IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 1, 2021, accepted November 23, 2021, date of publication December 30, 2021,
date of current version January 12, 2022.

Digital Object Identifier 10.1109/ACCESS.2021.3139684

Non-Decimated Wavelet Based Multi-Band Ear Recognition Using Principal Component Analysis

MATTHEW MARTIN ZARACHOFF, AKBAR SHEIKH-AKBARI, AND DOROTHY MONEKOSSO
School of Built Environment, Engineering and Computing, Leeds Beckett University, Leeds LS1 3HE, U.K.

Corresponding author: Matthew Martin Zarachoff (m.zarachoff4868@student.leedsbeckett.ac.uk)

This work was supported by the Innovate UK through Knowledge Transfer Partnership under Grant KTP 10304. The work of Matthew Martin Zarachoff was supported by Leeds Beckett University.

ABSTRACT Principal Component Analysis (PCA) has been successfully applied to many applications, including ear recognition. This paper presents a 2D Wavelet based Multi-Band Principal Component Analysis (2D-WMBPCA) ear recognition method, inspired by PCA based techniques for multispectral and hyperspectral images. The proposed 2D-WMBPCA method performs a 2D non-decimated wavelet transform on the input image, dividing it into its wavelet subbands. Each resulting subband is then divided into a number of frames based on its coefficient's values. The multi frame generation boundaries are calculated using either equal size or greedy hill climbing techniques. Conventional PCA is applied on each subband's resulting frames, yielding its eigenvectors, which are used for matching. The intersection of the energy of the eigenvectors and the total number of features for each subband shows the number of bands which yield the highest matching performance. Experimental results on the images of two benchmark ear datasets, called IITD II and USTB I, demonstrated that the proposed 2D-WMBPCA technique significantly outperforms Single Image PCA by up to 56.79% and the eigenfaces technique by up to 20.37% with respect to matching accuracy. Furthermore, the proposed technique achieves very competitive results to those of learning based techniques at a fraction of their computational time and without needing to be trained.

INDEX TERMS Biometrics, principal component analysis, image recognition, wavelet transforms.

I. INTRODUCTION

Ear recognition, a field within biometrics, concerns itself with the use of images of the ears to identify individuals. Much like fingerprints, ears are unique to an individual; even identical twins can have distinguishable ears [1]. Researchers have explored this topic extensively over the last two decades, investigating both the feature extraction and comparison of features of ear images [2], [3]. Successful feature extraction techniques in ear recognition include Principal Component Analysis (PCA) [2], [4]–[7], wavelet based [8], and neural network based methods [9]–[11]. Amongst these techniques, PCA has been found to be successful for both feature extraction [4]–[6] and feature reduction to reduce dimensionality of the data [2], [3]. In general, PCA based techniques operate by converting an image into a 1D vector and concatenating those vectors to form a 2D matrix. While some PCA based ear recognition techniques have been reported in the literature [2], [4], [5], [12], these techniques involve projecting ear images into a common eigenspace. In contrast, the authors previously introduced a single image multi-band PCA based technique for ear recognition in [13], inspired by hyperspectral PCA based techniques [14] such as Segmented PCA [15] and Folded PCA [16], which solely use the extracted principal components as features.

To the authors' knowledge, no similar techniques that utilize wavelets have been reported in the literature. This has inspired the authors to propose a single image, PCA based method for ear recognition, called 2D Wavelet based Multi-Band PCA (2D-WMBPCA), which was originally introduced in EUSIPCO in [17]. Unlike the aforementioned PCA based methods, the proposed technique does not require the images to be projected into a common eigenspace. Instead, the proposed technique performs a 2D non-decimated wavelet transform on the input image, dividing the image into its wavelet subbands. Each resulting subband is then split into a number of frames according to its coefficient values using either of two methods: equal size and greedy hill climbing. The proposed technique then applies the standard PCA method on each subband's resulting set

The associate editor coordinating the review of this manuscript and approving it for publication was Syed Islam.

of frames, extracting their principal components. These features are then used for recognition. Experimental results on the images of two benchmark ear image datasets, named IITD II [18] and USTB I [19], demonstrated that the proposed 2D-WMBPCA technique greatly outperforms both Single Image PCA and the eigenfaces technique. Experimental results also show that the proposed technique achieves its highest performance when the number of frames used is equal to the intersection of the energy of the eigenvectors and the total number of features. The rest of the paper is organized as follows: Section II gives an overview of current techniques for ear recognition and hyperspectral PCA algorithms, Section III introduces the proposed 2D-WMBPCA technique, Section IV illustrates the multiple frame generation techniques, Section V describes the benchmark datasets used, Section VI discusses the experimental results, and finally Section VII concludes the paper.

II. RELATED WORK

The proposed 2D-WMBPCA algorithm is among the first of its kind to bring hyperspectral based techniques to the field of single image ear recognition. Consequently, this section is divided into two sub-sections. A brief literature review on ear recognition techniques is first presented. This includes a discussion of PCA based techniques and the current state of the art algorithms for ear recognition, which are learning based. The second subsection covers PCA based hyperspectral image classification algorithms that serve as the inspiration for this research.

A. EAR RECOGNITION TECHNIQUES

Since its inception, ear recognition has often gone hand in hand with facial recognition, as the two fields operate on very similar kinds of data. Much research has been conducted in the last decades in both categories [2], [3], [20]–[22]. A summary of existing techniques for ear recognition is tabulated in Table 1.

Emeršič *et al.* proposed a taxonomy of ear recognition techniques in [2], yielding four categories: geometric, local, holistic, and hybrid. Geometric approaches to ear recognition seek to establish the geometry of an ear from its image. This can be performed using a number of methods including Voronoi diagrams [28], outer ear points [29], and relative points on the ear contour [30]. However, these techniques generally depend on edge detection, which can produce erroneous results in cases of partial occlusion and poor lighting. In contrast, local approaches encode mainly texture information by computing local descriptors. Some of these methods first detect keypoints within the image and calculate the local descriptor for each keypoint such as the Scale Invariant Feature Transform (SIFT) [31], [32] and Speeded-Up Robust Features (SURF) [33] methods. These algorithms allow for partial matching and are robust to partial occlusion because they only operate on keypoints, but global ear structure information is discarded. Other local techniques densely calculate the local descriptors of the entire image, such as wavelet [34], curvelet [35], Gabor filters [36], log-Gabor filters [37], local binary patterns [38], and histogram oriented gradients [39].

Principal Component Analysis (PCA) and other related methods are examples of holistic techniques that extract features using the entire face or ear image, where the resulting features are used for recognition. The first use of PCA for biometrics, known as the ''eigenfaces'' technique, was reported by Turk and Pentland in [7]. This method uses a training set to calculate the eigenvectors that span the ''eigenface'' space and then project facial images into the resulting space to create ''eigenfaces''. The Euclidean distances between the resulting projections are then used to find the best match for a query image. Victor *et al.* [4] later applied this method to both ear and face images. They reported that the ''eigenface'' technique produces a higher matching accuracy when applied to facial images than ear images, however, the matching accuracy for ear recognition using the ''eigenface'' method was high enough to warrant further study. A similar experiment was later conducted by Chang *et al.* [5], again on both face and ear images. They concluded that the difference in accuracy between ear and facial image recognition is not statistically significant. Yang *et al.* introduced a technique for face recognition called Two-Dimensional Principal Component Analysis (2DPCA) in [12]. Unlike conventional PCA [7], 2DPCA uses 2D image matrices rather than a concatenated set of 1D image vectors. Their technique calculates a covariance matrix directly from a set of training input images. The eigenvectors of this covariance matrix are then used to project the testing images and the resulting projections are classified using nearest neighbor classification. The authors applied their technique to three benchmark datasets and reported a superior matching accuracy when compared to traditional PCA and other linear based methods. Furthermore, their technique required much less computation time than traditional PCA. Querencias-Uceta *et al.* [6] examined various techniques to fine-tune conventional PCA for ear recognition to increase the matching accuracy using the ''eigenface'' method presented in [4]. They used both Euclidean distance and Eigendistance to find the best match. The authors concluded that Euclidean distance results in a higher accuracy. Furthermore, their results show that increasing the number of training images increases the achieved accuracy. However, the number of the enrollment images has no effect on accuracy of the technique.

PCA and other linear transformation techniques have also been used in conjunction with other feature extraction and classification methods to create hybrid classifiers [8], [9], [23]–[25], [27], [40]. A hybrid ear recognition method based on PCA and neural network was reported by Alaraj *et al.* in [23]. This method uses a traditional multilayer feed-forward neural network with images as the input and a target matrix, which is a binary matrix, to indicate correct matches for the output. They have reported that the achieved recognition accuracy of their method is a function of the used number of training images per individual. Moreover, they showed that the use of a greater number of eigenvectors

TABLE 1. Past and state of the art results in ear recognition.

Statistical Based Techniques				
Author	**Feature Extractor**	**Dataset**	**Train/Test%**	**Accuracy%**
[4]	PCA	Own	72/28	40
[5]	PCA	UND E	Various	71.6
[8]	Wavelet Transform and PCA	USTB II	50/50	90.5
Learning Based Techniques				
Author	**Feature Extractor**	**Classifier**	**Dataset**	**Train/Test%**	**Accuracy%**
[23]	PCA	Neural Network	UND	80/20	96
[24]	ICA	RBF Network	CP	66/34	94.1
[25]	LDA, SURF	Neural Network	Own	33/67	97
[9]	Neural Network	Pairwise SVM	USTB I	67/33	98
[26]	BSIF	SVM	IITD II	56/44	97.31
[27]	BSIF	SVM	USTB I	32/68	96.53

increases the achieved accuracy of the matching. Last but not least, their investigation determined that their algorithm generates superior results in terms of accuracy when using larger images.

The eigenvectors generated by PCA have previously been used successfully as features for ear recognition [8], [40]. Nosrati, Faez, and Faradji demonstrated the use of PCA on 2D wavelet coefficients of ear images in [8]. The authors used a 2D wavelet transform to extract three sets of features for each image, which are then summed into a single matrix. PCA was then performed on the resulting matrix and the eigenvectors were used for matching. The authors reported a superior matching accuracy with their technique compared to those of PCA and Independent Component Analysis (ICA). Zhang *et al.* [40] compared the use of PCA with ICA, a similar method to PCA, to extract features from the ear images. PCA is first applied to the images to reduce their dimensionality, generating their respective eigenvectors. The ICA transform is then applied to the reduced resulting eigenvectors, generating a linear representation of the eigenvectors of ears, which have minimum dependencies amongst its components. The resultant features are then classified via a three-layer Radial Basis Function (RBF) Network. They concluded that ICA based method outperforms the PCA based method in terms of the accuracy.

Galdámez *et al.* [25] used both Linear Discriminant Analysis (LDA) and Speeded-Up Robust Features (SURF) to extract features for ear recognition. LDA is similar transform to PCA except it utilizes Fisher's linear discriminant than eigenvectors to create earspace dataset. SURF is a scale and rotation invariant interest point detector and descriptor method. The results of both LDA and SURF feature extraction are then fed to two three-layer feed-forward neural networks. Results demonstrated that SURF method outperforms the LDA algorithm and that both techniques give superior results to that of the PCA based method. Omara *et al.* [9] reported a hybrid ear recognition technique using neural network, PCA and Support Vector Machine (SVM). They have applied VGG-M Net [41], a commonly used convolutional neural network for image recognition, on all input images, extracting image features that are then reduced through PCA. Both SVM and pairwise SVM were separately applied to the selected features to find the best match. Their results show that the application of pairwise SVM for finding the best match results in higher accuracy than that of the traditional SVM.

Benzaoui and Boukrouche [27] introduced a feature extraction technique for ear recognition. They created a grayscale image for each ear image color component in the RGB, HSV, and YCbCr color spaces, and they then applied three feature extraction methods called: Local Binary Patterns (LBP), Local Phase Quantization (LPQ) and Binarized Statistical Images Features (BSIF), to the resulting gray images, creating a histogram representation for each grayscale image. These three histograms, which represent the ear features, are concatenated and then fed to the SVM classifier. Their experimental results show that BSIF technique, in conjunction with the RGB color space, generate the highest accuracy in finding the best match compared to LPQ and LBP methods. Zhang *et al.* proposed a technique for ear recognition with cases of partial occlusion in [24]. Their proposed methodology uses a non-negative sparse representation of each input image, treating it as a signal representation of a linear additive combination of all the training ear signals. They also introduce an algorithm that solves the proposed system and is guaranteed to find the global minimum. The authors report a perfect classification of test ear images when the ear image is not occluded and superior performance to the state of the art algorithms when the ear is occluded.

Recently, Emeršič *et al.* established the Unconstrained Ear Recognition Challenge (UERC), a challenge devoted to ear recognition using images captured in the wild [42]. They created a dataset by first establishing a list of celebrity names and then using a web crawler to find their corresponding

images. The resulting images were then segmented using convolutional encoder-decoder networks and then manually screened. This process resulted in a total of 11,804 ear images of 3,706 subjects. Six techniques were submitted for the challenge and compared with two baseline approaches: one based on Local Binary Patters (LBP) [43] and the second a Convolution Neural Network (CNN) based on the 16-layer VGG architecture [44]. The results across all techniques submitted for the UERC show that the chainlet based technique outperformed all others, with a 90.4% Rank-1 accuracy and a 100% Rank-5 accuracy. For comparison, the LBP-baseline produced a 14.3% Rank-1 accuracy and a 28.6% Rank-5 accuracy, while the VGG-baseline accuracies were 18.8% and 37.5% respectively. However, the chainlet based method uses a handcrafted descriptor; the best learned descriptor method was Safdaii's convolutional neural network technique [42] that exhibited a Rank-1 accuracy of 38.5% and Rank-5 accuracy of 63.2%.

B. HYPERSPECTRAL PCA BASED CLASSIFICATION METHODS

Hyperspectral images consist of many small electromagnetic bands, where each band represents the scene as perceived at certain wavelengths. Each hyperspectral image can be imagined as a data cube with coordinates (x, y, λ), where x and y correspond to a 2D position and λ specifies the electromagnetic band. The most common technique used for capturing a hyperspectral image is spatial scanning, which uses a ''pushbroom'' method to capture a scene line by line [45]. It uses a prism or a grating to divide the one-dimensional input line into its spectral components, forming a two-dimensional pattern representing input light's wavelength versus position. The resulting light pattern's intensities are then captured by a two-dimensional photo-sensor, generating a spatial line within a hyperspectral image. This process is repeated to capture a full two-dimensional scene. As it can be seen, hyperspectral imaging systems acquire images in over one hundred contiguous spectral bands, where each band represents one of the hyperspectral image wavelength resolutions, while RGB and gray imaging systems captures images over three bands (red, green and blue) and a single band, respectively. This is due to the different photo sensors that these imaging systems use.

The use of orthogonal projections including PCA for feature extraction from hyperspectral images has been investigated in [14]–[16], [46], [47]. Harsanyi and Chang reported a method for dimensionality reduction and hyperspectral image classification in [14]. They project each pixel vector within the hyperspectral image onto an orthogonal subspace to reduce interfering spectral signatures. The projected pixel vectors are then projected onto a second basis to maximize their signal to noise ratios, and the resulting components are finally used for classification.

Several extensions of PCA for hyperspectral images have been reported in the literature [15], [16], [46], [47]. Jia and Richards introduced a PCA based technique called Segmented Principal Component Transformation (Segmented PCT) for hyperspectral image classification [15]. Their proposed method first generates a covariance matrix for each pixel vector in the hyperspectral image. Each resulting covariance matrix is then normalized to produce a correlation matrix, an example of which can be seen in Fig. 1. A number of subgroups within the correlation matrix is then chosen and highly correlated bands are separated into these subgroups. PCA is performed on each resulting subgroup to extract their features. The process can be repeated multiple times using the extracted features as the input for Segmented PCT until the desired dimensionality reduction is achieved. For their presented results, the authors opted to separate the data into three subgroups, which they noted would reduce the total computation time by 2/3 if all subgroups were of uniform size. They selected highly correlated bands by only considering coefficients with absolute values exceeding 0.5 to find the boundaries of each subgroup. The authors were able to reduce the number of features used for classification of two benchmark images (Jasper Ridge and Moffett Field [48]) from 196 to seven and from 187 to six, respectively. Despite such a drastic reduction, their reported classification accuracy is still quite high, with Jasper Ridge having 98.6% and Moffett Field having 97.0%.

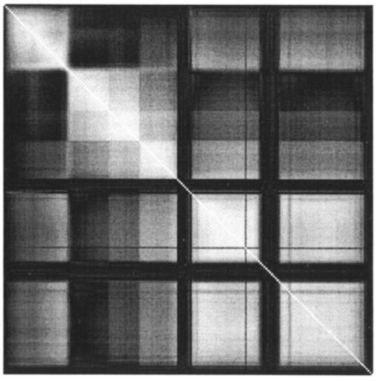

FIGURE 1. A 196 waveband correlation matrix for the Jasper Ridge dataset as presented in Jia and Richards [15].

While Segmented PCT offers savings in terms of computational cost when compared to PCA, it still requires the Eigen problem to be solved multiple times. Zabalza *et al.* proposed a similar PCA based technique called Folded-PCA for hyperspectral image classification in [16]. In Folded-PCA, each pixel vector within the hyperspectral image is used to create its own partial covariance matrix. This is accomplished by first folding the pixel vector into a 2D matrix. The pixel vector of length F, where F is the number of hyperspectral bands, is split into a number of groups H, where H is a hyperparameter that has been preselected. The resulting groups are then concatenated to form a 2D matrix for each pixel, as shown in Fig. 2.

Each resulting 2D matrix is then used to form a covariance matrix. The average of all resulting covariance matrices is

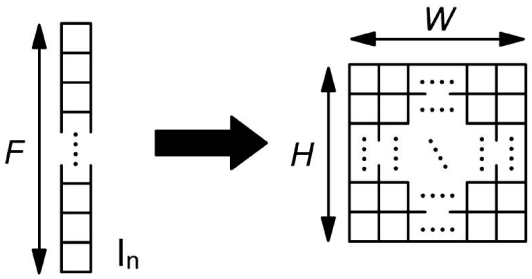

FIGURE 2. A pixel vector from a hyperspectral image is folded into a 2D matrix, as presented in Zabalza *et al.* [16].

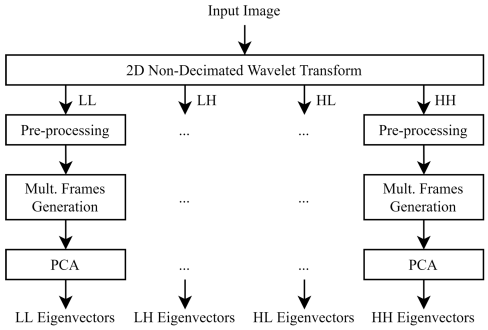

FIGURE 3. Block diagram of the proposed 2D-WMBPCA method.

then calculated and used as the covariance of the input hyperspectral image. PCA is then applied on the resulting overall covariance matrix, extracting its principal components, which are then used to classify the input image. The authors reported improved classification accuracy on two benchmark datasets (Indian Pines A and B [49]) when compared with both traditional PCA as well as Segmented PCT. On a third dataset, which is a SAR dataset [50] for target classification, classification accuracy was comparable for all three aforementioned PCA methods. Furthermore, the computation cost of Folded-PCA is approximately 10% and 7.65% lower on the Indian Pines datasets and the SAR dataset in comparison with that of standard PCA and Segmented PCT, respectively.

Holia and Thakar reported another PCA based technique, called Windowed PCA, to combine multiple views or modes for multi-focus and multi-modal images into a single focused image [51]. Their proposed algorithm divides the input images into a number of windows, where each set of windows corresponds to a particular region in all images and is referred to as a window block. It then applies standard PCA on each resulting window to extract its principal components. The resulting principal component vector associated with the largest eigenvalue is normalized. Every individual window within all input images is then multiplied by its resulting normalized vector. The resulting registered weighted windows within each window block are then summed. Finally, all window blocks are aggregated to form a registered image. The authors reported that the effectiveness of their algorithm increases with the number of windows. However, a larger number of windows increases the computation cost of the algorithm and the transition from one window to another window may be visible as the intensity value of the background may be different.

III. PROPOSED 2D-WMBPCA METHOD

A block diagram of the proposed 2D Wavelet based Multi-Band Principal Component Analysis (2D-WMBPCA) method is illustrated in Fig. 3. 2D-WMBPCA is inspired by state of the art PCA based techniques for hyperspectral images and wavelet based ear recognition algorithms. The proposed method first performs a 2D non-decimated discrete dyadic wavelet transform on the input ear image as introduced in [52], splitting the image into its four subbands (LL, LH, HL, HH). Each resulting subband is then pre-processed to improve its contrast, and is then fed to a multiple-frames generation algorithm, generating a number of frames based on the magnitude of the subband's coefficients. Finally, Principal Component Analysis (PCA) is applied to each resulting set of multiple-frames, extracting its eigenvectors. The eigenvectors from all four wavelet subbands are then concatenated and used for matching.

A. NON-DECIMATED WAVELET DECOMPOSITION

The 2D-WMBPCA algorithm begins by applying a 2D non-decimated wavelet transform on the input ear image e. A quadratic spline of compact support wavelet that is continuously differentiable, introduced in [53] which highlights the ear edges, is implemented using the non-orthogonal wavelet fast computation algorithm as seen in [52]. This decomposition splits the input image into four subbands, called Low-Low (LL), Low-High (LH), High-Low (HL), and High-High (HH). An example of the application of Mallat and Hwang's wavelet decomposition algorithm on a sample image is illustrated in Fig. 4.

B. SUBBAND PREPROCESSING

Coefficients within each subband s are then mapped to the [0,1] domain using (1):

$$p' = \frac{p - min(s)}{max(s) - min(s)} \tag{1}$$

where p represents an original coefficient in s and p' represents the corresponding mapped coefficient. Histogram equalization is then performed on the resulting subband coefficients to increase their contrast. To do so, the Probability Mass Function (PMF) of the resulting subband's coefficients is first calculated using (2):

$$P_X(x_k) = P(X = x_k) \quad \text{for } k = 0, 1, \ldots, 255 \tag{2}$$

where $X = x_1, x_2, \ldots, x_k$ represent the subbands' coefficients and $P_X(x_k)$ is the probability of coefficients in bin k. The resulting PDF is then used to calculate the Cumulative Distribution Function (CDF) of the subband using (3):

$$C_X(k) = P(X \leq x_k) \quad \text{for } k = 0, 1, \ldots, 255 \tag{3}$$

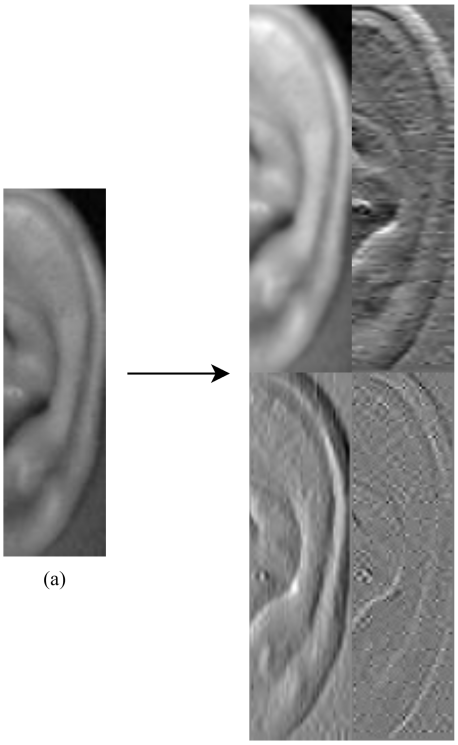

FIGURE 4. (a) A sample ear image from the IITD II dataset [18] (b) The non-decimated wavelet transform subbands (clockwise from upper left: LL, LH, HH, HL), where each subband's coefficients have been mapped to 0-255 for illustration purposes.

where $C_X(k)$ is the cumulative probability of $X \leq x_k$. Finally, all subband's coefficients are mapped to new values using the resulting CDFs. This improves the contrast of the subband's coefficients.

C. MULTIPLE FRAMES GENERATION

Various methods can be used by 2D Wavelet based Multi-Band PCA (2D-WMBPCA) to generate multiple frames from each input histogram equalized subband; two such methods are detailed in Section IV. In general, a multiple frame generation function f takes the form:

$$f : s \rightarrow F \tag{4}$$

where s is the input subband and $F = [f_1, f_2, \ldots, f_N]$ are the output frames.

D. PRINCIPAL COMPONENT ANALYSIS

For each frame $f \in F$, a mean adjusted frame f' is created using (5):

$$f' = f - \overline{f} \tag{5}$$

where $\overline{f}$ is the mean value of all coefficients in f. Every frame is then converted to a column wise vector, allowing F to be represented as a two dimensional matrix W. PCA is then performed using Singular Value Decomposition (SVD) on matrix W, creating the decomposition in (6):

$$W = U \Sigma V^T \tag{6}$$

where U is a unitary matrix and the columns of V are the orthonormal eigenvectors of the covariance matrix of W and Σ is a diagonal matrix of their respective eigenvalues. The eigenvectors form a basis for an eigenspace for each set of frames F. The resulting principal components in V are finally used for matching.

E. SUBBAND MATCHING

Let $M = [m_1, m_2, \ldots, m_{N-1}]$ be the set of principal components of a query image subband q. Furthermore, let $L = [l_1, l_2, \ldots, l_{N-1}]$ be the set of principal components of subband r, a subband of an image in the image dataset. The Euclidean distance D between q and r can be calculated using (7):

$$D = \sqrt{(\Sigma_n (m_n - l_n)^2)} \tag{7}$$

The resulting differences across all four subbands (LL, LH, HL, HH) are then summed to produce the total distance between the two images. The best match for query image q in the image database is the image with the lowest distance.

IV. MULTIPLE FRAME GENERATION TECHNIQUES

In this research, two frame generation techniques are presented.

A. EQUAL SIZE

The equal size multiple frame generation technique can be formulated as follows:

Assume s is the input subband and N is the number of desired frames to be generated from the subband s. The proposed algorithm uses $N - 1$ boundaries to split the input subband's coefficients into N target frames according to the coefficient values. Let $B = [b_1, b_2, \ldots, b_{N-1}]$ be the boundary values, calculated according to (8):

$$b_n = n/N \quad \text{for } n = 1, \ldots, (N-1) \tag{8}$$

The coefficients in each input subband s are divided into N target frames as follows:

1) Generate N frames of the same size of s and set their coefficients to zero. These frames are called: $F = [f_1, f_2, \ldots, f_N]$
2) Split the input subband's coefficients into different frames according to their values using the following ranges: $[0, b_1), [b_1, b_2), \ldots, [b_{N-1}, 1]$

It generates F frames from input subband, where these frames, in a sense, form a multispectral image. An illustration of the process to generate four frames can be seen in Fig. 5.

B. GREEDY HILL CLIMBING

The greedy hill climbing algorithm calculates the boundaries by iteratively running 2D-WMBPCA on a training set of images. The proposed algorithm initializes with a set of

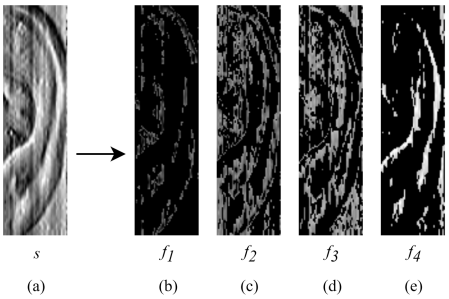

FIGURE 5. An example of multiple frames generation: a) Pre-processed input HL subband from an image in the IITD II dataset [18], resulting multiple frames b) f_1 c) f_2 d) f_3 e) f_4. The coefficients for frames b) and c) have been multiplied by 1.5 and 1.25 respectively for illustration purposes.

training images called input_images, set of boundaries called bnds, and a pre-selected step size κ. It begins with the LL subband and iteratively adds boundaries in a greedy manner until a local optimum is reached. The process is then repeated for the LH, HL, and HH bands respectively. Pseudocode for the algorithm can be seen in Algorithm 1.

A κ value of 0.05 was empirically determined as a compromise between performance, overfitting, and computation time for the results presented in this paper. Although this greedy hill climbing approach is not guaranteed to find the global optimum for boundary values, it produces sufficient results while simultaneously reducing computation time when compared to a brute force method.

V. BENCHMARK DATASETS

This investigation uses two benchmark ear image datasets: The Indian Institute of Technology Delhi II (IITD II) dataset [18] and the University of Science and Technology Beijing I (USTB I) dataset [19]. The IITD II dataset consists of 793 images of the right ear of 221 participants. Each participant was photographed between three and six times, with each image being of size 180 × 50 pixels and in 8-bit grayscale. For consistency, only the first three images for each individual are used in this research. The USTB I dataset consists of 180 images of the right ear of 60 participants, each of whom were photographed three times. The images of this dataset are 8-bit grayscale of size 150 × 80 and are tightly cropped; however, some of the images exhibit slight yaw and shearing. Examples from these two datasets can be seen in Fig. 6.

VI. EXPERIMENTAL RESULTS

To assess the performance of the proposed 2D Wavelet based Multi-Band PCA (2D-WMBPCA) technique and compare its performance against other ear recognition algorithms, the images of the aforementioned two ear image datasets were used. The proposed 2D-WMBPCA method using the two boundary selection algorithms described in Section IV, Single Image PCA, and the eigenfaces technique were applied to the images of both datasets. All of the experiments began

Algorithm 1: Pseudocode for the Greedy Hill Climbing Algorithm

Data: The training *input_images* to be used
Result: The *final_boundaries* (to be used on the testing set)

final_boundaries ← ∅, *top_percent* ← 0
current_boundary ← κ, *best_percent* ← 0,
optimum_boundary ← 0
if *current_boundary* < 1 **then**
  temp_boundaries ← *final_boundaries*
  temp_boundaries ← {*temp_boundaries*, *optimum_boundary*}
  matching_percent ← 2D-WMBPCA(*input_images*, *temp_boundaries*) **if** *matching_percent* > *best_percent* **then**
    best_percent ← *matching_percent*
    optimum_boundary ← *current_boundary*
  end
  current_boundary ← *current_boundary* + κ
else
  if *best_percent* > *top_percent* **then**
    top_percent ← *best_percent*
    final_boundaries ← {*final_boundaries*, *optimum_boundary*}
    current_boundary ← κ, *best_percent* ← 0
    optimum_boundary ← 0
  else
    return *final_boundaries*
  end
end

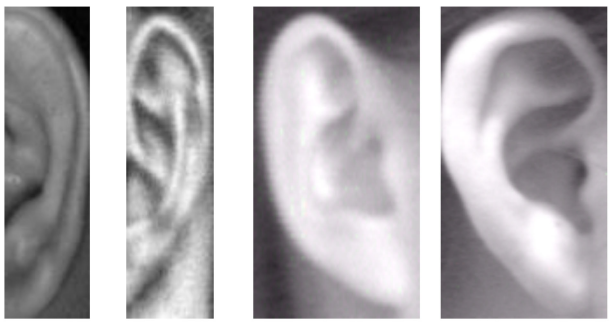

FIGURE 6. Sample images of two unique individuals from the IITD II dataset (a-b) [18]. Sample images of two unique individuals from the USTB I dataset (c-d) [19].

by selecting the first image of each subject to serve as a query set and the rest of the images to be a dataset. Given a particular query image, the Euclidean distance between this image and its correct corresponding image in the dataset is the lowest among all images, it is marked as Rank-1. Similarly, if the Euclidean distance between this image and its correct corresponding image in the dataset is within the lowest five

TABLE 2. Experimental results for single image PCA (%).

Dataset	Type of Match	
	Rank-1	*Rank-5*
IITD II	36.35	52.94
USTB I	45.00	70.00

TABLE 3. Experimental results for the Eigenfaces PCA method (%).

Dataset	Type of Match	
	Rank-1	*Rank-5*
IITD II	89.78	95.64
USTB I	75.93	90.74

distances, it is marked as Rank-5. This process is repeated for the second and third images for each individual, with the Rank-1 and Rank-5 accuracies averaged across all three trials.

A. EXPERIMENTAL RESULTS FOR THE SINGLE IMAGE PCA METHOD

To create results for the Single Image PCA method, standard PCA was applied to each original image individually. Their resulting eigenvectors were then compared using Euclidean distance. The results for the application of PCA on the images of the IITD II and USTB I datasets are presented in Table 2. From Table 2, it can be seen that the performance of PCA on the images of the IITD II dataset is lower than that of the USTB I dataset. This can be explained by the fact that some of the IITD II's images have a slight yaw, which causes slight occlusion.

B. EXPERIMENTAL RESULTS FOR THE EIGENFACES METHOD

To generate experimental results for the PCA based eigenfaces method [5], 10% of images of each ear dataset were used to calculate the eigenvectors. The remaining images were then projected along the resulting eigenvectors to create eigenears, which were compared using the Euclidean distance. The experimental results are tabulated in Table 3. From Table 3, it can be seen that the eigenfaces method vastly outperforms Single Image PCA on the images of both the IITD II and USTB I datasets. Interestingly, however, the eigenfaces method achieves higher accuracy on the images of IITD II dataset. This could be attributed to the wider diversity of ear images which IITD II contains compared to the images within the USTB I dataset.

C. EXPERIMENTAL RESULTS FOR THE PROPOSED 2D-WMBPCA USING EQUAL SIZE BOUNDARIES

The proposed 2D-WMBPCA method was applied to both images of the IITD II and USTB I datasets using two to ten frames of constant size, as discussed in Section IV-A. The number of correct matches was calculated for each set of frames. A subset of the results for both the IITD II and USTB I image datasets are tabulated in Tables 4 and 5.

TABLE 4. Rank-1 and Rank-5 matching accuracy (%) using equal size boundaries on the images of the IITD II [18] dataset.

Type of Match		Number of Bands
Rank-1	*Rank-5*	
90.28	94.64	2
94.14	**97.49**	3
92.80	97.32	4

TABLE 5. Rank-1 and Rank-5 matching accuracy (%) using equal size boundaries on the images of the USTB I [19] dataset.

Type of Match		Number of Bands
Rank-1	*Rank-5*	
93.83	97.53	3
96.30	**98.15**	**4**
96.30	**98.15**	**5**
94.44	98.15	6

From Tables 4 and 5, it can be seen that the proposed 2D-WMBPCA method significantly outperforms Single Image PCA on images from both the IITD II and USTB I datasets. From Table 2 and 3, it is evident that the Rank-1 accuracy of matching has been improved by 57.79% and 51.30% on the images of the IITD II dataset using three bins and images of the USTB I dataset using four bins when compared to Single Image PCA, respectively. Furthermore, the Rank-1 accuracy for 2D-WMBPCA on the IITD II and USTB I datasets increased by 4.36% and 20.37% when compared to eigenfaces, respectively. From the experiments presented in these two tables, it can be seen that the proposed 2D-WMBPCA method achieves its highest performance when using just three/four bands.

D. EXPERIMENTAL RESULTS FOR THE PROPOSED 2D-WMBPCA USING GREEDY HILL CLIMBING BASED BOUNDARIES

For this experiment, 2D-WMBPCA was performed using the greedy hill climbing based boundary selection method described in Section IV-B. κ values from 0.01 to 0.1 with a step size of 0.01 were tested on a 10% validation set. The value $\kappa = 0.05$ was chosen as a middle point between matching accuracy and computational complexity. Although this approach is not guaranteed to find the boundaries that globally maximize the matching accuracy, it produces sufficient results while simultaneously reducing computation time. The results are shown in Table 6. From Table 6, it can be seen that the greedy hill climbing based boundary selection method generates promising results on both datasets.

To compare the performance of the proposed 2D-WMBPCA using the equal size and greedy hill climbing techniques, a zoomed portion of the Cumulative Match

TABLE 6. Rank-1 and Rank-5 matching accuracy (%) using greedy hill climbing based boundary selection on the IITD II and USTB I datasets.

Dataset	Rank-1	Rank-5
IITD II	94.47	97.32
USTB I	97.53	97.15

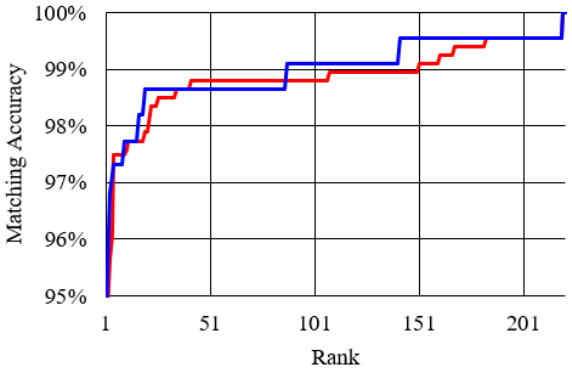

FIGURE 7. Region of interest of CMC curves for the equal size (red) and greedy hill climbing (blue) techniques for the IITD II dataset [18].

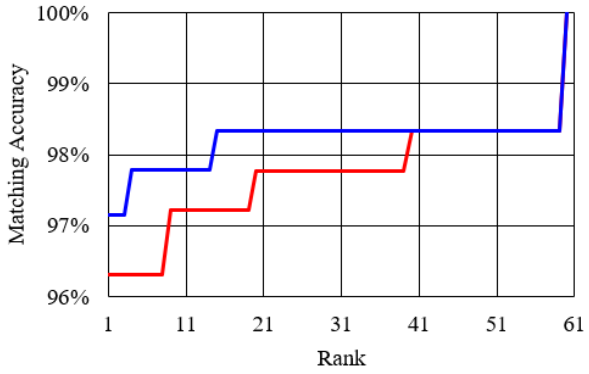

FIGURE 8. Region of interest of CMC curves for the equal size (red) and greedy hill climbing (blue) techniques for the USTB I dataset [19].

Curves (CMC) for the IITD II and USTB I datasets are presented in Fig. 7 and 8, respectively. From these figures, it can be easily seen that the greedy hill climbing technique generates more accurate matching than the equal size technique.

To compare the performance of the proposed technique with the state of the art PCA and learning based techniques, the Rank-1 experimental results of the proposed 2D-WMBPCA, Single Image PCA, eigenfaces [5], 2D-MBPCA [13], BSIF [26], GoogLeNet [10], ResNet18 and SVM [11], VGG-based Ensembles [54] and neural network and SVM based [9] techniques are tabulated in Table 7. From Table 7, it can be noted that the proposed 2D-WMBPCA technique significantly outperforms Single Image PCA, eigenfaces, and 2D-MBPCA methods. Furthermore, the proposed method gives competitive results compared to learning based techniques.

TABLE 7. Rank-1 results for the proposed technique and other PCA and learning based techniques (%).

Algorithm	Dataset	
	IITD II	*USTB I*
PCA based Techniques		
Single Image PCA	36.35	45.00
Eigenfaces [5]	89.78	75.93
2D-MBPCA [13]	92.76	96.11
Proposed Technique	**94.47**	**97.53**
Learning based Techniques		
BSIF and SVM [26]	97.31	-
GoogLeNet [10]	98.57	99.36
ResNet18 and SVM [11]	98.76	99.45
VGG-based Ensembles [54]	98.88	99.24
Neural Network and SVM [9]	-	98.30

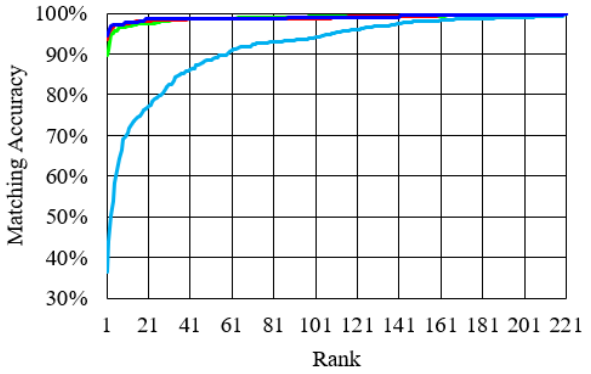

FIGURE 9. CMC curves for single image PCA (light blue), eigenfaces (green), 2D-MBPCA (red), and 2D-WMBPCA (dark blue) for the IITD II dataset [18].

A further comparison between the proposed 2D-WMBPCA technique and Single Image PCA, eigenfaces, and 2D-MBPCA is demonstrated using Cumulative Match Curves (CMC). The CMC curves for the IITD II dataset are shown Fig. 9 and its zoomed portion Fig. 10, while the CMC curves for the USTB I dataset are shown in Fig. 11 and its zoomed portion Fig. 12, respectively. From Fig. 9 and Fig. 11, it can be seen that the proposed 2D-WMBPCA and its anchor 2D-MBPCA significantly outperform both Single Image PCA and the eigenfaces technique. To enable the reader to differentiate between the proposed 2D-WMBPCA technique and its anchor 2D-MBPCA, a zoomed version of these two curves are shown in Fig. 10 and Fig. 12. From these two zoomed figures, it can be seen that the proposed 2D-WMBPCA technique gives higher accuracy to that of 2D-MBPCA.

E. JUSTIFICATION OF THE ACHIEVED PERFORMANCE

From the experimental results, it is clear that the proposed 2D-WMBPCA technique significantly outperforms other PCA based methods. This improvement can be explained by

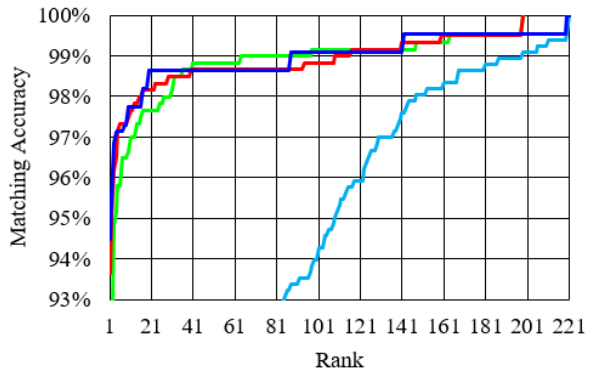

FIGURE 10. Region of interest of the CMC curves for Single Image PCA (light blue), eigenfaces (green), 2D-MBPCA (red), and 2D-WMBPCA (dark blue) for the IITD II dataset [18].

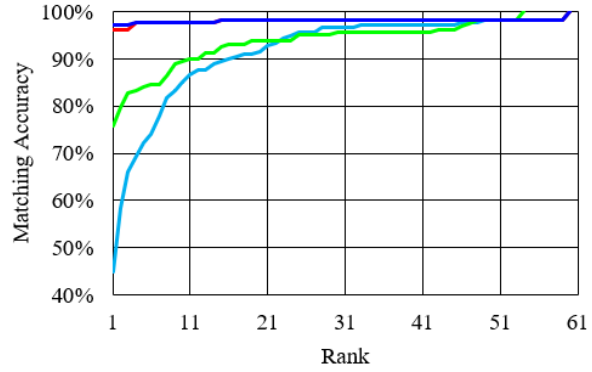

FIGURE 11. CMC curves for single Image PCA (light blue), eigenfaces (green), 2D-MBPCA (red), and 2D-WMBPCA (dark blue) for the USTB I dataset [19].

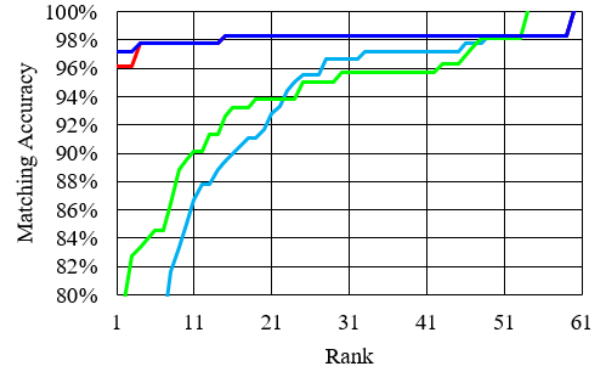

FIGURE 12. Region of interest of the CMC curves for Single Image PCA (light blue), eigenfaces (green), 2D-MBPCA (red), and 2D-WMBPCA (dark blue) for the USTB I dataset [19].

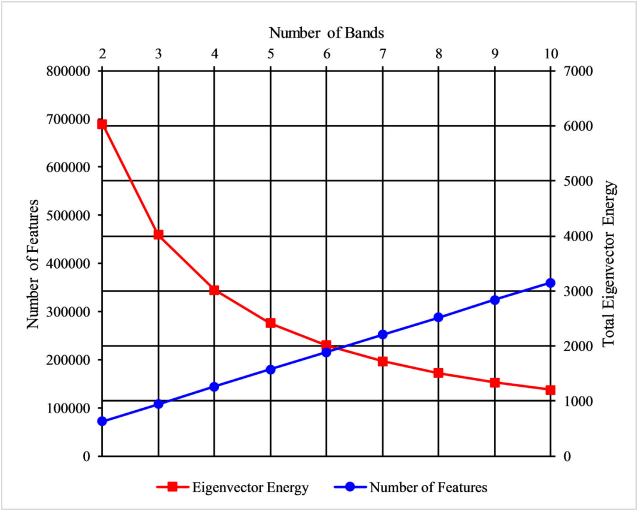

FIGURE 13. The number of features and total eigenvector energy versus the number of frames, where the intersection demonstrates the number of frames for maximum achievable performance, for the IITD II dataset [18].

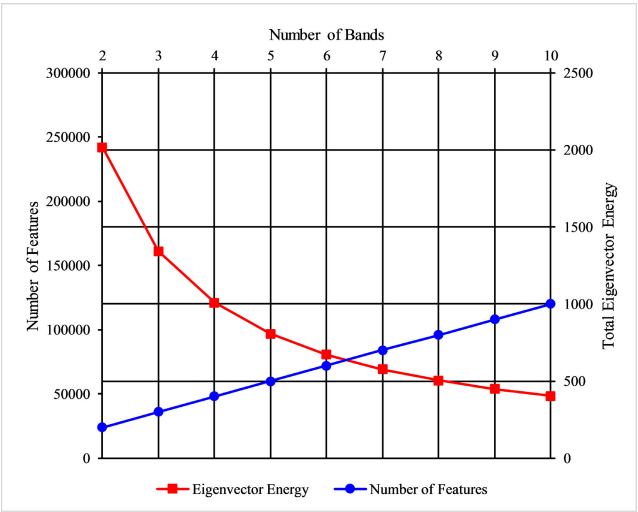

FIGURE 14. The number of features and total eigenvector energy versus the number of frames, where the intersection demonstrates the number of frames for maximum achievable performance, for the USTB I dataset [19].

the fact that the proposed technique expands the feature space by a factor of $b - 1$, where b is the number of frames (the number of features for Single Image PCA is $x * y$, while the number of features for 2D-WMBPCA is $x * y * (b - 1) * 4$, where the original image is of size $x * y$). Although increasing the number of frames within each image linearly increases the feature space, the effectiveness of the features is limited by the energy of individual eigenvectors. Consequently, there is a theoretical limitation on the maximum number of features, and thus the number of frames, that can be used for matching. This limitation is consistent with the experimental results in Section VI-D, where the matching performance of 2D-MBPCA first increases as the number of frames increases, reaching a maximum, and then decreases. To demonstrate this finding, the number of features and the total eigenvector energy for each number of frames were calculated for both datasets and are illustrated in Fig. 13 and 14.

From Fig. 13 and 14, it can be seen that as the number of bands increases, the total eigenvector energy decreases

inversely. The intersection of the Eigenvector Energy and Number of Features graphs occurs at approximately six bands in both figures. For both datasets, when using the greedy hill climbing technique, the HL subband required exactly six partitions; the most of any subband.

F. EXECUTION TIME

Ear recognition techniques are generally classified into two main categories: statistical- and learning-based techniques. Statistical based techniques, including PCA, Eigenfaces and the proposed 2D-WMBPCA algorithm, extract features (statistics) directly from the input image data and use these features to perform matching, while learning based techniques use a range of information including image data statistics and features, and other information, such as annotated image data, to train classifiers such as neural networks and support vector machines. The trained classifiers are then used to perform classification or matching. Consequently, learning based ear recognition algorithms are much more computationally expensive than their statistical based counterparts. In addition, their performance exhibits significant data dependency.

To compare the performance of the proposed 2D-WMBPCA algorithm to other statistical based methods, as well as the state of the art learning based techniques, 2D-WMBPCA, Single Image PCA, eigenfaces [5], 2D-MBPCA [13], BSIF [26], GoogLeNet [10], ResNet18 and SVM [11], VGG-based Ensembles [54] and neural network and SVM based [9] were implemented in MATLAB. The implemented algorithms were then executed on a Windows 10 personal computer equipped with a 7th generation Intel core i7 processor, an Nvidia GTX 1080 graphics card, and a 512 GB Toshiba NVMe solid-state drive (no other applications, updates or background programs were running during the computation). The average execution time for processing a query image using each algorithm (learning based techniques were already trained and their training time is not included in these measurements) was measured using 100 randomly selected query images from each dataset. The resulting measurements are tabulated in Table 8.

From Table 8, it can be seen that the proposed 2D-WMBPCA technique's execution time is significantly lower than the state of the art learning based methods, while generating very competitive matching performance. It is worth mentioning that the learning based methods require a training phase which is computationally intensive and has not been counted in the results presented in this table. However, the performance of the learning based techniques is highly data dependent and can significantly deteriorate when using cross-dataset validation, while the proposed 2D-WMBPCA method's performance has significantly less data dependency.

From Table 8, it can also be seen that the proposed 2D-WMBPCA technique's execution time is almost the same as Single Image PCA and 2D-MBPCA methods, while it outperforms both techniques. For example, the proposed technique takes 2.44% more time than Single Image PCA on the IITD II dataset, yet increases the Rank-1 accuracy by 58.12%. The eigenfaces technique is significantly faster than the proposed 2D-WMBPCA algorithm. This is due to the fact that 2D-WMBPCA performs PCA on each query image, whereas the eigenfaces method simply projects the query image along the pre-calculated eigenvectors. However, the performance of the proposed technique is significantly higher than the eigenfaces technique, e.g the Rank-1 accuracy for 2D-WMBPCA is 20.37% higher than the Rank-1 accuracy for the eigenfaces technique on the images of the USTB I dataset.

TABLE 8. Average execution time (milliseconds) of the proposed 2D-WMBPCA and the state of the art PCA based and learning based algorithms.

Algorithm	Dataset	
	IITD II	*USTB I*
PCA based Techniques		
Single Image PCA	13.55	12.16
Eigenfaces [5]	3.10	1.82
2D-MBPCA [13]	13.64	13.07
Proposed Technique	**13.88**	**13.21**
Learning based Techniques		
BSIF and SVM [26]	23.57	-
GoogLeNet [10]	22.88	21.59
ResNet18 and SVM [11]	24.24	23.88
VGG-based Ensembles [54]	23.51	22.79
Neural Network and SVM [9]	-	22.78

It is general knowledge that the performance of learning based techniques is dependent on their feature extraction techniques. The proposed 2D-WMBPCA algorithm generates significantly higher performance to those of statistical based techniques, which is due to the fact that 2D-WMBPCA extracts more eigenvectors with higher energy than other PCA based techniques. Therefore, the proposed 2D-WMBPCA technique has an inherent ability to further improve the performance of learning-based classification algorithms, including ear recognition techniques, when it is used as their primary feature extractor.

VII. CONCLUSION

In this paper, a non-decimated wavelet and PCA based ear recognition algorithm, called 2D Wavelet based Multi-Band PCA (2D-WMBPCA), was presented. The proposed algorithm performs a 2D non-decimated wavelet transform on the input image, dividing it into its subbands. Each resulting subband is then divided into a number of frames based on its coefficients. The standard PCA method is then applied on each subband's resulting frames, extracting their eigenvectors. The proposed technique uses the graph intersection of the number of resulting multiple frames features and their total eigenvector energies, which empirically has been shown to achieve the highest matching performance as the number of

frames for each subband. Experimental results on the images of two benchmark ear image datasets show that the proposed 2D-WMBPCA technique significantly outperforms the Single Image PCA, eigenfaces, and 2D-MBPCA methods. Furthermore, it generates competitive results to those of the state of the art learning based techniques at much reduced computational cost.

REFERENCES

[1] H. Nejati, Z. Li, T. Sim, E. Martinez-Marroquin, and D. Guo, "Wonder ears: Identification of identical twins from ear images," in *Proc. Int. Conf. Pattern Recognit. (ICPR)*, Nov. 2012, pp. 1201–1204.

[2] Ž. Emeršić, V. Štruc, and P. Peer, "Ear recognition: More than a survey," *Neurocomputing*, vol. 255, pp. 26–39, Sep. 2017. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S092523121730543X

[3] A. Pflug and C. Busch, "Ear biometrics: A survey of detection, feature extraction and recognition methods," *IET Biometrics*, vol. 1, no. 2, pp. 114–129, Jun. 2012.

[4] B. Victor, K. Bowyer, and S. Sarkar, "An evaluation of face and ear biometrics," in *Object Recognition Supported by User Interaction for Service Robots*, vol. 1. Quebec City, QC, Canada: IEEE, 2002, pp. 429–432.

[5] K. Chang, K. W. Bowyer, S. Sarkar, and B. Victor, "Comparison and combination of ear and face images in appearance-based biometrics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 8, pp. 1160–1165, Sep. 2003.

[6] D. Querencias-Uceta, B. Rios-Sanchez, and C. Sanchez-Avila, "Principal component analysis for ear-based biometric verification," in *Proc. Int. Carnahan Conf. Secur. Technol. (ICCST)*, Oct. 2017, pp. 1–6.

[7] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 1991, pp. 586–587.

[8] M. S. Nosrati, K. Faez, and F. Faradji, "Using 2D wavelet and principal component analysis for personal identification based on 2D ear structure," in *Proc. Int. Conf. Intell. Adv. Syst.*, Nov. 2007, pp. 616–620.

[9] I. Omara, X. Wu, H. Zhang, Y. Du, and W. Zuo, "Learning pairwise SVM on deep features for ear recognition," in *Proc. IEEE/ACIS 16th Int. Conf. Comput. Inf. Sci. (ICIS)*, May 2017, pp. 341–346.

[10] F. I. Eyiokur, D. Yaman, and H. K. Ekenel, "Domain adaptation for ear recognition using deep convolutional neural networks," *IET Biometrics*, vol. 7, no. 3, pp. 199–206, May 2018.

[11] S. Dodge, J. Mounsef, and L. Karam, "Unconstrained ear recognition using deep neural networks," *IET Biometrics*, vol. 7, no. 3, pp. 207–214, May 2018.

[12] J. Yang, D. Zhang, A. F. Frangi, and J.-Y. Yang, "Two-dimensional PCA: A new approach to appearance-based face representation and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 1, pp. 131–137, Jan. 2004.

[13] M. Zarachoff, A. Sheikh-Akbari, and D. Monekosso, "2D multi-band PCA and its application for ear recognition," in *Proc. IEEE Int. Conf. Imag. Syst. Techn. (IST)*, Oct. 2018, pp. 1–5.

[14] J. C. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 4, pp. 779–785, Jul. 1994.

[15] X. Jia and J. A. Richards, "Segmented principal components transformation for efficient hyperspectral remote-sensing image display and classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 1, pp. 538–542, Jan. 1999.

[16] J. Zabalza, J. Ren, M. Yang, Y. Zhang, J. Wang, S. Marshall, J. Han, "Novel folded-PCA for improved feature extraction and data reduction with hyperspectral imaging and SAR in remote sensing," *ISPRS J. Photogramm. Remote Sens.*, vol. 93, pp. 112–122, Jul. 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0924271614000938

[17] M. Zarachoff, A. Sheikh-Akbari, and D. Monekosso, "Single image ear recognition using wavelet-based multi-band PCA," in *Proc. 27th Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2019, pp. 1–4.

[18] *IIT Delhi Ear Database*. Accessed: Mar. 13, 2018. [Online]. Available: https://bit.ly/3rAPbWE

[19] *Ear Recoginition Laboratory at USTB*. Accessed: Jun. 25, 2018. [Online]. Available: http://www1.ustb.edu.cn/resb/en/index.htm

[20] L. Min and L. Song, "Face recognition based on PCA and 2DPCA with single image sample," in *Proc. 9th Web Inf. Syst. Appl. Conf.*, Nov. 2012, pp. 111–114.

[21] T. Barbu, "Unsupervised SIFT-based face recognition using an automatic hierarchical agglomerative clustering solution," *Proc. Comput. Sci.*, vol. 22, pp. 385–394, Jan. 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877050913009095

[22] S. Biswas and J. Sil, "An efficient face recognition method using contourlet and curvelet transform," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 32, no. 6, pp. 718–729, Jul. 2020. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1319157817300861

[23] M. Alaraj, J. Hou, and T. Fukami, "A neural network based human identification framework using ear images," in *Proc. IEEE Region 10 Conf. (TENCON)*, Nov. 2010, pp. 1595–1600.

[24] Y. Zhang, Z. Mu, L. Yuan, and C. Yu, "Ear verification under uncontrolled conditions with convolutional neural networks," *IET Biometrics*, vol. 7, no. 3, pp. 185–198, May 2018.

[25] P. L. Galdámez, A. G. Arrieta, and M. R. Ramón, "Ear recognition using a hybrid approach based on neural networks," in *Proc. 17th Int. Conf. Inf. Fusion (FUSION)*, Jul. 2014, pp. 1–6.

[26] A. Benzaoui, N. Hezil, and A. Boukrouche, "Identity recognition based on the external shape of the human ear," in *Proc. Int. Conf. Appl. Res. Comput. Sci. Eng. (ICAR)*, Oct. 2015, pp. 1–5.

[27] A. Benzaoui and A. Boukrouche, "Ear recognition using local color texture descriptors from one sample image per person," in *Proc. 4th Int. Conf. Control, Decis. Inf. Technol. (CoDIT)*, Apr. 2017, pp. 0827–0832.

[28] M. Burge and W. Burger, "Ear biometrics," in *Biometrics*. USA: Springer, 1996, pp. 273–285.

[29] B. Moreno, A. Sanchez, and J. F. Velez, "On the use of outer ear images for personal identification in security applications," in *Proc. IEEE 33rd Annu. Int. Carnahan Conf. Secur. Technol.*, Oct. 1999, pp. 469–476.

[30] M. Rahman, M. R. Islam, N. I. Bhuiyan, B. Ahmed, and A. Islam, "Person identification using ear biometrics," *Int. J. Comput., Internet Manage.*, vol. 15, no. 2, pp. 1–8, 2007.

[31] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Mar. 2004. [Online]. Available: https://doi.org/10.1023/B: VISI.0000029664.99615.94

[32] K. Dewi and T. Yahagi, "Ear photo recognition using scale invariant keypoints," in *Computational Intelligence*. San Francisco, CA, USA: ACTA Press, Nov. 2006, pp. 253–258.

[33] S. Prakash and P. Gupta, "An efficient ear recognition technique invariant to illumination and pose," *Telecommun. Syst.*, vol. 52, no. 3, pp. 1435–1448, Mar. 2013, doi: 10.1007/s11235-011-9621-2.

[34] Z. Hai-Long and M. Zhi-Chun, "Combining wavelet transform and orthogonal centroid algorithm for ear recognition," in *Proc. 2nd IEEE Int. Conf. Comput. Sci. Inf. Technol.*, Aug. 2009, pp. 228–231.

[35] A. Basit and M. Shoaib, "A human ear recognition method using nonlinear curvelet feature subspace," *Int. J. Comput. Math.*, vol. 91, no. 3, pp. 616–624, 2014, doi: 10.1080/00207160.2013.800194.

[36] A. Kumar and C. Wu, "Automated human identification using ear imaging," *Pattern Recognit.*, vol. 45, no. 3, pp. 956–968, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0031320311002706

[37] A. Kumar and D. Zhang, "Ear authentication using Log-Gabor wavelets," *Proc. SPIE*, vol. 6539, Apr. 2007, Art. no. 65390A.

[38] Y. Guo and Z. Xu, "Ear recognition using a new local matching approach," in *Proc. 15th IEEE Int. Conf. Image Process.*, Oct. 2008, pp. 289–292.

[39] N. Damer and B. Führer, "Ear recognition using multi-scale histogram of oriented gradients," in *Proc. 8th Int. Conf. Intell. Inf. Hiding Multimedia Signal Process.*, Jul. 2012, pp. 21–24.

[40] H.-J. Zhang, Z.-C. Mu, W. Qu, L.-M. Liu, and C.-Y. Zhang, "A novel approach for ear recognition based on ICA and RBF network," in *Proc. Int. Conf. Mach. Learn. Cybern.*, vol. 7, Aug. 2005, pp. 4511–4515.

[41] K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman, "Return of the devil in the details: Delving deep into convolutional nets," 2014, *arXiv:1405.3531*.

[42] Z. Emersic, D. Stepec, V. Struc, P. Peer, A. George, A. Ahmad, E. Omar, T. E. Boult, R. Safdaii, Y. Zhou, S. Zafeiriou, D. Yaman, F. I. Eyiokur, and H. K. Ekenel, "The unconstrained ear recognition challenge," in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Oct. 2017, pp. 715–724.

[43] M. Pietikäinen, A. Hadid, G. Zhao, and T. Ahonen, "Local binary patterns for still images," in *Computer Vision Using Local Binary Patterns* (Computational Imaging and Vision), M. Pietikäinen, A. Hadid, G. Zhao, and T. Ahonen, Eds. London, U.K.: Springer London, 2011, pp. 13–47.

[44] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[45] G. Lu and B. Fei, "Medical hyperspectral imaging: A review," *J. Biomed. Opt.*, vol. 19, no. 1, 2014, Art. no. 010901. [Online]. Available: https://www.spiedigitallibrary.org/journals/Journal-of-Biomedical-Optics/volume-19/issue-1/010901/Medical-hyperspectral-imaging-a-review/10.1117/1.JBO.19.1.010901.short
[46] R. Dianat and S. Kasaei, "Dimension reduction of optical remote sensing images via minimum change rate deviation method," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 1, pp. 198–206, Jan. 2010.
[47] J. Ren, J. Zabalza, S. Marshall, and J. Zheng, "Effective feature extraction and data reduction in remote sensing using hyperspectral imaging [applications corner]," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 149–154, Jul. 2014.
[48] G. Vane, R. O. Green, T. G. Chrien, H. T. Enmark, E. G. Hansen, and W. M. Porter, "The airborne visible/infrared imaging spectrometer (AVIRIS)," *Remote Sens. Environ.*, vol. 44, nos. 2–3, pp. 127–143, 1993. [Online]. Available: http://www.sciencedirect.com/science/article/pii/003442579390012M
[49] *Hyperspectral Remote Sensing Scenes–Grupo de Inteligencia Computacional (GIC)*. Accessed: Aug. 16, 2018. [Online]. Available: http://bit.ly/3qztsNw
[50] M. S. Andrić, B. P. Bondžulić, and B. M. Zrnić, "The database of radar echoes from various targets with spectral analysis," in *Proc. 10th Symp. Neural Netw. Appl. Elect. Eng.*, Sep. 2010, pp. 187–190.
[51] M. S. Holia and V. K. Thakar, "Image registration for multi focus and multi modal images using windowed PCA," in *Proc. IEEE Int. Advance Comput. Conf. (IACC)*, Feb. 2014, pp. 1104–1109.
[52] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.
[53] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 7, pp. 710–732, Jul. 1992.
[54] H. Alshazly, C. Linse, E. Barth, and T. Martinetz, "Ensembles of deep learning models and transfer learning for ear recognition," *Sensors*, vol. 19, no. 19, p. 4139, Sep. 2019. [Online]. Available: https://www.mdpi.com/1424-8220/19/19/4139

MATTHEW MARTIN ZARACHOFF received the B.S. degree in computer science and the B.S. degree in physics from Oklahoma State University, Stillwater, OK, USA, in 2014, and the M.S. degree in computer science from the University of California–Riverside, Riverside, CA, USA, in 2015. He is currently pursuing the Ph.D. degree with the School of Built Environment, Engineering and Computing, Leeds Beckett University. He has previously worked as a Quality Control Engineer at the aerospace industry. His research interests include image processing, machine learning, and computer vision.

AKBAR SHEIKH-AKBARI received the B.Sc. and M.Sc. (Hons.) degrees in electronic and electrical engineering and the Ph.D. degree in electronic and electrical engineering from Strathclyde University. After completing his Ph.D. degree, he joined Bristol University to work on an EPSRC project in stereo/multi-view video processing. He continued his career in the industry, working on real-time embedded video analytics systems. He is currently an Associate Professor (a Reader) with the School of Built Environment, Engineering and Computing, Leeds Beckett University. His main research interests include biometric identification techniques, colour constancy techniques, image source camera identification, hyperspectral image processing, standard and non-standard image/video codecs, such as H.264 and HEVC, multiview image/video processing, image resolution enhancement methods, assisted living technologies, compressive sensing, camera tracking using retro-reflective materials, and signal processing.

DOROTHY MONEKOSSO received the bachelor's degree in electronic engineering and the master's degree in satellite engineering, and the Ph.D. degree in space systems autonomy (autonomous spacecraft) from the Surrey Space Centre, in 2000. She is currently the Director of research with Leeds Beckett University. She began her career in space sector, developing on-board computers and control systems for spacecraft. She became interested in artificial intelligence. During her research at the Surrey Space Centre, she is interested in applying machine learning methods and techniques to spacecraft autonomy. She has led several digital health and assistive technologies projects funded by Innovate U.K., European Commission, Royal Academy of Engineering, DAIWA, and security projects funded by the USA Department of Homeland Security and co-led several security projects funded under the national and international program. She is currently Leeds Beckett PI for a H2020 Internet of Things Project MONICA: Management Of Networked IoT Wearables–Very Large Scale Demonstration of Cultural Societal Applications. She has authored and coauthored over 100 peer-reviewed publications in scientific journals. Her current research interest includes applications of machine learning to intelligent systems. On the basis of this work, she was awarded the Royal Academy of Engineering, Engineering Foresight Award, and spent a year at the Jet Propulsion Laboratory, Pasadena California.