Ear Recognition System using Radon Transform and Neural Network

C.MURUKESH¹, Dr.K.THANUSHKODI²
Department of EIE¹, Director²
Velammal Engineering College¹, Akshaya College of Engineering and Technology²
Chennai¹, Coimbatore²
India.
pcrmurukesh@gmail.com

Abstract: Ear recognition system is one among the many evolving cutting edge technologies in the field of security surveillance. This paper presents an ear recognition system based on the Radon transform combined with Principal Component Analysis (PCA) for feature extraction, and integration of Multi-class Linear Discriminant Analysis (LDA) and Self-Organizing Feature Maps (SOM) for classification. Radon transform is used to extract the directional features of an image by projection of an image matrix for different orientations. The experimental result shows that the verification of an ear recognition system tested on two different public ear databases is accurate and speed.

Key-Words: Radon Transform, Principal Component Analysis, Multi-class Linear Discriminant Analysis, Self-Organizing Feature Maps

1 Introduction

Ear recognition is a non-invasive, reliable and passive biometric system used to identify the person based on the physiological characteristics. Recently, human ear is becoming an emerging technology in the field of biometrics because it has some desirable properties such as uniqueness, universality, permanence, collectability and acceptability. Compared to other biometric traits such as finger vein, iris, fingerprint, the human ear can be easily captured using digital camera and is also less expensive.

Ear recognition system consists of different stages such as image acquisition, preprocessing, feature extraction and matching. In pre-processing stage a median filter smoothen the image by utilising the median of the neighbourhood to minimise salt-and-pepper noise. Edge detection is an important task in image processing to extract the edge points in an image. The classical edge detector uses specified edge patterns to match local image segments. The Haar wavelet edge detection which combines the coefficients of wavelets is suitable for singularity detection. Eigenears for the database images are determined and the features are extracted by the Radon transform and PCA. Radon transform is used to capture directional features in different orientations. Dimensionality reduction by PCA is used to represent an image in a lower dimensional feature vector space. PCA is a subspace projection technique widely used for prediction, redundancy removal, feature extraction and data compression. Holistic approach based ear recognition has gained high impact in extracting more discriminating features. Similarity approach, probabilistic approach and decision boundary approach are the three categories of classifiers used to obtain the recognised image. Decision-boundary approach based classifier is an important stage in the identification of images in extracting more discriminant features.

This paper explains two different classifiers namely Multi-class LDA and SOM which reduces time and increases the accuracy. SOM based Neural Network uses competitive learning approach and groups the input data into clusters which are, commonly used for unsupervised training. It is good for statistical or topological modelling of the data. This system is implemented to extract features from the image and train the images automatically. Considering all the above mentioned parameters, this paper proposes an efficient and computationally fast algorithm for ear recognition system. Fig.1 represents the flow diagram of ear recognition process.

The rest of the paper is organised as follows. Section II discusses some related work on ear recognition system. Section III presents ear pre-processing, edge detection and eigenears. Section IV describes feature extraction and matching for the proposed ear recognition system. Section V presents
the experimental results and section VI outlines the conclusion.

![Flow diagram of ear recognition process](image)

**Fig.1 The flow diagram of ear recognition process**

## 2 Related work

Human identification plays a vital role in biometrics for recognizing the right person for better security applications. Samuel Adebayo Daramola et al [1] presented a new approach for automatic ear recognition system using Wavelet Transform Decomposition and Back Propagation Network (BPN). Texture energy and edge density features are extracted separately from image blocks and fused together to form a feature vector. The image classification is done by using BPN Neural Network. David J. Hurley et al [2] developed force field transformation to extract ear features without loss of information and a classification rate of 99.2% has been achieved. Nazmeen Bibi Boodoo et al [3] used Karhunen-Loeve transform to select the more relevant features of face and ear images. The score level fusion done at decision level improves a recognition rate of 96%. Dattatray V. Jadhav et al [4] proposed Radon transform to capture directional features of face images in different orientations and enhance low frequency components. The Radon transform combined with Discrete Cosine Transform yields lower-dimensional feature vectors and also robustness to zero mean white noise. The work done by Shrikant Tiwari et al [5] for recognition of new born is by fusing ear images with soft biometric data, results in increasing the recognition accuracy. Changjun Zhou et al [6] have used Non-negative Matrix Factorization (NMF) on original images to obtain the residual face space. Fisher Linear Discriminant analysis (FLDA) is then applied to extract features from this space to scatter of all the samples within a class as small as possible. Haiyan Xu [7] has used the Gabor Wavelet to extract the ear features and finally, the features are utilized to train and test Support Vector Machines (SVM) for ear recognition. Atilla Ozmen et al [3] has proposed a new approach for edge detection, by combining steerable filters and Cellular Neural Networks (CNN) provides the direction of dominant orientation and desired edge is detected. Surya Prakash et al [9] presented an efficient and novel method to overcome the effect of illumination, poor contrast by three different image enhancement techniques and extract the local ear features using Speeded-Up Robust Features (SURF). The nearest neighbour classifiers are used to train the ear features. Fusion at score level carried out by weight sum rule significantly improves the recognition accuracy. Kshirsagar et al [10] explains a methodology for face recognition based on information theory approach of coding and decoding the face image. It uses PCA for feature extraction and Back Propagation Neural Network for classification. The goal is to implement the system (model) for a particular face and distinguish it from a large number of stored faces with some real-time variations as well. The eigenface approach uses Principal Component Analysis algorithm for the recognition of the images. It offers an efficient method to find the lower dimensional space. The main objective of this paper is to use Radon transform and PCA methods to detect and identify the features within an image which becomes a powerful tool for invariant under translation, rotation, and scaling.

## 3 Ear Pre-processing

### 3.1 Filtering

The purpose of filtering is to reduce noise and improve the visual quality of the image. Image filtering is useful in many applications, including smoothing, sharpening, removing noise, and edge detection.

### 3.2 Median Filter

It is a nonlinear digital filtering technique used to remove noise. In median filter, the value of the pixel is replaced by the median of the values of the pixels in around the given pixel, sorted in ascending order. A median is the value such that half of the values in...
the set are below and half are above the median (50 percentile). The median filtered ear image shown in the Fig.3 tends to preserve the sharpness of image edges while removing noise. It smoothened additive white noise and is also effective in removing impulses.

**3.3 Steerable Filter**

Steerable filters are oriented filters that allow synthesis of a filter at different orientation and phase to find out the ear objects clearly. The ear image can be detected by splitting the image into orientation sub-bands by basis filters. \( H(t_x;t_y) \) is said to be steerable at an arbitrary rotation \( \theta_a \), where \( h^{\theta_a}(t_x;t_y) \) is the rotated version of \( h(t_x;t_y) \) at \( \theta_a \) direction and \( k_i(\theta_a),1 \leq i \leq M \), are interpolation functions. The steerable filter coefficients obtained at different orientations yields the local direction of dominant orientation [11].

\[
h^{\theta_a}(t_x,t_y) = \sum_{i=1}^{M} k_i(\theta_a) h^{\theta_i}(t_x,t_y) \quad (1)
\]

Fig.4 illustrates the steerable features extracted from ear images to represent its shape.

**3.4 Edge Detection**

Edge detection is an important task in image pattern recognition and computer vision, particularly in the areas of feature detection and feature extraction. Edges in an image can be defined as local maxima of gradient in which brightness of the images changes abruptly. Edge detectors are mainly used to extract the edge points in an image and also to find the boundaries present in the ear images.

**3.4.1 Canny Edge Detection**

It is an edge detection operator that uses a multi-stage algorithm to detect a wide range of edges in images. The Canny design edge detector is based on three criteria. The first is to select a Gaussian function as a smoothing function to eliminate the noise. Next the detected edges are very near to true location of edges and finally there should be one response to a single edge. The Canny edge detector uses a filter based on the first derivative of a Gaussian, because it is optimal for step edges susceptible to white noise. The Gaussian filter convolves with the raw image. The result is a slightly blurred version of the original which is not affected by a single noisy pixel to any significant degree. The Gaussian mask is shown below.

\[
B = \frac{1}{59} \begin{pmatrix} 2 & 4 & 5 & 4 & 2 \\ 4 & 9 & 12 & 9 & 4 \\ 5 & 12 & 15 & 12 & 5 \\ 4 & 9 & 12 & 9 & 4 \\ 2 & 4 & 5 & 4 & 2 \end{pmatrix} A
\]

An edge in an image may point in a variety of directions. The Canny edge detection shown in Fig.5 uses four filters to detect horizontal, vertical and diagonal edges in the blurred image.

**3.4.2 Sobel Edge Detection**

Sobel operator consists of a pair of 3×3 convolution kernels. One kernel is simply the other rotated by 90°. The kernels are designed to respond maximally to edges running vertically and horizontally relative to the pixel grid and one kernel for each of the two perpendicular orientations. The sobel edge detection shown in Fig.6 makes it better for removing white noise. The kernels can be applied separately to the input image, to produce separate measurements of the gradient component in each orientation (\( G_x \) and \( G_y \)). The gradient magnitude is given by
\[ G = \sqrt{G_x^2 + G_y^2} \]  

(3)

**Fig.6 Sobel edge detected ear image**

### 3.4.3 Haar Wavelet Edge Detection

Wavelet analysis is similar to Fourier analysis and allows a target function over an interval to be represented in terms of an orthonormal function basis. The Haar wavelet transform is a convolution operation which decomposes an original image into four subbands and equivalent to passing an image through low pass and high pass filters. The 2x2 Haar matrix that is associated with the Haar wavelet is given below.

\[
H_2 = \begin{bmatrix}
1 & 1 \\
1 & -1
\end{bmatrix}
\]

(4)

Using the discrete wavelet transform, any sequence \((a_0, a_1, \ldots, a_n, a_{2n+1})\) of even length is transformed into a sequence of two component \(((a_0, a_1), \ldots, (a_{2n}, a_{2n+1}))\) vectors. If the right-multiplies of each vector with the matrix \(H_2\), gets \(((s_0, d_0), \ldots, (s_n, d_n))\) of one stage of the fast Haar wavelet transform. The sequences \(s\) and \(d\) are separated and continues transforming the sequence \(s\). If a sequence of length is a multiple of four, then the 4x4 Haar matrix is obtained by combining two stages of the fast Haar wavelet transform. Haar wavelet sub-band images shown in the Fig.7 illustrates that average components are detected by LL sub-bands, vertical, horizontal and diagonal edges are detected by LH, HL and HH sub-bands representing the level 1 decomposition of an image.

![Fig.7 Haar wavelet sub-band images](image)

### 3.5 Eigenears

Eigenear is an appearance-based method and widely used in ear recognition system due to its simplicity, speed and learning capability. The black and white image \(I(x, y)\) of size \(N \times N\) pixels is a matrix of 8-bit values may be considered as a vector of length \(N^2\) dimensional spaces. The eigenears are the principal components of a distribution of ears, or equivalently, the eigenvectors of the covariance matrix of the set of ear images. So a 128 x 128 pixel image can be represented as a point in a 16,384 dimensional space. Given a vector \(C\), the eigenvectors \(u\) and eigenvalues \(A\) of \(C\) satisfies. Let \(I_k\) represent the column vector of ear \(k\) obtained through lexicographical ordering of \(I_k(x, y)\). The mean normalized column vector for ear \(k\) is defined

\[ Cu = \lambda u \]  

(5)

Using (5) the equation

\[ (C - \lambda I)u = 0 \]  

(6)

Let \(I\) denotes the \(n \times n\) Identity matrix. By applying fundamental linear algebra, nontrivial solution exists if and only if

\[ \det(C - \lambda I) = 0 \]  

(7)

The determinant is evaluated to become a polynomial of degree \(n\). If \(C\) is \(n \times n\) matrix, then it exits \(n\) solutions or \(n\) roots of the characteristic polynomial. The \(n\) eigen values of \(C\) satisfying the equation

\[ Cu_i = u_i \]  

(8)

Eigenears are eigenvectors having all the eigenvalues that are distinct gives the efficient way to find lower dimensional space. The directions are
unique and will span on an n dimensional Euclidean space.

4 Feature Extractions and Matching

4.1. Radon Transform

The reduced features of an ear image can be formed using the Radon transform to extract the directional features of an image. The Radon transform is a line integral and computes the projection of an image matrix to detect the local features of an edge. The pixel intensity values in the given image are the projection of the image intensity along a radial line at different orientations. It transforms the 2-D image with lines and converts the rotation of \( f(x, y) \) into translation of \( R(r, \theta) \) and based on the parameterization of straight lines. The horizontal and vertical projection for an image matrix is shown in the Fig.16 Projections can be computed along any specified angle \( \theta \) and Radon space image for 0-179\(^o\) orientations.

The Radon transform of 2D function \( f(x, y) \) is defined as

\[
R(r, \theta) = \int \int f(x, y)(r - x \cos \theta - y \sin \theta) dx dy
\] (9)

Where \( r \) is the perpendicular distance of a line from the origin and \( \theta \) is the angle between the distance vector and X-axis [10]. The mean value of the white noise is zero and the Radon transform improves signal-to-noise ratio. The very important key feature of Radon transform is the ability to extract lines from noisy images. It is feasible to compute the Radon transform of any translated, rotated or scaled image, knowing the Radon transform of the original image and the parameters of the transformation applied to it.

4.2 PCA

The appearance shape model is used to extract the features of an image with respect to its shape. The PCA is used to reduce the large dimensionality of the data space (observed variables) to the smaller intrinsic dimensionality of feature space (independent variables), which are needed to describe the data economically [10][15]. The main idea of using PCA for ear recognition is to express the large 1-D vector of pixels constructed from 2-D ear image into the compact principal components of the feature space projection. Each ear image is considered as one single feature vector by concatenating the rows of pixels in the original image, resulting in a single row with \( r \times c \) elements.

A single matrix stores the training set of all images and each row of the matrix is an image. Eigen space is obtained by projection of the eigenvectors of the covariance matrix onto a dimensionally reduced space. \( P \) represents the pixel values.

\[
x_i = [p_1, ..., p_n]^T \quad i = 1, ..., M
\] (10)

The images are mean centered by subtracting the mean image from each image vector. Let \( m \) represent the mean image.

\[
m = \frac{1}{M} \sum_{k=1}^{M} X_k
\] (11)

The objective is to find a set of \( e_1 \) which has the largest possible projection.

\[
\omega_1 = x_1 - m
\] (12)

Average image has been calculated and then subtracted from the each original image. To find a set of \( M \) orthonormal vectors \( e_1 \), the quantity \( \lambda_1 \) is maximized with the orthonormality constraint.

\[
\lambda_1 = \frac{1}{M} \sum_{n=0}^{M} (e_1^T \sigma_k)
\] (13)

It has been shown that \( e_1 \) ’s and \( i \)’s are given by the eigenvectors and eigenvalues of the covariance matrix.

\[
C = WW^T
\] (14)

Each eigenvectors has the same dimensionality as the original image and also it can be seen as an image.

\[
WW^T(Wd_1) = \mu_1(Wd_1)
\] (15)

The Euclidean distance is given by

\[
\epsilon_K = \|\Omega - \Omega_K\|
\] (16)

The simplest method for determining which ear class provides the best description of an input ear image is to find the ear class of minimum Euclidean distance.
4.3 MULTI-CLASS LDA
Multi-class LDA with each individual in the database has a separate projection vector which discriminates the features from the rest of the individuals. It is class specific method and shapes the scatter in order to make it more reliable for classification and projects the image set to a lower dimension space using PCA [14]. The low dimensional vector of ear images is represented as \( x_k \), the mean value of class \( x_i \) as \( m_i \), and the mean value of all data as \( m \). The between class scatter matrix \( S_B \) can be calculated as

\[
S_B = \sum_{i=1}^{c} n_i (m_i - m)(m_i - m)^T
\]  

The number of classes is denoted as \( c \) and \( n_i \) represents the number of ear samples in class \( x_i \). The within class scatter matrix is defined as

\[
S_W = \sum_{i=1}^{c} \sum_{x_k \in C_i} (x_k - m_i)(x_k - m_i)^T
\]  

LDA tries to find a projection direction that maximizes the ratio of between-class scatter to within-class scatter.

\[
W_{FLD} = \arg \max_w \frac{W^T S_B W}{W^T S_W W} = [w_1, w_2, \ldots, w_n]
\]  

Optimal projection Matrix is described as

\[
W_{opt}^T = W_{fld}^T W_{pca}^T
\]

\[
W_{pca} = \arg \max_w \left| W^T S_W W \right|
\]

\( W_{opt} \) represents the optimal projection matrix. In its columns, it contains the generalized eigenvectors that correspond to the largest eigenvalues.

\[
S_B W_i = \lambda_i S_W W_i, i = 1, 2, \ldots, m
\]

If \( S_W \) is non-singular, \( W_{opt} \) can be calculated by simply computing eigenvectors of \( S^{-1} W_{SB} \). Instead of computing eigenvectors of \( S^{-1} W_{SB} \), it is preferable to diagonalize \( S_W \) and \( S_B \).

4.4 SELF ORGANIZING FEATURE MAPS
Self-Organizing Feature Maps (SOM) [16] which is proposed by Teuvo Kohonen is one of the well-known unsupervised learning algorithms in the field of Neural Networks for modelling the neurobiological behaviour of human brain. SOM as shown in fi are trained using self-learning networks to produce low dimensional representation of the training samples while preserving the topological properties of the input space.

SOM is typically a kind of competitive learning that only one neuron will fire after mutual competition of neurons. The principal goal of self-organizing maps is to transform an incoming signal pattern of arbitrary dimension into a one or two-dimensional discrete map, and to perform this transformation adaptively in a topologically ordered fashion [21].

SOM learn to recognize groups of similar input vectors in such a way that neurons physically near each other in the neuron layer respond to similar input vectors. They provide a quantization of the image samples into a topological space where inputs that are nearby in the original space are also nearby in the output space, thereby providing dimensionality reduction and invariance to minor changes in the image sample. Self-organizing maps are a single layer feed forward network where the output syntaxes are arranged in low dimensional (usually 2D or 3D) grid. Each input is connected to all output neurons. Attached to every neuron there is a weight vector with the same dimensionality as the input vectors. The number of input dimensions is usually a lot higher than the output grid dimension.
The input vector \( p \) is the row of pixels of the image. The \( \| \text{ndis} \| \) box accepts the input vector \( p \) and the input weight matrix \( IW \) produces a vector having \( S \) elements. The elements are the negative of the distances between the input vector and vectors \( IW \) formed from the rows of the input weight matrix. The competitive transfer function accepts a net input vector for a layer and returns neuron outputs of 0 for all neurons except for the winner, the neuron associated with the most positive element of net input \( n^1 \). The winner’s output is 1. The neuron whose weight vector is closest to the input vector has the least negative net input and wins the competition to output a 1. Thus the competitive transfer function produces a 1 for output element corresponding to \( i^\text{th} \) winning neuron. All other output elements in a 1 are 0.

A SOM network identifies a winning neuron using the same procedure as employed by a competitive layer. However, instead of updating only the winning neuron, all neurons within a certain neighbourhood of the winning neuron are updated using the Kohonen rule. Specifically, there is a need to adjust all such neurons as follows

\[
 w(q) = (1 - \alpha)w(q - 1) + \alpha p(q) \tag{23}
\]

The neighbourhood contains the indices for all of the neurons that lie within a radius \( d \) of the winning neuron.

\[
 N_i(d) = \{ j, d_{ij} \leq d \} \tag{24}
\]

Thus, when a vector \( p \) is presented, the weights of the winning neuron and its close neighbours move toward \( p \).

### 4.4.1 Algorithm for Self-Organizing Feature Map

- Assume output nodes are connected in an array usually 1 or 2 dimensional.
- Assume that the network is fully connected and all nodes in input layer are connected to all nodes in output layer.
- Randomly choose an input vector \( x \).
- Determine the “winning” output node, where \( w_i \) is the weight vector connecting the inputs to output node ‘i’.

\[
 |w_i - x| \leq |w_k - x| \forall k \tag{25}
\]

The weight updates for the winning node ‘i’ is

\[
 w_k(\text{new}) = w_k(\text{old}) + \mu N(j,k)(x - w_k) \tag{26}
\]

The neighbourhood function \((i, k)\) that has value 1 when \( i = k \) and falls off with the distance between units \( i \) and \( k \) in the output array. The units close to the winners as well as the winners itself, have their weights updated appreciably. Weights associated with the output nodes that are far away do not change significantly. It is here that the topological information is supplied. Neighbourhood output units receive similar updates and thus end up responding to nearby input patterns.

### 5 Experimental Results

The database samples collected for the purpose of this research have been procured using Logitech webcam. The database was also acquired from the Indian Institute of Technology, Delhi to evaluate the performance of the proposed ear recognition system.

Radon transform applied on ear images are used to compute its 2D Projection image for angles varying between 0° and 179°. The Radon transform of ear image shown in the Fig.14 has been used to compute the projection of an image along the given angles. The Radon operator maps the spatial domain \((m,n)\) to the projection domain \((s,\theta)\), in which each point corresponds to a line in the spatial domain. Conversely, each point in the spatial domain becomes a sine curve in the projection domain.
one dimensional image vectors. The SOM can be trained up to 500 iterations. The similarity gain is calculated and it is matched with the generated databases. Fig. 15 shows the SOM Neural Network training tool and implemented up to 200 iterations. Fig. 16 (a) and (b) shows the SOM layer weights and weight vectors.

The recognition performance of Table 1 shows that the proposed method improves in the identification accuracy and speed as compared with the other existing methods. The reconstructed image shown in Fig. 18 is almost same as the test image.

Table 1. Comparison performance of different methods using two databases

<table>
<thead>
<tr>
<th>S.No</th>
<th>Methods</th>
<th>Number of test images</th>
<th>Recognition rate (%)</th>
<th>Execution Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PCA-FLDA</td>
<td>50</td>
<td>91</td>
<td>High</td>
</tr>
<tr>
<td>2</td>
<td>PCA-BPN</td>
<td>50</td>
<td>96</td>
<td>Low</td>
</tr>
<tr>
<td>3</td>
<td>Proposed Method</td>
<td>50</td>
<td>98</td>
<td>High</td>
</tr>
</tbody>
</table>

Fig. 15 SOM Neural Network training tool

Fig. 16 (a) SOM layer weights

Fig. 16 (b) SOM weight vectors

Fig. 17 Test ear image of training set
6 CONCLUSION
In this paper, Radon transforms an efficient tool to detect the directional features of an image and PCA to reduce the overall dimensionality without the loss of information for ear recognition system is presented. The Radon transform is used to convert the rotation into translation, improves signal-to-noise ratio and robustness to zero mean white noise. Multi-Class LDA has been used as a powerful tool for extraction of discriminant features. SOM which proved to be highly accurate for recognizing a variety of ear images obtained at different instants under uniform light conditions. The experiment conducted on two databases shows that SOM combined with Radon transform and PCA improves the overall recognition performance within minimal process time. Compared with the other existing methods the proposed ear recognition system increases in recognition rate and speed.

References:
[8] Atilla Ozmen, Emir Tufan Akman,Koray kayabol, Osman Nuri Ucan, The Combination of Steerable Filters and CNN for Edge Detection Application, Department of Electrical and Electronics Engineering, University of Istanbul, Turkey.


