University of Pennsylvania
ScholarlyCommons

Statistics Papers

Wharton Faculty Research

6-1-1991

Automatic Recognition of Bone for X-Ray Bone Densitometry

Larry A. Shepp
University of Pennsylvania

Y. Vardi

J. Lazewatsky

James Libeau

Jay A. Stein

Follow this and additional works at: https://repository.upenn.edu/statistics_papers

 Part of the Computer Sciences Commons, and the Statistics and Probability Commons

Recommended Citation

Shepp, L. A., Vardi, Y., Lazewatsky, J., Libeau, J., & Stein, J. A. (1991). Automatic Recognition of Bone for X-Ray Bone Densitometry. *SPIE Proceedings, 1452* http://dx.doi.org/10.1117/12.45385

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/statistics_papers/400
For more information, please contact repository@pobox.upenn.edu.

Automatic Recognition of Bone for X-Ray Bone Densitometry

Abstract

We described a method for automatically identifying and separating pixels representing bone from those representing soft tissue in a dual- energy point-scanned projection radiograph of the abdomen. In order to achieve stable quantitative measurement of projected bone mineral density, a calibration using sample bone in regions containing only soft tissue must be performed. In addition, the projected area of bone must be measured. We show that, using an image with a realistically low noise, the histogram of pixel values exhibits a well-defined peak corresponding to the soft tissue region. A threshold at a fixed multiple of the calibration segment value readily separates bone from soft tissue in a wide variety of patient studies. Our technique, which is employed in the Hologic QDR-1000 Bone Densitometer, is rapid, robust, and significantly simpler than a conventional artificial intelligence approach using edge-detection to define objects and expert systems to recognize them.

Disciplines

Computer Sciences | Statistics and Probability

This journal article is available at ScholarlyCommons: https://repository.upenn.edu/statistics_papers/400

Automatic recognition of bone for x-ray bone densitometry

L. A. Shepp and Y. Vardi

AT&T Bell Laboratories
Murray Hill, New Jersey 07974

J. Lazewatsky, J. Libeau, and J. A. Stein

Hologic Inc.
Waltham, Massachusetts 02154

ABSTRACT

We describe a method for automatically identifying and separating pixels representing bone from those representing soft tissue in a dual-energy point-scanned projection radiograph of the abdomen. In order to achieve stable quantitative measurement of projected bone mineral density, a calibration using sample bone in regions containing only soft tissue must be performed. In addition, the projected area of bone must be measured. We show that, using an image with a realistically low noise, the histogram of pixel values exhibits a well-defined peak corresponding to the soft tissue region. A threshold at a fixed multiple of the calibration segment value readily separates bone from soft tissue in a wide variety of patient studies. Our technique, which is employed in the Hologic QDR-1000 Bone Densitometer™, is rapid, robust, and significantly simpler than a conventional artificial intelligence approach using edge-detection to define objects and expert systems to recognize them.

1. BONE DENSITOMETRY WITH γ-RAYS

A large fraction of the female population will experience a substantial loss of bone density and increased risk of fracture after menopause. The traditional method to measure bone density is to measure the attenuation of both a low energy (e_0) and a high energy (e_1) monochromatic γ-ray beam along a line l through the patient. The attenuation of a beam at energy e is given by the measured value

$$A(e) = \exp(-\mu B(e)L_B - \mu_T(e)L_T) \tag{1.1}$$

where $\mu_B(e)$ and $\mu_T(e)$ are the linear attenuation coefficients of bone and soft tissue at energy e, and L_B and L_T are the total lengths of bone and tissue respectively along l. Use of two values e_0 and e_1 for e gives two equations in the unknowns L_B and L_T which may then be determined and compared with standard values for a patient's weight and height. In practice, many lines $l = l_i$ are used and an overall bone mineral density is assigned to the patient corresponding to region of the spine. The region may be identified and duplicated on a second scan of the patient if an image is obtained and recorded.

2. OSTEODENSITOMETRY WITH X-RAYS

In principle, X-rays and γ-rays are identical, but while X-ray (electron-tube) sources have high intensity compared to γ-ray (isotope) sources, an advantage allowing decreased measurement times, it is difficult to obtain monochromatic X-ray beams at fixed energies e_0 and e_1. Instead, it is most convenient to use broad-band low and high energy X-ray beams which are centered about effective energy values e_0 and e_1. Moreover, each of e_0 and e_1 may vary slightly with time.

As can be seen,[1] X-rays can still be effectively used, provided that calibration/correction is done to remove drift and fluctuations in X-ray intensity. To analyze X-ray or γ-ray data bone images it is necessary to be able to distinguish automatically those pixels which contain bone from those which do not. Thus one must solve the pattern recognition problem of identifying bones automatically in a digital X-ray image such as that of Figure 1. To test the robustness of our methods, we formed Figure 1 as a composite of 4 different radiographs into a 256 × 256 digital image.

3. PATTERN RECOGNITION; ARTIFICIAL INTELLIGENCE VS. HISTOGRAM

Given an abdominal radiograph, such as in Figures 1-3, an untrained human observer can easily identify the bones. It is believed that the human uses understanding of skeletal anatomy and would not be confused by the horizontal streaks in Figures 1 and 2.

 0-8194-0551-5/91/$4.00

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/27/2016 Terms of Use: http://spiedigitallibrary.org/ss/TermsOfUse.aspx

The streaks present in these images are the result of the subject's breathing movement during the approximately five minutes necessary for the raster-scanned system to obtain the data. This feature is not normally present on a conventional single-shot radiograph, nor is it present in the image representing bone-mineral density, which is a linear combination of these low-energy and high-energy images. We first tried, somewhat feebly, to take an artificial intelligence approach to identify bones automatically by writing programmed rules understanding skeletal structure. An edge detector would first delineate blobs of higher density and then expert system inference would be used to eliminate blobs due to noise or jitter which would not have the right shape according to rules of skeletal anatomy. This method did not seem to be working, and was abandoned when a simpler approach was found to be successful, at least for low-noise cases such as Figure 3. The simpler approach described below was unsuccessful on Figures 1 and 2, and if it were necessary to identify bones in these Figures, perhaps an artificial intelligence approach as outlined above would be needed.

The specific algorithm described below is not the one actually used in the Hologic Bone Densitometer. Moreover, the subject of Figure 1 was a lean male rather than a postmenopausal female. There were also changes in hardware and sampling rate since this experiment was performed. Since our subject is a relatively lean male, the bone density is greater and the soft tissue thickness is lower than might be expected in most female subjects. This also means a smaller fraction of the projection is through fat which has a somewhat different coefficient of absorption than does muscular tissue. The combined result of these is that with a typically broader soft-tissue peak and a lower bone peak, the separation of the two peaks is likely to be considerably less pronounced than it is here. Even worse, in a typical elderly female patient, the bone peak in particular may be so low and broad as to be difficult to define as a peak. The resulting peak tends to look more like a single peak with some irregular training on the right side.

As a result the method suggested here must be modified for more difficult cases, but the general approach is similar. The mode of the first peak is the starting point to fix the lower limit for bone pixel values based upon a multiple (generally less than one) of the modal value of the corresponding peak in the histogram in which a bone-standard is interposed in the beam for calibration. Unfortunately, the bone peak is not reliable enough in a broad enough subset of the possible subjects to permit our simple method to extract the bone standard directly. But the histogram technique is employed in an essential way in the final product.

Figure 3 was obtained from Figures 1 and 2 by forming the normalized subtraction of the $n \times n$ images, (Figures 1 and 2 were obtained at two different X-ray kilovoltages – 70KVP and 140KVP)

$$x_i = D_i(e_1) - c\, D_i(e_0) = \log A_i(e_1) - c \log A_i(e_0)\,, \quad 1 \le i \le n \times n \tag{3.1}$$

where c is a constant, $c \approx \mu_T(e_1)/\mu_T(e_0)$ and where x_i is the gray level in the i^{th} pixel of Figure 3, and $D_i(e) = \log A_i(e)$ and $A_i(e)$ is the i^{th} measurement along the line l_i as in (1.1), $e = e_0, e_1$. Note that as a result of the hardware, Figures 1 and 2 already represent the logarithm of the projected X-ray attenuation. From (1.1) we see that

$$x_i = (c\,\mu_B(e_0) - \mu_B(e_1))L_B(i) + (c\,\mu_T(e_0) - \mu_T(e_1))L_T(i) \tag{3.2}$$

where $L_B(i)(L_T(i))$ is the length of the bone (tissue) along l_i. Note that the coefficient of $L_T(i)$ in (3.2) would vanish if the low and high energy beams were strictly monochromatic but c only represents an effective or average ratio of $\mu_T(e_1)/\mu_T(e_0)$. We use $c = 1.35$ or $c = 1.4$ empirically in the sequel.

Since the measurements (1.1) of $A_i(e_0)$ and $A_i(e_i)$ are made almost simultaneously, the patient motion noise present in Figures 1 and 2 nearly disappears after the subtraction (3.1) and is not visible in Figure 3.

We normalize x_i in (3.2) so that it has $g = 128$ integer (gray level) values, 0, 1, ..., $g-1$ and form the histogram, $h(k)$, $k = 0, 1, \ldots, g-1$ where $h(k)$ is the number of pixels i in which the normalized x_i has level k. Since every one of the $n \times n$ pixels must have some level k,

$$\sum_{k=1}^{g-1} h(k) = n \times n\,. \tag{3.3}$$

A table of the histogram of Figure 3 is given in Figure 4 and a plot is given in Figure 6. The $g = 128$ histogram values $h(k)$, $k = 0, \ldots, g-1$ are listed in Figure 4 starting in the third row, $h(0) = 271$, $h(1) = 11$, ..., $h(127) = 2$, the last column giving the last value of k in each row. The overall maximum of the unnormalized x_i was 1515.6, the minimum was zero and the ratio $c \approx \mu_T(e_0)$ used in (3.1) was $c = 1.35$ indicated in the first row of Figure 4. (Use of $c = 1.4$ gives similar results.) The overall maximum value, $\max_1$, of $h(k)$ occurs at $k = k_{m_1} = 53$ with $\max_1 = h(k_{m_1}) = 6185$, and the second overall maximum value $\max_2$ of $h(k)$ *which is separated* from k_{m_1} occurs at $k = k_{m_2} = 64$ with $\max_2 = h(k_{m_2}) = 2189$. Thus except for

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/27/2016 Terms of Use: http://spiedigitallibrary.org/ss/TermsOfUse.aspx

fluctuations near $k=0$ and $k=g-1$, the histogram values $h(k)$ are essentially bimodal, i.e. there are typically two local maxima, one corresponding to soft tissue centered at k_{m_1} and one to bone centered at k_{m_2} in this case. Note that if there were more bone pixels than soft tissue pixels the higher maximum of h might occur at the higher value of k. The values $k_1=26$ and $k_2=77$ in the second row of Figure 4 were determined, as the first values for which the sum of the $h(k)$ for $k \le k_1$ or for $k \ge k_2$ first exceeded a threshold t, where we took $t=512$ for convenience. Then the overall maximum k_{m_1} was required to lie in $k_1 \le k_{m_1} \le k_2$. Finally the second overall maximum k_{m_2} was required to lie in the union of the two intervals $k_1 \le k_{m_2} \le k_{m_1}-b$ and $k_{m_1}+b \le k_{m_2} \le k_2$ and be the largest value of $h(k)$ in this range. The value $b=5$ was used to keep k_{m_2} from lying too close to k_{m_1}. Of course the above procedure is ad hoc in the choice of $g=128$, in the choice of $t=512$, and in the choice of $b=5$, but it is clear that some values of these parameters ought to work.

By the above procedure, two local maxima k_{m_1} and k_{m_2} are defined. The larger of these, k_{high}, corresponds to the center of the bone pixels (the mode of the bone) while the smaller, k_{low}, corresponds to the center of the soft tissue pixels because bone is denser than soft tissue. Placing a threshold at

$$k_\tau = k_{\text{low}} + \tau(k_{\text{high}} - k_{\text{low}}) \tag{3.4}$$

where $\tau=\frac{1}{2}$ allows bone to be defined by $k>k_\tau$. Setting $\tau=\frac{1}{4}$ or $\frac{3}{4}$ allows one to favor bone or soft tissue in the decision. The different choices $\tau=\frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ give rise to the three sections of Figure 5 which are black-white, where white is placed in any pixel where the normalized $x_i=k>k_\tau$. It is clear from Figure 5 that the method works effectively.

The original image Figure 3 was sufficiently noise-free that no additional noise removal was needed in Figure 5. Certain noise could be removed by using an artificial intelligence or expert system approach; for example speckle noise or isolated white pixels cannot correspond to bones and can be removed. However there are limitations on how much noise is tolerable and indeed if the same procedure is tried using the histogram of Figure 1 or Figure 2 instead of the much-less-noisy subtracted image in Figure 3 our technique fails. The main reason that the histogram technique fails to find the right threshold in Figures 1 and 2 is that the threshold varies strongly among the composite parts. Even if one uses adaptive thresholding by using local histograms to define a threshold which varies with position the technique fails on Figures 1 and 2 because of the jitter due to patient motion (breathing) which is removed in the subtracted image (Figure 3). Thus our method has limitations and is only justified by its simplicity.

4. REFERENCES

1. D. B. Plewes and E. Vogelstein, "Exposure artifacts in raster scanned equalization radiography," *Med. Phys.*, vol. 11, pp. 158-165, 1984.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/27/2016 Terms of Use: http://spiedigitallibrary.org/ss/TermsOfUse.aspx

Figure 1. Low energy X-ray image of the spine. Note jitter due to patient motion. This is a composite of the spines of 4 patients, made for the purpose of testing robustness of the method.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/27/2016 Terms of Use: http://spiedigitallibrary.org/ss/TermsOfUse.aspx

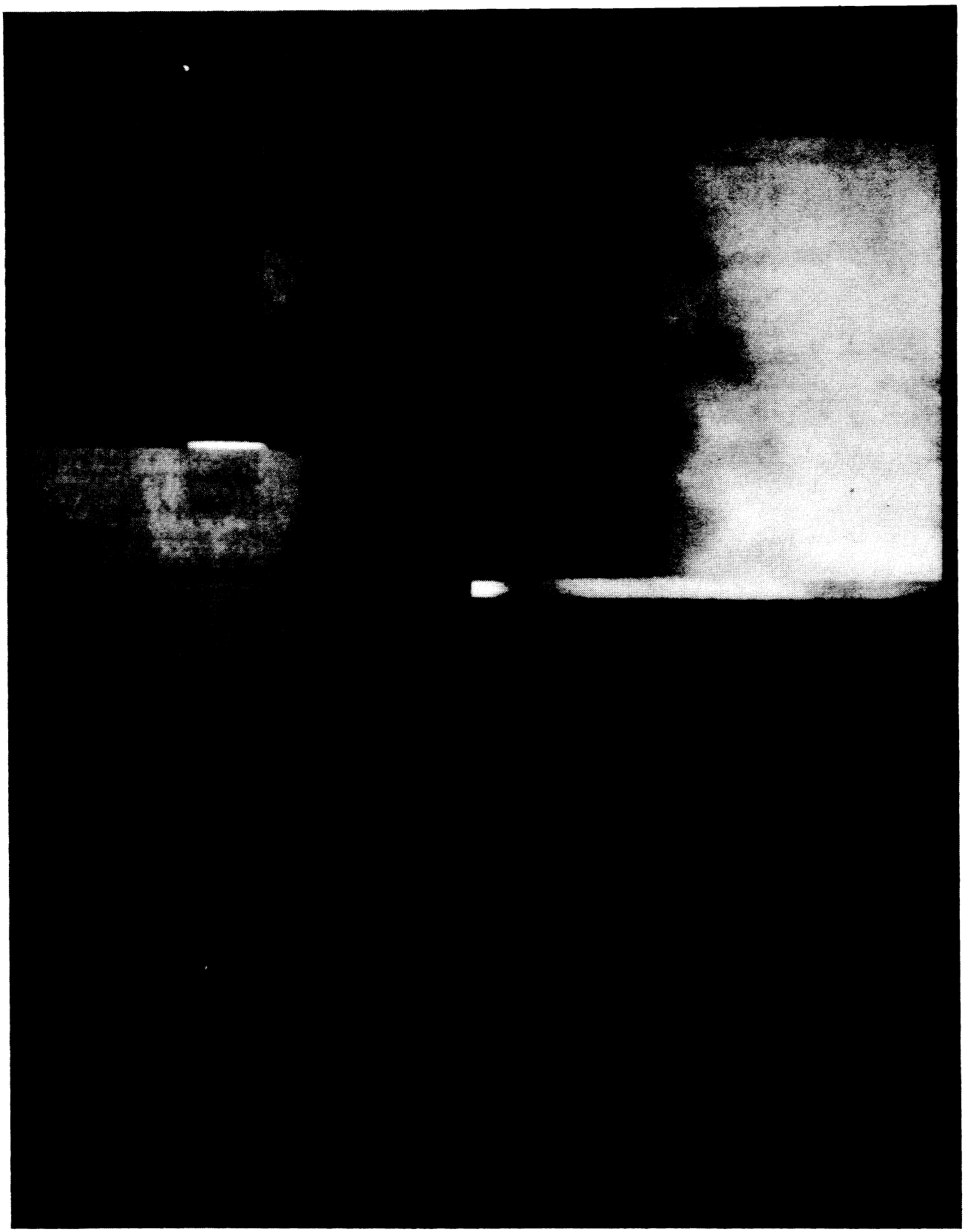


Figure 2. High energy X-ray image of the spine. Note jitter due to patient motion. This is a composite of the spines of 4 patients made for the purpose of testing robustness of the method.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/27/2016 Terms of Use: http://spiedigitallibrary.org/ss/TermsOfUse.aspx

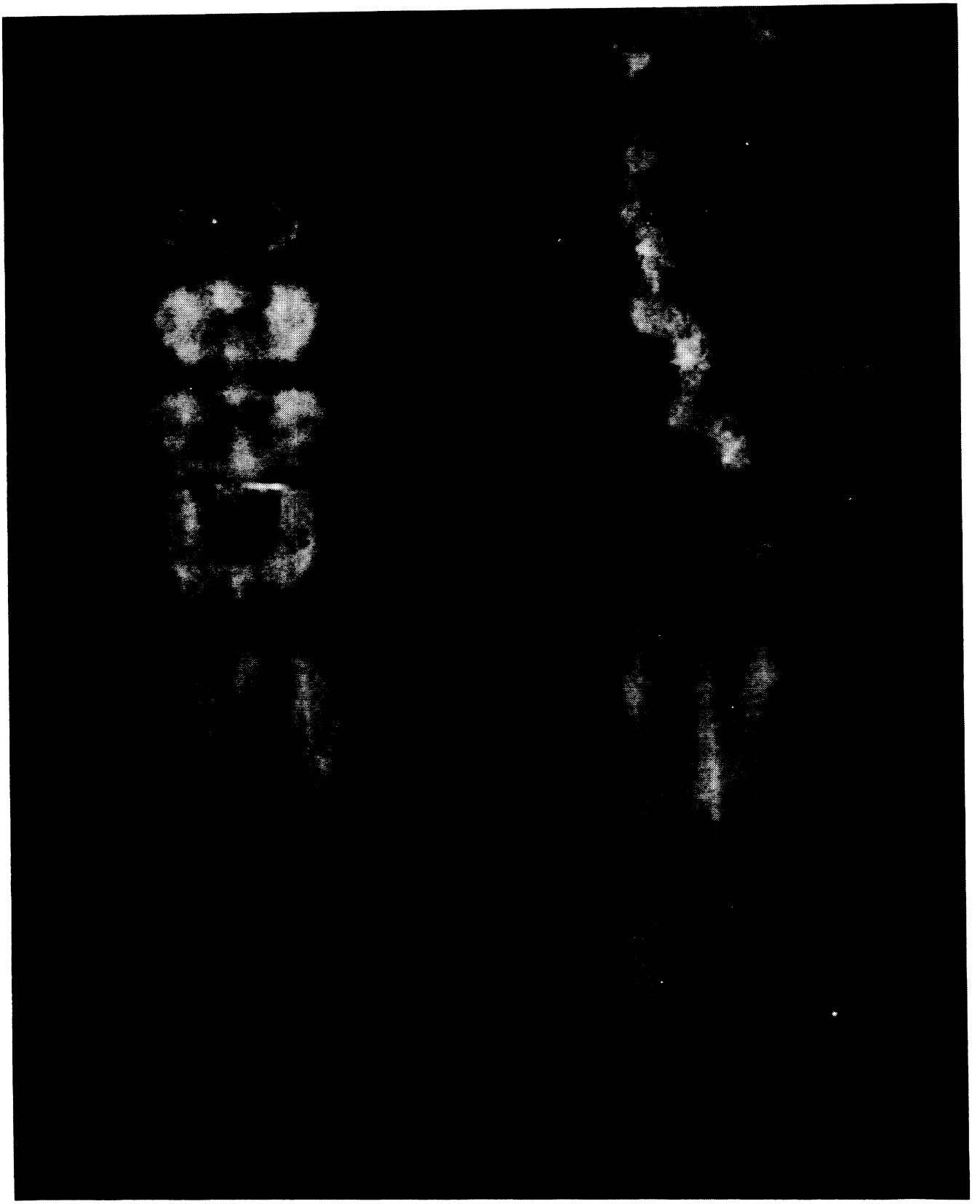


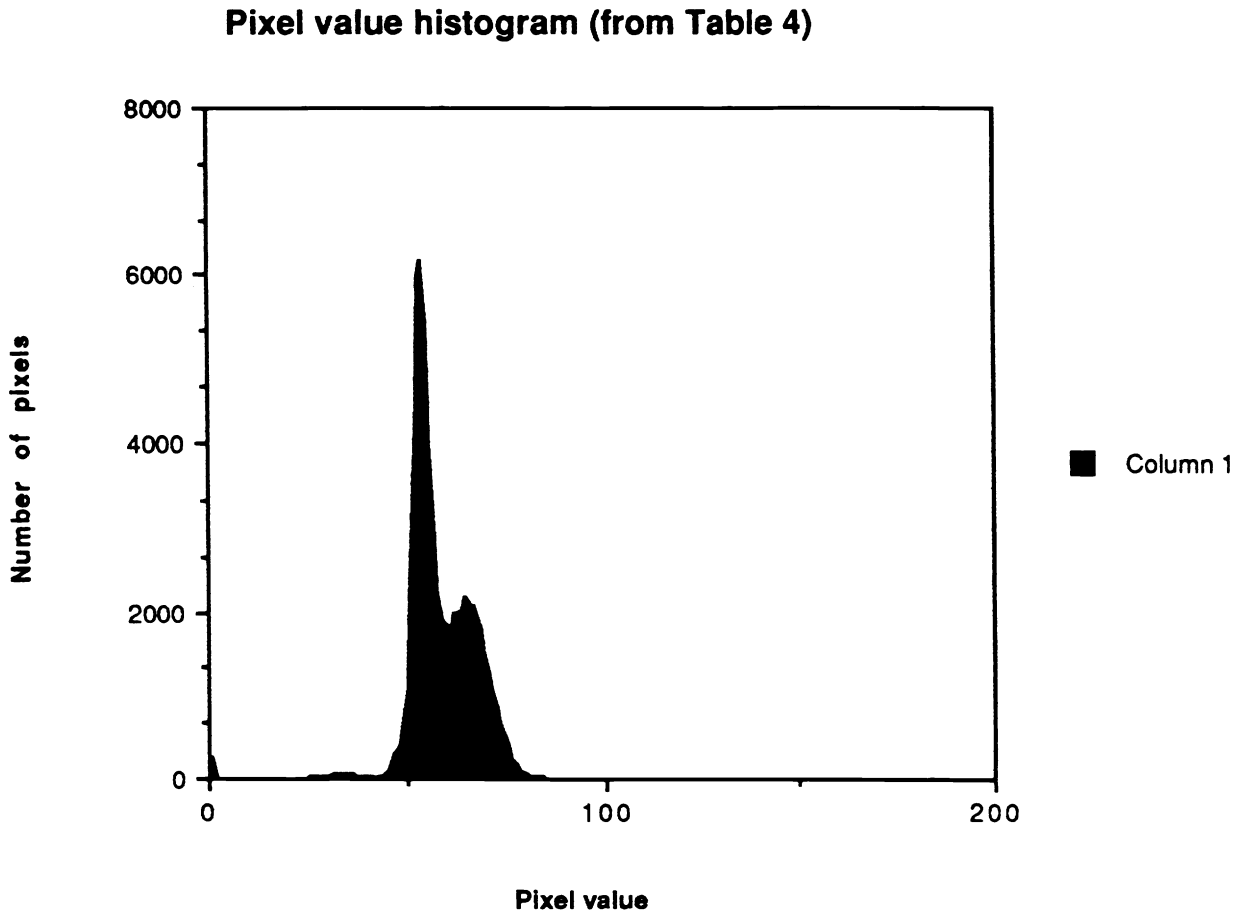
Figure 3. Subtraction of normalized low and high energy images. Note absence of jitter due to patient motion. This was the data we used to distinguish bone from tissue.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/27/2016 Terms of Use: http://spiedigitallibrary.org/ss/TermsOfUse.aspx

```
const,max,min = 1.350000 1515.599976 0.000000
k1 k2 km1 km2 max1 max2 26 77 53 64 6185 2189
 271 ,  11 ,   3 ,  12 ,| 3
   5 ,   5 ,   8 ,  10 ,| 7
   3 ,  10 ,   5 ,   7 ,| 11
  11 ,   9 ,   6 ,   9 ,| 15
   7 ,  11 ,   8 ,   6 ,| 19
  10 ,  12 ,  10 ,  15 ,| 23
  15 ,  17 ,  20 ,  20 ,| 27
  25 ,  39 ,  45 ,  51 ,| 31
  56 ,  67 ,  64 ,  63 ,| 35
  45 ,  47 ,  34 ,  20 ,| 39
  24 ,   7 ,  26 ,  39 ,| 43
  84 , 187 , 285 , 390 ,| 47
 606 , 1043 , 2361 , 4193 ,| 51
5994 , 6185 , 5409 , 3973 ,| 55
2928 , 2214 , 1892 , 1855 ,| 59
1786 , 1990 , 1983 , 2023 ,| 63
2189 , 2092 , 2075 , 1884 ,| 67
1796 , 1515 , 1241 , 1051 ,| 71
 833 , 639 , 491 , 391 ,| 75
 233 , 169 , 107 , 55 ,| 79
  40 ,  35 ,  20 ,  20 ,| 83
  10 ,   3 ,  13 ,   6 ,| 87
   3 ,   3 ,   5 ,   6 ,| 91
   5 ,   3 ,   1 ,   0 ,| 95
   6 ,   3 ,   2 ,   1 ,| 99
   1 ,   4 ,   1 ,   1 ,| 103
   2 ,   0 ,   2 ,   2 ,| 107
   0 ,   2 ,   1 ,   1 ,| 111
   1 ,   0 ,   0 ,   0 ,| 115
   0 ,   1 ,   0 ,   0 ,| 119
   0 ,   0 ,   0 ,   0 ,| 123
   0 ,   0 ,   0 ,   2 ,| 127
```

Table 4. Table of the histogram of the $n \times n$ image of Figure 5, $n=256$. The $g=128$ histogram values are listed beginning in the third line in sets of 4, hist(0)=271, hist(12)=1. The value of c in (3.1) is $c=1.35$. The maximum and minimum x_i are as shown in the first line. The values of k_1, k_2, k_{m_1}, k_{m_2}, max 1 and max 2 are as in the second line (see §3).

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/27/2016 Terms of Use: http://spiedigitallibrary.org/ss/TermsOfUse.aspx



Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/27/2016 Terms of Use: http://spiedigitallibrary.org/ss/TermsOfUse.aspx

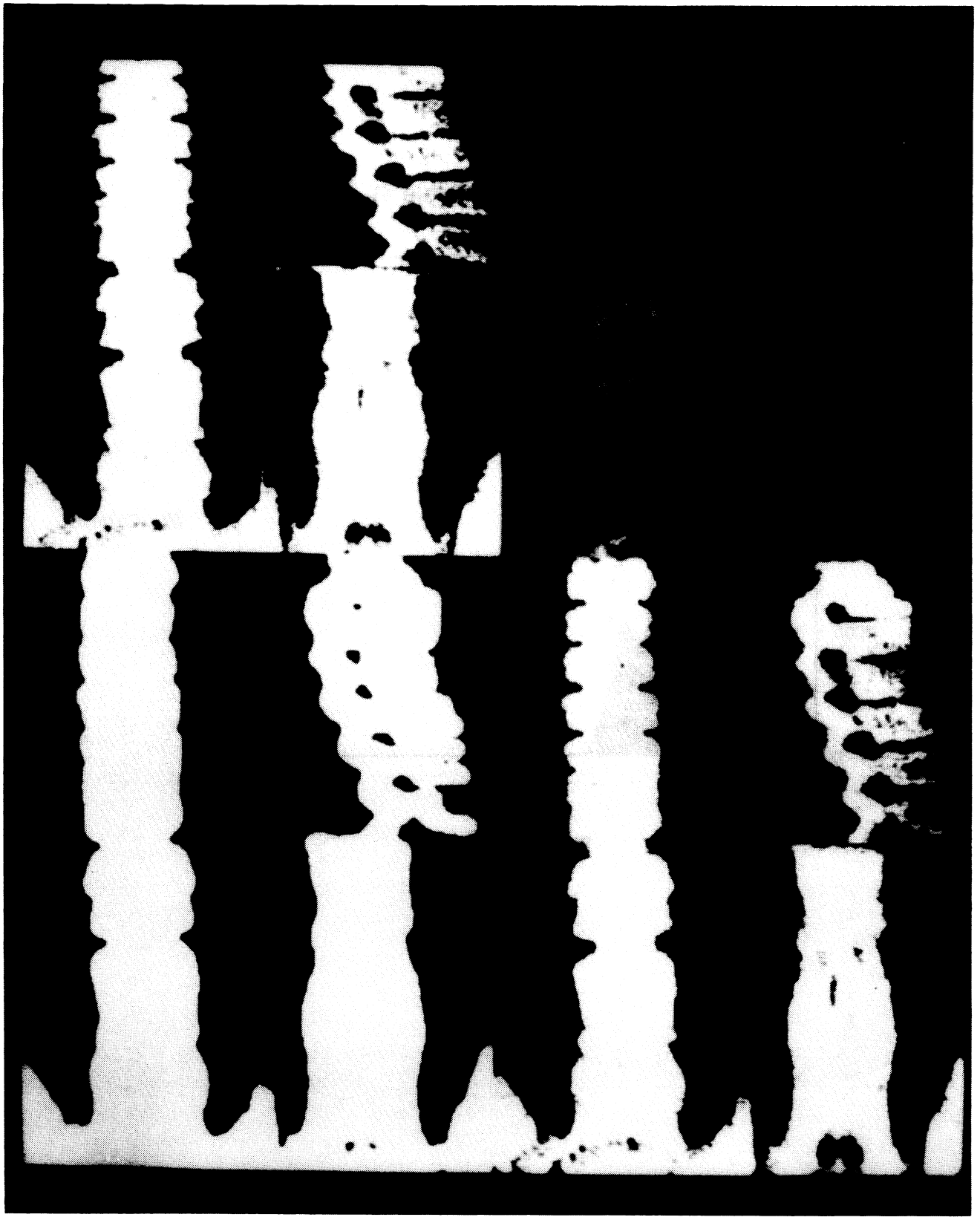


Figure 5. Upper right is Figure 3. Other three sections correspond to the threshold values placed at 1/4, 1/2, 3/4 of the distance from max_1 to max_2. This shows that bone can be distinguished from tissue automatically.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/27/2016 Terms of Use: http://spiedigitallibrary.org/ss/TermsOfUse.aspx