Recognizing Artificial Faces using Wavelet Based Adapted Median Binary Patterns

Abdallah A. Mohamed, Roman V. Yampolskiy
Computer Engineering and Computer Science
University of Louisville, Louisville, KY, 40292, USA
Email: {aamoha04, rvyamp01}@louisville.edu

Abstract

Recognizing avatar faces is a challenge and very important issue for terrorism and security experts. Recently some avatar face recognition techniques are proposed but they are still limited. In this paper, we propose a novel face recognition technique based on discrete wavelet transform and Adapted Median Binary Pattern (AMBP) operator to recognize avatar faces from different virtual worlds. The original LBP operator mainly thresholds pixels in a specific predetermined window based on the central pixel's value of that window. As a result the LBP operator becomes more sensitive to noise especially in near-uniform or flat area regions of an image. One way to reduce the effect of noise is to update the threshold automatically based on all pixels in the neighborhood using some simple statistical operations. Experiments conducted on two virtual world avatar face image datasets show that our technique performs better than original LBP, adapted LBP, Median Binary Pattern (MBP) and wavelet statistical adapted LBP in terms of accuracy.

Introduction

Face recognition is an automatic way to identify persons based on their faces' intrinsic characteristics (Zhenhua et al. 2010). It is one of the biometrics traits that received a great attention from many researchers during the past few decades because of its potential applications in a variety of civil and government-regulated domains. It usually involves: detecting the facial area, normalizing the detected faces, extracting facial features from appearance or facial geometry, and finally classifying facial images based on the extracted features (Mohamed, Gavrilova, and Yampolskiy 2012).

Face recognition however is not only concerned with recognizing human faces, but also with recognizing faces of non-biological entities or avatars (Gavrilova and Yampolskiy 2010). Quick investigation of Second Life and some other virtual worlds proves that these virtual worlds are populated by terrorist organizations such as Al-Qaeda and Second Life Liberation Army (Yampolskiy et al. 2012). Security experts believe that terrorist organizations used virtual worlds as an option to train their new members because of the presence of the USA army in Pakistan and Afghanistan where their previously training camps. Also, virtual worlds are very attractive for traditional criminals for conducting illegal gambling, fraud, identity theft and other traditional crimes. With the continuing growth of the virtual worlds' economies into billions of dollars and number of virtual crimes avatar recognition becomes a real problem for security experts and forensic investigators (Yampolskiy et al. 2012). There are a limited number of techniques that have been applied to recognize avatar faces such as what we performed in (Mohamed et al. 2011). These techniques are mainly based on applying Local Binary Pattern (LBP) descriptors for analyzing pure images.

LBP is sensitive to noise especially in near uniform or flat areas since it thresholds all pixels in the neighborhood based on the gray value of the central pixel of that neighborhood (Ahonen, Hadid, and Pietikainen 2006). One solution for this problem is to compute this threshold automatically from the available data. Therefore, we defined a new version of LBP called Statistical Adapted Local Binary Pattern (SALBP). The SALBP operator has a threshold computed automatically from the available data after redefining each pixel value in a local patch using a local weight. Applying the SALBP threshold mainly based on a scaling factor k where $0 \leq k \leq 1$ (Mohamed and Yampolskiy 2012). Changing the value of k has a great effect on the recognition rate. Therefore, the SALBP threshold is not completely automatically computed.

The main contribution in this paper is by defining a new LBP descriptor, Adapted Median Binary Pattern (AMBP) that has a completely automatically computed threshold to make LBP more robust to noise. The new computed threshold has no scaling factor or any other fixed value but it is a dynamic changeable threshold computed automatically from the available data. The efficacy of our proposed method is demonstrated by experiments on two different avatar face datasets from Second Life and Entropia Universe virtual worlds.

The rest of this paper is organized as follows; Section 2 briefly provides an introduction to discrete wavelet transform. In Section 3, an overview of the LBP is presented. Section 4 presents the median binary pattern (MBP) operator. The Adapted Median Binary Pattern (AMBP) operator is described in section 5. Section 6 explains the steps that we have to follow to apply the wavelet AMBP. Section 7 reports experimental results which followed by conclusions in Section 8.

Discrete Wavelet Transform

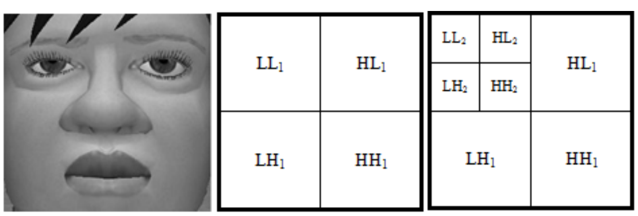
Discrete Wavelet Transform (DWT) is a widely used tool for image compression and texture classification because of its effective ability for multi-resolution decomposition analysis (Garcia, Zikos, and Tziritas 1998). It was also used to extract the essential features for avatar face recognition. In case of images we have to apply WT in two directions (row or horizontal direction and column or vertical direction) using four filters:

$$
\begin{aligned}
&\varphi(n_1, n_2) = \varphi(n_1)\,\varphi(n_2) \\
&\psi^H(n_1, n_2) = \psi(n_1)\,\varphi(n_2) \\
&\psi^V(n_1, n_2) = \varphi(n_1)\,\psi(n_2) \\
&\psi^D(n_1, n_2) = \psi(n_1)\,\psi(n_2)
\end{aligned}
\tag{1}
$$

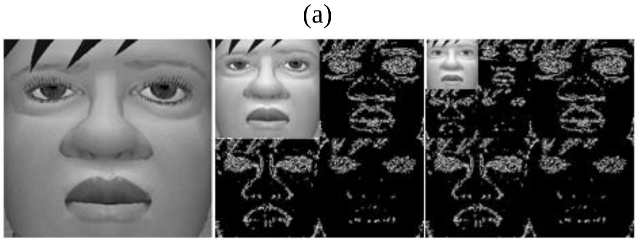
where n_1 is the horizontal direction and n_2 is the vertical direction, φ is the scaling function which is essentially a low pass filter, ψ is the wavelet function which is essentially a high pass filter, the product $\varphi(n_1)\ \psi(n_2)$ means applying the low pass filter in the horizontal direction and applying the high pass filter in the vertical direction, by the same way we can understand the meanings of all the four filters. In the second filter there is a super script H since there is a high pass filter applied on the horizontal direction, by the same way we can understand the superscripts V and D (Mazloom and Ayat 2008).

As a result of applying the four filters an image will be decomposed into four sub-bands LL, HL, LH and HH (see Fig. 1). The band LL represents an approximation to the original image while bands LH and HL represent respectively the changes of the image along the vertical and horizontal directions. The band HH records the high frequency component of the image.

The LL sub-band of the current decomposition level has to be decomposed to obtain the next level of decomposition.



(a)



(b)

Figure1. (a) Wavelet coefficient structure (b) Original image, One and two levels wavelet decomposition for an image.

Local Binary Pattern

LBP descriptor, proposed by Ojala et al., is a very simple and efficient local descriptor for describing textures (Ahonen, Hadid, and Pietikainen 2006). It labels an image pixels by thresholding the neighborhood of each pixel based on the central pixel value of that neighborhood (see Fig. 2), multiplied by powers of two and then added together to form the new value (label) for the central pixel. The output value of the LBP operator for a block of 3x3 pixels can be defined as follows (Ahonen, Hadid, and Pietikainen 2006):

$$LBP_{P,R} = \sum_{p=0}^{7} 2^p S(g_p - g_c) \tag{2}$$

where g_c corresponds to the gray value of the central pixel, g_p (p = 0,1,2,..,7) are the gray values of its surrounding 8 pixels and $S(g_p$ - $g_c)$ can be defined as follows (Ojala, Pietikainen, and Maenpaa 2002):

$$S(g_p - g_c) = \begin{cases} 1, & g_p - g_c \geq 0 \\ 0, & otherwise \end{cases} \tag{3}$$

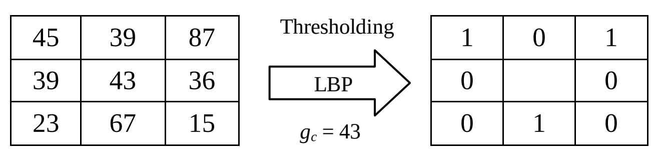
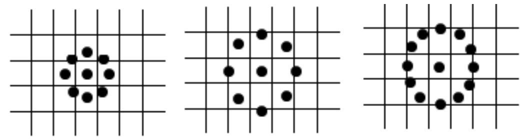


Figure 2. LBP operator applied over a 3x3 neighborhood.



(P=8, R=1) (P= 8, R=1.5) (P=12, R=1.5)

Figure 3. Three different LBP operators

Later new versions of LBP operator have been emerged as an extension to the original one and they used neighborhoods of different sizes to be able to deal with large scale structures that may be the representative features of some types of textures (See Fig. 3).

Median Binary Pattern

Instead of thresholding all pixels in the neighborhood against the central pixel value, the Median Binary Pattern (MBP) derives the localized binary pattern representation by thresholding against the neighborhood median value as in Fig. 4 (Hafiane, Seetharaman, and Zavidovique 2007).

45	39	87
39	43	36
23	67	15

Thesholding

MBP

Med = 39

1	1	1
1		0
0	1	0

Figure 4. Example of applying MBP over a 3x3 neighborhood

According to this new thresholding approach, the LBP operator and decision function defined by equations 2 and 3 respectively have to be changed into new format:

$$MBP = \sum_{p=0}^{7} 2^p S(a_p) \tag{4}$$

$$S(a_p) = \begin{cases} 1, & if \quad a_p \geq Med \\ 0, & otherwise \end{cases} \tag{5}$$

where a_p is the intensity value for any pixel p and *Med* is the median value for the neighborhood.

Adapted Median Binary Pattern (AMBP)

Thresholding all pixels in an image patch based on the central pixel value or the median value of this patch has a negative effect on how the LBP operator will deal with these pixels if they are noisy. One of the best solutions to overcome the high sensitivity of the LBP to noise is by adapting all pixels values in that patch according to the weight of each pixel in that patch.

Compute the Local Weight for Each Pixel

To compute the local weight for each pixel in any local patch of pixels we have to define the following equation (Mohamed and Yampolskiy 2012):

$$J = (g_c(i,j) - \sum_{q=1}^{P} w_q g_q(i,j))^2 \tag{6}$$

where g_c is the central pixel value for any local patch, g_q is any pixel value in that patch, and $\sum_{q=1}^{P} w_q = 1$

This equation minimizes the overall differences between the central pixel in any neighborhood and all pixels in that neighborhood. By deriving both sides of equation 4 with respect to w_p we get (Mohamed and Yampolskiy 2012):

$$\frac{\partial J}{\partial w_p} = -2 g_p(i,j)(g_c(i,j) - \sum_{q=1}^{P} w_q g_q(i,j)) = 0$$

then:

$$g_c(i,j) = \sum_{\substack{q=1 \\ q \neq p}}^{P} w_q g_q(i,j) + w_p g_p(i,j) \tag{7}$$

From equation 7 we can obtain the value of w_p using the following equation (Mohamed and Yampolskiy 2012):

$$w_p = \frac{g_c(i,j) - \sum_{\substack{q=1 \\ q \neq p}}^{P} w_q g_q(i,j)}{g_p(i,j)} \tag{8}$$

For more explanation about how we can compute w_p for different pixels we have to follow the following steps:

1- Initialization $w_p = 1/P$ for every $p = 1,2,\ldots,P$

2- Use the updated equation 8

Repeat

For $p = 1 : P$

Update w_p with the new value and each time when we compute the new value of w_p for any pixel use this value to compute the w_p for the next pixels.

end

AMBP Operator

By applying the previous steps, we can compute a local weight for each pixel. Therefore we can use these weights to compute a new value for each pixel and hence we can use the new pixels values to build the AMBP as follows:

$$AMBP = \sum_{p=0}^{7} 2^p S(g_p * w_p) \tag{9}$$

$$S(g_p * w_p) = \begin{cases} 1, & if \quad g_p * w_p \geq AMed \\ 0, & otherwise \end{cases} \tag{10}$$

where w_p is the local weight, g_p is the intensity value for pixel p and *AMed* is the adapted median value for the neighborhood. Figure 5 displays a comparison between LBP, SALBP and AMBP for the same local patch.

Wavelet Adapted Median Binary Pattern

In this section, we present the steps that we have to follow in order to apply our technique on different datasets.

Preprocessing Facial Image Dataset

We followed a set of preprocessing operations to improve the efficiency of applying AMBP operator to extract the images features. First, the input images are cropped to pure facial images by removing the background on the basis of each image should include its two eyes, mouth and nose.

Second, pure facial images have to be normalized and then finally decomposed using wavelet transform. Decomposing an image produces four coefficients. One of

63	68	42
64	64	27
61	95	83

threshold $q_c = 64$

0	1	0
1		0
0	1	1

LBP

65	69	42
66	64	27
60	95	83

threshold $q_c = 64$

1	1	0
1		0
0	1	1

Noise: LBP

63	68	42
64	64	27
61	95	83

7.875	8.625	6.375
8		3.375
7.625	11.875	10.375

1	1	0
1		0
1	1	1

SALBP with $k = 0.3$

65	69	42
66	64	27
60	95	83

8.125	8.625	5.875
8.25		3.375
7.5	11.875	10.375

1	1	0
1		0
1	1	1

Noise: SALBP with $k = 0.3$

63	68	42
64	64	27
61	95	83

7.875	8.625	6.375
8		3.375
7.625	11.875	10.375

0	1	0
1		0
0	1	1

AMBP

65	69	42
66	64	27
60	95	83

8.125	8.625	5.875
8.25		3.375
7.5	11.875	10.375

0	1	0
1		0
0	1	1

Noise: AMBP

Figure 5. Comparison of LBP, SALBP and AMBP

these coefficients, approximation coefficient, holds most of the original image features, so to obtain the next level of decomposition the approximation coefficient of the current level has to be decomposed. Therefore during our experiments we were concerned only with approximation coefficients obtained from Daubechies wavelet transform.

Extracting Features using AMBP

In this section we explain how we can compute the adaptive median binary code for each pixel. AMBP descriptor computes the new value for each pixel in an image as in the following steps:

- Starting from the first available local patch for an image compute the local weight for each pixel in that patch by using equation 8.
- Multiply the local weight its corresponding pixel to produce a new pixel value for each pixel in that local patch.
- Compute the updated median for the current local patch by using the new pixel values in that patch.
- Compute the AMBP binary code for the central pixel in that patch using equation 10. (See Fig. 5 as an example of how we can use the AMBP operator and its comparison with other techniques).
- The same steps have to be repeated to compute the adaptive median binary code for each pixel which will be transformed to its corresponding decimal representation.

Classification

We have used chi-square distance to compute the similarity distance between each input image and the training model as in the following equation:

$$D_{WAMBP}(X,Y) = \sum_{n=1}^{N} \frac{(X_n - Y_n)^2}{X_n + Y_n} \tag{11}$$

where X is the histogram of the testing image, Y is the training model histogram and N (N = 256 in this paper) is the number of histogram bins (Ahonen, Hadid, and Pietikainen 2006).

Experiments

In this section we present a briefly an overview of the datasets used in experiments (See Fig. 6) and the obtained results from these experiments.

Dataset Description and Experimental Setup

To evaluate the performance of our proposed technique, we applied it on two different avatar facial datasets.

The first dataset was acquired from Second Life (SL) virtual world avatar face dataset. This dataset contains 581 gray scale images with size 1280 x 1024 each to represent 83 different avatars (subjects). Each avatar subject has 7 different images for the same avatar with different frontal pose angle and facial expression.

The second dataset was collected from Entropia Universe (ENT) virtual world. ENT dataset contains 490 gray scale images with size 407 x 549 pixels. These images were organized in 98 subjects (avatars).

(a)

(b)

Figure 6. One subject of facial images from: a) Second Life dataset b) Entropia dataset

Each subject has 5 different images for the same avatar with different frontal angle and facial details.

To detect the facial area in each image automatically we used OpenCV with extended set of Haar-like features. The obtained results from manually cropping images were better than those obtained from automatically cropping images so by the end we decided to crop the facial part in each image manually.

Gaussian noise was added to both datasets. Each one of the two datasets was split in two independent sets. One set is used for training and the second set for testing. We used different sizes of both testing and training sets. We observed that selecting 3 images from each SL subject and 2 images from each ENT subject achieving the highest accuracy rate so we built our experiments based on that. All training images are randomly chosen while the rest are used for testing.

Experimental Results

We performed many tests to prove the efficiency of our proposed method in comparable to other methods. We performed these experiments using Daubechies wavelet transform families only but applying different discrete wavelet transform families may lead to different results. First we have to decide which family of Daubechies wavelet transforms is the best for dealing with our datasets. Table I summarizes the main tests carried out to retain the best wavelet family and its corresponding level of decomposition. From table I we can see that the best recognition rate for SL dataset can be achieved by Db5 wavelet family with the fourth level of decomposition and by Db7 with the seventh level of decomposition for the ENT dataset. Therefore we applied wavelets for both SALBP and AMBP with Db5 and the fourth level of decomposition for SL dataset and with Db7 and the seventh level of decomposition for ENT dataset.

Second we got the performance of original LBP, adapted LBP, MBP, WSALBP and WAMBP with different LBP operators applied on Gaussian noisy facial images from the two datasets (SL and ENT). We applied all techniques with R = 1, 2, 3 and P = 8, 16, 24. Changing the radius of the LBP operator or its size has a great effect on the performance of most of the techniques and within the same dataset, as we can see from figures 7 and 8.

Table I. Recognition rate for SL and ENT datasets

Wavelet Family	Level of Decomposition	Recognition Rate for SL	Recognition Rate for ENT
Db2	Level 1	72.29%	66.67%
	Level 2	73.49%	67.35%
	Level 3	76.51%	67.35%
	Level 4	78.92%	65.31%
	Level 5	84.38%	64.29%
	Level 6	78.01%	63.61%
Db4	Level 1	73.80%	67.69%
	Level 2	73.80%	68.39%
	Level 3	75.60%	70.75%
	Level 4	77.71%	71.77%
	Level 5	79.52%	68.39%
	Level 6	76.51%	66.67%
Db5	Level 1	74.10%	66.67%
	Level 2	76.51%	69.04%
	Level 3	77.71%	69.73%
	Level 4	**86.45%**	70.75%
	Level 5	82.83%	70.75%
	Level 6	77.71%	69.04%
Db7	Level 1	73.80%	68.39%
	Level 2	75.60%	69.73%
	Level 3	74.10%	70.75%
	Level 4	72.29%	72.45%
	Level 5	70.78%	**73.47%**
	Level 6	68.37%	69.04%
Db9	Level 1	68.37%	65.31%
	Level 2	72.29%	67.35%
	Level 3	74.10%	69.04%
	Level 4	76.51%	66.67%
	Level 5	75.60%	65.31%
	Level 6	73.80%	64.80%

The recognition rate of our method, WAMBP, is better than the recognition rate of all other compared techniques with almost all LBP operators by different values (see figures 7 and 8). For SL dataset, the recognition rate of WAMBP is greater than its closest competitor, MBP, by about 6% in average and the highest recognition rate is 86.45% that achieved with LBP (16, 2). For ENT dataset, the highest recognition rate is 73.47% that obtained with LBP (16, 3). The recognition rate of WAMBP in average is greater than its closest competitor, WSALBP, by about 5%. (See table II). Results obtained from figures 7 and 8 and table II prove the superiority of our technique when compared to other methods.

Table II. Average recognition rate for different algorithms

Dataset	LBP	ALBP	MBP	WSALBP	WAMBP
SL	40.29%	54.38%	69.65%	68.72%	75.91%
ENT	39.54%	50.19%	62.43%	62.54%	67.39%

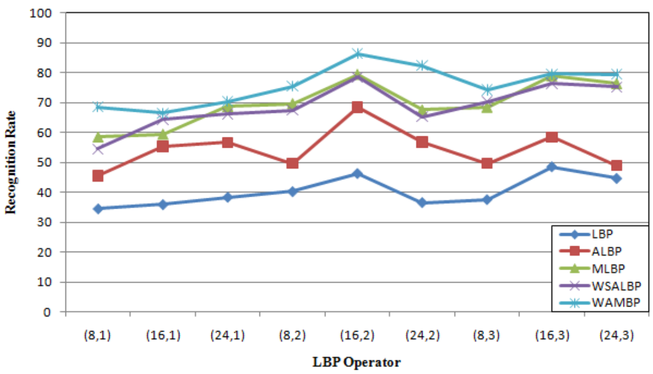


Figure 7. Recognition rate for different methods applied on noisy SL dataset.

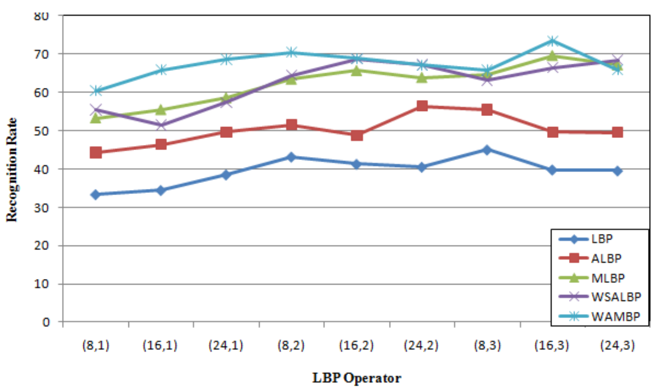


Figure 8. Recognition rate for different methods applied on noisy ENT dataset.

Conclusions

We proposed a novel LBP face recognition approach (WAMBP) in this paper. This approach based on combining the idea of wavelet transform with a new definition to the median local binary pattern to deal with one of the original LBP main limitations, its sensitivity to noise. Our approach computes the operator's threshold automatically from the available data and thresholds the new local pixel values according to its median. The effectiveness of this method is demonstrated on recognizing avatar faces from two noisy virtual world datasets. Compared to LBP ALBP, MLBP and WSALBP with LBP operators, our proposed technique improved the recognition rate for SL dataset by about 6% and by about 5% for ENT dataset. Redefine our approach using the concept of multi-scale definition is what we intend to apply in the future.

References

Zhenhua, G., Lei, Z., Zhang, D., and Xuanqin, M. 2010. Hierarchical multiscale LBP for face and palmprint recognition. 17th IEEE International Conference on Image Processing (ICIP), 4521-4524. Hong Kong.

Mohamed, A. A., Gavrilova, M. L., and Yampolskiy, R. V. 2012. Artificial Face Recognition Using Wavelet Adaptive LBP with Directional Statistical Features. International Conference on Cyberworlds (CW), 23-28. Darmstadt, Germany.

Gavrilova, M. L., and Yampolskiy, R. V. 2010. Applying Biometric Principles to Avatar Recognition. International Conference on Cyberworlds (CW), 179-186. Singapore.

Yampolskiy, R. V., Cho, G., Rosenthal, R., and Gavrilova, M. L. 2012. Experiments in Artimetrics: Avatar Face Recognition. Transactions on Computational Science 16:77-94.

Mohamed, A. A., D'Souza, D., Baili, N., and Yampolskiy, R. V. 2011. Avatar Face Recognition Using Wavelet Transform and Hierarchical Multi-scale LBP. 10th International Conference on Machine Learning and Applications and Workshops (ICMLA), 1:194-199. Honolulu, Hawaii.

Ahonen, T., Hadid, A., and Pietikainen, M. 2006. Face Description with Local Binary Patterns: Application to Face Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 28(12):2037-2041.

Mohamed, A. A., and Yampolskiy, R. V. 2012. Wavelet based Statistical Adapted Local Binary Patterns for Recognizing Avatar Faces. International Conference on Advanced Machine Learning and Applications, 92-101. Cairo, Egypt.

Garcia, C., Zikos, G., and Tziritas, G. 1998.A wavelet-based framework for face recognition. 5th European Conference on Computer Vision, 84-92. Freiburg, Germany.

Mazloom, M., and Ayat, S. 2008. Combinational Method for Face Recognition: Wavelet, PCA and ANN. International Conference on Digital Image Computing: Techniques and Applications (DICTA), 90-95. Canberra, Australia.

Ojala, T., Pietikainen, M., and Maenpaa, T. 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(7):971-987.

Hafiane, A., Seetharaman, G., and Zavidovique, B. 2007. Median Binary Pattern for Textures Classification. 4th International Conference on Image Analysis and Recognition, 387-398. Montreal, Canada.