EXTRACTION OF TONGUE CONTOUR IN REAL-TIME MAGNETIC RESONANCE IMAGING SEQUENCES

Dawei Zhang {dawei.zhang@nlpr.ia.ac.cn}, Minghao Yang, Jianhua Tao, Yang Wang, Bin Liu, Danish Bukhari
National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences

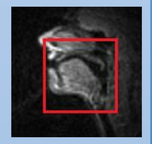
1. Problem Identification & Motivation

- Real-time magnetic resonance imaging (rtMRI) is becoming a practical tool in speech production research and language pathology observation.
- It is still a challenge to extract the tongue contour accurately in rtMRI sequences, because tongue is a flexible soft tissue and often touches other organs such as lips and upper mandible.
- Researchers have proposed many methods of extracting tongue contours from MRI data, however most of them are mainly dependent on complicated manual operation.
- This paper proposes a novel semi-automatic tongue contour extraction method from rtMRI sequences.

2. Initial Boundary Detection

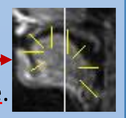
- Tongue movement range registration

  A magnetic resonance image usually contains many kinds of tissue boundaries and noise information. A predefined range helps to find the initial boundary points quickly in the whole image.



- Initial boundary points detection

  According to the characteristic of tongue shape, the registered region is firstly divided into two parts — the front part and the back part of tongue.



  In the front part, gradient directions of the tongue boundary points are followed as Bottom to Up, Right-Bottom to Left-Top, Right to Left and Right-Top to Left-Bottom;

-1	1	-1
-1	0	1
1	-1	1

  While in the back part, the directions are mainly Bottom to Up, Left-Bottom to Right-Top, Left to Right and Left-Top to Right-Bottom.

-1	1	-1
1	0	-1
1	-1	1

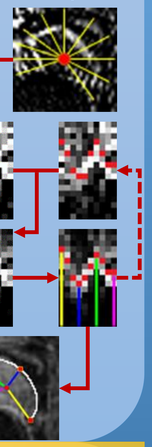
  So we use combined multi-directional Sobel operators to get equal gradient value on the four directions in each part. The image of initial boundary points is shown in the right figure.

3. Tongue Contour Extraction

- Boundary intensity map construction

  Gridlines are constructed around the central point of registered region and superimposed at a certain angle interval from the glottis to the glossodesmus. Then a boundary intensity map is constructed as following:

  a) the sectors between every two adjacent gridlines are labeled successively from the glottis to the glossodesmus.
  b) a boundary intensity map of $D \times N$ is built, where D is the maximum distance from the tongue boundary points to the central point and N is the total number of labeled sectors.
  c) along each arc of sectors, the maximum intensity of boundary points is selected as the map element.



- Searching for optimal boundary route

  A. Map modification with previous contour

$$M_{ij} = K_{ij} \times \exp\left(-\left(\frac{i - i_j}{\sigma_M}\right)^2\right)$$

  B. Column transition probabilities

$$T_{hi} = \exp\left(-\left(\frac{h - i}{\sigma_T}\right)^2\right)$$

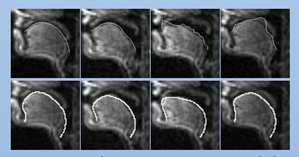
  C. Optimal boundary route

$$\boldsymbol{r}^* = \arg\max_{\boldsymbol{r}} P(\boldsymbol{r} | M, \sigma_T) = \arg\max_{\boldsymbol{r}} \prod_{j=1}^{N} T_{r(j-1)r(j)} M_{r(j)j}$$

  D. Obtaining tongue contour using B-spline approximation [*Yang, ICASSP 2013*]

4. Experiments

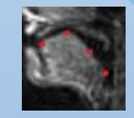
In the figure below, images in the first row are results of the baseline method [*Proctor, Interspeech 2010*], and the second row are the proposed method.



It demonstrates that the baseline method and the proposed method are both effective in most cases.

However, the baseline method may extract the unreasonable tongue contour when tongue touches other organ boundaries. While the proposed method works well in this case, which shows improved robustness in more different situations.

Key points on tongue tip (TT), tongue dorsum (TD), rear part of tongue body (TB) and tongue root (TR).



The table below shows the extraction errors of the key points in 50 frames respectively from Speaker M1 and F1. The extraction errors are defined as RMSE of Euclidean distance in pixels between the extracted and true key points.

Key points	Speaker M1	Speaker F1	Mean
TT	1.88	1.57	1.73
TD	0.41	0.51	0.46
TB	0.86	0.69	0.78
TR	0.07	0.13	0.10

To examine the extraction accuracy of tongue contour, 50 image frames are respectively taken from rtMRI sequences of ten different speakers. The errors of tongue contour extraction results using the baseline method and the proposed method are shown in the table below, which are the mean RMSE of the four tongue contour key points above.

Speaker	The baseline method	The proposed method
M1	1.06	0.81
M2	0.86	0.70
M3	0.57	0.63
M4	0.65	0.46
M5	1.16	0.92
F1	0.91	0.73
F2	1.02	0.69
F3	0.74	0.58
F4	0.80	0.55
F5	0.54	0.61
Mean	0.83	0.67

5. Conclusions

A novel method of semi-automatic tongue contour extraction in rtMRI sequences is proposed. Experiments show that the proposed method is of higher accuracy in tongue contour extraction, especially when some parts of tongue touch other organs. The proposed method could be extended to extract contours of other organs such as joints and backbones in MRI, Ultrasonic or X-ray images.