Dental Biometrics for Human Identification based on Neural Network and Image Properties in Periapical Radiographs

Payal¹ Mr. Sandeep Patil²
¹M.E. Scholar ²Associate Professor
¹,²Department of Electronics & Telecommunication Engineering
¹,²FET-SSTC, Bhilai (C.G.), India

Abstract— The Dental biometric system is used to identify deceased individual. Dental biometric system is used in forensic science. In this system AM radiograph is matched with PM radiograph to identify unidentified individual. Dental biometrics consists of four steps as: preprocessing of dental radiograph, segmentation, feature extraction and matching of AM and PM radiograph. Segmentation is a method used for feature extraction like shape and size of tooth. These features are used in matching of two radiographs and based on this matching, individuals can be identified. In this paper segmentation is used to extract single tooth and also for the dental work extraction. This paper presents matching of two radiographs based on properties and dental work. For the segmentation and matching we have developed algorithm.

Key words: Dental Radiographs, SOM Cluster, Neural Network

I. INTRODUCTION

A biometric system is used to identify individual. A biometric is a measurable physical characteristic which are reliable than a password. Many biometric systems are employed which work on the basis of image analysis. “Biometrics” is a general term used alternatively to describe a characteristic. The biometric systems are divided into two categories as: behavioural biometrics and physiological biometrics. In behavioural biometric systems a person is identified based on how he performs something. Physiological biometric characteristic are making a signature, walking, typing on a keyboard. In physiological biometric system a person is identified based on a unique characteristic of some body organ of that person. The typical examples of physiological biometric system are Iris biometry, face biometry, figure biometry, dental biometry etc.

Dental biometry used in forensic identification. This technique requires ante mortem and post-mortem radiographs. In this both radiograph are segmented and matched for identification of undefined victim. In this paper we are presenting dental radiograph segmentation and matching. In this paper we have extracted tooth and dental work as features by segmentation and used for matching.

A. Forensic Identification

Forensic identification is used for suspect identification and victim identification. Victim identification is done by physical biometrics. Dental radiograph can be used for victim identification based on dental evidences.

B. Types of Dental Radiographs

There are three types of dental radiograph (X-ray) (Figure 1). Periapical x-ray- It shows entire tooth, including crown, root and bone. Bitewing X-ray- Bitewing x-ray is taken at routine check-ups. Panoramic x-ray- It gives broader overview of entire dentition. It shows not only teeth also sinus, upper and lower jawbone.

Fig. 1: (a) Periapical Radiograph (b) Bitewing Radiograph (c) Panoramic Radiograph

A. Forensic Dental Biometry

Forensic dental biometrics used to identify unidentified victims. Automatic forensic identification utilizes dental radiograph. Biometrics can be classified in two category based on characteristic like behavioural and physical. Physical biometric represents iris, fingerprint, face recognition etc. Behavioural biometric represent voice, gait, signature and all behavioural traits of individual.

Evaluation of biometrics features requires characteristics such as universality, uniqueness, permanence, performance, collectability and acceptability. Forensic dental biometrics is used when no any other biometry is present. So it is very useful system to identify deceased human.

II. DENTAL IDENTIFICATION SYSTEM

The components of Dental identification system are:
1) Dental radiograph
2) Radiograph pre-processing
3) Feature Extraction
4) Clustering and Classification
5) Matching

In Dental identification system, dental radiographs of patients are collected for dental identification system. These radiographs firstly pre-processed for filter out unwanted background present with teeth using median filter. Then feature of radiographs are collected using GLCM that are texture features and image quality features. Clustering of
data is performed by SOM cluster and classification of data is performed using Levenberg-Marquardt neural networks. In matching stage features of query image try to match with database image if matching present then we can identify that individual. Algorithm is developed for this. Feature extracted from PM radiograph matched with each of database features and based on matching distance individual get identified.

Fig. 2: Block Diagram of Dental Identification System

Fig. 2 shows our dental identification system. By using this system we developed a dental biometric system for human identification.

III. PRE-PROCESSING AND FEATURE EXTRACTION OF RADIOGRAPH

A. Pre-processing of Image

Noise removing and resizing: here ‘salt and pepper’ noise of image is removed using 2-Dimension Median filter and resized in 256*256. Pre-processed dental radiograph is as shown in Figure 3(a).

B. Feature Extraction of Image

1) File Feature

The file feature contains the information fields like file name, file mod date, file size, format, format version, width, height, bit depth, color type, format signature, number of samples, coding method, coding process, and comment. To display information from the structure, for example coding method, we type ininfo function in the command window.

2) Texture Feature

An image texture is a set of metrics calculated in image processing designed to quantify the perceived texture of an image. Image texture gives us information about the spatial arrangement of color or intensities in an image or selected region of an image. Image textures can be artificially created or found in natural scenes captured in an image. Image textures are one way that can be used to help in segmentation or classification of images. For more accurate segmentation the most useful features are spatial frequency and an average grey level. To analyze an image texture in we are using statistical approach. Fig. 3(b) shows texture features of a radiograph image.

a) Entropy

\[ \text{Entropy} = - \sum_{i,j} g_{ij} \log g_{ij} \]

This statistic measures the disorder or complexity of an image. The entropy is large when the image is not texturally uniform and many GLCM elements have very small values. Complex textures tend to have high entropy. Entropy is strongly, but inversely correlated to energy.

b) Variance

\[ \text{Variance} = \sum_{i,j} (i - \mu)^2 g_{ij} \]

Where \( \mu \) is the mean of \( g_{ij} \).

This statistic is a measure of heterogeneity and is strongly correlated to first order statistical variable such as standard deviation. Variance increases when the gray level values differ from their mean.

c) Homogeneity

\[ \text{Homogeneity} = \sum_{i,j} \frac{1}{1 + (i - j)^2} g_{ij} \]

This statistic is also called as Inverse Difference Moment. It measures image homogeneity as it assumes larger values for smaller gray tone differences in pair elements. It is more sensitive to the presence of near diagonal elements in the GLCM. It has maximum value when all elements in the image are same. GLCM contrast and homogeneity are strongly, but inversely, correlated in terms of equivalent distribution in the pixel pairs population. It means homogeneity decreases if contrast increases while energy is kept constant.

d) Correlation

\[ \text{Correlation} = \frac{\sum_i \sum_j (i-\mu_x)(j-\mu_y) g_{ij}}{\sigma_x \sigma_y} \]

Where \( \mu_x \), \( \mu_y \), \( \sigma_x \), and \( \sigma_y \) are the means and standard deviations of \( g_x \) and \( g_y \).

The correlation feature is a measure of gray tone linear dependencies in the image. The rest of the textural features are secondary and derived from those listed above.

e) Skewness

Skewness is a measure of the asymmetry of the probability distribution of a real-valued random variable. The skewness value can be positive or negative, or even undefined. In terms of digital image processing, Darker and glossier surfaces tend to be more positively skewed than lighter and matte surfaces. Hence we can use skewness in making judgements about image surfaces.

3) Quality Feature

SSIM is used for measuring the similarity between two images. The SSIM index is a full reference metric; in other words, the measurement or prediction of image quality is based on an initial uncompressed or distortion-free image as reference. SSIM is designed to improve on traditional methods such as peak signal-to-noise ratio (PSNR) and mean squared error (MSE), which have proven to be inconsistent with human visual perception. Fig. 3(c) shows quality features of a radiograph image.

a) Mean Square Error (MSE)

MSE is cumulation square error between Original image and compressed image.

\[ \text{MSE} = \frac{1}{MN} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|f(i,j) - g(i,j)\|^2 \]

Where \( f \) – Pixel value of the original image, \( g \) – Pixel value of the compressed image, \( m \) – number of rows of the images and \( n \) – number of columns of the image and \( i \) represents the index of that row, \( n \) – number of columns of the image and \( j \) represents the index of that column.

b) Peak Signal to Noise Ratio (PSNR)

PSNR is the ration between Maximum possible values of signal power to that of the distorting noise.

\[ \text{PSNR} = 10 \cdot \log_{10} \left( \frac{255}{\sqrt{\text{MSE}}} \right) \]
supervised algorithm, although it does require more memory than other algorithms.

Levenberg-Marquardt algorithm is designed to approach second-order training speed without having to compute the Hessian matrix. When the performance function has the form of a sum of squares (as is typical in training feed forward networks), then the Hessian matrix can be approximated as

\[ H = J^T J \]  \hspace{1cm} (1)

And the gradient can be computed as

\[ g = J^T e \]  \hspace{1cm} (2)

Where J is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases, and e is a vector of network errors. The Jacobian matrix can be computed through a standard back-propagation technique that is much less complex than computing the Hessian matrix.

The Levenberg-Marquardt algorithm uses this approximation to the Hessian matrix in the following Newton-like update:

\[ x_{k+1} = x_k - [J^T J + \mu I]^{-1} J^T e \]  \hspace{1cm} (1.3)

When the scalar \( \mu \) is zero, this is just Newton's method, using the approximate Hessian matrix. When \( \mu \) is large, this becomes gradient descent with a small step size. Newton's method is faster and more accurate near an error minimum, so the aim is to shift toward Newton's method as quickly as possible. Thus, \( \mu \) is decreased after each successful step (reduction in performance function) and is increased only when a tentative step would increase the performance function. In this way, the performance function is always reduced at each iteration of the algorithm.

<table>
<thead>
<tr>
<th>1000</th>
<th>Maximum number of epochs to train</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Performance goal</td>
</tr>
<tr>
<td>6</td>
<td>Maximum validation failures</td>
</tr>
<tr>
<td>1e-7</td>
<td>Minimum performance gradient</td>
</tr>
<tr>
<td>0.001</td>
<td>Initial mu</td>
</tr>
<tr>
<td>0.1</td>
<td>mu decrease factor</td>
</tr>
<tr>
<td>10</td>
<td>mu increase factor</td>
</tr>
<tr>
<td>1e10</td>
<td>Maximum mu</td>
</tr>
<tr>
<td>25</td>
<td>Epochs between displays (NaN for no displays)</td>
</tr>
<tr>
<td>false</td>
<td>Generate command-line output</td>
</tr>
<tr>
<td>true</td>
<td>Show training GUI</td>
</tr>
<tr>
<td>inf</td>
<td>Maximum time to train in seconds</td>
</tr>
</tbody>
</table>

Table 2: Training Occurs According To Trainlm Training Parameters, Shown With Default Values

A. Self-organizing Map Cluster

Based on unsupervised learning, SOM for clustering data without knowing the class memberships of the input data provides a topology preserving mapping from the high dimensional space to map units. Map units, or neurons, usually form a two-dimensional lattice and thus the mapping high dimensional space onto a plane. The property of topology preserves the relative distance between the points that are near each other in the input space are mapped to nearby map units in the SOM that serve as a cluster analyzing tool of high-dimensional data. Also, the SOM has the capability to generalize which network can recognize or characterize inputs it has never encountered before. A new input is assimilated with the map unit it is mapped to.

1) SOM Algorithm

- Select output layer network topology
- Initialize current neighborhood distance, D(0), to a positive value
- Initialize weights from inputs to outputs to small random values
- Let \( t = 1 \)
- While computational bounds are not exceeded do
  a) Select an input sample \( i \)
  b) Compute the square of the Euclidean distance of from weight vectors \( w_j \) associated with each output node
     \[ \sum_{k=1}^{n} (i_{j,k} - w_{j,k}(t))^2 \]
  c) Select output node \( j^* \) that has weight vector with minimum value from step 2
  d) Update weights to all nodes within a topological distance given by D(t) from \( j^* \), using the weight update rule:
     \[ w_{j}(t+1) = w_{j}(t) + \eta(t) (i_j - w_{j}(t)) \quad 0 < \eta(t) \leq \eta(t-1) \leq 1 \]
  e) Increment \( t \)
End while

<table>
<thead>
<tr>
<th>1000</th>
<th>Maximum number of epochs to train</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>Epochs between displays (NaN for no displays)</td>
</tr>
<tr>
<td>false</td>
<td>Generate command-line output</td>
</tr>
<tr>
<td>true</td>
<td>Show training GUI</td>
</tr>
<tr>
<td>inf</td>
<td>Maximum time to train in seconds</td>
</tr>
</tbody>
</table>

Table 1: Training Occurs According To Trainlm Training Parameters, Shown With Default Values

B. Levenberg-Marquardt Neural Network

Trainlm is a network training function that updates weight and bias values according to Levenberg-Marquardt optimization. This is the one of fastest back-propagation algorithm and is highly recommended as a first-choice
V. MATCHING

Matching is a last stage of dental biometric system which finds out difference between two dental radiographs. Dental feature like entropy, standard deviation, correlation, skewness homogeneity and other 101 features(Intensity values) analysis is useful for shape analysis of objects. Comparison of the failure rate and time complexity between various segmentation algorithms performed before from literature survey is shown in Table 3.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Hajsaid, Nassar, Fahmy and Ammar Algorithm</th>
<th>Jain and Chen</th>
<th>Nomir and Abdel Morttaleb</th>
<th>Zhou and Abdel-Mottaleb</th>
<th>My Algorithm (Levenberg-Marquadt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Failure Rate</td>
<td>2.12%</td>
<td>2.61%</td>
<td>11.18%</td>
<td>3.47%</td>
<td>1.1%</td>
</tr>
<tr>
<td>Average Time Complexity (Sec)</td>
<td>4.614</td>
<td>5.658</td>
<td>57.251</td>
<td>7.323</td>
<td>14.3</td>
</tr>
</tbody>
</table>

Table 3: Comparison of the Failure Rate and Time Complexity between various Segmentation Algorithms

VI. CONCLUSION

Automated dental image segmentation algorithm that handles periapical dental images based on mathematical morphology. the proposed algorithm includes 1)Pre-processing of data using median filter, 2)Extraction of feature using GLCM 3)SOM clustering and 4) Classification by Neural Network(Levenberg-Marquardt algorithm) and 5) Matching to identify the human. The difficulties are image blurring, teeth interfering, image scan quality, and very low contrast between bones and teeth intensity.

The result shows 1) The Proposed algorithm is showing lowest failure rate. 2) The algorithm proposed have highest optimality 3) Time complexity is reduced.

REFERENCES


[9] Xin Li, Ayman Abaza, Diaa Eldin Nassar, and Hany Ammar, ," Fast And Accurate Segmentation of Dental X-Ray Records", Lane Department of Computer Science and Electrical Engg, Lane Department of computer science and electrical engg, Lane Department of computer science and electrical engg, Vol-2.6506-6109.


