Tooth Recognition in Dental Radiographs via Hu’s Moment Invariants

Nakintorn Pattanachai  
Department of Computer Engineering, Faculty of Engineering  
Chulalongkorn University, Bangkok 10330, Thailand

Nongluk Covavisaruch  
Department of Computer Engineering  
Faculty of Engineering  
Chulalongkorn University, Bangkok 10330, Thailand

Chanjira Sinthanayothin  
National Electronics and Computer Technology Center  
National Science and Technology Development Agency  
Thailand Science Park, Pathumthani 12120, Thailand

Abstract—This paper presents a method to recognize a tooth in dental X-ray images using Hu’s moment invariants. The proposed method starts with digitizing dental radiographic films. Next, rectangular bounding boxes of target teeth in the X-ray images are manually specified. The process continues with enhancing the X-ray image with histogram equalization. The teeth are segmented with the assistance of Otsu’s method. The next step is to calculate Hu’s moment invariants that are used as the teeth’s features. Finally, tooth recognition is done by feature matching with Euclidian distance.

Keywords—tooth recognition; dental radiographic films; dental X-ray images; Hu's moment invariants

I. INTRODUCTION

Image recognition has played an important role in various application areas over the past 2-3 decades. For example, recognition in some manufacturing industries includes monitoring printed circuit board (PCB), checking packaged drugs, checking product packaging in production processes and quality inspection, etc. In biometrics, the recognition of iris, face, fingerprint, hand geometry, etc. has been the main task for personal identification and verification. In traffic control, license plate recognition and speed control using video are also widely utilized.

Tooth image recognition can be used to classify the type of tooth, such as molar, premolar, canine and incisor. In the case of natural disasters such as tsunami, dental data can help identifying individual’s identity. It is because teeth are strong and unique for each person. Moreover, unlike other human’s organs, teeth are not dissolved after one’s death.

Generally, human has two sets of teeth. The first one is called deciduous teeth, and the second one, permanent teeth. The roots of teeth are embedded in the mandible bone or the maxillary bone and are covered by gums.

A set of permanent teeth usually consists of thirty-two teeth that can be separated into each individual tooth in four quadrants. They are the maxillary (upper) right quadrant, the maxillary left quadrant, the mandibular (lower) left quadrant and the mandibular right quadrant consecutively as shown in figure 1. Each quadrant consists of eight teeth which are two incisors, one cuspid (canine), two bicusps (premolar) and three molars. Therefore, there are sixteen teeth in the maxilla and another sixteen in the mandible, for a total of thirty-two teeth.

Figure 1 Maxillary and mandibular arches divided into 4 quadrants. [1]

Dental radiographs or X-ray films are pictures of teeth, bones and surrounding soft tissues. There are two types of dental X-rays: intraoral (meaning the X-ray film is inside the mouth) and extraoral (meaning the X-ray film is outside the mouth). Different types of intraoral X-rays include periapical X-rays, bite-wing X-rays, and occlusal X-rays. Extraoral X-rays include panoramic X-rays, tomograms, cephalometric projections, sialography, and computed tomography (CT). [2] The panoramic X-rays show the entire mouth area including sinuses, jaw joints, teeth and surrounding bone. This work deals with the last type, panoramic X-rays, and the recognition problem is scoped to molar and premolar tooth recognition.

This paper is organized as follows: section II presents some related works. Section III introduces our method for tooth recognition. Section IV explains the experiments and their results. Discussions and conclusion are presented in Section V.

II. RELATED WORKS

This section presents the researches and an image feature, namely Hu’s moment invariants, that are related to this work. Many approaches for tooth recognition or identification have been proposed by using dental radiographs based on the tooth shape, teeth’s relative positions, missing tooth detection and dental work. Dental images in most teeth recognition or identification researches are in bitewing view.

In 2003, Anil K. Jain et al. [3] used a semi-automatic method to extract shapes of teeth. They simplified the shape
extraction by manually selecting rectangular region of a tooth and the center of the crown from a bitewing image. They computed the gradient image for tooth contour and matched the teeth by using the affine transformation with euclidean distance.

In 2005, Jindan Zhouc et al. [4] also used bitewing images and separated each tooth into crown and root. They developed a method to identify missing teeth areas as well as the shapes of teeth. Adaptive segmentation was used to extract the teeth contours. The result of this method held a precision of 95% in the top five most similar images.

In 2007, Omaima Nomir et al. [5] detected teeth shapes by using adaptive thresholding segmentation. They chose to represent each tooth by a set of features based upon the forcefield energy function and Fourier descriptors in the matching process. Euclidean distance and absolute distance were utilized for matching.

Hu’s moment invariants have been extensively used as the measures for image recognition for their invariance to image translation, scaling and rotation. In 2008, Guang-Yuan Zhang et al. [6] proposed a real-time eye detection method using Support Vector Machine (SVM) with Hu’s moment invariants. They used the video frame capture to binarize and heuristic rules to screen the contour. Then used them to find the Region Of Interest (ROI). They calculated Hu’s moment invariants of ROI and used SVM model for classification. This method achieved an average successful classification rate of 92.8%.

In 2010, Ungkarn Jarujareet et al. [7] used an ‘iris-blob map’ as a new feature for iris identification. The iris texture is enhanced with the Difference of Gaussian (DoG). The iris feature was a map consisting of iris blobs’ bounding rectangles and Hu’s moment invariants of the detected blobs. The EER was reported at 4.8%.

III. TOOTH RECOGNITION ALGORITHM

This research proposes five steps for the tooth recognition algorithm as illustrated in the diagram in figure 2. The first step is image acquisition. Next is histogram equalization in order to improve the image contrast. The third step is to segment tooth by using Otsu’s method and then logical AND the result with the histogram equalized image. The fourth step is to calculate the tooth’s features, Hu’s moment invariants, and the last step is feature matching.

A. Image Acquisition

The dental radiographs used in this research are in the form of panoramic X-ray films of size 10 x 12 . The X-ray films are scanned by a film scanner (EPSON Perfection V700 Photo) in order to transform the X-rays to digital forms, and saved in bitmap file format (BMP). Next, each target tooth, molars and premolars, in the image in panoramic view is manually cropped with a rectangular window as shown in figure 3. It is noted that parts of neighbor tooth/teeth might appear in the cropped images. Each individual tooth image is assigned with a number based on FDI world dental federation notation [8] as shown in figure 4 such as ‘11’ for upper (or maxillary) right central incisor, ‘21’ for upper (or maxillary) left central incisor, and ‘48’ for lower (or mandibular) right third molar, etc. and then saved into a database.

B. Image Enhancement

Most X-rays in this research have low contrast, hence histogram equalization is chosen for our dental X-ray image enhancement to get higher contrast.

Histogram equalization [9] is a technique for adjusting image intensities to enhance contrast by increasing the dynamic range of intensity given to pixels with the most probable intensity values. For a discrete function, let \( r_k \) represent the gray levels of the image \( f \) to be enhanced, ranging \([0, L - 1] \). The number of possible intensity values is \( L \). The probability of gray levels or intensity levels is \( n_k \) in an image and the probability density function of the image is as shown in equation 1.

\[
p_k(r_k) = \frac{n_k}{n} \quad k = 0, 1, 2, ..., L - 1
\]

Where \( n_k \) is the number of pixels with intensity \( r_k \),

\[ n \] is the total number of pixels in the image,

and \( L \) is the total number of possible gray levels in the image.

Histogram equalization rearranges the pixel intensity values by using cumulative histogram. The input pixel intensity of
histogram equalization is transformed to a new intensity level. The transformation function is the product of a cumulative histogram. New intensity level needs to be within the range of the intensity level. The transformation function can be written as shown in equation 2.

\[
s_k = I^{-1}(r_k) = (L - 1) \sum_{j=0}^{r_k} p_j = (L - 1) \sum_{j=0}^{\hat{r}_j} \frac{n_j}{n}
\]

(2)

where \( r_k \) is input intensity, 
\( s_k \) is processed intensity, 
\( n_j \) is the frequency of intensity \( j \), and 
\( n \) is the total number of image pixels.

The original gray scale image is enhanced by histogram equalization. In figure 5, figures (a) and (b) illustrate an original image and its histogram, and figures (c) and (d), the result and its histogram from histogram equalization.

\[ 
\begin{align*}
\text{Figure 5. The histogram equalization of tooth image.} \\
\text{(a) Original Image} & \quad \text{(b) Histogram Equalization} \\
\text{(c) Equalized Image} & \quad \text{(d) Otsu's Thresholded Image}
\end{align*}
\]

C. Image Segmentation

In this research, the tooth segmentation is an important task in tooth recognition because it separates teeth from other areas in dental X-ray images. As the dental X-ray images from histogram equalization consist of teeth’s areas being brighter than other areas, this paper proposes to use Otsu’s thresholding method to assist in teeth segmentation.

- Otsu’s Method

Otsu’s thresholding method operates directly on gray level histogram to calculate the probability of each intensity level. The histogram to be thresholded contains two classes of pixels or bimodal histogram. With Otsu’s method, the optimal threshold is the one that minimizes the weighted within-class variance turns out to be the same as that maximizes the values between-class variances [10]. The within-class variance and maximum value of the between-class variance are calculate with equations (3)-(4) as follows:

\[
\sigma^2_w(t) = \omega_g(t) \sigma^2_g - \omega_f(t) \sigma^2_f(t)
\]

(3)

\[
\sigma^2_b(t) = \omega_g(t) \sigma^2_g - \omega_f(t) \sigma^2_f(t)
\]

(4)

where \( \omega_g(t) = \sum_{i=0}^{t-1} p(i) \) is the probability of the intensities in the background (where intensity varies from 0 to \( t-1 \)), 
\( \omega_f(t) = \sum_{i=t}^{L-1} p(i) \) is the probability of the intensities in the foreground (where intensity varies from \( t \) to \( L-1 \)), 
\( \sigma^2_g \) is the intensity variance of the background, and 
\( \sigma^2_f \) is the intensity variance of the foreground.

Then the maximum value of the between-class variance for the optimal threshold is calculated with equation (4).

\[
\sigma^2_b(t) = \omega_g(t) \sigma^2_g - \omega_f(t) \sigma^2_f(t)
\]

(5)

where \( \mu_b(t) \) is the intensity mean of the background, and 
\( \mu_f(t) \) is the intensity mean of the foreground.

The input to Otsu’s method is the histogram equalized image as shown in figure 6 (a). The output from Otsu’s method is a binary image of area(s) of teeth and non-teeth as shown in figure 6 (b).

\[ 
\begin{align*}
\text{Figure 6. Example images in the image segmentation step.} \\
\text{(a) Original Image} & \quad \text{(b) Equalized Image} & \quad \text{(c) Otsu's Thresholded Image}
\end{align*}
\]


Our segmentation step continues to discard non-tooth areas in the image. It is done by compute the bitwise AND between the equalized image and the Otsu’s thresholded image. An example of AND result is shown in figure 6 (c).

D. Image Feature Extraction

This paper proposes to extract feature by using Hu’s moment invariants [12]. Hu’s moment invariants can be classified as a shape descriptor which is used in computer vision. It is based on the theory of algebraic invariants and derives to seven invariants. The basic idea is to describe objects by a set of measurable quantities called invariants and its invariant features on image translation, scaling and rotation. A set of seven Hu’s moment invariants is as shown in equations (4) - (11):

\[
H_1 = \eta_2,0 + \eta_{0,2}
\]

(5)

\[
H_2 = (\eta_{2,0} - \eta_{0,2})^2 + 4\eta_{1,1}
\]

(6)

\[
H_3 = (\eta_{3,0} - 3\eta_{1,2})^2 + 3(\eta_{2,1} - \eta_{0,3})^2
\]

(7)

\[
H_4 = (\eta_{3,0} + \eta_{1,2})^2 + (\eta_{2,1} + \eta_{0,3})^2
\]

(8)

\[
H_5 = (\eta_{3,0} - 3\eta_{1,2})[\eta_{2,0} + \eta_{1,2}][\eta_{2,0} + \eta_{2,1} - 3\eta_{2,1} + \eta_{0,3}]^2 + (3\eta_{2,1} - \eta_{0,3})[\eta_{2,1} + \eta_{0,3}][\eta_{2,0} + \eta_{1,2} + (\eta_{2,1} + \eta_{0,3})^2]
\]

(9)

\[
H_6 = \eta_{0,0} + \eta_{2,2} + \eta_{2,0} - \eta_{0,0} + \eta_{2,2} - \eta_{0,0} - \eta_{2,2} + 4\eta_{1,1}\eta_{0,0} + \eta_{2,1} + \eta_{0,3}
\]

(10)

\[
H_7 = (3\eta_{2,1} - \eta_{0,3})[\eta_{2,0} + \eta_{1,2} + \eta_{2,0} + \eta_{1,2} + \eta_{2,1} + \eta_{0,3}]
\]

(11)

where \( \eta_{pq} \) is a normalized central moment of order \( p+q \), and \( \{H_1, ..., H_7\} \) is a set of Hu’s moment invariants.

E. Feature Matching using Euclidean Distance

The Euclidean distance or Euclidean metric is a distance between two points. Deriving the Euclidean distance between
two data sets involves computing the square root of the sum of the squares of the differences between corresponding values. Equation 12 is used to calculate Euclidean distance.

\[ d_{(x,y)} = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \]  

(12)

Where \( x_i \) is data \( i \) to \( n \) of an instance \( (x) \), and \( y_i \) is the corresponding data \( i \) to \( n \) of another instance \( (y) \). Therefore, \( d_{(x,y)} \) is the Euclidean distance between \( x \) and \( y \).

IV. EXPERIMENTS AND RESULTS

A. Dental Radiographs Data

Our test dental data are scanned from the panoramic films from one male and four female patients. Each patient provides two films, one before and one after orthodontic treatment, hence a total of ten films. Five films are used as references, two films, one before and one after orthodontic treatment, from one male and four female patients. Each patient provides images and test images.

The images used in our experiments consist of thirty-eight tooth image pairs. A pair of tooth refers to the same corresponding teeth of the same patient in two different films. Only the molar and premolar are manually cropped from each panoramic image. Thirty-eight tooth images are used as ‘references’ and the corresponding thirty-eight tooth images are used as ‘test images’.

B. Matching Algorithm

In the feature matching process, Hu’s moment invariants of the test image are matched with those of references using Euclidean distance. The best matched has the minimum matching distance.

Table 1 illustrates the data from our experiments of thirty-eight test images; twenty-one images were correctly retrieved from rank 1 to rank 10. Four images were ranked first, three were ranked second, three were ranked third, two were ranked fourth, three were ranked fifth, one were ranked sixth, one were ranked seventh, two were ranked eighth, one were ranked ninth and one were ranked tenth.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Rank</th>
<th>Rank</th>
<th>Rank</th>
<th>Rank</th>
<th>Rank</th>
<th>Rank</th>
<th>Rank</th>
<th>Rank</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
</tbody>
</table>

IV. CONCLUSION AND DISCUSSION

This paper presents tooth recognition in dental radiographs using Hu’s moment invariants technique. The database archives the references images by manually cropped the dental radiographs, numbering teeth, teeth segmentation and image feature extraction. For testing data, the panoramic image segmentation technique is described to extract image features. The recognition is based on the features extracted from moment invariants of dental radiographs. For tooth matching, the test image can be searched in the database for the minimum distance between the feature vectors of references images and test image. The experimental results of tooth recognition by sum of the minimum difference of Hu’s moment invariants are presented. However, some images are difficult to apply using the proposed method because the images are very poor quality and blurred. For example of the images which the matching approach fail. In figure 7, figures (a) and (c) is the same tooth with difference contrast, and figures (b) and (d), the image segmented.

ACKNOWLEDGMENT

The authors would like to thank the Faculty of Dentistry, Chulalongkorn University and Fun-Dee clinic, Bangkok, Thailand for their great assistance in our research.

REFERENCES


[8] FDI World Dental Federation , Tour De Cointrin, Avenue Louis Casai 84, Case Postale 3, 1216 Cointrin-Geneve, Switzerland


