Determination age and gender with developed a novel algorithm in image
processing techniques by implementing to dental X-ray images

Emre Avuçlu1*, Fatih Başçıfıcı2

Abstract: Age determination process is very important in legal medicine and forensic sciences. After a massive disaster or a legal situation, identification may be required. Age determination process is requested from the forensic medicine institute for live or dead body. In such cases, age determination process must be done correctly. In this study, a database was created with a total of 1313 panoramic dental images to determine age and gender. The images in this database are stored in different 6 folders (M1, M2, M3, M4, M5, M6) by preprocessing with image processing techniques. The calculation of the morphological properties of teeth in this folder (height, area, perimeter, similarity, center of gravity, radius of teeth) are stored separately in XML files (XML1, XML2, XML3, XML4, XML5, XML6). In order to obtain better results in determining age and gender, a novel and originally developed algorithm has been applied to the trapezoidal images in the database. Morphological measurement comparisons were made by testing the teeth with height, area, perimeter, radius, similarity ratio, center of gravity information separately. Thanks to the new and originally developed algorithms, the age determination process was performed with an error of +-1 years in general. In addition to the age determination process, gender determination was also performed. In this study, a multi-disciplinary study was performed to determine the age and gender from dental x-ray images.

Key Words: Age and gender determination, legal medicine, forensic sciences, image processing techniques, panoramic dental x-ray images.

INTRODUCTION

One of the most important elements in the identification of forensic medicine is the age determination process for the person. All the characteristics of a person’s identity distinguish that person from other people. Age determination is done according to specific characteristics of the person; gender, height, body weight, hair, skin, eye color, fingerprint, bone and teeth [1]. In our daily lives, for many reasons (earthquake, flood, fire etc.), identification of a living or non-living body may be required [2]. Nowadays, the process of identification for people has gained an international status [3]. Determination of age in the forensic process is a very important issue in terms of criminal and civil law. Therefore, identification based on anatomical features and changes occurring during the life of the organism should be done with the least minimum mistakes [4, 5].

After some events (disaster, suicide, murder, etc.), the person’s body may not be recognized and his physical characteristics may not give any information for that person. In such cases, the only structure for identification is the human tooth. Teeth are not affected more than physical factors. Teeth give better results than other organs. For estimating age, in dental practice and in forensic science, atlases are used in which the developmental stages of teeth are depicted. This method is used in x-ray images taken from teeth found both in living individuals and in inanimate bodies [6, 7]. It has also been shown that information from teeth due to low metabolism give more accurate results than other organs [8,9]. After accidents, wars and natural disasters, is an important event.

1) Aksaray University, Department of Computer Technology and Computer Programming, Aksaray, Turkey
* Corresponding author: Tel: +90(382) 382 288 2025, E-mail: emreavuclu@aksaray.edu.tr
2) Selçuk University, Technology Faculty, Department of Computer Engineering, Selçuklu, Konya, Turkey
to determine the identity of the dead body. Race, gender, age, physical features are the identification parameters that are prioritized by science [10,11]. In the literature they performed different segmentation and identification processes on dental x-ray images [12,13]. In their study, they performed age estimation using dental x-ray images using different methods [14,15]. Age estimation was performed by measuring different areas [16], in young and children [17,18] in different areas, and [19,20] different areas depending on the skeletal size [21-24].

Firstly, the necessary information for the study was collected from specialist dentists and forensic scientists. In this study, image processing techniques were applied to tooth images obtained from panoramic dental X-ray images. Along with these techniques, a new algorithm has been developed to obtain more accurate and standardized results in age prediction. This algorithm is used to correct trapezoidal tooth images. With this application, the age determination in the forensic unit will be done automatically. In addition to the age determination process, gender determination was also performed.

**MATERIALS AND METHODS**

In this study, datasets were created first. All images were obtained from a single device at the same distance. So the resolution of the images is the same. Dental images consist of all persons aged between 4 and 63 years. Dental panoramic X-ray images used to create a dataset were obtained from different cities in Turkey. Thus, the dental database has been enriched in terms of diversity. In the universal numbering system, the numbering of the teeth starts by giving the number 1 in the right uppermost rear tooth (3.molar). The numbering process goes through the arcs and the left uppermost rear tooth (3rd molar) becomes 16. Images obtained by extracting boundary lines from panoramic dental X-ray images were recorded in the database with AGE_GENDER_TEETHNUMBER_COUNT information.

Pre-process operations are applied on the teeth for better results. As in the general methodology shown in Figure 1 below, the pre-process methods are applied first.

For example, the second tooth number 18 of a 23-year-old male is recorded as: 23_M_18_2. A dental name of someone who is a female gender is kept as F instead of M.

**Pre-Process Operations**

The input image is first converted to grayscale by grayscaling. Contrast stretching is done to improve the contrast of poorly contrasted pictures. The LevelsLinear filter performs linear correction of the RGB channels by matching the input ranges of the specified channels to the output ranges. ForegroundEnhance method has been developed to improve the success rate of the frontal strengthening method to ensure that the image density of the teeth are optimized. In this method, the processing sequence is as follows;

\[ \text{avg} = \frac{\sum_{i=0}^{255} H_i i}{\sum_{i=0}^{255} H_i} \]  

\[ T = \frac{\sum_{i=\text{avg}}^{255} H_i i}{\sum_{i=\text{avg}}^{255} H_i} \]  

For each pixel in the image, the comparison is made according to the \( T \) threshold value. Thus, the images of the teeth become more apparent.

\[ F[x,y] = \begin{cases} 0 & \text{if } F[x,y] > T \\ F[x,y] & \text{otherwise} \end{cases} \]  

![Figure 1. General methodology.](image)

![Figure 2. Zooming and freehand drawing.](image)
At the same time, as a result of the color reduction process, an unwanted loud noise may appear on the new image. To remove this, use the Median filter formula in Equation 4:

\[
F[x, y] = \text{median}[g[p, q]] \tag{4}
\]

Where \(g[p, q]\) refers to the convolution kernel. Canny edge detection operator is applied through 4 operations.

Typically, the Canny operator smoothens images with Gaussian filtering. The Gaussian function is calculated according to Equation 5 below.

\[
G(x, y) = \frac{1}{(2\pi\sigma^2)} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \tag{5}
\]

The following Equation 6 is applied on the smoothed picture.

\[
f_{\text{xy}}(x, y) = G(x, y) * f(x, y) \tag{6}
\]

Where \(\sigma\) is the Gaussian filtering parameters. \(\sigma\) controls the degree of noise emission.

Degrade size and direction are calculated: The partial derivative of the corrected image is calculated according to the first partial derivative of Equation 7-a and 7-b using the infinite difference of the \(x\) and \(y\) directions in 2x2 neighborhoods.

\[
p_x(x, y) = \frac{f_{\text{xy}}(x+1, y) - f_{\text{xy}}(x, y) + f_{\text{xy}}(x+1, y+1) - f_{\text{xy}}(x, y+1)}{2} \tag{7-a}
\]

\[
p_y(x, y) = \frac{f_{\text{xy}}(x, y+1) - f_{\text{xy}}(x, y) + f_{\text{xy}}(x+1, y+1) - f_{\text{xy}}(x+1, y)}{2} \tag{7-b}
\]

After calculating the partial derivative of the directions of \(x\) and \(y\), the direction of the slope \(\theta\) and the magnitude of the \(G\) slope is calculated according to Equation 8 and Equation 9.

\[
G(x, y) = \sqrt{(p_x(x, y))^2 + (p_y(x, y))^2} \tag{8}
\]

\[
\theta(x, y) = \arctan\left(\frac{p_x(x, y)}{p_y(x, y)}\right) \tag{9}
\]

Non-maxima suppression is applied to the gradient magnitude: If the gradient magnitude of \(G(x, y)\) at point \((x, y)\) is not greater than the two adjacent interpolations in direction of \(\theta(x, y)\), it is marked as non-edge point otherwise it is marked as edge point.

The double threshold (Low and High) algorithm is applied to detect and connect the edges [25].

Sobel edge detector is based on 1st partial derivatives. In the Sobel algorithm, two convolution kernels are used which are shown in the Table 1. One of these kernels finds horizontal edges and the other finds vertical edges [26].

\[
\sum_{m=0}^{n} \sum_{k=0}^{K} G_i = \sum_{m=0}^{n} \sum_{k=0}^{K} G_{y(i,k)} \tag{10}
\]

Where \(i, j\) is the pixel coordinate \(m\) and \(n\) is the matrix value. \(G(x)\) and \(G(y)\) are the values obtained from that convolution.

### Developed method for rotation processing

Teeth have been automatically corrected with the algorithm developed to ensure that the measurements to be made over the teeth are more accurate and standard. In Figure 3 below, krone base, root base and root starting point, distance between tooth roots (\(rL - rR\)) can be determined automatically.

![Figure 3. Tooth rotation.](image)

In Figure 3, the tooth image is rotated step by step through the cycle from 1 to 90 degrees from the \(x\) origin pointed to by green. At each degree step, the image goes to the calcPosValue() function and the average \(Y\) values of the points at the top of the tooth image is summed. The smaller this value is, either the lower region or the upper region (result / = topEdge.Count, average \(Y\) values). Since the tooth may be trapezoidal at most 90 degrees, it can not be a sub-region, so it is the upper region. At the same time, if the number of upper sloped points is greater, then (result + = topEdge.Count total number of points) is added to affect the result value. Algorithm-1 checks whether the tooth is in the correct position (perpendicular).

Algorithm-2 is a different sub-function in Algorithm-1. In each degree rotation step of the tooth, both the distance of the blue arrows descending from the \(Y\) axis in the upper surface and the distance of the white

---

### Table 1. Convolutions

<table>
<thead>
<tr>
<th>(Z_1)</th>
<th>(Z_2)</th>
<th>(Z_3)</th>
<th>(-1)</th>
<th>(-2)</th>
<th>(-1)</th>
<th>(-1)</th>
<th>(0)</th>
<th>(1)</th>
<th>(-1)</th>
<th>(0)</th>
<th>(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Z_4)</td>
<td>(Z_5)</td>
<td>(Z_6)</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
<td>(-1)</td>
<td>(-1)</td>
<td>(1)</td>
<td>(-1)</td>
<td>(0)</td>
<td>(1)</td>
</tr>
<tr>
<td>(Z_7)</td>
<td>(Z_8)</td>
<td>(Z_9)</td>
<td>(1)</td>
<td>(2)</td>
<td>(1)</td>
<td>(-1)</td>
<td>(0)</td>
<td>(1)</td>
<td>(-1)</td>
<td>(0)</td>
<td>(1)</td>
</tr>
</tbody>
</table>

\(G_x = (Z_7 + 2Z_8 + Z_9) - (Z_1 + 2Z_2 + Z_3)\)

\(G_y = (Z_3 + 2Z_6 + Z_9) - (Z_1 + 2Z_4 + Z_7)\)
The total number of white pixels in the object in the calculation process according to the area gives us the area. The area finding process starts from the point (0,0) of the scene, until the last white spot on the picture. In Figure 5 (a), the field of sample given in the illustration is 22 pixel units.

In the calculation according to the perimeter, the Canny edge extraction filter is applied to the image to find the sum of the edge dots. For the enlarged tooth shown in Figure 5 (b), Canny edge detection algorithm is applied. The sum of the edges is summed with the red arrows shown as an example. Thus, the perimeter of the tooth is calculated.

The center of gravity is located in the calculation process according to the center of gravity. The result is a point \( P(x, y) \). If there is more than one center of gravity, the Polygon center of gravity is calculated (some of the tooth images can be several pieces of thresholding result). Such teeth are more common in children. For each object in images (connected component), the center of gravity is calculated separately. This gives us more accurate results. Since the coordinate values of all the objects in this image are known, the following equation 11(a)-11(b) can be calculated by the function shown as F. In Figure 6, a sample with 11 corner points shows the polygon (the anomalous example is shown here) and the gravity center is calculated as \( G \).

The G in Figure 6 is the center of gravity and final gravity. For general polygons, the formula shown in equation (11) is used in the calculation of the center of gravity of an 11-point object to find the x and y components (\( y > = 0 \) and the center number of the \( n \) objects):

\[
F_x = \frac{\sum_{t=1}^{n} (x_{t+1} - x_t) (y_{t+1} + y_t)}{11A} \quad \text{(11-a)}
\]

\[
F_y = \frac{\sum_{t=1}^{n} (x_{t+1} - x_t) (y_{t+1} - y_t)}{11A} \quad \text{(11-b)}
\]
Where $A$ is the number of objects and $(x, y)$ is the coordinates of the center of gravity of the $i$th object.

The Tanimoto similarity function ($T_s$) is applied in the computation according to Similarity Rate. The bits in the same coordinate for the two pictures are processed by the logical “and” and “or” operators. In the two images named $A$ and $B$, show the bits of both images $A_i, B_i$. Accordingly, the Tanimoto similarity can be written as in the following equation 12-a or 12-b:

$$T_s (A,B) = \text{cmpAND} \sum_i (A_i = B_i \text{ and } A_i = 255), \text{ cmpORsum} (A_i = 255 \text{ or } B_i = 255),$$

Then $T_s (A,B) = \text{cmpAND} / \text{cmpOR}$ \hspace{1cm} (12-a)

$$T_s (A,B) = \sum_i A_i \land B_i / \sum_i A_i \lor B_i \hspace{1cm} (12-b)$$

In the radius calculation process, the object is drawn with rectangle as in Figure 7. Then the center of gravity of the tooth image is found. Finally, the radius is obtained by drawing a line from the center of gravity to the rectangle’s X-min region (last white pixel).

**Mirror Tooth Finding**

In dentistry, a tooth exhibits morphological (height, area, perimeter, etc.) characteristics similar to the mirror itself. In Figure 8 (a), the 22nd tooth mirror is 27th tooth (for adults). If a mirror tooth image is loaded into the developed application, it is processed on the nearest tooth with the same property as that tooth. Figure 8. (b) shows tooth mirrors for children.

**RESULTS**

The black lines in Figure 9 show the distance each tooth is to the tested tooth. Figure 9 show the distance to the test tooth according to the height of the teeth. The blue straight line is the measurement result for the tooth to be tested. This line is determined by the nearest tooth, the red rounded area, and the label is shown on the graph with the name.

Figure 10 shows the distance chart according to the similarity ratio of a different tooth image.

As can be seen from Figure 9 and 10, the most definite result is obtained by comparing according to the similarity ratio. The blue line is only seen in close proximity to their age group teeth. The other 1309 teeth are too far for the searched tooth.

**CONCLUSION**

In this study, the age estimation was carried out automatically in computer environment instead of tooth atlas used for age estimation in legal medicine and forensic science. A multi-disciplinary study was conducted to determine age and gender from dental x-ray images. In fact, in this work a dynamic expert system based process is carried out. The database is rebuilt by experts and the applied comparing methods dynamically respond according to the created database. The database of 1313 tooth images obtained from panoramic images was created manually. Image enhancement was performed...
by applying 6 methods firstly to the images. The dental image to be tested both age and gender determination compared to 6 different methods (M1, M2, M3, M4, M5, M6) and 6 different morphologically comparing (Height, area, perimeter, center of gravity, similarity, radius). In order to get better results, a new and original algorithm has been applied to trapezoid images. Correcting these images resulted in more accurate and standardized results. The results obtained are compared with other age estimation studies in the literature and are shown in Table 2.

Looking at Table 2, it can be seen that the best results are obtained with Comparing by perimeter (M4), Comparing by Gravity (M1, M2) and Comparing by Radius (M4). It should not be forgotten that the results obtained by changing the test image will be better in different methods. The number of methods and comparing areas have been diversified to achieve more accurate results.

In the comparison by perimeter and area option, images existing in the database are grouped according to their number. For example, if the comparison by area is selected (18_F_19_6 is the number of the last number), a result is obtained when the total area of these teeth are divided by 6. This result is compared with the area of the test image input from the outside. The results obtained are given in Table 3.

Looking at Table 3, age estimation is generally more precise in the average calculation process. The

![Figure 9. Results by height.](image)

![Figure 10. Results by similarities.](image)

Table 2. Comparison with other methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Error (SE:Standard Error,Y:Year)</th>
<th>Count of teeth used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kvaal Method [27]</td>
<td>±9.8 SE, ± 0.5-2.5 Y</td>
<td>100</td>
</tr>
<tr>
<td>Cameriere Method [28]</td>
<td>± 5 SE</td>
<td>100</td>
</tr>
<tr>
<td>Yang et al. Method [29]</td>
<td>± 8.3 SE</td>
<td>28</td>
</tr>
</tbody>
</table>

Proposed Method Results (RG:Real Gender EG: Estimation Gender, Y:Year, E:Error)

<table>
<thead>
<tr>
<th>Comparing by</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
<th>M5</th>
<th>M6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height</td>
<td>F</td>
<td>F</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Area</td>
<td>F</td>
<td>F</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Perimeter</td>
<td>F</td>
<td>F</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Gravity</td>
<td>F</td>
<td>M</td>
<td>M</td>
<td>F</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>Similarity</td>
<td>F</td>
<td>M</td>
<td>F</td>
<td>M</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>Radius</td>
<td>M</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>M</td>
<td>M</td>
</tr>
</tbody>
</table>

Table 3. Calculation by average values

| Results based on average calculation (RG:Real Gender EG: Estimation Gender, Y:Year, E:Error) |
|---------------------------------|----------------|----------------|----------------|----------------|----------------|----------------|
| Comparing by                    | M1 | M2 | M3 | M4 | M5 | M6 |
| Area                            | M  | F  | M  | F  | M  | M  |
| Perimeter                       | M  | F  | M  | F  | M  | M  |

417
Determination age and gender with developed a novel algorithm in image processing techniques

Conflict of interest. The authors declare that there is no conflict of interest.

Acknowledgements. This work is supported by the Selçuk University Scientific Research Projects Coordinatorship /Konya, Turkey.

References

13. Dincen İ. Adli tupta yaş tayininin dişlerinin muayenesi ile elle edilen bilgilerin değerlendirilmesi, birbirine tezi, Ege University Faculty of Medicine Department of Forensic Medicine 2015, İzmir.