Object Recognition Using Polar-Exponential Grid Technique

Hai-yan Zhang
(Colledge of Computer Science & Technology, Harbin Engineering University, Harbin 150001)
E-mail: zhyfarry@0451.com

Dong-mu Wang
(Beijing Simulation Center, Beijing 100854)

Ke-ou Song
(Colledge of Computer Science & Technology, Harbin Engineering University, Harbin 150001)

Abstract

A new polar-exponential grid technique for object recognition is proposed in this paper. An object in the rectangular coordinate is mapped into the polar-log coordinate, and the closed profile of the object is transformed a one dimension curve. The structure of the curve has rotate, scale, translate invariance, which can be used in object recognition. Better recognition result using the structure invariant is got.

Keywords: object recognition, Polar-exponential grid technique, Polar-log coordinate transform

1.Introduction

Object recognition is an important part in computer vision and image processing. The pivotal step of the object recognition is extracting the invariant of rotation, scale and translation. Hu early put forward the moment[1], verified a series of basic characters for the moment and proved that the moment has rotation, scale and translation invariance property. Thereafter, a lot of invariant are advised, such as affine invariant[2], perspective invariant[3] and inflexion invariant[4]. Furthermore, the profile of the object is changed into one dimension curve in polar-log coordinate by polar-exponential sampling, and the profile invariant can be easily extracted with curve matching. Because profile invariant translates along u axis or v axis when the object rotate or scale, profile invariant ensures rotation, scale invariance by performing Fourier transform[5]. It spends much time. A new method for the object recognition is proposed in this paper, which uses object's structure invariant in polar-log coordinate to recognize object. The object's structure invariant has rotation, scale and translation invariance. This method applies to single object without rolling and wringing, and also be the base to three-dimension object recognition by aspect graph[6].

2.Polar-exponential grid technique

Polar-exponential grid technique was studied in 1979[7], hardware structure suited parallel processing is proposed in 1985. Polar-exponential grid technique was selected as a topic at international conference of robots and machine vision in 1989. Because of the character of clarity in center and blur in periphery, the transform itself is compressed largely. Combining polar-exponential grid technique and compression standard can get the great compression ratio. Because of the transform has rotation and scale invariance property; it can be used in object tracking[8]. It also gains better result in face check-up and face tracking[9][10].

Optic nerve cells in man's eyes' retina distribute asymmetry. Optic nerve cells distribute thick in center, whose image has the biggest differentiate ratio, and decrease in exponential rule with distance to center increase. When man watch object, the center is clear and blur in peripheral. By moving eyes and neck make the object is in center, we can observe it clearly. Polar-exponential grid technique simulated eyes' characteristic performs polar-exponential grid sampling and polar-log coordinate mapping. The mapping principle is describe in figure 1:

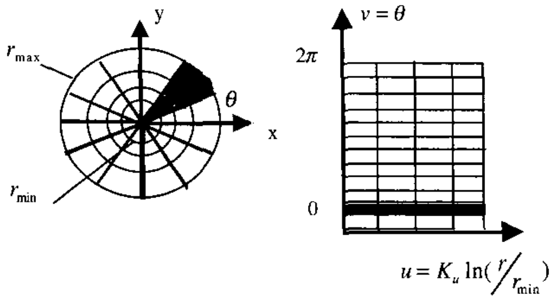


Figure 1. Polar-log coordinate transform

where, left image indicates that image in tangular coordinate is performed polar-exponential grid sampling, and right image indicates left image's equivalent plane afer it performs polar-log exponential transform. It makes out that grid cells close with center are smaller, and bigger and bigger with the distance to center increasing. Correspondingly, the image is blur and differentiate ratio decreases. Mapping expression is below:

$$u = k_u \ln(r/ r_{min}) \tag{1}$$

$$v = \theta = k_v arctg(y/x) \tag{2}$$

Where, k_u, k_v is distance differentiate ratio and angle differentiate ratio separately, in the paper, they are 40.23 and 20.4 separately. $r = \sqrt{x^2 + y^2}$ is the length between pixel and transform center, (x, y) is coordinate in rectangular coordinate. u, which is the horizontal axis in polar-log coordinate, is the length between pixel and transform center. v, which is the vertical axis in polar-log coordinate, is the angle between radius of pixel and horizontal axis. r_{min} is the least radius, and it is 10 pixel in the paper. r_{max} is the maximal radius, and it is defined the half of the minimal value between the image height and width.

Note: the center of polar-log coordinate transform is in the object's centroid; the traditional centroid arithmetic is used.

$$\bar{x} = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} x f(x,y)}{\sum_{i=1}^{M} \sum_{j=1}^{N} f(x,y)} \tag{3}$$

$$\bar{y} = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} y f(x,y)}{\sum_{i=1}^{M} \sum_{j=1}^{N} f(x,y)} \tag{4}$$

Where, *f(x,y)* is the binary image, and *M*, *N* is the image height and width separately.

3.Structure invariant

The structure invariant is defined in polar-log coordinate, which includes span of curve, ratio of curve area and distributing situation of curve.

3.1 Span of curve

An object in the rectangular coordinate is mapped into the polar-log coordinate, and the closed profile of the object is transformed a one dimension curve. Span of curve, k, is defined as difference between wave crest and trough. Wave crest and trough is the maximum u_{max} and minimum u_{min} in u axis separately. Span, wave crest and trough is defined in figure 2:


a. object b. span, wave crest and trough

Figure 2. Object span, wave crest and trough

The maximum pixel and the minimum pixel in u axis are (x, y), (x', y') separately before object rotate and scale. (kx_1, ky_1), (kx_1', ky_1') are the pixels after object rotate and scale. (x, y) and

(x_1, y_1), (x', y') and (x_1', y_1') satisfy formula (5) and (6):

$$\sqrt{x^2+y^2}=\sqrt{(x')^2+(y')^2} \tag{5}$$

$$\sqrt{x_1^2+y_1^2}=\sqrt{(x_1')^2+(y_1')^2} \tag{6}$$

If the coordinates in polar-log coordinate of (x, y) and (x', y') is $(u_{\max}, v)$, $(u_{\min}, v')$ separately, the coordinates is (u_0, v_0) and (u_1, v_1) after rotating and scale. Then:

$$\begin{aligned} u_0 &= k_u \ln(\sqrt{(kx_1)^2+(ky_1)^2}/r_{\min}) = k_u \ln(k\sqrt{x_1^2+y_1^2}/r_{\min}) \\ &= k_u \ln(k\sqrt{x^2+y^2}/r_{\min}) = k_u \ln k + k_u \ln(\sqrt{x^2+y^2}/r_{\min}) \\ &= k_u \ln k + u_{\max} \end{aligned} \tag{7}$$

$$\begin{aligned} u_1 &= k_u \ln(\sqrt{(kx_1')^2+(ky_1')^2}/r_{\min}) = k_u \ln(k\sqrt{(x_1')^2+(y_1')^2}/r_{\min}) \\ &= k_u \ln(k\sqrt{(x')^2+(y')^2}/r_{\min}) = k_u \ln k + k_u \ln(\sqrt{(x')^2+(y')^2}/r_{\min}) \\ &= k_u \ln k + u_{\min} \end{aligned} \tag{8}$$

Span before rotation and scale: $k = u_{\max} - u_{\min}$, if span after rotation and scale, then:

$$\begin{aligned} k' &= u_0 - u_1 = k_u \ln k + u_{\max} - (k_u \ln k + u_{\min}) \\ &= u_{\max} - u_{\min} = k \end{aligned} \tag{9}$$

It can be seen that span has rotation, scale and translation invariance.

3.2 Ratio of curve area

Ratio of curve area, ρ, is defined as below:

$$\rho = S/S_{all} \tag{10}$$

Where, S is the area between the wave crest and trough in u axis and S_{all} is the area of the rectangle whose length is the length of vertical axis and whose width is the span of curve. Then:

$$S_{all} = V \times (u_{\max} - u_{\min}) = V \times k = 2\pi k \tag{11}$$

The area said in the paper is the sum of pixel in the area.

It can be known from mapping theory that the size of image is fixed if the parameter is selected. It can be known from formula (9) that the value of k is invariable in spite of object rotation and scale, so the value of S_{all} is invariable. It can be known from mapping theory again that the profile curve of object in polar-log coordinate is invariable regardless of object rotation and scale, so the value of S is invariable. Thereby, ratio of area is invariable.

3.3 Distributing situation of curve

Curve is separated L regions in u axis, the regions is as below:

$$\begin{aligned} &[u_{\min}, u_{\min}+1/L), [u_{\min}+1/L, u_{\min}+2/L), \cdots, \\ &[u_{\max}-2/L, u_{\max}-1/L), [u_{\max}-1/L, u_{\max}] \end{aligned} \tag{12}$$

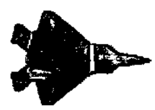
Distributing probability in i region P_i is:

$$P_i = S_i/S \tag{13}$$

Where, S is the same as formula (10), S_i is the area in i region, i.e. the sum of pixel in i region.

4.Experiment result

There are two plane models and one polygon as objects in the experiment. BP network is used to recognize and classify. The objects is in figure 3:



a

b

c

Figure 3. Three objects in the experiment

Structure invariant of object is composed of k, ρ, $P_i(i = 1,2,3,4,\ i.e.\ L = 4)$. Span is redefined as below:

$$k = k/U \tag{14}$$

Where, U is the image width in polar-log coordinate.

Each object changes pose every 15-degree in plane. There are 24 poses of each object. The scale of object is 1, 0.5, 0.8 and 2, so each object has 96 stylebooks. Three objects together have 288 stylebooks. In learning stage, each object selects 20 poses, together 60 poses. BP network has three layers. Structure invariant of each image is learned by BP network, and power value and threshold are gained. When 288 stylebooks are recognized by power value and threshold, only 4 stylebooks are not recognized. The recognition ratio of this method is 98.61%. The time spend is little than the method of profile invariant. Thereby, the recognition ability of this method is great.

5. Conclusion

Structure invariant is defined in polar-log coordinate in this paper. It is indicate from experiment that the good recognition result can be got by structure invariant. But this method combined profile invariant can correctly recognize the object that has the same span, close ratio of curve area and distributing of curve.

Reference

[1] Hu M K. Visual pattern recognition by moment invariant. IRE Trans Information Theory,1962,8:179-187P

[2] Chen Yan-xin, Qi Fei-hu. New affine invariants invariants for target recognition. Journal of Infrared Millimeter Waves. 1999,18(1): 61-66

[3] Dong Mi, Xu Zheng-wei, Wu Cheng-ke. Recognition of planer curves based on perspective invariants. Pattern Recognition & Artificial Intelligence.1995,8(2): 108-113

[4] Nirwan A, Edward J D. On detecting dominant points. Pattern recognition. 1991,24(5): 441-451

[5] Zhang Hai-yan, Wang Dong-mu, Song Ke-ou. 3D moving object recognition. CVRV2002. Hang Zhou, China, 2002,4

[6] Seiter M. and Waxman A.M. Adaptive 3D Object Recognition from Multiple Views. IEEE Trans. on Pattern Analysis and Machine Intelligence Vol.14.2, Feb.1992: 107-124P

[7] C.F.Hall and E.L.Hall, A non-linear model for the spatial characteristics of the human visual system, IEEE Trans: Systems Man,1979:161-167P

[8] Carl F.R.Weiman, Richard D.Juday. Tracking Algorithms Using Log-Polar Mapped Image Coordinates. SPIE Vol.1192 Intelligent Robots and Computer Vision VIII: Algorithms and Techniques (1989): 843-853P

[9] F. Jurie. A New Log-Polar Mapping for Space Variant Imaging Application to Face Detection and Tracking. Pattern Recognition. 1999:865-875P

[10] Michael Seibert, Allen M.Waxman. Recognizing Faces from Their Parts. SPIE vol.1611 Sensor Fusion IV(1991):129-140P