Indian Institute of Information Technology, Allahabad

Thesis Report

On

"Human Age Estimation Using Fingerprints"

Shweta Sharma

IWC2013018

Under the Guidance of:

Dr. Vijay Chaurasia

Assistant Professor

IIIT-Allahabad

17th July,2015

CANDIDATE’S DECLARATION

I do hereby declare that the work presented in this thesis entitled “Age group Classification and Estimation Using Curvelet features”, submitted in the partial fulfillment of the degree of Masters of Technology (M.Tech), in Wireless Communication and Computing at Indian Institute of Information Technology, Allahabad, is an authentic record of my original work carried out under the guidance of Dr.Vijay Kumar Chaurasia. Due acknowledgements have been made in the text of the thesis to all other material used. This thesis work has been done in full compliance with the requirements and constraints of the prescribed curriculum.

Place: Allahabad　　　　　　　　　　　　　　　　　　　　　　　Shweta Sharma

Date:　　　　　　　　　　　　　　　　　　　　　　　　　　　　IWC2013018

CERTIFICATE FROM SUPERVISOR

I do hereby recommend that the thesis work prepared under my supervision by Harshita Tiwari titled "Age group Estimation Using Curvelet features Extraction and Classification" be accepted in the partial fulfillment of the requirements of the degree of Master of Technology in Electronics and Communication Engineering for Examination.

Date:

Place: Allahabad

Dr.Vijay Kumar Chaurasia

Assistant Professor, IIITA

Countersigned by Dean(A) ________________________

CERTIFICATE OF APPROVAL

The forgoing thesis is hereby approved as a credible study in the area of Electronics and Communication Engineering and its allied areas carried out and presented in a manner satisfactory to warrant its acceptance as a prerequisite to the degree for which it has been submitted. It is understood that by this approval the undersigned do not necessarily endorse or approve any statement made, opinion expressed or conclusion drawn therein but approve the thesis only for the purpose for which it is submitted.

Signature of the Thesis Supervisor ________________________________

Countersigned by the Divisional Head________________________________

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my respected guide, Dr.Vijay Kumar Chaurasia, Assistant Professor, Department Of Electronics and Communication Engineering, IIIT Allahabad,for his support and help during my work. He has been a great source of knowledge and has inspired me with many ideas which were very useful for this research.His guidance, encouragement and patience has made this thesis possible.I am very grateful to the Head,Department of Information and Technology, IIIT Allahabad and all the other teachers and staff members of the department for their invaluable advice and assistance during my studies and project work.I extend my sincere gratitude to our Director,Prof. Somenath Biswas for his encouragement and support which removed all hurdles that came on the way of completion of my work.

Place: Allahabad Shweta Sharma

Date: M.Tech IIndYear, IIITA

PLAGIARISM REPORT

Human Age Estimation Using Fingerprints

By Shweta Sharma

FILE
TIME SUBMITTED 15-JUL-2015 03:08PM
SUBMISSION ID 555828942
WORD COUNT 5197
CHARACTER COUNT 28933
REPORT_FOR_PLAGUE.DOCX (2.01M)

ORIGINALITY REPORT

SIMILARITY INDEX	9%
INTERNET SOURCES	3%
PUBLICATIONS	7%
STUDENT PAPERS	2%

ABSTRACT

Every individual has unique fingerprints.Age estimation using fingerprints is done analyzing the minutiae features.To achieve good minutiae extraction in fingerprints with varying quality, preprocessing in form of image enhancement and binarization is first applied on fingerprints before they are evaluated.The issue of fingerprint feature extraction based on Curvelet transform and Singular Value Decomposition has been analyzed. The Curvelet Transform is a higher dimensional generalization of the Wavelet Transform.The Fingerprint classification is done by means of finding the Euclidean distance between the two corresponding feature vectors and hence matching is extremely fast.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURES PAGE NO.

OBJECTIVE

Digital - the universe of utilizations and administrations over PC systems or web is inclined to security issues. Instruments that are utilized for assaults are effortlessly accessible over the web and can be downloaded free of expense. This straightforward entry and ease of use, and secrecy that is effortlessly achievable by the aggressor are the underlying drivers of the issue. This makes specialists give more driving force to the field of digital security and give critical consideration regarding uproot the obstructions that come their approach to secure the present and future clients of web. Digital security itself is an immense zone covering security of everything that depends on web and PCs. Out of the quantity of issues, tyke security over digital world is one of the essential angles. Juvenile clients, for the most part youngsters, who are obscure to dim side of the digital world, are simple focuses for assailants. This implies in this manner that kids who are dynamic on web can turn out to be simple prey to aggressors. The United Nations General Assembly in the year 1989 has given its command and sanction the UN tradition on the Child Rights. Especially Article 17 gives critical regard for the insurance of kids from the material and data which may be harmful to their prosperity.

Age bunch estimation via robotized frameworks has been esteemed to be a testing issue by existing strategies since individuals from distinctive races have differing rates of maturing procedure. This is not just administered by the people groups' race and qualities however by numerous different variables, for example, working environment, living style, sociality and wellbeing condition to name a couple. Existing age bunch estimation systems use facial pictures for estimation and examination. In a circumstance when we oblige age bunch estimation and validation, fingerprints are the most suited biometric methodology to accomplish this errand. Age-bunch estimation through fingerprints can be of extraordinary use in ranges where we are not really worried about the character of a specific individual, yet we need to know under which age gather his or her age lies. It has been accounted for in a study done in the zone of fossil science that fingerprints contain discriminative elements in light of which we can separate in the middle of youngsters and grown-ups.

INTRODUCTION

The biometric techniques applied regarding fingerprints,faces and iris has been widely used for identification security.To prevent the identification fraud happening unprecedently in society,many techniques has been practiced like personel identification system or biometric verification considering various aspects of personel details.Under the same projection,fingerprint identification algorithms are implemented and extensively well performed and has been widely adopted for personel identities.Age estimation using fingerprint is a optimistic field that still remains unveiled.

Each individual has unique fingerprints nominated as its identity.A fingerprint is the representation of epidermis of finger.It constitutes unique pattern of straight or curved lines as ridges and valleys.The age can be determined using fingerprints by analysing the breadth or count of ridges.As the human body develops with the age, size of hand alongwith breadth of ridges increases but number of ridges remains unchanged.

History of fingerprint evolution:

In 1823,a man named Purkinje,was first to classify the fingerprints into 9 patterns.

The fingerprints were initially used by Chinese to save the records for business and to identify the children and mainly used fingerprint ink paper to obtain the impressions in the 19th century. In 1857,while working as a Chief Magistrate an Englishman,William Heschel,first time recorded the fingerprint to prevent the fraud for business documents.A doctor observed the impression of fingerprints left by a craftsman on the clay. So he chose to add to the unique finger impression grouping framework with the assistance of Charles Darwin.He suggested Galton to work with him and he accepted.Galton gathered about 8,000 different fingerprints to analyse.In 1892,he composed a book named ‘Fingerprints’ containing first outline about the fingerprint classification.Sir Edward Henry,using Henry Galton techniques,produced his own classification system based on the flow patterns,directions and other characterstics of fingerprint ridges.He derived equations from these characterstics to distinguish fingerprints of one person from others.

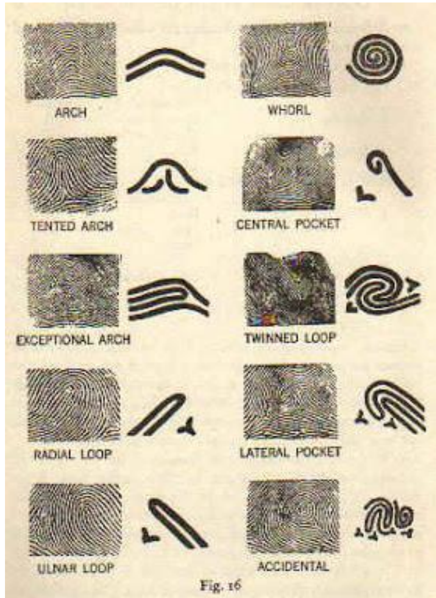


FIG 1.Patterns in fingerprints

Material Used for Fingerprinting:

Initially a chemical named Superglue (cyanoacrylate) was adopted to obtain the impressions of fingerprint.When it is heated in the presence of vapours,it sticks to the fingerprints left behind on the object and an whitish impression of fingerprints emerges on the object.The technique is known as Vacuum Metal Deposition (VMD).Superglue mainly constitutes gold film followed by zinc.

New technique that is used nowadays is to print the finger impressions using silver nitrate.When the silver nitrate reacts with chloride present in the oils of hand then forms silver chloride.When it is exposed to UV rays then it becomes visible.

New technique that is used to obtain the impression of fingerprints is to scan the fingerprint images through a scanner with optical or capacitive techniques.For this the scanner uses a light sensitive device named Charge Coupled Device and stores a digital image produced of fingerprints.

Types of Characteristics of fingerprints :

There are 3 types of features used to describe the information used for fingerprint characteristics.

Level I:

These features can be seen with naked eye.eg:cores in form of loops,whorls and arches & deltas that often occur on the sides of loops and whorls.

Level II:

These features include minutiae such as ridge bifurcations and ridge endings.Fingerprint variability exists due to relative positions of these features across fingers and individuals.These features are known as minutiae.

Level III:

They include the positions of sweat pores and ridge thickness.

Fingerprint processing as image :

An image is represented in the form of a matrix of pixels. The values of pixels are proportional to the brightness of the corresponding scene in the scene. The range of brightness depends on the type of values either gray or rgb.For gray,the range of pixel values varies from 0 to 255.

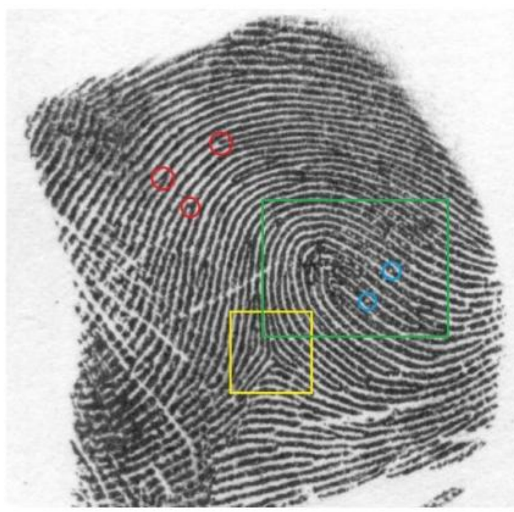


Figure Various image features commonly identified by expert examiners. Red circles indicate *minutiae* (ridge bifurcations or endings); blue circles indicate pores (they appear as small white dots along a ridge); the yellow square indicates the delta; the green rectangle indicates the core, in this case a leftward loop. doi:10.1371/journal.pone.0094617.g002

Former Fingerprint Matching Style :

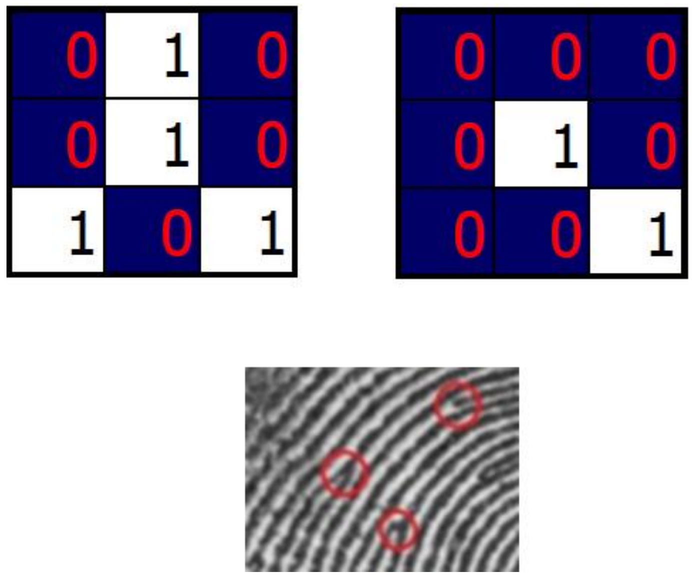


FIG 3.Matrix Representation of ridge bifurcation and ending

Minutiae can found out matching the template pattern of ridge ending and bifurcation as shown in fig.1(a) and (b) respectively.While finding the minutiae,its two relevant characteristics are obtained as of:position and orientation.These are saved as descriptor in the form of a matrix.

When minutiae matching is to be done,then the correlation is found out between the minutiae,by applying two constraints:

1)Euclidean distance between two minutiae point should be less than a threshold.

2)difference between their orientations should be in tolerant range.

Figure shown below illustrates the same process of finding whether the fingerprint pattern matches the input minutiae of fingerprint.

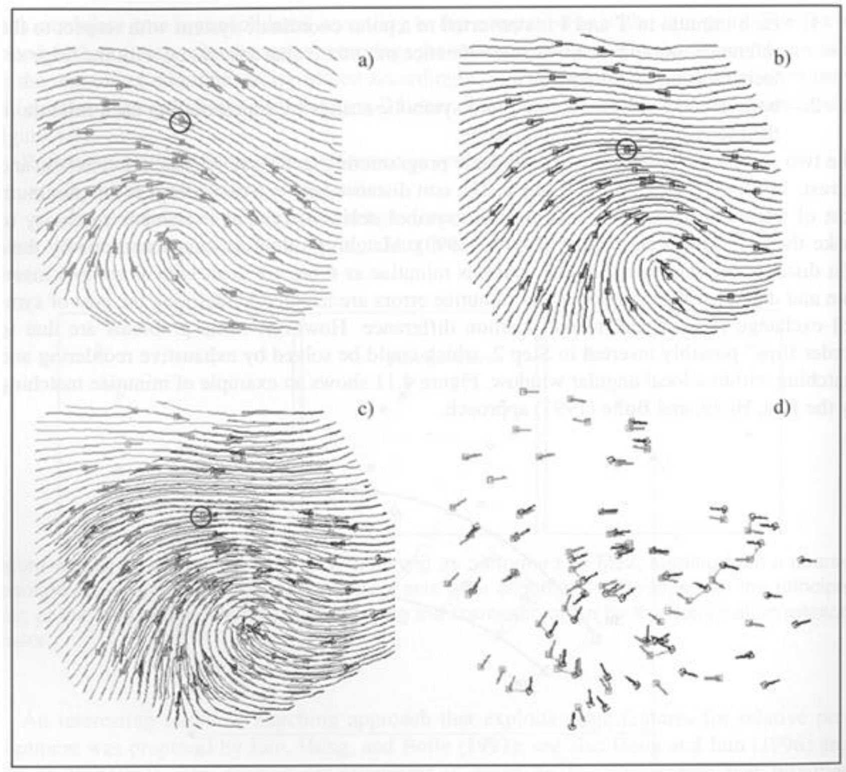


FIG 4. Minutiae Matching

LITERATURE SURVEY

Pavleen Thukral [1] and members have performed the work on automatic human age estimation using face images.For the purpose they have analysed the structure and appearance of faces using geometric and textural analysis of images.The analysis is covered in three steps of feature extraction,learning and classification for different age groups.Firstly,the features rely on the corner patterns of images as the landmarks of face like nose,eyes,lips etc.These landmarks are very perceptive to affine tranformation.So as a remedy the mean of all such points is calculated and average face is buit, then velocity vector is taken over these average faces.These vectors are further used for learning.In case of learning,a separate regression is needed to learn in order to study about every age.The method is like that with the estimation of weights,mean absolute error is aimed to be minimum and sequentially the weights are finallized for each age group.For classification,they have used five different classifiers as of SVC,Fisher Linear,Partial Least Square,KNN,and Naive Bayes and finally the classifier fusion is used to classification in different age groups.So the results are such that 1002 images of dataset containing 68 subjects are divided into groups like 0 to 15, 15 to 30, and 30+.and an accuracy of 70% is attained using classifier fusion.

In paper [2] the main technique used for fingerprint recognition is phase graph matching where the fingerprint patterns are designated as relational graphs.The graphs mainly consists of "core" and "deltas" as the relational attributes and various graph clustering policies like hierarchical or hashed databases is used to reduce the graph matching time.When using the graph to represent the object like fingerprint,vertices are used to represent regions or features like ridges,bifurcations or end points of fingerprints and edges for the relation between these regions.When comparing two fingerprints or graphs representing fingerprints,then in this paper they have applied clustering and then isomorphism on two graphs.For clustering they have manifested K-Nearest Neighbour algorithm with Euclidean distance.With the increase in size of graph the performance time for graph processing with KNN clustering improves in comparison to method graph processing without clustering.

The authors [3] have adopted a method to find ridges using eight different mask patterns.For this purpose they have divided the image into 8x8 size blocks.The gradients are to find out at each pixel and then orientation is obtained configuring the variation of principal axis.After binarization,filtering of the blocks and thinning,minutiae are extracted.Minutiae mainly include ridge endings and bifurcation.The authors have used the concept of Crossing Number Concept.Value of crossing number designates the number of 1-valued pixels in a 3x3 template. eg :- crossing no. of value one designates ridge ending while value three designates bifurcation.Previous techniques use minutiae but they are unable to track extreme variations in orientation and scaling. Besides they perform well on high quality images and not vice versa, which require more space to store.

In paper [4] the author has demonstrated about age discriminability according the face images.They have divided the experiment into five modules.Out of it,the most important are age manifold and robust regression.The main aim of age manifold is to find feature vector (or a matrix relation between a set of aligned face images and its representation).For the same,many methods has been illustrated like Principle Component Analysis(PCA)(to maximize the projected variance and reduce the dimensions accordingly),Locally Linear Embedding(LLE)(to exploit the local symmetries and obtain optimal weights for reconstruction).Another important module is Robust Regression.Robust regression module tries to minimize the error between estimated age and actual age using Support Vector Regression (as of linear and non-linear regression).Finally the results are locally improved by sliding the result up and down to reach the optimal result.The results of all schemes are compared according to mean absolute error.and the local adjustment techniques outperforms.

In [11] the age is classified according to two constraints as of personalized and temporal i.e. for each person a different age pattern is simulated and it is ordered according to the time in years. The patterns are ordered in an array in the order of increasing age,the blocks(age) for which patterns are missing are labelled differently and for these blocks the patterns are estimated using PCA.Aging features on face vary from person to person.So the conclusive subspace is lettered according to commonality between patterns of different persons as well as according to personal aging standard patterns.And the lacking pattern part of personel aging pattern is approximated by putting into practice broad aging pattern model and then further refined by updated personel aging patterns.

In [15],the author has evaluated the relation between age and fingerprints pattern.The age can be determined using fingerprints by analysing the breadth or count of ridges.As the human body develops with the age, size of hand alongwith breadth of ridges increases but number of ridges remains unchanged.Ridge breadth is measured from centre of one furrow to the centre of next furrow.A very high correlation between age and ridge breadth as well as height and ridge breadth is found.The age groups were carved up among 4 to 7,8 to 11,12 to 15,16 to 19,and 20 above.There was significant difference between average ridge breadths of mentioned age groups.The margin of error for the prediction is around 4.5 years.Finally,it is stated that the age groups can be classified by according to the average ridge breadths but gender difference cannot be significantly estimated on the same measures.

In this [6] paper direction resolving properties of curvelet provide very high identification precision even in case of high compressive images.The main drawback of minutiae recognition scheme is that preprocessing,registration and orientation flow estimation slower the process of recognition.An alternative approach for fingerprint identification can be image based, for instance wavelet and curvelet.Curvelet is helpful in determining the edge discontinuities in images.It divides the frequency plane into concentric rings that represent different.Each ring is partitioned into angular wedges which designate different orientations.To classify the feature vector obtained from curvelet,fuzy KNN has been used Data in a “boundary”region, i.e. samples that could be said to be a member of more than one class, may be classified differently depending on the distance measure used

A classification method of fingerprint quality based on neural network:

The learning process of backpropagation neural network algorithm mainly comprises of two stages:

i) Information Transmission:Calculation of output at each layer in the whole course of transferring input from input layer to hidden layer.

ii) Error Back Propagation: It modifies weight factors from the output layer forward step by step using the gradient descent method to minimize the output error signal.

A set of four dimensional features vectors are constructed based on four quality indexes as effective area, energy concentration, spatial consistency, directional contrast and linear weighted sum extracted from each fingerprint image.BP neural network to test the test set and compare its classification accuracy with the classification accuracy is used

.

A tutorial on Support Vector Machine:

In order to build computational classification models (or “classifiers”) that assign samples into two or more classes. Support vector machines (SVMs) is a binary classification algorithm offers a solution to this problem .In SVM,a linear decision surface (“hyperplane”) that can separate two classes and has the largest distance (i.e., largest “gap” or “margin”) between borderline elements of two classes(i.e.,“support vectors”);If such linear decision surface does not exist, the data is mapped into a much higher dimensional space (“feature space”) where the separating decision surface is found;The feature space is constructed via very clever mathematical projection (“kernel trick”).

Fingerprint Based location positioning using KNN classifier:

In this paper the on-line stage of fingerprint system is presented. Firstly, train sample selection decides the most likely potential position and the train sample set changes with the test sample. KNN has its drawback. Its performance is sensitive to the value of K, the metric one chooses to discriminate candidate points.By selection of train sample and giving different distance weights for every data points according to the similarity with query point can avoid setting a constant value; After the selective work, the value of K can be determined, K also changes with the number of train sample set. Then, the matching algorithm is performed. We mainly improve KNN algorithm by the calculation of feature, distance weights. No matter dynamic selection of K or weights calculation have positive effects on the performance of KNN algorithm.

FEATURE EXTRACTION

Fourier transform is a way of mapping signal into its constituent frequencies.Magnitude designates the share of each frequency component where phase designates as the timing or position of that frequency component in the whole image.The plot of magnitude of fourier transform of a signal is called the spectrum of the signal.

Low frequency defines that the data is changing slowly like a simple course object occupying the entire space.High frequency defines the rapidly changing data like a page of text.The position of frequency components depends on either they are low or high.Low frequency components lie near origin while high frequency components approach farther away.Average value of the image can be determined by finding the lowest frequency components,as low frequency components contain more image information than high frequency.The main aim of fourier transform is to represent the image in the form of sinusoidal components.For a NxN image,the 2-D DFT is acknowledged by:

$$F(k,l) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i,j) e^{-i2\Pi\left(\frac{ki}{N} + \frac{lj}{N}\right)}$$

Short Term Fourier Transform:

Heisenberg Uncertainty Principle:

In time-frequency diagram,there is an uncertainty between finding a time as well as a frequency resolution.The more accurate the time resolution is,lesser is the frequency resolution and vice versa.

The main disadvantage of fourier transform is that it cannot handle non-stationary signals.Non-stationary signals are those in which there are more than one frequency as shown in figure.So the signal is analysed after multiplying the signal with a window such that the signal seem stationary in the window and the same window is shifted to the whole signal.It can be formulated as:

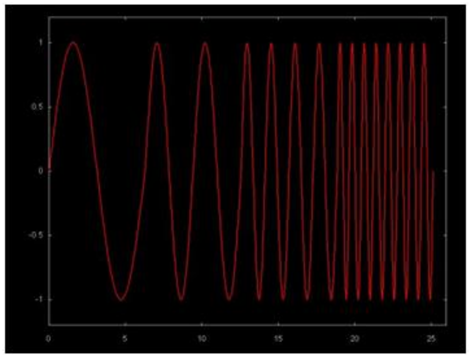


FIG 5.Non-stationary Signal

$$X(\tau,\omega) = \int_{-\infty}^{\infty} x(t)\omega^*(t-\tau)e^{-j\omega t}dt$$

Wavelet Transform:

Fourier transform permits decimation of time in the forward direction and decimation of frequency in the reverse direction,while wavelet allows decimation of time and frequency at the same time.Simultaneous decimation is helpful in explaining the frequency content of the whole image.i.e.the size of the window can be changed according to the significance of position and frequency as in case of STFT.As according to Heisenberg principle both cant be determined simultaneously therefore time resolution can be made significant at high frequency and vice versa.

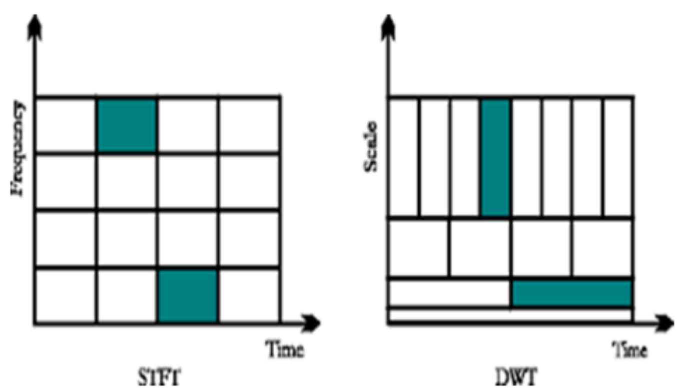


FIG 6.STFT V/s DWT

Discrete wavelet transform uses a space if basis function for image decomposition.In a 2D four combination of two filters:a high pass and a low pass are used defined as scaling function $\phi(x,y)$and a wavelet function $\psi(x,y)$.Four product produce the scaling function and three separable directional functions,accounting as a quaternary tree.

$$\Phi(x,y) = \phi(x)\phi(y) \quad (1)$$

$$\Psi_H(x,y) = \phi(y)\psi(y) \quad (2)$$

$$\Psi_V(x,y) = \psi(x)\phi(y) \quad (3)$$

$$\Psi_D(x,y) = \psi(x)\psi(y) \quad (4)$$

Thus the wavelet measure variations in three directions:horizontal,vertical and diagonal.These scaling and translating functions can be defined as:

$$\phi_{j,m,n} = 2^{j/2}\phi(2^j x - m, 2^j y - n) \quad (5)$$

$$\psi_{j,m,n} = 2^{j/2}\psi^i(2^j x - m, 2^j y - n) \quad (6)$$

i=H,V,D representing directions

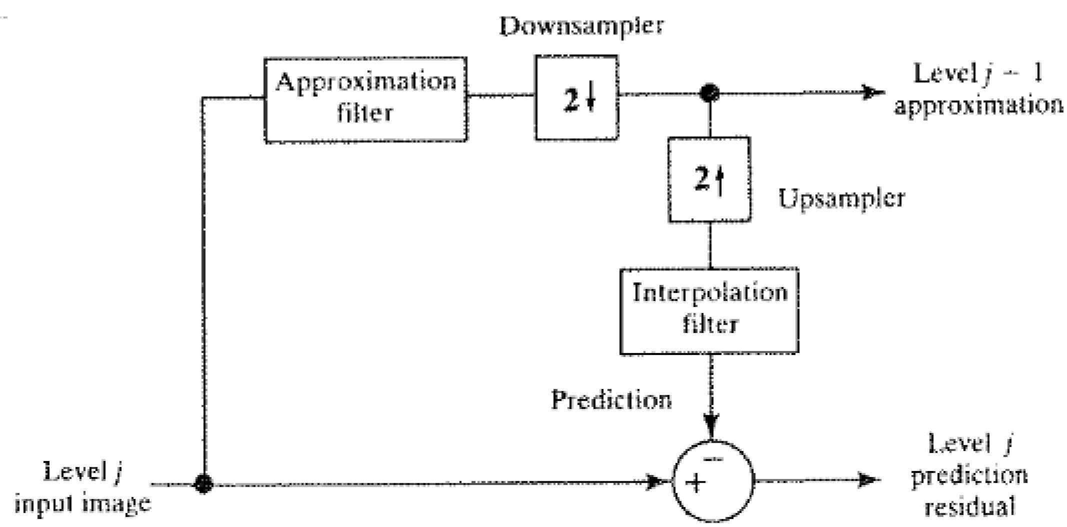


FIG 7.Block diagram for Wavelet Coefficient Calculation

Thus image can be disintegrated in multiple levels.At each level image is first downsampled by a factor 2 and then the difference between input image and approximation image can be saved as residual coefficients for reconstruction.

Curvelet Transform:

The wavelet that are used to compress image processing can decompose the image in four possible directions,namely horizontal,vertical and diagonal directions.In case of curves,there are lot wavelet coefficients required for representing the curves.As wavelets are efficient in representing the point level singularities.They can reveal the image features on both sides of edges, but not the features along them.Thus wavelet transform provides localisation and multiresolution but lacks on directionality and anisotropy.

1. Directionality means image can be decomposed in terms of orientation
2. Anisotropy implies representation of picture with the elements of prolonged shapes with diverse aspect proportion.

These features are attributed by curvelet transform and counterlet. The curvelet was introduced in the continuous domain through multi-scale filtering followed by a block ridgelet transform on each image.

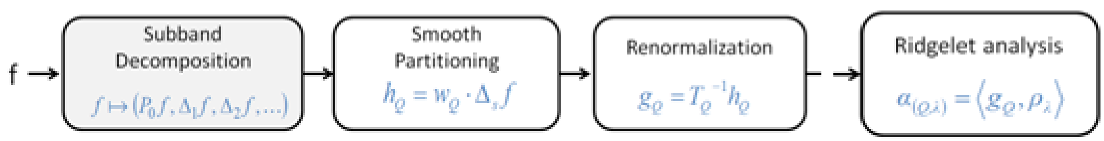


FIG 8.Curvelet Transform Steps

Curvelet transform can be analysed in four steps:

1. Subband Decomposition
2. Smooth Partitioning
3. Renormalization
4. Ridgelet Analysis

Sub-band Decomposition:

The image is filtered into subbands of frequencies to obtain multi resolution layers and the image can be reconstructed from the outputs of these subband filters as

$$f = P_0(P_0 f) + \sum_z \Delta_z(\Delta_z f)$$

where $\Delta_z f$ is the band-pass filter bank output and $P_0 f$ is the low pass filtered output of image f .
There is a relation between the filter banks of curvelet and wavelets as

- f can be decomposed into approximate,and other three directions
- $P_0 f$ can be obtained from approximate and any other directional output
- $\Delta_z f$ can be obtained by the combination of two directional filters.

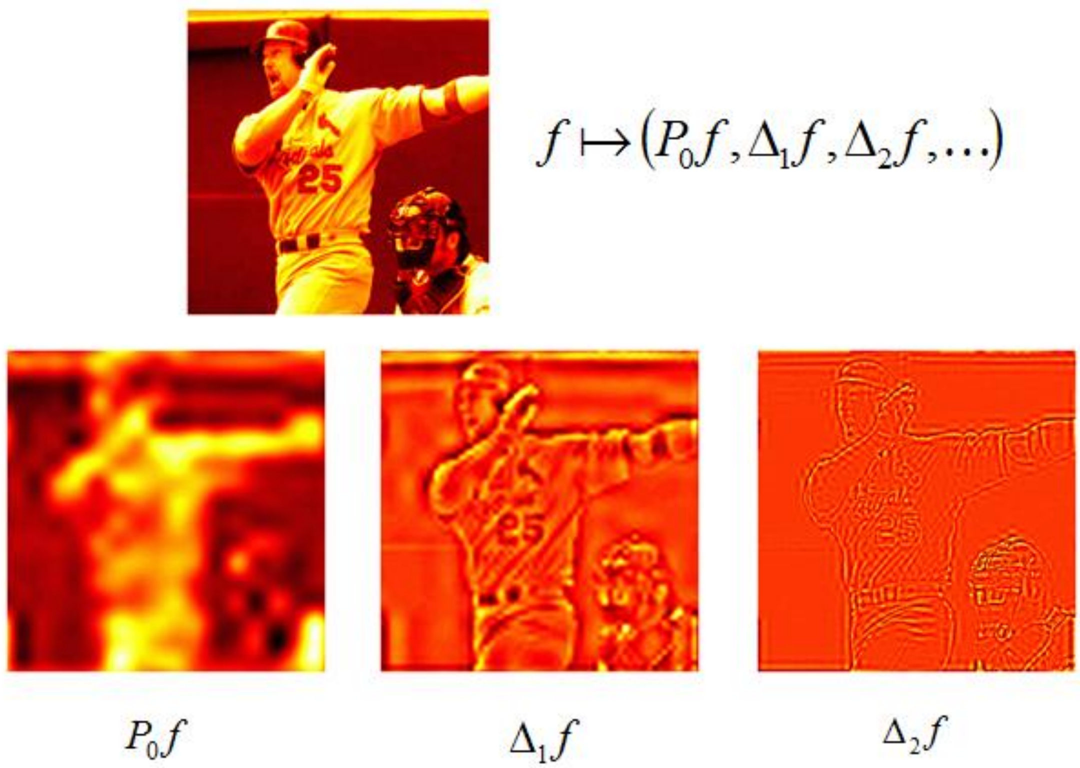


FIG 9.Subband Decomposition

Smooth Partitioning:

Each filtered image output from subband decompositioning is partitioned into smaller sections with the dimensions of $2^{-s} \times 2^{-s}$.and each block is smoothed so that each partition can be further analysed to obtain local curves and straight lines.

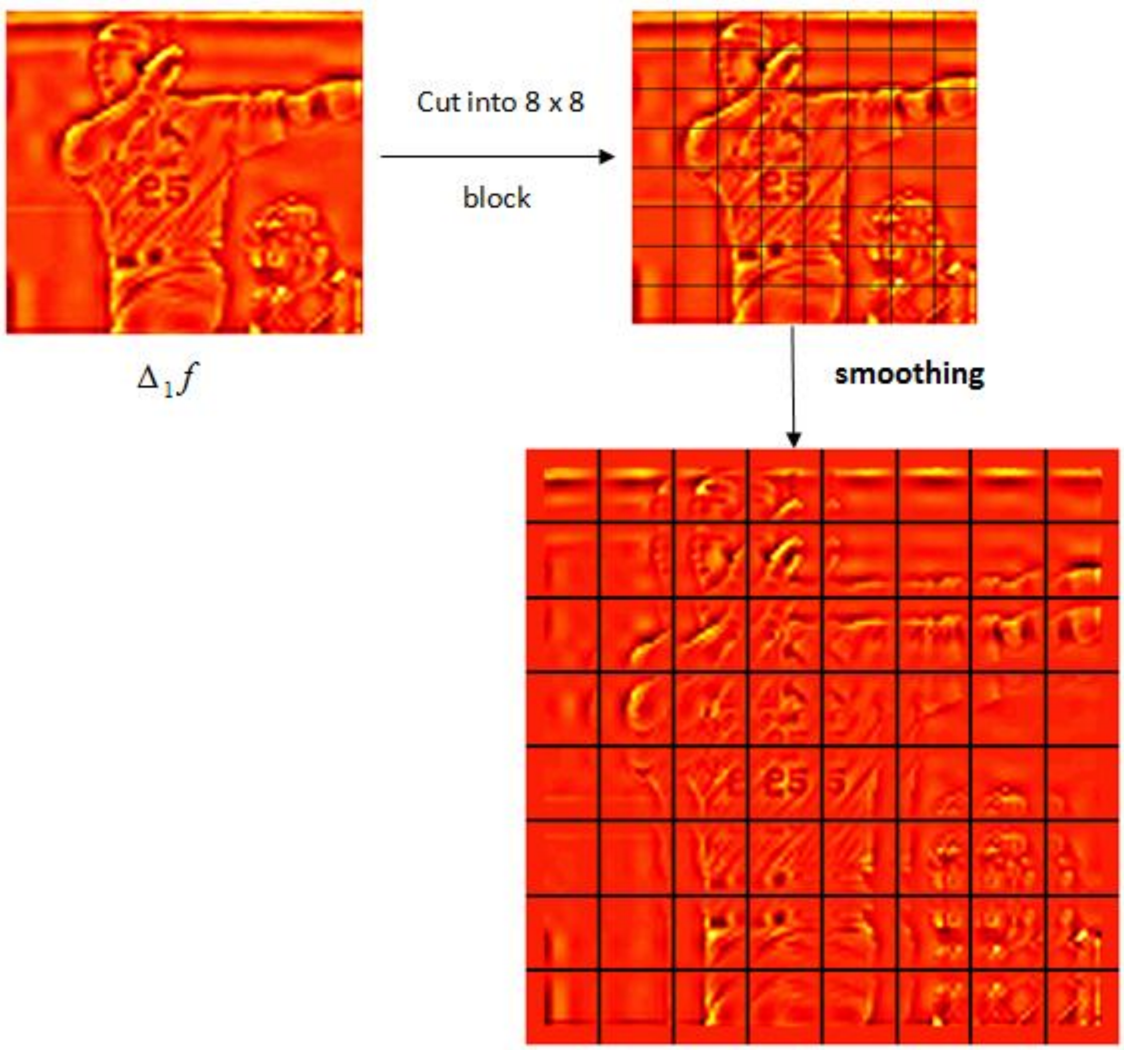


FIG 10.Smooth Partitioning

Each obtained block is normalised to unity.

Ridgelet Transform:

In case of wavelet transform,a lot of coefficients are needed to represent the edges.With the increase in coefficients,Mean Square Error also increases.To approach an optimum MSE,we need a few coefficients to represent the smooth surfaces or edges.Ridgelet transforms being introduced to represent smooth surfaces as well as edges.

Ridgelet transform of a bivariate function $f(x_1,x_2)$can be represented using the elements of form $s^{-1/2}\psi((x_1 cos\theta + x_2 sin\theta - t)/s)$where ψ is the wavelet transform s is the scaling parameter,t is the translation parameter and θ is the orientation.The ridgelets are constant along the ridge lines $x_1 cos\theta + x_2 sin\theta = t$.Many ridgelet coefficients superpose the lines.

Ridgelet transform coefficients can be characterized as :

$$F(s,t,\theta) = \int \psi_{s,t,\theta}(x) f(x) dx$$

And the function f can be reconstructed as

$$f(x) = \int_0^{2\pi} \int_{-\infty}^{\infty} \int_0^{\infty} F(s,t,\theta)\psi_{s,t,\theta}(x)\frac{ds}{s^3}dt\frac{d\theta}{4\pi}$$

Radon Transform:

Ridgelet transform can be computed as wavelet transform of the function in the radon domain.The radon transform of a function can be modelled as linear integral :

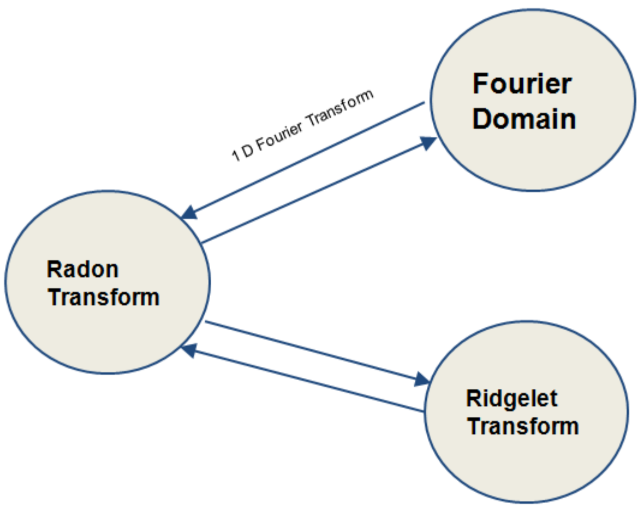
$$Rf(\theta,t) = \int f(x_1,x_2)[?][?](x_1 cos\theta + x_2 sin\theta - t)dx_1 dx_2$$

Ridgelet Transform

According to Projection Slice Theorem,1D fourier transform of a function on a specific position and orientation is equivalent to the values plotted on the line of the same defined position and orientation intersected on 2 dimensional fourier transform of the same function and strictly passing through origin.Thus Radon transform can be persuaded by implementing inverse FT of 2D FT confined to radial lines passing through origin. Following steps can be attributed to obtain discrete ridgelet transform:

1. Determine 2D FFT of function f and producing the output as a gridded data
2. The sampled values obtained on squared grid are placed on the polar grid restricted to radial lines passing through origin,using interpolation.
3. Determine 1D inverse fourier transform of the polar functional line obtained from previous step.

Frequency space is divided in polar coordinates where concentric circles represents scales and wedges represent orientation in that specific direction.



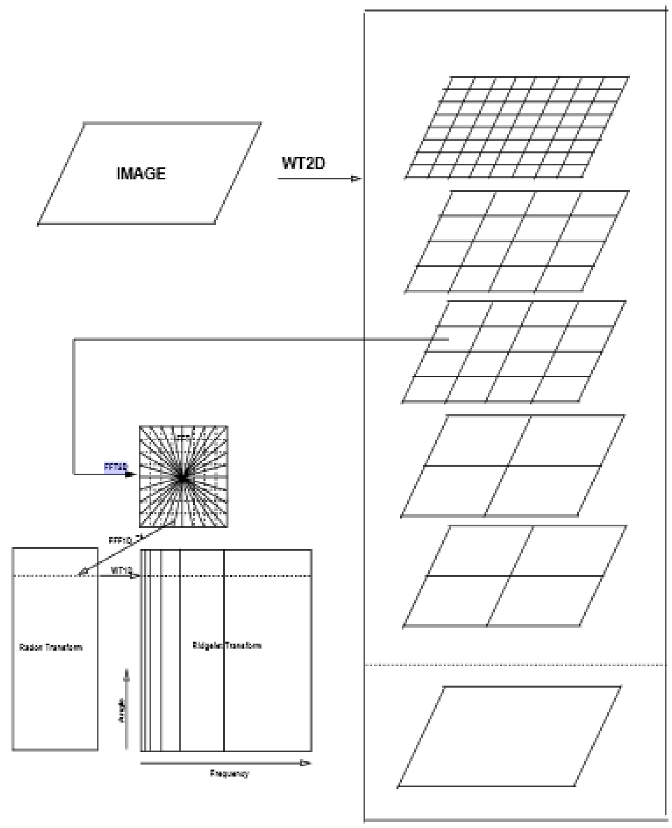


FIG 11.Relation between Radon,Fourier and Ridgelet Transform

FIG 12.Curvelet Transform

Converting to RectoPolar Grid :

Instead of using polar grid,rectopolar is prefered,As direct FFT can be implemented on the rectopolar grid.Though the lines are not uniformly distributed in angles but uniform in slope.Thus the transform can be designated as an array whose rows contain the projections and the radon transform can be obtained directly by applying 1D transform.

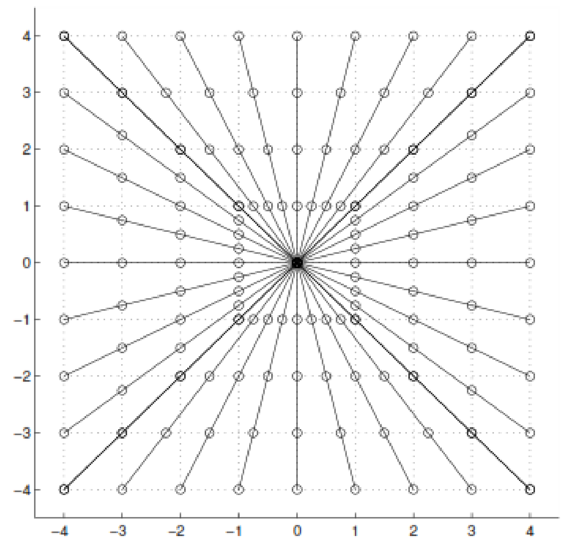


FIG 13.Recto-Polar Grid

These features are very well captured by curvelet transform[20] [21].The curvelet transform was carried out by performing multi-scale filtering in the continuous domain thereafter block ridgelet transform is carried out on each image[22]. Curvlet transforms provide us multi-scale and multiresolution transformation but for human age-group classification problem using fingerprint biometric we only use multiresolution properties of curvelet transform.

2G Curvelet Transform:

Due to mathematical complexity and quantitative analysis,1st generation ridgelet based curvelet transform is computationally cost.So the 2G curvelet transform is introduced in which a easy indexing in terms of scale,orientation and position is sufficient.The FDCT follows a parabolic scaling relation, which depicted at a scale of 2^{-j},having an envelope which is positioned along a sides of length $2^{-j/2}$ and width 2^{-j}.A faster, simpler and less redundant method is used as fast 2D discrete curvelet transform (FDCTs).Its takes the Cartesian array as input in the form f[t1; t2]; 0 t1; t2 < n and the coefficients c(j,l,k) Here (j, l, k); which is defined as below

$$c(j,l,k) = \sum_{0<t1,t2<n} f[t1,t2]\psi_{j,l,k}[t1,t2]$$

$\psi_{j,l,k}[t1,t2]$ depicts a curvelet transformation(digital) and Emmanuel Candes et al. reported two novel methods for fast discrete curvelet transform FDCTs.

1) Curvelets via wrapping ,

2) Curvelets via Nonequispaced FFT

Both the above algorithms shows same running time complexity as O(n^2 logn) for nX n Cartesian arrays,that is, n by n image.In comparison with FFT the computation time of FDCT via USFFT is 6-10 times more and also requires O(n^2) storage where storage,where n^2 is the total number of pixels in the image object.While doing curvelet transform initially the original image is decomposed into series of sub-bands to separate objects of different scales thereafter each and every sub-band is analyzed by using local block ridglet transform.The output of this algorithm is J + 1 sub-band matices having size nxn.

2D fourier transform is applied on f to get samples

1. At the onset we apply 2D FFT on an image object and obtain Fourier coefficients
2. Each sample is resampled in each scale and orientation to obtain $f[n_2,n_1 - tan\theta_l]$ where θ_l is the orientation and n_2,n_1 are the position of samples according to the window function to wrap around the origin.It is simply the re-indexing of the transformed output,to localise the samples in a rectangular grid in order to find the inverse transform.
3. Then the parabolic window $\widehat{U}$ called Cartesian window (as shown in figure 4) is multiplied with the sheared object $\widehat{f}$,this localizes $\widehat{f}$ close to the parallelogram with the orientation,which provides

$$f_{j,l,k}[n_2,n_1] = f[n_2,n_1 - tan\theta_l]\widehat{U}[n_2,n_1]$$

4. Inverse FT is applied to obtain curvelet coefficients.

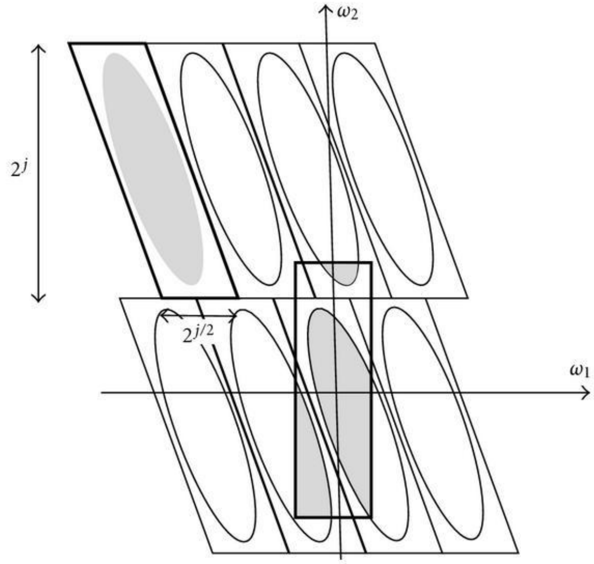


FIG 14.FDCT Coefficients Wrapping

$\widehat{U}$ [n1; n2] is the rectangular Cartesian window of length $L_{1,j}$ and width $L_{2,j}$.Further the length $L_{1,j}$ equals 2^j, because of the parabolic scaling, and width $L_{2,j}=2^{j/2}$ as shown in figure14.

P_j is defined as

$$P_j = (n1; n2) : n_{1,0} < n1 < n_{1,0} + L_{1,j}$$

$n_{2,0} < n2 < n_{2,0} + L_{2,j}$ where $(n_{1,0}, n_{2,0})$ is the index value of the pixel found at the bottom-left of the $\widehat{U}$ [n1, n2] , the rectangular Cartesian window. Thus with above steps of implementing FDCT using USFFT can simply be equated as

$$c^D(j,l,k) = f[n_2, n_1 - tan\theta_l]\widehat{U}[n_2, n_1]e^{-2\pi(k1n1/L1,j + k2n2/L1,j}$$

FEATURE CLASSIFICATION

The features that are extracted using curvelet are further classified into subcategories.These subcategories can be two or more in number.For illustration consider a set images are given and we have to distinguish the cat from other images including vehicles or dogs or buildings etc.Then we have to classify the images into two classes.The features that are obtained from feature extraction are utilized for this purpose.

Learning Stage:

Computer dont have brains or logics to distinguish automatically.So we need to make them learn on the basis of knowledge of features.Learining stage is of two types: supervised or unsupervised.supervised learning is like make the children learn about his surroundings by pointing the object and naming it.Supervised learning is to learn about objects by classifying the objects according to their commonity in features.

The classification process executes in two steps:

1. Training:the inputs are given which are already classified in the form of matrix along with their categories mentioned. Using the relation between features and classes,a function is derived
2. Testing:new inputs with unknown categories are input into classifier to run the function on these images to find their classes.

Neural Network:

Perceptrons:They are the basic unit of Neural Networks.In the specimen shown the perceptron has a handful of inputs, x_1, x_2, x_3. In commonplace it could take on more or fewer inputs. Rosenblatt professed a genuine rule to reckon the output. The weights, w_1, w_2,…, real numbers expressing the significance of the refined inputs to the output. The neuron's output, 0 or 1, is resolved by inevitably the weighted sum $\sum_j w_j x_j$ is less than or greater than some threshold value. Threshold is a just a real like the weights.Thus the logic can be formulated in equation as

$$\text{output} = \begin{cases} 0 \text{ if } \sum_j w_j x_j < \text{threshold} \\ 1 \text{ if } \sum_j w_j x_j > \text{threshold} \end{cases}$$

where b= - threshold

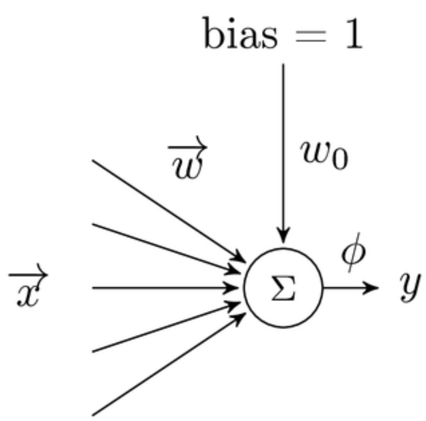


FIG 15.Perceptron

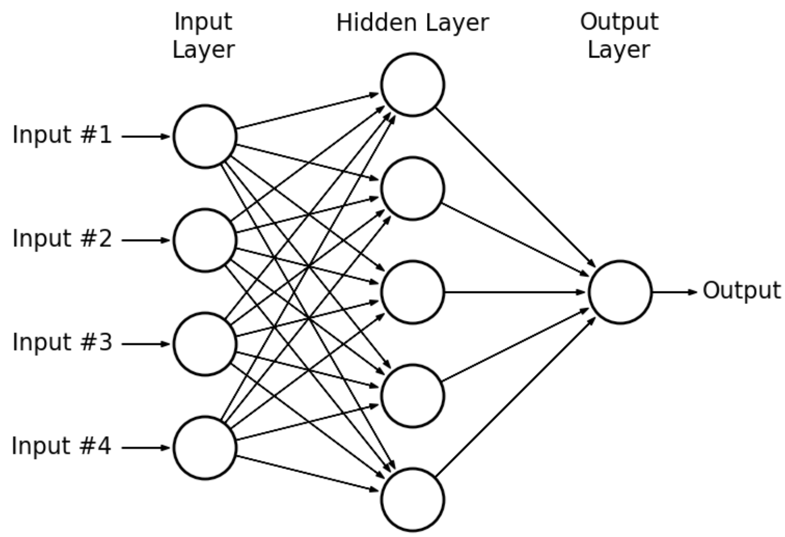


FIG 16.Neural Network

In order to formulate more complex problems layers are introduced and outer layers are to take in and out the inputs and outputs respectively and the layers that are used for computing the function only.Each layer contains one or more perceptrons.

Sigmoid function:

The output of a neural network can be drastically changed from 0 to 1 or 1 to 0 by a small change in the value of weight or bias.So it would be easier if the value band of input can be broadened instead of two discrete values.To get the input of value other than 0 or 1,we need to multiply the the output by a function to adjust the output accordingly.Thus if σ is the sigmoid function the output will be σ(wx+b).

$$z = \begin{cases} 1 & \text{if } \sigma(wx+b) > 0.5 \\ 0 & \text{if } \sigma(wx+b) < 0.5 \end{cases}$$

The idea of training is to minimize the error between actual output and desired output using stochastic gradient descent.The optimal value is when it attains global minimum so the weights are increased in small steps till optimal value is reached.

Backpropagation:

The cost function for a fixed training example value x can be written as :

$$C = \frac{1}{2}\sum(y - a_L)^2$$

K-Nearest Neighbour:

k-nearest neighbour is the unsupervised learning method in which the feautres are classified on the basis of their commonality.It is generally used method to classify objects that are closest to each other.Here k represent the majority number of votes to a class devoted by the neighbours.Neighbours are those nodes or objects where are already assigned their correct respective classes.It is a learning procedure in which the actual function is approached locally until no improvement in classification. In K-NN, an item is characterized by a greater part vote of its neighbors, with the article being doled out to the class most regular amongst its k closest neighbors (k is a positive whole number, ordinarily small).If k = 1, then the article is basically appointed to the class of its closest neighbor.

Fundamentally NN strategies accept the class of the closest occasion from x,as the class of an occurrence x. With a specific end goal to focus the closest occasion, NN strategies receive a separation metric that measures the closeness of example x to every single stored occurrence. Various distance metrics can be used, including the Euclidean. The KNN algorithm does not require any learning stage.They simply store the training set.When a new test sample is input,the distance of the test sample is determined and the class is assigned accordingly.

We can compute the distance between two scenarios using some distance function d(x, y), where x, y are scenarios composed of N features, such at x= $\{x_1, ..., x_N\}$, y=$\{y_1, ..., y_N\}$
Two distance functions are discussed in this summary:

- Ø Absolute distance:

$$d_A(x,y)= \sum_{i=1}^{N} |x_i - y_i|$$

- Ø Euclidean distance:

$$d_E(x,y)= \sum_{i=1}^{N} \sqrt{(x_i^2 - y_i^2)}$$

It is worth mentioning that when the set of training instances is large,k-NN based techniques contribute a high computational push to perform as a classifier.This happens on the grounds that for each new query (q) the entire training set should be gone by.. The idea of weighting features when utilizing a k-NN algorithm is used to give more significance, in the classification procedure, to relevant features. Consider a dataset containing m features ($f_1,f_2,...f_m$), where only one (f_j) is relevant for classifying instances. Two instances having the same f_j value, however, may be distant from each other in the m-dimensional space; this shows that non applicable features can assume an essential part in the classification process, dominating the distance measure. Attempting to optimize this problem, the use of weights associated with features is recommended to perform the calculation of the distance between them. KNN has been widely used for the pattern recognition, the reason behind the use of a KNN is that it is nonparametric (so does not require tuning), and is easy to implement

SINGULAR VALUE DECOMPOSITION

The main aim of SVD is to compute the eigenvalues so that the redundant values or rows of a matrix can be ignored. Principle Component is based on the same concept.

The SVD is the factorization of any (k x p) matrix into three matrices, each of which has important properties. That is, any rectangular matrix A of k rows by p columns can be factored into U, S and V,given as

$$A = U.S.V^{T}$$

Where $U = AA^{T}$

and $V = A^{T}A$

Here U and V are the lower triangular and upper triangular matrices. S matrix is the representational diagonal matrix with non zero values on the diagonal only and these values are equal to the rank of matrix A.Every value is square root of squares of eigenvalues of U and V.These values is placed in ascending order from left corner to right corner of the matrix. These values are stored in the form of Eigenvectors in the matrix S.

PROPOSED APPROACH

The approach we followed for estimating human age group using fingerprints is shown in the flow chart depicted in Fig.1.We extracted the discriminating features of fingerprints for training and testing sequences using Curvelet coefficients.The features so derived were having very hig dimensionality.So to remove this high dimensionality we projected the extracted features into principal component (PCA) subspace[17].At last we performed child identification or age group estimation using K-nearest neighbour(KNN)[18]classifier and arrive at the fingerprint age group.

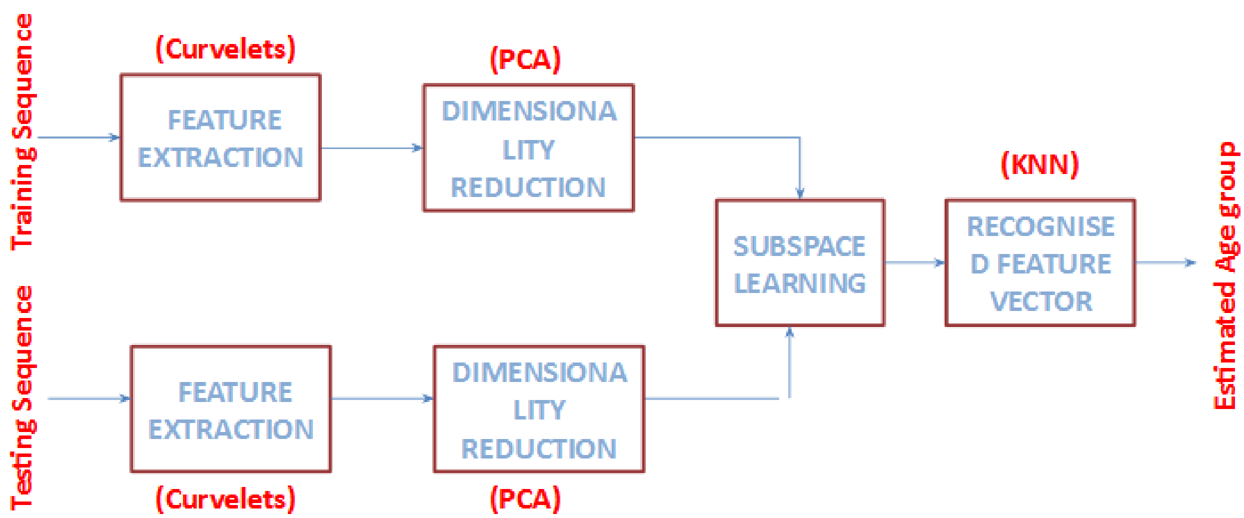


FIG 18.Methodology of Proposed Approach

Learning Stage:

Fusing the feature vectors obtained from SVD of extent 1x256 and from FDCT of extent 21x21, feature vector is formed of extent 1x697.

[Input]: fingerprint images X_i with known age class A

[Output]:Dimensionally reduced Training and Testing sequence matrix

Algorithm:

Step 1(Initialization): Let X_i be the image in Age class A

Step 2(Feature Extraction):Generate Curvelet coefficients using ridgelet transform for different orientation and scales .These coefficients will form the feature vector C for our image X_i

Step 3:Repeat above steps for each image X_i corresponding to age class A_i

Step 4(Dimension reduction & partition):After every Curvelet feature vector C_i of every image X_i in class A_i is collected as feature matrix. We partition the feature matrix into training sequence matrix and testing sequence matrix.Project both training sequence matrix and testing sequence matrix into PCA subspace.

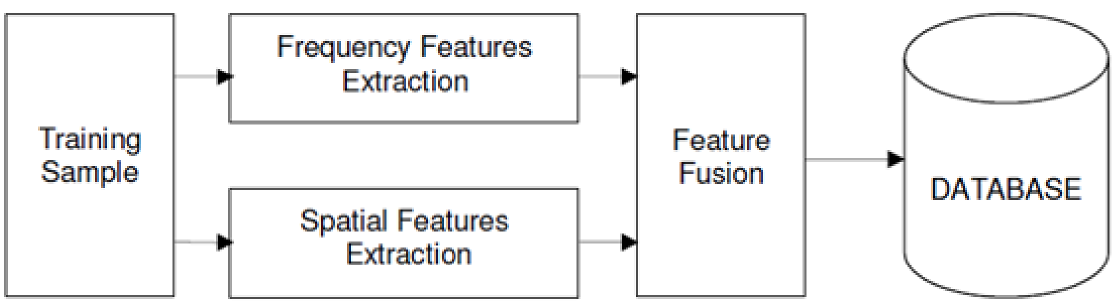


FIG 17.Learning Stage of the Proposed Classification System

Classification Stage:

[Input] unknown fingerprint and the feature database
[Output] the class of the fingerprint to which this unknown fingerprint is assigned

1) Decompose the given unknown fingerprint with 4 level scale high pas filtering and 8 diiferent orientation of FDCT.

2) Calculate the sub-band energy vector (E) using FDCT.

3) Reduce dimensions by finding the maximum eigen value with PCA.

4) Apply Neural Network and estimate the class of input fingerprint utilizing the matrix generated in learning stage.

DATABASE

Though there are various fingerprint databases available,they do not contain age varying discriminating features.For training and testing purpose we have developed in house database of digital fingerprints in three age groups:

- 6-10 ! (360 fingerprints of 600 dpi each)
- 10-14! (360 fingerprints of 600 dpi each)
- 14-18! (360 fingerprints of 600 dpi each)

360 fingerprint images are taken(180 males+180 females)for each age group.Fingerprints are numbered on the basis of following numbering scheme shown in Fig.7.

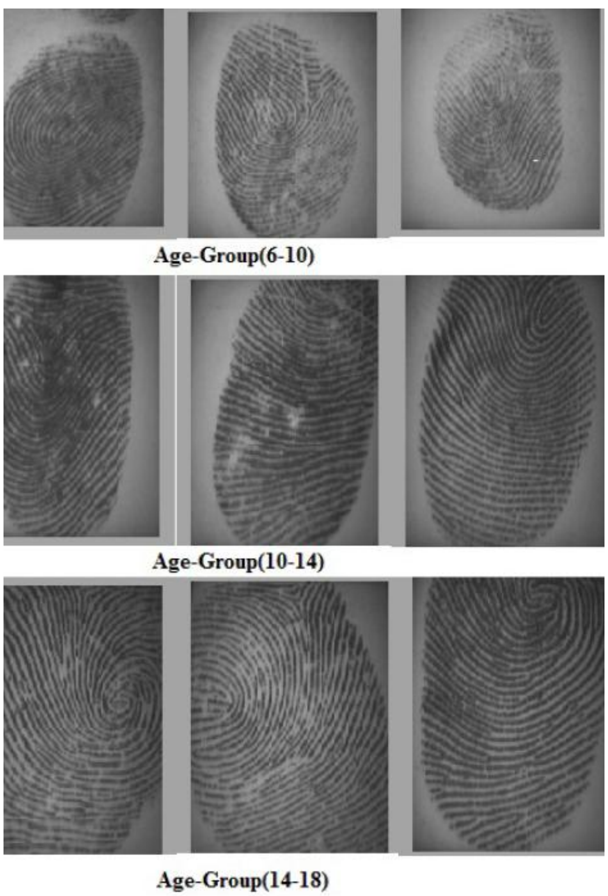


FIG 19.Database Illustration

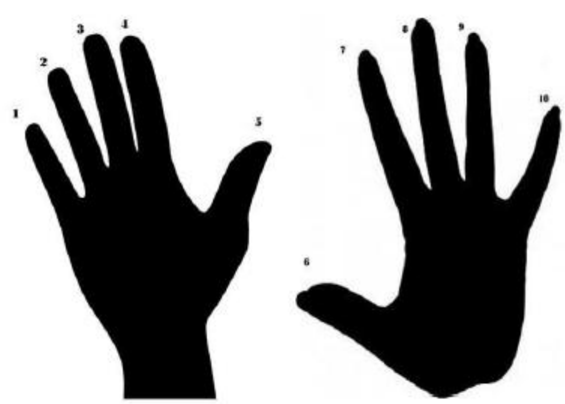


FIG 20.Numbering of Fingers

RESULTS

1. These are the images computed using curvelet coefficients at 5 different scales i.e. 5 different levels of different size and 8 different orientation with 40 degree as initial value.Image at first level shows the low pass filter or approximate value of every fingerprint.and the feature values that are used for classification is of 5th level.

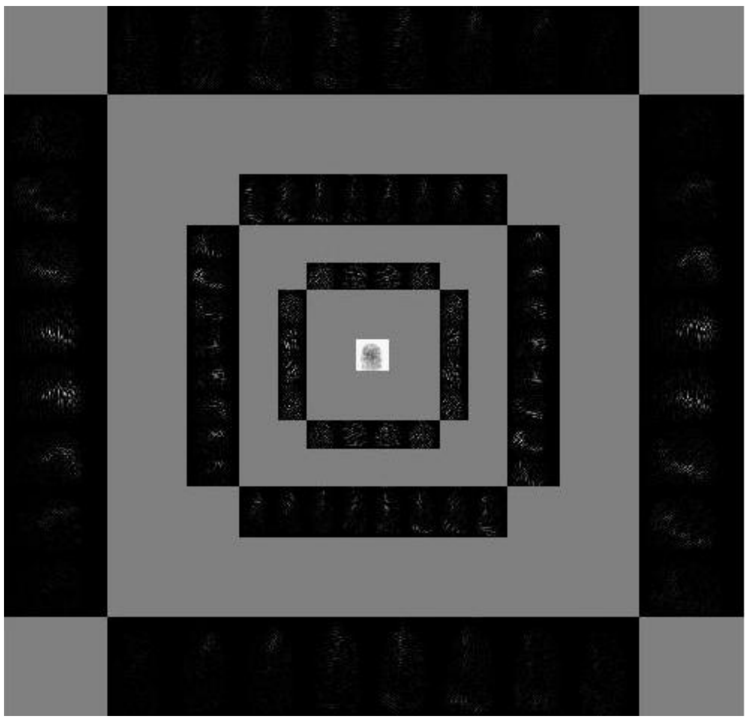


FIG 21.Curvelet Result.

2. 2^{nd} plot shows the confusion matrix.Confusion Matrix literally means to compute the percentage of samples which are confused in two or more classes.

where +ve is the class on which we want to focuss and -ve are the other classes.

Confusion Matrix	Predicted		
Abstract		-ve	+ve
	-ve	a	b
	+ve	c	d

- True Positive Rate = $\frac{d}{c+d}$
- False Positive Rate= $\frac{b}{a+b}$
- True Negative Rate= $\frac{a}{a+b}$
- False Negative Rate= $\frac{c}{c+d}$
- Precision= $\frac{d}{b+d}$
- Accuracy= $\frac{a+d}{a+b+c+d}$

FIG 19.Confusion Matrix

3. Receiver Operating Characteristics:

It is the plot of true positive rate against false positive rate.Accuracy depends on how well the test separates the group being tested into positive and negative.Accuracy can be computed as the area under ROC curve.

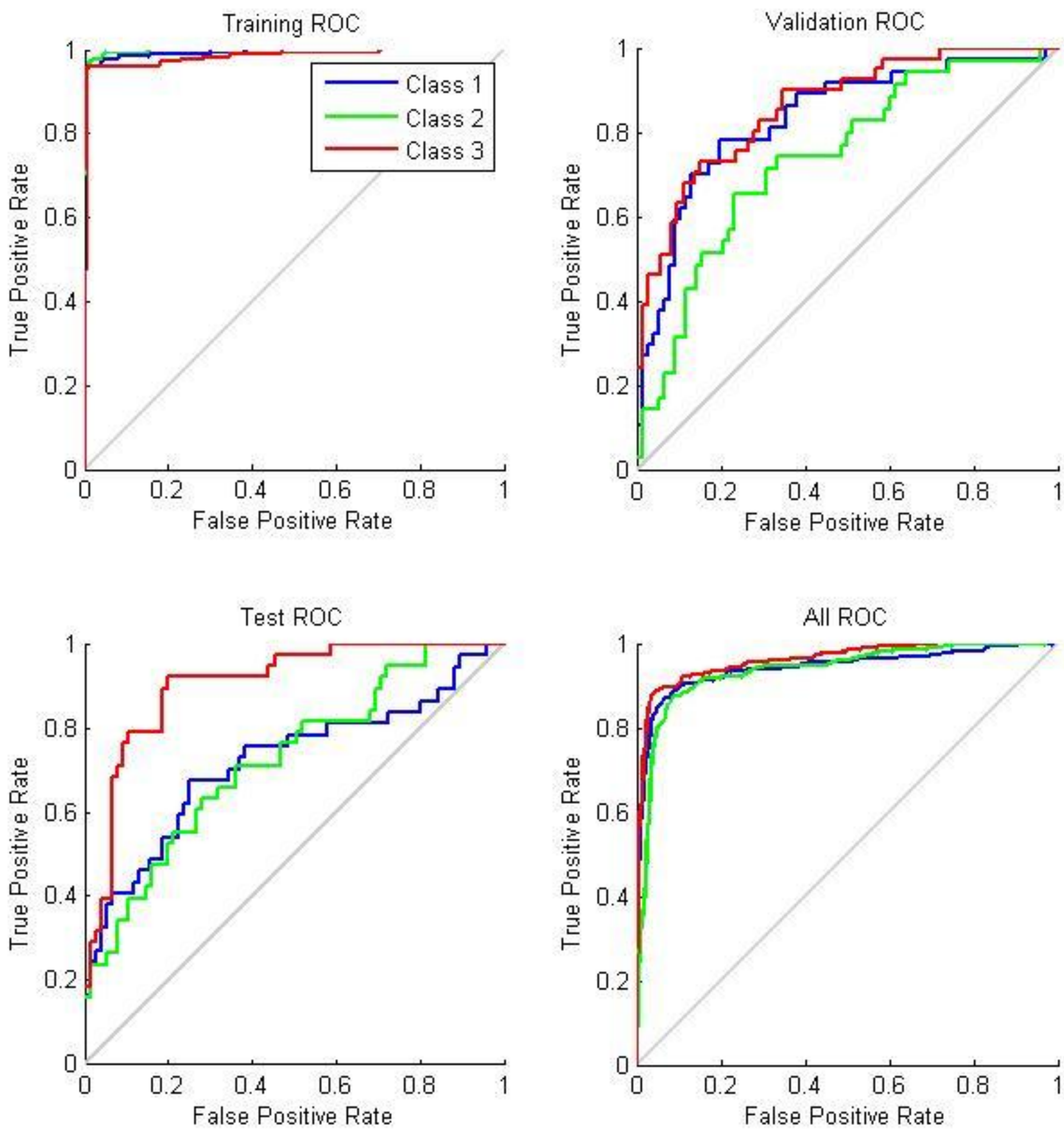


FIG 20.ROC Curve for Testing and Training Set

4. Mean Square Error:

Mean square error is the difference between the actual value and estimated value.According to the solution it can be concluded that as we train the system the MSE decreases.ANd during testing the best value attained is 0.14593 during the testing and validation.

An epoch is number of attempts the training vectors are utilized to update weights of the perceptrons in neural network. For group training the greater part of the training go through learning at the same time in one epoch to upgrade weights.

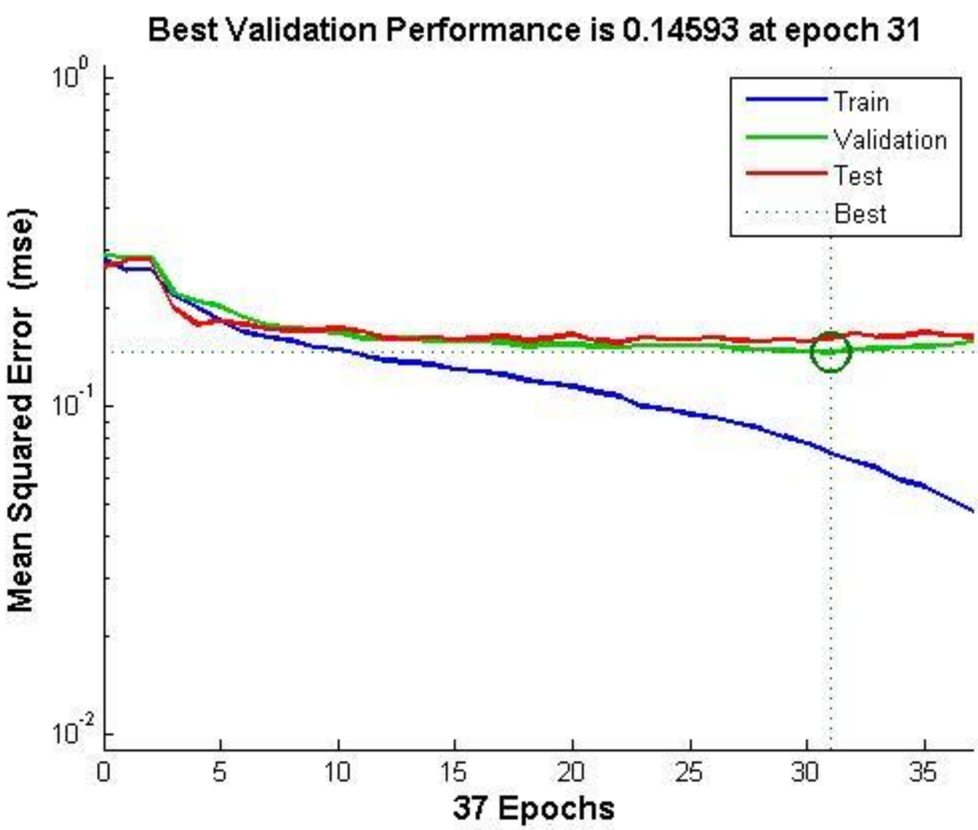


FIG 21.MSE Graph

REFERENCES

[1]:http://www.indiabroadband.net/india-broadband-telecom-news/11169-some-stastisstics-about-internet-users-india.html

[2] U. G. Assembly, “Convention on the rights of the child, 20 november 1989, united nations, treaty series, vol. 1577,” 1999.

[3] G. Guo, Y. Fu, C. R. Dyer, and T. S. Huang, “Image-based human age estimation by manifold learning and locally adjusted robust regression,”Image Processing, IEEE Transactions on, vol. 17, no. 7, pp. 1178–1188,2008.

[4] N. Ramanathan and R. Chellappa, “Face verification across age progression,”Image Processing, IEEE Transactions on, vol. 15, no. 11,pp. 3349 –3361, 2006.

[5] Wikipedia. Ecrm. [Online]. Available:http://en.wikipedia.org/wiki/ECRM

[6] J. E. Kloeppel. (2008) Step right up, let the computer lookat your face and tell you your age. [Online].Available:http://news.illinois.edu/news/08/0923age.html

[7] E. Patterson, A. Sethuram, M. Albert, K. Ricanek, and M. King,“Aspects of age variation in facial morphology affecting biometrics,”in Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007.First IEEE International Conference on. IEEE, 2007, pp. 1–6.

[8] A. Majumdar and R. K. Ward, �Fingerprint Recognition with Curvelet Features and Fuzzy KNN Classifier�, IASTED Conference on Signal and Image Processing 2008.

[9]K. Ricanek Jr and E. Boone, “The effect of normal adult aging on standard pca face recognition accuracy rates,” in Neural Networks, 2005.IJCNN’05. Proceedings. 2005 IEEE International Joint Conference on,vol. 4. IEEE, 2005, pp. 2018–2023.

[9] G. Guo, Y. Fu, T. S. Huang, and C. R. Dyer, “Locally adjusted robust regression for human age estimation,” in Applications of Computer Vision, 2008. WACV 2008. IEEE Workshop on. IEEE, 2008, pp.1–6.

[10] D. S. Berry, L. Z. McArthur et al., “Perceiving character in faces:The impact of age-related craniofacial changes on social perception,”Psychological bulletin, vol. 100, no. 1, pp. 3–18, 1986.

[11] A. Stone, “The aging process of the face & techniques of rejuvenation,” http://www.aaronstonemd. com/Facial Aging Rejuvenation.shtm,2010.

[12] A. Lanitis, C. J. Taylor, and T. F. Cootes, “Toward automatic simulation of aging effects on face images,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 24, no. 4, pp. 442–455,2002

[13] Y. Fu and T. S. Huang, “Human age estimation with regression on discriminative aging manifold,” Multimedia, IEEE Transactions on,vol. 10, no. 4, pp. 578–584, 2008.

[14] X. Geng, Z.-H. Zhou, and K. Smith-Miles, “Automatic age estimation based on facial aging patterns,”Pattern Analysis and Machine Intelligence,IEEE Transactions on, vol. 29, no. 12, pp. 2234–2240, 2007.

[15] K. A. Kamp, N. Timmerman, G. Lind, J. Graybill, and I. Natowsky,“Discovering childhood: Using fingerprints to find children in the archaeological record,” American Antiquity, pp. 309–315, 1999.

[16] I. Jolliffe, Principal component analysis. Wiley Online Library, 2005.

[17] N. S. Altman, “An introduction to kernel and nearest-neighbor nonparametric regression,” The American Statistician, vol. 46, no. 3, pp.175–185, 1992.

[18] A. Majumdar and A. Bhattacharya, “A comparative study in wavelets,
curvelets and contourlets as feature sets for pattern recognition.” Int. Arab J. Inf. Technol., vol. 6, no. 1, pp. 47–51, 2009.

[20] E. Candes, L. Demanet, D. Donoho, and L. Ying, “Fast discrete curvelet transforms,” Multiscale Modeling & Simulation, vol. 5, no. 3, pp. 861–899, 2006.

[21] E. J. Candes, “Ridgelets: theory and applications,” Ph.D. dissertation