'91 KACC 1991. 10. 22~24

Appplication of FFT to Orientation Determination

Dong-Min Woo* and Mignon Park**

* Dept. of Control and Instrumentation Eng., Myong Ji University
** Dept. of Electronic Eng., Yonsei University

ABSTRACT

In this paper, a new two dimensional processing method is presented, which can accurately determine the orientation of an part. Matching between the object and the model is performed in the frequency domain. The DFT of the object contour is decomposed to estimate the orientation of the object and to evaluate the similarity between the object and the model. In this context, this new approach is very robust with respect to noise and no preprocessing of the contour is required. Also, this method has many advantages over the conventional correlation technique. With only a few uniformly sampled points, this method can estimate the accurate orientation in an efficient manner even in a noisy environment.

I. Introduction

Image shape analysis is quite useful in a variety of situations such as charater recognition [1], target identification [2][3], classification of chromosomes [4], and industrial part identification and inspection [5][6]. An important topic in shape analysis is the pose estimation of objects in a scene. Especially in a real robotics applications, the precise position and orientation of parts need to be known so that they can be properly grasped. Most industry-oriented vision systems handle this topic in terms of boundary information of two dimensional shape. The current approaches are based on center of area and axis of the least moment of inertia[7], length of radial line from the center to the boundary[8] and θ-s transformation[6]. But most of these approaches have a problem in determining the orientation, since they are either not so accurate or too slow.

The approach described in this paper uses a contour representation scheme similar to Fourier descriptor method [9], which constructs invariant description of two-dimensional shape. In this paper, however, we decompose the DFT coefficients of the object into model components and noise components to estimate its orientation. Then noise components are minimized to estimate the optimal orientation in a least square sense. This new approach has many advantages over the spatial domain methods. With only a few sampled points on the contour, this menthod can generate a more accurate orientation than the conventional correlation techniques.

A block diagram of the method is shown in Fig. 1. Boundary contours of parts are extracted out of an image. A uniform sampling algorithm is used to extract uniformly sampled boundary contour to evaluate the DFT of the contour. The recognition scheme matches the object DFT with the model DFT to determine the identity, position and orientation. This new apprach is very robust with respect to noise and no preprocessing of the contour is required.

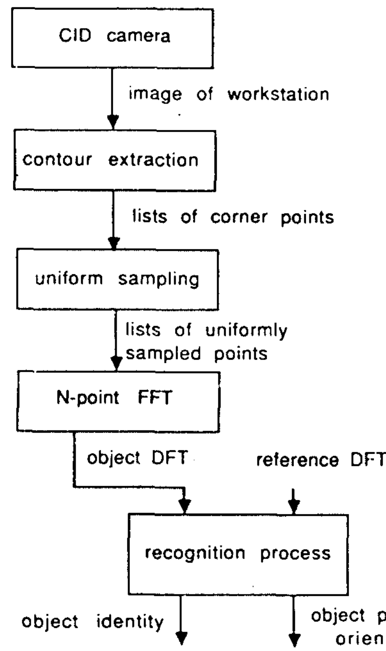

Fig. 1. Block diagram of the part-recognition system

II. Representation of the closed contour

In representing a contour in the frequency domain, we use a scheme similar to the Fourier descriptor. Consider a closed contour c in the complex plane, as shown in Fig. 2. Trace this contour once with uniform velocity v. Then, we will obtain a parametrized representation of c as a complex fuction $Z(t)$, with parameter, t. Choose v so that the time T required to traverse the contour is unity. If the contour is traced repeatedly, we get a periodic function, which may be expanded in a convergent Fourier series. We will define a Fourier descriptor of c to be the complex Fourier series expansion of $Z(t)$.

$$Z(t) = \sum_{n=\infty}^{\infty} A(n) \exp(j2\pi nt)$$

$$A(n) = \frac{1}{2} \int_{0}^{2\pi} Z(t) \exp(-j2\pi nt)\, dt.$$

This representation depends on both the shape of the contour and the starting point of $Z(t)$. Actually, c is taken from a digitized image, and thus $Z(t)$ is not available as a continous function. If $z(k)$ is a uniformly sampled version of $Z(t)$ of length N, the DFT provides the N lowest frequency coefficients $A(i)$.

$$z(k) = \sum_{i=0}^{N-1} A(i) \exp(j2\pi nk/N)$$

$$A(i) = \frac{1}{N} \sum_{k=0}^{N-1} z(k) \exp(-j2\pi nk/N).$$

The computation of this representation is fairly straightforward. The contour of the image is represented as a sequence of x-y coordinates which are derived from contour tracing. Here, we again remark that this contour be sampled uniformly. N uniformly sampled points, $z(k)$ for $k=0$ to $N-1$, are generated from a sequence of contour points by interpolation. Thus, the perimeter of the contour is computed first, and the contour is sampled at a spacing chosen to make the total number of samples, N, a prespecified number which is chosen to be a power of 2 for convenience. Then, the representation is computed by using an FFT implementation of the DFT of this sequence.

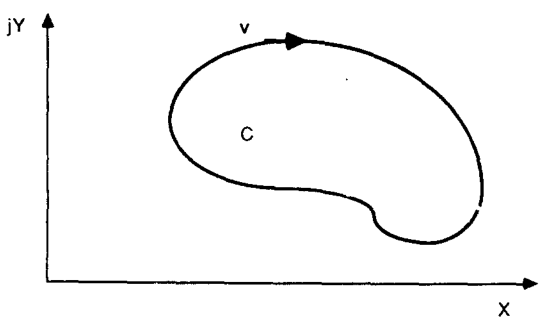

Fig. 2. Representation of the closed contour in the complex plane

III. Shape matching

We investigate the matching between the reference contour and the object contour, when the object contour and the reference contour are not exactly the same , but similar enough to be identified as being of the same shape. Also, we assume that the object contour is randomly positioned and oriented, and initialized from an arbitrary point.

Suppose that the object contour and the reference contour are represented as sequnces of uniformly sampled points, $a(k)$ and $m(k)$. Then the DFTs of the object contour and the reference contour are defined as $A(i)$ and $M(i)$, respectively. First, to normalize the position, we simply set the zero-coefficient of the DFT equal to zero. Then, the centroids of the object and reference contours are located at the origin of the coordinate system.

In the spatial domain, we decompose the object contour into the rotated and starting-point shifted version of reference contour $m'(k)$ and the error sequence $e(k)$. In this case, the magnitudes of the error sequence are the distances between the sequence $a(k)$ and $m'(k)$. The error is minimum when the normalized reference contour is correctly rotated and starting-point shifted to fit into the object contour. If we assume that the error sequence $e(k)$ is a zero mean Gaussian white noise sequence independent of the sequence $m'(k)$ the problem of "optimal" can be formulated as a least square problem. Define:

$$a(k) = m'(k) + e(k) \qquad \text{for } k = 0 \text{ to } N-1,$$

then the expectation of $\sum|e(k)|^2$ is minimized by proper choice of $m'(k)$ which is a contour synthesized by rotating the reference contour by θ_r and shifting the starting point by β points. We refer to $m'(k)$ as the optimally aligned version of the reference contour according to the objective of minimum squared error. $\sum|e(k)|^2$ is the summation of the squared distances between uniformly sampled points of the object contour and the optimally aligned version of the reference contour. As such, the matching reduced to the estimation of the optimal orientation θ_r and the evaluation of $\sum|e(k)|^2$., which is similarity measure between the object and the reference contour.

In the frequency domain, the decomposition is as follows:

$$A(i) = M'(i) + E(i) \qquad \text{for } i = 1 \text{ to } N-1,$$

where sequences, $M'(i)$ and $E(i)$, are the DFTs of sequences, $m'(k)$ and $e(k)$, respectively. Due to Parseval's relation, the above equation minimizes $\sum|E(i)|^2$. Using the properties of the DFT, we have the following relations:

$$|M'(i)| = |M(i)| \text{ and}$$
$$\arg[M'(i)] = \arg[M(i)] + \theta_r - 2\pi\beta i/N \qquad \text{for } i = 1 \text{ to } N-1,$$

where θ_r is the optimal rotation and β is the optimal starting point shift.

Since the starting point shift is in the range of $\{0,N-1\}$, we may define the optimal starting point shift angle as follows:

$$\theta_s = 2\pi\beta/N, \quad \text{then} \quad 0 \le \theta_s < 2\pi.$$

Then we have

$$\arg[M'(i)] = \arg[M(i)] + \theta_r - \theta_s\, i \qquad \text{for } i = 1 \text{ to } N-1.$$

Since $E(i)$ and $M'(i)$ are independent and orthogonal, $M'(i)$ can be interpreted as a minimum squared error approximation to $A(i)$. To evaluate the correct estimates of θ_r and θ_s, we have to select the two best approximations among N-1 approximations. Since $E(i)$ are the DFT coefficients of the white noise error sequence, the phase angle of $A(i)$ can be more accurately approximated by the phase angle of $M'(i)$, when the magnitudes of $A(i)$ is large for a certain spatial frequency i. In other words, the DFT coefficient, whose magnitude is large, contains relatively less noise than other coefficients.

To calculate the optimal orientation, two coefficients whose magnitudes of the object DFT are largest among the N-1 coefficients are selected as $A(f_1)$ and $A(f_2)$. Then, $M'(f_1)$ and $M'(f_2)$ can be approximated by $A(f_1)$ and $A(f_2)$, respectively and we have the following equations:

$$\arg[A(f_1)] = \arg[M(f_1)] + \theta r - \theta s\, f_1,$$
$$\arg[A(f_2)] = \arg[M(f_2)] + \theta r - \theta s\, f_2.$$

From the above equations, we can evaluate θ_r and θ_s. Since θ_r and θ_s are periodic with 2π, the solution is not necessarily unique. We have m solutions as follows:

$$\theta_s(j) = \theta_s(1) + \frac{2\pi\,(j-1)}{f_2 - f_1} \qquad \text{for } j = 1 \text{ to } m,$$

where $m = \min(\,|f_1-f_2|,\; N - |f_1-f_2|\,)$ and

$$\theta_s(1) = \frac{\arg[A(f_1)] - \arg[M(f_1)] - \arg[A(f_2)] + \arg[M(f_2)]}{f_2 - f_1}.$$

The corresponding solution of θ_r for each θ_s:

$$\theta_r(j) = \arg[A(f_1)] - \arg[M(f_1)] + \theta_s(j)\, f_1 \quad \text{for } j = 1 \text{ to } m.$$

The next step is to select the optimal orientation among the m solutions given above. Thus, we evaluate a similarity measure between object contour and reference contour. We choose the similarity measure to be defined as $\Sigma|e(k)|^2$. To calculate this, we calculate M'(i) using the given θ_r and θ_s.

$$\Sigma|e(k)|^2 = \Sigma\, |E(i)|^2 = \Sigma\, |\, A(i) - M'(i)\, |^2,$$

where $M'(i) = M(i)\exp\{j(\theta_r - \theta_s i)\}$ for i = 1 to N-1.

For each pair of θ_r and θ_s, we calculate a similarity measure. Among those, we certainly choose the least one. The optimal orientation is θ_r for the least similarity measure.

IV. Experimental results

The accuracy of the estimated orientation was tested for various numbers of uniformly sampled boundary points. Fig. 3 shows contours of the U-shhaped object with different numbers of contour sampling points. For this test, the object was relocated at several orientations (0, 30, ... , 330 degree) by robot manipulator. Contours were recognized using 4-point, 8-poit, 16-point, 32-point contour samples. Table 1 shows the results of the test.

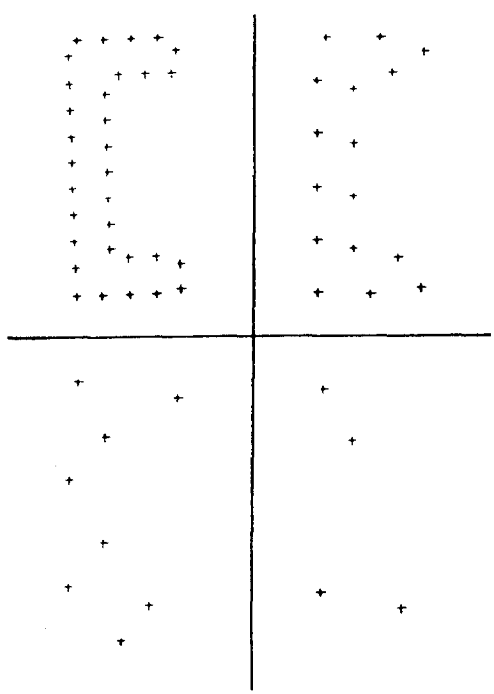

Fig. 3. Contours of the U-shaped object, uniformly sampled by 32, 16, 8, 4 samples

rotation	estimates			
	4-point	8-point	16-point	32-point
30	34.94	30.31	29.43	28.45
60	50.14	59.22	58.87	60.31
90	68.95	90.80	90.19	89.26
120	96.98	121.31	121.01	120.50
150	118.30	149.90	150.62	149.83
180	68.41	179.77	179.78	179.65
210	267.97	211.02	210.57	209.95
240	279.25	240.61	239.48	239.37
270	191.17	270.17	269.67	270.14
300	223.98	299.40	300.33	300.14
330	244.08	330.21	330.29	329.73
average error		0.54	0.50	0.43

Table. 1. Estimation of the orientation

The use of 4 boundary points did not provide any reasonale estimate of the orientation. But with 8 or more boundary points, reasonble estimate of orientation were obtained. Even with 8 points, the average orientation error is only about 0.54 degrees. With 16 and 32 points, the accuracy slightly increased. The accuracy of orientation is very important because it is directly related to the accuracy of the similarity measure.

Insensitivity of the estimation to the shape of object was tested by using an object with complex shape, shown in Fig. 4. For 8-point, 16-point, 32-point contour samples each average error of estimated orientation was shown as 3.21 degrees, 0.58 degrees and 0.43 degrees, respectively. Through this test, the performance of the algorithm was proved to be almost insensitive to the shape of an object.

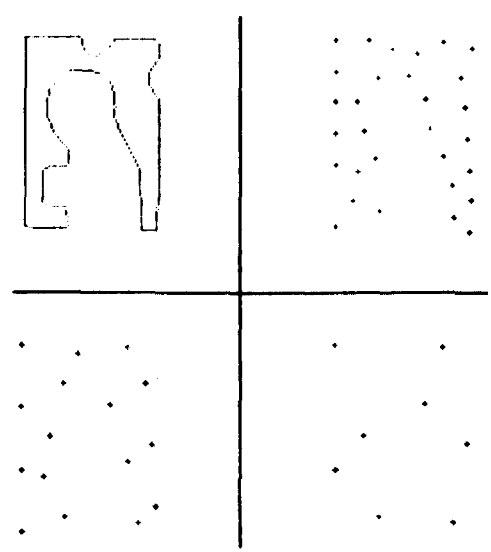

Fig. 4. Contour of the complex object and its uniform sampling by 32, 16, 8, 4 points

It should be noted that with small number of points the centroid of the contour is very sensitive to the selection of the uniformly sampled points and in turn, the object contour is not appropriately aligned with the model shape in the matching process. This results in the inaccuracy of the orientation.

V. Conclusion

To identify and locate parts, a new two-dimensional signal processing method was developed. Shape matching between the part and the model is performed in the Fourier domain. The DFT of the object contour is decomposed to estimate the orientation of the object and evaluate a similarity measure. In this context, this method appears to be insensitive to noise and does not require preprocessing of the boundary contour. Also, by using FFT algorithm which can be rapidly processed by DSP chip, the recognition speed can be very fast.

The experimental results indicate that this method is very accurate in estimating orientation. With only 8 uniformly sampled points, we estimated orientation more accurately than with conventional corelation methods in the spatial domain.

REFERENCES

[1] C. C. Tappert, C. Y. Suen and T. Wakara, "The State of Art in On-Line Handwritten Recognition," IEEE Trans. Pattern Anal. Machine Intell., vol. 12, no. 8, pp. 787-808, 1990.

[2] S. A. Dudani, K. J. Breeding, R. B. McGhee, "Aircraft Identification by Moment Invariants," IEEE Trans. Comput., vol. 26, no. 1, pp. 46-55, 1983.

[3] A. P. Reeves, R. J. Prokop, S. E. Andrews and F. Kuhl, "Three Dimensional Shape Analysis Using Moments and Fourier Descriptors," IEEE Trans. Pattern Anal. Machine Intell., vol. 10, no. 6, pp. 937-943, 1988.

[4] K. S. Fu, Syntactic Pattern Recognition and Application, Eaglewood Cliffs, NJ: Prentice-Hall, 1982.

[5] S. R. Dubois and F. H. Glanz, "An Autoregressive Model Approach to Two-Dimensional Shape Classification," IEEE Tran. Pattern Anal. Machine Intell., vol. 8, no. 1, pp. 55-66, 1986.

[6] W. A. Perkins, "A Model-Based Vision System for Industrial Parts," IEEE Trans. Comput., vol. 27, no. 2, pp. 126-143, 1978.

[7] S. W. Holland, L. Rossol and M. R. Ward, "Consight-I: A Vision-Controlled Robot System for Transfering Parts from Belt Conveyors," Computer Vision and Sensor-Based Robots, Plenum, pp. 81-100, 1979.

[8] M. Yachida and S. Tsuji, "A Versatile Machine Vision System for Complex Industrial Parts," IEEE Trans. Comput., vol. 26, no. 9, pp. 882-894, 1977.

[9] T. P. Wallace and P. A. Wintz, "An Efficient Three-Dimensional Aircraft Recognition Algorithm Using Normalized Fourier Descriptors," Computer Graphics and Image Processing, 13, pp. 99-126, 1980.