3D imaging from multipath temporal echoes: Supplementary Material

Alex Turpin^{1, \dagger ,*, Valentin Kapitany^{2, \dagger}, Jack Radford², Davide Rovelli²,}

Kevin Mitchell², Ashley Lyons², Ilya Starshynov², Daniele Faccio^{2*}

¹ School of Computing Science, University of Glasgow, Glasgow G12 8QQ, UK

² School of Physics & Astronomy, University of Glasgow, Glasgow G12 8QQ, UK

[†] These authors contributed equally to this work

(Dated: March 15, 2021)

Abstract

This document contains supplementary material information for the manuscript 3D imaging from multipath temporal echoes. The document is organised as follows: in Section I we demonstrate through a simple mathematical model that it is possible to locate the position of an object from the multipath temporal echo. In Section II we describe the physical model that we use to validate our approach, while full details on the numerical simulations can be found in Section III. Section V is devoted to describe the metrics used to analyse our approach. In Section IV we describe the deep neural network used for image retrieval and give details of the training. Finally, Section VI provides full details on the calculations we used to estimate the information carried by multipath echoes.

^{*} Corresponding author: alex.turpin@glasgow.ac.uk / daniele.faccio@glasgow.uk

I. LOCALISATION WITH MULTIPATH TEMPORAL ECHOES

We discuss first the following question: is it possible to estimate the position of an object by just looking at the wave echoes recorded with a time-resolving single-point sensor? Using two or more sensors would make the problem trivial by simply using triangulation and direct reflections (first echo) from the object [1]. However, restricting the problem to a single detector paired with unstructured illumination makes the problem ill-posed. We will show that this problem can be solved provided that we have access not just to the first echo from the object but also to secondary echoes originating from multiple bounces between the object and the room walls. Let us consider the simplest possible scenario in 2D, as depicted in Fig. 1.

The emitter and the detector are both located at the origin of coordinates, for simplicity. The scene is flash-illuminated with pulsed waves emitted within an equilateral triangle (translucent red), while we assume a point detector receives any wave incident at any angle and retrieves time-of-flight (ToF) information with a temporal resolution given by the a temporal impulse response function (IRF) that we can also control. We consider a scene only consisting of a mirror-like wall (perfect reflection) and a point-like scatterer at position (x_0, y_0) (i.e. we restrict to the 2D case), acting as our object of interest. Also for simplicity, we assume that the distance from the wall to the detector x_w is known. The scatterer (depicted with a green circle) is isotropic, i.e. it reflects waves in all directions within a circle. Our aim is to determine the position of the scatterer (x_0, y_0) from the wave echo (temporal information) without ambiguity by using as many bounces as needed.

By bounces we mean reflection events which a pulse wave suffers upon propagation. For instance, first wave echoes bounce only once and return to the detector. This is the case of direct reflections from the scatterer, shown with green arrows in Fig. 1. Waves can also bounce first on the scatterer, then on the wall, and then return to the sensor; this would be a 2-path case, shown with blue arrows in Fig. 1. 3-path wave echoes correspond to waves that reflect first on the object, then on the wall, return to the object, and from there to the detector (shown with red arrows in Fig. 1). As we assume specular reflection from the wall, these three cases described above are geometrically the only possible ones (apart from repetitions of object-wall reflections) where the waves will be reflected back to the detector, resulting in a time histogram similar to the one in Fig. 1. Note that the peaks on the



FIG. 1. Locating an object with time data and multipath echoes. By having some information about the object's environment, for instance the arrangement of a wall with respect to the emitter/detector, one can identify the object's coordinates with a simple set of equations and measurements of the arrival time of the pulsed waves that follow paths with 1, 2, and 3 bounces (presented in the form of a time histogram here). See text for further details.

time histogram have decreased amplitude for an increasing number of reflections because of non-perfect reflectivity (R < 1), as well as some finite width determined by the IRF of the detection system. 1-path events are necessarily the ones with smaller time-of-flight and therefore they will correspond to the peak at t_0 . With simple geometry, we then obtain that:

$$2\sqrt{x_0^2 + y_0^2} = ct_0,\tag{1}$$

which indicates that there are an infinite number of positions placed on a circle (depicted with green dashes in Fig. 1) where the object can be. 2-path events, corresponding to t_1 on the time histogram, provide further insight on the objects' position:

$$\sqrt{x_0^2 + y_0^2} + \sqrt{(x_w - x_0)^2 + (y_1 - y_0)^2} + \sqrt{x_w^2 + y_1^2} = ct_1,$$
(2)

at the cost of introducing a new unknown, y_1 , that does not allow us to place the object on the circle described by Eq. (1). Looking at the information that the 3-path event (corresponding to t_2) provides, we obtain the following identity:

$$2\sqrt{x_0^2 + y_0^2} + 2|x_0 - x_w| = ct_2,$$
(3)

which can be used together with Eq. (1) to locate the object. Eq. (2), corresponds to a new circle (red dashes) intersecting with the point on the wall at (x_w, y_0) and compatible

with the distance $|x_0 - x_w|$ and the measurement t_2 . The origin of this circle coincides with the coordinates of the object. Therefore, by placing the origin of this new circle (or solving the system of equations given by Eqs. (1) and (3)) on the green circle described by Eq. (1) allows obtaining (x_0, y_0) , which demonstrates that the scatterer position can be unequivocally determined with a single detector by simply using multipath time echoes up to 3 bounces, under some assumptions.

II. PHYSICAL MODEL OF OUR NUMERICAL SIMULATIONS

We use numerical simulations based on a physics-inspired model to test the validity of this multipath echo-based imaging approach, see Fig. 2 and Supplementary Video 1 [2].



FIG. 2. Physical system for our model for the 3D numerical simulations. (a) 3D visualisation of our physical system: a rectangular cuboid (blue) moves within a room. Pulsed waves are emitted at a random angle within azimuth and elevation in [67.5°, 67.5°]. The red line shows one example of the multipath reflections of a pulse, which eventually reaches the detector (red plane) that records its arrival time. (b) Collecting a high number of pulses allows creating a time histogram. (c) Colour-depth encoded 3D view of the scene.

Our scene consists of a closed room, which is a hollow rectangular cuboid, i.e. with X, Y, Z dimensions that are not necessarily equal. The room has uniform walls, ergo they interact with the waves all in the same way. Inside this room, we place an object (a smaller rectangular cuboid) that can move freely via translation along the floor-plane. We also consider an emitter of pulsed waves and a planar time-resolving single-pixel detector. We

consider that the probe pulse is emitted in all directions within azimuth and elevation angles θ and ϕ , respectively. In other words, we assume that the probe pulse can be decomposed spatially as the sum of a high number of individual pulsed plane-waves with a k-vector angle within $[-\theta, \theta]$ and $[-\phi, \phi]$. Each of these pulsed waves travel within the scene at a fixed speed, c, and are reflected by the room walls and the object.

Eventually, the waves hit the single-pixel detector, which acts like a stop-watch and provides the arrival time of the pulses. The probe pulses are emitted at a specific rate, η , which determines when the single-pixel detector timer is reset to zero: only echo waves returned within a time $1/\eta$ contribute to the time histogram. Non-perfect reflectivity losses in walls and objects can be introduced via a reflectivity factor R that provides the probability of the wave to be reflected, such that after N bounces, the probability that the wave has not been absorbed is $(1 - R)^N$. Although reflectivity can also depend on the angle of incidence of the wave with respect to the surface, we have not included this in our simulations.

Therefore, in practice the reflectivity of the materials at each wave regime limits the maximum number of bounces that the waves suffer before being absorbed. We also consider the diffusion of waves when they hit the scene walls and objects via a specularity factor, s, that accounts for the scattering solid angle at which waves can be reflected. Mirror-like surfaces have a scattering angle of 0 rad (s = 1), while so-called Lambertian surfaces have a scattering angle of π rad (s = 0). Surfaces between these two extreme cases are considered as glossy. Optical waves have $s \ll 1$, while GHz electromagnetic (e.g. RADAR) and kHz acoustic waves (both mm-wave) have typically $s \leq 1$, especially for the cases considered here, where the probe pulse wavelengths are much smaller than the dimensions of the objects [3–5]. Therefore, our physical model only distinguishes between waves at different regimes (e.g. optical, radio, or acoustic) via the reflectivity and specularity factors.

III. SIMULATIONS

Here we describe the numerical simulations and results we obtained based on the physical model, for a room with dimensions $4 \text{ m} \times 7 \text{ m} \times 7 \text{ m}$, with smooth, mirror-like walls of reflectivity R = 1 and specularity s = 1. In the room, we place a point-source emitter of probe pulses placed at (0.5, -1, 0.5) emitting waves within azimuthal and elevation angles ϕ and θ , both within $[-67.5^{\circ}, 67.5^{\circ}]$, at a pulse repetition rate $\eta = 10 \text{ MHz}$. Next to the

emitter, we placed a flat detector with dimensions $1 \text{ m} \times 1 \text{ m}$ in the YZ plane. In this room, we place a rectangular cuboid object with dimensions $1 \text{ m} \times 1 \text{ m} \times 5 \text{ m}$, with reflectivity R = 1 and specularity s = 1, which moves in the room by being translated in the XY plane such that its bottom is at Z = 0.

We control the number of allowed reflections per emitted pulse, which allows us to study the quality of the retrieved images for some set number of reflections. Every time histogram is obtained by emitting 10,000 pulsed plane waves per object position, following the physical model described above, and measuring the time-of-flight of the waves that return to the sensor. In order to speed up our simulations, we cropped out any pulses whose round trip time of flight would have been longer than the time between 2 pulses, or 100 ns - in practice this meant ignoring some portion of waves that reflected 8 or more times, as these histograms natively contained photons with > 100 ns arrival times.

To generate a simulated 3D image, we scanned the room from the camera position [0.5, -1, 0.5] over azimuth θ within $[-60^{\circ}, 100^{\circ}]$ and elevation ϕ within $[-80^{\circ}, 80^{\circ}]$. The asymmetry in the azimuthal angle was designed to rotate the field of view towards the bulk of the room, away from the leftside wall, for a better perspective.

IV. IMAGE RETRIEVAL ALGORITHM

In order to solve the inverse problem of providing an estimate of the scene from the time histogram generated by the waves echoes, we use a deep neural network algorithm.

First, we assemble a database of temporal histogram - depth image pairs, which form the inputs and labels of our supervised training scheme, using the numerical simulation described above. We created 2000 data pairs for training, and 100 for testing, per maximum number of scattering events. As we simulated single- to 10-path event scenarios, this gave a total of 20,000 and 1000 input-label pairs respectively. We set aside our testing data for later, to evaluate the neural networks.

In the second phase, the training data pairs are used to train the network. For each of the maximum-scattering-event-number scenarios, we trained separate networks on the corresponding data pairs, albeit with the same architecture. The neural network architecture, sketched in Fig. 3, is fully convolutional. It has an hourglass shape such as to force information through a bottleneck, which promotes the network to compress the input into some compact representation. For our downsampling blocks (DB), we have a series of convolutional layers with kernel size 7, strides = 2 and each convolutional layer is followed by a Rectified Linear Unit (ReLU) activation function. Between the bottleneck and the output layer, there are a series of up-sampling blocks (UB), which consist of 2-D up-sampling layers, 2-D convolutional layers with kernel size 5×5 and strides = (1,1), and a ReLU activation function.



FIG. 3. Deep neural network layout and operation. The input time histogram is passed through a series of downsampling blocks (DB), each of them consisting of a convolutional layer followed by a Rectified Linear Unit (ReLU) activation function. After 4 DB, we use a series of up-sampling blocks (UB), which consist of 2-D up-sampling layers, 2-D convolutional layers, and a ReLU activation function, being a depth-in-color encoded 3D image the output of the network.

We trained our neural networks using conventional machine learning methods, namely loss back-propagation via mini-batch gradient descent, implemented on a batch size of 100 using adaptive moment estimation (Adam [6]). Our loss function was pixelwise mean-squarederror (MSE) - see Eq.4. In order to prevent overtraining, we first validated the number of epochs for which to train the neural networks on 200 histogram-image pairs. In this way, we ascertained that the ideal number of training epochs increased as the maximum number of scattering events, from 110 epochs for single scattering event data to 350 for up-to-10 scattering event data. Simultaneously, the total training time of a single network increased, from 40 to 130 seconds respectively.

As stated in the main paper, for the RF and acoustic experiments we trained our neural networks on 9000 and 5000 temporal histogram-3D image pairs, and tested them on 865 and 500 respectively. The corresponding Supplementary Videos for our experimental results can be found in Ref. [2]. For these datasets, we implemented slightly different architectures compared to our simulated neural networks: we maintained the same form of up- and downsampling blocks, but each block had a different in- and output shape, and a different number of features. For the RF neural network, the input size (i.e. the number of bins in the temporal histogram) was 256, and the output size was 60×80 while the acoustic neural network was fed with inputs of size 9600, and 3D images of 64×64 pixels. For both the RF and acoustic neural networks, the number of features started at 64, increasing to 256 by the bottleneck, and then kept at 256 until the final layer. The increase in the number of convolutional features was chosen because the experimental scenes had a lot more variation than the simulated scenes (variety of objects, of various shapes and a range of reflectivities and specularities, uneven walls, etc.), and the corresponding histograms and images consequently showed more variability. Therefore the neural networks were designed to be able to correlate a greater number of input and output features. Otherwise, our training procedure was kept the same as for the simulated data, i.e. we used the same batch-size, loss, gradient descent optimizer, hardware, software, etc.

In the final phase of our approach and only after the deep learning algorithms are trained, we fed the latter with a single time histogram with the wave echo recording from the testing dataset, which provided a 3D image estimate of the scene.

In Fig. 4 we visualise the influence of the number of path events in our temporal echoes on the image retrieval, for one particular case from our test dataset (ground truth presented on the top image). As can be seen from the first and third columns, allowing waves to bounce more than once populates the time histogram in the horizontal axis, which increases its information content. This has a direct impact on the image reconstruction: the more bounces are considered, the closer the retrieved and ground truth images are. See Supplementary Video 2 [2] for reconstructions from our full test dataset. In general, it can be appreciated that 1-path events lead to ghosting in the reconstructed image, because of the





FIG. 4. Image reconstruction with different multipath events. We show the time histograms obtained for 1-... 10-path events and corresponding depth-images obtained with the image retrieval algorithm. The ground truth depth-image is shown on the top.

degeneracy problem outlined in Eq. (1). Adding more path events clearly allows a better image reconstruction of the scene, especially when the object is placed at the sides of the image.

Training was performed in Python 3.7.9 using TensorFlow 2.1.0 [7] and Keras [8], on an

NVIDIA GeForce RTX 2080 Ti GPU.

A. Performance on unseen individuals

It is an interesting question to address whether our technique can be trained just with one individual and tested on different individuals. To answer this question, we conducted an experimental test where we trained the image retrieval algorithm with data gathered using one individual, and then operated the algorithm with data from a different individual. Our results show that the algorithm is able to provide the general shape and depth of the individual (see Fig. 5). This allows us to successfully operate our technique on different individuals, while training only on one.

V. PERFORMANCE METRICS

Our reconstruction-quality metrics of choice are the pixelwise mean-squared-error (MSE) between reconstruction and ground truth, and the intersection over union (IOU) of the foregrounds of the reconstruction and ground truth. Formally, for reconstruction image R and ground truth image G, both of dimensions $M \times N$, containing pixel values R_{ij} and G_{ij} respectively, we can write MSE as:

$$MSE = \frac{1}{N \times M} \sum_{i=1}^{M} \sum_{j=1}^{N} (R_{ij} - G_{ij})^2$$
(4)

Having a low mean squared error means that on average, the depth value predicted at a random pixel value matches the true depth value well.

Intersection over union is a less common metric, which focuses on how well the shapes of the foreground objects are reconstructed, and is invariant to changes in the relative size of the foreground object within the field of view. To calculate IOU, we take a background mask, and compare our ground truth and prediction images to this mask; with this comparison, we binarise the ground truth and prediction images into foreground (variable) and background (static) pixels. This is shown in Fig. 6, in the bottom-right in greyscale. Their intersection, then, is simply the area of overlap, while the union is the combined area of all non-background regions.



FIG. 5. Performance of the technique with unseen individuals. First column shows the recorded time trace, second column shows the reconstructed image provided by the deep neural network (DNN), while the last column is the corresponding ground truth recorded with a ToF camera.

We can write this mathematically as follows: for background mask pixels M_{ij} , we binarise our reconstruction pixels R_{ij} and ground truth pixels G_{ij} to obtain \tilde{R}_{ij} and \tilde{G}_{ij} , where *i* and *j* are in used the same notation as in Eq. (4), as such:

$$\tilde{R}_{ij} = \begin{cases}
1 & (M_{ij} - R_{ij}) \ge \kappa r \\
0 & (M_{ij} - R_{ij}) < \kappa r
\end{cases}$$

$$\tilde{G}_{ij} = \begin{cases}
1 & (M_{ij} - G_{ij}) \ge \kappa g \\
0 & (M_{ij} - G_{ij}) < \kappa g,
\end{cases}$$
(5)

where r and g are the maximum pixel-wise differences between the mask and the reconstruction, and the mask and ground truth, respectively. The threshold κ was set at 0.5, 0.2 and 0.2 for the synthetic, RF and acoustic data respectively, as these values were found to consistently give visually good binarisation. Then, IOU is:

$$IOU = \frac{\sum_{i=1}^{80} \sum_{j=1}^{80} (\tilde{R}_{ij} \tilde{G}_{ij})}{\sum_{i=1}^{80} \sum_{j=1}^{80} \max\{\tilde{R}_{ij}, \tilde{G}_{ij}\}}.$$
(6)



FIG. 6. Visualisation of the meaning of the metric intersection over union. In the top left, we have a poor reconstruction, in the top right, its corresponding ground truth. We compare R and R to the mask shown in the bottom left, to identify the foreground regions. These regions, denoted as $\tilde{R} = 1$ and $\tilde{G} = 1$, are shown in light and dark grey respectively. Then, their intersection is the area of the region shown in black on the bottom right, and the union is the combined area of everything that is not white. Finally, the IOU is found according to Eq. (6).

A large intersection over union means the neural network reconstructs the foreground object largely at the same position as where it was in the ground truth image.

In Fig. 7 we compare the evolution of MSE and IOU as a function of the maximum number of scattering events, for our 100 testing 3D image-neural net reconstruction pairs. As stated in the main text, we trained 10 neural networks for each of the maximum scattering event scenarios, and averaged over the MSE and IOU obtained from the predictions of these 10 ANN copies on the test set. This was done to minimise the specificity of our predictions, and correspondingly, our performance metrics, on the starting configurations of our ANNs. Clearly, adding more bounces enhances the ability of the algorithm to retrieve images, which demonstrates the advantage of using waves that are not absorbed after the first bounce for such an imaging scheme. MSE, the loss metric of our neural network, improves up until 8 scattering events, and IOU up until 10. Neither MSE nor IOU improve homogeneously, however the local minima/maxima are probably attributed to the inherent randomness of our simulator as opposed to actual local maxima in the amount of information in the histograms.



FIG. 7. Mean mean-squared-error (MSE) and intersection over union (IOU) obtained from our 10 model replicas, each averaged over 100 histogram-3D image pairs, with increasing number of maximum scattering events.

VI. INFORMATION THEORY ANALYSIS

To quantify the gain in information when including an increasing number of bounces, we use the principle of Shannon entropy, joint entropy and mutual information derived from information theory [9–11]. Information is formally defined as the number of bits required to describe a reduction in uncertainty when observing a random variable X on a set \mathcal{X} at a value $x \in \mathcal{X}$ that occurs with probability p(x). This can be calculated by $-\log_2 p(x)$ and is commonly referred to as self-information or Shannon information. Furthermore, the expectation value of uncertainty reduction when observing variable X is known as the Shannon entropy:

$$H(X) = -\sum_{i=1}^{N} p(x_i) \log_2 p(x_i).$$
(7)

We first quantify the information of temporal data containing a signal due to a single bounce. In this context, we define the time of arrival of a photon as a random variable and neglect the variable number of counts at a given time-bin to simplify the calculation. Given 2000 examples of single bounce histograms from the stochastic model described above, we find identical temporal traces within the data; let us call the set of all unique traces $\tilde{\mathcal{X}}$. We then re-assign these traces to a new variable \tilde{X} , which can assume an integer label $\tilde{x} \in \tilde{\mathcal{X}}$. After normalising the distribution of \tilde{X} to a probability mass function, we then calculate entropy $H(\tilde{X})$ by Eq. (7). To assess what additional information is gained by including photons from a second bounce in the temporal-trace, we use the joint entropy $H(\tilde{X}, \tilde{Y})$ of the distribution of identical temporal-data for a single bounce \tilde{X} , and the equivalent distribution for data containing up to two bounces \tilde{Y} . The joint entropy is simply the Shannon entropy for a joint distribution of \tilde{X} and \tilde{Y} on \tilde{X} and \tilde{Y} respectively, and describes the expected uncertainty reduction when observing (\tilde{x}, \tilde{y}) where $\tilde{x} \in \tilde{\mathcal{X}}$ and $\tilde{y} \in \tilde{\mathcal{Y}}$. For context, the uncertainty reduction of observing a given temporal trace containing photons from one bounce can be increased when also considering the data from the second bounce. We find the joint probability distribution by the relation $p(\tilde{x}, \tilde{y}) = p(\tilde{y}|\tilde{x})p(\tilde{x})$ and use this to calculate joint entropy by:

$$H(\tilde{X}, \tilde{Y}) = -\sum_{j=1}^{M} \sum_{i=1}^{N} p(\tilde{x}_i, \tilde{y}_j) \log_2 p(\tilde{x}_i, \tilde{y}_j).$$
(8)

We then repeat the Joint Entropy calculation to compare the data containing $\langle n \rangle$ bounces and $\langle (n + 1) \rangle$ bounces, and present the gain in uncorrelated information when including photons from additional bounces using the definition of mutual information $MI(\tilde{X}; \tilde{Y})$ which describes the information shared by two random variables due to correlation:

$$MI(\tilde{X};\tilde{Y}) = H(\tilde{X}) + H(\tilde{Y}) - H(\tilde{X},\tilde{Y})$$
(9)

We rearrange Eq. (9) to find the additional uncorrelated information in \tilde{Y} , i.e. the mutual information $MI(\tilde{X};\tilde{Y})$ subtracted from the total information in \tilde{Y} or equivalently, $H(\tilde{X},\tilde{Y}) - H(\tilde{X})$. This gain in *uncorrelated* information $UI(\tilde{X};\tilde{Y})$ increases as photons which have experienced an increasing number of bounces are included in the data as shown in Fig. 2(a) of the main document. Clearly, there is a dramatic increase in information when including photons which have experienced two bounces compared with only a single bounce, however photons experiencing five bounces or more show no increase in information content for this data set. It is expected that any information about the 3D scene given by the knowledge of the number of counts at a given time-bin may further contribute the reconstructed image quality. The neural network algorithm used to reconstruct the 3D images can leverage both the time of arrival and the number of counts which may account for improvement of reconstructions when given data with photons experiencing five or more photons. If we consider the 3D image to be an input from an information theory point of view and the histogram to be the output, we hypothesise that adding more bounces beyond the actual Shannon entropy increases the redundancy of the transmitted data i.e. is similar to a better encoding scheme. This reduces the probability of a bit error when propagating down a noisy channel, in this case the stochastic sampling and discretisation performed at the radar/microphone. This hypothesis can be illustrated with the mirror room image from Fig. 2(b) in the main document: adding more bounces creates replicas of existing mirror images. These replicas do not add more information in a lossless/noiseless environment, but when the environment is sampled in a noisy way, the added redundancy can help in localising and identifying the object.

 S. Chan, R. E. Warburton, G. Gariepy, J. Leach, and D. Faccio, Opt. Express 25, 10109 (2017).

[2] "Supplementary Videos 1, 2, 3, and 4 can be found, respectively, in the following links:," https://www.youtube.com/watch?v=cjPkSo9kZuI&ab_channel= ExtremeLightGroupUniversityofGlasgow https://www.youtube.com/watch?v=e4Ywbveb6sw&ab_channel= ExtremeLightGroupUniversityofGlasgow https://www.youtube.com/watch?v=mD-Qqu6aCPw&ab_channel= ExtremeLightGroupUniversityofGlasgow https://www.youtube.com/watch?v=Ws_4WIkWuAw&ab_channel= ExtremeLightGroupUniversityofGlasgow.

- [3] F. Fuschini, S. Häfner, M. Zoli, R. Müller, E. Vitucci, D. Dupleich, M. Barbiroli, J. Luo,
 E. Schulz, V. Degli-Esposti, *et al.*, Journal of Infrared, Millimeter, and Terahertz Waves 38, 727 (2017).
- [4] A. A. Goulianos, A. L. Freire, T. Barratt, E. Mellios, P. Cain, M. Rumney, A. Nix, and M. Beach, in 2017 IEEE 86th Vehicular Technology Conference (VTC-Fall) (IEEE, 2017) pp.

1-5.

- [5] W. G. Neubauer, Acoustic reflection from surfaces and shapes (Naval Research Laboratory, 1986).
- [6] D. P. Kingma and J. Ba, arXiv preprint arXiv:1412.6980 (2014).
- [7] M. Abadi *et al*, "TensorFlow: Large-scale machine learning on heterogeneous systems," (2015).
- [8] F. Chollet, "keras," https://github.com/fchollet/keras (2015).
- [9] C. E. Shannon, The Bell system technical journal 27, 379 (1948).
- [10] D. J. C. MacKay, Information theory, inference, and learning algorithms (Cambridge University Press, Cambridge, 2003).
- [11] T. M. Cover and J. A. Thomas, *Elements of information theory* (Wiley-Interscience, 2006).