COMPUTERS AND BIOMEDICAL RESEARCH 19, 282–309 (1986)

Knowledge-Based Landmarking of Cephalograms[1]

A. D. LÉVY-MANDEL,* A. N. VENETSANOPOULOS,† AND J. K. TSOTSOS‡

*Department of Electrical Engineering, E. P. F. L., 1007 Lausanne, Switzerland; and Departments of †Electrical Engineering and ‡Computer Science, University of Toronto, Toronto, Ontario, Canada

Received June 13, 1985

Orthodontists have defined a certain number of characteristic points, or landmarks, on X-ray images of the human skull which are used to study growth or as a diagnostic aid. This work presents the first step toward an automatic extraction of these points. They are defined with respect to particular lines which are retrieved first. The original image is preprocessed with a prefiltering operator (median filter) followed by an edge detector (Mero–Vassy operator). A knowledge-based line-following algorithm is subsequently applied, involving a production system with organized sets of rules and a simple interpreter. The a priori knowledge implemented in the algorithm must take into account the fact that the lines represent biological shapes and can vary considerably from one patient to the next. The performance of the algorithm is judged with the help of objective quality criteria. Determination of the exact shapes of the lines allows the computation of the positions of the landmarks.

INTRODUCTION

Thomson was the first to propose a method for analyzing the growth of the head, in 1917 (*1*). His approach was global. Regularly spaced points were defined arbitrarily on the profile of the head, corresponding to the nodes of a square grid. The subsequent growth of the head resulted in a deformation of the grid, thereby highlighting the changes. Nowadays, however, the landmarking method is used by orthodontists. According to this method, characteristic points, or landmarks, are not arbitrarily defined as in the former approach, but are related to characteristic anatomical features. Since these points are relatively few (*36*) and are defined objectively, comparisons are possible on the basis of quantitative measurements. The constellation of points and its evolution in time help determine the appropriate orthodontic treatment.

In practice the basic material used by orthodontists is a lateral head X ray, or cephalogram, and the characteristic anatomical features appear either as lines

[1] The research reported in this paper was supported by Operating Grant A7397 of the Natural Sciences and Engineering Research Council of Canada.

0010-4809/86 $3.00

or as edges on a dark background. The orthodontist initially traces the relevant lines and edges on the X ray and then locates the characteristic points.

Up to now landmarking has been done manually. It is a long and tedious process in which the results can vary from one judge to the next, in proportions comparable to the variation of the position of the landmarks on two different cephalograms (*2*). Therefore a system performing the automatic computation of the landmarks was deemed highly useful for both the daily practice of orthodontists and the compilation of a large and reliable set of data, required as a statistical basis for scientific research on human growth. In particular, thousands of cephalometric X rays accumulated over numerous years at the Burlington Center of the University of Toronto have not yet been systematically exploited. The research presented in this paper aimed toward this goal. Figures 1a and b show two digitized cephalograms, called Lisa (a) and Diane (b). These cephalograms correspond to two young girls, 12 and 10 years of age, respectively. The X rays were digitized with a Matrox frame grabber, and quantized with 256×256 pixels, using 256 grey levels (8 bits). Lisa and Diane are two of the better cephalograms out of a series of many used in this research project.

The work presented here is a study of the automation of the landmarking process. The strategy adopted may be described as follows. Since the information is contained in the lines and edges, the first step is to highlight them by prefiltering and then to apply an edge detector to the digitized X ray. Subsequently the characteristic lines that define the landmarks are extracted from a background of noise and irrelevant lines, using a priori knowledge coded in the algorithm. Finally the computation of the landmarks is performed.

In this paper the general problem of the automation of the landmarking process is discussed in Section 2. Section 3 presents a brief literature review. In Section 4, the image quality criteria used to evaluate the X rays prior to the application of the algorithm are introduced. Prefiltering, edge detection, and extraction of the relevant lines followed by the computation of the positions of the landmarks, are presented in Sections 5 and 6. Finally, Section 7 presents some conclusions and suggestions for future research.

2. Automation of Landmarking

We first examine the general problem of automatic landmarking, along with its characteristics and difficulties. Thus, we develop our strategy in a step-by-step fashion.

Since our images represent biological shapes, they cannot be described in terms of shifted and rotated patterns that could be easily recognized by an algorithm. An infinite set of shapes, featuring a variety of different orientations, scales, and deformations that support the same significant features, is possible. This implies that our problem is a mixed detection, recognition, and estimation problem. In a classical detection problem the task is that of detecting one or more signals in the presence of noise. In a usual recognition problem, the task

FIG. 1. Digitized X rays of Lisa (a) and Diane (b).

is to classify the signal, without necessarily finding its exact shape. In an estimation problem, on the other hand, the task is that of retrieving the exact shape of the signal, without necessarily identifying it. In this application, in order to detect the relevant lines and estimate their exact shapes, it is necessary to recognize them. But it is impossible to recognize them without having detected at least some of them.

The task is complicated by the characteristics of the background in the image. The background is composed of white noise and of irrelevant edges and lines, highly correlated with the signal consisting of the relevant lines of more or less the same intensity. This means that the algorithm must contain some sort of a priori knowledge about the relevant features.

These features appear as both edges and lines of the original image. But since the information content of a line and of an edge is the same, as far as their shape is concerned, the detection problem is simplified if an edge detection is performed first on the image. The preprocessed image is then composed of a pattern of bright lines, some of which are significant, running across a dark and noisy background.

In addition to the noise originally present in the radiograph, the digitization noise alters the picture even more. The most annoying effect is a spurious grid that the digitizing camera adds to the homogeneous parts of the image. Consequently, part of the details that were clear on the original image disappears after digitization. In a way the simplification of the image resulting from the digitization can be considered beneficial, if only insignificant details disappear. However, in the problem at hand, some relevant lines also vanish in the process, which explains why it was not possible to retrieve some of the landmarks. Figure 2 shows an original X ray (Lisa). A comparison with its digitized version (Fig. 1a) illustrates the previous statement.

This loss of detail can also be due to a bad exposure or anomalies (e.g., filled cavities). Section 4 discusses how objective criteria can be developed to evaluate the quality of an image with respect to these defects.

Moreover, as mentioned above, an infinity of possible forms describes similar anatomical features. Hence the knowledge of these features cannot be described nor coded in terms of coordinates in a fixed frame of reference. We have chosen to describe it in terms of categories of knowledge, each corresponding to a characteristic of a line, as follows.

1. Position. The line must first be found; hence the algorithm must know where and how to look for it. As all lines have more or less fixed positions relative to one another, finding one line is made increasingly easy as the algorithm finds the others. It is, however, much easier to locate the exterior lines than those inside the head. There is, therefore, a natural partial ordering among them.

2. Number and characteristics of constituent segments. A line is modelled as formed of straight segments separated by abrupt changes of direction. Each segment has a number of characteristics:

FIG. 2. Original X ray (Lisa).

—its starting and ending points,
—its approximate length,
—its general orientation, and
—the characteristics of the noise around it. This corresponds to how well the line detaches itself from the background (whether it is much brighter than the background, how bright the intersecting lines are, etc).

3. Where the line starts and stops. This occurs because of a particular change of direction, when the edge of the image is reached, or else because the current pixel reaches a certain position relative to other features in the head. Some lines also stop because their intensity becomes too low.

Hence the three steps of the landmarking process are (1) a prefiltering followed by a process that transforms the picture into a pattern of lines; (2) the extraction from this pattern of the relevant lines, using the a priori knowledge presented above; and (3) the computation of the positions of the landmarks.

An outline of the head with the landmarks is shown on Figure 3. The points are defined with respect to the darker lines of the figure. The literature review that follows concerns the line extraction. The edge detectors used are abundantly described elsewhere (*3*).

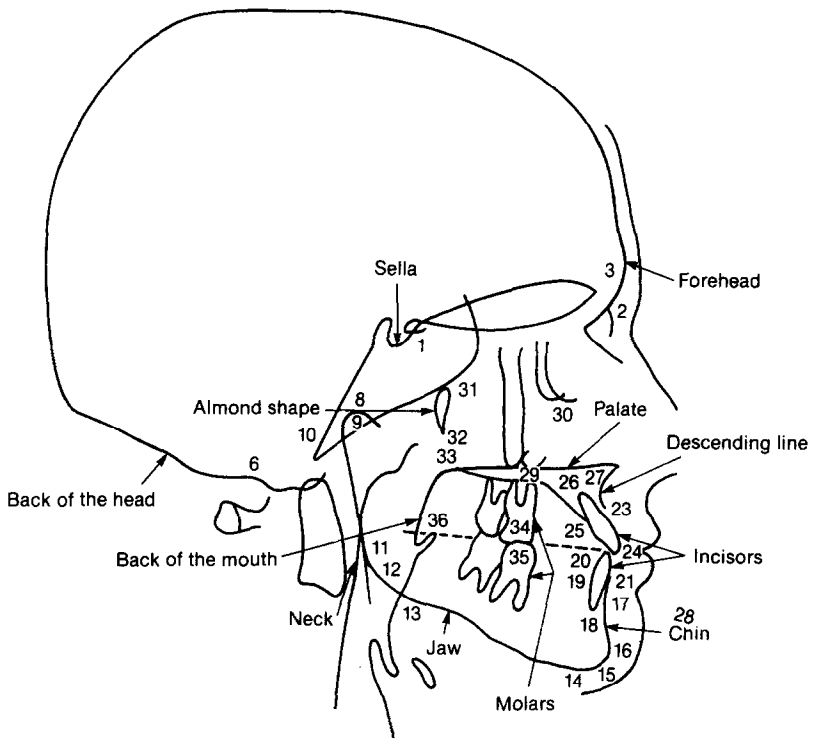

FIG. 3. Outline of the head with the landmarks and the names of the features.

3. LITERATURE REVIEW OF KNOWLEDGE-BASED LINE EXTRACTION TECHNIQUES

In a number of applications there is a need for an efficient use of a priori knowledge to help the extraction of the relevant lines. Most problems of image content analysis treated in the literature deal with recognition of hand-printed characters (especially Chinese, easily decomposable into building blocks) or analysis of aerial images for civil or military applications. Detection is more scarcely discussed. Two different methods described in the literature are heuristic edge detection and knowledge-based vision.

Heuristic edge detection. Kelly (*4*) defines a three-step plan using the edges of a smaller version of the image to find those of the original picture. These edges play the role of a priori knowledge and are used to eliminate parallel edges and as general orientation information. This method has been applied to the extraction of head contours in photographs. In the work of Ashkar and Modestino (*5*), the edge-enhanced image is transformed into a tree graph, where each node has seven successors (seven of its neighbors, the eighth one being the predecessor). An optimal path is found through the graph, according to the actual information about the preprocessed pixel (result of the edge operator), a posteriori knowledge (curvature of typical contours), and a priori knowledge (difference between the actual image and a prototype). The algorithm was

applied to lung and heart images. Finally, Gritton and Parrish (*6*) deform a standard set of lines progressively, comparing it to the actual image. This is the "bead chain algorithm" that was applied to liver cell images.

Knowledge-based vision methods. These methods are generally concerned with more complicated and less predictable scenes than the previous methods. Shirai (*7*) first analyzed a block world and developed an algorithm that finds its objects. When his method was applied to a more realistic model of the world (*8*), difficulties were encountered. In order to overcome them, he used a priori knowledge to build a "reference map" that guides an edge tracker; this map is updated as more objects are found. Ballard *et al.* (*9*), Russell (*10*), Levine (*11*), and Mackworth (*12*) all divide their representational state into three levels: a low level, containing the pixel-by-pixel information about the image, or an edge-detected version of it, or else a grossly segmented version; an intermediate level, with regions or edges labeled with a priori knowledge; and a high level, containing an interpretation of the scene analyzed. They insist on the necessity of an interaction between the model and the actual picture.

The algorithm described in this paper was developed using ideas from both the previously described strategies. The work of Kelly (*4*) is relevant here in the sense that the use of a particular plan for each patient's plate renders the search more efficient. However, another method was used to build up the plan faster than Kelly's. The a priori knowledge, consisting of the difference between a prototype and the actual radiograph, such as in the work of Ashkar and Modestino (*5*) or that of Gritton and Parrish (*6*), is not sufficient for complicated images such as cephalograms. On the other hand, Shirai's (*7, 8*) idea of a reference map was applied, and the information it contained was updated as lines were found. Finally, we opted for an interaction between model and data in the knowledge-based line extractor.

4. Objective Quality Criteria

Two radiographs can differ not only in the images they show, but also in their own quality. Some of the lines, clearly visible on one radiograph are hardly perceptible on another. Objective quality criteria are therefore necessary to decide whether or not an image is tractable, as well as to allow an evaluation of the performance of the algorithm. Two types of criteria are examined here: first, whether the image has been correctly exposed, and second, whether it is free from extraneous structures and improper positioning effects (anomalies).

4.1. Global Quality Criterion, Control of the Exposure

There is in fact a greater loss of information during the radiographic process than during the digitization of the radiograph. The exposure selects a limited range of densities of the skull that appear on the film, which functions as a compressor of the density range. The only parameter that can vary from one radiograph to the next is the quantity of radiation to which the patient is exposed. As a result, the only blurring that can occur, if the effects of scattered

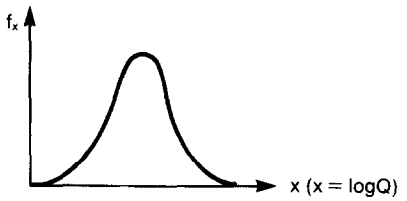

FIG. 4. $f_x(x)$.

and secondary radiation are neglected, is at the digitization stage. Moreover, there is no precise way to determine the exposure necessary for a particular patient. There are, therefore, instances when the exposure is wrong. Since the sensitivity of the eye is logarithmic, the characteristic curve of the radiographic film is to be understood as the logarithmic optical density after processing versus the logarithm of the incident density of X rays. Only under this representation is the compressor function of the film made clear: there is a linear region where a scale of gray shades matches a scale of X-ray densities, bordered by two saturation regions (under- and overexposure), yielding an S-shaped characteristic function. We are going to demonstrate that the exposure can be checked through the shape of the logarithmic histogram of the digitized image.

In this paper it is assumed that the logarithmic histogram of the X rays of all the patients' heads are the same, and that this histogram is bell shaped. Indeed, two human heads are sufficiently similar to be represented by the same density function f_x (Fig. 4). This hypothesis is verified by the experiments.

The logarithmic histogram of a cephalogram is given by the transform of $f_x(x)$ by the S-shaped curve that characterizes the film (Fig. 5). Applying the equation

$$f_y(y) = f_x(x)/|g(x)| \qquad [1]$$

where $x = g^{-1}(y)$, f_x is the bell-shaped curve, and g is the S-shaped curve characteristic of the film, approximated for simplicity as an arctangent function (Fig. 6), yields

$$f_y(y) = (l\pi/2m)(1 + \tan^2((\pi/2m)(y - <g>))f_x(x_0 - l\tan((\pi/2m)(y - <g>))) \qquad [2]$$

where $<g>$ = average grey, x_0 = corresponding exposure, and m, l = grey scale and exposure spans delimitating the linear part of the S curve.

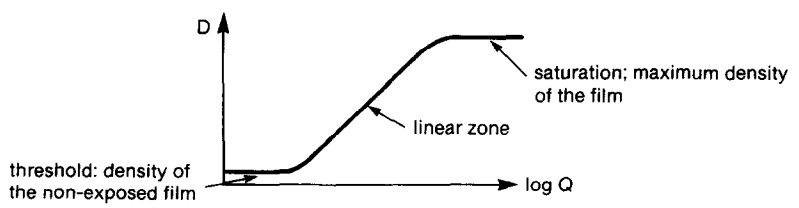

FIG. 5. $g(x)$.

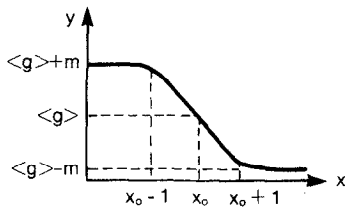

FIG. 6. Arctangent function.

$$|g'^{-1}(y)| = (l\pi/2m)(1 + \tan^2((\pi/2m)(y - <g>)). \quad [3]$$

This function goes to infinity for $y = <g> + -m$. Therefore the logarithmic histogram (f_y) of the pixels on a digitized X ray has two humps, one corresponding to white saturated regions with few details and the other to black saturated regions. If the X ray is correctly exposed, the two humps are of equal importance. If the first hump is larger, the picture is too dark, therefore overexposed, whereas if the second hump is larger, the picture is too light, thus underexposed. If the heights of the two humps are respectively called a and b, a simple criterion is the ratio $R = a/b$.

The shapes of the logarithmic histograms for Lisa and Diane can be found on Figs. 7a and b, respectively. The corresponding ratios are for Lisa, $R = 0.955$; and for Diane $R = 0.988$.

These two X rays are well exposed. On other X rays not shown here, the ratio R is very different from unity. On those, the desired information appears less clearly, and is severely degraded when the radiographs are digitized.

4.2. Presence of Anomalies

There are mainly three kinds of anomalies. Filled cavities, missing teeth, and the two profiles resulting from a patient whose head was not upright when X rayed. Cavities, when filled with a metal alloy, are nearly opaque to X rays. They therefore appear as bright spots, that have a tendency to "bloom" and obscure neighboring lines. The absence of a tooth is hard to detect. Teeth do not have very sharp edges, and it is difficult to distinguish between the shape of a tooth and that of a hole. Finally, the possibility of the presence of two shifted

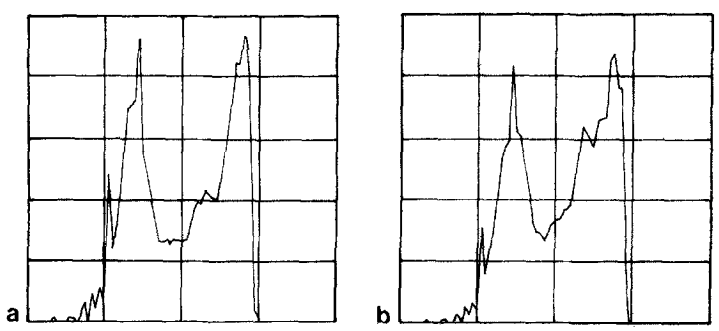

FIG. 7. Logarithmic histograms of Lisa (a) and Diane (b).

profiles has not been included in this first realization of the algorithm for the sake of simplicity.

In Diane, there are a few cavities and some instances of the two profiles. Lisa does not present any anomalies. We have also considered other X rays where anomalies were much more numerous, in which the algorithm makes more errors.

5. Edge Detection

5.1. Introduction

The two steps in the transformation of the digitized X ray into a picture with lines only are the following: first, a prefiltering stage, then an edge detection that performs the transformation.

5.2. Prefiltering

Prefiltering is used to improve the performance of the edge enhancer that it precedes. There are in fact two possibilities:

—enhance the edges without significantly enhancing the noise; and
—eliminate some of the noise while preserving the edge information.

Histogram equalization (*3*) and extremum sharpening (*13*) perform the first task. Median filtering (*13*), performs the second. After examining the results of these operators on a number of cephalograms, we decided to use median filtering, which produced superior results (*13*).

Median filtering replaces the value of the center pixel of the window with which the picture is scanned by the median of all the values inside the window. The median of a set of n numbers $x_1, x_2, \ldots, x_n$ is, for n odd, the number occupying the middle position after sorting of the values. Let x_i be a value in the window considered; its probability of apparition is called p_i; if n_i pixels have the value x_i in the window of size $m \times n$, then $p_i = n_i/(m \times n)$. Thus the median x_f of the sequence $[x_i]$ is the first value verifying

$$\sum_{i=-1}^{x_f} p_i \geq \tfrac{1}{2}. \qquad [4]$$

Since its result depends on all the values inside the window, and since the median always belongs to the set of values of the window, the median filter preserves the extremities of a ramp. Therefore it reduces impulse noise without blurring the contours. It also smoothes sharp projections or indentations that are small relative to window size, and destroys long and thin objects. This limits the possible size of the window. The result of a 3×3 median filtering is shown on Fig. 8. We see that the operator reduces spurious and granular-type noise, without blurring lines or edges.

FIG. 8. Result of the median filter.

5.3. *Edge Detection*

Four operators were tried and compared: Laplacian, two nonlinear gradient operators (Sobel and Prewitt), Kirsch, and Mero–Vassy (simplified Hueckel).

Laplacian operator. This is a linear approximation of the second derivative. On a 3 × 3 region, it is given by

$$L(x, y) = |4f(x, y) - (f(x, y + 1) + f(x, y - 1) + f(x + 1, y) + f(x - 1, y))|. \quad [5]$$

A representation of the 3 × 3 window can be found on Fig. 9a. The Laplacian is not overly sensitive to gradual intensity level changes, but enhances noise enormously.

Nonlinear gradient operators. These are given by the equations (*3*)

$$G(x, y) = [g_1(x)| + |g_2(y)| \quad [6]$$

where

a

x-1,y-1	x,y-1	x+1,y-1
x-1,y	x,y	x+1,y
x-1,y+1	x,y+1	x+1,y+1

b

A_1	A_2	A_3
A_8		A_4
A_7	A_6	A_5

FIG. 9. Representations of the 3 × 3 window (a, b).

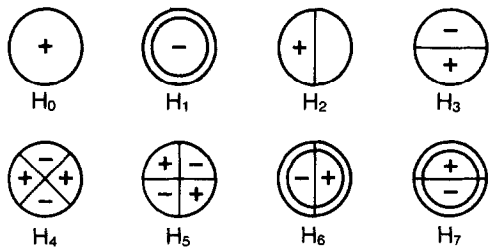

FIG. 10. Functions H_i.

$$g_1(x) = [f(x, y + 1) + kf(x + 1, y) + f(x + 1, y - 1)] - [f(x - 1, y) + f(x - 1, y - 1)],$$

and

$$g_2(y) = [f(x + 1, y + 1) + kf(x, y + 1) + f(x - 1, y + 1)] - [f(x, y - 1) + f(x - 1, y - 1)]. \quad [7]$$

When $k = 1$, G is the Prewitt operator, when $k = 2$, the Sobel operator, and when $k = sqrt(2)$, the operator is called isotropic.

Kirsch operator. This is also a nonlinear gradient operator. If the representation of the window is changed to that of Fig. 9b, the operator can be expressed as

$$K(x, y) = max[1, max(i = 0 - 7)[|S_i - 3T_i|] \quad [8]$$

where

$$S_i = A_i + A_{i+1} + A_{i+2}$$

$$T_i = A_{i+3} + A_{i+4} + A_{i+5} + A_{i+6} + A_{i+7}. \quad [9]$$

The indices are taken modulo 8.

Mero–Vassy operator. This is a simplified version of the Hueckel operator (*14*). Hueckel considers what a theoretical edge looks like within a circular region; he then finds how far each region of the image is from this configuration by using a best fit criterion such as the least-squares criterion. Two grey levels are supposed to exist in the circular neighborhood. Hueckel finds the location of the edge by considering a linear combination of 8 functions H_i (Fig. 10). The result of this operator is orientation invariant and relatively immune to noise; however, it is very costly. This is the reason why Mero and Vassy (*15*) proposed a simplified version of the Hueckel operator. It is obtained by reinterpreting it as a template-matching method with a linear combination of two square templates

$$\begin{vmatrix} 1 & -1 \\ 1 & -1 \end{vmatrix} \quad \begin{vmatrix} 1 & 1 \\ -1 & -1 \end{vmatrix}.$$

On a square area of dimensions 2×2, the operator is expressed by

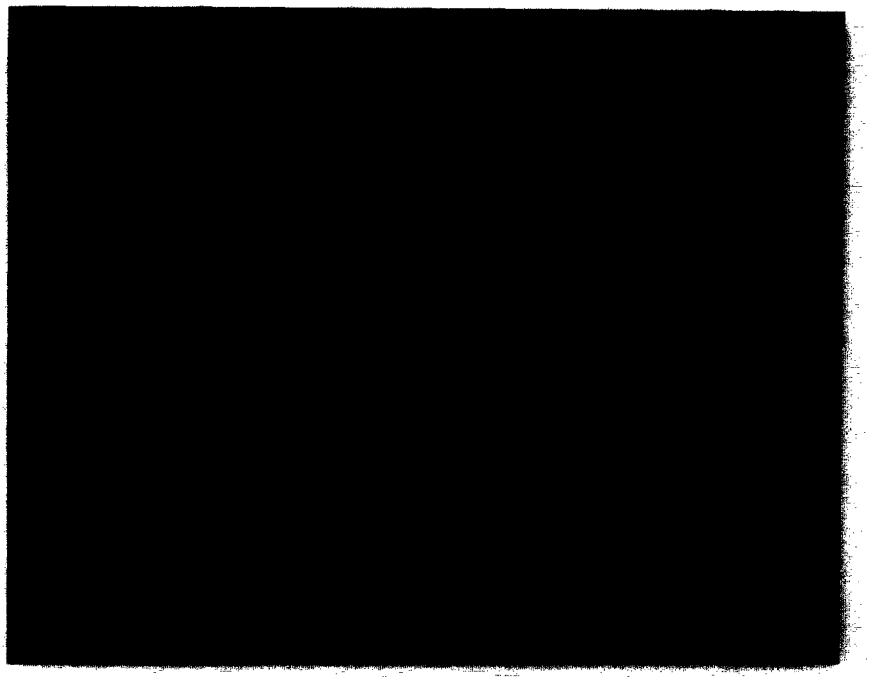

FIG. 11. Result of the Laplacian operator.

$$MV(x, y) = |f(x, y) + f(x, y + 1) - f(x + 1, y) - f(x + 1, y + 1)| + |f(x, y) + f(x + 1, y) - f(x, y + 1) - f(x + 1, y + 1)|. \quad [10]$$

5.4. *Results of the Edge Detectors*

No thresholding was used after the edge operators, in order to enable comparisons. All operators are applied to the median-filtered version of the X ray.

On Fig. 11, the poor results of the Laplacian are presented. They can be explained by the extreme sensitivity to noise of the operator. Figure 12 shows that the results of the Kirsch operator are very different from what is desired, i.e., a pattern of lines.

The Prewitt and Sobel operators give approximately the same result (Figs. 13 and 14). It is considerably better than those obtained previously. Edges are enhanced, thin lines are preserved and the important features are clearly visible. Finally, the result of the Mero–Vassy operator is shown (Fig. 15). It enhances noise slightly less than the Sobel and Prewitt operators, which is the reason that this operator was chosen for preprocessing.

6. EXTRACTION OF THE RELEVANT LINES AND COMPUTATION OF THE LANDMARKS

6.1. *Global Line Follower*

The basic line follower can be described as follows. The preprocessed image

FIG. 12. Result of the Kirsch operator.

FIG. 13. Result of the Prewitt operator.

FIG. 14. Result of the Sobel operator.

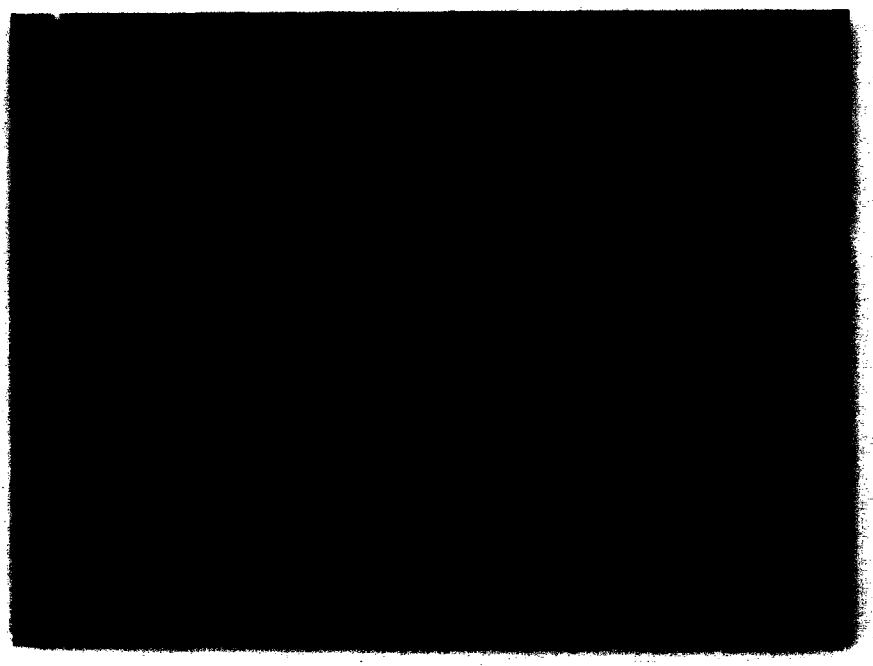

FIG. 15. Result of the Mero–Vassy operator.

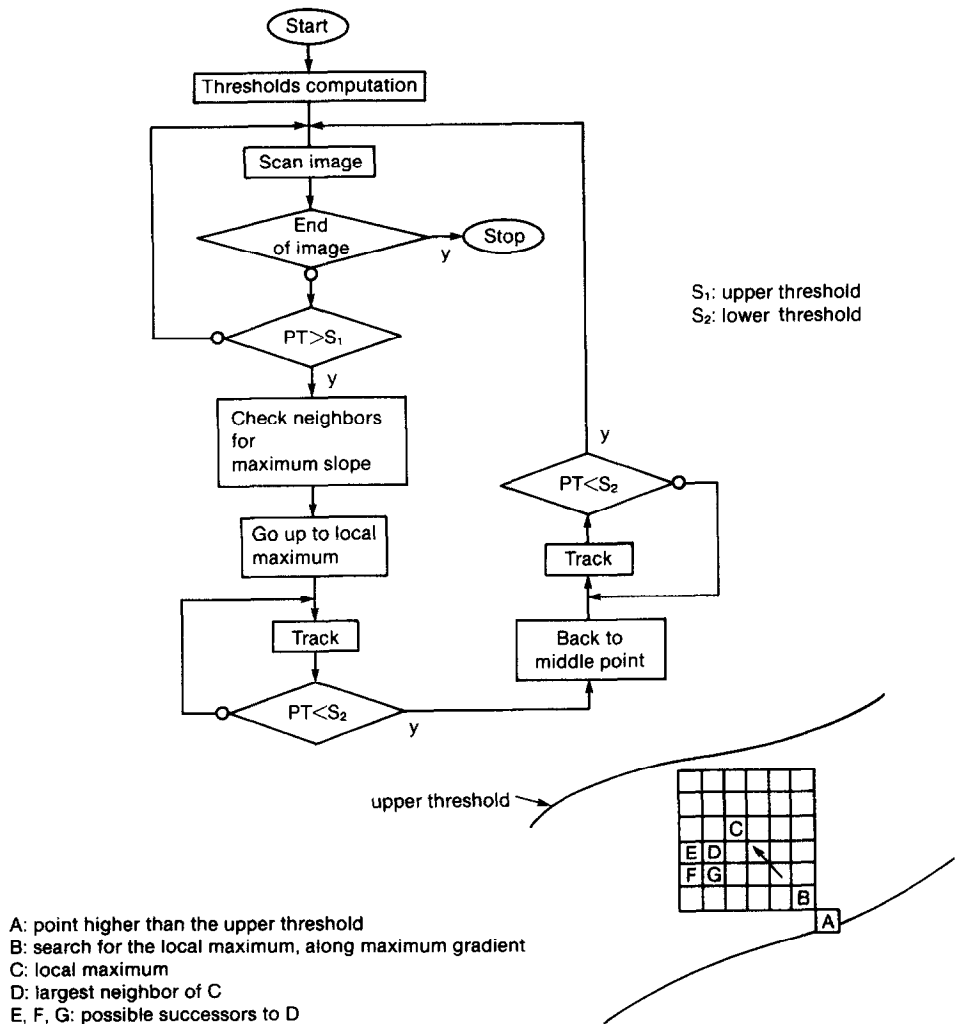

FIG. 16. Flowchart of the global line follower.

resembles a three-dimensional geographical map: the first two coordinates are in the plane of the image and the third corresponds to the grey level. Hence the lines are the ridges of mountains. The line follower, illustrated by Fig. 16, is composed of the operations that follow:

—Define two thresholds s_1 and s_2 (s_1—upper threshold, s_2—lower threshold).

—Scan the image line by line until the value of the current pixel is higher than the upper threshold (point A); the higher the threshold s_1, the less sensitive the process is to noise (with the risk of missing more and more lines).

—Check all of A's neighbors for the steepest slope direction (point B) and climb up to a local maximum (point C).

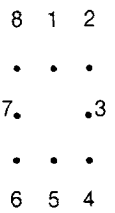

FIG. 17. Discrete directions in the image.

—Search in the direction given by C and its largest neighbor; the contour is followed at plus or minus 45 degrees the previous direction.

—Return to the starting point (C) when the current pixel has a level lower than the lower threshold, for the second half of the search; the threshold s_2 is rather low, so as not to create gaps in the lines.

—Resume scanning (return to Step 2).

This is a bidirectional tracking, where the algorithm finds a point on the line (not necessarily in the middle of the line) and tracks both halves, one after the other. This is convenient as most lines are dimmer at the extremities than in the middle. In the square grid used for sampling, each pixel has eight neighbors. There are therefore eight discrete directions in the image at the pixel level, labeled according to Fig. 17. An optimal computation of the thresholds can be found in (*13*). It is done on the basis of the histogram of each preprocessed image. The result of the global line follower is shown on Fig. 18. It is now

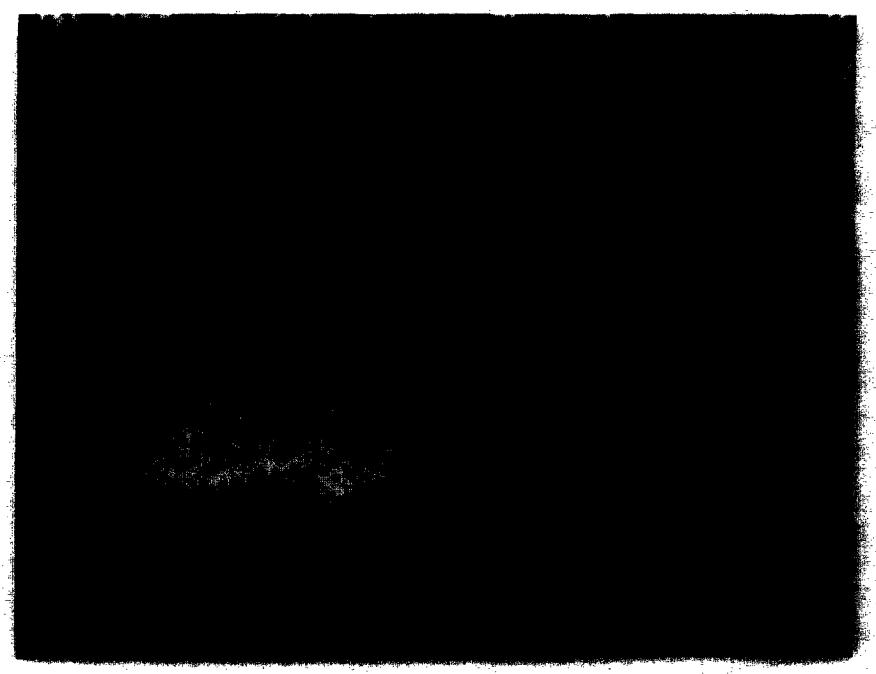

FIG. 18. Result of the global line follower.

possible to describe the assumptions about the data, as well as the knowledge-based algorithm.

Assumptions about the data. These are based on the criteria developed in Section 4. It is assumed first that the X ray is correctly exposed, i.e., that $R = a/b$ of the logarithmic histogram is approximately equal to unity. Then, it is assumed that the image does not present any of the anomalies previously described: filled cavities, missing teeth, or the two-shifted profiles that result from a patient with his head tilted.

Finally, some assumptions are made concerning the position of the head: it should be upright and placed in such a way that the spinal cord is completely enclosed in the bottom-right quadrant and the forehead in the upper left octant. The top of the head is not visible. These assumptions determine what knowledge should be included in the algorithm. What follows is a description of its general structure.

6.2. *Structure of the Knowledge-Based Line Follower*

A knowledge-based line tracker guided by a reference map is used, as in Shirai's work (*7, 8*). The map is initialized by a planning step that gives the gross proportions of the lines of this particular patient. The map is progressively updated as more lines are found. The initial plan uses knowledge about the likely positions of the lines. The rest of the knowledge is implemented in the tracking procedure itself.

The algorithm uses a production system based on the global line-following algorithm. A production system was chosen because it can be organized in a very natural fashion. As the landmarking problem has a modular structure well adapted to the formalism of a production system, the algorithm is easily written and read.

The algorithm traces all the lines in a predetermined order. The images are similar enough to assume that it is always the same lines that are clearer than the rest. The exact order is discussed later.

For each line, the algorithm tries to reach sequentially a number of goals. It fires sets of rules from the organized data base to achieve these goals. The modular sets of rules correspond to each of the categories of knowledge chosen previously.

1. The first set of rules is used to guide the search for the lines.
2. When the algorithm places itself on a line, for each segment there is a set of rules for start and stop, one for the way to choose the next pixel, including if there is an intersection, and one to decide if the chosen pixel is correct or has to be replaced, and by which other pixel.
3. A third set of rules establishes the stop conditions of the line.
4. The last set of rules is used to determine the correctness of the line and what parameter (one of the thresholds, for example) to modify before retracking it.

The interpreter is simple: it follows the logical path between these sets,

starting with Step 1, then applying 2 and 3 at each pixel, then 4 when the line is completely tracked. There is one rule for each action. The condition(s) to be satisfied are different for each line. The control strategy is mostly irrevocable, i.e., the line is traced without backtracking, but the result of the tracing is questioned and the line is possibly retraced, after modification of certain parameters. In the two following paragraphs, the algorithm is explained at the different levels of the lines and pixels.

Global line-follower. It is used as a kernel to which knowledge is added. The position and general orientation of the line are the basic elements of the a priori knowledge to be implemented at the line level. As stated earlier, the algorithm progressively expands an implicit model of the head: it starts with gross proportions, then the exact shapes of the lines are added one by one to the output. The algorithm finds line shapes that are used to modify the initial plan; this defines the order in which the lines have to be tracked, i.e., from the simple to the difficult ones, in the case of a lateral head X ray from the outside lines to those inside the head, as seen in Fig. 19.

Initial plan with the approximate proportions of the head. These gross proportions are computed with simple and fast operations like thresholding of projections and cuts; the locations of the spinal cord, of the back of the head, and of the jaw and the forehead are found.

Search for the lines. When the algorithm searches for them, the lines (cf., Fig. 3) are ordered in an inverted triangle as follows:

forehead back of the head jaw

sella chin back of the mouth

almond top of the mouth neck

descending line in the front

upper incisor

lower incisor

upper molar

lower molar

The lines are searched for on the above drawing to the right, then downward. Note that if some critical line at the base of the pyramid (as the line of the jaw, for example) is missing, no line can be found inside the head. For each line, the search proceeds as follows (Fig. 20).

—From the analysis of cuts and projections (gross proportions), find a point in the vicinity of the line; this point, called "area mid-point" (AMP), is defined as the center of a square area of size 10×10 pixels, through which it is guessed that the line runs.

—Scan the square and track on a short distance (typically 20 pixels) all the lines encountered.

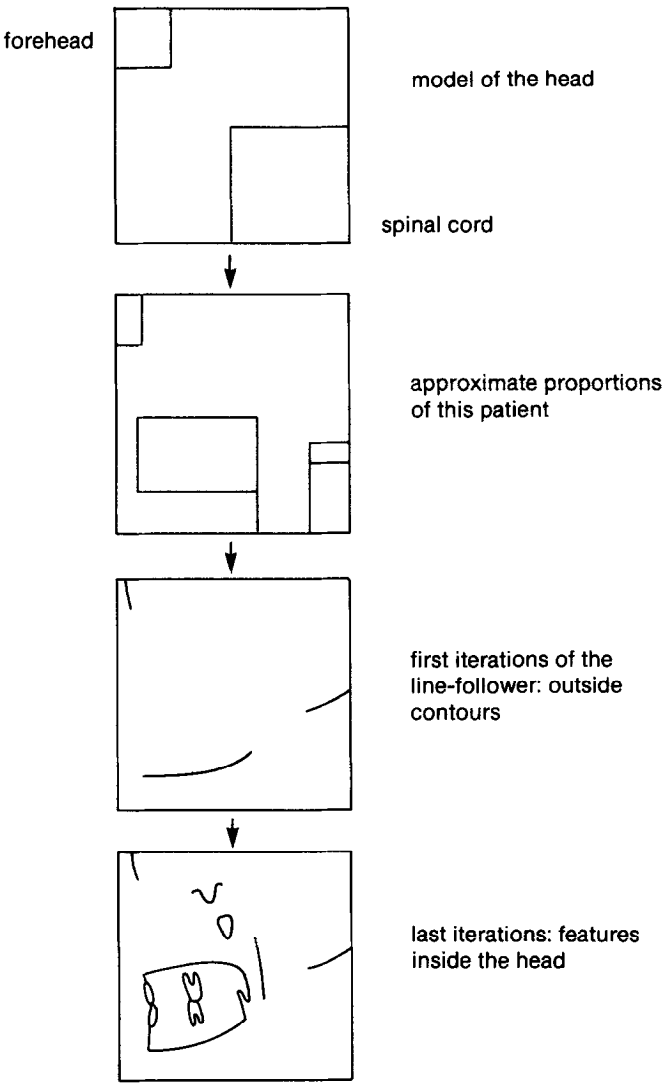

FIG. 19. Structure of the knowledge-based algorithm.

—Using arguments of relative position and according to the general orientation of the line, decide which line is the right one; in general this is not too difficult since the square is small and not many lines run through it.

—Track the line completely, using a priori knowledge.

This is a bidirectional tracking, using the basic global line follower previously described. The point found in the square defined by the AMP is on the line, which is therefore divided in two halves, both of which have to be tracked.

The computations of approximate dimensions allow definition of the AMPs for the outside contours. The knowledge gained by tracking the corresponding

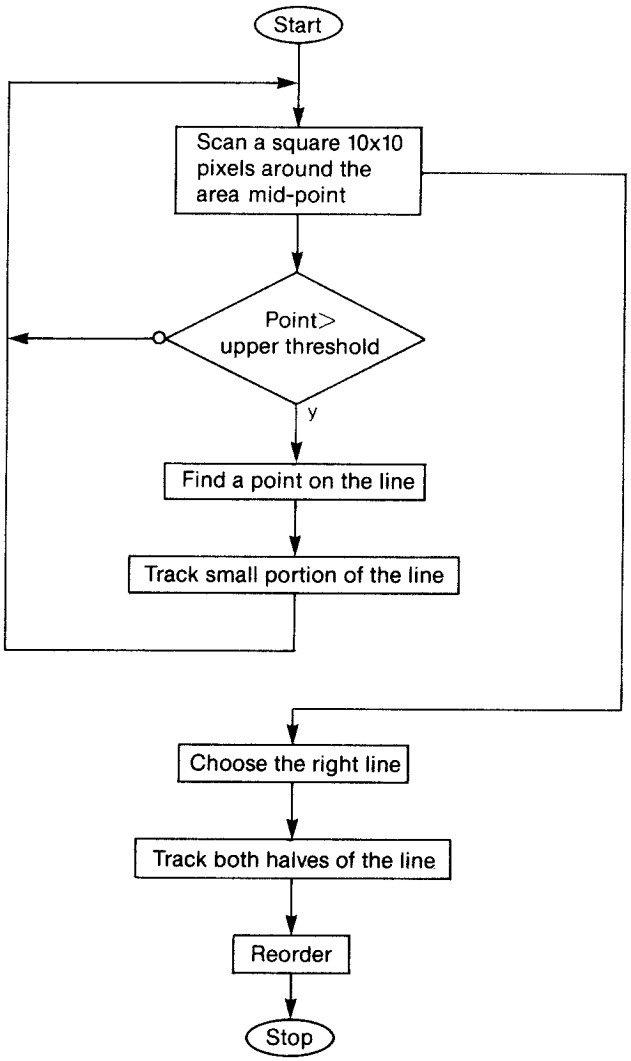

FIG. 20. Flowchart of the knowledge-based line follower.

features serves to compute the rest of the AMPs. This level of the description only takes into consideration a very small part of the knowledge, i.e., where the lines are situated in relation to one another and what general orientation they have around the area mid-point. While the algorithm is actually tracking each line, however, more knowledge is needed.

Knowledge-based tracking. The implicit model can be compared to the directions that a traveller could be given about the path he should take in a forest. At each step, he must compare his position with the directions, find the orientation of the next step while wondering about possible intersections, and decide when he has reached his goal.

The central part of the algorithm is the search for the next pixel on the line. It

is assumed that there are no hairpin turns in the lines. In any case they would be filtered out by the low-pass action of the median filter at the prefiltering stage. Therefore the derivative of the lines is continuous and the successor pixel may be searched among three of the eight elementary directions: if the direction between the previous and the current pixel is d, then the successors are the pixels which form the directions d, $d + 1 \bmod 8$, and $d - 1 \bmod 8$ with the current pixel.

Of course some criteria are involved in choosing the next pixel among the successors. We use a criterion derived from the maximum likelihood test. In the case of a binary signal, a detector observes the received signal r and computes the likelihood function

$$P(r|1)/P(r|0) \gtrless n. \qquad [11]$$

If the ratio is larger than the threshold n, the received signal is a 1; otherwise it is 0. In order to apply this equation exactly, it is necessary to know the analytical characteristics of the noise. Since the part of the background that contains the irrelevant lines cannot be expressed as a noise-like stochastic process, we use the similarity with the problem of a one-dimensional signal in Gaussian noise transforming the test into one comparing the intensities of the pixels.

The following three steps are thus performed:

—sorting the three candidates pixels by order of increasing intensity; calling the maximum intensity pixel *max* and the intermediate one, *next*;

—comparing the intensities of *max* and *next* and determining if there is an intersection (if (*next*/*max* > threshold) then intersection); and

—if there is an intersection, choosing the correct successor with appropriate rules; if there is no intersection, the successor is *max*.

There are psychophysiological reasons for choosing the comparison between *max* and *next* as a ratio rather than a difference. It has been shown that the eye is sensitive to the relative intensity of light, rather than to the absolute (Weber–Fechner law), thus it is likely that the decisions made by a human judge examining a radiograph are based on ratio comparisons.

The threshold corresponds to the "noise around the line" mentioned earlier. The parameter that represents it can vary between 0 and 1 and is assigned to different values according to how well the pattern at this particular segment can be distinguished from the background, as well as to the knowledge of anatomy. For example, it is increased in the vicinity of a change of direction.

Apart from the search for the next pixel, other control structures, which code other categories of knowledge, were implemented. The flowchart of Fig. 21 is a detail of the box labeled "track both halves of the line" in the flowchart of Fig. 20.

The first category is that of the knowledge about where each segment of the line starts and stops. The algorithm starts by initializing some parameters that correspond to the position of the current point on the line (the feature being divided into different segments, a variable is set indicating which segment the

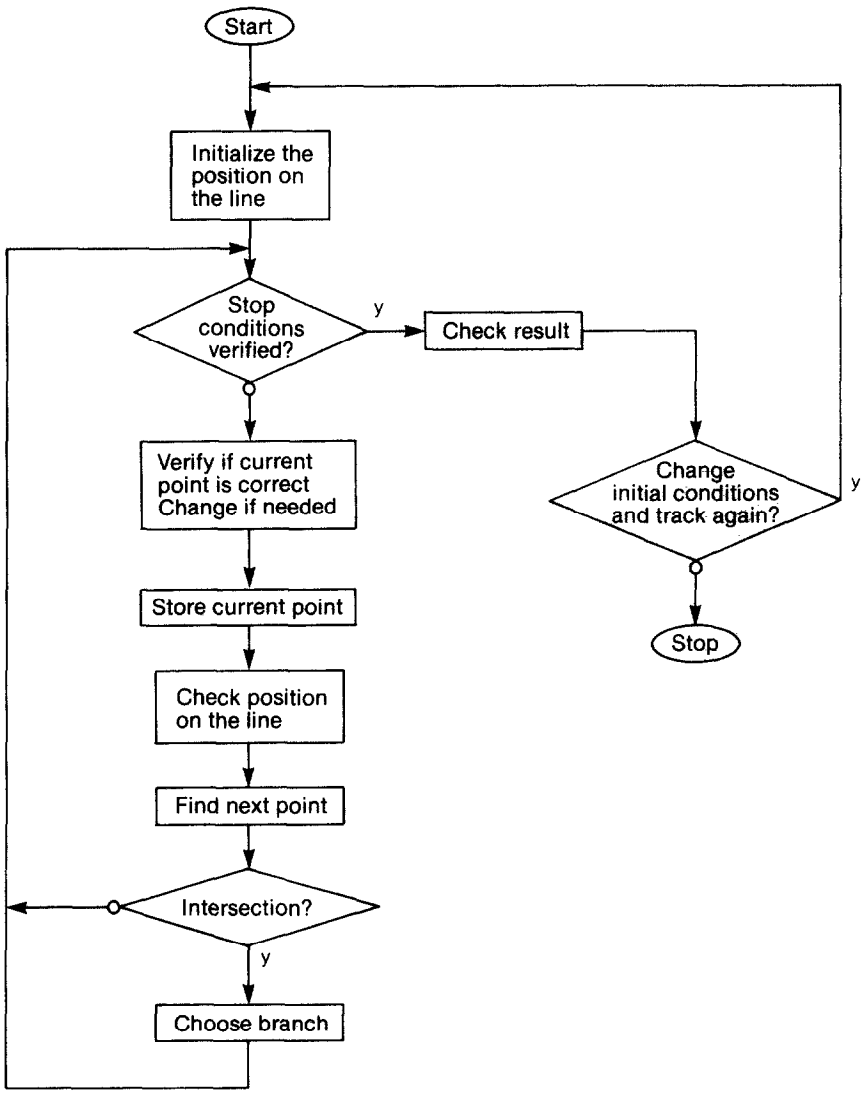

FIG. 21. Flowchart of the knowledge-based tracking.

first pixel belongs to); this is the segment currently being tracked. Then, while the line is tracked, these parameters are updated. This is particularly important for features with many different segments, such as sella.

In the case when the line that is being tracked merges with the background, the algorithm might make the wrong decision; this is the case when the likelihood test gives an answer that is around the threshold; no decision can be made on that basis, and a figurative coin is tossed to decide of the value of the received signal. When all the pixels of the neighborhood have very close inten-

sities, the test will give an isotropic result. In this case the algorithm forces it back to the likely direction of the segment.

Finally, the algorithm must decide when the line ends. This is either the beginning or the end of the line, as the tracking is bidirectional. The end of the line occurs because of a certain direction it takes, or because the current pixel's intensity is lower than the lower threshold.

When both halves of the line have been tracked and reordered, it is verified that the result is not grossly wrong, by checking the length and the orientation of the line. If it is incorrect, the algorithm changes the position of the area mid-point or changes the thresholds, and the tracking is repeated. The other solution is to replace the part that is incorrect with a reference segment. This part of the algorithm is the reason why the control strategy has been called mostly irrevocable.

What follows is a short example of the parameters and conditions for a segment of one of the features of the skull, the almond shape inside the head (Figs. 22a and b). The values of the parameters are chosen empirically. The area mid-point is found on the side of the feature (cross). For segment 1:

L: length at current pixel;

dir: direction to next pixel (see Fig. 17);

dirp: direction from previous pixel;

stop if $L > 10$ and dirp = 1 or 8 and dir = 2, or if $L > 13$ and dirp = 1 or 8.

sensitivity to intersections: 1/100;

if there is an intersection, take the outermost path.

if a pixel is wrong (dirp = 7 and dir = 7; horizontal path), replace it with that of direction 1.

6.3. Results of the Knowledge-Based Line Follower

As shown on Fig. 23, the results of the algorithm depend highly on the quality of the data: these results are much better, i.e., more complete, for Lisa than for Diane; this is in agreement with the results of the quality criteria.

6.4. Determination of the Positions of the Landmarks

Once the shapes of the lines have been computed, finding the position of the landmarks is straightforward. Indeed, they are related to the lines with simple geometrical definitions. For example, the almost shape inside the head (pterygomaxillary fissure, Fig. 22) defines two landmarks, numbers 31 and 3. Their definitions are the following:

31. Pterygomaxillary fissure (anterior): most anterior point of the pterygomaxillary fissure, with respect to the plane defined by the landmarks 1 (sella) and 2 (nasion), the S–Na plane.

32. Pterygomaxillary fissure (inferior): projected inferior tip of the pterygomaxillary fissure, with respect to the S–Na plane.

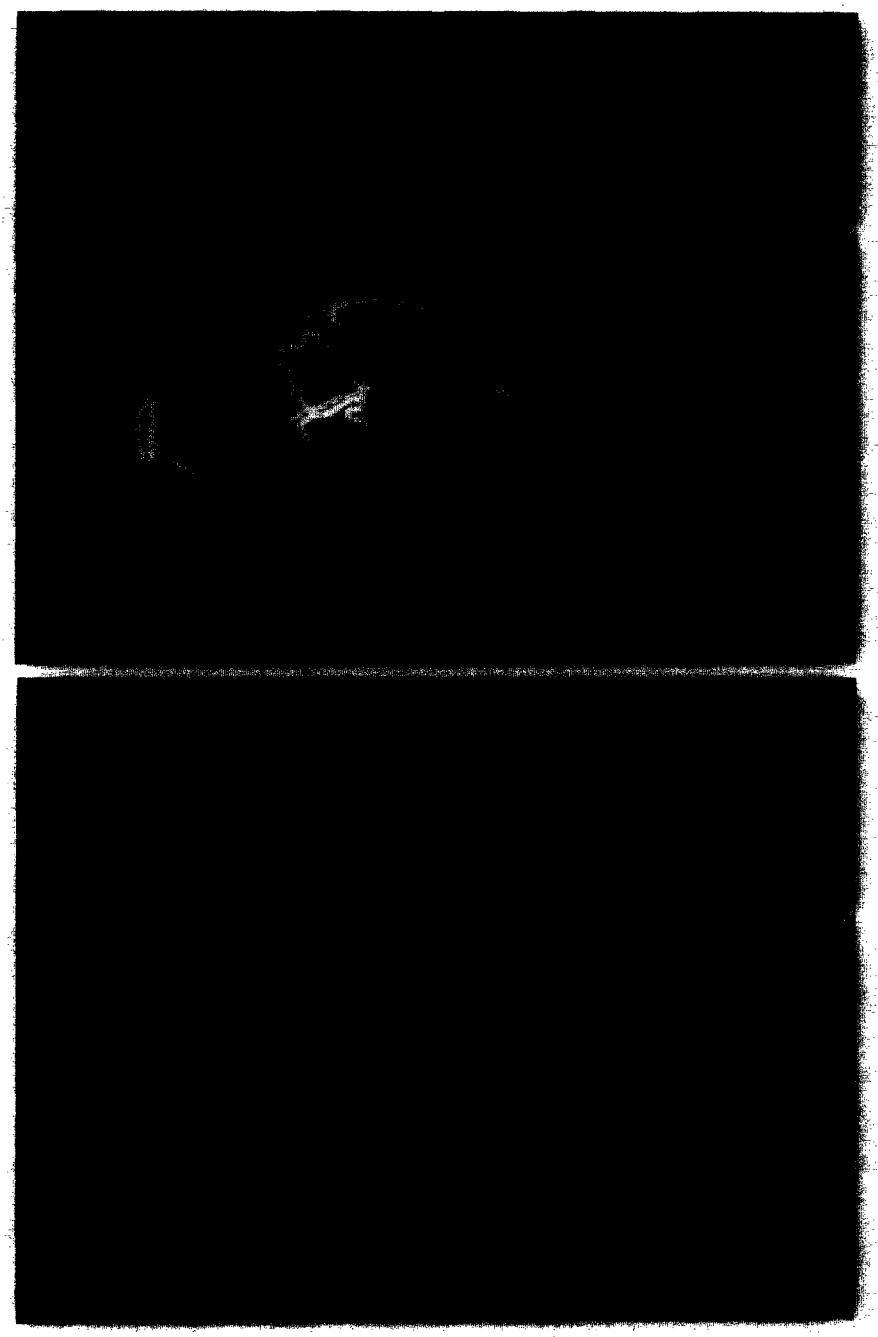

FIG. 22. Results of the knowledge-based line follower (a, b).

All the landmarks defined by lines present on the digitized version of the X ray have been found (Fig. 23).

FIG. 23. Lisa with the landmarks.

7. CONCLUSIONS

This work has shown that it is possible to automate the landmarking of cephalograms. We believe that the results of this first realization of the algorithm may depend more on the quality of the data than it is ultimably possible. Criteria have been developed to judge the performances of the algorithm as a function of the quality of the cephalograms.

As the landmarks are defined by a number of significant lines, the process is divided in three steps:

—transforming the image into a pattern of lines (line enhancement);
—retrieving the significant ones (line extraction); and
—finding the positions of the landmarks (landmarking).

The line-enhancement step is also separated in two parts: prefiltering, which prepares the image for the application of the edge detector, and the edge operator itself. For the prefiltering, a median filter was chosen due to its good noise reduction properties and its ability to preserve the contours in a satisfactory way. For the edge detector, the Mero–Vassy operator, a simplified version of the Hueckel (best fit) operator, was used.

The line extraction was performed with a knowledge-based line-tracking algorithm. This was based on a production system that uses a priori knowledge about the lines; the knowledge is divided in the following categories:

—the approximate location of the lines;

—the conditions under which it starts and stops; and
—the number of segments that constitute it, and for each segment:
  —the start and stop conditions,
  —its approximate length,
  —its general orientation, and
  —the characteristics of the noise around it.

The actual landmarking is straightforward, once the lines are found. The results of each of the steps were given. The performance of quality criteria and of the edge enhancement are satisfactory. They show that the assumption that the histograms of the patients' heads densities are bell shaped is realistic. The line extraction algorithm performs well on good-quality cephalograms. The positions of 23 out of the 36 landmarks can be determined (13 are not found because the lines that define them are not present on the digitized X rays).

In order to qualify for the name expert system, the robustness of the algorithm would have to be increased. The suggestions for future work that follow would be useful in this sense.

An early test for missing teeth, on the original X ray, would eliminate this restriction on the set of tractable radiographs.

Cavities could be detected because of their brightness; "educated guesses" could be made to replace obscured contours.

Registration of the image, i.e., normalization through geometric operations, could be implemented. This would eliminate some problems caused by small rotations or translations of the image relative to the model.

If more detailed cephalometrics were included in the algorithm, a more precise initial plan could be formed, according to age, sex, and race.

In order to deal with nonsuperimposed profiles, it could be assumed that all lines are duplicated; this assumption would have to be checked for all lines.

A multiple hypotheses scheme could be used that would choose between a number of lines the most likely candidate.

Finally, if the X ray was to be considered as one in a series, it would be possible to use the results from the X ray taken 6 months or a year earlier as the initial plan for the line follower.

REFERENCES

1. THOMSON, D'A. "On Growth and Form." Cambridge Univ. Press, 1979.
2. BAUMRIND, S., AND MILLER, D. M. Computer-aided head film analysis: The University of California San Francisco method. *Amer. J. Orthod.* **78,** 41 (1980).
3. FRIED, L. A. "Anatomy of the Head, Neck, Face and Jaw." Lea and Febiger, Philadelphia, 1980.
4. KELLY, M. D. Edge detection in pictures by computer using planning. *In* "Machine Intelligence" (B. Metzger and D. Michie, Eds.), pp. 397–409. Edinburgh Univ. Press, Edinburgh, 1970.
5. ASHKAR, G. P., AND MODESTINO, J. W. The contour extraction method with biomedical applications. *Comput. Graphics Image Processing* **7,** 331 (1970).

6. GRITTON, C. W. K., AND PARRISH, E. A. Boundary location from an initial plan: The bead chain algorithm. *IEEE Trans. Pattern Anal. Mach. Intelligence* **5,** 8 (January 1983).
7. SHIRAI, Y. Analysing intensity arrays using knowledge about scenes. *In* "The Psychology of Computer Vision" (P. H. Winston, Ed.), pp. 93–113. McGraw-Hill, New York, 1975.
8. SHIRAI, Y. Recognition of real-world objects using edge cue. *In* "Computer Vision Systems" (A. R. Hanson and E. M. Riseman, Eds.), pp. 353–362. Academic Press, New York, 1978.
9. BALLARD, D. H., BROWN, C. M., AND FELDMAN, J. A. An approach to knowledge-directed image analysis. *In* "Computer Vision Systems" (A. R. Hanson and E. M. Riseman, Eds.), pp. 271–281. Academic Press, New York, 1978.
10. RUSSELL, D. M. Where do I look now?: Modelling and inferring object locations by constraints. *In* "Proceedings of the IEEE Conference on Pattern Recognition and Image Processing," pp. 175–183, 1974.
11. LEVINE, M. D. A knowledge-based computer vision system. *In* "Computer Vision Systems" (A. R. Hanson and E. M. Riseman, Eds.), pp. 363–375. Academic Press, New York, 1978.
12. MACKWORTH, A. K. Vision research strategy: Black magic, metaphors, mechanisms, miniworlds, and maps. *In* "Computer Vision Systems" (A. R. Hanson and E. M. Riseman, Eds.), pp. 53–59. Academic Press, New York, 1978.
13. PUN, T. "Simplification Automatique de Scènes par Traitement Numérique d'Images en Vue d'une Restitution Tactile pour Handicapés de la Vue." Ph.D. Thesis, E.P.F., Lausanne, Switzerland, 1982.
14. HUECKEL, M. H. An operator which locates edges in digital pictures. *Assoc. Comput. Mach.* **18,** 113 (January 1971).
15. MERO, L., AND Z. VASSY, A simplified and fast version of the Hueckel operator for finding optimal edges in pictures. *In* "Proceedings of the 4th International Joint Conference on Artificial Intelligence, Tbilissi, Georgia, USSR, September 1975," pp. 650–655.