

Multilayered 3D LiDAR Image Construction Using Spatial Models in a Bayesian Framework

Sergio Hernandez-Marin, *Member, IEEE*, Andrew M. Wallace, and Gavin J. Gibson

Abstract—Standard 3D imaging systems process only a single return at each pixel from an assumed single opaque surface. However, there are situations when the laser return consists of multiple peaks due to the footprint of the beam impinging on a target with surfaces distributed in depth or with semitransparent surfaces. If all these returns are processed, a more informative multilayered 3D image is created. We propose a unified theory of pixel processing for Lidar data using a Bayesian approach that incorporates spatial constraints through a Markov Random Field with a Potts prior model. This allows us to model uncertainty about the underlying spatial process. To palliate some inherent deficiencies of this prior model, we also introduce two proposal distributions, one based on spatial mode jumping and the other on a spatial birth/death process. The different parameters of the several returns are estimated using reversible-jump Markov chain Monte Carlo (RJMCMC) techniques in combination with an adaptive strategy of delayed rejection to improve the estimates of the parameters.

Index Terms—3D reconstruction, burst illumination laser, delayed rejection, LiDAR, Markov random fields, photon counting, reversible jump MCMC, spatial constraints.

1 INTRODUCTION

THE goal of our work is a complete characterization of the surface geometry of 3D surfaces viewed by a laser 3D imaging system. Previously, we applied to single pixels a Bayesian statistical approach based on reversible-jump Markov chain Monte Carlo (RJMCMC) techniques [1] to assess the number, positions and amplitudes of the returned signals from target surfaces [2]. Since each pixel could record returns from multiple surfaces in its field of view, we employed a mixture distribution for each pixel, instead of a “global mixture” [3] that would include all the possible returns for all pixels.

However, most environments of interest contain spatial patterns that are better characterized by models that capture dependencies among pixels. Some pixel configurations are more likely than others since multiple returns can be concentrated in certain regions and completely absent in others. Natural images exhibit strong contextual dependencies in the form of spatial interactions among components. There are many reasons why this occurs, for example, when mapping the pixels to landscape patterns, if a pixel is identified as “water,” it will be most likely surrounded by the same class of pixels. Moreover, different parts of an object are related through geometric constraints. If such a spatial

interaction is well modeled, the classification accuracy can be improved [4].

The contribution of this paper is to include spatial dependencies between pixels within the RJMCMC framework, based on the number of elements of the mixture through the prior distribution. Furthermore, we incorporate two new moves within the RJMCMC algorithm, “spatial mode jumping” and a spatial birth/death process that incorporates spatial information in what is now a mixture of proposal distributions. We also employ a delayed rejection step [5] for variable-dimension setting in an RJMCMC algorithm. Hence, we incorporate spatial constraints in a multilayered image (i.e., an image in which each pixel has several surface returns) for the first time.

In the next section, we discuss why spatial contextual information is important in computer vision and pattern analysis. We also describe briefly two approaches to range measurement and depth imaging by Time-correlated Single Photon Counting (TCSPC) and by Burst Illumination Laser (BIL). In Section 3, we describe our research methodology using spatial constraints in a Bayesian framework. In Section 4, we show how the two proposed moves mentioned previously can be incorporated in an RJMCMC analysis of the posterior distribution obtained using a Bayesian approach. In Section 5, we present results obtained with this methodology. Finally, in Section 6, some conclusions are drawn.

2 BACKGROUND

One of the basic goals of computer vision is to produce a 3D description of any scene from the surrounding world, given its intensity/reflectance or range image(s). Such a description could then be used in a variety of tasks, e.g., terrain mapping, object recognition, medical image analysis, and robot navigation. Three-dimensional reconstruction techniques are typically divided into two families: range-based and intensity-based techniques. Each of these techniques has advantages and limitations [6].

• S. Hernandez-Marin 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: s.hernandezmarin@googlemail.com, 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 1 Aug. 2007; revised 5 Dec. 2007; accepted 29 Jan. 2008; published online 20 Feb. 2008.

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-2007-08-0471.

Digital Object Identifier no. 10.1109/TPAMI.2008.47.

TCSPC and BIL systems (see [2] for more details) are examples of 3D Lidar techniques that can acquire both depth and reflectance images of objects. Examples of subimages acquired by these techniques are shown in Figs. 12 and 13, respectively. Briefly, each pixel records a multivariate measurement that can be considered as an observed photon or intensity histogram which in turn is considered as a sample of a non-normalized statistical mixture distribution. In the histograms created, the returns come not only from the first surface encountered by the projected laser signal but also from subsequent surfaces in its path. There is little work to interpret multiple returns at the same pixel location, but Wagner et al. [7] have attempted to do this. However, there is no background or noise on the examples given in [7], and the pulses are very large. They do not give any explanation of how they explore the reversible, multiparameter space in fitting their Gaussian pulse shapes, this is a significant optimization problem that we address. Finally, their work does not include spatial context, which is a major factor in this paper.

Computer vision is therefore confronted with the inverse problem of recovering surfaces from images. Unfortunately, this is an ill-posed problem in a Hadamard sense [8] and providing a satisfactory solution is a hard issue. Many deterministic and stochastic approaches have been proposed when each pixel consists of a univariate measurement. The common factor in these approaches is the consideration of spatial dependencies among the different pixels. Standard regularization methods [9], [10] are deterministic approaches that restrict the number of possible solutions by imposing the smoothness constraint, which prevents the local behavior of the reconstruction from becoming too disjoint. Although standard regularization methods may lead to satisfactory solutions, they cannot deal with more general problems. Since these algorithms require that the search space has only one local minimum, they cannot deal with discontinuities, and stochastic approaches to the problem are needed. Further, fusion of information from multiple modules, e.g., intensity and range detectors, is not possible [11].

Spatial statistics provides a framework to cope with the problems associated with standard regularization methods and to quantify generic and a priori knowledge. Spatial statistics attempts to develop the best representation of the data set from which the maximum statistically robust information can be extracted. Within this framework, the Bayesian approach has played a predominant role in image analysis and reconstruction, pioneered by Grenander [12], Geman and Geman [13], and Besag [14]. In the Bayesian paradigm, a model of the response of the device, the likelihood, is combined with the prior distribution to obtain the posterior distribution of the quantity of interest. This prior distribution contains knowledge of the phenomenon under observation and information about the local regularities that are believed to exist. The Markov random field (MRF) [15] theory provides a convenient and consistent way to model context dependent entities such as image pixels. Although the posterior probability completely describes the state of certainty about any possible image, it is often necessary to select a single image as the result or reconstruction. One basic use of Bayesian image analysis is therefore to produce point estimates of pixel images. The most obvious Bayesian point estimate is that which maximizes the a posteriori probability [Maximum a posteriori (MAP)], however, if multimodality in the space of solutions is suspected, finding the global maximum is extremely difficult. Other ways to obtain estimates include Markov chain Monte Carlo (MCMC)

methods such as Gibbs sampling or simulated annealing [13], marginal posterior modes (MPM) [16] or deterministic methods such as iterated conditional modes (ICM) [14] or modified expectation maximization (MEM) [17]. Among all these methods, MCMC is the preferred approach for assessing the uncertainty in the estimates. Examples of surface reconstruction using the Bayesian approach are found in many applications such as single-photon emission computerized tomography (SPECT) [17] and ranging [18].

Many spatial problems, however, are inherently multivariate, in that more than one variable is typically measured at each spatial location. Multivariate data analysis allows users to display many different spatial data layers. Mardia [19] introduced a multivariate MRF model for image processing although, initially, this model received little attention due to computational difficulties. In recent years, interest has grown in the analysis of multivariate lattice data with special emphasis on remotely sensed data and more concretely in multispectral image processing [20] in applications of target detection and scene classification and segmentation. Multispectral imaging systems use several spectral bands to acquire data. Assuming we can treat this data as a collection of vector-valued observations on a two-dimensional (2D) lattice, a multivariate MRF (more concretely a multivariate Gaussian MRF) can be used to extract relevant information as proposed by Hazel [21]. However, as the number of bands increases, the data volumes are substantially increased and this makes it difficult to employ multivariate MRF.

In many situations, the observed multivariate data may arise from a time series of measurements. This is the case with TCSPC and BIL data. A simple approach to the problem would be to treat the time series of observations at a given image pixel as independent and identically distributed to that for all other pixels and to analyze them with an appropriate model to understand the underlying theory of the data points. To model such data, one can use a parametric approach [2], making use of appropriate mixture distributions and then use the techniques described therein to obtain estimates of the different parameters. However, such an approach may give suboptimal results because it neglects the correlations between the parameters in neighboring pixels. For instance, returns coming from the same surface are expected to have similar amplitudes, positions, and numbers of peaks. In a frequentist framework, applying EM algorithms to learning spatial models is not straightforward, as there may be context-specific constraints that are often difficult to incorporate. Some examples of these algorithms can be found in [22] and [23]. However, either they do not consider multivariate data or they have problems with an unknown number of dimensions.

Recently, spatial dependencies have been introduced in a mixture distribution to take into account spatial heterogeneity using a Bayesian approach [3], [24]. However, these consider that the different observations come from a unique mixture distribution in which the number of components and some of the parameters (maybe all) are variable. The spatial dependencies were introduced either through the weights of the mixture distribution [24] or through the different allocation variables [3], and posterior inference was performed using RJMCMC algorithms. Initially, one may consider applying a multivariate MRF to model the spatial dependencies among the different parameters of the different mixture distributions. However, it is not clear how

to use these structures in a variable-dimension setting, where the different peaks are continuously created or deleted to better explore the space of solutions.

3 BAYESIAN ANALYSIS OF MULTIPLE LIDAR RETURNS WITH SPATIAL CONSTRAINTS

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

3.1 Model

Consider a rectangular grid of pixels, labeled $m = 1, 2, \dots, N$ at each of which a multivariate measurement $\mathbf{y}^m = (y_1, y_2, \dots, y_t)$ is available. Each multivariate measurement is a histogram of photon counts or intensities in which y_i^m is the value recorded in channel i , $i = 1, \dots, t$. The exact functional forms of the photon count and the intensity histograms are unknown. To interpret both data, we follow initially the same approach as [2]. For the TCSPC data, we have used the parametric form of the expected temporal variation of the photon count distribution, employing four piecewise exponential functions. In this study, we assume that the shape parameters of the returned pulses are fixed and known from an instrumental response. In the case of the BIL system, we use a lookup table of an instrumental function to interpret the intensities histogram. This is acquired from the response of the BIL system to a Lambertian reflecting surface at a similar range to the object of interest.

Since the methodology developed to analyze the different histograms is the same, we use indistinctly the notation f_{op} when referring to the TCSPC operating model or the BIL instrumental response. Furthermore, assuming that several signals can be present in the same histogram and that these signals will be observed against a finite background level, B_m , whose expected value is considered as a constant across all the channels, the observed histogram in pixel m , y_i^m , can be considered as a sample of a statistical mixture distribution with density

$$\mathcal{F}_m(i, k_m, \beta_m, t_{0_m}, B_m) = \sum_{j=1}^{k_m} \beta_{mj} \cdot f_{op}(i, t_{0_{mj}}) + B_m, \quad (1)$$

where k_m is the number of peaks, $\beta_m = (\beta_{m1}, \beta_{m2}, \dots, \beta_{mk_m})$ is a vector of amplitude factors, and $t_{0_m} = (t_{0_{m1}}, t_{0_{m2}}, \dots, t_{0_{mk_m}})$ is a vector of the time of the peak maxima.

If the time resolution is sufficiently fine, the value y_i^m recorded can be considered to be a random sample of a Poisson distribution with intensity equal to $\mathcal{F}_m(i, k_m, \beta_m, t_{0_m}, B_m)$, which depends on the different model parameters

$$y_i^m \sim \text{Poisson}(\mathcal{F}_m(i, k_m, \beta_m, t_{0_m}, B_m)). \quad (2)$$

In order to build up a likelihood function, we make two further assumptions. First, the observations recorded in each channel i of the histogram are conditionally independent given the value of the parameters. Second, the spatial dependencies are included in Bayes' equation through parameters representing spatial interactions so that the different $\mathbf{Y} = (\mathbf{y}^1, \mathbf{y}^2, \dots, \mathbf{y}^N)$ are spatially independent given the parameters. The likelihood for the total array is expressed by the following equation

$$L(\mathbf{Y}|\mathbf{k}, \Omega, T, B) = \prod_{m=1}^N \prod_{i=1}^{i_{\max}} \left\{ e^{-\mathcal{F}_m(i, k_m, \beta_m, t_{0_m}, B_m)} \times \frac{\mathcal{F}_m(i, k_m, \beta_m, t_{0_m}, B_m)^{y_i^m}}{y_i^m!} \right\}, \quad (3)$$

where $\mathbf{k} = (k_1, \dots, k_N)$, $\Omega = (\beta_1, \dots, \beta_N)$, $T = (t_{0_1}, \dots, t_{0_N})$, and $B = (B_1, \dots, B_N)$.

3.2 Bayesian Inference with Spatial Constraints

The objective is inference about the parameters of (3) in order to obtain accurate estimates of the number of peaks, position, amplitude, and background of the returned signals. If we do not have any spatial contextual information, then these unknowns are regarded as drawn from appropriate independent prior distributions. The position parameter, $t_{0_{mj}}$, is drawn from a flat uniform distribution with support in the interval $[1, i_{\max}]$ to reflect in some way our complete prior ignorance of the true position of the peaks. For the prior distribution of scale parameters β_{mj} and B_m , we assume a gamma distribution since the support of our parameters is on \mathcal{R}^+ . The support of the prior for the number of peaks, k_m , is $[0, \dots, k_{\max}]$ with k_{\max} chosen to be suitably large. The full joint prior distribution when no spatial constraints are considered is therefore

$$\pi(k_m, \beta, t_0, B) = \frac{1}{k_{\max}} \left(\frac{1}{i_{\max}} \right)^k f_G(B_m|c, d) \prod_{j=1}^k f_G(\beta_{mj}|a, b), \quad (4)$$

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

In this work, we consider that observations that correspond to nearby locations are more likely to have similar numbers of peaks than observations from locations that are far apart. Therefore, we generalize the prior distribution given by (4), including a penalty function, the Potts model, which discourages adjacent pixels from having different numbers of peaks. The Potts model is a generalization of the Ising model explained in [3] in which the random variable is allowed to have more than two different values. The Potts model has been used previously in image processing applications and in disease mapping applications to model allocation variables. We follow the formulation [3] in which \mathbf{k} is modeled jointly

$$p_1(\mathbf{k}|\psi) \propto e^{\psi \cdot U(\mathbf{k})}, \quad (5)$$

where $U(\mathbf{k}) = \sum_{m \sim m'} I[k_m = k_{m'}]$ are the number of like-labeled neighboring pairs in the configuration $\mathbf{k} = (k_1, k_2, \dots, k_N)$. The parameter ψ is nonnegative and controls the amount of "smoothing," that is, $\psi = 0$ corresponds to a priori spatial independence of the number of peaks, and as $\psi \rightarrow \infty$, we favor patterns, where the numbers of peaks in neighboring pixels tend to be similar. In what follows, ψ is assumed known and will have a fixed value.

The full joint prior distribution incorporating spatial constraints can be modeled as

$$f(\mathbf{k}, \Omega, T, B) = p_1(\mathbf{k}|\psi) \times \prod_{m=1}^N \pi(k_m, \beta_m, t_{0_m}, B_m). \quad (6)$$

Using (3), (6) and Bayes' theorem the target distribution can be expressed as

$$\pi(k, \Omega, T, B|Y) \propto L(Y|k, \Omega, T, B)f(k, \Omega, T, B), \quad (7)$$

since we just have to know the posterior distribution up to a normalizing constant.

As simulating from the joint probability distribution is difficult, we propose to update the different parameters using their respective full conditional distributions. We define a MRF and use a 3×3 second-order neighborhood. The Hammersley-Clifford theorem [15] ensures that we can use the full conditional distribution of (5)

$$p_m(k_m|\dots) \propto \exp\left(\psi \sum_{m' \in \delta_m} I[k_m = k_{m'}]\right), \quad (8)$$

where δ_m denote the neighbors of m . In this way, a rather complex multivariate probability distribution of a MRF can be obtained by successive simulations from the full conditional distributions.

4 RJMCMC METHODOLOGY

The Bayesian models that are proposed in this work are too complex to be amenable to analytical calculations. We use RJMCMC techniques [3] to infer the number, positions, and amplitudes of the returned signals from target surfaces. This allows moves between state spaces with different dimensionality, i.e., different numbers of viewed surfaces, and moves when the dimension is fixed to improve the estimation of the parameters of such subspace, i.e., position (depth) and amplitude (reflectance) of surfaces. RJMCMC provides a natural method for updating beliefs in response to new information and incorporates prior knowledge into the analysis. Indeed, RJMCMC techniques allow us to explore the full posterior distribution of the parameters of the mixture distribution of (1), given the data values Y supplied by the different histograms. The parameters we seek are estimated from the values of a Markov chain whose limiting distribution is a target distribution π . In our case, π arises from the posterior distribution defined by (7), nevertheless, we use simpler versions of it involving only full conditional distributions. The Markov chain constructed involves moves of various types. These are

1. updates to parameters Ω , T , and B ,
2. random birth of a peak,
3. death of a peak,
4. random splitting of a peak into two peaks, and
5. merging of two peaks into a single peak.

To explore the different state spaces, we also incorporate a delayed rejection step in the RJMCMC algorithm, as proposed by Green and Mira [25]. As in the standard RJMCMC algorithms, our problem starts with the formulation of an intractable posterior distribution $\pi(\cdot)$ from which we want to sample and the construction of a Markov chain whose limiting distribution is this posterior distribution. In a Metropolis-Hastings update, such a Markov chain is generated by proposing new values for some or all of the variables drawn from a proposal distribution, and then, an accept/reject decision is made, depending on the generated values and on the form of the proposal and target distributions. In the delayed-rejection strategy, if a candidate

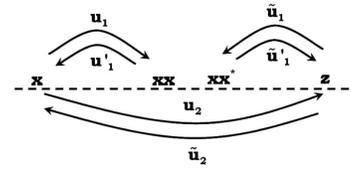


Fig. 1. Schematic adapted from that in [25] representing the variables in play.

move is rejected, we make another attempt to move using a second proposal instead of turning to the next transition. There is no restriction on the number of stages to be used although we shall use only two stages in this work to reduce the computational load.

To perform an RJMCMC with a delayed-rejection stage, three different acceptance probabilities have to be calculated. Using Green and Mira's notation [25] for general state spaces, to move from a state x to a state xx , we propose to draw a random number u_1 from a known density g_1 . The new stage xx is calculated as $xx = h_1^+(x, u_1)$. The reversed move, from xx to x , is performed by drawing u_1' from g_1' and calculating $x = h_1^-(xx, u_1')$, where h_1^+ and h_1^- are deterministic mappings. The acceptance probability of this first stage is given by

$$\alpha_1(x, xx) = \min\left\{1, \frac{\pi(xx)g_1'(u_1')}{\pi(x)g_1(u_1)} \left| \frac{\delta(xx, u_1')}{\delta(x, u_1)} \right| \right\}. \quad (9)$$

If this candidate move is rejected, a new move to state z is proposed in a similar way to the previous candidate move, this move is accepted or rejected with a probability $\alpha_2(x, z) = \min(1, A)$ where A is

$$A = \left\{ \frac{\pi(z)\tilde{g}_1(\tilde{u}_1)\tilde{g}_2(\tilde{u}_2)}{\pi(x)g_1(u_1)g_2(u_2)} \frac{\{1 - \alpha_1(z, xx^*)\}}{\{1 - \alpha_1(x, xx)\}} \left| \frac{\delta(z, \tilde{u}_1, \tilde{u}_2)}{\delta(x, u_1, u_2)} \right| \right\}, \quad (10)$$

and $z = h_2^+(x, u_1, u_2)$ and $x = h_2^-(z, \tilde{u}_1, \tilde{u}_2)$. The values \tilde{u}_1 , \tilde{u}_2 and u_2 are drawn from \tilde{g}_1 , \tilde{g}_2 , and g_2 , respectively.

To obtain the numerator of the second acceptance probability, $\alpha_2(x, z)$, we need to calculate a third acceptance probability, $\alpha_1(z, xx^*)$, corresponding to a fictional stationary Markov chain started in z , which proposes a move to xx at the first stage, rejects it, and accepts a second stage move to x . Although this chain is not really implemented, it is needed to ensure the reversibility condition and therefore preserve the stationary distribution. In the original work by Tierney and Mira [5], it was required that $xx = xx^*$. However, this assumption limited the application of the delayed-rejection algorithm to variable-dimension settings. Fortunately, this is not a required condition to preserve the stationary distribution and alternative values of xx^* can be devised. A schematic diagram of the moves adapted in [25] is presented in Fig. 1. The acceptance probability of this virtual stage is given by

$$\alpha_1(z, xx^*) = \min\left\{1, \frac{\pi(xx^*)\tilde{g}_1'(\tilde{u}_1')}{\pi(z)\tilde{g}_1(\tilde{u}_1)} \left| \frac{\delta(xx^*, \tilde{u}_1')}{\delta(z, \tilde{u}_1)} \right| \right\}, \quad (11)$$

where \tilde{u}_1' is such that $z = h_1^+(xx^*, \tilde{u}_1')$. These different stages do not destroy the Markovian property of the sampler, and therefore, the asymptotic Markov chain theory used for standard MCMC methods holds true. The advantage of this strategy is apparent if the optimal variance of the proposal is not constant throughout the state space. Furthermore, this stage is allowed to depend on the current state of the chain and

also on that just proposed and rejected. Therefore, we learn from the information obtained during the run of the chain.

4.1 Fixed-Dimension Moves

Assume we are at pixel m and the current state of the parameter variable set is ϕ , with $\phi = (k, \Omega, T, B)$. We propose a move n to a state $\phi' = (k', \Omega', T', B')$ drawn from the proposal probability density, $q_n(\phi, \phi')$. The proposed move is accepted with a probability given by (9). If the proposed move is accepted, the new value of the Markov chain becomes ϕ' , otherwise, we keep the current value ϕ . If the number of peaks does not change, i.e., a fixed dimension of the space of solutions, the acceptance probability of (9) reduces to the Metropolis-Hastings acceptance probability, and we perform a parameter updating using the full conditional distributions as a target distribution. In this work, the different parameters Ω , T , and B are updated separately, and as a consequence, each of them has a different acceptance probability. In our case, we draw the proposal of the amplitude, β'_{mj} , and the background, $B'_{m'}$, from the gamma distributions $\Gamma(1.5, \beta_{mj})$ and $\Gamma(1.5, B_m)$, respectively, as in [2]. The proposal for the position is more complicated since it involves spatial information from neighboring pixels.

4.1.1 Spatial Mode Jumping

In practice, the surfaces of the objects imaged by the system that have the same attributes tend to cluster. Hence, groups of pixels with the same number of peaks and similar parameter values are expected to occur together. Spatial interactions for positions are also expected to occur. However, only spatial interactions for the number of peaks are represented in the prior. Furthermore, the Potts model constrains the number of peaks without considering the relative positions and amplitudes of the current pixel with respect to its neighboring pixels. Since such contextual information is not available in the prior, we may force situations in which the state space is not explored properly. Therefore, we suggest a proposal distribution for position moves, which can exploit the fact that we expect data to show correlated positions. To this end, we incorporate a new move, called "spatial mode jumping," within the RJMCMC algorithm. The "spatial mode jumping" move proposes an update to the position of a peak in pixel m so that its position corresponds to a random perturbation of the position of a peak in pixel m' with some probability. In this way, we incorporate spatial information on the proposal distribution.

There are two different ways of incorporating this new move. On one hand, we may consider "spatial mode jumping" as a distinct move of the "position update" move. On the other hand, we may merge the "spatial mode jumping" move and the "position update" move into a single move. Both methodologies have their own advantages and disadvantages. An independent move is easier to implement, however, it is expected that a large number of rejections in the RJMCMC algorithm will be attained. If both moves are merged, a delayed-rejection step can be used every time we reject a "spatial mode jumping" move. This delayed-rejection step will propose an update of the current position parameter instead of a spatial mode jumping. The main disadvantage of this move is the increase of complexity of the acceptance probability. We use the second approach to present a uniform theory of spatial and nonspatial position update. The new

step is therefore based on proposing moves of the current peak in pixel m that depend on its current position and on the current positions of its pixel neighbors, δ_m . One could consider the whole array of pixels as the neighborhood. However, in our implementation, the size of this proposal is limited by the size of the second-order neighborhood of the Potts model. We use a mixture of proposal distributions that takes account of the current values of both the peaks of m and δ_m given by the following expression:

$$Q(\theta, \theta') = w_m \cdot q_m(\theta, \theta') + \sum_{l \in \delta_m} w_l \left(\sum_{t=1}^{k_{max}} \nu_{lt} \cdot q_t(\theta, \theta') \right), \quad (12)$$

where θ is the current estimate of the parameters, θ' are the proposed values, q are proposal distributions, k_{max} is a suitable large number of peaks, and w and ν are weights that satisfy $\sum_{l=1}^9 w_l = 1$ and $\sum_{t=1}^{k_{max}} \nu_{lt} = 1$, respectively. These weights define positive probabilities in such a way that at each step one of the proposals is selected according to these probabilities.

4.1.2 Computational Implementation

In this section, we describe how the "spatial mode jumping" move described in Section 4.1.1 has been implemented using an RJMCMC algorithm with delayed rejection. The key element of this move is the mixture of proposals given by (12). To calculate the acceptance probability, we need first to define the different probabilities involved in the mixture of proposals. Indeed, each component density of the mixture has to be sampled from some probability distribution to generate a proposed value for the Markov chain. The weights w will determine from which pixel we are going to "borrow" some information to propose a new value and depends on the current number of peaks. Assuming we are at pixel m , the following conditions apply:

1. the probability of choosing pixel i' is

$$w_{i'} = \frac{I[k^{i'} \neq 0]}{\sum_{s=1}^9 I[k^s \neq 0]},$$

where k^s is the number of peaks at pixel s , and $I[\cdot]$ is the indicator function,

2. if no peaks are available in the neighborhood to extract some spatial information, $w_m = 1$, $w_{i' \neq m} = 0$, and we propose updates using a random walk without spatial constraints.

On the other hand, the weights ν_l of a pixel l are chosen proportionately to the number of peaks on such pixel. In this work, uniformity on the choice of the peak is chosen for simplicity. More complex weighting schemes based on euclidean distance could be used. However, this will imply some kind of labeling of the different variables, which would complicate the analysis of the data.

Suppose we are updating the values of peak i at pixel m , and we choose pixel i' with some probability determined by $w_{i'}$. If $i' \neq i$, we choose peak j conditioned on i' , with probability $\nu_{i'j}$. The proposed position for peak i , $x x^i$, is drawn from the normal distribution $\mathcal{N}(x^j, \sigma_j)$, where x^j is the current position of the peak j at pixel i' . If no peak from a neighboring pixel is selected, we update the current estimate of the Markov chain using a standard random walk, $x x^i \sim \mathcal{N}(x^i, \sigma_i)$, where x^i is the current position of the

peak i at pixel m . Initially, σ_j can be as small as desired and different for each pixel. To simplify, we set $\sigma_j = \sigma_1, \forall j \neq i$ as we do not have any knowledge of the solution space. In general, we shall make $\sigma_i \gg \sigma_j$. Allowing σ large in the first stage of the algorithm allows better exploration of the state space. The acceptance probability of the first stage is given by $\alpha_1(x^i, xx^i) = \min\{1, A_1\}$ with A_1 equal to

$$A_1 = \frac{\left\{ \frac{\pi(xx^i) \left(w_m \cdot q'_m(xx^i) + \sum_{l \in \delta_m} w_l \left(\sum_{t=1}^{k_{max}} \nu_{lt} \cdot q'_t(xx^i) \right) \right)}{\pi(x^i) \left(w_m \cdot q_m(x^i) + \sum_{l \in \delta_m} w_l \left(\sum_{t=1}^{k_{max}} \nu_{lt} \cdot q_t(x^i) \right) \right)} \right\}}{(13)}$$

where $q_m(\cdot) = q'_m(\cdot) = \mathcal{N}(x^i, \sigma_i)$ and $q_{lt}(\cdot) = q'_t(\cdot) = \mathcal{N}(x^t, \sigma_t)$. This algorithm is valid as it is just a particular case of the Metropolis-Hastings algorithm.

For the delayed-rejection step, we propose to update the current position not considering spatial dependencies. It is also possible to perform a delayed-rejection step, which learns from the previous rejection and proposes another "spatial mode jumping" step. The procedure for the second stage is similar to that of the first stage, although using (10) and not considering the adjacent pixels. Indeed, the proposed value z^i is drawn from the following normal distribution $\mathcal{N}(x^i, \sigma_2)$. In other words, to move from x to z (and respectively from z to x), we perform a standard random walk with standard deviation equal to σ_2 . The value of σ_2 is in general small to allow better local exploration.

As stated before, to calculate the acceptance probability of the second stage, it is necessary to calculate the acceptance probability of the virtual step given by (11). The procedure followed is identical to that of the first stage but starting from state z instead of state xx^* . To simplify (and because we do not perform a change-of-dimension move), we assume that the virtual stage xx^* is equal to the previous rejected step xx . The acceptance probability is given by $\alpha_1(z^i, xx^i) = \min\{1, A_v\}$ with A_v equal to

$$A_v = \frac{\pi(xx^i) \cdot \left(w_m \cdot \tilde{q}'_m(xx^i) + \sum_{l \in \delta_m} w_l \left(\sum_{t=1}^{k_{max}} \nu_{lt} \cdot \tilde{q}'_t(xx^i) \right) \right)}{\pi(z^i) \cdot \left(w_m \cdot \tilde{q}_m(z^i) + \sum_{l \in \delta_m} w_l \left(\sum_{t=1}^{k_{max}} \nu_{lt} \cdot \tilde{q}_t(z^i) \right) \right)}, \quad (14)$$

where $\tilde{q}'_{1l} = \tilde{q}_{1l} = q'_{1l}$ and $\tilde{q}'_m = \tilde{q}_m = q'_m$. Thus, the acceptance probability of the second stage is $\alpha_2(x, z) = \min\{1, A_2\}$:

$$A_2 = \frac{\pi(z) \cdot \tilde{q}_2(\tilde{u}_2) \cdot \{1 - \alpha_1(z, xx)\}}{\pi(x) \cdot q_2(u_2) \cdot \{1 - \alpha_1(x, xx)\}} \times \frac{w_m \cdot \tilde{q}_m(z^i) + \sum_{l \in \delta_m} w_l \left(\sum_{t=1}^{k_{max}} \nu_{lt} \cdot \tilde{q}_t(z^i) \right)}{w_m \cdot q_m(x^i) + \sum_{l \in \delta_m} w_l \left(\sum_{t=1}^{k_{max}} \nu_{lt} \cdot q_t(x^i) \right)}, \quad (15)$$

where $\tilde{q}_2 = q_2 = \mathcal{N}(x^i, \sigma_2)$. Since the dimension-balancing holds in all the equations, this algorithm is a priori valid.

4.2 Variable-Dimension Moves

The birth, death, split, and merge moves imply changes in the parameter-space dimensionality (k changes). In the case of a birth or split move, the acceptance probability of (9) can be reduced to

$$\alpha_m^1(\phi, \phi') = \min \left\{ 1, \frac{\text{Target ratio}}{\pi(\phi|y)} \times \frac{\text{Proposal ratio}}{r_n(\phi)q(u)} \times \frac{\text{Jacobian}}{\left| \frac{\delta(\phi')}{\delta(\phi, u)} \right|} \right\}, \quad (16)$$

where $r_n(\cdot)$ is the probability of choosing move type n when in state ϕ , u is a vector of continuous random variables, which ensures the reversibility of the deterministic function $\phi' = h(\phi, u)$, which allows the move to a higher-dimensional space, and $q(\cdot)$ is the pdf of u . The Jacobian term joint to this deterministic function ensures the "dimension-balancing" condition, which allows reversible jumps between different state spaces. The Jacobian arises from the change of variable from (ϕ, u) to ϕ' . Note that the death and merge moves are obtained by inverting (16) with an appropriate change of index.

The split and merge moves are identical to those explained in [2]. For a split, a peak is chosen at random and divided into two peaks. The split is performed so that the two new peaks equate to the total mass of the peak split. For a merge, the initial peaks to be amalgamated must be within a temporal separation below a certain threshold, and mass balancing is again considered, that is, the total mass of the new peak must be equal to the addition of the total mass of the initial peaks. However, the birth and death process are different in the sense that they also include spatial information and a delayed-rejection step to improve mixing of the chain. For computational reasons, when using a delayed-rejection step, as will be explained in Section 4.2.2, we have split the birth move and the death move into two different kernels, respectively. As a consequence, we consider a birth move, called *0-birth step*,¹ to propose the birth of a peak in situations, where there is only background noise. The *0-birth step* does not incorporate a delayed-rejection step. In the remaining cases, the birth move uses a delayed-rejection step, so that is the reason we distinguish the different birth moves. A similar situation occurs for the death move. In this case, we consider a death move, called *0-death step*, to propose the death of a peak when only a single peak is present. Again, this move does not make use of a delayed-rejection step, whereas this step is used in the remaining cases. Since the different new kernels fulfill the detailed balance condition, as well as aperiodicity and irreducibility, the new mixture of kernels satisfies these conditions and the asymptotic theory of standard Markov chains holds.

Let $b_k^0, d_k^0, b_k, d_k, s_k$, and m_k be the probabilities of a *0-birth*, *0-death*, birth, death, split, and merge, respectively, where k is the current number of peaks. Further, assume that all of them are equally probable. The following conditions apply:

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

1. The 0 in *0-birth step* stands for the current number of peaks whereas in the *0-death step* case, it stands for the updated number of peaks if the move is accepted.

- If the current number of peaks, k , is equal to k_{\max} , only a death or a merge step is allowed.

4.2.1 Spatial Birth/Death Process

A way to penalize the existence of a small or large number of peaks in a histogram is by using a prior on the number of peaks. In general, this requires a very good knowledge of the problem and the use of very subjective priors. In applications with strong spatial dependencies among the different parameters, one includes this fact in the prior distribution, in our case, the Potts model. However, the information contained in the Potts model is not enough, the reason being the lack of information given by this model on which peak to be removed. To deal with this problem, we incorporate a spatial birth/death move into the previously described RJMCMC algorithm, which penalizes, in some way, the creation or removal of a peak given the information of the neighborhood of the pixel. Generally, for a birth, we propose values for the parameters of the new peak. In practice, in the case with nonspatial information, we propose values drawn from a gamma and a uniform distribution for the amplitude and the position of the peak, respectively. For a death, one of the current peaks is chosen at random to be removed. However, in the spatial birth process, the proposed values for the amplitude and the position of the new peak are drawn from two mixture distributions whose elements incorporate spatial information from neighboring pixels in a similar way to that of “spatial mode jumping.” Furthermore, the birth and death moves incorporate a delayed rejection step, which learns from the previous rejected value and therefore allows some kind of learning.

Assume we are at pixel m , the general expression for the mixture of proposals is similar to that of the “spatial mode jumping” move and is given by the following equation:

$$Q(\cdot) = w_m \cdot q_m^1(\cdot) + \sum_{l \in \delta_m} w_l \left(\sum_{t=1}^{k_{\max}} \nu_{lt} \cdot q_t^2(\cdot) \right), \quad (17)$$

where q^1 and q^2 are proposal distributions, k_{\max} is the maximum number of peaks, and w and ν are weights that satisfy $\sum_{l=1}^9 w_l = 1$ and $\sum_{t=1}^{k_{\max}} \nu_{lt} = 1$, respectively. The first term of (17) ensures that a birth move is going to be proposed when no peaks are available in the neighboring pixels, whereas the second term includes spatial information in the proposal distribution. The term q^1 represents a gamma distribution when proposing the amplitude of the new peak, whereas it represents a uniform distribution when proposing the position of such new peak. The term q^2 represents a normal distribution and includes spatial information in a similar spirit to that of the “spatial mode jumping” move. That is, the proposed position/amplitude is drawn from a normal distribution whose mean value is the current position/amplitude of a peak of a neighboring pixel. The different parameters of these distributions will be explained in the next section.

4.2.2 Computational Implementation

In this section, we describe how the “spatial mode jumping” move described in Section 4.2.1 has been implemented using an RJMCMC algorithm with delayed rejection. Key elements of this move are the mixtures of proposals given in a general form by (17). As in “spatial mode jumping,” each component density of the mixture has to be sampled according to some

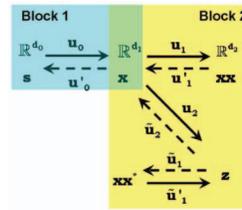


Fig. 2. Relationships among variables for delayed rejection in variable dimensions adapted from that in [25]. Different colors represent different blocks. Block 1 represents the 0-birth and 0-death stages. Block 2 represents the birth and death stages with delayed rejection. All these moves include spatial contextual information. The area where the two blocks are superimposed represents common variables for the moves.

probability to generate a proposed value for the Markov chain. This probability is given by the different weights employed in the mixtures of proposals. The procedure to choose a component of the mixture is identical to that explained in Section 4.1.2. As commented in previous sections, we have split the standard birth and death moves into four different moves 0-birth, 0-death, birth and death. This is because there are situations in which the delayed rejection step is not worth implementing, as in the case when we propose a death of a peak when only one peak is present. Indeed, in a standard death move, a peak is chosen to be deleted uniformly among the peaks contained in such pixel. If only a single peak is available, a delayed-rejection stage would imply persistent rejection of this move. Therefore, this move is computationally nonefficient. As the birth and death processes are needed in tandem, if one of them does not make use of a delayed rejection step, consequently, the other one does not make use of it either. Fig. 2 shows a schematic of the different stages and variables involved when the number of peaks is 0, 1, and 2. Block 2 corresponds to the standard birth and death moves with delayed rejection. The stages in block 2 can be replicated for larger dimensions. On the other hand, block 1 corresponds to the 0-birth and 0-death moves and is only used in transitions from 0 to 1 peak and 1 to 0 peaks.

Suppose we are updating the values of peak i at pixel m , and we choose pixel i' with some probability determined by $w_{i'}$. If $i' \neq i$, we choose peak j conditioned on i' , with probability $\nu_{i'j}$. We propose to have a birth in pixel i at position u_{1p} and with amplitude u_{1a} , where $u_{1p} \sim \mathcal{N}(x_p^{i'j}, \sigma_b)$, $u_{1a} \sim \mathcal{N}(x_a^{i'j}, \sigma_b)$, and $x_p^{i'j}$ and $x_a^{i'j}$ are the position and amplitude of peak j in pixel i' , respectively. An exception occurs when choosing the current pixel i , the proposed values for the amplitude and the position are drawn from the $\Gamma(1, (\frac{1}{i_{\max}}) \sum_{i=1}^{i_{\max}} g_i^m)$ and $\mathcal{U}(0, i_{\max})$ distributions, respectively, which do not incorporate spatial dependencies. This allows us to propose birth and death moves within a pixel in cases, where no peaks have been detected in the neighboring pixels. The acceptance probability of the first stage is given by $\alpha_1^i(k^i, x^i, k^i + 1, xx^i) = \min\{1, A_1\}$ with A_1 equal to

$$A_1 = \min \frac{\pi(k^i + 1, xx^i) \cdot d(k^i + 1) \cdot (1/(k^i + 1))}{\pi(k^i, x^i) \cdot b(k^i) \cdot Q_p(u_{1p}) \cdot Q_a(u_{1a})}, \quad (18)$$

where k^i is the current number of peaks in pixel i , $Q_a(\cdot)$ is the mixture of proposals of the amplitude, $Q_p(\cdot)$ is the mixture of proposals of the position, x^i and xx^i are the parameter vectors of the current and proposed state, respectively, and the

Jacobian term is 1. The death move acceptance probability is the reciprocal of the acceptance probability of the birth move with an appropriate change of index. The expression of this acceptance probability is valid for both block 1 and block 2 with an appropriate change of variables.

With reference to Fig. 2, the second stage corresponding to block 2 is calculated. Again, we need to calculate the acceptance probability of the virtual stage given by (11), which would correspond to a hypothetical death move if we were in state z . Thus, we propose to remove a peak in pixel i with probability $1/(k^i + 1)$, in which case, the acceptance probability is given by $\alpha_1(k^i + 1, z^i, k^i, xx^{*i}) = \min\{1, A_v\}$ with A_v equal to

$$A_v = \frac{\pi(k^i, xx^{*i}) \cdot b(k^i) \cdot Q_p(\tilde{u}_{1p}) \cdot Q_a(\tilde{u}_{1a})}{\pi(k^i + 1, z^i) \cdot d(k^i + 1) \cdot (1/(k^i + 1))}, \quad (19)$$

where \tilde{u}_{1a} and \tilde{u}_{1p} are, respectively, the amplitude and the position of the peak to be removed, z^i and xx^{*i} are the parameters vectors of the proposed state in the second stage and proposed virtual state, respectively, and the Jacobian term is 1.

It only remains to calculate the acceptance probability corresponding to the second stage given by (10). To propose a new birth, we use a rudimentary form of learning in which we avoid proposing peaks of adjacent pixels previously rejected in the first stage. In a similar way, if a peak has been proposed to be removed in a death move, and this move has been rejected, in the second stage, we propose to remove one of the remaining peaks. The complete expression for the second-stage acceptance probability for a birth move is $\alpha_1(k^i, x^i, k^i + 1, z^i) = \min\{1, A_2\}$, where A_2 is

$$A_2 = \frac{\pi(k^i + 1, z^i) \cdot d(k^i + 1) \cdot (1/(k^i + 1))}{\pi(k^i, x^i) \cdot b(k^i) \cdot Q_a(u_{1a}) \cdot Q_p(u_{1p}) \cdot Q_a(u_{2a}|u_{2a} \neq u_{1a})} \times \frac{(1/(k^i)) \cdot \{1 - \alpha_1(z, xx^*)\}}{Q_p(u_{2p}|u_{2p} \neq u_{1p}) \cdot \{1 - \alpha_1(x, xx^*)\}}, \quad (20)$$

and u_{2a} and u_{2p} are, respectively, the amplitude and the position of the peak. The Jacobian term is again equal to 1.

5 EXPERIMENTAL RESULTS

In this section, we illustrate the performance of the Bayesian model with spatial constraints developed on both synthetic and real data sets. All the results displayed correspond to runs of 200 sweeps of the algorithm after a burn in of 800 sweeps. The common constants to the previous approach without spatial constraints are set with the values previously specified [2], whereas the new unspecified constants are set as: $\sigma_i = 5$, $\sigma_1 = 0.01$, and $\sigma_b = 0.1$. As commented in previous sections, the parameters σ_1 and σ_b represent a small random perturbation of the parameters of a peak of a neighboring pixel. For that reason, small values of these parameters are envisaged. The value of the parameter σ_i has been chosen taking into account preliminary experiments. Different values are considered for the smoothing parameter ψ to observe the sensitivity of the results to variations of this parameter. The maximum number of peaks is set to be five for the experiments. We infer the number of peaks as that corresponding to the highest marginal posterior probability, $\hat{k} = \arg\max p(k|\mathbf{y})$. Once we determine the number of peaks,

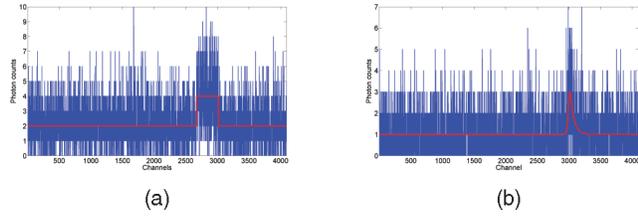


Fig. 3. Different clutter models used for the different synthetic experiments. Only a single pixel for each clutter model is shown. (1 channel ≈ 2.44 ps). (a) Clutter model 1. (b) Clutter model 2.

we extract estimates of the parameters from $p(\phi|k, \mathbf{y})$ by setting $p(\phi|k = \hat{k}, \mathbf{y})$. The values of the parameters are estimated as the mean values of the samples that correspond to such a parameter subspace. The different pixels are accessed randomly instead of using a raster scan of the image to avoid bias.

5.1 Synthetic Data

In this section, we report some simulation results that compare some aspects of the performance of the Bayesian model with spatial constraints developed in this paper to that in [2] without spatial constraints. We use several criteria for comparison of this model [3]. For simulated data, we know the ground truth of the different parameters of the mixture distribution denoted as F_{real} . For each pixel i , we calculate the posterior mean squared error (MSE), $MSE_i = E((F_{real}^i - F_{estimated}^i)^2|\mathbf{y})$, where $F_{estimated}$ is the estimate of the mixture distribution we have obtained. To summarize the performance over the whole array, we compute the root averaged MSE (RAMSE) defined as $RAMSE = (\sum_i MSE_i/N)^{1/2}$. Further, we have computed the deviance information criterion (DIC) described in [26] to compare both models using an approximate decision-theoretic justification. The DIC is composed of two terms: the posterior expected deviance, $E(D|\mathbf{y})$, and a penalty term p_D . In our case, D is the Poisson deviance and is given by $D = 2 \sum_i (y^i \log(y/F_{real}) - (y^i + F_{real}))$, as explained in [27]. p_D can be interpreted as the effective number of parameters in the model and is the difference between the posterior expectation of the deviance and the deviance at the posterior mean of the parameters and represents the model complexity. $E(D|\mathbf{y})$ represents the goodness of fit. Further information on this criterion can be found in the references previously mentioned.

First, we have simulated two 10×10 arrays of photon count histograms, and we have added to them random peaks coming from the two different clutter models shown in Fig. 3. The first clutter model is a rectangular signal that corresponds to uniform clutter of equivalent width to three adjacent peaks using the instrumental model of a single return. The second clutter model is constructed by means of the parametric form of the expected temporal variation of the photon count distribution explained in [2]. Figs. 7 and 8 display the simulated histograms corresponding to a subset of image pixels of the simulated images. In these arrays, each pixel has two different returns whose parameters are known. We have also added clutter to some pixels chosen at random. The clutter is randomly located between channels 2,500 and 4,000. For instance, this clutter can be observed in the pixel located at row 2 and column 1 in Fig. 7 and at row 2 and column 5 in Fig. 8. Figs. 4 and 5 show 2D histogram

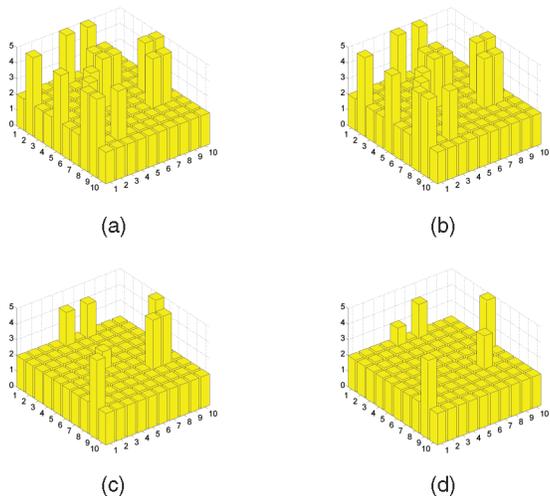


Fig. 4. Estimated number of peaks with and without spatial constraints from simulated data (whole array) for clutter model 1. (a) True position of the clutter (ground truth). (b) No spatial constraints. (c) $\psi = 0.5$. (d) $\psi = 10$.

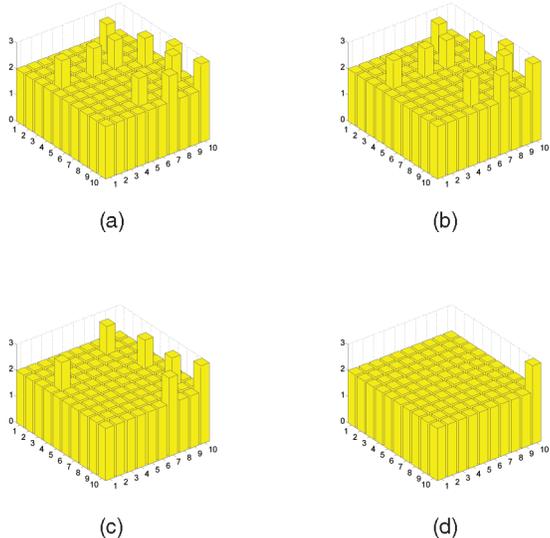


Fig. 5. Estimated number of peaks with and without spatial constraints from simulated data (whole array) for clutter model 2. (a) True position of the clutter (ground truth). (b) No spatial constraints. (c) $\psi = 0.5$. (d) $\psi = 10$.

representations of the estimated number of peaks when no spatial constraints or different values of the parameter ψ are used. It can be observed, when no spatial constraints are used, the unwanted clutter is not ruled out. As the value of ψ increases, there is an effective “smoothing” on the number of peaks, and all the pixels tend to have only two peaks. However, as in this particular case, in the case of the clutter model 1, detecting the exact number of peaks does not mean that we have detected the correct positions and amplitudes. In Fig. 6, we display two scatter plots of the channel position against the intensity value for the actual simulated value, and the values obtained with $\psi = 1$ and $\psi = 20$, corresponding to the different clutter models. Besides variations in the estimated values, it can be observed in Fig. 6a that some of the points accumulate between channel positions 2,500 and 3,000, that is, the area where the clutter was generated. As can be observed, one of the points corresponds to the case where $\psi = 20$, which means that one clutter return was preferred against the true simulated peak. In this case, we have a border

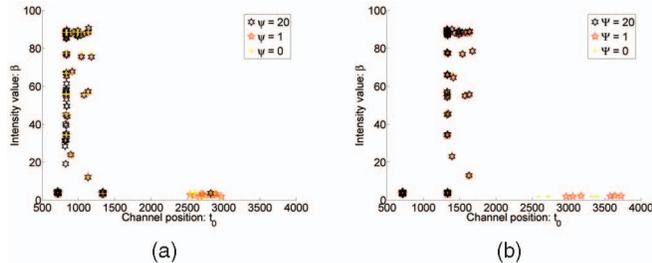


Fig. 6. Scatter plots representing channel position against intensity value for both clutter models. The actual simulated values ($\psi = 0$) and those obtained for the parameters $\psi = 1$ and $\psi = 20$ are represented. (1 channel ≈ 2.44 ps).

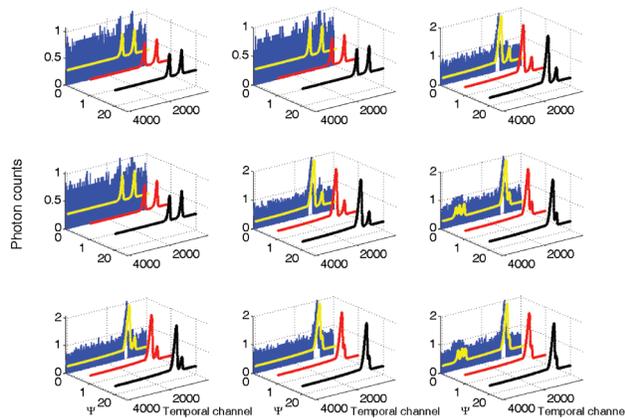


Fig. 7. Array of photon count histograms (in blue) corresponds to a 3×3 subimage of the whole array simulated using clutter model 1 and final fit without spatial constraints (red), with $\psi = 1$ (yellow) and $\psi = 20$ (black). The vertical axis represent photon counts, whereas the horizontal axis represent temporal channels and smoothing parameter.

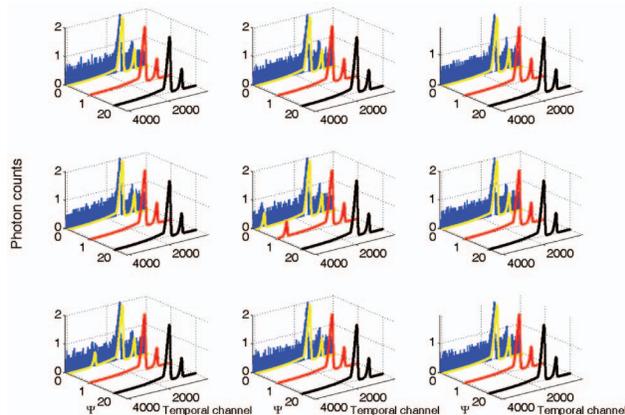


Fig. 8. Array of photon count histograms (in blue) corresponds to a 3×3 subimage of the whole array simulated using clutter model 2 and final fit without spatial constraints (red), with $\psi = 1$ (yellow) and $\psi = 20$ (black). The vertical axis represent photon counts whereas the horizontal axis represent temporal channels and smoothing parameter.

effect in which the pixels in the boundaries of the image have less information from neighboring pixels, which increases the uncertainty associated with estimates in these pixels.

The models with and without spatial constraints are compared on the basis of RAMSE, DIC, $E(D|y)$, and p_D for different values of the parameter ψ . The different results for these criteria are shown in Table 1. As can be seen from the table, the variation of RAMSE between models with spatial constraints is so small that it seems that all are perfectly

TABLE 1
Simulation Results Comparing RJMCMC with and without Spatial Constraints

Ψ	RAMSE		DIC		$E(D y)$		p_D	
	Clutter 1	Clutter 2	Clutter 1	Clutter 2	Clutter 1	Clutter 2	Clutter 1	Clutter 2
No constraints	0.316	0.435	463976.054	468361.072	462727.755	467785.984	1248.299	575.088
0.25	0.21	0.153	469304.319	468540.668	468897.615	468246.846	406.704	293.822
0.5	0.21	0.153	469647.117	468397.444	469154.308	468258.640	492.809	138.804
0.75	0.21	0.153	469760.579	468585.168	469228.735	468139.498	531.844	445.670
1	0.202	0.153	469361.138	468420.520	469290.940	468103.093	70.195	317.427
10	0.201	0.153	470330.748	469431.465	469912.312	469014.382	418.436	417.083
20	0.222	0.153	474119.728	469556.166	473731.597	469132.204	388.131	423.962

plausible. It seems that averaging over the whole array corrects the estimated error calculated by RAMSE since only some pixels present clutter and the remaining peaks are estimated accurately by the algorithm (in both versions with and without spatial constraints). If one accepts the DIC principle as the most adequate model measure since it is applicable to real data and no ground truth is needed, then the model without spatial constraints is the one which best explains the data. These results are consistent since the information criteria rely heavily on the data and do not consider spatial information. The clutter is therefore not

considered as an unwanted return, and the information criteria do not penalize spatial heterogeneities. A similar result is obtained for $E(D|y)$, whereas p_D indicates that the best model is that with $\psi = 1$ for clutter model 1 and $\psi = 0.5$ for clutter model 2.

5.2 TCSPC and BIL Images

In this experiment, we analyze TCSPC and BIL images that correspond to a pixel array of 50×50 histograms of photon counts and reflectance/intensity values, respectively. The TCSPC histograms (see Fig. 12) come from the structure shown in Fig. 9a. The area analyzed corresponds to that shown in the reflectance image in Fig. 9b. This data has been analyzed with the following values for the parameter $\psi = \{0.5, 5, 10\}$ of the Potts model defined by (5). The maximum number of peaks in this experiment was set to 10, and the initial number of peaks was selected randomly. Fig. 10 displays the estimated number of peaks for each value of the smoothing parameter using a color-coded image. As can be seen from the different figures, as the smoothing parameter increases, its value the algorithm tends to form clusters, that

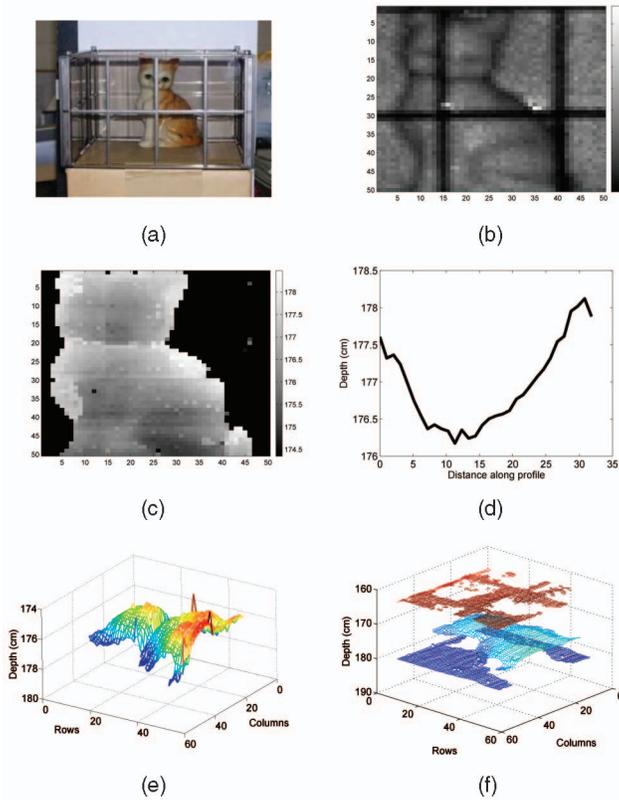


Fig. 9. (a) Image of complex multiple return structure composed of a toy cat and a semitransparent toy greenhouse. (b) Reflectance image obtained with TCSPC system, the unit is photon counts. (c) False color depth image of the toy cat. (d) Depth profile of the toy cat corresponding to row 30. (e) Mesh of the toy cat. (f) Multiple layer structure of the target, only three layers are represented to facilitate representation. A complete layer description may be seen in Fig. 11.

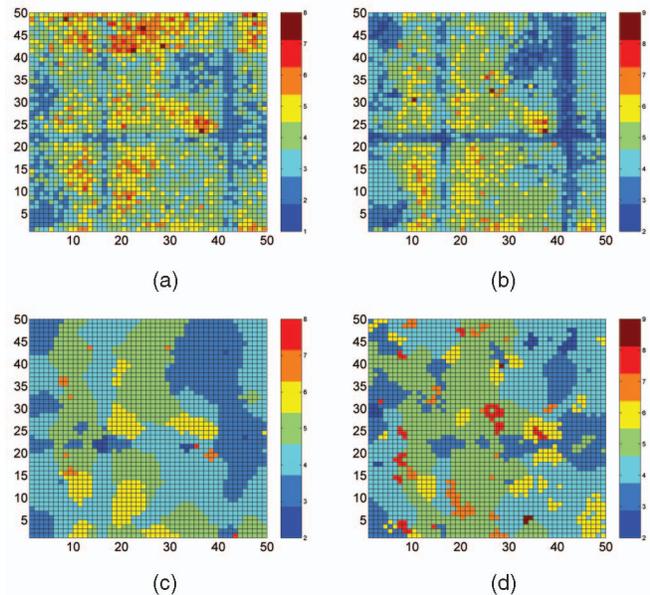


Fig. 10. Estimated number of peaks of the TCSPC data using RJMCMC with and without spatial constraints. (a) No spatial constraints. (b) $\psi = 0.5$. (c) $\psi = 5$. (d) $\psi = 10$.

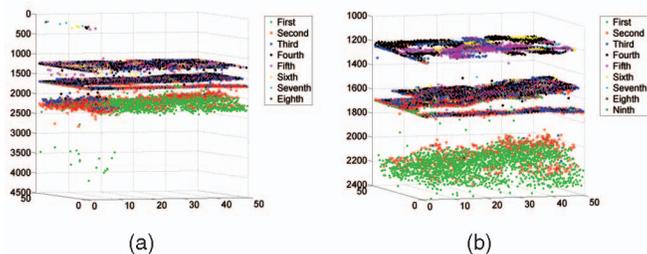


Fig. 11. Multiple-layer structure of the experiment in Fig. 9a with and without spatial constraints. (a) No spatial constraints. (b) $\psi = 10$.

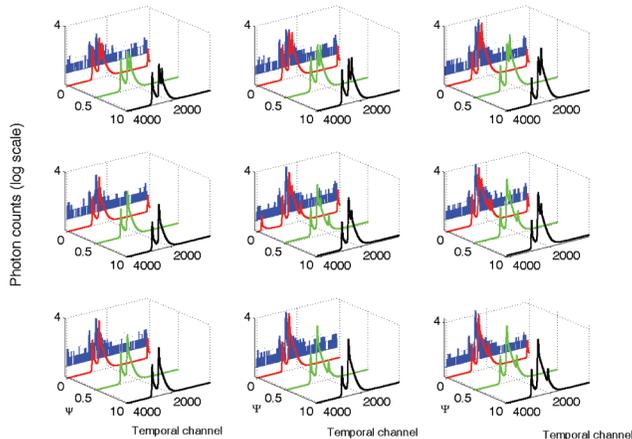


Fig. 12. Array of photon count histograms (in blue) of the structure of Fig. 9a and final fit without spatial constraints (red), with $\psi = 0.5$ (green) and $\psi = 10$ (black). Only nine pixels from an image of 2,500 pixels are represented. The vertical axis represent photon counts, whereas the horizontal axis represent temporal channels. The axis are represented in log scale for convenience.

is, areas with the same number of estimated peaks are connected by paths from neighbor to neighbor. Having seen the scene, initially, one may think that the number of peaks should be between 2 and 4. Indeed, we may have returns coming from the front glass and the toy cat and the front glass, the rear glass, and the wall, and finally, given the beam spread, returns may come from front glass, the toy cat, the rear glass, and the wall. However, this is not the case, as shown in the different figures. The “extra” number of peaks is due to the footprint of the beam impinging on a target with surfaces distributed in depth. Some examples of multilayered images are shown in Fig. 11, where all the returns are represented. When the smoothing parameter is equal to 10 (see Fig. 10d), the spatial interactions are so strong that the algorithm is not able to “kill” some of the unwanted returns. In this case, the posterior distribution is dominated by the prior information; so, the death move does not help to get rid of these peaks.

Fig. 9c shows the toy cat and its depth variation using a false color image, whereas Fig. 9d shows a depth profile of the previous image corresponding to row 30, where we can see that the algorithm captures the curvature of the object. Figs. 9e and 9f show a mesh reconstruction of the toy cat ($\psi = 5$) and a three layer representation of the complete target, where only three returns have been used to facilitate the display. The first layer (closer to the detector) represents the first glass of the toy greenhouse. Note that all the values are not available. This is due to three main factors. First, the total reflected energy of the toy greenhouse is small

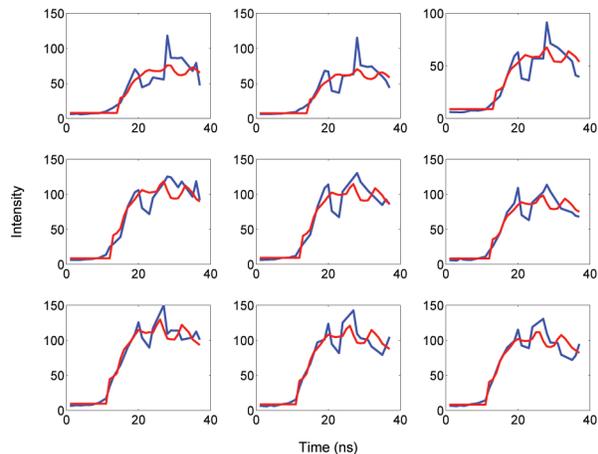


Fig. 13. Array of intensity/reflectance histograms (in blue) of the trig point and final fit using the proposed model. Only nine pixels from an image of 2,500 are represented. The vertical axis represent intensity values whereas the horizontal axis represents non-normalized temporal channels.

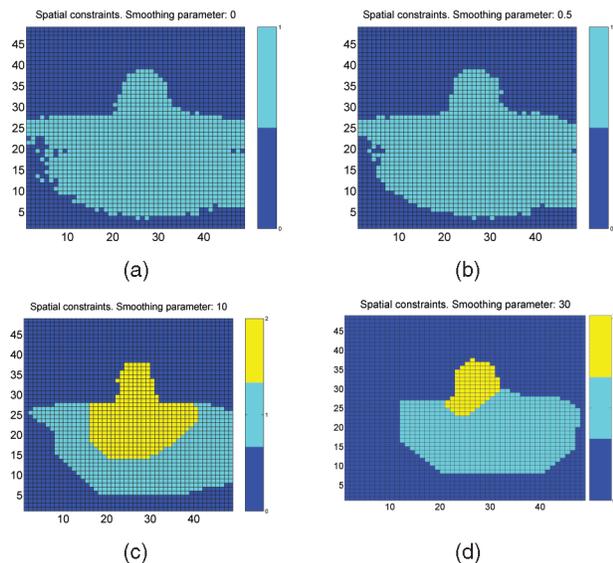


Fig. 14. Estimated number of peaks of the BIL data of the trig point of Fig. 15a.

(semitransparent surface). Second, the returns are well below the background noise. The acquisition time considered was very short and the incidence of the laser was not normal. Third, the prior on k is “oversmoothing” the number of peaks. The second layer corresponds to the toy cat, whereas the third layer corresponds to the second glass of the toy greenhouse. The distance between the pair of glass surfaces is estimated to be about 18.0 cm, which is consistent with the dimensions of the toy greenhouse.

The performance of the model is also illustrated on BIL data. The data set chosen is from a UK Ordinance Triangulation Point (trig point), imaged at approximately 6.6 km, shown in Fig. 15a. This data presents only one unique return, as can be seen in Fig. 13. Thus, we have tested this data under two different assumptions to better understand the behavior of the method. On one hand, only one peak or background was considered. The spatial information was incorporated through the proposal distributions. On the other hand, we allowed up to five possible returns in our calculations. Figs. 14a and 14b display the estimated number of peaks

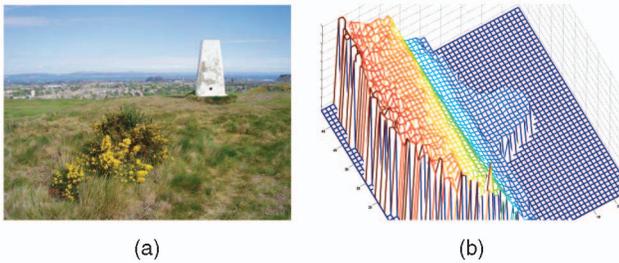


Fig. 15. Details from a trig point (distance ~ 6.6 km) and depth image obtained with a BIL system. (a) UK ordinance trig. point. (b) Mesh of a trig point.

when only background or 1 peak is allowed with a smoothing parameter of 0 and 0.5, respectively. Although both figures look similar, it can be appreciated that Fig. 14b helps to “smooth” the number of peaks in some places, where previously it was considered there was only background. If the smoothing parameter is increased, the results obtained are similar to those obtained with $\psi = 0.5$. Fig. 15b shows a depth reconstruction of the trig. point. Figs. 14c and 14d display the estimated number of peaks when considering up to five peaks, and $\psi = 10$ and $\psi = 30$, respectively. In the first case, it can be observed that three different clusters have been created, one corresponding to only background, another to 1 return, and the last one to two peaks. Since we are using a lookup table as the instrumental function to which we compare the returns, this does not take into account different effects such as the atmospheric turbulence and speckle. One peak is not enough, and the algorithm needs at least two peaks to explain the data in a Bayesian sense. In addition to this problem, when considering $\psi = 30$, a small portion of the image corresponding to the foreground is considered as background. As in the TCSPC case, the likelihood function is dominated by the prior, and therefore, the peaks cannot be added or removed.

6 CONCLUSION

We have described the development and application of RJMCMC techniques incorporating spatial contextual information to process time-of-flight data using both single photon counting and variable temporal gating to extract range measurements. This spatial information was incorporated on the prior distribution through an MRF on the number of peaks and on the proposal distribution through the development of two new moves: the spatial mode jumping and the spatial birth/death processes. Our results show that it is possible to resolve multiple returns and hence characterize objects distributed in 3D space and to extract the underlying spatial structure in the data processed. This spatial information can be used to reduce the influence of clutter of unknown origin by eliminating possible false returns. Further, we have described a delayed-rejection stage, which allows us to perform rudimentary learning and therefore improve the performance of this stage.

ACKNOWLEDGMENTS

The authors would like to acknowledge BAE SYSTEMS, SELEX S&AS, and the Royal Society for the support of this work.

REFERENCES

- [1] 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.
- [2] S. Hernandez-Marin, A.M. Wallace, and G.J. Gibson, “Bayesian Analysis of Lidar Signals with Multiple Returns,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 29, no. 12, pp. 2170-2180, Dec. 2007.
- [3] P.J. Green and S. Richardson, “Hidden Markov Models and Disease Mapping,” *J. Am. Statistical Assoc.*, vol. 97, pp. 1055-1070, 2002.
- [4] B. Tso and P.M. Mather, *Classification Methods for Remotely Sensed Data*. Taylor & Francis, 2001.
- [5] L. Tierney and A. Mira, “Some Adaptive Monte Carlo Methods for Bayesian Inference,” *Statistics in Medicine*, vol. 18, pp. 2507-2515, 1999.
- [6] P. Dias, V. Sequeira, F. Vaz, and J.G.M. Goncalves, “Combining Intensity and Range Images for 3D Modelling,” *Proc. Int’l Conf. Image Processing*, pp. 417-420, 2003.
- [7] W. Wagner, A. Ullrich, V. Ducic, T. Melzer, and N. Studnicka, “Gaussian Decomposition and Calibration of a Novel Small-Footprint Full-Waveform Digitising Airborne Laser Scanner,” *ISPRS J. Photogrammetry and Remote Sensing*, vol. 60, no. 2, pp. 100-112, Apr. 2006.
- [8] T. Poggio, V. Torre, and C. Koch, “Computational Vision and Regularization Theory,” *Nature*, vol. 317, no. 26, pp. 314-319, 1985.
- [9] J.L. Marroquin, “Surface Reconstruction Preserving Discontinuities,” technical report, Massachusetts Inst. of Technology, 1984.
- [10] P. Hall and D.M. Titterton, “On Some Smoothing Techniques Used in Image Restoration,” *J. Royal Statistical Soc. B*, vol. 48, no. 3, pp. 330-343, 1986.
- [11] J. Marroquin, S. Mitter, and T. Poggio, “Probabilistic Solution of Ill-Posed Problems in Computational Vision,” *J. Am. Statistical Assoc.*, vol. 82, no. 397, pp. 76-89, 1987.
- [12] U. Grenander, “Tutorial in Pattern Theory,” technical report, Division of Applied Math., Brown Univ., 1983.
- [13] S. Geman and D. Geman, “Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721-741, 1984.
- [14] J. Besag, “On the Statistical Analysis of Dirty Pictures,” *J. Royal Statistical Soc. B*, vol. 48, pp. 259-302, 1986.
- [15] J. Besag, “Spatial Interaction and the Statistical Analysis of Lattice Systems,” *J. Royal Statistical Soc. B*, vol. 36, pp. 192-236, 1974.
- [16] J. Besag, “Digital Image Processing: Towards Bayesian Image Analysis,” *J. Applied Statistics*, vol. 16, no. 3, pp. 395-407, 1989.
- [17] P.J. Green, “Bayesian Reconstructions from Emission Tomography Data Using a Modified EM Algorithm,” *IEEE Trans. Medical Imaging*, vol. 9, pp. 84-93, 1990.
- [18] G. Zhang and A.M. Wallace, “Physical Modeling and Combination of Range and Intensity Edge Data,” *CVGIP: Image Understanding*, vol. 58, no. 2, pp. 191-220, 1993.
- [19] K.V. Mardia, “Multi-Dimensional Multivariate Gaussian Markov Random Fields with Application to Image Processing,” *J. Multivariate Analysis*, vol. 24, no. 2, pp. 265-284, 1988.
- [20] H. Ren and C.I. Chang, “A Generalized Orthogonal Subspace Projection Approach to Unsupervised Multispectral Image Classification,” *IEEE Trans. Geoscience and Remote Sensing*, vol. 38, no. 6, pp. 2515-2528, 2000.
- [21] G. Hazel, “Multivariate Gaussian MRF for Multispectral Scene Segmentation and Anomaly Detection,” *IEEE Trans. Geoscience and Remote Sensing*, vol. 38, no. 3, pp. 1199-1211, 2000.
- [22] Y. Weiss and E.H. Adelson, “A Unified Mixture Framework for Motion Segmentation: Incorporating Spatial Coherence and Estimating the Number of Models,” *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 321-326, 1996.
- [23] A. Diplaros, T. Gevers, and N. Vlassis, “Skin Detection Using the EM Algorithm with Spatial Constraints,” *Proc. IEEE Int’l Conf. Systems, Man, and Cybernetics*, vol. 4, pp. 3071-3075, 2004.
- [24] C. Fernandez and P.J. Green, “Modelling Spatially Correlated Data via Mixtures: A Bayesian Approach,” *J. Royal Statistical Soc. B*, vol. 64, pp. 805-826, 2002.
- [25] P.J. Green and A. Mira, “Delayed Rejection in Reversible Jump Metropolis-Hastings,” *Biometrika*, vol. 88, pp. 1035-1053, 2001.
- [26] D.J. Spiegelhalter, N.G. Best, B.P. Carlin, and A. van der Linde, “Bayesian Measures of Model Complexity and Fit,” *J. Royal Statistical Soc. B*, vol. 64, no. 4, pp. 583-639, 2002.
- [27] P. McCullagh and J. Nelder, *Generalized Linear Models*, second ed. Chapman and Hall, 1989.



Sergio Hernandez-Marin received the degree in telecommunications engineering from both the Universidad Politecnica de Valencia (UPV), Spain, and the Ecole Nationale Supérieure des Telecommunications Bretagne (ENST-Br), France, in 2003, the DEA STT option optique (MSc degree in optoelectronics) from ENST-Br the same year, and the PhD degree from Heriot-Watt University, Scotland, in 2007. He is currently working on video analytics applications

at the R&D Department, Axis Communications AB, Sweden. His research interests include stochastic modeling in image processing, 3D computer vision, tracking and object recognition, and Monte Carlo methods. He is a member of the IEEE.



Andrew M. Wallace received the BSc and PhD degrees in electrical engineering from the University of Edinburgh in 1972 and 1976, respectively. He is a codirector of the Joint Research Institute in Signal and Image Processing, Edinburgh. His research interests include 3D computer vision, parallel and embedded visual processing, 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 fellow of the IET.



Gavin J. Gibson received the BSc degree in mathematics and the PhD degree in commutative algebra from Edinburgh University in 1982 and 1986, respectively. 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. He joined Biomathematics and Statistics Scotland in 1991 before moving to Heriot-Watt University as a 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 a fellow of the Royal Society of Edinburgh, the Institute of Mathematics and Its Applications, and the Royal Statistical Society.

He is a 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.**