194 S. Sukumar, D. Page, A. Gribok, A. Koschan, M. Abidi, D. Gorsich, and G. Gerhart, "Surface Shape Description of 3D Data from Under Vehicle Inspection Robot," Proc. SPIE Unmanned Ground Vehicle Technology VII, Vol. 5804, Orlando, FL, pp. 621-629, March 2005.

Surface Shape Description of 3D Data from Under Vehicle Inspection Robot

Sreenivas R. Sukumar[1], David L. Page[1], Andrei V. Gribok[1], Andreas F. Koschan[1], Mongi A. Abidi[1], David J. Gorsich[2] and Grant R. Gerhart[2]

[1] Imaging, Robotics, and Intelligent Systems Laboratory, The University of Tennessee Knoxville, TN, USA 37996-2100;

[2] U. S. Army RDECOM Tank-Automotive Research, Development and Engineering Center, Waren, MI, USA 48397-5000

ABSTRACT

Our research efforts focus on the deployment of 3D sensing capabilities to a multi-modal under vehicle inspection robot. In this paper, we outline the various design challenges towards the automation of the 3D scene modeling task. We employ laser-based range imaging techniques to extract the geometry of a vehicle's undercarriage and present our results after range integration. We perform shape analysis on the digitized triangle mesh models by segmenting them into smooth surface patches based on the curvedness of the surface. Using a region-growing procedure, we then obtain the patch adjacency. On each of these patches, we apply our definition of the curvature variation measure (CVM) as a descriptor of surface shape complexity. We base the information-theoretic CVM on shape curvature, and extract shape information as the entropic measure of curvature to represent a component as a graph network of patches. The CVM at the nodes of the graph describe the surface patch. We then demonstrate our algorithm with results on automotive components. With *apriori* manufacturer information about the CAD models in the undercarriage we approach the technical challenge of threat detection with our surface shape description algorithm on the laser scanned geometry.

Keywords: under vehicle inspection, laser range scanning, automotive component description, surface description, 3D surface feature.

1. INTRODUCTION

Under vehicle inspection has been traditionally accomplished through inspection personnel walking around the vehicle with a mirror at the end of a stick. The inspection personnel identify weapons, bombs and other security threats based on what they see on the mirror. The mirror-on-the-stick system allows only partial coverage under a vehicle and is restricted by ambient lighting during the day and requires a flash light in the night when there is very limited lighting. Such a system also places the inspecting personnel at an endangering risk of a possible detonation. As part of the Security Automation and Future Electromotive Robotics (SAFER) program [1] we are developing a robotic platform that deploys "sixth sense" sensors for threat assessment. In this paper, we present results on the idea of incorporating a 3D range sensor on a robot to identify suspicious objects. Our 3D sensing modality provides the flexibility of deployment without considering the ambient lighting. In addition, our system enhances the ease of visualization and the possibility of automating threat assessment.

In Figure 1, we show the robotic platform and the 3D geometry of the scene containing the muffler, shaft and the catalytic converter. With prior manufacturer's information on the components that make the undercarriage of the vehicle, threat

Further author information: Email: ssrangan@utk.edu ; phone: (865) 974-9213; fax: (865) 974-5459

Unmanned Ground Vehicle Technology VII, edited by Grant R. Gerhart, Charles M. Shoemaker, Douglas W. Gage, Proceedings of SPIE Vol. 5804 (SPIE, Bellingham, WA, 2005) · 0277-786X/05/$15 · doi: 10.1117/12.602930

detection can be approached as the computer vision task of designing a system that can capture the geometric structure of an object and using the shape information for intelligent interpretation.

Robotic Platform | Under Vehicle Scene | Sensed 3D data

Figure 1: Imparting 3D sensing capabilities to vehicle inspection robot.

To this end, we propose a curvature variation measure (CVM) as a surface feature on 3D datasets. We then extend the definition of the CVM to a curvature-based shape analysis algorithm to interpret surface shape information of the CAD models and sensed models alike. The idea is to identify objects that bear little or no similarity to the expected undercarriage components in the scene as potential threat objects.

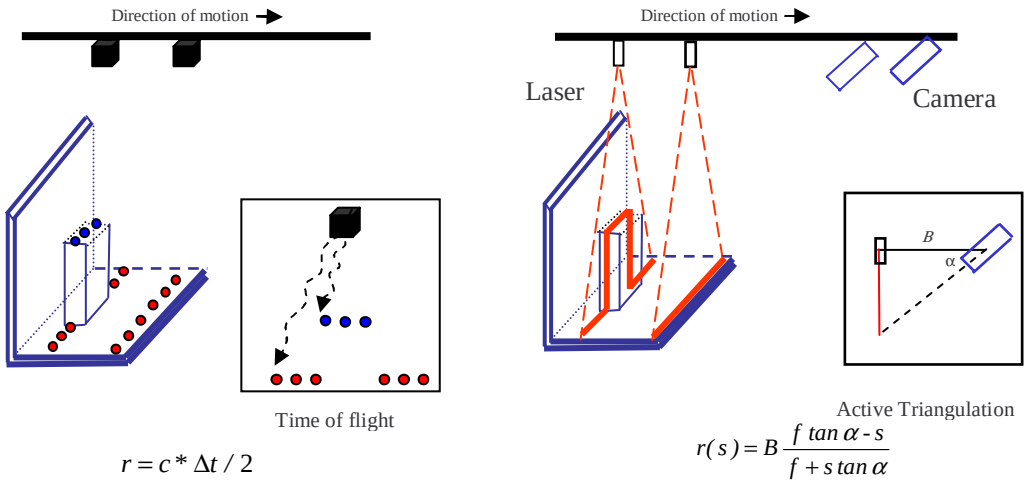
We have organized this paper as follows. We first describe in Section 2, the data collection system and the processing involved in visualizing the sensor data from the under vehicle inspection robot. In Section 3, we discuss our shape analysis algorithm on the polygonal mesh models that are a common representation of the CAD data and the laser scanners. We demonstrate experimental results in Section 4 and conclude with future work in Section 5.

2. UNDER VEHICLE 3D DATA COLLECTION

In this section, we address several challenges in the robotic data collection process underneath a vehicle. The most significant of those being the field of view of the scanning mechanism. The field of view underneath the vehicle is limited by the ground clearance and the large variation in size of the components that make up the scene. In addition, we also need to provide the flexibility to accommodate the variance in the ground clearance of a variety of cars and automobiles from different manufacturers. Typically, the ground clearance varies from a minimum of 10 centimeters (in compact cars) to up to 90 centimeters (in large trucks). The variability in the ground clearance, the resolution required for the shape analysis and the field of view requirements narrow our choice to time-of-flight systems and laser triangulation scanners over other state-of-the-art photographic measurement systems.

To make a beginning, we collect data using both the time-of-flight range finder and an active triangulation based system. We mount each of these sensors on a mobility platform and use it to build the 3D model of the scene assuming the motion trajectory of the sensing module to be linear.

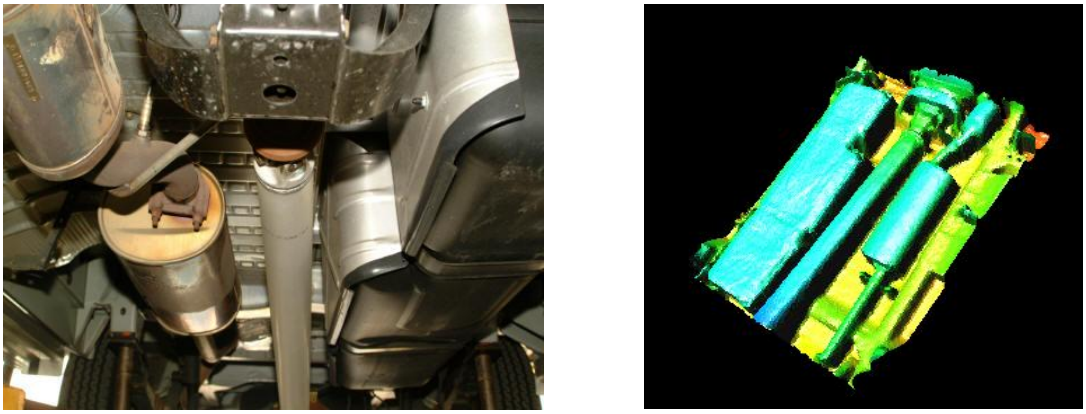
First, we discuss the time of flight system. The SICK time-of-flight scanner shoots out a laser pulse that gets reflected from objects within the 8 meters radius in the 180 degree field of view of the sensor. The built-in electronics detects the returned pulse and computes the distance based on the time taken for the laser pulse to reach the object and return to the detector in the sensor (Figure 2a). Our choice of this system is based on the fact that one pass under the vehicle is sufficient to visualize the entire 3D scene. We show the results from such a scanner in Figure 3. On the left we show the photograph of a scene underneath the vehicle (Dodge RAM) van and the corresponding 3D scene that we have imaged and rendered in 3D. The data seems sufficient for visualization and manual inspection. The accuracy of the 3D reconstruction appears to be in the order of a few centimeters. Though we are able to perceive the mechanical components such as the muffler, shaft and the catalytic converter these results do not seem to be provide the geometric accuracy required for the automation towards threat detection. The electronics of the scanner limits our data to that such a resolution and accuracy for the ground clearance underneath the vehicle.



(a) (b)

Figure 2: The principle behind time-of-flight scanners and the active triangulation based scanners. (a) The range (distance) is calculated using the time taken for the laser traveling at the speed of light to return to the sensor. (b) The range is a function of the geometry of the setup (baseline *B*, angle α), the sensor offset position of the laser in the camera operated at a focal length of *f*.

We now present our experience with using the IVP range scanner. The RANGER system consists of two parts: the range sensing camera and a low power laser. The sheet-of-light laser shoots on the target scene as shown in Figure 2(b) and the camera detects the laser as a single profile of the object surface shape parallel to the imaging plane. Such profiles are accumulated by moving the hardware setup relative to the scene traversing the area of interest. We use a pre-calibrated setup of the scanner. Although a powerful laser was used to counter ambient lighting, we could not compensate for spectral reflections since the metallic surfaces under a vehicle exhibit strong spectral reflection properties. A laser further complicates this problem as internal reflections lead to significant errors in range estimation. A promising approach for this problem involves the use of an optical filter tuned to the frequency of the powerful laser. The filter in front of the camera allows the data collection to isolate the initial reflection of the laser and thus improve the range sensing capabilities.

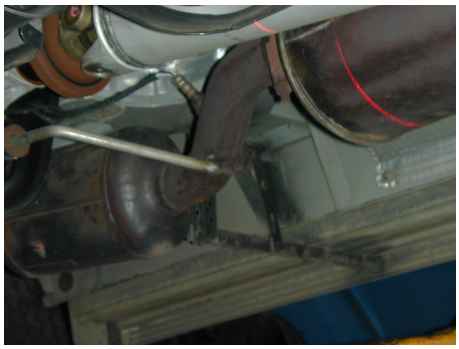


(a) (b)

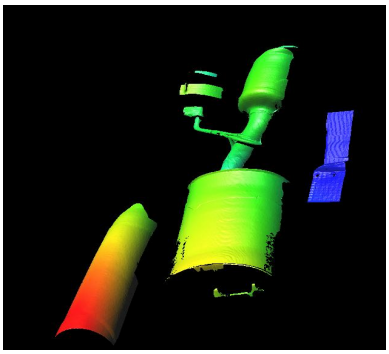
Figure 3: Imaging the under vehicle scene using the SICK time-of-flight scanner on a mobility platform. (a) The scene underneath a vehicle. (b) 3D scene imaged at an angular resolution of 0.5 degrees using the SICK laser range finder.

The other serious problem in acquiring complete 3D data from the robot is that of occlusions. The objects underneath a vehicle have different shapes and scales located at different depths. One object occludes another from the range sensor at the time of scanning. Our solution to fill up occlusions using multiple scans is a laborious one involving multiple calibration procedure iterations and redundant information. This large data set allows high fidelity geometry that other 3D sensors do not offer, but at the price of data redundancy and a potential data overload.

We present the results in Figure 4. Again, on the left we show the scene of interest and on the right the result rendered in 3D. We note that the output of this system appears to convey accurate geometric information that can be used for surface-based shape analysis. We have assumed a perfectly linear and smooth trajectory under the vehicle and ignored the jerks in the trajectory of the scanning mechanism.



(a)



(b)

Figure 4: Imaging the under vehicle scene by mounting the IVP active triangulation system on a mobility platform. (a) The scene of interest. (b) Rendered 3D scene.

With the 3D data, we now proceed to the computer vision task of formulating an algorithm in designing a system that can capture the geometric structure of an object and store the subsequent shape and topology information for interpretation.

3. THE CURVATURE VARIATION MEASURE

We begin this section with a brief summary of the state-of-the-art 3D features. Dorai and Jain [2] propose shape index and curvedness as surface features in their COSMOS framework for object description from range images. They represent un-occluded range images as maximal surface hyperbolic patches of constant shape index and formulate a shape spectrum on these patches. The definition of shape index and curvedness assumes the underlying topology of the range image. They also propose a graph-based system based on the connectivity information of the hyperbolic surface patch primitives towards object recognition. In direct application to triangle mesh datasets, we find shape signatures like the 3D Fourier Descriptors [3], Local feature histograms [4] and Alpha shapes [5]. Belongie et al. [6] propose the log histogram of shape contexts. The shape contexts that capture the distribution of all points with respect to a reference on the image boundary have been extended to 3D by Körtgen et al. in [7]. Khotanzad and Hong [8] have extended Duda and Hart [9] to include a subset of 3D moments as shift and rotation invariant point-based descriptors. Cybenko et al. [10] use second order moments, spherical kernel moment invariants, surface area, and other metrics as features. Another category of methods that involve intermediate surface description include spin images [11] and harmonic shape images [12]. While spin images are data level descriptors that encode the global characteristics of the surface as a collection of points in an object centered co-ordinate system, the harmonic shape images decompose an object into its harmonics using a spherical raster. Both these methods involve high computational matching. Histogram based methods reduce the cost of complex matching but sacrifice robustness since such methods generate features based on the statistical properties of the object. Besl [13] has considered the crease angle at each edge between two triangles in a mesh as a surface feature. Osada et al. [14] demonstrate shape classification and database search capabilities of their simple shape functions using the shape distributions. Most of these methods assume unoccluded and ideal 3D data.

As background, we proceed to the automation for the under vehicle inspection task that will take manufacturers CAD models and laser scanned data as the inputs. The laser scanned data from the robot is bound to have occlusions and cannot be assumed to be complete. We hence identify the need for a surface feature on 3D meshes for object description of partial data. To this end, we propose the curvature variation measure as a surface descriptor for a part-based shape representation scheme.

Our CVM algorithm is based on a curvature heuristic that smoothly varying curvature conveys very little shape information, while unpredictable variations in curvature attribute to the shape's feature availability towards description. We have chosen curvature because curvature is an invariant surface property that is not affected by the choice of the coordinate system, the position of the viewer, and the particular parameterization of the surface. In an attempt to define a measure for the variation we use information theory. The more likely the occurrence of an event, the lesser the information the event conveys. For example, consider a flat surface that has uniform curvature. The density of its curvature hence is a Kronecker delta function of strength one. The entropy of a Kronecker delta function is zero. This result implies that flat surfaces and spherical objects convey little or no significant shape information. We also note that spheres of different radii will also have zero shape information and argue that change in scale (radius) adds no extra shape to the object. Furthermore, as the variation in curvature increases, broader the density of curvature gets, higher the entropy and greater the complexity for shape description. The most complex shape hence would be a hypothetical shape that has randomly varying curvature at every point.

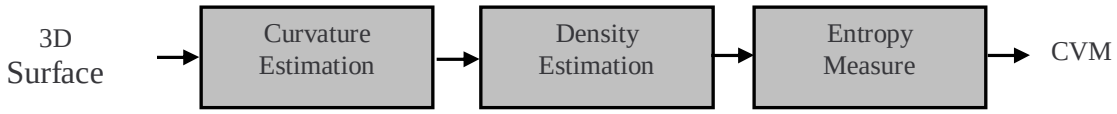


Figure 5: Block diagram of the CVM algorithm.

We formulate our CVM as shown in the block diagram in Figure 5 and discuss in detail the implementation issues on each of these blocks. We begin with the curvature estimation. For a given surface, the knowledge of curvature is sufficient information to define the shape. But surface curvature being an infinitesimal local feature lacks generalization to serve as a 3D feature in the computer vision context. Triangle mesh datasets such as the one that we discussed in Section 2 as the output of the laser scanners, are a piecewise smooth approximation to an underlying smooth surface. Several approaches to curvature estimation are summarized in [16].

As a tradeoff between computation speed and estimation accuracy, we have chosen the Gauss-Bonnet approach [17] to curvature estimation. This method uses the umbrella neighborhood of triangles immediately adjacent to a vertex to estimate the Gaussian curvature at that vertex. This method is also known as the loss of angle approach since we use the angles subtended by the edges emanating from the vertex of interest. If we consider a vertex v, then we can define the angle α_i for the corner of each triangle adjacent to v. From Gauss-Bonnet, we can estimate the Gaussian curvature, κ, of the underlying smooth surface at v as shown in Equation 1 where the summation is over the umbrella neighborhood and A is the accumulated area of the triangles around v.

$$\kappa = \frac{3}{A}\left(2\pi - \sum \alpha_i\right) \tag{1}$$

Now that we have computed curvature at each vertex of a mesh, our next step is to estimate the density function of curvature over the entire mesh. The simplest and perhaps the most common density estimation technique is the histogram, which Dorai and Jain [2] use for their shape spectrum. Although histograms are straightforward to implement, they do not provide accurate estimates. Additionally, histograms require selection of an origin and an appropriate bin width. Dependency on these user parameters reduces the confidence of our results when we later attempt to compute entropy from the histograms. To achieve a more robust solution, we make use of kernel density estimators (KDE) [18]. We use KDE as a tool to compute the density function p of the curvature values over the entire mesh. Consider Equation 2 where $\hat{p}$ is the estimate of p, n is the number of vertices in the mesh, h is the bandwidth of

interest, G is the kernel function and κ_i is the curvature at vertex v_i. We visualize KDE as a series of 'bumps' placed at each of the n estimates of curvature in the density space. The kernel function G determines the shape of these bumps while the bandwidth h determines their extent. With large data sets (n is large), the choice for G does not have a strong influence on the estimate. However, we recommend the Gaussian kernel although meshes provide large sample points:

$$\hat{p}(x) = \frac{1}{nh}\sum_{i=1}^{n} G\left(\frac{x-\kappa_i}{h}\right), \tag{2}$$

$$G(u) = \frac{1}{2\pi} e^{-\frac{u^2}{2}} \quad \text{such that} \quad \int_{-\infty}^{\infty} G(x)dx = 1. \tag{3}$$

The more significant parameter for accurate and stable estimation is not the kernel but the bandwidth h. Data driven bandwidth optimization approaches [19] such as the distribution scale methods, cross validation, L-stage, plug-in and advanced bootstrap methods theoretically aim to minimize the mean integrated square error between the actual density and the computed density. For our algorithm, we use the plug-in method for optimal bandwidth selection as this method provides useful results without the selection of user parameters. The plug-in method that uses Equation 4 for bandwidth selection reduces the computational overhead compared with commonly used cross validation methods.

$$h_{opt} = \left[\frac{243R(G)}{35\mu_2(G)^2 n}\right]^{\frac{1}{5}} \hat{\sigma}, \tag{4}$$

where $R(G) = \int G(t)^2 dt, \mu_2(G) = \int t^2 G(t) dt$ and $\hat{\sigma}$ is the absolute deviation of the curvature data κ_i.

We have used curvature estimates at each vertex to generate probability distribution curves. With these curves, we now formulate an information theoretic based on entropy to define the CVM. We argue that the amount of information that the curvature conveys about the surface can be quantified as a measure using Shannon's framework [15]. His definition of entropy for communication systems as a measure of uncertainty is the minimum number of bits required to encode a message. We do not apply Shannon's definition of entropy in that context but use the logarithmic relationship to measure the predictability of curvature considered as a random variable. Furthermore, we also have to deal with resolution and normalized measure space. The resolution of the mesh (the number of vertices n) decides the quantization of curvature. More the number of vertices, better the approximation to the infinitesimal curvature. We counter the asymptotic exponential interaction between the resolution and curvature by normalizing the Shannon's entropy measure as shown in Equation 5.

$$CVM = -\sum \hat{p}_i \log_n \hat{p}_i \tag{5}$$

The CVM of a 3D surface is the resolution normalized entropy of curvature that is indicative of the visual complexity associated with the surface. A high magnitude of the CVM specifies details of interest (feature availability) towards describing the object while small values of the CVM correspond to smooth flat surfaces. We use this definition of the CVM as a surface feature for object description that we describe in Section 4.

4. OBJECT DESCRIPTION USING THE CVM

We suggest an object description pipeline that first segments an object into reasonably sized patches and then specify the CVM measure for these patches. Our heuristic approach is to segment patches along crease discontinuities as these features should be excluded from the density estimates. Also, with man-made automotive components we note that the assumption about the existence of sharp edges can be justified. We use a simple region-growing method similar to [20] to segment the object along boundaries of high curvature that correspond to creases in the triangle mesh representation. During segmentation, we maintain patch adjacency information to create a graph where the nodes of the graph represent each segmented patch and the edges represent the connectedness of patches. The patches along with their orientation describe the global shape of the object while the CVM lends to describing the local variation within each of these

patches. The graph description of the 3D object will serve as the basis of object detection in a scene acquired from the robot. Similar objects will have similar graph description. Hence, objects with similar surfaces can be easily identified. In addition, the graph description serves as common platform for identifying objects in the partial data from the robot to the apriori manufacturer's CAD model.

We begin our experiments with 3D CAD models of automotive components such as the muffler, and the catalytic converter. In Figure 7, we show the mesh model as an assembly of segmented patches and its graph definition with the CVM describing each of the segmented patches.

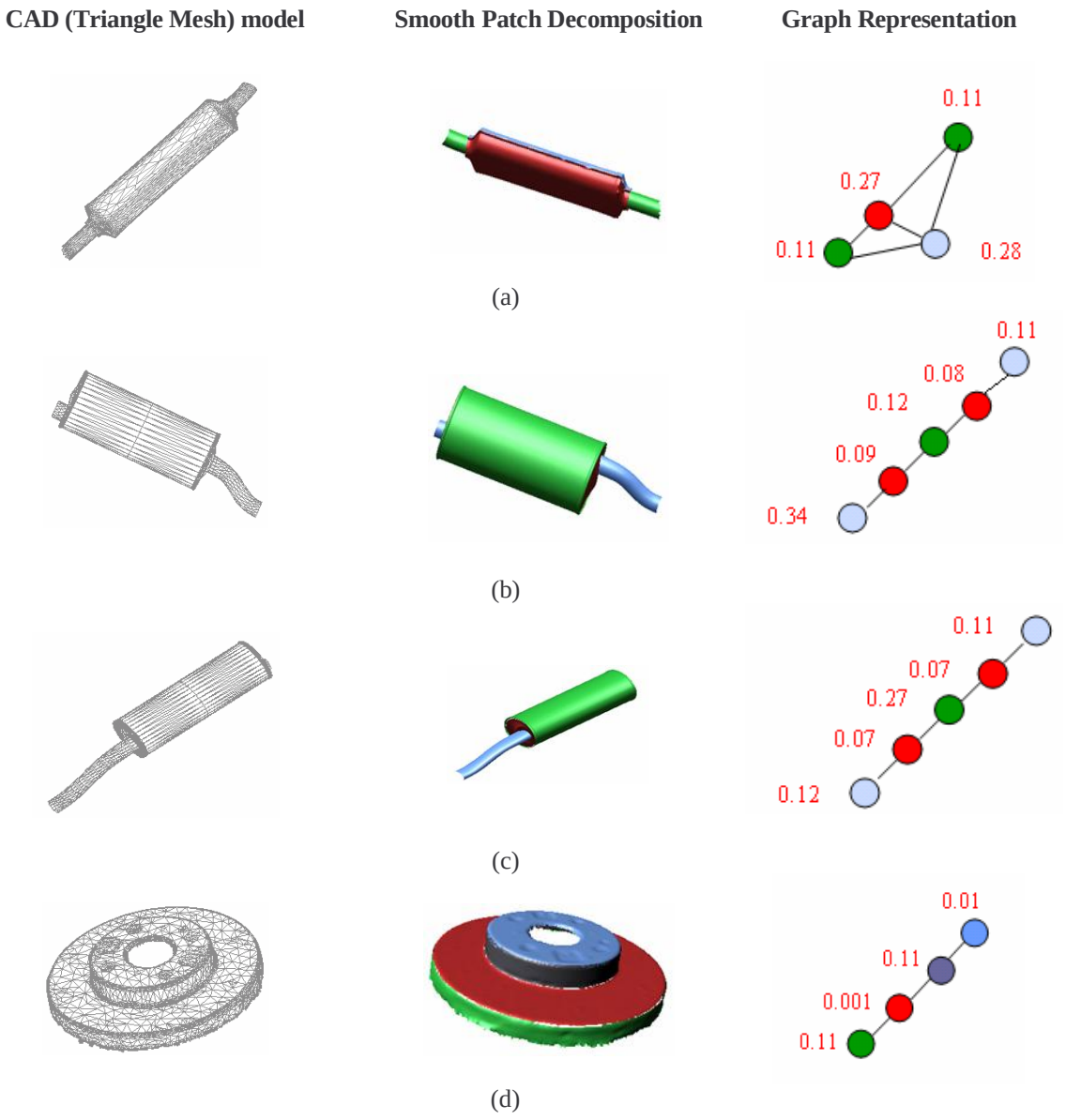
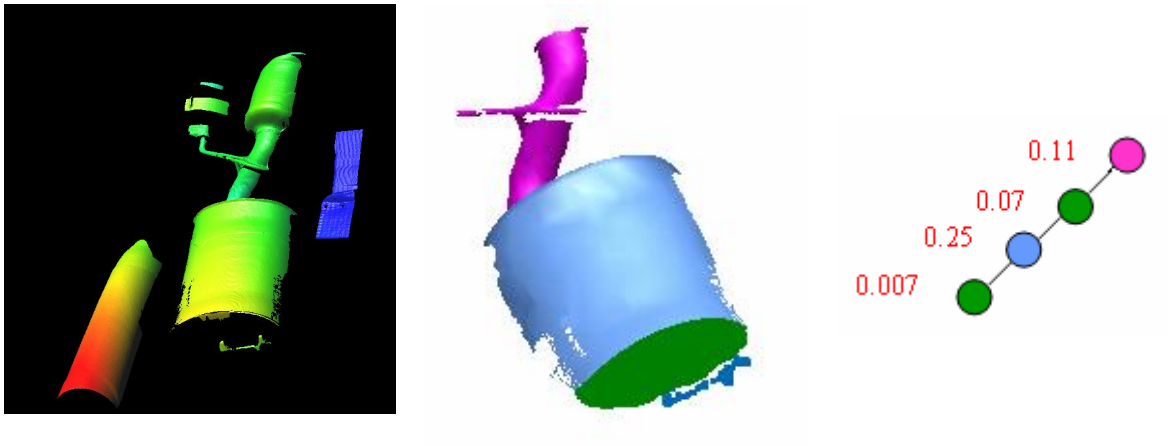


Figure 7: CVM graph results on automotive parts: 3D mesh model, smooth patch decomposition and graph representation of (a) catalytic converter, (b) Toyota muffler and (c) Dodge muffler (d) disc brake.

The graph that is the assembly of smooth patches also indirectly encodes the information about the topology of the patches. We note the descriptiveness of the CVM with a negligible value on almost flat surfaces and substantial shape information in certain descriptive unique patches. We also note that the deformed cylindrical surfaces of the disc brake and the muffler and its connecting pipes have close CVM's indicating the common underlying shape. On the other hand, we also observe the difference in the main body of the Toyota muffler and the Dodge muffler. Though both of them appear cylindrical, we are able to distinguish between the deformed cylinder in the Dodge muffler and the Toyota muffler. With these results, and the real data from the laser scanner, we now present results on the data from the IVP range scanner in Figure 8. We see that the muffler from the sensed data has components that match closely with the Dodge muffler's CAD model. At this point we disclose that the segmentation of the muffler from the real scene requires manual intervention because of the incomplete nature of the 3D data. Our future research efforts will target automating the scene analysis.



(a) (b) (c)

Figure 8: CVM graph results on an under vehicle scene. (a) 3D under vehicle scene. (b) Segmented muffler model. (c) Shape description of the muffler.

5. CONCLUSIONS

With focus on under vehicle inspection, we have collected 3D data by mounting laser range scanners on mobility bricks (robots). We have modeled the data as triangle meshes that are a common representation of digitized geometry used in the CAD and solid manufacturing process. On these meshes, we have proposed the CVM as a novel 3D surface feature and demonstrated the CVM towards patch-based object description. We do not claim contributions in terms of recognition as yet; but are encouraged by the results in this paper to test the recognition pipeline. Shapiro and Stockman [21] suggest two paradigms that use part (region) relationships to move away from a geometric definition of an object to a more symbolic one. Our algorithm benefits the creation of such a symbolic graph representation from a mesh representation. In addition, we plan to incorporate important scale information to support the scale-invariant CVM in the graph. We also would like to perform rigorous experiments on the effect of segmentation and graph matching before we claim the confidence and robustness required for a model-based threat object detection system.

ACKNOWLEDGEMENTS

We would like to thank Umayal Chidambaram, Santosh Katwal for their time and efforts in helping us with the data acquisition.

This work is supported by the University Research Program in Robotics under grant DOE-DE-FG02-86NE37968, by the DOD/RDECOM/NAC/ARC Program, R01-1344-18, by the US Army under grant Army-W56HC2V-04-C-0044, and by FAA/NSSA Program, R01-1344-48/49.

REFERENCES

1. D. Page, Y. Fougerolle, A. Koschan, A. Gribok, M. Abidi, D. Gorsich, and G. Gerhart, "SAFER Vehicle Inspection: A Multi-Modal Robotic Sensing Platform," in the Proc. SPIE Unmanned Ground Vehicle Technology VI, vol. 5422, Orlando, FL, USA, pp. 549-560, 2004.
2. C. Dorai and A. K. Jain, "COSMOS - a representation scheme for 3D free-form objects", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19,1115-1130,1997.
3. D. V. Vranic, "An Improvement of Rotation Invariant 3D Shape Descriptor Based on Functions on Concentric Spheres", in the Proc. of the IEEE International Conference on Image Processing, pp. 757-760, 2003.
4. G. Hetzel, B. Leibe , P. Levi, and B. Schiele, "3D Object Recognition from Range Images using Local Feature Histograms", in the Proc. of IEEE International Conference on Computer Vision and Pattern Recognition, pp. 394-399, 2001.
5. R. Ohbuchi, T. Minamitani, and T. Takei, "Shape Similarity Search of 3D Models by using Enhanced Shape Functions", in the Proc. of Theory and Practice in Computer Graphics, pp. 97-104, 2003.
6. S. Belongie, J. Malik, and J. Puzicha, "Shape Matching and Object Recognition Using Shape Contexts", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, pp. 509-522, 2002.
7. M. Körtgen, G. J. Park, M. Novotni, and R. Klein, "3D Shape Matching with 3D Shape Contexts", in the Proc. of the 7th Central European Seminar on Computer Graphics, 2003.
8. A. Khotanzad and Y. H. Hong, "Invariant image recognition by Zernike moments", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 12, pp. 489-497, 1990.
9. R. M. Duda and P. E. Hart, "Pattern Classification", John Wiley and Sons, New York, 1973.
10. G. Cybenko, A. Bhasin, and K. Cohen, "Pattern recognition of 3D CAD objects", Smart Engineering Systems Design, vol. 1, pp. 1-13, 1997.
11. A. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3D scenes", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, pp. 433-444, 1999.
12. D. Zhang and M. Hebert, "Harmonic Maps and Their Applications in Surface Matching", in the Proc. of IEEE Conference on Computer Vision and Pattern Recognition, vol. 2,pp. 524-530, 1999.
13. P. Besl, "Triangles as a primary representation: Object recognition in computer vision", in Lecture Notes in Computer Science, pp. 191-206, 1995.
14. R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin, "Shape Distributions", ACM Transactions on Graphics, vol. 21, pp. 807-832, 2002.
15. C. E. Shannon, "A mathematical theory of communication", The Bell System Technical Journal, vol. 27, pp. 379-423, 1948.
16. T. Surazhsky, E. Magid, O. Soldea, G. Elber, and E. Rivlin, "A comparison of gaussian and mean curvature estimation methods on triangle meshes", in the Proc. International Conf. on Robotics and Automation, (Taiwan), pp. 1021-1026, 2003.
17. C. Lin and M. J. Perry, "Shape description using surface triangulation", in the Proc. of the IEEE Workshop on Computer Vision: Representation and Control, pp. 38-43, 1982.
18. B. W. Silverman, "Density Estimation for Statistics and Analysis", London, UK: Chapman and Hall, 1986.
19. M. P. Wand and M. C. Jones, "Kernel Smoothing", London, UK: Chapman and Hall, 1995.
20. A. P. Mangan and R. T. Whitaker, "Partitioning 3D surface meshes using watershed segmentation", IEEE Transactions on Visual Computer Graphics, vol. 5, pp. 308-321, 1999.
21. L. G. Shapiro and G. C. Stockman, "Computer Vision", Prentice Hall, Inc., Upper Saddle River, NJ, 2001.