Automated contour segmentation for treatment planning: challenges and potentials

Minsong Cao, Ph.D.
Department of Radiation Oncology
University of California, Los Angeles
UCLA

Learning objectives

• To review the current state-of-the-art methods on automated contour segmentation for treatment planning
• To understand the challenges and potentials on automated contour segmentation

Disclosure

• None
Background

- Clinical contouring is critical for treatment planning:
  - directly impact dosimetry quality and clinical decision
  - time consuming and labor intensive
- Contour is one of the largest sources of dosimetric uncertainty
  - contour error and variation
  - quality of contouring:
    - spatial accuracy
    - dosimetric accuracy

Automated contour segmentation

- Seek to reduce time and inter-observer variability
- Clinical applications:
  - Standard treatment planning
  - Adaptive treatment planning
  - Motion tracking and gating
- Commercial products available, but not frequently used in clinical practice
- Conflict findings reported on contour accuracy and time saving

Automated segmentation methods

- Non prior-knowledge
  - Directly based on image voxel intensities and/or gradient
  - High contrast structures e.g. lung, bone, air cavity
- Prior-knowledge
  - Atlas based segmentation
  - Statistical model based segmentation: Shape (SSM) or Appearance (SAM)
  - Machine learning based segmentation
  - Hybrid segmentation
Atlas based segmentation

- Single-Atlas Selection
- Multi-Atlas Selection

New image → DIR

Performance of atlas based segmentation

- Quality of atlas images and reference contours
- Atlas selection strategy: robust metric
- No consensus on database size
- Multiatlas can improve robustness of segmentation
- Prone to topological error
- Voting scheme is crucial
- Combination of multimodality images (MRI and CT)


Atlas based segmentation - DIR

- Quality of segmentation highly relies on deformable image registration (DIR)
- Ground-truth is not available
- Many different approaches and transformation modes

Brock et al. Med. Phys. 44 (7), July 2017
**Statistical model based segmentation**

- Confine the segmented contours to anatomically plausible shape or appearance
- Require training dataset to characterize variation of shape or appearance of structure
- Fit the test image to the model based on image intensities, gradients, features etc.

[Image]

Bigino et al. Heart, 2016 0:1–6

---

**Machine-learning based segmentation**

- Outstanding performance in classification, detection, pattern recognition
- Automatically learn priors for structures or image context and tissue appearance
- Require training and significant computational resource
- Usually combined with shape model or atlas based methods

[Image]

Lustberg et al. Radio and Onc 126 (2018)

---

**Segmentation for adaptive planning**

- Intra-object segmentation for anatomy at two different time points
- Deformable image registration is the most popular method
- Time constraints require very robust and accurate segmentation

[Image]

• Literature review of segmentation and registration methods for adaptive cervical cancer treatment planning:
  • Landmark, rigid, B-spline, shape constrained B-spline registration
  • A average of 0.85 Dice similarity and mean surface distance of 2-4mm
  • The use of shape priors significantly improved segmentation accuracy

Evaluation of segmentation performance

• Geometric
  • Moment based
    • Center/Volume of structure
  • Overlap based
    • Dice similarity coefficient
  • Distance based
    • Average/maximum distance
    • Intra-observer variability

• Dosimetric
  • Dose optimization
  • Dosimetric metrics (DVHs)
  • Clinical decision

Sharp et al. Medical Physics, Vol. 41, No. 5, 2014

Geometry based evaluation

• Dice similarity index (DSI)
  • Insensitive to large structure
  • Insensitive to fine details
• Hausdorff distance (HD)
  • Sensitive to small regions
  • Usually use 95% percentile
  • May not correlate with each other
  • Do not relate to dosimetry!

\[ D = \frac{2|X \cap Y|}{|X| + |P|} \]
Inter-observer variability

- A single manual contour may not truly represent the ground truth
- Inter-observer and Intra-observer contour variations exist
- Consensus on contour definition is not always available
- Inter-observer variability should be used as benchmark to assess the accuracy and robustness of auto-segmentation

From geometry to dosimetry

Stiehl B et al. AAPM 2017
Auto-segmentation Challenge

- Allows assessment of state-of-the-art segmentation methods under unbiased and standardized circumstances:
  - The same datasets (training/testing)
  - The same evaluation metrics
- Head & Neck Auto-segmentation Challenge at MICCAI 2015 conference
  - Data from RTOG 0522 clinical trial
  - 25 datasets as training data
  - 10 datasets for off-site and 5 for on-site (2 hours) testing
  - 9 anatomical structures (brainstem, optical chiasm, mandible, parotid glands and submandibular glands)
Mandible:
• Exclusion of teeth
• Image artifacts from dental implant

<table>
<thead>
<tr>
<th>Team</th>
<th>SM</th>
<th>PR</th>
<th>UW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dice</td>
<td>0.95</td>
<td>0.765</td>
<td>0.708</td>
</tr>
<tr>
<td>95thHD(mm)</td>
<td>2.041</td>
<td>5.919</td>
<td>29.458</td>
</tr>
</tbody>
</table>

AAM - active appearance model
SSM - statistical shape model

Parotid glands:
• Large shape variation
• Poor soft tissue contrast
• Heterogeneous tissue including vessels and ducts

More on segmentation challenge
• AAPM 2017 Thoracic Auto-segmentation Challenge
• RTOG 1106 contouring atlas
• 36 training sets, 12 offline test and 12 live competition cases
• Intra-observer contour variability considered

Raudaschl et al.: Medical Physics, 44 (5), 2017

http://autocontouringchallenge.org
Summary

• Automated segmentation has shown promising performance in contouring for treatment planning

• Improvement on robustness, accuracy and throughput is still needed:
  • Consensus on contouring and benchmark database
  • Standardization of imaging acquisition; improvement of image quality; combination of multiple image modalities
  • Advancement in model and machine-learning based algorithms
  • Quality metrics and QA tools for spatial and dosimetric uncertainties
  • Effective translation from research to clinic with sufficient user training

Acknowledgement

• Dr. Dan Ruan
• Dr. Ke Sheng
• Brad Stiehl