

Clustering approach to bone and soft tissue segmentation of digital radiographic images of extremities

S. Kubilay Pakin

University of Rochester
Department of Electrical and Computer Engineering
Rochester, New York 14627

Roger S. Gaborski

Rochester Institute of Technology
Department of Computer Science
Rochester, New York 14653

Lori L. Barski

David H. Foos

Eastman Kodak Company
Health Imaging Research Laboratory
Rochester, New York 14650

Kevin J. Parker

University of Rochester
Department of Electrical and Computer Engineering
Rochester, New York 14627

Abstract. We present an algorithm for segmentation of computed radiography (CR) images of extremities into bone and soft tissue regions. The algorithm is region-based in which the regions are constructed using a region-growing procedure based on two different statistical tests. Following the region-growing process, a tissue classification method is employed. The purpose of the classification is to label each region as either bone or soft tissue. This binary classification goal is achieved by using a voting procedure that consists of the clustering of regions in each neighborhood system into two classes. The voting procedure provides a crucial compromise between the local and the global analysis of the image, which is necessary due to strong exposure variations seen on the imaging plate. Also, the existence of regions whose size is large enough such that exposure variations can be observed through them makes it necessary to use overlapping blocks during the classification. After the tissue classification step, the resulting bone and soft tissue regions are refined by fitting a second-order surface to each tissue, and reevaluating the label of each region according to the distance between the region and surfaces. The performance of the algorithm is tested on a variety of extremity images using manually segmented images as the gold standard. The experiments show that our algorithm provides a bone boundary with an average area overlap of 90% compared to the gold standard. © 2003 SPIE and IS&T. [DOI: 10.1117/1.1526846]

1 Introduction

Accurate diagnosis of disease states in medical radiographic images often depends on the detection of small, low-contrast details in the image. The visibility of these details is dependent on the film's tonescale curve. The tonescale response curve of traditional silver halide films is determined by the film's built-in chemical emulsion and the development process. The manufacturer attempts to optimize the film for a range of exposure techniques and exam types. Digital radiographic imaging systems record raw image data without the built-in tonescale curve and with a greater dynamic range. The increased dynamic range of the digital image, along with digital image processing techniques, enable the manipulation of the image before it is viewed by the radiologist. Segmentation of the region of interest (ROI) in the code value histogram is an important first step in performing image enhancement processing. Once identified, the range of code values corresponding to the ROI can be optimally mapped, via a human brightness perception model, to the film density response of a laser printer or to the luminance response of a CRT for visual interpretation. Robust algorithms exist for the segmentation of direct exposure and collimation regions in radiographic imagery.

Previous methods have focused on segmenting the code value histogram based on knowledge of the body part being

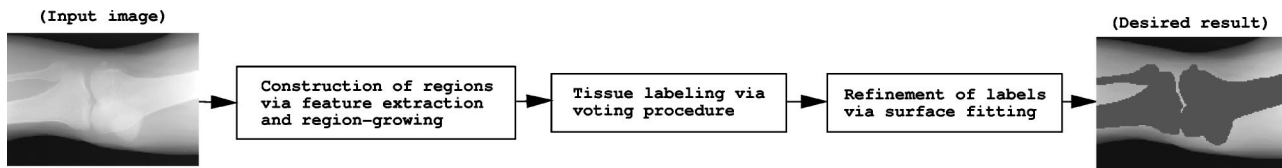


Fig. 1 Flowchart of the proposed algorithm.

imaged and the characteristics of the corresponding histogram.¹ For a particular body part, the histogram contains certain features (peaks, valleys, etc.) that can be used to map code values to anatomical features. This technique lacks robustness in certain exam types and exposure conditions. Another approach focused on segmenting the image into bone and tissue regions based on a neural network classifier using texture measures as features.² The time required to obtain the texture features makes this approach impractical. The definition of a robust and effective method to refine the estimate of the ROI by segmenting the bone and soft tissue components of the image has remained elusive.

The algorithm proposed in this paper can be considered as a hybrid technique that combines the features of region-based and histogram-based techniques.³ Histogram-based techniques follow the Bayesian approach for estimation of the label field by modeling the image as a sample from a Gaussian distributed random field. Also, the intensity values of the pixels are assumed to be independent random variables whose density function parameters (mean and variance values since they are Gaussian distributed) depend on the label of the pixel. The assumption of equally likely labels for every pixel results in the well-known K -means algorithm.⁴ There are also numerous papers in the literature that impose spatial constraints by modeling the label field as a Gibbs or Markov random field.⁵⁻⁷ In general, these algorithms provide better results than the K -means algorithm but they are also known for their significant computational burden. The main problem associated with the usage of histogram-based techniques in the context of the segmentation of computed radiography (CR) images is that CR images do not follow the image model that histogram-based techniques rely on due to the exposure variations.

This paper presents a method for segmentation of digital radiographic images into bone and soft tissue regions. A detailed description of our method is presented in Sec. 2. The experiments and performance evaluation of the method are provided in Sec. 3. Finally, the conclusions are drawn in Sec. 4.

2 Method

Segmentation techniques, in general, attempt to group pixels that have intensity-based similarities to provide a tessellation of the image. However, the purpose of medical image segmentation is to group pixels that belong to the same tissue type. The difficulty of medical image segmentation stems from the fact that the solutions of these two problems are not necessarily the same. Medical images frequently possess high-intensity variations throughout the regions that correspond to the same tissue type. Furthermore, two different tissues can share very similar intensity profiles. The former characteristic is the main problem with CR seg-

mentation and it is the reason for the poor performance of histogram-based techniques, such as K -means. To overcome this problem, we propose a clustering-based algorithm that is governed by the observation that the overlap between the intensity profiles of bone and soft tissue becomes insignificant when the image is analyzed locally. The flowchart of the proposed algorithm can be seen in Fig. 1. The second stage of the algorithm, which we call the voting procedure, performs the region-labeling task by local region clustering using the K -means algorithm with two classes. The effect of exposure variations is compensated for by processing the image locally using the neighborhood systems and subimages. Note that, the initial tessellation of the image is provided by the first stage of the algorithm, which is a region-growing approach, and it effectively introduces spatial constraints to the label field. The last stage of the algorithm refines the labels of the regions by fitting a second-order bivariate polynomial to each tissue and investigating the average distance between each region and the surface of the label to which they belong. The surface-fitting stage can also be conceptualized as a clustering scheme in which the regions are clustered into two classes using a different kind of metric than the usual ones (e.g., Mahalanobis or Euclidean distances) used in conventional clustering algorithms. From this point of view, the voting procedure provides the initial classification. Each of the stages of the algorithm are explained in detail in the following sections.

2.1 Construction of Regions

The first stage of the algorithm provides the tessellation of the image using a region-growing type technique. For this purpose, two features are computed for each pixel. The features are denoted as μ_i and σ_i , where i denotes the index of the pixel site, and they reflect the local characteristics around the pixel that they belong.

1. μ_i : The median value of the intensities of the pixels that are inside the neighborhood of pixel \mathbf{x}_i , assuming eight-connectivity.⁴ The region-growing algorithm depends on local differences that are suppressed to some extent in the sample mean computation. Because of this fact, the median value is used instead of the sample mean to avoid smoothing of region boundaries. More sophisticated schemes for feature extraction such as iterative methods that preserve line structures were avoided due to their high computational cost.⁸
2. σ_i : The standard deviation of the intensities of the pixels that are inside the neighborhood of pixel \mathbf{x}_i , assuming eight-connectivity.

The feature values are utilized for investigation of the statistical similarity of each adjacent pixel pair where adjacency is defined in an eight-connectivity sense. For the pixel pair $(\mathbf{x}_i, \mathbf{x}_j)$, the following statistical tests are utilized to determine whether they are sufficiently similar to be in the same region.

1. *F test*: The pixels $(\mathbf{x}_i, \mathbf{x}_j)$ and their corresponding neighbors are assumed to be the realizations of the random variables whose distributions are $N(\eta_i, \nu_i^2)$ and $N(\eta_j, \nu_j^2)$, respectively. The well-known *F* test, which derives from the fact that if $\nu_i = \nu_j$, then the RV σ_i^2/σ_j^2 has an *F* distribution with degrees of freedom $(K-1, K-1)$, where K is the size of both ensembles,⁹ is used to investigate the similarity of the variances of the distributions to test the hypothesis that both ensembles are indeed governed by the same distribution. Note that σ_i^2 and σ_j^2 are the sample variance values and have been computed in the feature extraction step. The *F* test formulation is as follows:

$$\frac{\sigma_i}{\sigma_j} < \gamma_F, \quad (1)$$

where γ_F is a predetermined threshold. Note that γ_F is always greater than 1, and the reciprocal of Eq. (1) must be considered if $\sigma_i < \sigma_j$. Variance is mainly a texture feature. The *F* test is the means of investigating whether two distributions have the same variance and attempts to determine whether or not 2 pixels belong to same tissue, assuming different tissue types possess different textures, hence different variances.

2. *Mahalanobis distance*: In addition to the *F* test, which performs a comparison based on the standard deviation values of both ensembles, we use the following two tests in which \mathbf{x}_i and \mathbf{x}_j are compared to the ensemble that consists of \mathbf{x}_j and \mathbf{x}_i and their neighbors, respectively,

$$\frac{|\mu_i - \mu_j|}{\sigma_i} < \gamma_M \text{ and } \frac{|\mu_i - \mu_j|}{\sigma_j} < \gamma_M, \quad (2)$$

where γ_M is a predetermined threshold.

For a particular adjacent pixel pair $(\mathbf{x}_i, \mathbf{x}_j)$, if all the preceding statistical tests are satisfied, that pixel pair is said to be connected. Following the investigation of the similarity of each adjacent pixel pair, regions that are defined as a maximal set of pixels all belonging to the same connected components are formed. During the stages that follow region growing, the algorithm basically attempts to label each region as either bone or soft tissue. Therefore, it is vital to adjust the thresholds γ_M and γ_F to construct regions whose pixels belong to only one type of tissue, otherwise some of the pixels will inevitably be assigned incorrect tissue labels. While high threshold values will cause leakage between regions that correspond to different tissue types, a value that is too low will result in over segmentation, which deteriorates the performance of the stages that follow region growing.

Region-growing type algorithms for image segmentation

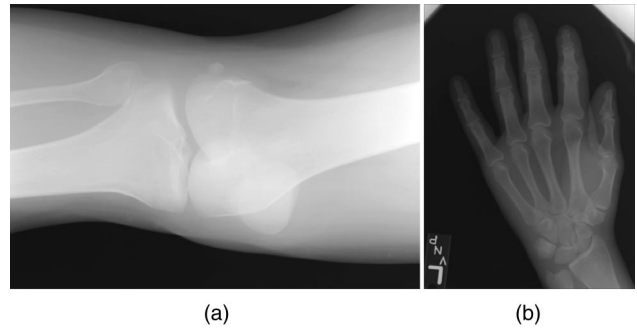


Fig. 2 Input images: (a) a knee scan and (b) a hand scan.

have been extensively investigated in the literature.¹⁰⁻¹² Our region-growing algorithm might be better understood by visualizing the image as an undirected adjacency graph, where each vertex corresponds to a pixel. According to this conceptualization, the vertex that corresponds to the pixel \mathbf{x}_i , is connected to vertices that correspond to pixels \mathbf{x}_j , where $\mathbf{x}_j \in N_i$, prior to the region-growing procedure. Note that N_i denotes the set of pixels that are in the neighborhood of \mathbf{x}_i , assuming eight-connectivity. The growing procedure will remove the arcs between pixels that are not statistically similar. After removal, each mutually exclusive subgraph will form a region. An advantage of this process is that there is no assumption about the number of regions.

Depending on a purely local intensity profile is one of the main drawbacks of the proposed region-growing scheme. One additional arc is adequate to connect two mutually exclusive subgraphs that form two adjacent regions. Because of this, low-contrast areas of the image are susceptible to leakage among regions. To reduce the possibility of leakage among regions that correspond to different tissue types, an edge map of the image (acquired by using Sobel operator with thinning) is utilized by prohibiting the connectivity of the pixels that lie on different sides of an edge contour. Note that the region construction process will not be adversely affected by spurious edges since (1) they will result in additional regions by breaking several arcs and (2) the main task of the region construction stage is to create a tessellation in which all regions are composed of one type of tissue. Hence, the oversegmentation caused by spurious edges in the edge map actually ensures a successful region construction stage. More complex edge detection schemes such as the Canny edge detector are avoided due to their computational cost.

Two different extremity images, a knee and a hand scan, were used as case studies. The input knee and hand images with corresponding region-growing results can be seen in Figs. 2(a), 2(b), 3(a), and 3(b), respectively (see Color Plate 1 for Fig. 3).

2.2 Voting Procedure

Following the region-growing stage, each region is assigned a tissue label using the voting procedure. Voting is a broad term and it has been used under several contexts in the literature. Examples include the Hough transform related grouping techniques,¹³ and the majority voting schemes used to combine the decisions of several classifiers.¹⁴ The procedure we propose assigns bone or soft

tissue votes to the regions that have been constructed in the previous step using local two-class clustering.

The exposure variations seen on the image plate prevent the digital radiographic images from having a bimodal histogram. The significant overlap between the histograms of bone and soft tissue intensities makes the use of global clustering algorithms inefficient. The voting procedure is essentially governed by the following observation: Because of the relatively high x-ray absorption of bone, locally, bone pixels are brighter than soft tissue pixels and the overlap between the intensity profiles of the two tissue types disappears. Exposure variations make these observations invalid for global analysis.

During the voting procedure, the image to be segmented is analyzed using overlapping blocks. The following notation is introduced to explain the voting procedure:

$\mathcal{R}_m = m'$ th region

$\mathcal{V}_m(\text{B}) =$ votes for the bone label in the m' th region

$\mathcal{V}_m(\text{S}) =$ votes for the soft tissue label in the m' th region

$\mathcal{B}_k = k'$ th block

$\mathcal{R}_{ik} = i'$ th region in the k' th block

$\mathcal{N}_{ik} =$ neighborhood system of \mathcal{R}_{ik} , which is composed of \mathcal{R}_{ik} and the regions that are adjacent to it

$\mathcal{R}_{jik} = j'$ th region in \mathcal{N}_{ik}

$\mathcal{L}_{jik} =$ label of \mathcal{R}_{jik} as the result of clustering of the regions in \mathcal{N}_{ik}

The algorithm of the voting procedure is as follows:

Algorithm voting.

For each \mathcal{B}_k

For each \mathcal{R}_{ik}

Compute the average intensities of $\{\mathcal{R}_{jik}\}$.

Cluster $\{\mathcal{R}_{jik}\}$ using K-means algorithm with 2 clusters.

For each \mathcal{R}_{jik}

If \mathcal{L}_{jik} is bone, increment $\mathcal{V}(\text{B})$ of corresponding \mathcal{R}_m .

If \mathcal{L}_{jik} is soft tissue, increment $\mathcal{V}(\text{S})$ of corresponding \mathcal{R}_m .

Since a particular region \mathcal{R}_m will appear in different blocks and neighborhood systems, it will acquire several votes. Specifically, \mathcal{R}_m will be processed in the neighborhood system of its own and in the neighborhood systems of adjacent regions. Also, \mathcal{R}_m will be processed inside the blocks in which a patch of \mathcal{R}_m (which is denoted as \mathcal{R}_{ik}) lies. After all the blocks and regions are processed, the tissue labels of the regions are determined as follows:

For each \mathcal{R}_m

If $\mathcal{V}_m(\text{B}) > \mathcal{V}_m(\text{S})$, \mathcal{R}_m is a bone region.

If $\mathcal{V}_m(\text{B}) < \mathcal{V}_m(\text{S})$, \mathcal{R}_m is a soft tissue region.

If $\mathcal{V}_m(\text{B}) = \mathcal{V}_m(\text{S})$, the label of \mathcal{R}_m is undecided.

The regions whose labels are undecided are reconsidered and their labels are assigned in the next stage, which is discussed in the Sec. 2.3.

Prior to the voting procedure, the regions that correspond to the background were detected using a simple thresholding scheme. The fact that the background has a relatively flat intensity pattern that does not overlap with

the intensity profile of tissues makes simple detection schemes applicable.

The preceding algorithm implicitly assumes that both tissue types exist in every neighborhood system. Even though this is generally the case for images that resemble the hand case study, it is not correct for images that resemble the knee case study. In such cases, the thickness of the soft tissue varies greatly in areas close to the tissue boundary, which causes a significant image gradient magnitude compared to other areas of the image. This fact results in densely clustered regions that belong to the same tissue type that can be observed in Fig. 3(a).

The adverse effect of this phenomenon on the performance of the voting procedure is obvious. When a neighborhood system consists of only one tissue type, which we call an ambiguous neighborhood system, some of the regions will inevitably acquire incorrect votes. If \mathcal{N}_{ik} happens to be an ambiguous system in an area where the regions that belong to the same tissue type are densely clustered, the neighborhood systems of \mathcal{R}_{jik} will also be ambiguous. Due to this fact, it will be highly unlikely for the systems of \mathcal{R}_{jik} to compensate for the inaccurate votes that are generated by the ambiguity of \mathcal{N}_{ik} . The following assumptions and observations have been made to construct a remedy for this problem.

1. *Observation:* In areas where regions from the same tissue type are densely clustered, the variation in the direction of image gradient vectors reduces considerably. This fact can be clearly seen in Fig. 4(a). Even though an easy way to quantify this characteristic is to compute the standard deviation of gradient directions in each region, the inherent periodicity of direction angles makes this computation nontrivial. Instead, the following quantity was computed for each region.

$$\Theta_m = \frac{1}{N_{\text{pair}}} \sum_{i \in \mathcal{R}_m} \sum_{j \in \mathcal{R}_m, N_i} \theta_{ij}, \quad (3)$$

where θ_{ij} is the angle between gradient vectors of adjacent pixels \mathbf{x}_i and \mathbf{x}_j and $\theta_{ij} < 180$ deg, and N_{pair} is the number of adjacent pixel pairs in region \mathcal{R}_m . The values of Θ_m , which is the average angle between the gradient vectors of adjacent pixel pairs in region \mathcal{R}_m , are depicted in Fig. 4(b). It is clear that the regions that contain the ambiguous neighborhood systems have lower values of Θ_m .

2. *Assumption:* If a neighborhood system is ambiguous, then the regions inside of it are soft tissue. Note that this assumption is governed by the observations on the results of the tessellation of the image provided by the region-growing stage.

The modified algorithm of the voting procedure, which makes use of Θ_m values, is as follows, where γ_Θ is a pre-determined threshold:

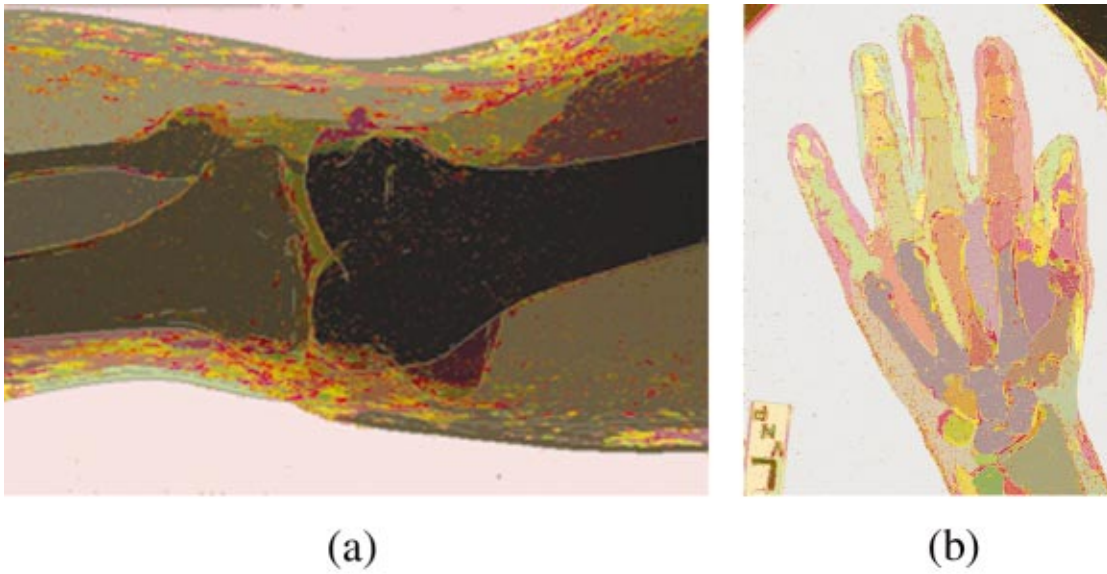


Fig. 3 Results of region growing for (a) knee and (b) hand images. Each color represents a different region.

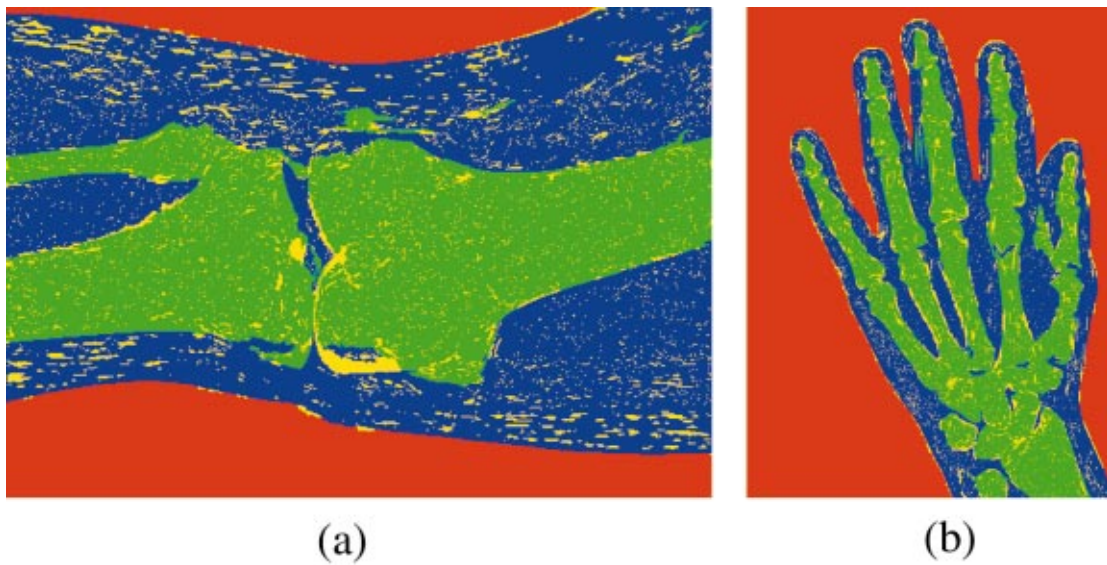


Fig. 5 Classification results at the end of the voting procedure for (a) knee and (b) hand images. The background, bone, soft tissue, and undecided regions are pseudocolored by using red, green, blue, and yellow, respectively.

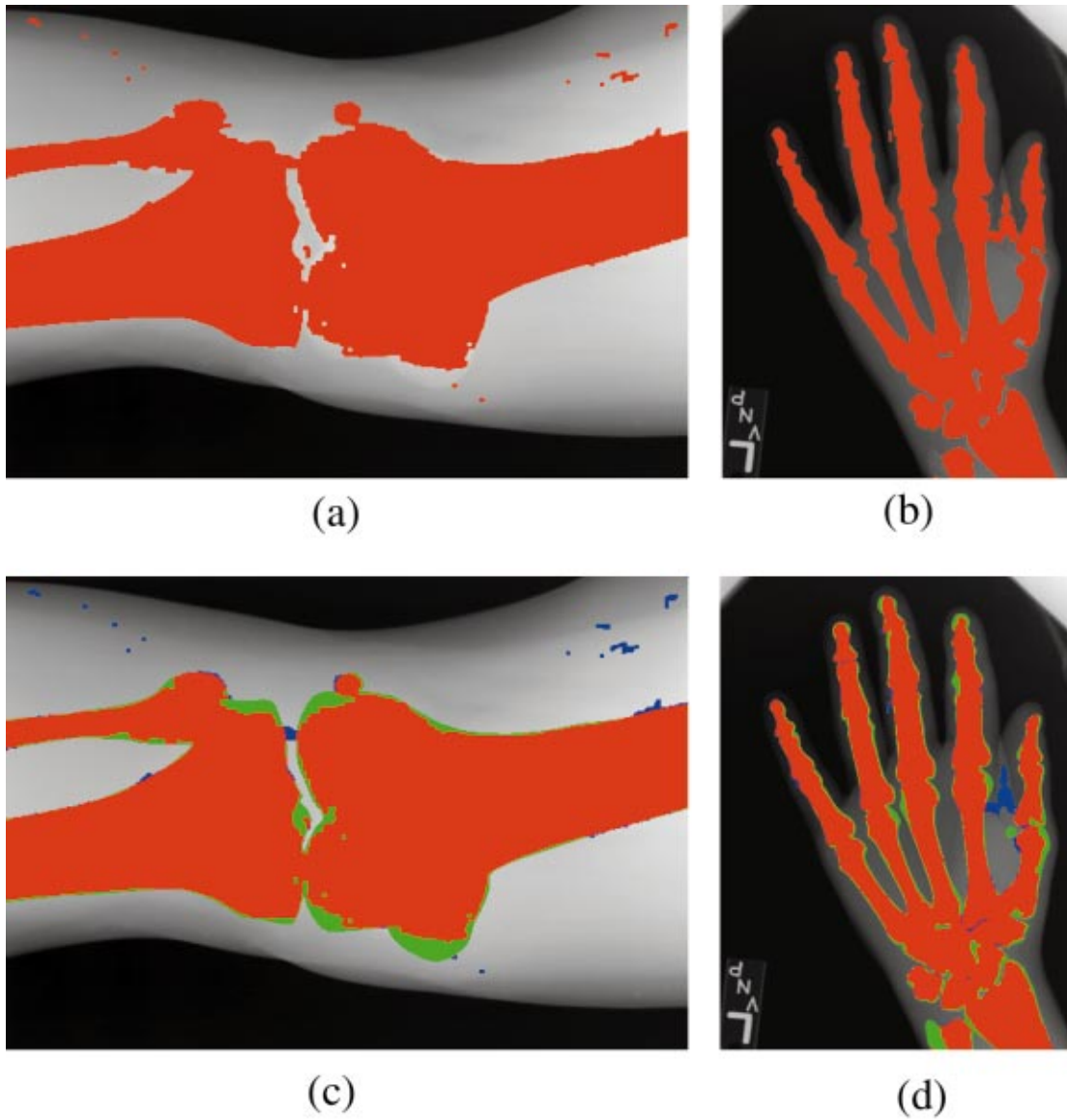


Fig. 6 Final segmentation of bone regions for (a) knee and (b) hand case studies and comparison of the final results with corresponding manual segmentation for (c) knee and (d) hand case studies.

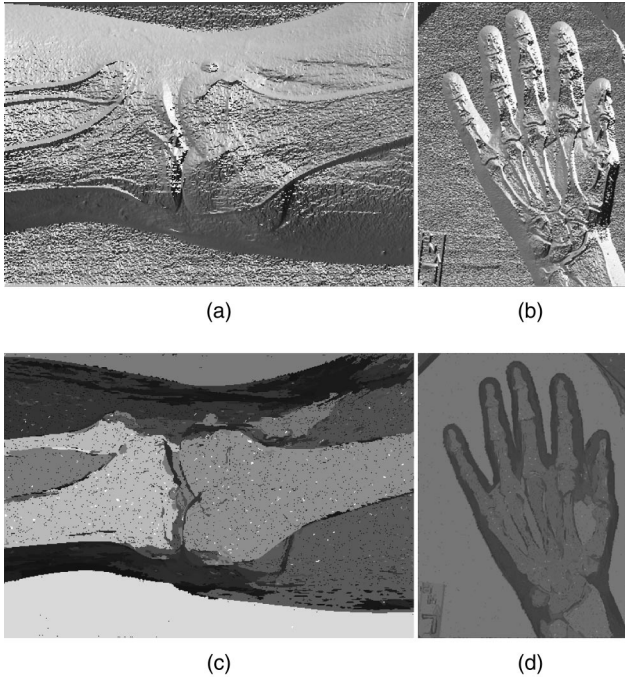


Fig. 4 (a) and (b) directions of image gradient vectors of each pixel are depicted via mapping [0 to 360 deg]→[0 to 255]. Due to the periodicity of angles, pure black and white colors actually correspond to the same direction in this figure. Angles are measured counterclockwise with respect to the positive horizontal axis. It is clear that the pixels which lie in soft tissue regions close to the tissue boundary have similar gradient directions. (c) and (d) Values Θ_m of regions.

Algorithm modified voting.

```

For each  $\mathcal{B}_k$ 
  For each  $\mathcal{R}_{ik}$ 
    Compute the average intensities of  $\{\mathcal{R}_{jik}\}$ .
    If  $\exists \mathcal{R}_{jik}$  such that  $\Theta_{ijk} > \gamma_\Theta$ 
      Cluster  $\{\mathcal{R}_{jik}\}$  using K-means algorithm with
      two clusters.
    For each  $\mathcal{R}_{jik}$ 
      If  $\mathcal{L}_{jik}$  is bone, increment  $\mathcal{V}(\text{B})$  of correspond-
      ing  $\mathcal{R}_m$ .
      If  $\mathcal{L}_{jik}$  is soft tissue, increment  $\mathcal{V}(\text{S})$  of corre-
      sponding  $\mathcal{R}_m$ .
  Else
    For each  $\mathcal{R}_{jik}$ 
      Increment  $\mathcal{V}(\text{S})$  of corresponding  $\mathcal{R}_m$ .
  
```

The algorithm attempts to determine whether the neighborhood system about to be processed is ambiguous by examining the Θ_m values of the regions that it includes. If the answer is yes, the soft tissue votes of all regions are incremented, otherwise regular clustering is employed. Note that, the modified voting algorithm does not search for neighborhood systems that consists of only bone regions. The preceding observation suggests that such a neighborhood system \mathcal{N}_{ik} will be surrounded by systems that includes both tissues, and the regions that acquire incorrect votes during processing of \mathcal{N}_{ik} will be compensated by the processing of surrounding systems. The classification re-

sults at the end of voting procedure for knee and hand case studies can be seen in Figs. 5(a), and 5(b), respectively (see Color Plate 1).

2.3 Refinement of Labels Via Surface Fitting

Even though the modified algorithm for the voting procedure addresses the problem regarding the areas of image where regions that belong to same type of tissue are densely clustered, such areas are still prone to generate misclassifications. Due to the misclassifications and the existence of unlabeled regions, the tissue labels are reevaluated by the last stage of the algorithm. This improves the segmentation result provided by the voting procedure. The last stage attempts to represent bone and soft tissue regions of the image with second-order bivariate polynomials whose coefficients are computed in a least-square (LS) sense. After the computation of the surfaces, root mean square (rms) fitting errors to both surfaces are computed for each region. Then, the label of a particular region is reversed provided that the rms error corresponding to its tissue type is greater than the other. The regions whose labels have been left undecided at the end of the voting procedure are assigned labels during the first iteration.

The algorithm of label refinement with surface fitting step is as follows, where $\epsilon_{\text{B},m}$ and $\epsilon_{\text{S},m}$ are defined as

$$\epsilon_{\text{B},m} = \left\{ \frac{1}{N_m} \sum_{i=1}^{N_m} [\mathcal{P}_{\text{B},k(i)}(i) - \mathcal{R}_m(i)]^2 \right\}^{1/2}, \quad (4)$$

$$\epsilon_{\text{S},m} = \left\{ \frac{1}{N_m} \sum_{i=1}^{N_m} [\mathcal{P}_{\text{S},k(i)}(i) - \mathcal{R}_m(i)]^2 \right\}^{1/2}, \quad (5)$$

where i is the index of the pixels that belong to region \mathcal{R}_m , $k(i)$ is the quadrant in which pixel with index i lies, N_m is the size of region \mathcal{R}_m , and $\mathcal{P}_{\text{B},k}$ and $\mathcal{P}_{\text{S},k}$ are the second-order bivariate polynomials for bone and soft tissue regions in the k 'th quadrant, respectively.

Algorithm refinement of labels via surface fitting.

```

Divide the image into four quadrants along
the principal axes of tissue regions.
Iterate the following steps until no label is changed.
  Compute the surfaces in each quadrant
  for bone,  $\mathcal{P}_{\text{B},k}$ , and soft tissue,  $\mathcal{P}_{\text{S},k}$ ,
  using LS fitting, for  $k = 1..4$ .
  For each  $\mathcal{R}_m$ 
    Compute  $\epsilon_{\text{B},m} = \|\mathcal{P}_{\text{B},k} - \mathcal{R}_m\|_{\text{RMS}}$  and  $\epsilon_{\text{S},m} = \|\mathcal{P}_{\text{S},k} - \mathcal{R}_m\|_{\text{RMS}}$ 
    If the label of  $\mathcal{R}_m$  is bone and  $\epsilon_{\text{B},m} > \epsilon_{\text{S},m}$ 
      Change the label to soft tissue
    If the label of  $\mathcal{R}_m$  is soft tissue and  $\epsilon_{\text{B},m} < \epsilon_{\text{S},m}$ 
      Change the label to bone
  
```

A drawback of the surface-fitting algorithm is that, since it depends on the LS data-fitting concept, it inherently assumes that a label might be incorrect only if the size of its corresponding region is relatively small. A large region has the ability to bend the surface to be fitted toward itself strongly due to the fact that LS data-fitting is very sensitive

Table 1 Summary of the experiments.

Case	Size	Time (s)	γ_M	γ_F	Accuracy (%)
13 (knee)	501×312	59	0.50	1.50	92.7
18 (wrist)	381×556	68	0.55	1.45	89.9
20 (knee)	606×416	91	0.50	1.50	92.5
22 (knee)	526×366	78	0.50	1.50	95.5
23 (knee)	566×341	77	0.50	1.50	96.0
25 (foot)	441×486	79	0.50	1.45	92.8
38 (foot)	596×441	90	0.60	1.50	92.3
47 (knee)	596×286	66	0.55	1.50	95.1
5 (wrist)	426×576	87	0.55	1.50	93.1
52 (ankle)	551×486	101	0.50	1.50	95.1
53 (hand)	451×326	49	0.55	1.45	90.7
55 (ankle)	606×379	80	0.50	1.50	93.8
6 (foot)	381×581	71	0.50	1.50	82.4
9 (ankle)	321×581	74	0.50	1.50	92.5

to noise. Hence, even if a relatively large region is misclassified during the voting procedure, its rms fitting error will turn out to be small and this will make the label correction impossible.

A common misclassification case in the images in which the bone tissue is clustered toward the center (such as knee) is that, some of relatively small regions that are close to the tissue boundary can be labeled erroneously. This is due to the fact that comparing the regions close to the tissue boundary with the bone tissue surface is effectively an extrapolation. To correct such misclassifications, a binary opening operation is applied to both bone and soft tissue regions after the surface fitting stage.

3 Experiments

The performance of the algorithm was tested on 14 extremity images, which included wrist, knee, hand, ankle, and foot exams. Images that were manually segmented under the supervision of a radiologist were used as the gold standard. The original images were in three different sizes, 2500×2048, 2048×2500, and 2392×1792, with 12 bits per pixel.

For the experiments, second level approximation coefficients of the wavelet decomposition which exploits the wavelet filter Daubechies4 were used.¹⁵ After the computation of the approximation coefficients, the resulting images, which are essentially the versions of the originals that are acquired by filtering and downsampling (by 4), are cropped to feed the algorithm with an image that is free of artifacts that occur toward the image boundaries. For each image used in the experiment, blocks of size 64×64 which overlap 32 pixels in all directions were used. Note that this overlapping scheme ensures the inclusion of a particular region, \mathcal{R}_m , in several subimages.

The final segmentation results of the knee and the hand case studies can be seen in Fig. 6 (see Color Plate 2). The red regions in Figs. 6(a) and 6(b) show the bone tissue detected by the algorithm. Also, visual comparison between

the final result and manual segmentation for knee and hand case studies is provided in Figs. 6(c) and 6(d), respectively. In these images, the blue regions indicate the pixels that were classified as bone by the algorithm but as soft tissue during manual segmentation, and vice versa for the green regions.

Table 1, the summary of the experiments, which contains the type of extremities, provides the sizes (after cropping) of the images, the computational time, and the values of thresholds for statistical tests. Note that the algorithm does not employ any kind of training scheme, thus, the parameter values have to be adjusted for each image. The values of parameters that are given in Table I show that for a set of images that possess statistical similarities the optimal values turn out to be very close to each other. This fact clearly makes the parameter tuning process easy since the parameter values of the first image in the set provide a very good initial estimate for the remaining images.

To evaluate the performance of the algorithm statistically, the labeling process can be viewed as a binary hypothesis testing problem. Note that such a view treats background detection as a trivial problem. If the events where a pixel belongs to bone and to soft tissue are assumed as the null and alternate hypothesis, respectively, the performance of the algorithm can be evaluated by computing the ratio of correct detections to the total number of decisions where being correct is defined with respect to the gold standard. In fact, this ratio is the well-known parameter accuracy, which is equal to

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}}, \quad (6)$$

where (TN+TP), FN, and FP are the number of correct decisions, false alarms [the green regions in Figs. 6(c) and 6(d)], and misses [the blue regions in Figs. 6(c) and 6(d)], respectively.

The experiments were conducted on an SGI Indigo2 workstation. We can see from Table 1 that the algorithm provides over 90% accuracy (except an outlier image, case 6) with reasonable computational time. The observer variability has to be taken into consideration when the accuracy results are interpreted. For this purpose, a particular hand image (case 5) is manually segmented 10 times and the accuracy values of those experiments are computed by assuming the original manual segmentation gold standard. The manual segmentations resulted in an average accuracy value of 96.53% with a standard deviation of 0.174. No attempt was made to streamline or optimize the code for a more rapid completion of the analysis.

4 Conclusions

We presented a novel algorithm for the segmentation of CR images of the extremities. The algorithm requires no supervision and performs the segmentation task with reasonable computational complexity. All the parameters of the algorithm, except the threshold values for statistical tests in the region-growing stage, γ_M and γ_F , were kept constant throughout the experiments. This fact essentially makes γ_M and γ_F the only values that must be adjusted for each scan in a set that consists of images that share the same statistical characteristics. Moreover, the experiments showed that the algorithm is robust in parameter space. In other words, the optimal parameter values of γ_M , and γ_F , which are found experimentally for the images that share common statistical properties, are close to each other. This property makes the adjustment of parameter values a relatively easy task.

We believe that the algorithm we propose can be useful for a variety of applications, including image enhancement and automatic recognition of x-ray exam type. Furthermore, the voting procedure can be applied to any classification problem provided that the observations on which the voting procedure depends are valid for the image of concern. Also, the extension of the voting procedure to more than two classes and alternate features is trivial. As future work, the application of the algorithm to 3D data sets and different modalities, and also using fuzzy schemes instead of hard clustering in the voting procedure are being considered.

Acknowledgments

We are grateful for support of this project by Eastman Kodak Health Imaging and the NYS Center for Electronic Imaging Systems. We would also like to thank Thomas Gaborski and Saara Totterman, MD, for their help in the acquisition of manual segmentations that were used in the segmentation evaluation and observer variability studies.

References

1. J. Capozzi and R. Schaetzing, "Method and apparatus for automatic tonescale generation in digital radiographic images," U.S. Patent 5,164,993 (1992).
2. L. Barski, R. Gaborski, and P. Anderson, "A neural network approach to the segmentation of digital radiographic images," in *ANNIE 93 Artificial Neural Networks in Engineering*, St. Louis, MO (1993).
3. K. Haris, S. N. Efstratiadis, N. Maglaveras, and A. K. Katsaggelos, "Hybrid image segmentation using watersheds and fast region merging," *IEEE Trans. Image Process.* **7**, 1684–1699 (1998).
4. M. Nadler and E. P. Smith, *Pattern Recognition Engineering*, John Wiley & Sons, New York (1993).
5. S. Geman and D. Geman, "Stochastic relaxation, gibbs distribution, and the bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-6**, 721–741 (1984).
6. J. Besag, "On the statistical analysis of dirty pictures," *J. R. Stat. Soc. Ser. B. Methodol.* **48**, 259–302 (1986).
7. T. N. Pappas, "An adaptive clustering algorithm for image segmentation," *IEEE Trans. Signal Process.* **40**, 901–914 (1992).
8. J. G. Tamez-Pena, "Four-dimensional reconstruction and visualization of complex musculoskeletal structures," PhD Thesis, Univ. of Rochester (1999).
9. V. K. Rohatgi, *An Introduction to Probability Theory and Mathematical Statistics*, John Wiley & Sons, New York (1976).
10. Y. L. Chang and X. Li, "Adaptive image region-growing," *IEEE Trans. Image Process.* **3**, 868–872 (1994).
11. R. Adams and L. Bischof, "Seeded region growing," *IEEE Trans. Pattern Anal. Mach. Intell.* **16**, 641–647 (1994).
12. P. K. Saha, J. K. Udupa, and D. Odhner, "Scale-based fuzzy connected image segmentation: theory, algorithms, and validation," *Comput. Vis. Image Underst.* **77**, 145–174 (2000).
13. G. L. Foresti, V. Murino, C. S. Regazzoni, and A. Trucco, "A voting-based approach for fast object recognition in underwater acoustic images," *IEEE J. Ocean. Eng.* **22**, 57–65 (1997).
14. L. Lam and C. Y. Suen, "Application of majority voting to pattern recognition: An analysis of ots behavior and performance," *IEEE Trans. Syst. Man Cybern.* **27**, 553–568 (1997).
15. I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.* **41**, 909–996 (1988).



S. Kubilay Pakin received his BS and MS degrees from the Department of Electrical and Electronics Engineering, Bilkent University, Ankara, both with highest honors, in 1995 and 1997, respectively. He is currently pursuing his PhD degree in electrical and computer engineering at University of Rochester, New York. During his MS work, he studied statistical methods for estimating the parameters of superimposed signals. His current research interests are medical imaging, 2-D and 3-D image segmentation, and deformable models. He is a member of IEEE.



Roger S. Gaborski is a professor and graduate chair in the Department of Computer Science, Rochester Institute of Technology, New York. His research interests include medical imaging, pattern recognition and machine learning. He received his PhD in electrical engineering from the University of Maryland.



Lori L. Barski is a research associate with Eastman Kodak Company's Health Imaging Research Laboratory (HIRL). She has a BS degree in mechanical engineering and a MS in mathematics and applied statistics from Rochester Institute of Technology and has worked in the field of digital image processing since 1985. Her current research focus is on image enhancement and image analysis for computed radiography (CR) and direct radiography (DR) imagery. She holds numerous patents related to document processing and medical imaging applications and has been a member of HIRL for 9 years.



David H. Foos is head of the Image Science Research Laboratory for Health Imaging at Eastman Kodak Company. He oversees a variety of research projects related to digital radiographic image processing and display. Mr. Foos joined Kodak in 1983 after graduating from Rensselaer Polytechnic Institute with an MS in physics. He has authored papers on medical image compression, image display processing for computed radiography, and diagnostic quality measurement.



Kevin J. Parker received the BS degree in engineering science, summa cum laude, from the State University of New York at Buffalo in 1976, and his MS and PhD degrees in 1978 and 1981 from the Massachusetts Institute of Technology for work in electrical engineering with a concentration in bioengineering. Dr. Parker is a professor of electrical and computer engineering, radiology, and bioengineering at the University of Rochester, where he has held positions since 1981. In 1998, Dr. Parker was named dean of the School

of Engineering and Applied Sciences at the University of Rochester. Dr. Parker has received awards from the National Institute of General Medical Sciences (1979), the Lilly Teaching Endowment (1982), the IBM Supercomputing Competition (1989), the World Federation of Ultrasound in Medicine and Biology (1991), and the Joseph P. Holmes Pioneer Award from the AIUM (1999). He is a member of the IEEE, the Acoustical Society of America (ASA), and the American Institute of Ultrasound in Medicine (AIUM). He was named a fellow of both the IEEE and the AIUM for his work in medical imaging and of the ASA for his work in acoustics. In addition, he was on the Board of Governors of the AIUM for a three-year term. Dr. Parker's research interests are in medical imaging, linear and nonlinear acoustics, and digital halftoning.