

Optical Engineering

OpticalEngineering.SPIEDigitalLibrary.org

Peak detection approaches for time-correlated single-photon counting three-dimensional lidar systems

Gustav Tolt
Christina Grönwall
Markus Henriksson

SPIE.

Gustav Tolt, Christina Grönwall, Markus Henriksson, "Peak detection approaches for time-correlated single-photon counting three-dimensional lidar systems," *Opt. Eng.* **57**(3), 031306 (2018), doi: 10.1117/1.OE.57.3.031306.

Peak detection approaches for time-correlated single-photon counting three-dimensional lidar systems

Gustav Tolt,^{a,*} Christina Grönwall,^{a,b} and Markus Henriksson^a

^aFOI (Swedish Defence Research Agency), Linköping, Sweden

^bLinköping University, Department of Electrical Engineering, Linköping, Sweden

Abstract. Time-correlated single-photon counting lidar provides very high-resolution range measurements, making the technology interesting for 3D imaging of objects behind foliage or other obscuration. We study six peak detection approaches and compare their performance from several perspectives: detection of double surfaces within the instantaneous field of view, range accuracy, performance under sparse sampling, and the number of outliers. The results presented are based on reference measurements of a characterization target. Special consideration is given to the possibility of resolving two surfaces closely separated in range within the field of view of a single pixel. An approach based on fitting a linear combination of impulse response functions to the collected data showed the best overall performance. © 2018 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.OE.57.3.031306]

Keywords: time-correlated single-photon counting; peak detection; matched filter; Lucy–Richardson; mixture of Gaussians; mixture of instrument response functions.

Paper 171381SS received Sep. 1, 2017; accepted for publication Jan. 16, 2018; published online Feb. 13, 2018.

1 Introduction

Time-of-flight (ToF) lidar is a powerful technique for collecting accurate three-dimensional (3-D) information about the surrounding area. After the emission of a laser pulse, the signal received at the detector is the aggregated contributions from all reflecting surfaces within the instantaneous field of view (IFOV) of the sensor. By performing time-resolved measurements, the distance to the individual surfaces can be measured and the data represented as a 3-D point cloud, which can be used for higher level tasks, such as object recognition or mapping. The range information improves the chance of separating and locating objects covering only part of the IFOV. The smaller the spatial separation between surfaces, the higher range resolution is needed. Time-correlated single-photon counting (TCSPC) is a lidar technology that can provide very high-range resolution data and hence separate closely spaced objects.¹ TCSPC lidar emits very weak laser pulses and detects the time of arrival of the first photon to be absorbed by the detector with very high temporal accuracy. By performing a large number of measurements, a histogram of arrival times is assembled, which represents the laser radar response of the scene within the IFOV. With a large enough number of measurements, detections will accumulate in the histogram bins corresponding to the range of a reflecting surface, making single surfaces appear as clear peaks above the shot noise of the background detections. The shape of the peak is determined by the instrument response function (IRF), which depends on factors such as the laser pulse width, detector timing jitter, and trigger jitter. Signals from closely spaced surfaces will overlap and may appear as a single, broadened peak in the histogram. In this article, we study signal processing approaches for detecting multiple surfaces at different ranges within one IFOV, with special emphasis on the performance for separating closely spaced

surfaces whose contributions to the histogram overlap. The application in mind is detection and segmentation of partly occluded objects, for example, objects behind vegetation. For this application, the TCSPC system must be able to detect and localize multiple surfaces at different ranges within one IFOV. The number of surfaces is unknown *a priori* and must be estimated in the signal processing.

Several techniques for surface detection in TCSPC lidar data have been proposed in the literature. Most of them try to minimize the number of detected photons, and hence the integration time needed to determine the distance to a single reflecting surface. Examples of algorithms to detect multiple reflecting surfaces have also been published. There are various approaches based on Markov chain Monte Carlo simulations reported in the literature.^{2–4} These algorithms were designed for low photon counts, i.e., short integration times or targets with low reflectance. Shin et al.⁵ proposed another peak detection algorithm (SPISTA) for low photon count but based on linear optimization. It was claimed⁵ that it outperforms a mixture of Gaussian (MoG)-based estimator that is using a greedy histogram-data-fitting strategy. There are also approaches based on local maximum detection and cross correlation,⁶ but these approaches are generally not intended for resolving multiple surfaces close to each other in range as the local maxima are distorted when surfaces are so close together that the IRFs overlap.

In our work, we assume longer integration times than what has been previously reported in the literature. The goal of this paper is to make true object surfaces stand out from the background noise. The increased signal-to-noise ratio (SNR) of the histograms is used to detect multiple closely spaced surfaces at different distances. The cost of the increased capability of detecting objects obscured by, e.g., vegetation is longer measurement time. The optimal balance between

*Address all correspondence to: Gustav Tolt, E-mail: gustav.tolt@foi.se

short integration times corresponding to high area coverage and high SNR allowing detection of multiple surfaces closely spaced in range is obviously application dependent.

In this work, we study a number of peak detection approaches and assess their performance from different perspectives; the possibility of detecting double surfaces within one IFOV, geometrical accuracy of the detections, the number of outliers, and how reduced signal quality affects the performance.

The remainder of this article is organized as follows: in the next section, a number of algorithms for preprocessing and peak detection are described, along with the tested peak detection approaches (mixes of the described algorithms). In Sec. 3, the setup of the tests is described, including a short description of the TCSPC system and the reference measurements performed. Section 4 describes the procedure for measuring range accuracy and in Sec. 5 results are presented. A discussion of the results is included in Sec. 6 and conclusions are given in Sec. 7.

2 Peak Detection Approaches

A fundamental requirement on a peak detection approach for our intended application is that no *a priori* assumption of the number of surfaces in the IFOV is made. It should also meet the following quality criteria:

1. High probability of detecting surfaces so that objects partially obscured by vegetation can be detected,
2. high-range precision so that geometrical properties of objects can be derived from the data collected, and
3. low false alarm rate, to prevent adding significant amounts of noise to already fragmented and spurious data collected in cluttered environments.

In this work, we use the term “approach” for denoting an algorithm for detecting and locating peaks in TCSPC histograms, possibly preceded by an algorithm for preprocessing such histograms.

Two types of preprocessing algorithms are considered: matched filter (MF) and Lucy–Richardson deconvolution (LR).⁷ Both algorithms use the impulse response function (IRF) of the measurement system but for completely different purposes. MF is used for suppressing noise and improving range precision and is obtained by convolving the TCSPC histogram with a time-reversed version of the IRF. Another advantage of MF is that the peak generated by a single surface after filtering is symmetric and hence more easily described by a fitting function. The purpose of the LR deconvolution of TCSPC data is to decompose histograms into modes corresponding to the individual surfaces. LR involves iterative convolutions with both the IRF and its time-reversed version, but still it is computationally fast provided that the number of iterations is kept low. A side effect of LR is that it amplifies noise rather than reducing it, which can be mitigated through imposing constraints on the deconvolved signal or removing spurious points in a postprocessing step.

For the problem of detecting peaks in histograms, the following algorithms are considered:

- Local maximum peak detection (LM),

- Mixture of Gaussian estimation using expectation maximization (EMGM),⁸
- Mixture of IRF estimation using linear minimization (SPISTA),⁵
- Mixture of IRF estimation using least squares minimization (IRF-LS),¹ and
- Eigenvalue analysis parameter estimation in the Fourier domain based on the theory of finite rate of innovation (FRI).⁹

The local maximum detection (LM) is performed in two steps. First, the range bins corresponding to LM values are located, using thresholds for height and prominence to reduce the number of false alarms. Then, the position of each peak is estimated with subsample accuracy through quadratic interpolation using the LM range bin and its immediate neighbors.

In the EMGM algorithm, data are modeled as a mixture of Gaussian basis functions. This is different compared with the other fitting methods investigated in this paper, which use representations of the IRF. The Gaussian basis functions allow for using computationally efficient standard implementations. The algorithm is based on expectation-maximization, which is an iterative approach for finding the maximum likelihood estimates of parameters in a statistical model. It has a nonconvex cost function, which means that only finding a local minimum can be guaranteed when used for fitting of multimodal distributions. Thus, good initialization is needed.

The SPISTA algorithm is essentially a reformulation of the multiple peak detection problems as an L1 norm problem with a global optimum. It models the input signal and background noise as Poisson distributions. In SPISTA, the single-photon detection statistics from multiple reflectors within a pixel is described as a multidepth profile and expressed as a sparse signal. The multidepth estimation problem from single-photon observations is formulated as a convex optimization problem by combining the statistics of photon detection data with sparsity of multidepth profiles. Linear minimization is known to be less affected by noise and outliers compared with least squares minimization and also computationally faster. SPISTA has a convex cost function, which is different from the EMGM and IRF-LS algorithms.

The IRF-LS algorithm minimizes the difference between the measured data and a linear combination of IRFs in a least square sense, recursively adding more surfaces until no further improvement is found. The approach is good at accurately detecting high-amplitude peaks with high resolution but is relatively computationally demanding and prone to adding low-amplitude surfaces from imperfections of the IRF description. Since, in our implementation, it never removes surfaces as some published MCMC-based methods do,³ it is more prone to overfitting. Furthermore, it has a nonconvex cost function.

The FRI algorithm reformulates the time-of-flight super-resolution problem as a parameter estimation problem. Based on the theory of finite rate of innovation (FRI) sampling,¹⁰ a bandlimited approximation of the sampling kernel function is used, which allows for formulating the problem as a linear least square problem in the frequency domain. The surface positions are determined through solving an eigenvalue problem, with a threshold applied on the eigenvalues

to determine the number of surfaces in the final result. The method is noniterative and hence computationally attractive. It has shown promising results for both superresolution and resolving overlapping peaks and was reported to provide better estimates than an MCMC-based method and at lower execution time.⁹

The algorithms are combined into the following peak detection approaches:

1. LM with LR preprocessing (LR-LM),
2. EMGM with MF preprocessing (EMGM-MF),
3. IRF-LS,
4. IRF-LS with MF preprocessing (IRF-LS-MF),
5. SPISTA, and
6. FRI with MF preprocessing (FRI-MF).

3 Tests

3.1 Lidar System

The TCSPC lidar data subject to peak detection was recorded using the system described in detail by Henriksson et al.¹ The lidar system uses a 1.55- μm fiber laser source with 22 ps pulse length and pulse repetition frequency of 8 MHz. The optical transceiver is a 500-mm focal length Cassegrain mirror telescope, where the laser light is emitted via a mirror in front of the central obscuration to provide a monostatic configuration. The collected photons are focused on a graded index fiber and, after bandpass filtering, focused on an InGaAs single-photon avalanche detector. The times of detections, measured relative to the emissions of laser pulses, were recorded with 32-ps precision using a Picoquant PicoHarp 300. The detector field of view was continuously swept row by row at 0.04 deg/s during the measurement. The signal is subsequently divided into pixels for small angular intervals during the postprocessing. Due to performing the measurements while sweeping, there will be a spatial broadening of objects, but, with small angular intervals for each pixel, this decrease in spatial resolution can be kept at an acceptable level.¹ The measurements are aimed at accurately finding targets in vegetation or behind camouflage and hence involve longer integration times than in Refs. 2–6. The effective integration time per pixel was 126 ms/pixel, which means that $\sim 10^6$ measurements were performed per pixel. The laser power, however, was low, meaning that on average 1120 detections were recorded per pixel. The background rate, caused by detector dark counts and solar illumination, was estimated to be 0.22 detections/(range bin \times pixel) or in total 870 detections/pixel. Approximately, 78% of the detections are thus uncorrelated to the laser and only contribute noise to the histograms.

The measured IRF was 392-ps full width at half maximum (FWHM). The fitting function described in Ref. 1 consisting of a Gaussian peak with exponential tails was used to describe the IRF in the signal processing. As seen in Fig. 1, the general shape is well described, but some high-frequency features near the peak leave a residual error.

3.2 Target Setup and Measurements

Measurements were performed against a reference target with holes behind which metal plates were placed at different

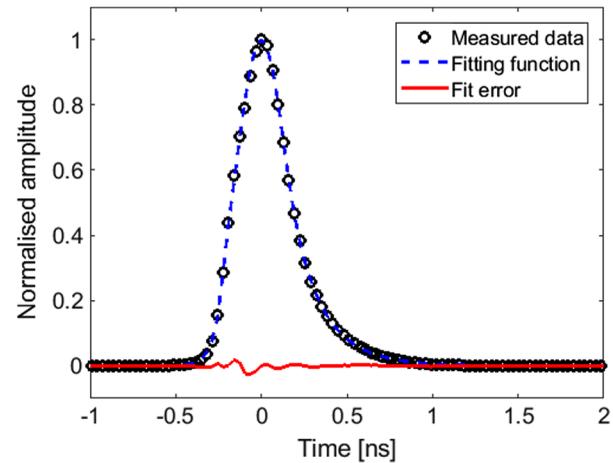


Fig. 1 The measured IRF and the fitting function used in the signal processing. The FWHM of the IRF is 392 ps.

distances: 33, 53, 73, and 93 mm (see Fig. 2). In the further analysis the front surface is designated as S_1 , and the back plates are designated as S_2 , S_3 , S_4 , and S_5 , in order of increasing distance from the front surface. The incidence angle deviated from the normal of the target surfaces with about 7-deg horizontally and 4-deg vertically, respectively.

Scanning across the target means that the histogram of every pixel corresponds to an IFOV in a different position on the target. We are, especially, interested in the cases where the IFOV covers the edge of a hole in the front plate so that the histogram is composed of photons reflected both from the front surface and plate behind the hole. The four different distances 33 mm (S_1 to S_2), 53 mm (S_1 to S_3), 73 mm (S_1 to S_4), and 93 mm (S_1 to S_5) allow us to investigate what distance between reflecting surfaces within the IFOV can be resolved. Figure 3 shows examples of collected histograms for a point at the edge of each hole, in addition to a model consisting of a linear combination of two IRFs fitted to the measured data. The histograms shown in Fig. 3 are created using a different dataset with higher SNR than the data used for peak detection analysis and only serve to illustrate the phenomenon. At different positions at the edges of the holes varying portions of the IFOV will be on the front and back plates, leading to varying proportions of the signal belonging to the first and second peaks in the histogram. The effects of different obscuration coverages or different reflectivities on the obscuration and target will have a similar effect. If the obscuration is caused by foliage or other



Fig. 2 The reference target.

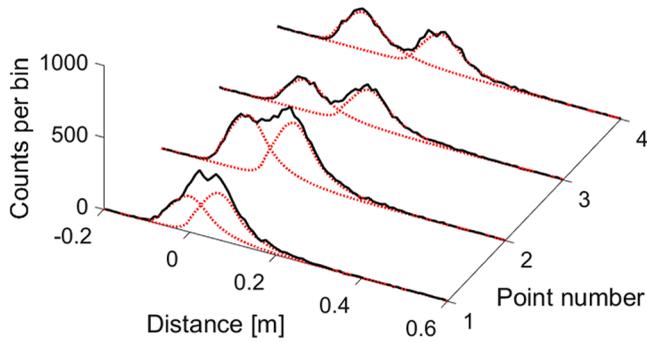


Fig. 3 Measurements at four different points at edges of each hole in the characterization target. The black lines are measured data and the red dotted lines are two fitted IRF peaks. Point number 1 includes surfaces S_1 and S_2 , point 2 S_1 and S_3 , point 3 S_1 and S_4 , and point 4 S_1 and S_5 .

vegetation, it will probably be in the form of several small surfaces with a distribution in range to the target.

3.3 Implementation Details

For the MF and LR algorithms, we use the IRF of the system defined in Henriksson et al.¹ Approaches 2, 4, and 6 (Sec. 2) are computed on MF preprocessed data, mainly to remove noise. The EGMG-MF is initialized using results from the LM algorithm when possible. When the LM algorithm does not detect any peaks, EGMG-MF is initialized with 1 to 5 randomly chosen peak locations. Also, approaches 3 and 4 are initialized using results from the LM algorithm.

To allow for a fair comparison of the approaches, detection thresholds were applied to filter out very weak points, as some algorithms are prone to adding very low amplitude false detections, whereas others have an internal thresholding mechanism built into the processing.

To study the effects of varying SNRs, the original data were also downsampled by randomly choosing 50% and 25% of detections. An example is shown in Fig. 4. Robustness to varying SNRs is important in practical situations as the dwell time on the target can change, or the signal strength can vary because of surface reflectivity differences.

4 Measuring Relative Range Accuracy

To measure the range accuracy, the detected peaks were converted into 3-D points in Euclidean space using the estimated range values in combination with the angle obtained through interpolation, assuming constant angular velocity between the start and stop angle values. The algorithms may introduce a slightly different bias to the absolute range estimates. However, we are only interested in the relative accuracy within the point cloud obtained by the respective approach. As a consequence, the relative position of the reference target in each of the respective datasets has to be determined first. This is achieved through fitting of a plane to the extracted 3-D data points. First, the orientation of the front plane is estimated using RANSAC¹¹-based plane fitting. For this purpose, a point cloud was selected after empiric testing with different parameters and visual inspection of both data and the resulting plane to ensure a good fit. The fitted plane was then used as an input to the assessment of performance for all other approaches, yet shifted in range to compensate for the differences in range bias.

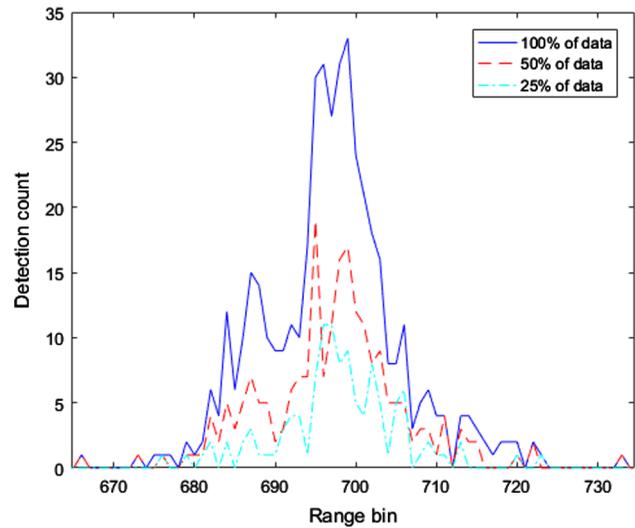


Fig. 4 Range histograms from a pixel within a double surface region computed with 100%, 50%, and 25% of the collected data, respectively.

The position of the front plane implicitly gives the position and orientation of the back plates, which are parallel to the front plate and at known distances (33, 53, 73, and 93 mm). Using the equation of the plane, the normal distance from each point to the front surface is then computed. Based on that, the points are segmented into five clusters corresponding to the front and the four back plates (Fig. 5).

A point is regarded as belonging to the nearest surface provided that the distance is small enough. This is to avoid having outliers far from the surfaces affecting the range accuracy estimation. We introduce a binary weight parameter w_{p,S_i}

$$w_{p,S_i} = \begin{cases} 1 & \text{if } [r_p - \hat{R}_{S_i}] < d_{\max} \\ 0 & \text{otherwise} \end{cases}, \quad (1)$$

where p is an extracted point, r_p is the range associated with p , S_i , $i = 1, \dots, 5$, is the surface number, and \hat{R}_{S_i} is the estimated range to S_i . After studying the typical distribution of range estimation errors, we set $d_{\max} = 0.03$ m. The range errors for the respective surfaces are quantified as a standard root mean square error (RMSE)

$$\text{RMSE}^{(S_i)} = \left[\frac{1}{|S_i|} \sum_{p \in S_i} w_{p,S_i} (r_p - \hat{R}_{S_i})^2 \right]^{1/2}, \quad (2)$$

where $|S_i|$ is the number of points associated with surface S_i (i.e., within range d_{\max}). Outliers are defined as points further away than d_{\max} from the nearest surface.

5 Results

In this section, we present the results obtained from the tests, studying the performance of the approaches regarding the criteria listed in Sec. 2 such as detection of double surfaces, range precision, and the number of outliers. Furthermore, we study how the performance of the respective approaches varies when reducing the signal strength.

Figure 6 shows the points along a line transect across the reference target detected with the respective approach.

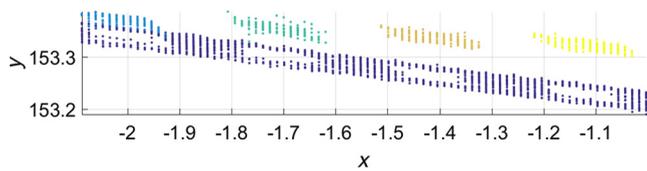


Fig. 5 Extracted points segmented according to their distance to the surfaces. S_1 : large front surface (dark blue), S_2 : surface nearest background (light blue), S_3 : surface farthest in back (yellow). Axes in meters.

Ideally, all points should lie on the solid lines representing the position of the planar plates of the reference target relative to the data. The color of the dots shows the estimated amplitude of the peak. All approaches except EMGM-MF produce the same overall, expected pattern with stronger peaks for single surfaces and weaker peaks in the double surface regions. Though, the LR-LM approach shows somewhat larger intensity variations between surfaces and is not as reliable in terms of drawing conclusions about the reflectivity of the surfaces. This is an expected effect of the LR preprocessing step as it amplifies noise.

The plot shows that none of the approaches accurately resolve the step to the left-most surface (S_2) situated 33 mm

behind the front plate but instead they all produce spurious points in between. We can also see that LR-LM and SPISTA cannot resolve the step to the second closest surface S_3 (at 53-mm distance) and that EMGM-MF still has difficulties at 73 mm. We also note that the range noise is higher for EMGM-MF and FRI-MF than for the other methods.

5.1 Detection of Multiple Surfaces

To measure how well the different approaches can resolve multiple surfaces within the IFOV, a set of regions around range discontinuities were selected manually by visual inspection of the histograms and identifying regions where signals coming from both surfaces could be discerned. Let $n_{DS,max}$ denote the expected maximal number of 3-D points in the selected double surface regions. After visual inspection of the data we set $n_{DS,max} = 224$, implicitly assuming that there is enough signal from both surfaces for the peak detection algorithm to be able to detect them. An in-depth analysis to assess that there really were strong enough signals for each of the histograms to allow for detection of the peaks was not performed, as it is not needed to make comparisons between the performances of the approaches.

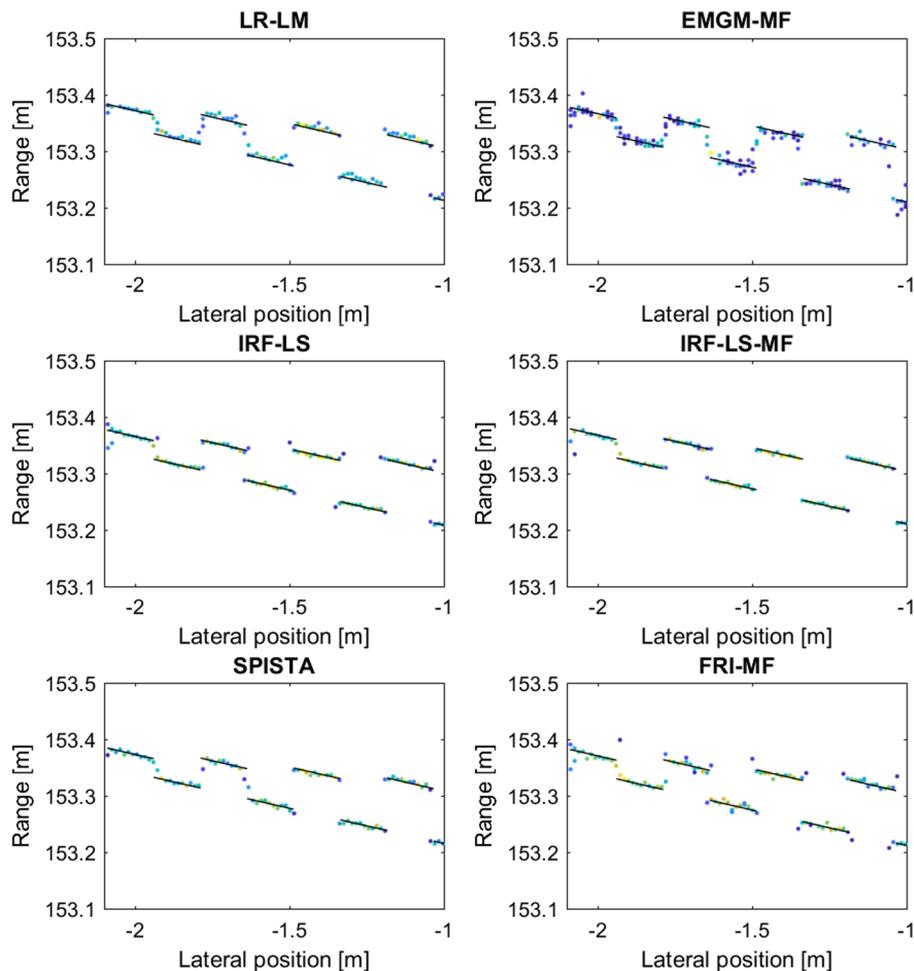


Fig. 6 The figure shows 2-D plots of points along a line transect across the reference target. The line segments show the position of the surfaces of the reference target relative to the data. Each dot corresponds to one detected peak, with dots on the same lateral position belonging to the same histogram. The color of the dots corresponds to the estimated peak amplitude, with the color scale going from blue (weak) via green to yellow (strong).

Table 1 Number of detected 3-D points for the respective surfaces in double surface regions. Only pixels with two detected surfaces that are not considered outliers are included. 100% of the data is used.

Approach	S_1	S_2	S_3	S_4	S_5	Perc. total
LR-LM	27	0	14	3	10	24.1
EMGM-MF	43	17	23	15	12	49.1
IRF-LS	63	1	21	19	22	56.3
IRF-LS-MF	47	0	17	17	11	41.1
SPISTA	27	0	1	14	12	24.1
FRI-MF	64	0	26	21	23	59.8

Table 2 RMSE for points in double surface regions (mm) using 100% of the data. Only pixels with two detected surfaces that are not considered outliers are included. Cases where no points are found are marked with dash.

Approach	S_1	S_2	S_3	S_4	S_5
LR-LM	11.9	—	15.3	6.4	16.8
EMGM-MF	12.8	6.1	10.7	15.1	11.5
IRF-LS	5.0	1.3	7.2	10.6	15.5
IRF-LS-MF	3.4	—	3.8	9.9	10.4
SPISTA	5.6	—	9.3	6.4	14.7
FRI-MF	5.6	—	7.5	11.8	15.4

The results obtained using histogram created with 100% of the collected data are shown in Tables 1 and 2. As our primary interest is detecting and correctly locating the correct number of surfaces, only pixels with two detected surfaces within a distance d_{\max} from the closest surface are considered in these tables. Table 1 shows the number of 3-D points detected for each surface in double surface regions and Table 2 shows the corresponding RMSE. Among the approaches considered, only EMGM-MF produces a significant number of detections at the 33 mm step of the S_1 to S_2 interface. However, from Fig. 6 it is evident that the step is not resolved with EMGM. Instead, the method has produced two detections at intermediate distances, which are so small that the detections are not considered to be outliers. In the later analysis (Sec. 5.3), it will become evident that a single surface is also commonly described as two detections with EMGM, and hence the two detections are not a robust indication of two surfaces.

A significant number of points on S_3 (53-mm distance) are detected with all approaches except SPISTA. The EMGM method detections are still spurious, which becomes evident when comparing the results with Fig. 6. The practical range resolution limit of the system is thus better than 53 mm, somewhat less than the 59 mm indicated by the 392 ps FWHM IRF. The minimum resolvable distance cannot be determined as the target had no steps between 33 and 53 mm and will also depend on the SNR.

Table 3 Number of detected double surface points using 50% of the data.

Approach	S_1	S_2	S_3	S_4	S_5	Perc. total
LR-LM	7	2	6	3	0	8.0
EMGM-MF	51	12	27	19	11	53.6
IRF-LS	72	2	24	21	25	64.3
IRF-LS-MF	33	1	10	9	13	29.5
SPISTA	29	0	2	15	12	25.9
FRI-MF	65	2	19	19	25	58.0

Table 4 RMSE for points in double surface regions (mm) using 50% of the data. Cases where no points were found are marked with a dash.

Approach	S_1	S_2	S_3	S_4	S_5
LR-LM	14.5	6.2	13.3	6.1	—
EMGM-MF	15.5	8.0	11.8	19.8	5.8
IRF-LS	6.2	2.0	6.8	12.5	14.8
IRF-LS-MF	4.2	4.6	4.3	7.0	14.5
SPISTA	5.1	—	6.6	7.3	14.5
FRI-MF	7.6	0.8	6.7	13.4	15.6

The performance of the IRF-LS and the FRI-MF approaches are comparable in terms of double surface detection rates and slightly exceeds that of the other approaches. As expected, IRF-LS-MF with its MF detects somewhat fewer double surfaces but, on the other hand, has a lower RMSE.

Ideally, the sum of points detected for S_2 to S_5 should be equal to that of S_1 , as all double surface regions include S_1 and there should be one detected point per surface. However, sometimes the two detected peaks lie closer to the same surface. This phenomenon is especially common for the EMGM-MF approach and also noticeable for the LR-LM at lower data rates (Tables 3 and 5).

When down-sampling data to emulate shorter integrations times, the performance of the approaches are affected differently (Tables 3–6). We note that the performance of IRF-LS is quite stable even when data quality is reduced down to 25%. Also, the SPISTA approach copes relatively well with this lower data rate. The performance of the LR-LM approach degrades significantly with decreasing signal quality. The EMGM-MF approach continues to indicate multiple surfaces at intermediate distances. In the tables, those show up as detections with large RMSE.

5.2 Range Accuracy for Single Surfaces

To measure the range accuracy for single surfaces, the RMSE was computed for a subset of points corresponding to parts of the front surface S_1 with no interference of signals from

Table 5 Number of detected double surface points using 25% of the data.

Approach	S_1	S_2	S_3	S_4	S_5	Perc. total
LR-LM	1	0	0	1	0	0.9
EMGM-MF	71	15	37	22	11	69.6
IRF-LS	64	3	25	18	22	58.3
IRF-LS-MF	8	0	2	8	0	8.0
SPISTA	34	0	6	16	12	30.3
FRI-MF	12	1	0	4	7	10.7

Table 6 RMSE for points in double surface regions (mm) using 25% of the data. Cases where no points were found are marked with dash.

Approach	S_1	S_2	S_3	S_4	S_5
LR-LM	5.6	—	—	15.6	—
EMGM-MF	14.3	8.9	11.4	14.9	13.6
IRF-LS	8.0	7.2	11.0	13.7	15.1
IRF-LS-MF	5.9	—	11.5	7.5	—
SPISTA	6.1	—	5.1	10.0	15.8
FRI-MF	8.1	4.6	—	7.1	14.9

the back plates. The results are shown in Table 7. The sample distance in the histograms corresponds to 4.8 mm, meaning that several of the approaches consistently give subsample accuracy.

From Table 7, we notice only minor differences between most of the approaches with LR-LM and EMGM-MF falling slightly behind. None of the approaches show a significant degradation in range precision as the signal is downsampled. The reason for the FRI-MF approach showing an improvement for decreasing data rate is that detected low-amplitude points with large error disappeared as the data were downsampled and the signal level fell beneath the threshold.

5.3 Detection Statistics

In Table 8, we show overall statistics for the approaches in terms of the distribution of the number of points detected on the surfaces and the number of outliers. The number of histograms resulting in one surface being detected with the respective approach (N_1) should be compared with the total number of histograms that correspond to a single surface, which is ~ 1400 . Analogously, the number of histograms for which two surfaces are detected is denoted by N_2 . The number of pixels in the edge regions, where two surfaces should be detectable in the histogram, is somewhere between 320 and 350. The number of histograms yielding more detected points (N_3 to N_5) should be zero.

The uncertainty regarding the ground truth statistics is due to the spatial sampling pattern across the reference target

Table 7 RMSE for single surface points on front plate (S_1) (mm) for different percentages of data used (100%, 50%, and 25%)

Approach	100 %	50 %	25 %
LR-LM	5.3	5.6	6.9
EMGM-MF	5.8	7.3	10.2
IRF-LS	1.5	2.0	3.6
IRF-LS-MF	1.6	2.0	2.6
SPISTA	2.3	2.5	3.4
FRI-MF	4.5	3.1	2.8

Table 8 The number of histograms resulting in 0 to 6 detected peaks (N_0 to N_6), the total number of detected 3-D points (NP_{tot}), and the number of outliers ($NP_{outliers}$), using 100% of the data.

Approach	N_0	N_1	N_2	N_3	N_4	N_5	NP_{tot}	$NP_{outliers}$
LR-LM	24	1541	154	6	0	0	1869	220
EMGM-MF	238	260	798	435	0	0	3161	79
IRF-LS	0	1518	234	3	0	0	1995	7
IRF-LS-MF	20	1542	163	4	3	1	1897	6
SPISTA	0	1642	94	0	0	0	1830	0
FRI-MF	20	1232	526	3	0	0	2293	19

being not perfectly aligned with the geometrical features of the target (e.g., edges between surfaces), which together with the extent of the laser footprint makes the signal level vary in double surface regions, so that some histograms may not be able to yield detections of two surfaces with any approach. N_0 denotes the number of histograms for which no points were found and $NP_{outliers}$ is the number of points lying farther than 3 cm from the nearest surfaces. The total number of detected 3-D points (NP_{tot}) should lie somewhere between 2040 and 2100 points; the exact number is not needed to identify qualitative differences in performance between approaches.

The EMGM-MF approach is prone to overfitting and hence often detects more peaks than that are actual surfaces. The low number of detected single peaks (the N_1 column) shows that a vast majority of the pixels that should result in single peaks, given the shape of the reference target, instead end up being split up in multiple peaks lying close to each other. The main cause of this is that the EMGM-MF approach tries to describe the non-Gaussian IRF (after MF) with a Gaussian function, often leading to a better fit using the sum of several Gaussians corresponding to several closely spaced peaks even if there is actually only one surface. Such detected point pairs or triplets could be merged into single detections in a postprocessing step but also that would be very likely to affect the double peaks stemming from actual multiple surfaces. The FRI-MF approach shows similar tendencies albeit to a lower extent,

again because the IRF is not exactly described by the truncated Fourier series.

The performance of the LR-LM stands out with its large number of outliers. As can be seen in Fig. 6, it is partly because it produces ghost points in between two surfaces but also a significant number of points around the surface but farther away than 3 cm, which we chose as threshold for attributing a point to a surface.

The SPISTA approach misses no detections but does not perform as well as IRF-LS and IRF-LS-MF approaches in terms of resolving double surfaces. From the results, we conclude that IRF-LS produces the best overall detection result, in that it detects the most double surfaces, has no missed detections, and only a few outliers.

6 Discussion

Parameter values and thresholds in the peak detection approaches affect the number of correctly detected points, the number of outliers, and RMSE values. In this work, the values were chosen empirically rather than with an exhaustive analysis of the relation between parameter settings, system parameters, detection probability, false alarm rate, and accuracy. For some approaches, the number of outliers can easily be decreased by increasing the minimum signal strength threshold but that would impair the ability to find weak surfaces. Analogously, the thresholds could be adjusted to find more multiple surfaces at the expense of adding noise points. The FRI-MF approach may suffer in resolution performance because of too hard truncation of the Fourier series describing the IRF. Including more components has the cost of more false alarms and longer execution time. Finding the optimal parameters to maximize performance under different circumstances and with different requirements will be part of the next phase of our work. Nevertheless, we notice that for the IRF-LS, IRF-LS-MF, and SPISTA approaches, it was possible to find parameters that give a good trade-off between detection, false alarm rate, accuracy, and signal quality, which we consider a strength in itself. In cases with a high contrast between the highest peak and the threshold for the lowest peak, which is counted as a surface, the IRF-LS algorithm has been found to add extra peaks to compensate for imperfections in the description of the IRF. Because the threshold was set higher compared with the peak value than the error in the IRF description ($\sim 3\%$) this did not generate any false alarms here.

The low performance for the EMGM-MF approach is due to the fact that an MoG is not a good model for the wide peaks resulting after the MF preprocessing step. Shin⁵ stated that SPISTA would have better performance than an MoG approach, which is the case also in our study.

We are ultimately aiming for peak detection that runs in real time to allow for human-in-the-loop experiments in realistic scenarios. The current implementations of the algorithms are made in Matlab with varying degrees of optimization, and it is not relevant to report execution time at this stage. However, we notice that the approaches involving least square optimization with recursive addition of components (IRF-LS and IRF-LS-MF) are noticeably slower than the other approaches.

In this work, we have focused on the range accuracy of the detected points. When creating a point cloud, the 3-D coordinates of the detected points are defined by the viewing

angle of the sensor and not the actual lateral position of the surfaces within the IFOV. A surface may be detected as soon as it is within the footprint of the laser beam and as a consequence, objects typically appear wider than they really are. This effect cannot be mitigated by analyzing each range histogram individually. Instead, it requires taking the data in the vicinity of the point into account, either through post-processing of the extracted 3-D point cloud or by including neighboring histograms in the peak detection process. Future work will include efforts to address this problem. Similar effects would appear from cross talk between pixels in an array detector.

The experiments used for this investigation were performed with a maximum of two discrete surfaces within the IFOV corresponding to one pixel, to have control over the true distances between peaks. One interesting application is targets behind foliage, where the obscuration consists of several small objects at varying distances to the target surface. The results shown here should still be valid for the minimum distance between different surfaces, which is resolvable, regardless of whether there are two or more surfaces. The important factors to accurately measure the distance to the target surface are thus the number of photons detected reflected off the target surface and the distance from the last obscuring object to the target surface. The distance between different obscuring objects will impact whether they can be resolved or not, but this is often of little or no importance.

In this work, we assumed a static scene and a stationary sensor, which allows us to create and analyze range data as one-dimensional (1-D) histograms. Adding platform motion requires a different approach to handle the collected data. Future work will address efficient data structures for accumulation and analysis of TCSPC data as well as advanced positioning techniques for accurate 3-D registration of TCSPC data acquired with a moving sensor. This will probably need significantly higher data collection rate than the results shown here. Higher laser power that gives a detection probability of up to 50% for every laser pulse, which has been shown to be the optimum power to detect an obscured surface,¹² could be one solution, which may also require a faster scanning mechanism. Array detectors perform multiple measurements for every laser pulse may be another solution. Higher data collection rate will also allow for small field of regard 3-D imaging of moving objects.

7 Conclusions

Of the approaches discussed herein, IRF-LS is our top candidate for peak detection in TCSPC data. It has the best performance regarding detection of multiple surfaces, which gives good range of accuracy and a low false alarm rate. However, it is computationally demanding in its current implementation and depending on what degree of optimization is achievable, other approaches may be considered in practical applications. Among the faster approaches, SPISTA gives the best results. Hence, a combination of several approaches is worthwhile investigating, where an advanced approach may be applied when multiple surfaces are present or when high accuracy is needed, and a less computationally demanding technique may be used to process the signal received from single surfaces. The width of a peak in a histogram might serve as an indicator of which approach to apply.

The processing of the data from the TCSPC lidar system with 392 ps FWHM IRF, indicating resolution of 59-mm, could with several methods resolve contributions from two surfaces at 53-mm range difference but not separate the contributions at 33-mm range difference.

References

1. M. Henriksson et al., "Continuously scanning time-correlated single photon counting single pixel 3D lidar," *Opt. Eng.* **56**(3), 031204 (2017).
2. Y. Altmann et al., "Lidar waveform-based analysis of depth images constructed using sparse single-photon data," *IEEE Trans. Image Process.* **25**(5), 1935–1946 (2016).
3. A. M. Wallace et al., "Full waveform analysis for long-range 3D imaging laser radar," *EURASIP J. Adv. Signal Process.* **2010**, 896708 (2010).
4. W. Yin et al., "Approach for LIDAR signals with multiple returns," *Appl. Opt.* **53**(30), 6963–6969 (2014).
5. D. Shin et al., "Computational multi-depth single-photon imaging," *Opt. Express* **24**(3), 1873–1888 (2016).
6. A. McCarthy et al., "Kilometer-range, high resolution depth imaging via 1560 nm wavelength single-photon detection," *Opt. Express* **21**(7), 8904–8915 (2013).
7. M. Bertero et al., "Image deblurring with Poisson data: from cells to galaxies," *Inverse Prob.* **25**(12), 123006 (2009).
8. C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York (2006).
9. A. Bhandari, A. M. Wallace, and R. Raskar, "Super-resolved time-of-flight sensing via FRI sampling theory," in *IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, IEEE (2016).
10. M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Trans. Signal Process.* **50**(6), 1417–1428 (2002).
11. M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM* **24**(6), 381–395 (1981).
12. P. Jonsson et al., "Experimental evaluation of penetration capabilities of a Geiger-mode APD array laser radar system," *Proc. SPIE* **10434**, 1043405 (2017).

Gustav Tolt received his MSc degree in engineering physics from Chalmers University of Technology, Gothenburg, in 2000 and his PhD degree in industrial measurement technology from Örebro University in 2005. His research interests include lidar data processing, 3-D feature extraction, and object detection.

Christina Grönwall received her MSc degree in computer science 1992 and her PhD in automatic control in 2006. Currently, she is working as a deputy research director at the sensor informatics group, Swedish Defence Research Agency (FOI), Sweden. Her research interests are in signal processing, geometric fitting, system modeling, and performance analysis of measurement systems, and particularly laser radar systems. She is an adjunct associate professor at the Automatic Control Group, Linköping University, Sweden.

Markus Henriksson is a senior scientist at the Swedish Defence Research Agency (FOI). He received his MSc degree in engineering physics from Chalmers University of Technology, Gothenburg, Sweden in 2002 and his PhD degree in applied physics from the Royal Institute of Technology, Stockholm, Sweden, in 2010. His current research interests include single-photon counting lidar, active imaging, and laser propagation in turbulence.