

Recognition of bones from X-rays of the hand†

STANLEY A. KWABWE‡, SANKAR K. PAL‡ and ROBERT A. KING§

Algorithms for the computer recognition of bones in an X-ray of the hand and wrist are described by analysing age-related changes that take place with growth in the bones. Since a straightforward structural description is inadequate because of these changes, the paper considers a shape description technique based on linear measurements (axial length, width and location) from a polygonal approximation of the bones. The recognition algorithm also needs strong *a priori* knowledge of the structure of the hand. System effectiveness is demonstrated on some sections of two typical X-rays of 10-12-year-old boys when histogram equalization, thresholding and edge detection, and octal code generation techniques are included in the preprocessing algorithm.

1. Introduction

This paper describes a method for the recognition of bones in X-rays of the hand. The results are from a project 'Identification of skeletal maturity from X-rays of the hand and wrist', under development in the digital communication section of Imperial College of Science and Technology, London.

During the growth of a child, the bones of the hand and wrist go through age-related changes; these provide invariant event sequences that occur in all individuals and cover the entire developmental age span (Tanner *et al.* 1975). Each bone develops to a reasonably constant final shape and a sequence of recognizable stages can be defined along its developmental path, providing a suitable invariant sequence. The radius, ulna, metacarpals and phalanges of the hand provide in all 28 such sequences, with events in one or another sequence occurring at almost all stages of development.

Physicians analyse X-rays of the left hand and use the event sequences resulting from bone growth as a measure of maturity and age. (A sketch of the bones of the left hand is given in Fig. 1.) To measure maturity, the evidence from these sequences is combined using a variety of rules, weights and scoring methods. Each analysis is time-consuming and complicated.

The use of a computer for the analysis and determination of maturity is thus facilitated by the existence of these evaluation rules and has the advantages of allowing larger numbers of radiographs to be analysed while reducing the subjectivity of the analysis. There is very little variability between X-rays as a result of the standardization of radiographic techniques.

The three tasks associated with this problem are:

- (1) study of the radiograph and recognition of the specific bones and their location;

Received 24 November 1983.

† This work has been supported by the Medical Research Council, U.K.

‡ Electronics and Communication Sciences Unit, Indian Statistical Institute, 203 Barrackpore Trunk Road, Calcutta 700035, India. On leave with the Electrical Engineering Department, Imperial College, London SW7, England.

§ Electrical Engineering Department, Imperial College, London SW7, England.

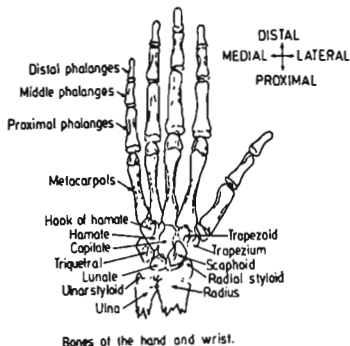


Figure 1. Sketch of hand bones (from Tanner *et al.* 1975).

- (2) processing of images of different bones and extracting their primitives for description;
- (3) determining the stage of development for each bone and combining the evidence from all the bones to arrive at a predicted age.

We are concerned here with the first task. The results on the second task have already been reported (Pal and King 1983, Pal *et al.* 1983 a). The last part is a subject of further research work, and will be reported in due course.

A simple and computationally cheap recognition method is presented to identify each bone and store its location coordinates. The algorithm is immune to changes in the orientation of the hand and the shape changes that occur during growth. Recent workers on similar problems have considered approaches based on analyses using transform or spatial domain techniques (Pavlidis 1978). Both statistical and syntactic (structural) pattern recognizers have been used, the former being associated with transform domain techniques, while the latter is associated with spatial domain techniques.

Young (1972) measured certain parameters of a cell image and used them to classify blood cells. Davis and Anderson (1981) used a combination of hierarchical constraint processes and syntactic pattern analysis to classify airplane silhouettes. Ledley and Ruddle (1966) used shape description and certain other parameters to detect chromosome abnormalities. Others have used graph related approaches (Shaw 1969, Moayer and Fu 1976). Pavlidis (1968) used length and width of strokes to recognize typewritten characters. Automata have been used to recognize the features of extensively described contours (Davis and Anderson 1981, Gonzalez and Thomason 1978).

In all these cases, the shape to be recognized was generally constant; however, bones go through drastic shape changes with age so that a detailed shape description is not useful for more than one age. One relatively invariant feature in the radiograph is the overall bone structure in the hand.

We have used shape description based on simple length and width measurements

together with some information on the structural location of the bones within the scene. Two basic pattern classes are defined: 'long and thin' and 'short and fat'. The classification is mainly based on these two classes. To differentiate between bones in the same class, each has a label corresponding to its structural location in the scene. Hierarchical ordering is used to identify bones with a similar structure, leading to an overall tree structure. Preprocessing algorithms include histogram equalization, thresholding and edge detection, chain code generation and polygonal approximation. The feature extraction algorithm, on the other hand, extracts three features, namely axial length, width and the location of the bones which, in turn, become the input to the classifier.

The block diagram of such a recognition system is shown in Fig. 2. Algorithms used in the different blocks are explained in the following sections. The performance of the system is demonstrated when different sections of X-ray images of the hand and wrist of 10-12-year-old boys are considered as input.

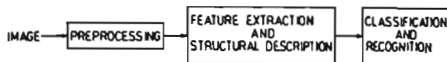


Figure 2. Block diagram of recognition system.

2. Preprocessing

Preprocessing involves several stages to produce a much reduced image from which features are extracted. A polygonal approximation is obtained from the edge-detected image, and recognition algorithms are thus applied to an image with fewer points.

2.1. Histogram equalization (Gonzalez and Wintz 1977, Pal and King 1981)

This well-established technique extends the dynamic range of the image, improving the contrast between low-level and high-level regions, as well as helping to standardize the dynamic range for a large number of radiographs.

2.2. Thresholding and edge detection

The contrast intensification resulting from histogram equalization techniques makes it easier to intensify the contrast of the image by defining its thresholds. Thresholds are selected automatically without reference to the histograms) using the 'index of fuzziness' and 'entropy' of an image (De Luca and Termini 1972, Pal 1982, Pal *et al.* 1983 b).

Having determined the threshold corresponding to a bone-flesh transition, it is used to produce a two-tone version of the image for the following edge-detection operation. Edge detection is achieved using 'max' or 'min' rules. The operator

$$g^* = |g - \max \{g_i\}| \quad (1)$$

or

$$g^* = |g - \min \{g_i\}| \quad (2)$$

(where g is the old intensity of the current pixel, g^* the new intensity of the current pixel, and g_i ($i = 1, 2, 3, 4$) the intensity of the four nearest neighbours to the current pixel), acts on a window covering the four nearest neighbours.

2.3 Chain-code generation and polygonal approximation

A chain-code description is generated in terms of eight basic primitive directions, shown in Fig. 3. An edge follower moves along each of the edge-detected bones and evaluates the direction code according to the above primitives over every 10 pixels; coordinates associated with each direction change are stored to form the polygon approximation. Slight direction variations within 22.5° of a code are taken as no change. Sharp changes in less than 10 pixels are, however, detected. The chain-code description used to generate the polygonal approximation helps in smoothing it, and in determining the orientation of the bone.

In smoothing, very short links are removed and neighbouring links in the same direction are replaced by one link so as to reduce the number of points in the polygonal approximation. Thus the $M \times N$ image is replaced by a string of n codes and coordinates that have all the information required for recognition. Figure 4(a) shows such a polygonal approximation of a finger bone. Figures 4(b) and (c) show its code description starting from 'a', and the number of pixels (length) corresponding to each code respectively.

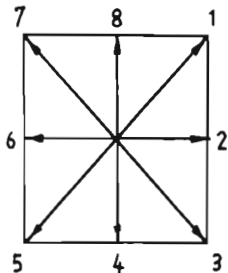


Figure 3. Eight primitive directions.

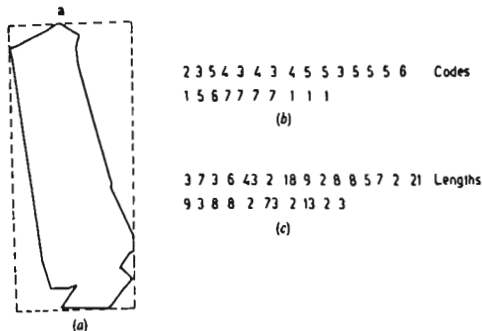


Figure 4. (a) Polygonal approximation. (b) Code descriptions. (c) Lengths of codes.

3. Feature extraction and structural description

The bones of the hand may be categorized into (a) long and thin, to include the phalanges, metacarpals, radius and ulna, and (b) short and fat, to include the carpals and most of the epiphyses of the long bones during skeletal development. Consequently, 'short and fat' bones have a ratio of length to width r less than or equal to 1, while the 'long and thin' bones have $r > 1$.

The second aspect of feature extraction is to determine whether the feature under consideration is a single bone or two or more fused or overlapping bones. The system is immune to the fusing that could occur naturally, because this is included in the structural information; overlaps may be detected by studying the few codes in the smoothed string description to find direction changes caused by them. The string description of a single bone will tend to progress smoothly so as to give a closed figure with a total angle change of 360° . Sharp reverse changes in these codes indicate direction changes caused by overlap and may be counted for each bone.

Three features are necessary for the recognition of each bone:

- (1) its axial length
- (2) its width
- (3) its location in the hand.

To determine the axial length of the bone from its polygonal approximation, the following algorithm is employed.

Step 1. The orientation of the bone is determined by considering the histogram of edge directions. The algorithm used to determine the orientation of the bones may be illustrated by considering the polygonal approximation of the middle phalanges of the first finger as shown in Fig. 4. To determine the orientation, the number of pixels for each code are added up and a histogram is obtained showing each code and the number of pixels in the polygonal approximation that belong to this code. The codes are divided into two groups. The codes 1-4 are in one group and the codes 5-8 are in the other group, and these two groups are complementary. The mode in each group gives the dominant direction on each side of the bone, and hence the orientation.

Step 2. The orientation of the lines linking every point with every other point in the approximation are determined.

Step 3. The link which is longest and also has the orientation nearest to that of the bone is taken as the axis link, and its length is the bone length.

To determine the width, the following algorithm is used.

Step 1. Determine the lengths of those links that are nearly at 90° to the axis.

Step 2. The longest link that is nearest to normal to the axis gives the width of the bone.

The coordinates of the polygonal approximation are studied, and for each bone the extreme points are noted so as to obtain the 'box' that contains the bones (Fig. 4(e)). This is stored as a bone feature and will be the input to the next stage to give the location of the bone. The position in the structure of the hand is determined when all the bones in the scene have been represented by their polygonal approximations and their position relative to other bones has been examined.

A tree structure is used for an overall description of the hand, as shown in Fig. 5. Using this tree structure, a determination is first made as to which of the three main regions the bone belongs. The fingers and thumb are differentiated from the rest because they have bones stacked in a row and there are five such stacks. The thumb has one 'long' bone less than the fingers so it is identified and used to orient the hand. Other fingers are counted off from this.

The middle hand region contains circular bones that are identified by the fingers nearest to them. Each bone has a structural label—a typical description would be '2nd bone along middle finger'. At this stage the three regions have been differentiated and the separate fingers identified. This is followed by identification of individual bones.

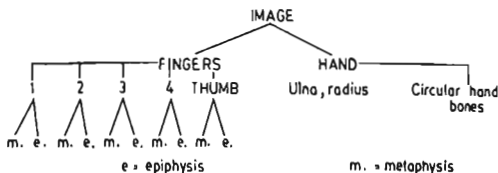


Figure 5. Tree structure.

4. Recognition

The final recognition algorithm applies both the structural description and class to obtain the last 'twig' of the tree structure. It is at this point that most *a priori* knowledge of the structure is applied, and descriptions of each finger and the other hand regions are required.

A normal finger bone will have between 4 and 8 bones depending on the age. In the very young all the bones are separate, whereas the metaphyses and epiphyses fuse in later ages. The range of possibilities for finger bones is shown in Fig. 6. The recognition algorithm for a finger bone is shown in Fig. 7.

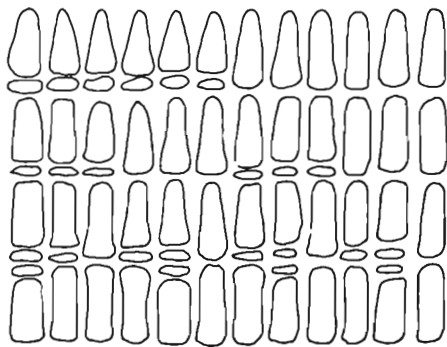


Figure 6. Range of possibilities for finger bones.

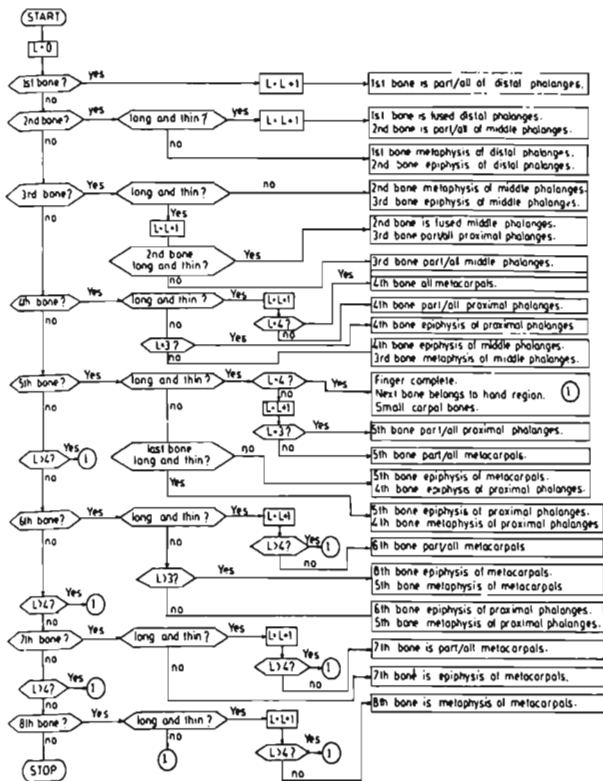


Figure 7. Recognition algorithm.

5. Results

The X-ray shown in Fig. 1 was digitized using a dynamic range of 0–255 levels and stored on magnetic tape in a 1024×2048 pixel matrix. For the results here, smaller image sections of varying dimensions were extracted from the tape and stored on disk in the CDC6500 computer that was used as the processing system. The test images were selected to provide the whole range of conditions the recognition system might encounter. Figure 8 shows two such fingers (first and third) which contain both classes of bone and various overlap/fuse conditions. Each of the images was processed as discussed in § 2 to produce their respective polygonal approxi-

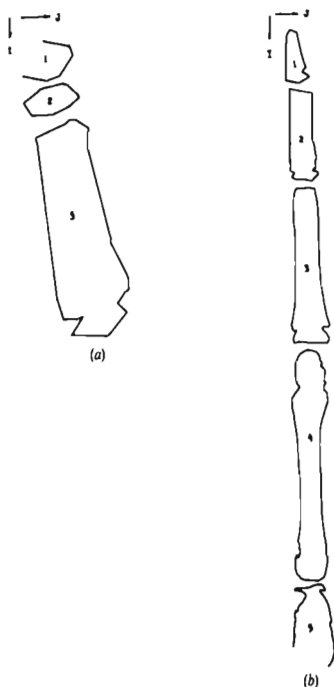


Figure 8. (a) First finger. (b) Third finger.

mations. The third finger, for example, initially occupied an 838×150 pixel matrix, but after polygonal approximation it required only 320 points to describe all the bones.

The recognition system looks at each of the bones and uses the structural information to determine their order along the finger. The features of each bone (position, length, width, class) are determined and printed. The system also determines whether the whole bone is present, or just part of the bone, by following the contour and determining if the starting and end points (coordinates) of the description are the same. A provisional identification is made using the flowchart shown in Fig. 7, which is applicable to the finger bones. A similar algorithm is derived for the circular bones in the carpal region; this is a relatively simpler task, but in the context of application this region was not of interest. In many cases (first bone in Fig. 8 (a) and last bone in Fig. 8 (b)) an identification can still be made even if the bone under consideration is incomplete. The results of the recognition system for these two fingers are shown in Tables 1 and 2 (Kwabwe 1981).

-
- (i) BONE 1 ENCLOSED BY I=1 TO I=20 AND J=1 TO J=27
SMOOTHED OCTAL CODED CHAIN IS 2 3 5 6 8
LENGTH OF BONE IS 21.0950, WIDTH OF BONE IS 27.8568
CLASS: SHORT AND FAT
THIS IS AN INCOMPLETE BONE
- (ii) BONE 2 ENCLOSED BY I=20 TO I=35 AND J=1 TO J=31
SMOOTHED OCTAL CODED CHAIN IS 1 3 5 6 7
LENGTH OF BONE IS 16.1245, WIDTH OF BONE IS 29.0689
CLASS: SHORT AND FAT
THIS IS A COMPLETE BONE
THEREFORE, 2ND BONE IS EPIPHYSIS OF DISTAL PHALANGES AND 1ST BONE IS THE METAPHYSIS
- (iii) BONE 3 ENCLOSED BY I=37 TO I=138 AND J=7 TO J=56
SMOOTHED OCTAL CODED CHAIN IS 2 3 4
3 4 5 3 5 6 1 6 7 8 1
LENGTH OF BONE IS 101.6071, WIDTH OF BONE IS 27.4591
CLASS: LONG AND THIN
THIS IS A COMPLETE BONE
THEREFORE 3RD BONE PART OR ALL OF MIDDLE PHALANGES

TIME USED IS=0.05700

1.034 CP SECONDS EXECUTION TIME

Table 1. Results for finger 1 in Fig. 8 (a).

Table 2.

-
- (i) BONE 1 ENCLOSED BY I=13 TO I=97 AND J=64 TO J=107
SMOOTHED OCTAL CODED CHAIN IS 3 5 3
5 6 8 1
LENGTH OF BONE IS 84.1724, WIDTH OF BONE IS 42.5793
CLASS: LONG AND THIN
THIS IS A COMPLETE BONE
- (ii) BONE 2 ENCLOSED BY I=100 TO I=246 AND J=69 TO J=125
SMOOTHED OCTAL CODED CHAIN IS
8 7 1 7 8 7 8 2 4 3 4 3 5 4 5 2 5
LENGTH OF BONE IS 146.2942, WIDTH OF BONE IS 50.0000
CLASS: LONG AND THIN

THIS IS A COMPLETE BONE
 THEREFORE 1ST BONE IS FUSED DISTAL
 PHALANGES AND 2ND BONE IS PART/ALL
 MIDDLE PHALANGES

(iii) BONE 3 ENCLOSED BY I=255 TO I=492 AND
 J=73 TO J=142
 SMOOTHED OCTAL CODED CHAIN IS 1 8 1 7
 8 2 3 4 3 4 3 4 3 5 6 3 5 6 2 6
 LENGTH OF BONE IS 237.8991, WIDTH OF BONE IS
 57.4282
 CLASS: LONG AND THIN

THIS IS A COMPLETE BONE
 THEREFORE 2ND BONE IS FUSED MIDDLE
 PHALANGES AND 3RD BONE IS PART/ALL
 PROXIMAL PHALANGES

(iv) BONE 4 ENCLOSED BY I=502 TO I=848 AND
 J=17 TO J=77
 SMOOTHED OCTAL CODED CHAIN IS 8 1 7
 8 1 2 3 4 3 5 3 5 3 4 5 4 4 5 4 3 4 7 4 3 2 1
 8 7 8 7 8 7 8 7
 CLASS: LONG AND THIN
 LENGTH OF BONE IS 283.6725, WIDTH OF BONE IS
 109.1146

THIS IS A COMPLETE BONE
 THEREFORE 4TH BONE IS PART/ALL META-
 CARPALS

(v) BONE 5 ENCLOSED BY I=943 TO I=983 AND
 J=73 TO J=90
 SMOOTHED OCTAL CODED CHAIN IS 1 8 1 7
 1 8 7 2 7 6 5 6 3 5 4 5 3 4 8 4
 LENGTH OF BONE IS 116.6062, WIDTH OF BONE IS
 44.8219
 CLASS: LONG AND THIN

THIS IS AN INCOMPLETE BONE
 THEREFORE 5TH BONE IS A SMALL CARPAL
 BONE

TIME USED IS=0.06200

2.003 CP SECONDS EXECUTION TIME

Table 2. Results for finger 3 in Fig. 8 (b).

6. Conclusions

The emphasis in this work was on identification of the long finger bones with epiphyses where relevant, and the radius and ulna. The carpals are not of significance in assessing skeletal maturity but may be dealt with in a similar manner.

Histogram equalization, grey-level thresholding using an index of fuzziness or fuzzy entropy, contour extraction using 'max' or 'min' rules, and encoding of the contours through octal codes, are found to be suitable in order to extract the features for final recognition. Although results corresponding only to phalanges are presented here, the algorithm can be easily and automatically extended to other regions

The smoothed polygonal approximation, together with *a priori* knowledge of the structure of the hand, forms an 'index' to the X-ray, and specific bones can be extracted as required.

ACKNOWLEDGMENTS

Provision of X-ray films by Professor J. M. Tanner and the valuable cooperation of Dr. A. A. Hashim are gratefully acknowledged.

REFERENCES

- DAVIS, L. S., and ANDERSON, T. C., 1981, *I.E.E.E. Trans. Patt. Anal. Mach. Intell.*, **3**, 265.
DE LUCA, A., and TERMINI, S., 1972, *Inf. Control*, **20**, 301.
GONZALEZ, R., and THOMASON, M., 1978, *Syntactic Pattern Recognition: an Introduction*, (Reading, Mass.: Addison-Wesley).
GONZALEZ, R. C., and WINTZ, P., 1977, *Digital Image Processing* (Reading, Mass.: Addison-Wesley).
KWABWE, S. A., 1981, An analysis of X-rays to locate the sites of specific bones, M.Sc. thesis, Imperial College, London.
LEDLEY, R. S., and RUDDLE, F. H., 1966, *Scient. Am.*, **214**, 40.
MOAYER, M., and FU, K. S., 1976, *I.E.E.E. Trans. Comput.*, **25**, 262.
PAL, S. K., 1982, *I.E.E.E. Trans. Patt. Anal. Mach. Intell.*, **4**, 204.
PAL, S. K., and KING, R. A., 1981, *Electron. Lett.*, **17**, 862; 1983, *I.E.E.E. Trans. Patt. Anal. Mach. Intell.*, **5**, 69.
PAL, S. K., KING, R. A., and HASHIM, A. A., 1983 a, *I.E.E.E. Trans. Syst., Man Cybernet.*, **13**, 94; 1983 b, *Patt. Recog. Lett.*, **1**, 141.
PAVLIDIS, T., 1968, *Inf. Control*, **12**, 526; 1978, *Comput. Graph. Image Process.*, **7**, 243.
SHAW, A. C., 1969, *Inf. Control*, **14**, 9.
TANNER, J. M., WHITEHOUSE, R. H., MARSHALL, W. A., HEALEY, M., and GOLDSTEIN, H., 1975, *Assessment of Skeletal Maturity and Prediction of Adult Height (T.W.2 Method)* (London: Academic Press).
YOUNG, I. T., 1972, *I.E.E.E. Trans. biomed. Eng.*, **19**, 291.