A Capsule Robot Attitude Transformation Perception Method based on Intestinal Fold Features

Chengcheng Yan[1], Yanping Hu[2]

School of Mechanical Engineering, Dalian University of Technology, Dalian 116024, PR China
ccyan_dlut[at]163.com

Abstract: ***As the world medical level of ascension, the traditional endoscopy has can't satisfy people's needs, using tiny robots enter human body bowel check is becoming a new way of examination. Active robot has become a research hotspot due to its controllable attitude. In this paper, an attitude localization method based on the features of intestinal folds is proposed, and the contour of intestinal tract is obtained by light compensation and adaptive threshold segmentation. The contour HU moment is used for coarse matching, and the shape context is used for fine matching to obtain the matching relationship of contour point pairs. The constraint equation is constructed from the contour centroid and the minimum enclose rectangular centroid angle to limit the search range of point pairs. Finally, the attitude transformation matrix of the robot is solved by using the optimal matching relation and multi-view geometric constraints.***

Keywords: Fold characteristics; HU moment; Shape context; Attitude recognition; Active robot

1. Introduction

The human intestines are prone to various diseases, but the existing endoscopy is not only complicated in procedures, but also brings great pain to patients.Using capsule robot to enter the human body for examination has become a new mainstream way [1-2].At present, capsule robots can be divided into two types: active and passive. The passive type mostly relies on the force between itself and the intestine to walk passively. It is easy to miss the diseased area and its position and posture are not controlled [3].Although the active capsule robot can realize its own attitude positioning, most of the existing active capsule robots adopt the mode of magnetic sensor array, which is extremely dependent on the hardware conditions. However, the visual-based pose localization method is limited by the intestinal environment, such as the difficulty in extracting feature points. How to design a targeted algorithm based on intestinal characteristics has become the main task at present.

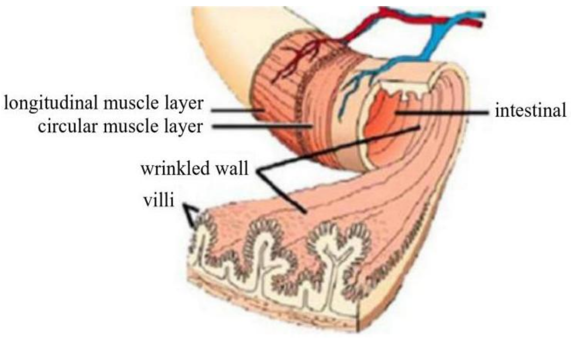


Figure 1: Structure of human intestinal wall[4]

The structure of human intestinal tract is shown in Figure1. The villi and wrinkled wall features are the most abundant inside the intestinal wall, which are reflected in the point features and contour features respectively in the image. Due to the uniform color of the intestinal wall and the lack of gray difference, point features are often difficult to extract in the absence of uniform illumination. In contrast, the contour features are richer and involve more pixels than point features, and have better robustness. Therefore, this paper selects the intestinal plica feature as the visual location feature, and on this basis, proposes a method to realize the self-pose recognition of capsule robot based on the wireless transmission image.

The method is divided into three modules: image processing module, contour matching module and attitude solving module. In the image processing module, Gaussian filter is used to remove image noise and illumination compensation is used to reduce the influence of uneven illumination. Based on this, adaptive threshold is used to segment and extract contour edge. The contour matching module uses shape histogram to describe the spatial position distribution relationship of the pixel points distributed on the contour, and then uses chi-square distribution cost function to match contour point pairs, and abstractly transforms contour matching into binary graph matching with weights. The attitude solving module uses the obtained contour matching relation and combines with the multi-view geometric constraints to solve the transformation matrix.

2. Intestinal feature information extraction

In human intestinal wall, intestinal fold is the most intuitive feature, and the fold can be represented as edge information by contour extraction, which can be represented as a collection of discrete points in digital images. Compared with a single pixel-level feature, edges reflect more feature information and do not change with the change of shooting position, which is an important method for image description [5].

2.1 Uneven illumination compensation

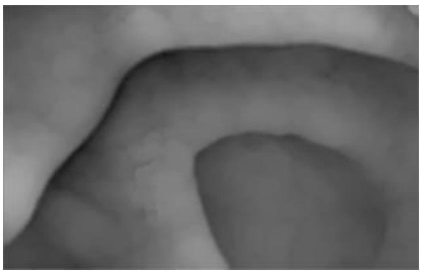
The situation in human intestinal is complex and changeable. Under the influence of uneven light, the information of intestinal fold is weakened and the image edge information is missing. Therefore, it is necessary to make light compensation for the obtained image before extracting the information of intestinal fold. Existing illumination compensation methods can be roughly divided into three categories: gray scale change method based on gray histogram, homomorphic filtering method based on reflection model, and Retinex frequency change method [6-8]. The homomorphic filtering method based on reflection model and Retinex frequency change method both convert the image to logarithmic domain processing, and need to adjust the filtering function according to the specific image, which cannot be self-adaptive. The grayscale change method based on grayscale histogram can enhance the image while retaining the original grayscale gradient information, which is more consistent with human visual characteristics. In order to make better use of the local gray level of each area of the image, this paper makes light compensation for the obtained image blocks in different regions. The specific methods are as follows:

- Calculate the overall average gray scale of the intestinal image, marked as *GrayScale*
- Divide the image into $N \times M$ blocks according to the preset blockSize, calculate the average grayscale of each block, and get the $N \times M$ average grayscale matrix M
- Calculate the difference matrix between the average gray matrix M and the average gray GrayScale, marked as E
- Bicubic interpolation matrix E,restore to the same gray scale compensation matrix C as the original image size
- The compensated image $I=R-C$

2.2 Contour extraction

The obvious folds contour in the intestine can be obtained by adaptive threshold segmentation after light compensation. In order to simplify the following pose solving part, this paper further defines the contour area threshold to screen out the main folds.

(a) The intestinal images



(b) illumination-compensated image

(c) Contour extraction

Figure 2: The intestinal images

3. Shape Context (SC) matching contours

In order to match the contour point pairs to solve the posture transformation matrix, the correspondence between the contours of the folds in the intestine needs to be obtained first. In this paper, contour matching is divided into two parts: coarse matching and fine matching. Coarse matching completes the overall matching between contour pairs, and fine matching uses the shape similarity of contour points as the criterion to match contour points.

3.1 Contour coarse matching

There are many mature schemes for overall matching between contours, which can be roughly divided into two types: contour curve-based matching and contour-enclosed area matching. The description method based on contour curve has better ability to describe the curve, but the description method based on contour region takes less time. In the rough matching process of intestinal contour, we mainly consider the execution efficiency of the algorithm. Therefore, Hu moment is constructed for the obtained contour in this paper to meet the requirement that the geometric moment of the same contour remains unchanged under different shooting angles. The moment of the contour can be obtained by using the HuMoments function in *OpenCV*.

3.2 Contour fine matching

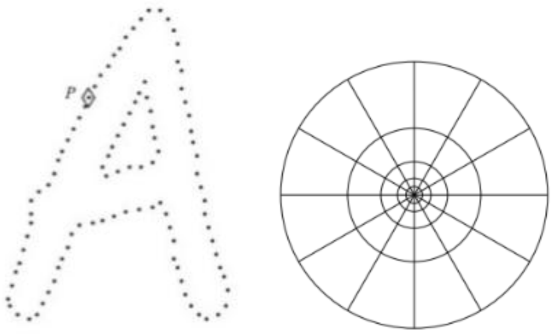
Shape context is a method proposed by Belongie et al. [9] to describe curves according to the spatial distribution of contour points from a statistical point of view. The shape histogram at the description point represents the spatial position relationship between the description point and other contour points, and takes into account both the local information of the description point and the global information of the overall contour.The method of calculating the shape histogram at the description point is as follows: establish a polar coordinate system with the description point as the origin, divide it into fan-shaped areas at equal angles, further divide each fan-shaped area into a region by equal logarithmic distance, and finally caculate the shape histogram of each area, where the shape histogram is defined as follows:

$$h_i(k) = \#\{(p - q_i) \in bin(k) \mid p \in P, p \neq q_i\} \quad (1)$$

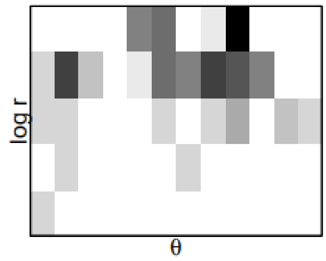
Where bin is defined in log-polar space.

Paper ID: SR21121141514 DOI: 10.21275/SR21121141514

International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2019): 7.583



(a)Set of sampling points (b) log-polar space



(c) Shape histogram at point *P*

Figure 2: Shape Context

For points C_i and C_j on any two contours, they themselves obey the $\chi 2$distribution, and the matching cost function is:

$$C_{ij} = C(p_i, q_j) = \frac{1}{2}\sum_{k=1}^{K}\frac{[h_i(k)-h_j(k)]^2}{h_i(k)+h_j(k)} \qquad (2)$$

Where $h_i(k)$ and $h_j(k)$ are the size of the shape histogram of the point and in the k-th block, C_{ij} representing the matching cost between the point and. The higher the matching cost is, the more dissimilar the shape is. The goal of contour matching is to make as many point pairs on the contour matched as possible. Further considering the shape similarity between matching points, it can be abstractly described as binary graph matching with weights, where the weight is the matching degree between point pairs. The extended *Kuhn-Munkres* algorithm of Hungarian algorithm is used, and its time complexity is $O(n^3)$.

3.3 Contour point pair constraint

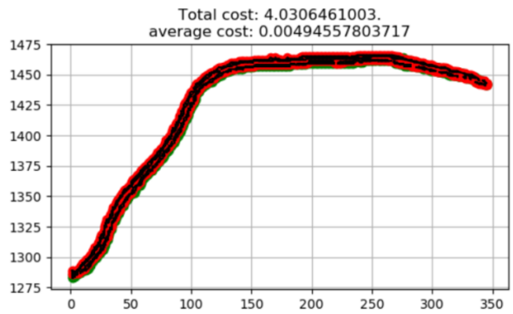
The traditional shape histogram matching between point pairs adopts the method of violent matching. For a point located on the contour, all possible matching points are traversed on another contour. Considering that there is a certain transformation relationship between two corresponding contours, the search range of point pair can be reduced by the constraint equation from the perspective of the whole contours.

Let *R* and *S* be the shooting results of the same intestinal fold obtained by rough matching under different viewing angles respectively. Point r_0 and point s_0 are the centroid of the contour *R* and *S*, the distance between r^0 and s^0 represents the movement translation $t=(tx,ty)^T$ of the overall contour.After de-centrifying contour *R* and contour *S* respectively, the existing rotation changes were considered.Point r^1 and point s^1 are the centroid of the smallest enclosing rectangle of contour *R* and contour *S* respectively. Its position remains unchanged when the image is scaled. The angle between vectors r^0r^1 and s^0s^1 represents the rotation angle θ of the overall contour.In consideration of the actual observation error, the threshold radius *ra* is defined, and the point set within the radius *ra* of each contour point is taken as the potential matching point set, and *ra* can be adjusted according to the actual situation.

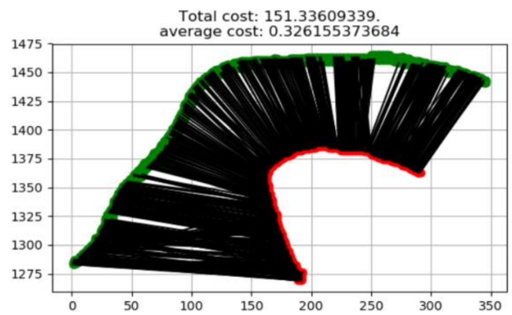
$$\begin{pmatrix} x_{(1)} \\ y_{(1)} \\ 1 \end{pmatrix} = \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_{(2)} \\ y_{(2)} \\ 1 \end{pmatrix} + \begin{pmatrix} \delta_x \\ \delta_y \\ 1 \end{pmatrix} \qquad (3)$$

where $R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, $t = \begin{pmatrix} t_x \\ t_y \end{pmatrix}$, $\delta_x{}^2 + \delta_y{}^2 \le ra^2$

Figure3 shows the partially extracted contour matching results. The lower the average matching cost between point pairs, the more similar the context shape is.



(a) matching contour



(b) dismatched contour

Figure 3: Cost of contour pair matching (partial images)

4. Attitude solution

After the matching relationship between the contour point pairs has been obtained, the recovery of the camera's front and back motion can be considered as a 2D-2D pose transformation problem, which can be solved by combining the knowledge of polar oppostion geometry.The polar geometric constraints [10] indicate that the same physical space point Q is on the same plane as the points Q_1 and Q_2 on the normalized plane before and after the camera attitude transformation, and the constraint equation is:

$$Q_2^T E Q_1 = 0 \quad (4)$$

Where E is called the essential matrix, Q_1 and Q_2 are the coordinates of the point on the normalized plane of the camera respectively, $Q_1 = K^{-1}q_1$, $Q_2 = K^{-1}q_2$, K is the internal parameter matrix of the camera, and q_1 and q_2 are pixel coordinates respectively.The scale equivalence ensures that the essential matrix E has five degrees of freedom and can be solved by at least five pairs of calibration points. However, in practical application, in order to ensure the computational efficiency, only the linear property of the polar constraint is usually considered and the eight-point method is used to solve the matrix[11]. *RANSAC* (random sampling consistency) is used to optimize the matching relationship of the obtained contour points, and the essential matrix E can be obtained by using the FindEssntialMatrix function in *OpenCV*. After *SVD* decomposition, the rotation matrix R and the translation vector t are obtained.

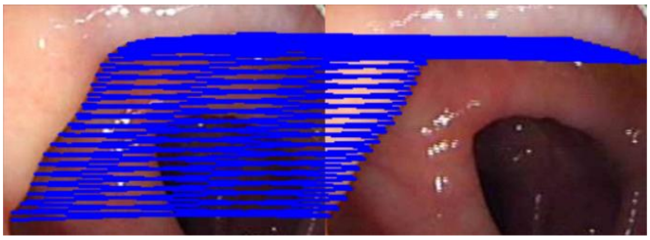


Figure 4: Contour point pair matching relationship

5. Summary

Aiming at the random and uncontrolled attitude of active robots in the human intestine, this paper proposes a new attitude transformation method of capsule robot based on the rich plica feature sensing in the intestine, which lays a foundation for the active closed-loop control of the robot.After analyzing the internal features of the intestine, this method extracts the most obvious folds as edge features.On this basis, the Hu moments are constructed to coarsely match contour pairs, and the constraint equations are constructed from the overall contour to reduce the calculation complexity of traditional shape context.After obtaining the matching relationship between the point pairs on the contour, combined with the epipolar geometric constraints, the essential matrix is solved by the eight-point method, and finally the camera attitude change rotation matrix R and movement vector t are obtained.

References

[1] CHAN J,CHAO A,CHEUNG V,et al."Gastrointestinal disease burden and mortality:A public hospital-based study from 2005 to 2014",Journal of Gastroenterology and Hepatology,34.124-131,2018

[2] CARAVACAAM,PIESTUNR."Single multimode fiberendoscope",Optics Express,25.1656-1665,2017

[3] LIENGS,WUMS,CHENCN,et al."Feasibility and safety of a novel magnetic-assisted capsule endoscope system in a preliminary examination for upper gastrointestinal tract",Surgical Endoscopy,32.1937-1944,2017

[4] Hu Haiyan."Research on semi-autonmous colonoscopic robot system",Harbin Instituteof Technology.1-127,2011

[5] Cai H,Zhu F."Shape matching method based on convex hull and multiscale integral features under affine transformation",Journal of Computer-Aided Design&Computer Graphics,29.269-278,2017

[6] YANG Ching-chung."Image enhancement by modified contrast-stretching manipulation",Optics&Laser Technology,38.196-201,2006

[7] Kim K,Kim S,Kim K S."Effective image enhancement techniques for fog-affected indoor and outdoor images",IET Image Processing,12.465-471,2018

[8] Land E H."The Retinex theory of color vision",Scientific American,237.108-128,1977

[9] Belongie S,Malik J,Puzicha J."Shape matching and object recognition using shape contexts",IEEE Transcations on Pattern Analysis and Maching Intelligence,24.509-522,2002

[10] RICHARD HARTLEY A.Multiple view geometry in computer vision(2nd Edition),Cambridge University Press,Cambridge,2004

[11] HARTLEY R.I."In defense of the eight-point algorithm",IEEE Transactions on Pattern Analysis and Maching Intelligence,19.580-593,1997