STOCHASTIC EM ALGORITHM FOR FAST ANALYSIS OF SINGLE WAVEFORM MULTI-SPECTRAL LIDAR DATA

Q. Legros, S. McLaughlin, Y. Altmann[*]

School of Engineering and Physical Sciences
Heriot Watt University
Edinburgh, UK

S. Meignen[†]

Laboratoire Jean Kuntzmann
University Grenoble Alpes
Grenoble, France

ABSTRACT

This paper addresses the problem of estimating spectral and range profiles from single-photon Lidar waveforms associated with single surfaces in presence of an unknown background. A single Lidar waveform per pixel is considered, whereby a single detector is used to acquire information simultaneously at multiple wavelengths. A novel Bayesian approach is developed to perform the estimation of model parameters in a reduced computational time. This is achieved by transforming an EM-based algorithm recently proposed into a stochastic EM algorithm, which is computationally more attractive. The reconstruction performance and computational complexity of our approach are assessed through a series of experiments using synthetic data under different observation scenarios. The obtained results demonstrate a significant speed-up compared to the state-of-the-art method, without significant degradation of the estimation quality.

Index Terms— Multispectral imaging, 3D imaging, Single-photon Lidar, Bayesian estimation.

1. INTRODUCTION

Light detection and ranging (Lidar) systems have received considerable interest recently for their ability to reconstruct a 3D scene at high resolution [1, 2]. By sending light pulses onto a target and measuring the time-of-flight (ToF) of the reflected photons, it is possible to estimate the depth profile of an entire scene. More precisely, the ToFs of the recorded photons are gathered to form histograms of photon detection events. Additional detection events caused by ambient illumination and detector dark counts are generally uniformly distributed over the histogram. Single-band Lidar (SBL) allows the reconstruction of high-resolution 3D scenes, but it only captures one intensity value per pixel, associated with the laser source wavelength. Multispectral Lidar (MSL) circumvents such an issue, as a single imaging system [3] is used to collect simultaneously spectral and spatial features [4]. However, the acquisition of such signals requires an overall longer acquisition time (sequential acquisition) or more expensive imaging systems (parallel acquisition) [5, 6].

A recent single-waveform MSL (SW-MSL) approach has been developed in [7], to reconstruct the depth and spectral profiles while keeping the acquisition time and complexity of the imaging system of the same order as SBL. It consists of registering a single histogram containing spectral information from multiple wavelengths for each spatial location, via wavelength-time coding (see Fig 1). Although the temporal delays between wavelengths are known, the overall waveform shape varies as the amplitude of its different peaks depends on the spectral signature of the surface observed. In [7], a full Bayesian method was proposed to analyze SW-MSL signals, which suffers from a prohibitive computational time (about 15 hours to reconstruct a single 3D scene of 200×200 pixels and 4 wavelengths). In [8], we recently proposed a faster alternative where the model parameters are estimated sequentially. More precisely, a mixture model was used as it naturally fits Expectation-Maximization (EM) based algorithms. The resulting algorithm first estimates the mixture weights and subsequently the depth profile. The spectral signatures are computed from the mixture weights using a post-processing step.

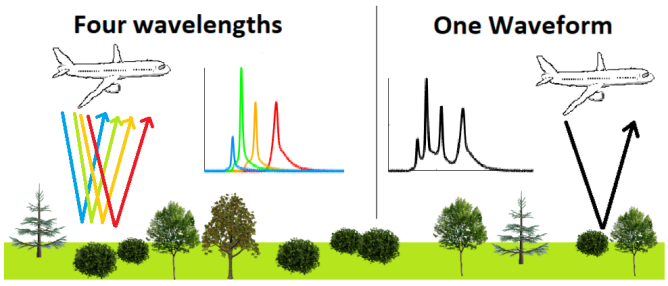

Fig. 1: Example of acquisition of four MSL waveforms, using one wavelength per waveform, and SW-MSL capturing simultaneously information from four wavelengths.

[*]This work was supported by the Royal Academy of Engineering under the Research Fellowship scheme RF201617/16/31 and by the Engineering and Physical Sciences Research Council (EPSRC) (grant EP/T00097X/1)

[†]This work was also supported by the IDEX of University Grenoble Alpes.

In this paper, we extend the work presented in [8] in two aspects. First, we use the traditional Poisson observation noise model and aim at directly estimating the spectral response of each surface, without resorting to the mixtures of distributions used in [8]. The reflectivity profiles are directly assigned prior models, without using additional transforms. Second, we propose a more computationally attractive algorithm, namely a stochastic EM algorithm [9, 10], to estimate the model parameters from single-photon Lidar waveforms. We also provide an additional solution to further accelerate the estimation process, by sub-sampling the grid of admissible depth position. We show that this new method yields a significant speed-up without degradation of the estimation performance when compared to our previous methods.

The remainder of this paper is organized as follows. Section 2 introduces the problem addressed in this work. In Section 3 the new SW-MSL observation model and the different prior distributions associated with the unknown model parameters are defined. The estimation strategy and computational improvement proposal are discussed in Section 4, and the performance of the proposed method is then assessed in Section 5 on synthetic SW-MSL data. Conclusions and future work are finally reported in Section 6.

2. PROBLEM FORMULATION

Consider a 2D observation array $\boldsymbol{Y}$ of SW-MSL signals $\boldsymbol{y}_n = [\boldsymbol{Y}]_{n,:} = [y_{n,1}, \ldots, y_{n,T}]^\top$ of size $N \times T$, where N is the number of pixels or spatial locations and T is the length of the ToF histograms. The element $y_{n,t}$ is the photon count in the nth pixel in the tth temporal bin. The spectral response of the object in the nth pixel is denoted by $\boldsymbol{r}_n = [r_{n,1}, \ldots, r_{n,L}]^\top \in \mathbb{R}_+^L$, where L stands for the number of spectral bands, and $b_n \geq 0$ gathers all the background contributions in that pixel. In the low-flux regime, and in the presence of a single surface per pixel, $y_{n,t}$ follows the Poisson distribution [7]

$$y_{n,t}|\boldsymbol{r}_n, b_n, t_n \sim \mathcal{P}\left(b_n + \sum_{\ell=1}^{L} r_{n,\ell} g_\ell(t - t_n)\right), \quad (1)$$

where t_n is the characteristic ToF of photons emitted by the laser source and reflected by an object at distance d_n of the sensor and $\{g_\ell(.)\}_\ell$ are the instrumental/impulse response functions (IRFs) associated with the L wavelengths, and whose shape can differ between spectral channels. These IRFs are assumed to be known as they are generally estimated during the system calibration. Moreover, the integral $G_\ell = \sum_{t=1}^{T} g_\ell(t - t_n)$ is constant for any admissible t_n.

The problem addressed in this paper is a computationally attractive estimation of $\boldsymbol{t} = \{t_n\}_n$ and $(\boldsymbol{R}, \boldsymbol{b}) = (\{\boldsymbol{r}_n\}_n, \{b_n\}_n)$ in presence of possibly high background illumination.

3. BAYESIAN MODEL

3.1. Likelihood

Assuming that the photon counts in all bins and histograms are mutually independent (conditioned on there mean), the joint likelihood can be expressed as

$$p(\boldsymbol{Y}|\boldsymbol{R}, \boldsymbol{b}, \boldsymbol{t}) = \prod_n p(\boldsymbol{y}_n|\boldsymbol{r}_n, t_n, b_n). \quad (2)$$

Since maximum likelihood estimation (MLE) based on Eq. (2) is generally sensitive to noise, the problem has to be regularized in order to perform robust estimation as discussed below.

3.2. Prior model

Natural scenes present depth and spectral profiles which are often spatially correlated. As in previous work, e.g., [7, 11–13] we model separately the depth and reflectivity profiles and neglect potential correlation between them to keep the inference process tractable.

Depth prior model: The range profile of the target is assumed to belong to a discrete grid of admissible positions $\mathbb{T} = \{t_{min}, \ldots, t_{max}\}$ such that $1 < t_{min} < t_{max} < T$ as in [14, 15]. As in [11, 12], we define a prior model for $\boldsymbol{t}$ which preserves sharp edges, i.e., we use the following total-variation (TV) Markov random field (MRF),

$$p(\boldsymbol{t}|\epsilon) = \exp\left[-\epsilon \|\boldsymbol{t}\|_{TV}\right], \quad (3)$$

where the fixed (user-defined) hyperparameter ϵ, which mostly depends on the range resolution of the single-photon detector, determines the level of correlation between the depth parameters of neighboring pixels. Here we fixed $\epsilon = 0.05$ for all the experiments performed in this paper.

Reflectivity prior model: The reflectivity profile is also assumed to present spatial correlation as neighboring pixels are likely to belong to the same surface. We propose a hierarchical model based on the gamma distribution as in [16]. In a similar fashion to [8] we assume that the pixels can be clustered into C distinct and known classes and that the pixels within each class share the same prior model. The joint prior model for $\boldsymbol{R}$ is then given by

$$f(\boldsymbol{R}|\boldsymbol{\Theta}) = \prod_{c=1}^{C} \prod_{n \in I_c} \prod_{\ell=1}^{L} \mathcal{G}(r_{n,\ell}; k_{c,\ell}, \theta_{c,\ell}), \quad (4)$$

where $\mathcal{G}(\cdot; k_{c,\ell}, \theta_{c,\ell})$ is the probability density function of the gamma distribution with respective shape and scale parameters $(k_{c,\ell}, \theta_{c,\ell})$, $\boldsymbol{\Theta} = \{k_{c,\ell}, \theta_{c,\ell}\}_{c,\ell}$ and I_c is the set of pixel indices in the cth group.

Weakly informative independent truncated gamma (resp. inverse-gamma) hyper-prior models are finally assigned to the

shape (resp. scale) parameters as follows

$$f(\boldsymbol{\Theta}) = \prod_{c,l} \mathcal{G}(k_{c,\ell}; 2, 0.5)\mathbf{1}_{(1,\infty)}(k_{c,\ell})\mathcal{IG}(\theta_{n,\ell}; 1.01, 0.5). \quad (5)$$

In practice the C classes are often unknown. Here they are set as in [8] by applying patch-based k-means clustering on $\boldsymbol{R}$ after a few iterations of the algorithm using a single class. In all the experiments performed in this work, the segmentation was done on the third iteration of the algorithm, as the current estimates were sufficiently informative to cluster the reflectivity image in our experiments, and patches of size (3×3) are clustered in $C = 7$ classes.

Background prior model: The background image may present spatial structure but can also be significantly different from $\boldsymbol{R}$. In the high-background regime, the likelihood is relatively informative and the background image $\boldsymbol{b}$ does not need to be strongly regularized. For this reason, we simply assign a separable prior (with respect of $\boldsymbol{b}$)

$$f(\boldsymbol{b}|\alpha, \beta) = \prod_{n} \mathcal{G}(b_n|\alpha, \beta), \quad (6)$$

and (α, β) is assigned the same prior model $f(\alpha, \beta)$ as in (5).

4. ESTIMATION STRATEGY

Using the likelihood and prior models defined in Section 3, we obtain

$$f(\boldsymbol{R}, \boldsymbol{b}, \boldsymbol{t}, \boldsymbol{\Phi}|\boldsymbol{Y}) \propto p(\boldsymbol{Y}|\boldsymbol{R}, \boldsymbol{b}, \boldsymbol{t})p(\boldsymbol{t})f(\boldsymbol{R}, \boldsymbol{b}|\boldsymbol{\Phi})f(\boldsymbol{\Phi}), \quad (7)$$

where $\boldsymbol{\Phi} = \{\boldsymbol{\Theta}, \alpha, \beta\}$, $f(\boldsymbol{R}, \boldsymbol{b}|\boldsymbol{\Phi}) = f(\boldsymbol{R}|\boldsymbol{\Theta})f(\boldsymbol{b}|\alpha, \beta)$ and where all the fixed model parameters have been omitted in the conditional distributions to simplify notations.

The likelihood in Eq.(2) is multimodal (because of $\boldsymbol{t}$) and the joint estimation of $\boldsymbol{t}$ and $(\boldsymbol{R}, \boldsymbol{b}, \boldsymbol{\Phi})$ is challenging within a reasonable computational time. Thus, we propose as in [8,13] to decompose the estimation process into successive problems, by first considering $\boldsymbol{t}$ as a nuisance parameter which is marginalized to simplify the estimation of $(\boldsymbol{R}, \boldsymbol{b}, \boldsymbol{\Phi})$ such as

$$\begin{aligned}(\hat{\boldsymbol{R}}, \hat{\boldsymbol{b}}, \hat{\boldsymbol{\Phi}}) &= \underset{\boldsymbol{R},\boldsymbol{b},\boldsymbol{\Phi}}{\operatorname{argmax}} \sum_{\boldsymbol{t}} f(\boldsymbol{R}, \boldsymbol{b}, \boldsymbol{t}, \boldsymbol{\Phi}|\boldsymbol{Y}) \\ &= \underset{\boldsymbol{R},\boldsymbol{b},\boldsymbol{\Phi}}{\operatorname{argmax}} f(\boldsymbol{R}, \boldsymbol{b}, \boldsymbol{\Phi}|\boldsymbol{Y}). \quad (8)\end{aligned}$$

The depth $\boldsymbol{t}$ is then estimated via marginal maximum a posteriori (MMAP) estimation using the estimates from Eq. (8)

$$\hat{t}_n = \underset{t_n}{\operatorname{argmax}} \sum_{\boldsymbol{t}_{\setminus n}} p(\boldsymbol{t}|\boldsymbol{Y}, \hat{\boldsymbol{R}}, \hat{\boldsymbol{b}}, \hat{\boldsymbol{\Phi}}), \quad \forall n, \quad (9)$$

where $\tilde{\boldsymbol{t}}_{\setminus n}$ stands for the vector $\tilde{\boldsymbol{t}}$ whose nth element has been removed. The next sections detail how Eqs. (8) and (9) are obtained.

4.1. Estimation of $(\boldsymbol{R}, \boldsymbol{b}, \boldsymbol{\Phi})$

The estimation of $(\boldsymbol{R}, \boldsymbol{b}, \boldsymbol{\Phi})$ can be performed as in [13] using a stochastic EM (SEM) algorithm capable of handling the intractable expectation step induced by the MRF prior in (3). However, here we adopt a generalized SEM approach to obtain a more computationally attractive inference process. The main modifications with respect to the algorithm presented in [8] are twofold. Let $(\boldsymbol{R}^i, \boldsymbol{b}^i, \boldsymbol{\Phi}^i)$ be the current estimate of $(\boldsymbol{R}, \boldsymbol{b}, \boldsymbol{\Phi})$. Instead of using a mean field-like approximation [17] in the E-step, we simulate a set of N_t samples $\{\bar{\boldsymbol{t}}^1, \ldots, \bar{\boldsymbol{t}}^{N_t}\}$ via Markov chain Monte Carlo (MCMC) sampling from $p(\boldsymbol{t}|\boldsymbol{Y}, \boldsymbol{R}^i, \boldsymbol{b}^i, \boldsymbol{\Phi}^i)$. Note that at each iteration, the Gibbs sampler used is hot-started with the samples generated at the previous iteration to improve the convergence speed. Sampling from $p(\boldsymbol{t}|\boldsymbol{Y}, \boldsymbol{R}^i, \boldsymbol{b}^i, \boldsymbol{\Phi}^i)$ is, in addition to the likelihood computation, one of the most computationally expensive step of the algorithm, even if the sampler is initialised properly, as it requires inverting a cumulative density function (CDF) defined on $\mathbb{T}$ for each pixel. To reduce this cost, it is possible to downsample $\mathbb{T}$ by a factor T_s and compute the CDFs on a coarser grid without affecting significantly the quality of the N_t simulated depth samples as well as the overall estimation quality.

After generation of the N_t samples, the cost function to be maximized in the M-step becomes
$Q(\boldsymbol{R}, \boldsymbol{b}, \boldsymbol{\Phi}|\boldsymbol{R}^i, \boldsymbol{b}^i, \boldsymbol{\Phi}^i) =$

$$\frac{1}{N_t} \sum_{m=1}^{N_t} \left[\log p(\boldsymbol{Y}|\boldsymbol{R}, \boldsymbol{b}, \bar{\boldsymbol{t}}^m)\right] + \log f(\boldsymbol{R}, \boldsymbol{b}|\boldsymbol{\Phi})f(\boldsymbol{\Phi}). \quad (10)$$

Instead of exactly maximizing this function at each iteration, in a generalized EM fashion [9], we used a few iterations of gradient ascent (resp. Newton-Raphson) to update $(\boldsymbol{R}, \boldsymbol{b})$ (resp. $\boldsymbol{\Phi}$) sequentially. A gradient ascent was preferred to second order methods for $(\boldsymbol{R}, \boldsymbol{b})$ to avoid the inversion of Hessian matrices, which are heavy to compute and whose complexity depends on L. We fixed the gradient step size to 10^{-3} as it gave the best trade-off between speed and performance in our experiments. When the number of signal photons is not to small compared to the background counts, our generalized SEM algorithm converges quickly and an average of 5 iterations are usually sufficient for the successive estimates to oscillate around fixed values. We assumed the convergence is reached when the relative error between two successive estimates of $\boldsymbol{R}$ falls below 10^{-2}. The final estimates $(\hat{\boldsymbol{R}}, \hat{\boldsymbol{b}}, \hat{\boldsymbol{\Phi}})$ are obtained by averaging the next 5 successive estimates of $(\boldsymbol{R}, \boldsymbol{b}), \boldsymbol{\Phi}$ after convergence.

4.2. Depth estimation

After convergence of generalized SEM algorithm proposed to estimate $(\boldsymbol{R}, \boldsymbol{b}, \boldsymbol{\Phi})$, the final step of the algorithm consists of estimating $\boldsymbol{t}$ using (9). A MMAP estimates is computed from a variational approximation of $p(\boldsymbol{t}|\boldsymbol{Y}, \hat{\boldsymbol{R}}, \hat{\boldsymbol{b}}, \hat{\boldsymbol{\Phi}})$ using

the samples generated during the last iteration of the generalized SEM algorithm. As such solution is computationally expensive for the reasons mentioned above, we fixed $N_t = 1$ for all the experiments unless otherwise stated. As in [8], this approximation can be expressed as a product of independent distributions over the N pixels, whose maximum can be computed easily.

5. RESULTS

We assess the performance of the proposed approach using synthetic SW-MSL data, generated with $N = 200 \times 200$ pixels depth and reflectivity profiles obtained from real data in [7], with $L = 4$ wavelengths ($473, 532, 589$ and 640nm). The number of temporal bins is $T = 1500$ (bin width of 2ps) and a spatially constant background was added to all of the waveforms. We set the admissible temporal positions to $[t_{min}, t_{max}] = [301, 900]$ to ensure the integrals of the IRFs remain constant over the admissible object range. Different scenarios have been reproduced to assess the performance of the proposed method through both the mean signal photon counts (MSC) and the average signal to background ratio (SBR) defined as

$$\text{MSC} = \frac{1}{N}\sum_{n,\ell} r_{n,\ell} G_\ell, \quad \text{SBR} = \frac{1}{N}\sum_{n=1}^{N} \frac{1}{T b_n} \sum_{\ell=1}^{L} r_{n,\ell} G_\ell.$$

The proposed approach, referred as C-Gamma (cluster-Gamma), is compared to the C-Dirichlet method proposed in [8]. To assess the speed-up/performance trade-off when sub-sampling the admissible depth grid or when estimating the depth using more samples, we considered several versions of our proposed method with a downsampling factor $T_s \in [4, 10]$. We also performed experiments with C-Gamma where $N_t = 100$ samples are generated in the S-step. These approaches are referred as C-Gamma$\{T_s, N_t\}$, when either T_s or N_t differ from 1.

We assess the reflectivity estimation quality using the mean squared error (MSE) defined as $\frac{1}{N}\sum_{n=1}^{N} \|\boldsymbol{r}_n - \hat{\boldsymbol{r}}_n\|_2^2$, where $\boldsymbol{r}_n$ (resp. $\hat{\boldsymbol{r}}_n$) is the actual (resp. estimated) spectral response of the nth pixel. The ranging performance is then quantified using the CDF of the depth absolute error $h_n = |t_n - \hat{t}_n|, \forall n$, where $\hat{t}_n$ is the estimate of the actual depth parameter t_n in the nth pixel. The reconstruction performance of the reflectivity (resp. depth profile) obtained with $[10, 44]$ MSC photons are depicted in Fig.2 (resp. Fig.3). Note that modifying T_s does not significantly change the reflectivity reconstruction performance for $N_t = 1$, so C-Gamma$\{4, 1\}$ is omitted in Fig. 2.

Increasing T_s implies estimating the reflectivity with a coarse depth profile at each iteration, but as T_s remains lower than the delays between each IRF, and more importantly as long as the resolution of the coarse grid resolution stays smaller than the width of the IRFs, this does not impact $\hat{\boldsymbol{R}}$ as shown in Fig. 2. Although this affect the final depth reconstruction performance, the loss is not important for $T_s \leq 4$.

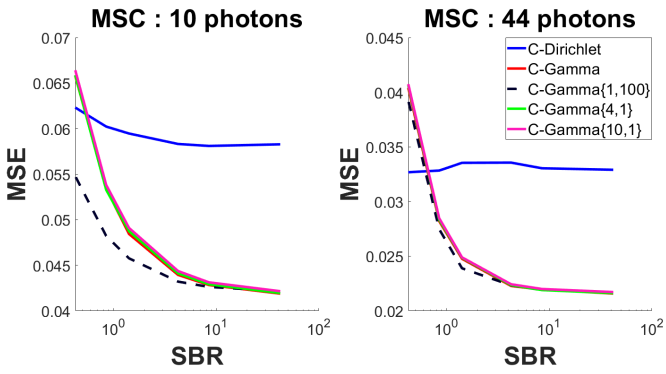

Fig. 2: MSE of $\boldsymbol{R}$ obtained with the different competing methods as function of the SBR. The left (resp. right) subplot corresponds to MSC equals to 10 (resp. 44) photons.

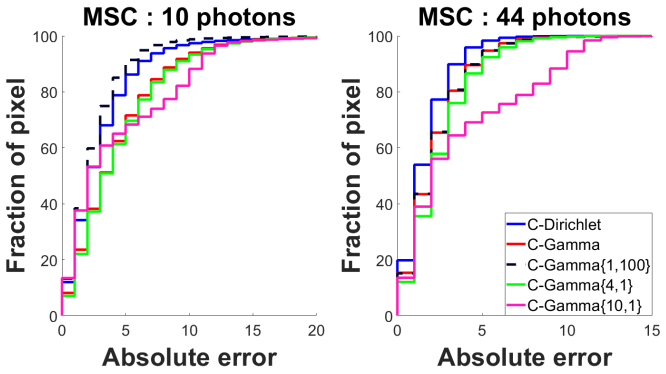

Fig. 3: CDFs of the depth absolute error, with high background (SBR $= 0.426$). The left (resp. right) subplot corresponds to MSC equals to 10 (resp. 44) photons.

The reflectivity reconstruction performance of C-Gamma and C-Gamma$\{1, 100\}$ differ only at low SBR in short imaging cases, explained by the accurate depth reconstruction of C-Gamma$\{1, 100\}$ in the same scenario that outperforms C-Dirichlet. Although the performance of C-Gamma$\{1, 100\}$ are better than those of C-Gamma, they differ only slightly and the high complexity of C-Gamma$\{1, 100\}$ makes it unattractive. The similarity between those results show that the proposed approach does not require perfect depth sampling at each iteration, as the depth profile converges quickly to a satisfactory estimate. Although the depth estimate with C-Dirichlet is more accurate than that of C-Gamma, their performance remain close and C-Gamma enables better reflectivity estimation for SBR≥ 0.426. Moreover, the loss is balanced by the significantly lower computational cost of our approach.

Finally, Table 1 reports the computational time associated

	10/∞	10/0.046	44/∞	44/0.426
C-Gamma	104s	132s	350s	493s
C-Gamma$\{4, 1\}$	43s	49s	55s	69s
C-Gamma$\{10, 1\}$	34s	36s	38s	42s
C-Dirichlet	530s	592s	629s	662s

Table 1: Computational cost of the competing approaches for synthetic SW-MSL data analysis. In the top row, the first (resp. second) value stands for the MSC (resp. SBR).

with the proposed model, compared to the method in [8]. The computational time of C-Gamma$\{1, 100\}$ is omitted here as it is computationally unattractive. All the experiments we performed in this work have been run on Matlab R2019a with a intel i7-8700k CPU @ 3.70GHz. Although the computational gain is implementation dependent and could still be improved via parallel programming, we noticed that a significant speed-up is obtained by the proposed method when compared to [8], for all illumination scenarios. In general, C-Gamma provides similar depth as the method in [8] but more accurate spectral signatures reconstruction and is about 10 times faster.

6. CONCLUSION

In this paper, a novel algorithm has been presented to estimate the spectral and depth profile of a target from SW-MSL return. The method works in presence of non negligible background photon level and for short acquisition scenarios (2.5 signal photons per pixel per wavelength, on average). The alternative estimation of the model parameters simplifies the inference process and makes it computationally tractable. Moreover, the non-local C-Gamma prior model, similarly built as in [8], allows to split the problem pixel wise, and the concatenation of the admissible depth position reduces markedly the tractability of the likelihood computation. While our new approach benefits from a significantly reduced computational cost, it is limited by the presence of high background level. Moreover, the starved photon regime has not been studied in our experiments. Thus, further studies should focus on adapting the method to extreme acquisitions scenarios.

7. REFERENCES

[1] T. Bosch, "Laser ranging: a critical review of usual techniques for distance measurement," *Opt. Eng.*, vol. 40, no. 1, 2001.

[2] P. Vines, K. Kuzmenko, J. Kirdoda, D. Dumas, M. Mirza, R. Millar, D. Paul, and G. Buller, "High performance planar Ge-on-Si single-photon avalanche diode detectors," *Nature Communications*, Sept. 2018.

[3] P. S. Chhabra, A. Maccarone, A. McCarthy, A. M. Wallace, and G. S. Buller, "Analysis of foliage penetrating photon counting Lidar data for underwater mine counter measures," in *MTS/IEEE Oceans Conference*, June 2017.

[4] X. Zhao, S. Shi, J. Yang, W. Gong, J. Sun, B. Chen, K. Guo, and B. Chen, "Active 3D imaging of vegetation based on multi-wavelength fluorescence Lidar," *Sensors*, vol. 20, no. 3, 2020.

[5] W. Gong, J. Sun, S. Shi, J. Yang, L. Du, B. Zhu, and S. Song, "Investigating the potential of using the spatial and spectral information of multispectral Lidar for object classification," *Sensors*, vol. 15, no. 9, pp. 21 989–22 002, Sept. 2015.

[6] J. Manzanera, A. Garc, C. Pascual, R. Gimeno, S. Mart, T. Tokola, and R. Valbuena, "Fusion of airborne Lidar and multispectral sensors reveals synergic capabilities in forest structure characterization," *GIScience & Remote Sensing*, vol. 53, Sept. 2016.

[7] X. Ren, Y. Altmann, R. Tobin, A. Mccarthy, S. Mclaughlin, and G. S. Buller, "Wavelength-time coding for multispectral 3D imaging using single-photon Lidar," *Opt. Exp.*, vol. 26, no. 23, pp. 30 146–30 161, Nov. 2018.

[8] Q. Legros, S. Meignen, S. McLaughlin, and Y. Altmann, "EM-based approach to 3D reconstruction from single-waveform multispectral Lidar data," *arXiv preprint arXiv:1912.06092*, 2019.

[9] G. McLachlan and T. Krishnan, *The EM algorithm and extensions*. John Wiley & Sons, 2007, vol. 382.

[10] G. Celeux, D. Chauveau, and J. Diebolt, "Stochastic versions of the EM algorithm: an experimental study in the mixture case," *Journal of Stat. Comp. and Simulation*, vol. 55, no. 4, pp. 287–314, 1996.

[11] Y. Altmann, A. Maccarone, A. McCarthy, G. Newstadt, G. S. Buller, S. McLaughlin, and A. Hero, "Robust spectral unmixing of sparse multispectral Lidar waveforms using gamma Markov random fields," *IEEE Trans. Comput. Imaging*, vol. 3, no. 4, pp. 658–670, Dec. 2017.

[12] R. Tobin, Y. Altmann, X. Ren, A. McCarthy, R. A. Lamb, S. McLaughlin, and G. S. Buller, "Comparative study of sampling strategies for sparse photon multispectral Lidar imaging: towards mosaic filter arrays," *Journal of Optics*, vol. 19, no. 9, p. 094006, Aug. 2017.

[13] Y. Altmann and S. McLaughlin, "Range estimation from single-photon Lidar data using a stochastic EM approach," in *26th European Signal Processing Conference (EUSIPCO)*, Rome, Italy, Sept. 2018, pp. 1112–1116.

[14] J. Rapp and V. K. Goyal, "A few photons among many: unmixing signal and noise for photon-efficient active imaging," *IEEE Trans. Comput. Imaging*, vol. 3, no. 3, pp. 445–459, Sept. 2017.

[15] Y. Altmann, X. Ren, A. McCarthy, G. S. Buller, and S. McLaughlin, "Lidar waveform-based analysis of depth images constructed using sparse single-photon data," *IEEE Trans. Image Processing*, vol. 25, no. 5, pp. 1935–1946, May 2016.

[16] Y. Altmann, A. Maccarone, A. McCarthy, S. McLaughlin, and G. Buller, "Spectral classification of sparse photon depth images," *Optics Express*, vol. 26, no. 5, pp. 5514–5530, Feb. 2018.

[17] G. Celeux, F. Forbes, and N. Peyrard, "EM procedures using mean field-like approximations for Markov model-based image segmentation," *Pattern Recognition*, vol. 36, no. 1, pp. 131 – 144, 2003.