COLOR-GUIDED ENHANCEMENT OF AIRBORNE LASER SCANNING DATA

Göksu Keskin, Wolfgang Gross and Wolfgang Middelmann

Fraunhofer Institute of Optronics, System Technologies and Image Exploitation

ABSTRACT

This paper suggests using color-guided depth enhancement algorithms of computer vision to improve the resolution of airborne laser scanning (ALS) point clouds for remote sensing applications. We use co-registered high resolution color images with nadir view to enhance the ALS data; and perform quantitative evaluation in form of RMSE considering the whole depth image as well as the depth discontinuities only. Investigated methods include joint bilateral filtering, Markov Random Field (MRF) optimization with first and second order smoothness terms, and anisotropic diffusion. RMSE results on discontinuities indicate that detail improvement performance of the selected methods on the depth discontinuities is not on a satisfactory level for airborne data. Anisotropic diffusion and MRF optimization are promising to provide better results with further adjustments on the smoothness terms.

Index Terms— lidar depth enhancement, MRF optimization, JBF, anisotropic diffusion

1. INTRODUCTION

Airborne laser scanning (ALS) provides a fast way of obtaining high accuracy point clouds for mapping purposes. Capturing georeferenced 3D information requires little effort using available ALS systems [1]. However, ALS point clouds have limited density, thus these systems are not preferred when high-density 3D data is required. In comparison to ALS, digital photogrammetry can achieve a higher point density at similar altitudes, hence is nowadays preferable to ALS when high level of detail is required [2]. Nevertheless, dense stereo reconstruction is an exhaustive process with high algorithmic complexity, which increases the time required to obtain the 3D data from hours to days [2]. To overcome this dilemma of calculation time versus detail level, we suggest combining ALS point clouds with high resolution imagery.

From nadir viewpoint, ALS point clouds can be perceived as low resolution depth maps for the registered color images. A fusion of these two data sources then could follow a simple assumption: Color and depth inconsistencies tend to co-occur, i.e. when observed from a single viewpoint, discontinuities in depth data are often accompanied by changes in color. Thus, fusing color and depth information into a higher density depth model should be possible. In the field of computer vision, this problem is called color-guided depth enhancement. Yang [3] suggests an iterative joint bilateral filter which performs upsampling on low resolution depth maps using a single reference image to a factor of 100 within several minutes. The reason behind choosing a bilateral filter is its edge-preserving properties [4]. Diebel and Thrun [5] express the depth map and the image in a two-layer graph and suggest an optimization solution using Markov Random Fields (MRF). The MRF has a data term and a smoothness term with the goal of incorporating both data conformity and color-guided discontinuity with appropriate weights in the optimization. Further work on MRF-based depth enhancement reinterprets this approach by adding more complex data and smoothness terms to increase the output quality, especially around depth discontinuities [6] [7]. Liu [8] suggests the available depth values to be modeled as a heat map and formulated the depth enhancement as a linear anisotropic heat diffusion problem.

The goal of this work is to provide a quantitative comparison of available algorithmic approaches to the color-guided depth enhancement problem on airborne remote sensing data. Guided by the research trends in computer vision, we choose to compare four approaches: (1) Joint bilateral filtering, (2) original MRF, (3) MRF with a second order smoothness prior and (4) anisotropic diffusion. We use the Vaihingen dataset [9] to evaluate different approaches and perform RMSE calculations using reference depth maps.

This paper is constructed as follows: Algorithmic foundations of the selected methods are presented in Section 2; followed by the details and results of our evaluation scheme in Section 3. Section 4 presents our conclusions.

2. DEPTH ENHANCEMENT METHODS IN RANGE SENSING

This section presents the methods we tested to perform depth enhancement with a given low-resolution input depth map and a co-registered high resolution image. In the following, we denote high resolution image with X, input depth map with Z, and high resolution depth map estimate with Y. Pixel locations are denoted by a two-element vector $\boldsymbol{p} = [\, p_0, p_1 \,]$, with $\boldsymbol{q} = [\, q_0, q_1 \,]$ in the local neighborhood of $\boldsymbol{p}$. Local neighborhood size depends on the algorithm in use. $x_{\boldsymbol{p}}$ and $y_{\boldsymbol{p}}$

are pixel values at the location $\boldsymbol{p}$ of the color image and the depth map, respectively. $z_p \in L$ are depth measurements with L denoting the locations of available measurements.

2.1. Joint Bilateral Filter

Given an initial low-resolution depth map Z and a high-resolution image X, Yang [3] uses an iterative filtering scheme to refine the low-resolution depth map. First, the input depth map is upsampled to the goal resolution e.g. via bilinear interpolation. Then a cost volume C of depth probabilities is created for each potential depth candidate k over a search range K, between k and the current depth map Y:

$$C(\boldsymbol{p},k) = \min(\alpha * K,\ (y_p - k)^2) \quad (1)$$

Here, α is a constant. Each slice of the cost volume is then filtered using a joint bilateral filter that is calculated from the color image. The filter coefficients are calculated as follows:

$$F(x_q) = f_c(x_q) f_s(x_q), \quad (2)$$

$$f_s = \exp\left(-\frac{|p-q|}{\gamma_s}\right), \quad (3)$$

$$f_c = \exp\left(-\frac{|x_p - x_q|}{\gamma_c}\right). \quad (4)$$

f_c is the color component and f_s is the spatial component. The constant γ_s in the spatial component f_s determines the filter size; whereas the impact of color difference is controlled by constant γ_c. The joint bilateral filter behaves like soft color segmentation in super-resolution calculation and aggregates the depth probabilities based on color difference. After filtering the cost slices, a minimum cost is selected by simple subpixel estimation to reduce quantization effects. The cost generation and the filtering are repeated iteratively to reach the optimal solution.

2.2. MRF Optimization

This depth enhancement method treats image and depth map as a connected two-layer graph in which five types of nodes exist: (1) low resolution range measurements Z, (2) high resolution image pixels X, (3) depth discontinuity u, (4) image gradient w and (5) estimated depth map Y. Depth discontinuity and image gradient fuse the image information into the estimation of Y [5]. The MRF is described as a weighted combination of the depth measurement potential Ψ and the depth smoothness prior Φ_1:

$$p(y \mid x, z) = \frac{1}{z} \exp(-\tfrac{1}{2}(\Psi + \Phi_1)), \quad (5)$$

$$\Psi = \textstyle\sum_{p \in L} \beta\, (y_p - z_p)^2, \quad (6)$$

$$\Phi_1 = \textstyle\sum_p \sum_{q \in N(p)} w_{pq} (y_p - y_q)^2. \quad (7)$$

Z is a normalizer; and β is a constant penalty for the difference between estimated and actual values when available. $N(\boldsymbol{p})$ is the 4-neighborhood of pixel location $\boldsymbol{p}$. In the depth smoothness prior, w_{pq} are the weighting factors providing a connection between image and depth values:

$$w_{pq} = exp(-c\, \|x_p - x_q\|_2^2). \quad (8)$$

Here, c is a key constant for penalizing the smoothness around the edges. Equation (5) is solved using a generic conjugate gradient algorithm such as in [10] to obtain the upsampling result.

A reformulation of MRF-based depth enhancement uses a second-order smoothness term. The optimization problem is reformulated as follows [6]:

$$y = \text{argmin}\{\lambda_1\lambda_2\Phi_1 + \lambda_1(1-\lambda_2)\Phi_2 + (1-\lambda_1)\Psi\}. \quad (9)$$

Here, Φ_1 is the first order smoothness prior in (7); Φ_2 is the second order smoothness term. Ψ is the depth measurement potential in (6); λ_1 and λ_2 are weighting factors. Second order smoothness is the key term in this formulation, based on the assumption that surface gradients vary smoothly in the absence of discontinuities. With this assumption, [6] formulates a prediction in the local neighborhood of each pixel:

$$\hat{Y} = PY. \quad (10)$$

The prediction matrix P is defined as a linear combination of neighborhood interpolations and extrapolations. Weights in the linear combination are constructed to prefer interpolation between adjacent color values with little or no difference and favor direction of change in case of a color discontinuity. This enables a better adaptation to discontinuities, since a larger neighborhood is considered in case of large color differences. For a detailed formulation of the construction of the prediction matrix, we refer the reader to [6]. Furthermore, the formulation in [6] reduces the problem to a linear system of equations, which results in a faster calculation than iterative approaches.

2.3. Anisotropic Diffusion

Depth enhancement using anisotropic diffusion [8] treats known depth values as heat sources from which depth diffuses to regions of unknown depth. The diffusion in this case is guided by color and takes place in image pixel domain. The Gaussian color difference measure (8) is used to form diffusion conductance between two neighboring image pixels. (8) implies that depth values will diffuse faster in case of high color similarity. Unknown depth values are estimated solving an equation system of form A$\mathbf{y}$ = $\mathbf{b}$, with A and b defined as follows:

$$A(i,j) = \begin{cases} I(i,j) & \text{if } \boldsymbol{p} \in L \\ I(i,j) - w(i,j) & \text{otherwise} \end{cases} \quad (11)$$

$$b(i) = \begin{cases} z_p & \text{if } \boldsymbol{p} \in L \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

I is the identity matrix, $w_{i,j}$ denotes the Gaussian color difference. The indices i and j denote the vectorized equivalent of pixel locations $\boldsymbol{p}$ and $\boldsymbol{q}$, respectively. $\mathbf{y}$ and $\mathbf{b}$ are considered as vectors with N elements, with N denoting number of image pixels, whereas A is a sparse matrix of size NxN.

3. EXPERIMENTS

3.1. Dataset

We use a subset of the Vaihingen dataset [9], which includes high resolution color images and a digital surface model (DSM) with 10 cm ground sampling distance. We selected a 1000x1000 pixels residential town area, which can be seen in Fig. 1. We used the DSM image as our depth map on a rectangular grid that is perfectly registered to image pixels; and performed downsampling using scaling factors that correspond to point densities 14pts/m^2 (7x), 10pts/m^2 (10x) and 5pts/m^2 (20x). To assure that the resulting depth maps resembled lidar point clouds, we used a pseudo-random grid for downsampling.

The input for the enhancement algorithms is then a simple interpolation of these depth maps except for anisotropic diffusion, which expects only the available set of depth measurements.

3.2. Evaluation

With the goal of restoring the downsampled depth maps to their original resolution, we test the enhancement quality using the algorithms presented in Section 2 and evaluate in form of RMSE between the high-resolution DSM and enhanced depth values. An additional discontinuity-RMSE (d-RMSE) is calculated on the areas around large discontinuities such as transitions from building roofs to roads. We define adjacent pixels with discontinuities larger than 2 meters as “edge pixels”. After determining these edge pixels in the DSM, we perform morphological dilation with 2.5 pixels diameter to obtain a mask for calculating the d-RMSE. Since the methods already favor simple interpolation on smooth surfaces and claim to perform edge-preserving enhancement on discontinuities, d-RMSE aims to evaluate the latter.

Table 1 presents the quantitative evaluation results. As the scaling factor increased, the resulting depth maps included more errors. In comparison to RMSE, d-RMSE performances are considerably poorer. Fig. 2 presents a visual assessment of these errors in form of absolute differences, zoomed into a smaller area for clarity. For all methods, error extents around large discontinuities are often

(a) (b)

Figure 1. Subset of Vaihingen dataset [14] used in experiments: (a) Color image, (b) DSM.

higher than 2 meters, which can be critical for certain classification tasks. On the other hand, the methods perform well in the areas with smooth changes or no changes in depth. An exception to that is joint bilateral filtering, which resulted in higher amounts of errors partially in smooth surfaces as well.

Anisotropic diffusion resulted in the best error performance, followed closely by MRF with a second order smoothness prior. Both methods recover discontinuities better than JBF. JBF tends to create a gradient around depth discontinuities which appears like a blur-effect, resulting in higher d-RMSE values.

4. DISCUSSION AND FUTURE WORK

We evaluated four different approaches to the color-guided depth enhancement problem in airborne data with a nadir view. The key measure of our evaluation was the d-RMSE, which is calculated only around depth discontinuities. Our results indicate that while smooth surfaces are well restored, enhancement around discontinuities require improvement to be used in remote sensing applications. MRF with a second order smoothness term and anisotropic diffusion are promising in addressing these issues since they aim at giving more weight to the smoothness term around color discontinuities. In addition, joint bilateral filtering can be supplemented with a better global optimization approach to avoid blurring the depth edges in a similar way to MRF-based methods. This would allow making use of the shape recovery performance of the joint bilateral filter.

Table 1. Quantitative Evaluation Results

	7x		**10x**		**20x**		**Average**	
	RMSE	***d-RMSE***	***RMSE***	***d-RMSE***	***RMSE***	***d-RMSE***	***RMSE***	***d-RMSE***
JBF	0.3328	1.0654	0.3472	1.1059	0.3971	1.2302	0.3590	1.1338
MRF	0.2933	0.9674	0.3156	1.0319	0.3839	1.2150	0.3309	1.0714
MRF-2	0.2749	0.9180	0.3000	0.9912	0.3763	1.2011	0.3171	1.0368
AD	0.2707	0.9116	0.2971	0.9864	0.3792	1.2140	0.3157	1.0373

JBF: Joint bilateral filter, MRF: Original MRF, MRF-2: MRF with second order smoothness term, AD: anisotropic diffusion.

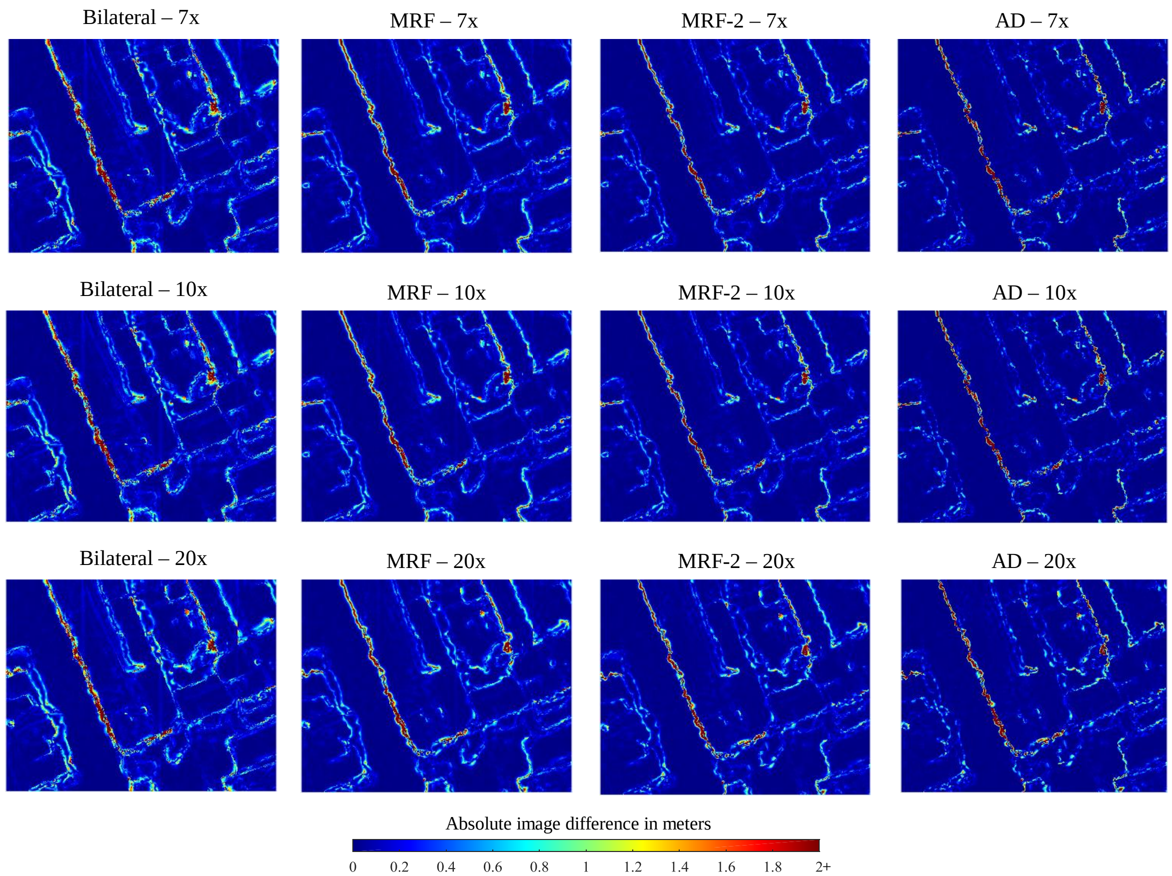

Figure 2. Absolute image differences on the same image subset for comparison. Joint bilateral filtering (JBF) and original MRF optimization (MRF) performed noticeably worse in comparison to MRF with second order smoothness prior (MRF-2) and anisotropic diffusion (AD). Input depth maps with lower resolution resulted in increased errors in smooth surfaces (right column).

Among the discussed methods, anisotropic diffusion is the only method that does not alter the given depth values and recovers the depth values locally from given “seeds”. Depending on the measurement accuracy, this can also lead to error propagation. We suggest investigating the impact of measurement errors in depth enhancement on both image and 3D measurements. Furthermore, for time critical applications, it is also important to make a quantitative evaluation of the runtime and hardware requirements.

5. REFERENCES

[1] E.P. Baltsavias, “A comparison between photogrammetry and laser scanning,” ISPRS Journal of photogrammetry and Remote Sensing, vol. 54, no. 2, pp. 83-94, 1999.

[2] F. Leberl, A. Irschara, T. Pock, P. Meixner, M. Gruber, S. Scholz, and A. Wiechert, “Point clouds: lidar versus 3D vision,” *Photogrammetric Engineering & Remote Sensing*, vol. 76, no. 10, pp. 1123-1134, 2010.

[3] Q. Yang, R. Yang, J. Davis and D. Nistér, “Spatial-depth super resolution for range images,” in *CVPR*, pp. 1-8, 2007.

[4] C. Tomasi and R. Manduchi, “Bilateral filtering for gray and color images,” in *ICCV*, pp. 839-846, 1998.

[5] J. Diebel and S. Thrun, “An application of markov random fields to range sensing,” in *Proc. of Conf. on Neural Information Processing Systems*, vol. 18, pp. 291-298, 2005.

[6] A. Harrison and P. Newman, “Image and sparse laser fusion for dense scene reconstruction,” *Field and Service Robotics*, vol. 62, pp. 219-228, 2010.

[7] S. Lu, X. Ren and F. Liu, “Depth enhancement via low-rank matrix completion,” in *CVPR*, pp. 4321-4328, 2014.

[8] J. Liu and X. Gong, “Guided depth enhancement via anisotropic diffusion,” in *14th Pacific-Rim Conf. on Multimedia*, pp. 408-417, December 2013.

[9] M. Cramer, “The DGPF-test on digital airborne camera evaluation – overview and test design”, *Photogrammetrie – Fernerkundung – Geoinformation*, vol. 2, pp. 73-82, May 2010.

[10] W.H. Press, *Numerical Recipes in C: The Art of Scientific Computing*, New York: Cambridge University Press, 1988, ch. 10.