Research Article

Wavelet-Based Multiscale Adaptive LBP with Directional Statistical Features for Recognizing Artificial Faces

Abdallah A. Mohamed¹, ² and Roman V. Yampolskiy¹

¹ Computer Engineering and Computer Science, University of Louisville, Louisville, KY 40292, USA
² Department of Mathematics, Menoufia University, Shebin El-Kom, Menoufia 32511, Egypt

Correspondence should be addressed to Abdallah A. Mohamed, abdmoham@umail.iu.edu

Received 21 August 2012; Accepted 29 September 2012

Academic Editors: E. Y. Du and D. P. Mukherjee

Copyright © 2012 A. A. Mohamed and R. V. Yampolskiy. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Recognizing avatar faces is a very important issue for the security of virtual worlds. In this paper, a novel face recognition technique based on the wavelet transform and the multiscale representation of the adaptive local binary pattern (ALBP) with directional statistical features is proposed to increase the accuracy rate of recognizing avatars in different virtual worlds. The proposed technique consists of three stages: preprocessing, feature extraction, and recognition. In the preprocessing and feature extraction stages, wavelet decomposition is used to enhance the common features of the same subject of images and the multiscale ALBP (MALBP) is used to extract representative features from each facial image. Then, in the recognition stage the wavelet MALBP (WMALBP) histogram similarity with statistical features of each test image and each class model is used within the nearest neighbor classifier to improve the classification accuracy of the WMALBP. Experiments conducted on two virtual world avatar face image datasets show that our technique performs better than LBP, PCA, multiscale local binary pattern, ALBP, and ALBP with directional statistical features (ALBPF) in terms of the accuracy and the time required to classify each facial image to its subject.

1. Introduction

Biometrics is the study of methods of recognizing humans based on their behavioral and physical characteristics or traits [1]. Face recognition 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 initial image normalization, preparing an image for feature extraction by detecting the face in that image, extracting facial features from appearance or facial geometry, and finally classifying facial images based on extracted features.

Face recognition, however, is not only concerned with recognizing human faces, but also with recognizing faces of nonbiological entities or avatars. To address the need for a decentralized, affordable, automatic, fast, secure, reliable, and accurate means of identity authentication for avatars, the concept of artimetrics has emerged [2, 3]. Artimetrics is a new area of study concerned with visual and behavioral recognition and identity verification of intelligent software agents, domestic and industrial robots, virtual world avatars, and other nonbiological entities [2, 3]. People often complain about the insufficient security system in the Second Life which motivates our research on security in virtual worlds [2, 4].

Extracting discriminant information from a facial image is one of the key components for any face recognition system [1]. There are many different algorithms proposed in the past to extract features, such as principal component analysis (PCA) [5], linear discriminant analysis (LDA) [6], local binary pattern (LBP) [7–10], multi-scale local binary pattern [11], and local binary pattern with statistical features [12].

LBP operator has proven itself as a powerful texture descriptor providing excellent results in terms of accuracy in many applications such as motion detection, image retrieval, remote sensing, and biomedical image analysis. Among all these applications, LBP method has shown its superiority in recognizing faces [1]. LBP is one of the most popular local feature-based methods. It was first proposed by Ojala et al.
Section 8. Section 6 presents our proposed technique. Section 7 reports with directional statistical features is described in Section 5. The adaptive local binary pattern (ALBP). The adaptive local binary pattern with wavelet transform to recognize avatar faces. Mohamed et al. applied discrete wavelet transform with adapted local binary patterns with direction statistical features to recognize avatar faces [15].

In this paper, we propose a new face recognition technique to recognize avatar faces from different virtual worlds. In this approach, we combine the discrete wavelet transform with the multi-scale adaptive local binary pattern with directional statistical feature (MALBP) operators. It is wellknown that applying multisresolution analysis for face recognition has very impressive results, but at the same time the computational complexity of the system and the features dimensions are very high [11]. To overcome these problems, we decomposed all facial images using discrete wavelet transform into a specific level of decomposition with the MALBPF. So, our proposed technique uses wavelet transform to enhance the common features of the same class of facial images to improve the recognition performance. Also, it computes the mean and the standard deviation of the local absolute difference between each pixel and its neighbors (in a specific patch of pixels) within the adaptive local binary pattern (ALBP) and the nearest neighbor classifier to improve the accuracy rate. The efficacy of our proposed method is demonstrated by the experiments on two different avatar 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 wavelet decomposition. In Section 3, an overview of the LBP with directional statistical features is presented. Section 4 presents the adaptive local binary pattern (ALBP). The adaptive local binary pattern with directional statistical features is described in Section 5. Section 6 presents our proposed technique. Section 7 reports experimental results which are followed by conclusions in Section 8.

2. Review of Wavelet Transform

Wavelet transform (WT) is a popular tool for analyzing images in a variety of signal and image processing applications including multi-resolution analysis, computer vision, and graphics. It provides multiresolution representation of the image which can analyze image variation at different scales. Many articles have discussed its mathematical background and advantages [16]. WT can be applied in image decomposition for many reasons [16] as follows.

(i) WT reduces the computational complexity of the system by producing lower resolution images (subimages) instead of operating on the original images with much higher resolution. For example, applying WT to reduce the resolution of an image from size $128 \times 128$ to size $32 \times 32$ will reduce the computational load by a factor of 16.

(ii) WT decomposes images into subimages corresponding to different frequency ranges, and this can lead to reduction in the computational overhead of the system.

(iii) Using WT allows obtaining the local information in different domains (space and frequency) while Fourier’s decomposition concerns only global information in the frequency domain. Thus it supports both spatial and frequency characteristics of an image at the same time.

WT decomposes facial images into approximate, horizontal, vertical, and diagonal coefficients. Approximate coefficient of one level is repeatedly decomposed into the four coefficients of the next level of decomposition. The process goes on until you find the required level of decomposition. Decomposing an image with the first level of WT provides four subbands LL, HL, LH, and HH.

The subband LL represents the approximation coefficient of the wavelet decomposition and it has the low-frequency information of the face image [17]. This information includes the common features of the same class. The other subbands represent the detailed coefficients of the wavelet decomposition and they have most of the high-frequency information of the face image. This information includes local changes of face image such as illumination and facial expression. To improve recognition performance we have to enhance the common features of the same class and remove changes. So, during our experiments we considered only the approximation images. Decomposing an image with two scales will give us seven subbands [16]: LL, HL, LH, HH, HL, LH, and HH, as in Figure 1.

3. Local Binary Pattern (LBP) with Directional Statistical Features

3.1. LBP Operator. LBP operator, proposed by Ojala et al. [13], is a very simple and efficient local descriptor for describing textures. It labels the pixels of an image by thresholding the pixels in a certain neighborhood of each pixel with its center value, multiplied by powers of two and then added together to form the new value (label) for the center pixel [9]. The output value of the LBP operator for a block of $3 \times 3$ pixels can be defined as follows [9]:

$$LBP_{p,r} = \sum_{p=0}^{7} 2^p s(g_p - g_c),$$ (1)
where $g_c$ corresponds to the gray value of the central pixel, $g_p$ \((p = 0, 1, 2, \ldots, 7)\) are the gray values of its surrounding 8 pixels, and $S(g_p - g_c)$ can be defined as follows:

$$
S(g_p - g_c) = \begin{cases} 
1, & g_p \geq g_c, \\
0, & \text{otherwise}.
\end{cases} \tag{2}
$$

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 \([7, 18]\).

The neighborhood of each pixel within an image can be either in a circular or square pattern (Figure 2 gives examples of circular neighborhoods for different sizes of the neighborhood and the radius values). In the following the notation \((P, R)\) will be used as indication of neighborhood configurations. $P$ represents the number of pixels in the neighborhood and $R$ represents the radius of the neighborhood.

One of the most important and successful extensions to the basic LBP operator is the uniform LBP (ULBP). An LBP is called uniform if the binary pattern contains at most two different transitions from 0 to 1 or 1 to 0 when the binary string is viewed as a circular bit string \([7]\). For example, 11000011, 00111110, and 10000011 are uniform patterns.

3.2. LBP Histogram. Suppose the given image is of size $N \times M$. To represent the whole texture image after computing the LBP pattern value for each pixel in that image, a histogram is built using \([18]\)

$$
H_i = \sum_{x,y} I(\text{LBP}(x, y) = i), \quad i = 0, 1, \ldots, n - 1, \tag{3}
$$

where $n = 2^P$ is the number of different labels produced by the LBP operator and $I(A)$ is a decision function with value 1 if the event $A$ is true and 0 otherwise.

The LBP histogram dissimilarity between a test samples $X$ and a class model $Y$ is computed using the chi-square distance:

$$
D_{LBP}(X, Y) = \sum_{n=1}^{N} \frac{(X_n - Y_n)^2}{X_n + Y_n}. \tag{4}
$$

3.3. LBP with Directional Statistical Features. Suppose that a given image is of size $N \times M$. Let $g_c$ be its central pixel and $g_p$ be its circular neighbors, where $p = 0, 1, \ldots, P - 1$. The mean ($\mu_p$) and the standard deviation ($\sigma_p$) of the local difference $|g_c - g_p|$ can be computed using the following two equations \([12]\):

$$
\mu_p = \frac{\sum_{i=1}^{N} \sum_{j=1}^{M} \left| g_c(i, j) - g_p(i, j) \right|}{(M \times N)}, \tag{5}
$$

$$
\sigma_p = \sqrt{\frac{\sum_{i=1}^{N} \sum_{j=1}^{M} \left( \left| g_c(i, j) - g_p(i, j) \right| - \mu_p \right)^2}{(M \times N)}}.
$$

$\mu_p$ and $\sigma_p$ represent the first-order and the second-order directional statistics of the local difference $|g_c - g_p|$ along orientation $2\pi p / P$ \([12]\). The vector $\vec{\mu}_X$ refers to the mean vector and $\vec{\sigma}_X$ refers to the standard deviation (std) vector.

The two vectors represent the directional statistical features of the local difference $|g_c - g_p|$, and they carry useful information for image discrimination that can be used to define the weighted LBP dissimilarity. Let $\vec{\mu}_Y$ and $\vec{\sigma}_Y$ refer to the directional statistical feature vectors for a sample test image $X$ while $\vec{\mu}_T$ and $\vec{\sigma}_T$ refer to the two vectors for a class
A new parameter called weight ($w_p$) called adaptive LBP (ALBP), to reduce the estimation error a new version of the LBP was proposed by Guo et al.,

4. Adaptive Local Binary Pattern (ALBP)

The directional statistical feature vectors can be used to improve the classification performance of an image by minimizing the variations of the mean and the std of the directional difference along different orientations. To this end a new version of the LBP was proposed by Guo et al., called adaptive LBP (ALBP), to reduce the estimation error of local difference between each pixel and its neighbors [12]. A new parameter called weight ($w_p$) is defined in the LBP equation, and so the new definition of the LBP equation will have the following form [12, 21]:

$$ALBP_{P,R} = \sum_{p=0}^{P-1} 2^p (p \ast w_p - g_p),$$  

where the objective function to compute the weight $w_p$ is as follows:

$$J = \sum_{i=1}^{N} \sum_{j=1}^{M} (g_c(i,j) - w \ast g_p(i,j))^2,$$

the target is to minimize the directional difference $|g_c - w_p \ast g_p|$, to this end we have to derive (9) with respect to $w$ and assign the derivation to zero as follows:

$$\frac{\partial f}{\partial w} = -2 \sum_{i=1}^{N} \sum_{j=1}^{M} \left( g_p(i,j) (g_c(i,j) - w \ast g_p(i,j)) \right) = 0,$$

so we get

$$w \sum_{i=1}^{N} \sum_{j=1}^{M} g_p(i,j) g_p(i,j) = \sum_{i=1}^{N} \sum_{j=1}^{M} g_p(i,j) g_c(i,j),$$

$$w = \frac{\sum_{i=1}^{N} \sum_{j=1}^{M} g_p(i,j) g_c(i,j)}{\sum_{i=1}^{N} \sum_{j=1}^{M} g_p(i,j) g_p(i,j)},$$

from (12) we get:

$$w_p = \frac{\mu_p}{\mu_p^2 + \sigma_p^2},$$

where $\mu_p = [g_p(1,1);g_p(1,2);\ldots;g_p(N,M)]$ is a column vector that contains all possible values of any pixel $g_p(i,j)$, $N \times M$ is the size of an image, and $\bar{g}_p = [g_p(1,1);g_p(1,2);\ldots;g_p(N,M)]$ is the corresponding vector for all $g_p(i,j)$ pixels.

Let $\bar{w} = [w_0,w_1,\ldots,w_{P-1}]$ refer to the ALBP weight vector. We have to note that each weight $w_p$ is computed along one orientation $2\pi p/P$ for the whole image.

5. ALBP with Directional Statistical Features

By using the ALBP weight, the directional statistics equations (5) can be changed to [12]

$$\mu_p = \sum_{i=1}^{N} \sum_{j=1}^{M} \frac{|g_c(i,j) - g_p(i,j) \ast w_p|}{(M \ast N)},$$

$$\sigma_p = \sqrt{\sum_{i=1}^{N} \sum_{j=1}^{M} \left(\frac{|g_c(i,j) - g_p(i,j) \ast w_p|}{(M \ast N)} - \mu_p\right)^2}.$$  

Based on the ALBP weight $w_p$, we have three vectors $\bar{\mu}$, $\bar{\sigma}$, $\bar{\mu}_X$, and $\bar{\sigma}_Y$. Similar to the normalized distance between $\bar{\mu}$ and $\bar{\mu}_X$, and $\bar{\sigma}$ and $\bar{\sigma}_Y$ we can define the normalized distance between $\bar{w}_X$ and $\bar{w}_Y$ as

$$d_w = \sum_{p=0}^{P-1} \frac{|\bar{w}_X(p) - \bar{w}_Y(p)|}{(P \ast e_w)},$$
where $\sigma$ is the standard deviation of $\hat{w}$ from training samples images [12, 20].

The weighted ALBP dissimilarity with statistical features using $d_\mu$, $d_\sigma$, and $d_w$ can be defined as

$$
D_{ALBP}^W(X, Y) = D_{ALBP}(X, Y) \ast \left( 1 + c_1 - c_1 \ast \exp \left( \frac{-d_\mu}{c_2} \right) \right) \\
\ast \left( 1 + c_1 - c_1 \ast \exp \left( \frac{-d_\sigma}{c_2} \right) \right) \\
\ast \left( 1 + c_1 - c_1 \ast \exp \left( \frac{-d_w}{c_2} \right) \right),
$$

(16)

where $D_{ALBP}(X, Y)$ is the ALBP histogram dissimilarity [12].

6. Wavelet-Based Multiscale ALBP with Directional Statistical Features

We have presented a general ALBP operator in Section 4 for extracting the facial images features using a single-scale circular symmetric neighbor set of $P$ pixels placed on a circle of radius $R$ with the weight parameter $w_p$. By altering $P$ and $R$ and combining the resulted images, a multiresolution representation can be obtained. However the main problem associated with the multiresolution analysis is the high dimensionality of the representation. There are some approaches to overcome this problem. One of these approaches minimizes the redundant information by applying feature selection techniques [22]. Another method reduces the dimensionality of the multiresolution representation by combining the multi-scale local binary pattern representation with linear discriminant analysis (LDA) to extract the features [11]. We propose another method to reduce the dimensionality by decomposing an image into a specific level of decomposition and then using the resulted approximation image for extracting the features.

6.1. Our Approach. In our approach, we combine Daubechies wavelet transform with the multi-scale adaptive local binary pattern representation. Decomposing an image with the first level of decomposition will produce four subbands (see Figure 1(a)). The subband LL represents a coarser approximation to the original image. The subbands HL and LH represent, respectively, the changes of the image along the horizontal and the vertical directions. The subband HH records the higher-frequency component of the image [16]. The decomposition, can be further extended to obtain the next levels of decomposition. Any one of the previous four subbands can be analyzed to obtain a higher level of decomposition. Images generally are very rich in the low-frequency contents so to obtain the next level of decomposition we have to analyze the LL subband of the current level of decomposition. We have to repeat the same steps until we reach to the required level of decomposition. This level differs from one dataset to another, and it is obtainable by practicing.

The adaptive local binary pattern operators at $L$ scales are then applied to the approximation facial image. This generates a new gray level code for each pixel at every resolution. The resulting ALBP images are divided into nonoverlapping subregions, $T_1, T_2, \ldots, T_q$, where $q$ is the number of sub-regions. The set of histograms computed at different scales for the same sub-region provides regional information about that region, and they have to be concatenated into a single histogram. This single histogram represents the final multiresolution regional face descriptor for this region.

Concatenating the final multiresolution regional face histogram for each region will form the final multiresolution face histogram for the whole facial image. By using the weighted ALBP dissimilarity with statistical features defined by (16) with the nearest neighborhood classifier for the histograms of both training and testing images, we can classify each image to its class.

7. Experiments

In this section, we verify the performance of the proposed algorithm on two different types of datasets: the first type is the Second Life (SL) dataset and the second is the Entropia Universe (ENT) dataset. Figure 3 gives an example of two subjects from each dataset. The proposed method is compared with single-scale LBP, traditional multi-scale LBP (MLBP), ALBP, and ALBP with directional statistical features (ALBPF).

7.1. Experimental Setup. To evaluate our proposed technique compared to other techniques, we have used two facial image datasets. The first dataset (SL) contains 581 gray images [23] with size $1280 \times 1024$ each to represent 83 different avatars. Each avatar subject has 7 different images for the same avatar with different frontal pose angle (front, far left, mid left, far right, mid right, top, and bottom) and facial expression.

The second dataset was collected from Entropia (ENT) Universe virtual world [24]. ENT dataset contains 490 grayscale images with size $407 \times 549$ pixels. These images were organized in 98 subjects (avatars). Each subject has different
5 images for the same avatar with different frontal angle and facial details (wearing a mask or not).

The facial part of each image in SL and ENT datasets was manually cropped from the original images based on the location of the two eyes, the mouth, and the nose. The new size of each facial image in SL dataset is 260 × 260 pixels while in ENT dataset each facial image was resized to the size of 180 × 180 pixels.

The performance of our method was affected by four parameters. The first one is the wavelet decomposition level. During our experiments we applied different levels of the Daubechies wavelet transform on both datasets. We figured out that the performance of our technique differs from one dataset to another and within the same dataset based on the decomposition level. So choosing the required level of decomposition is based on the dataset itself (see Figure 4). After applying the fourth level of decomposition on each facial image in the SL dataset, the resolution of each facial image was reduced to be 16 × 16 and after applying the third level of decomposition on each facial image in the ENT dataset, the resolution of each facial image was reduced to be 22 × 22. The second one is the circular neighborhood size $P$. Choosing a large size for the neighborhood increases the length of the histogram and then slows down the computation of the dissimilarity measure. Choosing a small size for the neighborhood size may lead to loss of information. So during our experiments, we have chosen a neighborhood of size $P = 8, 16$. The third parameter is the number of multi-scale operators. Using small number of operators cannot provide sufficient information about the facial images, also using large radius value reduces the size of the corresponding ALBP images. So in our experiments we have selected $L = 10$ which means that we have used 10 LBP operators to represent each facial image with $P = 8, 16$ and $R = 1, 2, 3, 4, 5$. The fourth parameter is the number of the facial image subregions $q$. Dividing the facial image into a large number of small sub-regions increases the computation time and may reduce the system accuracy while dividing the facial image into a small number of large subregions increases the loss of spatial information [11]. In our experiments each facial image has been divided into $q \times q$ nonoverlapping rectangle size sub-regions while the best value of $q$ is obtained by practicing.

7.2. Experimental Results. In order to gain better understanding on whether using wavelet transform with MALBP with directional statistical features (MALBPF) is advantageous or not, we compared WMALBPF with ALBPF, ALBP, MLBP, and LBP. First we got the average of recognition rate of WMALBPF, ALBPF, ALBP, MLBP, and LBP using the ten different LBP operators and with different number of regions (see Figure 4) over the SL and ENT datasets.

In this experiment the training sets were built by selecting the first image from each class of the SL and the ENT dataset while the rest are used for testing. The results showed that the average recognition rate of using WMALBPF is better than the average recognition rate of using the other methods with almost all values of $q$ and within the two datasets.

The recognition rate on average using WMALBPF is greater than that of its closest competitor, which is MLBP for SL datasets and ALBPF for ENT dataset, by about 7% and 3%, respectively. Comparing to other methods using wavelet transform with the MALBPf improves the recognition rate up to some point, and after that the recognition rate starts to be reduced based on the window size. As expected the recognition rate is reduced with large window (sub-region) size because of the loss of information. Based on the dataset itself and from Figure 5, it is shown that our technique provides a high and a robust average recognition rate especially when $8 \geq q > 4$ in case of the SL dataset and when $9 \geq q \geq 5$ in case of the ENT dataset.

The results showed also that not only the recognition rate of using WMALBPF is better than that of the other methods but also the time required to classify each input facial image to its class in case of using WMALBPF is less than that in the other methods with different LBP operators (see Table 1 for an example of the average time required in seconds to classify each facial image in the SL dataset to its class). This is an expected result since one of the main reasons of using wavelet decomposition in face recognition systems is that it reduces the computational complexity and overhead of the system and so the system can run faster.

It is very clear from Figure 5 and Table 1 that our proposed method can achieve better result than the other algorithms in terms of accuracy and classifying time.

8. Conclusion

In this paper, a novel LBP face recognition approach (WMALBPF) is proposed based on wavelet transform and adaptative local binary pattern with directional statistical features. The effectiveness of this method is demonstrated.
on recognizing faces from two different virtual worlds. Compared with ALBPF, ALBP, MLBP, and LBP with different LBP operators and different number of sub-regions \( q \), our proposed technique improved the recognition rate of the SL and ENT datasets by about 7% and 3% respectively. Also the time required by our technique to classify each input facial image to its class is less than the time in case of other methods.

Applying a hierarchical multi-scale definition to this approach may lead to better accuracy rate, and this is what we intend to try in the future.

**References**


