Artificial Intelligence In Medicine 90 (2018) 34–41

Contents lists available at ScienceDirect

Artificial Intelligence In Medicine

journal homepage: www.elsevier.com/locate/artmed

Abdominal, multi-organ, auto-contouring method for online adaptive magnetic resonance guided radiotherapy: An intelligent, multi-level fusion approach☆

Fan Liang[a,b,c], Pengjiang Qian[d], Kuan-Hao Su[a,b], Atallah Baydoun[e,f,g], Asha Leisser[a,b,h], Steven Van Hedent[a,b,i], Jung-Wen Kuo[a,b], Kaifa Zhao[d], Parag Parikh[j], Yonggang Lu[j], Bryan J. Traughber[b,k,l], Raymond F. Muzic Jr[a,b,g,m,*]

[a] Department of Radiology, Case Western Reserve University School of Medicine, Cleveland, OH, USA
[b] Case Center for Imaging Research, University Hospitals Case Medical Center, Case Western Reserve University, Cleveland, OH, USA
[c] Tianjin Key Laboratory of Information Sensing & Intelligent Control, Tianjin University of Technology and Education, Tianjin, China
[d] School of Digital Media, Jiangnan University, Wuxi, Jiangsu, China
[e] Department of Internal Medicine, Case Western Reserve University School of Medicine, Cleveland, OH, USA
[f] Department of Internal Medicine, Louis Stokes VA Medical Center, Cleveland, OH, USA
[g] Department of Biomedical Engineering, Case Western Reserve University, Cleveland, OH, USA
[h] Department of Biomedical Imaging and Image-guided Therapy, Medical University of Vienna, Vienna, Austria
[i] Department of Radiology, UZ Brussel (VUB), Brussels, Belgium
[j] Department of Radiation Oncology, Washington University School of Medicine, St. Louis, MO, USA
[k] Department of Radiation Oncology, Case Western Reserve University School of Medicine, Cleveland, OH, USA
[l] Department of Radiation Oncology, University Hospitals Seidman Cancer Center, Cleveland, OH, USA
[m] Department of Radiology, University Hospitals Cleveland Medical Center, Cleveland, OH, USA

ARTICLE INFO

Keywords:
Auto-Contouring
Machine learning
Adaptive radiotherapy
Image-guided
Radiotherapy

ABSTRACT

Background: Manual contouring remains the most laborious task in radiation therapy planning and is a major barrier to implementing routine Magnetic Resonance Imaging (MRI) Guided Adaptive Radiation Therapy (MR-ART). To address this, we propose a new artificial intelligence-based, auto-contouring method for abdominal MR-ART modeled after human brain cognition for manual contouring.

Methods/Materials: Our algorithm is based on two types of information flow, i.e. top-down and bottom-up. Top-down information is derived from simulation MR images. It grossly delineates the object based on its high-level information class by transferring the initial planning contours onto daily images. Bottom-up information is derived from pixel data by a supervised, self-adaptive, active learning based support vector machine. It uses low-level pixel features, such as intensity and location, to distinguish each target boundary from the background. The final result is obtained by fusing top-down and bottom-up outputs in a unified framework through artificial intelligence fusion. For evaluation, we used a dataset of four patients with locally advanced pancreatic cancer treated with MR-ART using a clinical system (MRIdian, Viewray, Oakwood Village, OH, USA). Each set included the simulation MRI and onboard T1 MRI corresponding to a randomly selected treatment session. Each MRI had 144 axial slices of 266 × 266 pixels. Using the Dice Similarity Index (DSI) and the Hausdorff Distance Index (HDI), we compared the manual and automated contours for the liver, left and right kidneys, and the spinal cord.

Results: The average auto-segmentation time was two minutes per set. Visually, the automatic and manual contours were similar. Fused results achieved better accuracy than either the bottom-up or top-down method alone. The DSI values were above 0.86. The spinal canal contours yielded a low HDI value.

Conclusion: With a DSI significantly higher than the usually reported 0.7, our novel algorithm yields a high

☆ Meeting presentation: A preliminary report of this research was submitted as an oral presentation for the Radiological Society of North America (RSNA) 2017 annual meeting.

* Corresponding author at: Radiology, Mail Stop: BSH: 5056, University Hospitals Cleveland Medical Center, 11100 Euclid Avenue, Cleveland, OH, 44106, USA.
E-mail addresses: bachelormd10@163.com (F. Liang), qianpjiang@jiangnan.edu.cn (P. Qian), kuan-hao.su@case.edu (K.-H. Su), atallah.baydoun@case.edu (A. Baydoun), asha.leisser@meduniwien.ac.at (A. Leisser), steven.vanhedent@case.edu (S. Van Hedent), jung-wen.kuo@case.edu (J.-W. Kuo), zhaokaifa@qq.com (K. Zhao), pparikh@radonc.wustl.edu (P. Parikh), yonggang.lu@wustl.edu (Y. Lu), bryan.traughber@case.edu (B.J. Traughber), raymond.muzic@case.edu (R.F. Muzic).

https://doi.org/10.1016/j.artmed.2018.07.001
Received 16 June 2017; Received in revised form 6 June 2018; Accepted 6 July 2018

segmentation accuracy. To our knowledge, this is the first fully automated contouring approach using T1 MRI images for adaptive radiotherapy.

1. Introduction

Adaptive radiation therapy (ART) is a radiation treatment process that modifies the treatment plan based on patient-specific functional and anatomic changes during the course of radiation therapy. ART accounts for the daily geometric variation in patient's anatomy resulting from patient's setup, physiological changes, and treatment response. ART may enhance the efficacy via enabling high treatment dose to clinical target volumes (CTVs) while minimizing morbidities by sparing organs at risk (OARs) [1]. There are three timescales for ART, i.e. offline between treatments, online immediately prior to a treatment, and in real time during treatment [2]. During online ART, a plan is generated for each radiation therapy session based on the image data acquired immediately prior to the session while the patient is on the treatment table. On-board imaging is thus required for online ART and with advanced technology, online, MRI-guided, adaptive radiation therapy (MR-ART) has been implemented in clinical and research settings [3,4]. Compared to computed tomography-guided ART (CT-ART), MR-ART provides two, main advantages, i.e. better soft tissue contrast leading to more accurate anatomical delineation [5,6] and further sparing of OARs.

Manual contouring, the delineation of CTVs and OARs, remains the most laborious and time-consuming task in traditional radiation therapy and is a significant hindrance to ART. Moreover, manual contouring is subject to intra- and inter-operator variation. These two features of manual contouring make it impractical for routine clinical use in online MR-ART and very difficult in the research setting. In addition, the waiting time on a treatment table is usually uncomfortable for many cancer patients who have significant morbidities. Recently published case series reported the successful implementation of MR-ART for different abdominopelvic malignancies, using a manual [1] or interactive segmentation algorithm [7]. Different image registration techniques do not meet the clinical needs for adaptive radiotherapy and thus require further manual adjustment of the contours to account for the daily variations. Therefore, an optimal auto-contouring algorithm for online MR-ART is needed in order to efficiently generate new contours each treatment day, and thus minimizing the time gap between image acquisition and the implementation of the corresponding radiation therapy plan.

Currently available, auto-contouring algorithms can generally be classified into two approaches, i.e. model-free and knowledge-based methods. Model-free methods are based on the analysis of image content and properties such as voxel intensities or gradient analysis. These methods include graph cuts [8], watershed [9], and adaptive thresholding [10] as well as region growing [11]. Model-free methods may roughly identify the organ boundaries, but are sensitive to image noise and artifacts [12]. Knowledge-based methods address these limitations using prior knowledge of the morphology of anatomical structures or the appearance of organs in different types of imaging modalities in order to improve the robustness and accuracy of the boundary determination. The knowledge-based methods include three sub-categories, i.e. atlas-, model-, and machine learning-based methods [13]. Atlas-based segmentation is defined as the process of performing segmentation on a new dataset using the knowledge of prior segmentation, i.e. a dataset from one or more patients that has the structures of interest already labeled [14]]. Many commercially available software such as ABAS (CMS-Elekta, Stockholm, Sweden), MIM (MIMVista corp, Cleveland, OH), and VelocityAI (Velocity Medical Systems, Atlanta, Georgia) use atlas-based segmentation in adaptive radiotherapy when based on one re-planning CT acquired during the radiotherapy course [15]. Disadvantages of this method include the large variation in intensity, contrast medium, and geometry between the atlas and the image to be segmented which is also subject to significant image artifacts and inter-patient differences [16]. Model-based organ delineation includes prior knowledge of the organ shape combined with organ-specific parameters, such as intensity range, gradient magnitude, and direction [17]. However, this method requires two, manual steps, i.e. close initialization to the target anatomy with its corresponding model shape according to a drag-and-drop operation and non-rigid manual deformations [16]. Alternatives to overcoming these undesired properties include machine learning methods using voxel-based features to train, for example, neural networks, boosting trees, the support vector machine (SVM), random forests models or ensembles of classifiers [13] in order to find tissue boundaries in new data based on the inferred function determined from the labeled training data. Unfortunately, these algorithms are generally computationally complex and time-consuming and impractical for online and real time MR-ART [18].

Inspired by the human contouring processing model, we developed a novel methodology for auto-contouring that is designed to support online MR-ART. We couple top-down and bottom-up methods to achieve a multiple-level, hybrid method integrating low-level features to object level models in addition to simulating the human fusion

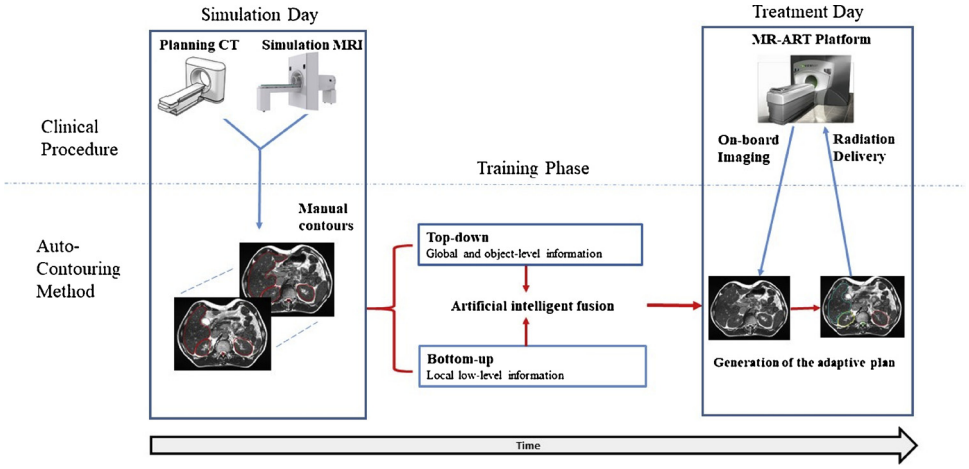

Fig. 1. Schematic representation of the auto-contouring algorithm.

process. To our knowledge, no such strategy in auto-contouring has been previously considered. We report the initial results in four abdominal organs and note that our method can be generalized for other organs. The outline of the algorithm along the clinical procedure timeline is presented in Fig. 1.

2. Methods/Materials

2.1. MR-ART clinical procedure

Before initiation of radiation therapy treatment, each patient undergoes simulation computed tomography (CT) in accordance with standard clinical practice. In addition, a simulation MRI is acquired using the MR-ART platform. CT and MRI images are then transferred to a third-party software (Pinnacle, version 9.0, Philips Medical Systems, Madison, WI) for rigid registration and contouring by the treating radiation oncologist. Simulation CT is selected as the primary image for manual contouring of the CTV and OARs. MRI may be used if needed for clear visualization of the tumor volume. Contours drawn on simulation CT image are transferred afterwards to the simulation MRI. Manual editing of these contours will be done, as needed, by the treating radiation oncologist and used in the rest of the clinical procedure. Online MR-ART is then delivered after acquiring a daily MRI before each treatment while the patient is on the treatment table. Contours are transferred to the daily MRI after deformable registration. One or more physicians, sometimes using a “divide-and-conquer” approach, compare the on-board, treatment-day MRI to the simulation images and then attempt to rapidly and accurately adjust the CTV and OARs while the patient remains on the table. A daily optimized plan is then generated using new contours while keeping the original CTV dose and OAR constraints. The MRI sequence used on the simulation and treatment day is the trueFISP pulse sequence. T1-weighted images are generated using a TR of 2 ms (ms) and a TE of 0.86 ms.

2.2. Computational framework

Our method is modeled after the human information processing and decision-making for contouring. Theoretically, there are three, basic concepts that underlie human cognition, i.e. granulation, organization, and causing [19]. Granulation involves decomposition of the whole into parts, organization involves integration of parts into the whole, and causing involves the association of causes with effects [19]. Based on these concepts, we can postulate that the human brain would proceed by granulation to decompose an image into different organs and by organization to relate each pixel to a specific organ. In most cases, human reasoning in manual contouring is guided by prior knowledge and contours are conceptualized in an indistinct sense.

In order to understand the image context and adjust the boundaries in complicated clinical scenarios, the software agent will be provided with prior knowledge regarding the anatomy and appearance of the organs. With this method two sets of information are integrated, i.e. granulation-related top-down and organization-related bottom-up. The top-down information grossly delineates the object boundary based on high-level information of the object class obtained from rendering the initial contours to the daily image. The bottom-up information labels each image pixel as object or background using low-level information obtained from learning the relationship between pixel values and local textures according to the manually contoured simulation MRI. As detailed below, the information is then fused in a unified framework using an artificial intelligence mechanism and toward a more refined segmentation, while simultaneously estimating the pixel-level labeling and the organ boundary localization.

2.2.1. Top-down pathway

In this part, the simulation-day contours are used as a non-parametric organ model reference for daily plan adjustment. The geometric boundaries of relatively rigid structures, such as the spinal canal, will be registered for time-saving purposes. In order to reinforce the registration results, two, feature descriptors are used, i.e. blob-type, Speeded Up, Robust Features (SURF) [20] and corner-type, Binary, Robust, Invariant, Scalable key points (BRISK) features [21]. While the local structural information has a more significant shape context in the image, multiple pairs are useful when good matches cannot be readily obtained between simulation and treatment day image sets using a

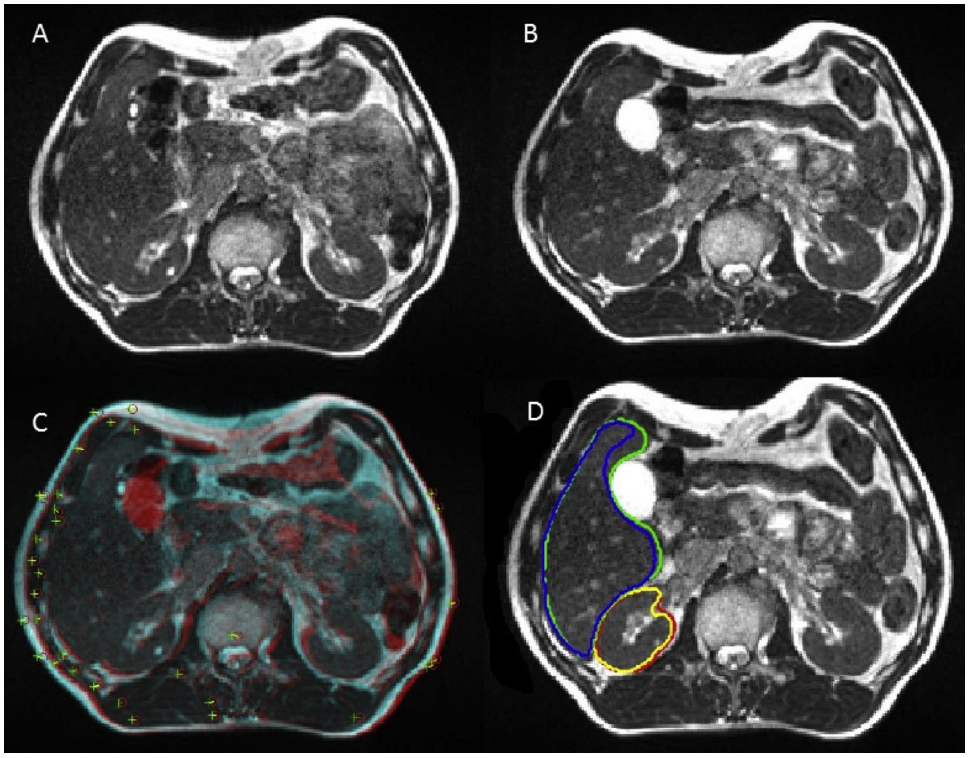

Fig. 2. Feature-based registration steps. Axial MR images from (A) simulation, (B): on-board, treatment-day, (C) extracted matching features on image B, and (D) registered image using the RANSAC algorithm. Legend: + Feature points on the Simulation MRI; o Feature points on the MRI treatment day; Blue and Yellow contours: Automatic contours; Green and Red contours: Ground truth. (For interpretation of the references to colour in this figure legend and text, the reader is referred to the web version of this article).

single-feature detector.

The feature-based registration between the contoured simulation and treatment-day images is done using the following steps: (1) Detect and extract the matching features between two images; (2) Find matched features and localize them in individual images; (3) Use the RANSAC algorithm [22] to estimate the transformation matrix to co-register image pairs; and (4) Transform the images. The registration steps are illustrated in Fig. 2. In Fig. 2(D), the blue contour of the liver and the yellow contour of the right kidney are the results of the automatic top-down method. The green contour of the liver and the yellow contour of the right kidney are manual contours defined by radiation oncologist or radiologist on the MR images and labeled as ground truth to be used as a reference evaluating the automatic contours.

2.2.2. Bottom-up pathway

In this section, we propose a machine learning algorithm for organ boundary detection. It uses self-adaptive active learning based classification. The voxelwise, data-driven classification method utilizes low-level information. A decision of a boundary point is made independently at each location in the image by automatic recursive selection of the correctly labeled class based on the output of the probabilistic support vector machine (SVM). The details of the proposed algorithm are listed in the following sections.

2.2.2.1. Problem formulation. Given an image set **I,** the medical image contouring task is to determine the boundary, **B,** of a specific component of interest, **C,** that can represent a tumor or a specific organ in each slice. i and j refers respectively to the row and column indices of the image matrix. We define the$S_{\mathbf{C}}^{k}(i,j)$ as the function associated with **C** that gives 1 if the image voxel (i, j) located in the k^{th} slice belongs to **C** and gives 0 if not.

2.2.2.2. Extracting features from MRI. Differing from the top-down pathway in which high-level information is enlisted, we exclusively use low-level information for the bottom-up classification. Specifically, we used two types of features: the low-level image appearance features, which are of central voxel value, and the 3D location features, which are three–dimensional, spatial coordinates representing the individual anatomic location of each voxel. The voxel intensity feature provides fundamental, low-level contrast information and is widely used for image segmentation. Since intensity features are often insufficient for accurate boundary detection because surrounding tissue may have the same intensity, we added the spatial coordinate of each voxel to depict the anatomic locations of different target tissues. Finally, the low-level intensity and the location features of each voxel were combined into one feature vector for the classification.

2.2.2.3. Probabilistic SVM. Given a training dataset, $\mathbf{X} = \{\mathbf{x}_i \in \mathbf{R}^d,\ i = 1,\ ...,l\}$, the formulation of conventional SVM can be represented as:

$$\begin{aligned} &\min_{\boldsymbol{\alpha}\in\mathbf{R}^l, b\in\mathbf{R}, \xi_i\in\mathbf{R}} \left(\tfrac{1}{l}\textstyle\sum_{i=1}^{l}\xi_i + \gamma\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{K}\boldsymbol{\alpha}\right), \\ &\text{s.t. } y_i\left(\textstyle\sum_{j=1}^{l}\alpha^j K(\mathbf{x}_i,\mathbf{x}_j) + b\right) \geq 1-\xi_i, \quad i = 1,...,l, \\ &\quad \xi_i \geq 0, \quad i = 1,...,l, \end{aligned} \tag{1}$$

where $y_i \in \{1, -1\}$ is the label of $\mathbf{x}_i$, $\mathbf{K}$ is the $l \times l$Gram matrix with $K_{ij} = K(\mathbf{x}_i, \mathbf{x}_j)$ and $K(.,.)$ being a kernel function, and $\gamma > 0$ is the regularization coefficient.

Based on the Karush-Kuhn-Tucker (KKT) conditions [23], the dual form of (1) can be derived as:

$$\begin{aligned} &\max_{\boldsymbol{\beta}\in\mathbf{R}^l}\left(\tfrac{1}{l}\textstyle\sum_{i=1}^{l}\beta_i - \tfrac{1}{2}\boldsymbol{\beta}^{\mathrm{T}}\mathbf{Q}\boldsymbol{\beta}\right), \\ &\text{s.t. } \textstyle\sum_{i=1}^{l}\beta_i y_i = 0, \\ &\quad 0 \leq \beta_i \leq \tfrac{1}{l}, \quad i = 1,...,l, \end{aligned} \tag{2}$$

where $\boldsymbol{\beta} = [\beta_1,...,\beta_l]^{\mathrm{T}} \in \mathbf{R}^l$are the Lagrange multipliers, $\mathbf{Q} = \mathbf{Y}(\mathbf{K}/2\gamma)\mathbf{Y}$ $\mathbf{Y} = \mathrm{diag}(y_1,\ ...,y_l)$, and diag(.) signifies the generating function of the diagonal matrix.

Via the solution $\boldsymbol{\beta}^*$of (2), the original solution of (1) can be obtained:

$$\boldsymbol{\alpha}^* = \mathbf{Y}\boldsymbol{\beta}^*/2\gamma \tag{3}$$

$$b^* = \frac{1}{l}\sum_{i=1}^{l}\left(y_i - \sum_{j=1}^{l}\alpha_j^* K(\mathbf{x}_i, \mathbf{x}_j)\right) \tag{4}$$

Then the final classifying decision function of conventional SVM can be expressed as:

$$f(\mathbf{x}) = \sum_{i=1}^{l}\alpha_i^* K(\mathbf{x}, \mathbf{x}_i) + b^* \tag{5}$$

It is clear that the decision function of the standard SVM in the form of (5) outputs a specific value rather than a probability. Alternatively, probabilistic SVM attempts to generate a probability that indicates the confidence level of a data point belonging to a class. To this end, Platt scaling [24] is commonly used to give a probabilistic $\mathrm{P}(y = 1 \text{ or } -1 \mid x)$ output in terms of the specific decision value of the standard SVM, i.e. (5). Specifically, it produces the probability estimates:

$$\mathrm{P}(y = 1x) = \frac{1}{1 + \exp(Af(x) + B)} \tag{6}$$

in which f(x) is generated by (5) and A and B are two, scalar parameters estimated by minimizing the negative log likelihood of training data using their labels and decision values. Complete details of the mathematical estimation have been published elsewhere [24].

2.2.2.4. Self-Adaptive, active learning classification algorithm (SAALC). Active learning is the process with which a machine learning algorithm achieves higher accuracy with fewer training labels than traditional learning approaches via choosing representative examples from the training set in a recursive fashion to improve learning efficiency [25]. Conventionally, the active learning uses human input to validate the segmentation results. Instead, we introduce the SAALC algorithm which recursively trains the SVM classifier by gradually including example data points which cannot be definitively classified with regard to the outputs of the probabilistic SVM to reduce the uncertainty of the prediction. The fully automatic active learner attempts to achieve high accuracy using the probability output of SVM in the training stage, and thereby maximizing the prediction accuracy in the segmentation testing stage. Details are as follows.

For every voxel (i, j) in the k^{th} slice in the target MR image, with the nomenclature$p_{ij_q}^{k}$ denoting the probability of the voxel to be classified into the q^{th} category, we will assign d probabilities corresponding, respectively, to d categories by means of the probabilistic SVM. Here, d will be set to 2 for a two organ classifier, e.g., the spinal cord and background, and to 3 for the three organ classifier. The eventual affiliation of the voxel is determined by the maximum probability $p_{ij\_max}^{k} = \max(p_{ij\_1}^{k}, p_{ij\_2}^{k}, ...,p_{ij\_d}^{k})$.

Both the contoured simulation and treatment-day images are used in the SAALC algorithm. Specifically, the training process only uses the simulation images, the contours defined on the simulated images, and the treatment-day images. The treatment-day contours are not used for the training but instead are used for the final performance evaluation. We first extract the low-level information, i.e., the intensity and the location features, from the simulation and treatment-day images to constitute two corresponding MR feature data sets, signified as $\mathbf{I}_s$ and $\mathbf{I}_t$, respectively. Then, the algorithm proceeds stepwise:

1) Randomly select l=2000 examples from each class in $\mathbf{I}_s$ with labels, i.e., the voxel subset of any l = 2000 voxels from given organs or

background in the contoured simulation images, to constitute the initial training set **T**;
2) Based on the current training set **T**, use the method of probabilistic SVM to learn a classifier *Cr*;
3) For each example in $\mathbf{I}_t$, i.e., any voxel (i, j) in the k^{th} slice in the treatment-day image, compute the maximum probability $p^k_{ij\_max} = \max(p^k_{ij\_1}, p^k_{ij\_2}, ..., p^k_{ij\_d})$, via the classifier *Cr* in step 2), to ascertain which class the voxel most likely belongs to;
4) Compare the $p^k_{ij\_max}$ with the threshold τ and obtain the uncertain subset **U** from $\mathbf{I}_t$ in which all the calculated maximum probabilities are smaller than the threshold;
5) Use a 30% sampling ratio to sample the uncertain subset **U** and get n uncertain examples;
6) For each of the n uncertain examples in $\mathbf{I}_t$, use the nearest neighbor strategy to find a matching example from $\mathbf{I}_s$ excluding the ones in the training set **T**, and we thereby get n new significant examples for active learning;
7) Update the training set **T** by adding the just achieved n significant examples for active learning;
8) Repeat steps 2) – 7) until the desired prediction performance is obtained;
9) Based on the principle of maximum probability, predict the organ/tissue class of each voxel in $\mathbf{I}_t$ using the final probabilistic SVM classifiers obtained using the recursive active learning procedure.

The desired prediction performance in step 8) refers to the case in which the difference in accuracy prediction between two consecutive active learning iterations is less than 2%. The threshold $\tau$to find the uncertain subset **U** is optimized within the range of 0.55 to 0.65 with a step size of 0.02, using the grid search. Within this range, the value that yielded the high classification accuracy is retained.

2.2.3. Fusion process

The process of human reasoning integrates different types of information for segmentation of organs and tumors. According to that, we might infer that only the top-down knowledge is useful because it outputs a global view of multiple organs. However, the top-down knowledge is not able to capture all the details of the image data due to the inconsistency of the tumor and organ positions between the radiation treatment sessions. On the other hand, the bottom-up method is limited by the lack of object-level knowledge especially, for the multi-organ auto-contouring. In fact, the contours can be refined by adding the bottom-up approach using low-level cues to integrate learning information from the previously segmented simulation MRI images, and thus achieving accurate segmentation for successive treatment-day images of the same patient. In this context, we fuse both the top-down and bottom-up information in a way similar to the human reasoning according to the variation probability of each organ. When we compare the constancy of each organ location in the different MR images, the rigid structure with high contrast seems to present less of a challenge than that of the non-rigid structures. The osseous structure, spinal canal in this case, is the most consistent one. In contrast, soft-tissue organs such as liver may vary in size, shape, and position. As for the kidneys, the predicted respiratory motion effect can be compensated for by feature-based registration using the top-down method.

In the human understanding and reasoning system, knowledge could be represented by using verbally formulated rules to form a natural language expression. The general form is If-Then rules. Considering the concrete variation property of each organ, we specify the variable knowledge: spinal canal and kidneys are considered as rigid and liver is considered as deformable. For the deformable organ, we use the polar coordination to fuse top-down and bottom-up information. Before fusion, a five points moving average low-pass filtering was applied to the bottom-up contours. By forming a weighted-average of the pixels in a neighborhood orientation, high frequency noise was reduced and the contours were smoothened. We first compute the center of the mass of the each contour and designate it as the zero reference coordination. The radii are then selected and combined with regard to the certain angle regarding the coordination from the bottom-up method result (for example, 0 degrees to 90 degrees) and the remaining angle from the top-down method result (91 degrees to 359 degrees). Later, the fused radii set is transformed back into Cartesian coordination.

The if-then fusion rules for each organ boundary using the top-down method or from the bottom-up method or using the fused one can be described as follows: 1) If **B** is of a rigid organ, the top-down result is selected as the final result. 2) If **B** is of a deformable organ, then the fused one is selected as the final result.

2.3. Experiment

For evaluation purpose, we tested our auto-contouring methods on a dataset of four patients with non-resectable, abdominal malignancies and who were enrolled in a randomized, controlled, phase II study (NCT02950025) [26] evaluating MR-ART using a commercially available MR-ART clinical system (MRIdian, Viewray, Oakwood Village, OH, USA). Each dataset included the simulation MRI and on-board T1 MRI corresponding to a randomly selected treatment session. In accordance with procedures listed in an Institutional Review Board-approved protocol, data were retrospectively obtained and de-identified by the treating physician prior to inclusion in our dataset. Each MRI dataset was composed of 144 image slices and with a resolution of 266 × 266 pixels. The pixel spacing was 1.5 mm × 1.5 mm within the axial plane and 3 mm between the planes. Four organs were considered, i.e. liver, right kidney, left kidney, and spinal canal. The ground truth contours were manually segmented for the four organs and the skin in order to serve as the background.

The simulation dataset was used as a training set, and each daily dataset was used as a testing set. Three classifiers were used for each patient at each day treatment processing. In this experiment, five categories were defined: liver, left kidney, right kidney, spinal canal, and the background. Background was defined as the area limited by the skin contours excluding the four previously mentioned organ. We arranged the five classes into three packages for the SVM-based, supervised classification, i.e. (a) liver, right kidney, and skin, (b) left kidney and skin, (c) and the spinal canal and skin. The method was implemented using MATLAB 2016b (The MathWorks, Natick, MA, USA).

We compared the automatically- to the manually-generated contours using two, overlapping metrics, i.e. the Dice similarity index (DSI) and the Hausdorff distance index (HDI). The DSI has been used to validate segmentation algorithms in various clinical settings. It computes twice the fraction of the overlapping area between two contours divided by the union of the areas covered by the contours. A value of 1.0 indicates two contours being in exact agreement. However, the DSI is a global similarity score and its value is compromised when there is a significant local difference between segments with relatively small volume differences [27]. Therefore, we also calculated the HDI which correlates more accurately with the distance between the two contour boundaries by accounting for the maximum of all the distances from a point in the predicted contour to the closest point in the ground-truth contour [28].

3. Results and discussion

The algorithm starts the training session after the planning manual contoured images are obtained. In our study, the number of iteration was set to 3 according to our preliminary evaluation that showed the difference in accuracy prediction between the second and third active learning iterations would be less than 2%. Running on an Intel Xeon®E5-2643 3.40 GHz CPU, 64GB of memory, 64-bit Windows 7 operating system (Microsoft, Redmond, WA, USA), the typical computation time for one patient in the training stage, i.e. for training three

classifiers, was approximately two hours. Given trained classifiers, the average segmentation time of a new image during the daily treatment session was approximately two minutes.

Fig. 3 displays an axial, T1-weighted MRI with manual and automatically generated contours of the four organs. The manual contours in this image were drawn by radiation oncologist or radiologist and labeled as ground truth to be used as a reference evaluating the automatic contours. By visual inspection, the contours are similar to ones drawn by physicians. The average DSI and HDI are shown in Table 1. The fused, artificial intelligence method achieved better accuracy than the bottom-up and top-down methods alone. The bottom-up method had better DSI for the left kidney, while the top-down method had better DSI for the right kidney. Among rigid organs, the HDI was higher using the bottom-up method compared to the top-down method. The HDI for the liver was higher with the top-down method compared to the bottom-up method.

The online MR-ART platform combines a radiation delivery to an advanced MRI acquisition system. The implementation of this hardware technology for ART in a routine clinical workflow faces the limitation of a significantly time–consuming, manual contouring task. While manual segmentation is adequate for defining organs, it remains poor for yielding consistent, confident distinction of organ boundaries at some tissue interfaces, even among clinical experts [29]. Compared with an average of nine minutes with manual contouring [30], our algorithm represents a solution for the time and accuracy barriers and, therefore, contributes to the eventual consideration of the use of adaptive radiation therapy.

A DSI above 0.7 usually suggests a good overlap [23]. Except for the spinal canal, DSI values obtained using this method were above 0.82 and, therefore, our results demonstrate a high accuracy with a minimal computational time for both soft and rigid organs. In fact, the spinal canal shape consists of a longitudinal cylinder of small diameter. Owing to the way the DSI is defined, structures having a high surface area-to-volume ratio tend to have an inappropriately low DSI. For such structures, the overlap is better reflected by the HDI which is the lowest for the spinal canal. As the two kidneys have a similar anatomical shape, it was thus expected to have similar contouring scores. The score difference between the two kidneys can be explained, in part, by the local difference in the anatomical neighborhood and the relative proximity of the right kidney to the medial liver edges. Due to similar regional tissue

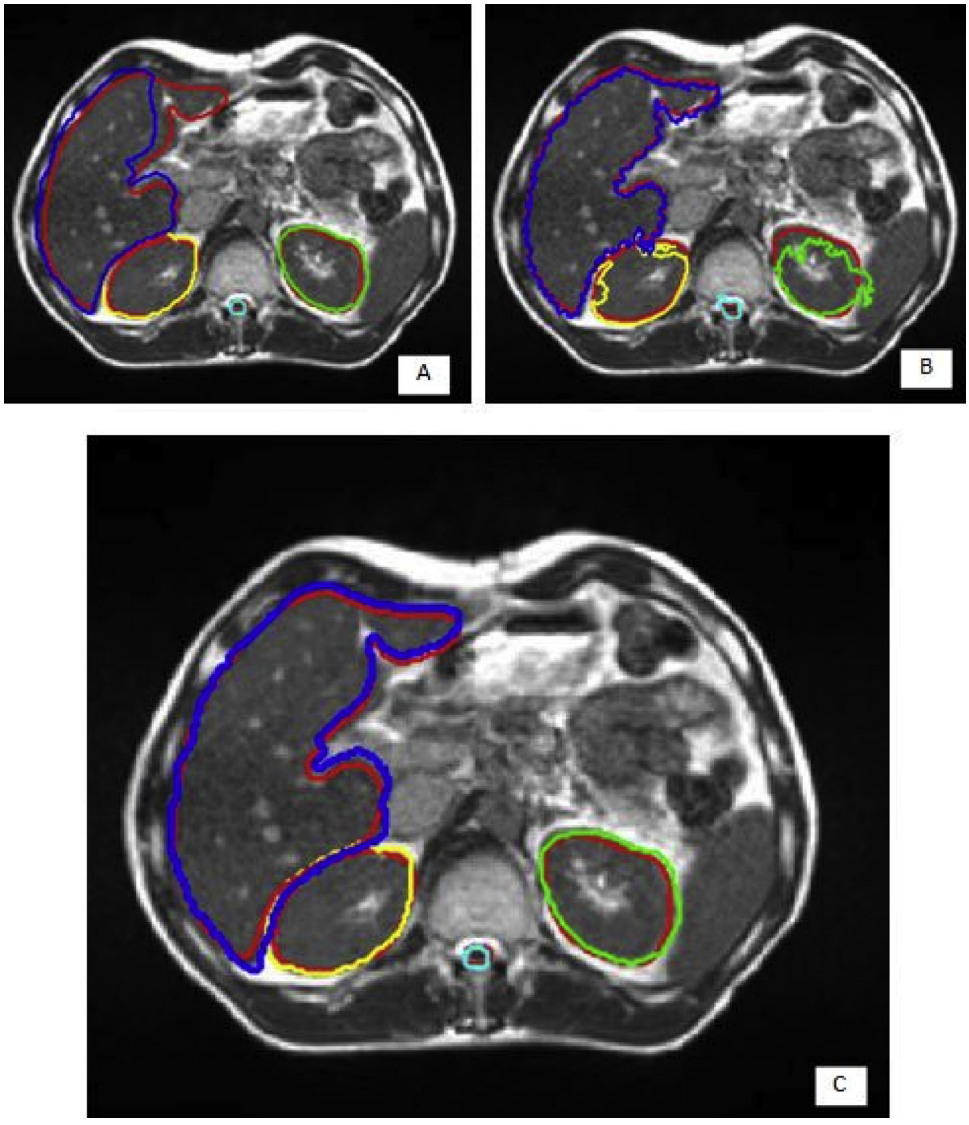

Fig. 3. Abdominal MRI transverse image showing the auto contouring results for liver, kidneys, and spinal canal obtained by the top-down (A) and bottom-up (B), and the fused top-down and bottom-up (C) methods. **Red contours:** manual contours used as ground truth. **Blue:** Auto-contours of liver. **Yellow:** Auto-contours of the right kidney. **Green:** Auto-contours of the left kidney. **Cyan:** Auto-contours of the spinal canal. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Table 1
Average Dice similarity index (DSI) and Hausdorff distance index (HDI) for kidneys, liver, and spinal canal using the bottom-up, top-down, and fused top-down and bottom-up methods.

	DSI			HDI			
	Right kidney	Left Kidney	Liver	Right Kidney	Left Kidney	Spinal Canal	Liver
Top-down Method	0.829 ± 0.02	0.898 ± 0.02	0.845 ± 0.03	6.226 ± 1.0	3.919 ± 0.2	2.753 ± 0.1	15.770 ± 1.0
Bottom-up method	0.863 ± 0.03	0.876 ± 0.03	0.951 ± 0.04	10.400 ± 1.0	13.727 ± 1.0	5.545 ± 0.7	11.232 ± 1.0
Fused Method	0.863 ± 0.01	0.898 ± 0.01	0.972 ± 0.01	6.226 ± 1.0	3.919 ± 0.4	2.753 ± 0.5	7.232 ± 0.8

intensities, multiple shapes, and considerable difference in liver size throughout treatment, it is difficult for only one method to completely solve the liver segmentation problem. From the above results, we can conclude that the bottom-up method was robust, even though the liver's shape was variable. In contrast, the top-down method could not get ideal results in these situations. The reliability of the bottom-up method can be partially related to computing both the position-oriented intensity information and the intensity information. Moreover, we can visually ascertain that the boundary between the liver and the right kidney is difficult to identify using only the bottom-up method because the features in that area are almost the same for both organs. However, it is easily detected using the top-down method through the registered contouring due to its stable property.

From the computational point of view, the current MR image pixel size and dimension achieve a satisfactory balance between resolution, accuracy, and computational time. The final step of our proposed method, just like human thinking, is likely to choose the results from the top-down method more than the results from the bottom-up method. Using our machine-learning-based, bottom-up method, our realization is arguably a more challenging test than the leave-one-out criterion [31] as one patient simulation MR imaging dataset is exploited in training and more separate daily patient MR scans for the testing. We could add more features from the planning day CT in the future for better supervised learning. Comparative study of the other machine-learning methods, such as random forest or the boosted probabilistic trees, could be supplemented with the SVM bottom-up method for better detail information extraction.

To our knowledge, this is the first fully automated segmentation approach using T1-weighted MR images for adaptive radiotherapy. From the technical point of view, our algorithm circumvents the limitation of poor tissue contrast due to image acquisition over a hybrid MR-ART platform using previous knowledge from planning CT and MRI. The minimal computational time coupled with high contouring accuracy makes, at this stage, our method suitable for supplementing the physician with a quick, automated plan that can be further manually adjusted while a patient is on board of the radiation therapy platform. While in this study the combination of low- and high-level features was only used, our method can be generalized to other imaging modalities by including mid-level features such as entropy value, Gabor texture [32], histogram of oriented gradient (HOG) [33] and local binary pattern (LBP) [34].

4. Conclusion

The clinical validation of a successful automated image segmentation algorithm is an essential step toward the application of a rapid workflow in MR-ART. By considering multiple features, our novel algorithm yields a high accuracy in delineating the liver, the kidney, and the spinal canal on T1 MRI acquired on-board of radiation therapy platforms. Our method can be implemented for semi-automated segmentation in MR-ART. Future studies should validate small bowel, stomach, and CTV contouring in order to achieve instantaneous, fully automated re-planning in abdominal MR-ART.

Funding

This research was partially supported by the National Natural Science Foundation of China under Grants 61603276 and 61772241, the Tianjin Natural Science Foundation of China under Grant 14JCQNJC04300, and the Fundamental Research Funds for the Central Universities of China under Grant JUSRP51614A.

Acknowledgment

The authors are deeply grateful to Bonnie Hami, M.A. for her invaluable editorial assistance.

References

[1] Acharya S, Fischer-Valuck BW, Kashani R, et al. Online magnetic resonance image guided adaptive radiation therapy: first clinical applications. Int J Radiat Oncol Biol Phys 2016;94:394–403.
[2] Timmerman RD, Xing L. Image-guided and adaptive radiation therapy. Lippincott Williams & Wilkins; 2012.
[3] Mutic S, Dempsey JF. The ViewRay system: magnetic resonance–guided and controlled radiotherapy. Semin Radiat Oncol 2014;24:196–9. Elsevier.
[4] Raaymakers B, Lagendijk J, Overweg J, et al. Integrating a 1.5 T MRI scanner with a 6 MV accelerator: proof of concept. Phys Med Biol 2009;54:N229.
[5] Devic S. MRI simulation for radiotherapy treatment planning. Med Phys 2012;39:6701–11.
[6] Khan FM, Gerbi BJ. Treatment planning in radiation oncology. Wolters Kluwer Health/Lippincott Williams & Wilkins; 2012.
[7] Lu Y, Chen I, Kashani R, et al. SU-C-BRA-01: interactive auto-segmentation for bowel in online adaptive mri-guided radiation therapy by using a multi-region labeling algorithm. Med Phys 2016;43:3320–1.
[8] Boykov YY, Jolly M-P. Interactive graph cuts for optimal boundary & region segmentation of objects in ND images. Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on. Vol 1: IEEE. 2001. p. 105–12.
[9] Mangan AP, Whitaker RT. Partitioning 3D surface meshes using watershed segmentation. IEEE Trans Vis Comput Graph 1999;5:308–21.
[10] Weszka JS. A survey of threshold selection techniques. Comput Graph Image Process 1978;7:259–65.
[11] Day E, Betler J, Parda D, et al. A region growing method for tumor volume segmentation on PET images for rectal and anal cancer patients. Med Phys 2009;36:4349–58.
[12] Li D, Liu L, Chen J, et al. Augmenting atlas-based liver segmentation for radiotherapy treatment planning by incorporating image features proximal to the atlas contours. Phys Med Biol 2016;62:272.
[13] Sharp G, Fritscher KD, Pekar V, et al. Vision 20/20: perspectives on automated image segmentation for radiotherapy. Med Phys 2014;41.
[14] Han X, Hoogeman MS, Levendag PC, et al. Atlas-based auto-segmentation of head and neck CT images. International Conference on medical image computing and computer-assisted intervention. Springer; 2008. p. 434–41.
[15] La Macchia M, Fellin F, Amichetti M, et al. Systematic evaluation of three different commercial software solutions for automatic segmentation for adaptive therapy in head-and-neck, prostate and pleural cancer. Radiat Oncol 2012;7:160.
[16] Qazi AA, Pekar V, Kim J, Xie J, Breen SL, Jaffray DA. Auto-segmentation of normal and target structures in head and neck CT images: a feature-driven model-based approach. Med Phys 2011;38:6160–70.
[17] Pekar V, McNutt TR, Kaus MR. Automated model-based organ delineation for radiotherapy planning in prostatic region. Int J Radiat Oncol Biol Phys 2004;60:973–80.
[18] Steinkraus D, Buck I, Simard P. Using GPUs for machine learning algorithms. Document Analysis and Recognition, 2005. Proceedings. Eighth International Conference on: IEEE. 2005. p. 1115–20.
[19] Zadeh LA. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. Fuzzy Sets Syst 1997;90:111–27.
[20] Bay H, Tuytelaars T, Van Gool L. Surf: speeded up robust features. European conference on computer vision. Springer; 2006. p. 404–17.
[21] Leutenegger S, Chli M, Siegwart RY. BRISK: binary robust invariant scalable

keypoints. Computer Vision (ICCV), 2011 IEEE International Conference on: IEEE. 2011. p. 2548–55.
[22] Dung L-R, Huang C-M, Wu Y-Y. Implementation of RANSAC algorithm for feature-based image registration. J Comput Commun 2013;1:46.
[23] Zou KH, Warfield SK, Bharatha A, et al. Statistical validation of image segmentation quality based on a spatial overlap index 1: scientific reports. Acad Radiol 2004;11:178–89.
[24] Chang C-C, Lin C-J. LIBSVM: a library for support vector machines. ACM Trans Intell Syst Technol (TIST). 2011;2:27.
[25] Cohn DA, Ghahramani Z, Jordan MI. Active learning with statistical models. J Artif Intell Res 1996;4:129–45.
[26] Daily Online Adaptation Versus Localization for MRI-Guided SBRT for Unresectable Primary or Oligometastatic Abdominal Malignancies: https://ClinicalTrials.gov/show/NCT02950025.
[27] Kim H, Monroe JI, Lo S, et al. Quantitative evaluation of image segmentation incorporating medical consideration functions. Med Phys 2015;42:3013–23.
[28] Rote G. Computing the minimum Hausdorff distance between two point sets on a line under translation. Inf Process Lett 1991;38:123–7.
[29] Hoyte L, Ye W, Brubaker L, et al. Segmentations of MRI images of the female pelvic floor: a study of inter-and intra-reader reliability. J Magn Reson Imaging 2011;33:684–91.
[30] Henke L, Olsen J, Green O, et al. Online adaptive magnetic resonance-guided (OAMR)-stereotactic body radiation therapy for abdominal malignancies: prospective dosimetric results from a phase 1 trial. Int J Radiat Oncol Biol Phys 2016;96:S222.
[31] Wolz R, Chu C, Misawa K, Fujiwara M, Mori K, Rueckert D. Automated abdominal multi-organ segmentation with subject-specific atlas generation. IEEE Trans Med Imaging 2013;32:1723–30.
[32] Depeursinge A, Foncubierta-Rodriguez A, Van De Ville D, Müller H. Three-dimensional solid texture analysis in biomedical imaging: review and opportunities. Med Image Anal 2014;18:176–96.
[33] Erdt M, Knapp O, Drechsler K, Wesarg S. Region detection in medical images using HOG classifiers and a body landmark network. Medical imaging 2013: computer-aided diagnosis. Vol 8670: International society for optics and photonics. 2013:867004.
[34] Ojala T, Pietikäinen M, Harwood D. A comparative study of texture measures with classification based on featured distributions. Pattern Recognit 1996;29:51–9.