2004 International Conference on Image Processing (ICIP)

THREE-DIMENSIONAL INTRINSIC SHAPE

Victor H. S. Ha

Digital Media Solutions Lab
Samsung Information Systems America
3345 Michelson Dr, Irvine, CA 92612
v.ha@samsung.com

José M. F. Moura

Electrical and Computer Engineering
Carnegie Mellon University
5000 Forbes Ave, Pittsburgh, PA 15213
moura@ece.cmu.edu

ABSTRACT

In this paper, we define the *intrinsic shape* of an object in three-dimensional (3D) space as the shape invariant to affine-permutation geometric shape distortions. We present an algorithm that blindly recovers from an arbitrarily affine-permutation distorted shape its intrinsic shape. This algorithm referred to as 3D BLAISER (BLind Algorithm for Intrinsic ShapE Recovery) executes as one of its steps 3D PRA (Point-based Reorientation Algorithm). 3D PRA reorients (blindly) rotated versions of the same object so that they become exactly aligned. 3D PRA extends 2D PRA to 3D space, but is much more complicated due to the multiple axes of rotation and the associated fold numbers. We describe the algorithms of 3D BLAISER and 3D PRA in detail.

1. INTRODUCTION

The affine distortion model is widely used by the image processing and machine vision communities because it provides a good approximation to the actual imaging process with a manageable complexity. The affine model describes geometric shape distortions such as translation, rotation, reflection, uniform and non-uniform scaling, and skewing. Further to these distortions, the order by which the feature/pixel points are scanned by the input device is usually unknown—we refer to this as a *permutation* distortion. The *intrinsic shape* is the shape that remains after these distortions—affine and permutation—have been factored out, i.e., the shape of the object that is invariant to the combined affine-permutation distortions. This concept is very useful in detection, classification, and identification of objects from distorted images of their shapes.

This paper presents a three-dimensional (3D) extension of the 2D intrinsic shape and 2D shape orientation presented in [1] and [2], respectively. As with the 2D case, the most critical step of the 3D BLAISER, the algorithm that reduces any affine-permutation distorted shape to its intrinsic shape, is the shape reorientation algorithm referred to as 3D PRA (point-based reorientation algorithm). Even though the main concept of 3D PRA is very similar to its 2D equivalence, the 3D version is much more complicated because 3D shapes may have multiple axes of rotation with a different fold number associated with each of them. The 3D orientation problem has been studied in its own right for many years now, [3], [4], [5], but these methods often work only with edges and surfaces, are limited to a certain class of shapes, or require matching with stored models. The algorithm that we present in this paper works directly with arbitrary 3D shapes with no a priori knowledge or pre-computed models.

The paper is organized as follows. Section 2 describes our models for 3D shapes and affine-permutation shape distortions. Then, it introduces the notion of intrinsic shape in 3D. In section 3, we present 3D BLAISER. Section 4 focuses on the orientation of 3D shapes and 3D PRA. Section 5 summarizes the paper.

2. MODEL

Shapes: A 3D shape consisting of N feature/pixel points is described by a $3 \times N$ configuration matrix

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & \dots & x_N \\ y_1 & y_2 & \dots & y_N \\ z_1 & z_2 & \dots & z_N \end{bmatrix}. \qquad (1)$$

Each column in $\mathbf{X}$ represents the location of a point in the reference coordinates with x-, y-, and z-axes. The configuration space $\mathcal{X}$ contains all 3D patterns of N points in 3D space, modulo trivial configurations with repeated points.

Shape Distortions: Two configuration matrices $\mathbf{X}$ and $\mathbf{X}^d$ are affine distorted from each other if they are related as

$$\mathbf{X}^d = \mathbf{A}\mathbf{X} + \mathbf{1}^T \otimes \delta \qquad (2)$$

where the linear distortion matrix $\mathbf{A}$ is a 3×3 invertible matrix of real numbers and the translation vector δ is a 3×1 vector of real numbers. Note that $\otimes$ is the Kronecker product [6] and $\mathbf{1}$ is a $N \times 1$ vector of ones. We now add to the model a $N \times N$ orthogonal matrix $\mathbf{P}$ representing the permutation distortions. Then, the two shapes $\mathbf{X}^d$ and $\mathbf{X}$ are related by an affine-permutation distortion if

$$\mathbf{X}^d = (\mathbf{A}\mathbf{X} + \mathbf{1}^T \otimes \delta)\mathbf{P} = \mathbf{A}\mathbf{X}\mathbf{P} + \mathbf{1}^T \otimes \delta. \qquad (3)$$

In vec notation, the above equation is written as

$$\mathbf{x}^d = (\mathbf{P}^T \otimes \mathbf{A})\,\mathbf{x} + \mathbf{1} \otimes \delta \qquad (4)$$

where $\mathbf{x}^d = vec\,\mathbf{X}^d$ and $\mathbf{x} = vec\,\mathbf{X}$. Note that the matrix $\mathbf{P}^T \otimes \mathbf{A}$ is invertible.

Intrinsic Shape: We define the intrinsic shape $\mathbf{S}$ of an object as the shape that is invariant to affine-permutation distortions. The appropriate abstract setting to formalize this notion of invariance is the framework of group theory. We extend here to 3D the set-up that we presented for 2D in [1]. The group of interest is the affine-permutation group,

$$\mathcal{A} = \left\{a = \left(\mathbf{P^T} \otimes \mathbf{A}, \mathbf{1} \otimes \delta\right)\right\}. \qquad (5)$$

Consider now the action ζ of the group $\mathcal{A}$ on the configuration space $\mathcal{X}$. That is,

$$\zeta : \mathcal{A} \times \mathcal{X} \quad \longrightarrow \quad \mathcal{X}.$$

The group element $a = \left(\mathbf{P^T} \otimes \mathbf{A}, \mathbf{1} \otimes \delta\right) \in \mathcal{A}$ acts on the configuration matrix $\mathbf{X}$ as

$$\begin{aligned} a\mathbf{X} &= \mathbf{AXP} + \mathbf{1}^T \otimes \delta \qquad (6) \\ &= \left(\mathbf{P^T} \otimes \mathbf{A}\right) vec\,\mathbf{X} + \mathbf{1} \otimes \delta. \end{aligned}$$

Two configuration matrices $\mathbf{X}_1$ and $\mathbf{X}_2$ are in the equivalence relation $\mathbf{X}_1 \equiv_{\mathcal{A}} \mathbf{X}_2$ defined by ζ if and only if there exists $a \in \mathcal{A}$ such that $\zeta(a, \mathbf{X}_1) = \mathbf{X}_2$. By the invertibility of the group element $a \in \mathcal{A}$, we have also that $\zeta\left(a^{-1}, \mathbf{X}_2\right) = \mathbf{X}_1$. An equivalence relation partitions the configuration space $\mathcal{X}$ into disjoint equivalence classes $\mathbf{C}_{\mathcal{A}}$, also referred to as orbits. The set $\mathcal{C}_{\mathcal{A}}$ of orbits is

$$\mathcal{C}_{\mathcal{A}} = \{\mathbf{C}_{\mathcal{A}} \subset \mathcal{X} : \forall \mathbf{X}_1, \mathbf{X}_2 \in \mathbf{C}_{\mathcal{A}}, \mathbf{X}_1 \equiv_{\mathcal{A}} \mathbf{X}_2\}. \qquad (7)$$

We formalize the concept of intrinsic shape as the canonical representative $\mathbf{S}$ of the orbit $\mathbf{C}_{\mathcal{A}}$. Of course, this definition makes sense only if it can be defined uniquely and if there is a way to recover it from any other element in the orbit. This is considered in the next section.

3. 3D BLAISER

In this section, we develop BLAISER (BLind Algorithm for Intrinsic ShapE Recovery) that recovers from any $\mathbf{X} \in \mathbf{C}_{\mathcal{A}}$ the corresponding intrinsic shape $\mathbf{S} \in \mathbf{C}_{\mathcal{A}}$. The algorithm is *blind* because it finds $\mathbf{S}$ by working only from the given shape $\mathbf{X}$ and no additional information is used by BLAISER. We first define the intrinsic shape.

Definition 1 *The intrinsic shape $\mathbf{S} \in \mathbf{C}_{\mathcal{A}}$ is defined uniquely by the following four properties:*

1. *The center of mass of the shape is located at the origin O of the reference coordinate system*
2. *The matrix outer product $\mathbf{SS}^T$ is the 3D identity matrix $\mathbf{I}$*
3. *The reorientation point (see subsection 4) falls on the 2D plane $x = 0$, $y \geq 0$, and $z \geq 0$*
4. *The columns of $\mathbf{S}$ are ordered in ascending z coordinate values, then in ascending y coordinate values for the columns with the same z values, finally in ascending x coordinate values for the columns with the same y and z values*

Definition 1 is one out of other possible ways of defining the intrinsic shape of the object.

Theorem 1 *The following four steps reduce any member $\mathbf{X}$ of an orbit $\mathbf{C}_{\mathcal{A}}$ to its canonical representative, or intrinsic shape, $\mathbf{S} \in \mathbf{C}_{\mathcal{A}}$:*

1. *Centering:* $\mathbf{X} = \mathbf{ASP} + \mathbf{1}^T \otimes \delta \longrightarrow \mathbf{X}^c = \mathbf{ASP}$
2. *Reshaping:* $\mathbf{X}^c = \mathbf{ASP} \longrightarrow \mathbf{X}^s = \mathbf{USP}$
3. *Reorientation:* $\mathbf{X}^s = \mathbf{USP} \longrightarrow \mathbf{X}^u = \mathbf{SP}$
4. *Sorting:* $\mathbf{X}^u = \mathbf{SP} \longrightarrow \mathbf{X}^r = \mathbf{S}$

Note that $\mathbf{U}$ is a 3×3 orthogonal matrix representing a 3D rotation in 3D space. It represents the orientation ambiguity of 3D shapes.

BLAISER is the algorithm that carries out the four steps in Theorem 1. The proof of Theorem 1 is omitted here due to lack of space. We explain below the centering, reshaping, and sorting operations. The reorientation step is discussed in the next section. Step 1 centers the shape at the origin O of the coordinate system, i.e., the shape is rigidly translated so that its center of gravity becomes the origin O of the coordinate system. Step 2 reshapes and rescales the centered shape to its *n*ormalized shape. Step 2 is accomplished by applying the "compacting" algorithm. That is, given a centered 3D shape $\mathbf{X}^c$, we multiply a 3×3 reshaping matrix $\mathbf{W}$ as in $\mathbf{X}^s = \mathbf{WX}^c = \mathbf{USP}$ where $\mathbf{W} = \mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{V}^T$ and $\mathbf{X}^c\mathbf{X}^{cT} = \mathbf{V\Lambda V}^T$ by a singular value decomposition. The outer product of the normalized shape $\mathbf{X}^s$ then becomes an identity matrix, i.e., $\mathbf{X}^s\mathbf{X}^{sT} = \mathbf{I}$. After the reorientation step, the configuration $\mathbf{X}^u = \mathbf{SP}$ is reduced to $\mathbf{X}^r$ by sorting its columns as described in Definition 1 (4). The result is defined to be the intrinsic shape $\mathbf{S}$. We consider next how to remove the orientational ambiguity $\mathbf{U}$ from $\mathbf{X}^s$ in the reorientation step.

4. 3D PRA

Determination of the orientation of 3D shapes is an important research topic in application areas such as machine vision, computer graphics, and medical imaging. In the shape reorientation step, the patterns extracted from the image are brought consistently to the same orientation so that further processing, such as detection, recognition, and matching, can proceed.

In this section, we introduce a new algorithm referred to as 3D PRA (point-based reorienation algorithm), which is a 3D extension of 2D PRA presented in [2]. In 3D, determining the rotational symmetry involves identifying the axes of rotation and the associated fold numbers. Reflections are considered irrelevant for 3D shapes. Instead, the direction of the axis of rotation becomes an issue — for example, one must distinguish the positive z-axis (out of the page) from the negative z-axis (into the page).

PRA works with shapes that are described by a configuration of feature/pixel points. The main idea of 3D PRA is to identify a unique point in the shape, referred to as the *r*eorientation point, and use this point to bring the shape to its normalized orientation. Given a 3D input shape that has been centered and reshaped by BLAISER, PRA first identifies the axes of rotation and associated fold numbers of the shape. Then, it uses these clues to locate the reorientation point of the shape. The shape is finally rotated to its normalized orientation using the reorientation point. We start by introducing the terminology used in 3D PRA.

4.1. Terminology

Rotational Symmetry A 3D shape has k-fold rotational symmetry with respect to the given axis of rotation if it is indistinguishable after being rotated about its center of mass by an angle $\frac{2n\pi}{k}$, $n = 1, \cdots, k$. The fold number of a shape is the largest integer k for which the shape is k-fold rotational symmetric. The fold number measures the degree of rotational symmetry of a shape, e.g., the square is 4-fold symmetric.

Spherical Coordinates A shape $\mathbf{X}$ consisting of N points is represented in spherical coordinates $\{\rho, \phi, \psi\}$ where ρ is the radius from the coordinates origin, ϕ is the polar angle measured from the z–axis with $0 \leq \phi \leq \pi$, and ψ is the azimuth angle in the x-y plane ($z = 0$) measured from the positive x-axis towards positive y-axis with $0 \leq \psi < 2\pi$.

Sub-shapes A shape is decomposed into a set of sub-shapes. Each sub-shape is composed of feature/pixel points located at the same distance from the center of the shape. A sub-shape is obtained by sorting the columns of the matrix $\mathbf{X}_\theta^\rho$ in decreasing order of the ρ_i. Then, columns with the same value of ρ are grouped together as a sub-shape $\mathbf{O}_j$, $j = 1 \ldots R$ where N_j is the number of points contained in the sub-shape $\mathbf{O}_j$ and R is the total number of distinct values of ρ in the shape. In 2D, a sub-shape is a set of points located on a ring centered at the coordinate origin. In 3D, a sub-shape is a set of points located on a sphere centered at the origin, numbered from the outermost one that has the largest radius.

Axis of Rotation An axis of rotation $\vec{r} = [1\ \phi\ \psi]^T$ is defined as a unit vector starting from the coordinate origin. The axes of rotation $\vec{r}$ and $-\vec{r}$ are considered to represent the same axis when the rotational symmetry of the shape is concerned. For the shape reorientation process, however, these two axes are distinguished from each other to bring the shape to the unique orientation in 3D space.

Rotational Distance The rotational distance $d_\theta^{\vec{r}}$ between two points p^1 and p^2 is defined as the angle $\angle p_\perp^1 O p_\perp^2$ where $p_\perp^i$ is the orthogonal projection of the point p^i on the plane that is perpendicular to the rotation axis $\vec{r}$ and O is the coordinate origin.

List of Angles The list of angles $\mathbf{Z}$ is generated by measuring the angles between each rotationally consecutive points in the shape. In the 3D case, two versions of the list of angles are generated depending on the sorting order of the shape. In the first version denoted as $\mathbf{Z}^a$, the shape is sorted in descending order of the azimuth angle ψ, and again in ascending order of the polar angle ϕ among those with the same value of ψ, and finally in descending order of the radius ρ among those with the same values of ψ and ϕ. The difference in the second version denoted as $\mathbf{Z}^d$ is that the polar angle ϕ is sorted in descending order. The superscripts a and d stand for ascending and descending, respectively. Once the shape is sorted, we measure the angles $\Delta\psi_{i,j} = \psi_i - \psi_j$, starting from an arbitrary point in the shape. If $\Delta\psi_{i,j} = 0$ between two points p_i and p_j in the sorted order, we enter the angle $\Delta\phi_{i,j}$ to the list. If both $\Delta\psi_{i,j} = 0$ and $\Delta\phi_{i,j} = 0$, we enter the difference $\Delta\rho_{i,j}$ to the list.

Fundamental List of Angles The fundamental list of angles $\mathbf{L}$ is a non-periodic segment of the list of angles $\mathbf{Z}$. Given the fold number k of the shape, the fundamental list of angles is obtained by collecting N/k consecutive elements from the list of angles. There are also two versions of the fundamental lists of angles $\mathbf{L}^a$ and $\mathbf{L}^d$ depending on the sorting order of the corresponding list of angles $\mathbf{Z}^a$ and $\mathbf{Z}^d$.

4.2. Algorithm

The reorientation point is the most critical concept for the correct operation of 3D PRA. PRA performs the orientation normalization by (i) first choosing one of the axes of rotation and its associated fold number, (ii) rotating the shape until the chosen axis of rotation coincides with the positive z-axis, (iii) identifying the unique reorientation point of the shape, and finally (iv) rotating the shape until the reorientation point falls on the plane $x = 0$, $y \geq 0$, and $z \geq 0$. The steps of 3D PRA are shown below.

Given a 3D shape with N feature/pixel points,

1. Identify the axes of rotation and the associated fold numbers
2. Select one of the axes of rotation by choosing the axis associated with the largest fold number
3. Rotate the shape so that the chosen axis of rotation coincides with the positive z-axis
4. Generate lists of angles $\mathbf{Z}^a$ and $\mathbf{Z}^d$
5. Extract corresponding fundamental lists of angles $\mathbf{L}^a$ and $\mathbf{L}^d$ from $\mathbf{Z}^a$ and $\mathbf{Z}^d$, respectively
6. Identify the reorientation point p_o of the shape using the fundamental lists of angles $\mathbf{L}^a$ and $\mathbf{L}^d$
7. Rotate the shape so that the reorientation point falls on the plane $x = 0$, $y \geq 0$, and $z \geq 0$

We now discuss each step in detail.

Step (1) Axis of rotation and fold number: A 3D shape may have a multiple set of rotation axes and associated fold numbers. In this step, we identify all of these rotation axes. We first decompose the shape into a set of sub-shapes $\mathbf{O}_j$, $j = 1...R$. When a k-fold symmetric sub-shape is rotated by an angle $2\pi/k$ about the corresponding axis of rotation, each point in the shape either stays at the same location or moves to the location that was previously occupied by another point. That is, for any sub-shape consisting of N points, there are at most N^2 distinct pairs to be tested for the existence of a rotation axis, i.e., $p_i \rightarrow p_j$ for i, $j = 1 \ldots N$. Therefore, the entire set of rotation axes can be found by exhaustively testing these candidates. Fortunately, the number of candidates can be greatly reduced, allowing an early termination. A sub-shape with N points can only have k-fold rotational symmetry where $k = 1, \ldots, N$. Thus, the angular distance between any two points that form an actual rotation axis must be one of $2\pi/k$ rad. We can eliminate those candidate pairs with the angular distances that are not in the set $\Theta = \{2\pi/k\}$ for $k = 1 \ldots N$.

The fold number of each sub-shape is computed from the list of angles generated for the sub-shape. The sub-shape, with the constant radius ρ, is sorted in increasing order of ϕ and then again in decreasing order of ψ among those with the same value of ϕ. At each value of ϕ, a sub-list of angles is generated by measuring the angles $\Delta\psi$. Each sub-list is periodic with the periodicity d_i. Finally, the fold number of the sub-shape is determined as the greatest common denominator of $\{d_i\}$, $i = 1...D$ where D is the total number of sub-lists in the sub-shape.

Once all axes of rotation and associated fold numbers are identified for each sub-shape, the rotation axes of the overall shape are the ones common to all sub-shapes. The fold numbers of the overall shape are computed as the greatest common denominator of the sub-shapes' fold numbers corresponding to each of the common axis of rotation. The outline of the algorithm implementing Step (1) is shown below.

1. Decompose the shape into sub-shapes, $\mathbf{O}_j, j = 1 \cdots R$
2. For each sub-shape with N_j points, generate the set $\Theta = \{2\pi/k\}$ of angular distances where $k = 1, \ldots, N_j$. Execute the following loop:

   for $n = 1 \ldots N_j,\ m = 1 \ldots N_j$

   - Compute the rotational distance $d_\theta^{\vec{r}}$ for each pair of points (p_n, p_m)
   - If the rotational distance $d_\theta^{\vec{r}}$ belongs to the set Θ, move every point in the sub-shape by the same rotational distance about the axis of rotation $\vec{r}$
   - If every point stays at the same location or moves to the location previously occupied by another point (i.e., the sub-shape appears unchanged), store the axis of rotation $\vec{r}$, and computed the associated fold number

   end for
3. The axes of rotation for the overall shape are obtained by collecting the axes of rotation that are common to all of its sub-shapes
4. The associated fold numbers are computed as the greatest common denominator of the sub-shapes' fold numbers with each of the common axis of rotation

Step (2)~(3) Selecting an axis of rotation for the shape: Given the set of rotation axes and associated fold numbers, we need to choose one that is uniquely identifiable for the shape. Since more distinguishing characteristics of the shape are represented by a larger fold number, we select the rotation axis that is associated with the largest fold number. [1]

Step (4)~(5) Fundamental list of angles: Starting from an arbitrary point in $\mathbf{X}$, we measure the angles $\Delta\phi$ and generate the lists of angles $\mathbf{Z}^a$ and $\mathbf{Z}^d$. From these lists of angles, we extract a pair of fundamental lists of angles, $\mathbf{L}^a$ and $\mathbf{L}^d$. The fundamental list of angles $\mathbf{L}^a$ is generated by extracting N/k successive elements from the list of angles $\mathbf{Z}^a$ where N is the total number of feature/pixel points and k is the largest fold number of the overall shape. The list $\mathbf{L}^d$ is generated similarly from $\mathbf{Z}^d$.

Step (6)~(7) Reorientation point: We identify the reorientation point p_o for the shape using the two fundamental lists of angles, $\mathbf{L}^a$ and $\mathbf{L}^d$. The shape reorientation process is then completed by rotating the shape with respect to the reorientation point. First, we identify the point p_o^a from the fundamental list of angles $\mathbf{L}^a$ and p_o^d from the $\mathbf{L}^d$. The reorientation point p_o is chosen as one of (p_o^a, p_o^d). If $p_o = p_o^a$, the orientation normalization is completed by rotating the shape about the axis of rotation until p_o falls on the 2D plane $x = 0$, $y \geq 0$. If $p_o = p_o^d$, we rotate the shape by π rad so that the rotation axis $\vec{r}$ moves to $-\vec{r}$, i.e., the upper and lower hemispheres are swapped. Then, we rotate the shape again so that p_o falls on the 2D plane $x = 0$, $y \geq 0$. The reorientation points p_o^a, p_o^d, and p_o are identified in the same way as done in 2D PRA. First, the element with the maximum value is chosen from the fundamental list of angles. When there are more than one such element, we compare the magnitudes of the neighboring elements in the lists. See [2].

[1] We are extending 3D PRA to include the case when the rotation axis associated with the largest fold number is not unique.

5. CONCLUSION

This paper defines the *intrinsic shape* of a 3D object, which is invariant to affine-permutation distortions. BLAISER is the algorithm that recovers from any affine-permutation distorted shape its intrinsic shape. 3D PRA, which is the critical step of 3D BLAISER, provides an efficient approach to the 3D shape orientation problem. The performance of 3D PRA is affected by the variation in the locations of the points due to finite resolution of the input/display devices, background noise, or erroneously added/deleted points. We are currently extending to 3D the robust version of 2D PRA in [2].

6. REFERENCES

[1] V. H. S. Ha and J. M. F. Moura, "Intrinsic shape," in *Proceedigns of the 36th Asilomar Conference on Signals, Systems and Computers*, Monterey, CA, November 2002, vol. 1, pp. 139–143.

[2] V. H. S. Ha and J. M. F. Moura, "Efficient 2D shape orientation," in *Proceedings of the IEEE International Conference on Image Processing*, September 2003, vol. 1, pp. 225–228.

[3] Reza Safaee-Rad, Ivo Tchoukanov, Kenneth Carless Smith, and Bensiyon Benhabib, "Three-dimensional location estimation of circular features for machine vision," *IEEE Transactions on Robotics and Automation*, vol. 8, no. 5, pp. 624–640, October 1992.

[4] A. Adán, C. Cerrada, and V. Feliu, "Automatic orienting of 3D shapes by using a new data structure for object modeling," in *Proceedings of the 1999 IEEE International Conference on Robotics and Automation*, 1999, pp. 2881–2886.

[5] M. A. Magnor, "Geometry-based automatic object localization and 3D pose detection," in *Proceedings of the Fifth IEEE Southwest Symposium on Image Analysis and Interpretation*, April 2002, pp. 144–147.

[6] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus*, John Wiley & Sons, 1988, Chapter 2: Kronecker products, the vec operator and the Moore-Penrose inverse.