On-line Supplement for
IEEE Transactions on Biomedical Engineering paper entitled
"Methods for 2D and 3D Endobronchial Ultrasound Image Segmentation"
by Xiaonan Zang, Rebecca Bascom, Christopher R. Gilbert,
Jennifer W. Toth, and William E. Higgins

Section III supplemental material:

This section provides additional parameter-sensitivity tests and illustrative examples for the proposed 2D and 3D segmentation methods. Reference [1] provides complete detail.

Additional Results for Section III.A — Below, we discuss the sensitivity of the four-step 2D EBUS segmentation method to parameter and method variations, as mentioned in Section III.A.

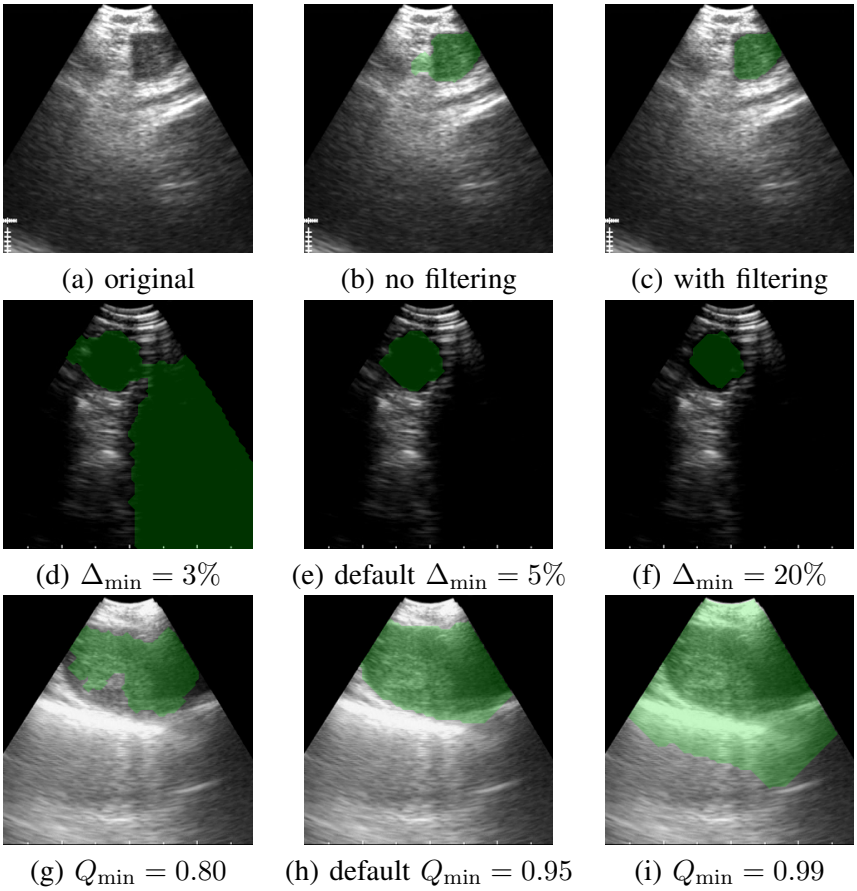
Beginning with the Image Filtering step, a test run on a 12-ROI subset drawn from 6 human cases (5 lymph nodes, 4 pulmonary arteries, 2 azygos veins, and 1 aorta) revealed that the mean Dice index of segmented ROIs improved from 81.1%±12.7% (no filtering) to 88.3%±4.5% (with filtering), with the poorest segmentation's Dice index rising from 51.4% to 80.3%. Fig. 9, row 1, illustrates how, without filtering, high noise and uncertain boundary segments result in excessive segmentation leakage.

Regarding Seed Selection, we performed a study to determine if a seed's location affects the segmentation result. In the study, an experienced technician interactively selected "reasonable" seeds in 2 separate trials for 19 ROIs and noted how the segmentations varied over these trials with respect to the ground truth (a "reasonable" seed is one that is somewhat centered within an ROI). Overall, this resulted in only a 1% standard-deviation variation on average for an ROI. Also, over the complete 19-ROI test set, the interactively selected seeds gave an aggregate Dice index of 86.2%±6.2%. As a comparison, automatic seed-selection gave an aggregate Dice index = 89.1%±5.2% for the same 19-ROI set.

Next, regarding the SVM seed classifier, we used 20 consecutive ROIs in our ROI set to train the classifier. One could argue that the classifier is then biased toward this preselected training data. However, the Dice index segmentation performance for these 20 ROIs was 90.2%±5.9% (range: [74.6%, 96.8%]) versus 90.0%±4.9% (range: [74.6%, 96.8%]) for the complete 52-ROI set. In addition, the 3D segmentation method, tested on data differing from the SVM training set, performed effectively. Hence, little bias appears to exist, and we noted minimal effect of the classifier on the final segmentation output.

Regarding the fast-marching process of the Initial ROI Segmentation step, two parameters warrant attention: d, which influences the gradient range emphasized by the sigmoid function (10), and $\Delta_{\min}$ in **Algorithm 1**. For the sensitivity tests, we drew upon the same 12-ROI subset used for testing Image Filtering. Table I measures the sensitivity to d. Overall, the range [66.5, 101.5] gave similar results, with the default $d = 76.5$ providing the best results. In general, a low d causes the sigmoid function to unnecessarily emphasize small intensity changes within an ROI, which potentially leads to under-segmentation; a high d, on the other hand, causes the sigmoid function to omit weak ROI boundaries, which may result in over-segmentation. Also, cleaner ROIs tended to be less sensitive to variations in d than noisier ROIs.

We next tested the impact of $\Delta_{\min}$, as shown in Table II and Fig. 9, row 2. We noted a satisfactory operating range of [5%, 10%] with the best results achieved by default $\Delta_{\min} = 5\%$. Also, low/high $\Delta_{\min}$ resulted in over/under-segmentation, with cleaner ROIs again exhibiting less sensitivity to parameter changes. Because the fast-marching process grows ROIs from the selected seeds based on the gradient magnitude, it tends to extract the majority of pixels belonging to more homogeneous ROIs. But, because EBUS ROIs often have incomplete boundaries and high noise, fast marching is susceptible to overgrowth,



(a) original (b) no filtering (c) with filtering

(d) $\Delta_{\min} = 3\%$ (e) default $\Delta_{\min} = 5\%$ (f) $\Delta_{\min} = 20\%$

(g) $Q_{\min} = 0.80$ (h) default $Q_{\min} = 0.95$ (i) $Q_{\min} = 0.99$

Fig. 9. Impact of parameter and method variations in 2D EBUS segmentation. In all examples, segmented regions appear green. Row 1 — image filtering; (a-c) depicts results for a station 7 lymph node, case 21405.139. Row 2 — impact of $\Delta_{\min}$ of the fast-marching process; (d-f) depict segmentations for various $\Delta_{\min}$ for an azygos vein, case 21405.113. Row 3 — impact of $Q_{\min}$ on ROI finalization; (g-i) show segmentations for various $Q_{\min}$ for a station 4R lymph node, case 21405.116.

as exemplified by Fig. 9. Hence, we chose our default values of d and $\Delta_{\min}$ to err on the side of under-segmentation.

Finally, regarding ROI Finalization, we tested the sensitivity to variations in $Q_{\min}$ and $R_{\min}$ for the 12-ROI test set used previously. When $Q_{\min}$ was varied over the range [0.10, 0.99], TPF monotonically increased from 67.9±16.9 to 92.7±6.5, while FDR monotonically increased from 1.3±1.0 to 29.2±37.4 (Table III). Hence, low $Q_{\min}$ can result in under-segmentation, while high $Q_{\min}$ can produce excessive leakage, as shown by Fig. 9, row 3. The default $Q_{\min} = 0.95$ provided the best results.

$R_{\min}$ was found to have less impact than $Q_{\min}$, with leakage being the greatest concern. When $R_{\min}$ was varied over the range [0.0%, 20.0%], FDR monotonically decreased from 18.6±30.2 to 4.3±3.5 (Table IV). Thus, low $R_{\min}$ causes region growing to be too liberal, resulting in possible excessive leakage through small broken-boundary sub-regions. We found that $R_{\min}$ operated effectively in the range [3%, 15%], with 3% being our chosen default. We noted that ROI finalization particularly helps fill in the initial ROI segmentation of noisier ROIs, such as lymph nodes. In particular, for the 12-ROI test set, the Dice segmentation metric increased from 72.7±9.4 after initial ROI segmentation to 86.9±6.4 after ROI finalization. However, because leakage can be severe in EBUS segmentation, we picked parameters that tended toward conservative segmentations [1].

TABLE I

SENSITIVITY OF 2D EBUS SEGMENTATION TO VARIATIONS IN SIGMOID PARAMETER d IN (10) FOR THE FAST-MARCHING METHOD. THE DICE METRIC MEAN±SD AND [MIN, MAX] RANGE ARE GIVEN OVER A 12-ROI TEST SET.

d	Dice	
51.5	74.8±22.1	[29.4, 94.9]
56.5	77.3±22.3	[29.4, 95.4]
61.5	82.9±17.5	[29.4, 95.3]
66.5	88.1± 7.2	[70.7, 94.9]
71.5	89.3± 5.3	[80.6, 95.9]
76.5	90.9± 4.1	[82.5, 96.3]
81.5	90.5± 4.8	[81.3, 96.0]
86.5	89.9± 5.1	[80.6, 96.5]
91.5	90.4± 4.3	[82.3, 94.8]
96.5	90.2± 4.8	[81.3, 95.8]
101.5	90.3± 5.0	[80.6, 96.0]
127.5	65.3±31.6	[19.7, 95.3]

TABLE II

SENSITIVITY OF 2D EBUS SEGMENTATION TO VARIATIONS IN $\Delta_{\min}$ OF **Algorithm 1** FOR INITIAL ROI SEGMENTATION. $\Delta_{\min}$ IS PRESENTED IN $\%$. DICE MEAN±SD AND [MIN, MAX] RANGE ARE GIVEN OVER A 12-ROI TEST SET.

$\Delta_{\min}$	Dice	
0	83.8±22.5	[11.6, 96.5]
1	83.7±22.5	[11.6, 96.3]
3	83.5±22.3	[11.9, 96.6]
5	90.9± 4.1	[82.5, 96.3]
7	89.0± 5.5	[80.6, 96.6]
9	89.5± 5.7	[80.8, 96.3]
10	88.8± 6.2	[78.1, 95.9]
20	85.1± 6.8	[73.2, 93.4]

TABLE III

SENSITIVITY OF 2D EBUS SEGMENTATION TO VARIATIONS IN $Q_{\min}$ IN THE ROI FINALIZATION STEP. $Q_{\min}$ DETERMINES THE CORRESPONDING VALUE OF $\delta_{\min}$ PER (14) IN THE MAIN MANUSCRIPT. DICE MEAN±SD AND [MIN, MAX] RANGE, TPF, AND FDR ARE GIVEN OVER A 12-ROI TEST SET.

$Q_{\min}$	$\delta_{\min}$	Dice		TPF		FDR	
0.10	0.13	79.0±14.2	[38.2, 91.4]	67.9±16.9	[23.6, 85.2]	1.3±1.0	[0.0, 2.6]
0.20	0.25	80.2±12.8	[45.0, 92.1]	69.3±16.0	[29.1, 86.5]	1.4±1.0	[0.0, 2.7]
0.40	0.52	82.3±10.1	[56.1, 92.1]	71.9±13.6	[39, 86.6]	1.4±1.1	[0.0, 2.9]
0.60	0.84	84.5±9.0	[60.9, 92.6]	75.4±12.6	[44.6, 89.1]	2.3±1.8	[0.0, 6.2]
0.80	1.28	86.7±8.3	[64.8, 93.6]	79.3±12.5	[48.9, 94.6]	3.0±2.4	[0.0, 7.4]
0.90	1.64	88.4±6.2	[76.0, 94.5]	82.6±10.3	[63.4, 98.6]	4.1±3.8	[0.3, 12]
0.95	1.96	90.9±4.1	[82.5, 96.3]	85.7±9.3	[70.1, 98.6]	4.8±4.4	[0.3, 13.9]
0.98	2.33	79.9±26.0	[14.4, 96.1]	89.9±7.8	[78.5, 100.0]	18.3±30.9	[0.3, 92.3]
0.99	2.56	71.7±31.7	[13.2, 95.8]	92.7±6.5	[80.2, 100.0]	29.2±37.4	[0.5, 92.9]

TABLE IV

SENSITIVITY OF 2D EBUS SEGMENTATION METHOD TO VARIATIONS IN $R_{\min}$ FOR THE ROI FINALIZATION STAGE. $R_{\min}$ IS PRESENTED IN $\%$. DICE MEAN±SD AND [MIN, MAX] RANGE, TPF, AND FDR ARE GIVEN OVER A 12-ROI TEST SET.

$R_{\min}$	Dice		TPF		FDR	
0	79.0±25.5	[10.4, 96.4]	88.2±9.8	[71.4, 100.0]	18.6±30.2	[0.3, 94.5]
1	85.9±14.5	[41.2, 96.0]	87.7±9.6	[71.4, 100.0]	11.5±19.2	[0.3, 74.0]
3	90.9±4.1	[82.5,96.3]	85.9±8.3	[71.4, 97.4]	5.1±3.5	[0.3, 13.9]
5	89.6±5.1	[80.0, 95.8]	84.9±7.7	[71.4, 94.8]	4.9±3.6	[0.3, 14.1]
7	89.2±4.8	[80.0, 95.3]	84.1±6.8	[71.4, 93.2]	4.7±3.7	[0.3, 14.1]
9	88.7±4.5	[79.8, 93.8]	82.9±6.2	[71.4, 90.6]	4.5±3.5	[0.3, 12.7]
10	88.7±4.5	[79.8, 93.8]	82.9±6.2	[71.4, 90.6]	4.5±3.5	[0.3, 12.7]
15	88.1±4.3	[79.8, 93.8]	81.8±5.7	[71.4, 89.2]	4.4±3.5	[0.3, 12.7]
20	87.6±4.5	[79.8, 93.8]	80.9±6.2	[71.0, 89.2]	4.3±3.5	[0.3, 12.7]

Additional Results for Section III.B — Below, we discuss the sensitivity of the 3D EBUS segmentation method to parameter variations.

Tables V-VIII give sample sensitivity results based on tests run on two sequences from case 21405.116. Sequence 1 vividly depicts a station-4R lymph node throughout its length. Sequence 2 is more challenging, focusing on the left PA, which often appears with missing boundary components and which fluctuates in size across the sequence from large, then small, and finally large again.

Time step τ in relation (26) can be varied widely over the valid range [0, 0.25], but anomalous results occur for large time steps (Table V). Hence, in keeping with our strategy of erring on the side of being conservative, we selected $\tau = 0.10$ as a default. Regarding ϵ in speed function $\widehat{F}$, we found a usable range of [5, 12] for ϵ, with a low value producing excessive segmentation leakage, while a high value completely stalled the segmentation process (Table VI). Hence, our default choice $\epsilon = 8$.

Parameter η, which influences $\widehat{F}$'s advection term in (23), was tested over the range [0, 20] and proved to have a robust operating range [5, 20] (Table VII). Nullifying the term via $\eta = 0.0$, however, resulted in segmentation leakage. Finally, per Table VIII, $\mathcal{E}_{\min}$, tested over the range [0.001, 0.64], also could be used over a wide operating range [0.01, 0.08], with overly small values prematurely halting the segmentation process, while large values resulted in overgrowth. Thus, in line with the default chosen by others in related applications, we chose the default $\mathcal{E}_{\min}$ = 0.02 [2]. Fig. 10 illustrates the impact of unsatisfactory parameter choices.

TABLE V

SENSITIVITY OF 3D EBUS SEGMENTATION METHOD TO VARIATIONS IN τ FOR THE GEODESIC LEVEL-SET PROCESS. DICE MEAN±SD AND [MIN, MAX] RANGE MEASURE AGGREGATE SEGMENTATION ACCURACY AS COMPARED TO THE GROUND-TRUTH OVER ALL FRAMES IN A GIVEN SEQUENCE.

τ	4R node		Left PA	
0.05	89.8±5.1	[82.9, 92.4]	79.8±6.8	[71.5, 89.7]
0.1	90.0±5.8	[81.8, 92.2]	81.4±7.2	[70.6, 89.7]
0.15	90.0±6.0	[81.1, 92.1]	81.9±6.9	[71.4, 89.7]
0.2	89.7±5.6	[81.8, 92.1]	81.9±7	[71.2, 89.5]
0.25	89.8±6.3	[79.7, 92.3]	15.0±33.8	[0.0, 89.3]

TABLE VI

SENSITIVITY OF 3D EBUS SEGMENTATION TO VARIATIONS IN ϵ FOR THE GEODESIC LEVEL-SET PROCESS.

ϵ	4R node		left PA	
0	90.4±4.6	[84.3, 92.2]	58.7±15	[30.7, 89.7]
4	90.0±5.2	[82.9, 92.3]	82.1±7.1	[71.4, 89.7]
8	90.0±5.8	[81.8, 92.2]	81.4±7.2	[70.6, 89.7]
12	90.0±7.1	[78.9, 92.1]	82.1±7.1	[71.4, 89.7]
16	88.0±8.7	[73.5, 92.2]	13.7±31.6	[0.0, 89.7]

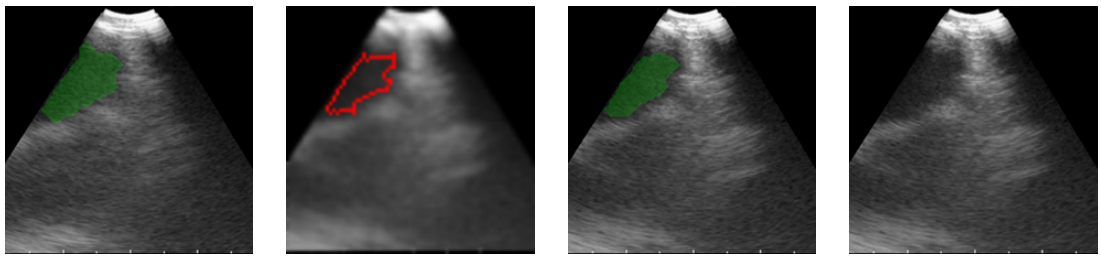
TABLE VII

SENSITIVITY OF 3D EBUS SEGMENTATION TO VARIATIONS IN η FOR THE GEODESIC LEVEL-SET PROCESS.

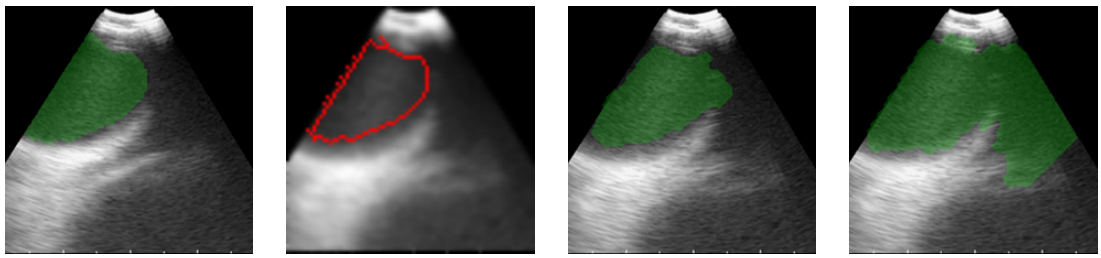
η	4R node		left PA	
0	90.6±4.3	[85.8, 93.2]	61.6±15.1	[29.9, 89.7]
5	89.9±5.2	[82.7, 92.4]	79.5±7.2	[71.0, 89.8]
10	90.0±5.8	[81.8, 92.2]	81.4±7.2	[70.6, 89.7]
15	89.9±5.6	[81.7, 92.1]	81.2±7.2	[71.2, 89.7]
20	89.9±6.0	[81.1, 92.2]	81.5±7.1	[70.6, 89.7]

TABLE VIII
SENSITIVITY OF 3D EBUS SEGMENTATION TO VARIATIONS IN $\mathcal{E}_{\min}$ FOR THE GEODESIC LEVEL-SET PROCESS.

$\mathcal{E}_{\min}$	4R node		left PA	
0.001	89.6±5.9	[80.1, 92.4]	14.7±34.0	[0.0, 89.7]
0.01	90.0±5.8	[81.7, 92.2]	78.7±7.5	[70.9, 89.7]
0.02	90.0±5.8	[81.8, 92.2]	81.2±7.5	[70.6, 89.7]
0.04	89.8±5.2	[82.9, 92.3]	79.2±7.4	[70.7, 89.7]
0.08	90.2±5.2	[83.3, 92.8]	79.5±7.2	[70.7, 90.5]
0.16	90.6±5.0	[84.3, 92.6]	79.5±7.8	[71.3, 90.4]
0.32	89.5±5.9	[80.5, 92.6]	80.2±9.4	[68.8, 94.0]
0.64	47.1±12.8	[37.3, 89.3]	56.9±13.1	[30.3, 91.0]



(a) $\mathbf{R}_8$ (b) $\mathbf{C}$ (c) Correct $\mathbf{R}_9$ (d) Incorrect $\mathbf{R}_9$



(e) $\mathbf{R}_2$ (f) $\mathbf{C}$ (g) Correct $\mathbf{R}_3$ (h) Incorrect $\mathbf{R}_3$

Fig. 10. Impact of parameters τ, ϵ, η, and $\mathcal{E}_{\min}$ on 3D EBUS segmentation (case 21405.116, sequence 3, left PA). Each row depicts previously segmented frame $\mathbf{R}_{l-1}$ (done correctly), contour $\mathbf{C}$ for initiating the geodesic level-set process, and a pair of incorrect and correct segmentation results for $\mathbf{R}_l$. The top row considers frame 9 with: (c) $\tau = 0.10$ and (d) $\tau = 0.25$. As shown in (d), the large time step caused the process to eliminate the ROI altogether; the same result occurred when $\epsilon = 16$ and $\mathcal{E}_{\min} = 0.001$. For the bottom row (frame 3), $\mathcal{E}_{\min} = 0.02$ leads to correct result (g). $\mathcal{E}_{\min} = 0.64$, however, leads to serious overgrowth, as shown in (h); a similar oversegmentation occurs for $\epsilon = 0.0$ or $\eta = 0.0$.

REFERENCES

[1] X. Zang, "EBUS/MDCT fusion for image-guided bronchoscopy," Ph.D. dissertation, The Pennsylvania State University, School of Electrical Engineering and Computer Science, 2015.

[2] H. J. Johnson, M. M. McCormick, L. Ibánez *et al.*, *The ITK Software Guide*. Kitware, 2014.