

# 2-Step Robust Vertebra Segmentation

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

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

e-mail: [jb.courbot, c.collet]@unistra.fr, edmond.rust@etu.unistra.fr

<sup>2</sup> Département CITI, CNRS, UMR 5157 SAMOVAR, 91011 Évry, France

e-mail: emmanuel.monfrini@it-sudparis.eu

**Abstract**— Knowledge of vertebra location, shape and orientation is crucial in many medical applications such as orthopedics or interventional procedures. The wide range of shapes, joint alterations and pathological cases encountered in an aging population makes automatic segmentation sometimes challenging. This paper presents a new automated vertebra segmentation method for 3D CT data which tackles these problems. This method has two consecutive main steps: first a new coarse-to-fine method produces a coarse shape of the vertebra, then a Hidden Markov Chain (HMC) segmentation using a specific volume transformation refine the segmentation. No shape prior is used thus allowing most frequent non-standard and pathological cases handling. We experiment this method on a set of standard vertebrae and on non-standard cases as encountered in daily practice. After expert validation, we show that our method is robust to shape and luminance changes, and provides correct segmentation for pathological cases.

**Keywords**— Clinical imagery, Automatic vertebra segmentation, Coarse-to-fine modeling, SLIC Clustering, Hidden Markov Chains.

## I. INTRODUCTION

A precise knowledge of vertebrae location, shape and orientation is essential for various medical application, such as spine lesions follow-up, multi-modality imaging, preoperative planning or interventional radiology treatments. Furthermore, some major concerns for segmentation are the varying neighborhood and shape in a single patient, and the possibility of pathological cases, affecting shape, size or region intensity. Hence, most of the published works about vertebra segmentation seem to be developed and evaluated in a young population in which vertebra are anatomically well separated (*e.g* [8], [9]), with a consequent lack of information about the robustness of the presented schemes toward daily practice cases.

Medical image segmentation can be divided in three types: the iconic, the texture-based and the edge-based methods [14]. The vast majority of recent works [8], [9], [10], [11], [13] on vertebrae segmentation are edge-based, relying mostly on deformable models to perform an adaptation of prior data to the vertebrae volume.

Nevertheless, two main key issues limit these works: (1) the algorithms use complex shape description, dramatically increasing global processing time and (2), methods were validated on a limited set of vertebrae in terms of scope (lumbar, thoracic or cervical) and healthiness (middle-aged patient, healthy cases). The proposed method overcomes these limitations, since it does not rely on prior data nor on complex shape

descriptors. To restrain the computation time, a fast coarse-to-fine pre-segmentation algorithm dropping voxel clusters from the data volume is proposed. Then a finer Hidden Markov Chain (HMC) voxel segmentation is processed, following a new volume-to-chain spiral transform.

Through this paper, the processed volumes will be obtained from the Computed Tomography (CT) imaging modality volumes interpolated to obtain isotropic voxels of size  $(0.98\text{mm})^3$ , 0.98mm being the most common slice thickness encountered in the tested volumes. We assume in the following that the vertebral volumes are properly isolated in bounding boxes, delimited by their inter-vertebral disk and corresponding mean planes. Many methods of *vertebra localization* can produce such delimitation [7], [8], [9]. The first segmentation step is built on the basis of a statistical testing of coherent voxel cluster, and is therefore robust to local and global luminance changes. The second step aims at discriminating the two classes in the remaining volume within a robust HMC framework and thus performs a consistent voxel-level segmentation. No shape prior were introduced in the algorithms and thus the method can deal with any type of standard vertebrae from lumbar to cervical as well as non-standard cases one can expect in clinical context.

Sections II and III detail the two main steps of the method. Section IV explains the experiments and the results, and a conclusion is given in Section V.

## II. COARSE-TO-FINE PRE-SEGMENTATION

In medical imaging, coarse-to-fine methods are mostly used to perform fast registration (see *e.g* [10]). Besides, image clustering [5] is a well-known tool to produce consistent high-level information. It is desirable to use both approaches to rapidly ensure a first accurate and consistent anatomical volume estimation. Thus a new algorithm is introduced to fulfill these requirements. It processes layer by layer iteratively in three main steps in a fashion summarized in Fig. 1.

### A. Layer Construction

This step isolates the external layer of voxels on which further processing will be applied. Let  $\hat{V}_{j-1}$  be the binary partially-segmented volume obtained at the step  $j - 1$ , or the initial volume  $V_0$  if the method starts, and let  $V_j$  be the layer

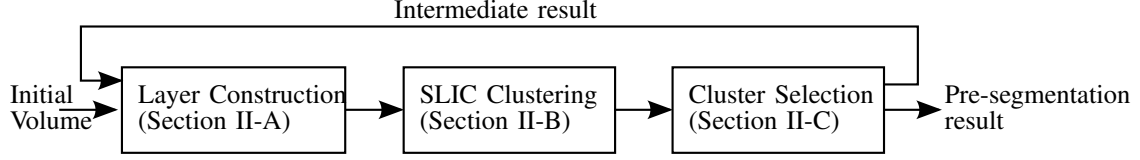


Fig. 1: Flowchart of the pre-segmentation step.

volume at step  $j$ . Then  $V_j$  is given by :

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

$V_0$  is the initial volume boundary, and the operators  $\oplus$  and  $\ominus$  stand for morphological dilatation and erosion respectively, with ball structuring elements  $R(\cdot)$  of diameter  $O_j$  and  $I_j$  for outer and inner thicknesses respectively.

### B. Layer Clustering

A clustering method related to the SLIC method proposed in [1] is developed. The authors present a clustering method for 2D color images, based on a mixed distance-luminance criterion. It is generalized here in the 3D gray-level case and we refer to it thereafter as "SLIC-3D". The spatial and luminance features are the 3D Euclidean distance and intensity range expressed in Housfield Units (HU) [15] respectively. A given voxel  $i$  is then represented by its spatial coordinates  $(x_i, y_i, z_i)$  and its luminance  $l_i$ . Given the centroid  $C_k = [x_k, y_k, z_k, l_k]^T$  of the  $k$  cluster, the mixed distance  $D_m$  combines the four features between  $C_k$  and any voxel  $i$ :

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

where  $d_s$  is the 3D Euclidean distance ;  $d_c$  is the luminance gap between  $C_k$  and  $i$ ;  $d_c(C_k, i) = |l_k - l_i|$ ;  $m$  balances the spatial and luminance features and  $S$  is the number of pixels a cluster should contain. The SLIC-3D procedure is described in Algorithm 1.

### C. Clusters Selection

A test is developed to assess if the layer clusters are part of the vertebra. The input of this test is the mean luminance  $l_k$  of each cluster  $k$ . An adaptive test is developed so as to ensure robust and consistent acceptance or rejection of clusters, and to be efficient for non-standard cases. The test is built on the basis of the *Statistical Region Merging* proposed in [12]. Whereas the authors process the test for a pixel pair set, here the clusters are tested with respect to a *reference luminance*  $l_0$  corresponding to the typical bone tissue luminance in CT scans. For a given cluster  $k$  let the *bone merging* predicate be:

$$\mathcal{P}_0(k) := |l_k - l_0| \leq b(k) \quad (3)$$

$b(\cdot)$  is a *merging threshold* defined [12] as :

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

where  $|k|$  is the number of voxels in  $k$ ,  $g$  is the gray level range, and  $\delta$  is the affordable probability of error for the predicate.  $Q$  stands for the expected number of underlying independent random variables for the current region. An *interiority* predicate is added as an alternative so that interior clusters remains in the result. Let  $v = (x_v, y_v, z_v)$  the remaining volume center, then for a given cluster  $k$  and any of its neighbor  $k'$  :

$$\mathcal{P}_I^k(k') := d_s(C_{k'}, v) \leq d_s(C_k, v) \quad (4)$$

where  $C_k$  and  $C_{k'}$  are the centroids of  $k$  and  $k'$  respectively, and  $d_s$  is the 3D Euclidean distance. The two predicates (3) and (4) are merged in the following *vertebral* predicate to test a cluster  $k$ :

$$\begin{aligned} \mathcal{P}(k) &:= \mathcal{P}_0(k) \\ \text{or } \mathcal{P}_I^k(k) \quad \forall k' \text{ such as } &\begin{cases} k \text{ and } k' \text{ are neighbor,} \\ \mathcal{P}(k') \text{ is valid.} \end{cases} \end{aligned} \quad (5)$$

### D. Model and Parameters

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

- The two first  $S$  parameter from SLIC-3D (Alg. 1) values will be higher than the latter ones, to process from coarse to fine. The size parameters from Eq. (1) will also be defined with  $S$ .
- The  $m$  parameter (Alg. 1) was chosen with decreasing values of  $m$  to exclude first spatially-coherent and then intensity-coherent clusters.
- The statistical parameters  $g$ ,  $Q$  and  $\delta$  (Pred. 3) can be produced automatically with the cluster to proceed.
- The reference intensity  $l_0$  (Pred. 3) is the typical intensity of bones in CT scans, always available in daily practice.
- The number of iterations  $J$  is given by  $J = \left\lceil \frac{h - S_1 - S_2}{S_j} \right\rceil$ , where  $h$  is the volume height,  $\lceil \cdot \rceil$  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. The entire pre-segmentation method that we called "carving" is summarized in Alg. 2 and illustrated in Fig. 2.

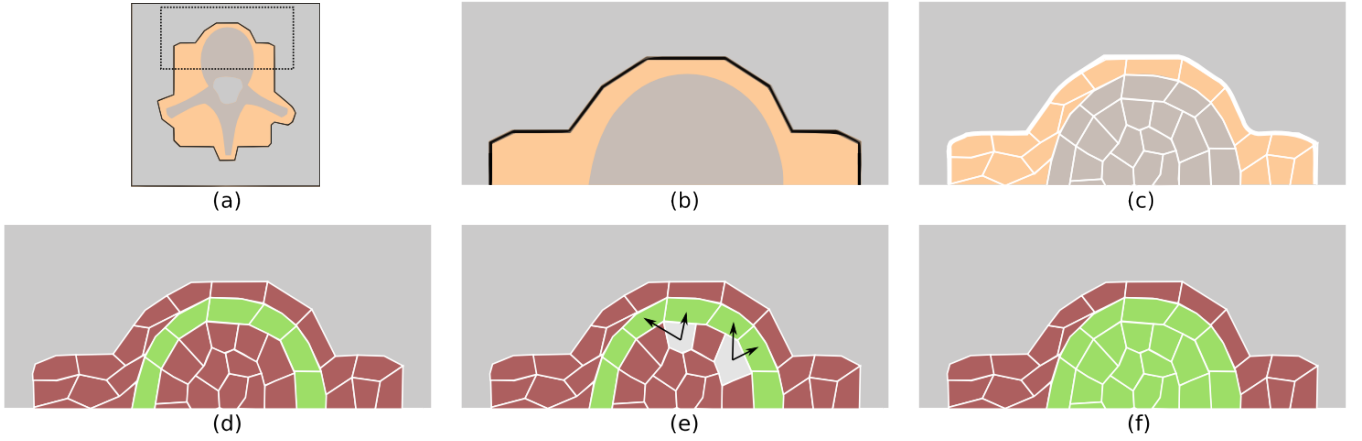


Fig. 2: Graphical summary of a pre-segmentation step. Gray regions are not proceeded at this step, orange regions are the current layer and include the brown vertebrae region. White limits represent cluster boundaries. (a) Whole slice with the zoom-in region (b) in the dotted limits. (c) Result from the SLIC-3D clustering. (d) Possible outcome of the selection step with only the bone merging predicate (3) (e) Neighborhood search example for the two light-gray clusters (f) Expected outcome of the vertebral predicate (5).

### III. HMC SEGMENTATION

The result from the carving step includes most of the underlying vertebrae and is smaller than the initial volume. It is transformed into a region of interest (ROI) built as a 10-mm-diameter ball morphological dilatation, to allow efficient statistical separation. The ROI will be processed in the Hidden Markov Chain (HMC) framework with a specific volume transformation, and with the Baum-Welch algorithm [2] in the unsupervised SEM framework [4] for segmentation. This section explains the specificity of this work, an application instance of the Baum-Welch and SEM algorithms can be found in [6]. A common transformation path is the Hilbert scan, which is known to be successful for transforming 2D images into chains (see *e.g.* [3], [6]). However, as it creates artifacts in segmentations, it does not always ensure a smooth result. Thus a new volume-to-chain transformation is introduced based on the ROI shape, and is explained in Fig. 3. The segmentation algorithm first transforms the volume along the spiral path. The mixture parameters are then estimated with the SEM algorithm. Its convergence is evaluated through Maximum Posterior Mode (MPM) estimates computed at each iteration. The segmented chain is estimated with the final MPM estimator. Finally the segmented volume is rebuilt along the initial chain path. Illustrative results are presented in Fig. 4. The gain of the HMC segmentation step is clear on the 3D interpolations: the results match our expectations of the vertebral volumes and does not include processing artifacts. Further extensive and comparative results are presented in the next section, as well as robustness examples.

### IV. RESULTS

In this section, the method performances are evaluated through a large set of standard cases as well as on specific

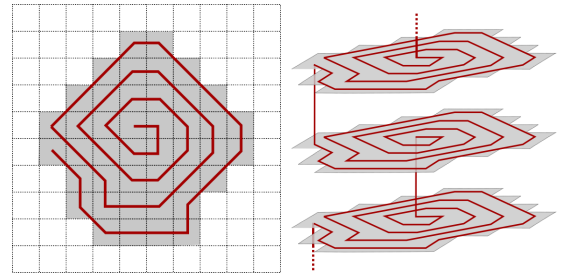


Fig. 3: Spiral transform illustration. The shaded regions represent the ROI sections and the red line follows the chain path. The spiral covers each slice by concentric perimeters, alternatively inward and outward. Left:  $10 \times 10$  pixel slice. Right: 3 consecutive slices.

non-standard and pathological cases.

#### A. Standard Cases: Quality Assessments

The method was evaluated on a set of 339 vertebral volumes from the whole spine of 15 consecutive patients in an oncologic tertiary center, excluding 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. Each volume was processed by the proposed method and a two-class K-means algorithm as a benchmark. The latter was selected as an instance of non-prior segmentation algorithm. Since it fails when air is present in the volume, a sub-sample without air has been created for accurate comparison. The results were visually inspected by an expert with respect to the following ranking:

- Excellent (100): the vertebrae is exactly delimited inside its bounding box.

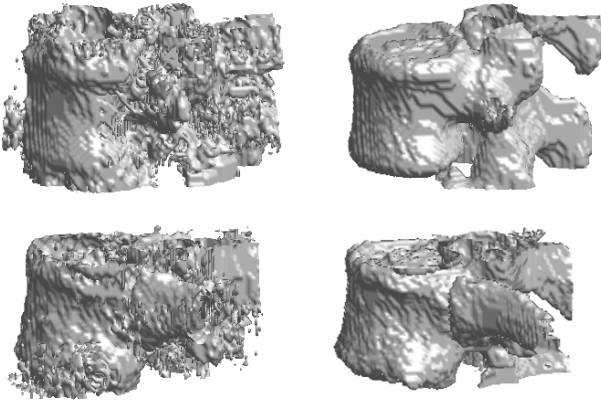


Fig. 4: 3D interpolations of the intermediate and final segmentation results. The first and second rows correspond to lumbar (L4) and thoracic (T11) vertebrae processing respectively. The first column represents the result of the Carving method (Alg.2) and the second column contains the results after the HMC segmentation (Section III). Note that the observed granularity corresponds to the voxel size, which is the minimal size addressed in this work.

- 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 failed to proceed.

Figure 5 illustrates the results obtained for the two methods on the same volume and Table 1 summarizes the results. Considering that good and excellent results both provide sufficient data for vertebrae segmentation and advanced treatment procedures, our method provides about 78% of successful results on the subsample in which K-means gives 72% of successful results, whereas on the full sample the method yields 67% of successful results (and K-Means returns 38%). As 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. The average

Table 1: Results on the subsample without air and on the full sample. This parting is made to provide accurate comparative results on the sub-sample. Average scores summarizes the essential outcomes.

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

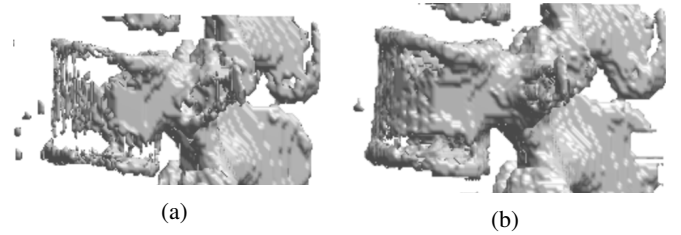


Fig. 5: Illustrative comparison between K-means segmentation (a) and the proposed method (b) for the L3 vertebra in a patient with marked osteoporosis. K-means segmentation was rated 50, whereas the proposed segmentation was rated 75.

processing time is 36 seconds by vertebra.

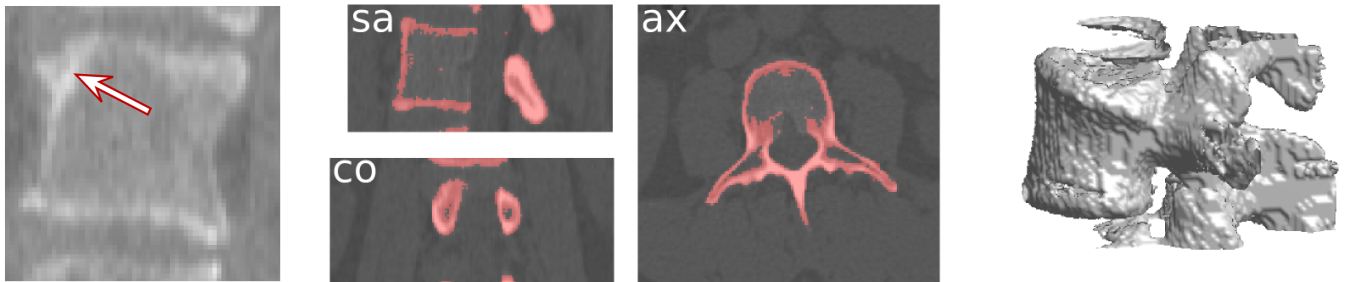
### B. Pathological Cases: Evaluating Robustness

In a clinical implementation perspective, robustness to the most frequent non-standard cases is mandatory. The two main key points are changes in shape and in intensity of the object to segment. They correspond to anatomical deformations and structural, 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. 6b). 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. 6d). 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 and was also retained (see Fig. 6f).

Fig. 6c, 6e and 6g provide the method outcomes on the three selected cases as well as the corresponding discussions. 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.

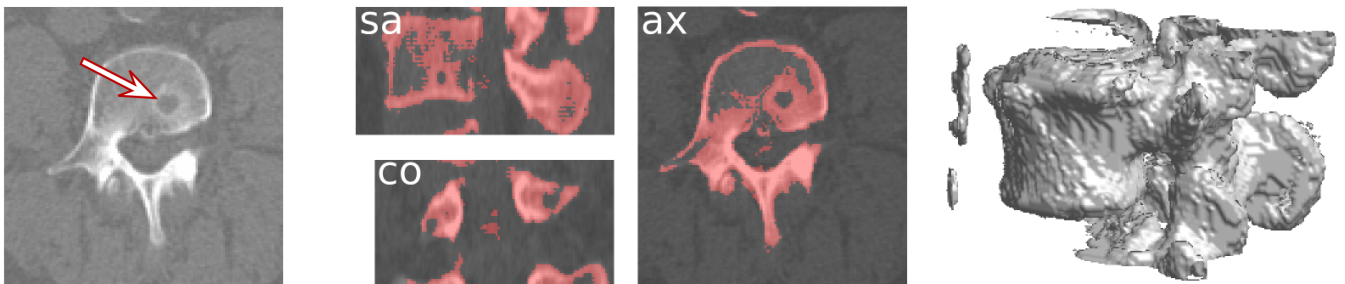


(a) Segmentation result for a standard L3 vertebra. The three sectionnal views are the volume sagittal, axial and coronal middle slices, the result is represented by its red superimposition. The segmentation clearly delimits the vertebra. Due to its high similarity to the external neighborhood, the inner vertebral body is segmented out.



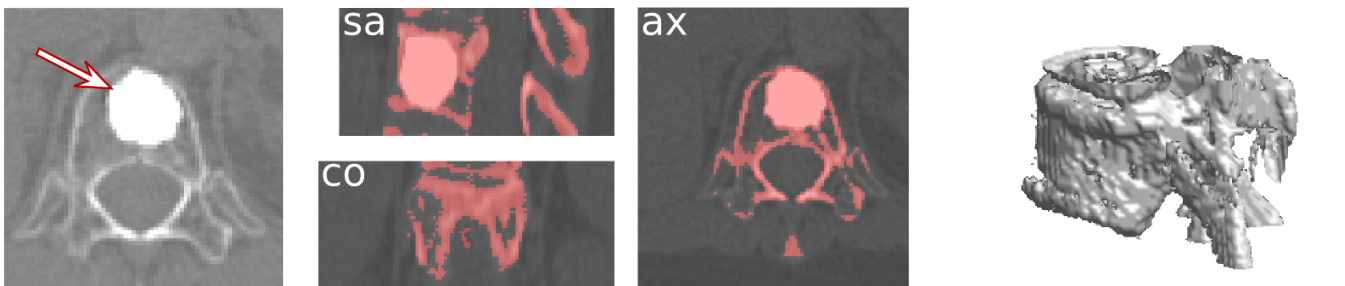
(b) Arthrosis in a L3 vertebra.

(c) The changes induced by arthrosis cleraly do not impede the segmentation result, which is similar to the result presented in Fig. 6a. Some errors appears from the 3D interpolation, however they are minor given the overall result.



(d) Hernia in a L4 vertebra.

(e) The hernia region is segmented out by the method since it differs from the bone in intensity, while the surrounding regions are included due to relative over-densities. The method yet performs correctly. Note also that some calcification were also segmented since they are close spatially and in intensity to the vertebra.



(f) Cementoplasty in a T12 vertebra.

(g) The result does not cover the full vertebra volume but does represent most of the underlying vertebra. It also shows that natural over-densities of lower range can be handled by our method, the cementoplasty being one of the most extreme case.

Fig. 6: Segmentation results in standard and pathological cases.

## V. CONCLUSION

Vertebrae segmentation remains a complex task, partly due to the variety of shape, aging-related modifications, and pathologic alterations. Nevertheless, the proposed method fulfills the requirements of a non-prior segmentation, which can be integrated into a spine processing procedure. This method could be more generally used for bone segmentation. It could also benefit from an automatic vertebra separation tool.

## ACKNOWLEDGMENT

The authors would like to thank Pr. A. Gangi, Head of the Imagery Department of the Hopitaux Universitaires de Strasbourg, for his interest in this work.

## REFERENCES

- [1] Achanta, Radhakrishna and Shaji, Appu and Smith, Kevin and Lucchi, Aurelien and Fua, Pascal and Susstrunk, Sabine. SLIC superpixels compared to state-of-the-art superpixel methods. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(11):2274–2282, 2012.
- [2] Baum, Leonard E and Petrie, Ted and Soules, George and Weiss, Norman. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *The annals of mathematical statistics*, 164–171, 1970.
- [3] Bricq, Stephanie and Collet, Christophe and Armspach, Jean-Paul. Unifying framework for multimodal brain MRI segmentation based on hidden Markov chains. *Medical image analysis*, 12(6):639–652, 2008.
- [4] Celeux, Gilles and Diebolt, Jean. L'algorithme SEM: un algorithme d'apprentissage probabiliste pour la reconnaissance de mélange de densités. *Revue de statistique appliquée*, 34(2):35–52, 1986.
- [5] Duda, Richard O and Hart, Peter E and Stork, David G. Pattern classification. *John Wiley & Sons*, 2012.
- [6] Fjortoft, Roger and Delignon, Yves and Pieczynski, Wojciech and Sigelle, Marc and Tupin, Florence. Unsupervised classification of radar images using hidden Markov chains and hidden Markov random fields. *Geoscience and Remote Sensing, IEEE Transactions on*, 41(3):675–686, 2003.
- [7] Glocker, Ben and Feulner, Johannes and Criminisi, Antonio and Haynor, David R and Konukoglu, Ender. Automatic localization and identification of vertebrae in arbitrary field-of-view CT scans. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI*, 2012.
- [8] Kim, Yiebin and Kim, Dongsung. A fully automatic vertebra segmentation method using 3D deformable fences. *Computerized Medical Imaging and Graphics*, 33(5):343–352, 2009.
- [9] Klinder, Tobias and Ostermann, Jörn and Ehm, Matthias and Franz, Astrid and Kneser, Reinhard and Lorenz, Cristian. Automated model-based vertebra detection, identification, and segmentation in CT images. *Medical image analysis*, 13(3):471–482, 2009.
- [10] Ma, Jun and Lu, Le and Zhan, Yiqiang and Zhou, Xiang and Salganicoff, Marcos and Krishnan, Arun. 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.
- [11] Mirzaalian, Hengameh and Wels, Michael and Heimann, Tobias and Kelm, B Michael and Suehling, Michael. Fast and robust 3D vertebra segmentation using statistical shape models. In *Engineering in Medicine and Biology Society (EMBC), 35th Annual International Conference of the IEEE*, 2013.
- [12] Nock, Richard and Nielsen, Frank. Statistical region merging. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(11):1452–1458, 2004.
- [13] Rasouljan, Abtin and Rohling, Robert and Abolmaesumi, Purang. Lumbar Spine Segmentation Using a Statistical Multi-Vertebrae Anatomical Shape+Pose Model. *Medical Imaging, IEEE Transactions on*, 32(10):1890:1900, 2013.
- [14] Sharma, Neeraj and Aggarwal, Lalit M. Automated medical image segmentation techniques. *Journal of medical physics/Association of Medical Physicists of India*, 35(1)3, 2010.

- [15] Sprawls, Perry. Physical principles of medical imaging. *Medical Physics Pub.*, 1995.

## Algorithm 1: SLIC - 3D

---

**Require:** Bounding box  $V$  obtained from a previous processing, parameters  $m$  and  $S$  from Eq. (2), 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  $v_i$ :  $L(i) \leftarrow -1$  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  $D_m(C_k, i)$  (2).

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

$L(i) = k$

$D(i) = D_m(C_k, i)$

**end if**

**end for**

**end for**

  Update the cluster centroids.

**end for**

---

## Algorithm 2: Carving

---

**Require:** Initial volume  $V_0$

**Ensure:** Binary pre-segmentation of  $V_0$

---

Compute the number of iterations  $J$ .

**for** each iteration  $j$  **do**

  Build  $V_j$ : (1) with  $O_j$  and  $I_j$ .

  Update  $g$ ,  $Q$  and  $\delta$  with the current values.

  Cluster  $V_t$ : Alg. 1 with  $S_j$  and  $m_j$ .

**for** each cluster  $k$  **do**

    Test  $\mathcal{P}(k)$ : (5) 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:  $I_j = S_j \forall j$ ,  $O_1 = 0$ ,  $O_j = (S_{j-1} + S_j)/2 \forall j > 1$ ,  $m_1 = 100$ , linearly decreasing until  $m_J = 70$ ,  $S_1 = 23\text{mm}$ ,  $S_2 = 17\text{mm}$ ,  $S_j = 11\text{mm} \forall j > 2$ ,  $g$  is set to the intensity range in the current  $V_t$ ,  $Q = 0.4 \times g$ ,  $l_0 = 1300 \text{HU}$ ,  $\delta = 1/|V_t|$

---