CVPR 2015
2015 IEEE Conference on Computer Vision and Pattern Recognition

Reviews For Paper

Paper ID 2205
Title Scene Segmentation with Dense Reconstruction From Monocular Video

---

Masked Reviewer ID: Assigned_Reviewer_1
Review:

Question	
Paper and Review Summary. Briefly describe the contributions of the paper to computer vision. Include a concise, bulleted list of the paper's main strengths and a concise, bulleted list of the paper's main weaknesses. Please keep these brief. You will elaborate on the pros/cons in the subsequent text boxes below.	This paper is proposing a method to segment 3d scene using a monocular camera. The model combines semgnetation cues over frames of different views with geometry (geodesic distance) from 3d reconstructed scene. Strenghths: - Accurate system - Novelty Weaknesses: - Readibility needs to be improved - Limited experimental evaluation and analysis
Paper Strengths. Please discuss the positive aspects of the paper. Be sure to comment on the paper's novelty, technical correctness, clarity and experimental evaluation. Notice that different papers may need different levels of evaluation: a theoretical paper may need no experiments, while a paper presenting a new approach to a known problem may require thorough comparisons to	- The method with combined cues from 2d segmentations from multiple views and geometry performs a lot better than its baseline with geometry only (Fig. 7) - Novel framework based on the Lambert-Ambient model.

existing methods. Also, please make sure to justify your comments in great detail. For example, if you think the paper is novel, not only say so, but also explain in detail why you think this is the case.	
Paper Weaknesses. Please discuss the negative aspects of the paper: lack of novelty or clarity, technical errors, insufficient experimental evaluation, etc. Please justify your comments in great detail. If you think the paper is not novel, explain why and give a reference to prior work. Keep in mind that novelty can take a number of forms; a paper may be novel in terms of the method, the theory, analysis for an existing problem, or the empirical evaluation. If you think there is an error in the paper, explain in detail why it is an error. If you think the experimental evaluation is insufficient, remember that theoretical results/ideas are essential to CVPR and that a theoretical paper	- Overall, the readibility of the paper needs to be improved. In many cases, mathematical equations are given before it is explained in a high-level. Similarly, it was hard to capture key idea from many of paragraphs. - The reviewer agree that there is no benchmark for comparison for 3d scene segmentation from monocular camera. However, wonder if it is possible to evaluate on 2d scene by reprojecting 3d segmentation into 2d. There would be more numbers of baseline methods (including [2] or gPb) to compare. - Line 123~125 argues that the system is realtime. However, fig 8 suggest that it takes around 8 seconds. Please clarify. - Figure 2 doesn't have enough description. E.g., what are those numbers? - Figure 3 doesn't have a caption to explain how this is drawn. Is this synthetic? Or, is this the statistic from the dataset? - (minor) typo: line 107, "adnd" -> "and", line 403 Eq (2.4) -> Eq (2) - (minor) Figure 7 is hard to read.

need not have experiments. It is *not* okay to reject a paper because it did not outperform other existing algorithms, especially if the theory is novel and interesting. It is also not reasonable to ask for comparisons with unpublished papers and papers published after the CVPR deadline.	
Preliminary Rating: This rating indicates to the area chair, to other reviewers, and to the authors, your current opinion on the paper. Please use 'Borderline' only if the author rebuttal and/or discussion might sway you in either direction.	Weak Reject
Preliminary Evaluation. Please explain your current rating on the paper. This explanation may include how you weight the importance of the various strengths and weaknesses you described above in Q1-Q3. Note, after the rebuttal period, you will be asked to submit a response to the rebuttal and a final rating.	Although the paper is dealing with an interesting problem and novel idea to solve, it has weakness in its writing and experimental validations.
Rebuttal	

Requests: Make a list of items you would like the authors to be sure to address in their rebuttal.	- Please describe Fig 2 & 3 more carefully, what it means and how there are drawn. - Please clarify the meaning of line 123-125. Is it a realtime system?
Confidence. Select: "Very Confident" to stress that you are absolutely sure about your conclusions (e.g., you are an expert who works in the paper's area), "Confident" to stress that you are mostly sure about your conclusions (e.g., you are not an expert but can distinguish good work from bad work in that area), and "Not Confident" to stress that that you feel some doubt about your conclusions. In the latter case, please provide details in your Preliminary Evaluation response.	Confident

Masked Reviewer ID: Assigned_Reviewer_2

Review:

Question	
Paper and Review Summary. Briefly describe the contributions of the paper to computer vision. Include a concise,	The paper presents a method to segment a dense meshed model of a scene, reconstructed from monocular video. The mesh is segmented by accumulating segmentations of individual images from the video sequence. The final segmentation of the mesh is based on affinities between mesh nodes based on a curvature-weighted geodesic distance between nodes, and the layer assignments from the image-based segmentation. The contributions are: - A dataset of scenes reconstructed from monocular video with segmentation labels - A method for combining segmentations from individual frames into a segmentation of the mesh of the scene - A feature which measures geodesic distance between nodes on a mesh

bulleted list of the paper's main strengths and a concise, bulleted list of the paper's main weaknesses. Please keep these brief. You will elaborate on the pros/cons in the subsequent text boxes below.	reconstructed from multiple images The strengths are: - The paper is well written and the methodology clear - The problem is under-explored and prior work in this area is lacking - The public dataset could be of some use to future research The key weaknesses are: - A lack of qualitative or quantitative comparison with any serious baseline measure (apart from a scaled-back version of their own algorithm) - It is not clear what the advantages are of the proposed method (image segmentation regularized on the mesh) over simpler alternatives (e.g. directly segmenting the mesh). - There is a lack of discussion of the requirements of the segmentation algorithm. This is related to the subjective nature of segmentation. - Some parts of the algorithm are complex yet have little motivation given for their complexity. Other details (e.g. parameter values) are omitted.
Paper Strengths. Please discuss the positive aspects of the paper. Be sure to comment on the paper's novelty, technical correctness, clarity and experimental evaluation. Notice that different papers may need different levels of evaluation: a theoretical paper may need no experiments, while a paper presenting a new approach to a known problem may require thorough comparisons to existing methods. Also, please make sure to justify your comments in great detail. For example, if you think the paper is novel, not only say so, but also explain in detail why you think this is the case.	- The paper explores a fairly novel area of segmentation of a mesh created from dense structure-from-motion. - The release of the dataset could be of some benefit to the wider community and would allow for other authors to compare their results with those from this paper (albeit with the caveats about labeling outlined below). - The introduction and related work sections are well written and appear to be fairly comprehensive. - The geodesic distance measure is well explained, in spite of its complexity. The intermediate results in Figure 6 give an indication that the geodesic distance measure could be a good cue for segmentation.
Paper	

Weaknesses. Please discuss the negative aspects of the paper: lack of novelty or clarity, technical errors, insufficient experimental evaluation, etc. Please justify your comments in great detail. If you think the paper is not novel, explain why and give a reference to prior work. Keep in mind that novelty can take a number of forms; a paper may be novel in terms of the method, the theory, analysis for an existing problem, or the empirical evaluation. If you think there is an error in the paper, explain in detail why it is an error. If you think the experimental evaluation is insufficient, remember that theoretical results/ideas are essential to CVPR and that a theoretical paper need not have experiments. It is *not* okay to reject a paper because it did not outperform other existing algorithms, especially if the theory is novel and interesting. It is also not reasonable to ask for comparisons	- A key weakness is the lack of qualitative and quantitative comparisons with baselines. The only baseline algorithm compared against is the a modification of the full algorithm presented in the paper. There are many algorithms for segmentation of RGB images, RGBD images and meshes which could have been compared against. Consider some of the mesh segmentation algorithms presented in Chen et al, A Benchmark for 3D Mesh Segmentation, SIGGRAPH 2009. At least one of these algorithms could be used on your generated meshes with little modification. Consider also 'Mesh Segmentation with Concavity-Aware Fields' (Au et al., TVCG 2012). - Some parts of the methodology are very complex. While this is not a problem in itself, there is sometimes no motivation in the paper for the introduction of these complexities. For example, equations 2.4, 2.5, 2.6 and 2.7 represent the affinities between mesh nodes, but only a small amount of justification is given for these formulae --- e.g. what is the benefit of using a Geman-McClure robust function? - The benefit of segmenting the images rather than the mesh is not immediately clear. The images are segmented using an existing technique for short-baseline video, ignoring the PTAM output and the meshed reconstruction of the scene. The individual image segmentations are then projected into the mesh using the known camera poses. These different segmentations are then combined. The benefits of doing this, rather than alternatives such as segmenting the mesh directly should be made clear through explanation and/or qualitative and quantitative evaluation. The proposed algorithm seems to suggest that two identical meshes reconstructed from different input video sequences could be given very different final segmentations. Is that true, and why is that desirable? - Segmentation is a subjective problem, as a scene could be segmented into regions comprising semantic objects, or parts of objects, or segments of similar sizes etc. These different possible segmentation requirements are not addressed at all in the paper. As the labeling of the dataset is provided by the authors, the desired type of segmentation should be discussed in some way. For example, on the ground truth, it appears also that two of the branches of the tree are segmented as one object, but the others are segmented separately. It is not clear that there is an obvious 'ground truth' for the sequences presented without a discussion of the desired aim of the segmentation. Potential improvements discussed on line 538 could also be valuable. - The authors state that they are not aware of benchmarks for scene segmentation from monocular video (line 532). The SUN3D dataset (http://sun3d.cs.princeton.edu/), while not directly designed for the problem in hand, comprises of an RGBD video with camera pose information and accompanying segmentations. Qualitative and quantitative results on a dataset such as this would help to convince reviewers of the strengths of the presented algorithm. - The values used for some parameters do not appear to be defined (e.g. \\alpha and \\epsilon). - As a more minor point, having the key for the images in Figure 9 and 10

with unpublished papers and papers published after the CVPR deadline.	buried in the text is frustrating and should be changed.
Preliminary Rating: This rating indicates to the area chair, to other reviewers, and to the authors, your current opinion on the paper. Please use 'Borderline' only if the author rebuttal and/or discussion might sway you in either direction.	Strong Reject
Preliminary Evaluation. Please explain your current rating on the paper. This explanation may include how you weight the importance of the various strengths and weaknesses you described above in Q1-Q3. Note, after the rebuttal period, you will be asked to submit a response to the rebuttal and a final rating.	The problem being tackled is currently under-explored, and the method presented could be a good attempt to solve the problem. However, it is impossible to be confident of the method's success without adequate evaluation, which is currently missing from the paper. Furthermore, the chosen method seems inadequately justified. The paper may be more suitable for acceptance if (a) more comprehensive evaluations are added showing an improvement above recent baselines on more 'fair' test datasets, and (b) comprehensive justification, in text and in quantitative evaluation, is given for the algorithms chosen.
Rebuttal Requests: Make a list of items you would like the authors to be sure to address in their rebuttal.	- What is the justification for using the method of [2] for segmentation rather than segmenting the mesh directly? - A discussion of why a mesh segmentation algorithm was not included as a baseline comparison - An explanation of the role the Geman-McClure robust function plays - A discussion of what requirements were desired when developing the segmentation algorithm. It would be useful to state the specific role this algorithm could play in a practical application area. - How do you hand-select the scale for the geometric scale used as your baseline comparison (line 568)?

Confidence. Select: "Very Confident" to stress that you are absolutely sure about your conclusions (e.g., you are an expert who works in the paper's area), "Confident" to stress that you are mostly sure about your conclusions (e.g., you are not an expert but can distinguish good work from bad work in that area), and "Not Confident" to stress that that you feel some doubt about your conclusions. In the latter case, please provide details in your Preliminary Evaluation response.	Confident

Masked Reviewer ID: Assigned_Reviewer_3

Review:

Question	
Paper and Review Summary. Briefly describe the contributions of the paper to computer vision. Include a concise, bulleted list of the paper's main strengths and a concise, bulleted list of the paper's main weaknesses. Please keep these brief. You will elaborate on the pros/cons in the subsequent text boxes below.	This paper presents a method to compute a dense 3D reconstruction from monocular images, in order to produce a segmentation of the static 3D scene: 1- The dense monocular reconstruction is computed in real-time with a combination of prior work. 2- Using a discretized mesh derived from the reconstructed scene, the scene segmentation problem is then represented as the task of aggregating/separating the nodes of the mesh. Affinities are computed between the nodes using an adaptive geodesic distance function between nodes, with edges drawn between points in adjacent voxels. Occlusion relationships between regions on the image plane are also used as constraints for segmenting the videos. 3- To make the segmentation based on pairwise geodesic distances tractable, the voxels are uniformly sampled, and the sparse segmentation is projected back to the full mesh to obtain the final segmentation.

Paper Strengths. Please discuss the positive aspects of the paper. Be sure to comment on the paper's novelty, technical correctness, clarity and experimental evaluation. Notice that different papers may need different levels of evaluation: a theoretical paper may need no experiments, while a paper presenting a new approach to a known problem may require thorough comparisons to existing methods. Also, please make sure to justify your comments in great detail. For example, if you think the paper is novel, not only say so, but also explain in detail why you think this is the case.	- The problem of 3D scene segmentation is of major interest in the computer vision community. - The adaptive geodesic distance function to compute affinities between nodes of the scene mesh is interesting.
Paper Weaknesses. Please discuss the negative aspects of the paper: lack of novelty or clarity, technical errors, insufficient experimental evaluation, etc. Please justify your comments in great detail. If you think the paper is not novel, explain why and give a reference to prior work. Keep in	- The key assumptions used in this approach are constraining and not very scalable: - “the only thing moving in the scene is the viewer”, l.130. This means that this approach works only for static scenes, which represent only a small

mind that novelty can take a number of forms; a paper may be novel in terms of the method, the theory, analysis for an existing problem, or the empirical evaluation. If you think there is an error in the paper, explain in detail why it is an error. If you think the experimental evaluation is insufficient, remember that theoretical results/ideas are essential to CVPR and that a theoretical paper need not have experiments. It is *not* okay to reject a paper because it did not outperform other existing algorithms, especially if the theory is novel and interesting. It is also not reasonable to ask for comparisons with unpublished papers and papers published after the CVPR deadline.	number of real-life scenarios. - "segmented regions in the images will be spatially consistent when back projected onto the scene", l.366. This is again not true in dynamic scenes. - There is a lack of details and motivation for the geodesic distance function, Sec. 2.4. This is regrettable, since this is the main novelty of the approach. The rest of the paper feels like a combination of other prior work, and not so much novel information is presented. - The approach feels convoluted: there are a lot of steps with many parameters that are left unexplained: - what are the s_i l.377 - how are the curvatures k1,k2 and eigenvalues v1,v2 in Eq.5 determined - why is Fig.5 not referenced in the text - The experimental evaluation is incomplete. There is absolutely no quantitative nor qualitative comparison to state of the art methods on the 3d segmentation task. For fair comparison, the authors should use a standard dataset, such as: - the NYU dataset, from N. Silberman, D. Hoiem, P. Kohli, and R. Fergus. Indoor segmentation and support inference from rgbd images. In ECCV 2012. - the TUM dataset, from J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers. A benchmark for the evaluation of rgb-d slam systems. In IROS 2012. - Sec. 3.2 should be revisited: - The second paragraph l.666-679 should come first, since it is very confusing to look at Fig.9 and Fig.10 without the context of what each row represents. - (left) l.589 should be (top) and (right) l.593 should be (bottom).
Preliminary Rating: This rating indicates to the area chair, to other reviewers, and to the authors, your current opinion on the paper. Please use 'Borderline' only if the author rebuttal and/or	Borderline

discussion might sway you in either direction.	
Preliminary Evaluation. Please explain your current rating on the paper. This explanation may include how you weight the importance of the various strengths and weaknesses you described above in Q1-Q3. Note, after the rebuttal period, you will be asked to submit a response to the rebuttal and a final rating.	This paper lacks novelty and because the experimental evaluation is incomplete with no comparison to the state of the art, it is not very strong. The approach is also limited to static scenes, which is not representative of most real-life scenarios. The writing in some sections is hard to read (e.g., l.501-502, Eq.5 and Eq.6) and some details are missing (definition of s_i l.377, values of \alpha and \epsilon Eq.7).
Rebuttal Requests: Make a list of items you would like the authors to be sure to address in their rebuttal.	- Address lack of comparison to state of the art, and use of standard dataset like NYU dataset. - Better illustrate and motivate the novel pieces of the paper. - Clarify the definition of the geodesic distance function, as well as Sec.3.2 in the experimental evaluation.
Confidence. Select: "Very Confident" to stress that you are absolutely sure about your conclusions (e.g., you are an expert who works in the paper's area), "Confident" to stress that you are mostly sure about your conclusions (e.g., you are not an expert but can distinguish good work from bad work in that area), and "Not Confident" to stress that that you feel some doubt about your	Confident

conclusions. In the latter case, please provide details in your Preliminary Evaluation response.	