

Bayesian Analysis of Lidar Signals with Multiple Returns

Sergio Hernández-Marín, *Student Member, IEEE*, Andrew M. Wallace, and Gavin J. Gibson

Abstract—Time-Correlated Single Photon Counting and Burst Illumination Laser data can be used for range profiling and target classification. In general, the problem is to analyze the response from a histogram of either photon counts or integrated intensities to assess the number, positions, and amplitudes of the reflected returns from object surfaces. The goal of our work is a complete characterization of the 3D surfaces viewed by the laser imaging system. The authors present a unified theory of pixel processing that is applicable to both approaches based on a Bayesian framework, which allows for careful and thorough treatment of all types of uncertainties associated with the data. We use reversible jump Markov chain Monte Carlo (RJMCMC) techniques to evaluate the posterior distribution of the parameters and to explore spaces with different dimensionality. Further, we use a delayed rejection step to allow the generated Markov chain to mix better through the use of different proposal distributions. The approach is demonstrated on simulated and real data, showing that the return parameters can be estimated to a high degree of accuracy. We also show some practical examples from both near and far-range depth imaging.

Index Terms—Three-dimensional reconstruction, burst illumination laser, delayed rejection, Lidar, photon counting, reversible jump MCMC.

1 INTRODUCTION

WE consider the processing of time-of-flight laser imaging detection and ranging (Lidar) signals to obtain accurate range measurement and 3D reconstruction of surfaces at ranges from a few meters to several kilometers. Normally, such systems process only an assumed single return from a single opaque surface. In recent years, interest has grown in situations in which the laser return consists of multiple peaks, for example, due to the footprint of the beam impinging on a target with surfaces distributed in depth or with semitransparent surfaces [1], [2], [3], [4]. The goal of our work is a complete characterization of the 3D surfaces viewed by the laser imaging system. Although we do apply our analyses to 3D image data, the algorithms we employ consider each pixel or scanned elements in isolation, that is, independently of adjacent elements in image space, as a first step in a fuller analysis that takes account of between pixel dependencies.

We make a number of contributions: First, we present a unified theory of pixel processing that is applicable to both Time-Correlated Single Photon Counting (TCSPC) and Burst Illumination Laser (BIL). To this end, we adopt a Bayesian statistical approach based on reversible jump Markov chain Monte Carlo (RJMCMC) techniques [5], [6] to assess the number, positions, and amplitudes of the returned signals

from target surfaces. The main advantages of the Bayesian approach are that it facilitates representation, takes a fuller account of the uncertainties related to models and parameter values, and provides a natural method for updating beliefs in response to new information and, therefore, to incorporate prior knowledge into the analysis. Second, we incorporate a delayed rejection step [7] in an RJMCMC algorithm that permits temporary local adjustments of the proposal distribution. In short, the delayed rejection step allows the Markov chain to mix better through the use of different proposal distributions and, therefore, to explore better the posterior distribution derived from the Bayesian approach.

In Section 2, we introduce the TCSPC and BIL approaches and discuss related work. In Section 3, we describe our research methodology using a Bayesian framework. In Section 4, we explain how RJMCMC can be used to approximate the full joint posterior distribution obtained in the previous section. In Section 5, we present the effectiveness of the algorithm developed through a series of different experiments using synthetic and real data. Finally, in Section 6, we present conclusions and possible directions for future work.

2 BACKGROUND AND RELATED WORK

We employ two principal approaches to range measurement and depth imaging, by TCSPC [8], [9] and by range gating or repeated BIL [10], [11]. With reference to Fig. 1, a pulsed laser is directed toward the target; the time-of-flight (TOF) of the laser to and from the target gives a measurement of distance since the speed of light in the medium is known. However, in the TCSPC method, the detector (a single photon avalanche photodiode (SPAD)) records the arrival of the first photon; repeated laser pulses result in a histogram of photon arrival times that can be processed to determine the target range and signature. In the BIL method, the detector (in our case, an electron-bombardment charge-coupled device (CCD))

• S. Hernández-Marín and A.M. Wallace are with the ERP Joint Research Institute Image and Signal Processing, School of Engineering and Physical Sciences, Heriot-Watt University, Riccarton, Edinburgh EH14 4AS, UK. E-mail: {snh3, a.m.wallace}@hw.ac.uk.

• G.J. Gibson is with the Maxwell Institute of Mathematical Sciences, Department of Actuarial Mathematics and Statistics, School of Mathematical and Computer Sciences, Heriot-Watt University, Riccarton, Edinburgh EH14 4AS, UK. E-mail: g.j.gibson@ma.hw.ac.uk.

Manuscript received 12 June 2006; revised 7 Dec. 2006; accepted 12 Feb. 2007; published online 27 Feb. 2007.

Recommended for acceptance by R. Klette.

For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-0436-0606. Digital Object Identifier no. 10.1109/TPAMI.2007.1122.

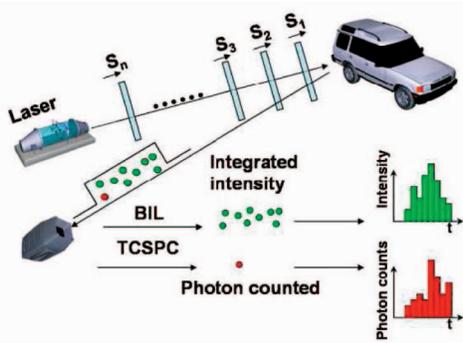


Fig. 1. TCSPC and BIL imaging principles.

integrates the received laser light over a time period (typically a few ns). Thus, the pulsed laser-CCD detector pair records the laser return from a distant slice of space. The range data is created from the multiple flashes by variable range gating. A histogram of integrated intensities is created with the time axis determined by the range gate delay. In either case, this pattern of recorded photon arrival times or gate delays is analyzed to find the distribution of depth and reflectance in the 3D scene.

To create a full 3D image (or “angle-angle-range” image) in the TCSPC method, the laser is normally scanned but there have been advances in the development of arrays of SPADs (or geiger-mode APDs) [9], [12]. In a BIL system, the CCD records a full 2D slice in 3D space at each range gate, so the 3D image is built from the repeated laser pulses as the range gate changes. The BIL images are acquired by a sensor of similar specification to that described by Duncan et al. [11] and consist of multiple frames that, in our case, have been typically acquired at 30 different range gate delays. The relative merits of the respective approaches are discussed in the references given. It is found generally that the TCSPC approach has greater sensitivity and accuracy of measurement than the range gated method, whereas the creation of the full 3D image is much faster in the BIL case.

As mentioned previously, current classical processing methods only consider a single return. In general, this single return corresponds to either the first pulse or the last pulse received. The peak detector is a simple threshold comparator or a matched filter. The system used therefore requires a reasonably high signal-to-noise ratio to be able to discriminate the return from the background noise. When the number of returns is increased, the approach chosen is that of implementing first/last pulse logic in the receiver electronics. In first pulse logic, this involves processing the first pulse detected, whereas in last pulse logic, this involves sampling and holding all the returns after the first but only recording the final return [1]. In practice, this fails because the pulse of interest is not obvious and more complex processing of the returns is needed. Schilling et al. [2] showed that the use of the entire multiple returns signal rather than a subset thereof provides more information about the scene of interest. For instance, in a defense application, the problem may be to detect a target under camouflage netting or foliage and in geosciences and remote sensing to produce a digital terrain model of a landscape or to estimate biomass parameters and the canopy structure of forested terrains [13].

Stochastic algorithms are computationally more intensive than methods such as matched filtering and threshold

detection and can take significantly longer to reach a solution. Nevertheless, they may give estimates of the range to higher resolution than the classical algorithms and may give indications of uncertainty. Typical processing methods calculate the Maximum Likelihood Estimates (MLE) of the parameters in the model within a frequentist framework [8]. One of the most used tools for finding the MLE is the Expectation-Maximization (EM) algorithm [14]. The main problem is that this algorithm can converge to a local maximum or a saddle point of the likelihood. Moreover, this methodology is sensitive to the initial estimates.

Previous statistical approaches to deal with the problem of multiple returns include nonparametric estimation using kernel-density estimation techniques and parametric estimation using the method of moments [15]. The former approach can fail to distinguish whether or not a return is composed of a single pulse broadened after passing through the atmosphere or several pulses, which are very close together. Therefore, these methods do not give any physical interpretation to the components of the mixture and may make inefficient use of data. They may fail to give accurate estimates when dealing with small samples and, in general, the maximum likelihood estimation gives a better representation of the true quantities.

Wax and Kailath [16] proposed detecting the number of signals in a multichannel time series using information theoretic criteria for model selection such as Akaike (AIC) and Schwartz and Rissanen’s Minimum Description Length (MDL). In practice, they calculated for every possible number of signals the MLE estimates of the parameters. The number of signals was determined by minimizing either the AIC or MDL criterion for each model. A Bayesian variant of the AIC is the Bayesian Information Criterion [17], which assumes equal priors on each model and noninformative priors on the parameters. Since the maximum number of signals must be fixed a priori, there is a risk of underestimation of the actual number of peaks and, therefore, of not exploring the space of solutions completely. In general, the MLE of the parameters is obtained using the EM algorithm. Celeux and Diebolt (see, for instance, [18]) proposed the use of a stochastic EM (SEM) that can be seen as random perturbation of the traditional EM algorithm. The SEM algorithm replaces the E step by incorporating a stochastic step to avoid entrapment in a local maxima of the likelihood function. A hybrid approach is found in Wallace et al. [19]; they used a deterministic nonparametric “bump-hunting” procedure that provided the number, amplitudes, and positions of the suspected returns. Once the initial estimates were obtained, they computed the MLE of the set of parameters that best explained the data. This methodology may fail when the number of returns increases or the returns are occluded by background noise.

3 ON THE BAYESIAN ANALYSIS OF MULTIPLE LIDAR RETURN

In this section, we describe how the detection problems described in Section 2 can be formalized in a Bayesian framework. Due to the similarity between the approaches, the methodology applied to TCSPC and BIL data is identical except for the representation of the underlying distribution, that is, the parametric model used.

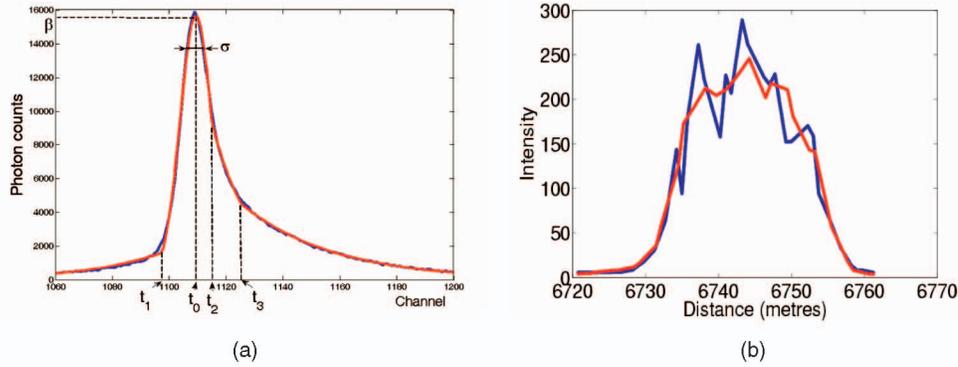


Fig. 2. (a) Histogram of photon counts of the real data (blue) and initial fit to obtain the shape parameters (red). The different parameters of the "operating" model are also shown. (b) Histogram of intensities (blue) of one pixel of the trig point in Fig. 7c and final fit from the RJMCMC estimation (red).

3.1 Data Modeling

The exact functional forms of the photon count and the intensity histograms are unknown. In the case of photon count data, we employ a nearest representation or operating model to describe the underlying distribution. This operating model is described in [20] and represents the parametric form of the expected temporal variation of the photon count distribution given by

$$f_{op}(i; p) = \beta \begin{cases} e^{-\frac{(t_1-t_0)^2}{2\sigma^2}} e^{\frac{(i-t_1)}{\tau_1}} & i < t_1 \\ e^{-\frac{(i-t_0)^2}{2\sigma^2}} & t_1 \leq i < t_2 \\ e^{-\frac{(t_2-t_0)^2}{2\sigma^2}} e^{\frac{(i-t_2)}{\tau_2}} & t_2 \leq i < t_3 \\ e^{-\frac{(t_2-t_0)^2}{2\sigma^2}} e^{\frac{(i-t_2)}{\tau_2}} e^{\frac{(i-t_3)}{\tau_3}} & i \geq t_3, \end{cases} \quad (1)$$

where $p = (\beta, \sigma, t_0, t_1, t_2, t_3, \tau_1, \tau_2, \tau_3)$ is the parameter set, β is an amplitude factor, t_0 is the time of the peak maximum, and t_1 , t_2 , and t_3 are the points at which the changeovers between functions occur, as shown in Fig. 2. In this study, we assumed that the shape parameters of the returned pulse are fixed and known from an instrumental response and, therefore, the only parameters subject to inference are the amplitude β and the time of the peak maximum t_0 .

In the case of the intensity histogram, we used a lookup table of an instrumental function, $w_{op}(i; \beta, t_0) = \beta \cdot w(i; t_0)$, instead of an operating model. In practice, this instrumental function is acquired from the response of the BIL system to a Lambertian reflecting surface at a similar range to the object of interest. Over this distribution, only two parameters were considered to be variable: The amplitude β and the time of the peak maximum t_0 .

However, there are many cases of several returns that may be overlapping in position and possibly weak in comparison with the noise or the background level, either because they have been observed at a distance or because the reflectance is low, for example, due to different alignment of the target and the sensor or the presence of obscurants on the target. Assuming several signals can be present in the same histogram and that these signals will be observed against a finite background level whose expected value is constant across all the channels, the observed photon (intensity) histogram, y , can be considered as a sample of a non-normalized statistical mixture distribution with density

$$F(i; k, \phi) = \sum_{j=1}^k f_{system}(i; \beta_j, t_{0j}) + B, \quad (2)$$

where k is the number of peaks, B is the background, and ϕ is the set of parameters of each signal and the background: $\phi = (\beta, t_0, B)$ with $\beta = (\beta_1, \beta_2, \dots, \beta_k)$ and $t_0 = (t_{01}, t_{02}, \dots, t_{0k})$. f_{system} is either $f_{op_j}(i; \beta_j, t_{0j})$ or $w_{op_j}(i; \beta_j, t_{0j})$ depending on the data we are analyzing.

If the time resolution is sufficiently fine, the value y_i recorded in each channel i can be considered as a random sample of a Poisson distribution with intensity $F(i; k, \phi)$, which depends on the model parameters

$$P(y_i | k, \phi) = e^{-F(i; k, \phi)} \frac{F(i; k, \phi)^{y_i}}{y_i!}. \quad (3)$$

Equation (3) represents the physics of sensor operation, where y_i is the number of photons (respectively, the intensity value in a BIL histogram) in channel i . Furthermore, assuming that the observations recorded in each channel i of the histogram are conditionally independent given the value of the parameters, the joint probability distribution of y is defined as

$$L(y | k, \phi) = \prod_{i=1}^{i_{max}} e^{-F(i; k, \phi)} \frac{F(i; k, \phi)^{y_i}}{y_i!}, \quad (4)$$

where $y = \{y_1, y_2, \dots, y_{i_{max}}\}$.

3.2 Bayesian Inference

The objective of this study is inference about the unknowns k and ϕ in order to obtain accurate estimates of the position, amplitude, number of peaks and background of the returned signals. These unknowns are assumed a priori independent of each other and drawn from appropriate independent prior distributions. The position parameter, t_{0j} , is drawn from a flat uniform distribution with support in the interval $[0, i_{max}]$ to reflect our complete prior ignorance of the true position of the peaks. Likewise, for the number of peaks, we take a uniform distribution on $[0, \dots, k_{max}]$ with k_{max} chosen to be suitably large. Converting the results obtained with this prior to those corresponding to other priors on k is straightforward [6]. For the prior distribution of scale parameters β_j and B , we assume a gamma distribution that is a natural choice in the statistics

community for parameters whose support is in \mathfrak{R}^+ . The full joint prior distribution can be modeled as

$$f(k, \phi) = \frac{1}{k_{max}} \left(\frac{1}{i_{max}} \right)^k f_G(B|c, d) \prod_{n=1}^k f_G(\beta_n|a, b), \quad (5)$$

where f_G is the probability density function (PDF) of a gamma distribution with mean $a \cdot b$ (respectively, $c \cdot d$) and variance equal to $a \cdot b^2$ ($c \cdot d^2$, respectively).

In the Bayesian paradigm, the information brought by the data, the likelihood, is combined with prior information and summarized in a probability distribution called the posterior distribution, our target distribution. This posterior distribution captures knowledge about the parameters given the data. Thus, using (4) and (5) and Bayes' theorem, the target distribution can be expressed as

$$\pi(k, \phi|y) = \frac{L(y|k, \phi)f(k, \phi)}{\int L(y|k, \phi)f(k, \phi)\delta(k, \phi)} \propto L(y|k, \phi)f(k, \phi). \quad (6)$$

In general, (6) is intractable and difficult to analyze. Powerful simulation algorithms such as MCMC allow posterior distributions to be explored and conclusions and inferences to be drawn directly from the sampled values.

4 RJMCMC METHODOLOGY

MCMC techniques enable simulation from an unknown distribution by embedding it as the limiting distribution of a Markov chain and simulating samples from that chain until it approaches equilibrium. When the state space is a set of a countable number of subspaces of different dimension, standard MCMC algorithms are not applicable and other algorithms are needed to explore all these different subspaces. The reversible jump MCMC approach [5] copes with this problem. This is an extension of the MCMC algorithm designed to allow jumps between subspaces of different dimension.

We follow an approach that is related to that described in [6] by constructing a Markov chain whose transitions involve changes to the number, positions, and amplitudes of peaks in the return signal. However, in contrast to that in [6], we do not associate each photon in the histogram with a particular component in a mixture model but treat the histogram as a discrete representation of a spatially heterogeneous Poisson process whose intensity is a linear superposition of the scaled and shifted templates of the form specified by (1). The transitions of the Markov chain we generate involved moves of various types. The different moves we allow are

1. updating the positions t_0 ,
2. updating the amplitudes β ,
3. updating the background B ,
4. random birth or death of a peak, and
5. random splitting of a peak into two peaks or merging of two peaks into a single peak.

One complete pass over these five moves is called a sweep.

4.1 Fixed Dimension Moves

Moves of type (1), (2), and (3) allow the posterior distribution to be explored in a state space with fixed dimension, that is, in this particular case, when the number of peaks does not vary, $k = k'$. To simplify notation, we write $\pi(k, \phi|y) = \pi(\phi|y)$ when k is fixed. Furthermore, to improve the rate of accepted values, the different parameters are updated separately

yielding three different acceptance probabilities, one for each parameter. Initially, when $k = 0$, moves (1) and (2) are not available since only background noise is present.

The task is to construct a Markov chain on a general state space with $\pi(\phi|y)$ as its limiting distribution. Once convergence toward this limiting distribution is achieved, we form a realization of the chain $\{\Phi^{(1)}, \Phi^{(2)}, \dots, \Phi^{(N)}\}$ and treat it as a random sample from $\pi(\phi|y)$. By means of the ergodic theorem, it is possible to use this realization of the chain to build histograms, estimate moments, calculate probabilities, and so forth. These realizations are sampled from a distribution called the transition kernel $K(\cdot, \cdot)$, which embodies the Markov property, that is, given the present state, the future state is conditionally independent on the past states. To allow the Markov chain to converge toward the limiting distribution, the transition kernel constructed has to satisfy three properties, that is, the transition kernel has to be aperiodic, irreducible, and invariant. A sufficient but not necessary condition for invariance is reversibility with respect to π . The detailed balance condition guarantees the reversibility condition

$$\pi(\phi|y)K(\phi, \phi') = \pi(\phi'|y)K(\phi', \phi). \quad (7)$$

In general, to construct a suitable transition kernel, $K(\cdot, \cdot)$, we can use the Gibbs sampler and the Metropolis-Hastings method. We have a set of variables $x = \{\phi_1, \phi_2, \dots, \phi_n\}$, and each method makes a random scan through different subsets of the space of solutions to update these values and, hence, to create a suitable Markov chain. The updated values are obtained in different ways depending on the methodology considered. The Gibbs sampler makes use of the full conditional distribution to draw the new values, whereas the Metropolis-Hastings method draws the proposed values from an arbitrary distribution $q(\cdot, \cdot)$. These values are accepted with probability $\alpha(\cdot, \cdot)$; otherwise, these values are rejected, and the existing values are retained. The Gibbs sampler is a variant of the Metropolis-Hastings algorithm where $\alpha(\cdot, \cdot) = 1$.

Since we are dealing with complex models, the Metropolis-Hastings paradigm is more appropriate than the Gibbs sampler. Indeed, when the dimension of the state space increases or π has a complex form, in general, it is not possible to sample from distributions unless standard distributions or conjugate priors are used. On the other hand, if the dimension of the space of solutions is not fixed, the procedure used in the Gibbs sampler is not meaningful since the interpretation of the different variables can be different depending on which subspace of state space we are in. In the Metropolis-Hastings approach, the (simplified) transition kernel can be written as

$$K(\phi, \phi') = q(\phi', \phi)\alpha(\phi, \phi') + \delta_\phi(\phi')w(\phi), \quad (8)$$

where $q(\phi', \phi)\alpha(\phi, \phi')$ is the probability of performing a move, $w(\phi) = (1 - \int \alpha(\phi, \phi_1)q(\phi, \phi_1)d\phi_1)$, δ_ϕ is a point mass at ϕ , and $\delta_\phi(\phi')w(\phi)$ is the probability of not moving from ϕ .

It can be shown (see Appendix, which can be found at <http://computer.org/tpami/archives.htm>) that the following equation satisfies (7):

$$\alpha(\phi, \phi') = \min \left\{ 1, \frac{\pi(\phi'|y)}{\pi(\phi|y)} \times \frac{q(\phi, \phi')}{q(\phi', \phi)} \right\}, \quad (9)$$

where $\alpha(x, x')$ is known as the acceptance probability.

Equation (9) will determine whether a candidate value is accepted or rejected, and it is a key element in the

generation of the Markov chain. There are no restrictions in the choice of the proposal distribution, $q(\phi, \phi')$. However, persistent rejection may indicate that, locally, the proposal distribution is badly calibrated to the target. Although this rejection is necessary to preserve the stationary distribution, if it is too high, we increase the correlation between consecutive samples of the generated chain.

The delayed rejection step improves the performance of RJMCMC methods to cope with this problem and is an alternative to careful offline tuning of state-dependent proposals. In our case, this methodology is used to better explore the posterior distribution of the different parameters since the range of values with which we work is very large.

Tierney and Mira [21] and, later, Green and Mira [7], for a variable-dimension approach, proposed that whenever a candidate move is rejected, we make another attempt to move using a second proposal instead of turning to the next transition. The proposed distribution at this second stage is allowed to depend on the previous rejected value, as well as the current position of the chain. Although the acceptance probability of the new candidate has to be adjusted to preserve the stationary distribution, this methodology does not destroy the Markovian property of the sampler. This process can be repeated indefinitely.

If the dimension is fixed, the methodology is the following: When at ϕ , we propose a new state ϕ' from a density $q_1(\phi, \phi')$. This move is accepted using a probability similar to that of (9). If the move is rejected, a second proposal, ϕ'' , is made from a density $q_2(\phi, \phi', \phi'')$, and the new acceptance probability is $\alpha(\phi, \phi'') = \min\{1, A\}$, where A is

$$A = \frac{\pi(\phi'')q_1(\phi'', \phi')q_2(\phi'', \phi', \phi)\{1 - \alpha_1(\phi'', \phi')\}}{\pi(\phi)q_1(\phi, \phi')q_2(\phi, \phi', \phi'')\{1 - \alpha_1(\phi, \phi')\}}. \quad (10)$$

In this work, we use only the delayed rejection algorithm when the dimension of the space of solutions is fixed as is the case with the parameter-updating step.

4.1.1 Updating the Positions

To update a position, our Markov chain explores the space of solutions using a random walk proposal. The proposal distribution is therefore

$$q(t'_0, t_0) = \prod_{j=1}^k q_N(t'_0 | t_0, \sigma_1), \quad (11)$$

where q_N denotes a normal PDF whose mean is equal to the current value of t_0 and with a standard deviation σ_1 , which is chosen according to the size of the data vector (TCSPC or BIL). The acceptance probability is $\alpha_1(t_0, t'_0) = \min\{1, A_1\}$ with A_1 equal to

$$A_1 = \frac{L(y|k, \beta, t'_0, B)}{L(y|k, \beta, t_0, B)}. \quad (12)$$

Since σ_1 is in general sufficiently large to handle efficient exploration of the range of t_0 , the Markov chain generated may suffer from slow mixing and persistent rejection. To increase the mixing of the chain and to refine the results more locally, we use a delayed rejection step.

For the second stage of the delayed rejection step, we also use a normal PDF as proposal with mean value equal to t_0 and standard deviation σ_2 . The acceptance probability is $\alpha_2(t_0, t''_0) = \min\{1, A_2\}$ with A_2 equal to

$$A_2 = \frac{\{1 - \alpha_1(t''_0, t'_0)\}}{\{1 - \alpha_1(t_0, t'_0)\}} \times \frac{L(y|k, \beta, t''_0, B)}{L(y|k, \beta, t_0, B)}. \quad (13)$$

In general, σ_2 is chosen several orders smaller than σ_1 to allow a more refined local exploration.

4.1.2 Updating the Amplitudes

To update the amplitudes, we use a gamma distribution denoted as q_G with shape parameter equal to 1.5 and scale parameter the current value of the Markov chain as follows:

$$q(\beta', \beta) = \prod_{j=1}^k q_G(\beta'_j | 1.5, \beta_j). \quad (14)$$

The acceptance probability $\alpha_1(\beta, \beta') = \min\{1, A_1\}$ with A_1 equal to

$$A_1 = \frac{L(y|k, \beta', t_0, B)}{L(y|k, \beta, t_0, B)} \times \prod_{j=1}^k \frac{f_G(\beta'_j | a, b) q_G(\beta_j | 1.5, \beta_j)}{f_G(\beta_j | a, b) q_G(\beta'_j | 1.5, \beta_j)}. \quad (15)$$

Since we are using a gamma distribution with scale parameter depending on the current value of the Markov chain, the rate of accepted values decreases as the current value increases. Indeed, if the current value is large, the gamma distribution will be more dispersed, and vice versa. To solve this problem, we apply again delayed rejection using as a second proposal a normal distribution with mean value equal to β and standard deviation σ_3 . Again, σ_3 is chosen to be relatively small compared to the support of the amplitude value. The second stage acceptance probability is $\alpha_1(\beta, \beta'') = \min\{1, A_2\}$ with A_2 equal to

$$A_2 = \frac{L(y|k, \beta'', t_0, B)}{L(y|k, \beta, t_0, B)} \times \frac{\{1 - \alpha_1(\beta'', \beta')\}}{\{1 - \alpha_1(\beta, \beta')\}} \times \prod_{j=1}^k \frac{f_G(\beta''_j | a, b) q_G(\beta'_j | 1.5, \beta'_j)}{f_G(\beta_j | a, b) q_G(\beta'_j | 1.5, \beta_j)}. \quad (16)$$

4.1.3 Updating the Background

To perform this move, we propose a new estimate from a gamma distribution, q_G with shape parameter equal to 1.5 and scale parameter the current value of the Markov chain. The acceptance probability is $\alpha_1(B, B') = \min\{1, A\}$ with A equal to

$$A = \frac{L(y|k, \beta, t_0, B')}{L(y|k, \beta, t_0, B)} \times \frac{f_G(B' | c, d) q_G(B | 1.5, B')}{f_G(B | c, d) q_G(B' | 1.5, B)}. \quad (17)$$

4.2 Variable Dimension Moves

In addition to learning about the parameters of a model, we are interested in comparing different models. Such models could have equal dimension and we may want to find which one explains better the data or we may have a mixture of distributions where the unknown is the number of elements of the mixture or, perhaps, a combination of both [22].

Our problem has a parameter space whose dimension is itself subject to inference, that is, the number of returns is unknown. Therefore, it is necessary to generalize the results obtained in the previous sections to explore the different parameter subspaces in a single simulation.

Using the Bayesian framework, the problem is now formulated as a joint probability distribution of a model indicator, $k \in \{1, 2, 3, \dots, K\}$, and a real stochastic vector of

parameters of possibly varying dimension, ϕ , given the data y , that is, $\pi(k, \phi|y)$. The parameter vector ϕ takes values in the set ξ that defines the general space of solutions, $\xi = \cup_K \xi_k$. Given $K = k$, the parameter vector ϕ can only take values in ξ_k . In general, the set ξ_k is taken to be $\xi_k \subseteq \mathbb{R}^{n_k}$, where n_k is the dimension of the space.

In order to explore the whole space, multiple types of move have to be set. A move type m might use transitions between ξ_i and ξ_j , $i, j \in K$. The new transition kernel has to take into account all the possible moves in such a way that the combination of all these moves will create an ergodic chain. Green [5] proposed the following transition kernel:

$$K(x, A) = \sum_M \int_A q_m(x, dx') \alpha_m(x, x') + s(x) \cdot I(x \in A), \quad (18)$$

where $I(\cdot)$ is the indicator function, and $s(x)$ is the probability of not moving from x .

Using this space of solutions, we propose a move m to $x' = (k', \phi_{k'})$ from $x = (k, \phi_k)$ with probability $j_m(x)$, where ϕ_k is a vector of parameters whose length depends on k . As stated by Green [5], given k' , it may be useful to generate the proposed $\phi_{k'}$ as a deterministic function, $\phi_{k'} = h(\phi_k, u)$, $h: \mathbb{R}^{n_k} \times \mathbb{R}^r \rightarrow \mathbb{R}^{n_{k'}}$, of the current value ϕ_k and a vector of r random components u from a known joint density g on \mathbb{R}^r . In a similar way, the reverse move (from x' to x) can be generated using another deterministic function, $\phi_k = h'(\phi_{k'}, u')$, $h': \mathbb{R}^{n_{k'}} \times \mathbb{R}^r \rightarrow \mathbb{R}^{n_k}$, where, given k , ϕ_k is the proposed parameter vector, and u' is a vector of r' random components from a known joint density g' on $\mathbb{R}^{r'}$. The transformation from $(k', \phi_{k'})$ to (k, ϕ_k) requires that

$$n_k + r = n_{k'} + r', \quad (19)$$

which is known as the ‘‘dimension-balancing’’ condition. As long as ‘‘dimension balancing’’ is satisfied, x and x' can have different dimensions. Moreover, this transformation must be a diffeomorphism, so the change of variable theorem can be used to preserve the ‘‘detailed balance’’ condition

$$\begin{aligned} \pi(x) j_m(x) g_m(u) \alpha_m(x, x') = \\ \pi(x') j_m(x') g_m(u') \alpha_m(x', x) \left| \frac{\partial(x', u')}{\partial(x, u)} \right|. \end{aligned} \quad (20)$$

The Jacobian term $\left| \frac{\partial(x', u')}{\partial(x, u)} \right|$ appears due to the deterministic transformation and is not really an inherent component of the RJMCMC algorithm. The acceptance probability is therefore

$$\alpha_m(x, x') = \min \left\{ 1, \frac{\pi(x') j_m(x') g_m(u') \left| \frac{\partial(x', u')}{\partial(x, u)} \right|}{\pi(x) j_m(x) g_m(u)} \right\}. \quad (21)$$

When $x = (k, \phi_k)$, all the moves cannot be available and, therefore, there is a nonzero probability that no move is attempted.

The birth/death and split/merge steps imply changes in the number of peaks (hence, the number of parameters) of our distribution and, therefore, the reversible jump mechanism is needed. In these two specific cases, the acceptance probability of (21) can be reduced to $\alpha_m(k, \phi, k', \phi') = \min\{1, A\}$, where A is

$$A = \frac{\pi(k', \phi'|y)}{\pi(k, \phi|y)} \times \frac{r_m(\phi')}{r_m(\phi)q(u)} \times \left| \frac{\partial(\phi')}{\partial(\phi, u)} \right|. \quad (22)$$

Let b_k, d_k, s_k , and m_k be the probabilities of a birth, death, split, and merge, respectively, where k is the current number of peaks. Further, assume all the moves are equally probable, the following conditions apply:

- $b_k + d_k + s_k + m_k = 1$.
- If the current number of peaks, k , is equal to 0, that is, only background noise is present, only a birth step is allowed.
- If the current number of peaks, k , is equal to 1, only a birth, death, or split step is allowed.
- If the current number of peaks, k , is equal to a predefined maximum number of peaks, k_{\max} , only a death or a merge step is allowed.

4.2.1 Birth/Death Move

For the birth/death move, we make a random choice between attempting to create a new peak or to delete an existing peak with probabilities b_k and d_k . For a birth, the value of the amplitude β_{k+1} and the position $t_{0_{k+1}}$ of the proposed new peak are drawn from

$$\beta_{k+1} \sim \Gamma \left(1, \left(\frac{1}{i_{\max}} \right) \sum_{i=1}^{i_{\max}} y_i \right), \quad t_{0_{k+1}} \sim U(0, i_{\max}). \quad (23)$$

The proposal ratio is therefore

$$\frac{d_{k+1}}{b_k \cdot (k+1)} \times \{g_1(\beta_{k+1}) \cdot g_2(t_{0_{k+1}})\}^{-1}, \quad (24)$$

where g_1 denotes a gamma PDF with shape parameter equal to 1 and scale parameter corresponding to the mean value of x , and g_2 denotes a uniform PDF on the interval $[0, i_{\max}]$.

From (22), (23), and (24), the expression for A for the birth case can be expressed as

$$\begin{aligned} A = \frac{L(y|k+1, \beta_1, \dots, \beta_{k+1}, t_0, \dots, t_{0_{k+1}}, B)}{L(y|k, \beta_1, \dots, \beta_k, t_0, \dots, t_0, B)} \times \frac{d_{k+1}}{b_k \cdot (k+1)} \\ \times \{g_1(\beta_{k+1})g_2(t_{0_{k+1}})\}^{-1} \times f_G(\beta_{k+1}|a, b) \times \frac{1}{i_{\max}}. \end{aligned} \quad (25)$$

For the corresponding death move, we make a random choice between any existing peaks, the acceptance probability is $\alpha_m(k, \phi, k', \phi') = \min\{1, A^{-1}\}$ with an appropriate change of index and $k' = k - 1$. For the birth and death step, the determinant of the Jacobian is equal to one.

4.2.2 Split/Merge Move

The split/merge move is more complicated. The choice of the deterministic function that links both moves is not straightforward since the ‘‘dimension balancing’’ condition has to be satisfied.

The split proposal begins by choosing a peak k_n at random to split into two new peaks, k_n'' and k_n' , with amplitude and position values equal to (β_n'', t_{0_n}'') and (β_n', t_{0_n}') , respectively, conforming to (26)

$$\begin{aligned} u_1 \sim U(0, 1), \quad \Delta \sim U(0, \Delta_1), \quad \text{where } \Delta_1 = (t_3 - t_0) \\ \beta_n'' = \beta_n u_1, \quad \beta_n' = \beta_n (1 - u_1) \\ t_{0_n}'' = t_{0_n} - u_1 \Delta, \quad t_{0_n}' = t_{0_n} + u_1 \Delta. \end{aligned} \quad (26)$$

The proposal ratio of (22) becomes

$$\frac{m_{k+1}}{s_k \cdot (k+1)} \times \{g_1(u_1)g_2(\Delta)\}^{-1}, \quad (27)$$

TABLE 1
Statistics for the Different Data Sets

Actual distance (cm)	Percentiles (cm)				Mean (cm)	Standard deviation (cm)	RMSE	Range of samples (cm)
	25%	50%	75%	97.50%				
1.7	1.338	1.418	1.492	1.976	1.462	0.169	0.029	0.770
3.2	3.255	3.290	3.325	3.359	3.281	0.059	0.010	0.316
5.2	5.082	5.086	5.089	5.099	5.086	0.006	0.011	0.033
7.2	7.051	7.053	7.055	7.058	7.053	0.003	0.015	0.015
9.2	9.106	9.109	9.110	9.113	9.108	0.003	0.009	0.015
11.2	11.090	11.092	11.094	11.098	11.092	0.003	0.011	0.014
13.2	13.153	13.155	13.157	13.161	13.155	0.003	0.005	0.014
15.2	15.254	15.256	15.257	15.261	15.255	0.003	0.006	0.014
17.2	17.240	17.242	17.243	17.246	17.242	0.003	0.004	0.011
19.2	19.238	19.239	19.241	19.244	19.239	0.003	0.004	0.014
21.2	21.235	21.237	21.239	21.243	21.237	0.003	0.004	0.013
26.2	26.353	26.354	26.356	26.359	26.355	0.002	0.016	0.011
31.2	31.305	31.306	31.308	31.311	31.306	0.002	0.011	0.011
36.2	36.203	36.204	36.206	36.209	36.204	0.003	0.001	0.012
41.2	41.198	41.199	41.201	41.204	41.200	0.002	0.000	0.011
51.2	51.110	51.112	51.113	51.117	51.112	0.002	0.009	0.011
61.2	61.017	61.018	61.020	61.024	61.018	0.003	0.018	0.012
71.2	70.990	70.991	70.993	70.998	70.991	0.003	0.021	0.014

where g_1 denotes the *uniform* PDF, $U(0, 1)$, and g_2 denotes the *uniform* PDF, $U(0, \Delta_1)$.

The acceptance probability of this move is $\alpha(k, \phi, k', \phi') = \min\{1, A\}$, where A is defined as

$$A = \frac{L(y|k+1, \beta', t'_0, B)}{L(y|k, \beta, t_0, B)} \times \frac{m_{k+1}}{s_k \cdot (k+1)} \times \{g_1(u_1)g_2(\Delta)\}^{-1} \times \frac{f_G(\beta'_n|a, b) \times f_G(\beta''_n|a, b)}{f_G(\beta_n|a, b)} \times \frac{1}{i_{max}} \times 2\beta_n u_1 \quad (28)$$

with $\beta' = \{\beta_1, \dots, \beta'_n, \beta''_n, \dots, \beta_k\}$, $t_0 = \{t_{0_1}, \dots, t'_{0_n}, t''_{0_n}, \dots, t_{0_k}\}$, $\beta = \{\beta_1, \dots, \beta_n, \dots, \beta_k\}$, $t_0 = \{t_{0_1}, \dots, t_{0_n}, \dots, t_{0_k}\}$, and $2\beta_n u_1$ is the Jacobian of the transformation from $(\beta_n, t_{0_n}, u_1, \Delta)$ to $(\beta'_n, t'_{0_n}, \beta''_n, t''_{0_n})$. To perform the merge move, we choose peaks that are adjacent in terms of Δ_1 . The acceptance probability is now $\alpha_m(k, \phi, k', \phi') = \min\{1, A^{-1}\}$ with the corresponding change of index and $k' = k - 1$. The following parameter restructuring is also needed:

$$\beta_n = \beta'_n + \beta''_n, \quad t_{0_n} = \frac{t'_{0_n} + t''_{0_n}}{2}. \quad (29)$$

5 EXPERIMENTAL RESULTS

The effectiveness of the RJMCMC algorithm in the processing of TOF signals is demonstrated for some synthetic and real data sets. In general, to process the data, we run the algorithm a determined number of sweeps to explore the full posterior distribution, and we discard all the samples corresponding to the burn-in. Then, we infer the number of peaks as that corresponding to the highest marginal posterior probability, $p(k|y)$. Once we determine the number of peaks, we extract estimates of the parameters from $p(\phi|k, y)$ by setting k equal to the estimated number of peaks. The values of the parameters are estimated as the mean values of the samples that

correspond to such a parameter subspace. The data sets are described for each experiment and a series of algorithm performance measurements are presented. The previously unspecified constants are set as $\sigma_1 = 1,000$, $\sigma_2 = 10$, $\sigma_3 = 50$, $a = 6$, $b = (\max(y)/2)/6$, $c = 1.0001$, and $d = 10,000$ unless otherwise specified. The values of a and b have been chosen to avoid the amplitude taking very small values coming from clutter of unknown origin, whereas the values of c and d have been chosen to represent vague prior knowledge of the background. Finally, the effectiveness of the RJMCMC algorithm for 3D ranging and imaging is demonstrated.

5.1 Algorithm Performance Measurements

The following experiment was performed to assess the ability of our algorithm to distinguish two surfaces separated by a finite distance. For that reason, two retro-reflecting corner cubes were placed at a distance of 330 m from the TCSPC receiver described in [23] and the distance between these surfaces was varied from 1.7 cm to 71.2 cm. To assess the performance of the algorithm, the following procedure was performed:

1. Run 100 independent Markov chains for a specific data set using RJMCMC.
2. For each Markov chain infer the number of peaks, positions, amplitudes, and background.
3. Estimate the distance between the detected peaks.
4. Once the distances for each Markov chain are computed, calculate different statistics.

The maximum number of peaks is constrained in this experiment to two peaks, allowing different scenarios where we detect only one peak or background noise. For each of the data sets, we report the results of 1,000 sweeps, following a burn-in period of 4,000 sweeps. Previous experiments have shown these values to be reliable. The different statistics obtained for this procedure are illustrated in Table 1. We use

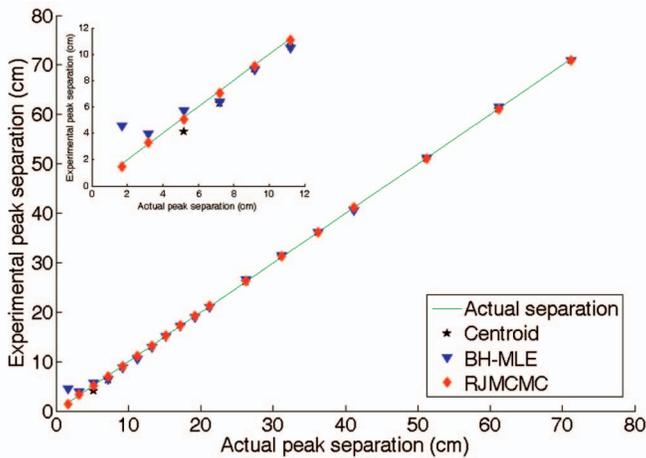


Fig. 3. Calculated separation using different methods versus actual separation, the green line shows the case of perfect agreement. The insert shows an expansion of the graph at the lower surface separations.

the standard deviation to measure the range precision, the root mean square error (RMSE) as a measure of the error relative to the actual distance and the range of the samples to measure the dispersion of the observations. The different percentiles are also calculated to assess where 50 percent of the observations lie and, therefore, to show the presence of any possible outliers since some of the data sets present a distribution of the distance that is not symmetric with respect to their mean value.

In Table 1, it can be observed that the values with more dispersion are the data sets corresponding to distances of 1.7 cm and 3.2 cm. These “large” values compared to those obtained with the other data sets are due to the time response of the sensor, which decreases the range resolution of the system. Indeed, as both surfaces approach, the peaks representing the position of the surfaces begin to merge, and therefore, it becomes increasingly difficult to calculate the separation of the peaks. Statistically, this is expressed as an increase of the range of the samples and the standard deviation in accordance to that in Table 1. As we increase the distance between both surfaces, the results obtained are more consistent with the expected values. The bias in the results is due to two main causes: 1) the accuracy of the ruled gauge is limited to 1 mm and 2) the operational model is not fully optimized for the sensor response. Nevertheless, the statistics calculated are quite stable and consistent even using relatively short data runs.

To fully explore the advantages of our Bayesian approach, we compare the RJMCMC algorithm developed with a basic centroid algorithm and the nonparametric bump-hunting and maximum likelihood estimation (BH-MLE) algorithm described by Wallace et al. [19]. These two algorithms are in essence deterministic as opposed to the stochastic nature of the RJMCMC algorithm. We use the mean value of the generated distances with RJMCMC as a benchmark against that which we compare the results obtained with centroid and BH-MLE. In Fig. 3, it can be observed that the centroid method is able to calculate with reasonable accuracy surface separations of down to 5.2 cm under the assumption of exactly two peaks present. However, below this value, this algorithm fails to detect two different surfaces and is not able to determine the number of peaks. On the other hand, the BH-MLE algorithm is able to calculate smaller separations between

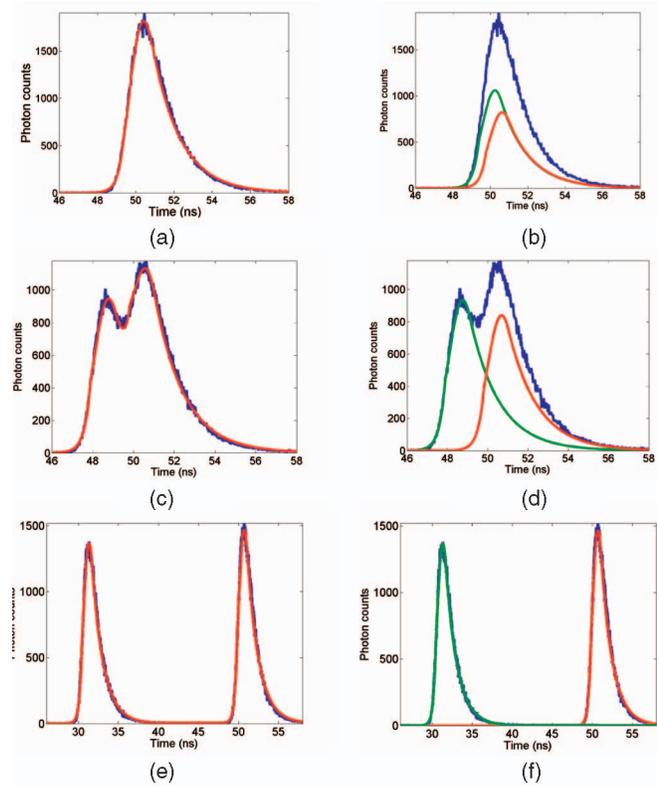


Fig. 4. Some results of the experiment in Section 5.1. Histogram of photon counts of the real data (blue) and final fit from RJMCMC estimation (red) in (a), (c), and (e). Vertical axis represents the number of detected photons, whereas horizontal axis represents time. In (b), (d), and (f) the first return detected is represented in red, whereas the green color represents the second return detected. (a) $d = 1.7$ cm. (b) $d = 1.7$ cm. (c) $d = 7.2$ cm. (d) $d = 7.2$ cm. (e) $d = 71.2$ cm. (f) $d = 71.2$ cm.

surfaces but fails to reasonably estimate the distance between peaks. Figs. 4a, 4c, and 4e show some examples of the final fit from RJMCMC estimation at several distances, whereas Figs. 4b, 4d, and 4f show the different returns detected. Examples of return fitting using BH-MLE algorithm can be found in [19].

5.2 Analysis of Multiple Returns

The purpose of this experiment is to assess the ability of our algorithm to detect multiple returns and show the ability to detect much smaller return amplitudes in comparison with the background level. Again, the results obtained are compared to those obtained with the BH-MLE algorithm. Two different data sets were considered, a synthetic and a real one. The synthetic data provides ground truth to determine whether or not the algorithm is able to infer the exact number of peaks. On the other hand, the real data provides situations where the number of returns is unknown, and the assumption of a unique return is wrong.

In this case, we set the value of the constant $b = 10,000$ to make the prior on the amplitude as vague as possible. The prior on k is taken as uniform on the integers $1, 2, \dots, k_{max} = 20$. For each data set, we compare the model with highest marginal posterior probability, $p(k|y)$, obtained with the RJMCMC algorithm to the best models chosen using AIC and BIC. Model selection using AIC or BIC works by calculating AIC/BIC scores for each candidate model and selecting the model with the lowest AIC/BIC as best. In our case, the number of peaks is considered as the model

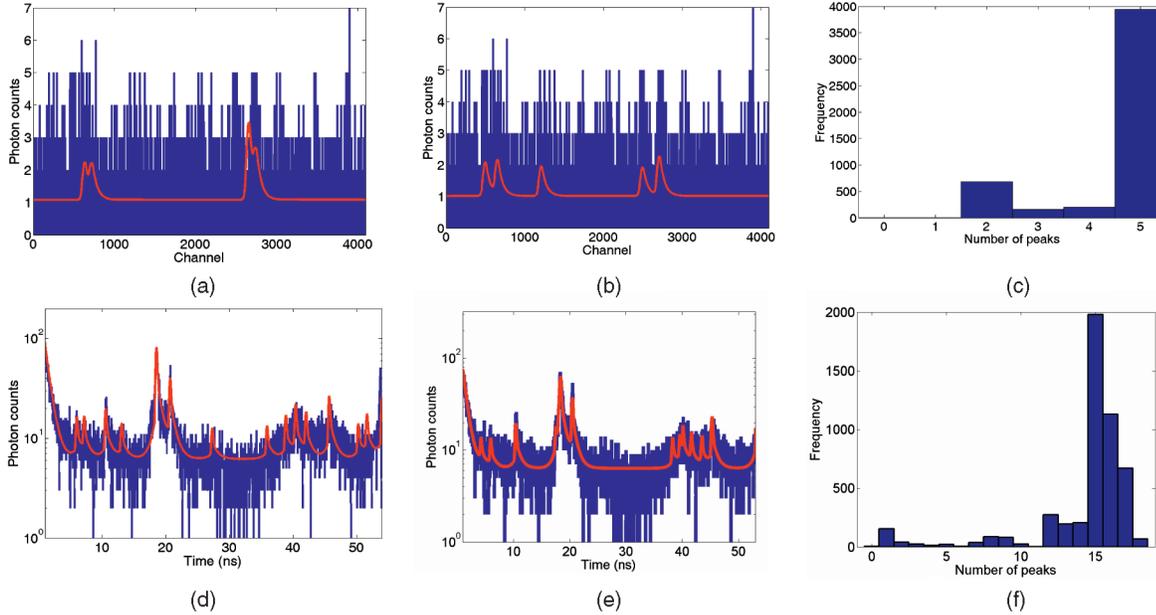


Fig. 5. Analysis of TCSPC data. Histogram of photon counts (blue) and final fit estimation (red) for synthetic (a), (b) and real (d), (e) data. (d) and (e) are represented on the log scale for illustrative purposes. (c) and (f) are histograms of the number of peaks for the synthetic and real data respectively obtained with RJMCMC. (a) BH-MLE. (b) RJMCMC. (c) $p(k|y)$, 5,000 iterations. (d) BH-MLE. (e) RJMCMC. (f) $p(k|y)$, 10,000 iterations.

TABLE 2
Simulated and Estimate Values

	Actual values		BH-MLE estimates		RJMCMC estimates	
	Position	Amplitude	Position	Amplitude	Position	Amplitude
1 st peak	500	1	635.621	1.171	496.326	1.07
2 nd peak	650	1	725	0.851	652.03	1.071
3 rd peak	1200	1	2656	1.267	1207.367	0.944
4 th peak	2500	1	2658	1.123	2494.939	0.909
5 th peak	2700	1	2748	1.015	2707.642	1.239
B	1		1.072		1.014	

indicator. Figs. 5a and 5b show a simulation of five returns using a real target signature as a model placed at arbitrary locations, whereas Figs. 5d and 5e show real data obtained from a complex target with multiple returns due to multiple and distributed or semitransparent surfaces from an optical lens assembly. The marginal posterior distributions of the number of peaks obtained with RJMCMC are shown in Figs. 5c and 5f.

Table 2 shows the actual values and the simulated ones obtained via BH-MLE and RJMCMC of the synthetic data set. For this data set, the different methods consider that the candidate model is that which presents five different peaks. However, the BH-MLE algorithm fails to estimate the true values of the different parameters. Relaxing some of the relevant thresholds used for peak detection helps to find the rest of the peaks; nevertheless, this implies an increase of the number of false returns as well. This shows that even when we do not have any a priori knowledge of the dimension of the parameter space, RJMCMC is able to infer the exact number of peaks, their amplitudes, and their positions. The small variations are due to the existence of strong noise in the original simulated photon count histogram. For the real data, however, the different methods consider different number of peaks and estimates for some of the parameters, as can be seen in Figs. 5d and 5e.

We have also calculated the AIC and BIC scores for the different candidate models using the RJMCMC estimates. In

the RJMCMC case, the highest marginal posterior probability, $p(y|k)$, determines that the number of peaks is 15, whereas the AIC and BIC criteria tend to be biased in favor of a more complicated model in which the number of peaks is 16. This can be explained by Lindley's Paradox [24] that states that it may happen that the log likelihood increases faster than the number of parameters, and therefore, the penalty against overfitting is not large enough.

RJMCMC algorithms are considered CPU-time consuming compared to simpler methods such as the centroid or BH-MLE. However, the complexity of these methods can be greatly reduced if we have a priori information about the signal context. Further, the estimates obtained with these algorithms are more accurate in both variable and fixed dimension as can be stated from the results here presented. Finally, RJMCMC methods allow us to represent the uncertainties inherent in any inference from the data, as shown in Table 1.

5.3 Three-Dimensional Ranging and Imaging

In this experiment, we acquired 3D images of a number of test objects of varying size at different distances, and we applied the RJMCMC method to both TCSPC Lidar data and range gated BIL data. Figs. 6a and 6b show, respectively, a toy fighter, and its corresponding depth image acquired using a TCSPC scanned system. The toy fighter is mounted on a

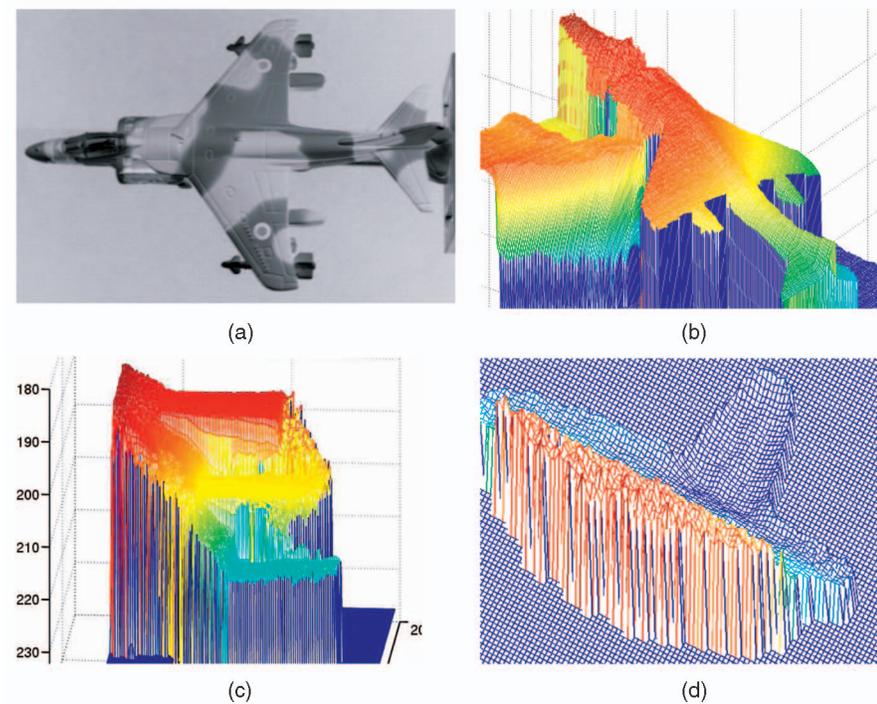


Fig. 6. Details from depth acquired images from a TCSPC system and a BIL system. (a) Image of a toy fighter. (b) Mesh of a toy fighter (distance ~ 3.5 m). (c) Mesh of three planar targets (distance ~ 200 m). (d) Mesh of a trig point (distance ~ 6.6 km).

clamp that is also visible in the depth image. Fig. 6c shows a meshed reconstruction of the BIL image of the three planar targets (shown in Fig. 7a) that gives some idea of the depth variation across the targets. The peaks in the depth data adjacent to the planes are caused by ground reflections before or beyond the planar targets. The mean depths of the three planar targets were measured (by the laser) as 192.05 m, 204.00 m, and 215.00 m, respectively, in comparison with a ruled separation of approximately 12 m between the targets, which is consistent with the results obtained.

Fig. 6d shows a meshed reconstruction from a BIL image sequence of a UK Ordnance Triangulation Point (trig.point) shown in Fig. 7b, located on top of a hill at several kilometers from the sensor. Reflections are also visible from the foreground below the target. Fig. 7c shows a single frame BIL image of the trig. point. Given the distance involved and the turbulence of the atmosphere between the sensor and the trig. point, the 3D mesh provides a very reasonable reconstruction of the 3D geometry of the stone and ground surfaces.

6 CONCLUSION

We have described the development and application of RJMCMC techniques, incorporating delayed rejection, to process time-of-flight data using both single photon counting and variable temporal gating to extract range measurements. The modeling and processing strategy we have presented has been used to explore posterior distributions in different dimensional state spaces, providing a unified treatment for Lidar data acquired from both TCSPC and BIL sensors. One of the main advantages of this strategy is that it allows the user to have uncertainty about the exact dimension of the problem. Indeed, the algorithm explores different dimensionality spaces and then determines the model, which is the closest to the unknown reality. Moreover, the algorithm is able to

infer the values of the parameters corresponding to the model selected. The experiments carried out have demonstrated the ability of this algorithm to determine the number, position, and amplitude of the returns under different signal to background conditions and when the distance between targets is relatively small. It is possible to resolve multiple returns and, hence, characterize objects distributed in 3D space. We anticipate that the results obtained from processing single pixels in isolation will be improved by consideration of interpixel constraints imposed in the focal plane for which algorithms are now being developed.

ACKNOWLEDGMENTS

The authors would like to acknowledge BAE Systems, Sensors and Airborne Systems (SELEX S&AS), and the Royal Society for the support of this work and Ryan Warburton and Gerald Buller (Heriot-Watt University) for providing the raw Time-Correlated Single Photon Counting (TCSPC) data of Section 5.1. The raw Burst Illumination Laser (BIL) data of this paper was provided by BAE Systems and SELEX S&AS.



Fig. 7. (a) Log-intensity single frame BIL image with standoff distance centered on the middle of three uniform, matt planar targets, situated approximately 12 m apart. (b) UK ordinance trig. point. (c) Log-intensity single frame BIL image of the trig. point (long range).

REFERENCES

- [1] M.J. Halmos and M.J.J. Asbrock et al., "3-D Flash Ladar at Raytheon," *Proc. SPIE Laser Radar Technology and Applications VI*, vol. 4377, pp. 84-97, 2001.
- [2] B.W. Schilling, B.N. Dallas, and G.C. Templeton et al., "Multiple-Return Laser Radar for Three-Dimensional Imaging through Obscurements," *Applied Optics*, vol. 41, no. 15, pp. 2791-2799, 2002.
- [3] S. Hernández-Marín, A.M. Wallace, and G.J. Gibson, "Markov Chain Monte Carlo Algorithms for 3D Ranging and Imaging," *Proc. IAPR Conf. Machine Vision Applications*, pp. 193-196, May 2005.
- [4] A.M. Wallace, P. Csakany, G.S. Buller, and A.C. Walker, "3D Imaging of Transparent Objects," *Proc. British Machine Vision Conf.*, pp. 466-475, 2000.
- [5] P.J. Green, "Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination," *Biometrika*, vol. 82, pp. 711-732, 1995.
- [6] S. Richardson and P.J. Green, "On Bayesian Analysis of Mixtures with an Unknown Number of Components," *J. Royal Statistical Soc. B*, vol. 59, pp. 731-792, 1997.
- [7] P.J. Green and A. Mira, "Delayed Rejection in Reversible Jump Metropolis-Hastings," *Biometrika*, vol. 88, pp. 1035-1053, 2001.
- [8] A.M. Wallace, G.S. Buller, and A.C. Walker, "3D Imaging and Ranging by Time-Correlated Single Photon Counting," *Computing and Control Eng. J.*, vol. 12, no. 4, pp. 157-168, 2001.
- [9] M.A. Albota, R.M. Heinrichs, and D.G. Kocher et al., "Three-Dimensional Imaging Laser Radar with a Photon-Counting Avalanche Photodiode Array and Microchip Laser," *Applied Optics*, vol. 41, no. 36, pp. 7671-7678, 2002.
- [10] J. Busck and H. Heiselberg, "Gated Viewing and High-Accuracy Three-Dimensional Laser Radar," *Applied Optics*, vol. 43, no. 24, pp. 4705-4710, 2004.
- [11] S. Duncan, J. Copley, G. Harvey, and D. Humphreys et al., "Advances in Laser Gated Imaging in Airborne Environment," *Proc. SPIE. Infrared Technology and Applications XXXII*, vol. 6206, pp. 4705-4710, Apr. 2006.
- [12] B.F. Aull, A.H. Loomis, and D.J. Young et al., "Geiger-Mode Avalanche Photodiodes for Three-Dimensional Imaging," *Lincoln Laboratory J.*, vol. 13, no. 2, pp. 335-348, 2002.
- [13] S.C. Popescu, R.H. Wynne, and R.F. Nelson, "Measuring Individual Tree Crown Diameter with Lidar and Assessing Its Influence on Estimating Forest Volume and Biomass," *Canadian J. Remote Sensing*, vol. 29, no. 5, pp. 564-577, 2003.
- [14] A.P. Dempster, N.M. Laird, and D. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm (with discussion)," *J. Royal Statistical Soc. B*, vol. 39, pp. 1-38, 1977.
- [15] D.M. Titterton, A.F.M. Smith, and U.E. Makov, *Statistical Analysis of Finite Mixture Distributions*. Wiley, 1985.
- [16] M. Wax and T. Kailath, "Detection of Signals by Information Theoretic Criteria," *IEEE Trans. Acoustic, Speech, and Signal Processing*, vol. 33, no. 2, pp. 387-392, 1985.
- [17] K.P. Burnham and D.R. Anderson, *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer, 2002.
- [18] W.R. Gilks, S. Richardson, and D.J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*. Chapman and Hall, 1995.
- [19] A.M. Wallace, R.C.W. Sung, and G.S. Buller et al., "Detecting and Characterising Returns in a Pulsed Ladar System," *IEE Proc. Vision Image Signal Processing*, vol. 153, no. 2, pp. 160-172, 2006.
- [20] S. Pellegrini, G. Buller, J. Smith, A. Wallace, and S. Cova, "Laser-Based Distance Measurement Using Picosecond Resolution TCSPC," *Measurement Science and Technology*, vol. 11, pp. 712-716, 2000.
- [21] L. Tierney and A. Mira, "Some Adaptive Monte Carlo Methods for Bayesian Inference," *Statistics in Medicine*, vol. 18, pp. 2507-2515, 1999.
- [22] R. Waagepetersen and D. Sorensen, "A Tutorial on Reversible Jump MCMC with a View toward QTL-Mapping," *Int'l Statistical Rev.*, vol. 69, pp. 49-61, 2001.
- [23] A.M. Wallace, G.S. Buller, and R.C.W. Sung et al., "Multi-Spectral Laser Detection and Ranging for Range Profiling and Surface Characterization," *J. Optics A: Pure and Applied Optics*, vol. 7, pp. S438-S444, 2005.
- [24] D.V. Lindley, "A Statistical Paradox," *Biometrika*, vol. 44, pp. 187-192, 1957.

[25] L. Tierney, "A Note on Metropolis-Hastings Kernels for General State Spaces," *The Annals of Applied Probability*, vol. 8, no. 1, pp. 1-9, 1998.

[26] P.H. Peskun, "Optimum Monte Carlo Sampling Using Markov Chains," *Biometrika*, vol. 60, pp. 607-612, 1973.



Sergio Hernández-Marín received the degree in telecommunications engineering from both the Universidad Politécnica de Valencia (UPV), Spain, and the Ecole Nationale Supérieure des Télécommunications Bretagne (ENST-Br), France, in 2003 and the DEA STT option optique (MSc degree in optoelectronics) from ENST-Br the same year. He is currently working toward the PhD degree in Bayesian analysis of Lidar signals using Reversible Jump Markov Chain Monte Carlo (RJMCMC) algorithms. His research interests include stochastic modeling in image processing, 3D computer vision, tracking and object recognition, and Monte Carlo methods. He is a student member of the IEEE and the IEICE.



Andrew M. Wallace received the BSc degree in electrical engineering in 1972 and the PhD degree in 1976 from Edinburgh University. Currently, he is the director of the Joint Research Institute for Signal and Image Processing, based in Edinburgh. His research interests include 3D computer vision, parallel and embedded visual processing, object recognition, and scene understanding. He has published widely on these topics and was a recent recipient of a Royal Society Industrial Fellowship to work on several aspects of 3D image analysis of image data collected from both time-correlated single photon counting and burst illumination imagery. He is a member of the IEE.



Gavin J. Gibson received the BSc degree in mathematics in 1982 and the PhD degree in commutative algebra in 1986 from Edinburgh University. After a year, as a visiting assistant professor at the University of Virginia, he returned to the United Kingdom, where he carried out research in digital signal processing with the Department of Electrical Engineering, Edinburgh University and, later, with Plessey Research Roke Manor. In 1991, he joined Biomathematics and Statistics Scotland before moving to Heriot-Watt University as professor of statistics in the School of Mathematical and Computer Sciences in 2000. His main research interests include Bayesian inference for stochastic models with applications in the study of epidemics and in signal processing. He is fellow of the Royal Society of Edinburgh, the Institute of Mathematics and Its Applications, and the Royal Statistical Society.

► For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.