PRESERVING TOPOGRAPHY IN 3D DATA COMPRESSION FOR SHAPE RECOGNITION

Reda Ezzat Fayek
Systems Design Engineering
University of Waterloo
Waterloo, Ontario N2L 3G1, CANADA
reda@watnow.uwaterloo.ca

Commision III, Working Group 3

KEY WORDS: Object Modeling, Recognition, Surface Triangulation, DEM/DTM Model.

ABSTRACT

This paper presents a powerful interaction of two 3D object modeling concepts with the goal of enhanced 3D object recognition from dense sensory range information. The first concept is the careful and subjective coarsening of a triangular mesh representation of the sensory data to reduce the complexity of the model matching during recognition. For this purpose, we use an iterative mesh decimation approach which reduces the number of vertices and triangles in the mesh while preserving its important features. The importance of a particular vertex or edge in the triangular mesh is determined using topographic concepts and local surface properties. This process is repeated to produce meshes at various controlled resolution levels of the same scene. This technique finally produces a coarser mesh with the same topography as the original data. Typical reductions in mesh complexity are about 90%. The other concept is the use of generic symbolic 3D object models. Such a model describes invariant relations between sections of the object's bounding surface and, therefore, is invariant to translation, rotation, scale variations and partial occlusion. The 3D object recognition task is therefore reduced to matching such an object model against subsets of the reduced meshes resulting from the coarsening phases. This model matching procedure is greatly simplified due to the abstraction capability of the coarsening technique and the flexibility of the generic models. We provide experimental results demonstrating the strength of this interaction.

1 INTRODUCTION

Triangular meshes have been frequently used for modeling general 3D curved surfaces (Boender *et al.*, 1994, Cheng *et al.*, 1988, Fua & Sander, 1991, García, 1994, George, 1991, Schroeder *et al.*, 1992). This spatial data representation is attractive for its simplicity and flexibility. The main advantage in using triangular meshes for modeling complex 3D surfaces and objects is that they can adapt to fit an arbitrary surface with any desired accuracy or abstraction. Nonetheless, the extensive usage of triangular meshes in modeling has been mostly confined to special-purpose visualization or scientific computation tasks (Rippa, 1992, Schroeder *et al.*, 1992). This was mainly due to concerns about storage, accuracy, and combinatorics. Recently, however, with the rapid increase in computing and storage capabilities made affordable, it is becoming more reasonable to adopt this representations for most 3D modeling applications. This is also encouraged by the frequent use of 3D sensors yielding massive amounts of data such as photogrammetry, stereo-imaging and laser scanners. Modeling such large data sets is not feasible without an abstraction and compression of the available huge data.

Furthermore, although parametric and analytical representations of curved surfaces (*e.g.,* NURBS) may be easier in design, triangular meshes are more suitable for modeling sensory 3D data. While mathematical manipulations of analytical representations facilitate the iterative design revisions of *generated* 3D surfaces (*e.g.,* mechanical parts, automotive design, artistic design, etc.), they provide no support for the automated symbolic reasoning about such 3D objects. In contrast, acquired 3D sensory data need to be effectively modeled in order to facilitate computer-based 3D scene interpretation. Triangular meshes provide both aspects of surface modeling, *i.e.,* accurate surface description from scattered points and support for symbolic, feature-based scene interpretation.

We use a general-purpose 3D representation based on irregular triangular meshes (ITM) (Fayek & Wong, 1994) for the modeling of large amounts of spatial sensory data expressed by digital elevation maps (DEM). The main advantage of these is the high 3D data compression which can be achieved while preserving surface topography (Fayek & Wong, 1995). Compression ratios of over 90% are achieved with different types of sensory range data yielding 3D models at various resolutions. Arbitrary generic 3D object models are defined and used to recognize objects and structures irrespective of the chosen resolution level. Experiments with range data from remote sensing, laser scanning and stereo-imaging were performed with consistent results.

2 TOPOGRAPHIC COMPRESSION OF TRIANGULAR MESH SPATIAL DATA

We are mostly concerned with the modeling of sensory range data for practical applications. This is opposed to the modeling of synthesized 3D models for visualization and animation. While model accuracy and intricate details are the main issues in the later research area, our main focus is in the abstraction and interpretation issues of such scenes. Therefore, while most of the work done on triangular mesh modeling is towards mesh refinement and reduction of approximation errors, our work tackles the mesh coarsening while preserving the topography; which is the reverse problem.

The strength of our approach in comparison with previous work (Chew, 1993, Rippa, 1992) lies in our efficient compression mechanism. More precisely, at each compression phase, the goal is to ensure the high quality of the preserved data and the insignificance of the discarded data. We achieve this target by the subjective selection of the information to preserve and the removal of all other redundant information. The majority of work done in triangular mesh modeling of 3D data, however, follows the following general pattern: *(i)* preselect a certain approximation error tolerance, *(ii)* start with an arbitrary initial triangular approximation of the data (possibly a very coarse one), *(iii)* incrementally insert additional points in the mesh to improve the mesh quality until the prescribed tolerance is met. This modeling technique is based on iterative mesh refinement to compress the available data. It is computationally expensive and requires a large number of iterations due to the small increments of mesh improvements associated with each step. Some recent techniques have been introduced (García, 1994) to avoid the excessively expensive global optimization steps required in traditional implementations of these techniques.

We adopt an alternate and arguably more efficient approach for 3D data compression based on topographic mesh coarsening. Previous triangular mesh coarsening techniques (Chen & Schmitt, 1994, De Floriani & Puppo, 1988, Schroeder *et al.*, 1992) repeatedly remove redundant data points, also improving the mesh quality by small incremental amounts at each step. In contrast, our technique relies on the automatic detection of important characteristic features in a triangulation of the initial data set. These features are compressed and preserved in the compressed mesh while all remaining points are candidates for deletion. Additional points are selected for inclusion in the coarser mesh to meet certain quality criteria. Finally, a constrained triangulation is built with all selected points and the compressed features.

Therefore, the initial step of triangulating the whole data set (which can be done in $O(n)$ time) is performed only once. Each subsequent compression iteration reduces the handled data by ratios depending on the imposed approximation accuracy. This is achieved by selecting vertices and edges in the initial mesh according to certain importance criteria based on the topographic features of interest in the modeled scene. For instance, points and edges of local curvature extrema (*e.g.*, peak, pit and saddle points and ridge and ravine edge segments) are important in describing the behavior of the 3D surface. Similarly, edges and vertices forming the boundary of the region are crucial to preserve in a coarser mesh. Our topographic mesh coarsening algorithm easily identifies such feature edges, and vertices belonging to them by simple analysis of the interfaces between adjacent triangles in the initial mesh. Connected chains of such feature edges are then coarsened by selecting key vertices and connecting these by single feature edges. Such coarsened topographic feature edges are embedded in the coarser triangular mesh together with the above mentioned topographic feature vertices by means of a constrained surface triangulation. Figure 1 illustrates an example of the coarsening of such feature edges. It starts from the triangular mesh representation of the object (Figure 1-*top left*), with the detected feature edges highlighted, to the identification and connection of the key feature vertices (Figure 1-*top right*), to the coarsened features (Figure 1-*bottom left*) and finally to the coarser mesh preserving those features (Figure 1-*bottom right*).

Although the approach of selecting important vertices in range data and embedding these in a triangular representation of the 3D scene was used by others (Zheng & Harashima, 1994), our emphasis on the topographic feature edges ensures the validity of the resulting mesh. If no such provision is made, the selected vertices can be triangulated with a variety of meshes not preserving the topographic features. Such a topographic coarsening process

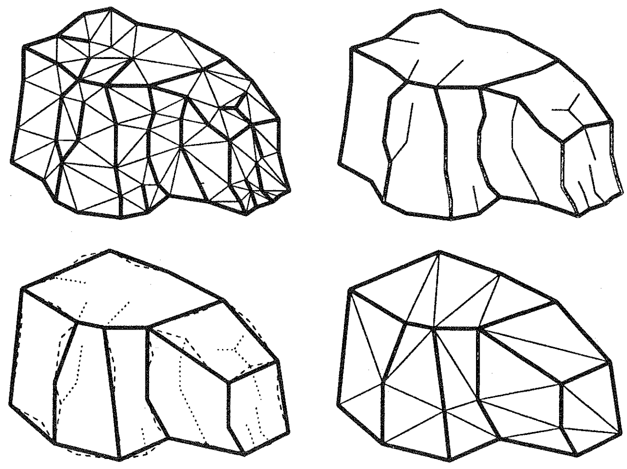


Figure 1: Coarsening of detected topographic feature edges. Original triangular mesh representation of an irregular 3D scene with detected feature edges ***(top left)***, feature vertices and connected chains of feature edges identified ***(top right)***, coarsened feature edges ***(bottom left)*** and coarser topographic triangulation of the same scene ***(bottom right)***.

can be further applied to the resulting coarser mesh to get even coarser representation of the 3D scene. Therefore, this technique acts as an abstraction mechanism for the original large set of sensory range data. The incremental coarsening steps are fully controlled through a user defined parameter.

We have implemented this technique and tested it with various data sets. Typical reduction figures for a single coarsening iteration are in the range of 50% to 70% for both real sensory data and simulated scenes. This is particularly true during the first few iterations performed on typical data sets of sensory range information; these are dense and possibly noisy scattered range measurements. Clearly, with progressive coarsening, these figures slowly decrease until they reach a minimum value when all the remaining vertices and edges in the mesh are crucial to preserve the topographic characteristics of the 3D scene. This approach, therefore, has many advantages among which we cite: *(i)* substantially high data compression ratios are achieved at each step, *(ii)* the compression preserves the topographic features of the scene, eliminating points contributing trivial and redundant information while keeping interesting points providing accurate scene description, *(iii)* multi-resolution versions of the data are inherently supported, and *(iv)* direct control over the mesh approximation accuracy.

3 MODELS FOR 3D OBJECT RECOGNITION

The problem of object recognition from sensory data invariably requires the comparison of an object model with the acquired data. Several types of models can be used ranging from explicit and rigid object models to flexible and symbolic ones. The modeling primitives and operators used in these models also vary in simplicity. The lack of a conceptual object model in techniques relying mostly on local operators prevents them from fully characterizing objects in the real world. Therefore, effective object recognition must incorporate model knowledge or context. Furthermore, the objects characteristics need to be unambiguously and completely encoded in the model. Otherwise, the task of extracting plausible model instances from the sensory data becomes a major undertaking that often requires complex symbolic interpretation. Strategies that support flexible models are best adapted to situations in which an initial guess for a model instance can be easily detected. The effectiveness of such methods depends strongly on the appropriateness of the modeling primitives used to initiate an object instance hypothesis. Therefore, methods based on flexible models are most adequate when well-known object models, described by a limited number of constraints, are available. Nonetheless, complex scenes may require models with multiple components; these components typically need to be symbolically combined in order to find a practical strategy for finding the optimal solution. These methods can work well in the presence of incomplete or noisy data, provided natural limitation on the size of the search space can be imposed (Suetens *et al.*, 1992).

For the purpose of object and structure recognition in 3D sensory data, we use generic 3D structural models. Each such model is expressed as a collection of conditions to be satisfied by a candidate subset of the data in order to be identified as an instance of the desired object. These conditions describe various properties of the object's structural components such as size, shape, connectivity, location, and orientation. These models are based on the abstraction of the object's bounding surface by a set of contiguous nearly-planar patches. For instance, a simple 3D cube would be trivially modeled by its six faces. While this technique clearly favors structured polyhedral objects, it is flexible enough to accommodate highly unstructured 3D scenes and forms if appropriate patches are used in the 3D object model. Figure 2 shows an illustrative example of a building model. Starting from a triangular mesh representation of the object of interest (Figure 2-*a*), such nearly-planar patches are easily extracted. They are characterized by the closures of topographic feature edges bounding them. Once identified, several numeric properties of such patches are readily computed such as surface, shape contour, normal direction and curvature behavior (*e.g.*, monotonous, oscillating, etc.). From these properties, several relations between the neighboring features are computed and used to characterize the boundary of the object of interest. A labeling of these patches, directed by domain knowledge and context information is used to complete the model (Figure 2-*b*).

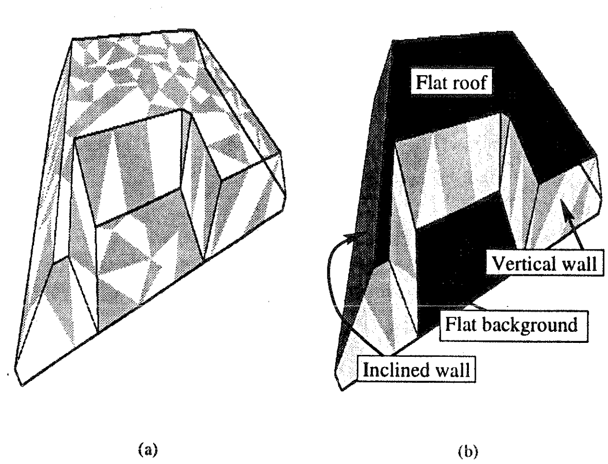


Figure 2: A simple 3D model of a building using nearly-planar patches. Topographic triangular mesh of an object *(left)* and its nearly-planar patch model *(right)*.

Due to the typical noise levels associated with real-world sensory data, parametric conditions are used as opposed to rigid constraints which would yield only exact matches to the fixed models. We, therefore, adopt a hypothesis generation and verification with relational, feature-based object model matching. Since our models are only structural parametric ones, they have minimal dependence on the resolution of the data. This is a direct result of the flexible symbolic and relational nature of the object models thus constructed. Additionally, the model is based on consistent relations between important patches on the object's boundary; these are preserved in various representations and resolution levels of the scene. Hence, such model is invariant to translation, rotation, scale, texture and partial occlusion. Therefore, objects can be identified at all intelligible resolution levels, yet more accurately at high resolutions. This allows fast and efficient identification of objects of interest in highly compressed data forms. This capability of flexible object recognition at any level of details is a desirable feature particularly with large amounts of data typical of dense range images since the size of the original data sets prevents any reasoning operations. However, when objects and structures are recognized at a certain compressed representation of the scene, they can be further selectively refined, if necessary, using local operations in higher resolution meshes.

This technique was fully implemented using a portable standard language, C, and visualization was performed using a general-purpose graphics library, OpenGL, thus making it fully portable to various computing platforms. Experiments involved substantially different types of 3D data: *(i)* natural terrain data sampled at 10 m intervals covering $10 \times 20\ km^2$ around Toronto, Ontario, Canada, *(ii)* suburban area of $240 \times 240\ m$ featuring 17 buildings and *(iii)* various sets of 3D range data using a simulated range acquisition sensor of indoor scenes cluttered with objects. The same feature detection and scene segmentation techniques were used for all experiments. Only the application-dependent generic models are different according to the objects to recognize. In the following two sections, we provide two experimental results that illustrate the behavior of this technique with various types of real-world sensory range data sets.

4 MODELING OUTDOOR NATURAL TERRAIN AND PLANETARY SCENES

The input to our system consists of sensory elevation data acquired by satellite remote sensing and various other photogrammetry techniques. The data is typically stored in the commonly used Digital Elevation Map (DEM) format. However, the large amounts of raw sensory data of such outdoor scenes hinder their effective use in reasoning and planning. Nonetheless, the prohibitively high cost of sensory information processing can be reduced by an appropriate compression technique which preserves the main characteristics of the scene. The irregular spatial distribution of the topographic features discourages uniform sampling of the terrain. Therefore, we approximate the acquired data by irregular triangular meshes. We also use our topographic mesh coarsening algorithm which preserve important terrain features. Such features are crucial in describing the scene and for various symbolic reasoning tasks such as autonomous visual navigation, path planning, geographical information systems (GIS) applications, site selection and many more.

Using the nearly-planar patches as modeling primitives, the detected local surface topographic features can be used to automatically segment the region into collections of nearly coplanar triangular patches. These patches are grouped using generic models to describe interesting, more abstract global scene features (*e.g.*, hills, valleys, mountains, plains, etc.) which provide a more abstract representation of the scene suitable for various reasoning tasks. The original DEM data consists of 1200×1200 elevation measurements sampled regularly at $1\ m.$ intervals around the West-most section of Lake Ontario. In Figure 3, the American shore is on the left side of the figure and the Canadian shore is on the right side. Figure 3 *(top)* illustrates an initial regular sub-sampling of the original triangular mesh representation of the above mentioned scene (uniquely for experimental reasons). The topographic mesh coarsening and scene feature detection and grouping reduces the storage requirements by several orders of magnitude. The resulting mesh is irregular in nature with more points concentrated around interesting regions with high feature density and much less points in flat regions.

Adopting the nearly-planar patches described above as modeling primitives, this scene was subsequently reduced to approximately 40 such patches. Generic models can be constructed for various global scene features to detect important symbolic entities. Figure 3 *(bottom)* shows the symbolic scene features identified using generic models based on collections of nearly-planar patches. Several topologic representations of this symbolic scene description can be formed (*e.g.*, topologic graphs, entity-relationship diagrams, etc.) to support practical symbolic reasoning and planning tasks.

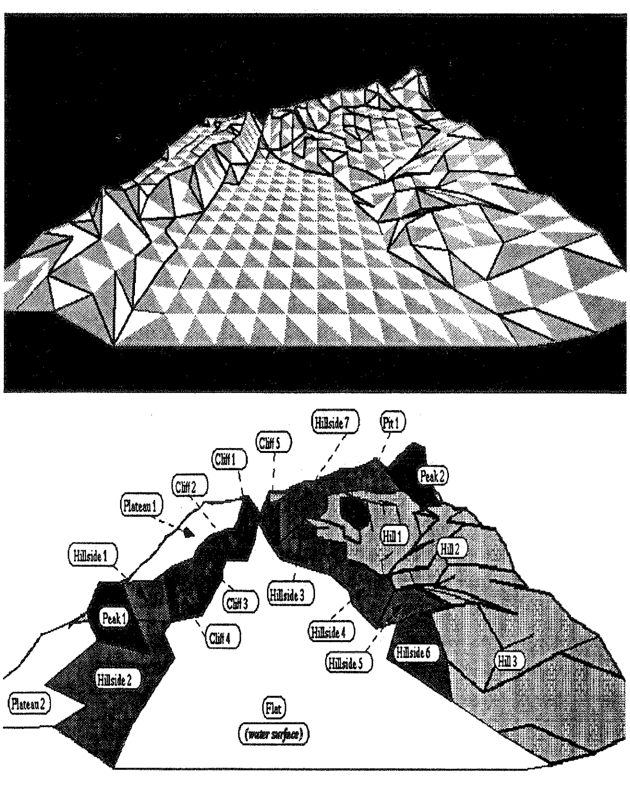


Figure 3: A triangular mesh representation of the terrain *(top)* with the detected nearly-planar patches *(bottom)*.

5 RECOGNITION OF MAN-MADE STRUCTURE IN AERIAL IMAGES

For the experimental validation of our terrain model, we use a set of geographic elevation data captured by aerial imaging and covering an area of roughly 240×240 meters with a one meter resolution. This sensory data covers a plain with several buildings of similar heights. For the experiments, we select subregions with various resolutions and sizes, and therefore, we generate test data covering a wide variety of scenes. For simplicity and without loss of generality, we use rectangular subsections of the terrain for the individual experiments. This choice does not bias nor affect the validity of the results. Figure 4 depicts the original range image of the entire region. The grey-scale format (Figure 4-*top*) is such as that the darker the pixels the higher their corresponding elevations. It also illustrates the detected houses (Figure 4-*bottom*) in a triangular mesh representing the full details of the scene.

Our building model is expressed in terms of a set of nearly-planar patches as described earlier. Some of these (corresponding to side walls) are far from horizontal and enclose other raised patches (corresponding to roofs) which are close to horizontal. Such a flexible model is very generic and able to extract numerous other objects such as buses and vans if they exist in the aerial scene. Therefore, we use some domain knowledge and context constrains on the object's dimensions to exclude spurious objects. If we were to recognize outdoor vehicles using this generic model, only the parameters of the constraints defining the model need to be changed.

We use a single generic model covering houses, apartment buildings, and large hangars. When such a structure is recognized, a set of derived parameters (*e.g.*, height, area of floor plan, volume, area of enclosing surface) are computed. They are used to distinguish the different objects using domain-specific knowledge (*e.g.*, the average size of a house compared to a high apartment building). The derived parameters are also used to reconstruct the detected objects' geometry if required. Table 5 provides the obtained results for the scene in question. The house labels are consistent with those labels in Figure 4-*top*. The values reported here were computed using the high resolution mesh sampled at 2 meters intervals. The center of gravity (C.O.G) of each detected house is reported in meters with respect to the origin at the top left corner in Figure 4. Surfaces are given in square meters.

We verified the capability of our man-made structure recognition system at various resolution levels of the available sensory range data. Therefore, we applied our topographic mesh coarsening algorithm mentioned earlier to the original dense mesh representing this scene. After several coarsening iterations, the number of vertices and triangles representing the same scene decreases significantly. The coarser mesh was then used to identify the same houses using the same model described above. The houses were detected with a high accuracy (within the allowable errors in mesh coarsening) and yielded almost identical results to the houses detected in the original dense mesh. Figure 6 illustrates a close-up to the coarse mesh in the neighborhood of the houses labeled 10 and 11 in Figure 4. It is clear from the table at the bottom of Figure 6 that the results obtained from such a coarser mesh are nearly identical to those obtained from the original high resolution mesh. The maximum error in the location of the house's center of gravity is 0.2 m only. This amounts to about than 0.083%.

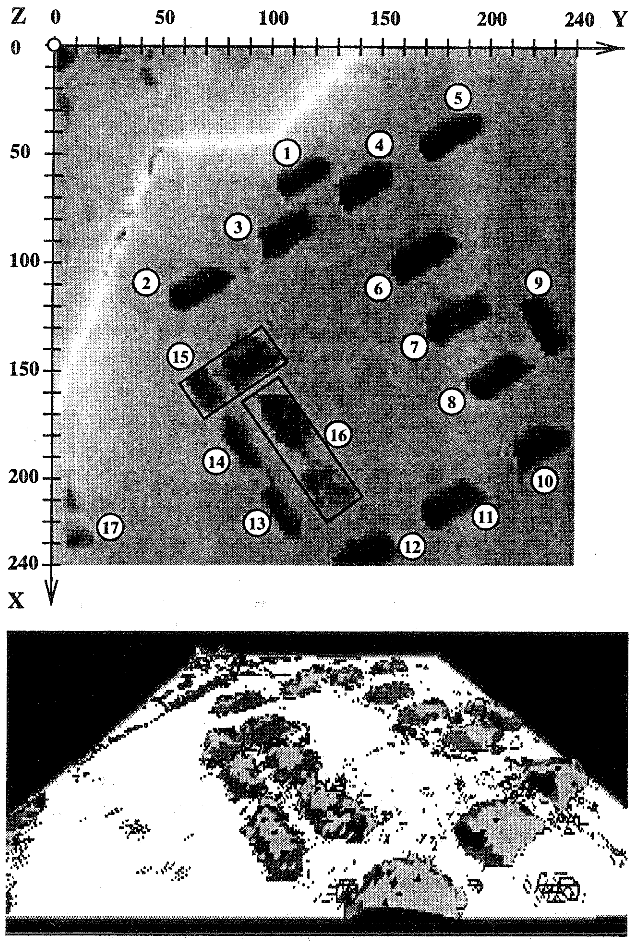
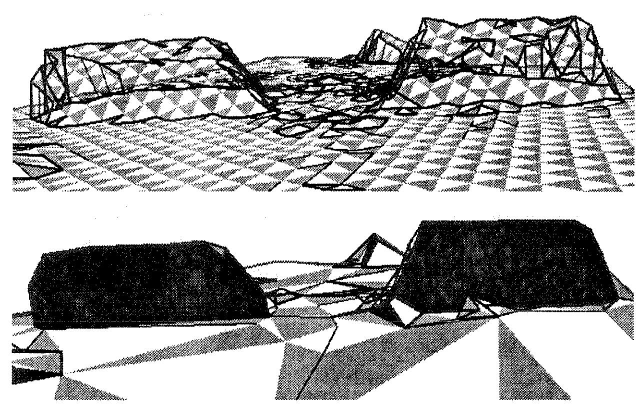


Figure 4: Houses in a suburban area: original dense range image *(top)*, detected houses in the triangular mesh *(bottom)*. [120×120 points, sampled at 2 m.]

#	Patches	Outside surface	Floor area	C.O.G.
1	51	618.9	409.9	(60.8,110.8,103.3)
2	59	825.7	512.5	(111.3,64.5,104.1)
3	57	818.9	511.6	(86.7,103.1,104.7)
4	76	835.0	514.0	(65.0,137.6,104.7)
5	47	860.1	535.1	(42.4,176.8,105.2)
6	78	929.3	596.1	(97.1,164.9,105.1)
7	80	932.6	634.2	(125.0,180.5,105.0)
8	90	895.4	590.0	(150.8,198.7,105.0)
9	85	821.3	534.7	(128.0,220.5,105.1)
10	77	915.7	604.4	(183.8,217.1,105.1)
11	63	885.7	588.4	(209.3,179.0,105.3)
12	58	631.2	436.3	(229.5,141.9,105.4)
13	98	609.7	400.3	(212.1,101.7,103.6)
14	91	617.3	415.9	(181.8,83.3,103.7)
15	165	1510.9	997.7	(149.7,79.7,104.4)
16	218	1599.0	1065.8	(183.7,111.2,104.5)
17	81	533.6	366.2	(218.0,9.1,101.4)

Table 5: Extracted properties of houses in a suburban area.



#	Patches	Outside surface	Floor area	C.O.G.
10	83	915.9	603.5	(183.8,217.1,105.1)
11	62	884.3	592.1	(209.3,178.9,105.3)

Figure 6: Extraction of houses: close-up in the original triangular mesh *(top mesh)* and the same detected houses in a coarser mesh with 82% less vertices *(bottom mesh)*.

From the experimental results described here we demonstrate that we could exploit our mesh topographic coarsening to reduce the computational cost of recognizing objects in complex 3D scenes using generic models based on nearly-planar patches. These results show that our triangular meshes, our topographic mesh coarsening, together with our generic symbolic object models, can be successfully used in the problem of 3D object recognition from real sensory data. Such data, gathered by remote sensing and photogrammetry techniques, is typically imperfect. Our approach yields accurate results while reducing the recognition complexity. This is possible through intelligent data compression and filtering by exploiting the triangular mesh topographic coarsening. If fast object recognition and identification is desired, the coarsened meshes are adequate. Their results can be used to initiate objects seeds in higher resolution meshes. For accurate object recognition, the high resolution meshes are used. If more accuracy is desired, the obtained results can serve as very good initial guesses for techniques using combined range and intensity information. It is our future goal to investigate this combination for precise 3D model synthesis from sensory data and reverse engineering of CAD models.

6 CONCLUSIONS AND FUTURE WORK

We presented a technique for the compression of huge sets of 3D sensory data. Unlike more general brute force data compression approaches, our technique is specifically concerned with the preservation of topographic features of the 3D scenes for the particular problem of shape recognition. For this purpose, we use an irregular triangular mesh representation of the scenes. From these, we identify and extract important topographic surface features which are preserved in coarser meshes representing the same scenes albeit at lower resolution levels. The topographic coarsening can be repeated several times with controlled steps preserving the same features. This method is both more computationally attractive and more informed than mesh refinement techniques. It simplifies the analysis of massive amounts of sensory range data that would, otherwise, be difficult to use.

At any resolution level, the topographic features can be used to segment the triangular mesh representing the scene into nearly-planar patches. Meanwhile, we also define generic models of 3D objects based on invariant relations between sections of the surface boundary of our objects of interest. These constructed models can be easily compared to the coarsened meshes of the 3D range data. Applications of this approach range from interpretation of remote sensing and photogrammetry data, digital terrain modeling (DTM), medical imaging, outdoor autonomous vehicle navigation, planetary terrain mapping, geographical information systems (GIS), CAD model reverse engineering.

In this context, we have identified several possible scenarios of interaction between triangular mesh modeling and coarsening for the purpose of fast and accurate object recognition. Although more effort is required to build a library of generic object models such as the example we presented, the initial results are very promising and indicate the generality and flexibility of this approach. Future work will be focused on the intelligent integration of visual intensity data with the available range data. We maintain that 3D object recognition is best done using range information. Nonetheless, such additional information would greatly guide the object recognition phases by clarifying possible ambiguities in interpreting range data which tends to be more noisy than intensity data.

REFERENCES

(Boender *et al.*, 1994) Boender, E., and Bronsvoort, W.F., and Post, F.H., 1994. Finite-element mesh generation from constructive-solid-geometry models. Computer-Aided Design, 26, pp. 379-392.

(Chen & Schmitt, 1994) Chen, X., and Schmitt, F., 1994. Surface modelling of range data by constrained triangulation. Computer-Aided Design, 26, pp. 632-645.

(Cheng *et al.*, 1988) Cheng, K., Idesawa, M., and Soma, T., 1988. Analysis and manipulation methods of geographic information. In: International Conference on Pattern Recognition, Rome, Italy, pp. 897-900.

(Chew, 1993) Chew, L.P., 1993. Guaranteed-quality mesh generation for curved surfaces. In: Ninth Annual Symposium on Computational Geometry, pp. 274-280.

(De Floriani & Puppo, 1988) De Floriani, L., and Puppo, E., 1988. Constrained Delaunay triangulation for multiresolution surface description. In: Ninth International Conference on Pattern Recognition, Rome, Italy, pp. 566-569.

(Fayek & Wong, 1994) Fayek, R.E., and Wong, A.K.C., 1994. Triangular mesh model for natural terrain. In: SPIE Conference on Intelligent Robots and Computer Vision XIII: Algorithms and Computer Vision, Boston, MA, 2353, pp. 86-95.

(Fayek & Wong, 1995) Fayek, R.E., and Wong, A.K.C., 1995. Hierarchical model for natural terrain using topographic triangular meshes. In: Third IASTED International Conference on Robotics and Manufacturing, Cancun, Mexico, pp. 252-255.

(Fua & Sander, 1991) Fua, P., and Sander, P.T., 1991. From points to surfaces. SPIE Geometric Methods in Computer Vision, 1570, pp. 286-296.

(García, 1994) , García, M.A., 1994. Reconstruction of visual surfaces from sparse data using parametric triangular approximants. In: IEEE International Conference on Image Processing, Austin, TX, pp. 750-754.

(George, 1991) George, P.L., 1991. Automatic Mesh Generation: Application to Finite Element Methods. John Wiley & Sons, Inc.

(Rippa, 1992) Rippa, S., 1992. Long and thin triangles can be good for linear interpolation. SIAM Journal of Numerical Analysis, pp. 257-270.

(Schroeder *et al.*, 1992) Schroeder, W.J., Zarge, J.A., and Lorensen, W.E., 1992. Decimation of triangle meshes. Computer Graphics, 26, pp. 65-70.

(Suetens *et al.*, 1992) Suetens, P., Fua, P., and Hanson, A.J., 1992. Computational strategies for object recognition. ACM Computing Surveys, 24(1), pp. 5-61.

(Zheng & Harashima, 1994) , Zheng, W., and Harashima, H., 1994. The automatic generation of 3D object model from range image. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP-94). pp. 517-520.