TARGET DETECTION FOR DEPTH IMAGING USING SPARSE SINGLE-PHOTON DATA

Yoann Altmann, Ximing Ren, Aongus McCarthy, Gerald S. Buller and Steve McLaughlin†*

Heriot-Watt University, School of Engineering and Physical Sciences, Edinburgh, United Kingdom.
{*Y.Altmann;xr5;A.McCarthy; G.S.Buller;S.McLaughlin*}*@hw.ac.uk*

ABSTRACT

This paper presents a new Bayesian model and associated algorithm for depth and intensity profiling using full waveforms from time-correlated single-photon counting (TCSPC) measurements when the photon count in very low. The model represents each Lidar waveform as an unknown constant background level, which is combined in the presence of a target, to a known impulse response weighted by the target intensity and finally corrupted by Poisson noise. The joint target detection and depth imaging problem is expressed as a pixel-wise model selection problem which is solved using Bayesian inference. A Reversible Jump Markov chain Monte Carlo algorithm is proposed to compute the Bayesian estimates of interest. Finally, the benefits of the methodology are demonstrated through a series of experiments using real data.

Index Terms— Full waveform Lidar, Poisson statistics, Bayesian estimation, Reversible Jump Markov Chain Monte Carlo, depth imaging.

1. INTRODUCTION

Time-of-flight laser ranging (Lidar) systems can be used to reconstruct 3-dimensional scenes in many applications, including automotive [1, 2, 3, 4], environmental sciences [5, 6], architectural engineering and defence [7, 8]. This challenging problem consists of illuminating the scene with a train of laser pulses and analysing the distribution of the photons received by the detector to infer the presence of objects as well as their range, and radiative properties (e.g., reflectivity, observation conditions,...). Using optical scanning systems, a histogram of time delays between the emitted pulses and the detected photon arrivals is usually recorded for each pixel, associated with a different region of the scene. In the presence of objects, the recorded photon histograms are classically decomposed into a series of peaks whose positions can be used to infer the distance of the objects present in each region of the scene and whose amplitudes provide information about the reflectance of the objects. In this paper, we investigate the target detection problem which consists of inferring the regions or pixels of the scene where objects are present. Moreover, we propose an algorithm for applications where the flux of detected photons is small and for which classical depth imaging methods [9] usually provide unsatisfactory results in terms of range and intensity estimation. In Lidar systems there is a trade-off between range, acquisition time and laser power levels. In this paper, we examine the sparse photon regime which occurs at the extremes of the trade-off, for example at very long ranges, and/or at very short acquisition times (e.g., for applications where very fast imaging is required (dynamic scene and/or sensor)). In this work, we assume that the targets in the scene of interest are composed of a single solid surface in each pixel. As in [10, 11], we consider the potential presence of two kinds of detector events: the photons originating from the illumination laser and scattered back from the target (if present); and the background detector events originating from ambient light and the "dark" events resulting from detector noise. The proposed method aims to estimate the respective contributions of the actual target (if any) and the background in the photon timing histograms. Following a classical Bayesian approach, as in [12, 13, 11], we express the target detection and identification problem as a pixel-wise model selection and estimation problem. More precisely, two observation models, conditioned on the presence or absence of a target (modelled by binary labels) are considered for each pixel. We then assign prior distributions to each model unknown parameters to include available information within the estimation procedure. The binary labels associated with the presence/absence of target are also assigned prior distributions accounting for spatial correlations. The classical Bayesian estimators associated with the resulting joint posterior cannot be easily computed due to the complexity of the model, in particular because the number of underlying parameters (number of pixels containing a target) is unknown and potentially large. To tackle this problem, a Reversible-Jump Markov chain Monte Carlo (RJ-MCMC) [14, 15] method is used to generate samples according to this posterior by allowing moves between different parameter spaces. More precisely, we construct an efficient RJ-MCMC algorithm that simultaneously infers the target presence and estimates the background levels and the target (if present) distances and intensity.

The remainder of this paper is organized as follows. Section 2 recalls the statistical models used for depth imaging using time-of-flight scanning sensors, based on TCSPC measurements. Section 3 presents the proposed hierarchical Bayesian model which accounts for the inherent spatial correlations between pixels. Section 4 briefly discusses the estimation of the model parameters including the detection labels using simulation methods. Simulation results conducted using an actual time-of-flight scanning sensor are presented and discussed in Section 5. Finally, conclusions and potential future work are reported in Section 6.

2. PROBLEM FORMULATION

We consider a set of $N_{\text{row}} \times N_{\text{col}}$ observed Lidar waveforms/pixels $\mathbf{y}_{i,j} = [y_{i,j,1}, \ldots, y_{i,j,T}]^T, (i,j) \in \{1, \ldots, N_{\text{row}}\} \times \{1, \ldots, N_{\text{col}}\}$ where T is the number of temporal (corresponding to range) bins. To be precise, $y_{i,j,t}$ is the photon count within the tth bin of the pixel or location (i,j). Let $z_{i,j} \in \{0,1\}$ be a binary variable associated with the presence ($z_{i,j} = 1$) or absence ($z_{i,j} = 0$) of target in the pixel (i,j). Conditioned on a target presence/absence, $y_{i,j,t}$ is assumed to

*Part of this work was supported by the EPSRC via grants EP/N003446/1, EP/K015338/1, EP/M01326X/1

†Part of this work was supported by the EPSRC via grant EP/J015180/1.

be drawn from one of the following Poisson distribution

$$y_{i,j,t}|z_{i,j}=0,\boldsymbol{\theta}^0_{i,j} \sim \mathcal{P}\left(b_{i,j}\right) \quad (1a)$$

$$y_{i,j,t}|z_{i,j}=1,\boldsymbol{\theta}^1_{i,j} \sim \mathcal{P}\left(r_{i,j}g_0\left(t-t_{i,j}\right)+b_{i,j}\right) \quad (1b)$$

where $b_{i,j}>0$ stands for the background and dark photon level (constant in all bins of a given pixel), $t_{i,j}$ is the position of an object (if present) at a given range from the sensor and $r_{i,j}$ its intensity. Note that the parameters of the models $\mathcal{M}_0^{(i,j)}$ and $\mathcal{M}_1^{(i,j)}$ in (1a) and (1b) are denoted by $\boldsymbol{\theta}^0_{i,j}=b_{i,j}\in\mathbb{R}^+$ and $\boldsymbol{\theta}^1_{i,j}=[r_{i,j},t_{i,j},b_{i,j}]\in\mathbb{R}^+\times\mathbb{T}\times\mathbb{R}^+$, respectively, with $\mathbb{T}=\{1,\ldots,T\}$. Moreover, $g_0(\cdot)>0$ in (1b) stands for the photon (instrumental) impulse response, which is assumed to be known (it can be estimated during the imaging system calibration)

The problem addressed in this paper consists of deciding, from the observed data gathered in the $N_{\text{row}}\times N_{\text{col}}\times T$ array $\mathbf{Y}$, whether a target is present ($\mathcal{M}_1^{(i,j)}$) or not ($\mathcal{M}_0^{(i,j)}$) in each pixel and of estimating the position and intensity of the targets (if any) present in the scene. Moreover, the background levels $b_{i,j}$ are also assumed to be unknown and need to be estimated. This problem can be seen as a pixel-wise model selection problem where the parameter space associated with each model is different. Precisely, under $\mathcal{M}_0^{(i,j)}$ (resp. $\mathcal{M}_1^{(i,j)}$), $\boldsymbol{\theta}^{(0)}_{i,j}\in\boldsymbol{\Theta}_0$ (resp. $\boldsymbol{\theta}^{(1)}_{i,j}\in\boldsymbol{\Theta}_1$) where $\boldsymbol{\Theta}_0=\mathbb{R}^+$ and $\boldsymbol{\Theta}_1=\mathbb{R}^+\times\mathbb{T}\times\boldsymbol{\Theta}_0$. To simplify notations, the unknown parameter vector for each pixel is noted $\boldsymbol{\theta}_{i,j}$ in the remainder of the paper when we do not specify whether it is included in $\boldsymbol{\Theta}_0$ or $\boldsymbol{\Theta}_1$. Estimating $\boldsymbol{\theta}_{i,j}$ is difficult using standard optimization methods since the dimensionality of the parameter vector depends on the underlying model. However, this model selection problem can be solved efficiently in a Bayesian framework by 1) performing inference for each pixel in the parameter space parameter space $\{\{0\}\times\boldsymbol{\Theta}_0\}\bigcup\{\{1\}\times\boldsymbol{\Theta}_1\}$, 2) incorporating relevant additional prior belief (through prior distributions) and 3) using RJ-MCMC methods adapted for problems whose finite dimensionality in unknown.

The next section presents the proposed Bayesian model and associated RJ-MCMC sampling strategy enabling cooperative target detection for depth imaging using single-photon data.

3. PROPOSED BAYESIAN MODEL

3.1. Parameter prior distributions

It has been shown in [11] that considering gamma priors for the background levels (under the two observation models) simplifies the sampling procedure by exploiting the likelihood and priors conjugacy. Moreover, due to the spatial organization of images, we expect the values of $b_{i,j}$ to vary smoothly from one pixel to another. In order to model this behaviour, we specify the background levels prior such that the resulting prior for the background matrix $\mathbf{B}$ such that $[\mathbf{B}]_{i,j}=b_{i,j}$ is a hidden gamma-MRF (GMRF) [16]. More precisely, we introduce an $(N_{\text{row}}+1)\times(N_{\text{col}}+1)$ auxiliary matrix $\mathbf{\Gamma}$ with elements $\gamma_{i,j}\in\mathbb{R}^+$ and define a bipartite conditional independence graph between $\mathbf{B}$ and $\mathbf{\Gamma}$ such that each $b_{i,j}$ is connected to four neighbour elements of $\mathbf{\Gamma}$ and vice-versa. We specify a GMRF prior for $\mathbf{B},\mathbf{\Gamma}$ [16], and obtain the following joint prior for $\mathbf{B},\mathbf{\Gamma}$

$$f(\mathbf{B},\mathbf{\Gamma}|\nu)=\prod_{((i,j),(i',j'))\in\mathcal{E}}\exp\left(\frac{-\nu b_{i,j}}{4\gamma_{i',j'}}\right)$$

$$\times\left[\frac{1}{G(\nu)}\prod_{(i,j)\in\mathcal{V}_{\mathbf{B}}}b_{i,j}^{(\nu-1)}\right]\times\left[\prod_{(i',j')\in\mathcal{V}_{\mathbf{\Gamma}}}\left(\gamma_{i',j'}\right)^{-(\nu+1)}\right] \quad (2)$$

where $\mathcal{V}_{\mathbf{\Gamma}}=\{1,\ldots,N_{\text{row}}+1\}\times\{1,\ldots,N_{\text{col}}+1\}$, $\mathcal{V}_{\mathbf{B}}=\{1,\ldots,N_{\text{row}}\}\times\{1,\ldots,N_{\text{col}}\}$, and the edge set $\mathcal{E}$ consists of pairs $((i,j),(i',j'))$ representing the connection between $b_{i,j}$ and $\gamma_{i',j'}$. It can be seen from (2) that $b_{i,j}|\mathbf{\Gamma},\nu\sim\mathcal{G}\left(\nu,\dfrac{\epsilon_{i,j}(\mathbf{\Gamma})}{\nu}\right)$ and

$$\gamma_{i,j}|\mathbf{B},\nu \sim \mathcal{IG}\left(\nu,\nu\xi_{i,j}(\mathbf{B})\right) \quad (3)$$

where $\epsilon_{i,j}(\mathbf{\Gamma})=4\left(\gamma_{i,j}^{-1}+\gamma_{i-1,j}^{-1}+\gamma_{i,j-1}^{-1}+\gamma_{i-1,j-1}^{-1}\right)^{-1}$ and $\xi_{i,j}(\mathbf{B})=\left(b_{i,j}+b_{i+1,j}+b_{i,j+1}+b_{i+1,j+1}\right)/4$. Notice that we denote explicitly the dependence on the value of ν, which here acts a regularization parameter that controls the amount of spatial smoothness enforced by the GMRF. For brevity, in this paper ν is assumed to be fixed but can be adjusted automatically during the inference procedure as in [11].

To reflect the absence of prior knowledge about the target ranges given $z_{i,j}=1$, we assign each possible target depth the following uniform prior $p(t_{i,j}=t)=\dfrac{1}{T},\quad t\in\mathbb{T}$. Note however that this prior can be adapted according to potential prior knowledge about the expected target depth distribution.

Accounting for potential spatial dependencies for the target intensities is more challenging than for the background levels as all pixels do not necessarily contain targets. Thus, considering fixed neighbourhood structures is not well adapted here. To reflect the lack of prior knowledge about the intensities of the target to be detected, we propose the following classical hierarchical model

$$r_{i,j}|\alpha,\beta \sim \mathcal{G}\left(\alpha,\beta\right),\quad\forall(i,j) \quad (4a)$$

$$\alpha|\alpha_1,\alpha_2 \sim \mathcal{G}\left(\alpha_1,\alpha_2\right) \quad (4b)$$

$$\beta|\beta_1,\beta_2 \sim \mathcal{IG}\left(\beta_1,\beta_2\right) \quad (4c)$$

where (α_1,α_2) and (β_1,β_2) are fixed parameters set to $(\alpha_1,\alpha_2)=(1.1,1)$ and $(\beta_1,\beta_2)=(1,1)$ to reflect the fact that the target intensities have a high probability to be in $(0,1)$. Indeed, the photon impulse response $g_0(\cdot)>0$ estimated during the imaging system calibration can be scaled appropriately using reference targets and acquisition times. Note that although (4) does not account for possibly spatially correlated target intensities, this model translates the prior belief that the potential target intensities share similar statistical properties (through α and β).

Finally, in a similar fashion to the background levels, it is often reasonable to expect the probability of a target to be present in a pixel to be related to the presence of targets in the neighbouring pixels (at least when considering targets larger than the spacing between pixels as considered in Section 5). To encode this prior belief, we attach the $N_{\text{row}}\times N_{\text{col}}$ detection label matrix $\mathbf{Z}$ ($[\mathbf{Z}]_{i,j}=z_{i,j}$) the following Ising model

$$f(\mathbf{Z}|c) = \frac{1}{G(c)}\exp\left[c\phi(\mathbf{Z})\right] \quad (5)$$

where $\phi(\mathbf{Z})=\sum_{i,j}\sum_{(i',j')\in\mathcal{V}_{i,j}}\delta\left(z_{i,j}-z_{i',j'}\right)$, $\delta(\cdot)$ denotes the Kronecker delta function, and $\mathcal{V}_{i,j}$ is the set of neighbours of pixel (i,j) (in this paper we consider an 8-neighbour structure). Moreover, c is an hyperparameter that controls the spatial granularity of the Ising model and $G(c)=\sum_{\mathbf{Z}\in(0;1)^{N_{\text{row}}\times N_{\text{col}}}}\exp\left[c\phi(\mathbf{Z})\right]$. In a similar fashion to ν, c is assumed to be fixed here but can be adjusted automatically using [17].

3.2. Joint posterior distribution

We can now specify the joint posterior distribution for $\mathbf{Z},\boldsymbol{\Theta}=\{\boldsymbol{\theta}_{i,j}\}_{i,j},\mathbf{\Gamma},\alpha$ and β given the observed waveforms $\mathbf{Y}$ and the value of hyperparameters $\alpha_1, \alpha_2, \beta_1, \beta_2, \nu$ and c. Note that for clarity the dependence of all distributions on the known fixed quantities $(\alpha_1, \alpha_2, \beta_1, \beta_2, \nu, c)$ is omitted in the remainder of the paper. Using Bayes theorem, the joint posterior distribution associated with the proposed Bayesian model is given by

$$f(\mathbf{Z}, \mathbf{\Theta}, \mathbf{\Gamma}, \alpha, \beta|\mathbf{Y}) \propto \left[\prod_{i,j} f(\mathbf{y}_{i,j}|z_{i,j}, \boldsymbol{\theta}_{i,j}) f(\boldsymbol{\theta}_{i,j}|\mathbf{Z}, \mathbf{\Gamma}, \alpha, \beta)\right] \times \quad f(\mathbf{Z}) f(\mathbf{\Gamma}) f(\alpha) f(\beta). \qquad (6)$$

4. BAYESIAN INFERENCE

4.1. Bayesian estimators

The Bayesian model defined in Section 3 specifies the joint posterior density for the unknown parameters $\mathbf{Z}, \mathbf{\Theta}, \mathbf{\Gamma}, \alpha$ and β given the observed data $\mathbf{Y}$ and the parameters ν and c. In this section we define suitable Bayesian estimators to summarize this knowledge and perform target detection. Here we consider the following coupled Bayesian estimators that are particularly suitable for model selection problems: the marginal maximum a posteriori (MMAP) estimator for the target presence labels

$$z_{i,j}^{MMAP} = \underset{z_{i,j} \in \{0,1\}}{\operatorname{argmax}} \; f(z_{i,j}|\mathbf{Y}, \hat{\nu}, c), \qquad (7)$$

and, conditionally on the estimated labels, 1) the minimum mean square error estimator of the background levels

$$b_{i,j}^{MMSE} = \mathrm{E}\left[b_{i,j}|z_{i,j} = \hat{z}_{i,j}^{MMAP}, \mathbf{Y}, \hat{\nu}, c\right], \qquad (8)$$

and 2) for the pixels for which $\hat{z}_{i,j}^{MMAP} = 1$, the minimum mean square error estimator of the target intensities and the marginal maximum a posteriori (MMAP) estimator of the target positions

$$\begin{aligned} r_{i,j}^{MMSE} &= \mathrm{E}\left[r_{i,j}|z_{i,j} = 1, \mathbf{Y}, \hat{\nu}, c\right] && (9) \\ t_{i,j}^{MMAP} &= \underset{t_{i,j} \in \mathbb{T}}{\operatorname{argmax}} \, f(t_{i,j}|z_{i,j} = 1, \mathbf{Y}, \hat{\nu}, c), && (10) \end{aligned}$$

Note that marginalising out the other unknowns (including α and β) in (7), (8), (9) and (10) automatically takes into account their uncertainty. Computing (7) to (10) is challenging because it requires having access to the univariate marginal densities of $z_{i,j}$ and the joint marginal densities of (among others) $(b_{i,j}, z_{i,j})$, which in turn require computing the posterior (6) and integrating it over a very high-dimensional space. Fortunately, these estimators can be efficiently approximated with arbitrarily large accuracy by Monte Carlo integration. This is why we propose to compute (7) to (10) by first using an MCMC method to generate samples asymptotically distributed according to (6), and subsequently using these samples to approximate the required marginal probabilities and expectations. Here we propose an RJ-MCMC method to simulate samples from (6), as this type of MCMC method is particularly suitable for models involving hidden Markov random fields and parameter spaces of varying of dimensions [18, Chap. 10,11]. The output of this algorithm are Markov chains of N_{MC} samples distributed according to the posterior distribution (6). The first N_{bi} samples of these chains correspond to the so-called *burn-in* transient period and should be discarded (the length of this period can be assessed visually from the chain plots or by computing convergence tests). The remaining $N_{\text{MC}} - N_{\text{bi}}$ of each chain are used to approximate the Bayesian estimators (7) to (10) as in [19, 11].

4.2. RJ-MCMC algorithm

The remainder of this section provides details about the main steps of the proposed method, summarised in Algo. 1.

ALGORITHM 1

Collaborative target detection

1: Fixed input parameters: Lidar impulse response $g_0(\cdot)$, $(\alpha_1, \alpha_2, \beta_2, \beta_2)$, number of burn-in iterations N_{bi}, total number of iterations N_{MC}
2: Initialization ($t = 0$)
- Set $\mathbf{Z}^{(0)}, \mathbf{\Theta}^{(0)}, \mathbf{\Gamma}^{(0)}, \alpha^{(0)}, \beta^{(0)}, \nu^{(0)}, c^{(0)}$

3: Iterations ($1 \leq t \leq N_{\text{MC}}$)
4: **for** $i = 1 : N_{\text{row}}$ **do**
5: **for** $j = 1 : N_{\text{col}}$ **do**
6: Update $(z_{i,j}^{(t)}, \boldsymbol{\theta}_{i,j}^{(t)})$ using RJ-MCMC
7: **end for**
8: **end for**
9: Sample $\mathbf{\Gamma}^{(t)} \sim f(\mathbf{\Gamma}|\mathbf{B}^{(t)}, \nu)$
10: Sample $\alpha^{(t)} \sim f(\alpha|\mathbf{Y}, \mathbf{Z}^{(t)}, \mathbf{\Theta}^{(t)}, \beta^{(t-1)})$
11: Sample $\beta^{(t)} \sim f(\beta|\mathbf{Y}, \mathbf{Z}^{(t)}, \mathbf{\Theta}^{(t)}, \alpha^{(t)})$
12: Set $t = t + 1$.

At each iteration, the elements of $\mathbf{Z}$ and $\{\boldsymbol{\theta}_{i,j}\}_{i,j}$ are updated sequentially, pixel by pixel (line 6 in Algo. 1). Precisely, with a given probability (0.5 here), for each pixel, $\boldsymbol{\theta}_{i,j}$ is updated while staying in the same parameter space (i.e., $\mathbf{\Theta}_0$ or $\mathbf{\Theta}_1$) and $z_{i,j}$ remains unchanged. This can be achieved using standard Gibbs updates. Otherwise (with probability 0.5), a move from the current model to the other is proposed. The generated candidate is then accepted with an appropriate probability ensuring that the resulting RJ-MCMC algorithm admits the target posterior distribution as invariant distribution (see [20] for details). The remaining parameters are updated using standard Gibbs and Metropolis-Hastings updates. Precisely, due to the conjugacy between (4a) and (4c), it can be easily shown that $f(\beta|\mathbf{Y}, \mathbf{Z}, \mathbf{\Theta}, \alpha) = f(\beta|\mathbf{Z}, \mathbf{\Theta}, \alpha)$ is an inverse-gamma distribution which is easy to sample from. Moreover, $f(\alpha|\mathbf{Z}, \mathbf{\Theta}, \beta)$ is strictly log-concave if $\alpha_1 \geq 1$ (see [21]). Consequently, α can be updated using standard Metropolis-Hastings updates or adaptive rejection sampling [22]. Here we used Metropolis-Hastings updates using a Gaussian random walk whose variance is adjusted during the early iterations of the sampler. Finally, by noting that $\mathbf{\Gamma}$ does not appear in (1a) nor (1b), sampling from its conditional distribution reduces to sampling from $f(\mathbf{\Gamma}|\mathbf{B}, \nu)$ in (3).

5. SIMULATION RESULTS

In this section, we assess the performance of the proposed method to reconstruct a depth image of a life-sized polystyrene head located at a stand-off distance of 325 meters from a TCSPC-based scanning sensor. The transceiver system and data acquisition hardware used for this work is broadly similar to that described in [23, 24, 25, 9, 11], which was previously developed at Heriot-Watt University. The measurements have been performed outdoors, on the Edinburgh Campus of Heriot-Watt University, under dry clear skies (around noon). Additional information about the acquisition process can be found in [20]. The acquisition time per pixel is 30ms (200×200 pixels). However, the data format of time-tagged events allows the construction of photon timing histograms associated with shorter acquisition times, after measurement, as the system records the time of arrival of each detected photon. Here, we evaluate our algorithm for acquisition times of 30ms, 3ms, 1ms, and 300μs per pixel. The average number of detected photons per pixel ranges from 5.6 to 554.6 for exposures of 300μs to 30ms per pixels. The instrumental impulse response $g_0(\cdot)$ is estimated from preliminary experiments by analysing the distribution of photons reflected from a standard commercial Lambertian scatterer. The proposed Bayesian algorithm has been applied with $N_{\text{MC}} = 1000$ iterations, including $N_{\text{bi}} = 300$ burn-in iterations and the hyperparameters (ν, c) have been adjusted from preliminary runs.

The proposed method is compared to the standard method used for depth imaging [9] and which is divided into two steps. The first step consists of estimating $t_{i,j}$ using cross-correlation between $\log(g_0(\cdot))$ and the photon histogram $\mathbf{y}_{i,j}$, which corresponds to log-match filtering or maximum likelihood (ML) estimation under background-free ($b_{i,j} = 0$) assumption. Once the estimated target distance or associated time-of-flight $\hat{t}_{i,j,\text{corr}}$ has been computed, the target intensity and the background level for each pixel are estimated using ML estimation and the likelihood (1b). Note that this likelihood is convex with respect to $(r_{i,j}, b_{i,j})$ with $r_{i,j} \geq 0, b_{i,j} \geq 0$ and that the ML estimates of $(r_{i,j}, b_{i,j})$ (conditioned on $\hat{t}_{i,j,\text{corr}}$)can be obtained efficiently using constrained convex optimization methods (here we used an ADMM method similar to [26]).

		π_{00}	π_{10}	π_{01}	π_{11}
3ms	X-corr	79.9	20.1	8.9	91.1
	Prop. algo.	99.9	0.01	10.8	89.2
1ms	X-corr	57.4	42.6	16.9	83.1
	Prop. algo.	99.9	0.01	18.6	81.4
0.3ms	X-corr	59.6	40.4	39.1	60.9
	Prop. algo.	99.9	0.01	20.4	79.6

Table 1. Empirical detection performance (prob. in %).

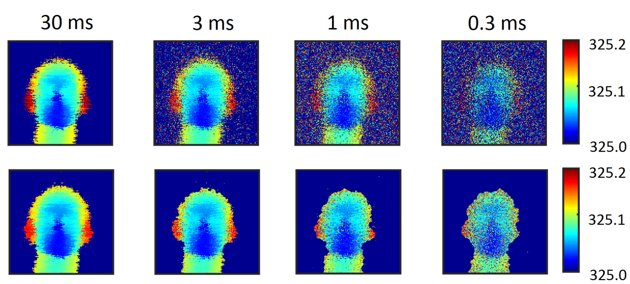

Fig. 1. Depth maps estimated by the standard method (top) and the proposed Bayesian algorithm (bottom). Distances are in meters.

The detection performance of the proposed and standard algorithms is quantitatively assessed by comparing their empirical specificity π_{00} (deciding $\mathcal{M}_0^{(i,j)}$ when $\mathcal{M}_0^{(i,j)}$ is true) and sensitivity π_{11} or equivalently their empirical probability of false alarm $\pi_{10} = 1 - \pi_{00}$ and of miss $\pi_{01} = 1 - \pi_{11}$. Although the standard method does not provide target detection results directly, it is possible to infer the target presence by thresholding the estimated intensity images. In all the results presented here, we set the threshold to $\eta = 0.1$, which corresponds to an estimated target intensity 10 times smaller than that of the Spectralon panel. Table 1 compares the detection performance of the standard and proposed methods, averaged over all pixels, for acquisition times of 3ms, 1ms and 0.3ms per pixel. The results obtained by the proposed method with an acquisition time of 30ms are used as ground truth here. This table shows that the performance of the two algorithms degrade when reducing the acquisition time. However, the proposed method (thanks to its target detection ability and the different special regularizations) generally provides lower probabilities of false alarm as well as less significant performance degradation than the standard method.

Fig. 1 compares the estimated depth maps obtained by the standard and the proposed methods. These results show that for large acquisition times, the two methods provide similar results. However, when the acquisition time decreases, the two methods start to fail in identifying the target positions. However, due to its better target detection ability, the method proposed here provides more reliable depth images as it can more accurately detect pixels not containing a surface.

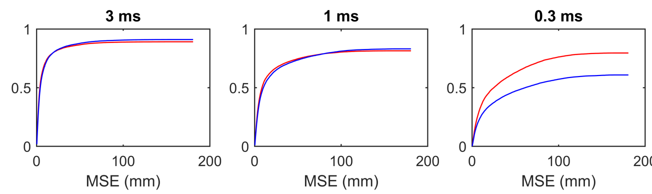

Fig. 2. Distance MSE cdfs provided by the standard (blue) and the proposed (red) methods for different exposures.

The performance of the two methods are quantitatively evaluated using the distance mean squared errors (MSEs) defined by $MSE(d_{i,j}) = \left\| \hat{d}_{i,j} - d_{i,j} \right\|_2^2$, where $\|\cdot\|_2$ denotes the ℓ_2-norm, $\hat{d}_{i,j}$ is the estimated value of $d_{i,j} = 3 \times 10^8 t_{i,j}/2$. Since the actual distances $\{d_{i,j}\}$ are unknown for the data sets considered, these values have been replaced by those estimated by the proposed method for the longest acquisition time (30ms). Fig. 2 depicts the cumulative density functions (cdfs) of the distance MSEs [11]. The steeper the curve, the better the depth estimation. As expected, the two methods performance generally degrades when reducing the exposure (the curves go lower and more gradually). Note that for each dataset, the cdfs are upper-bounded by the sensitivity π_{11} (see Table 1) of each method, which explains why the curves in Fig. 2 do not reach 1. For the longer exposures, the two methods present similar depth estimation performance for the pixels actually containing targets and the benefits of the proposed methods mainly rely on its target detection ability (lower π_{10}). For the shortest exposure, the proposed background prior model (2), together with the target presence model (5), enables a more accurate depth image (lower π_{10} and π_{01} and better depth estimation) than with the standard approach.

6. CONCLUSION

We presented a Bayesian algorithm for target detection and depth imaging using sparse single-photon data. The experiments conducted on real Lidar data demonstrate the ability of the proposed method to detect and identify targets observed under difficult observation conditions (high and spatially variable background levels, short acquisition times), with a better accuracy than the standard method. In contrast to [11], we didn't explicitly account for the possible correlations affecting the intensity and/or depth images. Although it is possible to apply the algorithm studied in [11] to refine the depth/intensity images after the target detection step, it would be interesting to extend the proposed model to capture additional parameter dependency. Finally, investigating less computationally intensive optimization-based alternatives (e.g., *Expectation-Maximization* methods) and extending the model to multiple targets are very interesting problems which are the subject of further investigations that will be reported in subsequent papers.

7. REFERENCES

[1] T. Ogawa and K. Takagi, "Lane recognition using on-vehicle lidar," in *IEEE Intelligent Vehicles Symposium*, 2006, pp. 540–545.

[2] P. Lindner and G. Wanielik, "3d lidar processing for vehicle safety and environment recognition," in *IEEE Workshop on Computational Intelligence in Vehicles and Vehicular Systems, 2009. CIVVS '09.*, March 2009, pp. 66–71.

[3] S. Matzka, A. M. Wallace, and Y. R. Petillot, "Efficient resource allocation for attentive automotive vision systems," *IEEE Trans. Intelligent Transportation System*, vol. 13, no. 2, pp. 859–872, June 2012.

[4] Y. Kang, C. Roh, S.-B. Suh, and B. Song, "A lidar-based decision-making method for road boundary detection using multiple kalman filters," *IEEE Trans. Industrial Electronics*, vol. 59, no. 11, pp. 4360–4368, Nov 2012.

[5] D. Ramirez, G. S. Buller, A. McCarthy, X. Ren, A. M. Wallace, S. Morak, C. Nichol, and I. Woodhouse, "Developing hyper-spectral lidar for structural and biochemical analysis of forest data," in *Advances in Geosciences, Proceedings of EARSEL Conference*, Mykonos, Greece, May 2012.

[6] T. Hakala, J. Suomalainen, S. Kaasalainen, and Y. Chen, "Full waveform hyper-spectral lidar for terrestrial laser scanning," *Opt. Express*, vol. 20, no. 7, pp. 7119–7127, Mar 2012.

[7] N. Cadalli, P. J. Shargo, D. C. Munson, Jr., and A. C. Singer, "Three-dimensional tomographic imaging of ocean mines from real and simulated lidar returns," in *Proc. SPIE, Ocean Optics: Remote Sensing and Underwater Imaging*, 2002, vol. 4488, pp. 155–166.

[8] J. Gao, J. Sun, J. Wei, and Q. Wang, "Research of underwater target detection using a slit streak tube imaging lidar," in *Acad. Int. Symp. Optoelectronics and Microelectronics Technology (AISOMT)*, Oct 2011, pp. 240–243.

[9] A. McCarthy, X. Ren, A. Della Frera, N. R. Gemmell, N. J. Krichel, C. Scarcella, A. Ruggeri, A. Tosi, and G. S. Buller, "Kilometer-range depth imaging at 1550 nm wavelength using an InGaAs/InP single-photon avalanche diode detector," *Opt. Express*, vol. 21, no. 19, pp. 22098–22113, Sept. 2013.

[10] A. Kirmani, D. Venkatraman, D. Shin, A. Colao, F. N. C. Wong, J. H. Shapiro, and V. K. Goyal, "First-photon imaging," *Science*, vol. 343, no. 6166, pp. 58–61, 2014.

[11] Y. Altmann, X. Ren, A. McCarthy, G. S. Buller, and S. McLaughlin, "Lidar waveform based analysis of depth images constructed using sparse single-photon data," *ArXiv e-prints*, July 2015.

[12] S. Hernandez-Marin, A. M. Wallace, and G. J. Gibson, "Bayesian analysis of lidar signals with multiple returns," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 29, no. 12, pp. 2170–2180, Dec 2007.

[13] A. M. Wallace, A McCarthy, C.J. Nichol, Ximing Ren, S. Morak, D. Martinez-Ramirez, IH. Woodhouse, and G. S. Buller, "Design and evaluation of multispectral lidar for the recovery of arboreal parameters," *IEEE Trans. Geoscience and Remote Sensing*, vol. 52, no. 8, pp. 4942–4954, Aug 2014.

[14] P. J. Green, "Reversible jump MCMC computation and Bayesian model determination," *Biometrika*, vol. 82, no. 4, pp. 711–732, Dec. 1995.

[15] C. Andrieu and A. Doucet, "Joint Bayesian model selection and estimation of noisy sinusoids via reversible jump MCMC," *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 19–37, Oct. 1999.

[16] O. Dikmen and AT. Cemgil, "Gamma markov random fields for audio source modeling," *IEEE Trans. Audio, Speech, Language Processing*, vol. 18, no. 3, pp. 589–601, March 2010.

[17] M. Pereyra, N. Whiteley, C. Andrieu, and J.-Y. Tourneret, "Maximum marginal likelihood estimation of the granularity coefficient of a Potts-Markov random field within an mcmc algorithm," in *Proc. IEEE-SP Workshop Stat. and Signal Processing*, Gold Coast, Australia, July 2014.

[18] Christian P. Robert and George Casella, *Monte Carlo Statistical Methods*, Springer-Verlag, New York, second edition, 2004.

[19] Y. Altmann, M. Pereyra, and J. Bioucas-Dias, "Collaborative sparse regression using spatially correlated supports - application to hyperspectral unmixing," *ArXiv e-prints*, Sept. 2014.

[20] Y. Altmann, X. Ren, A. McCarthy, G. S. Buller, and S. McLaughlin, "Target detection for depth imaging using sparse single-photon data," Tech. Rep., Heriot-Watt University, Sept. 2015, Available online, http://yoannaltmann.weebly.com/publications.html.

[21] M. Zhou, L. Hannah, D. Dunson, and L. Carin, "Beta-negative binomial process and poisson factor analysis," in *AISTATS 2012*. JMLR, 2011.

[22] W. R. Gilks and P. Wild, "Adaptive rejection sampling for gibbs sampling," *Applied Statistics*, vol. 41, pp. 337–348, 1992.

[23] A. McCarthy, R. J. Collins, N. J. Krichel, V. Fernández, A. M. Wallace, and G. S. Buller, "Long-range time-of-flight scanning sensor based on high-speed time-correlated single-photon counting," *Appl. Opt.*, vol. 48, no. 32, pp. 6241–6251, Nov. 2009.

[24] N. J. Krichel, A. McCarthy, and G. S. Buller, "Resolving range ambiguity in a photon counting depth imager operating at kilometer distances," *Opt. Express*, vol. 18, no. 9, pp. 9192–9206, April 2010.

[25] A. M. Wallace, J. Ye, N. J. Krichel, A. McCarthy, R. J. Collins, and G. S. Buller, "Full waveform analysis for long-range 3d imaging laser radar," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, no. 1, pp. 896708, 2010.

[26] M. Figueiredo and J. Bioucas-Dias, "Restoration of poissonian images using alternating direction optimization," *IEEE Trans. Image Processing*, vol. 19, no. 12, pp. 3133–3145, 2010.