Adaptive Three Dimensional Laser Scanning and Object Reconstruction

THEODORE LILAS, STEFANOS KOLLIAS
Dept. of Electrical and Computer Engineering
National Technical University of Athens
Zografou, Athens,
GREECE

Abstract:- Generating models of real objects is one of the central goals in computer vision. The aim of this paper is to present a method for deriving precise models of three-dimensional objects. The proposed approach efficiently combines techniques and algorithms used in several scientific fields closely related to computer vision. Stereo vision and photogrammetry are used in order to obtain accurate three-dimensional measurements. Methods of computational geometry based on Delaunay triangulations are used in order to generate a three dimensional model of the object. Pattern matching and segmentation techniques are used in order to evaluate the model and segment it. Path planning techniques are used for guiding the tool around the object in order to get all the necessary measurements to refine the model.

The outline of the proposed methodology is the following. Initially we calibrate the system. Then we use a robot arm to scan the object with a laser beam and obtain range data. Next we examine the surface of the object and identify areas where the object has not been sampled properly. Based on these results a new scanning strategy is derived. The new scanning operation reveals new information on non-accurately sampled areas. The resulting solid models are merged providing a more accurate model. The above process is repeated until the desired quality based on the structure of the object is reached.

Our approach is applied in cases where the structure of the object is known and simple, but also when the structure is complex. The proposed methodology has the following main characteristics: a) it adapts to object morphology, b) refines the initial estimation and c) provides a complete sequence of operations for deriving a three dimensional model.

Key-Words:- adaptive laser scanning, three dimensional object reconsuction CSCC'99 Proc.pp.5441-5446

1 Introduction

One of the central goals in computer vision is to generate models of real world objects. Several methods have been developed for deriving depth and shape of objects. These methods include stereo vision, shape from shading and techniques estimating motion and shape simultaneously. When it is required to precisely reconstruct an object, accurate range data are necessary. Methods, which sample the surface of the object and obtain range data, include laser triangulation, radar measurements, lens focus and interferometry. The sampled set of points on the surface of the object form the initial data, which are then processed by object reconstruction algorithms. In general methods for object reconstruction use (a) surface representations with polygonal meshes, or piecewise smooth surface approximations, and (b) volumetric object descriptions like 3D Delaunay triangulations and implicit solid modeling [3,5, 16].

2 Problem Formulation

Accurate object reconstruction has to overcome many difficulties depending on the structure of the object, its surface properties, its position-orientation and lighting conditions. The procedure has to handle noisy data, occlusions, accessibility problems and finally provide accurate results. Nevertheless, in several computer vision problems fusing sensor information and algorithms has given more accurate results compared to single approaches. For example, several ambiguities and ill-posed situations existing in a two-camera system are almost eliminated in a multicamera stereo system. Similarly, object recognition from image sequences provides better results compared to a single image approach[15]. Moreover addressing motion and structure estimation simultaneously and working on stereo image sequences provides better results in both estimation problems[9]. The framework of our approach is based on the above principle.

Our goal is to automatically scan an object and derive a precise model of it. Therefore we efficiently combine techniques and algorithms used in several scientific fields. Stereo vision and photogrammetry are used in order to obtain accurate three-dimensional measurements of the working area. Methods of computational geometry based on Delaunay triangulations are used in order to generate a three dimensional model of the object. Pattern matching and segmentation techniques are used in order to evaluate the model and segment it. Path planning techniques are used in order to guide the tool around the object in order to get all the necessary measurements to refine the model.

3 Problem Solution

The outline of the proposed methodology is the following. Initially we calibrate the system. Then we use a robot arm to scan the object with a laser beam and obtain range data. Next we examine the surface of the object and identify areas where the object has not been sampled properly. Based on these results a new scanning strategy is derived. The new scanning operation reveals new information on non-accurately sampled areas. The resulting solid models are merged providing a more accurate one. The above process is repeated until a desired quality level is reached. The following sections present the basic steps of our approach.

3.1 System Calibration

Depth perception can be derived by processing multiple images taken from different viewports. In order to extract depth it is necessary to know the relative position of the cameras, model the camera system and compensate any errors and distortions. The above problem is solved using camera calibration techniques [1].

The algorithm used estimates (a) the extrinsic parameters of each camera, describing its position and orientation, and (b) the intrinsic parameters describing its internal viewing characteristics. Calibration is based on viewing points on a calibration cube whose 3D coordinates are known with great accuracy. Although, it is possible to calibrate the camera using self-calibration methods, a special calibration pattern is necessary in order to achieve high accuracy. The perspective transformation matrix coefficients are computed solving the system of equations formulated by the image coordinates $p_i = (x_{1i}, x_{2i})$ of world points $P_i = (X_{1i}, X_{2i}, X_{3i})$ [7,8]. Since, the real image is quite different from the image derived by the pin hole camera model, we add non-linear terms to the perspective projection of a point in order to model distortions. The most important effect and is caused (a) by the radial distortion, which is modeled as: $x_{1d} = x_1 + ax_1(x_1^2 + x_2^2)$, $x_{2d} = x_2 + ax_2(x_1^2 + x_2^2)$ and (b) the decentering of the lens, which is modeled as:

$$x_{1d} = x_1 + ax_1(3x_1^2 + x_2^2) + 2bx_1x_2 \;,$$
$$x_{2d} = x_2 + ax_2(3x_1^2 + x_2^2) + 2ax_1x_2$$

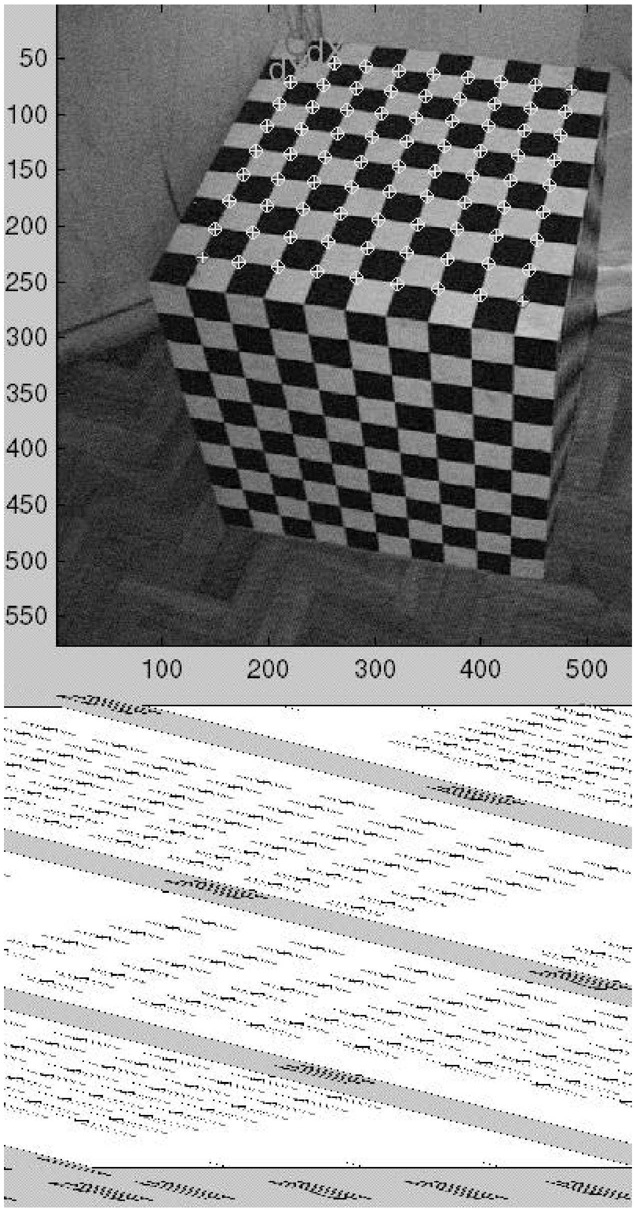

Fig.2 Distortion Compensation

Based on the calibration setup accurate estimation was achieved by modeling and compensating radial and decentering of the lens distortions using an iterative procedure[14].

3.2 Three Dimensional Range Data

The next step is to take 3D measurements of the object. Stereoscopic analysis has sometimes given poor results depending on the light conditions and workpiece texture. We were able to overcome these problems by scanning the object with a laser beam [13].

In the first scanning procedure, the scene is uniformly sampled using vertical laser beams. The robot arm, moves a laser beam along a regular grid in order to trace the object. Surface points on the object are detected with accuracy by taking into account the luminance gradient of the laser trace. The accuracy is much better than the size of the laser trace and the camera pixel.

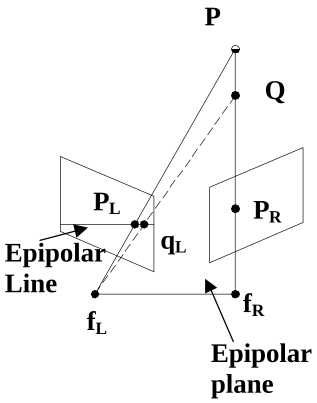

Fig.1 Epipolar Geometry

In order to speed up the matching operation we used rectification [12]. Rectification is a transformation of the images based on epipolar geometry, which enabled us to search for a match on the light stripe only in one dimension a not in two as it is generally required. This procedure transforms the image coordinates of each image plane so that pairs of conjugate epipolar lines become collinear and parallel to one of the image axis. Given the projection transformation matrix P_n, which rectifies the view, the linear transformation that maps the a point m_o of the actual projection view P_o into the point m_n of the rectified focal plane is of the form:

$$m_n = P_n P_o^{-1} m_o$$

Next, using inverse perspective on the calibrated cameras the image coordinates of the center of the laser trace are projected in the 3D world along the line of sight. The lines of sight from the different cameras, which see the trace, are intersected and the intersection coordinates gives the coordinates of an object point. However, due to noise and physical constraints, lines do not intersect in the geometrical sense so we compute the point, which is at minimum distance from the lines of sight.

As an alternative approach we used a laser scanner attached to a robotic arm. The sensor computes the coordinates of the traces based on the principle of triangulation. Then the coordinates given by the sensor are transformed to world coordinates by combining the transformation matrices between all the coordinate systems of the links of the system. Time stamps are used for synchronizing the measurements of the sensor with the robot position.

3.3 Model Reconstruction

The sampled set of points of the surface of the object forms the initial range data, which will be used to approximate the surface of the object. Reconstruction algorithms use triangular mesh representations or piecewise smooth functions for approximating the surface of the object [5]. Research is also focused in directly reconstructing the volume instead of the surface of the object [16]. Regarding

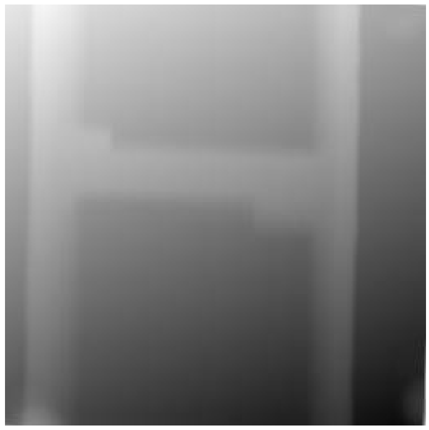

Fig.3 Range Data

function representation the problem of object reconstruction has two major approaches. The first deals with functions of the form z=f(x,y) representing range data and the second works with cloud points sets and implicit functions of the form f(x,y,z)=c [11]. In our approach each view consists of a set of consecutive measurements that form a function of two variables.

The range data go through the following processing steps: (a) their coordinates are transformed to the robot tool coordinate system, (b) Delaunay triangulation is performed on the (x, y) plane in order to assure mesh connectivity, (c) filtering is performed to remove outliers and smooth the surface, (d) the surface is processed and segmented into regions according to the quality measure of the range data. The results of this procedure are (a) regions of the surface that require additional scanning and (b) a solid model of the object as seen from this view point.

The first step is required in order to process the range data as a function of the form z=f(x,y). The range data of each view are taken using a constant orientation in the tool. Only the position of the tool changes, so we tranform each point to the coordinate system that the tool had in the beginning of the scanning procedure.

The next step is to perform Delaunay triangulation on the XY plane. Delaunay triangulation connects each point of set points to its natural neighbors and is the dual of the Voromoi diagram. The circle circumscribed about a Delaunay triangle has its center at the vertex of a Voronoi polygon. Using Delaunay triangulation we achieve mesh connectivity and avoid mesh intersection. Further more we are able to interpolate the surface using higher order functions to accurately describe curved areas.

Range data requires filtering. Filtering has to achieve the following: remove outliers, smooth the surface and preserve information. A neural network is used in order to filter out noise. The parametric function represented by a network with one hidden layer has the form $\hat{z} = s(ax + by + c)$ where $s(\ )$ is a non-linear activation function. Networks with more than one hidden layer can model complex surfaces using a variety of basis functions. The training algorithm is based on Back Propagation algorithm modified so that it can handle errors in the training data set [6]. Instead of minimizing the sum of squared errors, it minimizes a new function, which is adjusted with the progressively refined knowledge of noise in the data. The algorithm minimizes $\sum_{p=1}^{P} f_t(z_p - \hat{z}_p)$ where the shape of f_t resembles the tanh estimator whose derivative is given as :

$$y_t(r) = \begin{cases} r & |r| \le a_t \\ a \tanh(b\,(b_t - |r|))\,\mathrm{sgn}(r) & a_t < |r| \le b_t \\ 0 & |r| > b_t \end{cases}$$

When minor noise exists in the training data, the algorithm is similar to back propagation algorithm. However, when noise level increases the influence of noise on the learning process is significantly reduced, because the amount of weight adjustment at the output layer during leaning is proportional to $y_t(r)$ instead of the residual r. We need then to specify q, which is the percentage of bad data to be tolerated by the algorithm. Using the bootstrap method the confidence interval of the residuals is computed that separates the good from the bad data.

Then we have to identify areas where the scanning procedure should be repeated. For this reason we define a measure of quality of the sampling procedure according to which the surface is segmented. In particular, because the object is three dimensional and not an explicit function z=f(x,y) different criteria can be used for the range data quality. For example on areas that are occluded, the scanning process provides measurements, which form a region created by the extrusion of a curve. This region has data points only on its boundary. Another example is that areas of the object, which are aligned with the direction of the beam, provide very few measurements.

From the above we concluded on two criteria, regarding the quality of the range data. The

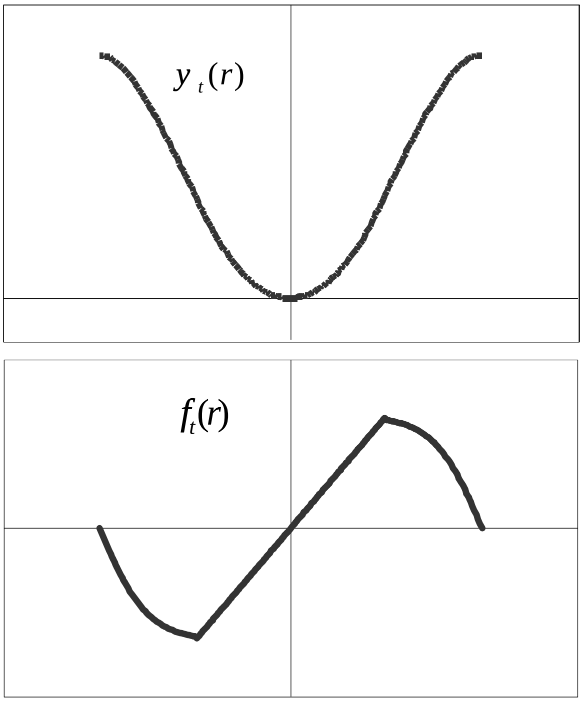

Fig. 4 Error Contribution Function

first one is the sampling density over the surface of the object. This is measured locally as the inverse of the area of the surface of the mesh triangles. The normal to the mesh triangles approximates the surface normal and is used in defining similar regions using region growing techniques. The second criterion is that areas where there is a significant change in the normal of the surface imply high curvature and therefore require more samples. The output of the segmentation process is areas where additional scanning is required and a principal scanning direction, which has the tool to follow for each region.

3.4 Path planning

The scanning procedure is handled by the robot arm, which must be programmed to move the laser beam along the area encircling the object. In the first stage the robot arm samples the object of the scene following a regular grid. The density of the grid is twice the sampling density we finally require on the surface of the object. The sampling frequency along the object surface depends on the orientation of the surface relative to the laser beam. Therefore, this procedure is not complete, because several areas of the object can be occluded or seen from a small angle. Therefore we need to derive additional scanning paths for the not accurate regions derived during the segmentation of the model. The normal to the surface of these regions is considered the appropriate orientation for the scanning tool. Based on these vectors we determine the new scanning orientation and positions. Then we examine whether these scanning orientations intersect with the object model already created and in this way assure that the region will be visible. Additionally we examine for collisions between the object and the tool of the robot arm [4]. In case the simulation system sees a collision a heuristic search is performed modifying the orientation of the tool. In case it is not possible to completely cover the area we accept the one with the best foreseen quality Q, which equals the sum of the

$$Q = \sum \underline{n} \bullet \underline{s}$$

dot products of the scanning orientation $\underline{s}$ with the normal of the surface $\underline{n}$ over the covered area.

Since the new scanning procedure will reveal new information on the area of the not accurately sampled region we do not need to cover the whole of the area in one step. The missing information will be covered in the next iteration if necessary.

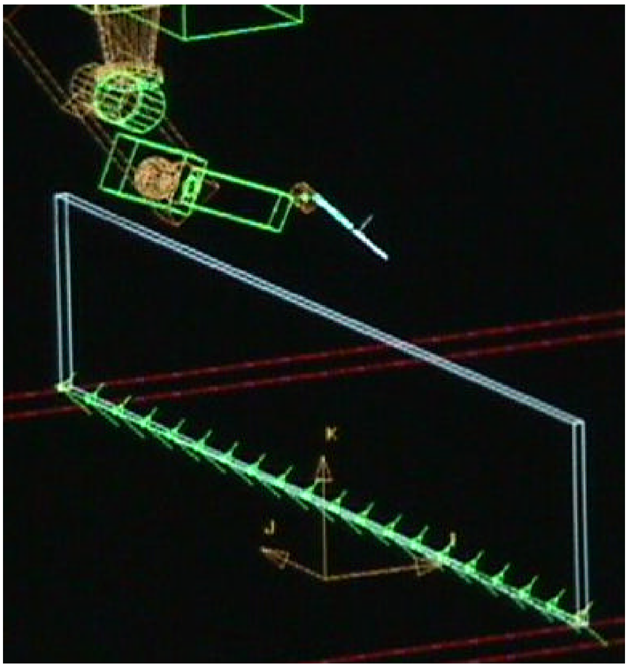

Fig.5 Path Simulation

3.5 Merging of three dimensional views

For each view of the scene a different solid model is created. The principle of the merging operation is to intersect the three dimensional models from each view and create the final model which is the common part of all views [13]. However due to errors in the mechanical joints of the robot arm, each model created by the laser scanner requires a small alignment. For this reason the scanner scans also three reference marks aside the object in order to exactly define its orientation and position and compensate for the above errors of misalignment.

Then the models from the different scanning orientations are aligned and intersected. Views are intersected inside the common bounding volume of each view. This happens because in some steps we do not scan the whole area encircling the object but only the part that is not clear. In general after the intersection of the solid models we perform smoothing operations, as described above, on the surface of the object in order to derive the final model.

3.6 Experiments

Experiments have been conducted on small and large objects. Large objects were scanned using only the fixed camera system, because the set up of the camera system can be adjusted according to the size of the objects and can be applied to small and large objects. The images were acquired from four CCD cameras. The range of the laser scanner is around 50 centimeters and small objects were scanned with more detail using the laser scanner. In order to scan objects at different angles, the tool of the robotic system, moved the laser in all required directions. The final accuracy of the laser scanner is considerably affected by the accuracy of the robot arm. These errors did not exist in the camera system, because is does not require measurements on moving parts. However, it requires accurate calibration. The calibration device is a cube with a chessboard pattern on each side. The image features were detected with 1/10th of a pixel accuracy. The calibration cube was placed around one meter away from each camera in order to model distortion parameters accurately along the field of view. When scanning large objects the calibration cube was used to calibrate the intrinsic parameters of the camera first. The extrinsic parameters were calibrated using eight reference marks on the area of the scanning operation.

Experimental results show that we overcame problems found in stereoscopy and laser scanning. The algorithms performed well under poor lightning conditions on matte objects and on shiny ones too, erroneous points were filtered out and the object was reconstructed accurately. The accuracy of the system was verified by taking external point measurements. The measurements were accurate to 1mm in the range of a 500 mm object and a 2-3mm for large objects of a few meters.

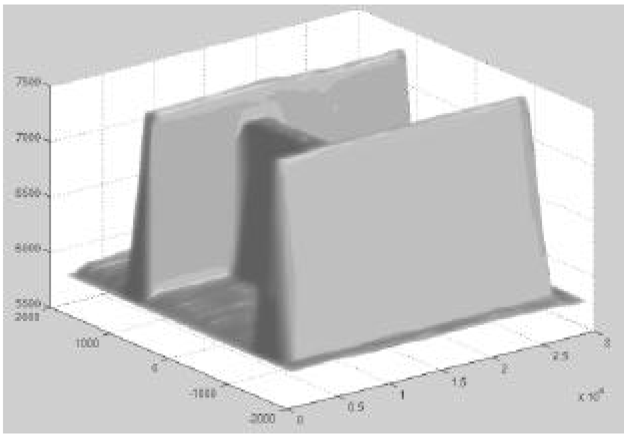

Fig.6 Object Model

4 Conclusions

Several fields related to robotic vision are integrated efficiently and form a complete system, which provides accurate results. This is achieved by fusing sensor data and by super sampling. Furthermore defects and distortions are compensated by software calibration and noise is filtered out adequately by the neural network. The algorithm adapts well to the structure of the objects of the scene and guides the robot arm efficiently. Although large bandwidth is required to handle and process the images the operation can be speed up significantly using appropriate hardware. Finally, we consider the procedure reliable enough so that it can be fully automated in the future.

References:

[1] R. Y. Tsai, "A versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses", *IEEE Journal of Robotics and Automation*, Vol. RA-3, No. 4, 1987, pp. 323-344.

[2] F. Ferrie, , J. Lagarde, and P. Whaite, "Recovery of volumetric object descriptions from laser rangefinder images", *In Proc. of Computer Vision -- ECCV '90*, 1990, pp. 387-396.

[3] J. Doi, and K. Koeda, "Three-dimensional reconstruction of solid models from multi-directional images and applications to industrial measuration", *In Proc. of SPIE, Close-Range Photogrammetry Meets Machine Vision*, 1990, pp. 564--571.

[4] L.P.Gewali, S. Ntafos and I.G.Tollis, 1990. "Path Planning in the Presence of Vertical Obstacles". *IEEE, Robotics and Automation*, Vol 6, No 3, pp.331-341.

[5] D.Terzopoulos and M Vasilescu, " Sampling and Reconstruction with adaptive Meshes", *In CVPR'91 IEEE Computer Vision and Pattern Recognition 1991*, pp 70- 75.

[6] D. Chen, R. Jain and B. Schunk, "Surface Reconstruction Using Neural Networks", *In CVPR'92 IEEE Computer Vision and Pattern Recognition 1992*, pp 815-817.

[7] O. Faugeras, "Three-Dimensional Computer Vision: A Geometric Viewpoint", *The MIT Press*, Cambridge, 1993.

[8] M. R. Shortis, T. A. Clarke, and T. Short, "A comparison of some techniques for the subpixel location of discrete target images", *In SPIE Vol. 2350 Videometrics III* ,1994, pp. 239-250.

[9] A. Murat Tekalp, "Digital Video Processing", *Prentice Hall*, 1995.

[10] D. Geiger, B. Ladendorf, and A. Yuille, "Occlusions and binocular stereo", *International Journal of Computer Vision*, April 1995, pp.211-226,.

[11] C.Lim, G.Turkiyyah, M. Ganter, D. Storti, "Implicit Reconstruction of Solids from Cloud Point Sets", ACM Proc. of the 3rd symposium on Solid modeling and applications,pp.393, 1995

[12] D. V. Papadimitriou and T. J. Dennis. "Epipolar line estimation and rectification for stereo images pairs", *IEEE Transactions on Image Processing*, April 1996, pp.672-676.

[13] Yang, Z.M., Wang, Y.F., "Error Analysis of 3D Shape Construction from Structured Lighting", *In PR(29), No. 2*, 1996, pp. 189-206.

[14] J. Heikkila, O. Silven, "Calibration procedure for short focal length off-the-shelf CCD cameras", Proc. of The 13th International Conference on Pattern Recognition. (1996) Vienna, Austria. pp. 166-170.

[15] A. Delopoulos and Y. Xirouhakis, "Robust Estimation of Motion and Shape based on Orthographic Projections of Rigid Objects", *Tenth IMDSP Workshop 98*, July 1998.

[16] L E.Mucke, "A Robust Implementation for Three-Dimensional Delaunay Triangulations", International Journal of Computational Geometry & Applications",vol.8, no2, pp.255-276,1998