IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 2, 2019, accepted August 6, 2019, date of publication August 15, 2019, date of current version September 11, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2935547

Volumetric Next Best View by 3D Occupancy Mapping Using Markov Chain Gibbs Sampler for Precise Manufacturing

LEI HOU[1], XIAOPENG CHEN[2], KUNYAN LAN[2], RUNE RASMUSSEN[1], AND JONATHAN ROBERTS[1], (Member, IEEE)

[1]Science and Engineering Faculty, Queensland University of Technology, Brisbane, QLD 4000, Australia
[2]Key Laboratory of Biomimetic Robots and Systems of Ministry of Education, School of Mechatronical Engineering, Beijing Institute of Technology, Beijing 100081, China

Corresponding author: Xiaopeng Chen (xpchen@bit.edu.cn)

ABSTRACT In this paper, we propose a model-free volumetric Next Best View (NBV) algorithm for accurate 3D reconstruction using a Markov Chain Monte Carlo method for high-mix-low-volume objects in manufacturing. The volumetric information gain based Next Best View algorithm can in real-time select the next optimal view that reveals the maximum uncertainty of the scanning environment with respect to a partially reconstructed 3D Occupancy map, without any priori knowledge of the target. Traditional Occupancy grid maps make two independence assumptions for computational tractability but suffer from the overconfident estimation of the occupancy probability for each voxel leading to less precise surface reconstructions. This paper proposes a special case of the Markov Chain Monte Carlo (MCMC) method, the Gibbs sampler, to accurately estimate the posterior occupancy probability of a voxel by randomly sampling from its high-dimensional full posterior occupancy probability given the entire volumetric map with respect to the forward sensor model with a Gaussian distribution. Numerical experiments validate the performance of the MCMC Gibbs sampler algorithm under the ROS-Industry framework to prove the accuracy of the reconstructed Occupancy map and the completeness of the registered point cloud. The proposed MCMC Occupancy mapping could be used to optimise the tuning parameters of the online NBV algorithms via the inverse sensor model to realise industry automation.

INDEX TERMS Active vision, Markov chain Monte Carlo, occupancy mapping, 3D reconstruction, viewpoint planning.

I. INTRODUCTION

With the development of industrial robotics and robotic vision, autonomous 3D reconstruction of one-of-a-kind objects for industrial manufacturing attracts more and more interest. The resultant 3D surface mesh can be used in path planning, inspection and quality analysis. To reconstruct the geometry surface of an object being scanned, a robot must position the vision sensor at different views until full surface coverage is achieved. Both coverage ratio and time efficiency are important for the scanning operation, because redundant partial point clouds degrade registration, whereas exhaustive exploration can be time-consuming.

The associate editor coordinating the review of this article and approving it for publication was Jingchang Huang.

The reconstruction problem has been intensively studied in active vision [1] for decades by Chen *et al.* [2], Scott [3] and Jing *et al.* [4], but a CAD model is required a priori for the computational geometry methods by Tarabanis *et al.* [5] or for the generate-and-test methods by Scott *et al.* [6] to solve the Set Covering Problem (SCP) with a resultant optimal sub-viewset. However, the model-based methods can only be implemented off-line and hence cannot be directly integrated into industrial automation. For the sake of Industry 4.0, as well as the model-free request for manufacturing, this paper proposes a fully autonomous Next Best View (NBV) solution under the open-source ROS-Industry framework Quigley *et al.* [7], improved by the Markov Chain Monte Carlo (MCMC) method, namely the Gibbs sampler, to generate highly accurate volumetric maps to guide the Next Best View algorithm to reconstruct surface geometry

via the metric of information entropy. The Gibbs sampler can precisely estimate the residual uncertainty in each voxel given the whole map and hence generates benchmark Occupancy maps with satisfactory accuracy for information gain based 3D reconstruction. The view that can reveal the maximum unobserved information of the scanning object w.r.t the partially reconstructed map is chosen as the next best view from the candidate viewset.

To adapt to the non-model based reconstruction tasks for manufacturing, the information that can be obtained from a view is not known a priori and needs to be estimated online during the reconstruction process. The Next Best View algorithm needs to reason about the difference between the observed and the unobserved volume based on the current 3D volumetric map to compute the information gain as described by Kriegel *et al.* [8] and Vasquez-Gomez *et al.* [9]. The use of Occupancy Grid Mapping (OGM) by discretizing the 3D environment into voxels and determining the occupancy state of each voxel in the representation of the world model was first introduced by Moravec and Elfes [10], Moravec [11]. The visibility of the object surface from a given view can be estimated by casting rays from the vision sensor frame into the voxel space upon the object surface according to the sampling model of the vision sensor [12]. The information entropy of each view is defined as the sum of the occupancy probability contained in each voxel covered by the rays within the frustum of the vision sensor [13].

The posterior occupancy probability of each voxel is calculated probabilistically in the Bayesian framework by estimating the residual uncertainty in the Occupancy map due to noisy sensor measurements and random reflection [14]. For computational efficiency, traditional Occupancy mapping methods make two independence assumptions that are strictly untrue in the real world but are still dominantly used even in the most recent model-free NBV algorithms by Isler *et al.* [15] and Daudelin and Campbell [16], namely the voxels in the occupancy map are independent of each other; and the scan measurements are conditionally independent of each other, which do not consider the correlation between proximate voxels especially for 3D vision sensors with wider frustum. Such simplifications expedite the occupancy computation in real-time but result in inaccurate maps that are inconsistent with the measurement data for complex objects, e.g. holes being ignored in the final model, not to mention the patch map under the camera frustum induces important dependency between the neighbouring voxels on edges. Consequently, the conditional independence assumption for scan measurements may yield suboptimal results and the two independence assumptions methods result in under or over-confident estimation of the occupancy probability of voxels in the map, especially near the boundary of objects.

To eliminate the incorrect independence assumptions, a variant of the Markov Chain Monte Carlo (MCMC) method, the Gibbs sampler [17], [18], is implemented for 3D Occupancy mapping in active vision by the Octomap library [19], to estimate accurate occupancy probability for each voxel in the volumetric NBV algorithms. The MCMC Gibbs sampler is a stochastic method mostly used in the Bayesian inference that composes of a set of conditional probabilities, as an alternative to the deterministic methods, e.g. the expectation maximisation algorithm (EM), to estimate the true posterior probability. The Metropolis algorithm [20], [21] and the improved Metropolis-Hastings algorithm [22], [23] are initial MCMC methods to generate a Markov chain from a distribution of interest and accept/reject the new sample according to an acceptance probability. However, the Metropolis algorithms may become inefficient in high dimensions if the resultant Markov chain diverges far from the expected distribution. The Gibbs sampler is especially useful in high dimensional sample space by generating samples from a multivariate joint distribution using the univariate full conditional probability distribution, as a series of one-dimensional random variable generation. The ideal scenario for Gibbs sampler happens when direct sampling is available for each full conditional probabilities. Even if the decomposition of the high-dimensional sampling space by the Gibbs sampler leads to intractable conditional distribution from which the random samples cannot be generated, namely the nonconjugate conditionals, several solutions have been well proposed, e.g. the Gibbs sampler by sampling importance re-sampling using the prior probabilities [24], the self-tuning Gibbs-stopper and the Griddy-Gibbs sampler [25] to generate the exact posterior, and the efficient Recycling Gibbs sampler [44] using all the auxiliary internal samples, thus making the Gibbs sampler the most practical MCMC method. Given that the posterior probability of the Bayesian update framework for our Occupancy mapping is in the context of the univariate conditional probability of each voxel given the state of all the other voxels that is actually the forward sensor model with Gaussian distribution, and that the integrated importance sampling and the rejection/acceptance algorithms may increase the computational burden as well as the extra parametrisation, only the standard Gibbs sampler is employed in our Occupancy mapping algorithm at the current stage to process tens of thousands of voxels.

This paper begins by reviewing the full Bayesian solution which is in fact computationally untraceable above one dimensional space. Our 3D Occupancy mapping method applies the statistical formulation of the forward sensor model to facilitate the computation of the full conditional probability by the MCMC Gibbs sampler under the Bayesian framework that can maintaining all the dependency among neighbouring voxels, thus solving the Occupancy mapping in the original high dimensions. The usefulness of the Gibbs sampler increases greatly as the dimension of the sampling space increases, e.g. the 3D Octree map representing millions of voxels in 3D environments. With the help of the uniformal voxels, the multi-resolution Octree data structure and the Gaussian 3D forward sensor model, the MCMC Gibbs sampling is directly extended to the 3D Occupancy mapping as a numerical solution in the active vision for the first time. It is validated that the Gibbs sampling method can accurately

estimate the true occupancy probability of each voxel by modeling the influence of its neighbouring voxels without any independence decomposition. The resultant Octree map can be used as a benchmark for 3D surface reconstruction, especially in the field of manufacturing, where the occupancy of the poorly mapped areas and the unmapped areas need to be taken into account for completeness. Furthermore, the MCMC method can also be used to improve the performance of the online two independence assumptions methods by optimising the tuning parameters of the update terms within the inverse sensor model off-line w.r.t to the cross entropy of the benchmark Octree map that is otherwise only exist for toy examples in simulation. The proposed MCMC Gibbs sampler algorithm is generic for all 3D robot mapping applications in ROS compatible formats, such as industrial manipulators, mobile robots and drones.

Numerical experiments are carried out with different initial conditions and prior probabilities, to validate the accuracy of the MCMC Gibbs Occupancy mapping based Next Best View algorithm. The contribution of this paper is as follows:

- Propose a 3D forward sensor model with Gaussian distribution as the full conditional probability for the Gibbs sampler to generate true probabilistic volumetric maps directly in high dimensional space without any independence decomposition.
- Propose a MCMC Gibbs sampler NBV by decomposing the high-dimensional sample space by the full conditional probability of each voxel, to generate benchmark 3D Occupancy maps, for accurate 3D surface reconstruction and optimisation.

II. RELATED WORK

The Next Best View research belongs to the active vision and has been researched into for decades as early as 1988 by Aloimonos *et al.* [1], Bajcsy [26], and Blake and Yuille [27]. There are three main applications making use of the Next Best View algorithms, namely: 3D reconstruction, robot exploration, and industrial manufacturing. In accordance with the wide range of applications, e.g. surface reconstruction, building surveillance and quality analysis, nowadays the observation targets are often not known a priori.

3D reconstruction was typically classified into model-based and non-model-based as described by Chen *et al.* [2] and Scott *et al.* [6]. The model-based methods require a prior CAD model of the object with precise knowledge about the size, the form or even the pose in the latest inspection algorithms by Jing and Shimada [28]. However, such detailed information for manufacturing objects at small businesses is not always available, let alone for the model-based methods the off-line view planning and the post-processing of point clouds alignment can be very time-consuming.

Non-model-based methods gradually become desirable for many real-world applications, e.g. historical scene reconstruction, scan-and-plan industry manufacturing, and the post-processing quality assessment. Based on the rapid development of computer vision algorithms and hardware upgrade, the non-model-based methods are further divided into the surface based and the volumetric based. For the surface-based approaches, the Next Best View is estimated in the form of the frontier voxels as described by Yamauchi [29], and latest by Quin *et al.* [30] and Monica and Aleotti [31] for unknown space exploration. Although these approaches can directly work on surface patches which compose the final output of the triangle mesh with high efficiency, the image processing and the visibility modeling can be rather complex, e.g. by using the B-Spline by Li and Liu [32].

The volumetric approach for the Next Best View algorithm becomes predominant with the development of the 3D robot mapping technologies, e.g. Occupancy Grid mapping (OGM). Each candidate view is evaluated by estimating the volumetric information contained in all the traversed voxels between the view origin and the first ending point upon the target surface. Information entropy is utilised to quantify the uncertainty in each voxel and to estimate the sum of the occupancy probability in all the traversed voxels from a view in the volumetric map. Merali and Barfoot [33]–[35] proposed several novel algorithms and mathematical models to improve the accuracy for the traditional 1D and 2D Occupancy grid maps built by laser range finders, namely the Patch Map and etc. Dhiman *et al.* [36] proposed a modern Maximum a Posterior (MAP) inference method over graphical models based on the forward sensor model, which was originally introduced by Thrun [37] with the expectation maximisation (EM) algorithm by performing hill-climbing for 2D maps, taking into account of the dependency between neighbouring voxels. Kaufman *et al.* [38] implemented a computational efficient inverse sensor model by the MATLAB simulation for a 2D laser range-finder by exactly using the cell occlusions on the Occupancy maps. With the development of the RGB-D cameras and the stereo vision [39], the sensor sampling model is extended from the narrow-beam sensors such as laser range-finder and sonar, to 3D vision sensors with limited ranges whereas at the faster frame rates. Iser *et al.* [15] implemented a probabilistic volumetric information gain formulation for active 3D reconstruction by mobile manipulation using the Octomap library. The exact inverse sensor model was further improved by classifying a group of different volumetric information (VI) formulations considering the occlusion and the proximity in complicated 3D environments. Daudelin and Campbell [16] further expanded a 3D reconstruction solution for mobile robots, eliminating the predetermined set of candidate views by using an RGBD-SLAM algorithm to autonomously navigate in a broader search space. Potthast and Sukhatme [40] introduced the Hidden Markov Model (HMM) to formalise the observation probability of voxels using an empirical state transition law in occluded environments. However, the two independence simplifications in all the above solutions result in an overconfident estimation of the occupancy probability of each voxel, especially on the boundary of the object which contains the richest geometry information for surface reconstruction.

The Gibbs sampler was formally introduced by Geman and Geman [17] with simulated annealing for Bayesian reconstruction of degraded images, but its origin can be traced back as early as [20], [22]. It was further shown by Gelfand and Smith [41] to be applicable to a variety of Bayesian inference problems. In image reconstruction [42], the Gibbs sampler can decompose the joint probability of high dimensional sample space into a series of one dimensional conditional probability of each unknown parameters in the image model. If decomposing the high dimensional space leads to intractable conditional distributions that cannot generate random samples, the adaptive rejection sampling [43] employing the envelope and the squeezing functions, the sampling importance re-sampling [24] using the prior probability to approximate the posterior probability, the component-wise Hastings algorithm [23] choosing the appropriate distribution automatically, and the recycling Gibbs sampler benefiting from the rest of the auxiliary samples can be integrated within the Gibbs sampler to draw from the non-conjugate conditional probabilities. Although the Markov random field image model was analysed to be a lattice system with Gibbs distribution to converge to the maximum a posterior (MAP) estimate and sufficient Gibbs sampler algorithms exist in 2D image processing, the Gibbs sampler has not been extended into the scope of high-dimensional 3D volumetric visual processing, e.g. Octree, 3D reconstruction, RGB-D cameras and etc.

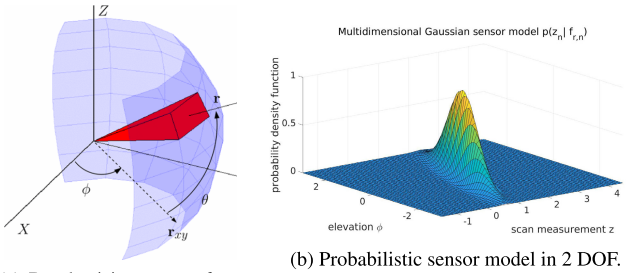
III. FULL BAYESIAN SOLUTION

Compared to the 'two assumptions' OGM methods, a full Bayesian solution without any independence assumptions can treat the residual uncertainty in each voxel in even noise-free environment. The full Bayesian solution uses the same Bayesian framework to update the occupancy probability for each voxel $P(m_i)$, given a set of scan measurements $Z_{1:t}$ and the corresponding robot pose $X_{1:t}$. Since the environment is divided into n_m voxels, there are $R = 2^{(n_m)}$ possible occupancy grids, which are represented by m^r, where $r = 1 \dots R$.

The forward sensor model $p(z_n|f_{r,n})$ is initially generalised by Thrun [37], and then specified by Pathak *et al.* [45] for 3D vision sensors. The probabilistic distribution for a Gaussian sensor along the measurement axis is,

$$p(z_n|f_{r,n}) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\eta_{hit}\frac{(z_n - z_{exp})^2}{2\sigma^2} - \eta_{max}\frac{(z_n - z_{max})^2}{2\sigma^2} - \eta_{rand}\log\frac{z_{max}^2}{2\sigma^2}\right\}, \quad (1)$$

The Gaussian noise variable centred at the range operator z_{exp} returns the radial distance of the centre of the voxel m_i from the sensor origin. The random measurement obeying the uniform distribution where z_{max} is the maximum sensor range. The scan measurement can also miss all targets as a maximum range reading. All the above different causes can be united together with respective weights.



(a) Depth vision sensor frustum in 3D space.

(b) Probabilistic sensor model in 2 DOF.

FIGURE 1. (a) Frustum parameterisation of depth vision sensor in 3D space. (b) Multidimensional gaussian distribution for a scan measurement in 2DOF for 3D visualisation.

The off-the-shelf depth vision sensors usually have a wider frustum including neighbouring voxels along the measurement axis, with Gaussian errors on orientations as shown in Figure 1 (a). A generic sensor model with two additional degrees of freedom, e.g. the azimuth (θ) and the altitude angle (ϕ) is in the closed-form as in [45],

$$p(z_n, \theta, \phi) = \frac{1}{2}\cdot\left\{1+\exp\left\{-\left[\frac{2\left((z_n - \bar{z}_n) + 2\sigma_{z_n}\right)}{\sigma_{z_n}} + \frac{(\theta-\bar{\theta})^2}{\sigma_\theta^2} + \frac{(\phi-\bar{\phi})^2}{\sigma_\phi^2}\right]\right\}\right\}^{-1} + \frac{1}{3}\cdot\exp\left\{-\left[\frac{(z_n-\bar{z}_n)^2}{\sigma_{z_n}^2} + \frac{(\theta-\bar{\theta})^2}{\sigma_\theta^2} + \frac{(\phi-\bar{\phi})^2}{\sigma_\phi^2}\right]\right\}, \quad (2)$$

A Gaussian sensor model in 2DOF (z and ϕ) with the joint multidimensional Gaussian distribution is illustrated in Figure 1 (b), whereas the higher dimension of the Gaussian variable can be illustrated by many sample points.

The full Bayesian algorithm can now be written in the form of the stochastic sensor model with Gaussian error long the measurement axis.

$$p(m_i|Z, X) = \frac{p(z_n|f_{r,n})p(m_i|Z_{1:n-1}, X_{1:n-1})}{\sum_{j=1}^{R} p(z_n|f_{r,n})p(m_j|Z_{1:n-1}, X_{1:n-1})}. \quad (3)$$

By computing the probability of each of the R maps, the occupancy probability of a voxel can be estimated by summing the map probabilities for all the maps in which the voxel of interest m_i is occupied.

$$P(m_k|Z_{1:t}, X_{1:t}) = \sum_{i=1}^{R} P(m_k|m_i)P(m_i|Z_{1:t}, X_{1:t}), \quad (4)$$

where $P(m_k|m_i) \in \{0, 1\}$.

The full Bayesian solution can calculate the probability of a whole map being occupied, rather than only considering a single voxel within a single ray, regardless of the correlation in its proximate voxels. However, computing the full Bayesian solution is exponential in the number of voxels in the map, n_m, namely $R = 2^{(n_m)}$ and is only feasible for 1D OGMs.

IV. MARKOV CHAIN MONTE CARLO

A. MCMC GIBBS SAMPLING

To obtain accurate Occupancy maps for 3D reconstruction, residual uncertainty in each voxel is required to accurately calculate the information gain for NBV. To approximate the accurate occupancy probability, an inference method by numerical sampling, known as the Monte Carlo technique [21] is used, specifically the Markov Chain Monte Carlo (MCMC) method, to sample from a complicated distribution, e.g. the full Bayesian solution. The MCMC method can satisfy the Markov property by generating a Markov chain that converges to the expected posterior probability. The MCMC method can be applied to automatic speech recognition, image restoration [17] and machine learning [46]. Both the Metropolis algorithm and the variant Metropolis-Hastings algorithm reply on the acceptance probability to choose the next sample. All the MCMC algorithms can generate a converged ergodic Markov chain by sampling from the invariant posterior probability.

The Gibbs sampler [18] is a special case of the Metropolis-Hastings algorithm, whose acceptance probability is always one, with better computational efficiency in high dimensional environments. The key idea of the Gibbs sampler is to iteratively update each component of the joint probability by sampling from its univariate conditional probability, given all the other components of the state, without knowing the probability density function (PDF) which can be complex or computationally intractable. In this paper, the full conditional probability of each component is available for the direct Gibbs sampler method. Consider the computational efficiency in 3D modeling, only the standard Gibbs sampler is employed at the current stage. After the stationary point, several converged samples are averaged to estimate any desirable statistics, e.g. the posterior occupancy probability of each voxel.

B. MCMC GIBBS SAMPLER FOR OCCUPANCY MAPPING

To use the MCMC Gibbs sampling for the Occupancy mapping, the occupancy probability for each voxel $P(m_i)$ needs to be calculated, given all the scan measurements $Z_{1:t}$, the corresponding robot poses $X_{1:t}$, and the occupancy probability of all the other voxels in the current Octree map $P(m_{\neg i})$. The new occupancy probability for each voxel is used to condition its next probability distribution. Computing the occupancy probability of each n_m voxels given the occupancy of all the other voxels in the map yields a sample of the full posterior probability. After several iterations, the resultant Markov chain will converge to a stationary distribution.

Please note that in each Gibbs iteration,

$$p(m|z,x) = p(m_i|z,x,m_{\neg i})p(m_{\neg i}|z,x) \tag{5}$$

$p(m_{\neg i}|z,x)$ is apparently invariant because $m_{\neg i}$ are kept the same in each iteration, and each Gibbs iteration samples from the univariate conditional probability $p(m_i|z,x,m_{\neg i})$. Therefore the joint distribution, or the full conditional probability, $p(m|z,x)$ is invariant in each sampling iteration, as well as the whole Gibbs sequence.

Inspired by Thrun's forward sensor model [37], the univariate conditional probability is computed as,

$$\begin{aligned} &P(m_i = 1|Z_t, X_t, m_{\neg i}) \\ &= \frac{P(m_i = 1|X_t, m_{\neg i})P(Z_t|X_t, m_{\neg i}, m_i = 1)}{P(Z_t|X_t, m_{\neg i})} \\ &= \frac{P(m_i = 1|m_{\neg i})\prod_{t=1}^{n_m} P(Z_t|X_t, m_{\neg i}, m_i = 1)}{P(Z_t|X_t, m_{\neg i})}. \end{aligned} \tag{6}$$

The term $P(m_i = 1|m_{\neg i})$ is the prior probability of the voxel being occupied that is actually an empirical constant. The other terms are the forward sensor model $P(Z_t|f_{r,t})$ as described in the full Bayesian solution and must be non-zero for all $Z_{1:t}$ and f. A Gaussian sensor model $\mathcal{N}(f, \sigma^2)$ is chosen specifically.

To compute the probability in the log-odds domain for better efficiency, from (6) its complementary probability can be deduced to be,

$$\begin{aligned} &P(m_i = 0|Z_t, X_t, m_{\neg i}) \\ &= \frac{P(m_i = 0|m_{\neg i})\prod_{t=1}^{n_m} P(Z_t|X_t, m_{\neg i}, m_i = 0)}{P(Z_t|X_t, m_{\neg i})}. \end{aligned} \tag{7}$$

So the log-odds of the conditional distribution is,

$$\begin{aligned} &\log(m_i|Z_t, X_t, m_{\neg i}) \\ &= \log\frac{P(m_i = 1|Z_t, X_t, m_{\neg i})}{P(m_i = 0|Z_t, X_t, m_{\neg i})} \\ &= \log(m_i|m_{\neg i}) \\ &\quad + \sum_{t=1}^{n_m}[\log P(Z_t|X_t, m_{\neg i}, m_i = 1) \\ &\quad - \log P(Z_t|X_t, m_{\neg i}, m_i = 0)]. \end{aligned} \tag{8}$$

The probability $p(m_k = 1|z,x,m_{\neg k})$ can conveniently be recovered from (8).

$$P(m_i = 1|Z_t, X_t, m_{\neg i}) = \frac{\exp(\log(m_i|Z_t, X_t, m_{\neg i}))}{1+\exp(\log(m_i|Z_t, X_t, m_{\neg i}))}, \tag{9}$$

The component $P(Z_t|X_t, m_{\neg i}, m_i)$ in (8) is equivalent to the forward sensor model $P(Z_t|f_{r,t})$ under either the occupied or the free state. The term $\log(m_i|m_{\neg i})$ in (8) is just the prior occupancy probability of each voxel that can be set to zero.

Algorithm 1 illustrates the MCMC Gibbs sampling method based on the Octree map and the formulation of the forward sensor model, which can return multiple samples from the full posterior occupancy probability of each voxel. After some *burn-in* process to discard the initial random samples before convergence, the remaining samples in the Markov chain are useful to compute various statistics about the full Bayesian solution, e.g. the occupancy probability of a voxel can be estimated by the average value over many samples.

V. VOLUMETRIC NEXT BEST VIEW

Compared to the ergodic optimal viewset employed by the model-based NBV, information gain (IG) in term of the probabilistic volumetric information, is the amount of information

Algorithm 1 MCMC Gibbs Sampling to Generate 3D Occupancy Mapping by a Forward Sensor Model

Result: occupancy probability of a particular voxel
Input : given $Z_{1:t}$, $X_{1:t}$, maxSamples; select current Octree as m^0; define m^r as $m^{(0)}$
Output: a desired set of samples from the full posterior probability
initialisation
/* burn-in first few unconverged samples */
for $j = 1$ **to** *maxSamples* **do**
  for $i = 1$ **to** $K \in$ *frustum* **do**
    $\text{elem}(2) = \sum_{n=1}^{N} [\log p(z_n|x_n, m_{\neg i}, m_i = 1)]$
    $\text{elem}(3) = \sum_{n=1}^{N} [\log p(z_n|x_n, m_{\neg i}, m_i = 0)]$
    $l(m_i|z, x, m_{\neg i}) = \text{zero} + \text{elem}(2) - \text{elem}(3)$
    Reset occupancy of m_i to its original value
    if $p(m_i = 1|z, x, m_{\neg i}) > threshold$ **then**
      Estimated posterior = p(occ)
    else
      Estimated posterior = p(free)
    end
    $p(m_i) =$ Estimated posterior
    /* Gibbs sampling full conditional */
    Random sample $m_i^{(j)}$ from $p(m_i)$
    Define m_i^r as $m_i^{(j)}$
  end
end

entropy that each candidate view can provide and thus serves as a metric to rank the Next Best View from the predetermined candidate viewset obtained by sampling redundantly from some blocking geometry like a semi-sphere. Next Best View is the one that can provide the maximum information gain mainly in entropy reduction.

A. ENTROPY BASED INFORMATION GAIN

Let R_v be the bundle of rays cast from the candidate view v from the origin of a vision sensor. The 3D environment is divided and represented by voxels, through which each ray traverses till hitting the first end-point on the surface of the scanning object. Each voxel traversed by each ray has an occupancy probability estimated by the MCMC Gibbs sampler based Occupancy mapping. The estimated information gain for the view v is $\mathcal{G}_v$ and the accumulated volumetric information along all the rays is $\mathcal{I}$, so

$$\mathcal{G}_v = \sum_{\forall r \in \mathcal{R}_v} \sum_{\forall x \in \mathcal{X}} \mathcal{I} \tag{10}$$

The volumetric information within a voxel is encoded as its residual uncertainty of the occupancy probability and is regarded as the information entropy.

$$\mathcal{I}(m_i|Z_{1:t}) = -P_o(m_i|Z_{1:t}) \ln P_o(m_i|Z_{1:t}) - \overline{P}_o(m_i|Z_{1:t}) \ln \overline{P}_o(m_i|Z_{1:t}), \tag{11}$$

where $P_o(m_i|Z_{1:t})$ is the occupancy probability of the voxel x, while $\overline{P}_o(m_i|Z_{1:t}) = 1 - P_o$ is its complement probability. Unknown voxels whose $P_o = 0.5$ have the highest entropy. The information gain of a candidate view is formulated by the changes of entropy in each occupied voxel,

$$\mathcal{I}(Z_{t+1}) = \sum_{i \in \mathcal{O}} [\mathcal{I}(m_i|Z_{1:t}) - \mathcal{I}(m_i|Z_{1:t+1})], \tag{12}$$

where Z_{t+1} is a new scan measurement.

B. NBV UTILITY FUNCTION

Considering the robot movement cost $\mathcal{C}_v$ of the robot positioning system along with the information gain $\mathcal{G}_v$, the object function $\mathcal{U}_v$ to evaluate the NBV for all the left candidate views at the current NBV reconstruction iteration is:

$$\mathcal{U}_v = \frac{\mathcal{G}_v}{\sum_{\mathcal{V}} \mathcal{G}} - \frac{\mathcal{C}_v}{\sum_{\mathcal{V}} \mathcal{C}} \tag{13}$$

The NBV v^* is the one with the maximal value $\mathcal{U}$ among all the candidate views.

$$v^* = \arg\max_{v} \mathcal{U}_v \tag{14}$$

The reconstruction procedure quits when the termination criterion is met, e.g. 20 iterations, or a specific surface coverage ratio, as illustrated in Algorithm 2.

C. MODULAR SYSTEM FRAMEWORK

The MCMC Gibbs NBV algorithm is evaluated within a modular ROS software framework that is generic to different robot platforms and vision sensor modalities. The software architecture is an extension of the work of [15], [47].

The main functional components in 3D reconstruction are listed below as illustrated by Figure 2:

- The 3D Vision Sensor module for point cloud acquisition.
- The Point Cloud Processing module to calculate the depth information.
- The 3D Occupancy Mapping module for point cloud integration, partial Octree generation and IG calculation.
- The NBV Planner module to receive information for the maximisation of the objective function and determine the NBV.
- The Robot Interface is the medium layer between the NBV Planner and the specific robot programming code.
- The Robot Driver translates the data from the Robot Interface for different robot platforms.

The iterative NBV steps mainly occur within the NBV Planner module, including the: the data retrieval from vision sensors; the processing of the depth information from stereo vision, photometry, or structured light; Octomap integration; and view point evaluation, as elaborated in Algorithm 2.

Algorithm 2 Volumetric Model-Free NBV

Data: Point clouds, Octree
Result: Octree map of unknown target
function RayCasting (M, p)
Input : Point cloud patch
Output: Ordered list of NBV evaluated by entropy
initialisation
while *true* **do**
  Path planning for view p
  Robot manipulation to view p
  z ← Take scan measurement
  M ← Update Octree map M by measurement z
  for $i = 1$ **to** n **do**
    if $v_i \in$ *Candidate Viewset* **then**
      RayCasting(M, v_i)
    else
      Delete v_i
    end
  end
  p ← NBV with maximal IG from v_n
  if *p meets terminal criterion* **then**
    Finish reconstruction
  else
    Go back to candidate view evaluation
  end
end

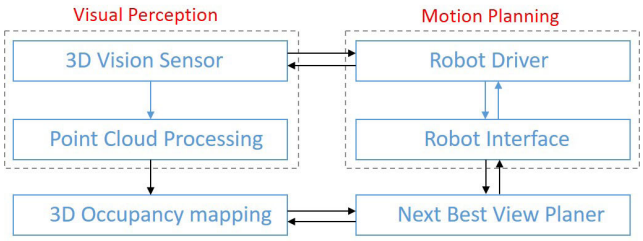


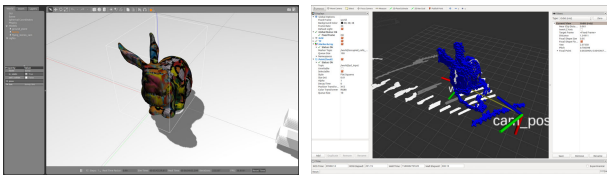
FIGURE 2. System framework: ROS functional modules and communication interfaces.

We validate the performance of the MCMC Gibbs sampling NBV by comparing it to the state of the art conventional NBV algorithms:

- Two independence assumptions based NBV: (a) occlusion aware, unobserved voxel, rear side voxel, rear side entropy, proximity count NBV by Delmerico *et al.* [48]; (b) combined VI weights from learning NBV by Isler *et al.* [15]; (c) average entropy NBV by Kriegel *et al.* [8]; and (d) area factor NBV by Vasquez-Gomez *et al.* [9].
- MCMC Gibbs sampling NBV that generates accurate benchmark 3D Octree maps for surface reconstruction.

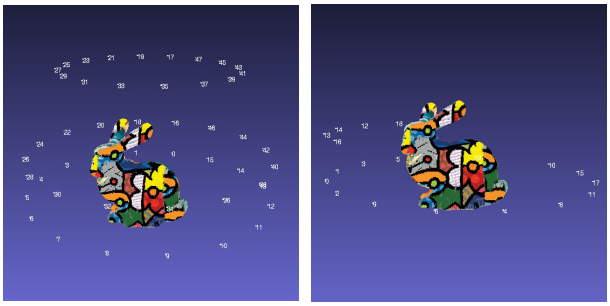
VI. EXPERIMENTS AND RESULTS

The proposed MCMC Gibbs sampler based NBV algorithm is proven to be effective theoretically, but the batch processing property prevents it from running online. At the current stage, the MCMC Gibbs sampler NBV algorithm is implemented by the ROS Gazebo simulation to generate a complete volumetric model of a scanning object that is unknown a priori. The Octree map is probabilistically updated in each Gibbs sampler iteration. The performance of the Gibbs sampler based Occupancy mapping is evaluated by the Kullback-Leibler divergence $D_{KL}(p||q)$ that can quantify the difference between two Occupancy maps. The 3D surface reconstruction is evaluated by the surface coverage between the reconstructed point cloud and the ground truth model. The efficiency of the NBV algorithms is evaluated by the entropy reduce trend within each iteration.



(a) Scanning object and vision sensor in Gazebo simulation (b) Octree map build-up in ROS Rviz

FIGURE 3. ROS gazebo simulation scene.

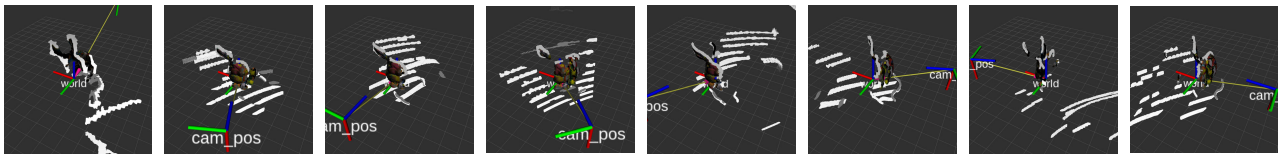


(a) Candidate views: 48 3DOF views facing the scanning object (b) NBV: 20 views selected by volumetric metric

FIGURE 4. Candidate views and selected NBVs.

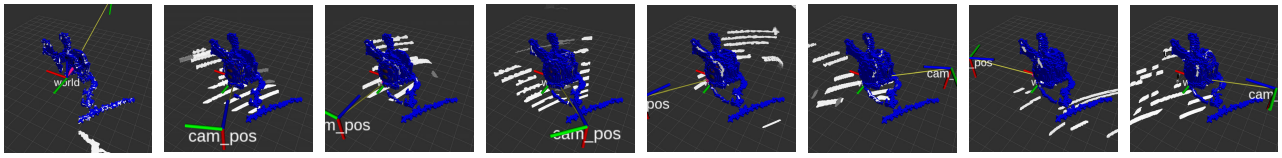
A. EXPERIMENTAL SETUP

The reconstruction simulation is implemented by ROS Gazebo as shown in Figure 3 in an uncluttered static environment. The scanning object is bounded within a 1.0m cube, with a set of 48 candidate views scattered uniformly facing the target frame, reducing the search space from full 6ODF to 3DOF in Figure 4. Please note that the generation of the optimal candidate viewset utilising the classic model-based NBV is beyond the scope of this paper. The eye-in-hand robot is simplified to be an ideal free-flying depth camera with 6DOF, Gaussian noise and 1.5m maximum range which can explore the searching space unconstrainedly, so that the performance of the NBV algorithm is evaluated independently. The on-line NBV evaluation and selection procedure based on the volumetric Octree map is illustrated in Figure 5. The reconstruction starts from a pre-determined view pose. In each NBV iteration, ray casting from each candidate views is used to check the visibility of the scanning object.



(a) Raycasting in NBV iteration 1 (b) Raycasting in NBV iteration 2 (c) Raycasting in NBV iteration 3 (d) Raycasting in NBV iteration 4 (e) Raycasting in NBV iteration 5 (f) Raycasting in NBV iteration 6 (g) Raycasting in NBV iteration 7 (h) Raycasting in NBV iteration 8

(i) Point Cloud in NBV iteration 1 (j) Point Cloud in NBV iteration 2 (k) Point Cloud in NBV iteration 3 (l) Point Cloud in NBV iteration 4 (m) Point Cloud in NBV iteration 5 (n) Point Cloud in NBV iteration 6 (o) Point Cloud in NBV iteration 7 (p) Point Cloud in NBV iteration 8



(q) Octree map in NBV iteration 1 (r) Octree map in NBV iteration 2 (s) Octree map in NBV iteration 3 (t) Octree map in NBV iteration 4 (u) Octree map in NBV iteration 5 (v) Octree map in NBV iteration 6 (w) Octree map in NBV iteration 7 (x) Octree map in NBV iteration 8

FIGURE 5. Volumetric NBV strategy: (a)-(h), ray casting from different candidate views in each NBV iteration to check the visibility of the scanning object. (i)-(p) captured point cloud patches from the next best views in each NBV iteration with post-processing. (q)-(x) local octree maps are generated from the point cloud patches and are registered to a global cartesian occupancy map in each NBV iteration incrementally.

The information gain of each candidate view is calculated based on the occupancy state of each voxel of the current volumetric model. The view with the maximal IG is chosen as the NBV iteratively. The depth visual processing is carried out by the `Pcl_ros`. The PCL module outputs a point cloud with post-processing, e.g. outlier removal, that is used by the Octomap to generate a local Octree in each NBV iteration. The resultant partial Octree maps are incrementally integrated into a single Octree map by global transformation at the resolution of 1cm.

The dataset of the scanning objects including the Stanford bunny, dragon, Armadillo and etc, as well as the multi view stereo (MVS) dataset at DTU as shown in Figure 6 (a)-(f). The MCMC Gibbs sampler NBV is implemented in the ROS-compatible format on a 12-core Intel i7-8700 CPU, 3.2GHz, 16G memory work station which uses 8 parallel threads to accelerate the visibility check by ray casting and the evaluation of candidate views.

The computational efficiency of the full Bayesian solution is exponential in the number of voxels in the map, $\mathcal{O}(2^K)$, whereas the time efficiency of the MCMC Gibbs sampler method is polynomial in the number of voxels, $\mathcal{O}(maxSamples \times K \times N)$. The efficiency of each Gibbs sampling iteration is approximately the time complexity of that of the traditional OGM, $\mathcal{O}(N \times F)$. Here, N stands for the set of scan measurements and F refers to the set of voxels under the coverage of the scan measurement. The simulation takes approximately 4 hours to build the Octree map by updating both the occupied voxels and the free voxels and takes about 40 minutes to build the map by just updating the occupied voxels, whereas the independence assumptions based method takes only about 12 minutes under the same situation. The conventional OGM method allows the map to be updated incrementally with each new scan measurement, and each update is linear in the number of voxels a scan measurement covers, $\mathcal{O}(F)$. Comparably the MCMC sampling method is a batch process that can only be calculated after the Markov chain gets converged. The terminal criterion is set to 20 iterations and it often takes about 10 iterations to become stationary. The timing results are concluded in Table 1.

TABLE 1. Timing results for each volumetric NBV algorithm, averaged over all trials, models, and views in gazebo simulated experiments.

Volumetric NBV algorithms	Avg. time per view [s]
Area Factor (Vasquez-Gomez et al. 2014)	3.7618
Average Entropy (Kriegel et al. 2015)	3.7714
Proximity Count (Isler et al. 2016)	5.2671
Combined VI (Isler et al. 2016)	5.2137
Occlusion Aware (Delmerico et al. 2018)	3.7798
Unobserved Voxel (Delmerico et al. 2018)	3.7786
Rear Side Voxel (Delmerico et al. 2018)	3.7709
Rear Side Entropy (Delmerico et al. 2018)	3.7889
MCMC Gibbs Sampler	144.42

It is worthwhile to note that although the two independence assumptions based NBV has good computational efficiency for online implementation, it cannot generate accurate maps for robot applications, whereas the gist of the MCMC Gibbs sampler based NBV is to produce benchmark maps for surface reconstruction, which will later be used to optimise the tuning parameters of the inverse sensor model of the

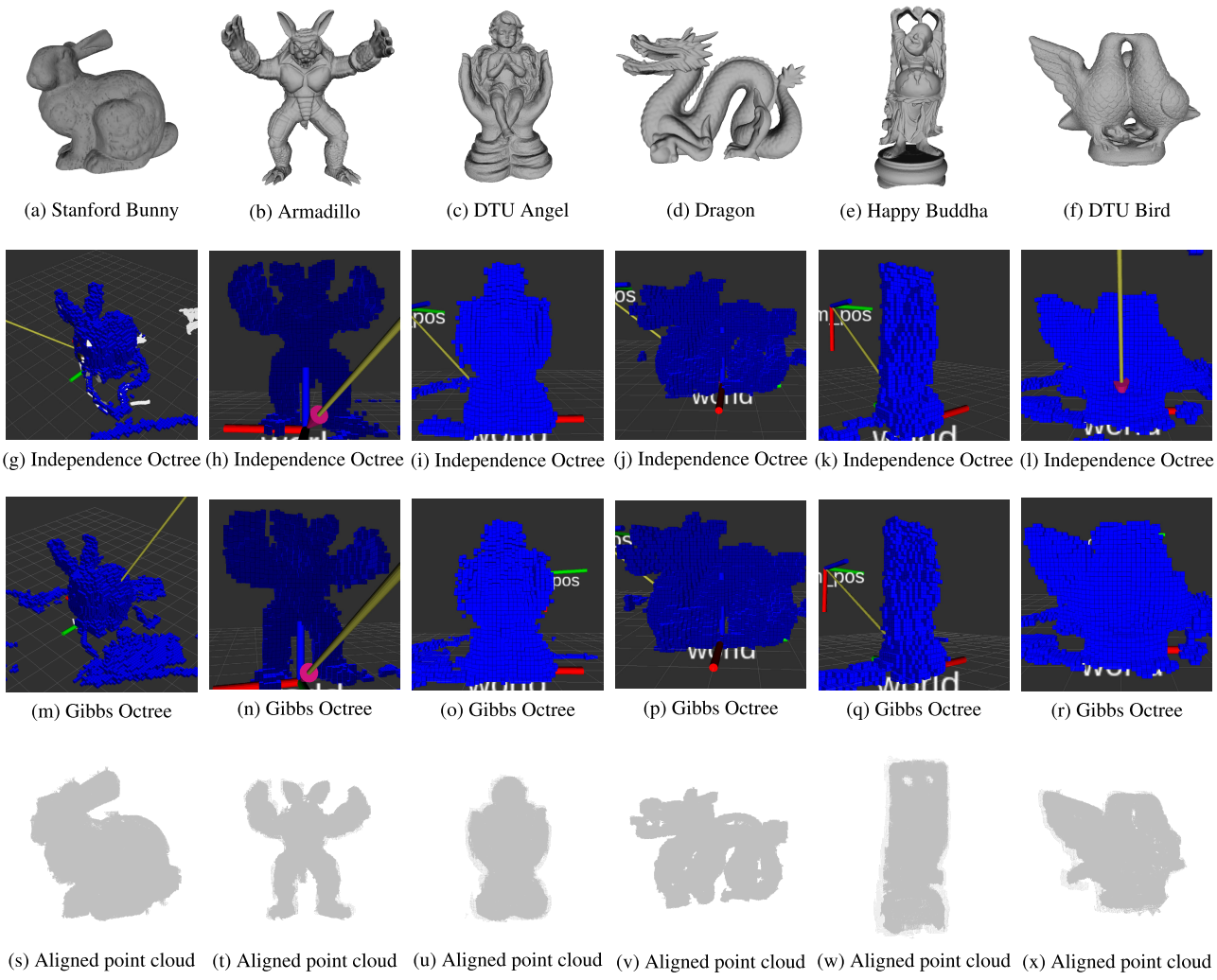


FIGURE 6. Reconstruction results by ROS gazebo:(a)-(f) synthetic model datasets. (g)-(l) octree maps by conventional NBV methods, e.g. proximity count. (m)-(r) smoothed octree maps by MCMC gibbss sampler method. (s)-(x) Globally aligned point clouds with gaussian white noise.

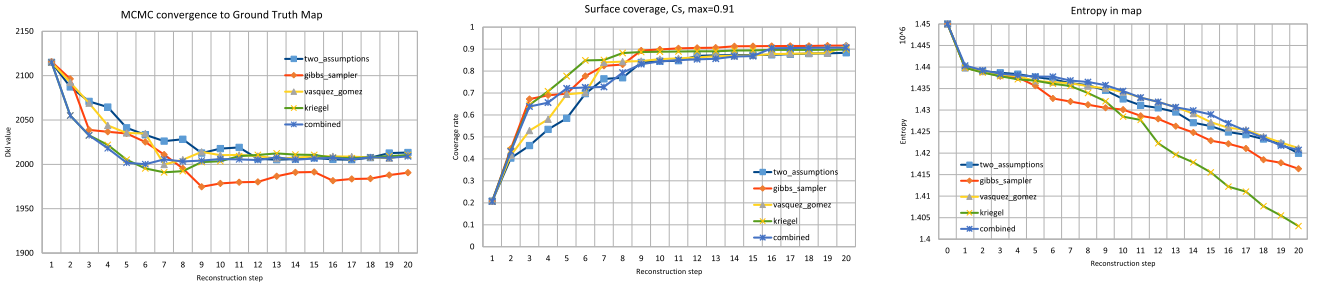
two independence assumptions based NBV by minimising the cross entropy between two maps by the Quasi-Newton method.

B. EVALUATION

The visibility is checked by casting rays on the object surface to update the volumetric map. The captured partial point cloud is processed, transformed and integrated into the global Octree map by the Octomap library. The resultant Octree maps by the two assumptions NBV of different models are shown in Figure 6 (g)-(l) but they suffer from incompleteness especially on boundaries in the final maps due to the inaccurate estimation of the residual uncertainty within each voxel such that the narrow openings are incorrectly regarded as closed and are left unattended if magnified. This undesirable phenomenon is even worse when the FOV of the high-end scanners become narrower to provide high-quality dense point clouds. By applying the Gibbs sampler smoothing framework, to the Occupancy mapping, the resultant Octree maps are distinctly complete and accurate under equivalent conditions, as shown in Figure 6 (m)-(r). Therefore the human intervene in industrial automation can be disregarded given the satisfactory complete reconstruction. The final point cloud is reconstructed by globally aligning the partial point clouds during each iteration, by assuming that there is no unexpected position error or camera calibration error in the ideal simulation environment, as shown in Figure 6 (s)-(x). Intuitively, the MCMC Gibbs sampler based NBV can provide more accurately volumetric maps, as well as more complete point clouds, specifically for high-end industrial 3D scanners with high resolution and limited frustum, e.g. the ATOS GOM and the Artec Eva.

C. KULLBACK-LEIBLER DIVERGENCE

To quantitatively show the difference between the probability distributions of the traditional Occupancy mapping methods and the Gibbs sampler Occupancy mapping method, the Kullback-Leibler divergence $D_{KL}(p\|q)$ is used as the



(a) Kullback-Leibler divergence (b) Surface coverage (c) Entropy reduce

FIGURE 7. Evaluation results of the MCMC gibbs sampler NBV compared to the state of the art NBV algorithms by isler, kriegel and vasquez-gomez: (a) the kullback-leibler divergence (D_{KL}) quantifies the difference between two probability distributions. (b) the surface coverage of the scanning objects. (c) the entropy reduce trend in each NBV iteration.

metric [49] , where p is the probability estimated by either the Gibbs sampler method or the traditional methods, while q is the probability of the ground truth map. The $D_{KL}(p_k \| q_k)$ is computed for each voxel m_k,

$$\begin{aligned} &\mathrm{D_{KL}}(p_k \| q_k) \\ &= \sum_{m_k=0}^{1} p(m_k) \ln \frac{p(m_k)}{q(m_k)} \\ &= \sum_{m_k=0}^{1} \ln\left(\frac{p(m_k)}{q(m_k)}\right) p(m_k) + \ln\left(\frac{1-p(m_k)}{1-q(m_k)}\right)(1-p(m_k)), \end{aligned} \tag{15}$$

and the sum over all the voxels is the D_{KL} between two Occupancy maps as shown in Figure 7 (a).

$$D_{KL}(p\|q) = \sum_{k=1}^{K} D_{KL}(p_k \| q_k) \tag{16}$$

The MCMC method will gradually converge around 10 iterations under the current setting. The Gibbs sampler method produces the lowest Kullback-Leibler divergence $D_{KL}(p\|q)$ as expected w.r.t the ground truth Octree map, compared to all the other faked two independence assumptions methods.

D. SURFACE COVERAGE

To quantify the 3D reconstruction performance in term of the surface coverage, the point cloud obtained in 3D reconstruction is compared with the point cloud generated from the ground truth dataset. For each point in the ground truth model, the closest point in the reconstructed point cloud is queried. If the point is closer than a registration threshold, e.g. $d_{reg} = 0.05cm$, the surface point of the original model is believed to have been observed. The surface coverage C_s is the percentage of the observed surface points, Pt_o, divided by the total number of surface points of the model, Pt_g:

$$\text{Surface coverage } C_s = \frac{\text{Observed surface points } Pt_o}{\text{Ground truth model points } Pt_g}. \tag{17}$$

The result in the current configuration is illustrated in Figure 7 (b) w.r.t different NBV algorithms and models. It is believed that the coverage ratio can be further improved by generating more accurate candidate viewsets by the Coverage Planning algorithms.

E. ENTROPY REDUCTION

To demonstrate the entropy decrease in each NBV iteration, a virtual bounding cube with the length of 1.28m around the scanning object is used to define the initial entropy as,

$$\text{Entropy in map} = \sum \text{Entropy of voxels within cube.} \tag{18}$$

The scanning object in the Gazebo simulation scene is adjusted to be approximately 0.5m in length within a 1m bounding box. The Octomap resolution is set to be 1cm so a cube with $(1.28 = 0.01 \times 2^7)$m length is the smallest initial cube for the Octree that can completely cover the reconstructed volumetric Occupancy map of the target. The maximal entropy within the bounding cube is when each voxel is initialised as unknown, with a default occupancy likelihood of $P(\text{unknown}) = 0.5$, therefore the initial maximal entropy is $-(2^7)^3 \log_2 0.5 = 2.097 \times 10^6$ Shannon. For the NBV exploration, only the maximum-likelihood Octree map containing either occupied or free voxels is evaluated for entropy reduce, where $p(free) = 0$ and $p(occ) = 1$. The reduce trend of entropy is illustrated in Figure 7 (c). The MCMC Gibbs sampler based NBV can select candidate views that more significantly decrease the environment uncertainty by estimating the residual uncertainty more accurately, compared to most of its conventional NBV counterparts.

Although the average entropy formulation by Kriegel *et al.* is the most effective at reducing the entropy in the map, the total entropy reduction in the map is not straightforwardly correlated to the accuracy of surface reconstruction because the possible occlusion is ignored, the free voxel update carries no new information and the same voxel may be observed multiple times to update its occupancy probability, not to mention in robot map applications the map accuracy and the

surface completeness tell more information about the real environment.

It is worthwhile noting that the classic model-based NBV algorithms that choose random views from a predetermined optimal viewset may be computationally less expensive than evaluating all the remaining candidate views in each NBV iteration. However, the performance is variant from time to time depending on the specific algorithms to sample the viewset.

VII. CONCLUSION AND FUTURE WORK

Although the two independence assumptions based NBV algorithms remain popular for decades, they are over-confident to generate true Occupancy maps of unknown environments compared to the ground truth map. We propose an MCMC Gibbs sampling algorithm for accurate 3D reconstruction, especially for precise manufacturing, which is used to sample from the complete posterior probability of the volumetric model for 3D reconstruction in polynomial time scale. The MCMC Gibbs sampler NBV can provide benchmark volumetric Occupancy maps along with accurate point clouds for further applications.

Future work includes evaluating the MCMC Gibbs sampler solution directly in 3D to optimise the parameters of the update terms within the inverse sensor model for online NBV deployed in industry automation. The optimised update term in the conventional two independence assumptions methods will better capture the residual uncertainty in 3D Occupancy maps, by the BFGS minimisation algorithm. Secondly, the voxel dilation and the Gaussian sampling will be used to generate near-optimal candidate view-sets to cover the surface of the scanning object with a higher coverage ratio. To improve the efficiency of visibility evaluation, hierarchical ray casting will be employed in the camera sampling process. Furthermore, by applying the *reinforcement learning*, specifically the Monte Carlo Tree Searching (MCTS), to dynamically generate the Next Best View in real-time, the pre-determined candidate viewset can be eliminated. Therefore the working range of the reconstruction system can be expanded to larger objects, e.g. drones for buildings. The current MCMC method ignores the information contained in the previous scan measurements, while the deep learning algorithm can use the previous information to help to generate the next candidate view. Reinforcement learning is expected to speed up the MCMC algorithm by an order of magnitude without the repeated evaluation.

ACKNOWLEDGMENT

The authors would like to thank the associate editor and the reviewers for their pertinent suggestions. Special thanks to Dr. J. Delmerico, Mr. S. Isler, and Dr. Rehman S. Merali for their technical support.

REFERENCES

[1] J. Aloimonos, I. Weiss, and A. Bandyopadhyay, "Active vision," *Int. J. Comput. Vis.*, vol. 1, no. 4, pp. 333–356, 1988. doi: 10.1007/BF00133571.

[2] S. Chen, Y. Li, and N. M. Kwok, "Active vision in robotic systems: A survey of recent developments," *Int. J. Robot. Res.*, vol. 30, no. 11, pp. 1343–1377, 2011. doi: 10.1177/0278364911410755.

[3] W. R. Scott, "Model-based view planning," *Mach. Vis. Appl.*, vol. 20, no. 1, pp. 47–69, Jan. 2009. doi: 10.1007/s00138-007-0110-2.

[4] W. Jing, J. Polden, W. Lin, and K. Shimada, "Sampling-based view planning for 3D visual coverage task with unmanned aerial vehicle," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2016, pp. 1808–1815.

[5] K. A. Tarabanis, R. Y. Tsai, and P. K. Allen, "The MVP sensor planning system for robotic vision tasks," *IEEE Trans. Robot. Autom.*, vol. 11, no. 1, pp. 72–85, Feb. 1995.

[6] W. R. Scott, G. Roth, and J.-F. Rivest, "View planning for automated three-dimensional object reconstruction and inspection," *ACM Comput. Surv.*, vol. 35, no. 1, pp. 64–96, Mar. 2003. doi: 10.1145/641865.641868.

[7] M. Quigley, K. Conley, B. P. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, "ROS: An open-source robot operating system," in *Proc. ICRA Workshop Open Source Softw.*, vol. 3, May 2009, p. 5.

[8] S. Kriegel, C. Rink, T. Bodenmüller, and M. Suppa, "Efficient next-best-scan planning for autonomous 3D surface reconstruction of unknown objects," *J. Real-Time Image Process.*, vol. 10, no. 4, pp. 611–631, Dec. 2015. doi: 10.1007/s11554-013-0386-6.

[9] J. I. Vasquez-Gomez, L. E. Sucar, R. Murrieta-Cid, and E. Lopez-Damian, "Volumetric next-best-view planning for 3D object reconstruction with positioning error," *Int. J. Adv. Robotic Syst.*, vol. 11, no. 10, p. 159, Oct. 2014. doi: 10.5772/58759.

[10] H. Moravec and A. Elfes, "High resolution maps from wide angle sonar," in *Proc. IEEE Int. Conf. Robot. Autom.*, vol. 2, Mar. 1985, pp. 116–121.

[11] H. P. Moravec, "Sensor fusion in certainty grids for mobile robots," *AI Mag.*, vol. 9, no. 2, pp. 61–74, 1988.

[12] J. I. Vasquez-Gomez, L. E. Sucar, and R. Murrieta-Cid, "Hierarchical ray tracing for fast volumetric next-best-view planning," in *Proc. Int. Conf. Comput. Robot Vis.*, May 2013, pp. 181–187.

[13] F. Bissmarck, M. Svensson, and G. Tolt, "Efficient algorithms for next best view evaluation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep./Oct. 2015, pp. 5876–5883.

[14] S. Thrun, W. Burgard, R. C. Arkin, and D. Fox, *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents Series)*. Cambridge, MA, USA: MIT Press, 2005.

[15] S. Isler, R. Sabzevari, J. Delmerico, and D. Scaramuzza, "An information gain formulation for active volumetric 3D reconstruction," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2016, pp. 3477–3484.

[16] J. Daudelin and M. Campbell, "An adaptable, probabilistic, next-best view algorithm for reconstruction of unknown 3-D objects," *IEEE Robot. Autom. Lett.*, vol. 2, no. 3, pp. 1540–1547, Jul. 2017.

[17] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984. doi: 10.1109/TPAMI.1984.4767596.

[18] G. Casella and E. I. George, "Explaining the Gibbs sampler," *Amer. Statist.*, vol. 46, no. 3, pp. 167–174, 1992. doi: 10.2307/2685208.

[19] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "OctoMap: An efficient probabilistic 3D mapping framework based on octrees," *Auto. Robots*, vol. 34, no. 3, pp. 189–206, Apr. 2013. [Online]. Available: http://octomap.github.com

[20] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *J. Chem. Phys.* vol. 21, no. 6, pp. 1087–1092, 1953. [Online]. Available: http://link.aip.org/link/?JCP/21/1087/1

[21] N. Metropolis and S. Ulam, "The Monte Carlo method," *J. Amer. Stat. Assoc.*, vol. 44, no. 247, pp. 335–341, Sep. 1949. [Online]. Available: http://www.jstor.org/stable/2280232

[22] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, no. 1, pp. 97–109, Apr. 1970.

[23] R. A. Levine, Z. Yu, W. G. Hanley, and J. J. Nitao, "Implementing componentwise Hastings algorithms," *Comput. Statist. Data Anal.*, vol. 48, no. 2, pp. 363–389, Feb. 2005.

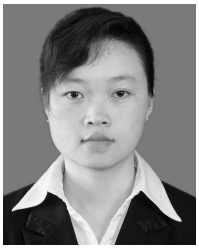
[24] K. R. Koch, "Gibbs sampler by sampling-importance-resampling," *J. Geodesy*, vol. 81, no. 9, pp. 581–591, Sep. 2007. doi: 10.1007/s00190-006-0121-1.

[25] C. Ritter and M. A. Tanner, "Facilitating the Gibbs sampler: The Gibbs stopper and the griddy-Gibbs sampler," *J. Amer. Stat. Assoc.*, vol. 87, no. 419, pp. 861–868, Sep. 1992. [Online]. Available: http://www.jstor.org/stable/2290225

[26] R. Bajcsy, "Active perception," *Proc. IEEE*, vol. 76, no. 8, pp. 966–1005, Aug. 1988.
[27] A. Blake and A. Yuille, Eds., *Active Vision*. Cambridge, MA, USA: MIT Press, 1993.
[28] W. Jing and K. Shimada, "Model-based view planning for building inspection and surveillance using voxel dilation, medial objects, and random-key genetic algorithm," *J. Comput. Des. Eng.*, vol. 5, no. 3, pp. 337–347, Jul. 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2288430017300866
[29] B. Yamauchi, "A frontier-based approach for autonomous exploration," in *Proc. IEEE Int. Symp. Comput. Intell. Robot. Automat.*, Jul. 1997, pp. 146–151.
[30] P. Quin, G. Paul, and D. Liu, "Experimental evaluation of nearest neighbor exploration approach in field environments," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 2, pp. 869–880, Apr. 2017.
[31] R. Monica and J. Aleotti, "Surfel-based next best view planning," *IEEE Robot. Autom. Lett.*, vol. 3, no. 4, pp. 3324–3331, Oct. 2018.
[32] Y. F. Li and Z. G. Liu, "Information entropy-based viewpoint planning for 3-D object reconstruction," *IEEE Trans. Robot.*, vol. 21, no. 3, pp. 324–337, Jun. 2005.
[33] R. S. Merali and T. D. Barfoot, "Patch MAP: A benchmark for occupancy grid algorithm evaluation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2012, pp. 3481–3488.
[34] R. S. Merali and T. D. Barfoot, "Occupancy grid mapping with Markov chain Monte Carlo Gibbs sampling," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2013, pp. 3183–3189.
[35] R. S. Merali and T. D. Barfoot, "Optimizing online occupancy grid mapping to capture the residual uncertainty," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May/Jun. 2014, pp. 6070–6076.
[36] V. Dhiman, A. Kundu, F. Dellaert, and J. J. Corso, "Modern MAP inference methods for accurate and fast occupancy grid mapping on higher order factor graphs," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May/Jun. 2014, pp. 2037–2044.
[37] S. Thrun, "Learning occupancy grid maps with forward sensor models," *Auto. Robots*, vol. 15, no. 2, pp. 111–127, Sep. 2003. doi: 10.1023/A:1025584807625.
[38] E. Kaufman, T. Lee, Z. Ai, and I. S. Moskowitz, "Bayesian occupancy grid mapping via an exact inverse sensor model," in *Proc. Amer. Control Conf. (ACC)*, Jul. 2016, pp. 5709–5715.
[39] D. Murray and J. J. Little, "Using real-time stereo vision for mobile robot navigation," *Auto. Robots*, vol. 8, no. 2, pp. 161–171, 2000. doi: 10.1023/A:1008987612352.
[40] C. Potthast and G. S. Sukhatme, "A probabilistic framework for next best view estimation in a cluttered environment," *J. Vis. Commun. Image Represent.*, vol. 25, no. 1, pp. 148–164, Jan. 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1047320313001387
[41] A. E. Gelfand and A. F. Smith, "Sampling-based approaches to calculating marginal densities," *J. Amer. Statist. Assoc.*, vol. 85, no. 410, pp. 398–409, 1990. [Online]. Available: http://www.jstor.org/stable/2289776
[42] S. Geman and D. E. McClure, "Statistical methods for tomographic image reconstruction," *Bull. Int. Stat. Inst.*, vol. 4, pp. 5–21, 1987.
[43] W. R. Gilks and P. Wild, "Adaptive rejection sampling for Gibbs sampling," *J. Roy. Stat. Soc. Ser. C, Appl. Statist.*, vol. 41, no. 2, pp. 337–348, Jun. 1992. [Online]. Available: http://www.jstor.org/stable/2347565
[44] L. Martino, V. Elvira, and G. Camps-Valls, "The recycling Gibbs sampler for efficient learning," *Digit. Signal Process.*, vol. 74, pp. 1–13, Mar. 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1051200417302750
[45] K. Pathak, A. Birk, J. Poppinga, and S. Schwertfeger, "3D forward sensor modeling and application to occupancy grid based sensor fusion," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct./Nov. 2007, pp. 2059–2064.
[46] W. Jing, C. F. Goh, M. Rajaraman, F. Gao, S. Park, Y. Liu, and K. Shimada, "A computational framework for automatic online path generation of robotic inspection tasks via coverage planning and reinforcement learning," *IEEE Access*, vol. 6, pp. 54854–54864, 2018.
[47] L. Torabi and K. Gupta, "An autonomous six-DOF eye-in-hand system for *in situ* 3D object modeling," *Int. J. Robot. Res.*, vol. 31, no. 1, pp. 82–100, Jan. 2012. doi: 10.1177/0278364911425836.
[48] J. Delmerico, S. Isler, R. Sabzevari, and D. Scaramuzza, "A comparison of volumetric information gain metrics for active 3D object reconstruction," *Auto. Robots*, vol. 42, no. 2, pp. 197–208, 2018. doi: 10.1007/s10514-017-9634-0.
[49] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Statist.*, vol. 22, no. 1, pp. 79–86, 1951. doi: 10.1214/aoms/1177729694.

LEI (HELEN) HOU received the degree in computer science from the China University of Mining and Technology, Beijing, China, in 2005, the master's degree from Tohoku University, Japan, and the Ph.D. degree from Tohoku University, in 2011. She is currently a Researcher with the Queensland University of Technology (QUT). Her research interests include computer vision and control engineering. She also developed skills in robotics.

XIAOPENG CHEN received the bachelor's degree in automation from the University of Science and Technology of China, in 2003, and the doctors degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Sciences, in 2008. Since 2008, he has been a Faculty Member with the Beijing Institute of Technology. He worked together with Prof. C. G. Atkeson as a Visiting Scholar with Carnegie Mellon University, from 2013 to 2014. He is currently an Associate Professor with the Beijing Institute of Technology. His research interests include robot vision and visual servo.



KUNYAN LAN graduated in electronic information engineering from the China University of Geosciences, Beijing, in 2007, and received the master's degree in pattern recognition and intelligent systems from the Beijing University of Posts and Telecommunications, in 2010. Her current research interest includes image processing.

RUNE RASMUSSEN joined the Institute of Future Environments (IFE), QUT, in 2017, as a Senior Software Engineer for the Research Engineering Facility. His interests are in solving and finding high-performing software solutions to problems in complex systems, which, in his current role, concerns solving software problems in robotics and autonomous systems. His current role allows him to specialize in systems involving the robotic operating systems (ROSs).

JONATHAN ROBERTS received the Honours degree in aerospace systems engineering and the Ph.D. degree from the University of Southampton, U.K., in 1991 and 1994, respectively. He was appointed as a Science Leader of Robotics with CSIRO, where he also led the Autonomous Systems Laboratory as a Research Director. He is a Chief Investigator for the Australian Centre for Robotic Vision and is currently a Professor in robotics with the Queensland University of Technology (QUT). His main research interests include field robotics and industry robotics.