DEPTH MAP RECONSTRUCTION USING COLOR-BASED REGION MERGING

Camilo Dorea and Ricardo L. de Queiroz

Department of Computer Science
University of Brasilia, DF, Brazil
Email: camilo@cic.unb.br, queiroz@ieee.org

ABSTRACT

This paper presents a novel depth map enhancement method which takes as inputs a single view and an initial depth estimate. A region-based framework is introduced wherein a color-based partition of the image is created and depth uncertainty areas are identified according to the alignment of detected depth discontinuities and region borders. A color-based homogeneity criterion is used to guide a constrained region merging process and reconstruct depth estimates within the uncertainty areas. Experimental results on publicly available test sequences illustrate the potential of the algorithm in significantly improving low quality depth estimates.

***Index Terms*—** depth map reconstruction, view rendering, region merging, segmentation

1. INTRODUCTION

The advent of 3D video services has also been accompanied by rising interest in the processing and generation of the images necessary for depth perception. In order to simultaneously provide multiple images of the same scene at different viewing angles, several video representation formats and processing techniques have adopted the use of depth maps. A depth map, associated to a particular view of the scene, is used to represent the distance of each point to a reference camera. The depth map results in a gray-scale image in which objects that are closer to the camera appear brighter while farther ones appear darker. Its structure contains generally smooth areas separated by discontinuities.

Given a particular view and associated depth map, with knowledge of depth range and camera calibration, rendering algorithms [1] may be used to synthesize other virtual views. Thus, depth maps allow view synthesis at arbitrary angles while occupying relatively little storage space. Nevertheless, usage of depth maps present a few disadvantages. Precise depth map computation, either through disparity estimation or with direct measurement equipment (e.g., light ranging systems), is a challenging and costly task. Inaccurate estimation of the discontinuities along the borders of objects at different depths can lead to disturbing rendering artifacts such as incomplete or disconnected objects. Although the presence of such artifacts may be mitigated when complementary information is available from more than one adjacent view, depth map accuracy remains essential to quality of the rendered view – and even more so if considering single-view formats such as video-plus-depth.

Several proposals have focused on the improvement of initial depth map estimates. In computer vision applications, segmentation-based approaches have been used to fill in missing depth information [2, 3]. Video images and a limited amount of range data are combined in an MRF framework to compute unknown depth values [2]. In [3] an inpainting algorithm allows the user to remove an object from the scene and then complete missing texture and depth information using stereo images. The previous approaches rely on multiple views to complete a user defined missing depth area. Another family of approaches is based on the design of non-linear, edge-preserving filters which incorporate image information while respecting depth map structure. The bilateral filter uses color information as weights to enhance discontinuous areas in low-resolution depth maps [4]. A similar reconstruction filter is applied in [5] to compensate for depth coding errors and recover object boundaries. These filtering approaches, however, depend on correct registration between the color image and depth map. Lastly, the technique in [6] uses an adaptive median filter and image and depth information from multiple views to post-process coded depth maps.

This paper proposes a novel segmentation-based, depth map reconstruction algorithm which takes as input a single view and an initial depth estimate. A region-based framework is introduced in which an initial color-based partition of the image is created. Based on this partition, areas of depth uncertainty are identified and reconstructed in accordance to region boundaries. This proposal extends and improves work initially presented in [7] by precisely identifying uncertainty areas without user intervention. Furthermore, performance is also evaluated based on the quality of the rendered views using objective metrics. Experimental results illustrate the potential of the algorithm in reconstructing severely distorted

This work has been partly supported by grants from FINATEC, FINEP and CNPq of the Brazilian Government.

depth maps. The technique may be applied, for instance, in enhancing low-resolution or coarse depth estimates originating from commercial range sensing devices when used in conjunction with a high-resolution color camera. Within such setups, the technique may also be used in adapting non-registered depth estimates to the current color view.

2. REGION-BASED FRAMEWORK

Region-based representations have several advantages over traditional pixel-based representations in image processing tasks. Regions can provide a meaningful support closer in scale to that of the decision level while greatly reducing the number of elements to be processed and accurately preserving edge information. Our region-based framework presents a methodology using an image partition. In subsequent sections we describe a specific implementation of this methodology.

In this work, region merging is used to form an image partition which will serve as a basis for our depth reconstruction algorithm, depicted in the diagram of Fig. 1. Creation of the image partition is guided by a color-based homogeneity criterion. The assumption is that depth discontinuities in the scene are a subset of the region borders present in the color-based image partition.

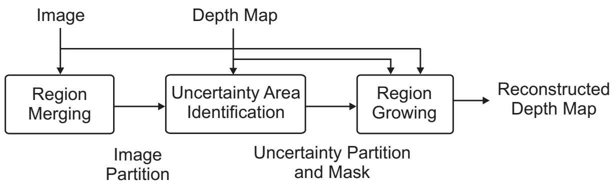


Fig. 1: Block diagram of the depth map reconstruction algorithm.

Once an appropriate image partition has been defined, uncertainty areas are identified as those which are not properly aligned to the relevant discontinuities of the input depth map. A connected component analysis is used to reformulate the image partition and produce an uncertainty mask.

Finally, the region growing block of Fig. 1 reconstructs depth within uncertainty areas. A color-based criterion is again used to guide a constrained merging process which reassigns depth values in accordance to region boundaries.

3. REGION MERGING

The region merging algorithm of [8] is adopted to create an image partition. The algorithm constructs a Region Adjacency Graph (RAG) where each node represents a region and links define connectivity between regions. The algorithm proceeds iteratively by removing a link from the RAG, merging the associated regions and updating the RAG to the new neighborhood configuration. A similarity measure between regions is defined to order the merging process until a termination criterion is reached.

The merging algorithm is initiated upon individual pixels and a color-based partition of an image, such as the one shown in Fig. 2, is created with a color-based similarity measure which accounts for area size as well as shape complexity. The color model $\mathbf{C}_R$ for a region R is a column vector containing the mean YUV color components over all pixels $p \in R$. A color similarity measure between regions R_i and R_j which also accounts for region area $|\cdot|$ is given by:

$$S_a(R_i, R_j) = |R_i| \left\| \mathbf{w}^{1/2} (\mathbf{C}_{R_i} - \mathbf{C}_{R_i \cup R_j}) \right\|^2 + |R_j| \left\| \mathbf{w}^{1/2} (\mathbf{C}_{R_j} - \mathbf{C}_{R_i \cup R_j}) \right\|^2 \quad (1)$$

where $\|\cdot\|$ is the Euclidean norm and the vector of weights $\mathbf{w}^{1/2} = [\sqrt{w_Y}\ \sqrt{w_U}\ \sqrt{w_V}]$ is such that $w_Y + w_U + w_V = 1$. The YUV space has been chosen for its convenience. Shape complexity of the resulting regions can be reduced by introducing an additional term into the color similarity metric. Assume, without loss of generality, that R_i is the region of smallest perimeter, $C(R_i, R_j)$ is the ratio of the increment in perimeter and the increment in area of $R_i \cup R_j$ relative to R_i and $cp(R_i, R_j)$ is the common perimeter between R_i and R_j. The final color-based similarity measure is

$$S(R_i, R_j) = \alpha S_a(R_i, R_j) + (1 - \alpha) \frac{C(R_i, R_j)}{cp(R_i, R_j)}. \quad (2)$$

Termination criteria for region merging can be the total number of regions or a global color modeling error, see [9] for further details. Note that the framework allows for alternate partition definition methods or similarity measures which may reinforce other image features.

(a) (b)

Fig. 2: (a) The *Ballet* sequence (view 1, frame 0) and (b) a color-based partition (typically 500 regions).

4. UNCERTAINTY AREA IDENTIFICATION

Depth estimation errors generally manifest themselves as displaced discontinuities within the depth map. The proper identification of uncertainty areas begins with the detection of relevant depth discontinuities. A Sobel operator is used to compute depth map gradients and these are thresholded against a minimum value δ. Fig. 3(b) shows the relevant discontinuities of the sample depth map of Fig. 3(a).

Depth discontinuities have been assumed to be a subset of the region borders in the color-based image partition. Thus,

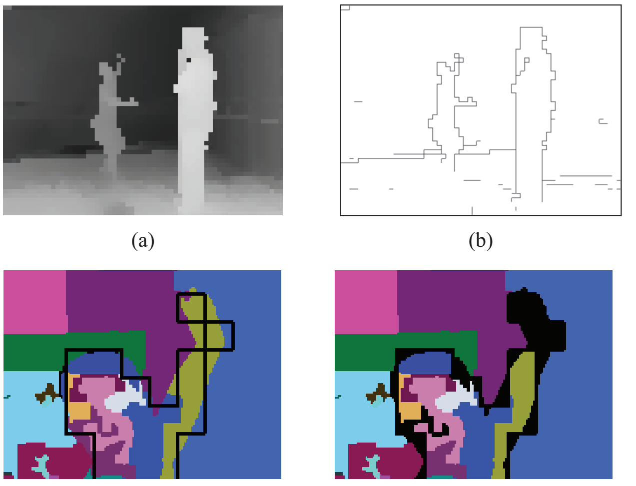


Fig. 3: (a) A corrupted depth map, (b) relevant depth discontinuities, detail crop of (c) color-based partition (in pseudo color) with relevant discontinuities (in black) and (d) uncertainty partition with mask.

uncertainty areas are identified within regions whose borders are not aligned with these discontinuities. For this purpose, the relevant discontinuities are superimposed on the color-based partition, as exemplified in Fig. 3(c), altering partition structure by either covering regions or splitting them. In order to preserve as many valid depth estimates as possible, the largest connected component of split regions retain their original labels while the remaining area, formed by one or more new regions, is classified as uncertain. A mask identifying uncertainty areas is defined by: (i) the regions covered by the discontinuity and (ii) the new regions created by splitting. The uncertainty partition is formed by incorporating the new regions derived from splitting into the original color-based partition. A sample mask superimposed on an uncertainty partition is presented in Fig. 3(d).

5. REGION GROWING AND DEPTH ASSIGNMENT

Depth map reconstruction is concluded with a region growing algorithm using the uncertainty partition and mask. Region growing is based on a merging procedure similar to that of section 3. The difference being that regions outside the uncertainty mask are not allowed to merge amongst themselves and regions within the uncertainty mask inherit the labels of the regions to which they merge. Merging is iterated until no uncertainty regions remain. In this manner, regions outside the mask grow over the uncertainty area.

Note that regions of homogeneous color are not necessarily homogeneous in depth. For example, the regions of the floor in Fig. 2(b) present gradual depth transitions. In this implementation, in order to obtain a more accurate depth assignment within the uncertainty area, region growing is initiated with a set of regions which are both color-homogeneous and of uniform depth. This set of regions is formed by taking the infimum of the uncertainty partition and a partition of uniform depth regions (obtained from the labeling of flat zones in the depth map). The infimum operation determines the intersection of the regions of both partitions, resulting in a set of color-homogeneous regions of uniform depth as depicted in Fig. 4(a).

Fig. 4: Detail crop of (a) color-homogeneous regions of uniform depth outside of uncertainty mask and (b) reconstructed depth map.

Region growing is ordered with the color-based similarity measure of (2). Once terminated, depth map pixels within the uncertainty area are assigned the depth values of their regions prior to growing. Fig. 4(b) presents the reconstructed depth map after depth assignment.

6. EXPERIMENTAL RESULTS

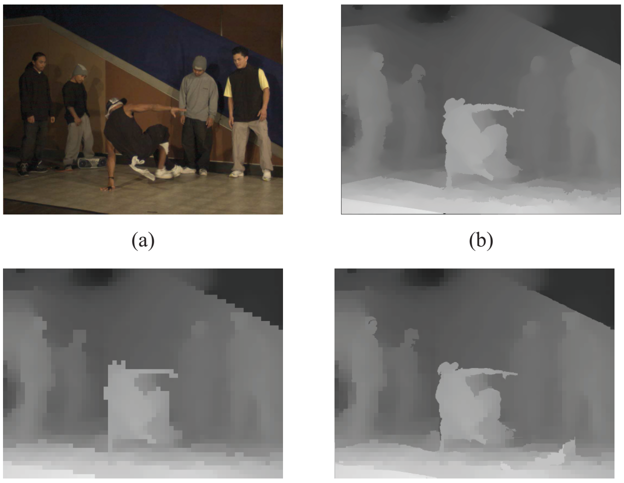
The proposed depth map reconstruction algorithm was tested on two publicly available data sets: *Ballet* and *Breakdancers* [10]. In all tests color-based partitions of 500 regions are used and weights $w_Y = w_U = w_V = 1/3$ and $\alpha = 0.25$. Relevant discontinuities are detected with $\delta = 10$.

The first set of experiments demonstrate depth reconstruction of severely corrupted depth maps. A blocking effect is generated by substituting depths in the originally provided maps for their median values in 16x16 blocks. The blocking simulates a possible lower-resolution depth measurement obtained from a range sensing device. For the *Ballet* image the corrupted map has been presented in Fig. 3(a). Original and reconstructed maps are shown in Fig. 5. For the *Breakdancers* image, original, corrupted and reconstructed maps are presented in Fig. 6. In both cases, the severe blocking effects present in the corrupted maps have been eliminated and depth discontinuities are well aligned with the object boundaries. A few minor holes and protrusions around foreground depth discontinuities in the reconstructed maps are reminiscent of the borders defined in the color-based partition construction.

The second set of experiments investigate the algorithm's potential for reconstructing non-registered depth maps. A depth map from *Breakdancers* view 3 is used as input and reconstructed based on the image of view 2. Results are presented in Fig. 7. In most areas, reconstructed depth discontinuities are properly aligned with the original depth discontinuities of view 2 (highlighted in red). The large inter-view distance relative to the size of thinner objects such as the dancer's left arm do not allow a complete recovery of depth.

(a) (b)

Fig. 5: (a) Original and (b) reconstructed depth maps for *Ballet*.



(a) (b) (c) (d)

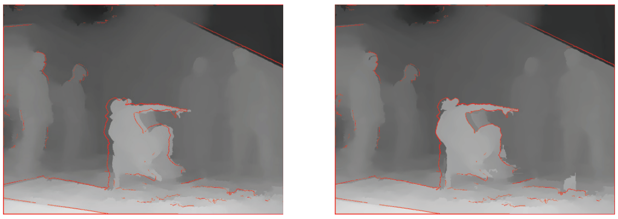
Fig. 6: (a) Image (b) original (c) corrupted and (d) reconstructed depth maps for *Breakdancers*.

Depth reconstruction performance is also evaluated in terms on the quality of rendered views. Depth maps are used to project the image onto a virtual camera. Valid projections areas are compared in PSNR to available ground-truth views. Table 1 summarizes these results indicating PSNR gains of the reconstructed maps obtained from corrupted input maps and non-registered input maps relative to the each of the respective input maps.

Depth map reconstruction improves view rendering in approximately 1.0 dB for the corrupted and non-registered input depth maps in *Ballet*. More importantly, rendered views from reconstructed maps maintain the integrity of foreground and background objects. In contrast, the large differences in alignment between object and corrupted or non-registered depth maps create severe distortions such as missing or split objects. PSNR gains for *Breakdancers* are modest and do not accompany the improvements in visual quality of the maps.

7. CONCLUSION

This paper presents a depth map reconstruction algorithm which takes as inputs a single view and an initial depth estimate. A region-based framework is introduced in which coarsely estimated depths are reconstructed in accordance to



(a) (b)

Fig. 7: (a) Depth map of view 3 and (b) reconstructed map for *Breakdancers* view 2. Original depth discontinuities of view 2 are superimposed in red for comparison.

Image & Input Depth Map	PSNR gain
Ballet & Corrupted	1.0 dB
Ballet & Non-registered	0.9 dB
Breakdancers & Corrupted	0.2 dB
Breakdancers & Non-registered	0.1 dB

Table 1: PSNR gains of rendered views by using reconstructed depth maps from corrupted and non-registered input maps.

the boundaries proposed by a color-based image partition. The algorithm is demonstrated on severely corrupted and non-registered input depth maps.

As future work we intend to exploit other region-based representations, such as partition hierarchies, and investigate non-uniform depth assignments using higher order models.

8. REFERENCES

[1] P. Kauff, N. Atzpadin, C. Fehn, M. Muller, O. Schreer, A. Smolic, and R. Tanger, “Depth map creation and image–based rendering for advanced 3DTV services providing interoperability and scalability,” *Signal Processing: Image Commun.*, vol. 22, no. 2, pp. 217–234, 2007.

[2] L. A. Torres-Mendez and G. Dudek, “Reconstruction of 3D models from intensity images and partial depth,” in *Proc. of American Assoc. for Artificial Intelligence*, 2004.

[3] L. Wang, H. Jin, R. Yang, and M. Gong, “Stereoscopic inpainting: Joint color nad depth completion from stereo images,” in *IEEE Proc. CVPR*, 2008.

[4] Q. Yang, R. Yang, J. Davis, and D. Nister, “Spatial-depth super resolution for range images,” in *IEEE Proc. CVPR*, 2007.

[5] K.-J. Oh, S. Yea, A. Vetro, and Y.-S. Ho, “Depth reconstruction filter and down/up sampling for depth coding in 3D video,” *IEEE Signal Proc. Letters*, vol. 16, no. 9, pp. 747–750, 2009.

[6] E. Ekmekcioglu, V. Velisavljevic, and S. T. Worrall, “Edge and motion adaptive median filetring for multi-view depth enhancement,” in *Proc. PCS*, 2009.

[7] C. Dorea and R. L. de Queiroz, “Depth map discontinuity correction for 3D video systems,” in *Intl. Telecom. Symp.*, 2010.

[8] L. Garrido and P. Salembier, “Region based analysis of video sequences with a general merging algorithm,” in *EUSIPCO98*.

[9] C. Dorea, M. Pardas, and F. Marques, “Trajectory tree as an object-oriented hierarchical representation for video,” *IEEE Trans. on Circ. Syst. Video Tech.*, vol. 19, no. 4, April 2009.

[10] C. L. Zitnick, S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, “High-quality video view interpolation using a layered representation,” *ACM SIGGRAPH and Trans. Graph*, vol. 23, pp. 600–608, 2004.