Evaluation of the Inter-observer Cardiac Chamber Contour Extraction versus a Level Set Algorithm

Diogo Roxo[1], José Silvestre Silva[1,2], Jaime B. Santos[3], Paula Martins[4], Eduardo Castela[4], and Rui Martins[5]

[1] Department of Physics, FCTUC, University of Coimbra, Portugal
[2] Instrumentation Center, FCTUC, University of Coimbra, Portugal
[3] Mechanical Engineering Center, FCTUC, University of Coimbra, Portugal
[4] Department of Pediatric Cardiology, Pediatric Hospital of Coimbra, Portugal
[5] HUC, Hospital of University of Coimbra, Largo Mota Pinto, Coimbra, Portugal
diogoroxo2@gmail.com, jsilva@ci.uc.pt, jaime@deec.uc.pt

Abstract. Segmentation of echocardiography images presents a great challenge because these images contain strong speckle noise and artifacts. Besides, most ultrasound segmentation methods are semi-automatic, requiring initial contour to be manually identified in the images. In this work, a level set algorithm based on the phase symmetry approach and on a new logarithmic based stopping function is used to extract simultaneously the four heart cavities in a fully automatic way. Then, those contours are compared with the ones obtained by four physicians to evaluate the performance, reliability and confidence for eventual clinical practice. That algorithm evaluation versus clinicians' performance is made using several metrics, namely Similarity Region, Hausdorff distance, Accuracy, Overlap, Sensitivity, and Specificity. We show that the proposed algorithm performs well, producing contours very similar to the physicians' ones with the advantage of being an automatic segmentation technique. The experimental work was based on echocardiography images of children.

Keywords: heart segmentation, echocardiographic images, phase symmetry, level set, similarity index.

1 Introduction

Medical ultrasonography is an important tool for imaging the heart structures, since it is noninvasive, safe, portable, and the images are available in real-time and the cost is low compared to other medical imaging techniques [1, 2]. Another reason for its success is that the information it provides is very helpful in clinical diagnosis of heart diseases, and also in emerging areas such as image-guided interventions [3, 4]. Anatomical quantitative information about the heart chambers is crucial for clinicians, namely wall motion, valves function, and volume estimation.

Today, most of the used methods are still semi-automatic requiring human interactivity to select the correct region of interest (ROI) [5-7]. Fully automatic segmentation [8] can reduce this intra-expert variability and provide clinical valuable information about the cardiac cavities and their respective volumes, working also as an important

M.M. Cruz-Cunha et al. (Eds.): CENTERIS 2011, Part III, CCIS 221, pp. 97–106, 2011.

step to achieve 3D heart reconstruction, which quality is greatly influenced by the segmentation accuracy. Most recent segmentation methods based in deformable models, as parametric models or geometric models including level-set, are used to identify not only heart regions [9-11], but also lungs [12-18], vascular or neural structures[19] and to assist the classification of pathologies [20, 21].

Active contours have been used for image segmentation and boundary tracking since the first introduction of snakes by Kass [22]. The basic idea is to start with initial boundary shapes represented in a form of closed curves, i.e. contours, and iteratively modify them by applying shrink/expansion operations according to the constraints of the image. Methods to accomplish those operations were introduced by Kass [22], and then reformulated in the context of partial differential equations [23] (PDE) by Caselles [24] using the level set framework.ï

Geometric active contour model was the first level set implemented active contour model for the image segmentation problem. It was simultaneously proposed by Caselles [24] and by Malladi [25]. This model is based on the theory of curve evolution and geometric flows. The geodesic active contour model proposed by Caselles [24] allows connecting classical "snakes" based on energy minimization and geometric active contours based on the theory of curve evolution. In order to improve the segmentation performance, the integration of edge and region based information sources using active contours has been proposed by a few authors. Paragios [26] proposed the geodesic active region model for supervised segmentation, integrating edge and region-based in an energy function. Zhang [27] also apply this association to achieve robust and accurate segmentation results. He tried to solve the problem from a more geometric perspective, describing the advantages of the model over Chan method [28] on medical images with complex backgrounds.

The use of level set theory has provided more flexibility and convenience in the implementation of active contours. Depending on the implementation scheme, active contours can use various properties used for other segmentation methods such as edges, statistics, and texture. In the present work, we describe the segmentation method based on level set that is capable to identify simultaneously the four heart chambers. The used phase symmetry approach, which acts in the frequency domain extracting low level features, is of great importance for the success of the algorithm [9, 10, 29]. The segmentation results are compared with those provided by the reference contours provided by four physicians, using several metrics of similarity.

The outline of the paper is as follows. In section 2, we describe the segmentation method and the metrics used for the evaluation. Then, the algorithms are applied to children echocardiographic images. The results and discussion are presented in section 3, and finally, the conclusions are formulated in section 4.

2 Methodology

2.1 Overview of the Segmentation Algorithm

The applied segmentation algorithm is composed by several phases. First, low-level features are extracted from the collected images (see figure 1) using a Phase-based Symmetry Detection, a procedure based on gradient/luminance information and on Log

Gabor wavelet by finding symmetric or partial symmetric components in the frequency space, allowing the enhancement of shapes in the images (see figure 1b) [29].

Second, a modified level set is used that minimizes the function by solving the correspondent partial differential equation. A higher dimensional function $\phi(x,t)$ is manipulated to evolve a contour (or surface) implicitly, where the zero level is used to extract the evolving contour $C = \{x \mid \phi(x) = 0\}$.

$$\frac{\partial \phi}{\partial t} = |\nabla \phi| \left[div\left(\frac{\nabla \phi}{\nabla \phi}\right) + v \right] P \tag{1}$$

The initial function ϕ_0 is a mask based on the Euclidean distance, where the central pixel has the biggest value and the remaining pixels decrease towards zero till they meet the image limits. Then, the function is updated at each iteration, based on shrink and expansion operations depending of the level set curvature and the stopping function P.

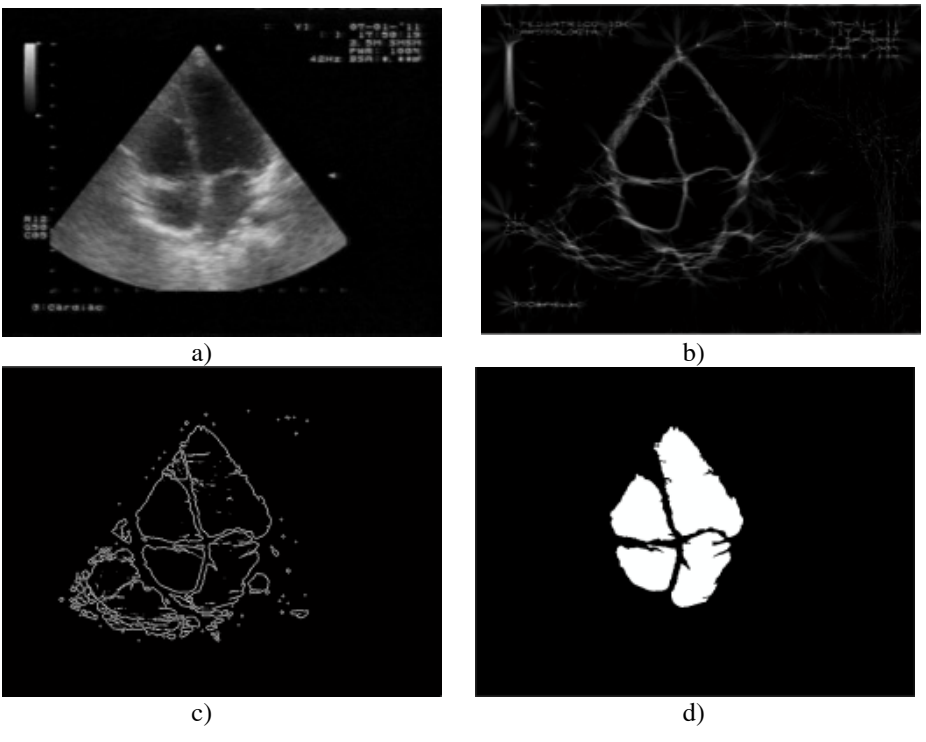


Fig. 1. Segmentation method: (a) original Image, (b) detecting low-level features, (c) applying the segmentation method, (d) post-processing

The stopping function P is represented by:

$$P = \log\left(\left|\frac{I - \varepsilon}{\gamma}\right| + 1\right) \tag{2}$$

where is the average intensity value of the image I and is the dynamic range of the region being analyzed. This stopping function of logarithmic variation performs adjustments to the regions in study until the convergence is reached. This happens after few iterations. The segmentation result is illustrated in figure 1c.

The last step consists in the post-processing, where all the smaller regions detected are removed by using morphological operations in order to obtain smoothed contours.

2.2 Validation Metrics

To evaluate the performance of the segmentation algorithm, we compare the extracted contours with the reference contours drawn by some physicians. For that goal the following six metrics are used.

Similarity Region (SR) is based on the difference between two regions [30, 31].

$$SR = 2\,\frac{A\,.\,B}{A\,+\,B} \tag{3}$$

The Similarity Region varies from zero to one where the value '0' corresponds to the total dissimilarity of the two regions, and the value '1' to the exact overlapping.

Hausdorff Distance (h) defines a distance between two given curves (A and B). Whenever an overlap between them occurs that distance is zero, otherwise the distance from each point in A to all points in B [32, 33] is calculated, and the smallest value is kept. Finally, the highest value is searched from the collected set of distances.

$$\mathrm{h}(\mathrm{A},\mathrm{B}) = \max_{\mathrm{a}\,\in\,\mathrm{A}} \{\ \min_{\mathrm{b}\,\in\,\mathrm{B}} \{\ \mathrm{d}(\mathrm{a},\mathrm{b})\ \}\ \} \tag{4}$$

The Hausdorff Distance does not have a defined variation range as the other metrics; the minimal and best value is zero (contours overlapping).

Accuracy, Overlap, Sensitivity, Specificity: These four measures are based on the relation between two regions defined by two contours A and B [34]. The pixels of both regions are classified as: N_{TP} = pixels present in both regions; N_{TN} = pixels absent in both regions; N_{FP} = pixels present in B, absent in A; N_{FN} = Pixels present in A, absent in B. Thus,

$$\mathrm{Accuracy} = \frac{\mathrm{N_{TP}} + \mathrm{N_{TN}}}{\mathrm{N_{TP}} + \mathrm{N_{TN}} + \mathrm{N_{FP}} + \mathrm{N_{FN}}} \tag{5}$$

$$\mathrm{Overlap} = \frac{\mathrm{N_{TP}}}{\mathrm{N_{TP}} + \mathrm{N_{FP}} + \mathrm{N_{FN}}} \tag{6}$$

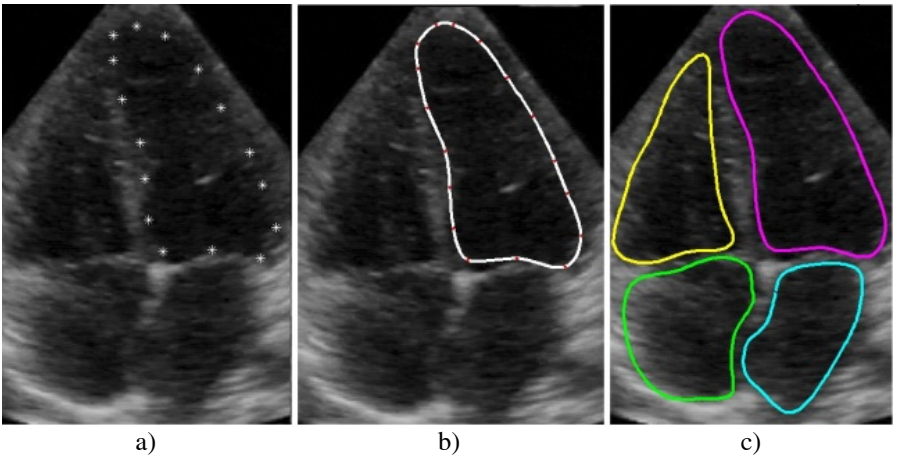
$$\mathrm{Sensitivity} = \frac{\mathrm{N_{TP}}}{\mathrm{N_{TP}} + \mathrm{N_{FN}}} \tag{7}$$

$$\mathrm{Specificity} = \frac{\mathrm{N_{TN}}}{\mathrm{N_{TN}} + \mathrm{N_{FP}}} \tag{8}$$

Accuracy is the ratio of correctly classified points in the ROI; Overlap gives the amount of intersection between A and B; Sensitivity gives the information of the pixels that are correctly classified and belong to the contour; Specificity measures the proportion of pixels that are correctly classified as not belonging to the contour. All of these metrics have a range from zero to one, where one represents the optimal value.

2.3 Dataset

The data set has 240 cardiac chamber regions extracted from sixty images with a resolution of 768×576 pixels, selected from several children echocardiographic videos. The reference cavity contours drawn by the experts, were obtained with the help of a graphical interface, then minimizing the inherent error associated to hand drawn contours and the following necessary digitalization. Figure 2a-c illustrates the procedures for the references delineation.



a) b) c)

Fig. 2. Reference contours acquisition: (a) points defined by the user, (b) cubic spline interpolation, (c) four cavity contours used as reference

3 Results and Discussion

This work consists of the performance analysis of all extracted contours from the dataset by the proposed algorithm, and the reference contours drawn by physicians. For that, two distinct inter-observer comparisons will be defined, namely Algorithm vs. Physician and Physician vs. Physician.

3.1 Comparison between Contours

The performance evaluation of the algorithm versus clinicians was accomplished by using the metrics described previously.

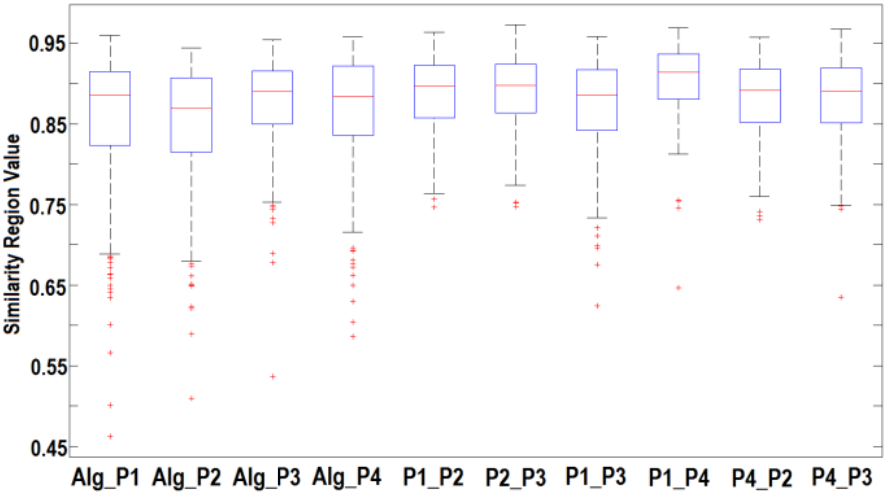


Fig. 3. Similarity Region Box-Plot Graphic: Alg - Algorithm; P1, P2, P3, P4 - four Physicians

Figure 3 shows the results for the Similarity Region metric. The median values for the first four columns that represent the Algorithm vs. Physician behavior are 0.89, 0.87, 0.89, and 0.88. The remaining median values are 0.90, 0.90, 0.89, 0.91, 0.89, 0.89 that are slightly superior to the previous ones, however the difference is very small (0.01 in average), then illustrating similar performances for the analyzed comparisons. It was also verified that the metrics Accuracy, Overlap, Sensitivity and Specificity show similar results.

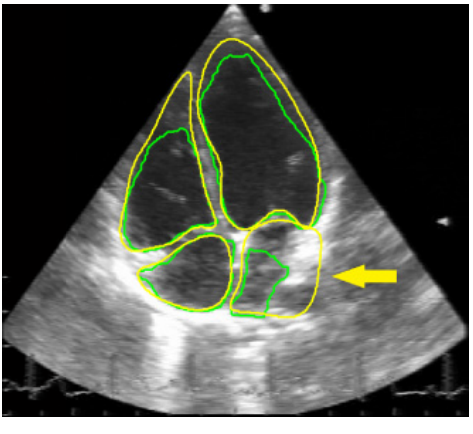


Fig. 4. Contour Outlier SR for Algorithm vs. Physician (Alg_P1)

From figure 3 it is observed the presence of a considerable number of outliers. In figure 4 it is depicted the worst outlier (Alg_P1) using the SR metric. The corresponding computed value was 0.46 dealing with the evaluation of the cavity identified by

the arrow. This worst case is a result of the fact that the boundaries are not clearly defined for this cavity, due to the presence of an artifact. That motivated the physician to draw the cavity region according to his knowledge and experience.

3.2 Single Cavity Analysis

A single cavity comparative study was also accomplished. Figure 5 shows the results related to the Left Ventricle (LV) using the Hausdorf Distance. Again, two comparative approaches were considered: Algorithm vs. Physician (first 4 columns) so called as set A and Physicians vs. Physicians (last 6 columns) referred as set B.

The calculated median values in set A are 17, 17, 13, 21 while in set B those values are 19, 19, 20, 21, 20, 17. These results lead us to conclude that the extracted contours by the algorithm are closer to the physician´s ones than what is verified when the performance of an expert is compared with the other experts. Comparing the 25th percentile in the interquartile range for the set A (values: 14, 13, 17, 9) and the set B (values: 15, 16, 16, 19, 16, 14) it is observed that set A has lower values meaning lower error, then better performance than it is provided by of a physician, using as reference the others physicians.

Similar conclusions are obtained comparing the 75th Percentile of set A (values: 21, 20, 25, 15) and set B (values: 21, 22, 27, 27, 24, 25). From the analysis, it can be concluded that the interquartile ranges for Algorithm vs. Physician present normally smaller dispersions than it is verified for Physician vs. Physician.

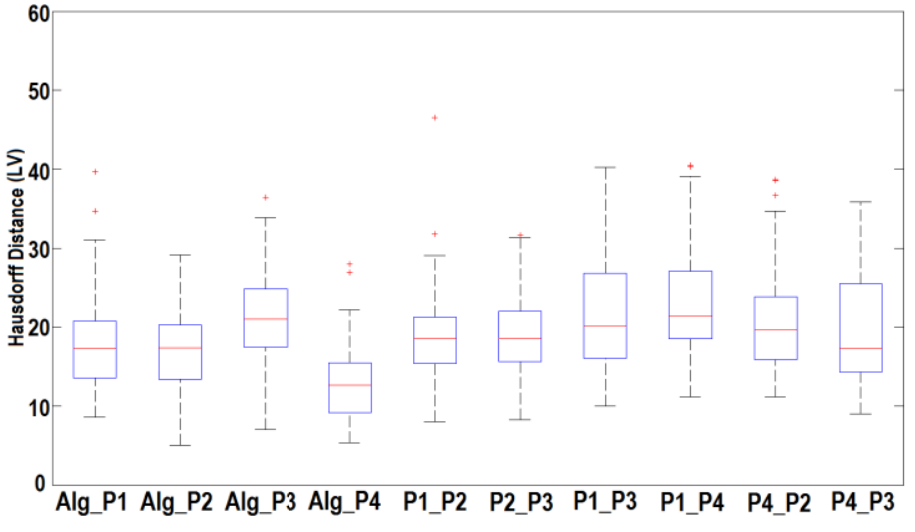


Fig. 5. Hausdorff Distance Box-Plot Graphic for Left Ventricle

The evaluation of the other heart cavities has led to similar results. Also, the same study was carried out for the metrics, Accuracy, Overlap, Sensitivity and Specificity, and the results agree with the ones produced by the Hausdorff metric.

4 Conclusions

Cardiac boundary extraction from echocardiographic images remains an important clinical challenge. We have proposed an improved algorithm for automatic contour extraction on apical long-axis four chamber view sequences of echocardiograms.

We tested the algorithm on a representative clinical dataset composed by two hundred and forty cavities and compared the results to contours manually sketched by four clinicians. That analysis was carried out by using some figures of merit, where the performances algorithm vs. physician and physician vs. physician were of concern.

It was observed that the algorithm provides results that are comparable to the inter-observer variability when the four cavities contour extraction are analyzed. Also, the algorithm performance is good when a single cavity is tested for algorithm vs. physician behavior, and presents superior results to the physician vs. physician achievements.

The positive results obtained with the present work, motivated the authors to proceed the development of new approaches for the simultaneous segmentation of the four cardiac chambers, thus assisting image-guided interventions and helping the experts in the clinical diagnosis, namely in the detection of congenital heart diseases.

References

1. Szabo, T.L.: Diagnostic Ultrasound Imaging: Inside out. Elsevier Academic Press (2004)
2. Noble, J.A.: Ultrasound image segmentation and tissue characterization. Institution of Mechanical Engineers, part H: Journal of Engineering in Medicine 224, 307–316 (2010)
3. Binder, T.: Three-Dimensional Echocardiography - Principles and Promises. Journal of Clinical and Basic Cardiology 5, 149–152 (2002)
4. Lang, R.M., Mor-Avi, V., Dent, J.M., et al.: Three-Dimensional Echocardiography: Is it Ready for Everyday Clinical Use? Journal of the American College of Cardiology 2, 114–117 (2009)
5. Suri, J.S., Setarehdan, S.K., Singh, S.: Advanced algorithmic approaches to medical image segmentation: state-of-the-art application in cardiology, neurology, mammography and pathology. Springer, New York (2002)
6. Jacob, G., Noble, J.A., Behrenbruch, C., et al.: A shape-space-based approach to tracking myocardial borders and quantifying regional left-ventricular function applied in echocardiography. IEEE Transactions on Medical Imaging 21, 226–238 (2002)
7. Noble, J.A., Boukerroui, D.: Ultrasound Image Segmentation: A Survey. IEEE Transactions on Medical Imaging 25, 987–1010 (2006)
8. Bosh, J.G., Mitchell, S.C., Lelieveldt, B.P.F., et al.: Automatic Segmentation of Echocardiographic Sequences by Active Appearance Motion Models. IEEE Transations on Medical Imaging 21, 1374–1383 (2002)
9. Antunes, S.G., Silva, J.S., Santos, J.B.: A New Level Set Based Segmentation Method for the Four Cardiac Chambers. In: V Iberian Conference on Information Systems and Technologies, Santiago de Compostela, Spain, vol. 1, pp. 173–178 (2010)
10. Antunes, S.G., Silva, J.S., Santos, J.B.: A Level Set Segmentation Method of the Four Heart Cavities in Pediatric Ultrasound Images. In: Campilho, A., Kamel, M. (eds.) ICIAR 2010, Part II. LNCS, vol. 6112, pp. 99–107. Springer, Heidelberg (2010)
11. Santos, J.B., Celorico, D., Varandas, J., et al.: Medical interface for echographic free-hand images. International Journal for Computational Vision and Biomechanics, 33–39 (2010)

12. Silva, J.S., Santos, B.S., Silva, A., et al.: A Level-Set Based Volumetric CT Segmentation Technique: A Case Study with Pulmonary Air Bubbles. In: Campilho, A.C., Kamel, M.S. (eds.) ICIAR 2004. LNCS, vol. 3212, pp. 68–75. Springer, Heidelberg (2004)
13. Silva, J.S., Silva, A., Santos, B.S., et al.: Detection and 3D representation of pulmonary air bubbles in HRCT volumes. In: SPIE Medical Imaging 2003: Physiology and Function: Methods, Systems, and Applications, USA, vol. 5031, pp. 430–439 (2003)
14. Silva, J.S., Cancela, J., Teixeira, L.: Intra-Patient Registration Methods for Thoracic CT Exams. In: Second International Conference on Bio-inspired System and Signal Processing, Porto, Portugal, pp. 285–290 (2009)
15. Silva, J.S., Silva, A., Santos, B.S.: Image denoising methods for tumor discrimination in high resolution computed tomography. Journal of Digital Imaging 24, 464–469 (2011)
16. Silva, J.S., Cancela, J., Teixeira, L.: Fast Volumetric Registration Method for Tumor Follow-Up in Pulmonary CT Exams. Journal of Applied Clinical Medical Physics 12, 362–375 (2011)
17. Cancela, J., Silva, J.S., Teixeira, L.: Fast Intra-Patient 3D Registration Method for Pulmonary CT Exams. In: 3rd Iberian Conference in Systems and Information Technologies, Vigo, Spain, vol. 1, pp. 539–543 (2008)
18. Silva, J.S., Silva, A., Santos, B.S.: A volumetric pulmonary CT segmentation method with applications in emphysema assessment. In: SPIE Medical Imaging 2006: Physiology, Function, and Structure from Medical Images, vol. 6143, pp. 885–896 (2006)
19. Ferreira, A., Morgado, A.M., Silva, J.S.: Automatic corneal nerves recognition for earlier diagnosis and follow-up of diabetic neuropathy. In: Campilho, A., Kamel, M. (eds.) ICIAR 2010, Part II. LNCS, vol. 6112, pp. 60–69. Springer, Heidelberg (2010)
20. Vasconcelos, V., Silva, J.S., Marques, L., et al.: Statistical Textural Features for Classification of Lung Emphysema in CT Images: A comparative study. In: V Iberian Conference on Information Systems and Technologies, Santiago de Compostela, Spain, vol. 1, pp. 496–500 (2010)
21. Ferreira, C., Santos, B.S., Silva, J.S., et al.: Comparison of a Segmentation Algorithm to Six Expert Imagiologists in Detecting Pulmonary Contours on X-ray CT Images. In: SPIE Medical Imaging 2003: Image Perception, Observer Performance and Technology Assessment, vol. 5034, pp. 347–358 (2003)
22. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. International Journal of Computer Vision 1, 321–331 (1988)
23. Osher, S., Sethian, J.A.: Fronts Propagation with Curvature Dependent Speed: Algorithms Based on Hamilton-Jacobi Formulations. Journal of Computational Physics 79, 12–49 (1988)
24. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic Active Contours. International Journal of Computer Vision 22, 61–79 (1997)
25. Malladi, R., Sethian, J.A., Vemuri, B.C.: Shape Modeling with Front Propagation: A Level Set Approach. IEEE Transactions on Pattern Analysis and Machine Intelligence 17, 158–175 (1995)
26. Paragios, N., Deriche, R.: Geodesic active regions: A new framework to deal with frame partition problems in computer vision. Journal of Visual Communication and Image Representation 13, 249–268 (2002)
27. Zhang, Y., Matuszewski, B.J., Shark, L.-K., et al.: Medical Image Segmentation Using New Hybrid Level-Set Method. In: Fifth International Conference Biomedical Visualization: Information Visualization in Medical and Biomedical Informatics, London, UK, pp. 71–76 (2008)

28. Chan, T.F., Vese, L.A.: Active Contours Without Edges. IEEE Transactions on Image Processing 10, 266–277 (2001)
29. Antunes, S.G., Silva, J.S., Santos, J.B., et al.: Phase Symmetry Approach Applied to Children Heart Chambers Segmentation: A Comparative Study. IEEE Transactions on Biomedical Engineering 58, 2264–2271 (2011)
30. Santos, B.S., Ferreira, C., Silva, J.S., et al.: Quantitative Evaluation of a Pulmonary Contour Segmentation Algorithm in X-ray Computed Tomography Images. Academic Radiology 11, 868–878 (2004)
31. Silva, A., Silva, J.S., Santos, B.S., et al.: Fast Pulmonary Contour Extraction in X-ray CT Images: A Methodology and Quality Assessment. In: SPIE - Medical Imaging: Physiology and Function from Multidimensional Images, USA, vol. 4321, pp. 216–224 (2001)
32. Lopez, C.A., Fernandez, M.M., Alzola, J.R.: Comments on: A Methodology for Evaluation of Boundary Detection Algorithms on Medical Images. IEEE Transactions on Medical Imaging 23, 658–660 (2004)
33. Chalana, V., Kim, Y.: A Methodology for Evaluation of Boundary Detection Algorithms on Medical Images. IEEE Transactions on Pattern Analysis and Machine Intelligence 16, 642–652 (1997)
34. Byrd, K.A., Zeng, J., Chouikha, M.: A Validation model for segmentation algorithms of digital mammography images. Journal of Applied Science & Engineering Technology 1, 41–50 (2007)