A spatial-temporal method to refine a depth image

In Yeop Jang* GIST

Min Ki Park, Seung Joo Lee, Sun Jong Jeong, Kyu Je Lim, Yong Hwi Kim, GIST

Kwan H. Lee GIST

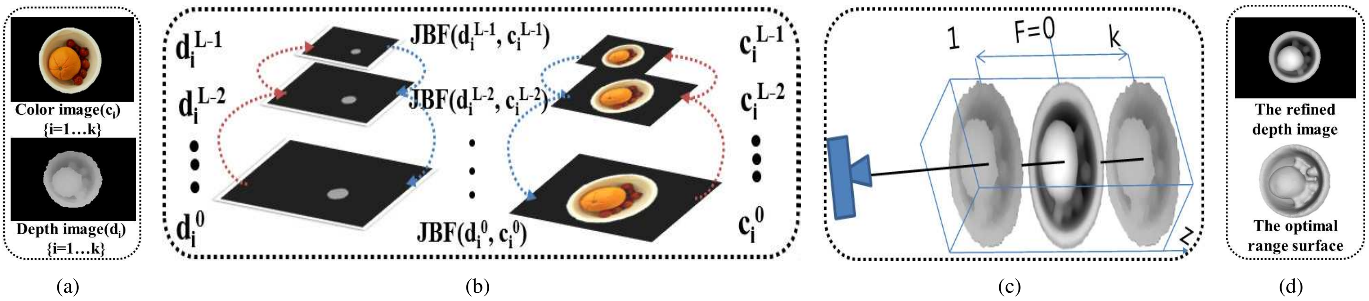

Figure 1: *(a) The input data from Kinect sensor. (b) The spatial filtering based on the pyramidal joint bilateral filter. (c) The temporal depth refinement based on the signed distance functions. (d) The optimal surface and the refined depth image.*

Keywords: Depth Image, Multi-Scale, Signed Distance Function

1 Introduction

In this poster, a spatial-temporal depth image refinement is proposed. First, we perform a spatial denoising by applying Joint Bilateral Filter(JBF) to an image pyramid iteratively. Second, we employ a signed distance function(SDF) [Curless 1996] to obtain an optimal depth image from the depth video in which the temporal noise occurs between video frames.

2 The proposed method

An image which contains objects of many sizes contains also the noises of many sizes. We perform a multi-scale image filtering to reduce the noises by constructing an image pyramid(Figure(b)). We apply JBF to the depth image pyramid iteratively from the smallest depth image with scaling up to the original size. In our joint bilateral filtering, the corresponding color image is used as additional information at each level. Our spatial filter can induce a pyramidal filtering classifying local noises and spread noises. Also, it improves the ambiguous depths of boundary regions of adjacent objects in the depth image by imposing a constraint based on the corresponding color information during the JBF filtering.

The pyramidal JBF is applied over k-frames. Now, we perform a temporal refinement to obtain an optimal depth image using the k filtered images. Even though the depth image is spatially filtered at each frame, the depth pixels suffer from flickered which occurs between frames. To resolve this, we employ a signed distance function which represents the signed distance of each point x to the nearest range surface along the line of sight to the depth camera. The refined k frames are converted the signed distance functions $f_1(x), ..., f_k(x)$ and they are integrated in a discrete voxel grid. First, a depth value is assigned to each voxel according to distance from the depth camera. Then, the pixels of depth image are projected to the voxel grid along their projection lines (see Figure 2). Next, the differences are computed between the depth of each voxel and the one projected from each depth pixel. The difference values mean the signed distance values for the depth image. Then, a SDF is defined by extracting the voxels having the minimum signed

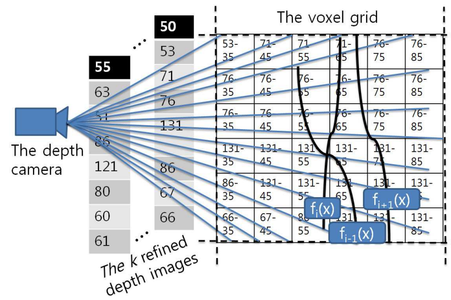

Figure 2: *The figure depicts how SDFs and their range surfaces are computed.*

distance value along the projection lines. Accordingly, k depth images converted as k SDFs and the SDFs construct k range surfaces (Figure 1(c), the black curves in the voxel grid of Figure 2) in the voxel grid. The range surfaces may be still interoperated since they are based on the flickered depth values. We finally compute a cumulative SDF $F(x)$ from all the SDFs, which extract an optimal range surface. The optimal surface can be obtained by satisfying '$F(x) = f_1(x) + .. + f_k(x) = 0$' in the least squares sense. Figure 1(d) shows the optimal depth image and the corresponding optimal range surface. It takes about 30 seconds to process a 640x480 depth image.

Conclusion In this poster, a depth image is spatially and temporally refined by applying a pyramidal JBF and the integration of SDFs.

Acknowledgements This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MEST) (No.2013031191).

References

CURLESS, V., LEVOY M. 1996. A volumetric method for building complex models from range images. In *Proceedings of of the 23rd annual conference on Computer graphics and interactive techniques*, 303–312.

*e-mail:inyeop@gist.ac.kr

SIGGRAPH 2013, July 21 – 25, 2013, Anaheim, California.

ACM 978-1-4503-2261-4/13/07