COMPACT LBP AND WLBP DESCRIPTOR WITH MAGNITUDE AND DIRECTION DIFFERENCE FOR FACE RECOGNITION

Soo-Chang Pei[1], Mei-Shuo Chen[1], Yi Yu[2], SuHua Tang[3], ChunLin Zhong[4]

[1]Graduate Institute of Communication Engineering, National Taiwan University
[2]National Institute of Informatics
[3]Graduate School of Informatics and Engineering, The University of Electro-Communications
[4]University of Science and Technology of China
pei@cc.ee.ntu.edu.tw, r03942051@ntu.edu.tw, yiyu@nii.ac.jp, shtang@uec.ac.jp,
chlzhong@mail.ustc.edu.tw

ABSTRACT

In this paper, we propose a novel descriptor for face recognition on grayscale images, depth images and 2D+depth images. It is a compact and effective descriptor computed from the magnitude and the direction difference. It can be concatenated with conventional descriptors such as well-known Local Binary Pattern (LBP) and Weber Local Binary Pattern (WLBP), to enhance their discrimination capability. To evaluate the performance of our descriptor, we conducted extensive experiments on three types of images using four different databases. The experimental results demonstrate the robustness and superiority of our approach, and the performances of our new descriptor surpass that without magnitude and direction difference. At the end, we further compare our descriptor with Convolution Neural Network (CNN) to show the compactness and effectiveness of the proposed approach.

Index Terms— Face Recognition, Magnitude and Direction Difference, Local Binary Pattern, Weber Local Descriptor, Convolution Neural Network

1. INTRODUCTION AND RELATED WORKS

In image processing, feature extraction is an indispensable step, which can extract the intrinsic content (information) from the images (data). A recent evaluation on various local descriptors for face recognition can be found in [16]. Among the popular face descriptors, LBP is considered to be a low-dimensional and high-efficient local face descriptor [1][2][3], and is denoted as follows:

$$LBP_{R,N}(x,y)=\sum_{p=0}^{N-1}s(g_p-g_c)2^p\ ,\quad s(x)=\begin{cases}1, x\geq 0\\ 0, x<0\end{cases}, \qquad (1)$$

where R is the radius, N is the number of neighbor pixels. g_c represents the grey level of the center pixel located at (x,y), and g_p corresponds to each pixel of the neighborhood.

Then, a histogram is extracted from the local binary patterns as a 2^N dimension descriptor, which offers more invariance to geometric translation. A subset of the local binary patterns called uniform LBP [4] is introduced to reduce the number of possible bins, where there are at most two bit transitions from 0 to 1 or vice versa in binary code. Fan Liu *et al.* proposed a new local descriptor, WLBP in [5], which is a combination of LBP and weber local descriptor (WLD) [12]. A WLBP consists of two components: differential excitation of weber local descriptor and the Uniform LBP, where the differential excitation extracts perception features by Weber's law.

Besides grayscale images, many researchers have investigated the LBP descriptor for depth images to achieve 3D face recognition recently. The authors in [10] proved the importance of magnitude difference in depth images. Yin *et al.* [11] proposed an extended LBP for 2D+3D face recognition named CLDP. They stressed on high order derivative, which can reveal the change trends between data.

In this paper, we propose to enhance LBP and WLBP with a new face descriptor, which, build on the new concept called "Magnitude and Direction Difference". We argue that the magnitude difference and the direction information both are indispensable features for face recognition. With the magnitude difference, we can derive a more discriminative facial representation. The second-order derivative pattern encodes the change of directions of the neighborhood, which reveals the trends of varieties. We implement the concept with a histogram of only 16 bins in the case of 8-neighbour and integrate it with LBP, Uniform LBP and WLBP to form new descriptors, which inherit the simplicity and efficiency of conventional methods.

Our paper is organized as follows: in Section 2, we present our new descriptor and explain how to enhance LBP and WLBP to improve face representation. In Section 3, we evaluate the proposed descriptor with extensive experiments. Finally, Section 4 provides the conclusion and future works.

2. PROPOSED DESCRIPTOR

The framework of the proposed approach is shown in Fig. 1, and is built on two parts: the sign part and the magnitude and direction difference part. The sign part is represented by a LBP, a Uniform LBP or a WLBP descriptor, and the magnitude and direction difference is extracted and denoted as a second histogram. Finally, the histogram of LBP or WLBP is followed by magnitude and direction difference to form the new descriptor.

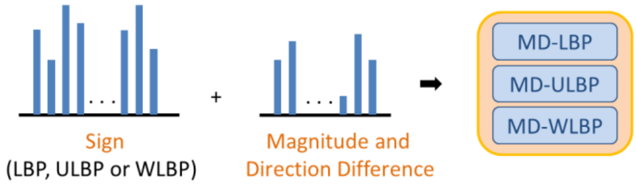

Fig. 1 The framework of magnitude and direction difference

2.1. The concept of magnitude and direction difference

Only the sign part (e.g., LBP) is actually incapable of describing more detailed information because it encodes the binary results of the first-order derivative among local neighbors by using a simple threshold function as in (1). In our new descriptor, we investigate the feasibility and effectiveness of using 2nd-order local patterns and further exploit a threshold of absolute magnitude (T_m).

The magnitude and direction can be defined as follow:

$$D(i) = \begin{cases} 1, & s'(g_i - g_c) \cdot s'(g_c - g_{i+4}) = 1 \\ 0, & otherwise \end{cases} \tag{2}$$

$$where\ s'(x) = \begin{cases} 1, x \geq T_m \\ 0, -T_m < x < T_m \\ -1, x \leq -T_m \end{cases} \tag{3}$$

T_m is the threshold value used for encoding the magnitude difference. A magnitude difference greater than T_m leads to 1 in s', and a magnitude difference less than -T_m leads to -1 in s', otherwise, encoded as 0. Finally, *D(i)* represents the combination of magnitude and direction. If the pixels monotonically increase or decrease in one direction with magnitude differences larger than the threshold value (T_m), $D(i)$ is encoded as 1; otherwise, it is encoded as 0. The main purpose of the magnitude and direction is to well-preserve the contours of face.

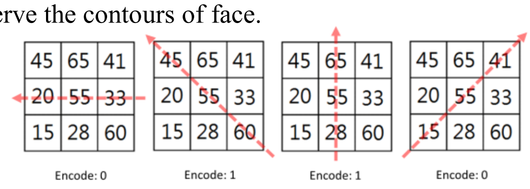

Fig. 2 An illustration of magnitude and direction difference

We consider four directions as the arrows in Fig.2, and set the threshold value (T_m) to 3. The first and the fourth bits are encoded to 0 because pixels do not continuously decrease or increase along the dotted lines. On the other hand, pixels along the second and third dotted lines monotonically decrease or increase and both magnitude differences are greater than the threshold value (eg. 55- 45 = 10 and 60 - 55 = 5). As a result, the binary code is encoded as "0110".

2.2. Proposed descriptor

Like other LBP-based descriptor, a spatial histogram is chosen to represent the magnitude and direction differences and concatenated after the sign part. According to the example in Fig. 2, the length of the binary string of *D*(i) is half of that of a LBP string. In the case of 8 neighbors, it generates a short histogram with only 16 bins, which is quite compact. The new descriptors are called magnitude and direction difference for LBP (MD-LBP), magnitude and direction difference for uniform LBP (MD-ULBP), and magnitude and direction difference for WLBP (MD-WLBP), respectively. These descriptors with additional magnitude and direction difference have both the advantages of the original descriptors and the 2nd-order derivative. The LBP descriptor is insensitive to monotonic gray-level variations caused by illumination. The Uniform LBP is introduced to reduce the number of possible bins and improve the recognition rate. The WLBP is a combination of the uniform LBP and the WLD. Finally, the integration of magnitude and 2nd-order derivative can reveal the trends of variations. Table 1 briefly shows the comparison of above approaches.

Table 1: A brief comparison of different descriptors

	Intrinsic Properties	Feature size (In case of 8-neighbor)
LBP [1] [2] [3]	Insensitive to illumination	256 bin
Uniform LBP [3]	Greatly reduce the feature size from LBP	59 bin
WLD [12]	According to human perception	50 bin
WLBP [5]	Combination of Uniform LBP and WLD increases discrimination	295 bin
MD-LBP*	LBP with Magnitude and Direction Difference	272 bin
MD-ULBP*	Uniform LBP with Magnitude and Direction Difference	75 bin
MD-WLBP*	WLBP with Magnitude and Direction Difference	311 bin

* Our new descriptors which contain the magnitude and direction difference can reveal more details

3. EXPERIMENTS

The face recognition experiments were performed on 4 public databases which are Japanese female facial expression (JAFFE) [7], TEXAS 3D Face Recognition Database (Texas 3DFRD) [8][9], The Olivetti Faces Dataset [13] and Yale Face Database [14]. In the experiments, we compare the conventional LBP [1][2][3], the Uniform LBP [3], WLBP [5] and with our MD enhanced descriptors. We

also compare the performance of our descriptor with CNN model. We conduct four different experiments to evaluate the proposed descriptors on three types of images, i.e., grayscales images, depth images and 2D+depth images.
The recognition rate is defined as follows:

$$Recognition\,Rate = \frac{Number\;of\;\;correct\;\;matchs}{Number\;\;of\;\;test\;\;images} \times 100\% \quad (4)$$

Since our feature has high discriminability, a simple classifier, the K-nearest neighbors method is used for the classification tasks. The recognition rates are obtained with a 10-cross-validation strategy.

3.1. Experimental result on the JAFFE database

The JAFFE database [7] is used to conduct the comparative experiments between the proposed descriptors and conventional descriptors on gray-level images using female faces with different expressions. Fig. 3 shows some results of different descriptors after extracting features. It is obvious to see that our method (g), magnitude and direction difference, can reveal the contours of the face clearly.

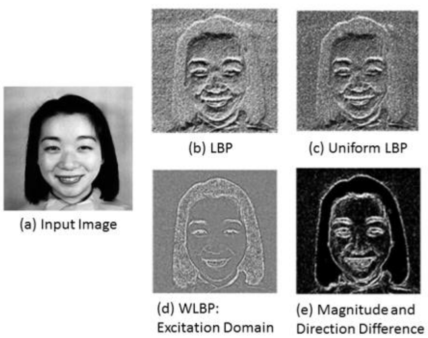

Fig. 3 Feature images after applying different descriptors. (a) Input image, (b) LBP, (c) Uniform LBP, (d) WLBP: Excitation Domain, (e) Magnitude and Direction Difference

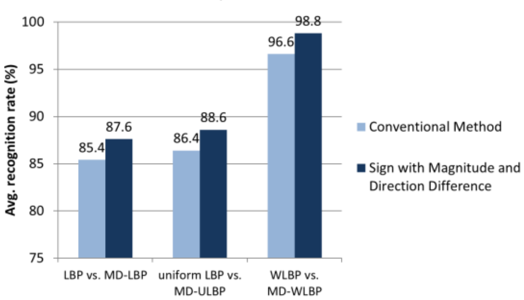

Fig. 4 The statistical chart shows a comparison between descriptors without/with our MD enhancement on the JAFFE database. Starting from the left is LBP, MD-LBP, uniform LBP, MD-uniform LBP, WLBP and MD-WLBP respectively. With the MD enhancement, the recognition rate is improved by 2.2% on average

In order to evaluate the proposed descriptor, we start by studying the impact of the threshold value. Due to the average difference between center pixel and its neighbors is about 7. As a result, we select 3 as the threshold value (T_m).

The experimental results using the K-nearest neighbors method shown in Fig. 4 demonstrates that the recognition rate is improved when taking magnitude and direction information into account. With the MD enhancement, the recognition rate is improved by 2.2% on average. Furthermore, the recognition rate of MD-WLBP can be as high as 98.8%.

3.2. Experiment on depth images of Texas 3DFRD

We further evaluate the performances of different descriptors on depth images with TEXAS 3DFRD [8][9]. Since the depth images are smoother than grayscale images and the intensity values of the neighboring pixels are close, WLBP is not suitable for depths images. As a result, we compare the LBP, Uniform LBP with MD-LBP and MD-ULBP. We select the quotient of two divided by average difference of center and its neighbors as the threshold value (T_m). The feature images are shown in Fig. 5. Our descriptor, magnitude and direction difference (f), shows salient variation especially the nose.

We can find in Fig. 6 that the conventional LBP has a recognition rate as low as 63.8% and uniform LBP also needs to be strengthened. With the magnitude and direction information, the recognition rate can be greatly improved, especially in the case of MD-ULBP. It achieves a recognition rate as high as 93.4%, which is the only descriptor exceeding 90%. Our idea consists in sign, magnitude and direction information which can represent the face depth images in more details and surpass the conventional descriptors.

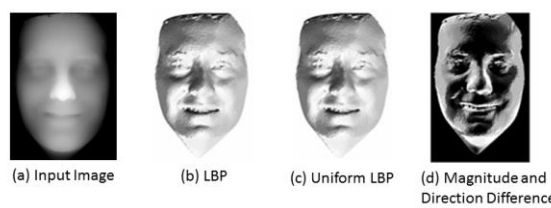

Fig. 5 Feature extraction of different descriptors (a) Input image, (b) LBP, (c) Uniform LBP, and (d) Magnitude and Direction Difference

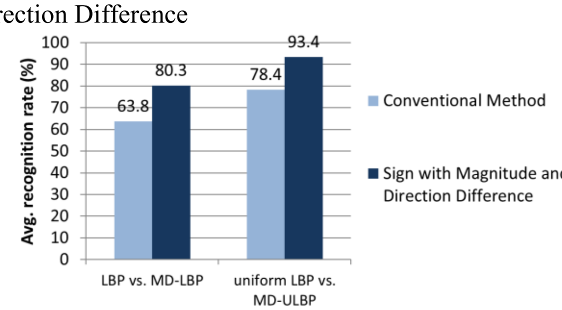

Fig. 6 The comparison of descriptors without/with magnitude and direction difference for depth images on the Texas 3D Face Recognition database shows that our new descriptors surpasses conventional methods. Furthermore, MD-ULBP is the most robust, which can achieve a recognition rate greater than 90%

3.3. Experiment on 2D + depth images of Texas 3DFRD

We also performed the experiment on 2D+depth images of the TEXAS 3D Face Recognition Database to prove that our descriptors can be used widely on 2D+depth face model.

As shown in Fig. 7, MD-ULBP achieves the best recognition rate. Once again, the proposed MD descriptor can effectively improve the recognition rate. The recognition rate of conventional LBP is 69.8%, while that of MD-LBP arises to 80.8%. The recognition rate of ULBP is 89.93% while that of MD-ULBP can reach 96%, which is the best result among all. Our proposed MD descriptor shows on average an improvement of 8.5%. Besides, our method is also superior to other LBP-based descriptors. We owe this to the superiority of retaining magnitude and direction difference information, which can provide detailed information of face contour.

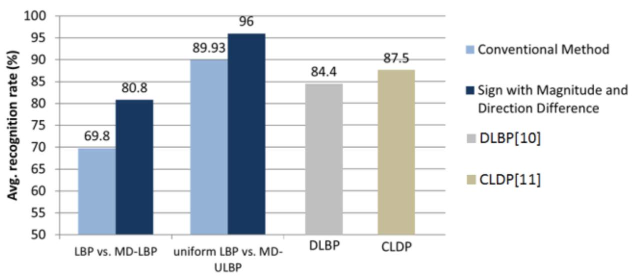

Fig. 7 Recognition rate of different 2D+depth face images recognition on Texas 3D Face Recognition database with K-nearest neighbor classifier

3.4. Comparative experiment on MD-WLBP and CNN

In order to prove the performance of our descriptor, we compare MD-WLBP, which achieves highest recognition rate in 2D experiment to Convolutional neural network (CNN) algorithm. Lots of researchers [15] have proved the performance of CNN in image processing and classification. Basing on the Theano framework [15], we build a CNN classification model, which contains 7 layers (one input layer, two convolutional layers, two pooling layers, one full connection layer and one output layer). The most important parameters in CNN model are set as: learning rate = 0.25, mini batch size = 20, the kernel of the first convolutional layer = 20, the kernel of the second convolutional layer = 40, pool size = (2,2), iteration = 2500.

We combine JAFFE [7], The Olivetti Faces Dataset [13], Texas 3DFRD [8] [9] and Yale Face Database [14] into a larger dataset. There are 80 identities and 10 images for each. 80% images are selected as training data randomly and the rest are testing images. Before face recognition task, we firstly detect the face of each image using Haar Feature-based Cascade Classifiers to crop the face portion.

The experimental result is shown in Fig. 8. We provide the scalability result of different number of images. When taking more identities into account, the accuracy rate of our descriptor decreases slightly, however, the processing time of CNN increase much faster than ours.

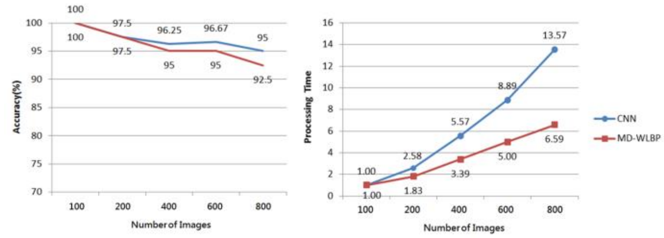

Fig. 8 (left) Accuracy rates of different number of images of MD-WLBP and CNN. (right) Concerning that different hardware and software may affect the processing time. The value represents the ratio of different number of images to 100 images of each descriptor. When taking more identities, the processing time of CNN increases quickly.

The advantages of our descriptor are simple, no training time consuming and we do not have to adjust many parameters such as learning rate and iteration of CNN model.

4. CONCLUSION AND FUTURE WORK

This paper presents a new descriptor, magnitude and direction difference, to enhance conventional LBP and WLBP descriptors for face recognition including grayscale, depth images and 2D+depth images. The descriptors with magnitude and direction difference surpass the conventional counterparts. This is due to the coherence of the magnitude and direction (2nd-order derivative) information can preserve the advantages of original methods and further can reveal the salient trends of variations between data. Furthermore, we compare our descriptor with CNN to further evaluate the applicability. Our method has an advantage in processing time and maintains over 90% accuracy.

The main contributions of this paper include:

1) The performances of our new descriptor surpass that without the magnitude and the direction information. Experiments conducted on various types of images show that models with magnitude and direction difference achieve better performances than conventional LBP and WLBP. This is due to the coherence of the conventional methods and further the magnitude and the direction difference.

2) The proposed descriptor is compact and effective. The binary string takes only half length of the neighborhood size. When exploit the histogram, its size is $2^{(R/2)}$ times smaller than conventional LBP. In addition, our descriptor increases much slower than CNN when considering more identities.

3) Our descriptor is widely applicable to different input resources. According to the extensive experiments, our method achieves the highest recognition rate on 2D grayscales images, depth images and also 3D images formed by 2D+depth images.

Our descriptor is simple, compact, well-preserving the face contour and achieves high accuracy for at least 80 identities. Later, we will conduct experiments on larger dataset and compare with more deep face descriptors [17]. We expect that the proposed descriptors are applicable to larger datasets and other object recognition tasks as well.

5. REFERENCE

[1] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions." *Pattern Recognition*, 29(1):51 –59, 1996.

[2] T. Ahonen, M. Pietikainen, A. Hadid, and T. Maenpaa, "Face recognition based on the appearance of local regions," In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, volume 3, pages 153–156 Vol.3, Aug 2004.

[3] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(12):2037–2041, Dec. 2006.

[4] A. Suruliandi, K. Meena, and R. R. Rose, " Local binary pattern and its derivatives for face recognition." *IET Computer Vision*, 6(5):480–488, Sept 2012.

[5] F. Liu, Z. Tang, and J. Tang, "WLBP: weber local binary pattern for local image description," *Neurocomputing*, 120:325–335, 2013.

[6] J. Chen, S. Shan, C. He, G. Zhao, M. Pietikainen, X. Chen, and W. Gao. Wld: A robust local image descriptor. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9):1705–1720, Sept 2010.

[7] The japanese female facial expression (jaffe) database. http://www.kasrl.org/jaffe.html

[8] S. Gupta, M. K. Markey, and A. C. Bovik. Texas 3d face recognitiondatabase. http://www.face-rec.org/databases/

[9] S. Gupta, M. K. Markey, and A. C. Bovik. Anthropometric 3d face recognition. *International Journal of Computer Vision*, 90(3):331–349, 2010.

[10] A. Aissaoui, J. Martinet, and C. Djeraba. Dlbp: A novel descriptor for depth image based face recognition. *In 2014 IEEE International Conference on Image Processing (ICIP)*, pages 298–302, Oct 2014.

[11] S. Yin, X. Dai, P. Ouyang, L. Liu, and S. Wei. A multi-modal face recognition method using complete local derivative patterns and depth maps. *Sensors*, 14(10):19561, 2014.

[12] J Chen, S. Shan, C. He, G. Zhao, M. Pietikainen, X. Chen, and W. Gao. Wld: A robust local image descriptor. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(9):1705–1720, Sept 2010.

[13] F. Samaria, A. Harter. Parameterisation of a Stochastic Model for Human Face Identification http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html

[14] http://vision.ucsd.edu/content/yale-face-database

[15] Theano: A Python framework for fast computation of mathematical expressions

[16] Bappaditya Mandal, Zhikai Wang, Liyuan Li, Ashraf A. Kassim. "Performance evaluation of local descriptors and distance measures on benchmarks and first-person-view videos for face identification," *Journal of Neurocomputing*, 2016.

[17] http://www.robots.ox.ac.uk/~vgg/software/vgg_face/