**PROJECT TITLE:** Automated Reconstruction of Three-Dimensional Sparse Pointe Cloud for Utility Infrastructure based on Convolutional Neural Network (CNN) and Structure from Motion (SfM)

**ABSTRACT:** 3D modeling of the existing structure can be a game-changer to save time and resources in the built infrastructure and facility management domain. Decision-makers can strategically utilize developed 3D models when inspecting, rehabilitating, and upgrading built infrastructure is required. This study develops an automated process, which consists of two steps: (1) extracting optimized sequence image segments of utility infrastructure from video data, and (2) performing information fusion for 3D sparse point clouds of that can be foundation for 3D dense point clouds and textured models of utility infrastructure. To extract image segments, the PIs will (1) detect targeted utility infrastructure based on Convolutional Neural Network (CNN) by developing rich training data set using monocular dashboard cameras, (2) develop an optimized crop function only for targeted utility infrastructure with Global Positioning System (GPS) data, and (3) convert video data to sequenced images by optimizing overlapping percentage and number of frames. In order to integrate information for 3D sparse point clouds, this study will (1) detect and match feature points of the optimized sequence image segments of utility infrastructure based on Scale-Invariant Feature Transform (SIFT), (2) solve fundamental matrices and bundle adjustment, and (3) validate quantitative and qualitative results by comparing with constructed 3D sparse points from surrounding video data of the utility infrastructure. The outcomes of this research lead toward the attainment of my long-term research goal, which automates the construction process for digital twin cities by creating effective data managing and processing algorithms and systems of any required built infrastructure.

**BACKGROUND AND SIGNIFICANCE:**

Digital twin, which is a digital replica of a living or non-living physical entity in the built environment, can play a critical role in saving time and resources for maintaining existing infrastructure [01]. Singapore has developed and experimented with the entire city’s digital twin to improve city life and sustainability [02]. One of the digital twin core steps for the built environment is 3D modeling and digital mapping of existing infrastructure. The biggest challenge associated with the digital twin era is that there is the large number of built infrastructure systems that need to be identified, inspected, and digitally mapped. It is not possible to manually perform these steps. For instance, there are 130 million electricity poles around the United States [03]. There remains, therefore, a ***critical need*** to develop automated processing methods to identify correctly, rigorously inspect, properly map, and effectively store a digital replica of existing infrastructure toward the digital twin era. In the absence of such methods, the promise of restoring and upgrading of existing infrastructure through digital twin will likely remain problematic.

My ***long-term goal*** of this research realm is to automate the process of mapping, labeling, and constructing digital twin of cities and counties in the United States by creating effective data managing and processing algorithms and systems of any necessary built infrastructure. The ***objective*** of this particular application, which is the next step toward the attainment of my long-term goal, is to construct 3D sparse point clouds of utility infrastructure using monocular cameras. The ***rationale*** for this research is that (1) computers should be able to precisely recognize targeted built infrastructure from the randomly collected image and video data, (2) to get a quality digital replica of infrastructure, the collected image and video data should be properly processed and optimized in terms of patterns, segments, and overlaps of contained information, and (3) reconstruction of 3D sparse point clouds can be a solid foundation to develop dense point clouds as well as 3D texture models of existing infrastructure. I am in a very competitive position to take on this three-month endeavor given my extensive previous research experience in photogrammetry, spatial modeling, and computer vision for built infrastructure. A total of two specific aims have been formulated to achieve the objective in this application. With the output of this research, the PIs plan to submit a proposal to NSF Civil Infrastructure Systems (CIS) program in Fall 2020.

**RESEARCH PLAN:**

Detailed below are the ***two specific aims*** that will be pursued to develop an automated reconstruction process of 3D sparse point cloud for utility infrastructure using dashboard cameras.

**Specific Aim #1: Optimized Sequence Image Segments of Utility Infrastructure**

* 1. **Introduction**: To create quality 3D sparse point clouds, it is vital to obtain good overlapped images of targeted utility infrastructure. A dashboard camera detects not only interested utility infrastructure, but also many uninterested objects such as vehicles and pedestrians during the object-detection process. A study needs to (1) snip video segments which contain interested utility infrastructure by automatically recognizing necessary objects and (2) convert to sequence images by optimizing percentage of overlapping areas between sequenced images. Figure 1 shows a result of a pilot study to detect a fire hydrant on Allisonville Road in Fishers.

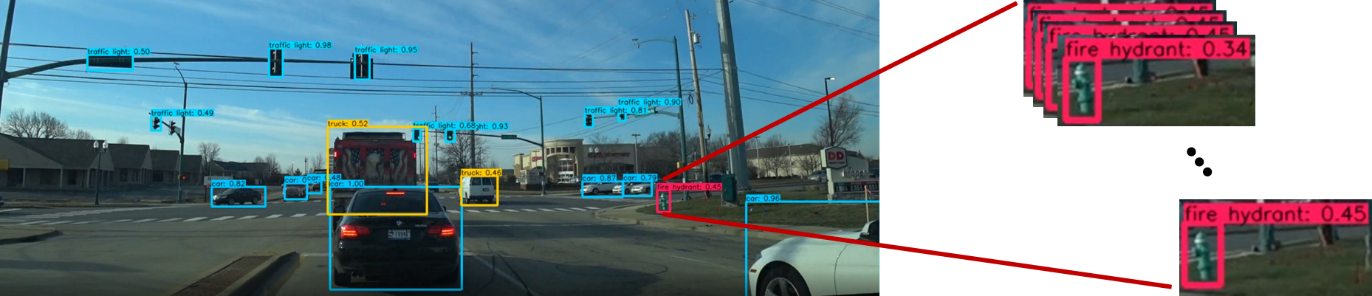


Figure Fire Hydrant Detection from a Pilot Study

* 1. **Research Design**: I plan to accomplish the proposed objectives by conducting the following studies:
     1. Detecting and extracting existing utility infrastructure based on Convolutional Neural Network (CNN) by developing rich training data set. The research team has purchased two dashboard cameras to collect image data to create rich training data set. We will use You Only Look Once (YOLO) object detection classifier [04] to detect fire hydrants located next to public roads. YOLO uses a single CNN to predict multiple bounding boxes and classficiation probabilities for those boxes.
     2. Developing optimized crop function for utility infrastructure with Global Positioning System (GPS) data. The research team will develop an optimized method to (1) snip the predicted bounding boxes of the utility infrastructure and (2) remove the background and noise of the boxes. This step will help to save data storage and enhance computational capacities. In addition, the purchased dashboard cameras have embedded GPS modules. We will develop a function to integrate GPS data into the predicted boxes, which can enhance the accuracy and precision of the 3D sparse point cloud in the later research step.
     3. Converting video data to sequenced images by optimizing a number of frames. As I mentioned earlier, it is important to have good overlapped sequencing images to get a good quality of the 3D sparse point cloud. It is all depending on what processing algorithm will be used, but a rule of thumb is higher than 85% of overlapping areas to get a good quality of the 3D sparse point cloud. During this step, the PIs will work on developing automated converting process from video data to sequenced images that have reasonable overlap percentage. The expected outcome of this step is tradeoff matrices among overlap percentage, computational processing time, and overall quality of 3D sparse point cloud.

**Specific Aim #2: Information Fusion for Automated 3D Sparse Point Cloud of Utility Infrastructure**

* 1. **Introduction**: This research step will develop a model to estimate 3D sparse points from optimized sequence 2D image segments of utility infrastructure based on Structure from Motion (SfM) [05] methodology. This study will validate the proposed model by comparing it with the 3D cloud points constructed from surrounding video data of utility infrastructure. Figure 2 shows a result of a pilot study to construct a 3D sparse point cloud of a fire hydrant using a mobile phone (surrounding video).

Figure Fire Hydrant 3D Sparse Point Cloud from a Pilot Study

* 1. **Research Design:** I plan to accomplish the objective of this aim by conducting the following studies:
     1. Detecting and matching feature points of the optimized sequence image segments of utility infrastructure based on Scale-Invariant Feature Transform (SIFT). The purpose of the SIFT algorithm is to simulate the multi-scale characteristics of images, and it uses Gaussian convolution to construct the scale space. This research, first, will process the image segments to construct different levels of scale and extract the SIFT feature in the segments. The study, then, will match extracted feature points, which consists of two steps (1) the nearest neighbor method to find the most matching point pairs and (2) filtering these matching point pairs. We will employ the threshold value of feature points from literature to validate the outcome of matching pairs.
     2. Solving fundamental matrix and bundle adjustment: Transforming from 2D to 3D data is the process of motion recovery. To recover motion information, the SfM algorithm will imitate how human eyes observe objects in the world using the fundamental matrix, which can estimate 3D coordinates of each feature point. After solved the fundamental matrix, we will differentiate back-project points and original points using the principle of the least square method. This step will mathematically find such a set of estimates of the matrix and the space points to minimize the mean squared error (MSE) between the back-projection points and the feature points.
     3. Validating quantitative and qualitative results by comparing with constructed 3D sparse points from surrounding video data of utility infrastructure: Two-sample image data sets of a fire hydrant (one from a dashboard camera and one from a mobile phone) will be compared to validate outcomes of the proposed model. The PIs will investigate the constructed 3D sparse point clouds of a fire hydrant qualitatively and quantitatively using Visual SfM and MySfM software, respectively. We will visually investigate the details of constructed clouds as well as summarize sizes of point clouds and consumed computational resources and time in tabloid format. Based on the expected research outcome over the summer, we will be able to continuously develop models to construct 3D dense point clouds of various utility infrastructure.

**REFERENCE:**

[01] Stojanovic, V., Trapp, M., Richter, R., Hagedorn, B., & Döllner, J. (2018). Towards The Generation of Digital Twins for Facility Management Based on 3D Point Clouds. In Proceeding of the 34th Annual ARCOM Conference (pp. 270-279).

[02] Tao, F., Sui, F., Liu, A., Qi, Q., Zhang, M., Song, B., ... & Nee, A. Y. C. (2019). Digital twin-driven product design framework. International Journal of Production Research, 57(12), 3935-3953.

[03] North American Wood Pole council (https://woodpoles.org/Why-Wood-Poles/How-Poles-Are-Made)

[04] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).

[05] Ullman, S. (1979). The interpretation of structure from motion. Proceedings of the Royal Society of London. Series B. Biological Sciences, 203(1153), 405-426.