Indian J. pure appl. Math., **25** (1&2); 95-111, January & February 1994

A SHAPE METRIC FOR 3-D OBJECTS*

DILIP K. BANERJEE, SWAPAN K. PARUI[1]

AND

D. DUTTA MAJUMDER

National Centre for Knowledge Based Computing, Indian Statistical Institute Calcutta 700 035

(Accepted 24 *June* 1993*)*

This paper examines the problem of constructing a similarity measure between shapes of 3-D objects. For this purpose a shape distance between 3-D objects is defined on the basis of certain characteristic planes of the objects. These planes normalize the orientation of the objects. After normalizing the position, size (which are straightforward) and orientation the shape distance is defined on the basis of the volume of mismatch between the two objects. A novel property of this distance is that it has a closed from solution and has all the metric properties.

1. INTRODUCTION

Recognition of 3-D objects has been a very important problem in computer vision. Besl and Jain[1] and others[2,3] give detailed surveys for object recongnition problems in two and three dimensions. The present paper deals with the shape of three dimensional objects and proposes a mehtod for quantifying shape similarity for them. This shape similarity measure may be used for classifying the shape of a new object as one of several known shapes, though the problem of defining shape similarity is much more fundamental than shape classification[4,5]. A 3-D object is defined to be a certain subset in the three dimensional space. Two objects have the same shape if and only if one is a translation, dilation and rotation of the other[6,7].

The approach taken here is an extension of an earlier work on 2-dimensional shape discrimination[8]. In this approach, for matching, an object has to be normalized in terms of position, size and orientation. Normalization of an object with respect to position and size can easily be achieved if its centre of mass and volume are considered. The main problem in shape matching involves normalization of the orientation of an object. The novelty in our approach is that the normalized orientation of an object comes as a closed form solution.

*This work has been partially supported by DoE/UNDP Project No. (IND/85/072).

[1]Electronics and Communication Sciences Unit.

The main issue here is that of shape similarity which is based on a volumetric representation of an object, namely voxel representation. Though it is well known that this representation is expensive in terms of computation time and storage space, it is quite useful, as can be seen later, in constructing a shape distance which has all the metric properties. (Our shape similarity measure is a strictly decreasing function of this shape distance). We do not intend to solve the problem of constructing the voxel representation of 3-D objects in this paper. There are several techniques available for this purpose[9,10,11,12]. We assume here that the voxel representation of a 3-D object is already available. We also assume that the objects are already segmented and single without occlusion.

We assume that the position and size of the objects are normalized in terms of their centre of mass and volume. We then define a shape distance which is the smallest volume of mismatch between two objects considering all possible relative orientations of theirs.

It can be seen that a closed form solution of this shape distance does not in general exist. So, we define another shape distance on the basis of the volume of mismatch which is a kind of approximation of the first but has a closed form solution. Both these distances satisfy all the metric properties. The first shape distance is not based on the normalized orientation of objects while the second one is. This normalized orientation of objects has a closed from solution. The normalization of the orientation is achieved on the basis of certain characteristic planes of 3-D objects. For two arbitrary objects, first these planes of one are aligned with those of the other and the volume of mismatch is found. This alignment is possible in four different orientations of one with respect to the other and the volume of mismatch is considered for each of these four orientations. Our shape distance is the minimum of these four volumes and at most five rotations in 3-D (which have closed form solutions) are needed to find this.

We have not come across any work on shape metric except in a paper by Shapiro and Haralick[13]. But our shape distance differs from theirs. Their definition is meant for shapes at a higher level of description while ours deals with shapes at a lower level. Also, their shape distance does not have a closed form solution while ours has. Moreover, our shape distance is rotation invariant while theirs is not.

Section 2 gives the formal definitions of 3-D objects and their shape. Section 3 provides a shape distance between 3-D objects considering all possible relative orientations. Section 4 gives the definitions of the characteristic planes of 3-D objects and a method of constructing them. Section 5 provides a shape distance between 3-D objects on the basis of the characteristic planes, which has a closed form solution. Section 6 gives the computational techniques for the shape distance in the digital case and some results on real life objects. Conclusions are given in Section 7.

2. Shape of 3-D Objects

Definition 2.1 — A subset A in $\mathbf{R}^3$ (3-dimensional Euclidean space) is called an object if

(i) A is compact,

(ii) Interior (A) is non-empty and connected,

(iii) Closure (Interior (A)) = A.

Notation — **F** is the class of all objects.

Definition 2.2 — For two subsets A and B in $\mathbf{R}^3$, B is a translation of A if there exist real numbers a, b, c such that

$$B = \{(x+a,\ y+b,\ z+c) : (x,y,z) \in A\}.$$

Definition 2.3 — For two subsets A and B in $\mathbf{R}^3$, B is a dilation of A if there exists a positive real k such that,

$$B = \{(kx, ky, kz) : (x,y,z) \in A\}.$$

Now any rotation of an object in $\mathbf{R}^3$ about the origin is determined by the corresponding rotation of the system of axes which in turn is specified by three angles α, β and τ in $[0, 2\pi)$ where α is the angle between the new and old x-axes, β is the angle between the new and old-y-axis and τ is the angle between the new and old z-axes. The rotations here are in the anticlockwise direction.

Definition 2.4 — B is a *rotation* of A around the origin if there exist α, β and τ in $[0, 2\pi)$ such that

$$B = \{P_3(\tau) P_2(\beta) P_1(\alpha) X : X' \in A\}$$

where

$$P_1(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \text{Cos}\,\alpha & -\text{Sin}\,\alpha \\ 0 & \text{Sin}\,\alpha & \text{Cos}\,\alpha \end{bmatrix}$$

$$P_2(\beta) = \begin{bmatrix} \text{Cos}\,\beta & 0 & -\text{Sin}\,\beta \\ 0 & 1 & 0 \\ \text{Sin}\,\beta & 0 & \text{Cos}\,\beta \end{bmatrix}$$

$$P_3(\tau) = \begin{bmatrix} \text{Cos}\,\tau & -\text{Sin}\,\tau & 0 \\ \text{Sin}\,\tau & \text{Cos}\,\tau & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and $X' = (x, y, z)$. P_1, P_2 and P_3 represent rotations around the x-, y- and z-axes respectively.

Proposition 2.1 — Each of translation, dilation and rotation defines an equivalence relation on **F**. These equivalence relations are denoted by R_1, R_2, R_3 respectively.

Definition 2.5 — R is a relation on **F** such that for A, B in **F**, there exist C and D in **F** such that (A, C) belongs to R_1, (C, D) belongs to R_2 and (D, B) belongs to R_3.

Proposition 2.2 — If A, B, C belonging to **F** are such that (A, B) $\in R_i$ and (B, C) $\in R_j$, then there exists a D in **F** such that (A, D) $\in R_j$ and (D, C) $\in R_i$ for $3 \geq i \neq j \geq 1$.

Proposition 2.3 — R is an equivalence relation on $\mathbf{F}$.

PROOF : *Reflexivity* : $(A, A) \in R$ since $(A, A) \in R_i$ for $i = 1, 2, 3$.

Symmetry : This will be proved by repeated applications of Propostition 2.2.

$$(A, B) \in R$$

$$\Rightarrow (A, C) \in R_1, \; (C, D) \in R_2 \text{ and } (D, B) \in R_3 \text{ for some } C, D \text{ in } \mathbf{F}$$

$$\Rightarrow (A, C_1) \in R_2, \; (C_1, D) \in R_1 \text{ and } (D, B) \in R_3 \text{ for some } C_1 \text{ in } \mathbf{F}$$

$$\Rightarrow (A, C_1) \in R_2, \; (C_1, D_1) \in R_3 \text{ and } (D_1, B) \in R_1 \text{ for some } D_1 \text{ in } \mathbf{F}$$

$$\Rightarrow (A, C_2) \in R_3, \; (C_2, D_1) \in R_2 \text{ and } (D_1, B) \in R_1 \text{ for some } C_2 \text{ in } \mathbf{F}$$

$$\Rightarrow (B, D_1) \in R_1, \; (D_1, C_2) \in R_2 \text{ and } (C_2, A) \text{ in } R_3$$

$$\Rightarrow (B, A) \in R.$$

Transitivity : $(A, B) \in R$ and $(B, C) \in R$

$$\Rightarrow (A, D) \in R_1, \; (D, E) \in R_2, \; (E, B) \in R_3 \text{ and}$$
$$(B, D_1) \in R_1, \; (D_1, E_1) \in R_2, \; (E_1, C) \in R_3$$
$$\text{for some } D, E, D_1 \text{ and } E_1 \text{ in } \mathbf{F}$$

$$\Rightarrow (A, D) \in R_1, \; (D, E) \in R_2, \; (E, B_1) \in R_1, \; (B_1, D_1) \in R_3,$$
$$(D_1, E_1) \in R_2, \; (E_1, C) \in R_3 \text{ for some } B_1 \text{ in } \mathbf{F}$$

$$\Rightarrow (A, D) \in R_1, \; (D, E_1) \in R_1, \; (E_1, B_1) \in R_2, \; (B_1, D_2) \in R_2,$$
$$(D_2, E_1) \in R_3, \; (E_1, C) \in R_3 \text{ for some } E_1, D_2 \text{ in } \mathbf{F}$$

$$\Rightarrow (A, E_1) \in R_1, \; (E_1, D_2) \in R_2, \; (D_2, C) \in R_3$$

$$\Rightarrow (A, C) \in R.$$

Definition 2.6 — The *shape* of an object is defined as an equivalence class generated by R on $\mathbf{F}$.

Notation — $\mathbf{S}$ is the family of all such equivalence classes, that is, of all shapes.

In other words, within each such equivalence class, for any two objects A and B, B can be obtained from A through translation, dilation and rotation. On the other hand, for any two objects, A, B from two different equivalence classes, B can never be obtained from A through these operations.

Notation — $\mathbf{F_1}$ is a subclass of $\mathbf{F}$ such that each object belonging to $\mathbf{F_1}$ has unit volume and has (0, 0, 0) as its centre of mass. That is, for A in $\mathbf{F_1}$,

$$\int_A da = 1 \text{ and } \int_A x\, da = \int_A y\, da = \int_A z\, da = 0,$$

where $\int_A$ means Riemann integration of the indicator function of Ar x, y, z and denote the coordinates of the points of A.

Notation — $\mathbf{S}_1$ is the family of the equivalence classes generated by R_3 on $\mathbf{F}_1$.

Remark : It is clear that for every equivalence class **C** of **S**, there is a unique $\mathbf{C}_1$ of $\mathbf{S}_1$ such that **C** contains $\mathbf{C}_1$. In fact, there is a one-to-one correspondence between **S** and $\mathbf{S}_1$. From now on we will be considering $\mathbf{S}_1$ instead of **S** since $\mathbf{S}_1$ contains all shapes of **S**.

3. Shape Distance

Definition 3.1 — D_1 is a distance function on $\mathbf{F}_1$ such that for A, B in $\mathbf{F}_1$,

$$D_1(A,B) = m_3 [(A-B) \bigcup (B-A)]$$

where m_3 is the Lebesgue measure in $\mathbf{R}^3$.

Proposition 3.1 — D_1 defines a metric on $\mathbf{F}_1$. That is, for A, B, C in $\mathbf{F}_1$,

(i) $D_1(A,B) \geq 0$

(ii) $D_1(A,B) = 0$ iff $A = B$

(iii) $D_1(A,B) = D_1(B,A)$

(iv) $D_1(A,B) + D_1(B,C) \geq D_1(A,C)$.

Proof : (i) Trivial.

(ii) For $A = B$, $D_1(A,B) = 0$.

Conversely, let $D_1(A,B) = 0$. Now, if $A \neq B$, then at least either of $(A - B)$ or $(B - A)$ is nonempty. Without loss of generality assume $A - B$ is nonempty. Note $A - B = A \bigcap B^C$ and $\text{Int}(A) - B = (\text{Int}(A)) \bigcap B^C$. Also, $\text{Int}(A)$ and B^C are open sets.

Now if E and F are disjoint open sets, then $\text{Cl}(E)$ and F are disjoint. (Int and Cl denote interior and closure respectively of a set). So $A - B$ is nonempty implies $\text{Int}(A) - B$ is nonempty, since $A = \text{Cl}(\text{Int}(A))$. But since $\text{Int}(A) - B$ is an open set, the volume of $\text{Int}(A) - B$ is positive. So, $m_3(A-B) > 0$ and hence $D_1(A,B) > 0$ which is a contradiction.

(iii) Trivial.

(iv) $m_3 [(A-B) \bigcup (B-A)] + m_3 [(B-C) \bigcup (C-B)]$

$$= m_3(A-B) + m_3(B-A) + m_3(B-C) + m_3(C-B)$$

$$= m_3 [(A-B) \bigcup (B-C)] + m_3 [(B-A) \bigcup (C-B)]$$

$$\geq m_3(A-C) + m_3(C-A)$$

(since $(A - C)$ and $(C - A)$ are subsets of $(A - B) \bigcup (B - C)$ and $(B - A) \bigcup (C - B)$ respectively)

$= m_3 [(A - C) \bigcup (C - A)]$

Hence $D_1 (A, B) + D_1 (B, C) \geq D_1 (B, C)$.

Notation : For any A in **F** and any angles $\alpha, \beta, \tau, A_{\alpha, \beta, \tau}$ denotes the rotation of A by α, β, τ around the origin. It is clear that A is in **F** implies $A_{\alpha, \beta, \tau}$ is in **F** and A is in $\mathbf{F_1}$ implies $A_{\alpha, \beta, \tau}$ is in $\mathbf{F_1}$.

Also, $(A_{\alpha, \beta, \tau})_{\Theta, \Phi\, \delta} = A_{(\alpha + \Theta), (\beta + \Phi), (\tau + \delta)}$

and $Ar = A_{2\pi, 2\pi, 2\pi} = A_{0, 0, 0}$.

Proposition 3.2 — $D_2 (A, B) = D_1 (A_{\alpha, \beta, \tau}, \beta_{\alpha, \beta, \tau})$ for any α, β and τ.

Definition 3.2 — D_2 is a distance function on $\mathbf{F_1}$ such that for A, B in $\mathbf{F_1}$,

$$D_2(A, B) = \inf_{\alpha, \beta, \tau} D_1 (A, B_{\alpha, \beta, \tau}).$$

Propostion 3.3 — $D_1 (A, B) = D_1 (A, B_{\alpha_1, \beta_1, \tau_1})$ for some $\alpha_1, \beta_1, \tau_1$ in $[0, 2\pi)$.

PROOF : Let $f(\alpha, \beta, \tau) = D_1 (A, B_{\alpha, \beta, \tau})$ which is a function defined on $V = [0, 2\pi] \times [0, 2\pi] \times [0, 2\pi]$. Since f is continuous and bounded in V, $W = \{f(\alpha, \beta, \tau) : (\alpha, \beta, \tau) \in V\}$ is a closed set. So $\inf_{\alpha, \beta, \tau} f(\alpha, \beta, \tau)$ belongs to W.

Thus, $\inf_{\alpha, \beta, \tau} f(\alpha, \beta, \tau) = f(\alpha_1, \beta_1, \tau_1)$ for some $\alpha_1, \beta_1, \tau_1$.

Propostion 3.4 — D_2 defines a metric on $\mathbf{S_1}$. That is, for A, B and C in $\mathbf{F_1}$,

(i) $D_2 (A, B) \geq 0$

(ii) $D_2 (A, B) = 0$ if and only if $A = B_{\alpha, \beta, \tau}$ for some α, β, τ.

(iii) $D_2 (A, B) = D_2 (B, A)$

(iv) $D_2 (A, B) + D_2 (B, C) \geq D_2 (A, C)$

PROOF : (i) Trivial.

(ii) Let $A = B_{\alpha, \beta, \tau}$. Then $D_1 (A, \beta_{\alpha, \beta, \tau}) = 0$. Hence $D_2 (A, B) = 0$. Conversely, let $D_2 (A, B) = 0$. So from Proposition 3.3, there exist α, β, τ such that $D_1 (A, \beta_{\alpha, \beta, \tau}) = 0$. Thus from Proposition 3.1, $A = B_{\alpha, \beta, \tau}$.

(iii)
$$\begin{aligned} D_2 (A, B) &= \inf_{\alpha, \beta, \tau} D_1 (A, B_{\alpha, \beta, \tau}) \\ &= \inf_{\alpha, \beta, \tau} D_1 (B_{\alpha, \beta, \tau}, A) \text{ (since } D_1 \text{ is symmetric)} \\ &= \inf_{\alpha, \beta, \tau} D_1 (B, A_{2\pi - \alpha, 2\pi - \beta, 2\pi - \tau}) \text{ (from Proposition 3.2)} \\ &= \inf_{\Theta, \Phi, \delta} D_1 (B, A_{\Theta, \Phi, \delta}) \\ &= D_2 (B, A). \end{aligned}$$

(iv) $D_2(A,B)+D_2(B,C) = D_1(A,B_{\alpha,\beta,\tau})+D_1(B,C_{\Theta,\Phi,\delta})$ for some α,β,τ and Θ,Φ,δ.

$= D_1(A,B_{\alpha,\beta,\tau})+D_1(B_{\alpha,\beta,\tau}, C_{\alpha+\Theta,\beta+\Phi,\tau+\delta})$ (from Proposition 3.2)

$\geq D_2(A, C_{\alpha+\Theta,\beta+\Phi,\tau+\delta})$ (from (iv) of Proposition 3.1)

$\geq D_2(A,C)$ as $D_2(A,C)$ is $\operatorname{Inf}_{\mu,\sigma,\phi} D_1(A,C_{\mu,\sigma,\phi})$.

Hence D_2 defines a metric on $\mathbf{S}_1$.

4. Characteristic Planes and Axes of an Object

In order to normalize the orientation of an object in **F**, three certain characteristic planes of the object are considered as follows :

Definition 4.1 — The *major plane* of an object is the one which minimizes the integral of squared perpendicular distances of the object points to the plane. (Its derivation is given later in the section).

Definition 4.2 — The *minor plane* is the one which maximizes the above integral of squared perpendicular distances.

The above two planes, it can be seen, pass through the centre of mass of the object ans are perpendicular to each other.

Definition 4.3 — The *intermediate plane* is the one that passes through the centre of mass of the object and is perpendicular to both the major and minor planes.

Definition 4.4 — The *major axis* of an object is the one which minimizes the integral of squared perpendicular distances of the object points to the line.

Definition 4.5 — The *minor axis* is the one which maximizes the above integral of squared perpendicular distances.

The above two lines pass through the centre of mass of the object and are perpendicular to each other.

Definition 4.6 — The *intermediate axis* is the one that passes through the centre of mass of the object and is perpendicular to both the major and minor axes.

A method of finding the major, minor and intermediate planes of an object on the basis of its points is given below.

The equation of a plane in 3 dimensions may be given by

$$x\cos\alpha + y\cos\beta + z\cos\tau - p = 0$$

where $\cos\alpha$, $\cos\beta$, $\cos\tau$ are the direction cosines of the normal to the plane from the origin and p is the length of that normal. The perpendicular distance from an arbitrary point (x, y, z) to the above plane is given by

$$| x\cos\alpha + y\cos\beta + z\cos\tau - p |.$$

The integral of the squared perpendicular distances of all points of object A to the plane is given by

$$\int_A (x \operatorname{Cos} \alpha + y \operatorname{Cos} \beta + z \operatorname{Cos} \tau - p)^2 \, da. \qquad \text{... (1)}$$

We would like to find a plane (characterized by β Cos α, Cos beta, Cos τ, *p*) which will minimize the above integral of squared perpendicular distances. The partial derivatives of this integral are used for minimization. We have the following restriction on the direction cosines :

$$\operatorname{Cos}^2 \alpha + \operatorname{Cos}^2 \beta + \operatorname{Cos}^2 \tau = 1.$$

When the above integral of squared perpedicular distances is minimized with respect to *p* we get

$$p = \bar{x} \operatorname{Cos}^2 \alpha + \bar{y} \operatorname{Cos}^2 \beta + \bar{z} \operatorname{Cos}^2 \tau \qquad \text{... (2)}$$

where $\bar{x} = \int_A x \, da, \; \bar{y} = \int_A y \, da, \; \bar{z} = \int_A z \, da.$

(Recall that $\int_A da = 1$).

Putting the above value of *p* in (1) we get the following :

$$\int_A [\, (x - \bar{x}) \operatorname{Cos} \alpha + (y - \bar{y}) \operatorname{Cos} \beta + (z - \bar{z}) \operatorname{Cos} \tau \,]^2 \, da \qquad \text{... (3)}$$

and minimizing (3) with respect to Cos α, Cos β, Cos τ, we get the following normal equations :

$$(S_{xx} + \lambda) \operatorname{Cos} \alpha + S_{xy} \operatorname{Cos} \beta + S_{xz} \operatorname{Cos} \tau = 0 \qquad \text{... (4)}$$

$$S_{xy} \operatorname{Cos} \alpha + (S_{yy} + \lambda) \operatorname{Cos} \beta + S_{yz} \operatorname{Cos} \tau = 0 \qquad \text{... (5)}$$

$$S_{xz} \operatorname{Cos} \alpha + S_{yz} \operatorname{Cos} \beta + (S_{zz} + \lambda) \operatorname{Cos} \tau = 0 \qquad \text{... (6)}$$

where λ is the Lagrange's multiplier and

$$S_{xx} = \int_A (x - \bar{x})^2 \, da, \quad S_{xy} = \int_A (x - \bar{x})(y - \bar{y}) \, da \text{ and so on.}$$

The set of equations (4), (5), (6) can be rewritten in a matrix form as $[S + \lambda I]\, \mathbf{e} = \mathbf{0}$ ... (7)

where S, the scatter matrix = $\begin{bmatrix} S_{xx} & S_{xy} & S_{xz} \\ S_{xy} & S_{yy} & S_{yz} \\ S_{xz} & S_{yz} & S_{zz} \end{bmatrix}$.

I is the identify matrix of order three and

$$\mathbf{e} = \begin{bmatrix} \operatorname{Cos} \alpha \\ \operatorname{Cos} \alpha \\ \operatorname{Cos} \tau \end{bmatrix} \text{ and } \mathbf{O} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

We have to find a solution for ***e*** from (7) with the restriction

$\mathrm{Cos}^2\alpha + \mathrm{Cos}^2\beta + \mathrm{Cos}^2\tau = 1$. For e to have any non null solution, $[S + \lambda I]$ must be singular i.e. $-\lambda$ is an eigen value of S. In fact, $|S + \lambda I| = 0$ gives a cubic equation in λ, and for each of the three solutions of λ, there will be a corresponding eigen vector as the solution of λ, there will be a corresponding eigen vector as the solution of e. Suppose the eigen values are l_i (i.e. $-\lambda_i$) and the corresponding eigen vectors are e_i ($i = 1, 2, 3$). Without loss of generality suppose, $l_1 \geq l_2 \geq l_3$. Then for (l_1, e_1), (3) will be minimized and for (l_3, e_3), (3) will be maximized. The major and minor planes will be determined by e_1 and e_3 respectively.

The characteristic planes of an object in **F** are invariant under rotation in the following sense.

Proposition 4.1 — If an object is rotated by (α, β, τ) in 3 dimensions, each of the major and minor planes of the object is rotated by the same angles (α, β, τ).

PROOF : The integral of squared perpendicular distances from all points of A to a plane characterized by Cos σ, Cos θ, Cos ϕ and passing through the centre of mass of A is given by :

$$\int_A [(x-\bar{x})\,\mathrm{Cos}\,\sigma + (y-\bar{y})\,\mathrm{Cos}\,\theta + (z-\bar{z})\,\mathrm{Cos}\,\phi]^2\, da.$$

Denoting $x_0 = \begin{bmatrix} x & -\bar{x} \\ y & -\bar{y} \\ z & -\bar{z} \end{bmatrix}$, $W_1 = \begin{bmatrix} \mathrm{Cos}\,\sigma \\ \mathrm{Cos}\,\theta \\ \mathrm{Cos}\,\phi \end{bmatrix}$

the above expression for integral of squared perpendicular distances

can be written as $f(X_0, W_1) = \int_A (X_0^1, W_1)^2\, da$.

If B is a rotation of A by (α, β, τ), then

$B = \{TX : X' \text{ in } A\}$, where T is the rotation matrix (which is, in fact, orthonormal). Then the integral of the squared perpendicular distances from all points of B to a plane passing through the centre of mass of B is

$$f(B, W_2) = \int_B [(TX_0)'\, W_2]^2\, db.$$

where $W_2 = \begin{bmatrix} \mathrm{Cos}\,\sigma' \\ \mathrm{Cos}\,\theta' \\ \mathrm{Cos}\,\phi' \end{bmatrix}$

characterizes the plane by the direction cosines.

But $f(B, W_2) = \int_A (X_0{}'T'\, W_2)^2\ |J(T)|\, da$

where $J(T)$ is the jacobian of the orthonormal matrix T.

Then $f(B, W_2) = \int_A (X_0{}'T'\, W_2)^2\, da$ as $|J(T)| = 1$

$$= \int_A (X_0' W_3)^2 \, da,$$

denoting $T' W_2$ by W_3 which is a direction cosine vector.

$$= f(A, W_3)$$

Now let W_{10} characterize the major plane of A,

i.e. $f(A, W_1)$ attains infimum at W_{10}. In other words,

$$\operatorname*{Inf}_{W_1} f(A, W_1) = f(A, W_{10}).$$

Now $f(B, W_2)$ is infimum iff $f(A, W_3)$ is infimum,

$$\text{i.e. } f(B, W_2) \text{ is infimum iff } \quad W_3 = W_{10}$$

$$\text{iff } T' W_2 = W_{10}$$

$$\text{iff } \quad W_2 = TW_{10}.$$

Hence, it is proved that if the object A has the major plane initially characterized by W_{10} and A is rotated by (α, β, τ), then the new major plane will also be rotated by the same angles (α, β, τ). Similarly, it can be shown that the same result holds for the minor plane also.

For some objects, the major and minor planes will not be unique (for example, spheres, cubes, regular polyhedrons etc.). These objects are called *regular* for which at least two of the eigen values are the same. From now on we will exclude such objects. Let $\mathbf{F}_2$ be the subclass of $\mathbf{F}_1$ such that all the three eigen values of each object in $\mathbf{F}_2$ are distintict and for each object in $\mathbf{F}_1 - \mathbf{F}_2$ at least two eigen values are the same. Let $\mathbf{S}_2$ be the set of all equivalence classes generated by R_3 on $\mathbf{F}_2$. Note that $\mathbf{S}_2$ is a proper subset of $\mathbf{S}_1$, the difference being the equivalence classes containing regular objects. In the next section, we will deal with the objects of $\mathbf{F}_2$ only.

5. Shape Distance on the Basis of the Major and Minor Planes

The distance function (of shape) $D_2(A, B)$ considers all possible rotations of the object B. The computation to find the value of D_2 in this case will be costly in terms of time. So we propose another definition of shape distance D_3 on the basis of the major plane and the minor plane for objects belonging to $\mathbf{F}_2$.

We normalize the orientation of the object A by rotating it about the origin so that the major plane of A coincides with the x-y plane and the minor plane of A coincides with the y-z plane. To achieve this desired orientation of A, say A_1, we need a 3-dimensional rotation matrix T_1 such that $A_1 = T_1 A$. Suppose, the same normalizing orientation of B is achieved by rotation matrix T_2 and $B_1 = T_2 B$. Now let E_1, E_2 and E_3 be the rotation matrices which rotate an object by 180° about the $x-$, $y-$ and $z-$ axes respectively. Now note that the major and minor planes of $A_1, E_1 A_1, E_2, A_1$, $E_3 A_1$, $B_1, E_1 B_1$, $E_2 B_1$ and $E_3 B_1$ are the x-y plane and the y-z plane respectively. That is, all these four orientations of each of A and B are normalized. In fact, there is no other orientation of A or B which is normalized.

Definition 5.1 — $D_3(A,B) = \underset{i,j=1,\dots,4}{\text{Min}} \; D_1(E_i A_1, \; E_j B_1)$ where $E_4 = I$.

Propostition 5.1 — $D_3(A,B) = \underset{j=1,\dots,4}{\text{Min}} \; D_1(A_1, \; E_j B_1)$

PROOF : Note that $E_iE_i = I$ for all $i = 1,\dots, 4$ and $E_iE_j = E_k$ for some k for all $i \neq j$. From Propostion 3.2,

$$D_1(E_iA_1, E_jB_1) = D_1(A_1, E_iE_jB_1) = D_1(A_1, E_kB_1).$$

Hence the proposition.

It can be seen that D_3 is unambiguously defined and one orientation of A and four orientations of B are needed to obtain D_3. Let $d_1 = D_1(A_1, B_1)$, $d_2 = D_1(A_1, E_1B_1)$, $d_3 = D_1(A_1, E_2B_1)$ and $d_4 = D_1(A_1, E_3B_1)$. So, $D_3(A,B) = \underset{i=1,\dots 4}{\text{Min}} \; d_i$

Proposition 5.2 — D_3 defines a metric on shapes in $\mathbf{S}_2$, that is, for A, B, C in $\mathbf{F}_2$,

(i) $D_3(A,B) \geq 0$

(ii) $D_3(A,B) = 0$ if and only if A is a rotation of B

(iii) $D_3(A,B) = D_3(B,A)$

(iv) $D_3(A,B) + D_3(B,C) \geq D_3(A,C)$

PROOF : (i) Trivial.

(ii) Let $D_3(A,B) = 0$ i.e. at least one of d_1, d_2, d_3, d_4 is zero.

Case 1 : $d_1 = 0$ i.e. $D_1(A_1, B_1) = 0$

Now $D_1(A_1, B_1) = 0 \Rightarrow A_1 = B_1$ (from Proposition 3.1)

$\Rightarrow T_1A = T_2B \Rightarrow A = (T_1)^{-1} T_2 B$

$\Rightarrow$ A is a rotation of B.

Case 2 : $d_2 = 0$ i.e. $D_1(A_1, E_1B_1) = 0$.

Now $D_1(A_1, E_1B_1) = 0 \Rightarrow A_1 = E_1 B_1 \Rightarrow T_1A = E_1 T_2 B$

$\Rightarrow A = (T_1)^{-1} E_1 T_2 B \Rightarrow$ A is a rotation of B.

Similarly, for the other two cases, namely, $d_3 = 0$, $d_4 = 0$, it can be shown that A is a rotation of B. Conversely, suppose A is a rotation of B. So, $A = TB$.

Now, $D_1(A, TB) = 0 \Rightarrow D_1(T_1 A, \; T_1 TB) = 0$

(Recall T_1 normalized the orientation of A and $A = TB$)

$$\Rightarrow D_1(A_1, T_2 B) = 0 \text{ (because } T_1 T \text{ can be taken as } T_2)$$
$$\Rightarrow \acute{D}_1(A_1, B_1) = 0 \Rightarrow D_3(A, B) = 0.$$

(iii) $D_3(A, B)$

$$= \text{Min } \{D_1(A_1, B_1), D_1(A_1, E_1, B_1), D_1(A_1, E_2 B_1), D_1(A_1, E_3 B_1)\}$$

$$= \text{Min } \{D_1(B_1, A_1), D_1(E_1 B_1, A_1), D_1(E_2 B_1, A_1), D_1(E_3 B_1, A_1)\}$$

$$= \text{Min } \{D_1(B_1, A_1), D_1(B_1, E_1 A_1), D_1(B_1, E_2 A_1), D_1(B_1, E_3 A_1)\}$$

$$= D_3(B, A)$$

(iv) Triangular inequality :

$$D_3(A, B) = \text{Min } \{D_1(A_1, B_1), D_1(A_1, E_1 B_1), D_1(A_1, E_2 B_1), D_1(A_1, E_3 B_1)\}$$

and

$$D_3(B, C) = \text{Min } \{D_1(B_1, C_1), D_1(B_1, E_1 C_1), D_1(B_1, E_2 C_1), D_1(B_1, E_3 C_1)\}$$

Case 1 : Let $D_3(A, B) = D_1(A_1, B_1)$ and $D_3(B, C) = D_1(B_1, C_1)$.

Then $D_3(A, B) + D_3(B, C) = D_1(A_1, B_1) + D_1(B_1, C_1)$

$$\geq D_1(A_1, C_1) \geq D_3(A, C).$$

Case 2 : Let $D_3(A, B) = D_1(A_1, E_3 B_1)$ and

$$D_3(B, C) = D_1(B_1, E_2 C_1).$$

Then $D_3(A, B) + D_3(B, C) = D_1(A_1, E_3 B_1) + D_1(B_1, E_2 C_1)$

$$= D_1(A_1, E_3 B_1) + D_1(E_3 B_1, E_1 C_1)$$

(since $E_3 E_2 = E_1$)

$$\geq D_1(A_1, E_1 C_1) \geq D_3(A, C).$$

Similarly, for the other cases, the triangular inequality can be proved. Hence D_3 defines a metric on $\mathbf{S}_2$.

Definition 5.2 — For two objects A, B in $\mathbf{F}_2$, the *shape similarity measure* is defined as

$$\mu(A, B) = 1 - D_3(A, B)/2 \qquad \text{... (8)}$$

Note that $0 \leq \mu \leq 1$ since $0 \leq D_3 \leq 2$. μ indicates how closely the shapes of two objects resemble each other.

6. Computational Techniques and Results

Here the volumetric representation of a 3-D object is given by a 3-D binary matrix $M(i, j, k)$, that is, $M(i, j, k)$ is either 0 or 1. The centre of mass of the object is computed as (i', j', k') where

$$i' = \Big(\sum_{M(i,j,k)=1} i\Big)/n, \quad j' = \Big(\sum_{M(i,j,k)=1} j\Big)/n, \quad k' = \Big(\sum_{M(i,j,k)=1} k\Big)/n$$

and $n = \sum_{M(i,j,k)=1} 1 =$ the volume of the object.

The centre of mass is needed to normalize the position of an object. To make the volumes of two objects equal, the smaller object is expanded with proper scaling factors. Rotation operation is needed to normalize the orientation of an object. The details of the algorithm used here for expansion and rotation of 3-D binary matrics (i.e. objects) are available elsewhere[14]. The factor of this expansion algorithm is arbitrary in the sense that it need not be an integer but can be any rational number so that an object with any volume can be reduced or enlarged to any other volume, Also, the factor of expansion along each of the three axes is the same so that the shape of the object remains undisturbed through expansion. The rotation algorithm deals not only with angles that are multiples of 45° but with any arbitrary angles in 3 dimensions and also preserves the shape of 3-D objects.

Suppose A and B are two 3-D objects whose shape distance (D_3) is to be computed. Let A and B have volumes m and n. Without loss of generality, suppose $m > n$. Now, A and B are translated to A_1 and B_1 respectively so that they have their centres of mass at (0, 0, 0). Then, B_1 is expanded to B_2 such that the volume of B_2 becomes (approximately) m. Now, A_1 and B_2 are rotated to A_2 and B_3 respetively such that the major planes of A_2 and B_3 coincide with the x-y plane and their minor planes with the y-z planes. Let

$$B_4 = E_1 B_3, \; B_5 = E_2 B_3, \; B_6 = E_3 B_3$$

For two binary matrices M and N of equal size, define $d'(M, N)$ as the number of voxels (i, j, k) such that $M(i, j, k)$ and $N(i, j, k)$ are unequal. Suppose, M_1, N_1, N_2, N_3, N_4 are the corresponding binary matrices of A_2, B_3, B_4, B_5 and B_6 respectively. Let

$$d(A, B) = \text{Min}\{d'(M_1, N_1), d'(M_1, N_2), d'(M_1, N_3), d'(M_1, N_4)\}/v \quad \dots (9)$$

where division by v is for normalization and the value of v is the average of the numbers of 1-voxels in M_1 and N_1. Note that d is the discrete form of D_3. Below we use the above techniques for shape discrimination of four real life objects.

We consider for our study a certain class of three dimensional objects, namely extruded objects. An extruded object can be generated by propagating its stable cross-section. A large number of industrial objects fall into this class. Four such industrial tools (Figure 1), namely, a flat file, a spanner, a tool bit and a V-block are studied for shape discrimination using the above techniques.

The three dimensional image for each of these objects is generated by capturing its stable cross-section and then propagating it along an axis perpendicular to the cross-section. The image of the stable cross-section of the flat file is of size 140 × 20 pixels. This image is propagated in 5 layers giving us a three dimensional image of size 140 × 20 × 5 voxels. The length, breadth and thickness of the objects, the size of the image of their stable cross sections, the size of their three dimensional images, and their volumes (in total number of object voxels) are shown in Table I. From Figure 1 and Table I it is clear that there is a relatively large variation in the relative dimensions of the V-block against the rest. Also, the volume of the V-block is significantly greater than the others.

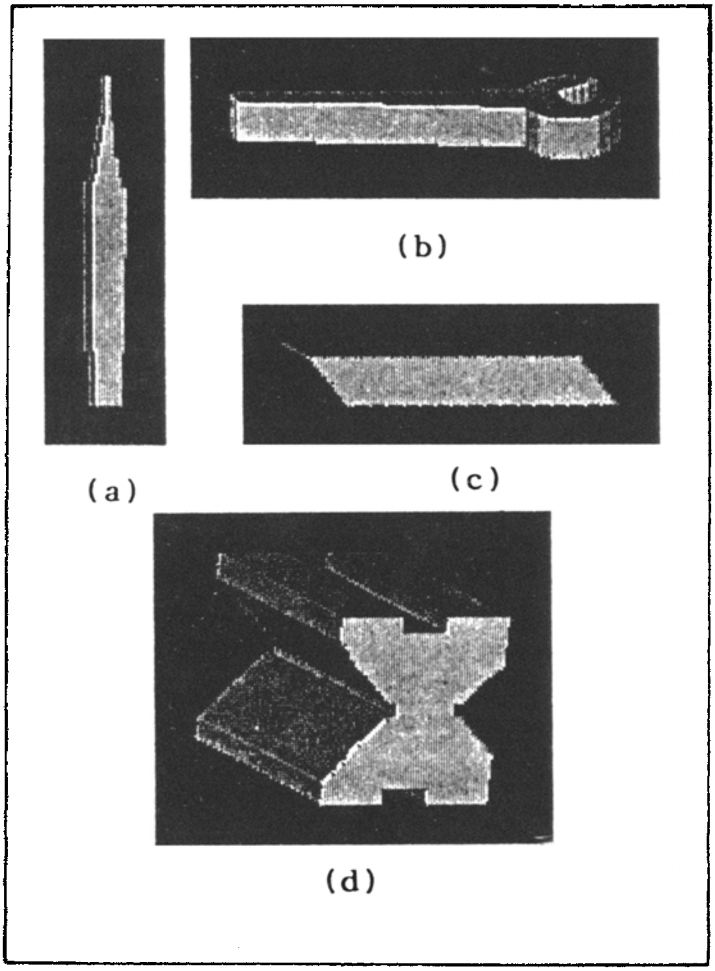

FIG. 1 Industrial objects a) Flat file b) Spanner c) Tool bit d) V-block

For shape matching the volumes of the smaller objects namely the flat file, the spanner and the tool bit are expanded by the factors shown in Table II. For example, the factor of expansion for the flat file is computed by taking the cube root of the ratio of the olume of the V-block (441774 voxels) and the volume of the flat file (15860 voxels). These factors of expansion are the same for each of the three dimensions to preserve the shape of the object. After expansion, their volumes become nearly equal to the volume of the V-block. The difference in volume of the expanded shapes and the volume of the V-block is due to discretization.

Table I Industrial Objects and their details

Objects	Length, Breadth and Thickness (in mm)	Size of cross section image (pixels)	Size of 3-D image (voxels)	Volume (voxels)
V-block	51, 39, 38	80 × 80	80 × 80 × 108	441774
Flat file	136, 12, 3	140 × 20	140 × 20 × 5	15860
Spanner	156, 46, 4	160 × 50	160 × 50 × 4	48960
Tool-bit	70, 13, 11	150 × 30	150 × 30 × 30	59532

Table II Factors of expansion and the expanded volumes

Objects	Factors of Expansion	Expanded Volume
V-block	1.000	441774
Flat file	3.031	438621
Spanner	2.082	460450
Tool bit	1.950	476256

Table III Shape distance (d) and shape similarity (μ)

	V-Block	Flat file	Spanner	Tool bit
V-block	-	1.813 .094	1.823 .088	1.674 .163
Flat file		-	1.522 .234	1.194 .403
Spanner			-	1.402 .299
Tool bit				-

Now, the 3D images of the V-block and the three expanded objects are shifted in a larger frame of size 430 × 110 × 110 so that each of these objects has their centre of gravity at (215, 55, 55) and then the images are rotated about the centre of gravity so that their major and minor planes coincide. The shape distance d in[9] and the shape similarity measure μ in[8] are shown in Table III. It can be seen that the shape distance between the V-block and any other object is quite large compared to the shape distances between other pairs. That is if we consider two clusters of objects in terms of their shapes, the V- block falls in one cluster and the other three in the other. This agrees with human visual perception. Finally, all of these objects can be discriminated on the basis of the shape distance.

7. Conclusions

A shape similarity measure μ has been defined for 3-D objects on the basis of their characteristic planes. An excellent property of these planes is that they normalize the orientation of the objects and have a close form solution. The measure μ can be useful in shape based object recognition. It can be seen that μ may not be suitable for shape description for complex objects. In such cases, the objects can be decomposed into relatiely simple parts and μ can be used for each of these parts[15, 16].

The shape distance D_1 can be useful for symmetry analysis of 3-D objects. It has been shown that if an object is symmetric about a plane P then P coincides with one of the characteristic planes of the object[17]. Suppose the reflections of an object A about its three characteristic planes are A_1, A_2 and A_3. Then

$$\text{dev}(A) = \text{Min}\{D_1(A, A_1),\ D_1(A, A_2),\ D_1(A, A_3)\}$$

gives a measure of deviation from symmetry. Clearly, $\text{dev}(A) = 0$ for symmetric A. Also, for a nearly symmetric object A if $\text{dev}\ (A) = D_1(A, A_i)$, then the characteristic plane corresponding to A_i gives an approximate plane of symmetry of A.

So far we have considered shape to be invariant under translation, dilation and rotation. The definition of shape can be extended by making shape invariant under reflection also. In fact,

$$D_4(A, B) = \text{Min}(D_3(A, B),\ D_3(A,\ \text{refl}(B))\}$$

is a shape distance which is invariant under reflection, where A and B are two 3-D objects and $\text{refl}(B)$ is the reflection of B around any plane.

As described earlier in the paper, the derivations of the characteristic planes involve computations of second order statistics. These computations involve all the voxles of the object and hence are of order $O(n^3)$, where n is the side length of the object along each of the three axes. These computations can possibly be reduced by considering only the surface voxels instead of all the voxels of the object. The order of complexity in that case will be $O(n^2)$. But finding the characteristic planes of an object on the basis of only the surface voxels has a drawback. These planes may be quite sensitive to noise that may be present on the surface of the object while the characteristic planes computed on the basis of all the voxels of the object are less so.

REFERENCES

1. P. J. Besl and R. C. Jain, Computing Surveys, **17** (1985) 75.
2. R. T. Chin and C. R. Dyer, Computing Surveys, **18** (1986) 67.
3. D. K. Banerjee, S. K. Parui and D. Dutta Majumder, in "Knowledge Based Computer Systems", Ed. S. Ramani et al. Springer Verlag, (1990), p. 365.
4. T. O. Binford, Personal Communication, (1992).
5. T. O. Binford, Proc. Darpa-Espirit Workshop on Invariance in Computer Vision, (1991).
6. S. K. Parui, "Some Studies in Analysis and Recognition of 2- Dimensional Shapes", Ph.D. Thesis, Indian Statistical Institute, Calcutta, (1984).
7. S. K. Parui and D. K. Banerjee, Tech. Report, TR/KBCS/1/90, Indian Statistical Institute, Calcutta, (1990).
8. S. K. Parui and D. Dutta Majumder, Proc. 7th Int. Conf. on Pat. Recog., Montreal, (1984), p. 72.
9. C. H. Chien and J. K. Aggarwal, Computer Vision, Graphics and Image Processing, **36** (1986) p. 256.
10. J. D. Boissonnat Computer Vision Graphics Image Processing, **44** (1988) p. 1.
11. H. Noborio, S. Fukuda and S. Arimoto, IEEE Trans. on Pattern Analysis and Machine Intelligence, **PAMI-10** (1988) p. 769.
12. P. Srinivasan, P. Liang and S. Hackwood, Pattern Recognition, **23** (1990) p. 843.
13. L. G. Shapiro and R. M. Haralick IEEE Trans. on Pattern Analysis and Machine Intelligence, **PAMI-7** (1985) p. 90.
14. S. K. Parui and D. K. Banerjee, Proc. 1988 IEEE SMC Int. Conference held in Shenyang and Beijing, China, (1988) p. 1031.
15. M. Brady and H. Asada, The International Journal of Robotic Research, **3** (1984) p. 36.
16. M. Brady, Personal Communication, (1992).
17. D. K. Banerjee and S. K. Parui, Tech. Report, TR/KBCS/2/92, Indian Statistical Institute, Calcutta, (1992).