Air-Tissue Boundary Segmentation in Real-Time Magnetic Resonance Imaging Video using Semantic Segmentation with Fully Convolutional Networks

Valliappan CA, Renuka Mannem and Prasanta Kumar Ghosh

SPIRE LAB
Electrical Engineering,
Indian Institute of Science (IISc), Bangalore, India

September 6, 2018
Section 1

1. Introduction

2. Proposed Method

3. Experiments and Results

4. Summary
Introduction

- **Goal**: Segmentation of the Air-Tissue Boundaries (ATBs)
Goal: Segmentation of the Air-Tissue Boundaries (ATBs)
Introduction

- **Goal**: Segmentation of the Air-Tissue Boundaries (ATBs)

- **Approach**: Semantic segmentation using Fully Convolutional Networks (FCN).
Motivation

Need for study

Understanding speech production.
Introduction

Motivation

Need for study
Understanding speech production.

Why rt-MRI?
Non-invasive technique\(^1\) unlike Electromagnetic articulograph (EMA), Ultra-sound, X-Ray.

\(^1\) E. Bresch et. al “Seeing speech: Capturing vocal tract shaping using real-time magnetic resonance imaging,” 2008.
Introduction

Dataset

- **USC-TIMIT** corpus
- **2-Female** (F1, F2) and **2-Male** (M1, M2).
- Subset: 16 Videos from each subject.
- Video: 23.18 fps
- Spacial resolution of $68 \times 68$. 

SPIRE LAB, IISc, Bangalore
Dataset

- **Manual annotation:**
  1. Complete ATBs
  2. Upper lip (UL)
  3. Lower lip (LL)
  4. Tongue base (AVR)
  5. Velum tip (VEL)
  6. Glottis begin (GLTB)

- Number of frames: 1462, 1270, 1642, 1399 for subjects F1, F2, M1, M2 respectively.

- Division of tissue regions into 3 masks.
Section 2

1. Introduction

2. Proposed Method

3. Experiments and Results

4. Summary
Proposed FCN based segmentation

1. **Three FCNs**: One FCN for each mask.

2. **FCN$_i$**: Does a given pixel belong to mask$_i$ or air cavity region?

---

Image Enhancement

**Proposed Method**

**Image filling**
1. Eliminate intra-region discontinuities
2. Morphological operations

**Image smoothing**
1. Eliminate jagged boundaries.
2. Moving average filter
Contour Prediction

- **Stage 1**: Canny edge detection
- **Stage 2**: Connecting edges via concave hull algorithm

Contour Prediction

- **Stage 1**: Canny edge detection
- **Stage 2**: Connecting edges via concave hull algorithm

Contour Pruning

- Proposed Method

Obtain upper and lower contours within the vocal tract.
Obtain **upper** and **lower** contours within the vocal tract
Proposed Method

Obtaining Upper Contour

![Diagram showing obtaining upper contour with markers UL, VEL, GLTB, and C1, C3, C1prun]
Proposed Method

Obtain Lower Contour
Proposed Method

Obtain Lower Contour
Proposed Method

Obtain Lower Contour
Proposed Method

Obtain Lower Contour

$2^{nd}$ order polynomial fit
Proposed Method

Proposed FCN based segmentation

Illustration of the steps in the proposed FCN based approach
Section 3

1. Introduction

2. Proposed Method

3. Experiments and Results

4. Summary
Experiments and Results

Experiments

Experimental Setup

- **4-fold** setup
- Training set: $\sim 2900$.
- Development & Test set: $\sim 1443$
- **120 epochs**, early stopping condition.
- Baseline scheme: Maeda grid-line (MG) based segmentation

---

1 - Kim et al., 2014.
Evaluation metrics

- **DTW distance**\(^1\): Measures the closeness of the estimated contour to the ground truth contour.

- **Pixel accuracy**\(^2\): To evaluate the performance of FCN.

---

### Experiments and Results

#### Experiments

**DTW distance**

<table>
<thead>
<tr>
<th></th>
<th>Lower ATB</th>
<th></th>
<th>Upper ATB</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$MG$</td>
<td>$FCN$</td>
<td>$MG$</td>
<td>$FCN$</td>
</tr>
<tr>
<td>F1</td>
<td>1.21 ± 0.21</td>
<td>1.00 ± 0.25</td>
<td>1.02 ± 0.19</td>
<td>0.91 ± 0.21</td>
</tr>
<tr>
<td>F2</td>
<td>1.28 ± 0.27</td>
<td>1.13 ± 0.31</td>
<td>1.24 ± 0.29</td>
<td>1.08 ± 0.19</td>
</tr>
<tr>
<td>M1</td>
<td>1.26 ± 0.60</td>
<td>1.17 ± 0.25</td>
<td>1.10 ± 0.20</td>
<td>1.02 ± 0.20</td>
</tr>
<tr>
<td>M2</td>
<td>1.35 ± 0.30</td>
<td>1.21 ± 0.23</td>
<td>1.19 ± 0.24</td>
<td>1.09 ± 0.21</td>
</tr>
<tr>
<td>Average</td>
<td>1.24 ± 0.35</td>
<td>1.13 ± 0.26</td>
<td>1.14 ± 0.23</td>
<td>1.03 ± 0.20</td>
</tr>
</tbody>
</table>

1 - Kim et al., 2014.
Experiments and Results

Experiments

DTW distance

<table>
<thead>
<tr>
<th></th>
<th>Lower ATB</th>
<th></th>
<th>Upper ATB</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SUB</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F1</td>
<td>MG</td>
<td>FCN</td>
<td>MG</td>
<td>FCN</td>
</tr>
<tr>
<td>F2</td>
<td>1.21 ± 0.21</td>
<td>1.00 ± 0.25</td>
<td>1.02 ± 0.19</td>
<td>0.91 ± 0.21</td>
</tr>
<tr>
<td>M1</td>
<td>1.28 ± 0.27</td>
<td>1.13 ± 0.31</td>
<td>1.24 ± 0.29</td>
<td>1.08 ± 0.19</td>
</tr>
<tr>
<td>M2</td>
<td>1.26 ± 0.60</td>
<td>1.17 ± 0.25</td>
<td>1.10 ± 0.20</td>
<td>1.02 ± 0.20</td>
</tr>
<tr>
<td>Average</td>
<td>1.35 ± 0.30</td>
<td>1.21 ± 0.23</td>
<td>1.19 ± 0.24</td>
<td>1.09 ± 0.21</td>
</tr>
</tbody>
</table>

Proposed method yields a lower DTW distance.

1 - Kim et al., 2014.
Experiments and Results

Experiments

Illustration

1 - Kim et al., 2014.
Experiments and Results

Experiments

Illustration

Proposed method performs comparable to or better than baseline

1 - Kim et al., 2014.
Complete Contours – Additional benefit

<table>
<thead>
<tr>
<th>SUB</th>
<th>$Mask_1$</th>
<th>$Mask_2$</th>
<th>$Mask_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>0.89 ± 0.11</td>
<td>1.05 ± 0.19</td>
<td>0.83 ± 0.11</td>
</tr>
<tr>
<td>F2</td>
<td>1.02 ± 0.17</td>
<td>1.12 ± 0.24</td>
<td>0.80 ± 0.10</td>
</tr>
<tr>
<td>M1</td>
<td>1.03 ± 0.21</td>
<td>1.37 ± 0.35</td>
<td>0.80 ± 0.09</td>
</tr>
<tr>
<td>M2</td>
<td>0.98 ± 0.09</td>
<td>1.01 ± 0.17</td>
<td>0.85 ± 0.10</td>
</tr>
</tbody>
</table>
Experiments and Results

Complete Contours – Additional benefit

<table>
<thead>
<tr>
<th>SUB</th>
<th>Mask₁</th>
<th>Mask₂</th>
<th>Mask₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>0.89 ± 0.11</td>
<td>1.05 ± 0.19</td>
<td>0.83 ± 0.11</td>
</tr>
<tr>
<td>F2</td>
<td>1.02 ± 0.17</td>
<td>1.12 ± 0.24</td>
<td>0.80 ± 0.10</td>
</tr>
<tr>
<td>M1</td>
<td>1.03 ± 0.21</td>
<td>1.37 ± 0.35</td>
<td>0.80 ± 0.09</td>
</tr>
<tr>
<td>M2</td>
<td>0.98 ± 0.09</td>
<td>1.01 ± 0.17</td>
<td>0.85 ± 0.10</td>
</tr>
</tbody>
</table>

Unique to the proposed method
Experiments and Results

Experiments

Pixel Accuracy

<table>
<thead>
<tr>
<th>SUB</th>
<th>Mask$_1$</th>
<th>Mask$_2$</th>
<th>Mask$_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>99.39</td>
<td>98.34</td>
<td>99.73</td>
</tr>
<tr>
<td>F2</td>
<td>99.20</td>
<td>98.14</td>
<td>99.75</td>
</tr>
<tr>
<td>M1</td>
<td>99.28</td>
<td>97.97</td>
<td>99.75</td>
</tr>
<tr>
<td>M2</td>
<td>99.32</td>
<td>98.09</td>
<td>99.70</td>
</tr>
</tbody>
</table>
Pixel Accuracy

<table>
<thead>
<tr>
<th>SUB</th>
<th>Mask₁</th>
<th>Mask₂</th>
<th>Mask₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>99.39</td>
<td>98.34</td>
<td>99.73</td>
</tr>
<tr>
<td>F2</td>
<td>99.20</td>
<td>98.14</td>
<td>99.75</td>
</tr>
<tr>
<td>M1</td>
<td>99.28</td>
<td>97.97</td>
<td>99.75</td>
</tr>
<tr>
<td>M2</td>
<td>99.32</td>
<td>98.09</td>
<td>99.70</td>
</tr>
</tbody>
</table>

Unique to the proposed method
Discussion

1. On an average $\sim 1\%$ pixels are being misclassified.
2. Misclassified pixels – boundary region: due to low resolution of the image.
3. Precision of annotation: 1 decimal place
4. Proposed method: Pixel level
Section 4

1. Introduction
2. Proposed Method
3. Experiments and Results
4. Summary
Conclusions
Conclusions

- Proposed method yields a better ATB than baseline – DTW distance

Future Directions

- Improve the performance using de-blurring technique and edge-detection algorithm.
Conclusions

- Proposed method yields a better ATB than baseline – DTW distance
- Complete contours – Novelty of the proposed method.
Conclusions

- Proposed method yields a better ATB than baseline – DTW distance
- Complete contours – Novelty of the proposed method.

Future Directions

Improve the performance using de-blurring technique and edge-detection algorithm.
The authors thank Pratiksha Trust for their support. Special credits to students those who helped us annotate this huge dataset.
Questions?
Proposed FCN based segmentation

Convex hull vs Concave hull

Convex Hull

Concave Hull