# Journal of Digital Imaging

VOL 14, NO 4 DECEMBER 2001


## Computer Automated Approach to the Extraction of Epiphyseal Regions in Hand Radiographs


By Ewa Pietka, Sylwia Pośpiech, Arkadiusz Gertych, Fei Cao, H.K. Huang, and V. Gilsanz



**Epiphyseal region is the most sensitive region to developmental changes of the skeletal system. Extraction of this area is the very first step in any computerized image analysis. In this report a fully automated analysis of a hand radiograph resulting in extraction of distal and middle regions of the II, III, and IV phalanx is presented. The processing is performed in 3 stages. First, the trend of background is removed from radiograph to obtain a binary hand mask. At this stage a labeling procedure is necessary to eliminate artifacts (markers). Then, II, III, and IV phalanges are identified in the binary image, and the phalangeal axes are drawn. Finally, the intensity profile along each phalangeal axis is analyzed, and, on its basis, distal and middle regions are located. The presented procedure is designed as a part of currently developed system for automatic bone age assessment; however, it also can be as a preprocessing step in other diseases the diagnoses of which may require a computer assistance.**





**KEY WORDS: skeletal maturity, bone age, computer assessment.**


STAGE OF EPIPHYSEAL development is the most relevant bony structure considered in an assessment of skeletal maturity. A left hand wrist radiograph is compared with patterns reflecting the bone age. A difference between the bone age and the chronological age indicates abnormalities in the skeletal development. The procedure also is used in monitoring the hormone therapy. Two methods are implemented clinically. One, very subjective, is based on a book atlas[1] and compares the entire hand image with patterns. Another one, more objective yet very time consuming, is the Tanner-Whitehouse (TW) method,[2] which compares selected regions of interest.

Several attempts have been made to computerize the epiphyseal analysis. One approach[3] uses a cross correlation as a measure of similarity between an analyzed region and a prototype. Yet, before the analysis, the radiologist has to mark manually the epiphysis and corresponding bones. Another approach[4] performs the analysis of the middle region of the III phalanx. The method is based on a point distribution model derived by several points marked manually on the bones. This model then is compared with pattern models describing each developmental stage. Investigators have mentioned an unsuccessful attempt to automatic location of regions.

Tanner and Gibbons[5] have developed a method based on the TW approach. The epiphyseal region is positioned beneath a video camera. An operator views the image on the computer screen and aligns it to the templates of the TW stage of development displayed on the screen. The closest match is chosen by comparing the Fourier transform coefficients of the region under analysis to the template regions. Another approach based on the TW method[6] performs an interactive image analysis. Patterns of all developmental stages are displayed on a screen, and the calculation is done automatically, yet still the regions are extracted manually.

All described methods require a manual extraction of the region of interest. In this study, a computerized approach to the extraction of epiphyseal regions of interest is presented. It is an extension of previous studies performed on the III phalanx only with no rotation of phalanges.[7] Currently, the II, III, and IV phalanx are analyzed, and the edges of extracted regions are aligned to the


*From the Silesian University of Technology, Institute of Electronics, Division of Biomedical Electronics, Gliwice, Poland, and Children's Hospital of Los Angeles, Department of Radiology, Los Angeles, CA.*

*Address reprint requests to Vicente Gilsanz, MD, Children's Hospital Los Angeles, Radiology Department, MS# 81, 4650 Sunset Blvd, Los Angeles, CA 90027.*





*0897-1889/01/1404-0006$35.00/0*

phalangeal axis. Because the procedure is intended to be used in unsupervised applications, error checking is introduced at various stage of the image analysis. A standard hand position in the image is assumed (ie, a left-hand wrist in the upright position).

The image analysis is performed in 3 phases. First, the entire hand area is extracted. Then, the phalangeal axes are marked. Finally, the epiphyseal regions are extracted.

## HAND LOCATION

The basic step of our analysis is to determine the location of the hand in the image. This aim is gained in 2 stages. First, the image is transformed into a binary mask, and simultaneously the image background is removed. Then, high intensity artifacts (labels, markers, tubes) are separated from the hand region.

To shorten the execution time at the preprocessing stage without the influencing the overall analysis, the image has been subsampled by a factor of 4. This image size often is used in radiology for the image preview.

### *Background Removal*

The background extraction is based on a histogram analysis. An overall image histogram of an input image indicates 2 areas (Fig 1). The low gray value peak corresponds to the background, whereas the high gray scale value peak reflects the soft and bony tissue. Bright markers and dark artifacts may appear as small additional peaks in peripheral high or low gray level areas, respectively. Because of the background nonuniformity, in most hand images the background yields a rather slow varying slope making it difficult or even impossible to correctly separate the hand area using a single threshold value. Thus, to estimate the background plane, the image is divided into 4 quarters, and a histogram of each area (called later a subhistogram) is searched for a peak locating the background (Fig 2). Because the ratio of the background peak to hand region peak amplitudes depends on the hand location within the image, the interval to be searched is delimited by the shape and location of the background peak in the main histogram.

The locations of background peaks: $h_{max1}$, $h_{max2}$, $h_{max3}$, $h_{max4}$, (Fig 3A) correspond to the average intensity level of background pixels $k_1$, $k_2$, $k_3$, $k_4$ in every quarter of the image. These 4 values are the basis for estimation of background in each image corner $bcg_1$, $bcg_2$, $bcg_3$, $bcg_4$. The bilinear interpolation is used.[8] The interpolated image $\hat{f}(x, y)$ is calculated by a linear combination of 4 nearest

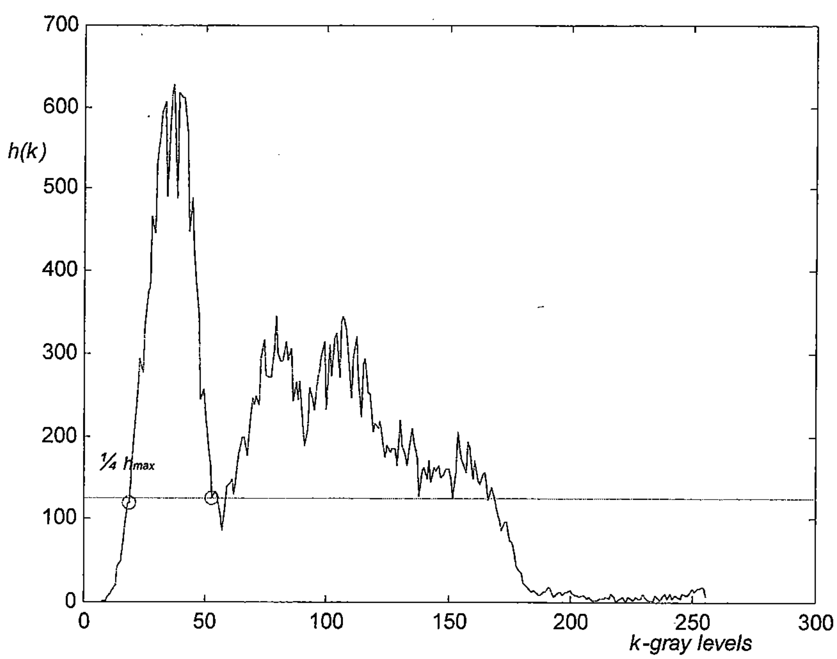


**Fig 1. The histogram of an input image.**

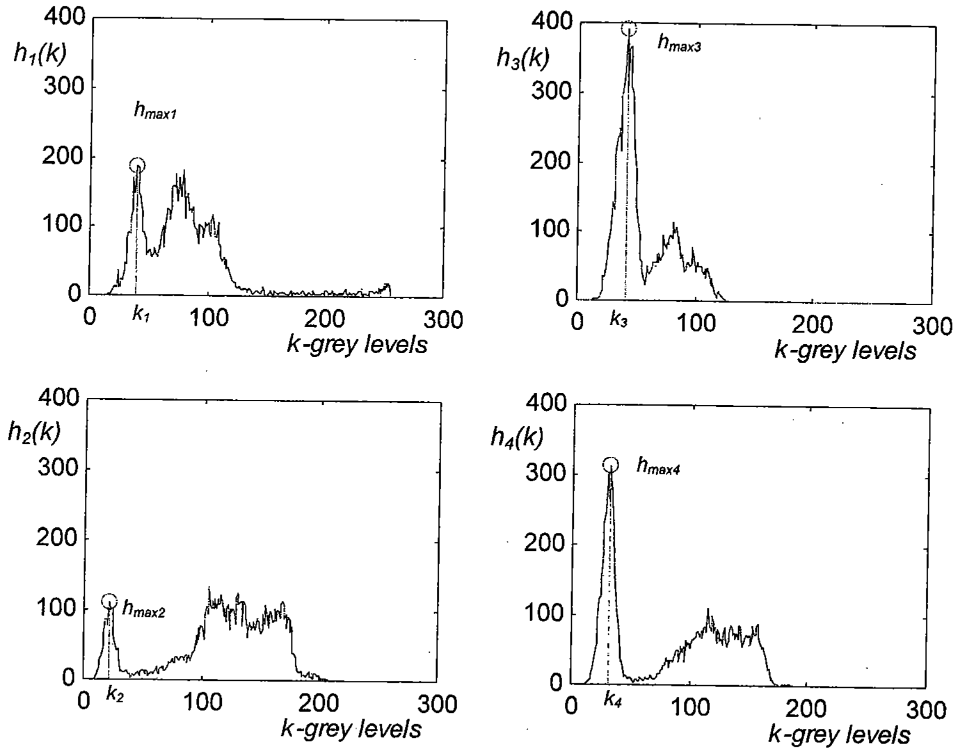


**Fig 2. Subhistograms of the input image.**

pixels $bcg_1$, $bcg_2$, $bcg_3$, $bcg_4$ of a given 2-dimensional data (Fig 3B). The interpolated image $\hat{f}(x,y)$ for $n_1T_1 \leq x \leq (n_1 + 1)T_1$ and $n_2T_2 \leq y \leq (n_2 + 1)T_1$ is given by:

$$
\begin{aligned}
f(x,y) = {} & (1 - Dx)(1 - Dy)f(n_1T_1,n_2T_2) \\
& + (1 - Dx)Dyf(n_1T_1,[n_2 + 1]T_2) \\
& + Dx(1 - Dy)f([n_1 + 1]T_1,n_2) \\
& + DxDyf([n_1 + 1]T_1,[n_2 + 1]T_2).
\end{aligned} \tag{1}
$$

where: $Dy = (y - n_2T_2)/T_2$, $Dx = (x - n_1T_1)/T_1$ The estimated background (Fig 4A) is subtracted from the hand image. The results are shown in (Fig 4B).

### *Hand Object Extraction*

The second stage of the image analysis yields a binary hand mask. After the background has been subtracted from the hand image, a histogram is found again. Because the background is now more

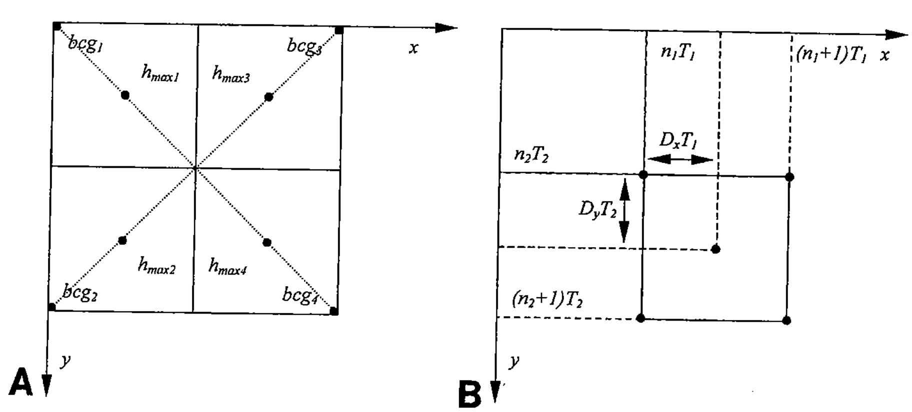


**Fig 3. Background analysis. (A) location of background peaks. (B) Estimation of the background plane.**

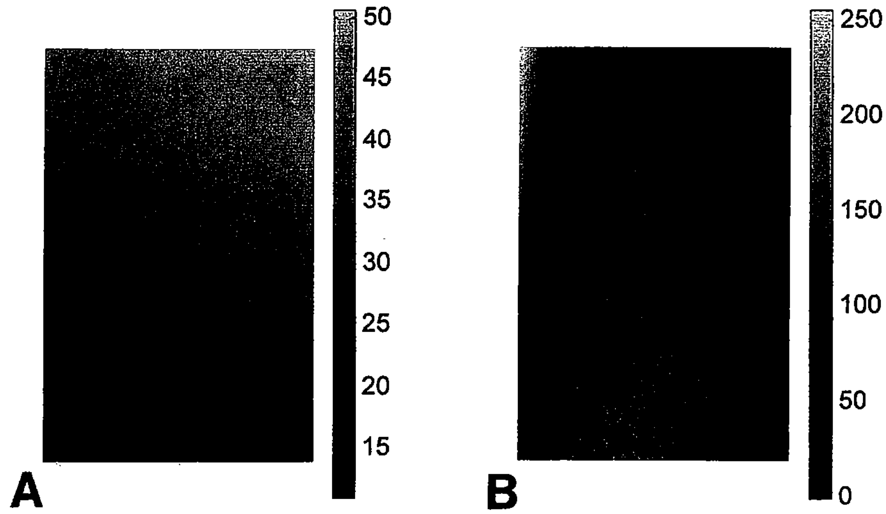


**Fig 4. Background removal. (A) Estimated image background. (B) Hand image after the background substraction.**

uniform, its peak is narrow and has a larger value than the background. To separate this peak, the triangle algorithm is applied. A line connecting the maximum value at the gray level $k_m$ and the value at the level $k_{max}$ is constructed (Fig 5A). The distance *d(k)* between the line and histogram points *h(k)* is computed for all gray levels from $k_m$ to $k_{max}$. The gray level value $k_t$ at the maximum distance

$$d(k_1) = max\ |h(k_i) - l(k_i)| \qquad (2)$$

yields the threshold value (Fig 5B).

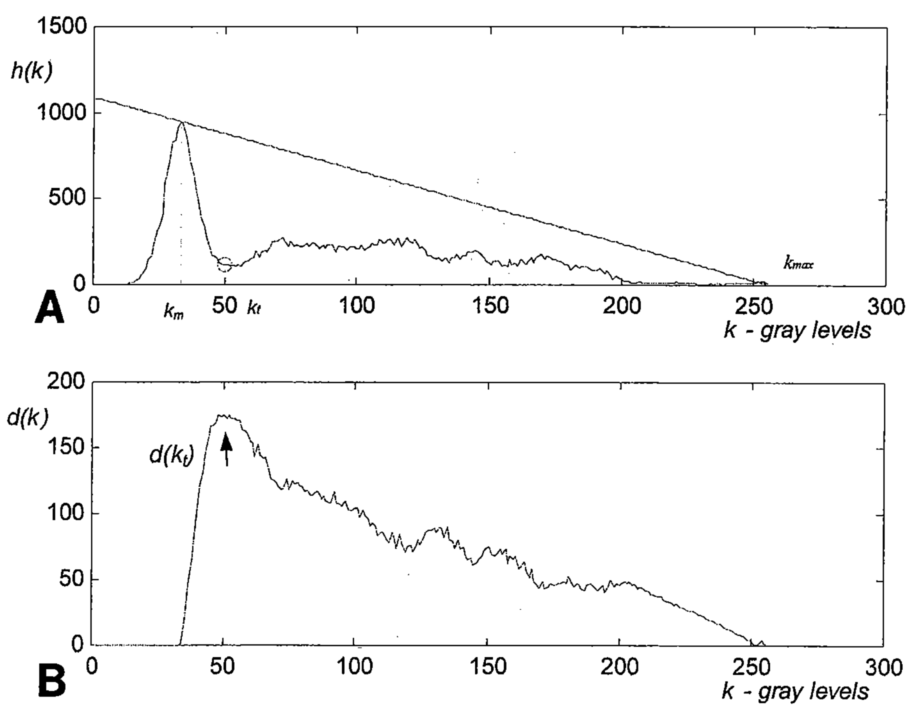


**Fig 5. Selection of the threshold value after background removal. (A) Line connecting the maximum value at the gray level $k_m$ and the value at the level $k_{max}$. (B) Thresholding value $d(k_t)$.**

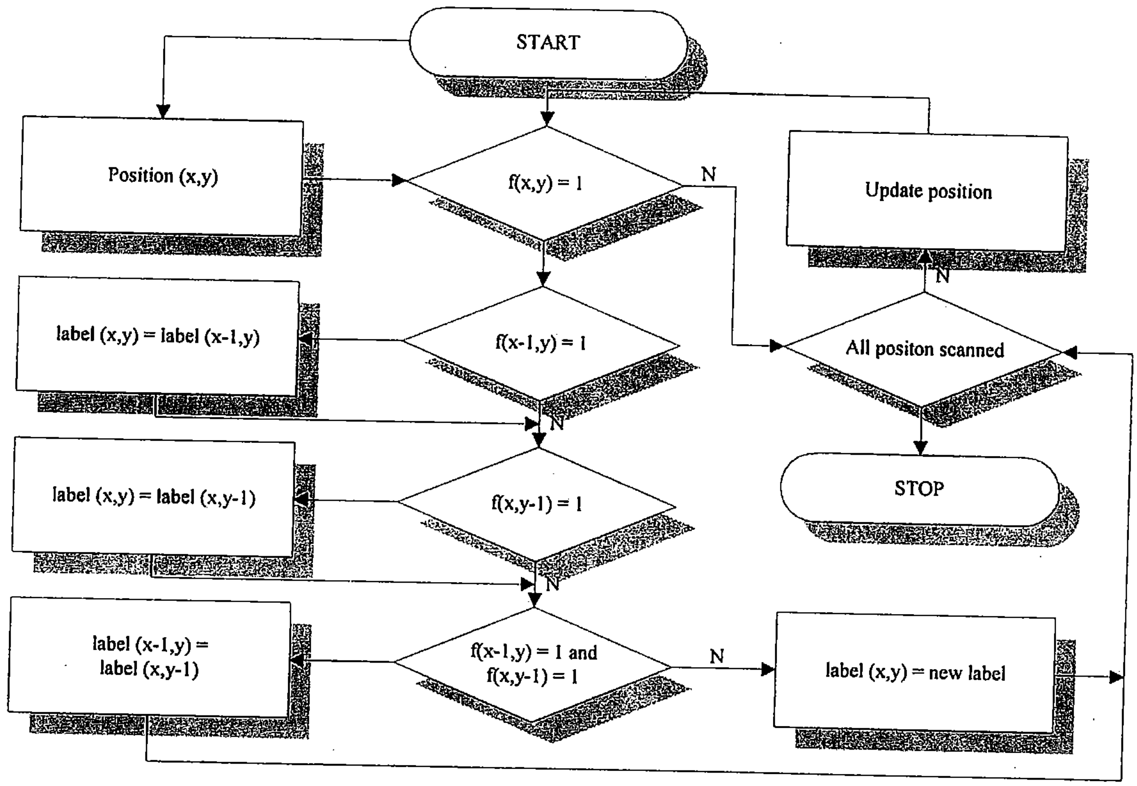


**Fig 6. The labeling procedure.**

After implementing a thresholding function, a binary image is obtained. Low pixel values correspond to the background, and high pixel values mark the hand area and other objects (if any). Morphologic operators remove isolated pixels, which increase the number of binary objects to be removed. Afterward a labeling procedure [9] is performed. Its goal is to find the greatest object (ie, hand area) in binary image (Fig 6). The binary image is scanned from bottom to top with 3 pixels window. If the pixel under consideration is equal to 1, it is compared with its 2 neighbors. If any of them belongs to the object, the same label is granted, otherwise a new label of a consecutive integer is assigned. The area of each object is found, and the largest value indicates the hand region. A more detailed description of labeling algorithm can be found in Haralick and Shapiro.[10]

## LOCATION OF PHALANGEAL AXES

After the hand background has been removed and the hand area has been extracted, the analysis of the phalanges may be performed. It starts with the location of the II, III, and IV, phalanx followed by the extraction of their axes.

### *Location of Phalanxes*

The thresholded image is scanned in the horizontal direction (Fig 7A) and an array of horizontal profiles is defined. To reduce the noise, each profile is an average of neighbor rows. Because a binary image is subjected to the image analysis, profiles contain high and low value levels. The low value level (LVL) reflects the background, whereas the hand location is marked by the high value level (HVL).

The array of horizontal profiles is searched for 2 fiducial rows (Fig 7A). First, the upper row marked by a profile of 3 to 5 HVLs is found. Then, it is shifted upward until only one HVL remains. It delimits the top of the phalangeal area and marks the tip of the III phalanx (and the first fiducial row). Two lower profiles containing HVL to the right and left of the III phalanx locate tips the II and IV phalanx, respectively. The second fiducial row is found by scanning the array downward until only 1

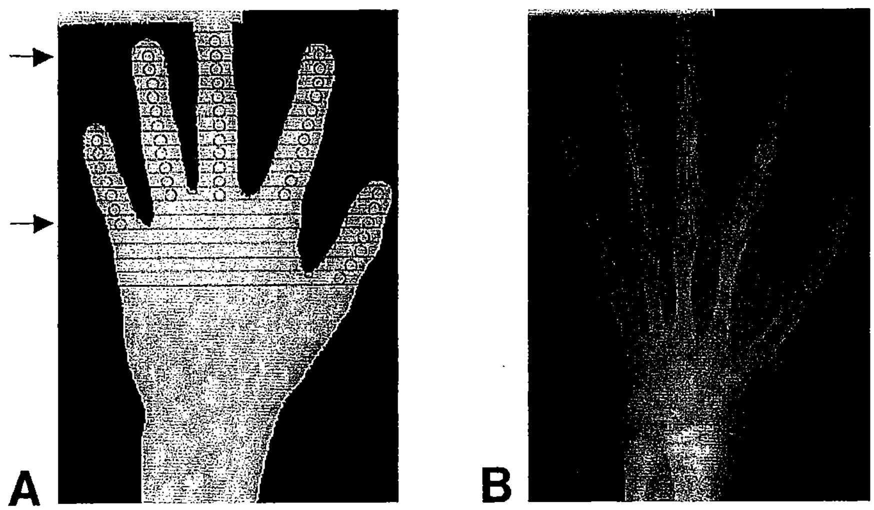


**Fig 7.** **(A) The thresholded hand image with profiles in phalangeal area, arrows point to fiducial rows, circles mark centers of high intensity sectors and are used for estimation of phalangeal axes. (B) original hand image with superimposed axes of the phalanges II, III, IV.**

HVL within the profile remains. It reflects the metacarpal area and delimits the lower edge of the phalangeal region to be analyzed.

*Extraction of Phalangeal Axes*

The extraction of phalangeal axes is performed by searching the array of profiles in the vertical direction. The central point of each HVL marks an axial point of the corresponding phalanx. As a result, 3 vectors are defined. Each of them contains the coordinates (ie, distance from the left image edge) of the axial points of one phalanx. These vectors are subjected to an interpolation procedure. Because fingers of most people exhibit a natural curvature, the third-order polynomial fit is used. (Figure 7B) shows axes of II, III, and IV phalanx superimposed on the hand image.

## LOCATION OF EPIPHYSEAL REGIONS

Location of epiphyseal regions of interest is performed on the image at an original resolution. The regions include epiphyses, lower part of metaphyses, and upper part of diaphyses.

To extract the epiphyseal region of interest (EROI) a modified version of an already developed, for a middle phalanx, procedure is applied. In the earlier version[7] a vector P representing intensity changes along phalanx axis has been scanned in a vertical direction. In the current function the axis is no longer assumed to be vertical, yet it is at any direction. Thus, elements of P are defined as an average of neighbor pixels perpendicular to the local direction of the axis (Fig 8):

$$P_j = \sum_{k=-n}^{k=n} x_{r_j + (c_j - c_j + k), c_j + k} \qquad (3)$$

where $p_j$ is the j-th element of the profile vector P, $r_j$, $c_j$ are row and column number of an j-th pixel of the axis, $x_{k,l}$ is intensity of pixel in location k, l, and $n$ is size of the neighborhood.

The first-order derivative (Fig 9) of the vector P is searched for 3 minima, which mark roughly the gap between epiphyses and metaphyses. To avoid 2 neighbor gaps within an area (the gap between diaphyses and epiphyses and between epiphyses and metaphyses) a sector that can include one minimum only is defined. This insensitivity sector

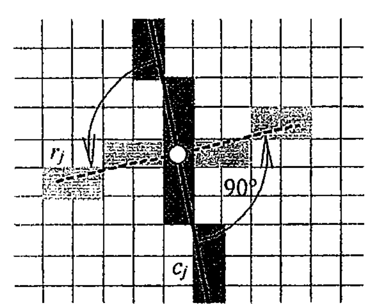


**Fig 8.** **Calculation of the profile along the axis as average of pixels in direction perpendicular to the axis.**

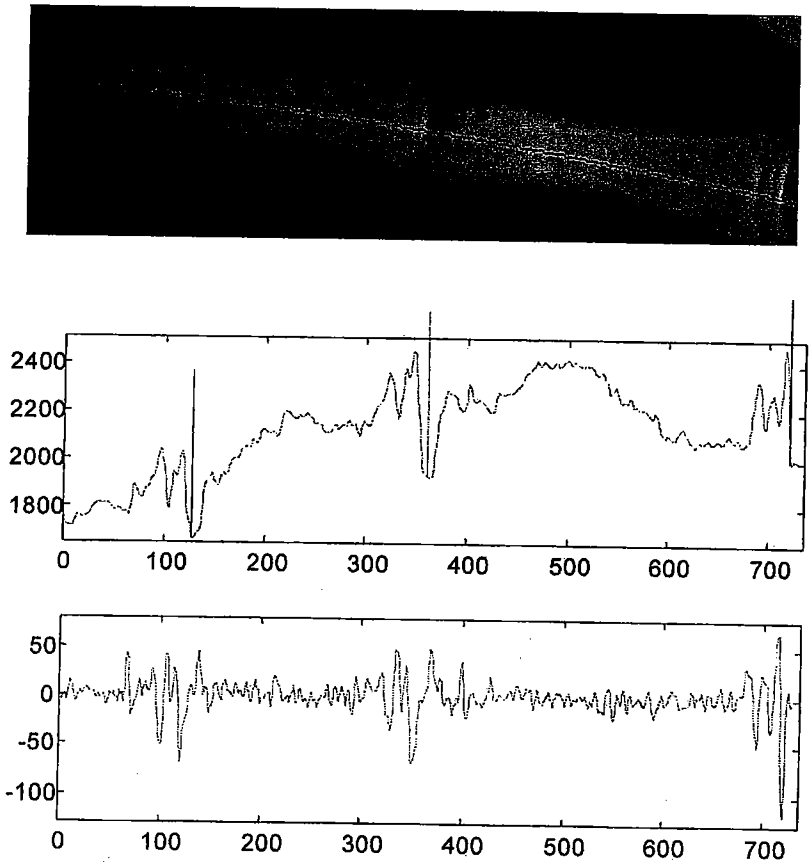


**Fig 9. Location of the phalangeal gaps. (Top) Middle phalanx and its axis. (Middle) Average profile along the axis (vertical lines mark local minima within gaps). (Bottom) Derivative of the above profile.**

should cover the entire sharp variation area surrounding the gap. Because the distance between gaps depends on the length of fingers, its size is assessed on the basis of the axis length. The lowest pixel values mark the final position of the gaps.

Once the gaps are located, the EROI edges are defined. The phalangeal width is estimated by the width of the HVI, and distance between gaps may be used for setting of the height. The current EROI size may be adjusted depending on image processing functions implemented at the stage of feature extraction. Location of extracted EMROI is shown in (Fig 10).

## RESULTS

The described methodology has been tested on a database of 130 clinical left hand-wrist radiographs. Films are digitized with a Lumiscan 200 digitizer to 2KB × 2KB × 12 bits. The test set includes radiographs of children between 1 and 18 years of age. All procedures have been implemented in Matlab 5.2 on a SUN workstation.

The hand extraction procedure has failed in 1% of images. In this case, the background has not been estimated correctly causing the suppression of soft tissue on the tip of phalanges. One radiograph of very bad quality (low contrast and many artifacts) has been detected by the error-checking routines and rejected from further analysis.

When considering the phanageal axes extraction, the ratio of failed image analysis cases increases to 7%. It includes the case mentioned above as well as hand images in which soft tissue of neighbor fingers are joined.

Regarding epiphyseal regions extraction, the number of units to be correctly identified increases 6 times (each hand image includes 6 EROI to be marked). In this case the number of failed region extraction decreases to 3%. This means that in many images only a single region is misextraced.

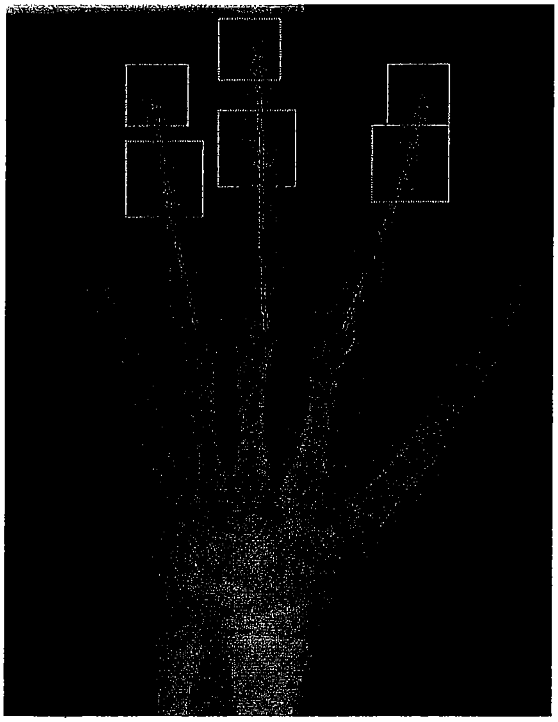

**Fig 10. Distal and middle epiphyseal regions of interest of phalanges II, III, IV.**

Most of the errors, resulting in shifting of the ROIs markers to the interior of bones, have been caused by low contrast images. In 2 cases the fingers have been bent, yielding higher intensity of the soft tissue around the gap. In a computer-assisted image analysis performed clinically, these regions require a manual adjustment.

## DISCUSSION

The main aim of the background removal procedure is to correctly determine the image background trend on the basis of the background peak location in each image quarter. When a nonfirmly closed cassette is exposed, overexposed spots or borders of a big area may lead to a wrong main background peak detection resulting in a wrong background estimation and removal. These artifacts may join or overlap the hand region. To decrease the sensitivity of the phalangeal axis detection to these type of failure, an interpolation procedure is applied. It reduces the axis deviation.

The algorithm is able to handle radiographs at a standard hand position (left hand, upright). Because a nonstandard hand position is likely to trigger the background removal error, an error detection procedure is implemented. It checks the presence of a handlike structure. If not shown, the analysis is suspended. The next generation of this module will handle the problem.

## CONCLUSION

Here we report a method of extraction of distal and middle regions of II, III, and IV phalanx in hand radiographs. Short processing time, high accuracy ratio, and build-in error checking procedures make this suitable for practical use in computer-aided bone age assessment. Because of the structured form of the package, procedure of hand extraction, axes extraction, and epiphseal regions location can be used independently. The improvement of image quality also would increase the ration of accurately extracted regions.

## REFERENCES

1. Greulich WW, Pyle SI: Radiographic Atlas of Skeletal Development of Hand Wrist (ed2). Stanford, CA, Stanford University Press, 1971

2. Tanner JM, Whitehouse RH: Assessment of Skeletal Maturity and Prediction of Adult Hight (TW2 Method). London, England, Academic Press, 1975

3. Dickhaus H, Habich R, Wastl S, et al: A PC-based system for bone age assessment. Proc EMBEC '99, Vienna, Austria, 1999

4. Al-Taani AT, Ricketts IW, Cairns AY: Classification of hand bones for bone age assessment. ICECS '96, 1996

5. Tanner JM, Gibbons RD: Automatic bone age measurement using computerized image analysis. J Pediatr Endocrinol 7:141-145, 1994

6. Albanese A, Hall C, Stanhope R: The use of a computerized method of bone age assessment in clinical practice. Horm Res 3:2-7, 1995

7. Pietka E, Gertych A, Pospiech S, et al: Computer-assisted bone age assessment: Image preprocessing and epiphyseal/metaphyseal ROI extraction. IEEE Trans Med Imag 20:715-729, 2001

8. Lim JS: Two Dimensional Signal & Image Processing. Englewood Cliffs, NJ, Prentice-Hall, 1990

9. Gonzalez RC, Woods RE: Digital Image Processing. Reading MA, Addison-Wesley, 1993, pp 42-43

10. Haralick RM, Shapiro LG: Computer and Robot Vision, Vol. 1. Reading MA, Addison-Wesley, 1992, pp 28-48