A USDOT NATIONAL UNIVERSITY TRANSPORTATION CENTER

Low-Cost 3D model acquisition for rapid accident investigation

Christoph Mertz (PI, ORCID: 0000-0001-7540-5211)
Tejas Khot, Wentao Yuan

FINAL RESEARCH REPORT

Contract # 69A3551747111

69A3551747111

DISCLAIMER

69A3551747111

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation's University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

69A3551747111

Introduction

Vehicular accidents are one of the great societal challenges. In some demographics, they are the leading cause of death. It is important that accidents are thoroughly investigated. The immediate reason is to find out who is responsible and liable for the damages. But just as important is it to collect information about accidents to determine if changes to vehicles, infrastructure or policy could prevent or mitigate future accidents. One set of tools in the investigation uses 3D models of the scene. They are usually acquired with laser scanners. However, they are very costly in time and money and are therefore only used in severe cases, like fatal accidents. In recent years digital cameras and computer vision algorithms have become so inexpensive, powerful and efficient that it is now possible to create 3D models from a set of digital images at a very low cost. In previous research [1] we have shown that one can do this for vehicle accidents and that these 3D models are useful for accident investigation. In this project we worked on shape completing, parts segmentation, unsupervised multi-view stereopsis, and 3D transformer networks. The detailed description and evaluation of these can be found in the publications [2,3,4]. The long-term goal is to have a set of tools that can automatically analyze an accident from a set of images. In this report we describe two steps towards that goal: Point cloud completion and parts separation (Figure 1).

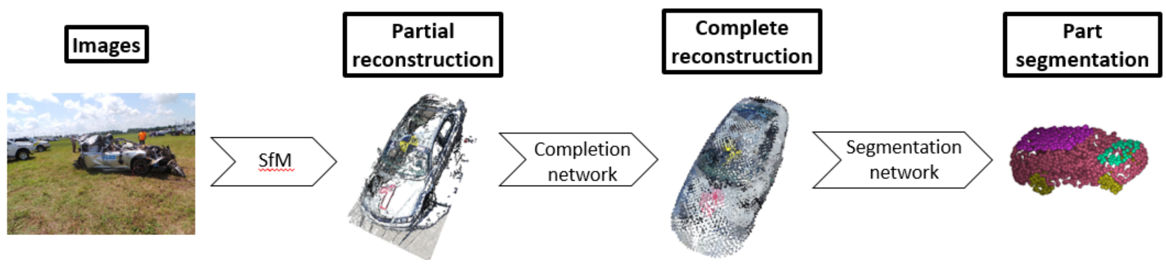

Figure 1 Final analysis pipeline: Starting with the raw images SfM is used for partial reconstruction. A completion network follows to get to the complete reconstruction. Finally, a segmentation network clusters the points into different parts of the vehicle.

Reconstruction and analysis pipeline

The first step in the pipeline is the **3D reconstruction** of the accident scene. On our website at [1] one can see many examples of 3D models of accidents. We are mostly using Colmap [5] to do the reconstruction because of its quality and ease of use. For a given scene (e.g. Figure 2 left) we take images all around the vehicle at different heights and angles. Colmap uses these images to create a sparse reconstruction and find the 6D location of the camera positions. Figure 2 (left middle) shows the camera positions in red and the sparse reconstruction as a point cloud. Next Colmap uses the know camera positions to do stereo vision on each image. The sum of all the stereo images becomes the dense reconstruction (Figure 2 right middle). Finally, one can build a mesh of the scene, also called a solid model (Figure 2 right). The dense point cloud is the input to the next steps of the analysis pipeline. We did some investigation to see if we can improve the dense reconstruction by using robust photometric consistency [4].

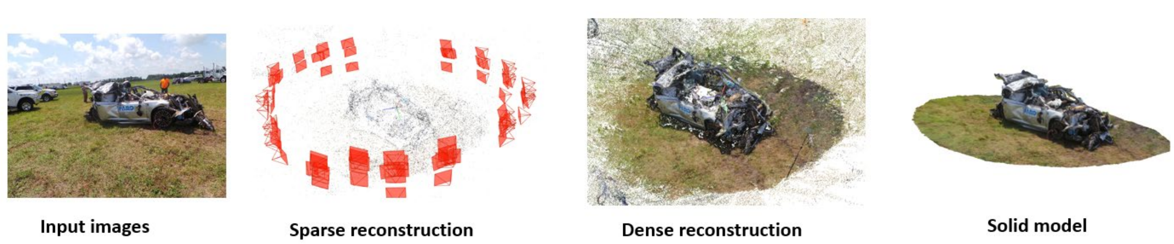

Figure 2 3D model reconstruction: Structure from motion uses the raw input images to create a sparse point cloud of the scene. It simultaneously estimates where the images were taken (indicated by red symbols). A stereo algorithm produces the dense reconstruction. Finally, the point cloud can be converted into a solid mesh.

Real-world point clouds are often incomplete. This might be because of occlusion, featureless surfaces, glare, insufficient number of images, people moving in the scene during data taking, or other errors. To better analyze the point cloud of an accident we want to be able to **complete the point cloud**. For that, we developed a point completion network [2]. It consists of an encoder and decoder, where the decoder first generates a coarse output and then a detailed output (Figure 3). To train our network, we use synthetic CAD models from ShapeNet [6] to create a large-scale dataset containing pairs of partial and complete point clouds. Specifically, we took 30974 models from 8 categories: airplane, cabinet, car, chair, lamp, sofa, table, vessel.

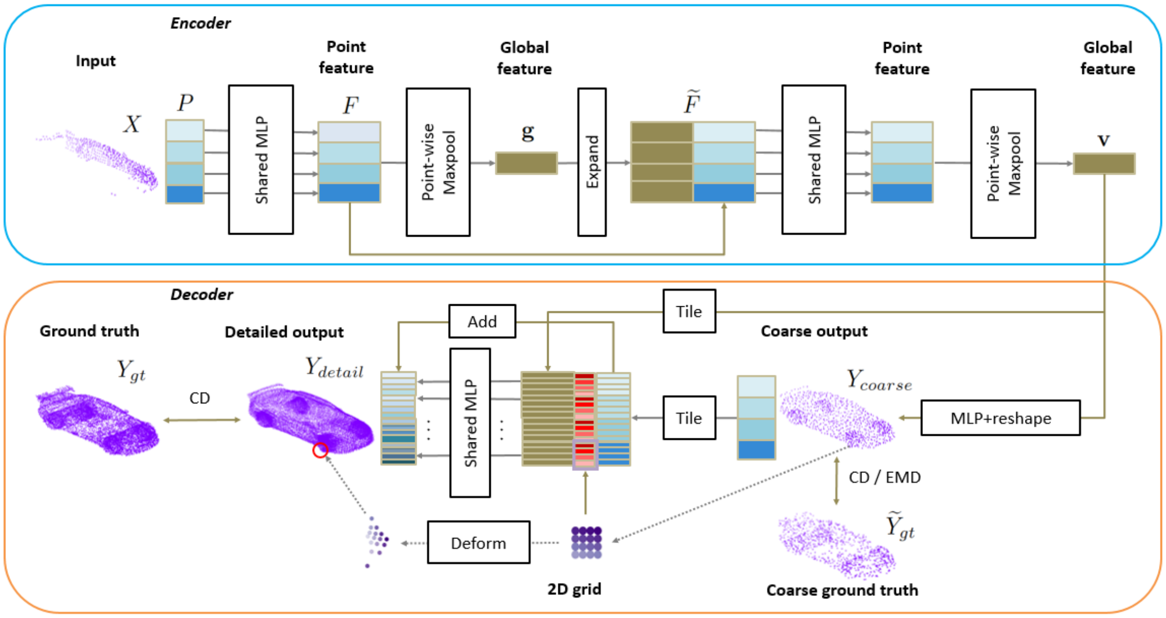

Figure 3 Completion Network Architecture: The encoder abstracts the input point cloud X as a feature vector v*. The decoder uses v to first generate a coarse output Y_{coarse} followed by a detailed output Y_{detail}. Each colored rectangle denotes a matrix row. Same color indicates same content.*

The result of shape completion on a car can be seen in Figure 4. Large parts of the original hood and roof are missing because these are areas of low texture. The completed model has points uniformly over all areas of the car.

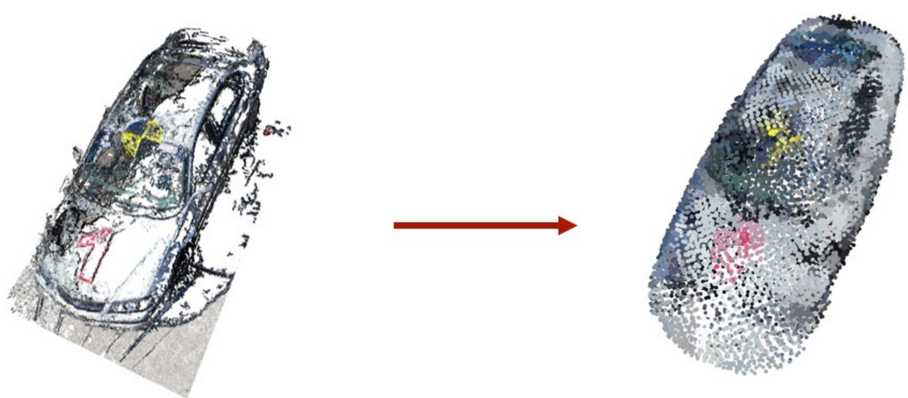

Figure 4 Shape completion example: Left is a dense reconstruction of a vehicle. One can see large parts of the hood and roof missing. On the right is the completed model, the points cover uniformly all areas of the vehicle.

The next step in the analysis pipeline is to **segment the vehicle into different parts**. The input point cloud has an arbitrary orientation. We align it to a common orientation by using an iterative transformer network (Figure 5). The segmentation network DGCNN [7] is then applied to segment the object into its parts. The network was trained with a dataset containing 16,881 shapes from 16 categories, annotated with 50 parts in total and 2-6 parts per category.
Figure 6 shows the successful application of this segmentation on four different objects: motorcycle, bicycle, pickup truck and limousine.

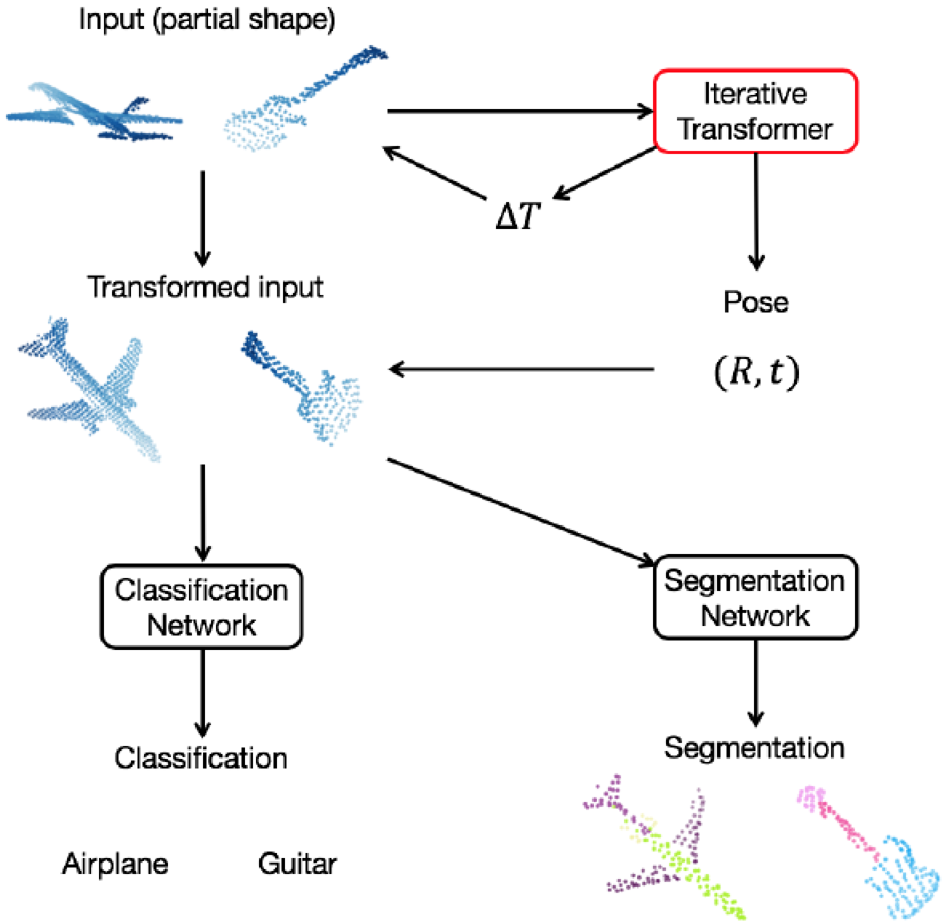

Figure 5 Iterative Transformer Network (IT-Net) predicts rigid transformations from point clouds in an iterative fashion. It can be used independently as a pose estimator or jointly with classification and segmentation networks.

Figure 6 Successful segmentation of point clouds. The motorcycle and bicycle are segmented into frame, saddle, and tire, the cars are segmented into wheels, roof, hood and body.

Conclusion and outlook

With deep learning, point cloud analysis has made great progress. In this project, we showed that they are now able to complete point clouds of objects that were only partially observed or which have missing parts because of glare or surfaces with little texture. In the next step of the analysis the point cloud of the object can be segmented into different parts.

There are several directions to be pursued and explored in future research. For one, the above networks should be trained with accident data. This is in principal straightforward, but in practice it will involve a lot of work. Point cloud analysis in general is a very active research field. Further progress is expected in the near future and these methods and algorithms should be modified and tuned for the specific purpose of accident investigation.

References

[1] http://www.cs.cmu.edu/~reconstruction/
[2] Yuan, W., Khot, T., Held, D., Mertz, C. and Hebert, M., 2018, September. Pcn: Point completion network. In 2018 International Conference on 3D Vision (3DV) (pp. 728-737). IEEE.
[3] Yuan, W., Held, D., Mertz, C. and Hebert, M., 2018. Iterative Transformer Network for 3D Point Cloud. arXiv preprint arXiv:1811.11209.
[4] Khot, T., Agrawal, S., Tulsiani, S., Mertz, C., Lucey, S. and Hebert, M., 2019. Learning Unsupervised Multi-View Stereopsis via Robust Photometric Consistency. arXiv preprint arXiv:1905.02706.
[5] https://colmap.github.io/
[6] Chang, A.X., Funkhouser, T., Guibas, L., Hanrahan, P., Huang, Q., Li, Z., Savarese, S., Savva, M., Song, S., Su, H. and Xiao, J., 2015. Shapenet: An information-rich 3d model repository. arXiv preprint arXiv:1512.03012.
[7] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, and J. M. Solomon. Dynamic graph cnn for learning on point clouds. arXiv preprint arXiv:1801.07829, 2018