The Breast 32 (2017) 44–52

Contents lists available at ScienceDirect

The Breast

journal homepage: www.elsevier.com/brst

Original article

Atlas-based segmentation in breast cancer radiotherapy: Evaluation of specific and generic-purpose atlases

Delia Ciardo [a], Marianna Alessandra Gerardi [a], Sabrina Vigorito [b], Anna Morra [a], Veronica Dell'acqua [a], Federico Javier Diaz [c], Federica Cattani [b], Paolo Zaffino [d], Rosalinda Ricotti [a], Maria Francesca Spadea [d], Marco Riboldi [e, f], Roberto Orecchia [g, h], Guido Baroni [e, f], Maria Cristina Leonardi [a, *, 1], Barbara Alicja Jereczek-Fossa [a, h, 1]

[a] *Department of Radiation Oncology, European Institute of Oncology, Milan, Italy*
[b] *Unit of Medical Physics, European Institute of Oncology, Milan, Italy*
[c] *Medical Radiation Oncology, Mevaterapia, Buenos Aires, Argentina*
[d] *Department of Experimental and Clinical Medicine, Magna Graecia University, Catanzaro, Italy*
[e] *Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy*
[f] *Bioengineering Unit, National Center for Oncologic Hadrontherapy, CNAO Foundation, Pavia, Italy*
[g] *Scientific Direction, European Institute of Oncology, Milan, Italy*
[h] *Department of Oncology and Hemato-oncology, University of Milan, Milan, Italy*

A R T I C L E I N F O

Article history:
Received 7 July 2016
Received in revised form
21 November 2016
Accepted 18 December 2016

Keywords:
Atlas-based segmentation
Breast cancer radiotherapy
Automatic contouring
STAPLE contours

A B S T R A C T

Objectives: Atlas-based automatic segmentation (ABAS) addresses the challenges of accuracy and reliability in manual segmentation. We aim to evaluate the contribution of specific-purpose in ABAS of breast cancer (BC) patients with respect to generic-purpose libraries.

Materials and methods: One generic-purpose and 9 specific-purpose libraries, stratified according to type of surgery and size of thorax circumference, were obtained from the computed tomography of 200 BC patients. Keywords about contralateral breast volume and presence of breast expander/prostheses were recorded. ABAS was validated on 47 independent patients, considering manual segmentation from scratch as reference. Five ABAS datasets were obtained, testing single-ABAS and multi-ABAS with simultaneous truth and performance level estimation (STAPLE). Center of mass distance (CMD), average Hausdorff distance (AHD) and Dice similarity coefficient (DSC) between corresponding ABAS and manual structures were evaluated and statistically significant differences between different surgeries, structures and ABAS strategies were investigated.

Results: Statistically significant differences between patients who underwent different surgery were found, with superior results for conservative-surgery group, and between different structures were observed: ABAS of heart, lungs, kidneys and liver was satisfactory (median values: CMD<2 mm, DSC≥0.80, AHD<1.5 mm), whereas chest wall, breast and spinal cord obtained moderate performance (median values: 2 mm ≤ CMD<5 mm, 0.60 ≤ DSC<0.80, 1.5 mm ≤ AHD<4 mm) and esophagus, stomach, brachial plexus and supraclavicular nodes obtained poor performance (median CMD≥5 mm, DSC<0.60, AHD≥4 mm). The application of STAPLE algorithm generally yields higher performance and the use of keywords improves results for breast ABAS.

Conclusion: The homogeneity in the selection of atlases based on multiple anatomical and clinical features and the use of specific-purpose libraries can improve ABAS performance with respect to generic-purpose libraries.

1. Introduction

Over the last years, the combined improvements in diagnostic imaging, segmentation techniques, dose calculation, dose delivery

* Corresponding author. Department of Radiation Oncology, European Institute of Oncology, Via Ripamonti 435, 20141, Milan, Italy.
E-mail address: cristina.leonardi@ieo.it (M.C. Leonardi).
[1] Co-last author.

http://dx.doi.org/10.1016/j.breast.2016.12.010
0960-9776/

Abbreviations

RT	radiation therapy
ABAS	atlas-based automatic segmentation
BC	breast cancer
CTV	clinical target volume
CT	computed tomography
SCV	supraclavicular nodes
OAR	organs at risk
VF	vector field
R_{ES}	radius of the equivalent sphere
Vol_{CB}	volume of the contralateral breast
STAPLE	simultaneous truth and performance level estimation
CMD	center of mass distance
AHD	average Hausdorff distance
DSC	Dice similarity coefficient

and quality assurance have increased the accuracy of radiation therapy (RT) [1]. Such developments usually imply increasingly complex and therefore time-consuming processes and an intensified workload for medical and non-medical staff.

As the goal of RT is to irradiate the tumor volume avoiding neighboring organs to prevent acute and late toxicity, one of the key points in the treatment planning process is the segmentation, namely the process of labeling image voxels with anatomical and biological meaningful labels [2]. In particular, the widespread practice of more conformal irradiation techniques, which maximize normal tissue sparing and improve cosmetic outcomes, requires a more accurate and time consuming segmentation of contours [3,4]. Moreover, a large number of organs at risk should be considered to take into account the low-dose bath in intensity-modulated RT and the related – still investigational – late oncogenetic effect [5,6]. Still, manual contouring is a slow time-consuming process, prone to errors and inter- and intra-operator variability, which is difficult to quantify [7–13], and this influences negatively the reliability of dose distributions and therapeutic outcomes comparison between different studies and institutions.

The recent introduction of fully or semi-automatic atlas-based segmentation techniques aims to address the challenges of accuracy and reliability of contouring.

The general impression from previous studies reporting the efficacy of atlas-based automatic segmentation (ABAS) in breast cancer (BC) patients is that the homogeneity among subjects and contours included in the library strongly influences the results of automatic contouring procedure [14]. In this frame, the aim of this study was to evaluate whether specific-purpose libraries featuring a large number of atlases and relying on homogeneity classification based on multiple selected anatomical and clinical features can improve ABAS performance in the segmentation of BC patient with respect to generic-purpose libraries. Performance of single-ABAS and multi-ABAS are also compared.

2. Materials and methods

2.1. Atlas based automatic segmentation

2.1.1. Theoretical overview

In general, an atlas is a model image segmented by an expert operator. In ABAS, a "library" of atlases is collected: the first step consists of selecting a template atlas to serve as reference for the non-rigid registration of all the remaining atlases populating the library, thus obtaining a dataset of deformation vector fields ($VF_{atlas} = \{V_1, V_2, \ldots, V_{n-1}\}$, where n is the number of atlases). When a new unlabeled computed tomography (CT) is given, it is registered to the template atlas and the resulting vector field (VF_{new}) is compared to those saved in the dataset VF_{atlas}. The atlas corresponding to the vector field V_i most similar to VF_{new} is identified as the best matching atlas and its contours are therefore propagated onto the unlabeled CT scan. With other words, the similarity between the unlabeled image with all those available as atlases (or with a specific subset of the library) is calculated and deployed in order to assign a proper label to the voxels of the unlabeled image. In single-ABAS, the atlas that maximizes the similarity index is non-rigidly registered with the new image, and the contours are propagated according to the deformation vector field resulting from the registration. In multi-ABAS, the information from more than one atlas is somehow combined to generate the automatic segmentation. One possible strategy to combine information from different atlases is the Simultaneous Truth and Performance Level Estimation (STAPLE) algorithm [15], which is an expected maximization algorithm that computes a probabilistic estimate of the true segmentation by weighting each segmentation on its estimated performance level. This method is often used as a reference standard segmentation for assessing performance of different algorithms and when an improved segmentation is needed.

A manual refinement is usually required, but with much less efforts with respect to a complete manual contouring from scratch on each CT scan [2].

2.1.2. Description of the implemented ABAS strategies

The creation of the libraries was performed by means of the commercial software suite MIM 6.1.7 (MIMvista Corp., Cleveland, US-OH). One generic-purpose atlas and 9 specific-purpose sub-libraries were created.

At the building of sub-libraries, two main subgroups were stratified as function of the type of surgery, namely post-conservative surgery BC patients (hereinafter referred as "conservative-surgery" group) versus post-mastectomy BC patients (hereinafter referred as "non-conservative-surgery" group). For non-conservative-surgery patients, the side of the tumor was also considered for patient stratification. In order to describe specific-purpose atlases, some anatomical features were selected. Since the most common indicators of body size were not available in our clinical dataset, a surrogate index was derived from thoracic circumference. This was obtained as the radius of the sphere equivalent to the volume of the axial slice at the sub-mammary fold level (R_{ES}). R_{ES} thresholds were identified as the 33rd and 66th percentiles of R_{ES} distribution, corresponding to 6.0 and 6.5 cm, respectively. Three sub-groups were therefore obtained by discriminating between *small* ($R_{ES} \leq 6.0$ cm), *medium* (6.0 cm < R_{ES}<6.5 cm), and *large* size ($R_{ES} \geq 6.5$ cm).

A further stratification was obtained by keywords describing the contra-lateral breast volume (Vol_{CB}) as "*small breast*" (S) and "*large breast*" (L), if it is inferior or superior to the median value of Vol_{CB} distribution (corresponding to 506 cm^3), respectively. Additional categorizing keywords were related to the presence of breast prosthesis (P) or expander (E). Therefore, there are 6 possible combinations of keywords (S, SP, SE, L, LP, LE) that the operator can chose when contouring the CT of a new patient.

2.2. Patients dataset

We collected 200 CT scans of BC patients, treated with adjuvant RT at the European Institute of Oncology (Milan, Italy) between January 2012 and December 2013. All patients gave written