The vision guidance and image processing of AGV

Tongqing Feng and Bin Jiao

Electrical Engineering College, Shanghai Dian Ji University, No. 300 blooms Road, Pudong New Area, Shanghai, China.
Email: 780862201@qq.com

Abstract. Firstly, the principle of AGV vision guidance is introduced and the deviation and deflection angle are measured by image coordinate system. The visual guidance image processing platform is introduced. In view of the fact that the AGV guidance image contains more noise, the image has already been smoothed by a statistical sorting. By using AGV sampling way to obtain image guidance, because the image has the best and different threshold segmentation points. In view of this situation, the method of two-dimensional maximum entropy image segmentation is used to solve the problem. We extract the foreground image in the target band by calculating the contour area method and obtain the centre line with the least square fitting algorithm. With the help of image and physical coordinates, we can obtain the guidance information.

1. Introduction
Visual navigation technology is a kind of advanced guidance mode, it has the advantages of intelligence, high reliability and flexibility. It can capture around more complete information, so it has wide application prospect in AGV navigation system. At present, the more common application is to install the camera on the AGV and use the partial vision navigation method. Through the on-board computer, the high-level decision-making tasks such as path identification and tracking control are completed. The production workshop of vision navigation AGV has less illumination change and better flatness of ground. In this condition, stability and real-time are two main performance indexes of vision system. Therefore, the faster the image processing speed is, the shorter the AGV will obtain the path information. Therefore, to improve the performance of AGV, it is necessary to establish an efficient image processing system.

2. Principle of Visual Navigation Technology
AGV vision guidance uptake the pavement images by the dynamic guiding camera. The pavement image collection is three-dimensional scene in a 2D space perspective [1]. It needs the inverse process to identify the information of the path in the process of machine vision and realize the measurement of partial difference and deflection angle. At present, the AGV vision system usually uses the corresponding point calibration method to obtain the path information of the image. The image coordinate system is one level of coordinates of the vision system [2]: It is divided into image pixel coordinate system and image physical coordinate system. The image pixel coordinate system is fixed on the image in pixels of the Cartesian coordinates, the origin is located at the upper left corner of the image. The unit of the image physical coordinates is millimetre and its origin is the intersection point and the image plane of the camera axis.

Firstly, we establish the image pixel coordinate system of $\mu-o-V$ and image physics coordinate system of $x-c-y$ image pixel coordinate system and the reference coordinate system is as shown in Figure 1.
Figure 1. Reference coordinate system

AGV travels along the Y axis negative direction, its trajectory is divided into straight line and arc segments. It is because that the ribbon width is 6cm, the camera view height is about 15cm and the arc track radius is larger than 1m, so the camera view range is relatively small. In order to simplify the problem, we can take the arc in the field of view as a straight line approximately. The deviation $e$ is the abscissa value of the intersection point between the fitted line and the X axis and the angle of departure $\theta$ is the angle between the fitted line and the positive axis of the Y axis.

3. The Information Processing of Guided Image

3.1. Two-dimensional Maximum Threshold Segmentation.
The two-dimensional maximum entropy threshold [3] method introduces the concept of entropy of the information theory. The image can be regarded as a two-dimensional gray function, which contains $N \times N$ pixels and the gray value of pixel is divided into $L$ grades. Firstly, the region gray of the original image is obtained and the target pixel and its adjacent pixels are used as the template. In the original image, the corresponding coordinate of the gray value of the pixel and its region gray mean constitute the data pair $(i, j)$. We set $n_{i,j}$ as the number of pixels of the point gray $i$ and the region gray $j$ and $p_{i,j}$ is the probability density. Therefore, we can have the Formula (1) as follow:

$$p_{i,j} = \frac{n_{i,j}}{N \times N}$$  \hspace{1cm} (1)

As shown in Figure 2, we take the point gray value and the region gray mean value as the coordinate axis to establish the two dimensional image of gray distribution[4].
Figure 2. The two-dimensional segmentation of image

If the segmentation threshold is located \((s, t)\), the two-dimensional image can be divided into four regions, in which the target pixels are distributed in the A region, the background pixels are distributed in the B region. The C region contains the boundary pixel points and the D region contains the noise signals. Therefore, in order to make the image target and background have good segmentation effect, the best entropy threshold should be determined by two-dimensional maximum entropy method in A and B regions. In order to make the partition entropy more additive, the posterior probabilities of A and B regions are normalized respectively:

\[
p_A = \sum_{i} \sum_{j} p_{i,j} \quad i = 1, 2, \ldots, s; \quad j = 1, 2, \ldots, t
\]

\[
p_B = \sum_{i} \sum_{j} p_{i,j} \quad i = s+1, s+2, \ldots, L; \quad j = t+1, t+2, \ldots, L
\]

The discrete two-dimensional entropy is defined as the Formula (4) as follow:

\[
H(A) = -\sum_{i} \sum_{j} p_{i,j} \log p_{i,j}
\]

Then the two-dimensional entropy of the A region and the B region are respectively shown in the Formula (5).

\[
H(A) = -\sum_{i} \sum_{j} \left( p_{i,j} / P_A \right) \log \left( p_{i,j} / P_A \right) \\
= -(1/P_A) \sum_{i} \sum_{j} \left( p_{i,j} \log p_{i,j} - p_{i,j} \log P_A \right) \\
= \log P_A + H_A / P_A
\]

\(H_A\) is shown in Equation (6):

\[
H_A = -\sum_{i} \sum_{j} p_{i,j} \log p_{i,j} \quad i = 1, 2, \ldots, s; \quad j = 1, 2, \ldots, t
\]

\[
H(B) = -\sum_{i} \sum_{j} \left( p_{i,j} / P_B \right) \log \left( p_{i,j} / P_B \right) = -(1/P_B) \sum_{i} \sum_{j} \left( p_{i,j} \log p_{i,j} - p_{i,j} \log P_B \right)
\]

The final \(H_B\) is shown in Equation (8):
\[ H_B = - \sum_i\sum_j p_{i,j} \lg p_{i,j} \quad i = s+1, s+2, \ldots, L; j = t+1, t+2, \ldots, L \]  

\[ P_B = 1 - P_A \quad , \quad H_B = H_L - H_A \]  

Form Equation (9), we have:

\[ H_B = \lg(1 - P_A) + (H_L - H_A) / (1 - P_A) \]  

The discriminant function of entropy is defined as:

\[ \phi(s,t) = H(A) + H(B) \]  

The best threshold is shown in Equation (12):

\[ \theta^{**} = \max \{ \theta(s,t) \} \]  

According to the above algorithm, the two-dimensional maximum entropy segmentation results are shown in Figure 3.

![Figure 3. Threshold segmentation image](image)

It can be seen that the two-dimensional maximum entropy threshold segmentation algorithm is more accurate for the calculation of the optimal threshold and it is suitable for processing the image in this paper. So, the foreground and the background of the image are separated completely and clearly.

3.2. Edge Detection Based on Canny Algorithm

After threshold segmentation, the guiding image hope that through an edge detection algorithm [5] to extract the image edge features of the guiding ribbon and sundries in the foreground object. The edge of the image corresponds to the local maximum of the first derivative and the detection of the weak edge of the image is easily disturbed by the noise signal. Therefore, we introduce canny edge detection [6] which uses double thresholds to detect strong edges and weak edges respectively.

The original input image is \( f(x, y) \). Firstly, the quasi Gauss function is used as a smoothing operation and the gradient of the smoothed \( g(x, y) \) is shown in Equation (13).

\[ \nabla g(x, y) = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial g}{\partial x} \\ \frac{\partial g}{\partial y} \end{bmatrix} \]  

By convolution operations, we have:
\[ \nabla g(x, y) = \nabla f(x, y) \otimes G(x, y) = f(x, y) \otimes \nabla G(x, y) \]  

(14)

Edge features are smoothed by the Gauss function and paid a fuzzy price. The width of the detected edge is increased and the fuzzy edge of the image is sharpened by NMS (Non-Maxima Suppression, NMS) [7]. The NMS method compares the gradient amplitudes of adjacent pixels to each other in the direction perpendicular to the edge and removes the gradient magnitude, which is smaller than the neighbourhood. According to this operation, the non-maxima of the gradient amplitude are removed and the edges become thin.

The NMS method suppresses the amplitude image and still contains many false edges caused by noise and fine texture, which can be removed by double threshold algorithm. Double threshold algorithm uses dual threshold \(T1\) and \(T2\) and \(T2 = 2T1\) on non-maximum suppression amplitude image. Two double threshold edge images \(G1[i, j]\) and \(G2[i, j]\) are obtained. The algorithm will constantly collect edges in \(G1[i, j]\) until all the gaps in the \(G2[i, j]\) are joined. Based on the above algorithm, canny edge detection results are obtained as shown in Figure 4.

![Figure 4. Edge detection of Canny](image)

4. Information Extraction of Path Guidance

According to the definition of the guidance parameters \(e\) and \(\theta\) and the fitting line \(L\) of the guidance path can reflect the path information more accurately. Therefore, the extraction of navigation information is realized by the aid of the image physics coordinate system and the equation of the fitting line in the image physical coordinate system is obtained. The least square method is used in the line fitting algorithm and the linear equation of the fitting line \(L\) in the image physical coordinate system is assumed to be shown [8]:

\[ y = ax + b \]  

(15)

The formula contains two undetermined parameters, \(a\) represents the intercept and \(b\) represents the slope. In this paper, 320 sets of data \((x_i, y_i), i=1,2,3,\ldots,320\) are obtained by averaging the coordinates of each row of pixels in the outline of the guided ribbon and the least squares method requires the weighted sum of squares of the deviation of the observed value \(y_i\) to be minimum:

\[ \sum_{i=1}^{320} \left[ y_i - (a + bx_i) \right]^2 \]  

(16)

The Equation (17) and Equation (18) is for the partial derivative of \(a, b\) respectively of above formula:

\[ \frac{\partial}{\partial a} \sum_{i=1}^{320} \left[ y_i - (a + bx_i) \right]^2 = -2i \sum_{i=1}^{320} (y_i - a - bx_i) = 0 \]  

(17)
\[
\frac{\partial}{\partial b} \sum_{i=1}^{320} \left[ y_i - (a + bx_i) \right]^2 = -2i = \sum_{i=1}^{320} (y_i - a - bx_i) \cdot x_i = 0 \tag{18}
\]

The Equation (19) can be obtained after sorting:

\[
\begin{align*}
320a + b \sum x_i &= \sum y_i \\
a \sum x_i^2 + b \sum x_i^2 &= \sum x_i y_i
\end{align*}
\tag{19}
\]

By solving the equation set, the best estimate of the parameters \(a\) and \(b\) of the fitted line can be obtained:

\[
\hat{a} = \frac{\sum x_i^2 (\sum y_i) - (\sum x_i)(\sum x_i y_i)}{320(\sum x_i^2) - (\sum x_i)^2} \tag{20}
\]

\[
\hat{b} = \frac{320(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{320(\sum x_i^2) - (\sum x_i)^2} \tag{21}
\]

According to the geometric relation \(\theta = \arctan(1/a)\), the \(e\) is shown in the Formula (22) by making \(y=0\).

\[
e = \frac{-b}{a} \tag{22}
\]

The results of path guidance information extraction are shown in Figure 5.

![Input image](image1.png) ![The image of linear fitting](image2.png)

**Figure 5.** The result of linear fitting

Therefore, we can have the final results that we want: \(e=17.763\) (unit: pixel), \(\theta=18.336^\circ\) (unit: angle).

5. Conclusion

Through the establishment of vision guided image processing platform, the camera provides original visual information. In the process of image processing, two-dimensional maximum threshold segmentation method, edge detection method and Canny least squares method make the original trajectory guidance information can be more clear and the deviation and deviation angle of the processing results more effective. It can effectively solve the problem of unstable operation and AGV track caused by other guidance. In general, we can determine the location of AGV, eliminate the error and improve the system stability and can effectively solve the actual operation of AGV. It can solve practical problems of the AGV.
6. References

[1] Peng W, Sun C and Zhang Z 2011 Linear pose estimation with a monocular vision system J. Chinese Journal of Scientific Instrument. 32 (5) 1126-31

[2] Mei Y C, Wang C L and Liao Q W 2009 Applied research of OpenCv-based camera calibration J. Computer Engineering & Design. 30(16) 3856-55

[3] Gao L, Yang S Y and Li H Q 2007 New unsupervised image segmentation via marker-based watershed J. Journal of Image & Graphics. 12(6) 1025-32.

[4] Luo Z, Liu H P, Hu X F and Xu W 2016 Research of Vision-Guided AGV Deviation-rectifying Algorithm J. Computer Simulation. 33 (1) 373-377

[5] Villagra J and Herrero P D 2012 a comparison of control techniques for robust docking maneuvers of an AGV J. IEEE Transactions on Control Systems Technology. 20(4) 1116-23.

[6] Jia Z, Balasuriya A and Challa S 2008 Sensor fusion-based visual target tracking for autonomous vehicles J. Artificial Life and Robotics. 12 317-328.

[7] Zhang J P, Lou P H, Qian X M and Wu X 2016 Research on precise positioning technology by multi-window and real-time measurement for visual navigation AGV J. Chinese Journal of Scientific Instrument. 37 1356-63

[8] Zhu L, Lin X and Wu W J 2016 Work piece positioning of industrial robot based on machine vision J. Journal of Chinese Computer Systems. 37(8) 1873-77