SCIENCE CHINA
Information Sciences

• LETTER •

November 2021, Vol. 64 219203:1–219203:2
https://doi.org/10.1007/s11432-020-2978-6

Center tracking for healthy and diseased cardia

Baiquan SU[1], Yunlai TENG[1], Zehao WANG[1], Yida HU[2], Shaolong KUANG[3], Ye WANG[4], Tingting ZHANG[5], Jie TANG[6*], Wei YAO[4*] & Ye ZONG[5*]

[1]*Medical Robotics Laboratory, School of Automation, Beijing University of Posts and Telecommunications, Beijing* 100876*, China;*
[2]*Brigham and Women's Hospital, Harvard Medical School, Boston* 02115*, USA;*
[3]*Robotics and Microsystems Center, School of Mechanical Engineering, Soochow University, Suzhou* 215021*, China;*
[4]*Gastroenterology Department, Peking University Third Hospital, Beijing* 100191*, China;*
[5]*Department of Gastroenterology, Beijing Friendship Hospital, Capital Medical University, Beijing* 100050*, China;*
[6]*Neurosurgery Department, Xuanwu Hospital, Capital Medical University, Beijing* 100053*, China*

Received 25 February 2020/Revised 30 April 2020/Accepted 29 June 2020/Published online 15 January 2021

Citation Su B Q, Teng Y L, Wang Z H, et al. Center tracking for healthy and diseased cardia. Sci China Inf Sci, 2021, 64(11): 219203, https://doi.org/10.1007/s11432-020-2978-6

Dear editor,

Autonomous gastrointestinal robots are the future form of traditional manually operated endoscopes for reducing adverse events, such as perforation of the gastrointestinal tract [1–3]. Cardia center detection is a key task for gastrointestinal robots to autonomously enter the stomach to fulfill inspection and therapeutic operations. However, none of the existing studies addresses the problem of finding the center of the cardia, i.e., an image guidance task serving the gastrointestinal robots for passing the cardia.

Detection of the cardia center is fulfilled by using the hybrid maximum pixel grayscale average variance algorithm and the reference point increment dilation method. An appropriate threshold is found to delineate the outline of the cardia, and then the dilation step is used to obtain higher accuracy and avoid big error provided by the modified Otsu's method for the scenarios like the cases in subfigure D in Figure A1 in Appendix A. Also, an appropriate dilation method is selected to obtain real-time image processing speed. The pipeline of the proposed method for finding the cardia center is shown in Figure A1, where all results are first obtained here.

Assume that a digital image is a matrix with $\mathscr{M}$ pixels by $\mathscr{N}$ pixels, and the pixels of the image are represented in $L \in N$ gray levels. The quantity of the pixels at the level i, $i \in [1, 2, \ldots, L]$ is denoted by n_i and the total number of the pixels of the image is denoted by $N = n_1 + n_2 + \cdots + n_L$. The probability at the level i is

$$p_i = n_i/N,\ p_i \leqslant 0, \quad \sum_{i=1}^{L \in N} p_i = 1. \tag{1}$$

Now suppose that we dichotomize the pixels into two classes C_0 and C_1 (background and objects, or vice versa) by a threshold at the level $k \in [1, 2, \ldots, L]$. C_0 denotes pixels with the levels $[1, 2, \ldots, k]$ and C_1 denotes the pixels with the levels $[k+1, k+2, \ldots, L]$. The occurrence probabilities of the classes C_0 and C_1 are given by

$$\omega_0 = \Pr(C_0) = \sum_{i=1}^{k} p_i = \omega(k), \tag{2}$$

and

$$\omega_1 = \Pr(C_1) = \sum_{i=k+1}^{L} p_i = 1 - \omega(k), \tag{3}$$

where $\omega(k)$ denotes the zeroth-order cumulative moments of the histogram up to the k-th level, and the probabilities of mean levels of the classes C_0 and C_1 are given by

$$\mu_0 = \sum_{i=1}^{k} i\Pr(i|C_0) = \sum_{i=1}^{k} ip_i/\omega_0 = \mu(k)/\omega(k), \tag{4}$$

and

$$\mu_1 = \sum_{i=k+1}^{L} i\Pr(i|C_1) = \sum_{i=k+1}^{L} ip_i/\omega_1 = \frac{\mu_T - \mu(k)}{1 - \omega(k)}, \tag{5}$$

where $\mu(k)$ denotes the first-order cumulative moments of the histogram up to the k-th level and is defined as

$$\mu(k) = \sum_{i=1}^{k} ip_i. \tag{6}$$

The total mean level of the original image is

$$\mu_T = \mu(L) = \sum_{i=1}^{L} ip_i. \tag{7}$$

The global optimal threshold is determined by maximizing the between-class variance in the classical Otsu's

* Corresponding author (email: tangjiett@163.com, ywei@yeah.net, zongye@medmail.com.cn)

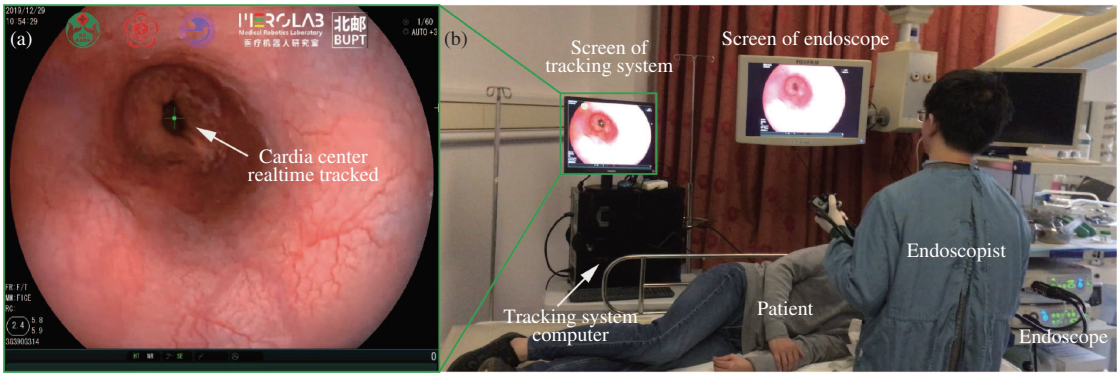

Figure 1 (Color online) Setup for real-time tracking of the cardia in a practical endoscopic screening procedure.

method, while in this study, the maximum pixel grayscale average variance [4] is employed to obtain the global optimal threshold. The pixel grayscale average variance between two areas is defined as the distance, i.e.,

$$U_i = \mu_T / N. \tag{8}$$

In the following, we present the procedure for finding the segmentation threshold which is denoted by Th_0. Firstly, the points with the grayscale value 0 or 255 are excluded from the following procedures. Secondly, an initial value is assigned to $Th_0 \in [0, 255]$. Thirdly, the pixel grayscale average variance S between target area and background area is defined as

$$S = \frac{(Th_0 - U_0)(U_1 - Th_0)}{(U_1 - U_0)^2}, \tag{9}$$

where U_0 and U_1 are obtained via (8). Fourthly, S is calculated according to (9) accompanying with the initial threshold Th_0 and the increment of Th_0 with the step length 1.0, and we obtain a pixel grayscale average variance set $\mathcal{S}$, i.e., $S \in \mathcal{S}$. Then, find the maximal value from the obtained set $\mathcal{S}$, i.e.,

$$S_{\max} = \max_{i \in K} S_i, \tag{10}$$

where K is the number of the elements of $\mathcal{S}$. Finally, we find the threshold $T\hat{h}_0$ corresponding to $S_{\max}$ via (10). Then the dilatation procedure, which is presented in Appendix B1, and the algorithm for calculation of cardia center, which is given in Appendix C1, are used to find the resultant cardia center.

Four cases of cardia shown in Figure D1 are selected to demonstrate the improvement of the detection time of the cardia center with the RPI method than the conventional dilatation method. Further, the position error of cardia center detection for various conditions is analyzed, and the errors are 6.58, 5.09, and 2.69 pixels between the center of ground truth and the center by the algorithm for side view closed orifice cardia, open orifice cardia, and FICE imaged cardia in Figure E2. Further, the proposed method is compared with seven algorithms to show its advantage, and the comparison is shown in Figure E1.

Four typical conditions of the normal healthy cardia are selected to demonstrate the effectiveness of the method, i.e., the simple and complicated folds, and the small and large open orifices. Also, the performance of the algorithm is verified by four typical lesions surrounding the cardia including reflux esophagitis, carcinoma of the gastric cardia, cardiac mucosal laceration syndrome and Barrett's esophagus. The detected center positions of the normal and diseased cardias are shown in Figures F1 and G1, respectively.

Video tracking experiment is also implemented. The cardia center tracking for the offline video is shown in Figure H1. We apply the method to a real endoscopy procedure to test the feasibility of the cardia center detection algorithm. We randomly select an endoscopy screening procedure on a thirty-six year-old male patient in the Sixth Examination Room, the Gastroenterology Department, Peking University Third Hospital. The cardia of the patient is healthy. The setup of the realtime cardia center tracking system is shown in Figure 1. The real-time tracking experiment is shown in Figure I1.

Conclusion. In this study, we develop an approach, i.e., the hybrid maximum pixel grayscale average variance and reference point increment dilation method, to find the center of the cardia. The task is one of the indispensable tasks for image guidance of an autonomous gastrointestinal robot to pass from the esophagus to the stomach through the cardia. The proposed method is generally effective for both normal cardia and diseased cardia neighbored with various kinds of common lesions.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 91748103, 61573208), Beijing Natural Science Foundation (Grant No. Z170001), and China Postdoctoral Science Foundation (Grant Nos. 2014M560985, 2015T80078). The authors gratefully acknowledge the insightful and very constructive comments from the anonymous reviewers.

Supporting information Appendixes A–I. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Paspatis G, Dumonceau J M, Barthet M, et al. Diagnosis and management of iatrogenic endoscopic perforations: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement. Endoscopy, 2014, 46: 693–711
2 Ben-Menachem T, Decker G A, Early D S, et al. Adverse events of upper GI endoscopy. Gastrointestinal Endoscopy, 2012, 76: 707–718
3 Yang G Z, Cambias J, Cleary K, et al. Medical robotics-regulatory, ethical, and legal considerations for increasing levels of autonomy. Sci Robot, 2017, 2: eaam8638
4 Sun H, Ma Y, Wu H, et al. An improved Otsu's method for CT image boundary contour extraction. In: Proceedings of 2016 IEEE International Conference on Imaging Systems and Techniques (IST), Chania, 2016. 493–497