

RESEARCH

Open Access



# Vertebra segmentation based on two-step refinement

Jean-Baptiste Courbot<sup>1\*</sup>, Edmond Rust<sup>1</sup>, Emmanuel Monfrini<sup>2</sup> and Christophe Collet<sup>1</sup>

\*Correspondence:

jb.courbot@unistra.fr

<sup>1</sup>ICube, Université de Strasbourg - CNRS, 67412 Illkirch, France

Full list of author information is available at the end of the article

## Abstract

Knowledge of vertebra location, shape, and orientation is crucial in many medical applications such as orthopedics or interventional procedures. Computed tomography (CT) offers a high contrast between bone and soft tissues, but automatic vertebra segmentation remains difficult. Hence, the wide range of shapes, aging, and degenerative joint disease alterations as well as the variety of pathological cases encountered in an aging population make automatic segmentation sometimes challenging. Besides, daily practice implies a need for affordable computation time. This paper aims to present a new automated vertebra segmentation method (using a first bounding box for initialization) for CT 3D data which tackles these problems. This method is based on two consecutive steps. The first one is a new coarse-to-fine method efficiently reducing the data amount to obtain a coarse shape of the vertebra. The second step consists in a hidden Markov chain (HMC) segmentation using a specific volume transformation within a Bayesian framework. Our method does not introduce any prior on the expected shape of the vertebra within the bounding box and thus deals with the most frequent pathological cases encountered in daily practice. We experiment this method on a set of standard lumbar, thoracic, and cervical vertebrae and on a public dataset, on pathological cases, and in a simple integration example. Quantitative and qualitative results show that our method is robust to changes in shapes and luminance and provides correct segmentation with respect to pathological cases.

**Keywords:** Clinical imagery, Automatic vertebra segmentation, Coarse-to-fine modeling, SLIC clustering, Hidden Markov chain

## Background

Primitive bone tumors such as osteoid osteoma, metastatic lesions, and degenerative disorders such as arthritis or vertebral body collapse and traumatic injuries can affect one or several vertebrae. Diagnosis and characterization of these spine lesions rely on medical imaging. Computed tomography (CT) is yet one of the first-line imaging procedures. This cross-sectional imaging technique discriminates tissues along their densities and allows a good contrast between bones, surrounding organs, and soft tissues. However, identification of vertebrae can be difficult. Even if vertebrae vary in shape and orientation along the spine, these modifications can be slight between two neighbor elements of the backbone, making assessment of the exact level sometimes challenging.

A precise knowledge of vertebrae location, shape, and orientation is however essential. Hence, an imaging follow-up of spine lesions requires a precise identification of

the affected levels and consequent reliable vertebrae identification. The same considerations are relevant in case of multi-modality imaging, that is to say supplementary spinal imaging procedures (e.g., bone scan with SPECT/CT, 18F-fluoride PET/CT) performed so as to allow a better characterization of lesions or tumor burden. This is even more crucial for preoperative planning and for interventional radiology treatments. Vertebra segmentation and identification is therefore a key issue for many medical applications.

Beside their ambiguous shapes and boundaries, one of the major concern in a segmentation perspective is the varying vertebrae neighborhood and shapes in a single patient, which led to the development of region-specific methods. Another important problem for clinical application is the eventuality of pathological cases, which is not always taken into account in previous works. This is challenging because of the wide range of diseases, e.g., on CT scans a spine lesion can affect the vertebra local shape (primitive tumor), global shape (scoliosis, fused vertebrae, degenerative disorders), or the intensity of some regions (hyper- or hypo-dense tumors). On top of that, one has to consider that a reliable vertebra segmentation method is one of the requirements needed to perform further advanced processing such as efficient image registration.

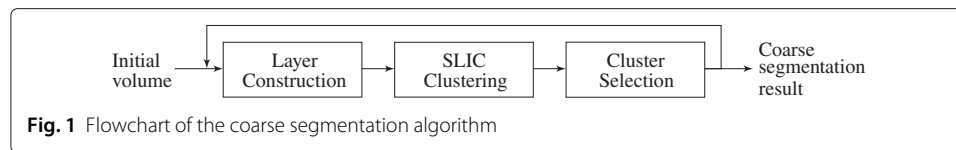
Medical image segmentation methods can be divided in three types: the iconic, the texture-based, and the edge-based methods [1]. Iconic methods rely directly on voxel intensities and include amplitude segmentation (e.g., thresholding) and region-based methods [2]. Texture-based methods rely on local operators [3] to describe and discriminate objects along their apparent texture. Edge-based methods use more abstract descriptors to constrain the shapes and boundaries. As mentioned in [4], vertebra segmentation is a challenging problem since vertebrae are inhomogeneous in intensity and texture and have complex shapes, which make traditional segmentation techniques inefficient to the problem. The vast majority of recent methods dedicated to vertebrae segmentation are edge-based and rely on deformable models performing an adaptation of prior data, such as templates or statistical atlases, to the vertebrae volume. For example, [5, 6] use a prior statistical shape model [7] as an initialization followed by a rigid or non-rigid registration, and in [4, 8, 9], the authors use a shape-constrained deformable model to fit a prior mesh into the data.

Nevertheless, two main key issues limit these works:

- (1) The algorithms use complex shape description and processing, which dramatically increase global processing time.
- (2) Methods are validated on a limited set of vertebrae in terms of scope (lumbar, thoracic, or cervical) and healthiness (middle-aged patient, healthy cases).

We assume in the following that vertebrae are properly isolated in bounding boxes, delimited roughly by their inter-vertebral disk and corresponding mean planes. Several methods of *vertebra localization* can produce such planes, such as the works presented in [4, 8, 10]. Since spine partition is not in the scope of this paper, the volume extractions are made manually. Therefore, this work focuses on the segmentation of vertebral elements contained in the volume, which may include parts of neighbor vertebrae.

Our method overcomes limitations listed above, since it does not rely on prior shape information or on complex shape descriptors. To restrain the computation time, we propose a coarse segmentation algorithm which drops voxel clusters from the data volume.



This first step of the segmentation is built on the basis of a coherent voxel cluster statistical testing and is therefore robust to local and global luminance change. This coarse segmentation step will be referred as “Carving”. The second step aims at discriminating the two classes in the remaining volume within a robust Hidden Markov Chain (HMC) framework and thus performs coherent voxel-level segmentation. No shape priors are introduced in the algorithms; thus, the method can deal with any type of standard vertebrae from lumbar to cervical as well as with most of the non-standard cases one can expect in clinical context. In the Sections “Coarse segmentation” and “Fine segmentation based on HMC modeling” the two-pass segmentation algorithm is described. The Section “Results” explains the experiments and the results obtained with the proposed method. A discussion on the results is given in the Section “Discussion and conclusion” as well as a conclusion on the method.

## Method

### Coarse segmentation

In the context of medical image processing, coarse-to-fine methods are mostly used to perform fast registration (see, e.g., [11] or [9]) as they reduce computation time. On the other hand, image clustering is a well-known tool grouping individual elements (*i.e.*, voxels) following a specific similarity criterion [12] (based on a given distance metric) and thus produces consistent high-level elements. In our case, it is desirable to combine both approaches to rapidly ensure a first accurate and consistent estimation of the anatomical vertebral volume. Therefore, a new algorithm is introduced to perform a coarse segmentation of the data volume within a previously delimited bounding box [4, 8, 10]. It processes the volume layer by layer iteratively, following three steps (see Fig. 1):

- The *layer construction* consists in selecting the external layer to be processed from the volume of interest in the current iteration.
- The *layer clustering* produces clustered voxels with a joint space-luminance criterion.
- The *cluster selection* tests if the clusters should be rejected or included in the final volume.

The three steps are repeated until the volume is completely processed within the initial bounding box.

### Layer construction

This step isolates the external layer of voxels on which the further processing will be applied. The first layer is defined by its depth  $I_1$  from the borders of the volume. Given the boundary from the previous step, the following layers cover both an inner part of depth  $I_j$  and an outer part of height  $O_j$ ,  $j > 1$ . The layer are isolated with mathematical morphology operator.

More precisely, let  $\hat{V}_{j-1}$  be the binary partially segmented volume obtained at the step  $(j - 1)$  and  $R(a)$  be a ball structuring element of radius  $a$ . The layer  $V_j$  at the  $j$ -th iteration is then defined as:

$$\begin{aligned} V_1 &= V_0 - V_0 \ominus R(I_1) \\ V_j &= V_{j-1} \oplus R(O_j) - V_{j-1} \ominus R(I_j) \quad \forall j > 1 \end{aligned} \tag{1}$$

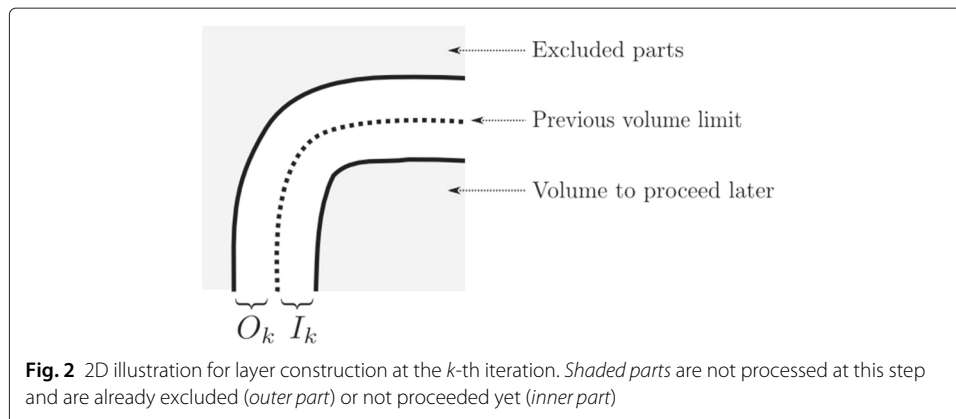
$V_0$  is the initial bounding box, and the operators  $\oplus$  and  $\ominus$  stand for morphological dilatation and erosion, respectively<sup>1</sup>. The external layer is used to avoid artifacts creation by processing the layer boundary regions at least twice. Figure 2 illustrates the layer construction step.

**Layer clustering**

We develop a clustering method based on the simple linear iterative clustering (SLIC) method proposed by Achanta et al. [13]. The authors presented a clustering method for color images we generalize in the 3D gray level case. Thereafter, it will be referred as “SLIC-3D”. In the color image case, a pixel  $i$  can be defined by its Cartesian coordinates  $(x_i, y_i)$  and the  $L * a * b$  intensities  $(l_i, a_i, b_i)$ . In [13], the authors combine the two representation spaces in one distance using two weighting parameters called  $m$  and  $S$ :  $m$  is used to balance the contributions of the color distance with respect to the Euclidean distance and  $S$  stands for the number of pixels a super-pixel is expected to contain. The SLIC algorithm proposed in [13] consists in clustering the pixels in order to approximately minimize for each pixel its combined distance to the cluster centroid.

We propose on a similar principle a clustering algorithm addressing 3D data. The Euclidean distance covers then a 3D space and the luminance of a voxel stands for the color channel. As CT data is processed, its intensity is expressed in Hounsfield Units (HU) corresponding to X-ray absorption ratio of organic tissues with respect to water [14]. A given voxel  $i$  is then represented by its spatial coordinates  $(x_i, y_i, z_i)$  and its luminance  $l_i$ . For a given voxel cluster  $k$  containing  $|k|$  elements, we define its centroid  $C_k$  as its mean value along the four features:

$$C_k = [x_k, y_k, z_k, l_k]^T = \frac{1}{|k|} \sum_{i \in k} [x_i, y_i, z_i, l_i]^T \tag{2}$$



Then, a mixed distance  $D_m$  combining the four features between a cluster centroid  $C_k$  and a voxel  $i$  is given by:

$$D_m(C_k, i) = \sqrt{\left(\frac{d_c(C_k, i)}{m}\right)^2 + \left(\frac{d_s(C_k, i)}{S}\right)^2} \quad (3)$$

where  $d_s$  is the 3D Euclidean distance and  $d_c$  is a luminance distance between  $C_k$  and  $i$ :

$$\begin{aligned} d_s(C_k, i) &= \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2 + (z_k - z_i)^2} \\ d_c(C_k, i) &= |l_k - l_i| \end{aligned} \quad (4)$$

The cluster centroids are initialized on a regular cubic grid of size  $S$ . For each cluster, the algorithm processes a cubic  $2S \times 2S \times 2S$  region centered on the centroid spatial coordinates. Each voxel in the region closer to the cluster centroid than to its current cluster centroid is then re-labeled. Finally, the cluster centroids are updated, and the procedure can be repeated for a few iterations. The SLIC-3D procedure is summarized in Algorithm 1.

---

#### Algorithm 1 SLIC—3D

---

**Require:** Bounding box  $V$  obtained from a previous processing, parameters  $m$  and  $S$  from Eq. (3), iteration number  $nItr$

**Ensure:** Label map  $L$ , pixel-centroid distance map  $D$

---

##### 1. INITIALIZATION

Place the cluster centroids  $C_k$  on a cubic grid of sample size  $S$ .

For each voxel  $i$ :  $L(i) \leftarrow -1$  and  $D(i) \leftarrow \infty$ .

##### 2. ITERATIVE PROCESSING

**for**  $Itr = 1 : nItr$  **do**

**for** each cluster centroid  $C_k$  **do**

**for** each voxel  $i$  in a  $(2S)^3$  cube centered on  $(x_k, y_k, z_k)$  **do**

      Compute  $D_m(C_k, i)$ : Eq. (3).

**if**  $D_m(C_k, i) \leq D(i)$  **then**

$L(i) \leftarrow k$

$D(i) \leftarrow D_m(C_k, i)$

**end if**

**end for**

**end for**

  Update the cluster centroids.

**end for**

---

#### Cluster selection

The clusters from a given layer need to be tested to assess if they belong to the vertebra volume. Since vertebrae are mainly bone tissue and have typical luminance in CT volume, the test is built with the mean luminance  $l_k$  of each cluster  $k$ . Thus, voxels are accepted or rejected by cluster, avoiding to deal with local irrelevant voxel variations. Furthermore, the test must be robust to changes in the vertebrae to proceed. Thus, simple test such as thresholding are not satisfying and a more elaborate procedure is needed. This is why an

adaptive test is developed to ensure robust and consistent clusters acceptance or rejection. The test is based on the *statistical region merging* approach, proposed by Nock and Nielsen [15]. Whereas the authors use the test for a pixel pair set, we propose to test the clusters with respect to a *reference luminance*  $l_0$  corresponding to the typical bone luminance in CT scans. For a given cluster  $k$ , let a first *bone merging* predicate be:

$$\mathcal{P}_0(k) : |l_k - l_0| \leq b(k) \tag{5}$$

where  $b(\cdot)$  is a *merging threshold* [15]:

$$b(k) = g \sqrt{\frac{1}{Q|k|} \ln\left(\frac{1}{\delta}\right)} \tag{6}$$

where  $|k|$  is the number of voxels in the cluster  $k$ ,  $g$  is the grey level range, and  $\delta$  is the acceptable probability of error for the predicate.  $Q$  stands for the expected number of underlying independent random variables (r.v.) for the current region, and according to [15], it allows to quantify its *statistical complexity*. The *bone merging* predicate can then be stated as “*accept the cluster  $k$  if  $|l_k - l_0| < b(k)$ ”.*

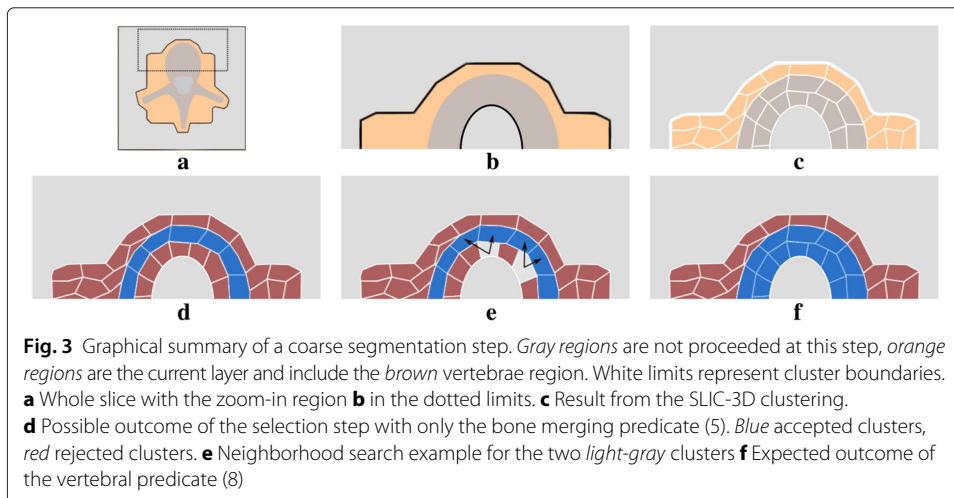
Furthermore, an alternative is added to the test: vertebrae being compact objects, each *interior* cluster is necessarily part of the result. A predicate is needed to assert for a given cluster, the *interiority* of its neighbors. We define as interior to the remaining volume any point closer to the remaining volume center  $v = (x_v, y_v, z_v)$  than the centroid of the super-voxel. Then, an *interiority* predicate is built for a given cluster  $k$  and any of its neighbor  $k'$ :

$$\mathcal{P}_I^k(k') : d_s(C_{k'}, v) \leq d_s(C_k, v) \tag{7}$$

where  $C_k$  and  $C_{k'}$  are the centroids of the clusters  $k$  and  $k'$ , respectively, and  $d_s$  is the Euclidean distance defined in (4). Finally, the two predicates (5) and (7) are combined in the following *vertebral* predicate to test a cluster  $k$ :

$$\mathcal{P}(k) : \mathcal{P}_0(k) \text{ or } \mathcal{P}_I^k(k) \forall k' \text{ such as } \begin{cases} k \text{ and } k' \text{ are neighbor,} \\ \mathcal{P}(k') \text{ is valid.} \end{cases} \tag{8}$$

This predicate allows an efficient selection of voxel cluster. Figure 3 illustrates the whole coarse segmentation step.



**Model and parameters**

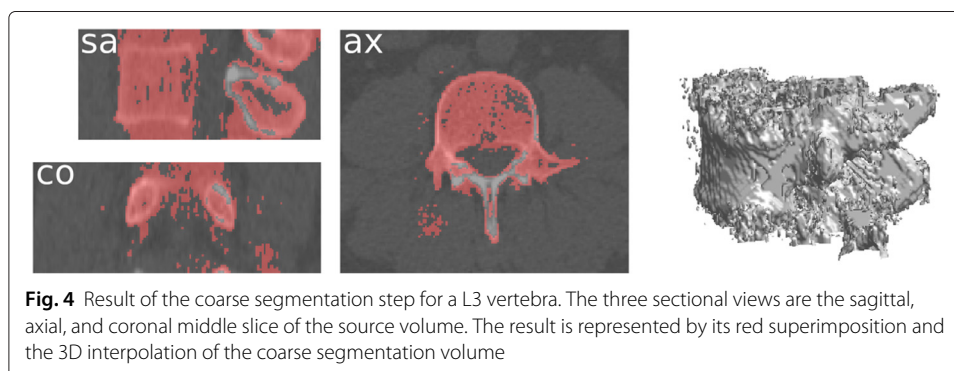
The model requires calibrating several parameters. They were evaluated on a first set of 12 lumbar thoracic and cervical vertebrae. Motivations are given below:

- Equation (1) defining the layer construction and the SLIC-3D method (Algorithm 1) both depend on a size parameter. We choose to define them only with the  $S$  parameter from SLIC-3D, meaning that we link the layer depth to the expected size of supervoxels. Thus, this parameter quantifies both the depth of the current layer and the scale of the supervoxel to exclude. We choose the two first values of  $S$  to be higher than the latter ones, as we process in a coarse-to-fine fashion.  $S$  is given in millimeters to ensure isotropy between axes and between scans.
- The  $m$  parameter used in clustering (Algorithm 1) mostly defines the clusters shape between a spatial regularity and an intensity regularity. We use decreasing values of  $m$  along iterations to exclude first spatially coherent and then intensity-coherent clusters.
- The statistical parameters  $g$ ,  $Q$ , and  $\delta$  used for cluster selection (Predicate 5) can be computed automatically on the basis of the cluster to proceed:  $g$  is the range of the current layer intensity and  $Q$  must be set lower than  $g$  to reduce the expected complexity of the bone merging predicate. The error probability  $\delta$  can be fixed arbitrarily, e.g., as the inverse of the cardinal of the cluster.
- The reference intensity  $I_0$  (Predicate 5) is the intensity of typical bone in CT scans and is provided by an expert.

The algorithm is built on the three steps previously detailed and processes the volume iteratively. For a volume of height  $h$ , the number of iterations  $J$  is given by<sup>2</sup>:

$$J = 2 + \left\lceil \frac{h - S_1 - S_2}{S_j} \right\rceil \tag{9}$$

where  $\lceil \cdot \rceil$  is the ceiling operator,  $S_1$ ,  $S_2$ , and  $S_j$  stands respectively for the two first values of  $S$  and its value for any iteration  $j > 2$ . The model is tolerant to parameter variations, as long as their order from coarse to fine is preserved. Note that variations of the  $S$  parameters may produce changes in terms of computation time, since it influences the number of iteration to proceed in (9). In all cases, we observed convergence towards similar results. The entire coarse segmentation method that we called Carving, as well as the parameters value, are summarized in Algorithm 2. Figure 4 illustrates the results obtained at this step.



---

**Algorithm 2** Carving

---

**Require:** Initial volume  $V_0$ **Ensure:** Binary coarse volume segmentation of  $V_0$ 

---

Compute the number of iterations  $J$ .**for** each iteration  $j < J$  **do**Build  $V_j$ : Eq. (1) with  $O_j$  and  $I_j$ .Update  $g$ ,  $Q$  and  $\delta$  with the current values.Cluster  $V_j$ : Alg. 1 with  $S_j$  and  $m_j$ .**for** each cluster  $k$  **do**Test  $\mathcal{P}(k)$ : Eq. (8) with  $g$ ,  $Q$ ,  $\delta$ ,  $l_0$ .**if**  $\mathcal{P}(k)$  is false **then**Exclude the voxels corresponding to  $k$  from the volume.**end if****end for****end for**

---

**Parameters values:** $O_j, I_j$   $I_j = S_j \forall j$ ,  $O_1 = 0$  and  $O_j = (S_{j-1} + S_j)/2 \forall j > 1$  $m_j$   $m_1 = 100$ , linear decrease until  $m_J = 70$  $S_j$   $S_1 = 23$  mm,  $S_2 = 17$  mm,  $S_j = 11$  mm  $\forall j \geq 2$  $g$  Intensity range in the current  $V_j$  $Q$   $0.4 \times g$  $l_0$  1300 HU $\delta$   $1/|V_j|$ 

---

The result of the coarse segmentation is very nice given the expectation: the first need is to reduce the data amount to proceed, which is efficiently done. The algorithm actually does more than data volume reduction since the results already have the shape of the underlying vertebrae. However, this step alone remains too coarse, and we have now to use a finer segmentation to perform a voxel-level classification of the remaining volume.

**Fine segmentation based on HMC modeling**

The coarse segmentation results obtained at the previous section are smaller than the initial volume and include most of the anatomical vertebrae volume. However, to allow an efficient final segmentation, we need to have in the volume enough voxels of the two classes to separate. Thus, a region of interest (ROI) is built based on the previous coarse result. It is a morphological dilation with a ball structuring element of radius 10 mm. In this section, this ROI will be processed, as it preserves the expected shape of the vertebra and includes enough non-vertebral voxels to allow automatic separation.

We are interested in a robust, voxel-wise segmentation method. The Bayesian framework meets these requirements and offers a consistent statistical modeling for the segmentation of an image into classes. When processing images or volumes, Hidden Markov Random Field (HMRF) [16] modeling often provides good results because the

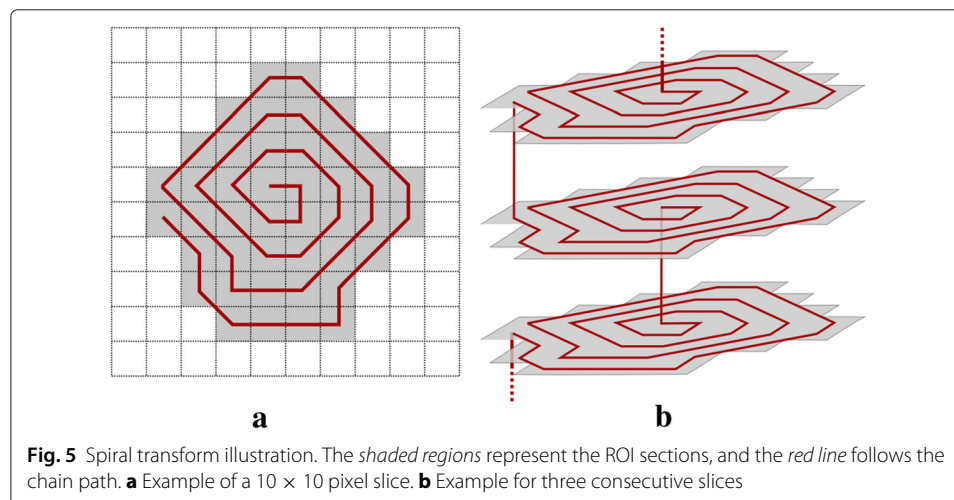


model do consider spatial relationships. However, using HMRF can be computationally time-consuming. This is mainly due to the sampling needed to perform estimations from an analytically unknown distribution. On the other hand, the classical HMC framework, while having the advantages of Bayesian segmentation, does not have the drawbacks of HMRF. It provides faster computations, and we use it with a specific volume transformation to preserve the most important spatial features. The Baum-Welch algorithm [17] is used for segmentation, based on parameters estimated with the Stochastic Expectation-Maximization (SEM) [18] method.

### **Volume transformation**

First of all, the 3D data needs to be transformed to obtain a one-dimensional chain. This point must be carefully considered, since while it permits fast computation, it introduces an artificial 1D order in 3D data and thus uses only 2 out of 26 neighbors for each voxel. A 3D volume can be transformed by sweeping each line, column, and row from first to last but this transformation induces too much distortion to the original data structure. Another alternative is the Hilbert curve [19], which is known to be successful for transforming 2D or 3D images into chains (see, e.g., [20–22]). The resulting chain is more spatially regular; however, it creates artifacts in the HMC segmentation because the chain requires having relatively few state transitions to produce a smooth segmentation estimate. The Hilbert curve path does not ensure this; therefore, in this section, a new volume-to-chain transformation is introduced, relying on the shape information obtained at the coarse segmentation step.

The volume is processed by slices. Given the symmetries of a vertebra, horizontal slices are retained: the axis of our spirals is axial<sup>3</sup>. For each slice, a spiral along concentric perimeters of the ROI section is built. The spiral path goes alternatively inward and outward, so that consecutive spiral extremities are spatially close from one slice to another. Algorithm 3 summarizes the process, which is illustrated in Fig. 5.



---

**Algorithm 3** Spiral transform of a ROI

---

**Require:** Binary ROI, initial volume

**Ensure:** Coordinates/values set of the spiral chain transform

---

Set the first spiral direction to "inward".

**for** Each horizontal slice of the ROI **do**

    Built concentric perimeters of the ROI in the slice.

    Sort the perimeters according to the current spiral direction.

    Add the current voxel coordinates/values set to the previously obtained data.

    Invert the current spiral direction.

**end for**

---

**Forward-backward algorithm**

This algorithm allows the computation of posterior densities required for segmentation. Let  $N$  be the length of the chain obtained from the spiral transform.  $\mathbf{X} = (X_1, \dots, X_N)$  and  $\mathbf{Y} = (Y_1, \dots, Y_N)$  are respectively the random variables sequences representing the spiral transformation of the class volume and the observed volume. We will note  $\mathbf{x} = (x_1, \dots, x_N)$  and  $\mathbf{y} = (y_1, \dots, y_N)$  their respective realizations. The class volume elements take their values in  $\Omega = \{\omega_0, \omega_1\}$  since we want to discriminate vertebral ( $\omega_1$ ) from non-vertebral ( $\omega_0$ ) elements. The observed volume voxels remain in Hounsfield Units, typically in a range of  $[-2000, 2000]$  HU. For clarity, we note  $p(x_n)$  and  $p(\mathbf{x})$  instead of  $p(X_n = x_n)$  and  $p(\mathbf{X} = \mathbf{x})$ , respectively, and likewise for the  $\mathbf{Y}$  process.

We assume that  $(\mathbf{X}, \mathbf{Y})$  is a HMC with independent noise (HMC-IN). The following properties are verified:

- $\mathbf{X}$  is a Markov chain:

$$p(\mathbf{x}) = p(x_1)p(x_2 | x_1) \dots p(x_N | x_{N-1}) \quad (10)$$

- The  $(Y_n)_{1 \leq n \leq N}$  are conditionally independent with respect to  $\mathbf{X}$ :

$$p(\mathbf{y} | \mathbf{x}) = \prod_{n=1}^N p(y_n | \mathbf{x}) \quad (11)$$

- The noise independence is verified :

$$p(y_n | \mathbf{x}) = p(y_n | x_n) \forall n \in \{1, \dots, N\} \quad (12)$$

- The previous points lead to the following expression for the joint  $(\mathbf{X}, \mathbf{Y})$  probability distribution:

$$p(\mathbf{x}, \mathbf{y}) = p(x_1)p(y_1 | x_1) \prod_{n=2}^N p(x_n | x_{n-1})p(y_n | x_n) \quad (13)$$

Then the classical *forward-backward* decomposition [17] of the posterior marginal probability yields:

$$\xi(x_n) = p(x_n | \mathbf{y}) = \frac{\alpha(x_n)\beta(x_n)}{\sum_{\omega \in \Omega} \alpha(\omega)\beta(\omega)} \forall n \in \{1, \dots, N\} \quad (14)$$

Where  $\alpha$  and  $\beta$  are respectively the *forward* and *backward* probabilities for  $n \in \{1, \dots, N\}$ :

$$\begin{aligned} \alpha(x_n) &= p(x_n, y_1, \dots, y_n) \\ \beta(x_n) &= p(y_{n+1}, \dots, y_N | x_n) \end{aligned} \tag{15}$$

Both  $\alpha$  and  $\beta$  can be computed using the following recursions:

$$\begin{aligned} \text{Initializations: } & \begin{cases} \alpha(x_1) = p(x_1, y_1) \\ \beta(x_N) = 1 \end{cases} \\ \text{Inductions:} & \\ \forall 1 \leq n \leq N - 1 : & \begin{cases} \alpha(x_{n+1}) = \left( \sum_{\omega \in \Omega} \alpha(\omega) p(x_{n+1} | \omega) \right) p(y_{n+1} | x_{n+1}) \\ \beta(x_n) = \sum_{\omega \in \Omega} \beta(\omega) p(\omega | x_n) p(y_{n+1} | \omega) \end{cases} \end{aligned} \tag{16}$$

The posterior marginal probabilities (14) can then be computed, and the maximum posterior mode (MPM) class estimation [23] yields:

$$\forall 1 \leq n \leq N \quad \hat{x}_n = \underset{\omega \in \Omega}{\arg \max} p(X_n = \omega | \mathbf{y}) \tag{17}$$

The distribution in (16) requires the knowledge of noise and model parameters. In an unsupervised segmentation framework, they must be estimated. An estimation method is reported in the next section.

### Parameter estimation

The parameters from Eq. (13) need to be estimated to perform the class estimation. We note:

$$\begin{aligned} p(X_1 = \omega_i) &= \pi_i \\ p(X_n = \omega_j | X_{n-1} = \omega_i) &= \pi_{ij} \quad \forall n \in \{1, \dots, N\} \end{aligned} \tag{18}$$

Note that the  $\pi$  parameters do not depend of  $n$  since we supposed the HMC to be homogeneous. We use the SEM algorithm [18] to perform the parameter estimation. SEM is chosen over its determinist counterpart EM [24] for robustness reasons, since the algorithm needs to efficiently deal with non-standard cases (e.g., noise, pathologies, artifacts) and to avoid local extrema convergence. We assume that we are in the case of a Gaussian mixture:

$$p(y_n | X_n = \omega_i) \sim \mathcal{N}(\mu_i, \sigma_i) \text{ with } i \in \{0, 1\} \text{ for } n \in \{1, \dots, N\} \tag{19}$$

The set of parameter to estimate is  $\Theta = \{\pi_{ij}, \pi_i, \mu_i, \sigma_i\}$  for  $i, j \in \{0, 1\}$ . With complete data  $(\mathbf{x}, \mathbf{y})$ , one can perform the estimation of  $\Theta$  with the maximum likelihood (ML) estimators. Complete data are however unavailable. This is why SEM iteratively provides simulations of  $\mathbf{x}$  along posterior distributions.

The SEM algorithm requires an accurate initialization to ensure fast parameter estimation convergence. At this step, we split the process: in some known cases, the volume can include air elements, which are clearly distinct from both soft tissues and bones. To avoid wrong class clustering, we use for the initialization a set of reference parameters obtained from other vertebra of the same patient without air in the neighborhood. This allows correct calibration with respect to the patient and the scanner, while avoiding wrong class clustering. In any other cases, we use a simple initialization where  $\mu_0 = 0.25$ ,  $\mu_1 = 0.75$ ,

and  $\sigma_0 = \sigma_1$  are estimated through the ML estimator from the whole sequence  $\mathbf{y}$  and  $\pi_{ij} = \pi_i = 0.5 \forall i, j \in \{0, 1\}$ . These initializations are chosen to ensure class separation and avoid reliance on other algorithm(s) convergence (e.g., the K-means algorithm [25]).

For the simulation (S) step, one needs to compute the posterior transition probabilities, given by:

$$\begin{aligned}
 p_{\Theta^k}(\mathbf{x} | \mathbf{y}) &= p(x_1 | \mathbf{y}) \prod_{n=2}^N p(x_n | x_{n-1}, \mathbf{y}) \\
 p(x_n | x_{n-1}, \mathbf{y}) &= \frac{p(x_n | x_{n-1})p(y_n | x_n)\beta(x_n)}{\sum_{\omega \in \Omega} p(\omega | x_{n-1})p(y_n | \omega)\beta(\omega)} \quad \forall 2 \leq n \leq N
 \end{aligned}
 \tag{20}$$

An additional step in the original SEM procedure is introduced to produce a convergence measure for the parameters estimate. Since the individual  $\pi$ ,  $\mu$ , and  $\sigma$  parameters differ in nature, we cannot use a direct Euclidean distance comparison between two consecutive estimations of  $\Theta$ . However, the forward-backward class estimation provides for a given  $\Theta$  and fixed observations  $\mathbf{y}$  a determinist result. The convergence between two SEM steps is then estimated through the variations between consecutive on-the-fly forward-backward estimates based on the consecutive parameter estimation. The convergence rate between two consecutive parameter estimations  $\Theta^k$  and  $\Theta^{k-1}$  is computed together with their corresponding forward-backward estimates  $\hat{\mathbf{x}}^k$  and  $\hat{\mathbf{x}}^{k-1}$ , respectively, as:

$$\epsilon = \frac{1}{N} \sum_{n=1}^N (\hat{x}_n^k - \hat{x}_n^{k-1}).
 \tag{21}$$

We assume that the SEM algorithm has converged when  $\epsilon < 1\%$ . This choice allows performing the algorithm in a small number of iterations: typically less than 15 iterations are needed. Setting a smaller  $\epsilon$  increases the global processing time and does not provide noticeable improvement to the result. The adapted SEM algorithm is summarized in Algorithm 4.

---

**Algorithm 4** Adapted SEM algorithm

---

**Require:** Observed data  $\mathbf{y}$ , initial parameters  $\Theta_{air}^0$  and  $\Theta^0$ , convergence threshold  $\epsilon$

**Ensure:** Parameters estimation  $\hat{\Theta}$

---

**if** Air is present in the volume **then**

Use  $\Theta_{air}^0$  as the initial parameters.

**else**

Use the default parameters  $\Theta^0$ .

**end if**

Initial estimation  $\mathbf{x}^0 = \hat{\mathbf{x}}^0$  with (17) and initial parameters.

**while**  $\epsilon < 1\%$  **do**

(M) Estimate  $\Theta^k$  with the ML estimator based on  $(\mathbf{x}^{k-1}, \mathbf{y})$ .

(E) Compute the posterior transition probabilities  $p_{\Theta^k}(\mathbf{x} | \mathbf{y})$  (20).

(S) Simulate  $\mathbf{x}^k$  according to  $p_{\Theta^k}(\mathbf{x} | \mathbf{y})$ .

Compute the current forward-backward estimation  $\hat{\mathbf{x}}^k$  (16) and the convergence ratio  $\epsilon$  (21).

**end while**

---

**HMC segmentation algorithm**

First, the segmentation algorithm transforms the volume along a spiral path adapted to the coarse shape. The parameters of the mixture are then estimated with the SEM algorithm. Once the mixture estimation is done, posterior marginal probabilities (14) can be computed to find the MPM estimation (17) of the observed vertebra. Finally, the segmented volume is re-built along the initial chain path. Algorithm 5 summarizes the segmentation procedure, and Fig. 6 illustrates the result for the segmentation following the coarse result from Fig. 4.

**Algorithm 5** Spiral HMC segmentation

---

**Require:** Coarse volume

---

**Ensure:** Complete segmentation  $\hat{V}$  of the vertebra

---

Produce the ROI with a 10mm disk dilatation of the coarse volume.

Transform the ROI into a chain: Algorithm 3.

Estimate the mixture parameter with SEM: Algorithm 4.

Compute the MPM estimation (17).

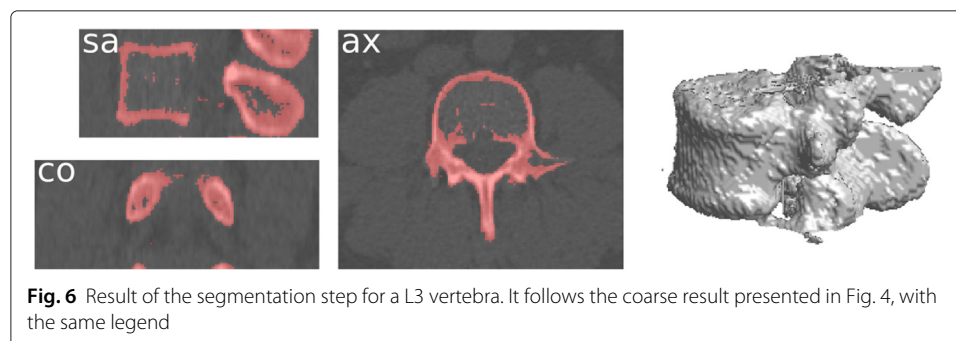
Transform of the chain into a volume along the inverse path.

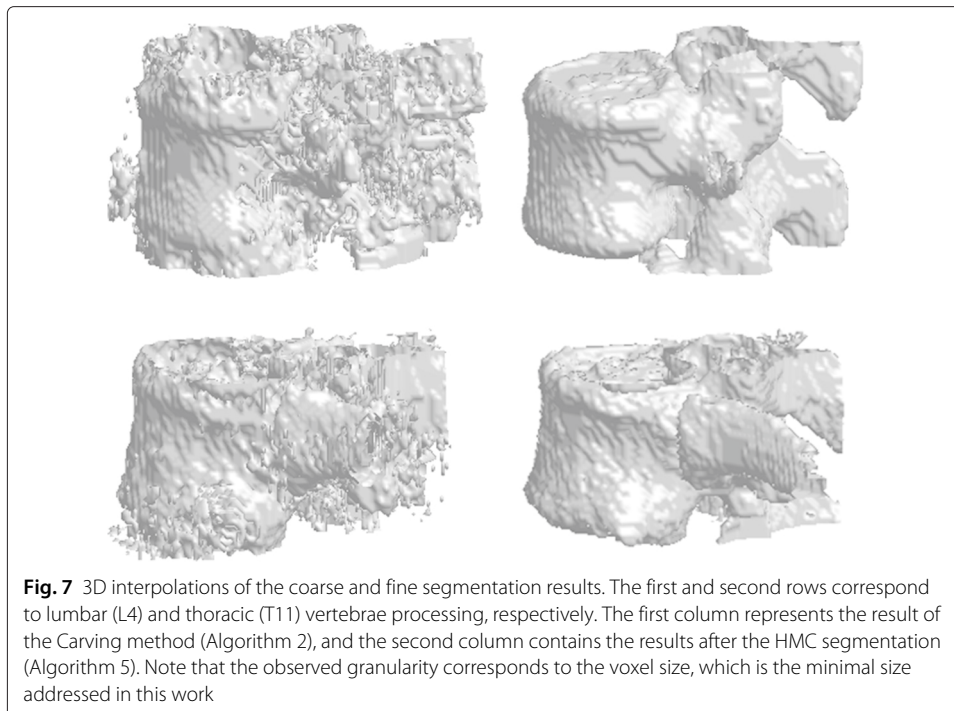
---

Complementary results are presented in Fig. 7. The gain of the HMC segmentation step is clear from the 3D interpolations: the results match our expectations of the vertebral volumes and does not include processing artifacts. Note that while the HMC excludes the inner part of the vertebral body for some lumbar vertebrae (e.g., the vertebra from Fig. 6), the border boundaries are well separated. This is not a key topic for a localization purpose, but other goals (e.g., biomechanical modelling) may require some post-processing in these cases. Further extensive and comparative results are presented in the next section, as well as robustness examples and whole-spine segmentations.

**Results**

In this section, the method performance is evaluated. First, the method is qualitatively evaluated on a set of 339 standard vertebrae acquired in daily practice. Then, quantitative results on manually segmented data are reported. Pathological cases make then the robustness evaluation possible. Finally, we provide simple integration examples with a segmentation of the full spine.





#### Standard cases: qualitative results

The method is evaluated on a set of vertebral volumes from the whole spine of 15 consecutive patients in an oncologic tertiary center, with exclusion of patients with bone tumors or metastatic spine involvement. Patients had a mean age of 63 and presented degenerative joint alterations and some osteoporotic changes, reflecting most of the situations encountered in daily practice; 339 vertebral volumes were extracted and evaluated, meaning that almost all patients' vertebrae were tested.

Each volume is applied successively Algorithm 2 for coarse segmentation and Algorithm 5 for fine segmentation. For the sake of comparison, the K-means algorithm is used as a benchmark since it has common ground with the proposed method: it processes the voxel intensities and has no prior on the volume shape. Since the two-class K-means classification fails when air is present in the volume, we use a sub-sample in which no air is present to provide a more accurate comparison.

For this data, there is no available ground truth (e.g., manual segmentation). Therefore, one must resort to qualitative evaluation. We define, in a similar fashion than in [4], the following ranking:

- *Excellent* (100): the vertebra is exactly delimited inside its bounding box.
- *Good* (75): most of the anatomical structure is covered, but some voxels are segmented out.
- *Bad* (50): the vertebra is recognizable but noticeable part are missing from the result.
- *Poor* (25): the vertebra is not recognizable enough.
- *Fail* (0): the segmentation fails to proceed.

The results were visually inspected by an expert with respect to these criteria. Table 1 summarizes the results obtained for the proposed method and the K-means method.

**Table 1** Results for the subsample without air and for the full sample. This parting provides accurate comparative results on the sub-sample

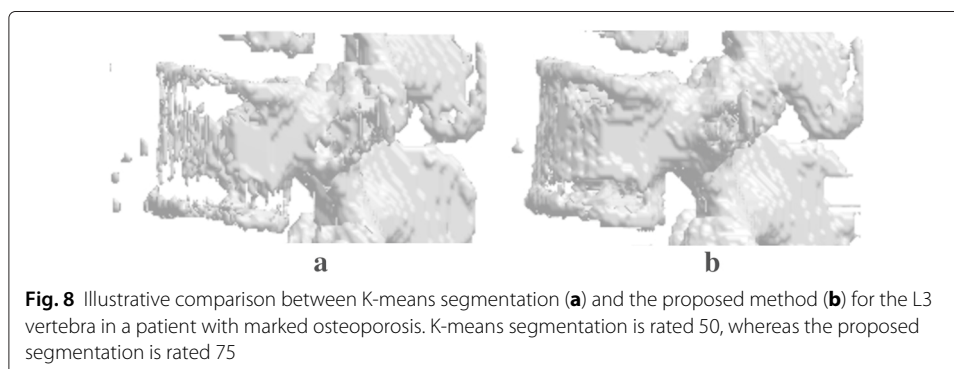
Grade (score)	Partial set: 178 volumes		Full set: 339 volumes
	Proposed method	K-means	Proposed method
Excellent (100)	75 (42.13 %)	46 (25.84 %)	98 (28.91 %)
Good (75)	64 (35.96 %)	83 (46.63 %)	129 (38.05 %)
Bad (50)	31 (17.42 %)	35 (19.66 %)	80 (23.60 %)
Poor (25)	6 (3.37 %)	9 (5.06 %)	29 (8.55 %)
Fail (0)	2 (1.12 %)	5 (2.81 %)	3 (0.88 %)
<b>Average score</b>	<b>78.65</b>	<b>71.91</b>	<b>71.39</b>

Considering that both good and excellent results provide sufficient data for vertebrae segmentation and further advanced processing, our method provides about 78 % of successful results on the subsample where K-means gives 72 % of successful results, whereas on the full sample, the method yields 67 % of successful result and the 2-class K-means fails on the remaining 161 volumes including air (yielding thus an average score of 37.76 % on the full sample). Keeping in mind that the sample originates from daily routines and includes a significant proportion of minor pathologies, these results are of significant interest for clinical use.

The algorithms were developed and tested using Matlab on an Intel i5 (2.6 GHz) on one core, without specific optimization of the code. The processing time is of 36 s by vertebra on average, with 10 s for the Carving step and 26 s for the HMC step. This processing time depends on the size of the vertebra to segment, the average total time being 71.4 s for lumbar vertebrae and 19.8 s for cervical vertebrae. For daily practice implementation, the use of C/C++ language is expected to provide a gain of a factor at least 10 to the processing time.

Our method yields satisfying results. Noteworthy, bad and poor results provided by the method are mainly encountered in levels with either pronounced degenerative joint diseases alterations or marked osteoporosis with consequent low contrast (see Fig. 8). In such cases, our method provides better results than the K-means segmentation, supporting the evidence that our proposal is more reliable in data reflecting daily practice. In only 6.5 % of the cases, the K-means segmentation provided better results than our method.

The results presented here cover standard volumes mostly encountered in practice, but did not yield voxel-wise error rate. The next section presents a quantitative evaluation of the method, with respect to manual segmentations.



### Quantitative results

When considering vertebral segmentation in CT images, there are few available complete dataset allowing a quantitative evaluation of a method. We use the dataset presented at the CSI 2014 challenge [26], available on the public SpineWeb platform [27]. This dataset contains 10 spine scans from a trauma center, acquired during daily clinical routine work. Patients were aged 16 to 35, and the scans covered thoracic and lumbar vertebrae in most cases. A total of 175 segmented volumes were extracted from the dataset. The performances are measured as the rate of correctly classified voxels (true positives and true negatives).

Figure 9 summarizes the results with respect to the provided vertebra segmentation. On average, the proposed method yields  $89.39 \pm 5.54$  % of correct segmentation. On the same dataset, the K-means segmentation provides  $81.65 \pm 12.23$  % correct segmentation. These results imply that our method provide both good results and a small variability on the output. Note that these results concern only the vertebra of interest: a correctly labeled neighboring vertebra is ignored in the measurements. An illustration of the ground truth compared to a segmentation result is provided in Fig. 10.

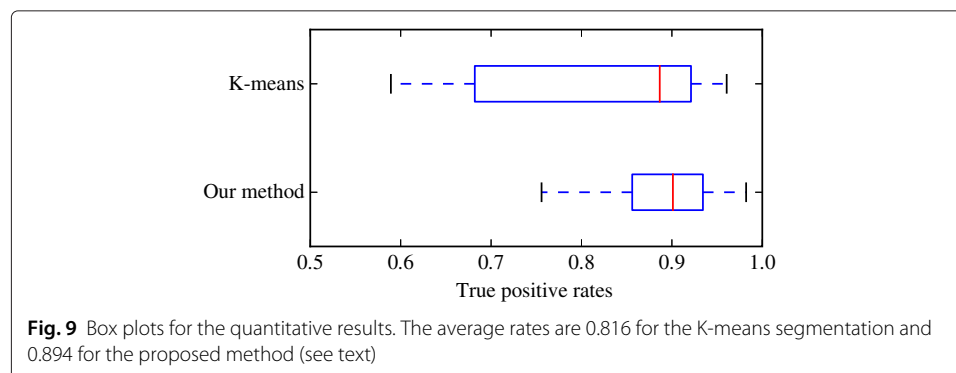
Errors are either false positive (type I error) or false negative (type II). False positive are most encountered in the presence of high-intensity elements, such as calcifications or ribs near thoracic vertebrae. On the other hand, false negatives are either missing voxels at the vertebrae boundary or missing voxels within the vertebral body.

These results are satisfying, given that most error sources are known and that specific post-processing could easily remove them. So far, the performances were evaluated on standard cases: the next section presents pathological cases that may be encountered in practice.

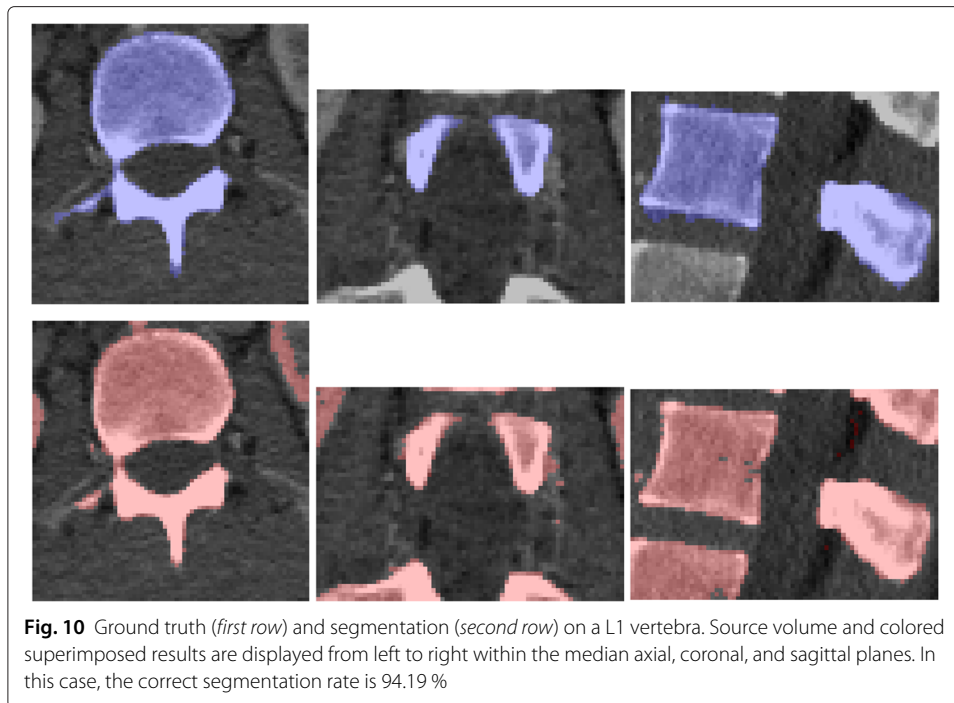
### Pathological cases: robustness evaluation

As our method performs correctly on standard case, it has to be evaluated in more difficult situations, namely pathological cases. Since we aim at a clinical implementation, the method robustness to the most frequent non-standard cases is indeed mandatory. We first briefly describe the selected cases and the corresponding challenges, then we provide their corresponding segmentation results and discussion.

The two main key points are changes in shape and in intensity of the object to segment. They correspond to anatomical and structural deformations, respectively. Structural

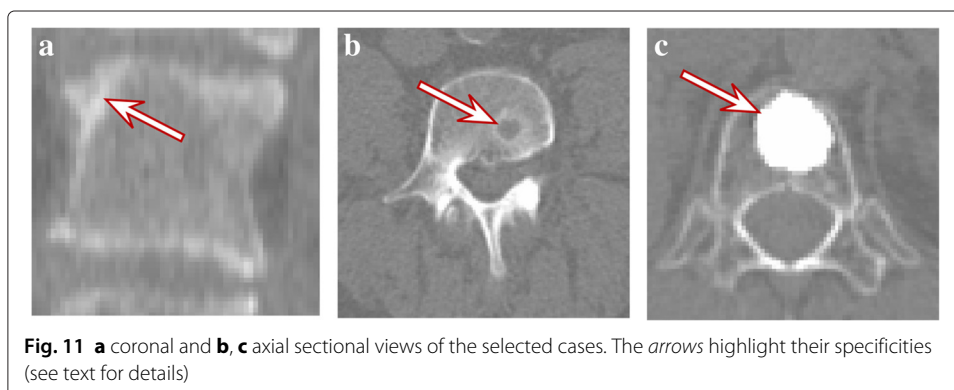






changes are related to alterations of bone and medullar matrix, with consequent modification of density and signal intensity in the CT volume. Changes of shape and density can be related to aging alterations. In particular, arthrosis is responsible of spine alterations in a general population, and we selected it as the first specific case (see Fig. 11a). The frequency and intensity of these modifications is in close relationship with age. After 40 years, hernia, osteophytes, and degenerative joint diseases are commonly encountered. We also selected a hernia case as an instance of common low-intensity structural alteration (see Fig. 11b).

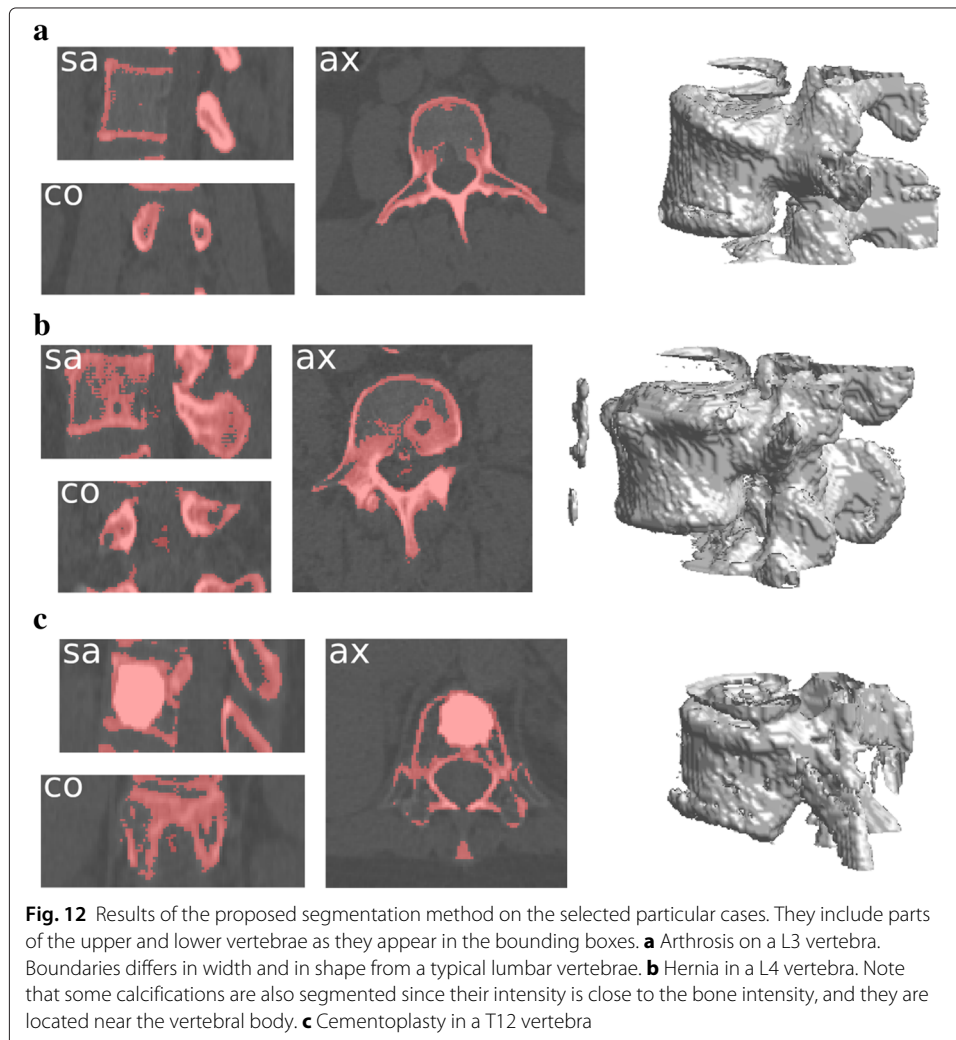
On the other hand, many pathologic conditions can lead to bone density variations. For instance, osteoblastic cancerous tumors will increase bone density. On the other side, osteolytic tumoral involvement is associated with bone destruction and is therefore seen as areas of decreased bone density. Finally, treatments—general treatments as chemotherapy or interventional treatments as cementoplasty—can induce bone density alterations. In particular, cementoplasty, which can be described as the interventional introduction



of artificial high-density material inside the vertebral body, represents an extreme case of overdensity. This is the third case we retained (see Fig. 11c).

Figure 12 provides the results on the three selected cases. They are discussed below:

- From Fig. 12a, it appears clearly that the changes of the vertebral body boundaries do not impede the segmentation result, which is similar to the result presented in Fig. 6. Some imprecisions appears in the 3D interpolation; however, they are minor given the overall result.
- The hernia case presents a region which can be seen as a vacuum in the bone material. This vacuum is segmented out by the method since it differs from the bone in intensity. However, this particular point does not prevent the method to perform the segmentation correctly. Noteworthy, some inner parts of the vertebral body are included in the segmentation, which is not the case for standard lumbar vertebrae (e.g., Figures 6,12a). This is due to the relative overdensity induced by the hernia in the surrounding material.
- Finally, the cementoplasty case represents a more challenging test. It produces indeed an almost homogeneous bright region inside the vertebral body, leading to a global



distortion of the data in comparison with the standard cases. Nevertheless, the proposed method succeeds in providing a correct result (Fig. 12c), which does not cover the full vertebra volume but does represent most of the underlying vertebra. It also shows that natural overdensities of lower range can be handled by our method, the cementoplasty being one of the most extreme cases.

The results presented here show that the proposed method is robust to some of the most frequent particular cases met in clinical context. Furthermore, as it provides a correct result for a challenging case, one can expect it to be robust to most of the lower-intensity specificities. We provide in the next section further results covering the range of all vertebrae for two patients.

### **Integration examples**

We present in this section examples of full-spine segmentations. Two CT scans acquired in clinical routine were selected: the first one does not present specificities in its spine and can be considered as a standard healthy case. The second one presents vertebral compression, which corresponds to a flattening of the inter-vertebral disk and occurs mostly with age. From each scan, we manually defined the vertebral bounding boxes so as to enclose the vertebral bodies. Each volume is processed separately and placed at its initial location. The final results are presented in Fig. 13.

First of all, one can notice on both results that the segmentations include some separation artifacts due to the delimitations of the volumes (initial bounding box). Thus, some vertebral elements have been segmented out from the total result, as they are initially badly delimited. Note also that some non-vertebral elements have been segmented in; this is in particular the case with the L1 vertebra from the second case which present calcifications (as in Fig. 12b). This also happens with surrounding bones, such as the ribs for thoracic vertebrae and the pelvis for the L5 vertebra. Nevertheless, the segmentation is not impacted by these inclusions and performs correctly.

From Fig. 13, the changes in vertebrae shape and size are clear within one patient and also between patients. Indeed, it is noticeable that the seven lower vertebrae shape differs between the two cases. The vertebral compression induces deformations in the observed vertebral bodies; thus, their shape does differ from prior expectations. Note also that the second case presents an arthrosis between T12 and L1, causing vertebral bodies to join. Despite these specificities, our method performs correctly on all vertebrae; all of the vertebrae sub-structure are clearly segmented.

The results show that the proposed method can be successfully integrated within a simple spine processing and thus can be used in more complex framework. Given the results from the previous sections, one can expect our method to perform well in most practical situations, regardless of the vertebra type, position, and specificity. Further discussion on the method is given in the next section.

### **Discussion and conclusion**

Vertebra segmentation is a challenging task. The wide range of shapes, the high rate of aging modifications, and the pathologic alterations frequently encountered in real cases explain the difficulties of an automatic segmentation in daily practice patients.



Bounding box initialization method can be obtained automatically using state-of-the-art techniques.

The presented method results fulfill the requirements of automatic bone segmentation prior to registration processes, with an affordable computational time.

## Endnotes

<sup>1</sup>Note that this operation is similar to a morphological gradient, with two different structuring elements instead of one.

<sup>2</sup>We use volume height since it is shorter than depth or width for the vertebral volumes we process.

<sup>3</sup>Experiments show that other perimeter-based path such as using a coronal axis slice by slice or using concentric helix leads to similar results.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

JBC developed the segmentation method, conducted quantitative and specific evaluations and wrote the manuscript. ER conducted the quantitative experiments, analyzed the results, and helped the writing of the manuscript. EM and CC helped design the method and the writing of the manuscript. All authors read and approved the final manuscript.

## Author details

<sup>1</sup>ICube, Université de Strasbourg - CNRS, 67412 Illkirch, France. <sup>2</sup>SAMOVAR, Département CITI, CNRS, 91011 Évry, France.

Received: 29 September 2015 Accepted: 27 June 2016

Published online: 26 July 2016

## References

- Sharma N, Aggarwal LM (2010) Automated medical image segmentation techniques. *J Med Physics/Assoc Med Phys India* 35(1):3
- Nyúl LG, Kanyó J, Máté E, Makay G, Balogh E, Fidrich M, Kuba A (2005) Method for automatically segmenting the spinal cord and canal from 3D CT images. In: *Computer Analysis of Images and Patterns*. Springer, pp 456–463
- Malik J, Belongie S, Leung T, Shi J (2001) Contour and texture analysis for image segmentation. *Int J Comput Vis* 43(1):7–27
- Kim Y, Kim D (2009) A fully automatic vertebra segmentation method using 3D deformable fences. *Comput Med Imaging Graph* 33(5):343–352
- Mirzaalian H, Wels M, Heimann T, Kelm BM, Suehling M (2013) Fast and robust 3D vertebra segmentation using statistical shape models. In: *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE, Osaka*. pp 3379–3382
- Rasoulian A, Rohling R, Abolmaesumi P (2013) Lumbar spine segmentation using a statistical multi-vertebrae anatomical shape + pose model. *IEEE Trans Med Imaging* 32(10):1890–1900
- Heimann T, Meinzer HP (2009) Statistical shape models for 3D medical image segmentation: a review. *Med Image Anal* 13(4):543–563
- Klinder T, Ostermann J, Ehm M, Franz A, Kneser R, Lorenz C (2009) Automated model-based vertebra detection, identification, and segmentation in CT images. *Med Image Anal* 13(3):471–482
- Ma J, Lu L, Zhan Y, Zhou X, Salganicoff M, Krishnan A (2010) Hierarchical segmentation and identification of thoracic vertebra using learning-based edge detection and coarse-to-fine deformable model. In: *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2010*. Springer, Beijing, pp 19–27
- Glocker B, Feulner J, Criminisi A, Haynor DR, Konukoglu E (2012) Automatic localization and identification of vertebrae in arbitrary field-of-view CT scans. In: *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2012*. Springer, Nice, pp 590–598
- Chen T, Huang TS, Yin W, Zhou XS (2005) A new coarse-to-fine framework for 3D brain MR image registration. In: *Computer Vision for Biomedical Image Applications*. Springer, Beijing, pp 114–124
- Duda RO, Hart PE, Stork DG (2012) *Pattern classification*. John Wiley & Sons
- Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Susstrunk S (2012) SLIC superpixels compared to state-of-the-art superpixel methods. *Pattern Anal Mach Intell IEEE Trans* 34(11):2274–2282
- Sprawls P. Jr. (1995) *Physical principles of medical imaging*. Aspen Publication, Rockville
- Nock R, Nielsen F (2004) Statistical region merging. *Pattern Anal Mach Intell IEEE Trans* 26(11):1452–1458
- Geman S, Geman D (1984) Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *Pattern Anal Mach Intell IEEE Trans* 6(6):721–741
- Baum LE, Petrie T, Soules G, Weiss N (1970) A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann Math Stat* 1:164–171
- Celeux G, Diebolt J (1986) L'algorithme SEM: un algorithme d'apprentissage probabiliste: pour la reconnaissance de mélange de densités. *Revue de statistique appliquée* 34(2):35–52
- Sagan H (1994) *Space-filling curves*, Vol. 18. Springer

20. Bricq S, Collet C, Armspach JP (2008) Unifying framework for multimodal brain MRI segmentation based on hidden Markov chains. *Med Image Anal* 12(6):639–652
21. Derrode S, Pieczynski W (2004) Signal and image segmentation using pairwise Markov chains. *Signal Process IEEE Trans* 52(9):2477–2489
22. Fjortoft R, Delignon Y, Pieczynski W, Sigelle M, Tupin F (2003) Unsupervised classification of radar images using hidden Markov chains and hidden Markov random fields. *Geosci Remote Sensing IEEE Trans* 41(3):675–686
23. Marroquin J, Mitter S, Poggio T (1987) Probabilistic solution of ill-posed problems in computational vision. *J Am Stat Assoc* 82(397):76–89
24. Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc Series B (Methodological)* 39:1–38
25. MacQueen J, et al. (1967) Some methods for classification and analysis of multivariate observations. In: *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1, Berkeley. pp 281–297
26. Yao J, Burns JE, Munoz H, Summers RM (2012) Detection of vertebral body fractures based on cortical shell unwrapping. In: *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2012*. Springer. pp 509–516
27. SpineWeb platform website. <http://spineweb.digitalimaginggroup.ca/>

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

---

Submit your next manuscript at ▶ [springeropen.com](http://springeropen.com)

---