3-0.9\} \) [20], some of these strategy even are contradictory to each other. Qin proposed a strategy adaptive method to select mutation type and parameters [19], in his method the mutation strategy and parameters are not fixed, they will evolve to fit the data set, history information and parameters are stored in memory like success strategy memory \( (S) \), fail strategy memory \( (L) \), and CR memory \( (CRM) \) to guide evolution process.

We assume that the Mallat decomposition is done for three layers, so that the individual in \( DE \) has the dimension of 10 \( (for \ HH1, HL1...LL3) \), and the individual could be designed as \( (e_1...e_{10}) \), each \( e_i \) denotes for the rescale factor for each Mallat frequency domain \( (e.g. HH1, HL1...LL3) \). The algorithm flow shows as follows

Step1: initialize the parameters: Let \( NP=50, \) and \( CR=0.5 \), iteration counter \( k=0 \), memory size \( LP=20 \), iteration stop times \( st=100 \), and the range of each dimension

\[
e_{\text{min}} = \{e_{\text{min}}^1, e_{\text{min}}^2...e_{\text{min}}^{10}\}, e_{\text{max}} = \{e_{\text{max}}^1, e_{\text{max}}^2...e_{\text{max}}^{10}\}
\]

Step2: initialize population: get a random number in each dimension range as the initial value of each dimension.

\[
e_i^{0,j} = e_i^{j} + R(0,1) \times (e_{\text{max}}^j - e_{\text{min}}^j) \quad i=1...50 \quad (14)
\]

Where \( e_i^{j} \) stands for the value of dimension \( j \) for individual \( i \) at generation \( k \), then initialize the probability of each mutation strategy being selected.

\[
P_m^0 = 0.5 \quad m=1..4 \quad (15)
\]

Where \( P_m^0 \) stands for the probability of mutation strategy \( m \) at generation \( k \)

Step3: produce differential vectors and then get next generation, there are two situations as follows

Situation1: \( k<=LP \)

1) Produce \( NP \) differential vector \( u_i \) base on each individual \( e_i \) for iteration \( k \):

Firstly, update \( CR \) and \( F \) for each individual \( e_i \) as \( (16) \).

\[
CR_i = N(0.5,0.1)
\]

\[
F_i = N(0.5,0.3)
\]

Where \( CR_i \) stands for the \( CR \) value for individual \( i \) under strategy \( j \), and \( F_i \) is the \( F \) value for individual \( i \).

Then, select mutation strategy with roulette selection method based on probability \( P_m^k \) and produce each differential-vector \( u_i \) under the chosen strategy as \( (8)-(11) \).

2) Evolve the next generation:

Firstly, rescale the Mallat coefficient with every individual \( e_i \) and differential vector \( u_i \) as \( (7) \) (here \( e_i \) and \( u_i \) function as \( \lambda \), and \( e_i^{j,k} \) and \( u_i^{j,k} \) function as \( \lambda_j \)), then inverse it back to spatial image \( I_i' \) and \( I_i'' \), where \( i=1...50 \)

Then, calculate the detail-energy \( f_{i'} \) and \( f_{i''} \) of reconstruction image \( I_i' \) and \( I_i'' \), then select the next individual \( i \) for generation \( k+1 \)

\[
e_i^{k+1} = \begin{cases} u_i^{j,k} & \text{if } f_i^{j,k} < f_i' \\ e_i^k, \text{other} & \end{cases} \quad (17)
\]

3) Update the memory:

Firstly, update the success strategy memory \( (S) \):

If an individual \( e_i \) or differential vector \( u_i \) is selected into next generation \( k+1 \) with mutation strategy \( z \), the success strategy memory would be updated as:

\[
S_z^k = S_z^k + 1 \quad (18)
\]

Secondly, update the fail strategy memory \( (L) \):

If an individual \( e_i \) or differential vector \( u_i \) is not selected into next generation \( k+1 \) with mutation strategy \( z \), the fail strategy memory would be updated as:

\[
L_z^k = L_z^k + 1 \quad (19)
\]

Lastly, update the CR memory \( (CRM) \):

If an individual \( e_i \) or differential vector \( u_i \) is selected into next generation \( k+1 \) with mutation strategy \( z \), \( CRM_z \) will be updated to \( CR_i^z \).

Situation1: \( k>LP \)

1) Produce \( NP \) differential vector \( u_i \) base on each individual \( e_i \) for iteration \( k \):

Firstly, update the probability of each mutation strategy being selected.

\[
P_i = \frac{P_m^k}{\sum_{j=1}^{NP} P_j} \quad \sum_{j=1}^{NP} 4S_i^{'} + 4L_i^{'} \quad i=1...4
\]

(20)
Then, update $CR$ and $F$ for each individual $e_i^k$

$$
CR_i^k = N(CR_{ai}, 0.1) \quad CR_i^k = \frac{\sum \text{CRM}_i}{LP} \\
F_i = N(0.5, 0.3)
$$

Lastly, select mutation strategy with roulette selection based on probability $P_m^k$ and produce each differential-vector $u_i^k$ under the chosen strategy as (8)-(11).

2) Evolve the next generation:
The operation is the same with the situation when $k <= LP$.

3) Update the memory:
Firstly, update the success strategy memory ($S$):
Calculate $S_i^k, i = 1.4$ with the same operation as the situation when $k <= LP$, then remove the old recode $S_i^k, i = 1.4$ from $S$ memory, and add the new recode $S_i^k, i = 1.4$ to the tail of $S$ memory.

Then, update the fail strategy memory ($L$):
Calculate $L_i^k, i = 1.4$ with the same operation as the situation when $k <= LP$, then remove the old recode $L_i^k, i = 1.4$ from $L$ memory, and add the new recode $L_i^k, i = 1.4$ to the tail of $L$ memory.

Lastly, update the CR memory ($CRM$):
The operation is the same with the situation when $k <= LP$.

Step 4: loop step 3 until the generation counter $k$ is larger than iteration stop criterion $st$, or the detail-energy does not increase any more.

6. Experiments

A. Database
The experiments are carried out on the Yale Face Database B and Extend Yale B database; here we only concentrated on images with illumination variations.

B. Experimental Scheme
In order to evaluate the performance of the proposed method, we tested a set of various real-world images, and to test the capacity of process image with noise, we designed two groups of face-image, group A is original images sampled from Yale B database, and group B is the noised images from group A with Gaussian coefficient 0.01, images are given in Fig.7.

When compared to other illumination normalization method, the experiment result showed as Fig.8.
Row one and three are the results of group A, and row two and four are the results of group B. The first image of each row is the result of LDCT which only zeros low-frequency coefficients, the second image is the result of LDCT which rescales low-frequency coefficients by a constant, the third image is the result of entropy based LDCT, and the last image is the result of LDCT-Wavelet approach which is introduced in this whole paper.

In group A (images without noise), the result of illumination normalization is significantly improved by entropy based LDCT compared to other two types of LDCT, and the detail could be enhanced somewhat.

In group B (images with noise), the effect of LDCT was concealed by the noise, what’s more, the noise was relatively strengthened to increase the entropy of image. Here, proposed wavelet method solved this problem by denoising and rescaling Mallat coefficients, the result was satisfactory.

We used Yale B and Extend Yale B database to test the accuracy of face recognition, the recognition method is nearest neighbor 0-1 classifier, and we did not consider the influence of facial expression and pose. The result shows as Table I.

<table>
<thead>
<tr>
<th>Method Type</th>
<th>Error Rate(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Normalization</td>
<td>YaleB</td>
</tr>
<tr>
<td></td>
<td>38.5</td>
</tr>
<tr>
<td>L-DCT (zero low-frequency coefficient)</td>
<td>5.8</td>
</tr>
<tr>
<td>L-DCT (rescale low-frequency coefficient with constant)</td>
<td>4.6</td>
</tr>
<tr>
<td>Proposed Method</td>
<td>2.1</td>
</tr>
</tbody>
</table>
To test the ability of operating noise image, we compared the error rate of face recognition of our method with LDCT method which rescale low-frequency coefficient with constant under different Gaussian noise (based on Yale B database), the result showed as Fig.9.

![Fig. 9 the error rate of face recognition](image)

**7. Conclusions**

LDCT is a kind of real-time pre-processing methods and requires no assumption on the light source and any prior information on 3-D face geometry, but traditional methods can not supply a criterion to guide how to rescale LDCT coefficients adaptively, in this paper, entropy based LDCT is proposed to compensate the shortage of traditional LDCT. When operating noised images, entropy based LDCT may get error normalization result by enlarging the noise, and traditional LDCT can not get the right result either. We use wavelet de-noise before entropy-based LDCT, and to compensate the loss caused by LDCT and enhance the detail of face we proposed a detail–energy maximization wavelet method to rescale Mallat coefficients in reason. Algorithm was tested on Yale B and Extend Yale B database, and results showed this proposed approach has outstanding advantage over other LDCT-based illumination normalization; it can control the error rate of face recognition at 2.1% with Yale B database and 9.3% with Extend Yale B database.

**References**


