Intensity Guided Depth Upsampling by Residual Interpolation

Yosuke Konno, Yusuke Monno, Daisuke Kiku, Masayuki Tanaka and Masatoshi Okutomi
Tokyo Institute of Technology

***Abstract* — In this paper, we propose a novel depth upsampling method by residual interpolation (RI) that uses both a low-resolution depth map and a high-resolution intensity image. Our method is an application of the RI to depth upsampling, where the upsampling is performed in a residual domain following its success in the field of image demosaicking. Experimental results demonstrate that our method preserves the sharpness of depth discontinuities and outperforms existing well-known methods such as joint bilateral upsampling and guided filtering.**

1 Introduction

The estimation of an accurate depth map is a highly-demanded task in many fields including computer vision and robotics. Various types of the depth map estimation systems such as a stereo camera, a laser range scanner, and a time-of-flight (ToF) sensor, have been designed. In particular, the ToF sensor is becoming increasingly popular after it is adopted by the Microsoft Kinect v2 sensor. Nevertheless, an inherent limitation of resolution for the obtained depth map keeps it away from various practical applications.

There is high demand to enhance the resolution of an acquired depth map. The enhancement algorithms can be roughly divided into two categories according to whether a high-resolution (HR) intensity image is used as a guide or not. In this paper, we focus on the case that we can involve the HR intensity image. One pioneer work in this category is the joint bilateral upsampling (JBU) [1] that applies the bilateral filter to the low-resolution (LR) depth map under the guidance of the HR intensity image. A guided filter (GF) can be also applied to the depth upsampling with the HR guide image as well as the joint bilateral upsampling.

In this paper, we propose the intensity guided depth upsampling method by residual interpolation (RI). The RI is a technique that has been originally proposed for image demosaicking [3, 4]. Following the success of the RI in image demosaicking, our method performs the upsampling in a residual domain, where the residual is defined as a difference between a tentatively estimated HR and the input LR depth values. The tentative estimate of the HR depth map is generated from the HR intensity image. In experiments, we demonstrate that our method preserves the sharpness of depth discontinuities and outperforms the JBU and the GF in terms of both quantitative and qualitative evaluations.

2 Proposed method

Based on the idea of the RI, the proposed method upsamples the input LR depth map in the residual domain, where the residual is defined as the difference between the tentatively estimated HR and the input LR depth values. Figure 1 shows the overall processing pipeline of our proposed method via the RI for the $\times 2$ upsampling operation. First, the tentative estimate of the HR depth map is generated from the HR intensity image by upsampling the input LR depth map. Second, the residuals are calculated by subtracting the tentatively estimated depth values from the corresponding input depth values. Then, the residuals are interpolated. Finally, the tentative estimates are added to the interpolated residuals to obtain output HR depth map. Our method is motivated from the fact that the residual map becomes flatly smooth when there is a correlation between the intensity guide image and the depth map, and that the smooth condition of the residual map generally improves the performance of the interpolation.

In the first step, we use the GF [2] for the upsampling operation to generate the tentative estimate of the HR depth map. The GF reconstructs the tentative estimate as a local linear combination of the intensity guide as

$$t_i = a_k I_i + b_k, \quad \forall i \in \omega_k, \tag{1}$$

where ω_k denotes a local image window centered at the pixel k, i represents a pixel location in the window, t_i is the tentative depth value, and I_i is the pixel value of the HR intensity guide. The local linear coefficients (a_k, b_k) are calculated by minimizing the following cost function.

$$E(a_k, b_k) = \sum_{i \in \omega_k} \left\{ (a_k I_i^M + b_k - p_i)^2 + \varepsilon a_k^2 \right\}, \tag{2}$$

where I_i^M is the pixel value of the masked HR intensity guide and p_i is the corresponding LR depth value. The second term is a regularization term and ε is a regularization parameter.

Now, there are multiple values per a single pixel by the overlap of each local linear approximations. Instead of uniformly averaging them as in [2], we introduce a weighted averaging of the obtained linear coefficients as

$$\bar{a}_k = \frac{\sum_{i \in \omega_k} W_i a_i}{\sum_{i \in \omega_k} W_i}, \quad \bar{b}_k = \frac{\sum_{i \in \omega_k} W_i b_i}{\sum_{i \in \omega_k} W_i}, \tag{3}$$

where the weight is determined by the cost of the GF as

$$W_i = \frac{1}{\max\left(\frac{1}{|\omega_i|} \sum_{j \in \omega_i} (a_i I_j^M + b_i - p_j)^2, \ \delta \right)}, \tag{4}$$

where δ is a threshold parameter to avoid division by zero. The tentative HR depth map is finally calculated as

$$t_k = \bar{a}_k I_k + \bar{b}_k. \tag{5}$$

After generating the tentative estimate, the residuals are calculated and interpolated. Although our method can include any sophisticated interpolation algorithms, we use the bicubic interpolation here for its simplicity. Finally, the output HR depth map is obtained by adding the tentative estimate to the interpolated residual map.

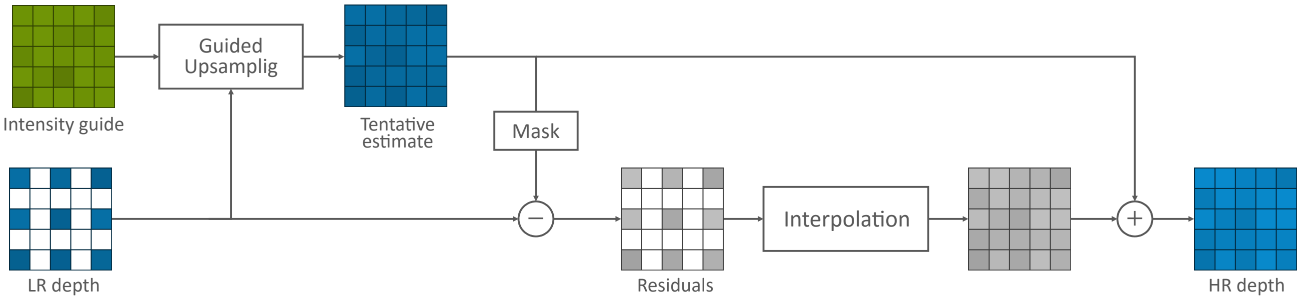
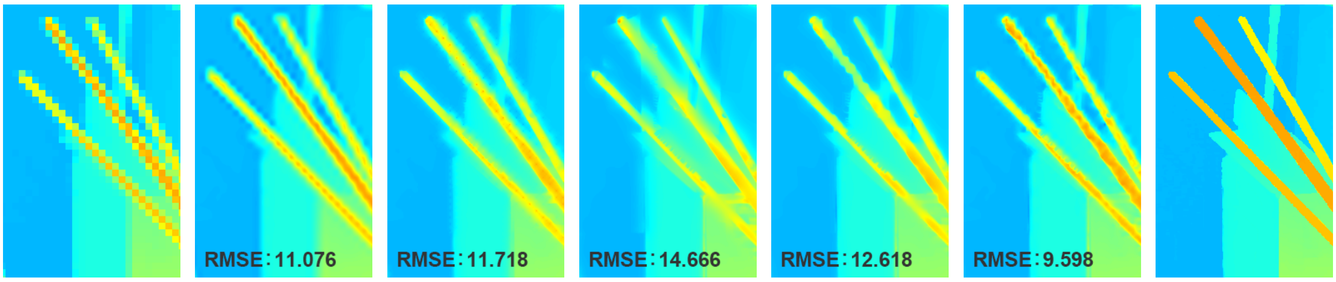


Fig. 1 The overall flow of our proposed method.



(a) Input (b) Bicubic (c) JBU [1] (d) GF [2] (e) GF+wei. (f) Proposed (g) Ground truth

Fig. 2 The visual comparison of $\times 8$ upsampled depth maps. (RMSE values in the above images are calculated on those clipped area.)

3 Experimental results

In experiments, the performance of our method[1] is evaluated for the upsampling operation with the scaling factor = 2, 4, 8, and 16 on the Middlebury benchmark dataset [5] used in [6]. The input LR depth map was synthesized by subsampling the ground-truth HR depth map after Gaussian blurring. We iteratively performs the $\times 2$ upsampling operation when the scale factor is larger than 2. The window size and the regularization parameter ε in (2) are set to 5×5 and 0.001, respectively.

We compare our method with bicubic interpolation, the JBU [1], the GF [2] and the GF with the weighted average in (4). Table 1 shows the average RMSE errors of three images, namely *Art*, *Books*, and *Moebius*. One can see that our method yields the lowest RMSE errors at all upsampling scales. Figure 2 shows the visual comparison of upsampled depth maps at a snippet of *Art*. In the three sticks in the front, one can see that our method can correctly reconstruct the depth values compared with the others. Our method can also suppress some kind of bleeding artifacts that bicubic interpolation yields.

Table 1 The average RMSE on the Middlebury dataset.

Scale factor	Bicubic	JBU [1]	GF [2]	GF+ wei.	Proposed
$\times 2$	1.878	1.915	2.101	1.820	**1.793**
$\times 4$	2.421	2.519	2.915	2.197	**2.111**
$\times 8$	3.372	3.463	4.187	3.153	**2.852**
$\times 16$	4.888	4.981	6.051	4.707	**4.197**

4 Conclusion

In this paper, we have proposed the novel depth upsampling method by the RI that uses both the LR depth map and the HR intensity image. Our method performs the upsampling in the residual domain, where the residual is defined as the difference between the tentatively estimated HR and the input LR depth values. The tentative estimate of the HR depth map is effectively generated from the HR intensity image. The experimental results demonstrate that our method outperforms the JBU and the GF, which are well-known depth upsampling methods, in terms of both quantitative and qualitative evaluations.

[1] http://www.ok.ctrl.titech.ac.jp/res/DSR/RI.html

References

[1] J. Kopf, M. F. Cohen, D. Lischinski, and M.Uyttendaele, "Joint bilateral upsampling." *ACM TOG*, vol. 26, no. 3, 2007.

[2] K. He, J. Sun, and X. Tang, "Guided image filtering." In *ECCV*, pp. 1-14, 2010.

[3] D. Kiku, Y. Monno, M. Tanaka, and M. Okutomi, "Residual interpolation for color image demosaicking." In *ICIP*, pp. 2304-2308, 2013.

[4] Y. Monno, D. Kiku, S. Kikuchi, M. Tanaka and M. Okutomi, "Multispectral Demosaicking with Novel Guide Image Generation and Residual Interpolation." In *ICIP*, pp.645-649, 2014.

[5] D. Scharstein and C. Pal, "Learning conditional random fields for stereo." In *CVPR*. 2007.

[6] D. Ferstl, C. Reinbacher, R. Ranftl, M. Ruether, and H. Bischof, "Image guided depth upsampling using anisotropic total generalized variation." In *ICCV*, 2013.