Chapter 8
Information Entropy Based Planning

In this chapter, we present an approach with information entropy based sensor planning for reconstruction of freeform surfaces of 3D objects. To achieve the reconstruction, the object is first sliced into a series of cross-section curves, with each curve to be reconstructed by a closed B-spline curve. In the framework of Bayesian statistics, we propose an improved Bayesian information criterion (BIC) for determining the B-spline model complexity. Then, we analyze the uncertainty of the model using entropy as the measurement. Based on this analysis, we predict the information gain for each cross section curve for the next measurement. After predicting the information gain of each curve, we obtain the information change for all the B-spline models. This information gain is then mapped into the view space. The viewpoint that contains maximal information gain about the object is selected as the Next Best View. Experimental results show successful implementation of the proposed view planning method for digitization and reconstruction of freeform objects.

8.1 Overview

This chapter presents an information entropy based viewpoint planning method for the digitization and reconstruction of a 3D freeform object. The object is sliced into a set of cross section curves and a closed B-spline curve is used to reconstruct each cross section curve by fitting to partial data points. An information criterion is developed for selecting the B-spline model structure. Based on the selected B-spline model, we use information entropy as the uncertainty measure of the B-spline model and analyze the uncertainty of each B-spline cross section curve to predict the information gain for new measurements to be taken. As a result, we can obtain the prediction of the information gain about the object. The information gain is then mapped to the view space. The view that has the maximal information gain on the object is then selected as the Next Best View (NBV). The proposed information entropy based viewpoint planning procedure is illustrated in Fig. 8.1.

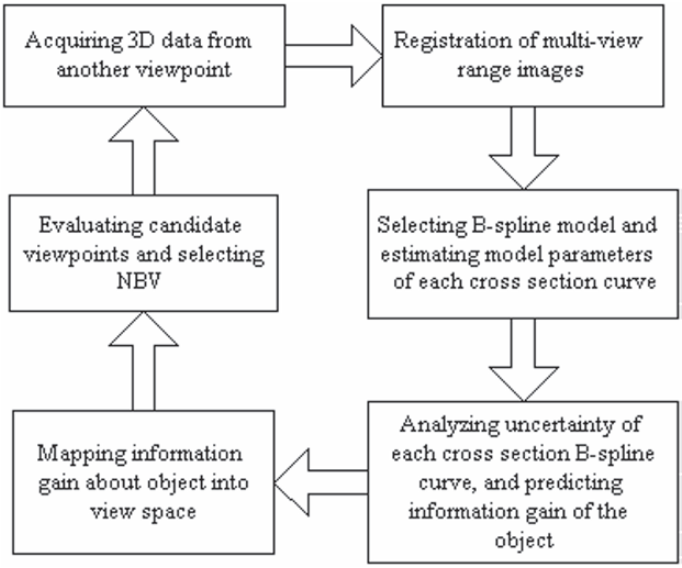


Fig. 8.1. Information entropy based viewpoint planning

This work is novel concerning the parameter estimation for the NBV problem. In contrast to Whaite's method (1997), here we analyze and reconstruct a B-spline model in the framework of Bayesian statistics. The B-spline model is more powerful in describing objects than a super-ellipsoid. In addition, we introduce the principle of model selection by which the proposed improved BIC criterion makes the B-spline model adaptable when newly acquired data are available. The rest of this chapter is organized as follows. In Sect. 8.2, we describe the reconstruction of cross section curves with closed B-splines and introduce the modified BIC for selecting a B-spline model structure. In Sect. 8.3, we define the information entropy of B-spline model to analyze its uncertainty and predict the information gain on an object. In Sect. 8.4, we evaluate the visibility of candidate viewpoints for selecting NBV. Finally, we present the experimental results in implementing the proposed method in Sect. 8.5 followed by conclusions in Sect. 8.6.

8.2 Model Representation

For object surface reconstructions, the 3D shape can be divided into a series of cross section curves each representing the local geometrical feature of the object. These cross section curves can be described by a set of parametric equations. For reconstruction purposes using parametric equations, the most common methods include spline functions (e.g. B-splines) (Fernand and Wang 1994), implicit polynomials and superquadrics (e.g. superellipsoids) (Whaite 1997). Compared

with implicit polynomials and superquadrics, B-splines have the following main advantages:

- *Smoothness and continuity*, which allows a curve to consist of a concatenation of curve segments, yet be treated as a single unit;
- *Built-in boundedness*, a property which is lacking in implicit or explicit polynomial representation whose zero set can shoot to infinity;
- *Parameterized representation*, which decouples the x, y coordinates to be treated separately.

8.2.1 Curve Approximation

Let a closed cubic B-spline curve consist of n+1 curve segments, defined by

$$\mathbf{p}(t) = \sum_{j=0}^{n+3} B_{j,4}(t) \cdot \mathbf{\Phi}_j, \tag{8.1}$$

where $\mathbf{p}(t) = [x(t), y(t)]$ is a point on the B-spline curve with location parameter t. $B_{j,4}(t)$ is the jth normalized cubic B-spline basis function defined over the following uniform knots vector

$$[u_{-3}, u_{-2}, u_{-1}, u_0, \ldots, \ldots, u_{n+4}] = [-3, -2, -1, 0, \ldots, n+4]. \tag{8.2}$$

The amplitude of $B_{j,4}(t)$ is in the range of (0.0, 1.0), and the support region of $B_{j,4}(t)$ is compact and nonzero for $t \in [u_j, u_{j+4}]$. $(\mathbf{\Phi}_j)_{j=0}^{n+3}$ are the cyclical control points which satisfy the following conditions

$$\mathbf{\Phi}_{n+1} = \mathbf{\Phi}_0, \ \mathbf{\Phi}_{n+2} = \mathbf{\Phi}_1, \ \mathbf{\Phi}_{n+3} = \mathbf{\Phi}_2 \tag{8.3}$$

For a set of m data points $\mathbf{r} = (\mathbf{r}_i)_{i=1}^m = ([x_i, y_i])_{i=1}^m$, let d^2 be the sum of the squared residual errors between the data points and their corresponding points on the B-spline curve, i.e.

$$d^2 = \sum_{i=1}^{m} \left\| \mathbf{r}_i - \mathbf{p}(t_i) \right\|^2 = \sum_{i=1}^{m} \left[\mathbf{r}_i - \sum_{j=0}^{n+3} \mathbf{B}_{j,4}(t_i) \cdot \mathbf{\Phi}_j \right]^2. \tag{8.4}$$

From the cyclical condition of control points in (8.3), there are only n+1 control points to be estimated. The LS estimation of the n+1 control points are obtained from the curve points by minimizing d^2 in (8.4) with respect to $\mathbf{\Phi} = [\mathbf{\Phi}_x^T, \mathbf{\Phi}_y^T]^T = [\Phi_{x0}, \ldots \Phi_{xn}, \Phi_{y0}, \ldots, \Phi_{yn}]^T$. By factorization of the B-spline, two separate solutions are obtained in the matrix as follows

$$\begin{cases} \mathbf{\Phi}_x = [\mathbf{B}^T\mathbf{B}]^{-1}\mathbf{B}^T\mathbf{x} \\ \mathbf{\Phi}_y = [\mathbf{B}^T\mathbf{B}]^{-1}\mathbf{B}^T y \end{cases} \tag{8.5}$$

where $\mathbf{x} = [x_1, \ldots, x_m]^T$, $\mathbf{y} = [y_1, \ldots, y_m]^T$,

$$\mathbf{B} = \begin{bmatrix} \overline{B}^1_{0,4} + \overline{B}^1_{n+1,4} & \overline{B}^1_{1,4} + \overline{B}^1_{n+2,4} & \overline{B}^1_{2,4} + \overline{B}^1_{n+3,4} & \cdots & \overline{B}^1_{n,4} \\ \overline{B}^2_{0,4} + \overline{B}^2_{n+1,4} & \overline{B}^2_{1,4} + \overline{B}^2_{n+2,4} & \overline{B}^2_{2,4} + \overline{B}^2_{n+3,4} & \cdots & \overline{B}^2_{n,4} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \overline{B}^m_{0,4} + \overline{B}^m_{n+1,4} & \overline{B}^m_{1,4} + \overline{B}^m_{n+2,4} & \overline{B}^m_{2,4} + \overline{B}^m_{n+3,4} & \cdots & \overline{B}^m_{n,4} \end{bmatrix},$$

and $\overline{B}^i_{j,4} = B_{j,4}(t_i)$.

Here, we adopt the chord length method, which is the most popular one, for the parameterization of the B-spline. The chord length L of a curve is calculated as follows

$$L = \sum_{i=2}^{m+1} \left\| \mathbf{r}_i - \mathbf{r}_{i-1} \right\| \tag{8.6}$$

where $\mathbf{r}_{m+1} = \mathbf{r}_1$ for a closed curve. The t_i associated with the point q_i is given as

$$t_i = t_{i-1} + \frac{\left\| \mathbf{r}_i - \mathbf{r}_{i-1} \right\|}{L} \cdot t_{\max} \tag{8.7}$$

where $t_1 = 0$ and $t_{max} = n + 1$.

8.2.2 Improved BIC Criterion

It is known that for a given set of measurement data, there exists a model of optimal complexity corresponding to the smallest prediction (generalization) error for further data. The complexity of a B-spline model of a surface is related to its control point (parameter) number (Fernand and Wang 1994). If the B-spline model is too complicated, the approximated B-spline surface tends to over-fit noisy measurement data. If the model is too simple, then it is not capable of fitting the measurement data, making the approximation results under-fitted. The problem of finding an appropriate model, referred to as model selection, is important for achieving a high level generalization capability. Model selection has been studied from various standpoints in the field of statistics, including information statistics, Bayesian statistics, and structural risk minimization. The Bayesian approach (Djuric 1998, Torr 2002) is perhaps the most general and most powerful model selection method. Based on posterior model probabilities, the Bayesian approach

estimates a probability distribution over an ensemble of models. The prediction is accomplished by averaging over the ensemble of models. Accordingly, the uncertainty of the models is taken into account, and complex models with more degrees of freedom are penalized.

Given a set of models $\{M_k, k = 1, 2, \ldots, k_{\max}\}$and data $\mathbf{r}$, the Bayesian approach selects the model with the largest posterior probability. The posterior probability of model M_k is

$$p(M_k \mid \mathbf{r}) = \frac{p(\mathbf{r} \mid M_k)p(M_k)}{\sum_{L=1}^{k_{\max}} p(\mathbf{r} \mid M_L)p(M_L)} \tag{8.8}$$

where $p(\mathbf{r} \mid M_k)$ is the integrated likelihood of model M_k and $p(M_k)$ is the prior probability of model M_k. To find the model with the largest posterior probability, we evaluate $p(M_k \mid \mathbf{r})$ for $k = 1, 2, \ldots, k_{\max}$ and select the model that has the maximum $p(M_k \mid \mathbf{r})$, that is

$$\begin{aligned} M &= \arg \max_{M_k, k=1,\ldots,k_{\max}} \{p(M_k \mid \mathbf{r})\} \\ &= \arg \max_{M_k, k=1,\ldots,k_{\max}} \left\{ \frac{p(\mathbf{r} \mid M_k)p(M_k)}{\sum_{L=1}^{k_{\max}} p(\mathbf{r} \mid M_L)p(M_L)} \right\}. \end{aligned} \tag{8.9}$$

Here, we assume that the models have the same likelyhood a priori, so that $p(M_k) = 1/k_{\max}$, ($k = 1, 2, \ldots, k_{\max}$). Therefore, the model selection in (8.8) will not be affected by $p(M_k)$. This is also the case with $\sum_{L=1}^{k_{\max}} p(r \mid M_L)\, p(M_L)$ since it is not a function of M_k. Consequently, we can ignore the factors $p(M_k)$ and $\sum_{L=1}^{k_{\max}} p(r \mid M_L)\, p(M_L)$ in computing the model criteria. Equation (8.9) then becomes

$$M = \arg \max_{M_k, k=1,\ldots,k_{\max}} \{p(\mathbf{r} \mid M_k)\}. \tag{8.10}$$

To calculate the posterior probability of model M_k, we need to evaluate the marginal density of data for each model $p(\mathbf{r} \mid M_k)$, which requires multidimensional integration

$$p(\mathbf{r} \mid M_k) = \int_{\mathbf{\Phi}_k} p(\mathbf{r} \mid \mathbf{\Phi}_k, M_k)p(\mathbf{\Phi}_k \mid M_k)\mathrm{d}\mathbf{\Phi}_k \tag{8.11}$$

where $\mathbf{\Phi}_k$ is the parameter vector for model M_k, $p(\mathbf{r} \mid \mathbf{\Phi}_k, M_k)$ is the likelihood and $p(\mathbf{\Phi}_k \mid M_k)$ is the prior distribution for model M_k.

In practice, calculating the multidimensional integration is very hard, especially for obtaining a closed-form analytical solution. The research in this area has resulted in many approximation methods for achieving this. The Laplace's approximation method for the integration appears to be a simple one and has

become a standard method for calculating the integration of multi-variable Gaussians (Torr 2002). This yields

$$\begin{aligned} p(\mathbf{r} \mid M_k) &= \int_{\mathbf{\Phi}_k} p(\mathbf{r} \mid \mathbf{\Phi}_k, M_k)\, p(\mathbf{\Phi}_k \mid M_k)\, \mathrm{d}\mathbf{\Phi}_k \\ &\cong (2\pi)^{d_k/2} \mid H(\hat{\mathbf{\Phi}}_k) \mid^{-1/2} \; p(\mathbf{r} \mid \hat{\mathbf{\Phi}}_k, M_k)\, p(\hat{\mathbf{\Phi}}_k \mid M_k) \end{aligned} \tag{8.12}$$

where $\hat{\mathbf{\Phi}}_k$ is the maximum likelihood estimate of $\mathbf{\Phi}_k$, d_k denotes the number of parameters (control points for B-spline model) in model M_k, and $H(\hat{\mathbf{\Phi}}_k)$ is the Hessian matrix of $-\log p(\mathbf{r} \mid \mathbf{\Phi}_k, M_k)$ evaluated at $\hat{\mathbf{\Phi}}_k$,

$$H(\hat{\mathbf{\Phi}}_k) = -\frac{\partial^2 \log p(\mathbf{r} \mid \mathbf{\Phi}_k, M_k)}{\partial \mathbf{\Phi}_k \partial \mathbf{\Phi}_k^T} \Bigg|_{\mathbf{\Phi}_k = \hat{\mathbf{\Phi}}_k} \tag{8.13}$$

This approximation is particularly good when the likelihood function is highly peaked around $\hat{\mathbf{\Phi}}_k$. This is usually the case when the number of data samples is large. Neglecting the terms of $p(\hat{\mathbf{\Phi}}_k \mid M_k)$ and using log in the calculation, the posterior probability of model M_k becomes

$$M = \arg \max_{M_k, k=1,\ldots,k_{\max}} \left\{ \log p(\mathbf{r} \mid \hat{\mathbf{\Phi}}_k, M_k) - \frac{1}{2} \log \mid H(\hat{\mathbf{\Phi}}_k) \mid \right\} \tag{8.14}$$

The likelihood function $p(\mathbf{r} \mid \hat{\mathbf{\Phi}}_k, M_k)$ of a closed B-spline cross section curve can be factored into x and y components as

$$p(\mathbf{r} \mid \hat{\mathbf{\Phi}}_k, M_k) = p(\mathbf{x} \mid \hat{\mathbf{\Phi}}_{kx}, M_k) \cdot p(\mathbf{y} \mid \hat{\mathbf{\Phi}}_{ky}, M_k) \tag{8.15}$$

where $\hat{\mathbf{\Phi}}_{kx}$ and $\hat{\mathbf{\Phi}}_{ky}$ can be calculated by (8.5) .

Consider the x component. Assuming that the residual error sequence is zero mean and white Gaussian with variance $\sigma_{kx}^2(\hat{\mathbf{\Phi}}_{kx})$, we have the following likelihood function

$$\begin{aligned} p(\mathbf{x} \mid \hat{\mathbf{\Phi}}_{kx}, M_k) = &\left(\frac{1}{2\pi\sigma_{kx}^2(\hat{\mathbf{\Phi}}_{kx})} \right)^{m/2} \\ &\exp\left\{ -\frac{1}{2\sigma_{kx}^2(\hat{\mathbf{\Phi}}_{kx})} \sum_{k=0}^{m-1} [x_k - \mathbf{B}_k \hat{\mathbf{\Phi}}_{kx}]^2 \right\} \end{aligned} \tag{8.16}$$

with $\sigma_{kx}^2(\hat{\mathbf{\Phi}}_{kx}, M_k)$ estimated by

$$\hat{\sigma}_{kx}^2(\hat{\mathbf{\Phi}}_{kx}) = \frac{1}{m}\sum_{k=0}^{m-1}[x_k - \mathbf{B}_k\hat{\mathbf{\Phi}}_{kx}]^2 \tag{8.17}$$

Similarly, the likelihood function of the *y* component can also be obtained. The corresponding Hessian matrix $\hat{H}_k$ of $-\log p(\mathbf{r} \mid \mathbf{\Phi}_k, M_k)$ evaluated at $\hat{\mathbf{\Phi}}_k$ is

$$H(\hat{\mathbf{\Phi}}_k) = \begin{bmatrix} \dfrac{\mathbf{B}^T\mathbf{B}}{\hat{\sigma}_{kx}^2(\hat{\mathbf{\Phi}}_{kx})} & \mathbf{0} \\ \mathbf{0} & \dfrac{\mathbf{B}^T\mathbf{B}}{\hat{\sigma}_{ky}^2(\hat{\mathbf{\Phi}}_{ky})} \end{bmatrix} \tag{8.18}$$

Approximating $\frac{1}{2}\log|H(\hat{\mathbf{\Phi}}_k)|$ by the asymptotic expected value of Hessian $\frac{1}{2}(d_{kx}+d_{ky})\log(m)$, we can obtain the Bayesian information criterion (BIC) for selecting the structure of the B-spline curve

$$M = \arg\max_{M_k, k=1,\ldots,k_{\max}} \begin{Bmatrix} -\dfrac{m}{2}\log\hat{\sigma}_{kx}^2(\hat{\mathbf{\Phi}}_{kx}) - \dfrac{m}{2}\log\hat{\sigma}_{ky}^2(\hat{\mathbf{\Phi}}_{ky}) \\ -\dfrac{1}{2}(d_{kx}+d_{ky})\log(m) \end{Bmatrix} \tag{8.19}$$

where d_{kx} and d_{ky} are the number of control points in the *x* and *y* directions respectively, *m* is the number of data points.

In the conventional BIC criterion as shown in (8.19), the first two terms measure the estimation accuracy of the B-spline model. In general, the variance $\hat{\sigma}_k^2$ estimated from (8.17), tends to decrease with the increase in the number of control points. The smaller the variance value in $\hat{\sigma}_k^2$, the bigger the value of the first two terms (as the variance is much smaller than one) and therefore the higher the order (i.e. the more control points) of the model resulting from (8.19). However, if too many control points are used, the B-spline model will over-fit noisy data points. An over-fitted B-spline model will have a poor generalization capability. Model selection thus should achieve a proper tradeoff between the approximation accuracy and the number of control points of the B-spline model. With a conventional BIC criterion, the same data set is used for estimating both the control points of the B-spline model and the variances. Thus the first two terms in (8.19) cannot detect the occurrence of over fitting in the B-spline model selected. In theory, the third term in (8.19) could penalize over-fitting as it appears directly proportional to the number of control points used. In practice, however, we note from our experience

that the effect of this penalty term is insignificant compared with that of the first two terms. As a result, the conventional BIC criterion is rather insensitive to the occurrence of over-fitting and tends to select more control points in the B-spline model to approximate the data point, which normally results in a model with poor generalization capability.

The reason for the occurrence of over-fitting in a conventional BIC criterion lies in the way the variances σ_{kx}^2 and σ_{ky}^2 are obtained. A reliable estimate of σ_{kx}^2 and σ_{ky}^2 should be based on re-sampling of the data. In other words, the generalization capability of a B-spline model should be validated using another set of data points rather than the same data used in obtaining the model. To achieve this, we divide the available data into two sets: a training sample and a prediction sample. The training sample is used only for model estimation, whereas the prediction sample is used only for estimating the data noise σ_{kx}^2 and σ_{ky}^2. For a candidate B-spline model M_k with d_{kx} and d_{ky} control points in x and y directions, the BIC in (8.19) is thus evaluated via the following two steps:

1. Estimate the model parameter $\hat{\mathbf{\Phi}}_k$ using the training sample by (8.5);
2. Estimate the data noise σ_k^2 using the prediction sample by (8.17).

If the model $\hat{\mathbf{\Phi}}_k$ fitted to the training data is valid, then the estimated variance $\hat{\sigma}_k^2$ from the prediction sample should also be a valid estimate of the data noise. If the variance $\hat{\sigma}_k^2$ found from the prediction sample is unexpectedly large, we have reasons to believe that the candidate model fits the data badly. It can be seen that the data noise $\hat{\sigma}_k^2$ estimated from the prediction sample will thus be more sensitive to the quality of the model than the one directly estimated from the training sample, as the variance σ_k^2 estimated from the prediction sample also has the capability of detecting the occurrence of over-fitting.

8.3 Expected Error

In Sect. 8.2, we described our approach to model selection and parameter estimation in the framework of Bayesian statistics. In this section, we will discuss how the same framework for B-spline curve approximation relates to the task of selecting the NBV for acquiring new data. For simplification of the description, we will replace $\mathbf{\Phi}_k$ by $\mathbf{\Phi}$ to show that we are dealing with the selected "best" B-spline model with d_{kx} and d_{ky} control points. To obtain the approximate B-spline model, we will predict the distribution of the information gain on the model's parameter $\mathbf{\Phi}$ along each cross section curve. A measure of the information gain will be designed whose expected value will be maximal when the new measurement data are acquired. The measurement is based on Shannon's entropy whose properties make it a sensible information measure here. We will describe the information entropy of the B-spline model and how to use it to achieve maximal information gain on the parameters of the B-spline model $\mathbf{\Phi}$.

8.3.1 Information Entropy of a B-Spline Model

Given $\mathbf{\Phi}$ and the data points $\mathbf{r} = (\mathbf{r}_i)_{i=1}^m$ which are assumed to be statistically independent, with Gaussian noise of zero mean and variance σ^2, the joint probability of $\mathbf{r} = (\mathbf{r}_i)_{i=1}^m$ is

$$p(\mathbf{r} \mid \mathbf{\Phi}) = \frac{1}{(2\pi\sigma^2)^{m/2}} \cdot \exp[-\frac{1}{2\sigma^2}(\mathbf{r} - \mathbf{B} \cdot \mathbf{\Phi})^T (\mathbf{r} - \mathbf{B} \cdot \mathbf{\Phi})] \tag{8.20}$$

Equation (8.20) has an asymptotic approximation representation defined by Subrahmonia et al. (1996)

$$p(\mathbf{r} \mid \mathbf{\Phi}) \approx p(\mathbf{r} \mid \hat{\mathbf{\Phi}}) \exp[-\frac{1}{2}(\mathbf{\Phi} - \hat{\mathbf{\Phi}})^T H_m (\mathbf{\Phi} - \hat{\mathbf{\Phi}})] \tag{8.21}$$

where $\hat{\mathbf{\Phi}}_k$ is the maximum likelihood estimation of $\mathbf{\Phi}$ given the data points and H_m is the Hessian matrix of $-\log p(\mathbf{r} \mid \mathbf{\Phi})$ evaluated at $\hat{\mathbf{\Phi}}$ given data points $\mathbf{r} = (\mathbf{r}_i)_{i=1}^m$.

The posteriori distribution $p(\mathbf{\Phi} \mid \mathbf{r})$ of the given data is approximately proportional to

$$p(\mathbf{\Phi} \mid \mathbf{r}) \approx p(\mathbf{r} \mid \hat{\mathbf{\Phi}}) \cdot \exp[-\frac{1}{2}(\mathbf{\Phi} - \hat{\mathbf{\Phi}})^T H_m (\mathbf{\Phi} - \hat{\mathbf{\Phi}})] p(\mathbf{\Phi}) \tag{8.22}$$

where the $p(\mathbf{\Phi})$ is the priori probability of the B-spline model parameters. If the priori has a Guassian distribution with mean $\hat{\mathbf{\Phi}}$ and covariance H_{m}^{-1}, we have

$$p(\mathbf{\Phi} \mid \mathbf{r}) \propto \exp[-\frac{1}{2}(\mathbf{\Phi} - \hat{\mathbf{\Phi}})^T H_m (\mathbf{\Phi} - \hat{\mathbf{\Phi}})] \tag{8.23}$$

From Shannon's information entropy, the conditional entropy of $p(\mathbf{\Phi} \mid \mathbf{r})$ is defined by

$$E_m(\mathbf{\Phi}) = \int p(\mathbf{\Phi} \mid \mathbf{r}) \cdot \log p(\mathbf{\Phi} \mid \mathbf{r}) \, \mathrm{d}\mathbf{\Phi} \tag{8.24}$$

If $p(\mathbf{\Phi} \mid \mathbf{r})$ obeys Guassian distribution, the corresponding entropy is Mackay (1991)

$$E_m = \Delta + \frac{1}{2} \log(\det H_m^{-1} \tag{8.25}$$

where Δ is a constant.

The entropy in (8.25) measures the information about the B-spline model parameters, given data points $(\mathbf{r}_1,\ldots,\mathbf{r}_m)$. The more information about $\mathbf{\Phi}$, the smaller the entropy will be. In this work, we use the entropy in (8.25) as the measurement of the uncertainty of the model parameter $\mathbf{\Phi}$. Thus, to minimize E_m, we will make $\det(H_{\mathrm{m}}^{-1})$ as small as possible.

8.3.2 Information Gain

In order to predict the distribution of the information gain, we assume a new data point $\mathbf{r}_{m+1}$ collected along a contour. The potential information gain is determined by incorporating the new data point $\mathbf{r}_{m+1}$. If we move the new point $\mathbf{r}_{m+1}$ along the contour, the distribution of the potential information gain along the whole contour can be obtained. Now, we will derive the relationship between the information gain and the new data point $\mathbf{r}_{m+1}$.

Assume that a new data point $\mathbf{r}_{m+1}$ has been collected. Let $P(\mathbf{\Phi} \mid \mathbf{r}_1,\ldots,\mathbf{r}_m,\mathbf{r}_{m+1})$ be the probability distribution of model parameter $\mathbf{\Phi}$ after a new point $\mathbf{r}_{\mathrm{m+1}}$ is added. Its corresponding entropy is $E_{m+1}=\Delta+\frac{1}{2}\log(\det \hat{H}_{m+1}^{-1})$. The information gain then is

$$\Delta E = E_m - E_{m+1} = \frac{1}{2}\log\frac{\det H_m^{-1}}{\det H_{m+1}^{-1}} \tag{8.26}$$

From (8.18), the new data point $\mathbf{r}_{m+1}$ will incrementally update the Hessian matrix as follows

$$H_{m+1} \approx H_m + \begin{bmatrix} \frac{1}{\sigma_x^2}\cdot \overline{B}_{m+1}^T \overline{B}_{m+1} & \mathbf{0} \\ \mathbf{0} & \frac{1}{\sigma_y^2}\cdot \overline{B}_{m+1}^T \overline{B}_{m+1} \end{bmatrix} \tag{8.27}$$

where $\hat{\sigma}_{m+1}^2 \approx \hat{\sigma}_m^2$. $\overline{B}_{m+1}$ is defined by

$$\overline{B}_{m+1} = [\overline{B}_{0,4}^{m+1} + \overline{B}_{n+1,4}^{m+1}, \overline{B}_{1,4}^{m+1} + \overline{B}_{n+2,4}^{m+1}, \ldots, \overline{B}_{n,4}^{m+1}]$$

The determinant of H_{m+1}

$$\det H_{m+1} \approx \det[\mathbf{I} + \begin{bmatrix} \frac{1}{\hat{\sigma}_x^2} \cdot \overline{B}_{m+1}^T \overline{B}_{m+1} & \mathbf{0} \\ \mathbf{0} & \frac{1}{\hat{\sigma}_y^2} \cdot \overline{B}_{m+1}^T \overline{B}_{m+1} \end{bmatrix} H_m^{-1}] \cdot \det H_m$$

can be simplified to

$$\det H_{m+1} \approx (1 + \overline{B}_{m+1} \cdot [\mathbf{B}^T \mathbf{B}]^{-1} \cdot \overline{B}_{m+1}^T)^2 \cdot \det H_m \tag{8.28}$$

Since $\det H^{-1} = 1/\det H$, (8.26) can be simplified to

$$\Delta E = \log(1 + \overline{B}_{m+1} \cdot [\mathbf{B}^T \mathbf{B}]^{-1} \cdot \overline{B}_{m+1}^T) \tag{8.29}$$

Assuming that the new additional data point $\mathbf{r}_{m+1}$ travels along the contour, the resulting potential information gain of the B-spline model will change according to (8.29). In order to reduce the uncertainty of the model, we would like to have the new data point at such a location that the potential information gain attainable is largest. Therefore, after reconstructing the section curve by fitting partial data acquired from previous viewpoints, the Next Best Viewpoint should be selected as the one that senses those new data points which yield the largest possible potential information gain for the B-spline model.

8.4 View Planning

A view space is a set of 3D positions where the sensor (vision system) takes measurements. We assume that the 3D object is within the field of view and the depth of view of the vision system. The optical settings of the vision system are fixed. Based on these assumptions, the parameters of the vision system to be planned are the viewing pose of the sensor. In this section, the candidate viewpoints are represented in a spherical viewing space. The view space is usually a continuous spherical surface. To reduce the number of viewpoints used in practice, we discretize the surface by using the icosahedron method. In addition, we assume that the view space is centered around the object, and its radius is equal to an a priori specified distance from the sensor to the object. As shown in Fig. 8.2, since the optical axis of the sensor passes through the center of the object, the viewpoint can be represented by pan-tilt angles ϕ ([–180°, 180°]) and θ ([–90°, 90°]).

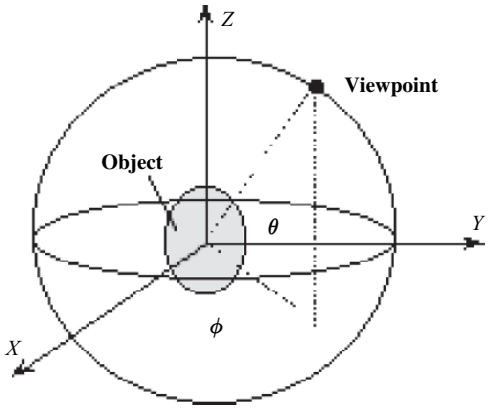


Fig. 8.2. Viewpoint representation

According to the representation of the viewing space, the fundamental task in the view planning here is to obtain the visibility regions in the viewing space that contain the candidate viewpoints where the missing information about the 3D object can be obtained without occlusions. The NBV should be the viewpoint that can give maximum information about the object.

With the above view space representation, we can now map the predicted information gain to the view space for viewpoint planning. For a viewpoint $v(\theta, \phi)$, we say one data point on the object is visible if the angle between its normal and the view direction is smaller than a breakdown angle α of the sensor. The view space V_k for each data point $\mathbf{r}_k$ ($k = 1, 2, \ldots$) is the set of all possible viewpoints that can see $\mathbf{r}_k$. The view space V_k can be calculated via the following procedure:

1. Calculating the normal vector $\mathbf{n}_k$ of a point $\mathbf{r}_k$ ($k = 1, 2, \ldots$) on the object, using a least square error fitting of a 3×3 local surface patch in its neighborhood.
2. Extracting viewpoints from which $\mathbf{r}_k$ is visible. These viewpoints are denoted as view space V_k.

After the view space V_k ($k = 1, 2, \ldots$) is extracted, we construct a measurement matrix **M.** The components $m_{k,j}$ of an l-by-w measurement matrix are given as

$$m_{k,j} = \begin{cases} \langle \mathbf{n}_k \cdot \mathbf{v}_j \rangle & \text{if } \mathbf{r}_k \text{ is visible to } v_j \\ 0 & \text{otherwise} \end{cases} \tag{8.30}$$

where $\mathbf{v}_j$ is the direction vector of viewpoint v_j.

Then, for each view $v(\theta, \phi)$, we define a global measure of the information gain $I(\theta, \phi)$ as the criterion to be summed over all visible surface points seen under this view of the sensor. $I(\theta, \phi)$ is defined by

$$I_j(\theta_j, \phi_j) = \sum_{k \in R_j} m_{k,j} \cdot \Delta E_k \tag{8.31}$$

where ΔE_k is the information gain at surface point $\mathbf{r}_k$, which is weighted by $m_{k,j}$.

Therefore, the Next Best View (θ^*, ϕ^*) is one that maximizes the information gain function of $I(\theta, \phi)$

$$(\theta^*, \phi^*) = \max_{\theta_j,\, \phi_j} I_j(\theta_j, \phi_j) \tag{8.32}$$

8.5 Experiments

8.5.1 Setup

The information entropy based viewpoint planning algorithm is implemented as part of the work for 3D object reconstruction. The setup of a general 3D shape measurement system is schematically shown in Fig. 8.3. The sensor mounted on a robot consists of a projector that projects structured light onto the object and a CCD camera that captures the image of the illuminated object surface (Li and Liu 2003). This range sensor can give depth information of the scanned surface of an object in the form of a "data cloud". In the current implementation, the object is placed on a stationary platform. The robot has 6 DOF and is able to take a measurement of the object from any viewing pose specified within its work space. The modeling process for a 3D object consists of a sequence of four repeated steps: acquiring data on the object surface from a viewpoint, registering the acquired data, integrating the new data with a partial model and determining the NBV. This cycle will be repeated until the NBV terminates.

To slice the acquired "data cloud", we define an interval distance between cross section curves in a certain direction (e.g. the z direction) and project the data in the neighborhood of each cross section curve onto the plane on which the cross section curve lies. The preprocessing results of the 3D "data cloud" are shown in Fig. 8.4. Here the interval between two cross section curves was set at 1.5 mm and the neighborhood of each cross section curve is set at 0.2 mm.

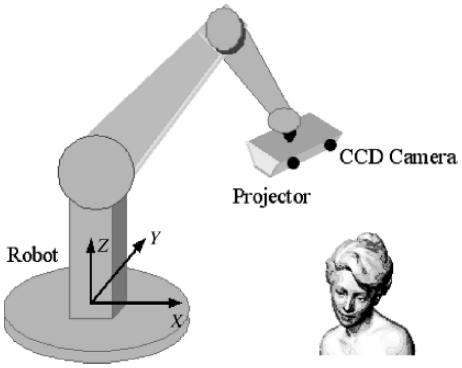


Fig. 8.3. System setup

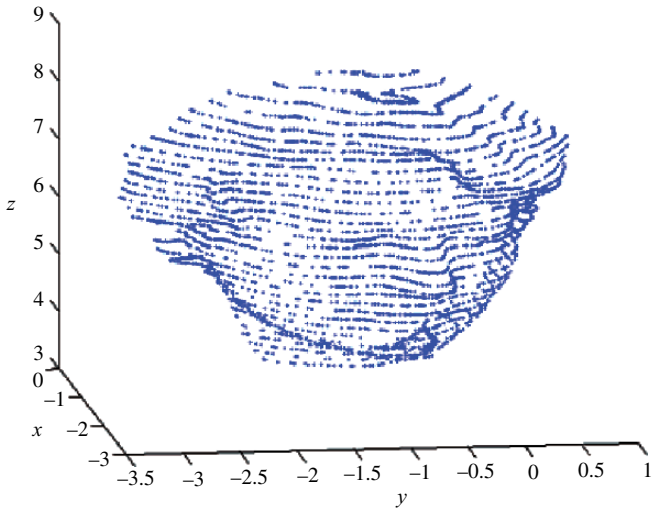


Fig. 8.4. Cross section curves after preprocessing

These points projected onto a cross section curve were distributed randomly. They need to be sorted out before the curve reconstruction can be performed. For each section curve, these projected points were transformed into the polar coordinate system. The phase angle was used to sort these points. To reconstruct these cross section curves via B-splines, we need to select an appropriate model structure first. The model selection is important for automated 3D modeling, to account for the data already acquired and to avoid over-fitting of the model.

8.5.2 Model Selection

In this section, the improved BIC criterion proposed will be used to select the B-spline model to represent the cross section curves. Two cross section curves from a series of sliced cross section curves will be used as examples to demonstrate the effectiveness of our approach. To evaluate the selected models, the following performance indexes are used:

– *Model complexity*, which refers to the number of control points of the B-spline model;
– *Estimation accuracy*, which is defined as the MSE (mean squared errors) between the actual data points and the reconstructed model chosen by a selection criterion.

The model complexity and estimation accuracy provide insights into the appropriateness of model fitting (i.e. over-fitting or under-fitting). In the current implementation, a uniform B-spline is used for reconstructing the cross section curves whose control points are uniformly distributed in the interval between the two end points of the curve in the parameter space. In selecting the model for a cross section curve, the number of control points is iteratively incremented by one from the initial

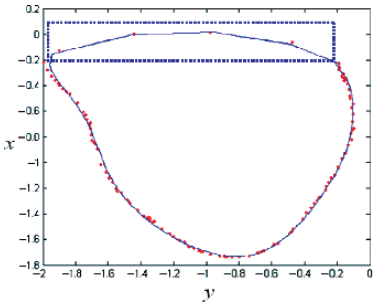
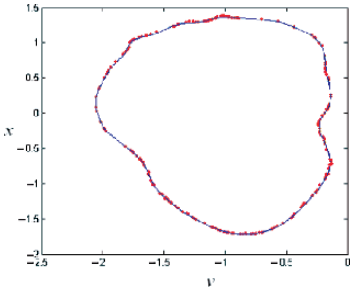
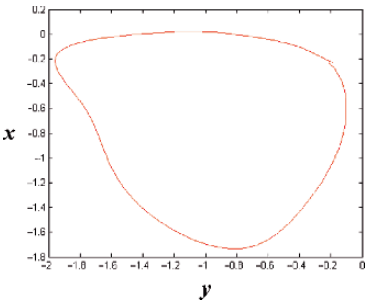
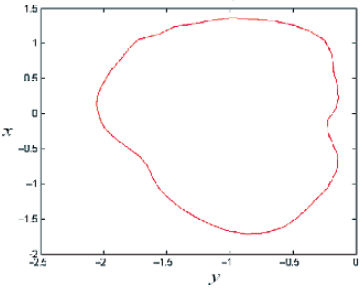
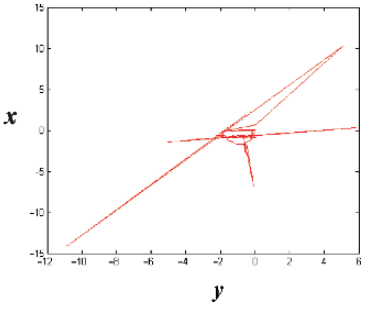
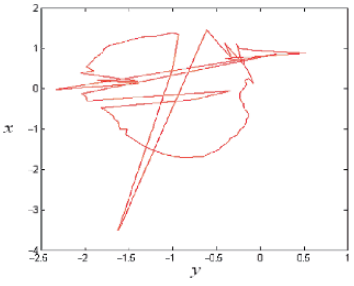
minimum number, while the corresponding BIC value is evaluated using (8.19). The minimum number of control points of a B-spline model is normally set at six here.

We first conducted experiments with only partial data of an object surface acquired by our range sensor from the first view. The object is the head of a statue as shown in Fig. 8.4. For each of the cross section curves, some data points were available for its reconstruction. Here we describe the modeling process via an example in reconstructing one cross section curve. To implement our improved BIC, the available data were first divided into two parts: a training sample set and a prediction sample set. The training sample set was used to estimate the parameters of a candidate B-spline model by (8.5), followed by the estimation of the variance σ_{kx}^2 and σ_{ky}^2 by (8.17) using the prediction sample set. The corresponding BIC value for each of the candidate B-spline models was evaluated by (8.19). The model with maximum BIC value was selected as the optimal one to approximate the data points, giving the resulting model complexity of 9. This model was then verified by using another set of data on the same cross section. The resulting curve is given in the second row in Table 8.1. The estimation accuracy, which is the mean squared errors between the actual data points and the reconstructed model, was found to be 0.0406 mm. As a comparison, the conventional BIC was also applied to the same curve. However, all the 240 data points were used in selecting the model via evaluating the BIC value by (8.19), giving the selected model complexity of 150. Again using another set of data (the same set as used in the above verification), this model was verified, with the resulting curve given in the third row in Table 8.1. The estimation accuracy in this case was found to be 1.8015 mm. This large error shows that the conventional BIC results in over-fitted approximation for the whole curve via the partial data. This illustrates the limitation of the conventional BIC criterion: its insensitivity to over-fitting. Note that in Table 8.1, the scales of the figures are set differently, in order to show the resulting errors in the reconstructed curves by different criteria which are significantly different in magnitudes. Similar phenomena were observed for other cross section curves. Here only the results for one curve are given in Table 8.1. In practical implementation, some physical constraints need to be given. For example, due to self-occlusion, the back of the object will not be visible from the first view. Some points were thus defined between the two end points of the available cross section data to limit the range of the occluded part of the object. It is useful and reasonable to confine the occluded part of the object within the range of the two end points of the available data beyond which the part would actually become visible to the current view. These defined points are highlighted in the "blue box" in the figures in the first row in Table 8.1.

The second experiment was conducted with the one where complete data of a surface were available. The procedures in reconstructing the cross section curves were the same as those in the first experiment. For each section curve, verifications of the models reconstructed by the two methods (our improved BIC and conventional BIC) were again conducted using another set of data (different from that used for reconstructing the model) on the same curve, with the results listed in the second column in Table 8.1. From the results, it is observed that even with complete data for a curve, the conventional BIC still results in an over-fitted model

as seen in the large errors in the verification, while our improved BIC method can reconstruct these cross section curves satisfactorily. With more data available in this experiment, the complexities of the selected models increased using both selection criteria. Yet, the conventional BIC performed poorly with apparent over-fitting in its reconstructed models.

Table 8.1. Comparison of the results of our improved BIC with conventional BIC

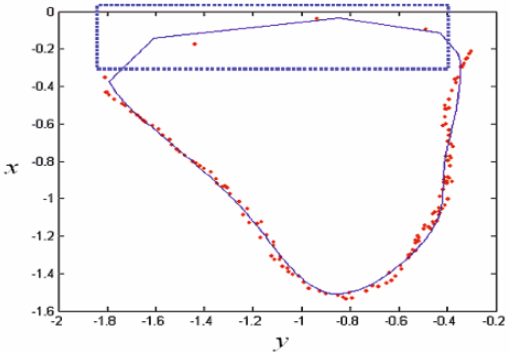
	In the case of partial data available	In the case of complete data available
Cross section data		
Verification results by our improved BIC	Model complexity: 9 Estimation accuracy: 0.0406 	Model complexity: 53 Estimation accuracy: 0.0049 
Verification results by conventional BIC	Model complexity: 150 Estimation accuracy: 1.8015 	Model complexity: 147 Estimation accuracy: 0.3811 

8.5.3 Determining the NBV

In the above section, we showed how our improved BIC criterion selects the B-spline model for the reconstruction of cross section curves. In this section, we will analyze the uncertainty of the B-spline model selected by our improved BIC for each cross section curve, and predict the information gain of the model along each curve using (8.29). Based on this analysis, we then map the information gain onto the view space. The view with maximum information gain is selected as the NBV. Then the vision sensor can take another measurement from the NBV to update the B-spline model. We will take one cross section curve as an example to illustrate the process in determining the NBV.

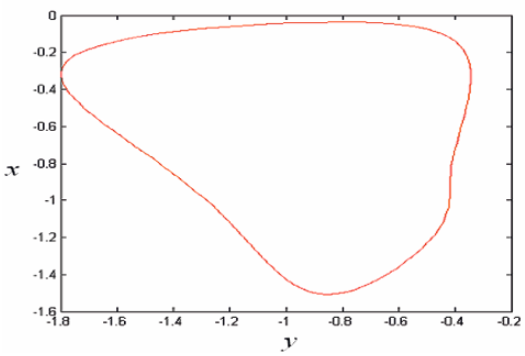
8.5.3.1 Determining the First NBV

First, we take the measurement from an arbitrary initial viewpoint to acquire the first part of data of the unknown object. The data points on one of the cross section curves are shown in Fig. 8.5a. The "blue box" in Fig. 8.5a contains the points to confine the range of the occluded part of the object. Since these points are few in number, their effects on the predicted information gain of the B-spline model can be ignored. Figure 8.5b is the reconstructed B-spline model using the partial data acquired from the first viewpoint. This model is a rough approximation for the whole cross section curve. Using this model, we predict the potential information along the reconstructed curve. As shown in Fig. 8.5c, the place on the curve where the data are missing (the missing part) corresponds to a high-potential information gain. This indicates that the occluded part should be given high priority in the next measurement. Note that the information gain (in Fig. 8.5c) is given in the parameter space of the B-spline curve here.

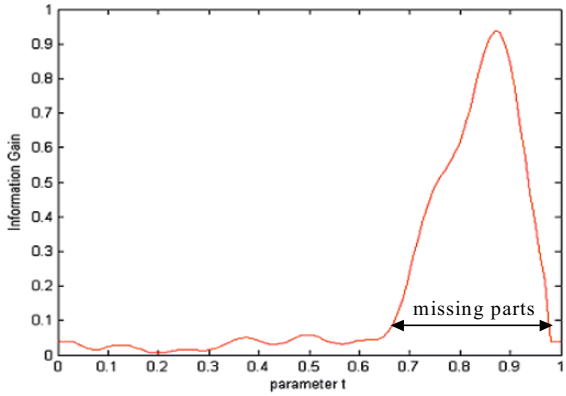


(a) Data on a cross section curve acquired from the first view

Fig. 8.5. Reconstruction of cross section curve and predicted potential information gain under the first viewpoint



(b) Reconstructed B-spline Curve



(c) The potential information gain

Fig. 8.5. (Continued)

Following the above procedure, each cross section curve is reconstructed in a B-spline model, with the corresponding information gain obtained. Here each cross section curve is considered to be equally important, so that we can normalize the predicted information gain for each of the cross section curves covered by the current view. Figure 8.6 shows all the cross section curves reconstructed from the 3D data points taken from the first viewpoint.

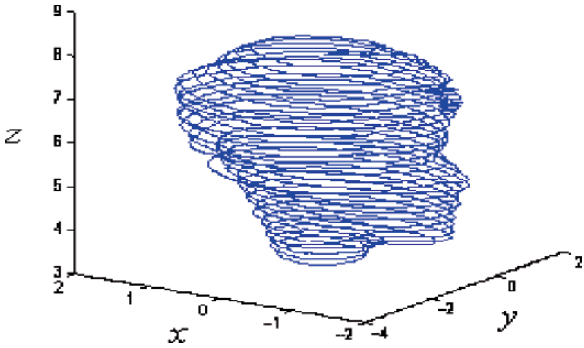


Fig. 8.6. The reconstructed cross section curves

In the above reconstruction, since only the data points from the first viewpoint are available, the obtained B-spline model cannot describe the whole object accurately. Yet, it enables us to obtain a rough shape and the information gain about the object. Based on the reconstructed partial model, we then map the predicted information gain onto the view space. As a result, we can obtain the relationship between the predicted information gain about the object and the viewpoints, which is also referred to as "View Space Visibility". As shown in Fig. 8.7, the viewpoint at [–3.0°, 107°] has the maximum information gain and is thus selected as the NBV.

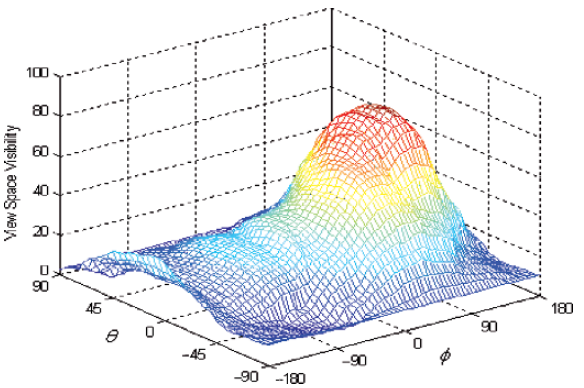
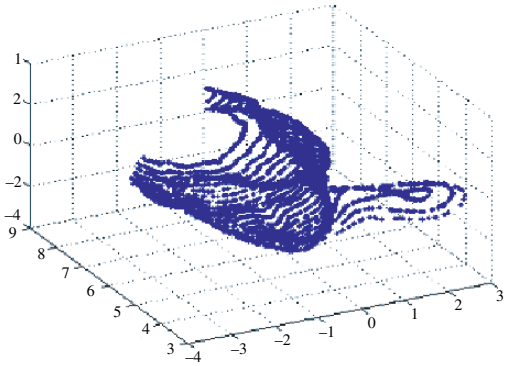


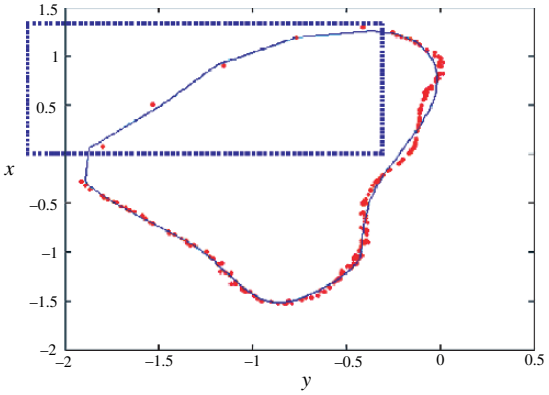
Fig. 8.7. "View Space Visibility" for the first NBV

8.5.3.2 Determining Further NBVs

After the first NBV was selected, the robot was commanded to move the vision sensor to this viewpoint to take new measurements. The newly acquired data were then sliced and registered, to yield the data acquired from the first two viewpoints as shown in Fig. 8.8a.

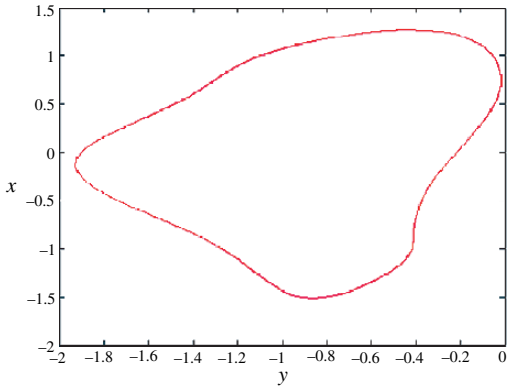


(a) Data acquired from the first two viewpoints after slicing

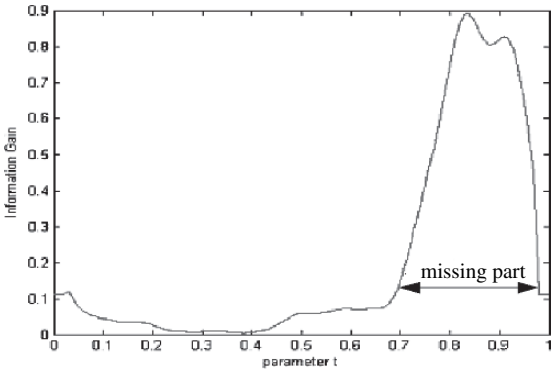


(b) Data on a cross section acquired from the first two viewpoints

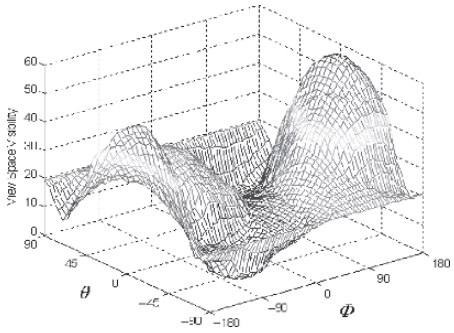
Fig. 8.8. The process of determining the second NBV



(c) Reconstructed B-spline Curve based on the first two viewpoints



(d) The information gain based on the first two viewpoints

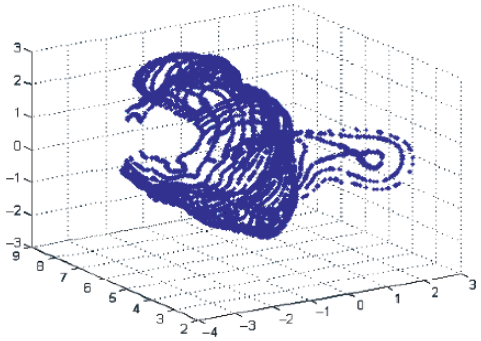


(e) "View Space Visibility" for determining the second NBV

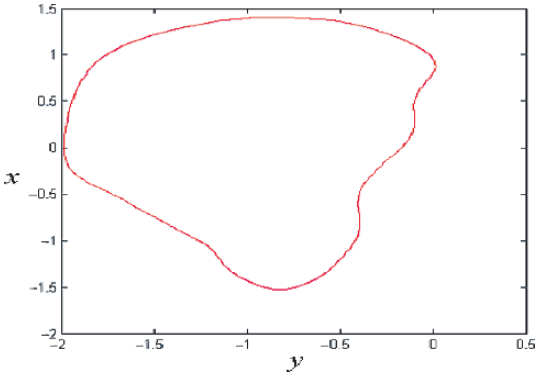
Fig. 8.8. (Continued)

Using the available data, model selection and information gain prediction were performed following the same procedures as described above. For an example cross section shown in Fig. 8.8b, the newly reconstructed curve is given in Fig. 8.8c and the updated information gain is given in Fig. 8.8d. The predicted information gains for all the cross section curves were then mapped onto the view space, to give the updated view space visibility (shown in Fig. 8.8e) for determining the second NBV. From this view space visibility map, the second NBV was selected at [5°, 160°].

The above described procedures in determining the NBV and acquiring new data are repeated for subsequent NBVs. The procedures and results in determining the third NBVs are given in Fig. 8.9. Each time when new data are available from the new viewpoint, the corresponding cross section curves (e.g. the curve in Fig. 8.5b) are updated (as shown in Figs. 8.8c and 8.9b). The prediction of the information gain is also updated at each new viewpoint, as seen in Figs. 8.8d and 8.9c. As a result of the updated "View Space Visibility" evaluation at the second NBVs (see Fig. 8.9d), the third NBV was selected at [7°, –10°].

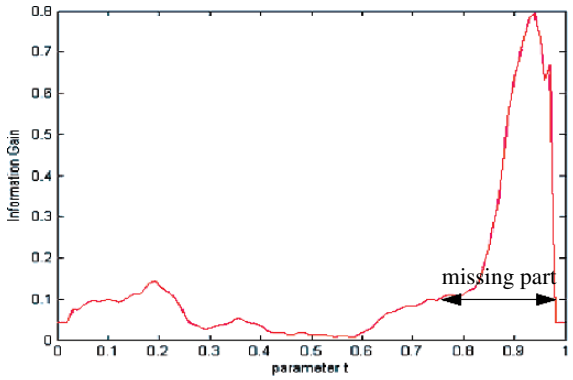


(a) Data acquired from the first three viewpoints

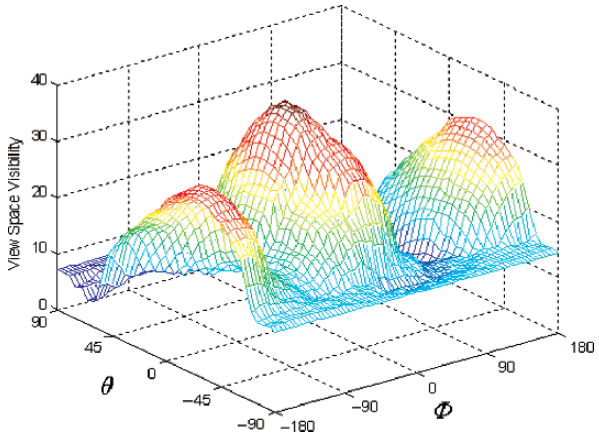


(b) Reconstructed B-spline Curve based on the first three viewpoints

Fig. 8.9. The process in determining the third NBV



(c) The information gain based on the first three viewpoints

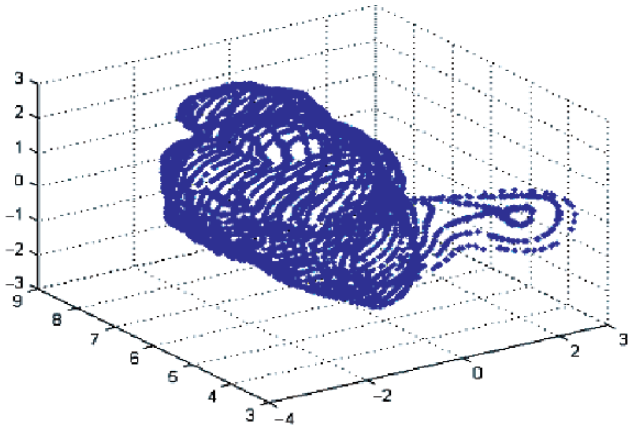


(d) "View Space Visibility" for determining the third NBV

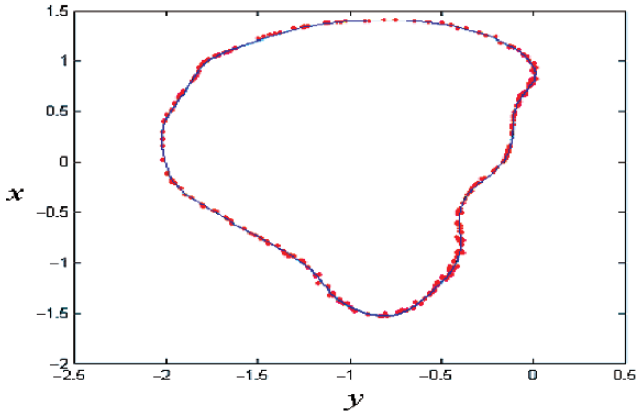
Fig. 8.9. (Continued)

8.5.3.3 Complete Reconstruction

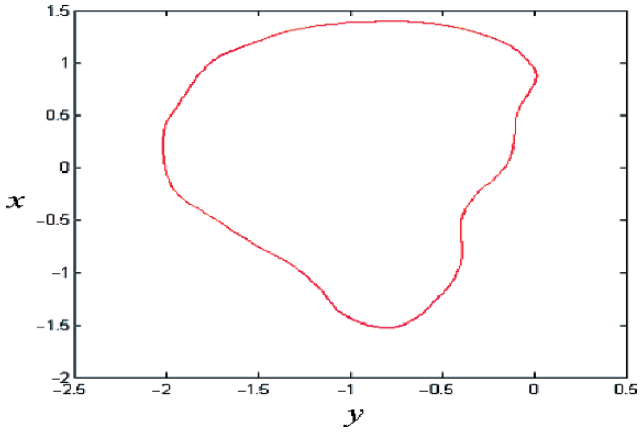
After the third NBV is determined, we obtained the complete data about the object as shown in Fig. 8.10a. The complete data points and final reconstruction result of a cross section curve are shown in Fig. 8.10b and c respectively.



(a) Data acquired from the first four viewpoints

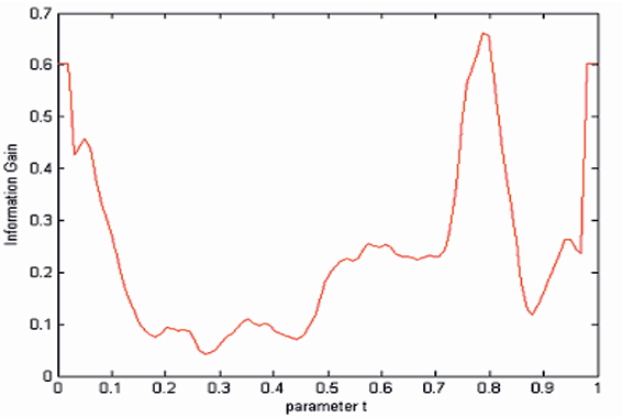


(b) Data on a cross section curve acquired from the first four viewpoints



(c) Reconstruction result of a cross section curve based on the first four viewpoints

Fig. 8.10. Reconstruction of a cross section curve and information gain



(d) The information gain based on the first four viewpoints

Fig. 8.10. (Continued)

As shown in Figs. 8.5c, 8.8d and 8.9c, the information gain has an outstanding peak on the part where the 3D data are missing. This peak will become less and less outstanding with the increase of the 3D data available from new viewpoints. When complete data on these cross section curves are obtained (as from the third NBV here), the peak in the information gain becomes non-apparent and appears more "noise" like (as seen in Fig. 8.10d), which indicates that there are no apparent missing data or occluded parts on the object surface. The disappearance of the peak (significant decrease in the peak value) in the information gain was used as the termination condition in automated planning of the NBVs.

From the experiment results, it is observed that the reconstructed model complexity tends to increase with the availability of additional data, which indicates that the model can describe the previously unknown object in more and more details as new measurements are taken. At the same time, the uncertainty about the object decreases gradually. The results for a typical cross section curve are shown in Table 8.2. The finally reconstructed model is visualized in Fig. 8.11. The final reconstruction accuracy evaluated using MSE between the actual data points and the reconstructed cross section B-spline curves was 0.0061, which is quite satisfactory.

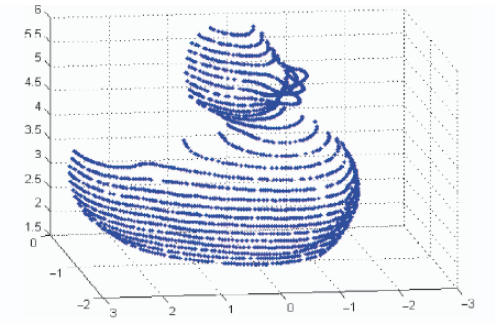
Table 8.2. The results of view planning for the statue

Next best view	1st viewpoint	1st NBV	2nd NBV	3rd NBV
Model complexity	7	11	22	26
Entropy of B-spline model	–15.81	–16.23	–18.95	–19.79

Fig. 8.11. The finial reconstruction result of the statue

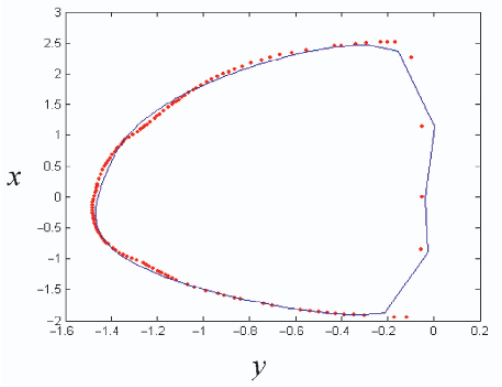
8.5.4 Another Example

Another experiment was conducted using a model of a duck. For simplicity, we only give the results (in Fig. 8.12) to show the procedures of determining the first NBV.

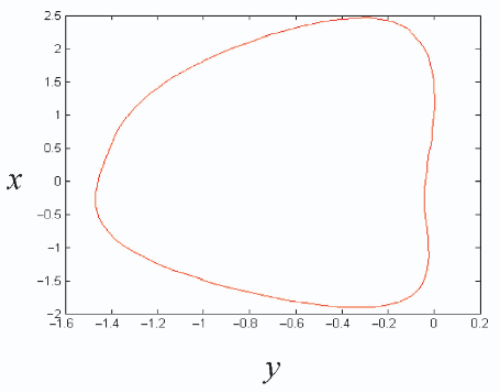


(a) Data acquired from the first viewpoint of the duck model

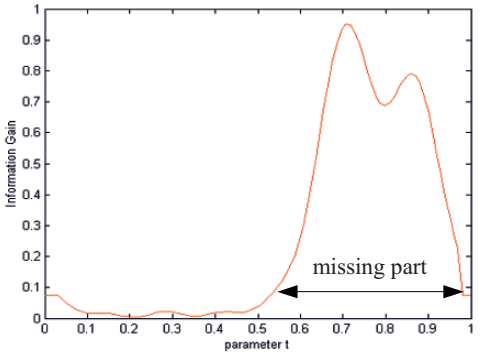
Fig. 8.12. Reconstruction of cross section curves and predicted information gain



(b) Data on a cross section acquired from the first viewpoints

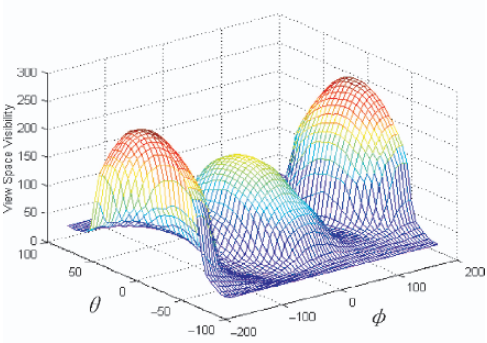


(c) Reconstructed B-spline curve



(d) The information gain

Fig. 8.12. (Continued)



(e) “View Space Visibility” for the first NBV

Fig. 8.12. (Continued)

The viewpoint [0°, 175°] with maximum information gain was selected as the NBV. The procedures of determining other NBVs are the same as those described in the above section. In this example, three viewpoints in total were needed to reconstruct the duck model. The results in view planning for a typical cross section curve are shown in Table 8.3. The accuracy of the finally reconstructed object surface is 0.0076. The reconstructed object is shown in Fig. 8.13. It is observed that the model complexity for the finally reconstructed cross section curve (68) here is higher that that for the example curve (26) in the previous experiment. This is due to the difference in the shapes from the actual data points. The shape of the former curve (partly given in Fig. 8.12b) is simpler and smoother than the latter (Fig. 8.10b). A higher complexity in the selected model indicates the higher level of confidence in the reconstruction for a simpler shape. For a complex shape, a lower complexity in the selected model gives it stronger ability in preventing over-fitting the data, which is of particular importance for NBV planning.

Table 8.3. The results in view planning for the duck model

Next best view	1st viewpoint	1st NBV	2nd NBV
Model complexity	7	35	68
Entropy of B-spline model	–14.26	–20.23	–30.56

Fig. 8.13. The finial reconstruction result of the duck model

8.6 Summary

In this chapter, we presented a novel viewpoint planning method by incrementally reducing the uncertainties of the reconstructed models. With this method, the object's surface is first decomposed into a set of relatively simple cross section curves, with each to be reconstructed by a set of closed B-spline curves. Then the uncertainties of the B-spline models are analyzed with the information entropy as the measurement of the uncertainty for guiding the selection of the next best view. The information gain of the set of cross section B-spline models is predicted and mapped onto the view space. The viewpoint with maximum visibility is selected as the Next Best View. In addition, an improved BIC criterion is proposed for the model selection. With this new criterion, the acquired data points are divided into two parts: one for estimating the B-spline model parameters and the other for estimating the data noise. The re-sampling of the data enables a reliable estimate of data noise, since the generalization capability of a B-spline model should be validated using another set of data points rather than those used for the approximation. Compared with the conversional BIC criterion, the model selected with our improved BIC criterion is more sensitive to over-fitting and thus has a better generalization capability which is particularly important for NBV planning.