Pacific Graphics (2012)
C. Bregler, P. Sander, and M. Wimmer (Editors)

Short Papers

User-Assisted Disparity Maps

Hsin-Yi Chen[1], Yi-Shan Lin[1], I-Chao Shen[2], Sheng-Jie Luo[1], Wen-Huang Cheng[2], Bing-Yu Chen[1]

[1]National Taiwan University, Taiwan
[2]Research Center for Information Technology Innovation, Academia Sinica, Taiwan

Abstract

In this paper, we present an intuitive and efficient system for correcting the artifacts and noises caused by imperfect disparity estimation. Disparity estimation is a very important and basic process for almost all stereoscopic 3D applications. Although there are many disparity estimation methods, none of them are perfect and suitable for any arbitrary case. Given a defect disparity map estimated from a stereoscopic image pair, our system incorporates some simple user hints and rich information derived from sparse feature matching as guides for disparity correction. We also utilize 3D connectivity prior and planarity assumption to overcome the challenge presented by textureless surfaces and occluded regions. By providing an intuitive user interface with integrated stereo visualization, our system allows novel users to correct their own disparity maps easily.

Categories and Subject Descriptors (according to ACM CCS):
I.4.8 [Image Processing and Computer Vision]: Scene Analysis—Stereo; H.5.2 [Information Interfaces and Presentation]: User Interfaces—Graphical user interfaces

1. Introduction

Over these decades, many two-view stereo methods for estimating the disparity maps have been proposed [SS02]. However, the accuracy of such automatic methods is still not sufficient for a wide variety of input stereoscopic images. There are several reasons which cause imperfect disparity estimation: (1) matching ambiguities in the regions with large texture-less or repetitive patterns; (2) unmatched regions due to occlusions; (3) wrong point correspondences due to color discrepancy between the stereoscopic image pair; (4) erroneous depth values inferred from the scene with reflections, e.g., mirrors, lakes, etc. Therefore, even applying the best stereo method to real scenes cannot yield perfect disparity estimation. To obtain high quality disparity maps, manually crafting the disparity maps is unavoidable. Current workflow in industry consists of two steps. First, a reviewer uses red-cyan anaglyph images to verify the disparity content. He segments the inaccurate regions and describe the appropriate surface properties. Second, the disparity map is repainted by an artist accordingly. Due to different interpretations of visual perception between the artist and reviewer, it takes lots of time for complex refinement. Costly iterations taken in accessing precise disparity and burdens placed to whom performing the manual editing or tuning are the problems in current solutions. In this paper, we provide an intuitive and efficient system for correcting the artifacts and noises of the disparity maps. It eases the burden of manual correction and achieves better disparity maps than which are estimated automatically. The main contributions of our work are threefold:

1. We designed simple, specific and well-defined tasks for users and utilize the efficiency of computer to resolve many challenging real scene cases.
2. Our framework corrects the errors caused by manual disparity assignments or automatic disparity estimation. To the best of our knowledge, our work is the first one which focuses on the disparity correction given a stereoscopic image pair.
3. Combining interactive selection and instant visual feedback in 3D, we speed up the current work-flow of manual processes with high-quality results.

2. Related Work

Disparity Estimation. Our approach is mostly inspired by the use of Ground Control Points (GCPs) [WY11]. In the stereo literature, GCPs refer to the high confident matches. The use of the GCPs eliminates both the need for biasing the process towards a smoothness solution and the task of selecting critical prior probabilities describing image forma-

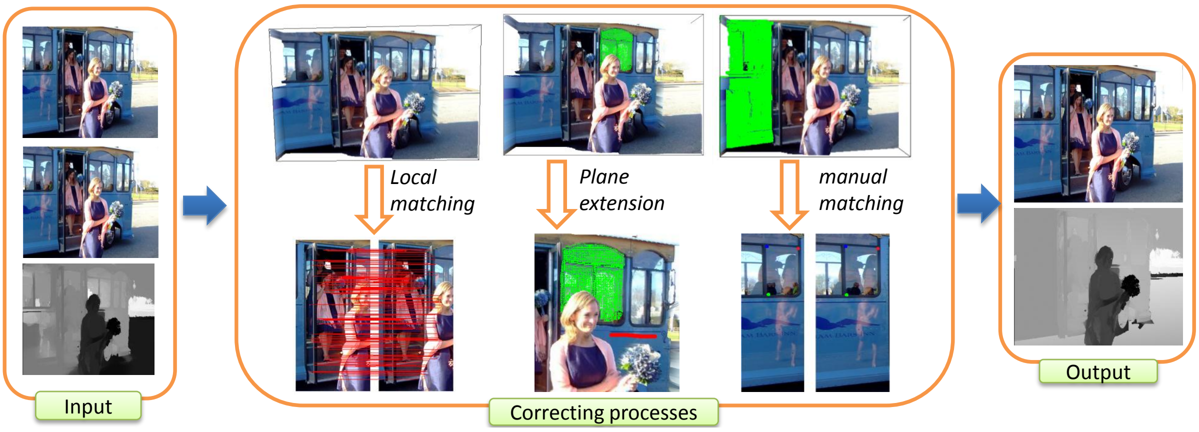


Figure 1: *An overview of the User-assisted Disparity Maps system. The input is a stereoscopic image pair with defect disparity map (Left column). In the disparity correcting processes, users first select the incorrect regions (shown in green), and then perform the correction by (1) local matching propagation for high curvature regions; (2) plane extension (reference plane is derived by the red stoke); (3) manual matching for the surface with reflection. Right column shows the corrected disparity map.*

tion. Experimental results demonstrate the advantages of utilizing a sparse set of points with reliable depths.

Disparity Correction. To correct an inaccurate disparity map, two types of approaches are commonly used: (1) trial-and-error parameter tuning and (2) precise per-pixel assignment. The primary difficulty of the first approach lies in how to settle the appropriate parameters for a specific image. In the second approach, users usually use a 2D image editing tool like “Adobe Photoshop” to manipulate the disparity map as the same flow as in 2D image editing. It is laborious to decide and assign precise disparity values for novel users. Park *et al.* [PKT*11] described a framework to perform up-sampling on depth maps captured from a low-resolution and noisy 3D time-of-flight camera. It allows users to modify the depth map by giving precise depth values, though it is not very easy to use.

3. System Overview

Figure 1 illustrates an overview of our system. Given a stereoscopic image pair and a (defect) disparity map, we first reconstruct its 3D scene. Combining original 2D image and the 3D scene representation enables users to check incorrect pixels and verify the corrected results intuitively. Once the users find incorrect regions, they can segment these regions interactively in the 2D view or directly select them in the 3D view. The problematic regions can be simply classified into the following three categories:

Correction by local matching. In rich-texture or high-curvature regions, we use the information of extracted sparse features to form the soft constraints via the adaptive propagation technique [LLW04]. Further optimized disparity map is obtained by global stereo matching framework [SZJ09]. (details in Section 4.1)

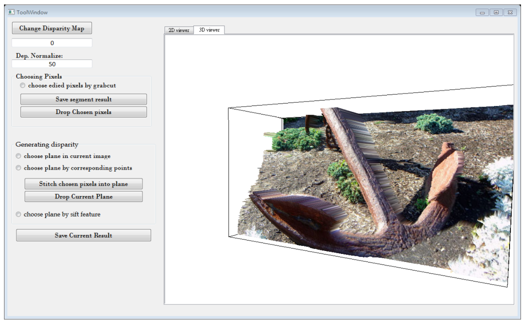
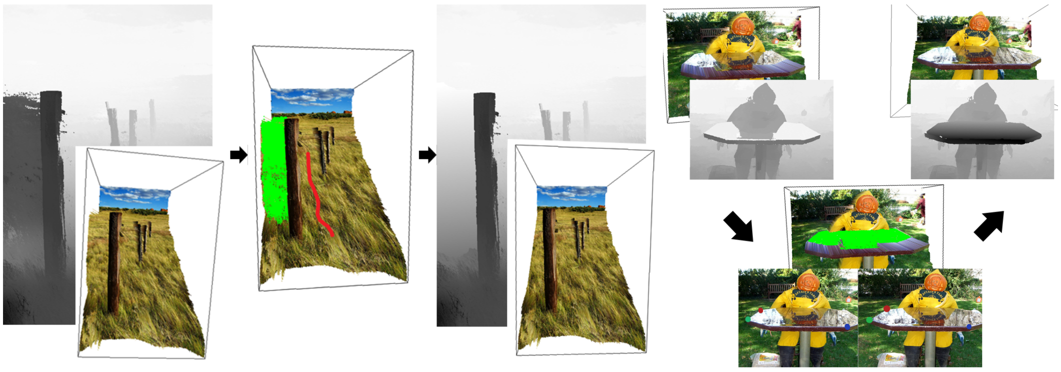


Figure 2: *The user interface of our system. Besides showing the image pair (i.e., the 2D view) directly, our system also provides a 3D view which is an integrated stereo visualization.*

Correction by plane extension. In texture-less or occluded regions, users can choose a reference plane to fit them, so that the incorrect regions can be corrected by extending the reference plane. We assume that the texture-less region is approximately planar. (details in Section 4.2)

Correction by manual matching. For non-Lambertian or color inconsistent regions, we let users specify three or more corresponding features in the image pair. The best fitting plane for the regions can then be constructed and the incorrect disparities can be corrected using least square fitting. (details in Section 4.3)

Finally, the system renders a novel view using the corrected disparity map, and the users can perform these operations iteratively using our interface which is shown in Figure 2.



(a) (b)

Figure 3: *(a) An example work-flow of correction by plane extension. (b) An example work-flow for correcting the surface with reflectance by manual matching.*

4. Disparity Correction

Our method supports three alternative ways for users: correction by local matching, correction by plane extension, and correction by manual matching.

4.1. Correction by local matching

For rich-textured areas and highly-curved surfaces, we adopt GCPs (ground control points) as a guide for matching. Users only need to indicate the regions requiring correction. Our system will perform local correction to the effected regions automatically. This work-flow is demonstrated in Figure 4. The GCPs can be derived by stable sparse feature matching, e.g., [Low99]. Their influences during the global inference in the Markov Random Field can be modeled in a principle way [WY11].

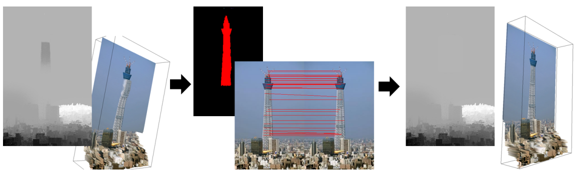


Figure 4: *An example of correction by local matching. From left to right: a defect disparity map and its corresponding 3D view, the incorrect regions selected by users and the corrected result.*

4.2. Correction by plane extension

For texture-less or occluded regions, after indicating the regions requiring correction, the users can assign its reference plane by simply scribbling on the image. The work-flow is illustrated by the sequence of images shown in Figure 3(a). Since each part of these regions is actually connected in 3D, by assuming there is no large depth discontinuity in the regions with homogeneous color, we impose the assumption: all 3D points in each texture-less region lies on the same plane in 2D. If the users indicate some parts of the regions with correct disparities, such reference regions can extend their disparities to other regions with incorrect disparities. The disparities in each region obey the same plane equation. To satisfy these assumptions, our system applies plane fitting methods to the selected area.

4.3. Correction by manual matching

For the objects / regions with reflectance property (e.g. lake, glass, etc.), which violate the assumption of photo-consistency, we require users to choose at least three corresponding point pairs to construct the fitting plane. Since these regions are largely planar, a planar model is sufficiently to model the object's geometry. Figure 3(b) gives an illustration of this work-flow.

Finally, only the user-selected region(s) is affected by the local correction. The discontinuity on the boundary introduced by the corrected result is fixed by the Poisson-based blending [PGB03].

5. Experiment Result

5.1. Analysis

We evaluated our method on the Middlebury data set [SS02] as well as a number of general outdoor scenes. Figure 5 demonstrates several results generated by the proposed method. The top row of Figure 5(a) is the estimation result obtained from [MSJ*11] which is ranked as the best algorithm in the Middlebury at this moment. Pixels colored in black indicate that their disparity errors exceed 1 pixel when comparing to the ground truth. With *correction by plane fitting* in our system, the users can select the defect regions indicated by the red rectangles and the corresponding reference planes. The local corrected result is shown in the top row of Figure 5(b). The second and third rows of Figure 5(a) are the estimation results obtained using [SZJ09]. With the *correction by local matching* in our system, the users can select the defect region(s) indicated by the red rectangle(s), and the local corrected result is shown in the second row of Figure 5(b). Also, with the *correction by manual matching* in our system, it only requires the users to select the defect region(s) indicated by the red rectangle(s), and indicate corresponding feature pairs in the image pair. The local corrected result is shown in the third row of Figure 5(b). These cases demonstrate that our method can correct the defects in user-selected regions.

We also evaluated our method on stereo image pairs collected from personal digital photos. The corrected results are shown in Figure 3. The table shown in Figure 3(b) has perfect specular effects, and it is difficult for traditional formulations to estimate accurate disparity map. After selecting those defect regions, users can reconstruct their original geometry by plane fitting via simply indicating the corners of these objects as corresponding pairs.

5.2. Limitation

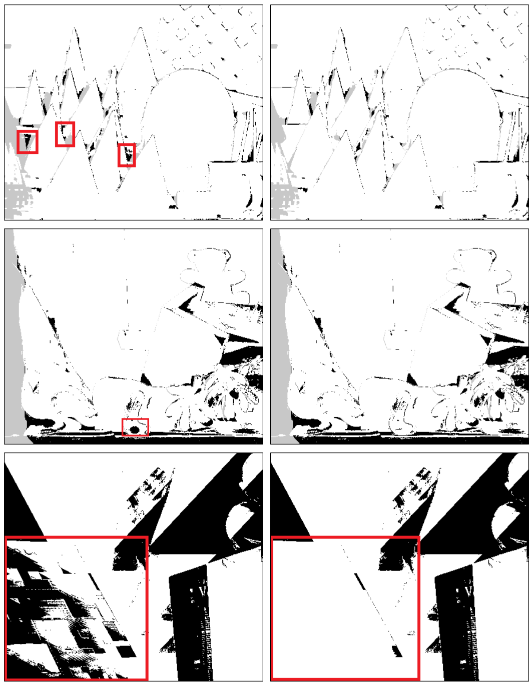
There are still some limitations in our method. For the cases without matched feature pairs, obtaining correct disparity by local matching is still challenging. Plane extension cannot solve the case when isolated objects and objects appearing only in one view may not have any reference region with correct disparity. Besides, objects with complex structure may not be solved by manual matching because the planar assumption would not be satisfied.

6. Discussion and conclusion

We have presented a novel work-flow that allows users to correct the disparity maps easily. We also provide a useful interface that eases the process of current work-flow for disparity correction. Many interesting directions for future work remains. We would like to investigate how to develop more operations with different scene constraints using this system.

7. Acknowledgments

We thank the anonymous reviewers for their constructive comments. Special thanks also go to flickr users: Jerry3D and pinboke for sharing their stereoscopic images. This research was supported in part by the National Science Council of Taiwan under grant NSC100-2622-E-002-016-CC2.



(a) (b)

Figure 5: *The bad pixel images produced by previous method (a) and the corrected bad pixel images (b).*

References

[LLW04] LEVIN A., LISCHINSKI D., WEISS Y.: Colorization using optimization. *ACM TOG 23*, 3 (2004), 689–694. 2

[Low99] LOWE D.: Object recognition from local scale-invariant features. In *Proc. IEEE ICCV '99* (1999), vol. 2, pp. 1150–1157. 3

[MSJ*11] MEI X., SUN X., JIAO S., WANG H., ZHANG X.: On building an accurate stereo matching system on graphics hardware. In *Proc. IEEE ICCV '11 Workshop on GPU in Computer Vision Applications* (2011), pp. 467–474. 3

[PGB03] PÉREZ P., GANGNET M., BLAKE A.: Poisson image editing. *ACM TOG 22*, 3 (2003), 313–318. 3

[PKT*11] PARK J., KIM H., TAI Y.-W., BROWN M., KWEON I.: High quality depth map upsampling for 3D-TOF cameras. In *Proc. IEEE ICCV '11* (2011), pp. 1623–1630. 2

[SS02] SCHARSTEIN D., SZELSKI R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithm. *IJCV 47*, 1-3 (2002), 7–42. 1, 3

[SZJ09] SMITH B. M., ZHANG L., JIN H.: Stereo matching with nonparametric smoothness priors in feature space. In *Proc. IEEE CVPR '09* (2009), pp. 485–492. 2, 4

[WY11] WANG L., YANG R.: Global stereo matching leveraged by sparse ground control points. In *Proc. IEEE CVPR '11* (2011), pp. 3033–3040. 1, 3