Pretty Point Clouds

Vilma Ylvén Oscar Montelin

Lund University | Faculty of Engineering

Have you ever been a security guard and looked at some security footage with the burning question of "How tall are those people anyway, and how close to that fence is that car actually?" *Probably not, but if you have, the solution is using a LiDAR, which is a sort of laser camera, able to measure depth and distances. However, the issue with a LiDAR is that it outputs a set of coordinates that are most easily visualised as a 3D cluster of colourless points, also known as a point cloud. And if you've ever seen such a point cloud, you know that you now will have more questions than before. These point clouds are inherently unclear to the human eye, which is terrible when you want to start integrating LiDARs into consumer technology, and so, we decided to see what we could do about it. With one LiDAR, one normal camera, some surface reconstruction methods, and a neural network we managed to make a 3D environment that you will want to live in!*

Introduction

A LiDAR is more or less a high frequency radar. Instead of sending out microwaves it sends out laser beams. By measuring the time it takes for the beams to return, it can measure distances in a space with a very high accuracy and produce a 3D environment made up of points - a point cloud.

The point clouds have no colour and can at times be rather sparse, which can make it hard to see what is going on for the human eye. In our thesis we have explored ways of helping humans with this by also using information from a normal video camera. In Figure 1 we can see what the raw data from the camera and the LiDAR looks like.

Figure 1. Camera video (left) and LiDAR point cloud (right) captured for a roundabout.

LiDAR technology is only in it's toddler years when it comes to use cases, and better ways to visualise point clouds would open up many doors for the future. Hopefully LiDAR and normal security cameras can work side by side in the future. A security guard can get much more information of they know exactly where things are. LiDAR can provide that solution, but if the guard cannot see what is happening in the point cloud, it is of extremely little use.

Adding Colour

The easiest way of adding colour to the points is to give each point a colour value based on the corresponding point's colour in the camera image. In Figure 2 we have done just that, instantly making the scene a bit clearer! Perhaps you can now see what parts are the buildings in the back of the scene and what points make up the centre of the roundabout?

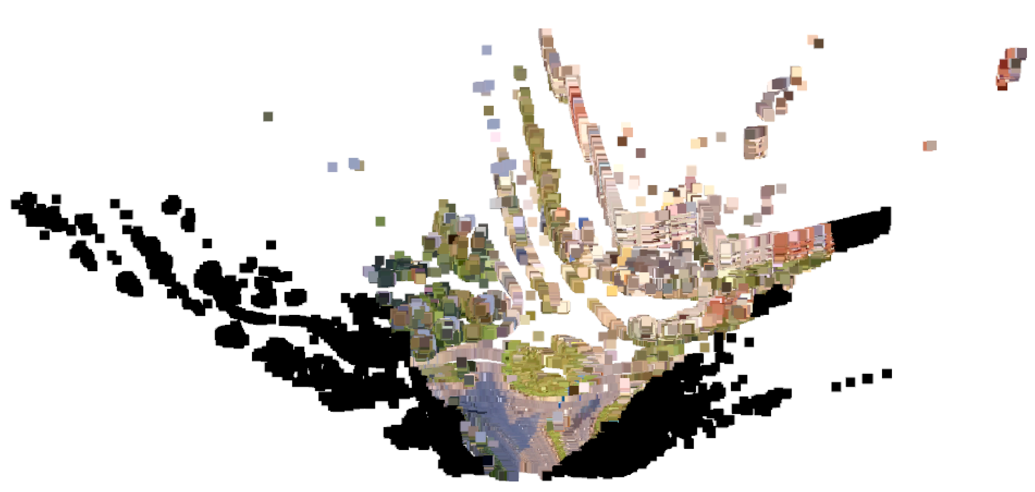

Figure 2. Point cloud with colour added. The black point had no corresponding point in the camera image.

Accumulating Point Clouds Over Time

Objects like cars don't change shape over time, they only move in a space, which was an opportunity too excellent not to make use of! We thought this was a perfect way to get as much information from a recording as possible by using every instance of the same object.

We achieved this by processing the point clouds from the LiDAR by using algorithms to detect movement and to decide which of the moving points belonged to the same objects. With this we could then track each separate object throughout an entire recording.

The gist of it is that we cut out an object from every frame and then pasted all of these cutouts on top of each other and align them, creating one big point cloud for every moving object. In Figure 3 you can see some of these objects and what they looked like in the video. Much prettier than the unclear blob in Figure 1!

Figure 3. Coloured point clouds, accumulated over time, for three different vehicles and their appearance in the video.

These object could also later be placed back into a background model created by adding together a bunch of LiDAR frames and removing the points that correspond to the moving objects.

Training a Neural Network

NeRFs, or Neural Radiance Fields [1], are a neural network that can be trained on a sequence of images of an object from different angles to represent the object in 3D. Point-Based Neural Radiance Fields (Point-NeRF) [2] accomplishes this by storing the features, such as colour and reflectance, in discrete points. This idea matched so well up with the point clouds we have constructed that we wanted to try to apply it! Hopefully, it was going to result in even prettier and more accurate 3D models of our objects.

We gave the network a colourless accumulated point cloud and trained it on the corresponding video sequence. Essentially we let the network "learn" what colour each point should be from every angle based on what colour they are from the angles that we have seen and are able to provide to it. This is done to see how well the neural network could represent the object in comparison to just having the colouring be done as in Figure 1. In Figure 4 we can see that comparison made on a point cloud of a car.

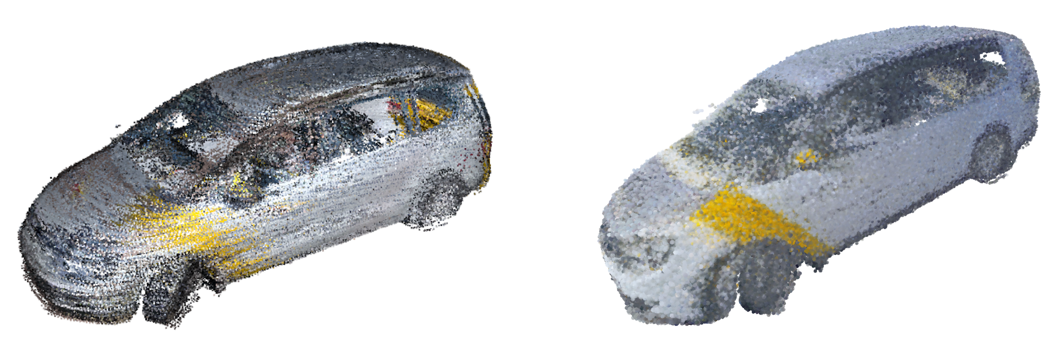

Figure 4. A coloured point cloud, accumulated over time (left) and the result of training a Point-NeRF network on the colourless accumulated point cloud and the video sequence (right).

The results were, in fact, very pretty, but they also took a long time to generate. While the entire process up to getting the coloured accumulated point cloud object took mere minutes, training a neural network took hours to days. Whether or not this is a cost worth paying is of course up to the specific use case. Sometimes, maybe prioritising good enough and fast is important, and other times presenting a dazzling high quality 3D model might be the ultimate goal. We leave that up to the future!

Creating Surfaces

To avoid the lack of clarity of discrete points we also tried to recreate the actual surfaces in the scene based on the points. We found that this was most important to do for the background, as our brains really like to see the ground as a surface. In Figure 5 we have done just that by using a couple of adjacent points to estimate the direction of a potential surface in each point, and then letting the computer find a surface that fits all of these the best.

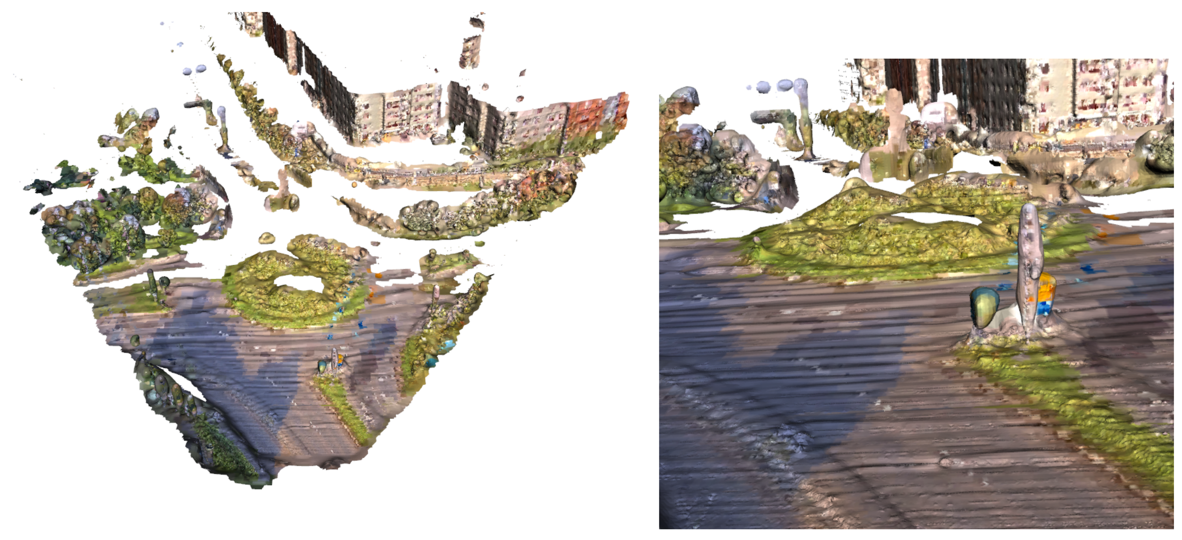

Figure 5. A version of Figure 2 where the background surface has been recreated.

We did the same thing for a couple of cars, but we found that they didn't benefit nearly as much from it and that it sometimes even made the visual result worse depending on the point cloud. Thus, we stuck with our pretty pretty point cloud cars!

A Complete Scene

Putting a surface model of a background together with a coloured accumulated point cloud model of a moving car results in Figure 6!

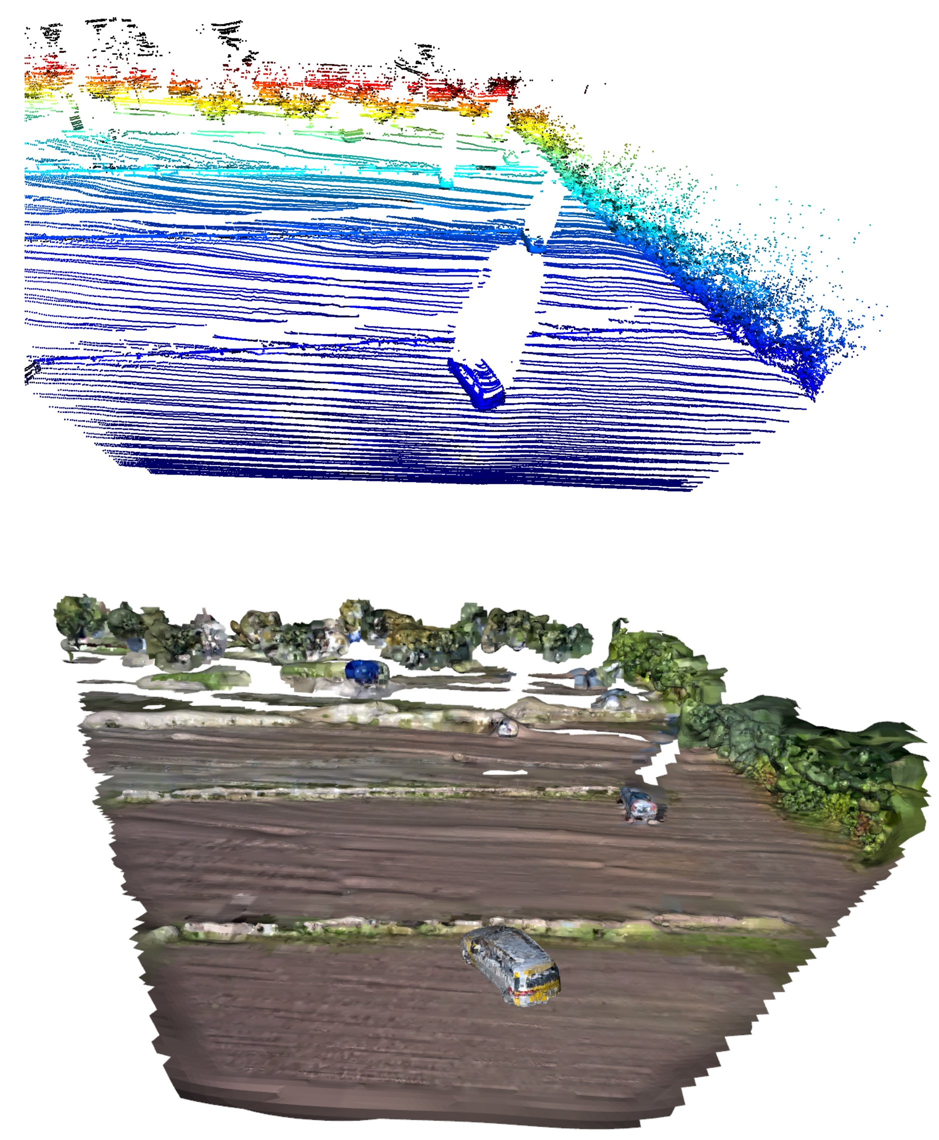

Figure 6. Raw LiDAR data for the scene (above) and the reconstruction with a surface background and an accumulated point cloud car (below).

The Future

The most exciting part of this thesis has been looking towards the future. We think that versions of the work we have done really could be helpful in getting LiDAR technology off the ground.

Perhaps, when LiDARs aren't so ridiculously expensive, someone can expand this technology to fuse the data from multiple LiDARs and cameras. Just putting one on either side of a road would, for example, yield accurate full 3D models for every car passing by!

Maybe someone will develop a neural network especially constructed for LiDAR data, and perhaps it will be fast enough to run in real time. The possibilities are endless and we can't wait to see what happens within the field in the future!

References

[1] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng.
Nerf: Representing scenes as neural radiance fields for view synthesis.
In *ECCV*, 2020.

[2] Qiangeng Xu, Zexiang Xu, Julien Philip, Sai Bi, Zhixin Shu, Kalyan Sunkavalli, and Ulrich Neumann.
Point-nerf: Point-based neural radiance fields.
In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5438–5448, 2022.