Robust Restoration of Sparse Multidimensional Single-Photon LiDAR Images

Abderrahim Halimi, *Member, IEEE*, Rachael Tobin, Aongus McCarthy, *Member, IEEE*, Jose Bioucas-Dias, *Fellow, IEEE*, Stephen McLaughlin, *Fellow, IEEE*, and Gerald S. Buller

***Abstract*—The challenges of real world applications of the laser detection and ranging (Lidar) three-dimensional (3-D) imaging require specialized algorithms. In this paper, a new reconstruction algorithm for single-photon 3-D Lidar images is presented that can deal with multiple tasks. For example, when the return signal contains multiple peaks due to imaging semitransparent surfaces, or when imaging through obscurants such as scattering media. A generalization to the multidimensional case, including multispectral and multitemporal 3-D images, is also provided. The approach is based on the minimization of a cost function accounting for Poissonian observations of the single-photon data, the nonlocal spatial correlations between pixels and the small number of depth layers inside the observed range window. An alternating direction method of multipliers that offers good convergence properties is used to solve this minimization problem. The resulting algorithm is validated on synthetic and real data and in challenging realistic scenarios including sparse photon regimes for fast imaging, the presence of high background due to obscurants, and the joint processing of multispectral and/or multitemporal data.**

***Index Terms*—Three-dimensional (3-D) imaging, Lidar, multispectral imaging, multitemporal imaging, obscurants, image restoration, ADMM, non-local total variation, collaborative sparsity.**

I. Introduction

THREE-DIMENSIONAL (3-D)-imaging using Lidar systems has generated significant interest from the scientific community in recent years. This is due to its ability to provide rich and high-resolution information regarding the depth profile and reflectivity of observed targets, which can be combined with other imaging modalities such as Radar or Sonar to improve the navigation performance of autonomous vehicles, etc. Lidar systems work by emitting laser pulses and recording the arrival times of the reflected photons using for example a time-correlated single-photon counting (TCSPC) module. A histogram of photon counts with respect to (w.r.t.) time-of-flight is constructed for each beam location corresponding to a pixel (see Fig. 1). This operation is repeated for different beam locations to build a cube of data containing the 3D information about the target. As shown in Fig. 1, the reflected photons from a given target are gathered into a peak whose location and amplitude are related to the depth and reflectivity of the observed target. Extracting this information from the histograms provides depth and reflectivity images of the target, i.e., its 3D image. Note that it is also possible to expand dimensions in the Lidar system by for example (i) varying the wavelength of the laser pulses, which allows the acquisition of multispectral 3D images [1], [2], (ii) acquiring successive frames, leading to 3D videos or multitemporal 3D images.

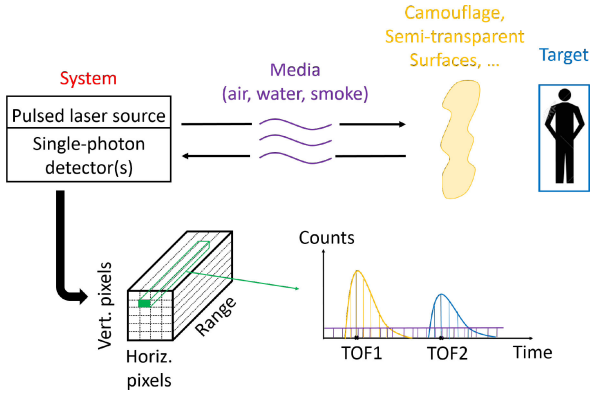


Fig. 1. Schematic description of 3-D Lidar imaging in presence of obscurants or camouflages.

In this paper, we are interested in challenging imaging scenarios encountered in real world applications including the photon starved regime when reducing the acquisition times or for long-range imaging [3], [4]; the presence of multiple peaks due to imaging through semi-transparent surfaces or when the laser beam covers many depth surfaces [5], [6]; a high background level due to imaging through obscurants [7]–[9], and the multidimensional case due to multispectral or multitemporal 3D imaging (see Fig. 1 for some illustrations). Some of these challenges have been addressed in the literature. The authors of [10] and [11] proposed Markov chain Monte-Carlo (MCMC) algorithms leading to promising results for the restoration of

Manuscript received November 25, 2018; revised March 18, 2019 and May 23, 2019; accepted July 7, 2019. Date of publication July 25, 2019; date of current version January 13, 2020. This work was supported in part by the U.K. Royal Academy of Engineering under the Research Fellowship Scheme (RF/201718/17128), in part by the EPSRC Grants EP/J015180/1, EP/N003446/1, and EP/M01326X/1, and in part by FCT/MEC through Portuguese national funds and co-funded by FEDER—PT2020 partnership agreement under project UID/EEA/50008/2019. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Orazio Gallo. *(Corresponding author: Abderrahim Halimi.)*

A. Halimi, R. Tobin, A. McCarthy, S. McLaughlin, and G. S. Buller are with the School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, U.K. (e-mail: a.halimi@hw.ac.uk; rt2@hw.ac.uk; a.mccarthy@hw.ac.uk; s.mclaughlin@hw.ac.uk; g.s.buller@hw.ac.uk).

J. Bioucas-Dias is with the Instituto de Telecomunicações and Instituto Superior Técnico, Universidade de Lisboa, Lisboa 1649-004, Portugal (e-mail: bioucas@lx.it.pt).

Digital Object Identifier 10.1109/TCI.2019.2929918

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

sparse Lidar data and multilayered data, respectively. However, these algorithms are computationally complex which might limit their use in practical situations (see for example [12] for a computational comparison between an optimization algorithm and the MCMC algorithm proposed in [10]). An alternative algorithm was recently proposed in [5] which considers a convex formulation coupled with an ℓ_1 sparsity promoting regularizer. This approach takes into account the Poisson statistics of the data and assumes the sparsity of the received photons. However, it does not account for the target continuity of surfaces and may lead to false detections when imaging through obscurants. The latter problem has also been studied in several papers as in [8], [13], which considered spatial correlation between pixels to remove false detections.

This paper proposes a solution which addresses all of these challenges by considering an optimization algorithm that reconstructs 3D scenes while taking into account prior knowledge about the observed targets. Target reconstruction is obtained by minimizing a convex function composed of a data fidelity and regularization terms. The former is based on the Poisson statistics of the observed photon counts and models the presence of multiple peaks and background noise in each histogram. With regards to the regularization terms, we first assume the presence of spatial correlation across pixels due to the spatial extent of the observed object, which is introduced using a convex non-local total variation (TV) regularizer [14]. Thanks to the fine depth resolution and the large observed range window, we also assume that the number of depth layers is lower than the number of available time bins, which is introduced using a collaborative sparse prior (group-sparsity) [15]–[17]. This leads to a convex cost function that is easily generalized to multidimensional cases as multispectral or multitemporal imaging. To provide a fast solution, the cost function is minimized using a new variant of an alternating direction method of multipliers (ADMM) algorithm [18] which has shown good results in several applications [12], [19], [20]. The new algorithm is tested on synthetic and real data showing promising results, when compared to state-of-art algorithms, to solve the previously described real world challenges.

The paper is organized as follows. The observation model and motivation of the proposed approach are described in Section II. The proposed formulation for the regularized problem and the estimation algorithm are presented in Sections III and IV. Section V analyses the proposed algorithm's performance when considering synthetic data with known ground-truth. Results on real data are presented in Section VI. Conclusions and future work are finally reported in Section VII.

II. Problem Formulation

A. Observation Model for 3D Lidar Imaging

A Lidar system operates by sending light pulses and detecting the reflected photons and their time of flight from the target, where each illumination can cause up to one photon detection. This operation can be repeated for each pixel location when using a scanning system (e.g., raster scan systems [21], [22]) or by directly acquiring an array of pixels (e.g., array based systems [23]). For both approaches, a histogram can be constructed for each pixel by representing the number of received counts with respect to their time of arrival. More precisely, the Lidar observation can be gathered in the matrix $\boldsymbol{y}_{n,t}$ which represents the number of photon counts within the tth bin of the nth pixel, where $n \in \{1, \ldots, N\}$, $t \in \{1, \ldots, K\}$, T is the timing resolution of the system and N, K are the number of pixels and time bins, respectively. If the laser beam is reflected by a single surface, the histogram will contain a single peak whose amplitude and position are related to the target's reflectivity and depth, respectively. However, when the observed scene contains obscurants, semi-transparent or scattering surfaces, or the laser beam covers many depth surfaces, the returned signal may contain multiple peaks, located at distances related to the observed depths (see Fig. 1).

This paper deals with the latter case, i.e., restoration of Lidar data in presence of multi-peaks, and a high level of noise due to imaging through obscurants or a scattering environment. For these data, the observed photon counts $\boldsymbol{y}_{n,t}$ are distributed according to a Poisson distribution $\mathcal{P}(.)$ as follows [10], [24]

$$y_{n,t} \sim \mathcal{P}\left(s_{n,t}\right) \tag{1}$$

where

$$s_{n,t} = \sum_{m=1}^{M_n} \left[r_{n,m} g_0\left(t - k_{n,m}T\right)\right] + b_n \tag{2}$$

and M_n is the number of layers in the nth pixel, $k_{n,m} \geq 0$ is the range of the mth object from the sensor (related to its depth), $r_{n,m} \geq 0$ is the reflectivity of the mth target, $b_n \geq 0$ denotes the background and dark counts of the detector, and g_0 represents the system impulse response (SIR) assumed to be known from a calibration step. The discrete time version of (2), when considering K time bins, can be expressed as a linear system as follows [5]

$$\boldsymbol{s}_n = \boldsymbol{G}\boldsymbol{x}_n \tag{3}$$

where $\boldsymbol{G} = [\boldsymbol{g}_1, \ldots, \boldsymbol{g}_K, \boldsymbol{1}_{K\times 1}]$ is a $K \times (K+1)$ matrix gathering shifted impulse responses, $\boldsymbol{1}_{i\times j}$ denotes the $(i \times j)$ matrix of 1, $\boldsymbol{g}_i = [g_0(T - iT), g_0(2T - iT), \ldots, g_0(KT - iT)]^\top$ is a $(K \times 1)$ vector representing the discrete impulse response centered at iT and $\boldsymbol{x}_n$ is a $(K+1) \times 1$ vector whose value are zero except for $\boldsymbol{x}_n(k_{n,m}) = r_{n,m}, \forall m$, and $\boldsymbol{x}_n(K+1) = b_n$. Using (3), straightforward computations show that the negative-log-likelihood associated with the discrete observations $\boldsymbol{y}_{n,k} \sim \mathcal{P}[(\boldsymbol{G}\boldsymbol{x}_n)_k]$ is given by

$$\mathcal{L}_n\left(\boldsymbol{x}_n\right) = \mathcal{H}_n\left(\boldsymbol{G}\boldsymbol{x}_n\right) \tag{4}$$

where $\mathcal{H}_n : \mathbb{R}^K \to \mathbb{R} \cup \{-\infty, +\infty\}$ is given by

$$\mathcal{H}_n\left(\boldsymbol{z}\right) = \sum_{k=1}^{K} \left\{ z_k - y_{n,k} \log\left[z_k^{(+)}\right] + i_{\mathbb{R}_+}\left(z_k\right) \right\} \tag{5}$$

where $z_k^{(+)} = \max\{0, z_k\}$ and $i_{\mathbb{R}_+}(x)$ is the indicator function that imposes non-negativity ($i_{\mathbb{R}_+}(x) = 0$ if $x \geq 0$ and $+\infty$ otherwise).

Finally, denoting $\boldsymbol{Y}$ (resp. $\boldsymbol{X}$) is a $K \times N$ (resp. $(K+1) \times N$) matrix gathering the vectors $\boldsymbol{y}_n$ (resp. $\boldsymbol{x}_n$), and assuming

that $y_{n,t}, \forall n, t$ are conditionally independent given $\boldsymbol{X}$, leads to the following negative-log of the joint likelihood

$$\mathcal{L}(\boldsymbol{X}) = -\log\left[P(\boldsymbol{Y}|\boldsymbol{X})\right] = \sum_{n} \mathcal{L}_n(\boldsymbol{x}_n). \tag{6}$$

The goal is then to estimate the sparse matrix $\boldsymbol{X}$, where the positions and values of the non-zero elements correspond to the target depths and intensities, respectively.

B. Generalization to Multidimensional Data

In addition to the rich 3D spatial information, Lidar systems can be generalized to acquire other dimensions leading to high dimensional data. For example, this might include the following: the acquisition of multispectral or polarimetric 3D images, which allow target identification using reflectivity or texture information; and the acquisition of multitemporal 3D images to capture time varying phenomena. While it is common to assume a fixed scene while acquiring a 3D images, the model associated with more dimensions (i.e., multispectral, temporal or polarized images) should take into account fluctuations between successive measures as the other dimensions might be acquired at different time instants or under a different configuration. Under this assumption, the general model will be given by

$$y_{n,t,d} \sim \mathcal{P}\left(s_{n,t,d}\right) \tag{7}$$

where

$$s_{n,t,d} = \sum_{m=1}^{M_n} \left[r_{n,m,d} g_0^d\left(t - k_{n,m,d}T\right)\right] + b_{n,d} \tag{8}$$

where it should be noted that the target parameters (reflectivities and depths) depend on the dth dimension to account for the target, system and environment variations between successive images. Similarly, we allow the system impulse response to vary with respect to d as for the case of multi-wavelength imaging. The matrix expression associated with the nth pixel and dth image is given by

$$\boldsymbol{s}_{n,d} = \boldsymbol{G}_d \boldsymbol{x}_{n,d}. \tag{9}$$

The goal is then to estimate the sparse matrices $\boldsymbol{X}_d$, while considering that the targets are observed by different dimensions $d \in \{1, \ldots, D\}$.

C. Challenges/Motivations for the Proposed Algorithm

This paper considers three major challenges whose solution will help to expand the use of Lidar systems to environmental, medical, or industrial applications. The first is related to the sparse photon regime that is encountered when reducing the acquisition times (i.e., fast imaging [10], [12], [13]) or for long range imaging (i.e., imaging at several kilometres [3]). In such cases, the constructed histograms will contain a reduced number of photons or, in the extreme case, no photons at all which corresponds to empty pixels (i.e., pixels with no photon counts). The second challenge appears when imaging through obscurants (e.g. camouflage or semi-transparent surfaces [5], [11], [25]) or a scattering media (e.g. water, fog [7], [26]). This leads to the presence of multiple peaks or a high level of background noise in the pixel's histograms, which might affect the detection of the target's peak especially when using classical single-peak based algorithms (such as cross-correlation with the system impulse response). The third challenge is related to the need to extend the 3D Lidar data to higher dimensions by acquiring several wavelengths for 3D multispectral imaging, or different time instants for 3D videos to capture time varying phenomena or moving objects. These challenges limit the use of classical single-peak based approaches and require the development of specialized algorithms to estimate $\boldsymbol{X}$ (resp. $\boldsymbol{X}_d$) while taking into account these extreme imaging scenarios.

III. Regularized Problem

This section presents the proposed regularized problem to estimate the restored point cloud matrix $\boldsymbol{X}$. To this end, we adopt an optimization approach that minimizes a regularized data fidelity cost function. More precisely, considering that the observed data is Poissonian distributed, then the data fidelity term $\mathcal{L}(\boldsymbol{X})$ is given by (6). Estimating the matrix $\boldsymbol{X}$ is an ill-posed inverse problem that requires the introduction of prior knowledge (or regularization terms) related to the target depths and reflectivities. The latter should be carefully chosen as they introduce some estimation biases regarding the observed data. With these considerations in mind, we propose to solve the following optimization problem

$$\mathcal{C}(\boldsymbol{X}) = \mathcal{L}(\boldsymbol{X}) + i_{\mathbb{R}_+}(\boldsymbol{X}) + \underbrace{\tau_1 \phi_1(\boldsymbol{X}) + \tau_2 \phi_2(\boldsymbol{X})}_{\text{regularization terms}} \tag{10}$$

where $\tau_1 > 0$, $\tau_2 > 0$ are two regularization parameters, $i_{\mathbb{R}_+}(\boldsymbol{X}) = \sum_{n,k} i_{\mathbb{R}_+}(x_{n,k})$ and ϕ_1, ϕ_2 are two regularization functions associated with the depth and reflectivity, respectively. These terms will account for two properties of Lidar data: (i) the detected photons associated with a target are generally clustered inside the cube of histograms in contrast to the background counts that spread over the full cube, (ii) the number of detected photons associated with a target presents spatial correlations, i.e., 3D points of a target often show similar count values whether it is locally (i.e., neighbour pixels) or non-locally (i.e., non neighbour pixels with similar target reflectivities). Based on these properties, we define the regularization functions ϕ_1, ϕ_2 as detailed in the next sub-sections.

A. Priors on the Support: Depth Regularization

This section presents the regularization term associated with the data support. As highlighted in [5], [27], 3D Lidar data is sparse especially in the low acquisition time regime. However, considering sparsity alone does not help separate the target's returns from those due to the background noise. In [25], an $\ell_{2,1}$ mixed norm was considered to promote the presence of sparse depth clusters in the full image. This regularization showed good results for scenes presenting objects well separated in depth, however, it might be inappropriate for targets with distributed depths which will occupy the full depth histogram when considering all the pixels. A possible strategy to deal with this is to impose this depth localization locally, i.e., for each small patch of pixels. In this paper, we choose to combine the priors in [5], [27]

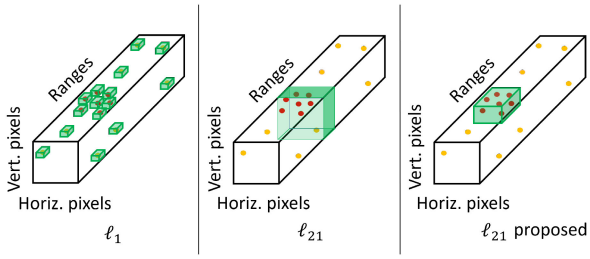


Fig. 2. Illustrative examples of the effect of different support regularizations, i.e., (left) effect of the ℓ_1-based regularization in [5], [27], (middle) the $\ell_{2,1}$-based regularization in [25], and (right) the $\ell_{2,1}$-based regularization proposed in this paper. The black cubes represent the obtained histograms, the dots represent the detected photons (red for the target and yellow for the background returns) and the green cubes represent the detected supports promoted by the regularizations.

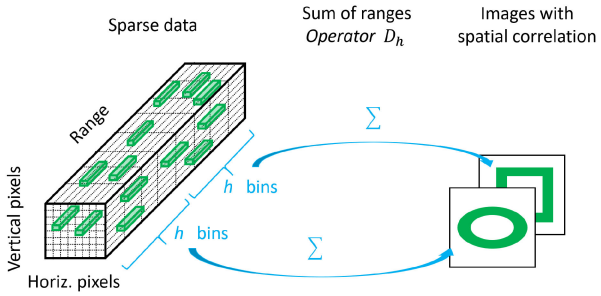


Fig. 3. Illustrative example of the range down-sampling benefits to impose spatial correlations.

and [25] to benefit from their advantages and provide a solution to their limitations mentioned above. To separate the noise from the signal return, we assume sparsity on a down-sampled image. This will detect clusters of returns in the cube that are probably due to the target, while it will eliminate the isolated counts due to noise (see Fig. 2). To achieve this, we consider an $\ell_{2,1}$ mixed norm that impose collaborative sparsity [15], [17], i.e., sparsity on small cubes obtained by grouping local pixels and depth bins. The ℓ_2 is first applied to sum the quadratic returns of each small cube, then, an ℓ_1 norm is applied to the resulting small cubes to promote their sparsity. Formally, the obtained regularization term is given by

$$\phi_1(\boldsymbol{X}) = ||\text{diag}(\boldsymbol{v})\boldsymbol{K}\boldsymbol{F}\boldsymbol{X}(:)||_{2,1} \tag{11}$$

where $\boldsymbol{X}(:) \in \mathbb{R}^{(K+1)N\times 1}$ denotes the vectorization of the matrix $\boldsymbol{X}$, $\boldsymbol{F} \in \mathbb{R}^{KN\times(K+1)N}$ is a matrix that selects the first K rows of $\boldsymbol{X}$ and discard the background row, $\text{diag}(\boldsymbol{v})\boldsymbol{K}$ can be gathered in $\boldsymbol{K}_v : \mathbb{R}^{KN\times 1} \to \mathbb{R}^{S_b\times N_B}$ which is a linear operator that provides an $S_b \times N_B$ matrix as an output where its column gathers the S_b elements of a bloc of size $S_b = (r_b \times c_b \times t_b)$ and N_B denotes the number of these blocs in the data cube and $\boldsymbol{v} \in \mathbb{R}^{N_B\times 1}$ contains weights for each bloc. This can be formulated as follows

$$\boldsymbol{K}_v\boldsymbol{F}\boldsymbol{X}(:) = \begin{bmatrix} v_1K_1\boldsymbol{F}\boldsymbol{X}(:) & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \vdots \\ \boldsymbol{0} & \cdots & v_{N_B}K_{N_B}\boldsymbol{F}\boldsymbol{X}(:) \end{bmatrix} \tag{12}$$

with $K_i \in \mathbb{R}^{S_B\times KN}, \forall i \in \{1,\ldots,N_B\}$ is a matrix that selects the ith bloc and $v_i > 0$ is a predefined weight associated with the ith bloc. Using (12), one can express $\phi_1(\boldsymbol{X})$ as follows

$$\phi_1(\boldsymbol{X}) = \sum_{i=1}^{N_B} v_i \sqrt{\left(\sum_{(t,n)\in\psi_i} x_{n,t}^2\right)} \tag{13}$$

where ψ_i contains the pixel and time bin indices of the ith bloc.

B. Priors on the Counts: Regularized Intensity

This section presents the regularization associated with the count values, i.e., target reflectivities. To restore the observed count values, the proposed term will mainly exploit the spatial correlation between count values of the images. Several denoising strategies have previously been proposed to account for spatial correlation and we distinguish between local [14], [20] and non-local [28]–[30] approaches. The latter have shown promising results to restore natural images (especially textured images) and represent a key ingredient in most state-of-the-art algorithms [28]–[30]. While local algorithms only use neighbourhood pixels to restore a given pixel, non-local approaches also exploit the information contained in distant pixels that have similar characteristics, which lead to a better denoising thanks to a wider averaging. In this paper, we consider an NL approach to restore the intensity of the Lidar data obtained in a photon-staved regime. In this case, the histograms are sparse and prevent the direct application of spatial correlation regularizations. To solve this problem, the spatial regularization is performed on a range down-sampled image that benefits from the clustered target's photon, thus, improving the separation between the target features and the background noise (see Fig. 3 for an illustrative example of the benefit of the range down-sampling). This idea of using down-sampled images (spatially or in depth) in photon-starved regime has shown good results in the restoration of sparse data as highlighted in [13], [25]. Note that the down-sampling is only performed in the depth dimension which has a very high resolution (i.e., millimetre resolution even at long ranges), while the non-local regularization term will restore the spatially corrupted data due to the noise or the absence of detected photons. To achieve this, each pixel is connected to the other similar pixels in the image (through the consideration of similarity weights $\boldsymbol{W}$) and will benefit from their information to improve its estimate, as illustrated in Fig. 4. Under these considerations, the proposed regularization term can be mathematically expressed as

$$\phi_2(\boldsymbol{X}) = ||\text{diag}(\boldsymbol{w})\boldsymbol{H}\boldsymbol{D}_h\boldsymbol{F}\boldsymbol{X}(:)||_F^2 \tag{14}$$

where $\boldsymbol{D}_h \in \mathbb{R}^{K_hN\times KN}$ is a matrix summing the photon counts of each $\#h$ successive time bins as illustrated in Fig. 3, K_h is the integer part of the division K/h, and $\boldsymbol{H} \in \mathbb{R}^{n_dK_hN\times K_hN}$ is a block-circulant-circulant-block matrix which computes

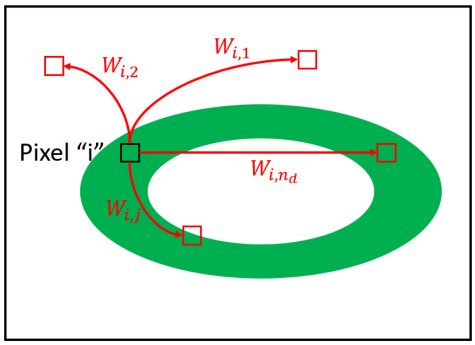


Fig. 4. Illustrative example of the connection of a given pixel i with n_d other pixels. For this image, the weights can be ordered as follows $w_{i,1} = w_{i,2} < w_{i,j} < w_{i,n_d}$ where the more similar is the connection, the higher is its weight.

weighted differences between each pixel and other n_d pixels located in a predefined field. More precisely, the operator $\text{diag}(\boldsymbol{w})\boldsymbol{H} = \boldsymbol{H}_w : \mathbb{R}^{K_h N} \to \mathbb{R}^{n_d K_h N}$ performs the following operation

$$||\boldsymbol{H}_w \boldsymbol{z}||_F^2 = \sum_{i=1}^{n_d} \sum_{n=1}^{N} \sum_{\ell=1}^{K_h} w_{i,n}^2 \left(H_i^{\text{Diff}} \boldsymbol{z}_\ell\right)^2 \Big|_n \tag{15}$$

where $\boldsymbol{z}_\ell \in \mathbb{R}^N$ denotes the ℓth column of a matrix $\boldsymbol{Z} \in \mathbb{R}^{N \times K_h}$ built from $\boldsymbol{z} \in \mathbb{R}^{K_h N}$ as follows $\boldsymbol{Z} = [\boldsymbol{z}_1, ..., \boldsymbol{z}_{K_h}]$, $H_i^{\text{Diff}} \in \mathbb{R}^{N \times N}$ computes the difference between each pixel and that located at the ith direction (the algorithms consider n_d predefined directions or shifts), and $w_{i,n}^2$ are the weights associated with the nth pixel and ith direction. For simplicity, we treat the matrices H_i^{Diff} with periodic boundary conditions as cyclic convolutions. This non-local total variation can also be expressed in matrix form as follows

$$||\boldsymbol{H}_w \boldsymbol{z}||_F^2 = \left|\left| \begin{array}{c} \left\{\mathbb{I}_{K_h} \otimes \left[\text{diag}(\boldsymbol{w}_{1,:}) H_1^{\text{Diff}}\right]\right\} \boldsymbol{z} \\ \vdots \\ \left\{\mathbb{I}_{K_h} \otimes \left[\text{diag}(\boldsymbol{w}_{n_d,:}) H_{n_d}^{\text{Diff}}\right]\right\} \boldsymbol{z} \end{array} \right|\right|_F^2 \tag{16}$$

where $\otimes$ denotes the Kronecker product and $\mathbb{I}_{K_h}$ is the identity matrix, $\boldsymbol{w}_{i:} \in \mathbb{R}^N$ are the weights associated with the ith direction, and diag is the diagonalization operator.

C. Choice of the Weights

The regularization terms introduce two weighting vectors $\boldsymbol{w}, \boldsymbol{v}$ that should reflect our prior knowledge regarding possible correlations between spatial regions of the image and/or the expected depth of the target. Such information can be extracted from complementary imaging modalities of the same scene such as a 2D panchromatic image to determine pixel's correlations, or radar image to approximate the depth. In the presence of such complementary information, it is possible for the proposed algorithm to combine the information obtained by these modalities, i.e., it will perform a fusion task. Note also that setting equal weights for a local neighbourhood (e.g., 4 neighbourhood structure) will lead to the common local total-variation approach. In this paper, we do not use other imaging modalities, and only use the Lidar data to fix these weights. Akin to [31], [32], the weights $\boldsymbol{w}$ are chosen to reduce the spatial smoothing across discontinuities. The weights $\boldsymbol{v}$ are chosen to enforce removing isolated photons and promote the presence of a target in regions presenting clustered photons. However, setting accurate weights require the use of a relatively clean image to improve the restoration results of the overall algorithm. Such an image is not available as we observe a noisy, sparse and temporally convolved Lidar image, however, it can be approximated as follows. A spatial low pass filter is applied to the histogram data (by using a square window of size $\sqrt{n_d} \times \sqrt{n_d}$ pixels) to fill empty pixels/histograms and reduce the noise as already used in [13], [33], [34]. A simple method will then be used to extract k_p depth and reflectivity estimates from each pixel. More precisely, a classical cross-correlation method (see Section V-B for a detailed description of this method) is used on each pixel to extract the depth and reflectivity parameters of a first peak. The counts associated with this peak are then removed from the histogram and the classical algorithm is used again to estimate a new peak. This procedure is repeated k_p times or until no-peak is detected (in which case reflectivity is put to zero). Intensity images are then obtained by multiplying the reflectivity images by the sum of the impulse response's counts. The weights $\boldsymbol{w}$ being related to intensity, they can be fixed as follows (see [31])

$$w_{ij} = \max\left[0.5, \exp\left(-\frac{|I_i - I_j|}{\sigma_w}\right)\right] \tag{17}$$

where I denotes the sum of the k_p intensity images, $\sigma_w = 0.1$ is a fixed coefficient and max(.) denotes the maximum operator that truncates the coefficients to 0.5 (this empirically improves the results as shown in [31]). Similarly, the weights $\boldsymbol{v}$ are fixed by considering the k_p depth and intensity images. For this, we first reconstruct a data cube $\tilde{\boldsymbol{y}}$ by associating to each depth location obtained from downsampled histograms the corresponding intensity value. Second, the resulting reconstructed blocs with higher intensities are assigned lower coefficient values $\boldsymbol{v}$, as follows

$$v_i = \max\left[0.5, \exp\left(-\frac{\sum_{(t,n)\in\psi_i} \tilde{y}_{n,t}}{\sigma_w}\right)\right]. \tag{18}$$

Under these considerations, the weights $\boldsymbol{v}$ enforce more sparsity regularization on isolated photons due to background, than on clustered ones related to signal. Note finally that k_p has a limited effect on the final result of the algorithm as it is only used to set the weights (it is not related to the final estimated number of peaks) and most counts are generally contained in the first two peaks. We choose in the following to fix $k_p = 2$ since this provides a reduced computational complexity.

D. Generalization to Multidimensional Data

The proposed approach is general and can be easily extended to account for the presence of multidimensional data (e.g. multitemporal, multispectral, polarimetric data either separately or jointly). Note first that each dimension of this data can be processed independently. However, this is not optimal as these dimensions might have complementary information, thus, the

need for a joint processing algorithm, that will help downstream applications (such as object tracking) operating in one dimension by bringing information from other dimensions. Joint processing requires, however, the adoption of some additional assumptions for optimal exploitation of these data. We therefore assume that the multidimensional data corresponds to the same scene with a slight movement of the observed object or the camera as considered in many studies [35]–[38]. Under this assumption, we might assume that the support of the downsampled images is almost the same leading to

$$\Phi_1\left(\tilde{\boldsymbol{X}}\right) = \sum_{i=1}^{N_B} v_i \sqrt{\left(\sum_{d=1}^{D} \sum_{(t,n)\in\psi_i} x_{n,d,t}^2\right)} \tag{19}$$

$$= \left|\left| \begin{array}{c} \text{diag}(\boldsymbol{v})\boldsymbol{K}\boldsymbol{F}\boldsymbol{X}_1(:) \\ \vdots \\ \text{diag}(\boldsymbol{v})\boldsymbol{K}\boldsymbol{F}\boldsymbol{X}_D(:) \end{array} \right|\right|_{2,1}. \tag{20}$$

where $\tilde{(.)}$ denotes the matrix that gathers the data or operations for all D dimensions. Similarly, the intensity regularization can be assumed as

$$\Phi_2\left(\tilde{\boldsymbol{X}}\right) = \sum_{d=1}^{D} ||\boldsymbol{H}_w \boldsymbol{D}_h \boldsymbol{F}\boldsymbol{X}_d(:)||_F^2. \tag{21}$$

Note first that spatial correlation is promoted between the pixels belonging to the same dimension, as different dimensions might have different intensity responses (e.g. multispectral imaging). Note also that this term introduces correlation between dimensions as we assume the same weights $\boldsymbol{w}$ (that can obtained using all dimensions) for all of them. However, the independent case can be easily obtained by associating a different vector $\boldsymbol{w}_d$ to each dimension. Note finally that the weights can be fixed using the procedure described in the previous section. This can be done by considering all dimensions or by only considering the more informative dimensions, if available, as will be seen in the results.

For simplicity of notation, we rewrite the obtained regularization terms Φ_1, Φ_2 using new operators as follows

$$\Phi_1(\boldsymbol{X}) = ||\text{diag}(\boldsymbol{v})\tilde{\boldsymbol{K}}\tilde{\boldsymbol{F}}\tilde{\boldsymbol{X}}(:)||_{2,1} \tag{22}$$

$$\Phi_2(\boldsymbol{X}) = ||\tilde{\boldsymbol{H}}_w \tilde{\boldsymbol{D}}_h \tilde{\boldsymbol{F}}\tilde{\boldsymbol{X}}(:)||_F^2. \tag{23}$$

IV. The Estimation Algorithm

A. The ADMM Algorithm

Consider the optimization problem

$$\underset{\boldsymbol{X},\boldsymbol{C}}{\text{argmin}}\, g(\boldsymbol{C}), \text{ subject to } \boldsymbol{A}\boldsymbol{X} + \boldsymbol{B}\boldsymbol{C} = 0 \tag{24}$$

where $\boldsymbol{X} \in \mathbb{R}^{(K+1)\times N}$, $g(.)$ is a closed, proper, convex function, and $\boldsymbol{A}$, $\boldsymbol{B}$ are arbitrary matrices. The ADMM algorithm consists first in computing the augmented Lagrangian for problem (24), as follows

$$\mathcal{AL}(\boldsymbol{X},\boldsymbol{C},\boldsymbol{J}) = g(\boldsymbol{C}) + \frac{\mu}{2}||\boldsymbol{A}\boldsymbol{X} + \boldsymbol{B}\boldsymbol{C} - \boldsymbol{J}||_F^2 \tag{25}$$

Algorithm 1: ADMM for (24).

1: Initialization
2: Initialize $\boldsymbol{X}^{(0)}, \boldsymbol{C}_j^{(0)}, \boldsymbol{J}_j^{(0)}, \forall j, \mu > 0$.
3: Set $i \leftarrow 0$, conv$\leftarrow 0$
4: **while** conv$= 0$ **do**
5: $\quad \boldsymbol{X}^{(i+1)} \leftarrow \underset{\boldsymbol{X}}{\text{argmin}}\, \mathcal{AL}\left(\boldsymbol{X}, \boldsymbol{C}^{(i)}, \boldsymbol{J}^{(i)}\right)$
6: $\quad \boldsymbol{C}^{(i+1)} \leftarrow \underset{\boldsymbol{C}}{\text{argmin}}\, \mathcal{AL}\left(\boldsymbol{X}^{(i+1)}, \boldsymbol{C}, \boldsymbol{J}^{(i)}\right)$
7: $\quad \boldsymbol{J}^{(i+1)} \leftarrow \boldsymbol{J}^{(i)} - \boldsymbol{A}\boldsymbol{X}^{(i+1)} - \boldsymbol{B}\boldsymbol{C}^{(i+1)}$
8: $\quad$ conv$\leftarrow 1$, if the stopping criterion is satisfied.
9: **end while**

where μ is a positive constant, and $\boldsymbol{J}/\mu$ denotes the Lagrange multipliers associated with the constraint $\boldsymbol{A}\boldsymbol{X} + \boldsymbol{B}\boldsymbol{C} = 0$. As a second step, the algorithm optimizes $\mathcal{AL}$ sequentially with respect to $\boldsymbol{X}$ and $\boldsymbol{C}$, and then updates the Lagrange multipliers as shown in Algorithm 1.

Algorithm 1 converges when the function g is closed, proper, and convex and $\boldsymbol{A}$ is full column rank [39, Theorem 1]. The latter theorem also states that the sequence $\boldsymbol{X}^{(i)}$ converges to a solution of (24), for any $\mu > 0$, if it has a non-empty set of solutions. If (24) does not have a solution, then at least one of the sequences $\boldsymbol{X}^{(i)}$ or $\boldsymbol{J}^{(i)}$ diverges. Note that the details of the steps of Algorithm 1 are not provided for brevity, however, they reduce to the solution of a linear system of equations (line 5), the computation of Moreau proximity operators [40] (line 6), and the updating of the Lagrange multipliers (line 7). The convergence speed of the algorithm is affected by the parameter μ, that has been updated using the adaptive procedure described in [17], [18]. This procedure keeps the ratio between the ADMM primal and dual residual norms within a given positive interval, as they both converge to zero. The algorithm is stopped when these residual norms are lower than a given threshold [18]. The interested reader is invited to read [18] for more details regarding the ADMM algorithm and [19], [41], [42] for its application to solve inverse imaging problems.

B. Proposed Algorithm

This section presents the optimization problem considered for estimating the matrix of interest $\tilde{\boldsymbol{X}}$. Using the same notation as in (24), problem (10) can be expressed as follows

$$g(\boldsymbol{C}) = \mathcal{H}(\boldsymbol{C}_1) + i_{\mathbb{R}_+}(\boldsymbol{C}_2) + \tau_1||\text{diag}(\boldsymbol{v})\boldsymbol{C}_3||_{2,1} + \tau_2||\text{diag}(\boldsymbol{w})\boldsymbol{C}_5||_F^2 \tag{26}$$

with $\boldsymbol{C}_1 = \tilde{\boldsymbol{G}}\tilde{\boldsymbol{X}}$, $\boldsymbol{C}_2 = \tilde{\boldsymbol{X}}$, $\boldsymbol{C}_3 = \tilde{\boldsymbol{K}}\tilde{\boldsymbol{F}}\tilde{\boldsymbol{X}}(:)$, $\boldsymbol{C}_4 = \tilde{\boldsymbol{D}}_h\tilde{\boldsymbol{F}}\tilde{\boldsymbol{X}}(:)$, $\boldsymbol{C}_5 = \tilde{\boldsymbol{H}}\boldsymbol{C}_4$ leading to $\boldsymbol{A} = [\tilde{\boldsymbol{G}}, \mathbb{I}, \tilde{\boldsymbol{K}}\tilde{\boldsymbol{F}}, \tilde{\boldsymbol{D}}_h\tilde{\boldsymbol{F}}, \boldsymbol{0}]^\top$ and

$$\boldsymbol{B} = \begin{bmatrix} -\mathbb{I} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & -\mathbb{I} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & -\mathbb{I} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & -\mathbb{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \tilde{\boldsymbol{H}} & -\mathbb{I} \end{bmatrix} \tag{27}$$

where $\mathbb{I}$ (resp. $\mathbf{0}$) denotes the identity matrix (resp. a vector of zeros) of adequate size. Note that the use of two variables $\boldsymbol{C}_4$ and $\boldsymbol{C}_5$ allows decoupling of the optimization in the spatial domain from the optimization in the time (or range) domain (since $\tilde{\boldsymbol{D}}_h\tilde{\boldsymbol{F}}$ operates on the range dimension and $\tilde{\boldsymbol{H}}$ on the spatial one). This operation reduces the computational cost associated with the resulting optimization problems (see (29)) and has been previously used for processing hyperspectral imagery in [20]. For this problem, the matrix $\boldsymbol{A}$ is full column rank. This matrix and the properties of $g(.)$ ensure the algorithm convergence. Moreover, to accelerate convergence, the algorithm is initialized using a relatively clean image $\boldsymbol{X}^{(0)} = \tilde{\boldsymbol{Y}}$, where $\tilde{\boldsymbol{Y}}$ has been defined in Section III-C and used to set the weights. Finally, the optimization problems shown in line 5 and 6 of algo. 1, and resulting from the minimization of the $\mathcal{AL}$ given by

$$\begin{aligned}
&\mathcal{AL}\left(\tilde{\boldsymbol{X}}, \boldsymbol{C}, \boldsymbol{J}\right) \\
&\quad = \mathcal{H}\left(\boldsymbol{C}_1\right) + i_{\mathbb{R}_+}\left(\boldsymbol{C}_2\right) + \tau_1 ||\text{diag}(\boldsymbol{v})\boldsymbol{C}_3||_{2,1} \\
&\qquad + \tau_2 ||\text{diag}(\boldsymbol{w})\boldsymbol{C}_5||_F^2 + \frac{\mu}{2}\left(||\tilde{\boldsymbol{G}}\tilde{\boldsymbol{X}} - \boldsymbol{C}_1 - \boldsymbol{J}_1||_F^2\right. \\
&\qquad + ||\tilde{\boldsymbol{X}} - \boldsymbol{C}_2 - \boldsymbol{J}_2||_F^2 + ||\tilde{\boldsymbol{K}}\tilde{\boldsymbol{F}}\tilde{\boldsymbol{X}} - \boldsymbol{C}_3 - \boldsymbol{J}_3||_F^2 \\
&\qquad + ||\tilde{\boldsymbol{D}}_h\tilde{\boldsymbol{F}}\tilde{\boldsymbol{X}} - \boldsymbol{C}_4 - \boldsymbol{J}_4||_F^2 \\
&\qquad \left. + ||\tilde{\boldsymbol{H}}\boldsymbol{C}_4 - \boldsymbol{C}_5 - \boldsymbol{J}_5||_F^2\right). \qquad (28)
\end{aligned}$$

admit analytical solutions as follows (see also [12], [19], [20] for more details regarding similar optimization problems)

$$\begin{aligned}
\tilde{\boldsymbol{X}}^{(i+1)} &\leftarrow \left\{\tilde{\boldsymbol{G}}^\top\tilde{\boldsymbol{G}} + \mathbb{I} + \tilde{\boldsymbol{F}}^\top\tilde{\boldsymbol{K}}^\top\tilde{\boldsymbol{K}}\tilde{\boldsymbol{F}} + \tilde{\boldsymbol{F}}^\top\tilde{\boldsymbol{D}}_h^\top\tilde{\boldsymbol{D}}_h\tilde{\boldsymbol{F}}\right\}^{-1} \\
&\quad \times \left\{\tilde{\boldsymbol{G}}^\top\xi_1^{(i)} + \xi_2^{(i)} + \tilde{\boldsymbol{F}}^\top\tilde{\boldsymbol{K}}^\top\xi_3^{(i)} + \tilde{\boldsymbol{F}}^\top\tilde{\boldsymbol{D}}_h^\top\xi_4^{(i)}\right\} \\
c_{1,k,n,d}^{(i+1)} &\leftarrow \frac{1}{2}\left[z_{k,n,d} - \frac{1}{\mu} + \sqrt{\left[z_{k,n,d} - \frac{1}{\mu}\right]^2 + 4\frac{y_{k,n,d}}{\mu}}\right] \\
\boldsymbol{C}_2^{(i+1)} &\leftarrow \max\left\{\tilde{\boldsymbol{X}}^{(i)} - \tilde{\boldsymbol{J}}_2^{(i)}, 0\right\} \\
\boldsymbol{c}_{3,n,t}^{(i+1)} &\leftarrow \text{vect-soft}\left(\boldsymbol{J}_{3,n,t}^{(i)} - \tilde{\boldsymbol{K}}\tilde{\boldsymbol{F}}\tilde{\boldsymbol{x}}_{n,t}^{(i)}, \frac{\tau_1 v_j}{\mu}\right), \forall n, t \in \psi_j \\
\boldsymbol{C}_4^{(i+1)} &\leftarrow \left(\tilde{\boldsymbol{H}}^\top\tilde{\boldsymbol{H}} + \mathbb{I}\right)^{-1}\left[\tilde{\boldsymbol{D}}_h\tilde{\boldsymbol{F}}\boldsymbol{X}^{(i)} - \boldsymbol{J}_4^{(i)} + \tilde{\boldsymbol{H}}^\top\xi_5\right] \\
\boldsymbol{C}_5^{(i+1)} &\leftarrow \text{diag}\left[\frac{\mu}{2\tau_2(\boldsymbol{w}\odot\boldsymbol{w}) + \mu}\right]\left[\tilde{\boldsymbol{H}}\boldsymbol{C}_4^{(i)} - \boldsymbol{J}_5^{(i)}\right] \qquad (29)
\end{aligned}$$

where $\xi_j = \boldsymbol{C}_j^{(i)} + \boldsymbol{J}_j^{(i)}$ for $j \in \{1,2,3,5\}$, $z_{k,n,d} = (\boldsymbol{G}_d\boldsymbol{x}_{n,d}^{(i)})_k - \boldsymbol{J}_{1,k,n,d}^{(i)}$, $\odot$ denotes the term-wise product, $\text{soft}(\boldsymbol{X}, \frac{\tau}{\mu}) = \text{sign}(\boldsymbol{X}) \odot \max\{|\boldsymbol{X}| - \frac{\tau}{\mu}, 0\}$ denotes the soft threshold operator, $\text{vect-soft}(\boldsymbol{x}, \frac{\tau}{\mu}) = \boldsymbol{x}(\frac{\max\{||\boldsymbol{x}||_2 - \frac{\tau}{\mu}, 0\}}{\max\{||\boldsymbol{x}||_2 - \frac{\tau}{\mu}, 0\} + \frac{\tau}{\mu}})$ is the vect-soft-threshold operator, and $|.|$, sign(.), max(.) are the element-wise operators corresponding to the absolute value, the sign function and the maximum operator, respectively. Note that the required matrix inversion to update $\tilde{\boldsymbol{X}}$ is done once outside the iterative loop. The two other matrix inversions involve a diagonal and Fourier diagonalizable matrices whose inverses can be efficiently computed.

V. Results on Synthetic Data

A. Evaluation Metrics

The performance of the proposed algorithm is evaluated qualitatively by visual inspection of the estimates and quantitatively using different metrics. For histograms showing a single peak, the depth results are quantitatively evaluated using the root mean square error defined by $\text{RMSE} = \sqrt{\frac{1}{N}||\boldsymbol{d}^{\text{ref}} - \widehat{\boldsymbol{d}}||^2}$ (as used in [13]), and the reflectivity using the signal to reconstruction error $\text{SRE} = 10\log_{10}(\frac{||\boldsymbol{r}^{\text{ref}}||^2}{||\boldsymbol{r}^{\text{ref}} - \widehat{\boldsymbol{r}}||^2})$ criterion (as used in [12] and similarly to NMSE measure adopted in [33]), where $\boldsymbol{x}^{\text{ref}}$ (resp. $\widehat{\boldsymbol{x}}$) is the reference (resp. estimated) depth or reflectivity image, and the higher SRE (in dB) the better. For those algorithms not using spatial correlation, the reflectivity and depth of empty pixels are replaced by the average of available pixels before evaluating RMSE and SRE. Note that the reference depth and reflectivity images are available for synthetic data. For real data, we consider images obtained with the highest acquisition times and under clear environment conditions as references. In presence of multiple peaks, we consider the point detection criteria introduced in [33]. This includes the percentage of true detections as a function of the distance τ, where a true detection occurs if an estimated point of a given nth pixel has a reference point in its surrounding such that $|\hat{d}_n - d_n^{\text{ref}}| \leq \tau$. The estimated points that can not be assigned to any true point at a distance of τ are considered as false detections.

B. Comparison Algorithms

The performance of the proposed algorithm is compared to several state-of-the-art algorithms depending on the considered scenario. For scenes having only one peak per pixel, the algorithm is compared with:

- the classical algorithm (denoted Class.) which estimates depth using the maximum of matched filtered histograms by the impulse response, i.e., $\hat{k}_n = \text{argmax}_{k_n} \sum_{t=1}^{T} g_0(k_nT - t)y_{n,t}$, $\forall n$. Reflectivity is obtained by summing the recorded photons around the peak normalized by the number of counts in the impulse response as follows $\hat{r}_n = \frac{\sum_{t=t_l}^{t_h} y_{nt}}{\sum_{t=1}^{K} g_0(t - \hat{k}_nT)}$, $\forall n$, where $t_l = \max(1, \hat{k}_n - t_{le})$, $t_h = \min(K, \hat{k}_n + t_{te})$, and t_{le}, t_{te} represent the length of the leading and trailing edge of the impulse response, respectively. This algorithm is denoted by Class. when applied to noisy data and by BF-classical or BFC when applied to background-free data.
- RDI-TV algorithm [12] which computes the maximum-a-posteriori estimates of the depth and reflectivity images while considering a total-variation regularization term. This algorithm assumes the presence of one peak and known corrupted pixel positions. The algorithm in [43]

assumes similar regularizations and we expect its performance to be of the same order of magnitude as for RDI-TV.

- TV-ℓ_{21} algorithm [25] which generalizes RDI-TV by accounting for the presence of multiple peaks and is designed to reconstruct scenes presenting objects well separated in the depth dimension, as for the studied mannequin face real target.
- Unmixing algorithm (denoted UA) [13] which is the state-of-the-art algorithm in presence of one surface per pixel and high background levels. The algorithm also assumes that the target lies inside the observation window and some known parameters from a calibration step.

In presence of multiple surfaces per pixel, we compare with

- TN-Class.: which is the classical algorithm with known true number of peaks per pixel. In this case, the classical algorithm is iteratively applied to each pixel to estimate the parameters of all peaks. At each iteration, a peak is estimated using Class. and then removed from histogram by enforcing $y_{nt} = 0$ for $t \in [t_l, t_h]$. The Class. algorithm is then applied on the resulting histogram to estimate the next peak.
- SPISTA [5]: This algorithm minimizes a cost function that accounts for the data Poisson statistics and enforces an ℓ_1-regularization term to recover the peaks. In this paper, the minimization is performed using an ADMM algorithm as for the proposed algorithm. Comparing with this algorithm highlights the benefit of considering spatial correlation between peaks which is not considered by SPISTA.
- MANIPOP [33]: This algorithm considers a Bayesian strategy coupled with a reversible jump MCMC algorithm to obtain clean point cloud estimates. It assumes the presence of multiple peaks and considers local spatial correlations for restoration. Comparisons with this algorithm highlight the robustness of our algorithm to noise as a result of the non-local spatial correlations.

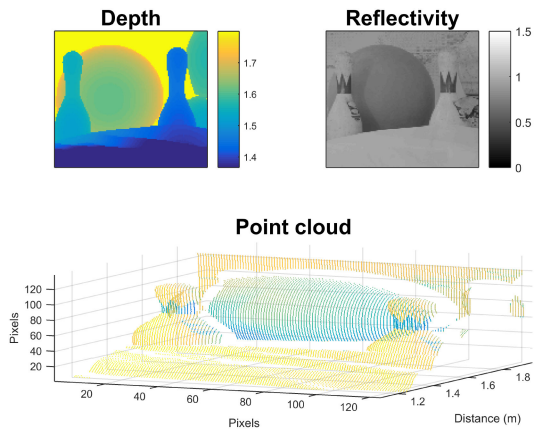


Fig. 5. Synthetic image (123×139 pixels) of a bowling scene. (Top-left) Depth map, (top-right) reflectivity and (bottom) point cloud combining depth and reflectivity.

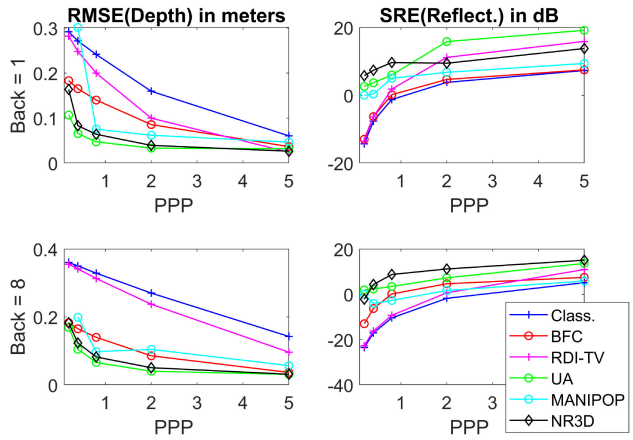


Fig. 6. Quantitative evaluation of different algorithms for reflectivity and depth estimation on simulated data sets with different signal photon per pixel on average and two background levels (i.e., different SBR values).

C. Robustness to Background and Sparse Photon Regime

This section evaluates the performance of the proposed algorithm using synthetic bowling data generated from a real physical scene and provided in the Middlebury dataset[1] [13], [44]. The depth and reflectivity images in Fig. 5 have been used to generate a $123 \times 139 \times 300$ data cube following model (1) while considering a real impulse response $g(.)$ measured from our imaging system which has a leading-edge of 10 bins and trailing-edge of 70 bins (the same impulse response will be considered for all simulated synthetic data). Different signal levels have been considered to obtain an average of signal photons-per-pixel (PPP) ranging from 0.2 ppp to 25 ppp. To simulate variations in the signal-to-background ratio (SBR) the background level was varied from $b_n = 1$, $\forall n$ to $b_n = 8$, $\forall n$ leading to SBR values in the interval $[0.025, 25]$ (we emphasise here that the PPP levels are related to signal or target counts and that the count levels of the observed histograms $\boldsymbol{Y}$ is denoted by YPPP and can be obtained by summing PPP with the background level b_n of the data, i.e., YPPP = PPP+b_n). The proposed algorithm, denoted multidimensional-Nonlocal Restoration of 3D (M-NR3D) images or just NR3D when applied to a single data cube, is run using $(r_b, c_b, t_b, h, n_d) = (4, 4, 50, 5, 9)$ (which are empirically chosen depending on the data size) and is compared to several algorithms. The NR3D regularization parameters are manually selected to provide best performance when testing nine log-spaced values inside the following intervals $\tau_1 \in [0.001, 1000]$ and $\tau_2 = [0.0005, 10]$. Fig. 6 shows the obtained depth RMSE and reflectivity SRE for six algorithms on data generated with different PPP and SBR levels, where we considered the main peak for NR3D and MANIPOP. Regarding depth, NR3D and UA provides best performance with a slight advantage to UA as it estimates less parameters (UA only assumes one peak). RDI-TV and MANIPOP algorithms are sensitive to background noise as they only consider local correlations and this leads to intermediate depth restoration performance. As expected, BFC provides better performance than Class. where

[1][Online]. Available: http://vision.middlebury.edu/stereo/data/

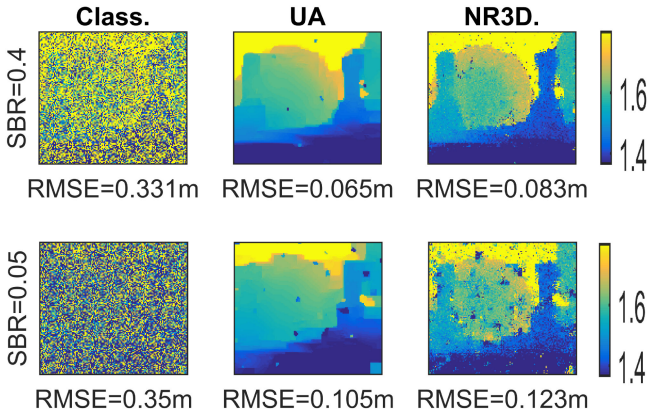


Fig. 7. Simulated depth results (123 × 139 pixels) for Bowling scene [16] at 0.4 signal photon-per-pixel (ppp) and background fixed to (top row) 1 and (bottom row) 8.

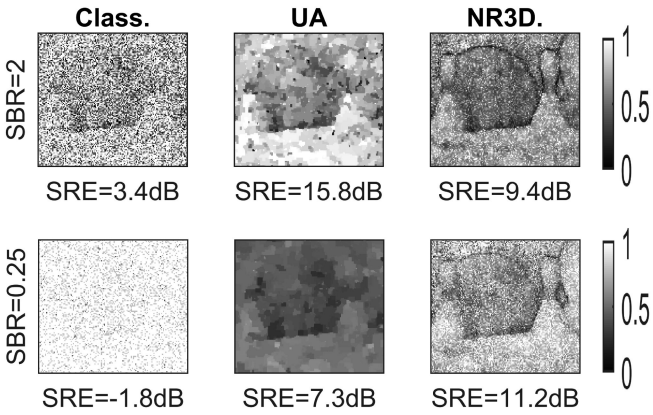


Fig. 8. Simulated intensity results (123 × 139 pixels) for Bowling scene [16] at 2 signal photon-per-pixel (ppp) and background fixed to (top row) 1 and (bottom row) 8.

these two algorithms do not consider spatial correlation which explain their limited performance. Regarding reflectivity, NR3D showed more robustness to photon sparsity and to the presence of background as highlighted in the right column of Fig. 6. Note that negative reflectivity SREs are due to the presence of corrupted pixels for reduced PPP. Fig. 7 shows example of estimated depth maps by Class., UA and NR3D algorithms for two background levels. These figures show that the classical algorithm performs poorly and is sensitive to low PPP levels and presence of background noise. Although UA shows better RMSE values, it can be seen that the obtained images are blurry and that NR3D shows a better separation between the scene components. Fig. 8 shows estimated reflectivity maps for different SBR levels. Since Class. algorithm does not consider spatial correlation, it provides noisy estimates. UA algorithm estimates better maps but NR3D provides best maps where the scene objects are recognizable under these extreme conditions. Note finally that the average computational times to process a data cube was 44s for RDI-TV, 18 s for UA, 132 s for MANIPOP and 119 s for NR3D, however, better results could be obtained for the proposed strategy by adopting parallel computations as

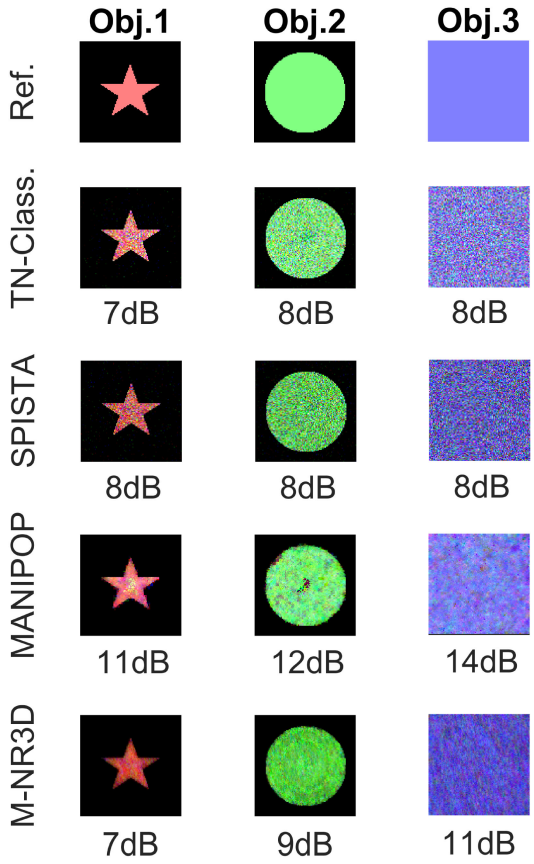


Fig. 9. Reflectivities for synthetic data. From top to bottom: Reference images used to create synthetic data, TN-Class., SPISTA, MANIPOP, and M-NR3D. The reflectivity SRE in dB are provided bellow images.

for UA. These results confirm the robustness of the proposed strategy in presence of empty pixels (due to low acquisition times or long-range measurements) or a high background level (due to obscurants).

D. Restoration of Multidepth-Multispectral Targets

This section highlights M-NR3D's ability to deal with multidimensional data obtained using several wavelengths in presence of multiple peaks (note that a multitemporal case for a static scene in presence of obscurants will be presented when considering real data). The synthetic data (of size $(x, y, t) = 100 \times 100 \times 300$) contains three geometrical objects of different colours (see Fig. 9 (top-row)) which are located at different depths. To assess the proposed strategy and weights setting, the objects have been selected to have different geometrical properties including different angles (as for the ball) and sharp-narrow edges (as for the star). This is a simulated scene where the objects are considered to be semi-transparent as for stained glasses, allowing the laser pulses to see behind them, leading to the presence of multiple peaks in some pixels as indicated in Fig. 10 (left). We assume that this scene is observed using three wavelengths corresponding to red, green and blue leading to three data cubes as shown in Fig. 11 (top-row). Note that, for each wavelength, higher reflectivity values (i.e., photon counts) are

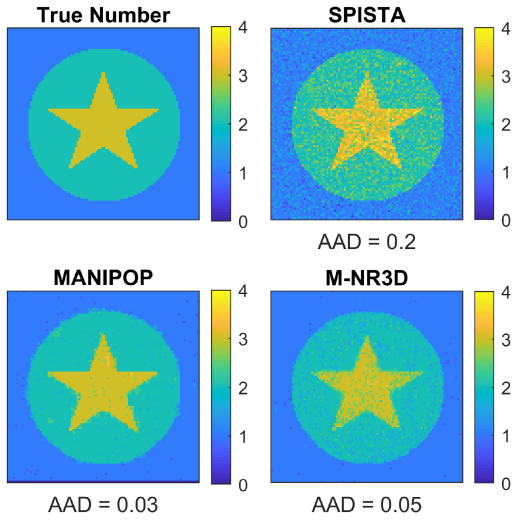


Fig. 10. Number of peaks per pixel averaged through the RGB dimensions. (top-left) true number, (top-right) estimated using SPISTA, (bottom-left) estimated using MANIPOP, and (bottom-right) estimated using M-NR3D. The average of absolute differences (AAD) between the true and estimated numbers are provided bellow images.

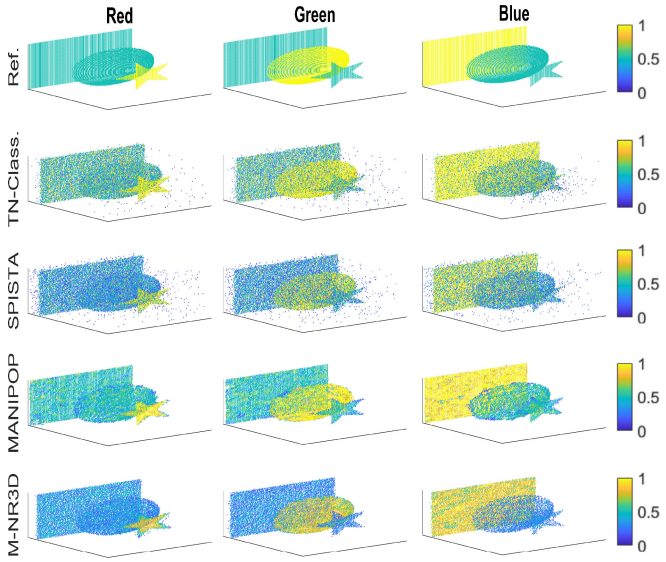


Fig. 11. 3D-point cloud corresponding to three colours (i.e., wavelengths) for synthetic data. (First row) reference depth and reflectivity cubes, (Second row) classical algorithm with true number of peaks, (Third row) SPISTA, (Fourth row) joint estimation using M-NR3D.

obtained from the geometrical object that has the same colour. As previously, the histogram data was generated using model (1), a real impulse response $g(.)$, a background level $b_n = 1, \forall n$ and a signal PPP equal to 10 corresponding to an SBR$= 10$. The three data cubes have been independently processed using the TN-classical, the SPISTA and MANIPOP algorithms. The proposed M-NR3D algorithm has also been used to jointly process the three cubes to improve performance with the following parameters $(r_b, c_b, t_b, h, n_d) = (2, 2, 10, 20, 9)$. Fig. 10 shows the estimated number of peaks per pixel for SPISTA, MANIPOP

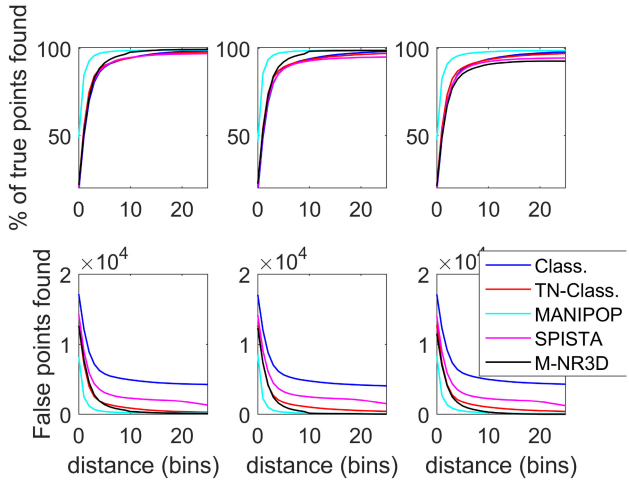


Fig. 12. Top: Percentage of true detections for different algorithms as a function of maximum distance (in bins). Bottom: Number of false detections. (Left) red dimension, (middle) green dimension, (right) blue dimension.

and M-NR3D after averaging through the RGB dimensions. This figure shows best estimates for MANIPOP and M-NR3D which is confirmed by computing the average of absolute differences between the true and estimated numbers leading to 0.2 for SPISTA and 0.03 for MANIPOP and 0.05 for M-NR3D. Fig. 9 presents the obtained reflectivity results for the three objects. The classical and SPISTA algorithms provide noisy images while both MANIPOP and M-NR3D results are cleaner thanks to imposing spatial correlations. The four algorithms also provide depth estimates for each wavelength as represented in Fig. 11. As expected, the TN-classical and SPISTA algorithms show noisy results since they do not account for spatial correlations, nor correlations between wavelengths. In contrast, M-NR3D and MANIPOP present cleaner point clouds with best performance for the latter. These results are quantitatively confirmed in Fig. 12 which provides the probabilities of true and false detections for the algorithms (the results of Class. algorithm are obtained by applying it to three depth gated parts including at most one peak each). This figure highlights the superiority of MANIPOP (that is optimized to deal with multiple-peaks) followed by the proposed algorithm. Finally, these results confirm the ability of M-NR3D to use spatial and spectral correlations to restore multispectral data in presence of multilayered objects.

VI. Results on Real Data: Multitemporal Imaging Through Obscurants

The proposed algorithm is validated on real data under challenging scenarios including photon starved regime and high background levels due to the presence of obscurants. The scene considered, consists of a life-sized polystyrene head as shown in Fig. 13 (left), and was set in a fog chamber (of dimensions $26 \times 2.5 \times 2.3$ meters) located in an indoor facility at the French-German Research Institute of Saint-Louis (ISL). The target was located at a distance of 21.5 m from the sensor, and 5 m inside the fog chamber and was acquired in March

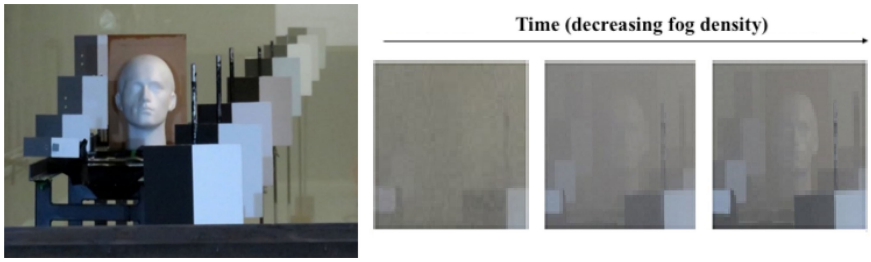


Fig. 13. Picture of the mannequin target: (left) in air, (right) three successive images in presence of a decreasing fog density (acquired with a gated camera as indicated in [45]).

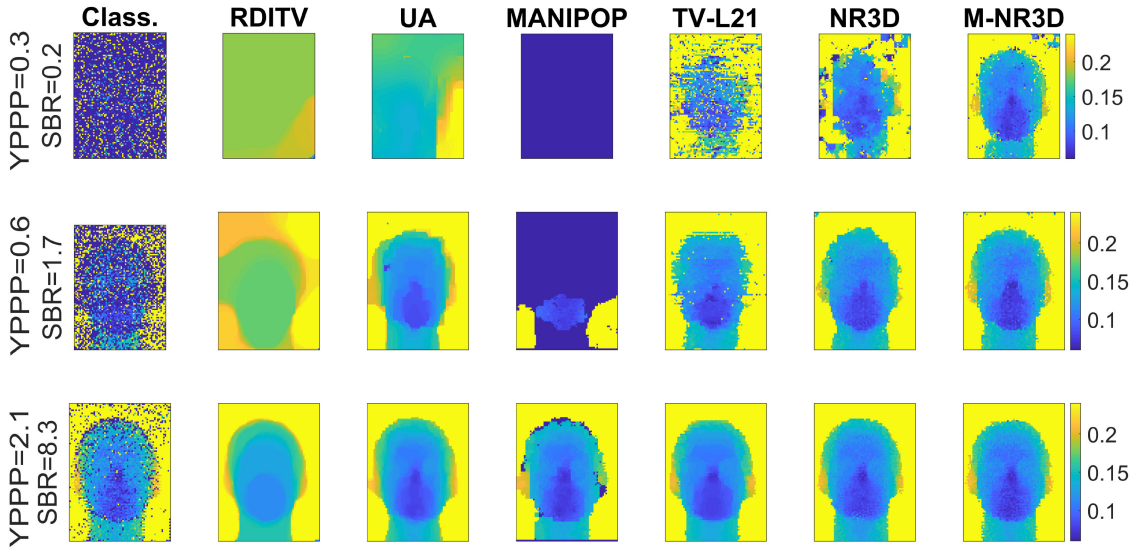


Fig. 14. Depth maps (92×67 pixels) of a life-sized polystyrene head acquired under different level of water fog. From left to right: classical, RDI-TV, UA, MANIPOP, TV-ℓ_{21}, proposed NR3D and M-NR3D algorithms. The acquisition time per pixel is 187 μs.

2017. The scanning imaging system operates at a wavelength of 1550 nm, and uses a Peltier cooled InGaAs/InP single-photon avalanche diode (SPAD) detector. This imaging system uses a pulsed laser with a 15.6 MHz repetition rate. Only a maximum of one photon event can be recorded for each laser pulse. The time between the emission of the outgoing laser pulse and the recorded photon event is recorded and a histogram of photon counts is formed by repeating this operation at each pixel location for a given acquisition time. The reader is invited to see [9], [45] for more details regarding the imaging system and the effect of obscurants on the background level. To study the effect of different fog levels, the chamber was filled with water fog with a high density, then a succession of 3D images were taken as the fog density decreased. In this paper, we study three successive images showing a decreasing level of fog (i.e., decreasing background level) as represented in Fig. 13 (right)). These data present decreasing attenuation lengths (AL) levels with respect to time given by AL$=3.8$, AL$=3.1$ and AL$=2.8$ as indicated in [9], where one attenuation length is defined as the distance after which the transmitted light power is reduced to $1/e$ of its initial value. The data cube were acquired every 60 s, where each image contains 92×67 pixels, and was acquired during 18.5 seconds (i.e., 3 ms acquisition time per pixel). However, since the TCSPC system [21], [22] delivers time-tagged data, this allows the construction of histograms with lower acquisition times to study the photon starved case. Therefore, we build new data by considering 0.75 ms acquisition times per pixel (4.6 s full image acquisition time) and 0.1875 ms acquisition times per pixel (1.15 s full image acquisition time). The data is processed using seven algorithms, namely: Class., RDI-TV, UA, MANIPOP, TV-ℓ_{21}, NR3D and M-NR3D algorithms. All algorithms have been run with different hyperparameters and we only show those providing best performance. For NR3D and M-NR3D, we considered $(r_b, c_b, t_b, h, n_d) = (3, 3, 5, 20, 64)$ and tested nine log-spaced values inside the following intervals $\tau_1 \in [0.001, 1000]$ and $\tau_2 = [0.0005, 10]$. We first focus on the images obtained at the lowest acquisition times (i.e., 0.1875 ms acquisition times per pixel) which represent extreme conditions in terms of SBR and YPPP. Figs. 14 and 15 show the depth images and point clouds obtained for this lowest acquisition time (0.1875 ms acquisition times per pixel) using different algorithms for three time instants, where the fog density decreases from top to bottom rows. Except Class. and RDI-TV that are sensitive to background, all the other algorithms provided

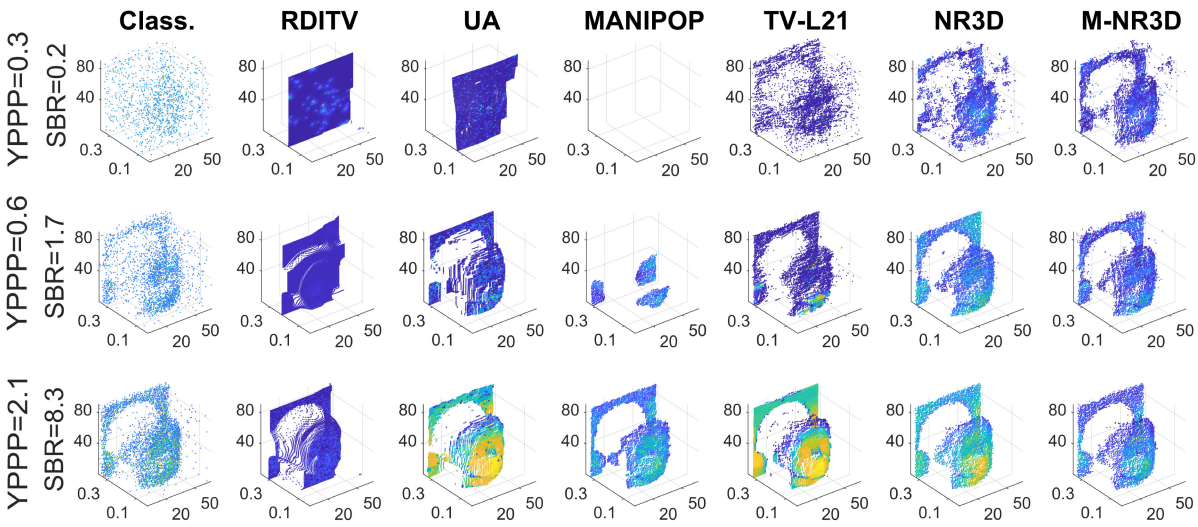


Fig. 15. Cloud points (92×67 pixels) of a life-sized polystyrene head acquired under different level of water fog. From left to right: classical, RDI-TV, UA, MANIPOP, TV-ℓ_{21}, proposed NR3D and M-NR3D algorithms. Normalized reflectivity is colour coded between 0 and 1. The acquisition time per pixel is 187 μs.

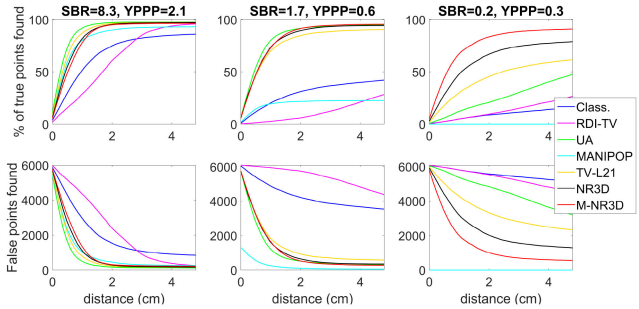


Fig. 16. Percentage of (top) true and (bottom) false detections of different algorithms on a life-sized polystyrene head acquired under different level of water fog. The acquisition time per pixel is 187 μs.

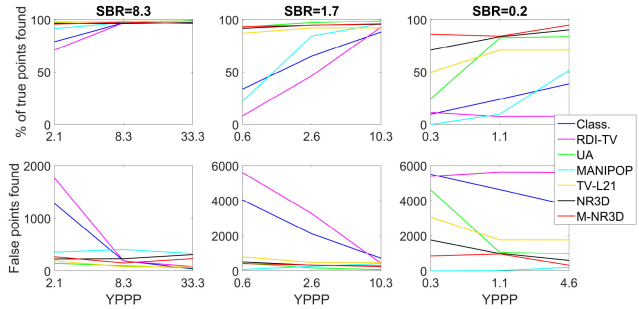


Fig. 17. Percentage of (top) true and (bottom) false detections of different algorithms on a life-sized polystyrene head acquired under different level of water fog (i.e. SBR) and YPPP (i.e., acquisition times).

good results in presence of a low fog density. However, the advantage of both NR3D and M-NR3D with respect to Class. RDI-TV, TV-ℓ_{21}, MANIPOP and UA becomes clear in presence of a dense fog (see top images) since they better exploit the non-local spatial correlation of the histograms. The multidimensional NR3D performs the best under extreme conditions (see top row of Figs. 14 and 15) as it accounts for the temporal evolution of the scene and processes the three datasets jointly. Fig. 16 shows the probabilities of true and false detections of the different algorithms as a function of distance. This figure confirms previous results, where it shows similar performance of the proposed algorithm and UA for low SBR (i.e., background) and better results by M-NR3D for very low SBR. A quantitative evaluation was also performed when considering the higher acquisition times 0.75 ms and 3 ms per pixel. Fig. 17 represents the probabilities of true and false detections at fixed distance ($\tau = 2.4$ cm) when considering different fog levels (columns) and different acquisitions times (x-axis of each sub-figure). Again, MNR3D and NR3D show best performance for extreme scenarios while they perform similarly to UA for higher SBRs or

TABLE I
AVERAGE COMPUTATIONAL TIMES IN SECONDS OF RDI-TV, UA, MANIPOP, TV-ℓ_{21}, NR3D AND M-NR3D FOR THE MANNEQUIN FACE SCENES IN FOG FOR DIFFERENT ACQUISITION (ACQ.) TIMES

Acq. time per pixel	RDI-TV	UA	MANIPOP	TV-ℓ_{21}	NR3D	M-NR3D
3ms	87	8	49	377	63	56
0.75ms	100	6	45	376	50	69
0.187ms	145	5	42	368	30	160

YPPP. These results highlight the interest of M-NR3D and the importance of the joint processing of the multidimensional data. Table I finally shows the averaged computational times of the different algorithms. It can be observed that both RDI-TV and M-NR3D time increases with the data sparsity while MANIPOP and NR3D present opposite behaviour. TV-ℓ_{21} presents the highest processing times while UA is the fastest algorithm as it uses parallel processing. Finally, Table II summarises the main properties and performance of the studied algorithms.

TABLE II
SCOPE OF THE STUDIED ALGORITHMS

	Class.	RDI-TV	UA	TN-Class	SPISTA	TV-ℓ_{21}	MANIPOP	NR3D	M-NR3D
Single surface per pixel	✓	✓	✓	✓	✓	✓	✓	✓	✓
Multiple surfaces per pixel				✓	✓	✓	✓	✓	✓
Background noise			✓		✓	✓	✓	✓	✓
High Background noise			✓			✓		✓	✓
Sparse case (< 1 ppp)		✓	✓			✓	✓	✓	✓
Multispectral									✓
Multitemporal (3D videos)									✓

VII. CONCLUSION

This paper presented a new algorithm to restore multidimensional 3D Lidar images obtained with several wavelengths or at different time instants under challenging cases. The latter are related to the starved photon regime, multilayered objects or a high background level due to obscurants. In addition to the Poisson statistics of the data, the proposed method accounts for available prior knowledge related to the presence of non-local spatial correlation between pixels and the small number of depth layers inside the observed range window. A convex cost function was defined and minimized using an ADMM algorithm that has good convergence properties. The proposed formulation and algorithm showed good restoration results when processing simulated and real images representing different scenarios, i.e., sparse regime, multilayered, multispectral and multitemporal data, and imaging through fog. Despite the good performance of the algorithm, there are still points to improve in future work. For example, the algorithm complexity is related to the size of the data cube and it would be interesting to study other approaches that scale well for very high dimensional data. Considering distributed algorithms and sophisticated strategies to learn non-local correlations [46] are also interesting to perform real time processing. The study of other imaging scenarios such as moving objects in presence of obscurants, and/or the use of complementary information from other sensors (e.g., radar) to learn the weights are of significant interest and will be considered in future studies. Finally, including measures of uncertainty about the inferred parameters using the strategy in [47] will also be investigated in the future.

ACKNOWLEDGMENT

The authors would like to thank F. Christnacher and M. Laurenzis (from Research Institute of Saint-Louis) for their help in acquiring the data for the imaging through obscurants and the editor/reviewers for very helpful comments.

REFERENCES

[1] Y. Altmann *et al.*, "Robust spectral unmixing of sparse multispectral lidar waveforms using gamma Markov random fields," *IEEE Trans. Comput. Imag.*, vol. 3, no. 4, pp. 658–670, Dec. 2017.

[2] Y. Altmann, A. Maccarone, A. Halimi, A. McCarthy, G. Buller, and S. McLaughlin, "Efficient range estimation and material quantification from multispectral lidar waveforms," in *Proc. Sensor Signal Process. Defence*, Sep. 2016, pp. 1–5.

[3] A. M. Pawlikowska, A. Halimi, R. A. Lamb, and G. S. Buller, "Single-photon three-dimensional imaging at up to 10 kilometers range," *Opt. Express*, vol. 25, no. 10, pp. 11 919–11 931, May 2017.

[4] D. B. Lindell, M. O'Toole, and G. Wetzstein, "Single-photon 3-D imaging with deep sensor fusion," *ACM Trans. Graph.*, vol. 37, no. 4, Jul. 2018, Art. no. 113.

[5] D. Shin, F. Xu, F. N. C. Wong, J. H. Shapiro, and V. K. Goyal, "Computational multi-depth single-photon imaging," *Opt. Express*, vol. 24, no. 3, pp. 1873–1888, Feb. 2016.

[6] R. Tobin *et al.*, "Long-range depth profiling of camouflaged targets using single-photon detection," *Opt. Eng.*, vol. 57, 2017, Art. no. 031303.

[7] A. Halimi, A. Maccarone, A. McCarthy, S. McLaughlin, and G. S. Buller, "Object depth profile and reflectivity restoration from sparse single-photon data acquired in underwater environments," *IEEE Trans. Comput. Imag.*, vol. 3, no. 3, pp. 472–484, Sep. 2017.

[8] G. Satat, M. Tancik, and R. Raskar, "Towards photography through realistic fog," in *Proc. IEEE Int. Conf. Comput. Photography*, 2018, pp. 1–10.

[9] R. Tobin, A. Halimi, A. McCarthy, M. Laurenzis, F. Christnacher, and G. S. Buller, "Depth imaging through obscurants using time-correlated single-photon counting," *Proc. SPIE,* vol. 10659, 2018, Art. no. 106590S.

[10] 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 Process.*, vol. 25, no. 5, pp. 1935–1946, Mar. 2015.

[11] S. Hernandez-Marin, A. M. Wallace, and G. J. Gibson, "Multilayered 3D lidar image construction using spatial models in a Bayesian framework," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 6, pp. 1028–1040, Jun. 2008.

[12] A. Halimi *et al.*, "Restoration of intensity and depth images constructed using sparse single-photon data," in *Proc. 24th Eur. Signal Process. Conf.*, 2016, pp. 86–90.

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

[14] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, no. 1/4, pp. 259–268, Nov. 1992.

[15] P. Sprechmann, I. Ramirez, G. Sapiro, and Y. C. Eldar, "C-Hilasso: A collaborative hierarchical sparse modeling framework," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4183–4198, Sep. 2011.

[16] H. K. Aggarwal and A. Majumdar, "Hyperspectral unmixing in the presence of mixed noise using joint-sparsity and total variation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 9, pp. 4257–4266, Sep. 2016.

[17] M. D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Collaborative sparse regression for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 341–354, Jan. 2014.

[18] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.

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

[20] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Total variation spatial regularization for sparse hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4484–4502, Nov. 2012.

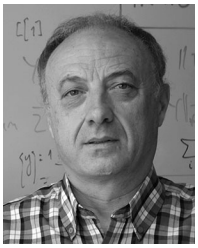
[21] A. McCarthy *et al.*, "Kilometer-range depth imaging at 1550 nm wavelength using an InGaAs/InP single-photon avalanche diode detector," *Opt. Express*, vol. 21, no. 19, pp. 22 098–22 113, Sep. 2013.

[22] A. M. Wallace, J. Ye, N. Krichel, A. McCarthy, R. Collins, and G. S. Buller, "Full waveform analysis for long-range 3D imaging laser radar," *EURASIP J. Adv. Signal Process.*, vol. 2010, no. 1, Dec. 2010, Art. no. 896708.
[23] X. Ren *et al.*, "High-resolution depth profiling using a range-gated CMOS SPAD quanta image sensor," *Opt. Express*, vol. 26, no. 5, pp. 5541–5557, Mar. 2018.
[24] S. Hernandez-Marin, A. Wallace, and G. Gibson, "Bayesian analysis of Lidar signals with multiple returns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 12, pp. 2170–2180, Dec. 2007.
[25] A. Halimi, R. Tobin, A. McCarthy, S. McLaughlin, and G. S. Buller, "Restoration of multilayered single-photon 3D lidar images," in *Proc. 25th Eur. Signal Process. Conf.*, 2017, pp. 708–712.
[26] F. Christnacher, S. Schertzer, N. Metzger, E. Bacher, M. Laurenzis, and R. Habermacher, "Influence of gating and of the gate shape on the penetration capacity of range-gated active imaging in scattering environments," *Opt. Express*, vol. 23, no. 26, pp. 32 897–32 908, Dec. 2015.
[27] D. Shin, J. H. Shapiro, and V. K. Goyal, "Computational single-photon depth imaging without transverse regularization," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2016, pp. 973–977.
[28] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 490–530, 2005.
[29] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
[30] J. Salmon, Z. Harmany, C.-A. Deledalle, and R. Willett, "Poisson noise reduction with non-local PCA," *J. Math. Imag. Vis.*, vol. 48, no. 2, pp. 279–294, 2014.
[31] C. Lanaras, J. Bioucas-Dias, E. Baltsavias, and K. Schindler, "Super-resolution of multispectral multiresolution images from a single sensor," in *Proc. Large Scale Comput. Vis. Remote Sens. Imagery*, 2017, pp. 20–28.
[32] J. Yang, X. Ye, K. Li, C. Hou, and Y. Wang, "Color-guided depth recovery from RGB-D data using an adaptive autoregressive model," *IEEE Trans. Image Process.*, vol. 23, no. 8, pp. 3443–3458, Aug. 2014.
[33] J. Tachella *et al.*, "Bayesian 3-D reconstruction of complex scenes from single-photon lidar data," *SIAM J. Imag. Sci.*, vol. 12, no. 1, pp. 521–550, 2019.
[34] W. Marais and R. Willett, "Proximal-gradient methods for poisson image reconstruction with bm3D-based regularization," in *Proc. IEEE 7th Int. Workshop Comput. Adv. Multi-Sensor Adaptive Process.*, 2017, pp. 183–187.
[35] A. Halimi, N. Dobigeon, J. Y. Toumeret, S. McLaughlin, and P. Honeine, "Unmixing multitemporal hyperspectral images accounting for endmember variability," in *Proc. 23rd Eur. Signal Process. Conf.*, Aug. 2015, pp. 1656–1660.
[36] S. Henrot, J. Chanussot, and C. Jutten, "Dynamical spectral unmixing of multitemporal hyperspectral images," *IEEE Trans. Image Process.*, vol. 25, no. 7, pp. 3219–3232, Jul. 2016.
[37] P. A. G. Thouvenin, N. Dobigeon, and J. Y. Tourneret, "A hierarchical Bayesian model accounting for end member variability and abrupt spectral changes to unmix multitemporal hyperspectral images," *IEEE Trans. Comput. Imag.*, vol. 4, no. 1, pp. 32–45, Mar. 2017.
[38] Y. Zhang, M. P. Edgar, B. Sun, N. Radwell, G. M. Gibson, and M. J. Padgett, "3D single-pixel video," *J. Opt.*, vol. 18, no. 3, Feb. 2016, Art. no. 035203.
[39] J. Eckstein and D. P. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Math. Program.*, vol. 55, no. 1, pp. 293–318, 1992.
[40] P. L. Combettes and J.-C. Pesquet, *Proximal Splitting Methods in Signal Processing*. New York, NY, USA: Springer, 2011, pp. 185–212.
[41] M. Afonso, J. Bioucas-Dias, and M. Figueiredo, "An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681–695, Mar. 2011.
[42] A. Halimi, J. M. Bioucas-Dias, N. Dobigeon, G. S. Buller, and S. McLaughlin, "Fast hyperspectral unmixing in presence of nonlinearity or mismodeling effects," *IEEE Trans. Comput. Imag.*, vol. 3, no. 2, pp. 146–159, Jun. 2017.
[43] K. Yan, L. Lifei, D. Xuejie, Z. Tongyi, L. Dongjian, and Z. Wei, "Photon-limited depth and reflectivity imaging with sparsity regularization," *Opt. Commun.*, vol. 392, pp. 25–30, 2017.
[44] D. Scharstein and C. Pal, "Learning conditional random fields for stereo," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.
[45] R. Tobin, A. Halimi, A. McCarthy, M. Laurenzis, F. Christnacher, and G. S. Buller, "Three-dimensional single-photon imaging through obscurants," *Opt. Express*, vol. 27, no. 4, pp. 4590–4611, Feb. 2019.
[46] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman, "PatchMatch: A randomized correspondence algorithm for structural image editing," *ACM Trans. Graph.*, vol. 28, no. 3, Aug. 2009, Art. no. 24.
[47] M. Pereyra, "Maximum-a-posteriori estimation with Bayesian confidence regions," *SIAM J. Imag. Sci.*, vol. 10, no. 1, pp. 285–302, 2017.

Abderrahim Halimi (S'11–M'14) received the Eng. degree in electronics from the National Polytechnic School of Algiers, El-Harrach, Algeria, in 2009, and the M.Sc. and Ph.D. degrees in signal processing from Institut National Polytechnique de Toulouse, Toulouse, France, in 2010 and 2013, respectively. He was a Postdoctoral Research Associate from 2013 to 2018, and is currently an Assistant Professor and Royal Academy of Engineering Research Fellow with the School of Engineering and Physical Sciences, Heriot-Watt University. His research activities focus on statistical signal and image processing, with a particular interest in Bayesian inverse problems with applications to remote sensing (hyperspectral imaging, satellite altimetry), single-photon depth imaging and medical imaging.

Rachael Tobin received the B.Sc. (Hons.) degree in physics from Heriot-Watt University, Edinburgh, U.K, in 2015. She began her DSTL sponsored Ph.D. degree with Heriot-Watt University under the supervision of Prof. Gerald Buller in 2015. She is currently a Postdoctoral Research Associate with the Single-Photon Group, Institute of Photonics and Quantum Sciences, Heriot-Watt University. She is currently a doctoral candidate, expecting to graduate in 2019. Her current research interests include single-photon LiDAR in extreme environments and imaging through obscurants at low light levels.

Aongus McCarthy (M'03) received the B.Sc. degree from University College Galway, Galway, Ireland, in 1989, the Diploma degree in electronics engineering from the Institute of Technology, Carlow, Ireland, in 1990, the B.Sc. degree in physical optoelectronics from Essex University, Essex, U.K., in 1991, and the Ph.D. degree in physics from Heriot-Watt University, Edinburgh, U.K., in 2002. He worked in industry from 1992 to 1997 as a Design Team Leader on the development of a thermal transfer printing system. He is currently a Research Fellow with the School of Engineering and Physical Sciences, Heriot-Watt University. His research interests include optical and optomechanical system design, time-of-flight depth imaging, single-photon counting technologies, and microscope systems. He is a member of the Optical Society of America and the IEEE Photonics Society.



José Bioucas-Dias (S'87–M'95–SM'15–F'17) received the EE, M.Sc., Ph.D., and Habilitation degrees in electrical and computer engineering from Instituto Superior Técnico (IST), Universidade Técnica de Lisboa (now Universidade de Lisboa), Lisbon, Portugal, in 1985, 1991, 1995, and 2007, respectively. Since 1995, he has been with the Department of Electrical and Computer Engineering, IST, where he is a Professor and teaches inverse problems in imaging and electric communications. He is also a Senior Researcher with the Pattern and Image Analysis group of the Instituto de Telecomunicacoes, which is a private non-profit research institution.

His research interests include inverse problems, signal and image processing, pattern recognition, optimization, and remote sensing. He has introduced scientific contributions in the areas of imaging inverse problems, statistical image processing, optimization, phase estimation, phase unwrapping, and in various imaging applications, such as hyperspectral and radar imaging. He was included in Thomson Reuters' Highly Cited Researchers 2015 and 2018 lists and was the recipient of the IEEE GRSS David Landgrebe Award for 2017.

Stephen McLaughlin (F'11) was born in Clydebank, Scotland, in 1960. He received the B.Sc. degree in electronics and electrical engineering from the University of Glasgow, Glasgow, U.K., in 1981 and the Ph.D. degree from the University of Edinburgh, Edinburgh, U.K., in 1990. From 1981 to 1984, he was a Development Engineer in industry. In 1986, he joined the University of Edinburgh where he studied the performance of linear adaptive algorithms in high noise and nonstationary environments. In 1988, he joined the academic staff at Edinburgh, and from 1991 until 2001 he held a Royal Society University Research Fellowship to study nonlinear signal processing techniques. In 2002, he was awarded a Personal Chair in Electronic Communication Systems, University of Edinburgh. In October 2011, he joined Heriot-Watt University as a Professor of Signal Processing and Head of the School of Engineering and Physical Sciences. His research interests lie in the fields of adaptive signal processing and nonlinear dynamical systems theory and their applications to biomedical, energy and communication systems. He is a Fellow of the Royal Academy of Engineering, of the Royal Society of Edinburgh, of the Institute of Engineering and Technology.

Gerald S. Buller received the B.Sc. (Hons.) degree in natural philosophy from the University of Glasgow, Glasgow, U.K., in 1986 and the Ph.D. degree in physics from Heriot-Watt University, Edinburgh, U.K., in 1989. He is currently a Professor of physics with Heriot-Watt University. In 2002, he cofounded Helia Photonics Ltd., based in Livingston, U.K. In 2015, he was the recipient an EPSRC Established Career Fellowship in Quantum Technology to research sparse photon imaging. His current research interests include single-photon detection methods, particularly at infrared wavelengths, and applications involving the use of single photons. These applications include single-photon depth imaging and novel quantum communications protocols. He is a Fellow of the Royal Society of Edinburgh, the U.K. Institute of Physics, and the Optical Society of America.