Proceedings of the 5th World Congress on Intelligent Control
and Automation, June 15-19, 2004, Hangzhou, P.R. China

第五届全球智能控制与自动化大会
2004 年 6 月 15 - 19 日, 中国杭州

3D Object Reconstruction from Multiple Controlled Viewpoints

S. Y. Chen and Y. F. Li*
Department of Manufacturing Engineering and Engineering Management,
City University of Hong Kong, Kowloon, Hong Kong
(*Author for correspondence, Email: meyfli@cityu.edu.hk)

Abstract— **To reconstruct the complete model of a 3D target by robot vision, multiple viewpoints often need to be planned to obtain sufficient 3D surfaces for integration. This paper presents a novel method of viewpoint planning for incrementally building the models of unknown objects or environments by an active vision system. This method is based on the concept of trend surface, which is the regional feature of a surface for describing the global tendency of change. A new mathematical model is proposed for predicting the unknown area of the object surface. A criterion is defined to determine the exploration direction and a unique surface model is established by analyzing the surface curvature. Then an algorithm is developed for determining the sensor pose which satisfies the many placement constraints such as resolution, focus, and field of view. Experiments are carried out to demonstrate the proposed method.**

Index Terms - Sensor planning; Viewpoint; Object reconstruction; 3D modeling; Trend surface; Surface prediction

I. INTRODUCTION

For 3D object reconstruction, a strategy with controlled viewpoints determines each subsequent vantage point and offers the benefit of reducing or eliminating the manual work. The research on actively moving a vision sensor for modeling objects has been active for more than ten years, since 1988 when an early attempt was made in this aspect by Cowan and Kovesi [1]. Among the previous approaches to the modeling problem, "occlusion" and uncertainty have been strongly associated with viewpoint planning for a period. Kutulakos *et al.* [2] utilized the changes in the boundary between sensed surface and occlusion with respect to sensor position to recover shape. A similar histogram-based technique was used by Maver and Bajcsy [3] to find the viewing vector that would illuminate the most edge features derived from occluded regions. Whaite and Ferrie [4] used a sensor model to evaluate the efficacy of the imaging process over a set of discrete orientations by ray-casting: the sensor orientation that would hypothetically best improve the model is selected for the next view. The work by Pito [5] removed the need to ray-cast from every possible sensor location by determining a subset of positions that would improve the current model.

On the next best view (NBV) problem for incremental object modeling, two issues were addressed [6] to determine the next-best-viewpoint: 1) a uniform tessellation of the spherical space and its mapping onto the 2D array; 2) incremental updating computations from evaluating viewpoints as the NBV. Pito *et al.* [5] presented a solution for planning a depth camera in the process of digitizing unknown parts. Arbel *et al.* [8] showed how entropy maps can be used to guide an active observer along an optimal trajectory and how a gaze-planning strategy can be formulated by using entropy minimization as a basis for choosing a next best view. A NBV will reveal the greatest quantity of previously unknown scene information [9]. These is the uncertainty driven approach to the NBV problem where maximizing the information gain for the next view is set as the goal in the view point planning [10]. Reed *et al.* [11] determined the visibility volume, which is the volume of space within which a sensor has an unobstructed view of a particular target. Assuming polyhedral and cylinder objects, a technique was proposed to solve the NBV problem via a depth-first search algorithm [12].

To this end, two distinct methods have been widely used: the weighted function method and tessellated space approach. The former [11], [13], [14] employs a function that combines several components representing the placement constraints. This method is usually used in model-based planning tasks [15], [16]. The latter method is mainly for object modeling tasks. The object surface is partitioned as void surface, seen surface, unseen surface, and uncertain surface. The working space is also partitioned into void volume and viewing volume. An algorithm is then developed to plan a sequence of viewpoints so that the whole object can be sampled. This method is effective for dealing with some small and simple objects, but it is difficult to model a large and complex object, e.g. an environment with many concave areas, as it cannot solve the problems of occlusion constraint.

Therefore, previous efforts were often made on finding the best next views by volumetric analysis or occlusion as a guide. The viewpoint planning method developed in this paper is an effective strategy for generating a sequence of viewing poses for optimal completion of a task. It involves decision of exploration direction and determination of the next view. The trend surface is proposed as the cue to predict the unknown portion of an object or environment and the next best viewpoint is determined by the expected surface. The viewpoint determined in such a way is predictably best.

This work was fully supported by a grant from the Research Grants Council of Hong Kong [Project No. CityU1049/00E] and City University of Hong Kong (Project No. 7001431).

II. TREND ANALYSIS AND SURFACE PREDICTION

A. Trend Analysis

Surface trend describes the global shape of a surface and trend surface analysis is a global method for processing spatial data. Mathematically, a mapped surface can be separated into two components - that of the trend and the residuals from the trend. The trend is the regional feature of a surface, and the residuals are the local fluctuations of high frequent features.

Trend surface analysis is often used for fitting and interpolating regression surfaces in three dimensions as smoothed representation of area data. It is assumed that the spatial distribution of a particular phenomenon can be represented by some form of continuous surface, usually a defined geometric function. The observed spatial pattern can be regarded as the sum of such a surface and a "random", or local, term. The surface is a function of two orthogonal coordinate axes which can be represented by

$$z = f(x, y) + e, \tag{1}$$

in which the variable z at the point (x, y) is a function of the coordinate axes, plus the error term e. This expression is the generalized form of the General Linear Model (GLM), which is the basis of most trend methods.

The function $f(x, y)$ is usually expanded or approximated by various terms to generate polynomial equations. To develop complex, smoothed equations for geophysical data by expanding the summation term of the General Linear Model, the relationship between standard multivariate regression analyses and trend methods can be defined. This expansion was performed by incorporating power terms and cross-products of the x and y coordinates. For an n-order three-dimensional surface, the form of the power series is given by

$$f(u,v) = \sum_{i=0}^{n} \sum_{j=0}^{i} b_{ij} u^j v^{i-j} , \tag{2}$$

where u and v are the coordinates on an arbitrary orthogonal reference system, b_{ij} is the constant coefficient of the surface (b_{00} is the surface base).

The trend part is very helpful for predicting the unseen part of an object or environment and is thus used for determining the next viewpoint in this paper. The residuals (local features) do not affect viewpoint planning much, but they should be filtered out during the image processing.

Let a single surface $\boldsymbol{M}$ be split into two parts, $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$,

$$\boldsymbol{M} = \boldsymbol{M}_1 \cup \boldsymbol{M}_2. \tag{3}$$

If the surface $\boldsymbol{M}$ changes smoothly, both the trends of $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ should be approximately equal to the trend of $\boldsymbol{M}$, i.e.

$$\mathit{Trend}(\mathbf{M}) \approx \mathit{Trend}(\mathbf{M}_1) \approx \mathit{Trend}(\mathbf{M}_2). \tag{4}$$

Suppose the vision agent has already captured a part of the surface, say $\boldsymbol{M}_1$, but $\boldsymbol{M}_2$ remains unknown. Then by computing the surface trend of $\boldsymbol{M}_1$, the surface shape of $\boldsymbol{M}_2$ can be predicted. In this paper, we will not use (1) or (2) directly as the trend model for surface prediction, since it relies on interpreting regression of the known area. Instead, we will develop a new mathematical model for describing the surface trend, thus emphasizing on the prediction of the unknown area.

B. Exploration Direction

With the partially known model, the pose of a next view is decided according to the known information. Here two steps are used to make this decision. The first step is to determine the exploration direction and the second is to determine the sensor pose in the space.

Except for surface edges and object boundaries, since the curvature of trend surface changes smoothly, the unknown part of object surface can be predicted by analyzing the curvature tendency of the known surface. Assume that the known part is located in the center of the scene and its surrounding areas are unknown. Since only one direction, called exploration direction, can be chosen for planning the next viewpoint, it may be determined as the area where the surface is most smoothed or with the lowest surface order. The reason is that the trend surface can predict the unknown area accurately where the surface has a low order. Fig. 1 illustrates the selection criterion of the exploration direction.

The surface order is determined according to (2) with the same fitting error. To avoid computation of surface fitting, alternatively we may just approximately compute the integral value of the curvatures in a small area, i.e.

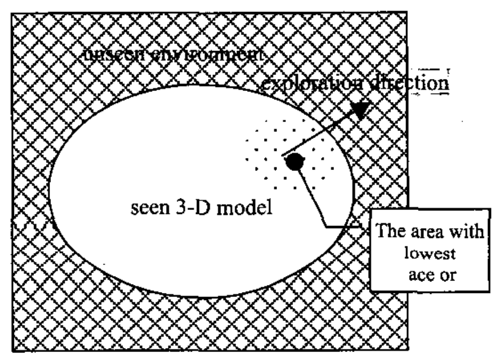

Fig. 1 Exploration of object surface

$$n_{\text{order}}(u, v) \approx \iint_{x,y \in S(u,v)} k_{\min}(x,y)\,dx\,dy, \tag{5}$$

where $S(u, v)$ is the neighborhood area of point (u, v) and $k_{\min}(x, y)$ is the curvature at point (x, y) along a specific direction.

It is only necessary to compute the surface orders in the areas near to the boundary of the known surface. The surface order in the center area of the known model does not affect the exploration direction. After the minimum surface order is obtained, i.e. $n_{\min} = \min\{n_{\text{order}}(u, v)\}$, the exploration direction is decided to be along the direction outside the unknown area.

C. Surface Prediction

There are different curvatures for a surface point along different directions, although the principal curvature and Gaussian curvature are the most frequently used ones. Without loss of generality, we may describe the mathematical formula

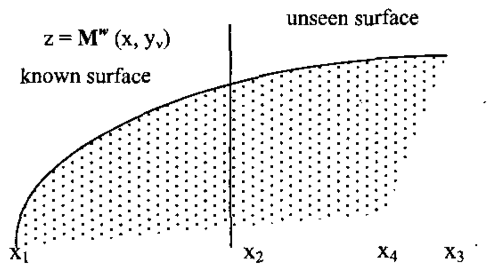

Fig. 2 A curve on the sectional plane

along the horizontal direction. Using a vertical sectional plane which is parallel to x-axis, at $y=y_v$, to cut through the 3-D surface, we get a surface curve (see Fig. 2),

$$z_v = f_{yv}(x) \cdot \tag{6}$$

The curvature of this curve is

$$\boldsymbol{k} = \boldsymbol{k}_{yv}(x) = z_v'' / [1+(z_v')^2]^{3/2}. \tag{7}$$

Let $\boldsymbol{X}_k = [x_1, x_2]$ be the domain of the known part of the surface curve. To predict the unseen surface, we use a linear regression of x on $\boldsymbol{k}$ and get a fitted curve $\boldsymbol{C}_v$ for approximating the curvature tendency on the curve $\mathbf{z}_v$. Hence,

$$c_v(x) = a\,x + b,\ x \in [x_1, x_3], \tag{8}$$

where $[x_1, x_3]$ is the whole domain including the known area and unseen area, i.e. $[x_1, x_3] = [x_1, x_2] \cup [x_2, x_3]$. The two parameters a and b are fitted by the known part of surface curve, i.e.,

$$a = \frac{6(x_2 + x_1)\int_{x1}^{x2} k(x, y_v)dx - 12\int_{x1}^{x2} xk(x, y_v)dx}{(x_1 - x_2)^3}, \tag{9}$$

and

$$b = [\int_{x1}^{x2} k(x, y_v)dx - \frac{1}{2}(x_2^2 - x_1^2)a]/(x_2 - x_1) \cdot \tag{10}$$

Given a threshold k_{max}, the curvature in the unseen area is expected to be:

$$c(x, y_v) = \begin{cases} ax + b, & c(x, y_v) < k_{max} \\ k_{max}, & c(x, y_v) \geq k_{max} \end{cases},\ x \in [x_2, x_4], \tag{11}$$

where $x_2 < x_4 < x_3$ is a domain for satisfying the constraint that the object surface will be in the Field-of-View (FOV) of the sensor. Then the surface curve in the unseen part of the object will be a solution of the following equation:

$$\|z''\| / [1+(z')^2]^{3/2} - c(x, y_v) = 0. \tag{12}$$

The solution of this differential equation is:

$$z = \pm\int \frac{(ax^2 + 2bx + 2C_1)}{\sqrt{4 - (ax^2 + 2bx + 2C_1)^2}}dx + C_2$$

$$\text{when } x < (k_{max} - b)/a, \tag{13}$$

$$\text{or } z = \sqrt{k_{max}^{\ 2} - (x - C_3)^2} + C_4 \text{ when } x \geq (k_{max} - b)/a, \tag{14}$$

where C_i is differential constants which can be determined according to boundary conditions, such as $z(x_2) = z_2$ and $z'(x_2) = z_2'$. The sign "+" or "-" can also be determined by the known part of the surface curve (convex or concave). Since the predicted curve is based on the analysis of the tendency of known area, it is called the *trend curve.*

III. Parameters of Controlled Viewpoints

To determine the next viewpoint is to specify the sensor's placement parameters as well as to satisfy the placement constraints. The placement parameters include sensor's position (x, y, z) and sensor orientation (α, β, γ). The placement constraints usually include visibility, focus, field of view, viewing angle, resolution, overlap, occlusion, and some operational constraints such as kinematic reachability of the sensor pose and robot-environment collision. Let the resolution constraint be

$$r_p = 2\sqrt{(x_p - x_2)^2 + [z_p - z(x_2, y_v)]^2}\tan(\frac{\Omega}{2})/N < r_{max}, \tag{15}$$

where N is the pixel number on a scanning line of the digital image and Ω is the angle of view.

To satisfy the constraints of sensor placement on resolution and FOV, the parameter x_4 in (11) is determined by a searching algorithm:

Step 1. Obtain the numerical solution of (13) or (14) and stores the pairs (x_i, z_i) in an array. $x_i = x_2 + (i-1)*w_x$, where w_x is the pixel length in x-direction;
Step 2. For each pair, test the satisfaction of the resolution constraint (15) and eliminate the points that fail;
Step 3. Let $x_4 = \max\{x_i\}$.

Algorithm 1 Determination of boundary position

The mid-point of such a trend curve is:

$$\mathbf{Q}_{m,yv} = (x_v, y_v, z_{m,v}),\ z_{m,v} = f(x_v, y_v),\ x_v = \frac{x_2 + x_4}{2}. \tag{16}$$

By moving the V-V plane to different positions, in the domain of $-Y_{fov} < y_v < +Y_{fov}$, we get a series of surface curves. Connecting the mid-point of each such curve, we get a new curve:

$$\boldsymbol{L}_{i+1} = \boldsymbol{L}(\mathbf{Q}_{m,yv}, y_v),\ -Y_{fov} < y_v < +Y_{fov}. \tag{17}$$

Calculating its centroid, the position of the reference point (i.e. the new scene center) is obtained:

$$\mathbf{O}_{i+1}:\ x^o{}_{i+1} = \frac{\int_{L_{i+1}} xLdl}{\int_{L_{i+1}} Ldl},\quad y^o{}_{i+1} = \frac{\int_{L_{i+1}} yLdl}{\int_{L_{i+1}} Ldl},\quad \text{and}$$

$$z^o{}_{i+1} = \frac{\int_{L_{i+1}} zLdl}{\int_{L_{i+1}} Ldl}. \tag{18}$$

where (i+1) denotes the next view pose.

Now the position of the eye point and the viewing direction can be determined. To achieve the maximum viewing angle (i.e. the angle between the viewing direction and the surface tangent) for minimizing the reconstruction uncertainty, the viewing direction is the inverse of the average normal on the predicted surface, that is,

$$\boldsymbol{I}_{i+1} = -\frac{\iint N(x,y)dxdy}{\iint dxdy} = -\frac{\iint (\mu\mathbf{i}, \nu\mathbf{j}, \kappa\mathbf{k})(x,y)dxdy}{\iint dxdy}$$

$$= (-\mu_{i+1}\boldsymbol{i}, -\nu_{i+1}\boldsymbol{j}, -\kappa_{i+1}\boldsymbol{k}), \qquad (19)$$

where $\mu(x,y) = \frac{\partial f}{\partial x}$, $\nu(x,y) = \frac{\partial f}{\partial y}$, $\kappa(x,y) = -1$, and $N(x, y)$ is the surface normal on point (x, y, z).

The sensor's position $\boldsymbol{P}_{i+1}(x, y, z)$ for the next viewpoint is planned as a solution of the following set of equations:

$$\begin{cases} \dfrac{x^o_{i+1} - x^p_{i+1}}{\| O_{i+1} - P_{i+1} \|_2} = \dfrac{-\mu_{i+1}}{\| I_{i+1} \|_2} \\ \dfrac{y^o_{i+1} - y^p_{i+1}}{\| O_{i+1} - P_{i+1} \|_2} = \dfrac{-\nu_{i+1}}{\| I_{i+1} \|_2} \\ \| O_{i+1} - P_{i+1} \|_2 - \dfrac{r_{\max} N}{2\tan\frac{\Omega}{2}} = c_{cmp}, \quad c_{cmp} \ge 0 \end{cases} \qquad (20)$$

where c_{cmp} is a positive constant for compensating the depth value range, Ω is the sensor's Angle-of-View, $\| I_{i+1} \|_2 = \sqrt{\mu_{i+1}^2 + \nu_{i+1}^2 + \kappa_{i+1}^2}$, and $\|\boldsymbol{O}_{i+1} - \boldsymbol{P}_{i+1}\|_2$ is the distance between point $\boldsymbol{O}_{i+1}$ and $\boldsymbol{P}_{i+1}$.

Finally, the placement parameters of the vision sensor is described as a vector:

$$\boldsymbol{P}_{i+1} = (x^p_{i+1}, y^p_{i+1}, z^p_{i+1}, \mu_{i+1}, \nu_{i+1}, \kappa_{i+1}). \qquad (21)$$

This placement vector is based on the local coordinate system. It needs to be converted to the world coordinate system by multiplying a coefficient matrix.

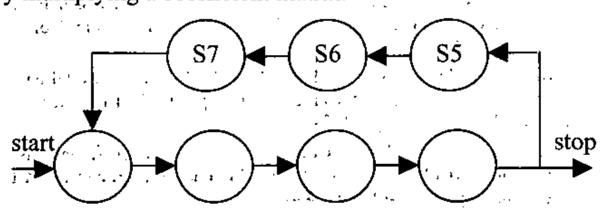

Fig. 3 The repetitive modeling process

Finally the iterative modeling process is illustrated in Fig. 3, in which the symbols Si (i=1, 2, ..., 7) represent the states of:

S1: Acquisition of a view
S2: Reconstruction of the 3-D local model
S3: Registration and fusion with the global model
S4: Model analysis and checking complete conditions
S5: Ranking for selection of exploration directions
S6: Computing trend surface and determining next view
S7: Moving the robot to the new viewpoint.

IV. EXPERIMENTS

A. Modification for Digital Image Processing

Several experiments were carried in our laboratory for construction of object models. The range data are obtained by a structured light system [7] set up in this research, which mainly consists of a DLP projector and a CCD camera. The projector is connected to a computer and is controlled to generate some gray-encoded stripe-light patterns for 3D reconstruction. The camera has a 1-inch sensor and a 25mm lens. The pixel number on a scan line is N=1024, and the resolution constraint is r_{max}=0.85mm. This structured light system was installed on a 6DOF robot to reach any arbitrary position in the workspace according to the setting parameters.

Fig. 4 is an object as the typical example here for illustration of model construction. The first view is assumed to be taken from the top view. To determine a next view for acquiring some unseen information of the object, we used trend surface method and developed a program to compute the expected surface curves. Then the trend is computed and the next viewpoint is determined.

The object model was incrementally built by four views. Each view acquired a new surface and it was integrated with the existing ones to form a partial model. The exploration direction and sensor placement were determined by the proposed method. Some intermediate computation results are shown Fig. 5 to Fig. 9. For example, after a partial model was integrated from the first two views (Fig. 5), an exploration direction was decided and several trend curves are computed to predict the unknown part of the surface as in Fig. 6 and Fig. 7. Then the next viewpoint was determined as in Fig. 8. 3D surface acquisition and integration were followed to increase the information of the target (Fig. 9). The Fig. 10 illustrates the final result of the 3D model and the planned four viewpoints in the space.

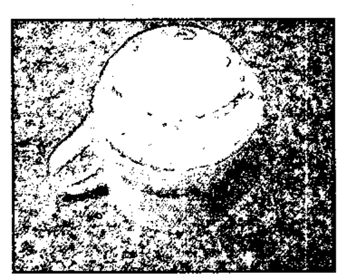
Fig. 4 The object to be reconstructed

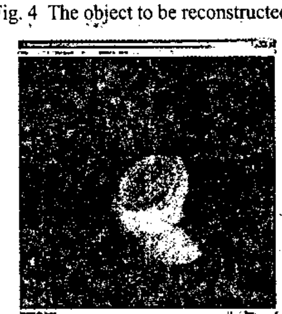
Fig. 5 The partial model integrated from the first two views

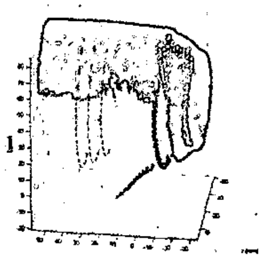

Fig. 6 Trend curves in the second step

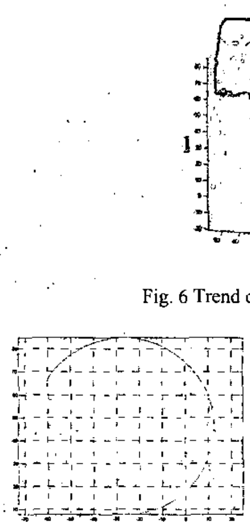

Fig. 7 One of the trend curve.

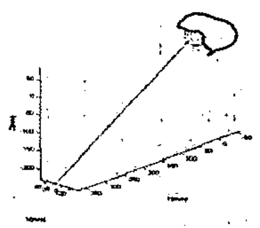

Fig. 8 The next viewpoint.

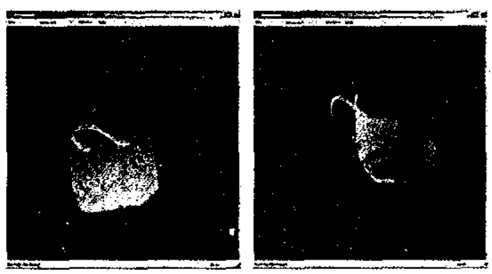

Fig. 9 The 3D surface was obtained from the third view and integrated with all known views.

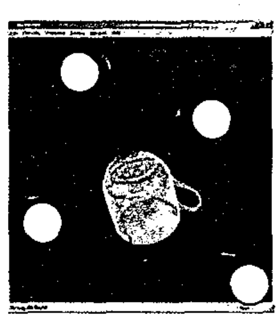

Fig. 10 The complete model and planned viewpoints

V. CONCLUSION

In this paper, we proposed a trend surface model which was used to predict the unseen part of objects or environments. In this way, the next viewpoint can be determined for on-line acquisition of the range data until the whole structure of the object or environment is reconstructed. The trend surfaces and trend curves were computed from the curvature tendency. While determining the next viewpoint, multiple sensor placement constraints were considered in this paper, such as resolution, field-of-view, and viewing angle. The analysis shows that the trend model can accurately predict the unknown surface of the object if the surface is composed of 1st-order and 2nd-order curves and surfaces. Experiments were carried out to demonstrate the method presented in this paper.

REFERENCES

[1] C. K. Cowan and P. D. Kovesi, "Automatic sensor placement from vision task requirements", IEEE Trans. on PAMI, Vol. 10, No. 3, May 1988, pp. 407-416.

[2] K.N. Kutulakos and C.R. Dyer, "Global surface reconstruction by purposive control of observer motion", IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, Seattle, 1994, pp. 331-338.

[3] J. Maver and R. Bajcsy, "Occlusions as a guide for planning the next view", IEEE Trans. on PAMI, Vol. 15, No. 5, May 1993, pp. 417 -433.

[4] P. Whaite and F.P. Ferrie, "Autonomous exploration: driven by uncertainty", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 19, No.3, March 1997, pp. 193-205.

[5] R. Pito, "A solution to the next best view problem for automated surface acquisition", IEEE Trans. on PAMI, Vol. 21 (10), 1999, pp. 1016-30.

[6] H. Zha, K. Morooka, T. Hasegawa, T. Nagata, "Active modeling of 3-D objects: planning on the next best pose (NBP) for acquiring range images", Int. Conf. on Recent Advances in 3-D Digital Imaging and Modeling, Ottawa, Ontario, Canada, May 1997, pp. 68 –75.

[7] Y. F. Li and S. Chen, Automatic Recalibration of an Active Structured Light Vision System, ***IEEE** Transactions on Robotics and Automation,* Vol. 19, No. 2, April 2003, pp. 259-268.

[8] T. Arbel and F.P. Ferrie, "Viewpoint selection by navigation through entropy maps", Seventh IEEE Int. Conf. on Computer Vision, vol. 1, Kerkyra, Greece, 1999, pp. 248-254.

[9] J.E. Banta and M.A. Abidi, "Autonomous placement of a range sensor for acquisition of optimal 3-D models", 22nd IEEE Int. Conf. on Industrial Electronics, Control, and Instrumentation, Taipei, Taiwan, Aug. 1996, vol. 3, pp. 1583-1588.

[10] Y. F. Li and Z. Liu, Uncertainty-Driven Viewpoint Planning for 3D Object Measurements, *Proc. **IEEE** International Conference on Robotics and Automation*, Taipei, Taiwan, pp.127-132, Sept. 2003.

[11] M. Reed, P. Allen, "Constraint-based sensor planning for scene modeling", IEEE Int. Symposium on Computational Intelligence in Robotics and Automation, Monterey, California, 1999, pp. 131-136.

[12] E. Marchand and F. Chaumette, "Active vision for complete scene reconstruction and exploration", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 21, no. 1, Jan. 1999, pp. 65-72.

[13] B. Triggs, C. Laugier, "Automatic camera placement for robot vision tasks", IEEE Int. Conf. on Robotics and Automation, Nagoya, 1995, pp. 1732-1737.

[14] K. A. Tarabanis, P. K. Allen, R. Y. Tsai, "A survey of sensor planning in computer vision", IEEE Trans. on Robotics and Automation, Vol. 11, No.1, 1995, pp. 86-104.

[15] E. Trucco, M. Umasuthan, A. M. Wallace, V. Roberto, "Model-based planning of optimal sensor placements for inspection", IEEE Trans. on RA, Vol. 13, No.2, 1997, pp. 182-194.

[16] S.Y. Chen and Y.F. Li, "Automatic Sensor Placement for Model-Based Robot Vision", IEEE Trans. on Systems, Man and Cybernetics, Part B, Vol. 34, No. 1, pp. 393-408, Feb. 2004.