ACM SIGGRAPH

SIS Conference Reporting

241 - Submission Reviews - By Question

Generated: July 14, 2009 02:23:52 EDT

Reviewer	Response
Submission Information:	
Id:	papers_0241
Title:	Lightweight 3D Modeling of Urban Buildings From Range Data
1) Briefly describe the paper and its contribution to computer graphics and interactive techniques. Please give your assessment of the scope and magnitude of the paper's contribution.	
Reviewer #60:	This paper presents an approach for reconstructing fairly light weight 3D models of urban buildings from a dense point cloud obtained from laser range scanners. The main idea in the proposed approach, is that the 3D model can be generated from multiple 2D cross-sectional contours of the building from a series of extrusion and tapering operations. The approach proceeds by partitioning the input point cloud (pre-registered range scans with principal orthogonal directions recovered) into horizontal slabs. 2D polygonal curve reconstruction is performed on the subset of points in a horizontal slab which are projected to a suitable 2D domain. Symmetry is utilized in this hole-filling step. Next adjacent contours called 'keyslices' are compared to detect portions of it that can be merged into an extrusion surface across keyslices and also tapered extrusion for non-vertical surface patches. Experiments are performed on a single dataset. The dense scan of a building consisting of 14 million points is converted into a compact 3D model while retaining the prominent geometric features in the building without over-simplification.
Reviewer #8:	The paper presents a new approach to reconstruct the geometric model of a building from dense laser-scan point data. The key idea is to partition the points into slabs (horizontal intervals), and project the points in each slab to form a 2D image. Symmetry detection and line-beautification are used to fill in missing data within each slab. Some of the slabs which the largest changes are identified as "keyslices". These keyslice contours are stitched together to form an output triangle mesh. Finally, special processing is used to detect angled planar roofs.
Reviewer #1:	The paper presents a light weight framework for generating models of urban buildings from range data. The work is based on the observation that many building can be modeled as a combination of extrusion and tapering. The algorithm treats the problem of model reconstruction for urban buildings as a 2.5D problem, divides the scan into a number of volumetric slabs, converts each projected layer into a polyline, and then reconstructs the model using a mixture of extrusion and tapering. I found the idea simple and interesting. The motivation is similar to generation of procedural models using facades (more later). The current method has some limitations regarding the types of buildings it can model. Conceptually it is similar to architectural modeling using floorplans [Yin et al. 2009], where the 'plans' are obtained by slicing the input model into vertical slabs.
Reviewer #78:	The paper describes a system to reconstruct building models with a low polygon count from range scans.
Reviewer #90:	This paper addresses a very important problem: convert an overwhelming amount of data from range scanning to lightweight polygonal models. The proposed method cut the range point set into slabs, project each slabs onto a 2D slice, vertorize the boundaries, extrude to form a model and refine by tapering. The proposed method is a step-by-step system. The assessment of contribution and major drawbacks are discussed below in "7) Explain your rating".
2) Is the exposition clear? How could it be improved?	
Reviewer #60:	Yes.
Reviewer #8:	The paper is generally well written. The video is excellent. Figure 12 (a) and (b) use different viewpoints, which is confusing. The correct definition of Hausdorff distance uses MAX, not SUM as in Equation 3. See http://en.wikipedia.org/wiki/Hausdorff_distance . The current d_H is reasonable but is not Hausdorff. line 257: "problem" is unclear. line 270-271: unclear.

Reviewer	Response
Reviewer #1:	The paper is fairly well written, and easy to follow. There are a few parts which are unclear, but Figure 2 already explains how the procedure works. The rest of the details are fleshed out later. In Equation (1), why are the slabs normalized? It seems to be just for convenience. This is probably more for ease of implementation, rather than any other reason. Section 4.2, the assumption about only 2D translations can severely constrain the approach. This will manage to find only symmetry aligned to the axes. In cases, like castles, etc. such an assumption can be easily violated (for example, rotated towers). Symmetry based hole filing may still leave significant gaps that can produce artifacts after BPA. In the present formulation, there is no coupling between adjacent slices. Using information (as prior) across adjacent slices can help in filling in large holes (due to occlusion, etc.). Can equation (3) lead to high values in case of missing points (in the slice for I_r), leading to additional slices? Figure 7 is unclear (from the caption). This part should be clearly described, though I think I understand what is being done (lines 251-265). Curvature information -- maybe we get this information directly from the ball radius being used (BVP). Figure 9b, there are some kinks which appears as non-manifold edges. How is this resolved during extrusion? It may be a good idea to try to preserve corners during merging of small lines. Also, instead of the proposed way of combining small segments, a least squares based fitting can be more stable. Figure 11: Is this the top view or the side view? Figure 13: This figure is hard to see (the deviation map). Lines 459-460, which model are simplified using qslim? For completeness it will be good to mention how to obtain a model is obtained directly from the point cloud (Figure 15b).
Reviewer #78:	The exposition is ok.
Reviewer #90:	Generally clear. No picture taken by normal camera is shown for the scene in Figure 12. People don't know what it is. An odeum?

3) Are the references adequate? List any additional references that are needed.

Reviewer	Response
Reviewer #60:	A comparison with Qslim was done but a variational method like Steiner et al 2004 will potentially do much better than Qslim. This method also will not have the horizontal bias that this method will probably exhibit (see discussion below). @inproceedings{1015817, author = {Cohen-Steiner, David and Alliez, Pierre and Desbrun, Mathieu}, title = {Variational shape approximation}, booktitle = {SIGGRAPH '04: ACM SIGGRAPH 2004 Papers}, year = {2004}, pages = {905--914}, location = {Los Angeles, California}, doi = {http://doi.acm.org/10.1145/1186562.1015817}, publisher = {ACM}, address = {New York, NY, USA}, }
Reviewer #8:	Image-based procedural modeling of facades P Müller, G Zeng, P Wonka, L Van Gool - ACM Transactions on Graphics, 2007 - vision.ee.ethz.ch That paper presents an alternative but related approach which uses images and user assistance to construct a mesh that has similar grid structure to that here. There is a large body of work on "surfaces from contours" -- see http://scholar.google.com/scholar?q=surfaces+from+contours . It appears quite relevant here since the slab slices are effectively contours.
Reviewer #1:	The following paper is very related and should be mentioned. @article{Mueller:2007:PMF, author = {Pascal Müller and Gang Zeng and Peter Wonka and Luc Van Gool}, title = {Image-based Procedural Modeling of Facades}, booktitle = {Proceedings of ACM SIGGRAPH 2007 / ACM Transactions on Graphics}, year = {2007}, volume = {26}, number = {3}, publisher = {ACM Press}, } Also there is some related work in CT scanning where body parts are modeled in a similar fashion using slices.

Reviewer	Response
Reviewer #78:	The references are ok in general. However, the paper does not do a good job in presenting, structuring, and recognizing the problems that are tackled. If the paper would do a better job in classifying the methodology and the problem, more papers will then have to be cited. I believe that the authors can go more in the simplification direction, or more in the fitting shapes to point clouds directions (see link below).
Reviewer #90:	Adequate, with minor problems, such as: What is qslim in Line 450-458? Add a citation? Surface Simplification Using Quadric Error Metrics, Michael Garland and Paul Heckbert, SIGGRAPH 97? If qslim is refer to the software of this paper, I would suggest the authors to compare with more recent approaches rather than this 12 year old paper. Confusing reference: direction rectification is done by [Liu et al. 2006] or [Liu and Stamos 2005]? Line 181-182 said [Liu and Stamos 2005], while the 5-th line of section 4.2 said [Liu et al. 2006], which contradicts each other.

4) Could the work be reproduced from the information in the paper? Are all important algorithmic or system details discussed adequately? Are the limitations and drawbacks of the work clear?

Reviewer	Response
Reviewer #60:	Yes, it should be fairly straight forward.
Reviewer #8:	The line beautification scheme in lines 320-327 could use additional detail; it was unclear to me. Would it be useful to prefer lines that are axis-aligned (with respect to the symmetry axis)? How exactly is the surface created between adjacent keyslices? Line 339 explains "the space between each pair of keyslices can be interpolated by the lower keyslice". Does this mean that there is no continuous join with the upper keyslice? Does this result in horizontal gaps in the surface, i.e. lack of watertightness? For the roof in Figure 10, we see that a horizontal surface is created at the top of the main building and below the 8 turrets. When is such a horizontal surface created, and how does it connect with the vertical surfaces? Is the resulting surface a manifold? All these issues are related to the "surfaces from contours" problem. In this work the "branching" case is handled by introduction of horizontal surfaces, which is interesting. This needs a lot more detail. Limitations are not explicitly discussed.
Reviewer #1:	Yes, the work can be reproduced with some experimentation, though several important details are missing. Additionally please see section 2. Limitations are not properly discussed, specially failure cases. I can see the method failing to take advantage of non-axis aligned symmetry. Also the segmentations into U_i-s can suffer when using simple axis-aligned boxes. More later.
Reviewer #78:	I think the work can be reproduces. The results are not sufficient and it is not really clear how well the method works compared to other ideas.
Reviewer #90:	Impossible to reproduce due to missing details. For example: Line 275: The combination of Hausdorff distance measurement and curvature inference... How to combine? Line 325: We integrate the BPA and HT methods by first applying a dilation operation on I using 8-connected neighbors to get the dilation image ... How does the result of BPA fit into this procedure? For the description, we can only know that HT used the dilation of the image I, which is the original image defined in Section 4.2. I guess that the author wants to define I as the result of BPA by removing outlier points? or by linking the points in the circular list. Please clarify this issue. Dilation by what kernel size? Line 379 As before, the 3D data points inside the range of HR are projected along both left-right (X axis) and face-inside (Z axis) directions. Then, the keyslice detection is carried out based on the Hausdorff distance similarity measurement for both directions. Which direction first? X or Z? Why? How to get the segmentation as shown in Figure 11? Experiment: Slices 2D image pixel resolution = ? How sensitive to image pixel resolution? More missing details are listed below.

5) Please rate this paper on a continuous scale from 1 to 5, where:
1 = Definitely reject. I would protest strongly if it's accepted.
2 = Probably reject. I would argue against this paper.
3 = Possibly accept, but only if others champion it.
4 = Probably accept. I would argue for this paper.

Reviewer	Response
5 = Definitely accept. I would protest strongly if it's not accepted. Please base your rating on the paper as it was submitted.	
Reviewer #60:	2.75
Reviewer #8:	3.0
Reviewer #1:	2.7
Reviewer #78:	2.9
Reviewer #90:	2.1
6) Please rate your expertise in the subject area of the paper on a continuous scale from 1 to 3, where: 1=Tyro, 2=Journeyman, 3=Expert.	
Reviewer #60:	2.0
Reviewer #8:	2.0
Reviewer #1:	2.4
Reviewer #78:	2.0
Reviewer #90:	3.0
7) Explain your rating by discussing the strengths and weaknesses of the submission. Include suggestions for improvement and publication alternatives, if appropriate. Be thorough – your explanation will be of highest importance for any committee discussion of the paper and will be used by the authors to improve their work. Be fair – the authors spent a lot of effort to prepare their submission, and your evaluation will be forwarded to them during the rebuttal period.	
Reviewer #60:	It it not clear to me whether the proposed method, which is based on existing work on 2D shape simplification, is the best way to approach the 3D reconstruction problem for modeling urban buildings. While the decomposition of the point cloud to horizontal slabs does address the memory issue, solving independent 2D curve reconstruction problems only takes horizontal smoothness into account. However this fails to take into account smoothness in the vertical direction. An early commitment is made while solving the 2D problem which could potentially lead to a bias for models which resemble a stack of slabs esp. for surfaces which contain curvature in a vertical plane. A comparison with the method of "Variational Shape Approximation" from Steiner et al. 2004 seems like a more appropriate comparison rather than the comparison with Qslim shown in the paper. The experiments in this paper are limited to one dataset which does not exhibit some of the complexities that come up in urban modeling. The input point cloud is also extremely dense and high quality in this case, and the experiments described in the paper does not convince me whether the method will work as well for noisy, less dense range scans with larger holes. Some more experiments on a wider set of datasets would be more convincing.
Reviewer #8:	The approach looks like an improvement over last year’s work by Xiao et al. in that it is able to capture much more (nonplanar) detail in the facades. In that sense, it is similar to the image-based approach of [Muller et al 2007], but here it is automated and uses 3D points to infer depth of features. The approach of forming a surface between beautified keyslices looks a reasonable one. The results look excellent. It is surprising that it works so well given that each slab is treated completely independently, i.e. without trying to exploit shared features across similar slabs. The associated (and crucial) "surfaces from contours" solution should be better described. A limitation of the approach is that it currently does not detect and simplify large planar regions across successive keyslices, for instance the vertical wall region between windows on different floors. My impression is that a post-processing pass using mesh simplification should quite effective in fixing this. The paper is unfairly critical of mesh simplification. The poor quality of the simplified mesh in Figure 15 is due to the presence of topological noise in the original surface (namely lots of holes). Simplification should applied to the model of Figure 15d. Preserving "sharpness" can be achieved using subsequent simplification algorithms that consider attributes (such as normals) defined over the surface.
Reviewer #1:	I liked the general idea of the paper. I would have given it higher points, but the current execution made some choices which can restrict the type of building that can be handled. The problem considered in the paper is very important, and relevant. The proposed procedure is reasonable, and I see quite some potential for such an approach. However, in the current form the paper is not complete. Some details are glanced over, while some of the figures can be improved. The results section is rather limited with tests results presented for only a couple of cases. Regarding comparison, a suitable competitor can be Image based Facade modeling system. (In additions to comments in previous sections.) The model can lead to bad reconstructions for ornamental patterns like floral patterns found on church facades. A zoom in such a level will be interesting to note. The control parameters (the two sigma parameters) even when set to a very fine resolution value, will lead to missing or distorted features. The method will perform well at

Reviewer	Response
Reviewer #78:	coarse resolutions (like Fig 15c), however when we increase the resolution, the method will stop giving acceptable results when the extrusion/tapering assumption breaks down (there is also restriction due to resolution of the input scan). I find it sub-optimal that the method does not make use of redundancy across cross-sections. If such information is carefully exploited then the reconstructions can be significantly improved. The running time is a bit on the high side. I am curious about how sensitive is the method with respect to finding the major axis (line 182-183). Slight error can potentially lead to staggered reconstruction (think of the tower of Pisa). Overall, I liked the idea, but find the paper can be significantly improved. The problem considered in the paper is important, and such a solution may be a good compromise. The paper is not up to the conference standard, and may require a major rewrite, with suitable comparison with existing solutions (image facade paper for example). The topic of the paper is very important and the research question is important for a large audience. The paper is good and has a contribution to the field. In my point of view this contribution is not at the level of Siggraph Asia though. The paper could be accepted at a less competitive venue, but I believe it would need substantial changes to the methodology to be at Siggraph Asia level. My main criticism is the following: The paper breaks down the overall problem in several sub-problems, but none of the sub-problems are studied very systematically. I think the paper should break down the overall problem into better recognized sub-problems and provide solutions to those. The paper somehow creates its own ad-hoc problems and then provides ad-hoc solutions. Overall, I think the approach is good enough, so that I cannot directly name an alternative solution that is clearly better. The system would be an interesting and maybe robust starting point. Therefore, I give a 2.9, indicating that the method is good overall, but not at the required level yet. Some more detailed discussion below. Problems: I do not understand the section 4.2 Hole Filling and what problem is beeing solved here. It seems to me that the algorithm looks for one (global) symmetry line, i.e. one global reflective symmetry, in the model and mirrors all point along this line. This is not a solution to a general hole filling problem. Could you please explain in the rebuttal how that works on the example in figure 12 that depicts the interior scan? Where can you put a symmetry plane in this model? Could you please also explain the following. If windows are the problem and create missing data point, will not the same areas be missing in the mirrored second half of the building? The key slice idea in section 5.1 is a very interesting simplification of the problem, but it is not systematic enough in my point of view. The solution is assuming that the building has a lot of coherence between horizontal cross sections. That means many features with a small extent in the up direction will be removed by the process. It seems to me that this idea can potentially remove or disfigure many small elements. For complex facades this assumption could very well be too simple. I am not really sure that this approach is appropriate for the problem. Also the idea of global detection of key slices is too risky for me. I would argue that a more local computations is necessary for this problem. If we globally compare two cross sections, I do not think we can decide between global noise and a small local feature somewhere. 5.2 seems to produce good results and I also like the method. However, the method is only in 2d and there is a large volume of research that reconstructs surfaces from point clouds. I am not sure how this 2d method fits into the state-of-the art based on the description on the paper. For example, the project "point cloud processing by primitive shapes" also has several strong examples. http://cg.cs.uni-bonn.de/en/projects/point-cloud-processing-with-primitive-shapes/ I think the paper needs to do a better job to put this problem in the context of existing "fitting shapes to point clouds" literature. 5.3 is another version of the more general fitting shapes to point clouds problem.

Reviewer	Response

I am concerned that the paper and video only contain a single real example (plus some interior scan, but that is mainly one wall). How can I know that the parameters of the algorithm work for a larger class of models? How do I know that the algorithm itself works for a larger class of models? What is the class of models that the system works for? What is a failure example?

Overall, the quality is good, but not excellent. There are also some questions about the comparison.
The paper mixes two existing, but somewhat unrelated problems together. First, the problem is surface reconstruction from point clouds. Second, the problem is surface simplification. If these two problems are solved separately, it would be important to use a higher quality solution to the reconstruction problem than the one that was used to generate figure 15 a and b.
How about using your model with all individual slices as input and compare your keyslice version against the many slice + qslim version in addition to the comparison that is there now.

Details:
Figure 1 right has too much compression. It is necessary to zoom in using a factor 1600 to be able to see the details.

As a rule of thumb, one can choose
dHD = MAXfdH(I; Ir); dH(Ir; I)g as the Hausdorff distance.

I do not think this is an appropriate formulation.
I am not sure what that should mean.
Do you claim that the Hausdorff distance is defined like that or do you claim that as a simplification?

What is the purpose of formula 4 and the accompaning paragraph. Why is this 2d to 3d transformation not trivial?

Reviewer #90:

This paper focuses on very important open question. The motivation is good. However, it is premature for publication due to several serious problems:

1. This paper claim to be an automatic method (such as Line 48 and 83). However, it is not. Human efforts are needed. For example, human efforts are needed to define inner and outer boundaries to delete points belong to other objects. Line 191: To tackle these issues, we define inner and outer bounding boxes for the building to clip away unrelated scene objects. Line 213: which can be obtained through user input. Line 406: The chairs and some other fine details were manually culled. If a large manual interaction time is needed for each building, given the fact that the model quality is low (discussed below), why do we take the trouble to use this approach instead of modeling the building by hands in software such as Google Sketchup, where a skill worker only need one or two minutes?

2. The proposed method relies on many parameters. For example:
a) parameters□ \sigma in Line 205,
b) threshold \tao_d in Line 247,
c) radius r in Line 296,
d) threshold \tao_r in Line 312,
e) a new set of threshold parameters needed in Line 379-381.
The parameters are no-trivial to determine and vary from one case to another since they directly affected by the scope of the building. For example, a small house need a set of parameters. A tall or large building need another set of parameters. A range scanning for a building at 200 meters distance need a set of parameters, which scanning at 100 meters distance need another set of parameters. This makes the method very difficult to be useful in practice.

3. Very limited demos.
The paper demos only on two scenes. Therefore, it is difficult to judge the efficiency and effectiveness of the proposed method.

4. Low result quality.
The result is far from visual pleasing. In Figure 15, we can see that no matter (c) or (d) is not good at all, given the fact the range scanning can provide a great detail. In (c), we can see that the mesh of the wall has many artifacts. It is even worse than just giving a flat plane. Lots of important details are lost, which makes the structure of the building become unrealist. For example, the major pattern of windows in (c) and (d) can not be preserved. An intruded rectangle for a window with simple mesh is even much better than a complex shape shown in (d).

5. Insufficient and unfair Comparison.
Only one result is compared with qslim (a 12 year old paper?). Maybe qslim is bad for

Reviewer	Response
	this only example, it is better in other cases. No information about the parameters of qslim is used. Maybe some parameter tuning for qslim can produce much better result. On the other hand, how about the result if we take a two-step approach: first fit a surface, and them simplify them? For example, fit a surface by Poisson (such as Poisson Surface Reconstruction, Kazhdan, SGP2006) and simplify with progressive mesh. 6. Robustness issue. It seems that a small inaccuracy in one step will cause totally fail in the following steps. For example, if the rectification for x,y,z axis is non-perfect, the following steps become very difficult, such as symmetric detection. Need to manually cull some points for the method to work in Line 405-406. What happen if no manually cull is used? If symmetric axis can not be found perfectly, the hole filing and later process seems impossible to produce reasonable results. ... 7. Over-complicated mesh for applications. The produced mesh is still complicated for web application claimed by the authors. For example, if we have 400 buildings in a small town such as the one in the paper loaded into Google Earth, it is very unlikely that Google Earth can download the buildings and render them at a reasonable time. Please provide the evaluation of download and rending time when concluding that such method is suitable for web application. 8. Many technical problems: Only line symmetry, no rotation symmetry? Equation 2 only defines the objective function. How to optimize to obtain L? How many L detected? An ideal rectangle has at least 2 Ls. An ideal square has 4 Ls. Non-perfect symmetry? how about a small difference, possibly due to the error of scanning or registration? No symmetry => No hole fill? Share symmetry between slabs? Symmetry detection together of more slabs is more robust? Holes will affect Hausdorff distance in Equation 3? Circular ball or square? Lost of push and pull structure of building in Figure 8. Maybe translating a square is better? Line 371: The structure containing tapered sub-structures will show a wide and uniform distribution of keyslice images. This contradicts the facts that the keyslice is detected by local maxima of curvature in Line 275. Figure 7 exactly shows a good example that tapered roof has less keyslices.