IEEE Pacific Rim Conference on Communications, Computers and Signal Processing
June 1st - 2nd, 1989

MORPHOLOGICAL SKELETON TRANSFORMS FOR DETERMINING POSITION AND ORIENTATION OF PRE-MARKED OBJECTS

Z. Zhou, K. C. Smith* †, B. Benhabib †, R. Safaee-Rad †*
Computer Integrated Manufacturing Laboratory
† Department of Mechanical Engineering
* Department of Electrical Engineering
University of Toronto
Toronto, Ontario M5S 1A4

Abstract

In this paper, the feasibility of a pre-marking scheme for 3-D object recognition is presented. The proposed scheme is based on the assumption that an object can be modeled by a small number of its distinct 2-D perspective projections, referred to as "standard-views". Circular markers are used to identify these views by determining their surface normals passing through their centers. Since a circular marker would always appear as an ellipse on all acquired images, its shape can be related to the orientation of the surface, on which it is located, with respect to the camera's optical axis. The surface normal of the marker can be determined by analysing the geometrical features of its acquired pseudo-ellipse image using morphological skeleton transforms. The position of the marker, on the other hand, has to be determined by acquiring two images from different viewing-angles. The experimental results illustrate that the specific pseudo-Euclidean skeleton transform used in this paper can accurately determine the features of the ellipse to allow the succesful aplication of the proposed pre-marking scheme.

1. Introduction

Although many 3-D recognition techniques and algorithms have been reported in the literature, most are either computationally too expensive or too inflexible to be applied in real-time. The complexity of the problem arises from the fact that a vision system has to search for the desired information about an object within the impoverished image data collected by the camera. A pre-marking scheme can greatly reduce the complexity of the search process by using a set of pre-designed markers on the surfaces of an object, so that the desired information, such as position, surface normal, identity, etc., about the object can be inferred from the limited amount of data provided by the markers. The markers should have simple geometries, such that they can be readyly distinguished from the object features. Thus, analyzing these markers in an image is much simpler and faster than dealing with the original shape of the object.

The concept of pre-marking has been successfuly used in industrial inspection, [1,2], as well as various other fields, including biomedical research, [3], and space research, [4]. Thus, the pre-marking concept will not be further discussed herein. In this paper a new pre-marking scheme based on circular markers is presented for 3-D object recognition in industrial environments. In the next section, 3-D object modeling by several distinct perspective projections is discussed as an introduction to the pre-marking reasoning. In section 3, the morphological skeleton transform is discussed. In section 4, the pre-marking scheme based on circular markers is presented. The detection process of circular markers using morphological skeleton transform is discussed in section 5. Section 6 presents the conclusions of the research.

2. 3-D Object Modeling

In the proposed 3-D object recognition technique, an object is modeled using only a small number of 2-D distinct perspective views. These are referred to as "standard views", each with a corresponding "standard view-axis". For successful recognition, the input image of an object should be one its standard views. Thus, the camera's optical axis has to aligned with one of the standard view-axes of the object to acquire a standard view, Figure 1. The matching process is, then, performed between the acquired 2-D standard view of the object under consideration and the library of 2-D standard views of a set of objects.

To enable the vision system to acquire standard views, standard view-axes must be pre-defined. This can be accomplished by defining a local surface normal for each distinct view of an object. These local surface normals can be defined by adding non-functional features to an object, preferably, during the manufacturing process - pre-marking. Thus, in the context of the proposed object recognition method, pre-marking is used for defining local surface normals, and the corresponding standard view-axes.

This paper presents a pre-marking scheme based on circular markers. Circular geometry has been selected due to the following reasons: 1) many industrial parts have holes, which can be used as natural markers; 2) a circle has symmetry with respect to its norm; and, 3) circles have been shown to have the property of higher image location accuracy.

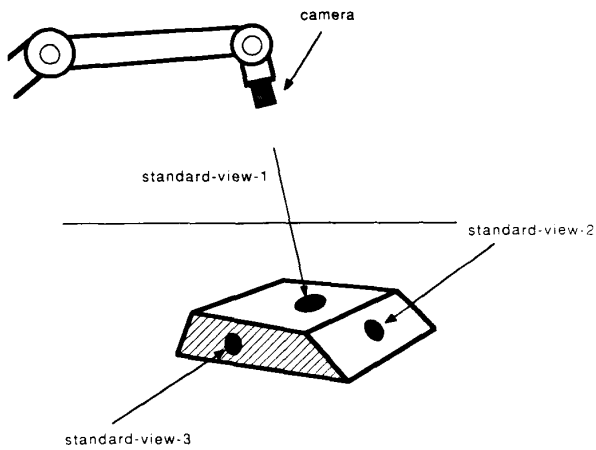


Figure 1. Standard-views of a pre-marked object.

3. Morphological Skeleton Transform Algorithms

Mathematical morphology is a study of geometrical shapes and forms using set-oriented approaches, [5]. In mathematical morphology, a shape in a digital binary image is represented by a subset of discrete space Z^2. It is analyzed and processed by different morphological transformations, each of which is either a set operation, such as set union, intersection and subtraction, or the one derived from set operations. The basic morphological transformations are erosion and dilation. Both of them use a simpler set called structuring element to analyze the original shape. For example, erosion of the shape examines whether a moving structuring element fits inside its interior, while the dilation actually finds all those pixels where the moving structuring element hits the original shape. By varying the size and the shape of the structuring element, morphological transformations

CH2691-4/89-0000- 301 $1.00 © 1989 IEEE

provide a systematic approach to a quantitative study of complex shapes.

One of the most useful morphological transformation is the *Morphological Skeleton Transform (MST)* which reduces a shape into a graph-like description while preserving all its information, [6]. The MST is a very efficient representation since it only uses a small subset of pixels of the original shape. Moreover, it emphasizes the geometrical structure of the shape so that the detection of certain features becomes much easier.

The MST $SK(X)$ of a shape X is computed by the following algorithm [7],

$$\left.\begin{aligned} SK(X) &= \bigcup_{n=0}^{N} S_n(X) \\ S_n(X) &= [X \ominus B(n)] / [X \ominus B(n) \circ B_n(1)] \\ N &= \max\{n: X \ominus B(n) \neq \varnothing\} \end{aligned}\right\} \quad (1)$$

where operators $\ominus$ and $\circ$ represent erosion and opening operations respectively. Set $B(n)$ is a structuring element of size n, where $B_n(1)$ is a structuring element of unit size. Each skeleton pixel x in $SK(X)$ is associated with a positive integer referred to as the skeleton function $skf(x)$ of x, which is equal to the morphological distance from x to the shape's boundary. In practice, $skf(x)$ of any skeleton pixel can be found by

$$skf(x) = n, \qquad \text{for } x \in S_n(X), \qquad (n = 0, 1, .., N). \quad (2)$$

The MST is classified by the types of its structuring elements $B(n)$ and $B_n(1)$. The most common type is the USD skeleton (based on uniform-step-distance), whose structuring elements are defined by

$$\left.\begin{aligned} B(n) &= nB = B \oplus B \oplus \cdots \oplus B \quad (n \text{ times}) \\ B_n(1) &= B, \qquad\qquad n = 1, 2, \cdots \end{aligned}\right\} \quad (3)$$

where B is a bound discrete set of unit-size. The $B(n)$ in Eq. (3) is composed by an n-fold dilation of set B. Several examples of the most commonly used structuring element B are shown in Figure 2. The second class of MST is the PUSD skeleton (based on periodically-uniform-step distance), which uses multiple unit-size structuring elements $B_1, B_2, .., B_m$ in the composition of $B(n)$

$$B(n) = \begin{cases} B_1 \oplus B_2 \oplus \cdots \oplus B_m, n \le m, \\ pB(m) \oplus B(q), n = pm + q > m, \quad q < m \\ B_1 \subset B_2 \subset \cdots \subset B_m, \end{cases} \quad (4a)$$

while its unit-size SE $B_n(1)$ varies with size n,

$$B_n(1) = \left\{ B_i: \quad i = (n-1) \bmod m + 1 \right\}. \quad (4b)$$

An important example of PUSD skeleton is the one which uses $B_1 = DIAMOND$, $B_2 = SQUARE$ and $m = 2$, which is also known as the skeleton based on octagonal distance.

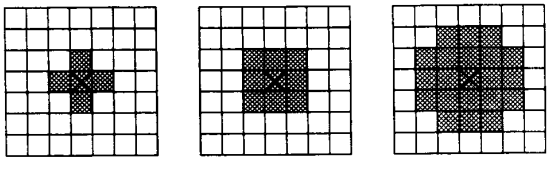


Figure 2. Examples of commonly-used unit-size structuring elements.

Although the PUSD skeleton, based on octagonal distance for example, can be made less sensitive to the rotation of a shape than the USD skeleton does, none of them is satisfactory enough for rotation-invariant applications. Therefore, a third class of MST, namely the pseudo-Euclidean skeleton, has been propoesd in [11]. The pseudo-Euclidean skeleton uses a set of quasi-circular discrete sets as its structuring elements B(n). These quasi-circular structuring elements are composed in such a way that each of them can be decomposed into a set of simple sets called rotation sets, each of which contains only two or three pixels including the origin. The expression for a general quasi-circular SE $B(n)$ is generated by

$$B(n) = a_1 P(1) \oplus a_2 P(2) \oplus \cdots \oplus a_k P(k) \oplus b_1 B_1 \oplus b_2 D_1, \quad (n = 0, 1, 2, 3, \cdots), \quad (5)$$

where sets $P(1),...,P(k)$ consist from various types of rotation sets, and a_k, b_1, and, b_2 are non-negative integer coefficients. Sets B_1 and D_1 are the unit-size sets DIAMOND and SQUARE respectively, Figure 2. The $B_n(1)$ is defined by

$$B_n(1) = \begin{cases} B_1, & n = 4k,\ 4k+1,\ 4k+3, \\ D_1 & n = 4k+2. \quad (k > 0). \end{cases} \quad (6)$$

Since the structuring elements are composed by rotation sets through dilation, the algorithm of Eq. (1) can be implemented iteratively, where the total computational cost is reduced. The shape of the resulting structuring elements $B(n)$ are good approximations to circular disks of different sizes, so that the pseudo-Euclidean skeleton has good rotation-invariant property. The composition of these quasi-circular structuring elements is beyond the scope of this paper. It is only pointed out here that the pseudo-Euclidean skeleton has the best rotation property among existing MST algorithms.

4. Circular Markers

The most relevant property of circular markers in the context of our research is the rotation-invariancy, in the sense that the projection of a circular plane at any viewing angle would always be a pseudo-ellipse. Therefore, the surface normal of a marker can be determined from its image, by extracting and analysing the features of the ellipse, since the shape of an ellipse depends on the orientation of the surface on which it is located.

Let the X-Y plane in Figure 3 be the image plane, and the Z axis be the optical axis of the camera, where a circular marker is modeled by the intersection of the sphere

$$x^2 + y^2 + (z - z_p)^2 = R^2$$

and the plane

$$Ax + By + C(z - z_p) = 0.$$

The surface normal at the center of the marker is then equal to that of the plane, $\mathbf{n} = (A, B, C)$.

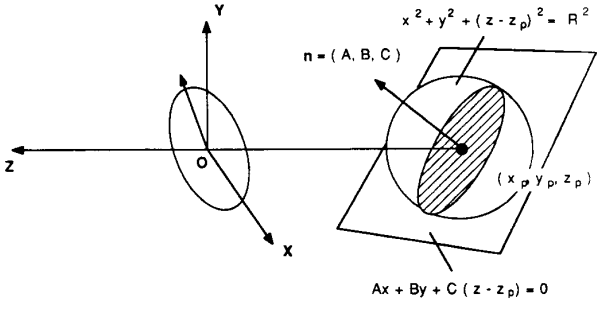


Figure 3. Surface normal.

The projection of the marker on the X-Y plane can be described as follows, assuming orthographic projection, since the size of the marker is much smaller compared to its distance from the camera,

$$\frac{C^2+A^2}{C^2R^2}x^2+\frac{C^2+B^2}{C^2R^2}y^2+\frac{2AB}{C^2R^2}xy=1. \tag{7}$$

Equation (7) defines a rotated ellipse on the image plane, Figure 4, whose rotation angle θ is related to the surface normal by

$$\tan 2\theta=\frac{2AB}{A^2-B^2}. \tag{8}$$

If the coordinate system of the image plane is rotated according to

$$\begin{bmatrix}X\\Y\end{bmatrix}=\begin{bmatrix}\cos\theta & -sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}U\\Y\end{bmatrix}, \tag{9}$$

then, the equation of the ellipse with respect to the new coordinate system U-V becomes

$$\frac{U^2}{M^2}+\frac{V^2}{N^2}=1 \tag{10a}$$

where

$$\left.\begin{aligned}M&=R\left[1+\frac{(A\cos\theta+B\sin\theta)^2}{C^2}\right]^{-\frac{1}{2}}\\ N&=R\left[1+\frac{(A\sin\theta-B\cos\theta)^2}{C^2}\right]^{-\frac{1}{2}}\end{aligned}\right\} \tag{10b}$$

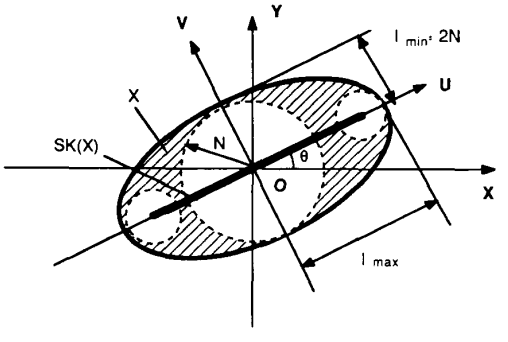


Figure 4. Skeleton representation of the projection of a circular marker.

M and N are the lengths of the ellipse axes collinear with the axes U and V respectively. However, the angle θ should be further defined so that no ambiguity results from solving Eq. (10) for the surface normal.

When an ellipse is perceived on an image, its major axis can be distinguished from its minor axis by comparing their lengths. It is therefore convenient to define the rotation angle θ as the angle from the horizontal axis of the image plane to the direction of the major axis of the ellipse. In this way, the major axis of the ellipse is always in the direction of the U axis, Figure 4. Hence, Eq. (8) reduces to

$$A=-B\tan\theta. \tag{11}$$

Substituting Eq. (11) into Eq. (10), the following expression can be obtained,

$$C=\frac{|A\sin\theta-B\cos\theta|}{\sqrt{\left(\frac{M}{N}\right)^2-1}}. \tag{12}$$

Both the rotation angle θ and the ratio M/N, can be measured from the ellipse in the image. Considering also that A, B, C are directional cosines of $\mathbf{n}$, and that

$$A^2+B^2+C^2=1,$$

the surface normal $\mathbf{n}$ can be determined using Eq.'s (11) and (12).

The position of the marker can be determined by acquiring two distinct images. In Figure 5, point $P(x_p,y_p,z_p)$ is defined as the center of a circular marker. The camera first acquires an image at position A, and then, it is moved to position B to acquire a second image. The offset from position A to B is defined as $(\Delta x,\Delta y,\Delta z)$. To simplify the analysis, the offset Δz can be set to zero. If the focus length f of the camera is known, the position of the center of the marker can then be defined by

$$\left.\begin{aligned}z_p&=f-\frac{f\Delta x}{u_a-u_b}\\ x_p&=\frac{u_a\Delta x}{u_a-u_b}\\ y_p&=\frac{v_a\Delta y}{v_a-v_b}\end{aligned}\right\} \tag{13}$$

where (u_a,v_a) and (u_b,v_b) are the coordinates of the centers of the ellipses perceived from the images acquired from positions A and B respectively.

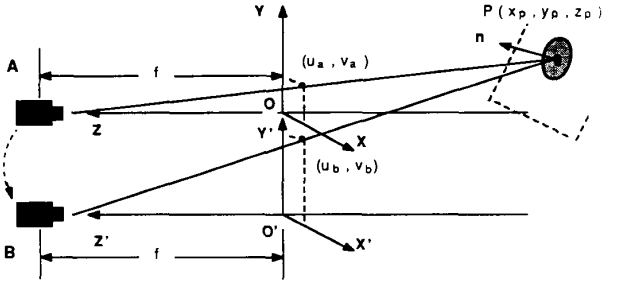


Figure 5. Depth information from two images.

5. Detection of Circular Markers Using Morphological Skeleton Transforms

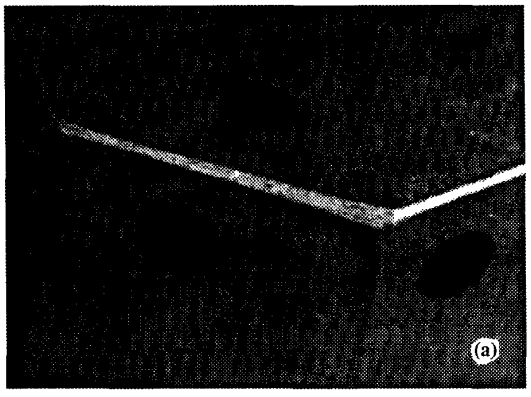
The previous section stated that, in order to determine a surface normal and the position of an object, at least the following three features of an ellipse should be known: its center, its rotation angle θ, and its major-axis-length/minor-axis-length ratio. In this section, MSTs are proposed for extracting these features from an image of an ellipse. Although these features can also be extracted by other methods, the approach presented herein is more robust to noise and potentially faster than other methods.

The real image in Figure 6(a) shows an object with three visible surfaces, each of which is pre-marked by a circular marker. In order to make the markers prominent in the image, the colors and the shades of the markers should be carefully chosen. Based on the histogram of Figure 6(a), a simple thresholding process resulted in the binary image shown in Figure 6(b). Some noise can been seen in the image due to the shades which resulted from the non-critical lighting. It is likely that the noise patterns are inevitable even if a more sophisticated thresholding process would have been used. Therefore, the next step is to detect the existence of the ellipses and to separate them from the noise patterns, which can be accomplished based on MST representation.

The MST representation of Figure 6(b) is displayed in Figure 6(c), in which the pseudo-Euclidean skeleton transform is used to

ensure accurate measurements despite possible rotation. Thus, the ellipses as well as the noise patterns are simultaneously reduced into their own axial descriptions. The ellipses in Figure 6(b) are mutually disconnected, so are their MST representations in Figure 6(c). The MST of an ellipse tends to be a straight line coincident with the major axis of the ellipse. Also, the skeleton function of a skeleton pixel along the MST of an ellipse displays a very regular distribution; its value increases as its position gets closer towards the center. The maximum skeleton function value is defined to represent the mid-most pixel. On the other hand, the MST of the noise patterns are perceived either as irregular curves, or as straight lines whose skeleton function do not follow the above distribution. In fact, it can be proved in theory that only elliptic shapes have straight-line skeletons with the above described skeleton function distribution. In addition, most noise patterns are so small that they can easily be eliminated by examing their maximum skeleton values. Hence, detecting ellipses from other patterns is much easier in the MST domain than in the binary image. A search algorithm was developed, which can remove noise patterns and leave the ellipses intact for further analysis.

Once the noise patterns have been discarded, the features of the remaining ellipses can be extracted. The process can be described by the following example: Let shape X be one of the ellipses found in Figure 6(b) and $SK(X)$ be its MST after the removal of noise. The center of this ellipse can be obtained by computing the center of gravity of the skeleton subset having maximum skeleton function. If N is the maximum value of the skeleton function of $SK(X)$, and there are m skeleton pixels in $S_N(X)$, then the coordinates (u,v) of the center of the ellipse can be determined by



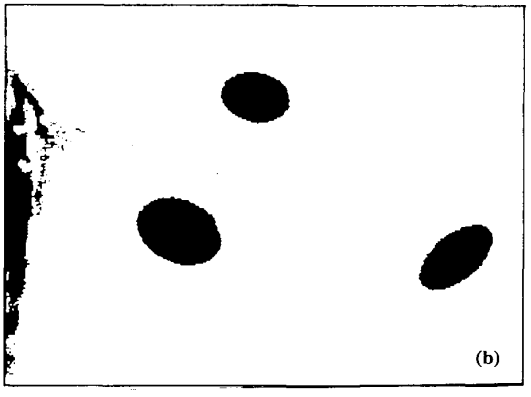
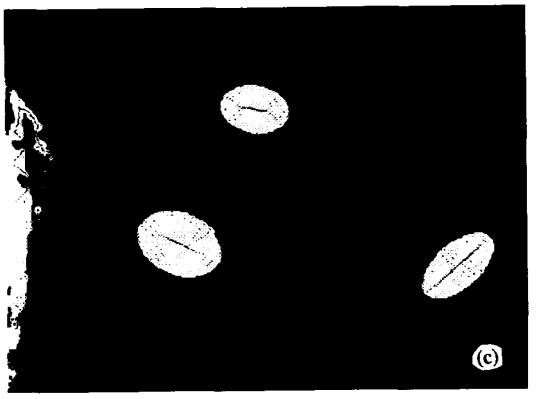


Figure 6. Detection and analysis of a pre-marked object
(a) Original image of a pre-marked object with three visible surfaces;
(b) After thresholding;



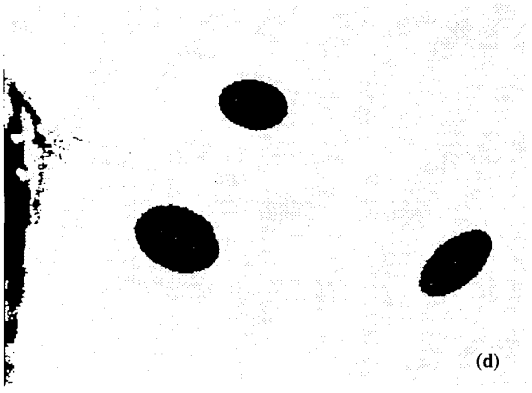


Figure 6. (continued)
(c) Pseudo-Euclidean skeleton of (b);
(d) Detection of center and end pixels for each marker.

$$u = \frac{1}{m} \sum_{(x,y)\in S_N(X)} x, \qquad v = \frac{1}{m} \sum_{(x,y)\in S_N(X)} y. \tag{14}$$

In many cases, the accuracy of Eq. (14) can be improved as more skeleton subsets are taken into consideration during the averaging process of Eq. (14). Figure 6(d) shows the positions of the center pixels determined for the three visible markers.

The ratio of major-axis-length/minor-axis-length defined as,

$$r(X) = \frac{l_{\max}}{l_{\min}} \tag{15}$$

can be determined once $l_{\max}$, the length of the major axis, and $l_{\min}$, the length of the minor axis, have been determined. From Figure 4 it can be observed that the length of the minor axis is

$$l_{\min} = 2N. \tag{16}$$

In order to determine $l_{\max}$, on the other hand, the two end pixels of the major axis have to be identified. The end pixels end_1 and end_2 of $SK(X)$ are defined as the skeleton pixels which have the largest distances from the center pixel (u,v) in the opposite directions. A search algorithm has been devised to determine end_1 and end_2 by comparing the distance of each skeleton pixel in $SK(X)$ to the center pixel found in Eq. (14). The pair of end pixels for the three markers in Figure 6(a) determined by the above search algorithm are shown in Figure 6(d). Once the pair of end pixels haven been determined, $end_1 = (x_{end1}, y_{end1})$ and $end_2 = (x_{end2}, y_{end2})$, then the length of major axis is obtained by

$$l_{max} = \sqrt{(x_{end1} - x_{end2})^2 + (y_{end1} - y_{end2})^2} + skf(end_1) + skf(end_2). \quad (17)$$

The angle θ between the major axis and the horizontal image axis can be calculated from the coordinates of the above two end pixels by

$$\theta = \tan^{-1} \frac{y_{end2} - y_{end1}}{x_{end2} - x_{end1}} \quad (18)$$

if $x_{end2} \geq x_{end1}$, and vice versa.

6. Discussion of Results

The MST and the feature extraction algorithms were implemented in C programming language and tested on images of a pre-marked object using Sun 3 Workstation. Figure 6(a) shows part of a pre-marked object with three visible surfaces. Each surface is pre-marked by a circular marker of equal size, which is perceived as an ellipse. Due to the distance variation from the camera, the sizes of these markers appear different. However, the assumption of orthographic projection still remains valid since the important parameter, namely the ratio of the major-axis-length/minor-axis-length, is only decided by the relative length measurement.

The image in Figure 6(a) was thresholded into a binary image of Figure 6(b), and its pseudo-Euclidean skeleton was computed, Figure 6(c). A noise filtering process was performed on the above obtained skeleton to remove the noise and spurious skeleton branches resulted from the rotation. The algorithm eliminated all patterns but the three skeletons of the three markers. Then the center pixels and the pair of end pixels were computed, Figure 6(d). The features of the three markers are presented in a tabular form in Table 1. For comparison purposes, the same set of features was also computed using the PUSD skeleton, where the results are also listed in Table 1. The features obtained using the pseudo-Euclidean skeleton are more accurate than those obtained using the PUSD skeleton.

Table 1. Features of circular markers.

Features	Marker 1		Marker 2		Marker 3	
	Pseudo	PUSD	Pseudo	PUSD	Pseudo	PUSD
l_{max}	104.7	99.1	81.9	81.5	105.2	102.0
l_{min}	74.0	68.0	58.0	54.0	56.0	50.0
Center Pixel	(211,294)	(208,292)	(306,126)	(306,126)	(553,325)	(554,324)
End_1	(233,306)	(230,304)	(321,129)	(330,131)	(582,300)	(582,301)
End_2	(193,286)	(192,286)	(285,121)	(283,124)	(524,351)	(531,346)
r	1.42	1.46	1.41	1.51	1.88	2.04
θ	-26.6^o	-25.3^o	-12.5^o	-8.5^o	41.3^o	41.4^o

6. Conclusions

It has been shown that the pre-marking scheme based on circular markers can facilitate the 3-D recognition process of a object, especially when the object is defined by a selected set of standard-views, wher each standard-view is represented by a marker. The surface normals of the object can be easily extracted from its image by analyzing the features of the ellipses. Morphological skeleton transform provides a powerful tool in analyzing markers and extracting features of the markers. Among the various MST algorithms considered, the pseudo-Euclidean skeleton yields the most accurate results, since it is the least sensitive to rotation.

References

[1] M. L. Baird, "Future directions of Industrial Applications of Pattern Recognition (panel discussion)", *Proc., Fourth International Joint Conference on Pattern Recognition,* Kyoto, pp. 1146, 1978.

[2] J. W. Bales, and L. K. Barker, "Marking Parts to Aid Robot Vision", NASA Technical Paper 1819, 1981.

[3] R. Safaee-Rad, E. Shwedyk, and A. O. Quanbury, "Functional Human Arm Motion Study with a New 3D Measurement System (VCR-PIPEZ-PC)", *Proc., IEEE-EMBS Ninth Annual Conference,* pp. 401-411, 1987.

[4] H. F. L. Pinkney, and C. I. Perratt, "A Flexible Machine Vision Guidance System for 3-Dimensional Control Tasks", *Advances in CAD/CAM and Robotics, NRC of Canada Contributions",* 1987.

[5] J. Serra, *Image analysis and Mathematical Morphology,* Academic Press, New York, 1982.

[6] P. A. Maragos, and R. W. Schafer, "Morphological Skeleton Representation and Coding of Binary Images", *IEEE, Trans. on Acoustics, Speech Signal Processing,* Vol. ASSP-34, No. 5, pp. 1228-1244, 1986.

[7] Z. Zhou, and A. N. Venetsanopoulos, "Pseudo-Euclidean Morphological Skeleton Transform for Machine Vision", *Proc., IEEE, Int. Conf. on Acoustics, Speech Signal Processing,* Glasgow, Scotland, In Print, May 1989.

Acknowledgements

This research was partially supported by the Manufacturing Research Corporation of Ontario (MRCO).