

Fast MRF Optimization with Application to Depth Reconstruction

Qifeng Chen
Stanford University

Vladlen Koltun
Adobe Research

Abstract

We describe a simple and fast algorithm for optimizing Markov random fields over images. The algorithm performs block coordinate descent by optimally updating a horizontal or vertical line in each step. While the algorithm is not as accurate as state-of-the-art MRF solvers on traditional benchmark problems, it is trivially parallelizable and produces competitive results in a fraction of a second. As an application, we develop an approach to increasing the accuracy of consumer depth cameras. The presented algorithm enables high-resolution MRF optimization at multiple frames per second and substantially increases the accuracy of the produced range images.

Fast MRF Optimization

Consider an image of size $W \times H$. For each pixel $p = (x, y)$ in the image, our goal is to assign it a label $l_p = l_{(x,y)}$. For simplicity of exposition, assume that the labels are natural numbers from 0 to $L - 1$. We also restrict the exposition to first-order models over the 4-connected pixel grid, although the presented algorithm can be generalized beyond this setting. The MRF optimization problem can be expressed as follows:

$$\min_{\mathbf{l}} \mathbf{E}(\mathbf{l}) = \min_{\mathbf{l}} (\gamma \mathbf{E}_{\text{data}}(\mathbf{l}) + \mathbf{E}_{\text{reg}}(\mathbf{l})),$$

$$\mathbf{E}_{\text{reg}}(\mathbf{l}) = \sum_{\{p,q\} \in \mathcal{N}} w_{pq} \rho(\ell_p - \ell_q).$$

Block coordinate descent

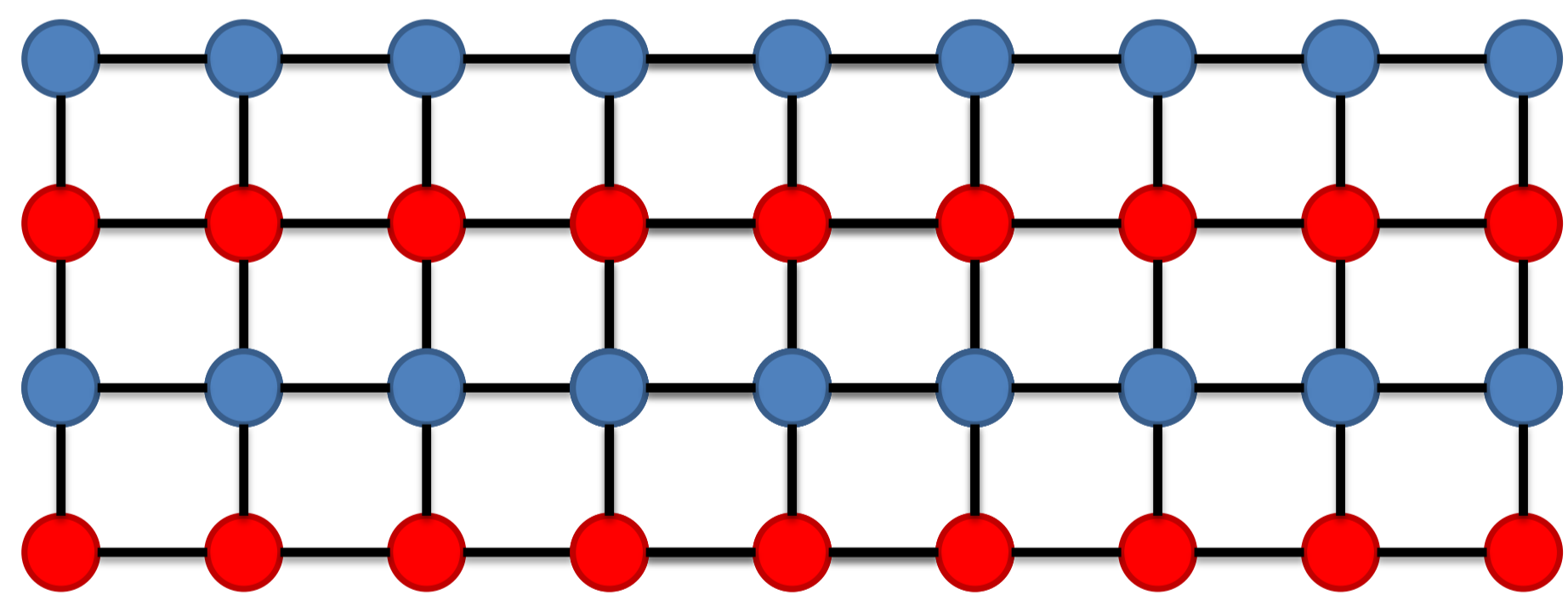


Figure 1: All the even (blue) or odd (red) rows can be processed in parallel.

Our algorithm optimizes the objective $\mathbf{E}(\mathbf{l})$ by block coordinate descent. In each step, \mathbf{l} is fixed outside a complete horizontal or vertical line in the image. The optimal label assignment along this line is computed by dynamic programming, given the fixed assignments along the other lines. We process the lines in four blocks: even rows, odd rows, even columns, odd columns. This enables trivial parallelization.

A naive application of dynamic programming would have complexity $O(VL^2)$ in each step, where V is the number of pixels in the line. However, as observed by Felzenszwalb

and Huttenlocher, this can be accelerated to $O(VL)$ without loss of optimality [2]. The Felzenszwalb-Huttenlocher acceleration scheme was presented in the context of belief propagation and was formulated for a number of specific penalty functions ρ . We give a general formulation that clarifies that this acceleration scheme is applicable to all penalty functions that satisfy a certain combinatorial criterion. In particular, the acceleration scheme applies to all convex penalties, non-convex robust penalties such as the Lorentzian and the generalized Charbonnier, and to truncated variants of these functions.

Depth Reconstruction

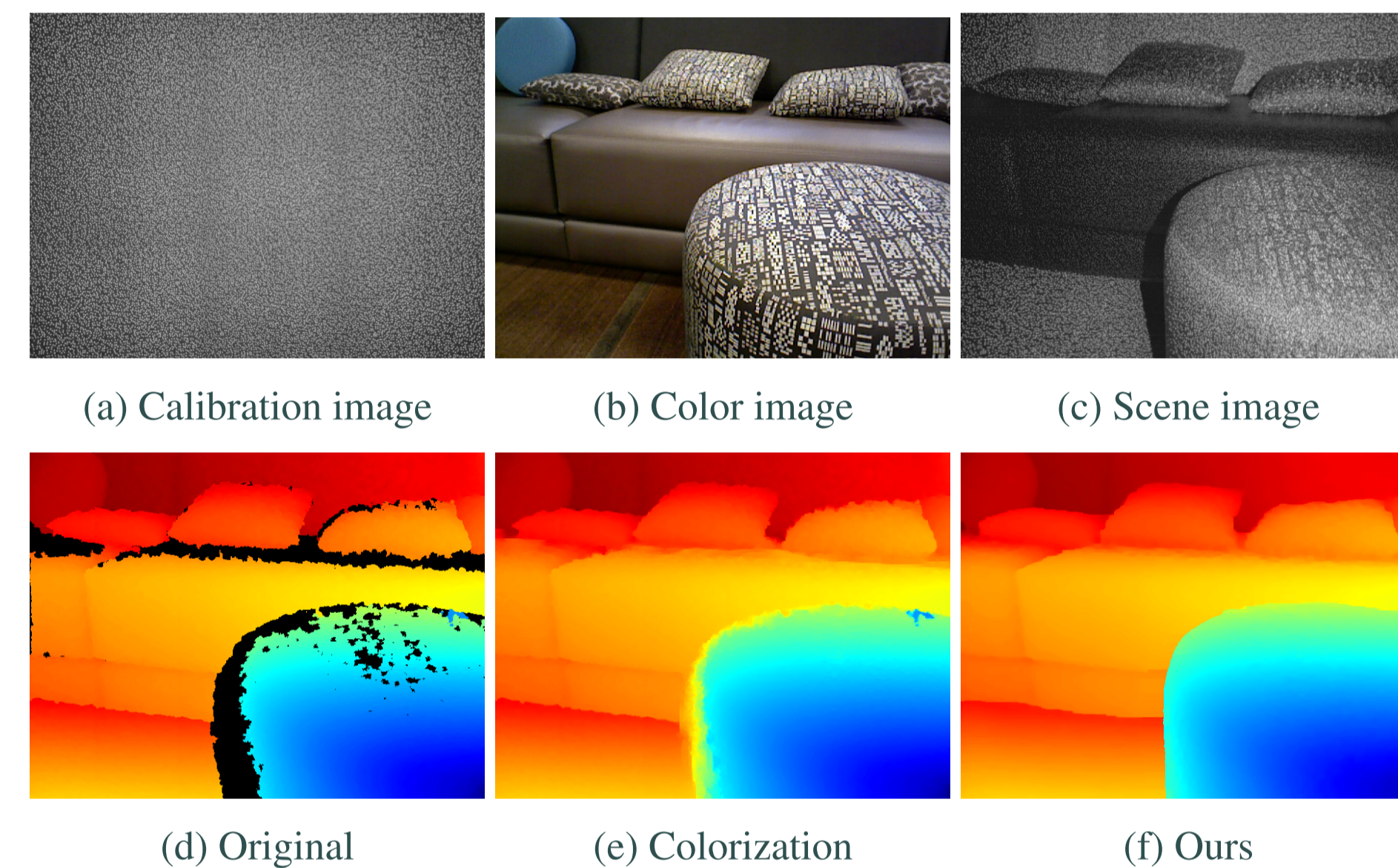


Figure 2: (a) Infrared image of a speckle pattern projected onto a planar surface, (b) color image of an indoor scene, (c) infrared image of the same speckle pattern projected onto the scene, (d) range image produced by the camera, (e) rectification of the range image by a commonly used approach [4], (f) range image produced by our approach.

We develop a depth reconstruction approach that operates directly on the infrared speckle pattern image produced by consumer depth cameras. These cameras perform triangulation on a pair of infrared images. Both images are of a fixed speckle pattern projected by the sensor: one is of the pattern projected onto the current scene and the other is a fixed calibration image of the pattern projected onto a flat plane. The goal is to estimate the disparity between the two images for every visible point in the scene.

Results

We evaluate the presented block coordinate descent (BCD) algorithm on the benchmark problems of Szeliski et al. [6] that have large ordered label sets. The results are shown in Figure 3 (left). The BCD algorithm is not as accurate at convergence as TRW, but is much faster. Figure 3 (right) shows the objective value of each algorithm at convergence as a function of the truncation distance d and the corresponding truncated penalty is defined as $\rho'(a) = \min(\rho(a), \rho(d))$.

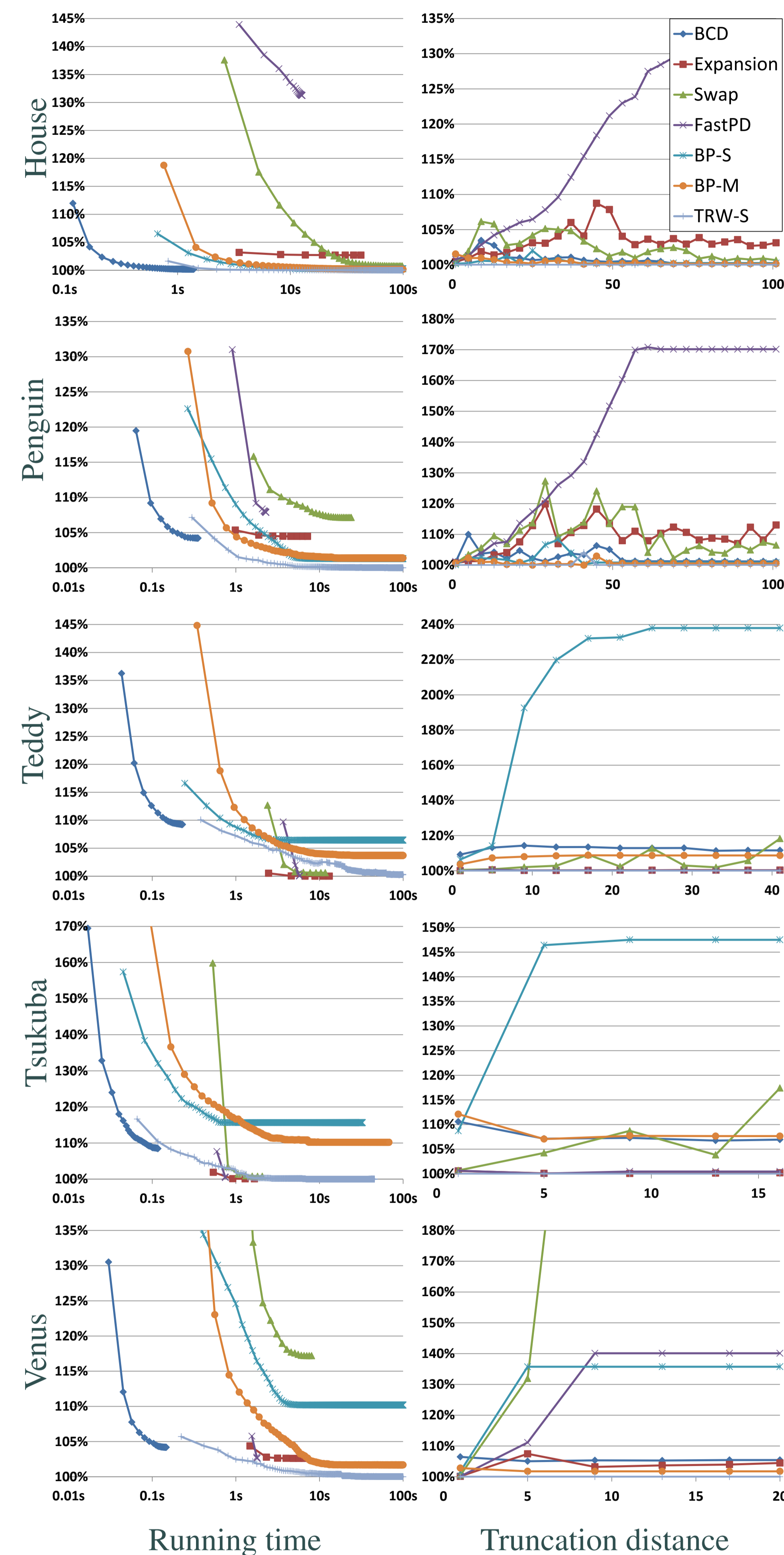


Figure 3: Left. The presented BCD algorithm is compared to alternative MRF optimization algorithms on benchmark problems.



Figure 4: Ground truth models acquired using an industrial 3D scanner. From left to right: angel, arch, fox, gargoyle, and lion.

We also compare the performance of the BCD algorithm to other MRF optimization algorithms on the depth reconstruction problem; the results are shown in Table 1. Finally, in Table 2 we provide an extensive comparison of our depth reconstruction approach to range rectification algorithms advocated in prior work.

	time	angel	arch	fox	garg.	lion	mean
Expansion	112 s	4.45	3.74	3.39	4.94	4.81	4.26
Swap	23 m	4.27	3.51	3.20	4.77	4.51	4.05
TRW-S	83 m	3.71	3.39	3.06	4.49	4.24	3.78
BCD	0.31 s	3.70	2.99	3.08	4.64	4.42	3.77

Table 1: Metric accuracy (in millimeters) of different MRF optimization algorithms on the depth reconstruction problem.

	time	angel	arch	fox	garg.	lion	mean
Original		4.76	3.77	3.55	5.22	5.29	4.52
BF	5.2 s	4.48	3.70	3.27	5.15	4.73	4.27
JBF	1.8 s	4.61	3.61	3.57	5.24	4.91	4.39
Colorization	22 s	4.35	3.45	3.36	5.54	4.57	4.26
DT	3.1 s	4.39	3.47	3.39	5.55	4.57	4.27
Nonlocal	44 s	4.56	3.55	3.50	5.65	4.81	4.42
FoE	79 m	4.22	3.50	3.31	4.95	4.50	4.10
Ours	0.31 s	3.70	2.99	3.08	4.64	4.42	3.77

Table 2: Metric accuracy (in millimeters) achieved by range image rectification algorithms and by our approach. The top row shows the accuracy of the original range image. The following rows report the accuracy achieved by bilateral filtering (BF), joint bilateral filtering (JBF), colorization [4], the MRF model of Diebel and Thrun [1] (DT), the nonlocal filtering approach of Park et al. [5], the Fields of Experts model of Herrera et al. [3] (FoE), and our approach.

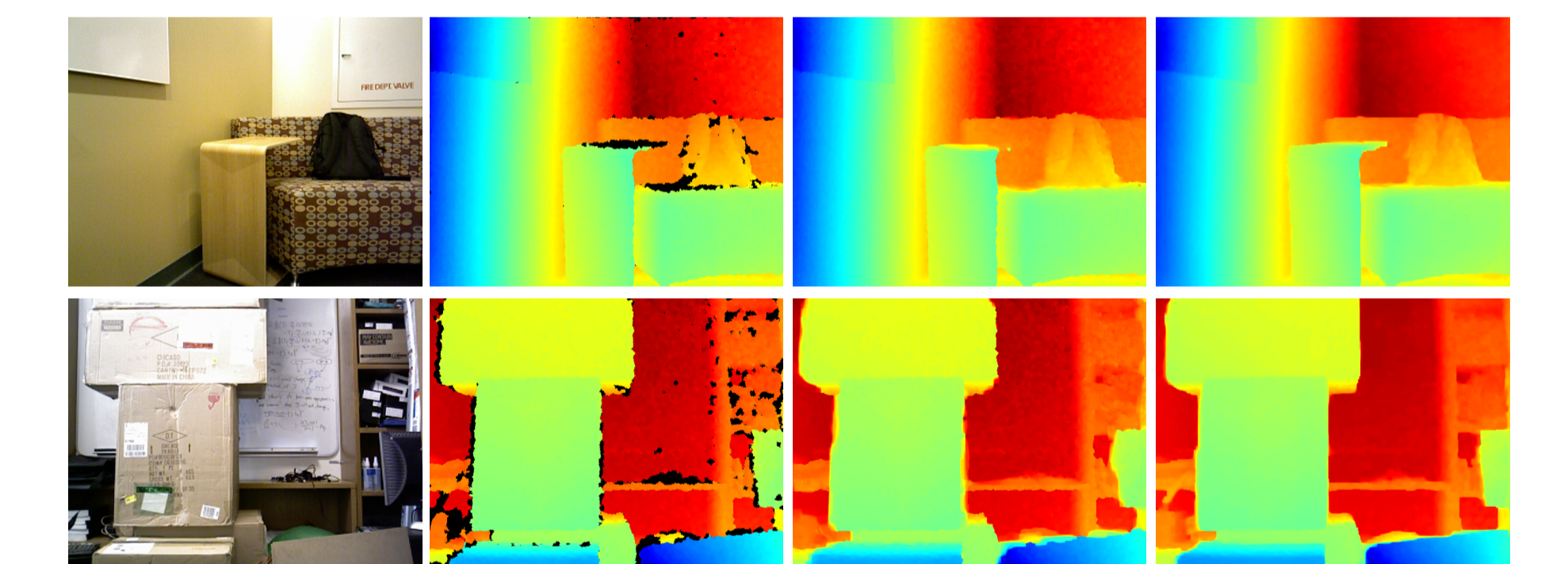


Figure 5: Qualitative results for depth reconstruction.

References

- [1] James Diebel and Sebastian Thrun. An application of Markov random fields to range sensing. In *NIPS*, 2005.
- [2] Pedro F. Felzenszwalb and Daniel P. Huttenlocher. Efficient belief propagation for early vision. *IJCV*, 70(1), 2006.
- [3] Daniel Herrera, Juho Kannala, Peter Sturm, and Janne Heikkilä. A learned joint depth and intensity prior using Markov random fields. In *3DV*, 2013.
- [4] Anat Levin, Dani Lischinski, and Yair Weiss. Colorization using optimization. *ACM Trans. Graph.*, 23(3), 2004.
- [5] Jaesik Park, Hyeonwoo Kim, Yu-Wing Tai, Michael S. Brown, and In-So Kweon. High quality depth map upsampling for 3D-TOF cameras. In *ICCV*, 2011.
- [6] Richard Szeliski, Ramin Zabih, Daniel Scharstein, Olga Veksler, Vladimir Kolmogorov, Aseem Agarwala, Marshall F. Tappen, and Carsten Rother. A comparative study of energy minimization methods for Markov random fields with smoothness-based priors. *PAMI*, 30(6), 2008.