Pose estimation of two-pose 3D models using the base tangent plane and stability constraints

Soon-Yong Park and Murali Subbarao

Department of Electrical and Computer Engineering
State University of New York at Stony Brook
Stony Brook, NY11794-2350, USA
Email: parksy@ece.sunysb.edu

Abstract

A pose estimation technique is presented to determine coarse registration parameters between two 3D models of an object. The models are reconstructed by merging multi-view range images of two different poses of the object, which is arbitrarily placed on a turntable. The result of pose estimation will facilitate registration refinement and integration of two 3D models. Registration refinement of two 3D models usually requires a rough estimate of pose between them. However, obtaining such estimate is a difficult problem and computationally expensive. We introduce a simple pose estimation technique based on matching tangent planes of a 3D model with the base tangent plane (BTP) which is invariant for a vision system. In order to reduce computation time, we employ geometric constraints to find consistent and stable tangent planes (STP). A STP is a plane on which the object can rest in a stable state on a turntable. By matching the base tangent plane of one 3D model to a stable tangent plane of the other model, we derive a pose transformation matrix and register the models to a common coordinate system. We find the best-matching pose transformation matrix which minimizes a pose error beween two models. Experimental results on several real objects are presented to demonstrate the effectiveness of the method.

1 Introduction

In *3D Computer Vision*, determining the pose of a 3D model is a challenging problem without *a priori* knowledge of the model. In this paper, we consider a pose estimation problem between two models which are reconstructed from two poses of an object. The result of pose estimation will facilitate registration and integration of two 3D models to reconstruct a complete 3D model of the object.

A 3D model of an object can be reconstructed by a computer vision technique. One common technique is empolying a fixed range sensor and a turntable to obtain multi-view range images, and registering and integrating them into a complete 3D model [6]. However, such a system has a problem with reconstructing a complete 3D model. Given a single pose of the object, there will be constraints on viewing direction of the range sensor. For example, the top or the bottom of an object may not be visible from the sensor, and there may be reconstructed an incomplete 3D model of the object. Therefore, it is necessary to acquire all visible surfaces to reconstruct a complete 3D model. One approach is placing the object on the table in two different poses, acquiring corresponding 3D models, registering, and integrating them into a complete model. Our previous work shows registration and integration of two pose models of several real objects [13].

Registration refinement of multi-view range images is facilitated by initial calibration parameters of a vision system [2, 4]. However, without *a priori* knowledge of transformation parameters between two pose models, registration of the pose models is very difficult. Suppose we place an object in an upright pose on the turntable for the first 3D model reconstruction, and turn the object and place it on its side for the second model. Then there are 6 degrees of freedom between the two models, and registration refinement is very difficult without an initial rough estimate of the pose.

There have been many investigations on pose estimation for multiple 3D models [1, 3, 12]. In this paper, we introduce a novel pose estimation technique of two 3D models. The proposed technique finds a stable tangent plane (STP) on a 3D model which can be transformed to the base tangent plane (BTP) of the other model and *vice versa*. The BTP is the flat plane of the turntable on which the object is placed. For a rigid object, the BTP is a tangent plane of the object's outer surface. The BTP is fixed or invariant for a given turntable vision system irrespective of the object's pose. When an object rests in a stable state on the flat turntable top, the BTP coincides with an STP of the object. Therefore, we match the BTP of the first pose to an STP of the second pose, and simultaneously match the BTP of the second pose to an STP of the first pose. The best match that minimizes a volumetric error between the two poses is used to estimate the transformation matrix between the two poses.

The procedures of estimating the pose between two models are as follows. We find a set of tangent planes of each pose model using the Extended Gaussian Image (EGI) of the model [8, 9]. We refine tangent planes by interpolating their supporting vertices. By employing geometric constraints on the tangent planes, we reduce the number of tangent planes and obtain consistent and stable tangent planes (STP) for each pose model. For each possible match of a BTP and a STP between the two poses, we derive a pose transformation matrix. Then we register one 3D model to the other's coordinate system, and measure a volumetric pose error between them. The best matching STP is determined from each model which minimizes the pose error. Experimental results show that our technique estimates a reliable initial pose for further registration refinement of the 3D models.

2 3D Model reconstruction

A 3D model is reconstructed by integrating multiple range images [5, 13] of an object. Range images of the object are reconstructed by using a stereo-vision technique. The vision system is initially calibrated by Tsai's calibration algorithm [14]. In order to reconstruct the first 3D model, we place the object in an upright pose on a turntable and acquire multiple range images by turning the table. They are then registered by using the calibration parameters, and integrated to obtain the first 3D model. Then, we turn the object and place it on its side on the table. And we obtain another range image set and reconstruct the second 3D model. Figure 1 shows a schematic diagram of our vision system.

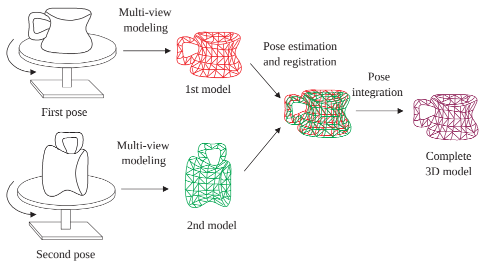


Figure 1: Two poses models are reconstructed, registered, and integrated into a complete 3D model. Pose estimation is needed for the registration refinement

3 Methodology

3.1 Base tangent plane

When we place a rigid object on the flat (planar) top of a turntable, the object rests with its outer surface touching the table top. The planar table top will be a tangent plane of the object's surface. We call the planar table top the base tangent plane (BTP) of the turntable. The BTP is a global tangent plane in the sense that it will not intersect the object's volume anywhere (in contrast, a local tangent plane may intersect the object's volume at a point far from the point of tangency). The BTP is invariant with respect to the object's pose and the world coordinate system. The following are some characteristics useful in pose estimation.

- The BTP of the turntable is a global tangent plane of the object.
- There exists a unique tangent plane of the first pose model which corresponds to the BTP of the second pose, which is also a tangent plane of the second model.

Suppose an object is placed on the turntable with two different poses, $Pose1$ and $Pose2$ as shown in Figure 2. Then there is only one plane normal $\hat{n}_{T1}$ in the first pose which matches the normal of the BTP, $\hat{n}_B$ in the second pose. Similarly, there is a normal vector $\hat{n}_{T2}$ in $Pose2$, which matches $\hat{n}_B$

in $Pose1$. Because $\hat{n}_B$ is a common and invariant vector in the vision system, we can estimate a rotation matrix using $\hat{n}_{T1}$ and $\hat{n}_{T2}$.

Let P_{12} be a transformation matrix from $Pose1$ to $Pose2$. By aligning $\hat{n}_{T1}$ with $\hat{n}_B$, we estimate a rotation matrix except one degree of freedom along the axis of $\hat{n}_B$. Then, by aligning a transformed vector $P'_{12}\hat{n}_B$ with $\hat{n}_{T2}$, we estimate the rotation matrix. Translation is estimated by Center of Mass (CoM) of two models. We construct a new coordinate system for each tangent plane T_1 and T_2. It will be desribed in the next section. And the rotation matrix is estimated by transforming the coordinate system T_1 and T_2 to the common coordinate system of the base B.

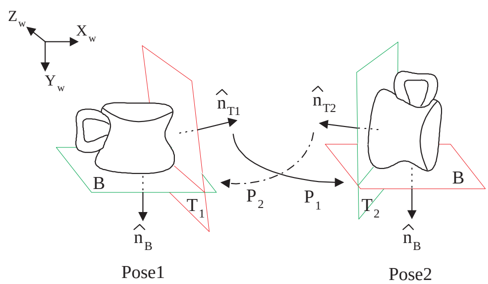


Figure 2: Matching methology. A tangent plane of the first pose $\hat{n}_{T1}$ uniquely matches with the base plane of the second pose $\hat{n}_B$, and *vise versa*.

Consequently, our technique finds tangent planes T_1 and T_2 for each model, which minimize a cost fuction, or a volumetric pose error, between two models. The pose error is estimated in terms of a SSD (Sum of Square Difference) error between two models. Suppose a vertex $\mathbf{v}_{1i}$ on $Pose1$ corresponds to another vertex $\mathbf{v}_{2i}$ on $Pose2$. If there is a transformation P_{12} from $Pose1$ to $Pose2$, then the pose error is measured by

$$\sum{}^2 = \sum_{i=0}^{K} \|P_{12}\mathbf{v}_{1i} - \mathbf{v}_{2i}\|^2, \qquad (1)$$

where, K is the number of vertices in $Pose1$, but it is usually down-sampled to increase computation speed.

3.2 Stability constraints

The surfaces of the object are represented by a finite number of triangle meshes [10]. Therefore, there will be a finite number of tangent planes for each mesh model. Let a set of tangent planes of $Pose1$ be $\mathcal{T}_1$, and $\mathcal{T}_2$ be another set of tangent planes of $Pose2$. If we assign a tangent plane for every vertex, however, measuring the cost fuction for all combinations of two sets $\mathcal{T}_1$ and $\mathcal{T}_2$ is computationally expensive.

In order to reduce computational complexity, we remove some local or unstable tangent planes by employing geometric constraints. A key idea for employing the constraints is that the object is placed on the turntable in a stable pose and the turntable is horizontal.

1. *Base plane constraint* : An object is placed on the BTP which is one of the global tangent planes of the object. This BTP does not intersect the object's volume.
2. *Stability constraint* : The BTP of the turntable is horizontal and the object is in a stable pose. Therefore, the projection of the CoM to a STP is always inside the convex hull of its supporting vertices.
3. *Height constraint* : If two pose models are correctly registered, their heights will be very similar (It may not be the same, because of noise).

Based on the constraints above, we consider only stable tangent planes (STPs) on the model. These constraints greatly reduce the number of tangent planes and the computation time for pose matching.

4 Pose estimation

First we list the steps in our pose estimation algorithm and then provide additional details.

1. For both pose models, obtain an EGI using vertex normals of the 3D model.
2. For each face of the tesselated sphere, determine whether the tangent plane corresponding to the face is a candidate for matching with the base plane of the other pose. The tangent plane is a valid candidate in all cases except the following. If the orientation histogram count of the face is zero, or the angle between the face

normal and the vertical axis is smaller than a threshold, then the tangent plane is not a candidate. The latter condition is due to the assumption that the relative rotation between the two poses is greater than a threshold angle.

3. Reject any tangent plane T which is intersecting its own volume by checking the signed distance from each voxel to the plane.
4. Reject any tangent plane T for which the projection of the CoM of the model is outside the convex hull of its supporting vertices.
5. Reject any tangent plane T if the height difference bewteen the model transformed by P and the fixed model is greater than a threshold δ_H.
6. Find two matching planes from both models by measuring the pose error between the two registered models.

4.1 Finding tangent plane

In the first step, an EGI of a pose model is constructed and tangent planes on a tesselated Gaussian sphere are obtained. Suppose $\hat{n}_f$ is the normal vector of a tesselated polygon and $\hat{n}_T$ is the normal vector of an associated tangent plane, where $T = Ax + By + Cz + D$. Because of the first constraint, we assume that the tangent plane passes through a vertex whose projection distance to $\hat{n}_f$ is the maximum. We call this vertex v_m a *frontier vertex*. Then we initialize the plane $T(A, B, C, D)$ using the *frontier vertex* and the face normal.

However, because we search for a STP which can support the object as a base plane, we refine the tangent plane T by employing some supporting vertices. Supporting vertices of a tangent plane must have similar normal vectors and be close to the plane. Therefore we find some vertices whose normal vectors $\hat{n}_v$ and their dot product, $\hat{n}_v \cdot \hat{n}_f$ are less than a threshold $cos(\theta_G)$ as shown in Figure 3. Ideally, supporting vertices should be on the same tangent plane so that we can find at least 3 vertices as supporting vertices. However, due to inherent errors on model's surface, no vertex is exactly on the plane of T. Instead, we search for supporting vertices which are close to the plane and refine the plane T.

As shown in Figure 4(a), we select a vertex v_i as one of the supporting vertices, if its relative distance to the plane, D_i/D_{mi}, is less than a threshold. Here, D_i is the distance from v_i to the plane and D_{mi} is the distance from the *frontier vertex* to V_i.

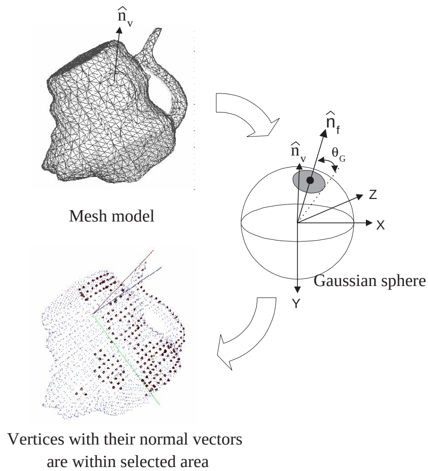


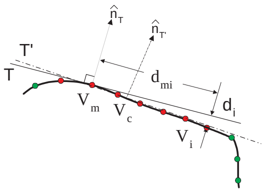
Figure 3: Initializing a tangent plane

In order to avoid selecting vertices only in a very small region (it may produce an unstable plane), we pick only one vertex from a triangle face when mutiple vertices of the face meet the condition. After finding a set of supporting vertices $\mathcal{V}_\mathbf{s}$, we move the origin of the plane to V_c, the centroid of $\mathcal{V}_\mathbf{s}$, and refine the normal vector $\hat{n}_{T'}$ by averaging normal vectors of $\mathcal{V}_\mathbf{s}$.

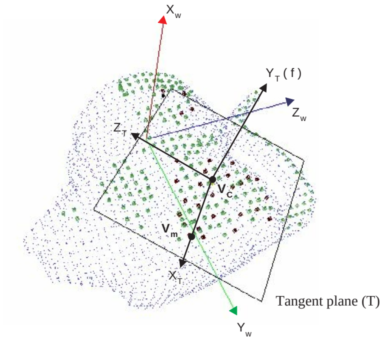
A new coodinate system is then generated for the tangent plane in order to obtain a transfomration matrix to the reference coordinate system. We set the normal vector $\hat{n}_T$ (let T denote a refined tangent plane from now on) to the Y axis as shown in Figure 4(b), because it is convenient to match with the Y_W of the reference coordinate system. Y_W is also the vector normal of the base plane. For X axis, a vector from V_c to the projection of V_m to T is normalized. And the Z axis is set according to the right hand coordinate system. In Figure 4(b), the coordinate system of a tangent plane T is shown. A 3D model is represented as point clouds, the green (light-grey) dots are vertices whose normal vectors are within angle θ_G from $\hat{n}_f$, and the red (dark-grey) dots are all supporting vertices selected from green vertices.

4.2 Finding stable tangent plane (STP)

For two 3D models, we find all tangent planes using EGIs. Each tangent plane T consists of sup-



(a)



(b)

Figure 4: Construction of a tangent plane on a 3D model. (a) Tangent plane T is refined by their supporting vertices. (b) Green (light grey) dots are all vertices with their normals within a threshold angle θ_G. But only red dots (dark grey) consist of supporting points on the tangent plane.

all vertices to determine if a tangent plane intersects the volume. It can be easily checked by measuring the dot product $\hat{v} \cdot \hat{n}_T$ is greater than a parameter D of the plane. Considering errors on model's surfaces, we set a threshold to $D + \delta_I$.

Next constraint is stability of the tangent plane, because we are finding a tangent plane which can be the BTP on which the object can be placed. A simple algorithm is employed to check its stability. Given the Center of Mass (CoM) of a 3D model, its projection to tangent plane T, CoM_T should be inside the convex hull of projections of all supporting vertices. The object will be unstable and fall over if CoM_T is outside the convex hull as shown in Figure 5.

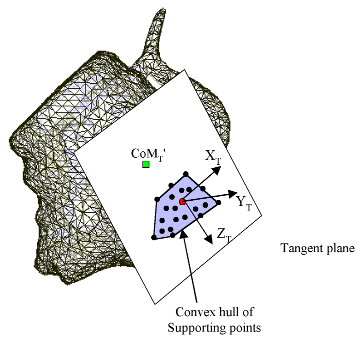


Figure 5: An unstable tangent plane. The projection of CoM to the plane is outside convex hull of supporting vertices.

porting vertices and has its own coordinate system. Using tangent planes from each 3D model, we can estimate a pose transformation between two models. However, there are a large number of tangent planes on each model, and it is computationally very expensive to find the matching planes by measuring the cost funtion for every combination of tangent planes. If there are N and M tangent planes on each model, we need $O(NM)$ computations for estimation. Therefore, it is necessary to reduce the number of planes to decrease computation time.

This section describes finding STPs based on constaraints we described earlier. First of all, a tangent plane T whose number of supporting vertices is less than 3 is rejected, because it is not stable. The first constraint is that the base plane can not intersect the object's volume. Therefore we check

The last constraint is comparison of object's height. By transforming T using an initial matrix P and aligning the Y_T axis of T with the Y_W axis of the world coordinate system, we compare the height of the transformed model with that of the other model. We reject tangent planes when the height difference is greater than a threshold δ_H.

Figure 6 shows an example of finding STPs. We can see that the number of tangent planes is greatly reduced after rejecting inconsistent and unstable planes.

4.3 Matching tangent planes

Rejection of unstable and inconsistent tangent planes significantly reduces the number of tangent planes. The last step in pose estimation is finding a

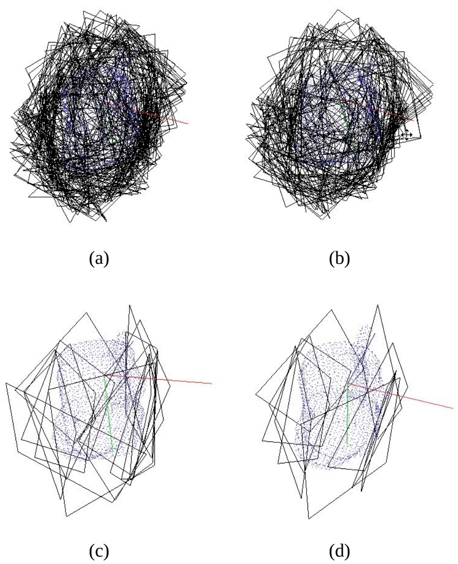


Figure 6: Tangent plane rejections based on geometric constraints. (a) Tangent planes without any constraint (b) After removing volume-intersecting planes (δ_I = 3mm) (c) After removing unstable planes (d) After height comparison (δ_H = 3mm)

tangent plane in each pose which matches two 3D models by minimizing a pose error. For every STP in $\mathcal{T}_\mathbf{1}$, we derive a transformation matrix P_{ij} using every STP in $\mathcal{T}_\mathbf{2}$, measure the pose error, and find two STPs which yield the best-matching. Let an STP in $\mathcal{T}_\mathbf{1}$ be T_1 and another STP in $\mathcal{T}_\mathbf{2}$ be T_2. The transfomration matrix from the first pose (O_1) to the second pose (O_2) is estimated by using T_1 and T_2. The transformation matrix first aligns T_1 with the base plane B

$$B' = {T_1}^{-1} B, \quad (2)$$

and T_2 with B' by rotating the model along the Y axis

$$T_2 = {R_y}^{-1} B'. \quad (3)$$

The coordinate system of B is the same as the world (or common) coordinate system. A rotation matrix R_y is aligns B' with the coordinate system T_2. It is a rotation along the Y axis and computed by

$$R_y = \begin{pmatrix} cos\theta & 0 & -sin\theta & 0 \\ 0 & 1 & 0 & 0 \\ sin\theta & 0 & cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$where, \begin{pmatrix} cos\theta \\ sin\theta \end{pmatrix} =$$

$$\begin{pmatrix} B'_x + B'_z & B'_z - B'_x \\ B'_x - B'_z & B'_x + B'_z \end{pmatrix}^{-1} \begin{pmatrix} T_{2x} + T_{2z} \\ T_{2x} - T_{2z} \end{pmatrix}. \quad (4)$$

Translation between two poses is decided by CoMs of the models. Center of mass of a 3D model is computed by a mass-computation technique [11]. Let the translation from the origin of the common coordinate system to each CoM be M_1 and M_2. Then the pose transformation matrix P_{12} from O_1 to O_2 is computed by

$$O'_1 = M_2 {R_y}^{-1} {T_1}^{-1} {M_1}^{-1} O_1 \quad (5)$$

$$= P_{12} O_1 \quad (6)$$

The best matching pose transformation P_{12} is estimated by minimizing the cost function as in Eq.(1) between two models,

$$\min_{\{T_1 \in \mathcal{T}_\mathbf{1}, T_2 \in \mathcal{T}_\mathbf{2}\}} \{\sum \|O_2 - P_{12} O_1\|^2\} \quad (7)$$

5 Experimental results

We tested our pose estimation technique on several real objects. Two 3D pose models are reconstructed by a volumetric modeling technique by rotating the turntable [13]. Figure 7(a) and (b) shows front images of two poses of an object “Monkey”. Figure 7(c) is an overlapping of two models in their original poses. After pose estimation, the second model is registered to the first model as shown in 7(d). Figure 7(e) and (f) show an integrated and a textured 3D model. The number of tangent planes at every rejection step is shown in Table 1. This object is polygonized with a voxel size of 4mm [10]. Table 2 shows threshold angle θ_G, threshold distances δ_H and δ_I, average pose error between two models, number of vertices, and total estimation time. For measuring pose error, we use Equation (1). From a vertex v_{1i} in O_1, we find the closest vertex v_{2i} in O_2 as the corresponding one. We sampled 50 vertices from a 3D model for the error measure. Figure 8 is the result of another object “PolarBear” and Figure 9 shows the result of an object “Pokemon”. All results show that our technique reliably estimates the pose of two models.

Table 1: Number of tangent planes. (a)Initial planes (b) Intersection constraint (c) Stability constraint (d) Height constraint

Constraints / Object		(a)	(b)	(c)	(d)
Monkey	pose1	186	141	18	8
	pose2	171	154	21	4
PolarBear	pose1	180	150	18	7
	pose2	167	144	11	4
Pokemon	pose1	205	173	8	2
	pose2	185	173	17	9

Table 2: Pose estimation error and estimation time ($D_i/D_{mi} = 0.1$)

Object	Monkey	Polarbear	Pokemon
Voxel size (*mm*)	4	4	2
cos(θ_G)	0.8	0.5	0.5
δ_I,δ_H (mm)	3.0	3.0	3.0
Avg. error (*mm*)	2.46	1.95	2.44
No. vertices	3294	5106	6180
Time (*sec*)	22.8	24.3	27.5

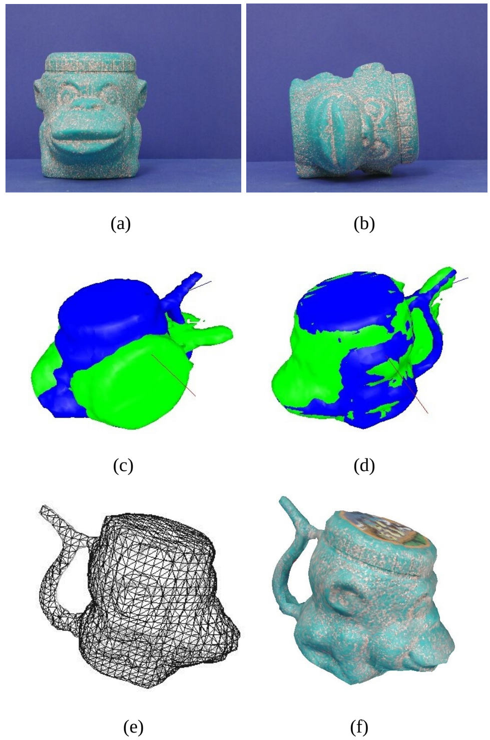


Figure 7: Pose estimation result for "Monkey" (a) Pose1 (b) Pose2 (c) Before estimation (d) After estimation (e) Integrated mesh (f) Texture mapping

6 Conclusions

Pose estimation bewteen two 3D models is estimated by matching global tangent planes with base planes. Two 3D models are reconstructed by multi-view modeling of two different poses of an object. In order to reconstruct a complete 3D model by registering and integrating two pose models, we estimate the pose between two models. Because the object always stands on one of its global tangent planes, we find a tangent plane from the first model, which matches the base plane of the second plane and *vice versa*.

Matching tangent plane with the BTP gives constraints that reject inconsistent and unstable tangent planes. We employ three constraints- base plane constraint, stability constraint, and object height constraint, to reject such tangent planes. Experimental results show a great decrease in the number of tangent planes and matching complexity. We use our pose estimation for an initial registration of two 3D models, refine the registration, and integrate 3D models. Experiments on real objects show that our technique is effective.

References

[1] G.Barequet and M.Sharir, "Partial surface and volume matching in three dimensions," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 19, no. 9, pp.929-948, 1997.

[2] R. Bergevin, M. Soucy, H. Gagnon, and D. Laurendeau, "Toward a general multi-view registration technique," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, no. 5, May. 1996.

[3] R. J. Campbell and P. J. Flynn, "A survey of free-form object representation and recognition techniques," *Computer Vision and Image Understanding*, vol.81, pp. 166-210, 2001.

[4] Y. Chen and G. Medioni, "Object modeling by registration of multiple range images," *Image and Vision Computing*, vol. 10, no. 3, pp.

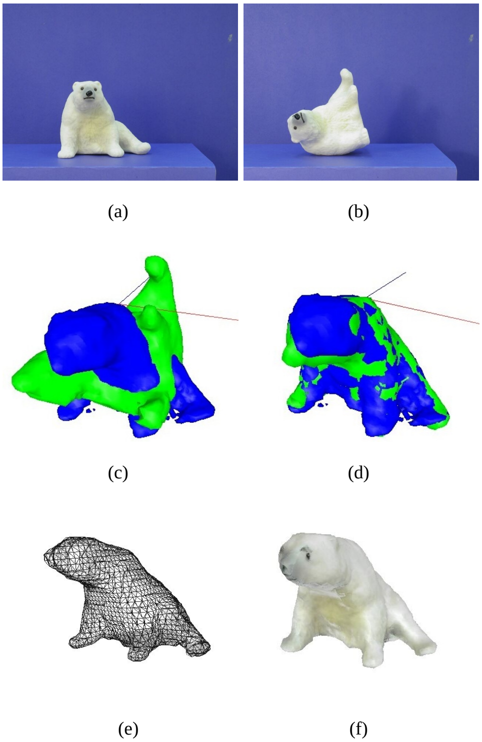


Figure 8: Pose estimation for “Polar Bear”

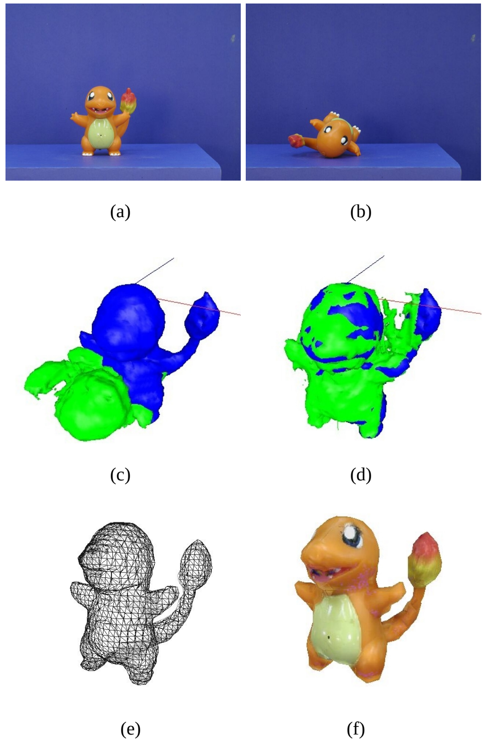


Figure 9: Pose estimation for “Pokemon”

145-155, 1992.

[5] B. Curless and M. Levoy, “A volumetric method for building complex models from range images,” In *Proceedings of SIGGRAPH*, pp. 303-312, 1996.

[6] A.W. Fitzgibbon, G. Cross, and A. Zisserman, “Automatic 3D model construction for turn-table sequences”, *European Workshop SMILE’98, Lecture Notes in Computer Science 1506*, pp. 155-170, 1998.

[7] A. Hilton, A.J. Stoddart, J. Illingworth, and T. Windeatt, “Reliable surface reconstruction from multiple range images,” *Proceedings. of the ECCV’96*, pp. 117-126, Springer-Verlag, 1996.

[8] K.Ikeuchi and M.Hebert, “Spherical representations: from EGI to SAI,” *Technical report CMU-CS-95-197*, Oct. 1995.

[9] S.B.Kang abd k.Ikeuchi, “3D object pose determination using complex EGI,” *Technical report CMU-RI-TR-90-18*, Oct. 1990.

[10] W.E. Lorensen and H.E. Cline, “Marching cubes: A high resolution 3D surface construction algorithm,” *Computer Graphics*, vol. 21, no. 4, July 1987.

[11] B.Mirtich, “Fast and accurate computation of polyhedral mass propertices,” *Journal of Graphics Tools*, 1996.

[12] M.Novotni and R.Klein, “A geometric approach to 3D object comparison,” In *Proceedings of International Conference on Shape Modeling and Applications(SMI2001)*, May 2001.

[13] S.Y. Park and M. Subbarao, “Automatic 3D model reconstruction using voxel coding and pose integration,” *International Conference on Image Processing*, Sep. 2002.

[14] R.Y. Tsai, “A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV camera and lenses,” *IEEE Journal of Robotics and Automation*, vol. 3, no. 4, Aug. 1987.