

AUTO-EXTRACTION OF URBAN FEATURES FROM VEHICLE-BORNE LASER DATA

Dinesh MANANDHAR, Ryosuke SHIBASAKI

Centre for Spatial Information Science
The University of Tokyo, Japan
dinesh@skl.iis.u-tokyo.ac.jp, shiba@skl.iis.u-tokyo.ac.jp

Working Group IV/7

KEY WORDS: Laser Data, Feature Extraction, Mobile Mapping, Texture Mapping, 3-D Modeling

ABSTRACT:

Laser mapping has become quite popular in recent days due to its capability of providing information directly in three dimensions. These days, we observe the emergence of new demands for 'along-street' 3-D data. 'Along-street' 3-D data provide a real or true visualization of the environment that is different from virtual reality. However, there still lacks a reliable and quick method of acquiring 3-D 'along-street' data of the urban area at higher resolution. Most of the vehicle-borne systems developed so far are based on stereo photographs as the main source of data. In order to overcome the problem of acquiring 3-D data in urban area reliably, quickly and at high resolution, we have developed a vehicle-borne laser mapping system (VLMS). The system is equipped with laser scanners and line cameras for data acquisition. The system will assist in building 3-D GIS database of urban areas.

In this paper we focus our discussions on auto-extraction of building faces from the laser data captured by VLMS. In an urban area we encounter many different types of features like buildings, roads, utility facilities, vehicles, trees and others. Geo-referenced raw laser data were used for extraction of urban features. The raw laser data were simply point cloud data in 3-D space without any other information like intensity value. We have classified the laser data into various classes like building faces, road surfaces, tunnels, utility poles, vehicles and trees. Our discussions will be basically limited to the extraction of building faces with some explanation on extraction of road surfaces, trees and tunnels.

1. INTRODUCTION

The vehicle-borne mobile mapping technology has been developed around late 80's. The development of the mobile mapping system became possible due to the availability of GPS signal to the civilian community. As per our knowledge, the vehicle-borne mobile mapping systems are based on image data acquired by CCD cameras. However, image-based device cannot provide automated extraction of 3-D features and in some cases it even fails due to the lack of correct identification of corresponding control points on the image. On the other hand, a laser scanner provides data directly in 3-D. Thus, automated extraction of 3-D features can be achieved by using laser scanner though it lacks texture information. The lack of texture in laser data can be complemented by acquiring image data simultaneously. This led us towards the development of the VLMS, which is capable of acquiring urban data ('along-street' data) quickly and reliably. This system is capable of observing the objects at closer range, thus giving greater details.

For reference of some of the image based systems, refer GPS-VanTM (Bossler et al, 1991), VISAT-Van (Schwarz et al, 1993, Li et al, 1994), TruckMapTM (Pottle, 1995), KiSS (Hock et al, 1995), GPS Vision (He, 1996) and GeoMaster (Tamura et al, 1998). Refer Manandhar and Shibasaki, 2000 for details about VLMS system description and calibration. Refer Manandhar and Shibasaki, 2001 for details about feature extraction from laser data that are not explained in this paper.

Feature extraction from range data is quite a complex task, as the range data contain nothing but only clouds of 3-D points.

The extraction is more complex as the observation is from a moving platform (vehicle). The outdoor observation from a moving platform captures all the points reflected from the features like building faces, road surface, trees, poles, cables, guard rails etc. Thus the data consists of a heterogeneous mixture of laser points. The challenging work is to automatically identify and classify these points reflected by each of these features and then finally make a 3-D model to represent the features. In this paper, we present algorithms to classify these points into different groups and then generation of 3-D models (planar faces) to represent building faces.

2. SYSTEM CONFIGURATION

The mobile mapping system is shown in figure 1. The system is called Vehicle-borne Laser Mapping System (VLMS). This system consists of six line cameras, three laser scanners and Hybrid Inertial Survey System (HISS) which consists of DGPS, INS and an electronic odometer for position and orientation data. The laser scanning frequency is 20Hz, which enables driving at a speed of 40km/hr or more (we have acquired data at 80km/hr) without any serious problem. Data can be continuously acquired as long as the long as the INS / GPS errors (especially the drift error) are within the allowable limit. The along track resolution and across track resolution depends on the vehicle speed and the distance from the sensor to the target respectively. The along track resolution varies from 14cm to 110cm for speed from 10km/hr to 80km/hr. The across track resolution varies from 2cm to 52cm for a range distance from 2m to 60m.

3. FEATURE EXTRACTION

We define feature extraction as the process of classifying the laser points reflected by the urban objects into different groups like building surface, road surface, ground surface (e.g. walking path), trees, poles, tunnel, parked vehicles etc.

The range data includes only clouds of points in three dimensions. Figure 2 shows a sample of range data. It has no other information like reflectance value etc. We would like to extract various features usually seen in the urban environment. Basically, we categorize these features into three classes viz: man-made, natural and dynamic (moving) / static (non-moving) features. The natural features are trees, bush etc. The man-made features are buildings, roads, utility facilities (poles etc), pavement, guardrails etc. The dynamic features include moving cars and pedestrians and the static features include parked vehicles (something which can be dynamic but at static state during the observation period). Our primary focus is on extraction of man-made features like building faces, trees and road surfaces.

outdoor mobile environment, we do not have control on viewing the objects as we like. Thus we get quite heterogeneous mixture of range points reflected by all the urban objects within the viewing angle of the scanner. This makes the identification of the objects very complex.

Figure 3 shows some of the laser scan lines that consist of range points reflected by buildings, trees and roads. Generally, we can observe three sets of geometric structure of the range data as shown in figure 3. Building faces reflect points in vertical linear fashion (not necessarily straight line), ground or road surface reflect points in horizontal linear fashion and natural objects like trees

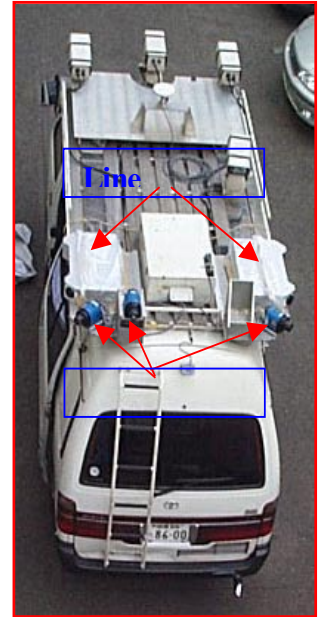


Figure 1: VLMS vehicle

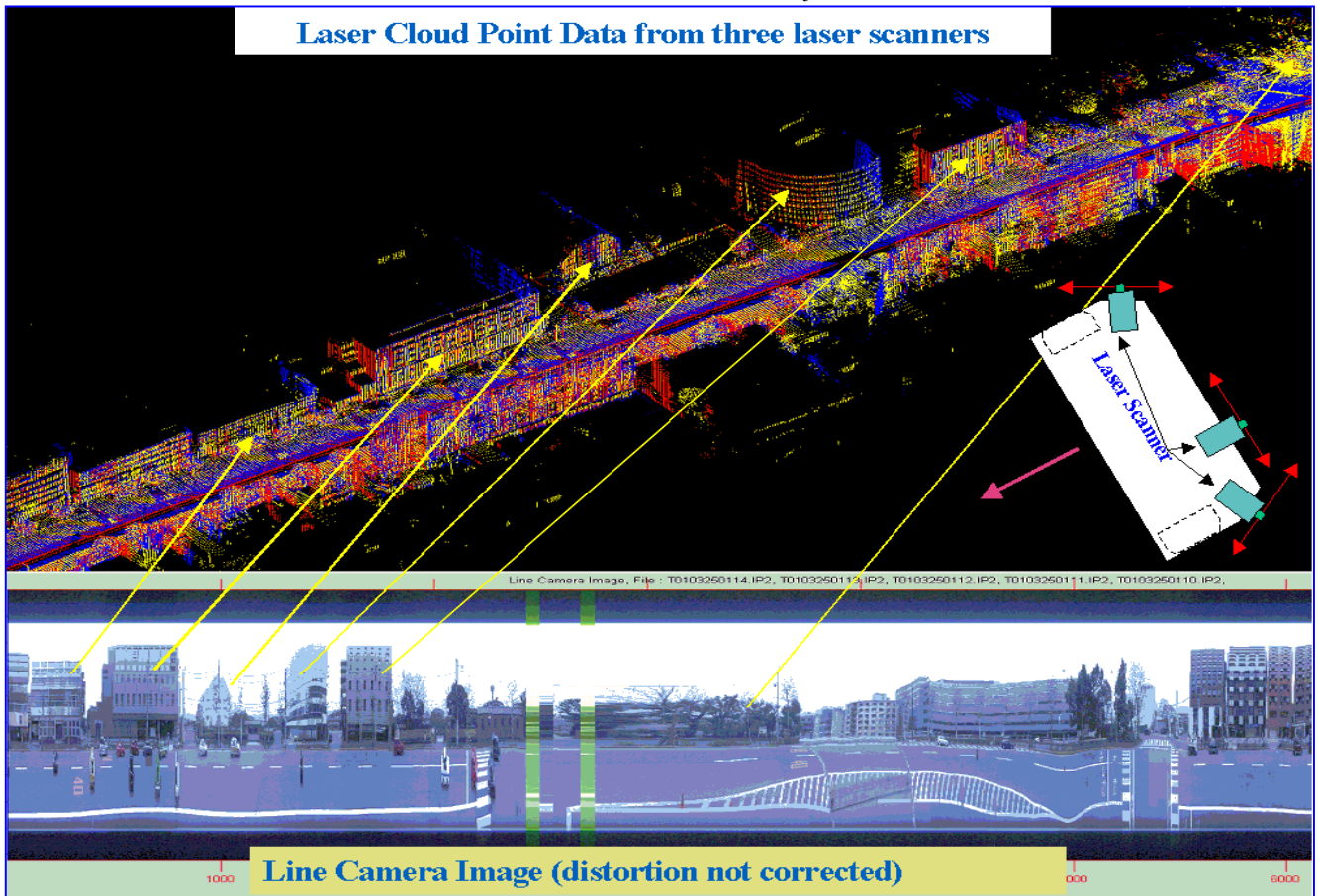


Figure 2: Point Laser Data from VLMS. Laser Data from three laser scanners are integrated and shown by red, blue and yellow point

3.1 Basic Concept of Feature Extraction

The laser scanner scans the objects around it's surroundings within 300° of angular coverage. As the vehicle moves forward, the scanner scans the objects sequentially at a frequency of about 20Hz. As the system is designed for

reflect points in scattered fashion. Thus a quick look of the data gives us rough indication of buildings, roads and trees. The road surface being the nearest surface to the scanner with respect to other objects, the reflected laser point density is higher compared to other objects. We use this information (geometric structure of points, their scatterness and point density) for feature extraction.

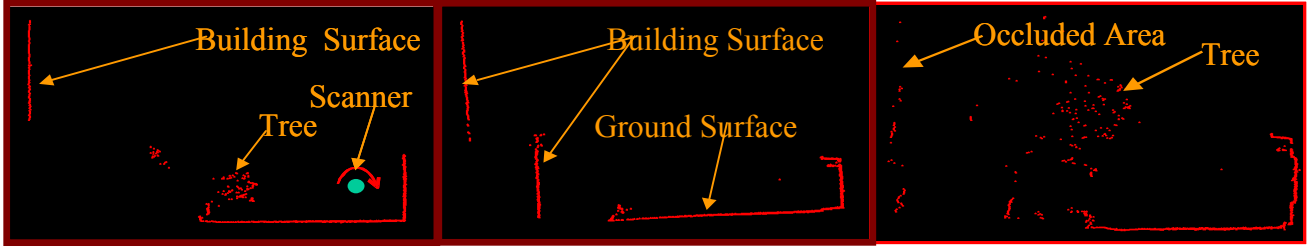


Figure 3: Individual scan lines of laser points. Laser points from natural features are scattered. Laser points from man-made features follow some geometric pattern.

3.2 Road Surface Extraction

Road surface being the nearest object relative to other objects like buildings and trees, the laser point density is also higher due to higher across track scanning resolution. A lot of range points are reflected from the road surface, as there are no obstructions between the scanner and the object. Since, the road surface height varies smoothly with a very small deviation (about 1% slope) along the width of the road, we get reflected range points that have small variation in height value (about 10~20cm).

Histogram analysis of height value is computed for every scan line with a bin width equal to 10cm. The bin with highest value (number of points) is taken and some other parameters to compensate the measurement errors are added to it. Then this value is taken as the threshold value to filter the laser points data are reflected from the road surface. Refer Manandhar & Shibasaki, 2001 for details about it and the results. Figure 4 shows the extracted road surface along with other features.

3.3 Scattered and non-Scattered Points Extraction

After identifying the laser points that belong to the road surface, the remaining points are analysed for classifying into two broad categories of scattered (natural features) and non-scattered (man-made) features. Points from natural features are scattered due to their irregular geometric shape, where as points from man-made features are generally not scattered as they have certain geometric shape. In order to analyse the scatterness of the points in three dimensions, we use range distance analysis given by equation 1.

$$d = (d_{n+1} - (d_{n+2} - d_n)/2) \quad (1)$$

where, d_n , d_{n+1} and d_{n+2} are the distance from the sensor position to the n^{th} , $(n+1)^{\text{th}}$ and $(n+2)^{\text{th}}$ range points on a scan line.

Every transition between the smooth value and scattered value indicates either different features or discontinuity in the same feature (in case of spike like transition) due to location of features at different distance from the sensor or due to partly occlusion.

In order to ease the feature classification, we have analysed the standard deviation (SD) of every cluster of points along the x, y and z coordinates separately. Our hypothesis is that, points from building surface will have much higher SD value

in z-coordinates compared to x and y coordinates. Similarly, a ground surface like footpath will have much higher SD value in x or y coordinates compared to z coordinate. Trees exhibit high SD values in all the three directions due to their scattered nature of range points. Point clusters that have smaller SD value on z coordinate (or less than the across track resolution) are classified as vertically oriented points. Point clusters that have higher SD value for either x or y are classified as horizontally oriented points. If the SD value is higher for all the three components then they are classified as scattered points, which generally belong to natural features like trees.

3.4 Building Face Extraction

The laser data as shown in figure 2 are processed scan line by scan line by analysing the geometric orientation of the points on each scan line. As mentioned earlier, the laser data are first processed for the extraction of road surface and then again analysed for identifying the natural and man-made (scattered and non-scattered) features by using the derivative analysis of the range distance. The man-made features exhibit certain geometric properties that are either linear, planar or follow some other geometric shape, but are not scattered. The derivative analysis identifies the points reflected by the man-made features as well as classifying these features into different geometric shape categories like vertically aligned or horizontally aligned points. Vertically aligned points are taken for extraction of the building faces. It is not necessary that the points should align strictly vertical, which is not possible in practice. One scan line may have more than one group of points that are vertically aligned, horizontally aligned or scattered. Thus we take every group of points, which are vertically aligned for the extraction of building faces.

Straight lines are fitted to each group of vertically aligned points on each scan lines. After fitting the line, any points that are beyond the threshold distance from the fitted line are rejected. The line is then fitted once again. This is necessary to remove any possible inclusion of the outliers in the fitting data. The line fitting is done by using a robust least square (RLS) method, which is based on the weighted least square (WLS) model. The weight is computed using bi-square weight function. Once the lines are fitted to each group of scattered points, a line connectivity table is generated. This table is generated by using the line-to-line vertical centre distance and direction vector of a pair of line between the line segments within the scan line as well as between the line segments to the next neighbouring scan line. This helps us to identify the line segments that can be grouped together within the scan line or to the neighbouring scan line to form a surface. This will also help us to separate two different types of surfaces like building face and guard rail or hoarding

boards as they are at different range distance as seen from the sensor position.

After creating the line connectivity table, we can simply take a pair of lines that can be connected together and form a patch surface from them. The results of this are shown in figure 6. The laser point data for this result is shown in figure 5. However, due to occlusion or glass windows, there are many empty spaces in the generated surface. When a laser beam pass through the glass window, it is either lost in the space (as the laser scanner can not receive the reflected signal within the prescribed time) or reflected by inner objects like blinds, curtains etc. In the later case, the laser points appear as scattered data and are classified into the natural class (trees) instead of man-made feature class. Since the surface is generated by using only a pair of line, it can generate non-planar surface as well. The problems are however, the empty spaces between the surfaces and irregular shape of the surface. This type of surface generation is not effective for texture mapping as we will have numerous small size surfaces. Due to the fitting errors and some other errors the surfaces that belong to a single building face might not align perfectly. The normal vector of each surface might be pointing at different direction. In order to rectify these problems we have further enhanced the algorithm as explained below.

All the line segments that can be grouped together are selected based on the information from the line connectivity table. Once the lines are identified, the points that are used to fit these lines are identified. Thus we will have a group of scattered points, which are used for fitting the lines. This process leads us a few steps backward again. Of course, we could have taken only the end points of the fitted line for fitting the plane. But, this will induce the fitting error of the straight lines as well. Next, we fit planar face to identify the building surface from these points using the well-known plane equation model given by equation 2.

$$Ax + By + Cz + D = 0 \quad (2)$$

The fitting is done to minimise sum of the squares of the distance from the points to the fitted plane rather than to minimize the sum of the squares of the fitting errors, given by equation 3. Equation 3 is solved to minimize the sum of the squares of the distances subject to the constraint given by equation 4.

$$\|d\| = [Ax + By + Cz + D = 0] / \sqrt{A^2 + B^2 + C^2} \quad (3)$$

$$A^2 + B^2 + C^2 = 1 \quad (4)$$

The Fitting model is solved by using a singular value decomposition method. The points are then projected back to the fitted plane using equation 5.

$$\begin{aligned} Px &= x - A(Ax + By + C + D) \\ Py &= y - A(Ax + By + C + D) \\ Pz &= z - A(Ax + By + C + D) \end{aligned} \quad (5)$$

In order to regenerate the planar surface, the centroid of the projected points is taken as the reference point for the parametric generation of the plane. The surface is generated using a parametric model of the plane equation given by equation 6.

$$X = X_0 + \alpha R + \beta S \quad (6)$$

The vectors R and S in equation 6 are two orthogonal vectors to the normal vector of the fitted plane and they lie on the fitted plane. Hence, the normal vector, vectors R and S are all orthogonal to each other. The parameters alpha and beta are any real numbers and are increased or decreased at a step of 0.5 till the regenerated surface exceeds the boundary defined the maximum and minimum coordinates of the point data. We assume that the building faces are planar in shape though this may not be true for a few buildings. There might be buildings with curved or irregular shape. Irregular shape means a mixture of planar face as well as curved face within a single front face. However, at the moment, for simplicity, we have considered only the planar faces, which can explain most of the building shapes.

Figure 7 shows the results of planar face fitting to laser point data that belong to the building face. The results show that the planar face thus generated can better represent the building faces. The results are better than the results shown in figure 6, which are obtained by drawing patches for a pair of connecting neighbouring lines. However, one of the difficult tasks is to determine the correct size of the planar face so that it can exactly represent each individual building face. For example, identification of correct building height is not so easy. The uppermost rows of laser points may not always represent the top of the building. Another difficulty is to identify the exact boundary between the two similar buildings. The algorithm will produce one single planar face for the buildings that are joined together at the same range distance from the sensor. Still, we can see the gaps between the planar faces that belong to the same building face. This is either due to lack of point data (probably from occlusion) or due to the scattered data (in case of glass windows). The occlusion may be by trees, pedestrians, vehicles and signboards or hoarding boards. Thus the algorithm developed so far though much improved than it's previous version, still lacks the capability of exact representation of the building faces. We will add these capabilities in the future version. One of the possible ways to overcome such problem is to analyse the normal vector of the fitted plane. If the neighbouring normal vectors point to the same direction (analyse the angle between the normal vectors) within a small threshold value and if the distance between the planes (horizontal distance) is very small (a few centimeters, around the across track resolution value), then those planes can be merged together to form a new and bigger planar face. Also, it is necessary to merge all the three data sets taken by three laser scanners. Since, they have different viewing angles, the

problems of identifying boundaries can be solved to some extent. Finally, the image data will be used for texture mapping. The texture will also help to determine the boundary between the buildings.

3.5 Extraction of other Features

We have also extracted other urban features like tunnels, utility poles and vehicles using only laser point data. The extraction of tunnels is quite successful. The extraction of vehicles are partly successful and needs further development in the algorithm. The extraction of poles though partly successful is limited by other factors as well. For example, the speed of the vehicle (VLMS) defines the along track resolution. If the vehicle speed is high (80km/hr or so), it is difficult to capture poles as the resolution is about 1.1 metre and normally the pole diameter is less than this. Thus, in order to capture the utility poles, the vehicle speed should not be above 40km/hr, which gives along track resolution of about 50cm. Refer Manandhar and Shibasaki, 2001 for details.

3.6 Conclusion

We have developed algorithms for the extraction of building faces from laser data taken by vehicle-borne laser mapping system. The results of the algorithm are satisfactory. It can successfully classify the laser points that belong to the building faces and generate planar faces to represent these building faces. However, the algorithm can not exactly determine the size and shape (this means to determine the shape of the building accurately within the accuracy of the data, which is around one metre) of the building face and cannot determine the boundary between the two similar buildings. Thus the results from this algorithm may be good enough for visualisation purpose rather than exact spatial analysis. We will include these capabilities in the future version.

3.7 References

References from Books:

Medioni, G., M.S. Lee, C.K. Tang, 2000, A Computational Framework for Ssegmentation and Grouping, Elsevir Publication, ISBN: 0 444 50353 6

Mikhail, E.M., 1976, Observations and Least Squares

References from Other Literature:

Bossler, J.D., Goad, C.C., Johnson, P.C., and Novak, K., 1991, GPS and GIS map the nation's highways, GeoInfo Systems, March 1991, p.27-37

He, G., 1996, Design of a mobile mapping system for GIS data collection, IAPRS, Vol. XXXI, part B2, pp.154-159

Hock, C., W. Caspary, H. Heister, J. Klemm, and H. Sternberg 1995, Architecture and design of the kinematic survey system KiSS, Proc. of the 3rd Int Workshop on High Precision Navigation, Stuttgart, April, p. 569-576

Hoover, A., Jean-Baptiste, G., Jiang X., J., Flynn, P.J., Bunke H., Goldgof, D., Bowyer K., A Comparison of Range Image Segmentation Algorithm, URL: <http://marathon.csee.usf.edu/range/seg-comp/SegComp.html>

Li, R. Mobile Mapping – An emerging technology for spatial data acquisition, <http://shoreline.eng.ohio-state.edu/ron/teaching/894a/paper1.htm>

Manandhar, D., Shibasaki, R., 2000, Geo-referencing of Multi Sensor Range Data for Vehicle-borne Laser Mapping System, 21st Asian Conference on Remote Sensing (ACRS), 4-8 December, Taipei, Vol 2, pp. 932-937

Manandhar, D., Shaibasaki, R., 2001, Proceedings of ACRS 2001–22nd Asian Conference on Remote Sensing, 5-9 November 2001, Singapore, Vol. 2, pp 1113 – 1118Pottle, D., 1995, New mobile mapping system seeds data acquisition, GIM, Vol 9, No. 9, p51-53

Schwarz, K.P., H.E. Martell, N.El. Sheimy, Ron Li, M.A. Chapman, and D. Cosandier 1993, VISAT – A mobile highway survey system of high accuracy. Proc. of the IEEE vehicle navigation and information systems conference, October 12-15, Ottawa, pp. 476-481

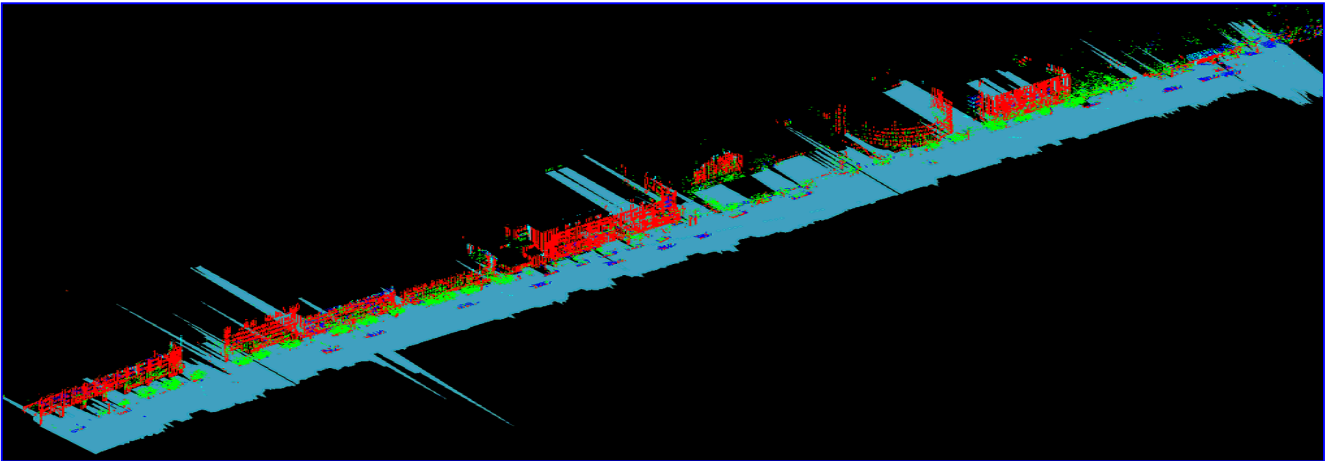


Figure 4: Classified laser point data. Red color shows the patches created from the fitted straight lines that describe building faces, cyan color represents the road surface, green color represents scattered points classified as trees and blue color represents vehicles

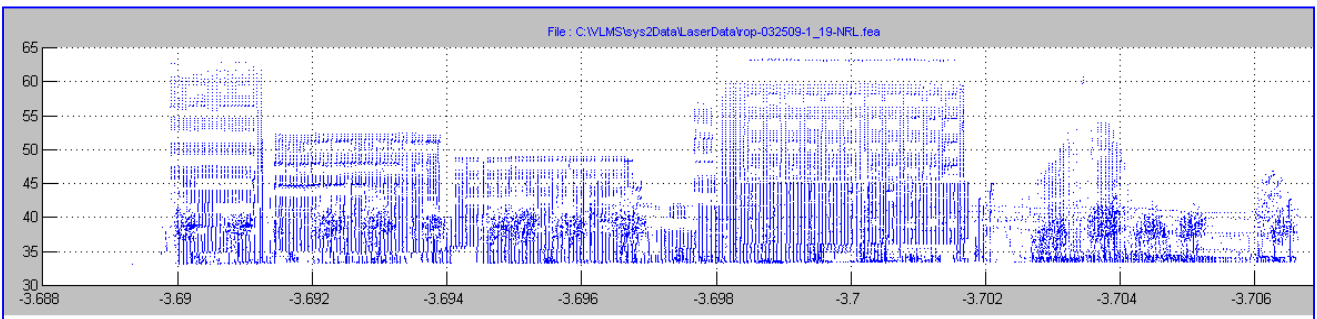


Figure 5: Laser point data as acquired by VLMS.

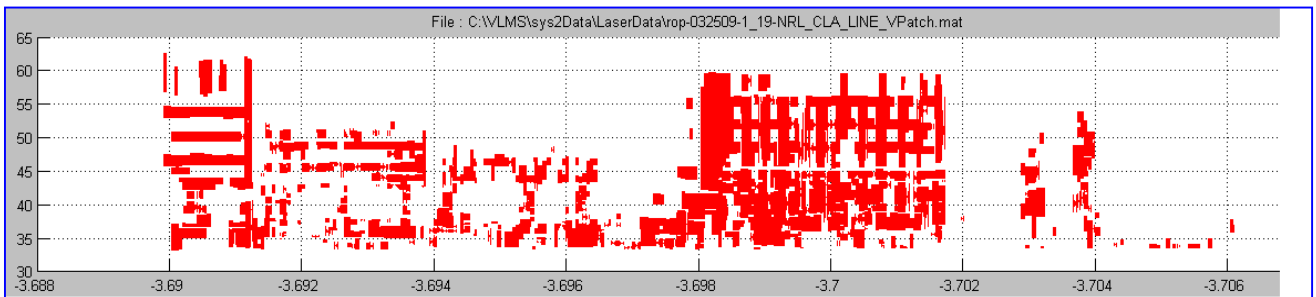


Figure 6: Patches to represent the Building Faces from classified data. The classified data are the points that are marked as vertically aligned

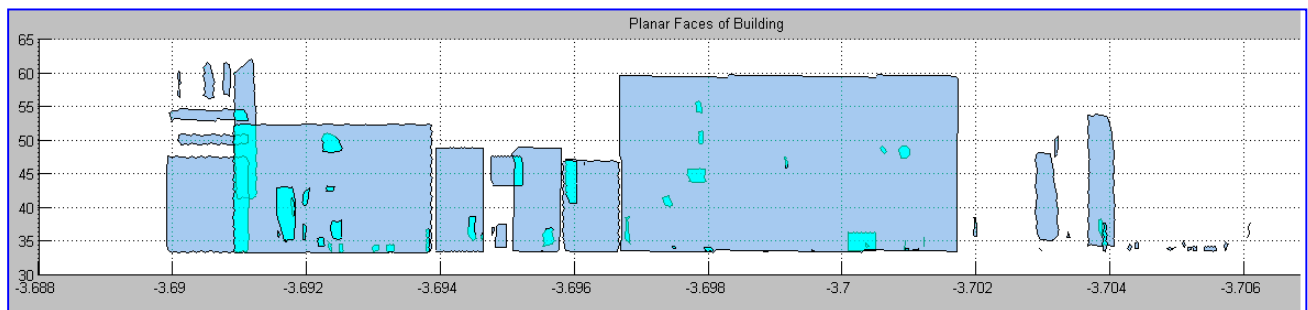


Figure 7: Planar Faces to represent the Building Faces from classified data. The classified data are points that are marked as vertically aligned