Simultaneous Stereo Video Deblurring and Scene Flow Estimation

Liyuan Pan, Yuchao Dai, Miaomiao Liu, Fatih Porikli
Northwestern Polytechnical University, Australian National University, Data61 CSIRO
panliyuan@mail.nwpu.edu.cn, {yuchao.dai, miaomiao.liu, fatih.porikli}@anu.edu.au

IEEE 2017 Conference on Computer Vision and Pattern Recognition
CVPR July 21-26 HONOLULU 2017

Introduction

Goal

To handle the blurs in **stereo videos** caused by the motion of the camera, objects, and large depth variations in a scene.

(a) Blur image (b) Kim [3] CVPR15 (c) Sellent [4] CVPR16 (d) Ours

Challenges

- Non-uniform blurred image **B**;
- Spatial-variant kernels $\mathbf{A}_m^{\mathbf{x}}$.

Contributions

- A novel joint optimization framework to simultaneously estimate the scene flow and deblur latent images for dynamic scenes;
- Based on the piece-wise planar assumption, we obtain a structured blur kernel model;
- Successfully handle complex real-world scenes depicting fast moving objects, camera motions, uncontrolled lighting conditions, and shadows.

Basic models

Blur model

Blur image **B** is integration of light intensity emitted from dynamic latent images **L** over the aperture time interval of the camera, the model is:

$$\mathbf{B}_m(\mathrm{x}) = \frac{1}{\tau}\int_{m-\frac{\tau}{2}}^{m+\frac{\tau}{2}} \mathbf{L}_m(\mathrm{x}+\mathrm{u}_m^{\mathrm{x}})\mathrm{d}_m = \mathbf{A}_{\boldsymbol{m}}^{\mathbf{x}}\mathbf{L}_m(\mathrm{x})$$

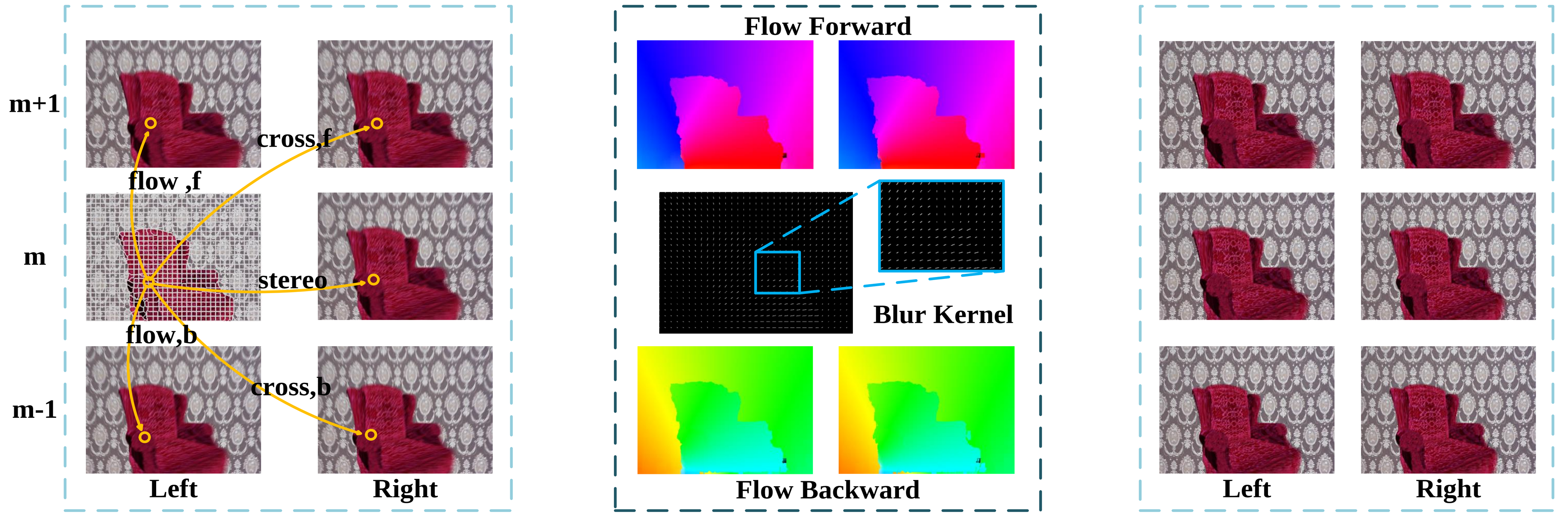
τ is the duty cycle, $\mathrm{u}_m^{\mathbf{x}}$ is the optical flow at pixel **x**, $\mathrm{u}_m^{\mathbf{x}} = \mathbf{H}_i \mathrm{x}_i - \mathrm{x}_i{}'$;

Piece-wise planar model

Each superpixel i is parameterized by a plane n and associated with an object k, the inheriting corresponding motion parameters is $o_k = (\mathbf{R}_k, \mathbf{t}_k)$. Given the parameters (o_k , $\mathrm{n}_{i,k}$), the homography defined for i as $\mathbf{H}_i = \mathbf{K}(\mathbf{R}_k - \mathbf{t}_k \mathrm{n}_{i,k})\mathrm{K}^{-1}$, where $\mathbf{K} \in \mathbb{R}^{3\times3}$ is the intrinsic matrix, $\mathbf{R}_k \in \mathbb{R}^{3\times3}$ is the rotation matrix and $\mathbf{t}_k \in \mathbb{R}^3$ is the translation vector.

Algorithm

Flow chart



Six Blur Images Estimate Blur Kernel Simultaneously Deblur Six Images

Formulation:

A single framework to jointly estimate the scene flow and deblur the images. Particularly, it is a discrete-continuous optimization problem:

$$\mathbf{E} = \underbrace{\sum_{i\in\mathcal{S}} \phi_i(\mathbf{n}_i, \mathbf{o}, \mathbf{L})}_{\text{data term}} + \underbrace{\sum_{i,j} \phi_{i,j}(\mathbf{n}_i, \mathbf{n}_j, \mathbf{o})}_{\text{scene flow smoothness term}} + \underbrace{\sum_{m} \phi_m(\mathbf{L})}_{\text{latent image regularisation}}$$

- Data term

Brightness constancy $\quad \emptyset_i^1(\mathrm{n}_{i,k}, o_k, \mathbf{L}) = \theta_1 |\mathbf{L}(\mathrm{x}) - \mathbf{L}^*(\mathbf{H}^*\mathrm{x})|_1$

Anchor point constraint $\quad \emptyset_i^2(\mathrm{n}_{i,k}, o_k) = \theta_2 \|\mathbf{H}^*\mathrm{x} - \mathrm{x}^*\|_2$

Blur image constraint $\quad \emptyset_i^3(\mathrm{n}_{i,k}, o_k, \mathbf{L}) = \theta_3 \sum\sum \left\|\partial_* \mathbf{A}_m(\mathrm{n}_{i,k}, o_k)\mathbf{L}_m - \partial_* \mathbf{B}_m\right\|_2^2$

(the superscript $*$ denote the direction.)

- Smoothness term

Compatibility of two superpixels i and j that share a common boundary by respecting the **depth discontinuities**; Neighbor superpixels orient to the same **direction**; motion boundaries are co-aligned with **disparity discontinuities**.

- Regularization Term

Total variation to suppress the noise in the latent image while preserving edges, and penalize spatial fluctuations. $\emptyset_m = |\nabla \mathbf{L}_m|$

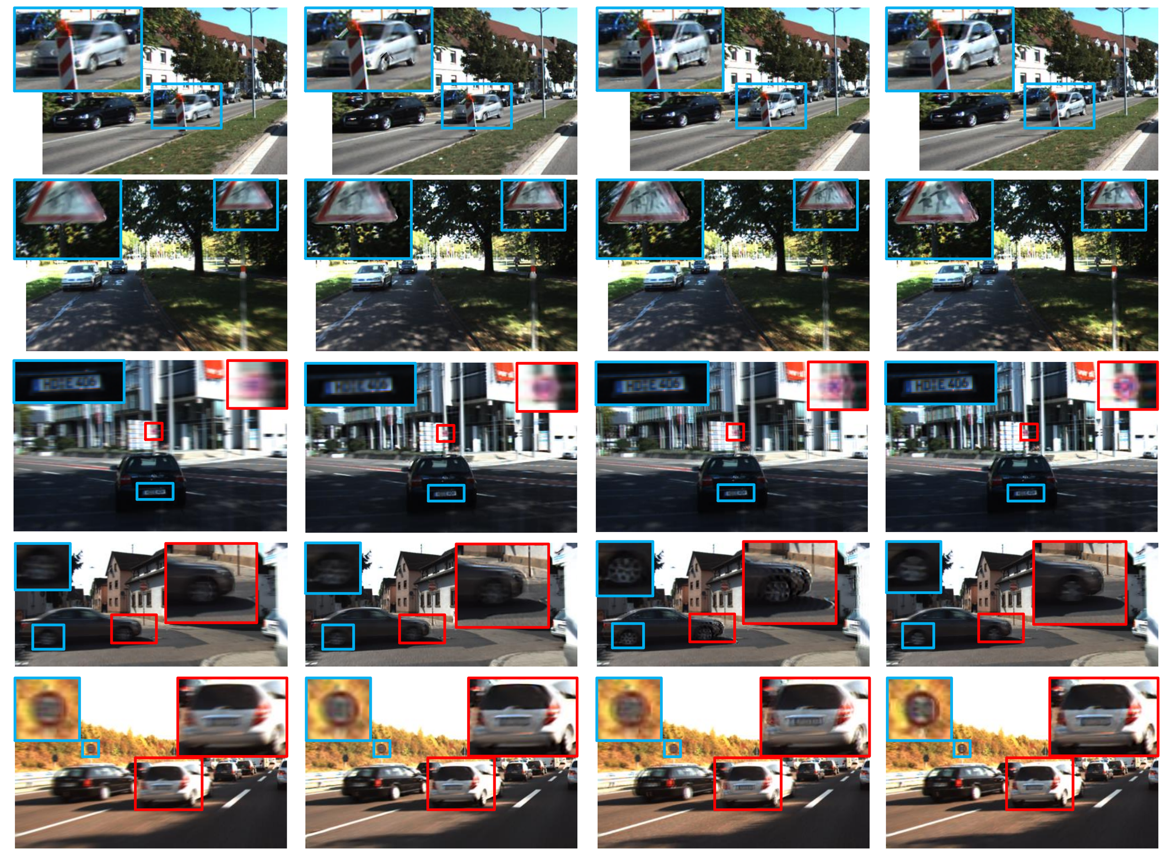
Solution:

Alternatively optimize the scene flow and latent images.

- Fix latent images, solve for scene flow -- Discrete-Continuous Optimization, solved with Tree-reweighted message passing
- Fix scene flow, solve for latent images -- Convex Optimization, solved with Primal-dual

Results

Results on KITTI.



(a) Blur image (b) Kim[3] (c) Sellent[4] (d)Ours

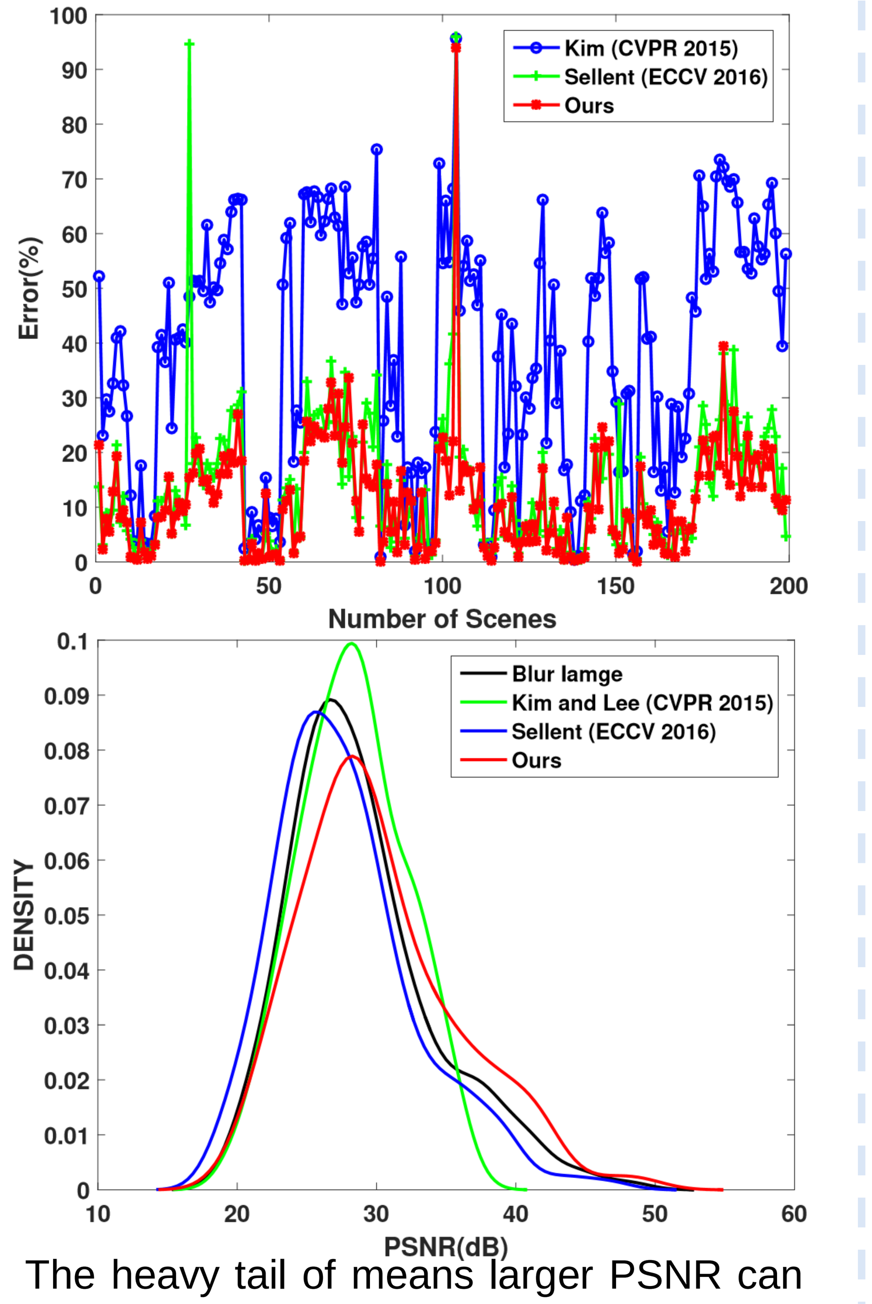
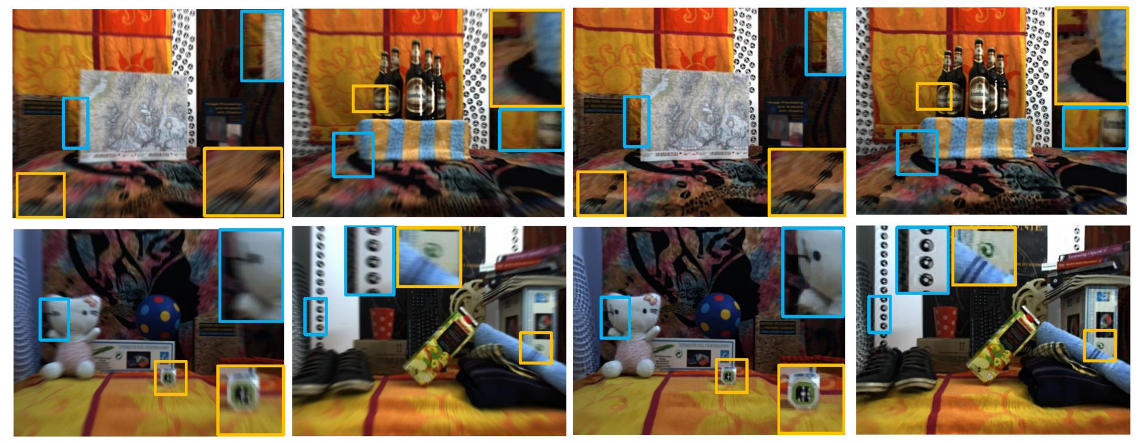


Table 1. Quantitative comparisons on the KITTI dataset.

KITTI Dataset		Disparity(%)		Flow(%)		PSNR(dB)	
		m	m+1	Left	Right	Left	Right
Vogal *et al.* [1]		8.20	8.50	13.62	14.59	/	/
Kim and Lee [3]		/	/	38.89	39.45	28.25	29.00
Sellent *et al.* [4]		8.20	8.50	13.62	14.59	27.75	28.52
Ours	2 Frames	7.02	8.55	11.44	19.34	**30.24**	**30.71**
	3 Frames	**6.82**	**8.36**	**10.01**	**11.45**	29.80	30.30

The heavy tail of means larger PSNR can be achieved using our method.

Results on [4] Dataset



(a) Blur Images (b) Our Results

Table 2. Performance comparisons on "Chair" [5]

Chair video		Disparity (%)	Flow Error(%)	PSNR (dB)
Menze [2]		1.17	9.33	/
Vogel [1]		1.34	2.13	/
Kim [3]		/	9.08	19.95
Sellent [4]		1.34	2.13	23.07
Ours	2 Frames	1.28	1.22	23.13
	3 Frames	**1.15**	**1.18**	**23.26**

Acknowledgement: This work was supported in part by China Scholarship Council (201506290130), Australian Research Council (ARC) grants (DP150104645, DE140100180), and Natural Science Foundation of China (61420106007, 61473230, 61135001). ,and Aviation fund of China(2014ZC53030).

[1] C. Vogel, et al. 3d scene flow estimation with a piecewise rigid scene model. CVPR 2015
[2] M. Menze and A. Geiger. Object scene flow for autonomous vehicles. CVPR 2015
[3] T Hyun Kim and K Mu Lee. Generalized video deblurring for dynamic scenes. CVPR 2015
[4] Anita Sellent, Carsten Rother, and Stefan Roth. Stereo video deblurring. ECCV 2016