

# A Variational Approach to Bone Segmentation in CT Images

Jeff Calder<sup>a</sup> and Amir M. Tahmasebi<sup>b</sup> and Abdol-Reza Mansouri<sup>a</sup>

<sup>a</sup>Dept. of Mathematics and Statistics, Queen's University, Kingston, ON, K7L 3N6, Canada;

<sup>b</sup>Rotman Research Institute, University of Toronto, Toronto, ON, M6A 2E1, Canada

## ABSTRACT

We present a variational approach for segmenting bone structures in Computed Tomography (CT) images. We introduce a novel functional on the space of image segmentations, and subsequently minimize this functional through a gradient descent partial differential equation. The functional we propose provides a measure of similarity of the intensity characteristics of the bone and tissue regions through a comparison of their cumulative distribution functions; minimizing this similarity measure therefore yields the maximal separation between the two regions. We perform the minimization of our proposed functional using level set partial differential equations; in addition to numerical stability, this yields topology independence, which is especially useful in the context of CT bone segmentation where a bone region may consist of several disjoint pieces. Finally, we present an extensive validation of our method against expert manual segmentation on CT images of the wrist, ankle, foot, and pelvis.

**Keywords:** Computed Tomography, Variational Methods, Bone Segmentation, Level Sets

## 1. INTRODUCTION

Segmentation of bone structures from Computed Tomography (CT) images has many applications to image-based computer assisted orthopedic surgery. Such segmentations can be very challenging to obtain since osseous tissue does not always yield readily distinguishable features from soft tissue regions in CT images.<sup>1-5</sup> Indeed, osseous tissue on the surface of bones (cortical bone) is very dense and strong, and, as a result, has a larger electromagnetic absorption coefficient than the weaker osseous tissue in the interior region of bones (cancellous bone). Hence, cortical bone appears as a thin bright rim surrounding the darker region corresponding to the cancellous bone on CT images. Complicating matters further, cortical bone contrast with soft tissue and with cancellous bone can be highly uneven for the same bone, leading to faint intensity boundaries between the regions. Furthermore, the trabeculae of cancellous bone induce a certain texture pattern in the interior of the bone regions leading to large intensity variations in regions corresponding to cancellous bone.

Manual segmentation of bone regions in CT images is tedious, time consuming, and subject to observer variability; as a result, there is a strong demand for automating the segmentation procedure. Despite considerable research on the segmentation problem, existing solutions to bone segmentation in CT images still remain inadequate for systematic deployment and operation in a clinical setting chiefly due to drawbacks such as difficulties in algorithm initialization and parameter settings, as well as a lack of robustness to image inhomogeneities or noise. In a recent survey, Wang et al.<sup>1</sup> have classified the existing segmentation techniques into four categories: intensity-based,<sup>3,6-8</sup> edge-based,<sup>4,9</sup> region-based,<sup>10,11</sup> and deformable.<sup>4,5,12</sup>

Intensity-based approaches refer to the simplest segmentation techniques that are based on local or global intensity thresholding. Thresholding approaches require homogeneity of the objects being segmented, which is a patently invalid assumption for bone segmentation. Bones, as described above consist of several inner structures, which causes the intensity of bone pixels to vary in a wide range and sometimes overlap with that of other tissue types such as fat and muscle. Edge-based techniques, on the other hand, try to extract bone contour points using edge-detection filters (such as the Canny edge filter<sup>13</sup>), and then attempt to reconstruct the boundary of the bone by suitably connecting together these boundary points.

In region-based techniques, on the other hand, it is global region information which guides the segmentation, and the goal is to partition the image into regions based on some given homogeneity condition. Initialized seeds

---

*Email addresses:* jcalder@umich.edu, atahmasebi@rotman-baycrest.on.ca, mansouri@mast.queensu.ca.

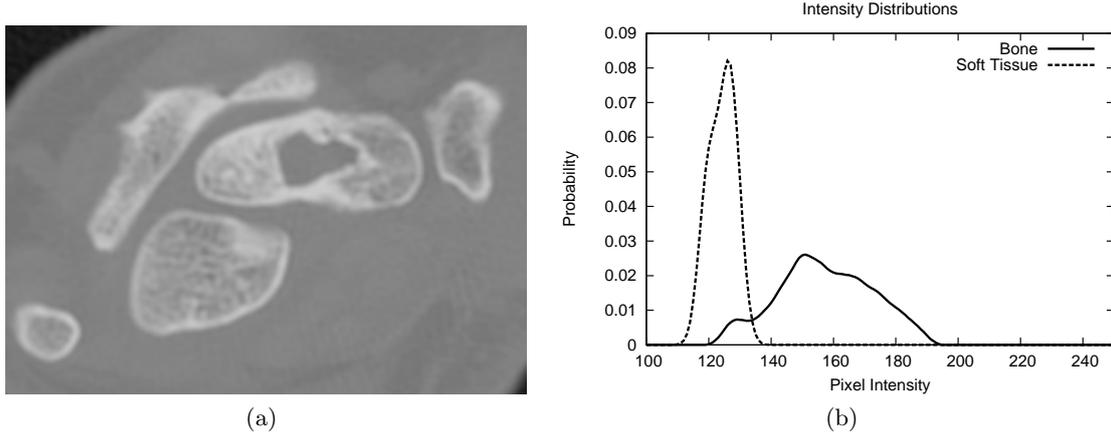


Figure 1. (a) Typical CT image (wrist), (b) Empirical intensity distributions of the soft tissue and bone regions obtained from manual segmentation.

grow by incorporating similar pixels, where similarity is defined by a set of statistical tests. Adams et al.<sup>11</sup> proposed an improvement to standard region growing by introducing a “competition” between growing regions through ordering of all candidate pixels according to some suitability condition. In addition to the fact that it requires seed initialization, their proposed technique makes no use of any information regarding the geometry of the object being segmented. This complete absence of any *a priori* geometric information on the region being segmented leads both to “leaks” through gaps or weak edges, and to the merging of bones that are located very close to each other.

Since the initial introduction of active contours by Kass et al.,<sup>12</sup> variational methods have played a key role in image segmentation, as they allow the design of segmentation algorithms that incorporate both the local features of edge-based algorithms and the global features of region-based algorithms. In the variational approach, the desired segmentation is formulated as the solution of a functional minimization problem. The functional to be minimized is a mapping from a set containing all allowed segmentations to the real line; that set varies depending on the assumptions that are made on the desired segmentation (*e.g.*, differentiability properties of the regions to be found). In addition, the functional minimization problem can often be recast as a Maximum A Posteriori (MAP) estimation problem, and, as a result, additional priors on the segmentation can be additively combined with the functional. The ease with which *a priori* information can be incorporated into segmentation algorithms in the variational approach is one of the key advantages of that approach. Such an advantage is not negligible in the context of bone and soft tissue segmentation in CT images, where the location of a given scan is often a good predictor of the geometry of the segmentation obtained. Indeed, many anatomical structures have distinct bone geometries, and since the anatomical location of the scan is known *a priori*, this information could be used to bias the segmentation towards those geometries.

In this work, we define a functional that can serve as the core functional for any variational approach to bone segmentation in CT images. The proposed functional can be easily augmented to include *a priori* information on region shape, region texture, or other characteristics, should they be available.<sup>14–17</sup> Incorporating further *a priori* information can only improve the quality of these segmentations.

## 2. PROBLEM FORMULATION AND DEFINITION OF FUNCTIONALS

Consider the CT image shown in figure 1(a); the soft tissue appears as an almost uniform dark gray background, while the osseous tissue appears as two relatively distinct parts: a cancellous core, appearing in the interior of the bone regions as a textured light gray pattern, and a cortical part, appearing as a bright rim near the boundary of the bone regions. Furthermore, as can be seen on the same image, the cortical bone does not appear with the same intensity on the same bone structure, resulting in widely varying contrast between the bone region and the soft tissue background. Figure 1(b) shows the empirical intensity distribution within the bone and soft tissue regions obtained by a manual segmentation of the CT image in figure 1(a); it is clear that

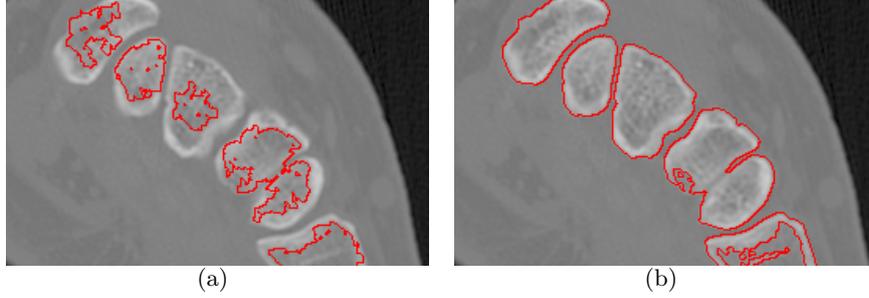


Figure 2. (a) Geodesic active contours and (b) Chan–Vese active contours applied to a CT image.

the empirical distribution of the bone region is quasi-bimodal (corresponding to cortical and cancellous bone tissue) and hence far from Gaussian. Furthermore, the supports of the two empirical distribution functions have substantial overlap, due to the fact that cancellous bone tissue and soft tissue can have relatively close absorption characteristics. It is these characteristics which make bone automatic segmentation in CT images such a challenging task.

## 2.1 Shortcomings of popular variational and level set based methods

One of the most widely used level set methods is the Chan–Vese active contours without edges.<sup>18</sup> The functional proposed there is given by

$$R \mapsto E[R; I] = \int_R (I(\mathbf{x}) - \mu_R)^2 d\mathbf{x} + \int_{R^c} (I(\mathbf{x}) - \mu_{R^c})^2 d\mathbf{x} + \lambda \oint_{\partial R} ds,$$

where  $\mu_R$  is the mean value of the image function  $I$  within the region  $R$ . This method has proven very popular and quite successful on certain types of images, however, as can be seen from the functional itself (see also the discussion in<sup>18</sup>), this functional is designed for an image which is piecewise constant with additive stationary white Gaussian noise. This is clearly not a valid assumption for CT images where extreme intensity inhomogeneity is typically observed between cancellous and cortical bone. Another popular level set approach is the geometric active contours approach:<sup>19,20</sup>

$$\frac{\partial u}{\partial t} = g(\mathbf{x}) \|\vec{\nabla} u\| \left( \operatorname{div} \left( \frac{\vec{\nabla} u}{\|\vec{\nabla} u\|} \right) + \nu \right),$$

where  $u$  is the level set function,  $\nu$  is a positive constant, and  $g$  is an “edge detection” function. A typical choice for  $g$  is  $g(\mathbf{x}) = 1/(1 + \|\vec{\nabla} \hat{I}\|^p)$ , where  $\hat{I}$  is a smoothed version of the image function  $I$  and  $p > 1$ . Of course, the above level set evolution is equivalent to the curve evolution equations:

$$\frac{\partial \vec{\gamma}}{\partial t} = g(\mathbf{x}) (\kappa - \nu) \vec{N}$$

where  $\kappa$  is the curvature of  $\vec{\gamma}$  and  $\vec{N}$  denotes the inward pointing unit normal vector field to  $\vec{\gamma}$ . It should be noted that this evolution equation is not necessarily the descent equation for a functional, and hence, depending on the choice of  $g$ , this approach may or may not correspond to a variational approach. The idea behind this method is that the contour will grow until reaching the edges of the object at which point  $g \approx 0$  and the evolution stops. This method is useful for extracting objects with strong edges and relatively smooth interiors. In the context of CT images, the boundary between bone and soft tissue is often far from being uniformly strong, and hence the assumption of a strong edge cannot be used as the sole guiding principle for bone segmentation. Furthermore, this evolution equation uses only local boundary information and does not make use of any global region information to supplement the local information.

In figure 2, we show examples of both of the above mentioned algorithms applied to a typical CT image. Note in 2(a) that the lack of homogeneity within the cancellous bone stops the region growing geodesic active contours method before the contour reaches the bone boundary. In 2(b) we note that many bone edges are missed and the segmentation is unable to fully separate two neighbouring bones.

## 2.2 Shortcomings of probability density function separation

Functionals incorporating some similarity measure between the probability density functions (PDF) of the region and background have recently been proposed for region tracking,<sup>21</sup> and also specifically for bone segmentation in CT images.<sup>22</sup> For the application of region tracking, the goal is to match an object with an *a priori* distribution whereas in segmentation, the goal is typically to maximally separate the intensity distributions of the region and background. Here, we analyze such a functional in the specific context of CT image segmentation and show that it has undesirable properties which make it too sensitive to local minima.

Let now  $\Omega \subset \mathbb{R}^2$  denote the image domain,  $I : \Omega \rightarrow \mathbb{R}$  the CT image function, and let  $R \subset \Omega$  denote a subset of the image domain, and  $R^c$  its complement. Consider the functional  $E_{pdf} : R \mapsto E_{pdf}[R; I]$  defined by:

$$E_{pdf}[R; I] = - \int_{\mathbb{R}} (p_R(z) - p_{R^c}(z))^2 dz, \quad (1)$$

where  $p_R$  denotes the empirical intensity distribution of  $I$  within  $R$ , and  $p_{R^c}$  the empirical intensity distribution of  $I$  within  $R^c$ . Minimizing  $E$  with respect to  $R$  is equivalent to finding the region which maximizes the  $L^2$  distance between the empirical distribution functions of its interior and exterior. Figure 3 shows the result of segmenting the CT image in figure 1(a) by minimizing the functional  $R \mapsto E_{pdf}[R; I] + \lambda E_{length}[R]$ , where  $E_{pdf}$  is given in equation (1), and  $E_{length}$ , given by  $R \mapsto E_{length}[R] = \oint_{\partial R} ds$  is a measure of the length of the boundary  $\partial R$  of  $R$ . Incorporating  $E_{length}$  in the overall functional has a regularizing effect on the segmentation obtained by penalizing regions with excessively “jagged” boundaries, and the degree of this regularization is controlled by the scalar  $\lambda > 0$ . Figure 3 demonstrates an important flaw in probability density function separation: the gradient descent on  $E_{pdf} + \lambda E_{length}$  is highly susceptible to local minima. In this case, the algorithm partially segments some bone edges, but also segments a slight variation in intensity within the soft tissue which is clearly undesirable.

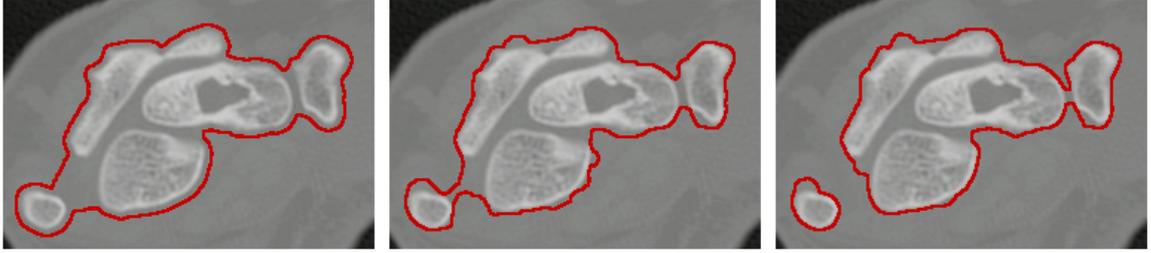
This is attributed to the fact that functionals based solely on probability distribution separation (such as the Kullback-Leibler or Bhattacharya distances) are minimized when the overlap in the support of the respective distributions is minimized. This is generally not a sufficient criterion for segmentation and any algorithm including a probability distribution separation term is likely to include auxiliary terms to compensate for this drawback. Although such addition of auxiliary terms could certainly bias the segmentation towards intensity boundaries,<sup>22</sup> our aim here is to illustrate and understand the basic limitations of probability density function separation. Truc et al.<sup>22</sup> use a convex combination of the Bhattacharya distance and a likelihood term. The likelihood term they use is optimal when the image is piecewise constant with additive Gaussian noise. In light of figure 1(b), we see that the background is well-approximated by a Gaussian distribution, but the bone tissue is clearly far from Gaussian. Thus, we need a solution that does not make any assumptions on the shape of the intensity distribution within the bone yet is not as susceptible to local minima as  $E_{pdf}$ .

## 2.3 Shortcomings of mean separation

Recall that the key drawback of the probability density function separation functional  $E_{pdf}$  is that as long as the probability density functions  $p_R$  and  $p_{R^c}$  have disjoint supports, the value of  $E_{pdf}[R; I]$  remains unchanged under shifting of  $p_R$  or  $p_{R^c}$  on the real axis; in other words, as long as  $p_R$  and  $p_{R^c}$  have disjoint supports, the value of  $E_{pdf}[R; I]$  is independent of the first moments of  $p_R$  and  $p_{R^c}$ . With this in mind, consider now the functional  $E_{mean} : R \mapsto E_{mean}[R; I]$ , first proposed by Yezzi et al.,<sup>23</sup> and defined by

$$E_{mean}[R; I] = -(\mu_R - \mu_{R^c})^2 \quad (2)$$

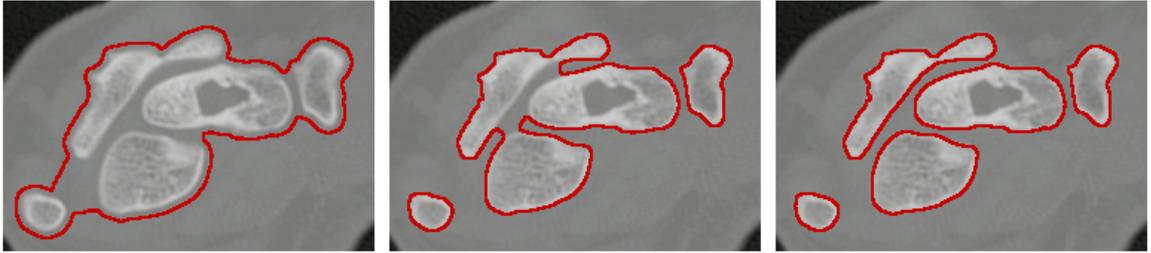
where  $\mu_R$  denotes the mean of the empirical intensity distribution of  $I$  within  $R$  and  $\mu_{R^c}$  that of  $R^c$ . Minimizing the functional  $R \mapsto E_{mean}[R; I] + \lambda E_{length}[R]$ , therefore yields a segmentation with maximal separation between empirical means of the two regions. The result of segmenting the CT image in figure 1(a) by minimizing the functional in Equation (2) is shown in figure 3. From figure 3, we can see that mean separation is less susceptible to local minima than PDF separation (ie: the segmentation is immediately driven towards the bone). However, the final segmentation contains only the brightest white pixels in the cortical region of the bone. This is because the functional is considering only the difference of means. Hence it is a function of only the first moment of



(a) Segmentation obtained by minimizing  $E_{pdf} + \lambda E_{length}$ .



(b) Segmentation obtained by minimizing  $E_{mean} + \lambda E_{length}$ .



(c) Segmentation obtained by minimizing  $E_{cdf} + \lambda E_{length}$ .

Figure 3. Results of segmentation under various functionals. From left to right: initial contour, intermediate contour, and final segmentation after convergence. All contours are shown in red.

the empirical distribution and does not consider any higher moments. One could imagine that if the boundaries of the bone were uniformly bright, then separation by means would find the correct segmentation (as a local minima). This is likely correct, but we cannot assume that the boundaries of bones are strong. In fact, it is quite common to see weak boundaries and this is a recurrent problem in CT image segmentation.<sup>1,5</sup> Again, similar to the case with probability density function separation, we could append terms (such as uniformity priors on the soft tissue region) to this functional to attempt to rectify the situation, but our goal is to study the behaviour of these functionals alone. Our proposed functional, described in the next section, will borrow features from both  $E_{mean}$  and  $E_{pdf}$  and in doing so, will benefit from the positive features of the respective functionals.

## 2.4 Proposed Functional

In light of the discussion in Section 2.2, we propose the following basic energy functional for bone/soft tissue segmentation in CT images:

$$R \mapsto E_{cdf}[R; I] = - \int_{\mathbb{R}} (c_R(z) - c_{R^c}(z))^2 dz, \quad (3)$$

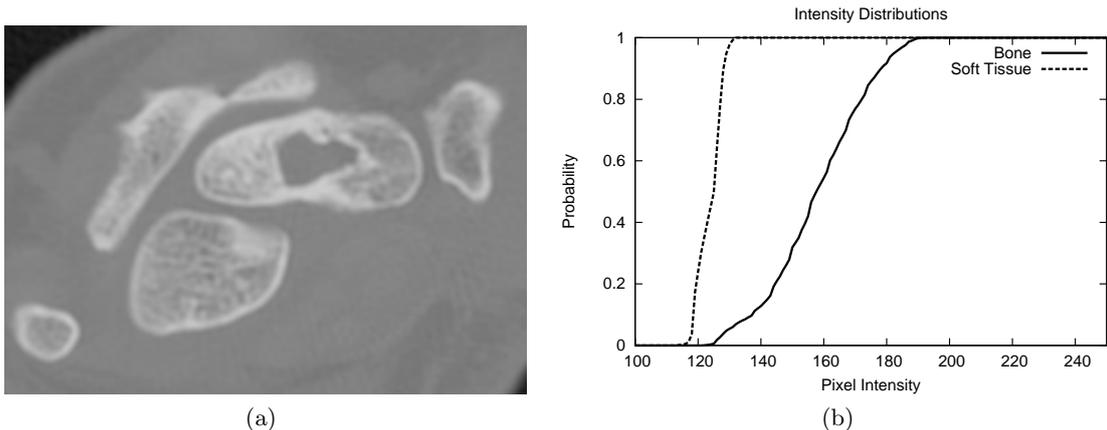


Figure 4. (a) Typical CT image, (b) Empirical cumulative intensity distributions of the soft tissue and bone regions obtained from manual segmentation.

where  $c_R : z \mapsto c_R(z)$  denotes the *cumulative* empirical distribution function of  $I$  within  $R$ , and  $c_{R^c}$  the cumulative empirical distribution function of  $I$  within  $R^c$ . More precisely, we have:

$$c_R(z) = \frac{\int_R \chi_z(I(x)) dx}{\int_R dx}, \quad \text{and} \quad c_{R^c}(z) = \frac{\int_{R^c} \chi_z(I(x)) dx}{\int_{R^c} dx}, \quad (4)$$

where  $\chi_z(t) = 1$  for  $t \leq z$  and  $\chi_z(t) = 0$  otherwise. Other variational approaches incorporating region statistics have been proposed recently. In the work by Chan et al.<sup>24</sup> local histograms are compared to the region and background histograms using the  $L^1$  distance between their cumulative distribution functions (CDF). This approach can be easily understood as an extension of Chan–Vese active contours;<sup>18</sup> the idea being to segment the image into regions such that the local histograms around each point best match the global histograms of the region. There is an implicit homogeneity assumption in this approach; *i.e.* the local histograms are assumed to be homogeneous within the object. Since CT images typically exhibit extreme intensity variations between cortical and cancellous bone this assumption is clearly invalid.

Truc et al.<sup>22</sup> propose to segment CT images with a convex combination of the Bhattacharya distance and a Chan–Vese energy, and Freedman et al.<sup>25</sup> propose to segment medical imagery by matching *a priori* distributions with a CDF energy. In the former, only the probability density functions (PDF) of the region and background are compared which has the drawback that two PDFs are considered maximally separated as soon as their supports have no overlap, regardless of the separations of their means. In the latter, an *a priori* probability distribution must be supplied to the segmentation algorithm.

As a remark, the functional proposed in (3) is related to the well-known Earth Mover’s Distance (EMD).<sup>24</sup> Informally, if we interpret each probability distribution as a pile of dirt, the EMD is defined to be the minimum amount of work required to turn one pile into the other. There is no closed form expression for the EMD and hence the corresponding minimization problem is not tractable in the variational framework. However, the EMD is closely related to the Wasserstein distance (first proposed for segmentation by Chan et al.<sup>24</sup>) which leads to a far more tractable minimization problem. The Wasserstein distance is the  $L^1$  distance between the cumulative distribution functions whereas we have chosen the squared  $L^2$  distance. Although our functional is not exactly the EMD, it is useful to think of it in this light as it has many of the same properties such as its dependence on the separation of means. We can see this in figure 4 which shows the cumulative distribution functions for the probability distribution functions shown in figure 1(b). Clearly, the  $L^2$  distance between the two cumulative distribution functions increases as the means of the two distribution functions get further and further separated; thus we expect this functional to retain some of the properties of  $E_{mean}$ , in particular, to be less susceptible to local minima. Furthermore, since our functional makes use of the entire empirical distribution inside and outside the region, we expect this functional to be better suited to resolving bone edges than  $E_{mean}$ . Figure 3 shows the

result of segmenting the CT image in figure 1(a) by minimizing the functional

$$R \mapsto E_{cdf}[R; I] + \lambda E_{length}[R];$$

This result suggests, in agreement with our previous analysis, that the functional  $E_{cdf}$  does not suffer from the main shortcomings of either  $E_{pdf}$  (susceptibility to local minima) or  $E_{mean}$  (dependence on the distributions only through their first moments). The functional  $E_{cdf}$  is a measure of the separation of the cumulative distribution functions of the intensity function within a region and its complement, and as such is a global measure; but it is only a relative measure on those two cumulative distributions, and no information about what the cumulative distribution inside or outside of a bone region should look like is included in that functional. Such *a priori* information on the cumulative distribution function inside and outside bone regions can be obtained by compiling statistics on bone and soft tissue regions, and can be easily added to  $E_{cdf}$ , as we have discussed previously. We shall not do this here as this is not our goal; rather, as explained before, our goal is to propose a very basic functional that nevertheless yields very accurate bone segmentations while making minimal assumptions on bone and soft tissue region characteristics, be they related to shape, texture, intensity, etc.; further assumptions can then, as needed, be incorporated to the core functional we shall propose.

As observed above, the functional  $E_{cdf}$  is a global measure of the region characteristics; we shall augment it additively by another functional in order to also make use of local characteristics of bone regions; in particular, we shall exploit the fact that bone regions often (but not always) have non-negligible image intensity boundaries, with inward-pointing intensity gradient vector fields. We will incorporate this local information by augmenting  $E_{cdf}$  with the functional  $E_{flux}$  given by

$$R \mapsto E_{flux}[R; I] = - \oint_{\partial R} \vec{\nabla} I \cdot \vec{N} ds \quad (5)$$

where  $\vec{N}$  denotes the unit inward normal vector field to  $\partial R$  and  $\vec{\nabla} I$  the gradient of the image function  $I$ . The term  $E_{flux}$  is called a flux maximizing term.<sup>26</sup> Combining  $E_{cdf}$  and  $E_{flux}$  with the standard penalty term  $E_{length}$  on the length of region boundaries yields the functional

$$R \mapsto E[R; I] = - \int_{\mathbb{R}} (c_R(z) - c_{R^c}(z))^2 dz - \alpha \oint_{\partial R} \vec{\nabla} I \cdot \vec{N} ds + \lambda \oint_{\partial R} ds \quad (6)$$

where  $\partial R$  denotes the boundary of  $R$ , and  $\lambda, \alpha \geq 0$  are weighting coefficients.

### 3. IMPLEMENTATION

We derive the Euler-Lagrange equations for (6) in the standard way and obtain the following level set evolution equations:

$$\begin{aligned} \frac{\partial u}{\partial t}(\vec{x}) &= \left[ 2 \left( \frac{1}{A_{R_{\vec{\gamma}}}} + \frac{1}{A_{R_{\vec{\gamma}}^c}} \right) \int_{\mathbb{R}} \chi_z(I(\vec{x})) (c_{R_{\vec{\gamma}}} - c_{R_{\vec{\gamma}}^c}) dz \right. \\ &\quad \left. - 2 \int_{\mathbb{R}} (c_{R_{\vec{\gamma}}} - c_{R_{\vec{\gamma}}^c}) \left( \frac{c_{R_{\vec{\gamma}}}}{A_{R_{\vec{\gamma}}}} + \frac{c_{R_{\vec{\gamma}}^c}}{A_{R_{\vec{\gamma}}^c}} \right) dz - \alpha \Delta I + \lambda \kappa \right] \|\vec{\nabla} u\| \end{aligned} \quad (7)$$

where  $u$  is the level set function,  $\vec{\gamma}$  is the zero level set of  $u$ ,  $\kappa$  is the curvature of  $\vec{\gamma}$ ,  $R_{\vec{\gamma}}$  is the region bounded by  $\vec{\gamma}$  and  $A_{R_{\vec{\gamma}}}$  and  $A_{R_{\vec{\gamma}}^c}$  denote the area of  $R_{\vec{\gamma}}$  and  $R_{\vec{\gamma}}^c$ , respectively. Our algorithm was implemented in C++ and tested on a laptop with a 2 GHz processor and 2 GB of RAM. The standard discretization of the level set equation (7) was used<sup>27</sup> and the time increment was set at each iteration based on the CFL condition to ensure the stability of the PDE.<sup>27, 28</sup> The algorithm was iterated until the area inside the region  $R_{\vec{\gamma}}$  stabilized indicating convergence of the segmentation procedure. The pseudocode of the level set method is shown in algorithm 1.

As with any segmentation algorithm, there are parameters which need to be adjusted to obtain a correct segmentation. In this case, there are two parameters  $\alpha$  and  $\lambda$  from equation (6). By inspection of the functional (6), it is clear that the gradients of  $E_{cdf}$ ,  $E_{flux}$  and the boundary penalty term will all have different orders

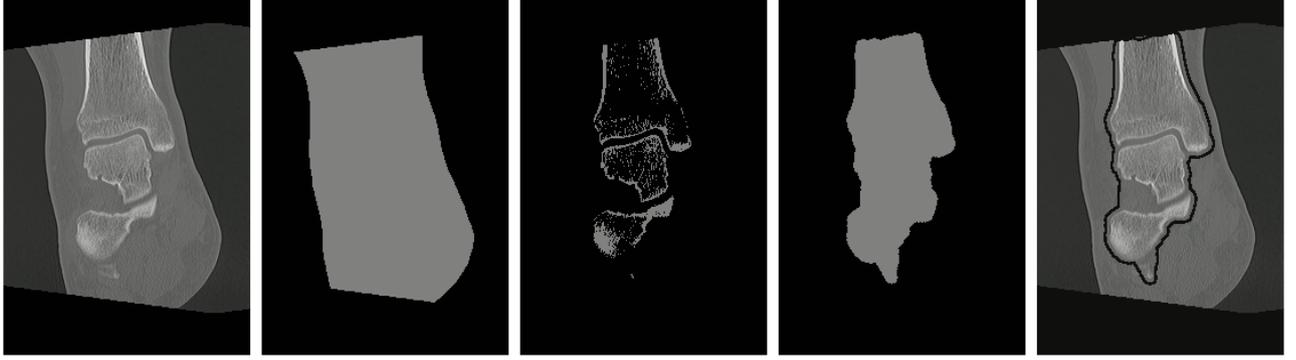


Figure 5. Overview of initialization algorithm, from left to right: Original image, region of interest mask, global thresholding, initialization mask, initialization contour.

of magnitude. In order to simplify the use of our algorithm, we decided on some scaling factors which act to normalize these terms so that the values of  $\alpha$  and  $\lambda$  can be set to reasonable values. The scaling factors for  $E_{cdf}$ ,  $E_{flux}$  and the length penalty were experimentally determined to be 12 500, 250 and 1 respectively.

The execution time of the proposed algorithm depends on many factors such as the resolution of the image, the size of the region of interest and the number of iterations required for convergence. Furthermore, the proposed functional is a function of the statistics of the image, and hence a well defined background or region of interest, to which the foreground's statistics will be compared, must be defined. It is desirable to choose a small, yet physically meaningful, region of interest surrounding the bone and initialize the zero level set as close to the desired segmentation as possible in order to avoid entrapment in spurious local minima of the functional. Such an initialization could certainly be done manually, however in order to have a fully automatic segmentation procedure, it is essential to have a fully automatic initialization procedure as well. In what follows, we detail a fully automatic initialization technique which determines the region of interest and initial contour so as to both ensure an accurate segmentation and reduce the execution time of the algorithm.

---

**Algorithm 1** Compute Segmentation

---

- 1:  $u :=$ Level Set Function
  - 2:  $I :=$ Image
  - 3:  $\Omega :=$  Image Domain
  - 4: Initialize Level Set Function  $u$
  - 5: **repeat**
  - 6:     Compute  $c_{R_{\bar{\gamma}}}, c_{R_{\bar{\gamma}}^c}$  using equation (4)
  - 7:     **for**  $p \in \Omega$  **do**
  - 8:         Compute  $\frac{\partial u}{\partial t}(p)$  from equation (7)
  - 9:         Keep track of the maximum  $\frac{\partial u}{\partial t}(p)$
  - 10:     **end for**
  - 11:      $\Delta t = 0.5/max$
  - 12:     **for**  $p \in \Omega$  **do**
  - 13:          $u(p) := u(p) + \Delta t \times \frac{\partial u}{\partial t}(p)$
  - 14:     **end for**
  - 15: **until** Convergence conditions met
- 

The initialization procedure is divided into three steps which are shown in figure 5. In the first step, the region of interest containing all the soft tissue and bone is determined. This is a relatively simple task that can be done with a global thresholding algorithm. Subsequently, the region containing the bones is roughly determined by classifying pixels by their intensity. The brightest 10% of the pixels are designated as bone; this figure is a safe upper bound, based on empirical observations, on the relative cortical bone area in a typical CT scan. The

last step of the initialization procedure is based on the contour reconstruction technique from.<sup>29</sup> Let  $S$  denote the set of bone pixels designated in step 2 and consider the curve evolution equations

$$\frac{\partial \vec{\gamma}}{\partial t} = - \left( \vec{\nabla} d(x, S) \cdot \vec{N} \right) \vec{N} \quad (8)$$

where  $d(\cdot, S) : \Omega \rightarrow \mathbb{R}$  is the distance function defined by  $x \mapsto d(\mathbf{x}, S) = \inf_{\mathbf{y} \in S} \|\mathbf{x} - \mathbf{y}\|$ . If we start with a curve  $\vec{\gamma}(0)$  containing all the relevant bone structures, then the stationary point of this curve evolution will be a curve which traces the outer most pixels in  $S$  with a linear interpolation between distinct bone structures (see figure 5). We dilate the contour obtained in this way by 5 pixels for a safety margin and use the resulting contour as the initialization for our algorithm. Using this initialization and the standard level set method which has complexity  $O(n^2)$ , our algorithm (including initialization) takes on average between 10 and 60 seconds to obtain a segmentation on a 2 GHz laptop with 2 GB of RAM. Although this is not prohibitively long, one should note that if real-time processing is desired for a specific application, there are numerous fast level set methods that can be used to improve the performance by an order of magnitude.<sup>29–31</sup>

## 4. EXPERIMENTAL RESULTS

We have conducted an extensive evaluation of our proposed algorithm for automatic segmentation of human bone structures from CT image data. This section describes the test data, the experimental results, and the validation methodology.

### 4.1 Test Data

We have gathered our test data from sets of previously evaluated data as well as publicly available CT datasets. The CT datasets used for experimentation and validation are:

1. Wrist dataset, kindly provided by MedIA laboratory, Queen’s University, Kingston, Canada: 46 slices of  $512 \times 512$  with the spacing of  $0.23 \times 0.23 \times 4 \text{ mm}^3$ ;
2. Foot dataset, publicly available as part of Virtual Animation of the Kinematics of the Human for Industrial, Educational and Research Purposes (VAKHUM), provided by University of Brussels, Belgium: 12 slices of  $512 \times 512$  and spacing of  $0.23 \times 0.23 \times 4 \text{ mm}^3$ ;
3. Ankle dataset, captured from a volunteer patient at Hotel Dieu Hospital, Kingston, Canada: 15 slices of  $512 \times 512$  and spacing of  $0.62 \times 0.62 \times 2 \text{ mm}^3$ ; and finally,
4. Iliac bone dataset also available from VAKHUM project: Iliac is the uppermost and the largest bone of the pelvis. The iliac dataset consists of 20 slices of  $512 \times 512$  and spacing of  $0.49 \times 0.49 \times 1 \text{ mm}^3$ .

The main reason we have selected such diverse datasets for evaluating our algorithm is to test our algorithm precisely for a wide range of bone structures and a correspondingly wide range of imaging conditions. Indeed, each specific bone structure presented here offers its own specific set of challenges: For example, the wrist and the foot bones are closely located to each other and consist of several bone pieces in each CT slice; the bone structures in ankle data show overlap in some horizontal CT slices. Furthermore, the ankle data we show here was captured from a patient who had a small fracture in the Talus bone and accurate segmentation of the tiny fractured pieces can be a very challenging. Finally, a few CT slices from the iliac data have weak edge connectivity, which make them quite challenging to segment even for the human eye.

### 4.2 Results

For the experiments, the values of  $\lambda$  and  $\alpha$  were empirically optimized for each dataset and the same values were used across all slices within the dataset. Experimentally, we have found that  $\lambda$  should vary linearly with  $\alpha$  and in most of the datasets we have chosen  $\lambda = 2\alpha$ . The optimal values for  $(\alpha, \lambda)$  are (40, 80), (2, 6), (30, 60) and (0.5, 1) for the wrist, ankle, foot, and iliac datasets respectively.

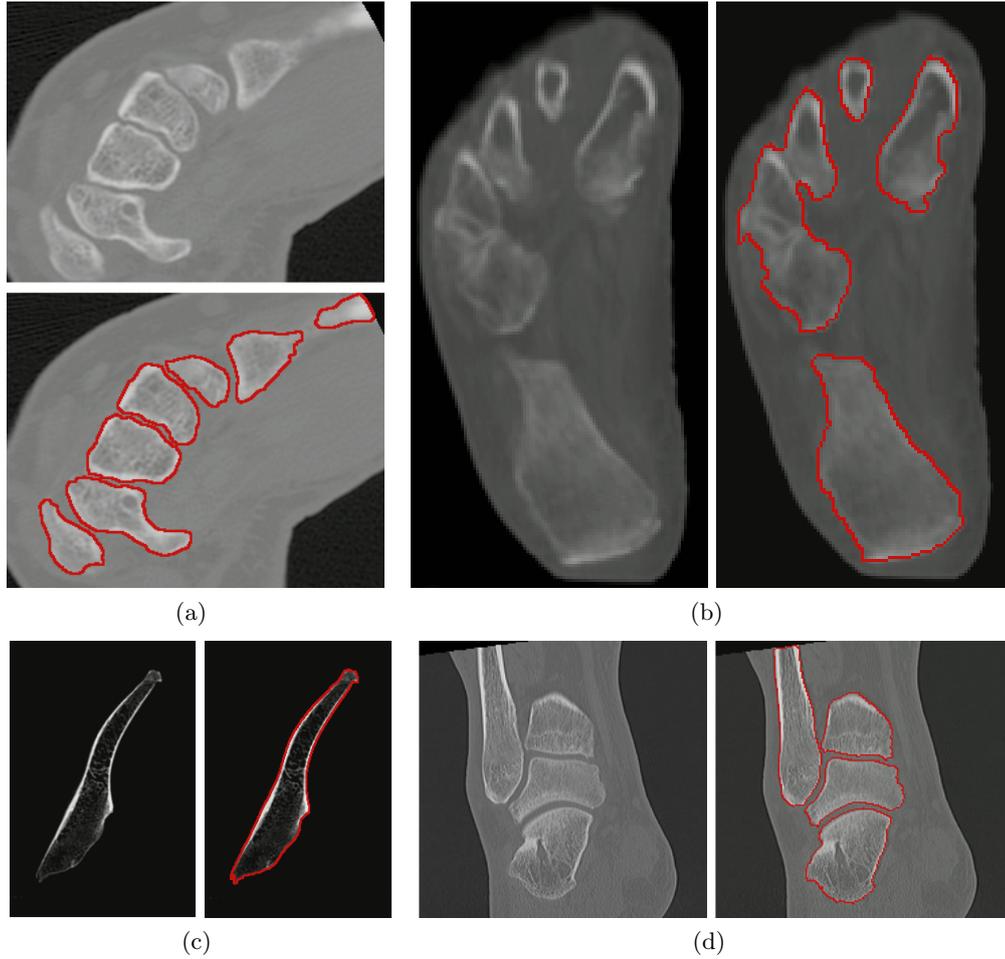


Figure 6. Automated segmentation of slices from (a) wrist dataset, (b) foot dataset, (c) iliac dataset, and (d) ankle dataset. Automatic segmentation is shown in red on the right side.

Figure 6 depicts an example of the proposed segmentation algorithm on each of the aforementioned datasets. We should note that the wrist CT data has the highest signal-to-noise ratio (SNR) of all the datasets. On the other hand, the foot and ankle datasets are of much lower quality than the wrist data, both in terms of resolution and SNR. In the ankle dataset, there is significant background noise in the soft tissue, which makes it hard to define clear bone contour edge points. Furthermore, this specific ankle data was captured from a patient who had a tiny fractured, isolated piece of bone located between the lower edge of Tibia and upper edge of Talus. Because of the size of the fracture (only a few pixels wide) and the proximity to the Tibia and Talus bones, such a fracture can be quite difficult for any segmentation algorithm to capture accurately. Figure 7 shows the corresponding slice from the ankle dataset along with our automatic segmentation. We can see that our algorithm is capable of fully extracting the boundaries of the fractured piece despite the small size and proximity to other bones. The foot dataset has many weak edges around bone contours, which are difficult for many segmentation algorithms to capture.<sup>1,5</sup> We can see these characteristics from the foot dataset in figure 6(b). The bone structures are difficult to quickly identify by eye and the cortical bone is in general very weak; both of which make the foot dataset one of the most challenging of the four datasets. These different examples illustrate the fact that our algorithm can perform well on considerably different types of CT images, with only minor adjustments of the two parameters  $\alpha$  and  $\lambda$ .

Since our algorithm is implemented on two dimensional slices and CT is inherently a three dimensional modality, it is certainly of interest to reconstruct a three dimensional model based on the automatic segmentation results, which can then be compared visually to the manual segmentation. Figure 8 shows the three dimensional

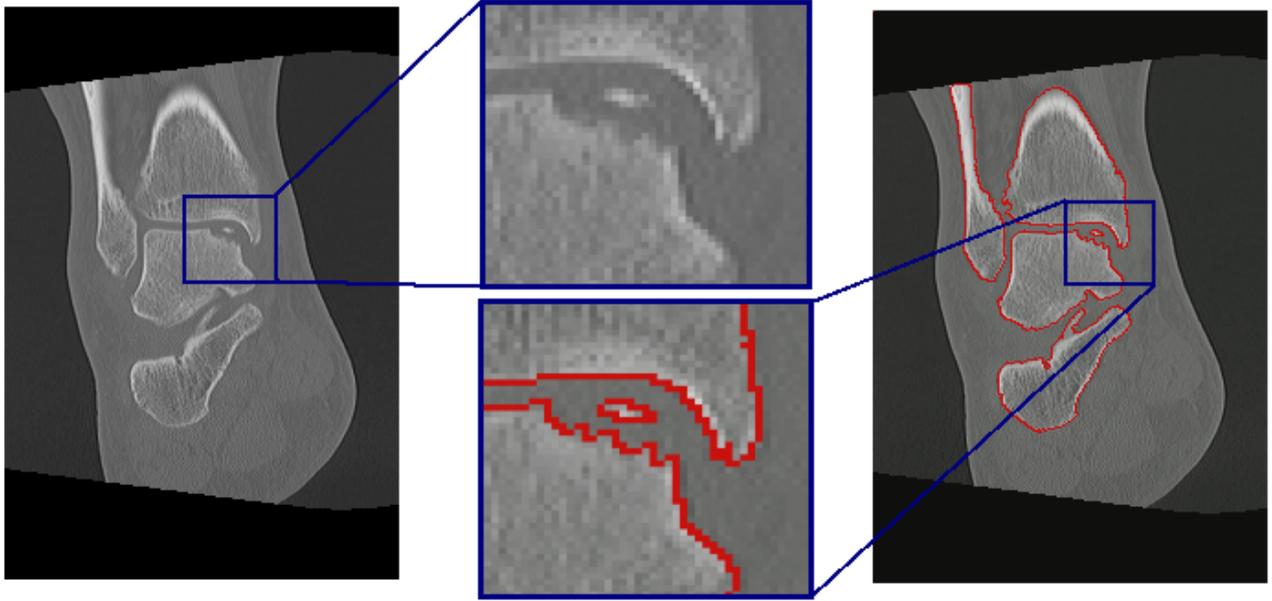


Figure 7. A slice from the ankle dataset with a small fracture in the Talus. Although the fractured piece of bone is only a few pixels wide, our algorithm is able to segment the fractured bone and clearly separate it from the Talus bone.

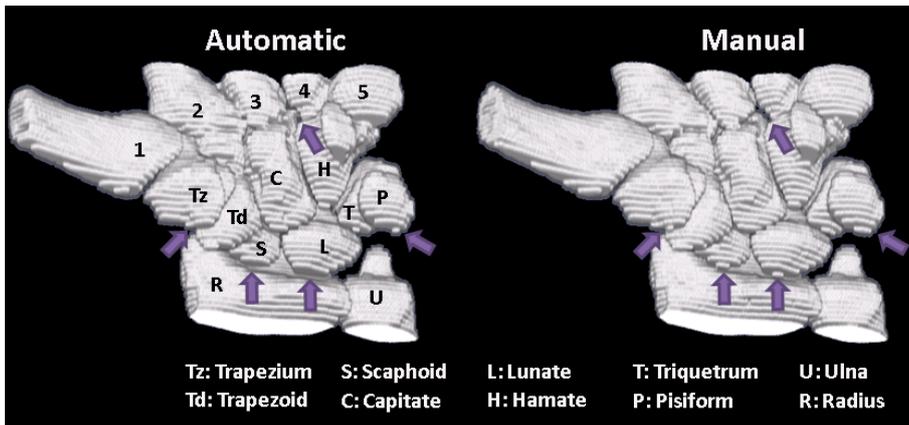


Figure 8. Comparison between 3D reconstructed model of the wrist using both manual and automatic segmented CT slices. All eight bone structures are labeled with abbreviated letters in the figure. The arrows highlight the residual difference between the 3-D reconstructions using our automatic segmentation algorithm and the gold standard manual segmentations.

model obtained by applying our segmentation algorithm to the wrist dataset. As can be seen in that three dimensional model, the eight bone structures that constitute the wrist<sup>32</sup> are closely matched by their counterparts in the three dimensional model obtained from the manually segmented data of the same dataset.

### 4.3 Validation

The four aforementioned CT datasets were manually segmented by an expert and the results were taken as the gold standard. We have carried out an extensive validation of our proposed algorithm against the manually segmented data. We have classified the segmented pixels into the four standard categories. The correctly segmented pixels of the bone structure are true positive (TP) pixels. The bone pixels that were not included in the segmented region are false negatives (FN), and pixels that are incorrectly segmented as bone pixels by the automatic segmentation technique are false positive (FP) pixels. Finally, the soft-tissue (or other non-bone)

pixels that are not included in the automatically segmented region are true negatives (TN). Using these counts of true/false positive/negative pixels, we define the following standard measures for comparing a segmentation to its ground truth: Precision Rate =  $TP/(TP+FP)$ , Sensitivity =  $TP/(TP+FN)$ , and Specificity =  $TN/(TN+FP)$ .

Dataset	Measure (mean±std) %		
	Precision Rate	Sensitivity	Specificity
Wrist	98.96 ± 1.69	95.97 ± 1.91	99.79 ± 0.31
Foot	96.15 ± 0.97	89.19 ± 4.87	98.41 ± 0.61
Ankle	99.74 ± 0.09	92.43 ± 1.31	99.87 ± 0.05
Iliac	96.38 ± 2.19	97.86 ± 0.27	97.87 ± 1.23

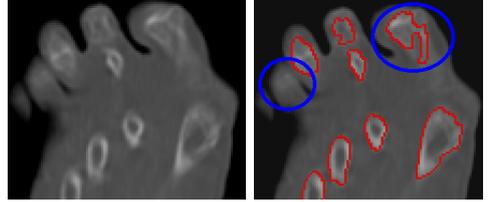


Figure 9. Comparison between automatic and manual segmentation images in terms of precision rate, sensitivity, and specificity in the form of (mean±std) percentage.

Figure 10. Example of segmentation error in a difficult CT image.

Figure 9 shows the precision rate, sensitivity, and specificity scores for each dataset. On the wrist, foot, and iliac datasets, our algorithm performs very well in terms of all three performance measures (precision, sensitivity and specificity). The foot dataset performs well in terms of precision and specificity, but has a lower score on sensitivity indicating that there were a significant number of false negatives on the foot dataset. However, because of the previously mentioned challenges with the foot dataset (*e.g.*, low resolution, faint cortical bone regions), such a score is expected given the quality of the data. Figure 10 illustrates one of the segmentation errors on a slice from the foot dataset. We can see that the errors generally occur when there is a fair amount of ambiguity over where the bone boundaries should be. In particular, in the top right corner of figure 10 it appears as if there are two bones which slightly overlap. Such conditions make it difficult for even an expert to segment such an image, so given the quality of the CT data, the performance of our algorithm is actually quite good on the foot dataset. Overall, these results validate the effectiveness of our algorithm and suggest that our algorithm could be indeed well-suited for use as the core functional in CT bone segmentation algorithms.

#### 4.4 Noise in CT Data

In any segmentation algorithm, it is important to analyze the effect of noise on the segmentation performance. Figure 11(a),(b),(c) shows the segmentation of a CT slice from the wrist dataset with directional noise induced by the imaging conditions under which the scan was taken. As can be seen in the figure, the segmentation still captures the essential structures, but is not able to separate some of the neighboring bones. This can be attributed to the fact that the strong spatial correlation in the noise can bias the  $E_{flux}$  term in our functional towards edge points induced by the noise and away from bone boundaries. This suggests that disabling the  $E_{flux}$  term could alleviate this problem; the result is shown in figure 11(d),(e),(f): with the  $E_{flux}$  term disabled, the segmentation is able to fully separate all the bones; however since the segmentation is not biased anymore by the  $E_{flux}$  term towards bone boundaries, the segmentation misses part of a bone with a very faint cortical bone boundary. As a rule, the  $E_{cdf}$  term, being a function of the global statistics of the image, is more immune to noise whereas the  $E_{flux}$  term, being a function of the local statistics of the image, is less immune to noise.

The previous observations also suggest that the segmentation performance should not decrease considerably under addition of uncorrelated noise. To verify this, we have tested our algorithm on a slice from the wrist dataset with additive stationary zero-mean white Gaussian noise. Figure 11 shows the results for two different noise levels;  $\sigma^2 = 100$  and  $\sigma^2 = 1000$  where  $\sigma^2$  is the variance of the noise and the intensity range is taken to be between 0 and 255. Since the  $E_{cdf}$  term is a global measure, it is largely immune to the effects of noise, so even in the presence of extreme noise (*i.e.*,  $\sigma^2 = 1000$ ) we obtain very similar segmentations.

A common source of noise in CT images is the presence of metal artifacts. Such artifacts cause a massive shift in the intensity distribution of the image making segmentation a challenging task. Figure 12 shows a CT slices with a metal artifact and the result of our segmentation algorithm. Although the segmentation is quite poor in the presence of this distortion, it should be noted that the functional we are proposing for CT image segmentation was not designed with such artifacts in mind, and hence one would not expect a good performance.

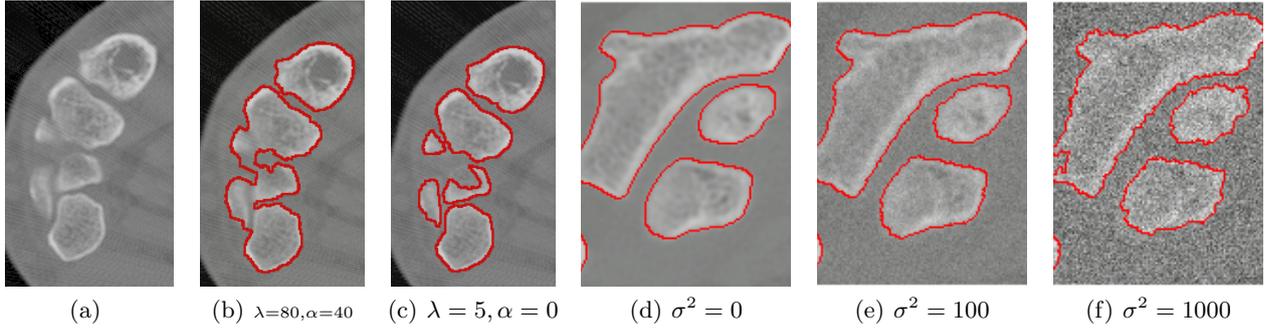


Figure 11. (a),(b),(c) Example segmentation on a noisy slice of the wrist dataset. (a) CT slice (b) segmentation with parameters chosen for the wrist dataset ( $\lambda = 80, \alpha = 40$ ) (c) segmentation without  $E_{flux}$  ( $\lambda = 5, \alpha = 0$ ). In the presence of directional noise, the segmentation is unable to fully separate neighbouring bones (b). This is mainly due to the  $E_{flux}$  term in the functional because when it is removed (c), the segmentation is able to separate all bones. However, because the segmentation is not biased by the  $E_{flux}$  term towards bone boundaries, the segmentation misses part of a bone with a very faint cortical boundary. (d),(e),(f) Example of a slice from the wrist dataset with synthetically added zero-mean white Gaussian noise with variance  $\sigma^2$ . In each segmentation the parameters  $\lambda = 15, \alpha = 0$  were used. Since  $E_{cdf}$  is a global measure, it is quite immune to noise, even with extreme noise (f), the segmentation is still very good.

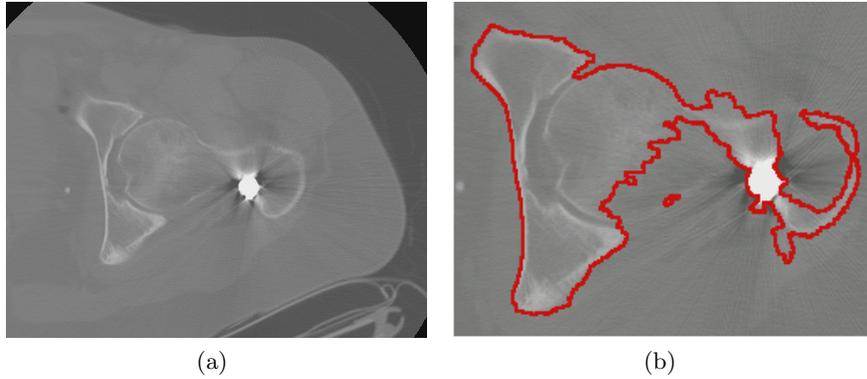


Figure 12. (a) Examples of a CT image with a distorting metal artifact and (b) the segmentation of (a) obtained via our proposed algorithm.

This failure could possibly be corrected by trying to detect a singularity in the intensity distributions (*i.e.*, the metal artifact) and masking it out in the image. However, the distortion caused by the presence of a metal artifact is observed throughout the entire slice and is not entirely contained in a vicinity of the metal artifact itself, so it is not clear whether or not such an approach would be fruitful.

## 5. CONCLUSION AND FUTURE WORK

We have presented a variational approach to segmenting bone and soft tissue regions in computed tomography images. The functional we propose provides a measure of similarity of the intensity characteristics of the bone and tissue regions based on the similarity of their respective cumulative distribution functions. We validated our approach by testing it on several CT databases and comparing the results to ground truth segmentations. The experimental results and validation show that our core functional is well-suited for segmentation of bone structures in CT images and there is much potential for this functional to be the core of future CT segmentation algorithms. We are currently investigating the following extensions to our algorithm:

1. Employing other image features along with intensity (such as image gradient, or higher order derivatives of the image such as the Hessian) to improve the accuracy of segmentation. Since the derivative of the

image at a point is a function of the pixel intensities in a neighbourhood of that point, the use of image derivatives as a feature can capture higher order correlations in the image.

2. Estimating the coefficients  $\alpha$  and  $\lambda$  from the image data. We suspect that there is a relationship between  $\alpha$ ,  $\lambda$  and the image contrast and such a relationship could provide a robust automatic method of determining these coefficients.
3. Incorporating shape priors to guide the segmentation based on the anatomical location of the scan: Such priors could be realized as additive extensions to our functional.
4. Since CT is inherently a 3D modality, we are currently investigating the direct extension of our functional to 3-D bone volume segmentation. Since a 3-D bone volume segmentation functional would make full use of all three spatial dimensions including a spatial regularization on the bone surface area, we expect to obtain better, more spatially coherent, segmentations which will likely alleviate some of the difficulties observed in 2-D segmentation.

### Acknowledgments

The authors would like to thank Serge Van Sint Jan from the University of Brussels for access to the CT images with the VAKHUM dataset and Professors P. Abolmaesumi and R. Ellis of the Queen's University Computer Science department for providing us with CT images for testing. A. Mansouri was partially supported by a research grant from the Natural Sciences and Engineering Research Council of Canada (NSERC).

### REFERENCES

- [1] Wang, L. I., Greenspan, M., and Ellis, R., "Validation of bone segmentation and improved 3-d registration using contour coherency in CT data," *Medical Imaging, IEEE Trans. on* **25**(3), 324–334 (2006).
- [2] Yao, W., Abolmaesumi, P., Greenspan, M., and Ellis, R. E., "An estimation/correction algorithm for detecting bone edges in CT images," *Medical Imaging, IEEE Trans. on* **24**(8), 997–1010 (2005).
- [3] Kang, Y., Engelke, K., and Kalender, W., "A new accurate and precise 3-D segmentation method for skeletal structures in volumetric CT data," *Medical Imaging, IEEE Trans. on* **22**(5), 586–598 (2003).
- [4] Pardo, X., Carreira, M., Mosquera, A., and Cabello, D., "A snake for CT image segmentation integrating region and edge information," *Image and Vision Computing* **19**, 461–475 (2001).
- [5] Sebastian, T. B., Tek, H., Crisco, J. J., and Kimia, B. B., "Segmentation of carpal bones from CT images using skeletally coupled deformable models," *Medical Image Analysis* **7**(1), 21–45 (2003).
- [6] Otsu, N., "A threshold selection method from gray-level histograms," *Systems, Man, and Cybernetics, Trans. on SMC-9*(1), 62–66 (1979).
- [7] Kohler, R., "A segmentation system based on thresholding," *Computer Graphics and Image Processing* **15**, 319–338 (1981).
- [8] Bernsen, J., "Dynamic thresholding of grey-level images," *8th International Conference on Pattern Recognition* (1986).
- [9] Grau, V., Mewes, U., Alcaniz, M., Kikinis, R., and Warfield, S. K., "Improved watershed transform for medical image segmentation using prior information," *Medical Imaging, IEEE Trans. on* **23**(4), 447–458 (2004).
- [10] Beveridge, J., Griffith, J., Kohler, R., Hanson, A., and Riseman, E., "Segmenting images using localized histograms and region merging," *International Journal of Computer Vision* **2**(3), 311–352 (1989).
- [11] Adams, R. and Bischof, L., "Seeded region growing," *Pattern Analysis and Machine Intelligence, IEEE Trans. on* **16**(6), 641–647 (1994).
- [12] Kass, M., Witkin, A., and Terzopoulos, D., "Snakes: Active contour models," *International Journal of Computer Vision* **17**(4) (1988).
- [13] Canny, J., "A computational approach to edge detection," *Pattern Analysis and Machine Intelligence, IEEE Trans. on* **8**(6), 679–714 (1986).

- [14] Tsai, A., Anthony Yezzi, J., Wells, W., Tempany, C., Tucker, D., Fan, A., and Grimson, W. E., “A shape-based approach to the segmentation of medical imagery using level sets,” *Image Processing, IEEE Trans. on* **22**(2), 137–154 (2003).
- [15] Chen, Y., Tagare, H. D., Thiruvenkadam, S., Huang, F., Wilson, D., Gopinath, K. S., Richard, Briggs, W., and Geiser, E. A., “Using prior shapes in geometric active contours in a variational framework,” *International Journal of Computer Vision* **50**, 315–328 (2002).
- [16] Rousson, M. and Paragios, N., “Shape priors for level set representations,” in [*European Conference on Computer Vision*], 78–92, Springer (2002).
- [17] Xie, J., Jiang, Y., and tat Tsui, H., “Segmentation of kidney from ultrasound images based on texture and shape priors,” *Medical Imaging, IEEE Trans. on* **24**(1), 45–57 (2005).
- [18] Chan, T. F. and Vese, L. A., “Active contours without edges,” *Image Processing, IEEE Trans. on* **10**(2), 266–277 (2001).
- [19] Caselles, V., Catte, F., Coll, T., and Dibos, F., “A geometric model for active contours in image processing,” *Numerische Mathematik* **66**(1), 1–31 (1993).
- [20] Malladi, R., Sethian, J. A., and Vemuri, B. C., “Shape modeling with front propagation: A level set approach,” *Pattern Analysis and Machine Intelligence, IEEE Trans. on* **17**, 158–175 (1995).
- [21] Freedman, D. and Zhang, T., “Active contours for tracking distributions,” *Image Processing, IEEE Trans. on* **13**(4), 518–526 (2004).
- [22] Truc, P. T. H., Lee, S., and Kim, T.-S., “A density distance augmented Chan-Vese active contour for CT bone segmentation,” in [*Proc. 30th International Conference IEEE Engineering in Medicine and Biology Society*], 482–485 (2008).
- [23] Yezzi, A. J., Tsai, A., and Willsky, A., “A statistical approach to snakes for bimodal and trimodal imagery,” in [*Proc. 7th IEEE International Convergence on Computer Vision*], **2**, 898–903 (1999).
- [24] Chan, T., Esedoglu, S., and Ni, K., “Histogram based segmentation using Wasserstein distances,” *Springer* **4485** (2008).
- [25] Freedman, D., Radke, R., Zhang, T., Jeong, Y., Lovelock, D., and Chen, G., “Model-based segmentation of medical imagery by matching distributions,” *Medical Imaging, IEEE Trans. on* **24**, 281–292 (march 2005).
- [26] Vasilevskiy, A. and Siddiqi, K., “Flux maximizing geometric flows,” *Pattern Analysis and Machine Intelligence, IEEE Trans. on* **24**(12), 1565–1578 (2002).
- [27] Sethian, J. A., [*Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science*], Cambridge University Press (1999).
- [28] Kimmel, R., [*Numerical Geometry of Images: Theory, Algorithms, and Applications*], SpringerVerlag (2003).
- [29] Whitaker, R. T., “A level-set approach to 3d reconstruction from range data,” *International Journal of Computer Vision* **29**, 203–231 (1998).
- [30] Adalsteinsson, D. and Sethian, J. A., “A fast level set method for propagating interfaces,” *Journal of Comp. Physics* **118**, 269–277 (1994).
- [31] Bernard, O., Friboulet, D., Thevenaz, P., and Unser, M., “Variational b-spline level-set: A linear filtering approach for fast deformable model evolution,” *Image Processing, IEEE Trans. on* **18**(6), 1179–1191 (2009).
- [32] Agur, A., “Grant’s atlas of anatomy,” *Twelfth edition* (2009).