Automatic Localization of Cephalometric Landmarks

V. Grau,*† M. Alcañiz,*‡ M. C. Juan,*§ C. Monserrat,*§ and C. Knoll*

*MedICLab, †Departamento de Ingeniería Electrónica, ‡Departamento de Expresión Gráfica en la Ingeniería (DEGI), and §Departamento de Sistemas Informáticos y Computación (DSIC), Universidad Politécnica de Valencia, 46022 Valencia, Spain

Received November 23, 2000; published online September 20, 2001

A system for automatic detection of cephalometric landmarks is presented. Landmark detection is carried out in two steps: a line detection module searches for significant, well-contrasted lines of the image, such as the jaw line or the nasal spine. The landmark detection module uses the lines located in the first module to determine the search areas and then applies a pattern detection algorithm, based on mathematical morphology techniques. Relations between landmarks and lines are determined by means of a training process. The system has been tested for the detection of 17 landmarks on 20 images: more than 90% of the landmarks are accurately identified.

Key Words: cephalometry; X-ray imaging; pattern matching; mathematical morphology.

1. INTRODUCTION

1.1. Cephalometry

The name cephalometry is given to the morphological study of all the structures present in a human head. This study is carried out through the measurement of distances and angles between significant structures. Comparison of these measurements with the standards accepted for the same age, race, and sex is a valuable tool for diagnosis, treatment monitoring, and prediction of orthodontic treatment results.

It was d’Arcy Thomson, in 1917 [1], who for the first time analyzed the growth of the head using anatomical landmarks. These landmarks were the nodes of a grid, so changes in their position resulted in deformations of the grid. His method, however, was more ambitious and was applied not only to quantify the effects of growth, but also to relate different individuals and even different species.

Broadbent [2] and later Brodie [3] applied a method based also in landmarks to quantify malocclusions and study their effects. Landmarks were defined using cephalometric radiographs, where some structures, in bony or soft tissue, had to be identified. The power of this approach comes when normal (standard) values are known for a specified measurement, at an individual of known age, race, and sex, so differences can be quantified and used for diagnosis.

In 1948, Downs [4] introduced the first cephalometric analysis method. Downs selected 10 angular measurements on lateral radiographs from a group of selected individuals, taking average values and giving them a clinical significance. Downs’ has been the basis for most methods used at present, such as the ones by Ricketts [5] and Steiner [6]. In Fig. 1, landmarks used in the Ricketts analysis are shown. Lines

1. INTRODUCTION

1.1. Cephalometry

The name cephalometry is given to the morphological study of all the structures present in a human head. This study is carried out through the measurement of distances and angles between significant structures. Comparison of these measurements with the standards accepted for the same age, race, and sex is a valuable tool for diagnosis, treatment monitoring, and prediction of orthodontic treatment results.

It was d’Arcy Thomson, in 1917 [1], who for the first time analyzed the growth of the head using anatomical landmarks. These landmarks were the nodes of a grid, so changes in their position resulted in deformations of the grid. His method, however, was more ambitious and was applied not only to quantify the effects of growth, but also to relate different individuals and even different species.

Broadbent [2] and later Brodie [3] applied a method based also in landmarks to quantify malocclusions and study their effects. Landmarks were defined using cephalometric radiographs, where some structures, in bony or soft tissue, had to be identified. The power of this approach comes when normal (standard) values are known for a specified measurement, at an individual of known age, race, and sex, so differences can be quantified and used for diagnosis.

In 1948, Downs [4] introduced the first cephalometric analysis method. Downs selected 10 angular measurements on lateral radiographs from a group of selected individuals, taking average values and giving them a clinical significance. Downs’ has been the basis for most methods used at present, such as the ones by Ricketts [5] and Steiner [6]. In Fig. 1, landmarks used in the Ricketts analysis are shown. Lines
Manual location of landmarks is a tedious and time-consuming task, which in the case of experienced orthodontists can take approximately 10–15 min. Several researchers have studied the problem of automatic landmark location. Two different approaches can be found in the literature: knowledge-based feature detection and pattern matching.

The first attempt to develop an automatic landmark detection system using knowledge-based operators was carried out by Lévy-Mandel et al. [9]. Their system consists of three stages: a prefiltering one, in which a median filter, an image equalization, and a final sharpening filter are applied to minimize noise and enhance the image; followed by an edge detection using the Meron-Vassy operator; and finally a line-following algorithm guided by a priori knowledge, introduced in the system by means of simple ad hoc criteria. In this case, only the significant lines were detected, not the exact position of the landmarks. Based on the work by Lévy-Mandel, Parthasaraty et al. [10] developed a more complete system, using a similar sequence of operators, in which the most relevant improvement is the inclusion of a resolution pyramid to reduce processing time. In this case, not only the lines, but also the landmarks were located, by using similar ad hoc criteria. Yan et al. [11] also improved the previous work [9] by continuing the cephalometric analysis even if any particular feature could not be found in a particular image. More sophisticated attempts have been carried out recently: Davis and Forsyth [12] presented a knowledge-based system, based on a blackboard architecture.

The solutions given previously [9–11] have the same flaw: due to the rigidity of the knowledge-based rules used, they are highly dependent on the quality of the input images. The approximation in Ref. [12], being more flexible, allows improvement of the results of previous systems. However, still in this case the structures are found directly by edge detection, which makes the system highly sensitive to poor quality of the image in the vicinity of landmarks. Another
The common disadvantage of knowledge-based detection systems is the difficulty for the user to add new landmarks to the calculation, for which new rules should be formulated.

The second approach to landmark detection is based on the use of pattern-matching techniques. Cardillo and Sid-Ahmed [13] used a pattern-matching algorithm, based on mathematical morphology, to locate directly the landmarks on the original images. To make detection more robust and accurate, the system searched for several structures, assured to maintain an almost exact geometrical relation to the landmark. However, these structures can only be determined for a small number of landmarks. Lately, Rudolph [14] developed a similar system, based on a different pattern-matching technique, spatial spectroscopy.

The main disadvantage of the systems that apply only a pattern detection step is that false detections can arise far from the expected landmark location, dramatically reducing the system accuracy, as was verified in our preliminary experiments.

Despite the efforts oriented in this direction, we can say that at present automatic landmarking methods are not used in clinical routine, due to the high failure rate and dependence on the quality of the images of existing systems. Location of landmarks is done currently on the image, using digitizer tablets, or on the computer on the digitized image, but always by hand [15].

In this paper, we present an automatic landmark detection system, which tries to combine the two mentioned approaches in the following way: a simple knowledge-based, edge detection step is first carried out to determine approximately the location of significant structures of the image. Only a few clear structures, which can be found easily even in poor quality images, are detected. The result is then used to define a small search area for each landmark, where a pattern-matching detector is used to determine its exact location. The use of search areas provides an important advantage: similar, but distant, structures do not affect the result, as they are not included in the search area, so the results are significantly improved. Furthermore, computation time required for the pattern-matching process is reduced because only a small fraction of the image is considered for each landmark.

2. SYSTEM OUTLINE

Previous works, as mentioned in the introduction, show two different possibilities for landmark location: edge detection techniques followed by identification of singular points on the edges, or pattern-matching techniques which detect directly the structures where landmarks are located. Both approaches have disadvantages: edge detection depends strongly on image quality; moreover, not all the landmarks are located directly on significant edges of the image; on the other hand, pattern matching is a slow technique and has the possibility of giving false alarms, wrong detection of structures similar to landmarks, but located on distant positions. To minimize these disadvantages, we have taken the two approaches and structured our system in two modules, which are applied sequentially to the digitized images:

- Line detection module. Using edge detection techniques, the system detects in the image a set of well-defined lines. These lines are identified depending on their location, length, and direction. A precise detection of the lines allows us to construct an initial structure of the images, which determines the zones in which each landmark will be searched.
- Point detection module. A pattern detection algorithm is used to detect the exact position of landmarks. To reduce processing time and eliminate the risk of false alarms, pattern matching is applied only in the zones determined by the line detection module, where it is possible that the landmark can be found.

3. IMAGE SCANNING

Because digital X-rays are not yet the generalized image-capturing technique for cephalometric images, our system must deal with digitized images from film radiographs. Due to the great variety of scanners that are currently available, it is impossible to provide a set of scanning parameters that assure the best quality for the images. On the other hand, it is necessary that the scanning resolution always be the same because the pattern-matching algorithm works by comparison of the image with a set of models corresponding to different landmarks. We have established a resolution of 33 pixels/cm, which provides a pixel size of approximately 0.3 × 0.3 mm, sufficient for accurate placement of the landmarks, and an image size usually about 500 × 500 pixels.

4. LINE DETECTION MODULE

4.1. Edge Detection

Several systems exist to detect contour lines in the image. Usually, they require smoothing and differentiation of the
Comparisons among them have provided the conclusion that no single edge detector is always the best: it depends on the image and the application [17]. One of the most popular approaches is detection of zero crossings of the Laplacian of Gaussian (LoG) [18], which relies on the property that edges always correspond to sign changes of the second derivative.

The LoG algorithm has an important advantage compared to other common techniques: closed edges are always obtained, even with the application of a wide Gaussian filter to minimize image noise. In this case, however, strong edge displacements can be expected. This drawback is not important in our case, in which we need the detected contours only to obtain an approximate search zone for landmarks, not to calculate their exact locations.

The width of the Gaussian must be carefully adjusted to eliminate only noise, not significant contours. A universal value cannot be given for every situation because it depends on the amount of noise present in the image, which depends on the image acquisition hardware. However, in our tests, we have found that values of sigma of around 2 mm perform well in most situations. In Fig. 3, we present the output of the edge detection step.

4.2. Contour Segmentation

As we stated in Section 4.1, and can be seen in Fig. 3, the result of zero crossings detection of the LoG is a set of closed contours, among which we can always find the reference lines we are looking for. However, these lines are usually connected to other structures of the image, so they must be separated from them before we can calculate their position and orientation. Segmentation of the contours is carried out to obtain a polygonal approximation of the curves, with the following algorithm [19]:

Given a contour \( P_i(x_i, y_i), i = 0 \ldots n \)

1. Algorithm starts from point \( i = 2 \).
2. Coordinate origin is moved to the initial point of the contour \( P_0 \).
3. We initialize \( f_i = 0 \).
4. Increments \( \Delta x_i = x_i - x_{i-1}, \Delta y_i = y_i - y_{i-1} \) are calculated.
5. \( f_i \) is updated according to the expression \( f_i = f_{i-1} + \Delta f_i \), where \( \Delta f_i = x_i \cdot \Delta y_i - y_i \cdot \Delta x_i \).
6. Length of the current segment \( L_i \) from the origin to the current point \((x_i, y_i)\) is calculated as \( L_i = (x_i^2 + y_i^2)^{1/2} \).
7. Condition \( |f_i| \leq T \cdot L_i \) is tested:
   - If true, \( i = i + 1 \); return to Step 4.
   - If false, a segment has been completed. This point is the initial point of the rest of the contour, return to Step 1.

The basis of the algorithm is that, in Step 5, \( \Delta f_i = x_i \cdot \Delta y_i - y_i \cdot \Delta x_i \) is calculated, which corresponds to the product \( |L_i \times \Delta f_i| = L_i \cdot d_i \), \( L_i \) is the length of the segment joining \( P_i \) and \( P_{i+1} \), and \( d_i \) is the distance from \( P_{i-1} \) to this segment. These quantities are shown in Fig. 4, where the area corresponding to \( \Delta f_i \) can be seen. Obviously, high deviations of the curve from a straight line will produce high values of \( \Delta f_i \), so points where \( \Delta f_i \) is high are good candidates for curve segmentation. Values of \( \Delta f_i \) are accumulated in the

FIG. 3. Results of the edge detection step.
not appear in anatomical structures, so in the real case the
behavior of the algorithm is completely satisfactory. We
have used a $T$ value of 3 mm, determined heuristically.

4.3. Segment Joining

After contour segmentation, we obtain from the original
image a set of straight lines, which correspond approximately
to significant structures in the head of the patient. However,
a problem has still to be solved: due to Gaussian filtering,
it is possible that a contour wrongly joins to an adjacent
one, as can be seen in Fig. 5. Contour segmentation solves
the problem of wrong joints, breaking the contours, but then
the new segments generated should be tested for joining.

For this problem, a special submodule has been incorpo-
rated to the system, which merges segments attending to
their location and direction. All possible segment pairs
$(s_1, s_2)$ are tested for their possible merging: the two nearest
extrema of $s_1$ and $s_2$ ($e_1, e_2$) are taken, as shown in Fig. 6.
Then the angles of the two segments at $e_1$ and $e_2$ are calcu-
lated, and the lines are prolonged, with the same angle, to
calculate the distances $d_1$ and $d_2$ between each extremum
and the prolongation of the segments. If each one of the two
distances is smaller than a threshold, then the contours are
joined, using a simple straight line between their ends. The
threshold value we have used is 3 mm, to be coherent with
the contour segmentation step.

Although the quality of the joined contour could be im-
proved using a finer approximation, rather than a straight
line, we must remember that the purpose of the line detection module is only to obtain a coarse modeling of the head structure. This approximation has proven to be sufficient for our needs.

4.4. Reference Line Detection

After segment joining, the system can detect the ones that correspond to the reference lines that are going to form the basis for landmark detection. In our case, we have chosen four lines that are visible in every image, always have a good contrast and are sufficient to provide a sketch for the structure of the patient’s head. These four lines are:

- The lower limit of the jaw
- The frontal bone contour
- The nasal spine
- The diagonal line that marks the endocranial part of the frontal bone, between Sella and the upper part of the orbit.

These contours have been marked on a radiograph in Fig. 7. As can be seen, all of them are approximately straight lines.

All segments extracted from the image must be tested to detect whether they correspond to one of the four special contours. Two characteristics are used for the test: position and angle (position of the center of the contour, and angle of the straight line that joins the extrema of the contour).

Position of the candidate line is tested by means of a search zone, a rectangle that the searched line must cross. To define the search zones, the jaw line is first detected using as search zone all the lower half of the image. Because the jaw is generally the most contrasting structure of the image, this detection is almost always correct. Then, the rest of the search zones are determined with respect to the jaw position. In this way, we obtain a search algorithm that is quite independent from the particular characteristics of the image. Relative positions of the other lines with respect to the jaw line have been calculated by the training process described in Section 6.

The angle of the candidate line is calculated and can take values from 0 to 360°. This means that a vertical line which separates a bright structure on the left from a darker structure on the right has an angle of 90°, while if the location of the structures is the inverse, the angle of the vertical line is 270°. This permits one to discriminate between contours that would be otherwise confused.

Both position and angle are tested on candidate contours by means of thresholds, so it is possible that more than one contour passes the position and angle test. In this case, we take the longer of these contours, as normally the rest are very short contours due to noise or small structures in the vicinity. In Fig. 8, detected contours are shown on the image from Fig. 3.
5. LANDMARK DETECTION MODULE

Once the interesting contours of the image have been detected, the exact position of landmarks must be calculated. For this purpose, a template-matching technique is used. Many techniques can be found in literature to perform this operation; we have chosen an algorithm based on mathematical morphology [20] for the short calculation time required (21). Furthermore, this algorithm has been successfully used previously [13] for landmark detection.

Detected contours are used to establish the search zones for the landmarks, so that false detections are avoided and searching time is minimized.

5.1. Landmark Localization Areas

Localization areas are defined with respect to the position and angle of the contours. Relative positions of the different landmarks with respect to the contours have been calculated through the training process described in section 6.

5.2. Basic Morphological Operations

Mathematical morphology techniques are powerful tools in the field of image analysis, with a strong theoretical basis [17]. The basic morphological operations are dilation and erosion, which, in gray level images, are defined:

\[
(A \oplus B)(x, y) = \max_{i,j}[A(x - i, y - j) + B(i, j)]
\]

\[
(A \ominus B)(x, y) = \min_{i,j}[A(x + i, y + j) - B(i, j)],
\]

where \(A(x, y)\) corresponds to the original image, and \(B(x, y)\) is the structural element applied.

5.3. Pattern Matching

In our case, mathematical morphology is used as a means for pattern detection [21]. Let \(A(x, y)\) represent the initial image, \(B(x, y)\) represent the mask corresponding to the element searched, and \(A' (x, y), B' (x, y)\) represent their complementaries (in our case with 256 gray levels, \(A'(x, y) = 256 - A(x, y)\)). We define a new function \(C(x, y)\) using the erosion operation:

\[
C(x, y) = (A' \ominus B')(x, y) - (A \ominus B)(x, y)
\]

which can be written as

\[
C(x, y) = \max_{i,j}[A(x + i, y + j) - B(i, j)] - \min_{i,j}[A(x + i, y + j) - B(i, j)].
\]

The point \((x, y)\) which gives a minimum for function \(C\) corresponds to the most likely location of the structure \(B(x, y)\) in image \(A(x, y)\). A small modification of the above equation is carried out to avoid the influence of noise in the final result. The difference \(\max - \min\) that is carried out is very sensitive to noise, so it has been replaced by variance calculation, which is more robust against noise while still giving a measurement of the spread.

In our case, the above expression, with the mentioned modification, is used as follows: \(A(x, y)\) corresponds to the section of the original image where the landmark will be searched, and \(B(x, y)\) is a model for the landmark, calculated by training, as explained in the following section. In Fig. 9, the original image is shown, along with a small image which corresponds to \(B(x, y)\) for the Menton landmark.

6. TRAINING

For the steps that have been described in Sections 4 and 5, important decisions must be taken that rely on the a priori knowledge about cephalographic images. The probable position of the other lines, needed for the step of reference line detection, described in Section 4.4, is calculated depending on the position of the jaw line. Also, to define the landmark localization areas depending on the location of detected lines, a priori knowledge must be used. The model \(B(x, y)\) for each landmark must also be known to apply the pattern matching operations. All this knowledge was obtained by a previous training process, based on a training set of 20 images that were processed as described below.

All the training steps were carried out by experienced orthodontists, for which a simple user interface was developed.

6.1. Location of Significant Lines

In the significant line detection step, the jaw line is searched in the lower half of the image. For the other three
FIG. 9. Location of Menton landmark and landmark model used in the search.

FIG. 10. Average errors between manual and automatically located landmarks.
significant lines, detection is not always so easy, so smaller search areas, relative to the jaw line, are defined to avoid false positives in distant positions.

To define these search areas, an initial edge detection, followed by a contour segmentation and segment joining process, was carried out in the test set images. The jaw line is then found automatically according to its length and angle. The other three lines were selected by hand from the set of detected contours. Relative positions and angles of selected lines relative to the jaw line were computed, and the extremum values of distance and angle, after an addition of extra 2 cm in each direction, for robustness reasons, were taken as the limits of the search areas.

6.2. Calculation of Landmark Localization Areas

As have been described in Section 5, landmark localization is carried out through a pattern-matching step which looks for a model of the landmark in a part of the image, called the landmark localization area.

To define the localization areas, all landmarks were manually located in the test set images. With the significant lines selected as described in Section 6.1, relations between lines and landmarks were calculated. Each landmark was independently studied, to identify interesting relations, depending on their standard deviations. Once identified, these relations were quantified, calculating their extremum values to obtain landmark localization areas for every image.

6.3. Calculation of Landmark Models

The model $B(x, y)$ for each landmark was defined starting with the landmarks manually located in the test set images. Centered in the landmarks, small zones $B_k(x, y)$ were extracted from each image. The final model $B(x, y)$ was calculated by averaging the extracted models.

7. RESULTS

The system was trained, as explained in Section 6, to find 17 cephalometric landmarks, selected from the analysis of Ricketts [5] and Steiner [6]: Gonion (Go), Antegonial (Ag), Point A or subspinal (A), Pogonion (Pg), Gnathion (Gn), Menton (Me), mandibullary central incisor crown, root and most vestibular point (Md1c, Md1r, Md1v), and their corresponding maxillary landmarks (Mx1c, Mx1r, Mx1v), Sella

![FIG. 11. Rate of landmarks successfully detected.](image-url)
TABLE 1
Comparison of Results with [13]

<table>
<thead>
<tr>
<th>Landmark</th>
<th>Our system (%)</th>
<th>Cardillo et al. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Go</td>
<td>85</td>
<td>61</td>
</tr>
<tr>
<td>A</td>
<td>95</td>
<td>77</td>
</tr>
<tr>
<td>Pg</td>
<td>95</td>
<td>97</td>
</tr>
<tr>
<td>Gn</td>
<td>90</td>
<td>100</td>
</tr>
<tr>
<td>Me</td>
<td>100</td>
<td>78</td>
</tr>
<tr>
<td>Md1c</td>
<td>90</td>
<td>64</td>
</tr>
<tr>
<td>Md1r</td>
<td>100</td>
<td>89</td>
</tr>
<tr>
<td>Mx1c</td>
<td>100</td>
<td>76</td>
</tr>
<tr>
<td>Mx1r</td>
<td>100</td>
<td>79</td>
</tr>
<tr>
<td>Se</td>
<td>65</td>
<td>53</td>
</tr>
<tr>
<td>PNS</td>
<td>80</td>
<td>71</td>
</tr>
<tr>
<td>ANS</td>
<td>90</td>
<td>68</td>
</tr>
<tr>
<td>Or</td>
<td>65</td>
<td>40</td>
</tr>
<tr>
<td>Na</td>
<td>85</td>
<td>83</td>
</tr>
<tr>
<td>Average:</td>
<td>88.6</td>
<td>74.0</td>
</tr>
</tbody>
</table>

TABLE 2
Comparison of Results with [10] and [12]

<table>
<thead>
<tr>
<th></th>
<th>Parthasarathy et al.</th>
<th>Davis et al.</th>
<th>Our system</th>
</tr>
</thead>
<tbody>
<tr>
<td>Success rate (%)</td>
<td>83</td>
<td>86.7</td>
<td>95.9</td>
</tr>
<tr>
<td>Deviation (mm)</td>
<td>2.06</td>
<td>2.42</td>
<td>1.1</td>
</tr>
</tbody>
</table>

8. CONCLUSION

A system was implemented for automatic localization of cephalometric landmarks. The system uses a line detection module and a landmark detection module, which are applied sequentially to the images. Results obtained range from acceptable (65%) for landmarks which are not always sufficiently contrasted, such as the lower side of the orbit or the Sella, to excellent (100%) for landmarks clearly defined such as the Menton or the maxillary incisor crown. The high success rate (90.3% for all landmarks) allows the system to be used routinely by clinicians, if they are provided a simple editing function to modify the landmark locations. Comparison of results with previous research was favorable to our system.

Due to the high variability of the images, these results are very difficult to improve without an improvement of the image capture. In this context, direct digital radiography would allow our system to interact more closely with the capturing process, adjusting its parameters and thus improving results.

REFERENCES