PROCEEDINGS OF SCIENCE

Vision-Based Object Fuzzy Recognition in Plain Industrial Environment by Contour Matching

Menglei Sang[1]
Department of Electrical Engineering, Zhejiang University
Hangzhou, Zhejiang P.R. China
E-mail: sangml@163.com

Wei Wei[2]
Department of Electrical Engineering, Zhejiang University
Hangzhou, Zhejiang P.R. China
E-mail: wwei@zju.edu.cn

Yonggang Peng[3]
Department of Electrical Engineering, Zhejiang University
Hangzhou, Zhejiang P.R. China
E-mail: pengyg@zju.edu.cn

The vision-based object recognition and location (VORL) technology can greatly increase the flexibility of production automation. To simplify some very complicated VORL systems while keeping the capacity of recogniting and locating many kinds of objects, the contour matching technology is introduced to the industrial VORL system consisting of object region detection, contour detection, contour matching, pose estimation and object location. To obtain the global optimal contour matching and decrease the computation complexity as much as possible, the dynamic programming with ordered contour points is adopted. Besides, based on the result of contour matching, the rotation angle can be estimated at a very little cost of computation complexity.

CENet2017
22-23 July, 2017
Shanghai, China

[1]Speaker
[2]This work was supported by the Significant Investigative Planned Project of Zhejiang Province of China (Project No.: 2017C01007).
[3]Correspongding Author

PoS(CENet2017)013

1.Introduction

The vision-based object recognition and location (VORL) technology can drastically increase the flexibility of production automation. With the imperious demand of real-time response and recoginition of many different objects in the field of industrial application, the industrial VORL system usually adopts template library trick which requires the control system possess strong computing power or offline classifier which needs a lot of data to model in advance [1]. However, in many occations, the industrial robot only needs to know the shape of the object and then much operation such as grasp can be done. In order to decrease the complexity of detecting the 3D shape in images, 2D contour of objects in images captured by fixed vision sensors can be used to implement the system of VORL. This article aims at proposing a novel online real-time vision-based object recognition and location algorithm frame based on contour matching, which could be used at many simple industrial sceneris and provide high recognition efficiency at a very low computation complexity.

2. Problem Specification

It can be assumed that the industrial robot would do the same operations on the objects with the same shape, such as picking and placing, and the objects with the same shape but different textures can be regarded as the same type of objects. Then the shape can be used to distinguish the objects. The vision system is required to belong to the eye-to-hand vesion system and vision sensors are positioned right above the working-platform where the objects would be shipped onto. The contour of one object extracted from the overhead view can express some primary information of its shape. According to the contour information of different objects, the industrial objects can be classfied and the industrial robot can do different operations on them. Besides, it should be noted that the problem solved by the algorithm frame described in this article is aiming at some universal industrial sceneries.

3. Contour Extraction

3.1 Contour Extraction Algorithm Analysis

The contour extraction lies in the core of the automatic object detection. Many contour extraction algorithms from complicated background feature high time complexity of more than one minute and can not fulfill the demand of real-time response. Considering the fact that in the industrial production line or some other industrail environment, the background of images that the vision sensors capture is changeless and relatively simple, then the background substraction method can be adopted to detect the object region. The output of the background substraction algorithm is just the images with monochrome background. The computation time of extracting contours from one monochrome static image containing object region is about several milliseconds and would not drag down the real-time performance of the whole VORL system [2]. Besides, when the objects that the industrial robot operates are moving at a low speed, their frame difference method can also be used to extract contours preliminarily and then some other contour extraction algorithms of monochrome background can be used to refine the initial contours. This article mainly describes the contour algorithm based on the background substraction.

3.2 Region Detection By Background Substraction

The background substraction is an important research branch in the computer vision field. Many algorithms have been designed to detect intruding regions and the computation complexity of some algorithms can fulfill the real-time response in the guarantee of accuracy, such as algorithms based on Mean and Variance Model, Gaussian Mixture Model, Fuzzy Model and so on [2]. Many of these background substraction algorithms make the preprocessed intensity images as the input. In order to intensify the boundary information of the intruding region, the intensity image with enhanced edge information is made as the input of the backgroud subground algorithm. Equation (3.1) describes the mixed input image, in which, I'(x,y) refers to the weighted sum of intensity and gradient of image I. α represents the weight of intensity in the sum, 1-α represents the gradient weight . In this article, simple background substraction based on the mean model is implemented to verify the feasibility of the proposed algorithm frame. The results can be seen in Figure 1.

$$I'(x,y)=\alpha\times I(x,y)+(1-\alpha)\times\sqrt{I_x^{\,2}+I_y^{\,2}} \quad (3.1)$$

Figure 1(a) shows the background image, Figure 1(b) shows image that contains objects, and Figure 1(d) is the result of background substraction.

(a) (b) (c)

Figure 1: The input and output images in the background substraction

3.3 Contour Extraction By Binary Image

The contours are a useful tool for shape analysis, object detection and recognition. The output images of background ground can be converted into the binary images and many algorithms of contour extraction from binary images with relatively low computation complexity can be used to extract contours. In order to provide convenience for the subsequent contour matching algorithm, here, the algorithm proposed in by Satoshi Suzuki which can ouput ordered contour points is adopted [3]. Figure 2(a) is the image containing possible object contours obtained from Figure 1(c) after deleting some too small contours, and Figure 2(b) is one template contour image.

(a) (b)

Figure 2: The contours extracted from images

4.Contour Matching

4.1 Contour Matching Algorithm Analysis

The contour matching problem mainly consists of two parts, choosing appropriate contour descriptor to construct the affinity matrix between two point sets and optimizing the global minimization. The contour descriptor usually meets the demand of rotation and scaling invariance. Considering the computation complexity and robustness, the mixed descriptor of relative distance and the included angle of two adjacent contour points to the contour centroid is established, which are respectively defined by Equations (4.1), (4.2) and (4.3) [5]. Besides, the assignment algorithms are often used to determine the best global matching, such as convex optimization, simulated annealing, linear integer programming, dynamic programming, etc. Considering the demand of real-time performance in available contour matching system and the performance of various assignment algorithms, the dynamic programming algorithm is adopted to solve this global optimization problem in one paper [4].

$$d_i=\frac{\sqrt{(x_i-x_c)^2+(y_i-y_c)^2}}{avg(d_i)} \tag{4.1}$$

$$a_i=\frac{\vec{P_{i-k}C}\cdot\vec{P_{i+k}C}}{|\vec{P_{i-k}C}||\vec{P_{i+k}C}|\times avg(a_i)} \tag{4.2}$$

$$c_i=\omega\times a_i+(1-\omega)\times d_i\,,\,w\in[0,1] \tag{4.3}$$

In Equation (4.1) and (4.2), C denotes the centroid of one contour, P_{i-k} and P_{i+k} are two adjacent contour points around point i.

4.2 Affine Matrix Construction

According to the descriptor in Equation (4.3), the descriptor vectors of two contours can be obtained, and the similarity of two points in these two contours, as also one element in the affinity matrix, can be defined in Equation (4.4) and (4.5).

$$c_{ij}=|c_i-c_j|\,,\,i=1,2,3,\ldots,n\,,\,j=1,2,3,\ldots,m \tag{4.4}$$

$$n\leq m \tag{4.5}$$

In Equation (4.4), n denotes the size of point set of the contour to be matehed, and m denotes the size of template contour point set. Equation (4.5) requests the point set of the template contour that should have much more points. Besides, in the contour matching process, the size of the template point set would be resized to be less than twice of the other contour point set.

4.3 Contour Matching Based on Dynamic Programming

The optimal matching between two contour point sets must preserve the circular order and the point set of contour obtained in Part 3.3 is in order exactly. Referred to Paper [4], dynamic programming can be used to optimize the matching cost while preserving the points matching order, which possesses a cubic computational complexity of O(n×m×(m-n+1)). And this optimization problem can be described by Equations (4.6), (4.7), (4.8), and (4.10).

$$min\ f=\sum_{i=1}^{n}\sum_{j=1}^{m}x_{ij}c_{ij} \tag{4.6}$$

$$x_{ij} = 1 \leftarrow if\ point\, i\ of\ one\, contour\ is\, mathing\ point\ j\ of\ the\, other\, contour\ , 0 \leftarrow otherwise \tag{4.7}$$

$$\sum_{j=1}^{m} x_{ij} = 1, i = 1,2,3,\ldots, n \tag{4.8}$$

$$\sum_{i=1}^{n} x_{ij} \leqslant 1, j = 1,2,3,\ldots, m \tag{4.9}$$

Based on the memory array used in the solving process of dynamic programming, the correspondence of points in two contour sets can be obtained by backward search. The mapping results can be shown by Figure 3.

(a) (b)

Figure 3: The points mapping images of contour matching

5. Similarity Judgement

The minimum sum cost value f of the matching points in two contours can be regarded as an indicator to reflect the similarity of two contours. However, the minimum sum cost is often related with the size of the contour to be matched, and the value defined in Equation (5.1) can be seen as a more effective similarity judgement index. Then one threshold value can be set to decide whether two contours should be regarded as the same type of contour and same kind object or not, which can be defined as Equation (5.2). In Figure 3, the matching index f' of Figure 1(c) with template Figure 2(b) are respectively 0.0558 and 0.0248.

$$f' = \frac{f_{min}}{n}, n \in N^{+} \tag{5.1}$$

$$is\, same = 1 \leftarrow f' < \epsilon, 0 \leftarrow otherwise \tag{5.2}$$

6. Pose Estimation and Disparity Computation

Since the platform where the industrial objects would be shipped onto is usually approximatively horizontal, then the key of pose estimation lies in the rotation angle estimation. In order to simplify the pose estimation, it can be assumed that the template contour lies in the reference position. According to the correspondence between the point sets of candidate contour and template contour, the rotation angle of candidate contour to template contour can be estimated by the average value of the rotation angle of two matching points in two contour sets, which is defined by Equations (6.1), (6.2) and (6.3). The final estimated value of the rotation angle can be obtained by Equation (6.4). In Equation (6.2), (6.3), point C and C' are respectively the centroid of candidate contour and template contour.

$$\beta_i = \arccos\left(\frac{v \cdot v'}{|v||v'|}\right) \tag{6.1}$$

$$v = (x_i - x_c, y_i - y_c) \tag{6.2}$$

$$v' = (x_j - x_{c'}, y_j - y_{c'}) \tag{6.3}$$

$$\bar{\beta} = \frac{1}{n} \sum_{i=1}^{n} \beta_i \tag{6.4}$$

According to the theory of stereo matching, as long as the disparity of the same point on one object has been obtained through two images captured by the binocular visual system, the location coordinate of that point in the camera physical coordinates can be fixed. The center of contour can be regarded as one representative point and used to obtain a disparity to estimate the coordinate of the object in the camera physical coordinates as well as in the robot's working coordinates.

7. Algorithm Frame Analysis

The algorithm frame of vision-based object recognition and the location described in this article highlights its the performance of online real-time recognition and location without constructing very complicated classifier or consuming valuable computation power of the industrial control computer.

7.1 Computation Complexity

The computation complexity of image preprocessing and the object region detection based on background substraction linearly depend on the size of images. The contour extraction from binary images has a polynomial computatin complexity of extracted points. The disparity computation of some special points is linearly relative with the width of images. And the computation complexity of pose estimation in this algorithm frame is linearly relative with the size of the contour point set. The most time-consuming phase is the contour maching. However, the optimization algorithm based on dynamic programming only has a cubic computational complexity, and appropriate sampling on the contours to be matched can effectively decrease the computation complexity of process solution. To sum up, the total computaion complexity of this algorithm frame is relatively low and can meet the demand of real-time response in the industrial control system.

7.2 Advantage and Disadvantage

The algorithm frame proposed herein exactly features the simplicity and on-line real-time performance. Although its frame can only be applied to some relatively simple industrial sceneries, its simplicity also signifies lower cost and simpler template. However, this algorithm frame mainly lies in its limited scope of applications and recognition of objects with similar shape.

8. Experiment Results

8.1 General Matching Results

Use the contour in Figure 2(b) as a template to match with some different contours, and the matching results can be described as the table below. It can be seen that the matching value in Table 1 is obviously greater than the matching value of contours in Figure 2(a) to be matched with the same contour template.

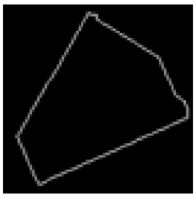	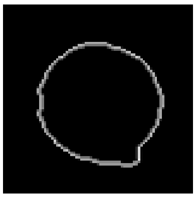	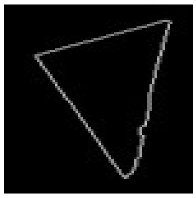	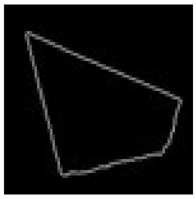
0.524	0.164	0.154	0.099

Table 1：Matching results of different contours with the same template contour

8.2 Computation Complexity

The computation complexity mainly lies in the matching of contour points, and the computation time is approximately 300ms for 76 points and 133 points in debug mode with the computer of 3.20GHz, 8GB RAM.

9. Conlusion

In this article, an algorithm frame of vision-based object recognition and location by using contour matching has been proposed, which can be applied in some simple industrail sceneries. is to combine many technologies especially contour matching in the implementation of an vision-based object recognitin and location system, and offer a simpler on-line real-time algorithm frame at the cost of losing the power of recogniting objects with the same shape. However, some more efforts shall be still devoted to improve robustness in the region detection and the subsequent processing of contour points, etc..

References

[1] Felzenszwalb P F, Girshick R B, McAllester D, et al. *Object detection with discriminatively trained part-based models*[J]. IEEE transactions on pattern analysis and machine intelligence, 2010, 32(9): 1627-1645.

[2] Sobral A, Vacavant A. *A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos*[J]. Computer Vision and Image Understanding, 2014, 122: 4-21.

[3] Satoshi Suzuki and others. *Topological structural analysis of digitized binary images by border following*. Computer Vision, Graphics, and Image Processing, 30(1):32–46, 1985.

[4] Oliveira F P M, Tavares J M R S. *Matching contours in images using curvature information and optimization based on dynamic programming*[J]. IEEE Latin America Transactions, 2009, 7(6).

[5] Amanatiadis A, Kaburlasos V G, Gasteratos A, et al. *Evaluation of shape descriptors for shape-based image retrieval*[J]. IET Image Processing, 2011, 5(5): 493-499.

[6] Estrada F J, Elder J H. *Multi-scale contour extraction based on natural image statistics*[C]//Computer Vision and Pattern Recognition Workshop, 2006. CVPRW'06. Conference on. IEEE, 2006: 183-183.