# LSP: A View Insensitive 3D Representation Method

Xiaobu Yuan
Department of Computer Science
Memorial University of Newfoundland
St. John's, Newfoundland, Canada A1C 5S7

## Abstract

*LSP(Localized Surface Primitive) is a method that provides view insensitive three-dimensional(3D) representations. In this method, 3D objects are represented as locally defined surface primitives, and their topographical relations are specified by homogeneous transformations. This paper first investigates surface locality, and then after introducing the idea of local coordinate systems, describes how localized surface primitives are defined and used to represent objects as well as applied in object recognition. The stability of this method is analyzed and tested with range images.*

## 1 Introduction

Object recognition is important in computer vision because it enables a vision system to understand its environment by matching observed objects with object models in the system. A 3D object model is a spatial representation described under a certain coordinate system whose relation with the universal coordinate system is specified by a translation vector $\vec{\omega} = (\sigma, \delta, \zeta)$ and a rotation vector $\vec{\theta} = (\alpha, \beta, \gamma)$ [3]. Therefore, if there are $N$ object models in the system, the model set $\mathcal{M}$ is a group of ordered (*object, translation, rotation*) triples

$$\mathcal{M} = \{(A_i, \vec{\omega}_i, \vec{\theta}_i)\}_{i=0}^{N-1}, \tag{1}$$

where $A_i$ is the representation for the $i$th object with position $\vec{\omega}_i$ and rotation $\vec{\theta}_i$.

At an arbitrary appearance of an object, it is a projection under a translation $\vec{\omega}'$ and a rotation $\vec{\theta}'$:

$$P(A_k, \vec{\omega}'_k, \vec{\theta}'_k). \tag{2}$$

Recognizing the object is to identify an inverse mapping on the model set $\mathcal{M}$ so that the following equation exists

$$(A_k, \vec{\omega}'_k, \vec{\theta}'_k) = P^{-1}(A_i, \vec{\omega}_i, \vec{\theta}_i) \tag{3}$$

where $(A_i, \vec{\omega}_i, \vec{\theta}_i) \in \mathcal{M}$, or to prove such an inverse mapping does not exist.

To identify one particular appearance from a combination up to a number of $N \times \mathcal{R}^6$ is theoretically impossible since $\mathcal{R}^6$ is a set of real numbers with six-degrees of freedom. Different methods have been proposed to solve the problem in one way or another[8, 5, 7, 10], but a resolution to make an one-to-one matching in Eq. 3 still needs more work. Introduced in the paper is a method that uses localized surface primitives to describe and recognize 3D objects, and therefore eliminate the effects of $\vec{\omega}'_k, \vec{\theta}'_k$ and $\vec{\omega}_i, \vec{\theta}_i$ in the matching defined in Eq. 3 so that the one-to-one matching becomes practical.

## 2 Localized Surface Primitives

A 3D quantity is *transformation invariant* if it remains the same under any rigid transformations. In geometric surface modeling methods[6], objects are represented with boundary surface primitives. Consequently, the invariance for such models shifts to *surface invariance*[1], i.e., the description of surface primitives and their relations should be invariant with respect to any transformations.

### 2.1 Surface locality

An object surface $S$ is composed by a set of surface points $\{(x_i, y_i, z_i)\}$, for $i$=*0, 1, ..., n-1*, in a universal reference system $\{F_U\}$. When observed in a viewing coordinate system $\{F_{\vec{\mu}}\}$, $S$ becomes another point set in the new system $S_{\vec{\mu}} = \{(u_i, v_i, w_i)\}$. Let the two transformation vectors relating $\{F_{\vec{\mu}}\}$ to $\{F_U\}$ be $\vec{\omega} = (\sigma, \delta, \zeta)$ and $\vec{\theta} = (\alpha, \beta, \gamma)$, surface $S_{\vec{\mu}}$ is then described under three parameter groups as $S(\vec{\mu}, \vec{\omega}, \vec{\theta})$. As parameters change from $(\vec{\mu}, \vec{\omega}, \vec{\theta})$ to $(\vec{\mu}', \vec{\omega}', \vec{\theta}')$ due to a changed viewing coordinate system $\{F_{\vec{\mu}'}\}$, each examined surface point $(u_k, v_k, w_k)$ is mapped in $\{F_{\vec{\mu}'}\}$ as an element $(u'_k, v'_k, w'_k)$ of $S_{\vec{\mu}'}$. Since the observed surface point set is now $S_{\vec{\mu}'}$, the surface description obtained from it has a different format,

$$S(\vec{\mu}', \vec{\omega}', \vec{\theta}') \neq S(\vec{\mu}, \vec{\omega}, \vec{\theta}). \tag{4}$$

Therefore, the choice of coordinate systems in which surface primitives are defined plays a key role in surface invariance.

Differential geometry has been used for image segmentation because view independent properties capture the spatial properties of surface shapes[7, 2]. Similarly, when surface invariance is concerned, the coordinate systems chosen to define object surfaces should

[1] A surface characteristic is transformation invariant if it does not change under any rigid transformations that don't affect the visibility of that surface, i.e., *invariant if visible*[2].

come from the surface characteristics that depend only on surface shapes. According to homogeneous operations, if $\mathbf{T}(\vec{\omega})$ is the translation matrix and $\mathbf{R}(\vec{\theta})$ is the rotation matrix for $\{F_{\vec{\mu}}\}$, the matrix product $\mathbf{R}(\vec{\theta})\mathbf{T}(\vec{\omega})$ relates $\{F_{\vec{\mu}}\}$ to $\{F_U\}$. The same is true that $\mathbf{R}(\vec{\theta}')\mathbf{T}(\vec{\omega}')$ relates $\{F_{\vec{\mu}'}\}$ to $\{F_U\}$. As a result, $\{F_{\vec{\mu}'}\}$ relates to $\{F_{\vec{\mu}}\}$ with a product of homogeneous transformation matrices,

$$\tau = \mathbf{T}^{-1}(\vec{\omega}')\mathbf{R}^{-1}(\vec{\theta}')\mathbf{R}(\vec{\theta})\mathbf{T}(\vec{\omega}).$$

If denoting $\tau$ as

$$\tau = \begin{bmatrix} t_{00} & t_{01} & t_{02} & t_{03} \\ t_{10} & t_{11} & t_{12} & t_{13} \\ t_{20} & t_{21} & t_{22} & t_{23} \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

the projected surface points on two image planes are related by a homogeneous operation,

$$\begin{bmatrix} u'_k \\ v'_k \\ w'_k \\ 1 \end{bmatrix} = \begin{bmatrix} t_{00} & t_{01} & t_{02} & t_{03} \\ t_{10} & t_{11} & t_{12} & t_{13} \\ t_{20} & t_{21} & t_{22} & t_{23} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_k \\ v_k \\ w_k \\ 1 \end{bmatrix} \quad (5)$$

Since a surface centroid is the average surface point, the two surface centroids $C_{\vec{\mu}}$ and $C_{\vec{\mu}'}$ for $S_{\vec{\mu}}$ and $S_{\vec{\mu}'}$ are then represented by the following equations,

$$C_{\vec{\mu}} = \begin{bmatrix} \bar{u} \\ \bar{v} \\ \bar{w} \end{bmatrix} = \begin{bmatrix} \frac{1}{n}\sum_{l=0}^{n-1} u_l \\ \frac{1}{n}\sum_{l=0}^{n-1} v_l \\ \frac{1}{n}\sum_{l=0}^{n-1} w_l \end{bmatrix}, \quad (6)$$

$$C_{\vec{\mu}'} = \begin{bmatrix} \bar{u}' \\ \bar{v}' \\ \bar{w}' \end{bmatrix} = \begin{bmatrix} \frac{1}{n}\sum_{l=0}^{n-1} u'_l \\ \frac{1}{n}\sum_{l=0}^{n-1} v'_l \\ \frac{1}{n}\sum_{l=0}^{n-1} w'_l \end{bmatrix} \quad (7)$$

Substituting Eq. 5 to Eq. 7 creates,

$$C_{\vec{\mu}'} = \begin{bmatrix} \frac{1}{n}\sum_{l=0}^{n-1}(t_{00}u_l + t_{01}v_l + t_{02}w_l + t_{03}) \\ \frac{1}{n}\sum_{l=0}^{n-1}(t_{10}u_l + t_{11}v_l + t_{12}w_l + t_{13}) \\ \frac{1}{n}\sum_{l=0}^{n-1}(t_{20}u_l + t_{21}v_l + t_{22}w_l + t_{23}) \end{bmatrix}$$

From Eq. 6, Eq. 7 can be further derived as,

$$\begin{bmatrix} \bar{u}' \\ \bar{v}' \\ \bar{w}' \\ 1 \end{bmatrix} = \begin{bmatrix} t_{00} & t_{01} & t_{02} & t_{03} \\ t_{10} & t_{11} & t_{12} & t_{13} \\ t_{20} & t_{21} & t_{22} & t_{23} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \bar{u} \\ \bar{v} \\ \bar{w} \\ 1 \end{bmatrix}. \quad (8)$$

Eq. 8 shows the one-to-one mapping relation of $C_{\vec{\mu}'}$ and $C_{\vec{\mu}}$. Especially, when $\{F_{\vec{\mu}}\}$ is the universal reference system, $C_{\vec{\mu}}$ is the real surface centroid of the surface primitive itself. Therefore, all the surface centroids of observed surfaces in different viewing coordinate systems are mapped from the same real surface centroid. When a coordinate system $\{F_S\}$ is set up on the surface $S$ so that its origin coincides with the

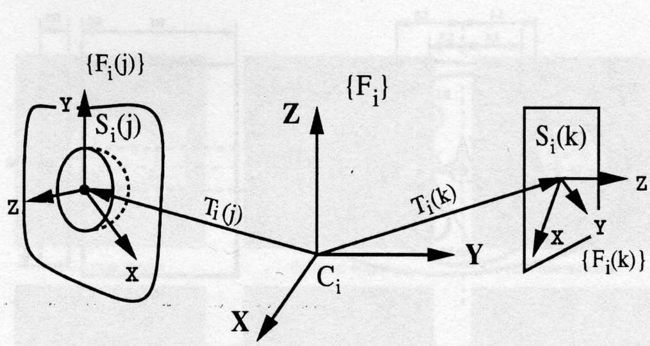


Figure 1: An Example of Local Coordinate Systems

surface centroid and its z axis points to the average surface normal direction, $\{F_S\}$ is the same on surface $S$ no matter it is observed in $\{F_{\vec{\mu}}\}$ or $\{F_{\vec{\mu}'}\}$. The locally defined frame $\{F_S\}$ is not only fixed at a viewpoint invariant location of the surface, but its z axis indicates the visible direction of that surface as well.

In addition, according to differential geometry, there are also some transformation invariant local spatial characteristics at the surface centroid, such as its maximum and minimum curvatures, and its spatial Euclidean distance to any point on the surface. When the x axis of $\{F_S\}$ is chosen as a direction determined by these characteristics, and y axis makes $\{F_S\}$ a right hand Cartesian system with the other two axes, the local coordinate system fixed on the surface is independent of viewing directions. If described under such a coordinate system, the description of surface $S$ inherits the view insensitivity too,

$$S(\vec{\mu}, \vec{\omega}, \vec{\theta}) = S(F_U) = S(\vec{\mu}', \vec{\omega}', \vec{\theta}'). \quad (9)$$

## 2.2 LSP models

Among the $N$ object models in $\mathcal{M}$, suppose the $i$th object $O_i$ has $m_i$ surface primitives $S_i(j)$, $0 \leq j \leq m_i - 1$ and $0 \leq i \leq N-1$. Object $O_i$ is described in an object coordinate system $\{F_i\}$ by its surface primitives, and each surface primitive $S_i(j)$ on $O_i$ is in turn described under a local coordinate system $\{F_i(j)\}$ especially established for $S_i(j)$, as shown in Fig. 1.

The local coordinate system $\{F_i(j)\}$ for a surface primitive $S_i(j)$ is a Cartesian system. Its z axis is the visible direction of $S_i(j)$, i.e., the average surface normal; the x axis is perpendicular to z in the direction of minimum/maximum curvature at surface centroid $C_i(j)$ for curved surfaces, or of the longest/shortest distance from $C_i(j)$ to edge points for planar surfaces; and $\mathrm{y} = \mathrm{x} \times \mathrm{z}$. The origin of $\{F_i(j)\}$ locates at $C_i(j)$ if $O_i$ is not a rotational sweeping object, otherwise, $\{F_i(j)\}$ originates at the average point along the symmetric axis in order to make the description simple. An object coordinate system $\{F_i\}$ for $O_i$ locates at the average origin point of all local systems. Since $\{F_i\}$ is used only for reference to relate local coordinate systems, its three axes can be set up randomly, e.g., the average directions of all local systems.

As an example, the body of a pencil sharpener is shown with parameter labels in Fig. 2. Even though

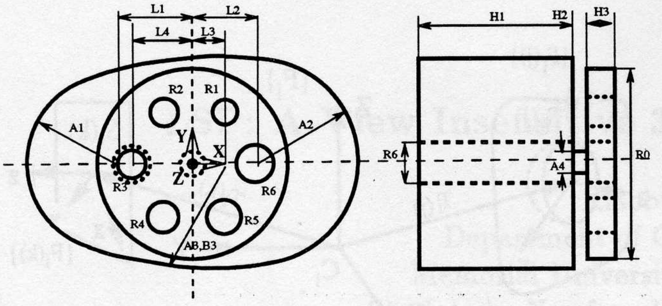


Figure 2: The LSP Model of a Pencil Sharpener

it has thirteen surface primitives, its shape is actually determined by two parts, a plate and a container. The front face of the plate is a big circle with six equally spaced circle holes whose centroids are $L_4$ apart from the center of the big circle, and the back face is the front face minus the touching pole circle. The radii of those circles are $R_0$, $R_1$, $R_2$, $R_3$, $R_4$, $R_5$, $R_6$, and $A_4$. Correspondingly, there are eight cylinder surfaces as thick as $H_3$ or $H_2$.

In comparison, the face of the container is specified by the joint area of two circles, which are $L_1$ and $L_2$ apart from the big circle center and have radii as $A_1$ and $A_2$, and an ellipse at $L_1$ with parameters $A_3$ and $B_3$. Because of the hole piercing though the container with a radius $R_6$ at $L_4$, the front face is the joint area minus the hole, and the back face has an additional subtraction of the touching pole circle. While the hole, which is as deep as $H_1$, is a cylinder surface, the side of the container can be specified as the area between the outside boundaries of the front and back faces.

If the object coordinate system is set up as in the left figure and the thirteen local coordinate systems are defined according to the previous discussion, as a LSP model, the object can be described as

$$\{\langle\tau(0), S(0)\rangle, \ \ldots \ \langle\tau(j), S(j)\rangle, \ \ldots \ \langle\tau(12), S(12)\rangle\},$$

where $\tau(j)$, $0 \leq j \leq 12$, is the transformation matrix relating $\{F_j\}$ to object system $\{F\}$ and $S(j)$ is the surface description of the $j$th surface primitive under $\{F_j\}$. For instance, under $\tau(0) = \mathbf{I}$ and $\tau(1) = T_z(-H_3/2)T_y(\sqrt{3}L_4/2)T_x(L_4/2)$, surfaces $S(0)$ and $S(1)$ are described as follows,

$$\begin{aligned} S(0) &= S_0 - (S_1 + S_2 + S_3 + S_4 + S_5 + S_6) \\ S(1) &= x_1^2 + y_1^2 - R_1^2, \quad -H_3/2 \leq z_1 \leq H_3/2 \end{aligned}$$

where $z_0 = 0$ and

$$\begin{aligned} S_0 &= x_0^2 + y_0^2 - R_0^2 \\ S_1 &= (x_0 + L_4/2)^2 + (y_0 + \sqrt{3}L_4/2)^2 - R_1^2 \\ S_2 &= (x_0 - L_4/2)^2 + (y_0 + \sqrt{3}L_4/2)^2 - R_2^2 \\ S_3 &= (x_0 - L_4)^2 + y_0^2 - R_3^2 \\ S_4 &= (x_0 - L_4/2)^2 + (y_0 - \sqrt{3}L_4/2)^2 - R_4^2 \\ S_5 &= (x_0 + L_4/2)^2 + (y_0 - \sqrt{3}L_4/2)^2 - R_5^2 \\ S_6 &= (x_0 + L_4)^2 + y_0^2 - R_6^2 \end{aligned}$$

## 2.3 LSP in object recognition

The major problem to identify an observed object from a set of object models is that the matching between the arbitrary appearance of that object with the model set has a six-degree freedom(Eq. 3). However, in the LSP method, objects are described by localized surface primitives. Therefore, the representation of object $O_i$ is no longer a transformation related triple $(A_i, \vec{\omega}_i, \vec{\theta}_i)$, but a set of view insensitive matrix-surface pairs

$$\{\langle\tau_i(0), S_i(0)\rangle, \ \ldots \ \langle\tau_i(j), S_i(j)\rangle, \ \ldots \ \langle\tau_i(m_i), S_i(m_i)\rangle\}.$$

As the same as in the model for the pencil sharpener, $S_i(j)$ is the localized $j$th surface primitive and $\tau_i(j)$ relates the local coordinate system $\{F_i(j)\}$ to $\{F_i\}$.

Object recognition with LSP models are accomplished by a three-step candidate selection including surface identification, topo-relation matching, and visibility checking, accompanied by a disambiguating process when necessary[11]. In the procedure, observed surfaces $S_k(j)$, $0 \leq j \leq m_k$ in a partial model are first compared with all the surfaces $S_i(l)$, $0 \leq l \leq m_i$, of the acquired models $M_i$ in the model set $\mathcal{M}$ (Eq. 1). Therefore, model $M_i$ can be simplified when only surface are considered,

$$\{S_i(0), \ S_i(1), \ \ldots \ S_i(j), \ \ldots \ S_i(m_i)\}.$$

Similarly, the observed object can be described as

$$\{S_k(0), \ S_k(1), \ \ldots \ S_k(j), \ \ldots \ S_k(m_k)\}.$$

As a result, the problem to match $(A_k, \vec{\omega}'_k, \vec{\theta}'_k)$ with all $(A_i, \vec{\omega}_i, \vec{\theta}_i)$ in $\mathcal{M}$ is changed to find the object $O_l$ that meets the condition:

$$\{S_k(0), \ \ldots \ S_k(m_k)\} \in \{S_l(0), \ \ldots \ S_l(m_l)\}. \tag{10}$$

Only when the condition is satisfied that the observed object can be considered as a counterpart of $O_l$. Otherwise, no matching exists.

# 3 Stability Analysis

As implied in Eq. 10, it is the identical descriptions of localized surface primitives that make the direct comparison possible. Therefore, the stability of these descriptions is the main concern in the analysis.

The description of a 3D feature is specified by a vector function together with a parameter vector,

$$f(\mathbf{x}, \mathbf{p}) = 0 \qquad \mathbf{x} \in \Re^3, \quad \mathbf{p} \in \Re^n. \tag{11}$$

For surface primitives, the vector function $f(\cdot)$ is a family of parameterized features, such as all quadratic surfaces. Any specific value $\mathbf{p}_i$ of $\mathbf{p}$ describes a particular instance in the feature family $f(\cdot)$, such as a sphere. Generally a geometric model consists of a set of parameter vectors with particular values to specify its geometric features. The stability of a vision system depends directly on the tolerance when matching these parameter values.

Geometric uncertainty is defined as multivariate Gaussian distributions[9, 4]. Given a mean vector $\mu$ and the covariance matrix $\Lambda$, the distribution predicts the probability of obtaining a particular parameter value $\mathbf{p}_i$ as

$$g(\mathbf{p}_i) = \frac{1}{\sqrt{(2\pi)^n \det(\Lambda)}} e^{-\frac{1}{2}(\mathbf{p}_i-\mu)^t \Lambda^{-1}(\mathbf{p}_i-\mu)}, \quad (12)$$

where $n$ is the dimension of the corresponding parameter vector p in Eq. 11. Suppose two parameter values, one from the model and another from observed features, $\mathbf{p}_i$ and $\mathbf{p}_j$ in a same feature family $f(\cdot)$ are obtained with the covariance matrices $\Lambda_i$ and $\Lambda_j$. The attempt to compare the similarity of $\mathbf{p}_i$ and $\mathbf{p}_j$ involves an uncertainty tolerance within a maximum allowable distance[4]. The distance is normalized by a joint covariance $\Lambda_{ij}$ as,

$$\mathsf{NormDist}(\mathbf{p}_i, \mathbf{p}_j, \Lambda_{\mathbf{ij}}) = (\mathbf{p}_i - \mathbf{p}_j)^t {\Lambda_{\mathbf{ij}}}^{-1}(\mathbf{p}_i - \mathbf{p}_j), \quad (13)$$

where $\Lambda_{ij} = \Lambda_i + \Lambda_j$. $\mathsf{NormDist}(\mathbf{p}_i, \mathbf{p}_j, \Lambda_{\mathbf{ij}})$ determines the likelihood if the two features are the same. As an example, for a 90% confidence, the distance must be no greater than 4, i.e., $\mathsf{NormDist}(\mathbf{p}_i, \mathbf{p}_j, \Lambda_{\mathbf{ij}}) \leq 4$, or in 1D case, $\mathsf{Abs}(p_i\text{-}p_j) \leq 2\sigma$.

In addition, when a geometric feature $\mathbf{p}_i$ is transformed to $\mathbf{p}_i'$ because of a change in description format or reference coordinate system, the new covariance matrix $\Lambda_i'$ can be calculated from the original $\Lambda_i$ by a Jacobian matrix that makes the transformation[9]. Suppose $\mathbf{J}_p$ and $\mathbf{J}_f$ are the Jacobians for the transformation from $\mathbf{p}_i$ to $\mathbf{p}_i'$ due to parameter and system changes respectively. The new covariance $\Lambda_i'$ for the changed parameter vector is

$$\Lambda_i^p = ({\mathbf{J}_p}^t)^{-1} \Lambda_i {\mathbf{J}_p}^{-1} \quad (14)$$

in the new description format, or

$$\Lambda_i^f = \mathbf{J}_f \Lambda_i {\mathbf{J}_f}^t \quad (15)$$

in the new coordinate system.

More specifically, the covariance matrix $\Lambda_i$ has three independent derivations when a geometric feature $S_i$ is obtained from an image. They are $\sigma_x$, $\sigma_y$, and $\sigma_z$ in the three axis directions. The uncertainty inherited to the localized description is

$$\Lambda_i' = ({\mathbf{J}_p}^t)^{-1} \mathbf{J}_f \Lambda_i {\mathbf{J}_f}^t {\mathbf{J}_p}^{-1} \quad (16)$$

As an example, let the image function for an observed surface $S$ be $h(\mathbf{x})$. Because of sensor noise, the image function is then corrupted with an additive random noise $\mathbf{n}(\mathbf{x})$. If there are $2^{N_b}$ range levels in an image and the resolutions is $2^{N_r} \times 2^{N_r}$, the final measurement of surface primitive $S$ is an integer array $\hat{h}(i,j)$ determined by the equation,

$$\hat{h}(i,j) = \hat{h}(\lfloor \mathbf{x} \rfloor_{N_r}) = \lfloor h(\mathbf{x}) + \mathbf{n}(\mathbf{x}) \rfloor_{N_b}. \quad (17)$$

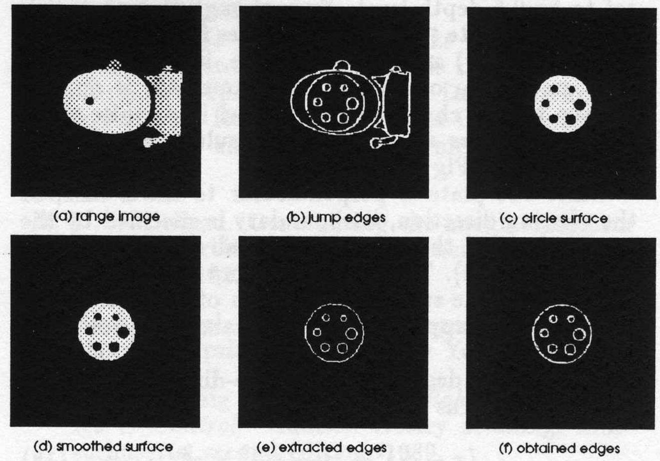


Figure 3: A Circle Surface from a Range Image

The derivation of $x$ and $y$ caused by the integration of $i$ and $j$ could be no greater than half a pixel, and the noise added to each pixel would not exceed 9.2 absolute range levels in practical applications[2]. It makes $\sigma_x \leq 0.5$, $\sigma_y \leq 0.5$, and $\sigma_z \leq 9.2$. For simplicity, suppose there is no need to transform formats or reference systems when comparing $S_i$ and $S_j$, i.e., $\Lambda_i' = \Lambda_i$ and $\Lambda_j' = \Lambda_j$. From Eq. 13, the distance between any parameter of $S_i$ and $S_j$ should be no greater than 4 for a 90% confidence,

$$\Delta^2 x \sigma_x + \Delta^2 y \sigma_y + \Delta^2 z \sigma_z = 0.5(\Delta^2 x + \Delta^2 y) + 9.2\Delta^2 z \leq 4.$$

If $\Delta^2 x$ and $\Delta^2 y$ take the maximum difference as 1.0, then the absolute value of $\Delta z$ must be no greater than 0.6 range level. Or if only quantification noise exist, i.e., $\sigma_z = 0.0$, the 2D distance of two features must be less than $\sqrt{8}$ pixels.

## 4 Experiment with Range Images

The LSP method has been tested with a pair of range images. It is a pencil sharpener viewed in two viewing directions with a 60° rotation. There are 256 lines and 256 columns in each image. The distance between pixels is 0.75mm, and the scale factor in z direction is 100.

A localized surface description is the description of a surface primitive under its own local coordinate system. Given the range image in Fig 3(a), it can be easily segmented into surfaces corresponding to regions in the image divided by jump edges. After applying a neighborhood operation on the original range image, which thresholds pixels that have a bigger than 4 depth difference with a neighbor pixels, surface segmentation is obtained as shown in Fig 3(b).

The surface description for each separated region is then approximated from its eroded center region, and the pixels that lie on the approximation surface form the corresponding surface patch. Working on the region within the big circle, a circle plate of the pencil sharpener is extracted(Fig 3(c)) when a threshold is set to be 0.5 depth level. Even though there is a tiny hole on the plate that actually comes from the original image(Fig 3(a)) as a result of noise, it can be deleted with smoothing operation. If an operator is applied on Fig 3(c) to change a blank pixel into white when all its neighbors are white, the result is a smoothed circle plate in Fig 3(d).

Since the plate is perpendicular to the z axis, or the viewing direction, its boundary is specified by the jump edges of the two-dimensional smoothed circle plate in Fig 3(d). The extracted jump edges are shown in Fig 3(e). The surface description of the circle plate is obtained by applying a Hough transformation[1] on the jump edges.

The general description of a two-dimensional circle has the format as

$$(x - c_1)^2 + (y - c_2)^2 = c_3^2 \tag{18}$$

where $c_1$ and $c_2$ are translation factors and $c_3$ is the radius. For any edge point, its $x$ and $y$ values are fixed. In the parameter space $< c_1, c_2, c_3 >$, every edge point is examined with different $c_1$, $c_2$, and $c_3$ according to Eq. 18. The approximated coefficients for circles on the plate are obtained from the clusters with the most edge points. More specifically, for a cluster $< c'_1, c'_2, c'_3 >$, $c_i$ is the sum of $m_i\, c'_i + n_i\,(c'_i - 1) + k_i(c'_i + 1)$ divided by $m_i + n_i + k_i$ for $i = 1, 2, 3$, where $m_i, n_i$, and $k_i$ are the number of edge points in parameter space at $c'_i$, $c'_i - 1$, and $c'_i + 1$ respectively. In such a way, seven circles $C_i$, $0 \le i \le 6$ are obtained with parameters shown as below. The image to overlap the approximated surface description with extracted jump edges is given in Fig 3(f).

$$\begin{array}{ll} C_0 = \{129.0,\ 128.1,\ 39.1\} & C_1 = \{141.0,\ 107.7,\ 4.1\} \\ C_2 = \{116.8,\ 108.2,\ 4.7\} & C_3 = \{104.7,\ 126.8,\ 5.0\} \\ C_4 = \{115.8,\ 148.2,\ 5.8\} & C_5 = \{140.1,\ 149.2,\ 7.1\} \\ C_6 = \{152.2,\ 128.9,\ 7.6\} & \end{array}$$

Taking the centroid of $C_0$ as the origin of the local coordinate system, z as local z, the direction connecting the two centroids from $C_0$ to the biggest inside circle $C_6$ as local x, and y = z × x, the localized description of the circle plate is obtained as the following equation,

$$S = S_0 - (S_1 + S_2 + S_3 + S_4 + S_5 + S_6) \tag{19}$$

where,

$$\begin{aligned} S_0 &= x^2 + y^2 - 39.1^2 \\ S_1 &= (x + 11.3)^2 + (y + 20.8)^2 - 4.1^2 \\ S_2 &= (x - 12.9)^2 + (y + 19.5)^2 - 4.7^2 \\ S_3 &= (x - 24.3)^2 + (y + 0.5)^2 - 5.0^2 \\ S_4 &= (x - 13.2)^2 + (y - 20.5)^2 - 5.8^2 \\ S_5 &= (x + 11.8)^2 + (y - 20.7)^2 - 7.1^2 \\ S_6 &= (x + 23.2)^2 + y^2 - 7.6^2 \end{aligned}$$

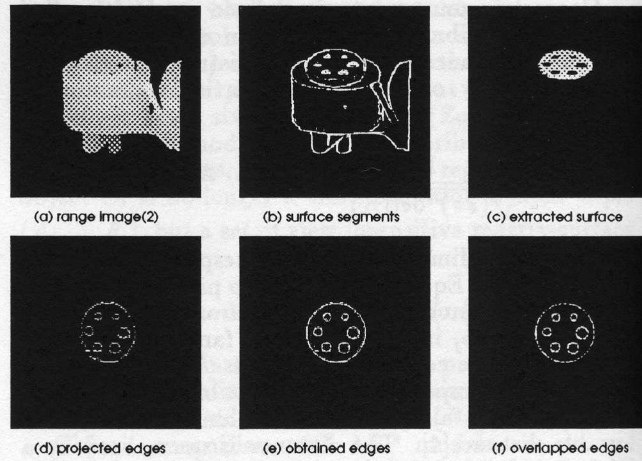


Figure 4: The Circle Surface from Another Image

In another image, the data is obtained after rotating the pencil sharpener for 60°(Fig 4(a)). Similar to what has been done with the first image, the circle plate is extracted(Fig 4(c)) when surface approximation is applied on the separated region that are segmented with jump and roof edges(Fig 4(b)). At this time, however, Hough transformation is performed on the edge points in the orientation of their average normal direction. In other words, these edge points are transformed so that surface description is approximated under the local z. The transformed edge points are shown in Fig 4(d), and the circles obtained by Hough transformation are shown in Fig 4(e) and listed below.

$$\begin{array}{ll} C'_0 = \{129.1,\ 128.0,\ 39.1\} & C'_1 = \{141.1,\ 107.0,\ 4.2\} \\ C'_2 = \{117.1,\ 107.2,\ 4.8\} & C'_3 = \{105.0,\ 127.1,\ 5.4\} \\ C'_4 = \{116.0,\ 146.7,\ 5.9\} & C'_5 = \{140.1,\ 147.7,\ 7.2\} \\ C'_6 = \{152.1,\ 127.9,\ 7.7\} & \end{array}$$

As a result, the localized surface description is given in the following equation when the local coordinate system, which sits at the centroid of the biggest circle $C'_0$, is correspondingly set as its z axis is the average normal direction, its x axis connects the centroids from $C'_0$ to $C'_6$, and y = z × x.

$$S' = S'_0 - (S'_1 + S'_2 + S'_3 + S'_4 + S'_5 + S'_6) \tag{20}$$

where,

$$\begin{aligned} S'_0 &= x^2 + y^2 - 39.1^2 \\ S'_1 &= (x + 12.1)^2 + (y + 20.9)^2 - 4.2^2 \\ S'_2 &= (x - 11.9)^2 + (y + 20.9)^2 - 4.8^2 \\ S'_3 &= (x - 24.1)^2 + (y + 1.0)^2 - 5.4^2 \\ S'_4 &= (x - 13.2)^2 + (y - 18.6)^2 - 5.9^2 \\ S'_5 &= (x + 10.9)^2 + (y - 19.7)^2 - 7.2^2 \\ S'_6 &= (x + 23.0)^2 + y^2 - 7.7^2 \end{aligned}$$

To compare the two localized surface descriptions between Eq. 19 and Eq. 20, each circle in $S'$ is checked with all the circles in $S$. The matched candidates are obtained as follows when the radius threshold is set to be 0.4.

$$\begin{array}{ll} \{P^{-1}(S'_0)\} = \{S_0\} & \{P^{-1}(S'_1)\} = \{S_1\} \\ \{P^{-1}(S'_2)\} = \{S_2, S_3\} & \{P^{-1}(S'_3)\} = \{S_3\} \\ \{P^{-1}(S'_4)\} = \{S_4\} & \{P^{-1}(S'_5)\} = \{S_5, S_6\} \\ \{P^{-1}(S'_6)\} = \{S_6\} & \end{array}$$

For those circles with multiple candidates, which are $S'_2$ and $S'_5$, their translation parameters are checked again. As thresholds are set to be 1.0 in x direction and 1.4 in y, or $\sqrt{3}$ Euclidean distance in total, the out of limit parameter difference between $S_3$ and $S'_2$ takes $S_3$ out of the candidate list of $S'_2$. It is the same to delete $S_6$ from $\{P^{-1}(S'_5)\}$. It makes the one to one matching between $S'$ and $S$.

$$\begin{aligned} P^{-1}(S') &= P^{-1}(\{S'_0, S'_1, S'_2, S'_3, S'_4, S'_5, S'_6\}) \\ &= \{S_0, S_1, S_2, S_3, S_4, S_5, S_6\} = S' \end{aligned}$$

Fig 4(f) shows the result when $S'$ and $S$ are overlapped together.

Because of their identical descriptions, surfaces defined in local coordinate systems are compared directly with each other even if they are extracted from different viewing directions. Under noisy practical conditions, approximated descriptions for a surface patch exhibit parameter changes when obtained from images of different viewpoints(Eq. 12). These parameter changes are theoretically related by the covariance matrix $\Lambda'_i$ in Eq. 16, and practically demonstrated in the experiment as parameter differences between Eq. 19 and Eq. 20.

Nevertheless, the result is quite encouraging. Despite the fact that the resolution of the second image is only half in y direction, the parameter difference is still within the $\sqrt{8}$ limitation for a 90% confidence. Of course, the half resolution obviously caused more difference in y axis, which has a maximum 1.4 in comparison to the maximum 1.0 in x axis. It was because the changed y components produced more effects on the matrix $\Lambda'_i$ than x. In terms of real objects, under the condition of the experiment, two circle plates will be recognized as the same only if the difference between circle radiuses is equal to or less than 0.3mm and the difference in translation is equal to or less than 1.05mm.

# 5 Conclusion

LSP is a method to obtain view insensitive 3D representations. By introducing local coordinate systems that are related with surface properties only, localized surface primitives are obtained for view insensitive object descriptions. In this way, observed objects can be compared directly with object models in spite of any spatial transformations. Fundamental ideas have been discussed in the paper. Further research will be concentrated on the application in complex objects.

# 6 Acknowledgment

I would like to express my thanks to Dr. Wayne Davis, my former supervisor, for his inspiration and help during my Ph.D study, and to Dr. C Archibald at the National Research Council of Canada for the range images he kindly provided to me.

# References

[1] D. H. Ballard. Generalizing the hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2):111–122, 1981.

[2] Paul. J. Besl. *Surfaces in Range Image Understanding.* Springer-Verlap, New York, 1988.

[3] John. J. Craig. *Introduction to Robotics: Mechanics & Control.* Addison-Wesley Reading, Massachusetts, second edition, 1989.

[4] James L. Crowley and Fano Ramparany. Mathematical tools for representing uncertainty in perception. In *Spatial Reasoning and Multi-sensor Fusion: Proceeding of the 1987 Workshop*, pages 293–302. New York: Academi, 1987.

[5] Ting-Jun Fan. *Describing and Recognizing 3D Objects Using Surface Properties.* Springer-Verlag, 1990.

[6] J. Lansdown. Methods of presentation of 3D structures. In *Geometric Modeling and Computer Graphics: Techniques and Applications*, pages 223–232. Technical Press, 1987.

[7] Ping Liang and John S. Todhunter. Representation and recognition of surface shapes in range images: A differential geometry approach. In *Computer Vision, Graphics, and Image Processing*, volume 50, pages 77–109. Acedemic Press, Inc., October 1990.

[8] M. Magee and M. Nathon. Spatial reasoning, sensor repositioning and disambiguation in 3D model based recognition. In *Spatial Reasoning and Multi-sensor Fusion: Proceeding of the 1987 Workshop*, pages 262–271. New York: Academi, 1987.

[9] R. Smith, M. Self, and P. Cheeseman. Uncertain geometry in robotics. In *Proceedings of the International Conference on Robotics and Automation*, pages 850–856, 1987.

[10] A.K.C. Wong and S.W. Lu. Recognition and shape synthesis of 3D objects based on attributed hypergraphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 279–290, March 1989.

[11] Xiaobu Yuan. *3D Reconstruction as An Automatic Modeling System.* Ph.D thesis, Department of Computing Science, University of Alberta, Canada, 1992.