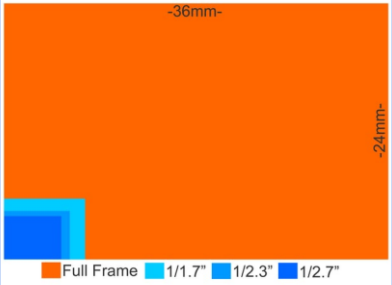
Multi and Single Image Approaches for Latent Image Recovery in Handheld Cameras

Nimisha T M
EE13D037

under the supervision of
Prof. A N Rajagopalan and Prof. R Aravind

Image Processing and Computer Vision Lab
IIT Madras

Main Issues Addressed



Smaller sensor size

Low resolution

Limited aperture size

Motion blur

Modes of Capture and Problems in Handheld Imaging

1 Video Mode

Modes of Capture and Problems in Handheld Imaging

1 Video Mode

2 Burst Mode

Modes of Capture and Problems in Handheld Imaging

1 Video Mode

2 Burst Mode

3 Image Mode

Modes of Capture and Problems in Handheld Imaging

1 Video Mode

2 Burst Mode

3 Image Mode

Increasing difficulty in restoration problems

Modes of Capture and Problems in Handheld Imaging

1 Video Mode

2 Burst Mode

3 Image Mode

Increasing difficulty in restoration problems

Problems Addressed:

Video deblurring for
panning-shots

Modes of Capture and Problems in Handheld Imaging

1 Video Mode

2 Burst Mode

3 Image Mode

Increasing difficulty in restoration problems

Problems Addressed:

Video deblurring for
panning-shots

Joint deblurring and
Super-resolution (SR)
from multiple frames

Modes of Capture and Problems in Handheld Imaging

1 Video Mode

2 Burst Mode

3 Image Mode

Increasing difficulty in restoration problems

Problems Addressed:

Video deblurring for panning-shots

Joint deblurring and Super-resolution (SR) from multiple frames

Single image deblurring

(1) Generating High Quality Panning-Shots from Videos

Problem statement: Given motion blurred videos generate panning-shots

Panning-shots

What are panning-shots?

- Moving object is sharp
- Background is blurred

How to Capture?

What are panning-shots?

- Moving object is sharp
- Background is blurred

How to capture:

- Perfect shutter speed and exposure settings
- Prior information about object motion to pan
- Follow the object in sync

Motivation

What are panning-shots?

- Moving object is sharp
- Background is blurred

Motivation:

- Getting a perfect shot is difficult
- The event might get over by the time settings are made
- A video based method can resolve this

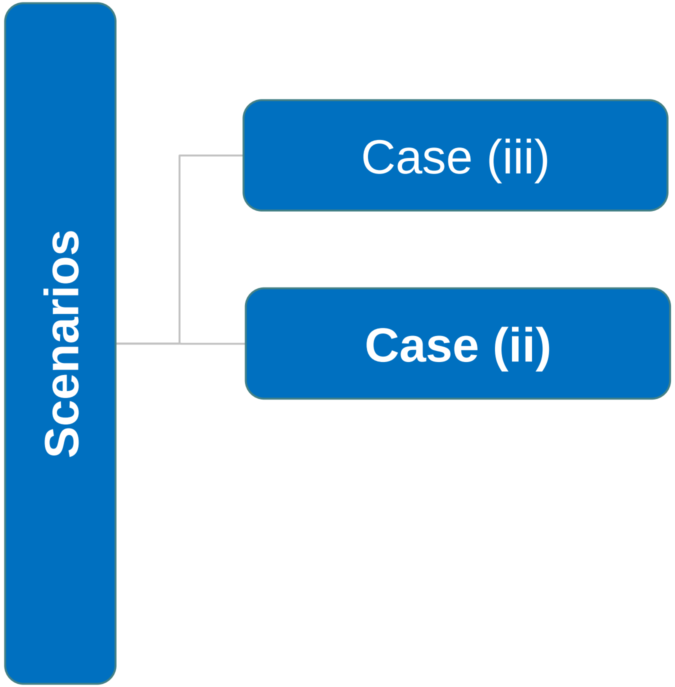
Scenarios Considered

Scenarios

Case (iii)

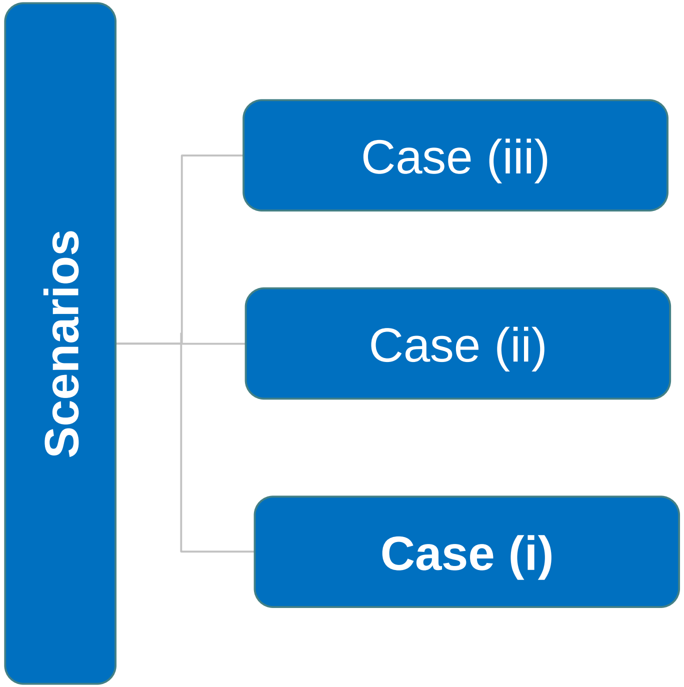
High shutter speed and frame rate/ object velocity very small: All frames are clear

Scenarios Considered



Moderately high shutter speed and frame rate but object velocity is higher: Foreground is blurred while background is clean.

Scenarios Considered



Lower shutter speed/ higher object velocity: Both background and foreground are blurred.

Related Works

Dynamic Video Deblurring

Kim and Mu CVPR 2015: Uses inter-intra frame motion dependency

Hyun Kim, T. and K. Mu Lee, Generalized video deblurring for dynamic scenes. CVPR 2015

Related Works

Dynamic Video Deblurring

Kim and Mu CVPR 2015: Uses inter-intra frame motion dependency

- Not true for general camera motion

Hyun Kim, T. and K. Mu Lee, Generalized video deblurring for dynamic scenes. CVPR 2015

Related Works

Dynamic Video Deblurring

Kim and Mu CVPR 2015: Uses inter-intra frame motion dependency

- Not true for general camera motion

Ma et al. CVPR 2015: Assumption of clean pixels in few frames

Hyun Kim, T. and K. Mu Lee, Generalized video deblurring for dynamic scenes. CVPR 2015
Ma, Z., R. Liao, X. Tao, L. Xu, J. Jia, and E. Wu, Handling motion blur in multi-frame super-resolution. CVPR 2015

Related Works

Dynamic Video Deblurring

Kim and Mu CVPR 2015: Uses inter-intra frame motion dependency

- Not true for general camera motion

Ma et al. CVPR 2015: Assumption of clean pixels in few frames

- All pixels can be blurred

Hyun Kim, T. and K. Mu Lee, Generalized video deblurring for dynamic scenes. CVPR 2015
Ma, Z., R. Liao, X. Tao, L. Xu, J. Jia, and E. Wu, Handling motion blur in multi-frame super-resolution. CVPR 2015

Related Works

Dynamic Video Deblurring

Kim and Mu CVPR 2015: Uses inter-intra frame motion dependency

- Not true for general camera motion

Ma et al. CVPR 2015: Assumption of clean pixels in few frames

- All pixels can be blurred

Nah et al. 2017, Kim et al. 2017: Deep learning based

Nah, S., T. H. Kim, and K. M. Lee, Deep multi-scale convolutional neural network for dynamic scene deblurring. CVPR 2017
Kim, T. H., K. M. Lee, B. Schölkopf, and M. Hirsch, Online video deblurring via dynamic temporal blending network. ICCV 2017

Related Works

Dynamic Video Deblurring

Kim and Mu CVPR 2015: Uses inter-intra frame motion dependency

- Not true for general camera motion

Ma et al. CVPR 2015: Assumption of clean pixels in few frames

- All pixels can be blurred

Nah et al. 2017, Kim et al. 2017: Deep learning based

Panning-shot generation

Liu et al. TOG 2014

- For track shot generation
- User intervention for segmentation
- Does not consider the cases with motion blurred frames

Liu, S., J. Wang, S. Cho, and P. Tan. Trackcam: 3d-aware tracking shots from consumer video. ACM TOG 2014

Related Works

Dynamic Video Deblurring

Kim and Mu CVPR 2015: Uses inter-intra frame motion dependency

- Not true for general camera motion

Ma et al. CVPR 2015: Assumption of clean pixels in few frames

- All pixels can be blurred

Nah et al. 2017, Kim et al. 2017: Deep learning based

Panning-shot generation

Liu et al. TOG 2014

- For track shot generation
- User intervention for segmentation
- Does not consider the cases with motion blurred frames

Our Method: Panning shot from videos with general camera motion

Generating High Quality Panning-Shots from Videos

- **Contributions:**
    - Multi frame background (BG) deblurring

Generating High Quality Panning-Shots from Videos

- **Contributions:**
    - Multi frame background (BG) deblurring
    - Non-blind deblurring of the foreground (FG)

Generating High Quality Panning-Shots from Videos

- **Contributions:**
  - Multi frame background (BG) deblurring
  - Non-blind deblurring of the foreground (FG)
  - Fully automated method for panning-shot generation from videos

Generating High Quality Panning-Shots from Videos

- **Contributions:**
  - Multi frame background (BG) deblurring
  - Non-blind deblurring of the foreground (FG)
  - Fully automated method for panning-shot generation from videos
  - Identify the case using a gradient approach

Generating High Quality Panning-Shots from Videos

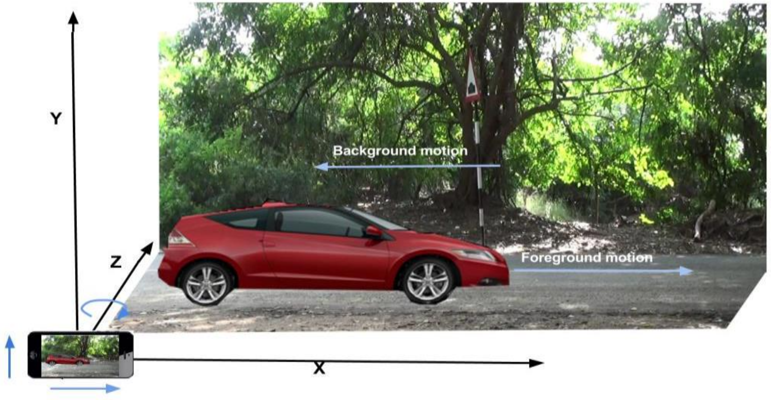
- **Contributions:**
    - Multi frame background (BG) deblurring
    - Non-blind deblurring of the foreground (FG)
    - Fully automated method for panning-shot generation from videos
    - Identify the case using a gradient approach
    - Deal with all the three cases

Generating High Quality Panning-Shots from Videos

- **Contributions:**
  - Multi frame background (BG) deblurring
  - Non-blind deblurring of the foreground (FG)
  - Fully automated method for panning-shot generation from videos
  - Identify the case using a gradient approach
  - Deal with all the three cases

- **Assumptions:**
  - Bilayered scene
  - Object moving parallel to camera plane
  - Camera motion –

    in plane rotation and translations



Proposed Approach

Aim : To synthesize panning-shots from blurred video frames $\{\mathbf{B}^k\}_{k=1}^N$

Proposed Approach

Aim : To synthesize panning-shots from blurred video frames $\{\mathbf{B}^k\}_{k=1}^N$

BG registration, FG motion (v) and relative depth (γ) estimation

Proposed Approach

Aim : To synthesize panning-shots from blurred video frames $\{\mathbf{B}^k\}_{k=1}^{N}$

BG registration, FG motion (v) and relative depth (γ) estimation

BG blur model:

$$\mathbf{b} = \sum_{\mathbf{c}_l \in \mathcal{C}} \omega_{\mathbf{c}_l} \mathbf{H}_{\mathbf{c}_l} \mathbf{f}$$

Proposed Approach

Aim : To synthesize panning-shots from blurred video frames $\{\mathbf{B}^k\}_{k=1}^N$

BG registration, FG motion (v) and relative depth (γ) estimation

BG blur model:

$$\mathbf{b} = \sum_{\mathbf{c}_l \in \mathcal{C}} \omega_{\mathbf{c}_l} \mathbf{H}_{\mathbf{c}_l} \mathbf{f}$$

Multi-frame BG deblurring

Estimate $(\boldsymbol{\omega})$ and clean BG

Proposed Approach

Aim : To synthesize panning-shots from blurred video frames $\{\mathbf{B}^k\}_{k=1}^N$

BG registration, FG motion (v) and relative depth (γ) estimation

BG blur model:

$$\mathbf{b} = \sum_{\mathbf{c}_l \in \mathcal{C}} \omega_{\mathbf{c}_l} \mathbf{H}_{\mathbf{c}_l} \mathbf{f}$$

Multi-frame BG deblurring

Estimate $(\boldsymbol{\omega})$ and clean BG

Estimate FG blur from
BG blur and object motion

Proposed Approach

Aim : To synthesize panning-shots from blurred video frames $\{\mathbf{B}^k\}_{k=1}^{N}$

BG registration, FG motion (v) and relative depth (γ) estimation

BG blur model:

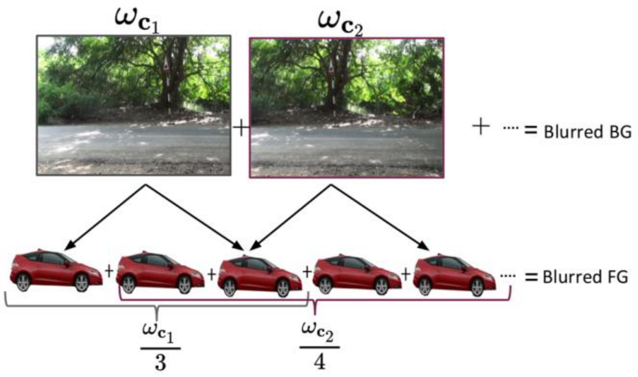
$$\mathbf{b} = \sum_{\mathbf{c}_l \in \mathcal{C}} \omega_{\mathbf{c}_l} \mathbf{H}_{\mathbf{c}_l} \mathbf{f}$$

FG blur model:

$$\mathbf{b}_{\mathrm{FG}} = \sum_{\mathbf{m}_p \in \mathcal{M}} w_{\mathbf{m}_p} \mathbf{H}_{\mathbf{m}_p} \mathbf{f}_{\mathrm{FG}}$$

Multi-frame BG deblurring

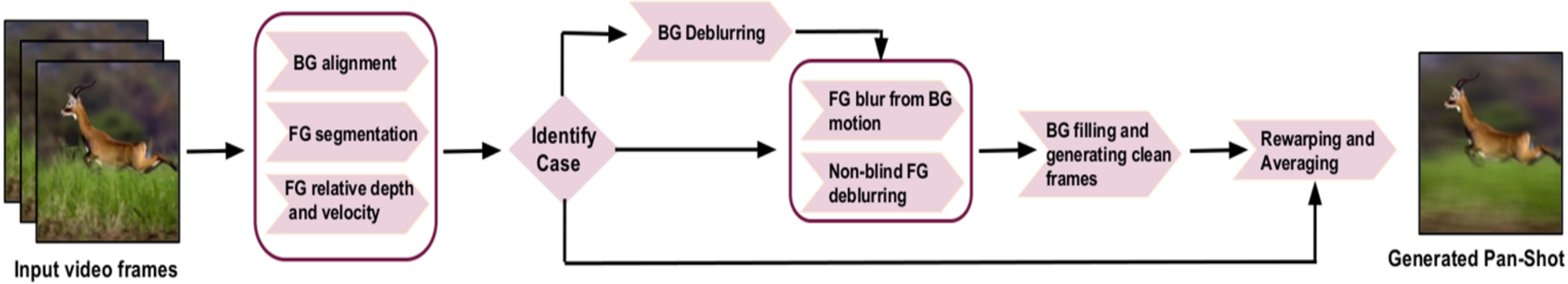
Estimate (ω) and clean BG



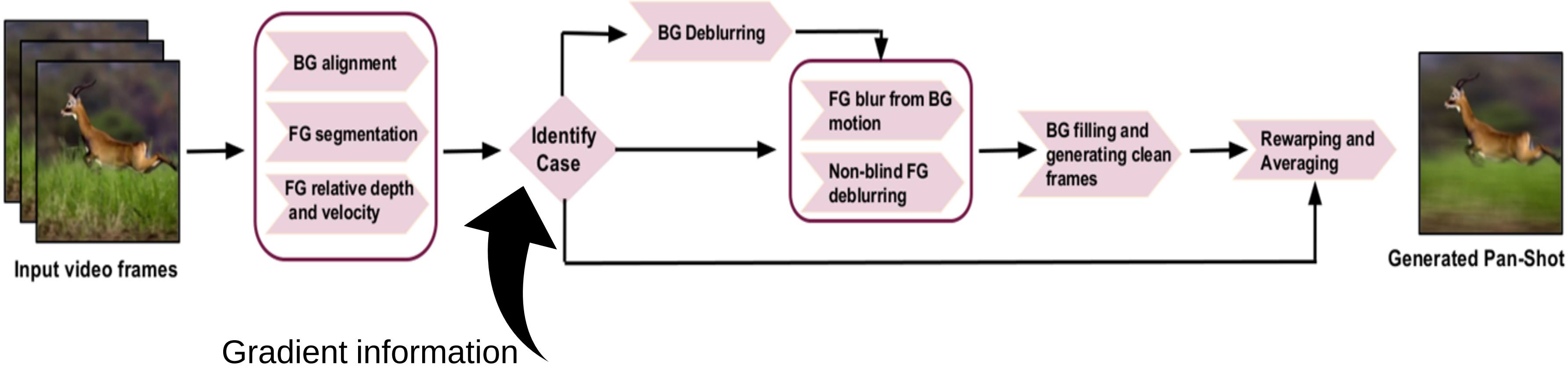
Estimate FG blur from
BG blur and object motion

Non-blind FG deblurring

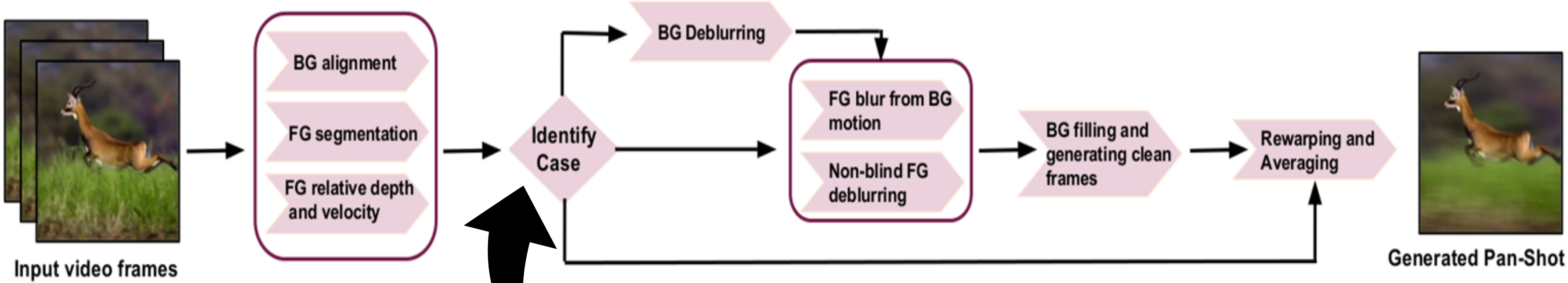
Proposed Approach- Flow chart



Proposed Approach- Flow chart

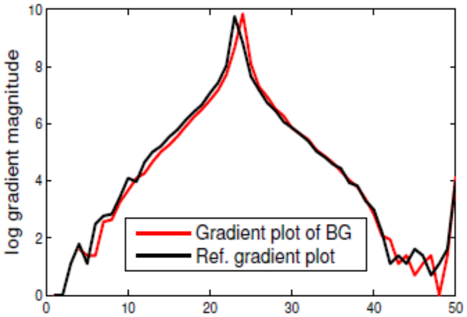


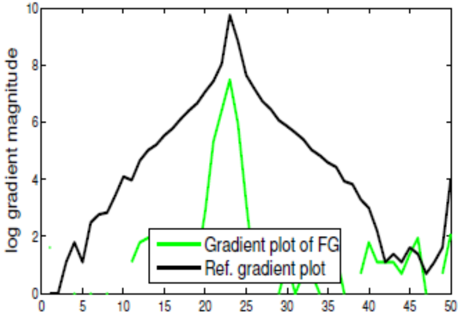
Proposed Approach- Flow chart



Gradient information

Case (ii)





Quantitative Results

PSNR: Peak Signal to Noise Ratio
SSIM: Structural SImilarity Index

PSNR/SSIM

Methods	Kim & Mu 2015	Ma et al. 2015	Nah et al. 2017	Ours
FG	26.14/0.8252			
BG	25.77/0.8762			


GT




Quantitative Results

Methods	Kim & Mu 2015	Ma et al. 2015	Nah et al. 2017	Ours
FG	26.14/0.8252	24.756/0.8036		
BG	25.77/0.8762	24.11/0.8407		

GT

Quantitative Results

Methods	Kim & Mu 2015	Ma et al. 2015	Nah et al. 2017	Ours
FG	26.14/0.8252	24.756/0.8036	27.12/0.8617	
BG	25.77/0.8762	24.11/0.8407	25.05/0.8127	

GT

Quantitative Results

Methods	Kim & Mu 2015	Ma et al. 2015	Nah et al. 2017	Ours
FG	26.14/0.8252	24.756/0.8036	27.12/0.8617	**27.35/0.8668**
BG	25.77/0.8762	24.11/0.8407	25.05/0.8127	**25.86/0.7979**

GT

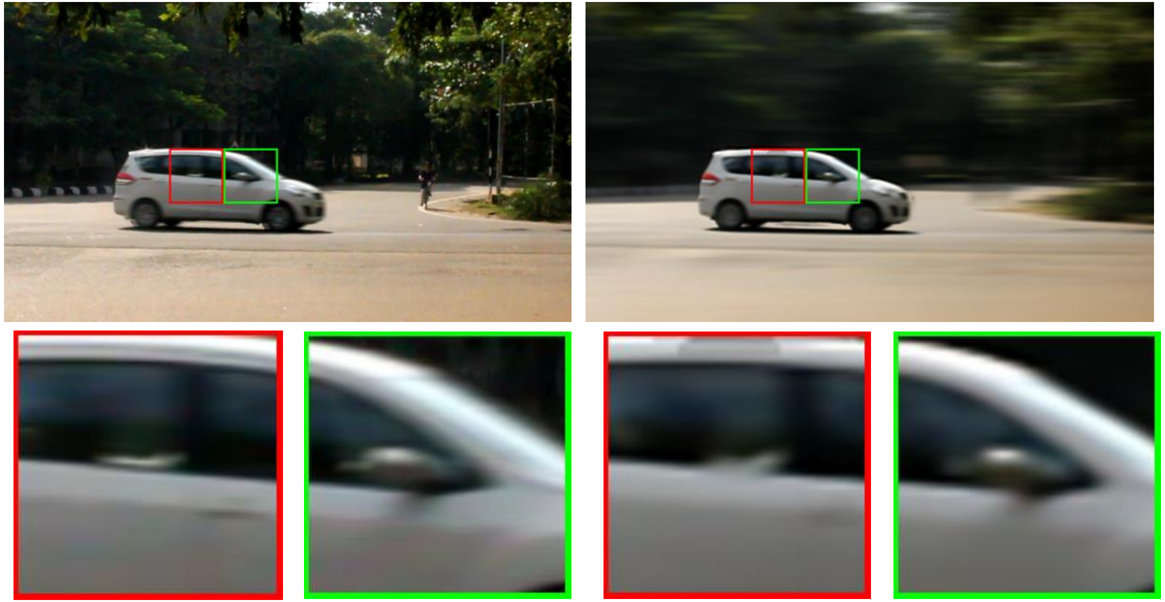
Qualitative Results

Input frame

Qualitative Results

Input frame

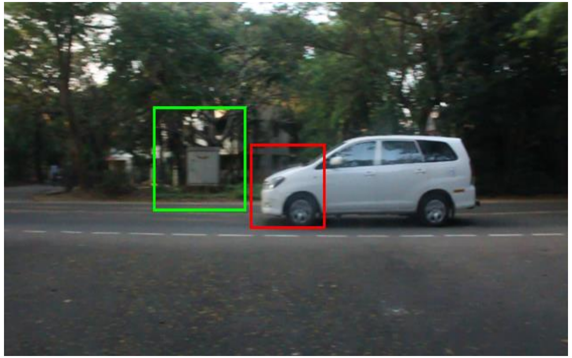
Kim & Mu CVPR 2015



Qualitative Results

Qualitative Results

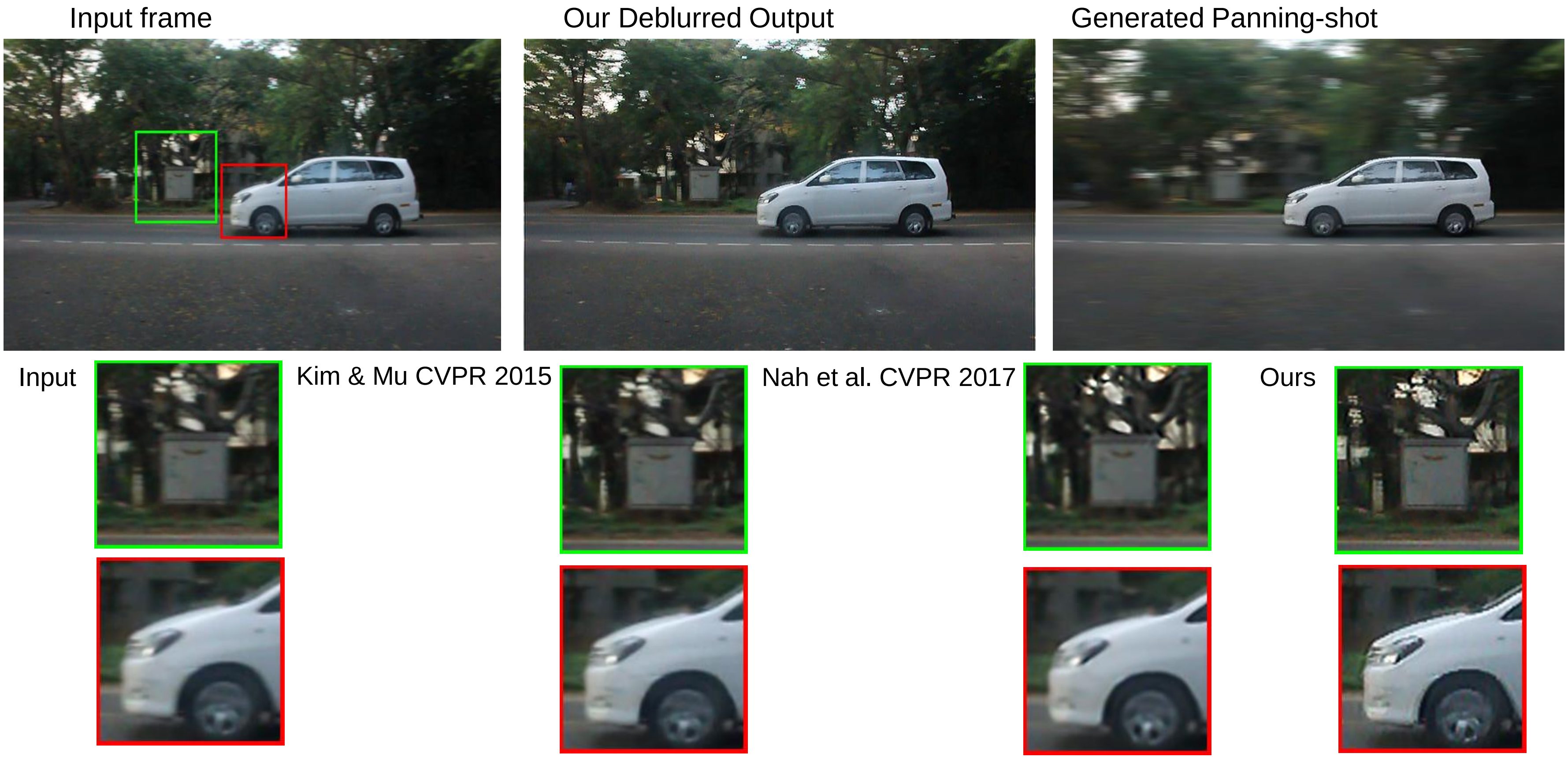
Input frame



Our Deblurred Output

Generated Panning-shot

Qualitative Results



T. M. Nimisha, A. N. Rajagopalan, and R. Aravind, “Generating High Quality Pan-Shots from Motion Blurred Videos,” Elsevier Journal: Computer Vision and Image Understanding (CVIU), vol. 171, pp. 20-33, June 2018.

(2) Multi-Shot Blind Super-resolution

Problem Statement: Joint deblurring and SR of 3D scenes from multiple motion blurred frames

Related Works

(1) SR from motion blur

Sroubek et al. TIP 2007 : Planar scene and Convolutional blur
Ma et al. CVPR 2015 : Planar scene and Space-varying blur
Assumes availability of clean pixels in frames

Sroubek, F., G. Cristobal, and J. Flusser . A unified approach to superresolution and multichannel blind deconvolution. TIP 2007

Related Works

(1) SR from motion blur

Sroubek et al. TIP 2007 : Planar scene and Convolutional blur

Ma et al. CVPR 2015 : Planar scene and Space-varying blur

Assumes availability of clean pixels in frames

(2) 3D SR

Mudenagudi et al. 2007, Bhavsar and Rajagopalan 2010: Estimates depth and SR image from clean frames

Mudenagudi, U., A. Gupta, L. Goel, A. Kushal, P. Kalra, and S. Banerjee, Super resolution of images of 3d scenes. ACCV 2007
Bhavsar, A. V. and A. N. Rajagopalan. Resolution enhancement in multi-image stereo. TPAMI 2010

Related Works

(1) SR from motion blur

Sroubek et al. TIP 2007 : Planar scene and Convolutional blur
Ma et al. CVPR 2015 : Planar scene and Space-varying blur
Assumes availability of clean pixels in frames

(2) 3D SR

Mudenagudi et al. 2007, Bhavsar and Rajagopalan 2010: Estimates depth and SR image from clean frames

Our Method: Joint deblurring and SR for space-variant blurred 3D scenes

Multi-Shot Blind Super-resolution

Contributions:

- Joint framework
  - Estimate latent High Resolution (HR) image
  - Depth map of a 3D scene
  - Camera motion from non-uniformly blurred Low Resolution (LR)observations
- Elegant patch-based approach to compute the global HR camera motion

Multi-Shot Blind Super-resolution

Contributions:

- Joint framework
  - Estimate latent High Resolution (HR) image
  - Depth map of a 3D scene
  - Camera motion from non-uniformly blurred Low Resolution (LR)observations
- Elegant patch-based approach to compute the global HR camera motion

Assumptions:

- Layered scene
- Camera motion- in plane translations and rotations
- Two frames with major translational motion
- No dynamic objects

Proposed Approach

Aim : Obtaining an HR frame and depth map from motion blurred LR observations $\{\mathbf{g}^k\}_{k=1}^{K}$

Planar Scene LR: $$\mathbf{g} = \mathbf{D}_\epsilon \left(\sum_{\mathbf{c}_l \in \mathcal{C}} \omega_{\mathbf{c}_l} \mathbf{H}_{\mathbf{c}_l} \mathbf{f} \right)$$

Proposed Approach

Aim : Obtaining an HR frame and depth map from motion blurred LR observations $\{\mathbf{g}^k\}_{k=1}^{K}$

Planar Scene LR: $$\mathbf{g} = \mathbf{D}_\epsilon \left(\sum_{\mathbf{c}_l \in \mathcal{C}} \omega_{\mathbf{c}_l} \mathbf{H}_{\mathbf{c}_l} \mathbf{f} \right)$$

3D Scene LR:

3D layered scene:

$$\mathbf{f} = \sum_{r=1}^{R} \mathbf{f}_r$$

Proposed Approach

Aim : Obtaining an HR frame and depth map from motion blurred LR observations $\{\mathbf{g}^k\}_{k=1}^{K}$

Planar Scene LR: $$\mathbf{g} = \mathbf{D}_\epsilon \left(\sum_{\mathbf{c}_l \in \mathcal{C}} \omega_{\mathbf{c}_l} \mathbf{H}_{\mathbf{c}_l} \mathbf{f} \right)$$

3D Scene LR:

3D layered scene:

$$\mathbf{f} = \sum_{r=1}^{R} \mathbf{f}_r$$

$$\mathbf{g} = \mathbf{D}_\epsilon \left(\sum_{\mathbf{c}_l \in \mathcal{C}} \omega_{\mathbf{c}_l} \left(\sum_{r=1}^{R} \mathbf{H}_{(\delta_r, \mathbf{c}_l)} \mathbf{f}_r \right) \right)$$

Proposed Approach

Aim : Obtaining an HR frame and depth map from motion blurred LR observations $\{\mathbf{g}^k\}_{k=1}^{K}$

Planar Scene LR: $$\mathbf{g} = \mathbf{D}_\epsilon \left(\sum_{\mathbf{c}_l \in \mathcal{C}} \omega_{\mathbf{c}_l} \mathbf{H}_{\mathbf{c}_l} \mathbf{f} \right)$$

3D Scene LR:

3D layered scene:

$$\mathbf{f} = \sum_{r=1}^{R} \mathbf{f}_r$$

$$\mathbf{g} = \mathbf{D}_\epsilon \left(\sum_{\mathbf{c}_l \in \mathcal{C}} \omega_{\mathbf{c}_l} \left(\sum_{r=1}^{R} \mathbf{H}_{(\delta_r, \mathbf{c}_l)} \mathbf{f}_r \right) \right)$$

Unknowns :

Camera motion

Proposed Approach

Aim : Obtaining an HR frame and depth map from motion blurred LR observations $\{\mathbf{g}^k\}_{k=1}^{K}$

Planar Scene LR: $$\mathbf{g} = \mathbf{D}_\epsilon \left(\sum_{\mathbf{c}_l \in \mathcal{C}} \omega_{\mathbf{c}_l} \mathbf{H}_{\mathbf{c}_l} \mathbf{f} \right)$$

3D Scene LR:

3D layered scene:

$$\mathbf{f} = \sum_{r=1}^{R} \mathbf{f}_r$$

$$\mathbf{g} = \mathbf{D}_\epsilon \left(\sum_{\mathbf{c}_l \in \mathcal{C}} \omega_{\mathbf{c}_l} \left(\sum_{r=1}^{R} \mathbf{H}_{(\delta_r, \mathbf{c}_l)} \mathbf{f}_r \right) \right)$$

Unknowns :

Camera motion

Depth

Proposed Approach

Aim : Obtaining an HR frame and depth map from motion blurred LR observations $\{\mathbf{g}^k\}_{k=1}^K$

Planar Scene LR: $$\mathbf{g} = \mathbf{D}_\epsilon \left(\sum_{\mathbf{c}_l \in \mathcal{C}} \omega_{\mathbf{c}_l} \mathbf{H}_{\mathbf{c}_l} \mathbf{f} \right)$$

3D Scene LR:

3D layered scene:

$$\mathbf{f} = \sum_{r=1}^{R} \mathbf{f}_r$$

$$\mathbf{g} = \mathbf{D}_\epsilon \left(\sum_{\mathbf{c}_l \in \mathcal{C}} \omega_{\mathbf{c}_l} \left(\sum_{r=1}^{R} \mathbf{H}_{(\delta_r, \mathbf{c}_l)} \mathbf{f}_r \right) \right)$$

Unknowns :

Camera motion

Depth

HR image

Proposed Approach

Aim : Obtaining an HR frame and depth map from motion blurred LR observations $\{\mathbf{g}^k\}_{k=1}^K$

- ❑ **Camera motion estimation**
  - Initial depth map
- ❑ **HR frame estimation**
- ❑ **Depth map refinement**

Proposed Approach

Aim : Obtaining an HR frame and depth map from motion blurred LR observations $\{\mathbf{g}^k\}_{k=1}^{K}$

- ❑ **Camera motion estimation**

  Initial depth map

- ❑ **HR frame estimation**
- ❑ **Depth map refinement**

Estimate initial depth with optical flow

Proposed Approach

Aim : Obtaining an HR frame and depth map from motion blurred LR observations $\{\mathbf{g}^k\}_{k=1}^{K}$

- **Camera motion estimation**
  - Initial depth map
- **HR frame estimation**
- **Depth map refinement**

Estimate initial depth with optical flow

Camera motion w.r.t ref depth from LR patches

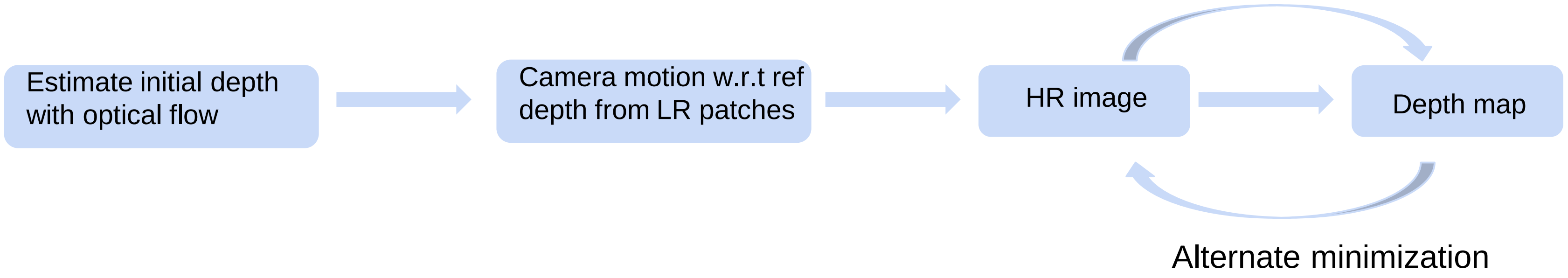
Proposed Approach

Aim : Obtaining an HR frame and depth map from motion blurred LR observations $\{\mathbf{g}^k\}_{k=1}^K$

- ❑ **Camera motion estimation**

  Initial depth map

- ❑ **HR frame estimation**
- ❑ **Depth map refinement**



Proposed Approach

Aim : Obtaining an HR frame and depth map from motion blurred LR observations $\{\mathbf{g}^k\}_{k=1}^K$

Camera motion estimation

- Initial Depth using Optical flow

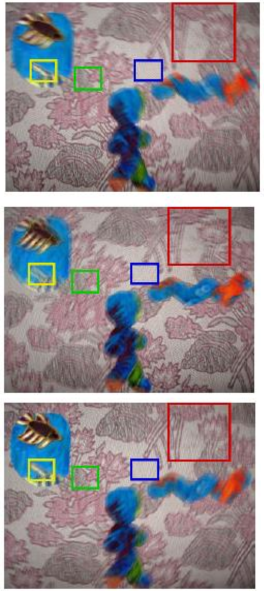
    Estimate optical flow between LR frames and choose the flow corresponding to least rotation

    This flow magnitude is taken as initial depth

- Pick a depth layer
- Estimate HR PSF’s at the picked layer using Sroubek et al. TIP 2007
- From HR PSF estimate HR global camera motion w.r.t the layer selected

Initial depth map

LR frames



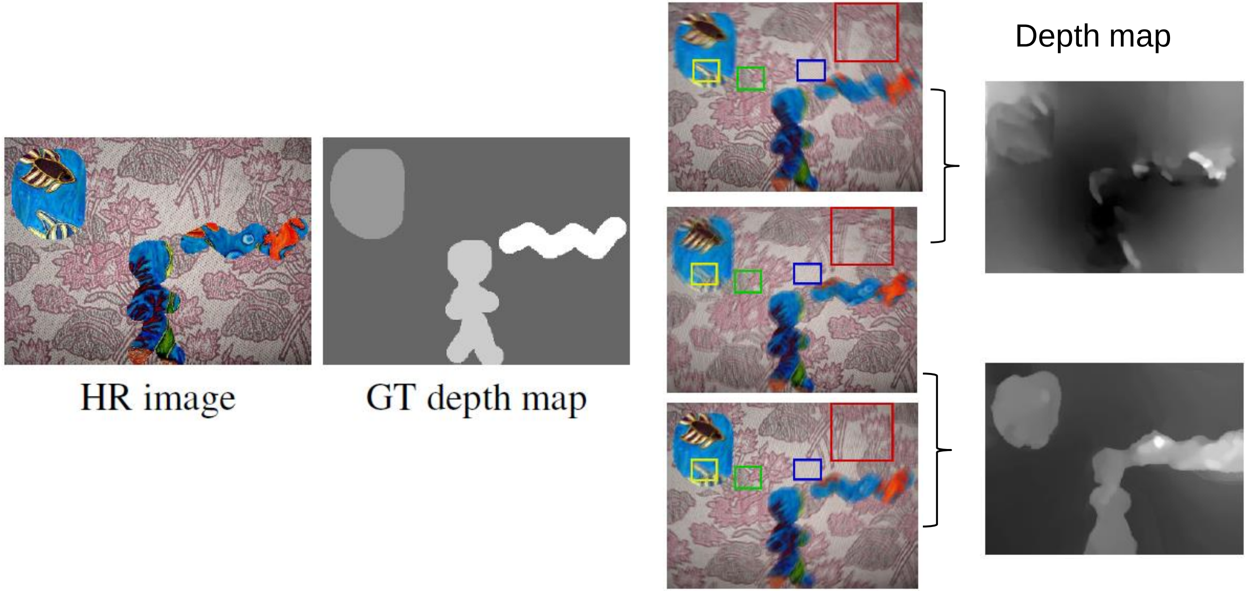
HR image

GT depth map

Initial depth map

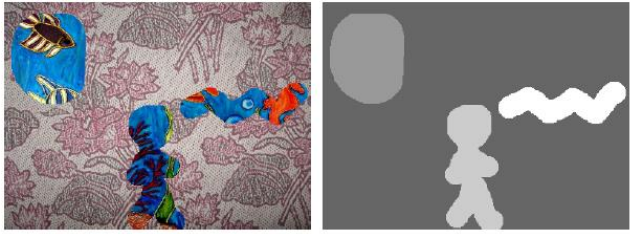
LR frames

Depth map



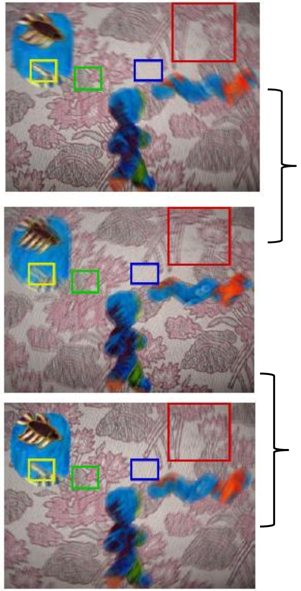
HR image GT depth map

Initial depth map

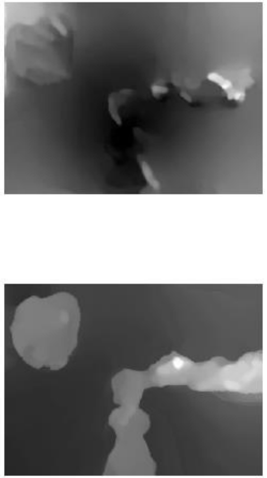


HR image GT depth map

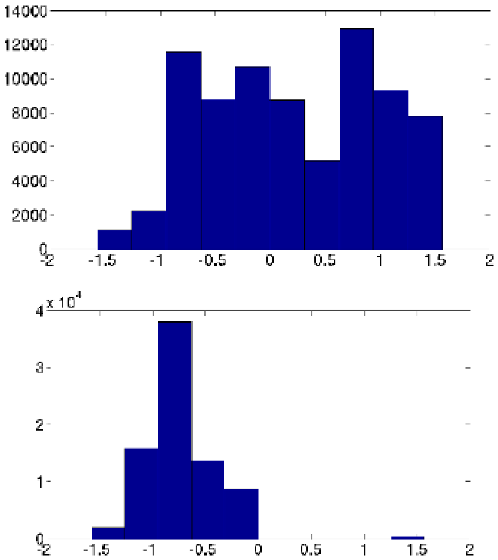
LR frames



Depth map



Histogram of phase of optical flow



Proposed Approach

Aim : Obtaining an HR frame and depth map from motion blurred LR observations $\{\mathbf{g}^k\}_{k=1}^{K}$

- **HR frame estimation**

$$E(\mathbf{f}) = \sum_{k=1}^{K} ||\mathbf{W}^k(\mathbf{D}_\epsilon \boldsymbol{\mathcal{H}}^k \mathbf{f} - \mathbf{g}^k)||_2^2 + \lambda \mathbf{f}^T \mathbf{L} \mathbf{f}$$

Binary modulating function for pixel visibility

Proposed Approach

Aim : Obtaining an HR frame and depth map from motion blurred LR observations $\{\mathbf{g}^k\}_{k=1}^{K}$

- **HR frame estimation**

$$E(\mathbf{f}) = \sum_{k=1}^{K} ||\mathbf{W}^k(\mathbf{D}_\epsilon \boldsymbol{\mathcal{H}}^k \mathbf{f} - \mathbf{g}^k)||_2^2 + \lambda \mathbf{f}^T \mathbf{L} \mathbf{f}$$

Formed from camera motion
and initial depth map

Proposed Approach

Aim : Obtaining an HR frame and depth map from motion blurred LR observations $\{\mathbf{g}^k\}_{k=1}^{K}$

- ❑ **HR frame estimation**

$$E\left(\mathbf{f}\right) = \sum_{k=1}^{K} ||\mathbf{W}^k(\mathbf{D}_\epsilon \boldsymbol{\mathcal{H}}^k \mathbf{f} - \mathbf{g}^k)||_2^2 + \lambda \mathbf{f}^T \mathbf{L} \mathbf{f}$$

Prior on image gradient

Proposed Approach

Aim : Obtaining an HR frame and depth map from motion blurred LR observations $\{\mathbf{g}^k\}_{k=1}^{K}$

❑ HR frame estimation

$$E(\mathbf{f}) = \sum_{k=1}^{K} ||\mathbf{W}^k(\mathbf{D}_\epsilon \boldsymbol{\mathcal{H}}^k \mathbf{f} - \mathbf{g}^k)||_2^2 + \lambda \mathbf{f}^T \mathbf{L} \mathbf{f}$$

Prior on image gradient

Solved using conjugate gradient method

Proposed Approach

Aim : Obtaining an HR frame and depth map from motion blurred LR observations $\{\mathbf{g}^k\}_{k=1}^{K}$

❑ **Depth map refinement**

$$E(\delta_{r_{\mathbf{y}}}) = \sum_{k=1}^{K} \Bigg(\mathbf{g}^k(\mathbf{x}) -$$

LR pixel location

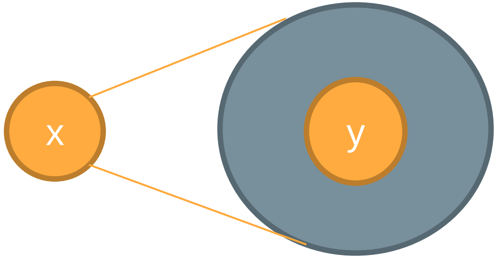
Proposed Approach

Aim : Obtaining an HR frame and depth map from motion blurred LR observations $\{\mathbf{g}^k\}_{k=1}^{K}$

- **Depth map refinement**

$$E(\delta_{r_{\mathbf{y}}}) = \sum_{k=1}^{K} \Bigg(\mathbf{g}^k(\mathbf{x}) -$$

LR pixel location



Proposed Approach

Aim : Obtaining an HR frame and depth map from motion blurred LR observations $\{\mathbf{g}^k\}_{k=1}^{K}$

- ❑ **Depth map refinement**

$$E(\delta_{r_{\mathbf{y}}}) = \sum_{k=1}^{K} \left(\mathbf{g}^k(\mathbf{x}) - \mathbf{D}_\epsilon \left(\left(\sum_{\mathbf{c}_l \in \mathcal{C}} \omega^k_{\mathbf{c}_l} \mathbf{H}^k_{(\delta_{r_{\mathbf{y}}}, \mathbf{c}_l)} \mathbf{f}^t \right)(\mathbf{y}) \right.\right.$$

Proposed Approach

Aim : Obtaining an HR frame and depth map from motion blurred LR observations $\{\mathbf{g}^k\}_{k=1}^{K}$

- ❑ **Depth map refinement**

$$E(\delta_{r_{\mathbf{y}}}) = \sum_{k=1}^{K} \left(\mathbf{g}^k(\mathbf{x}) - \mathbf{D}_\epsilon \left(\left(\sum_{\mathbf{c}_l \in \mathcal{C}} \omega_{\mathbf{c}_l}^k \mathbf{H}_{(\delta_{r_{\mathbf{y}}}, \mathbf{c}_l)}^k \mathbf{f}^t \right)(\mathbf{y}) + \sum_{\underline{\mathbf{y}} \in \mathcal{N}} \left(\sum_{\mathbf{c}_l \in \mathcal{C}} \omega_{\mathbf{c}_l}^k \mathbf{H}_{(\delta_{r_{\underline{\mathbf{y}}}}, \mathbf{c}_l)}^k \mathbf{f}^t \right)(\underline{\mathbf{y}}) \right) \right)$$

HR pixel neighbourhood

Proposed Approach

Aim : Obtaining an HR frame and depth map from motion blurred LR observations $\{\mathbf{g}^k\}_{k=1}^K$

❑ **Depth map refinement**

$$E(\delta_{r_\mathbf{y}}) = \sum_{k=1}^{K}\left(\mathbf{g}^k(\mathbf{x}) - \mathbf{D}_\epsilon\left(\left(\sum_{\mathbf{c}_l \in \mathcal{C}} \omega_{\mathbf{c}_l}^k \mathbf{H}_{(\delta_{r_\mathbf{y}},\mathbf{c}_l)}^k \mathbf{f}^t\right)(\mathbf{y}) + \sum_{\underline{\mathbf{y}} \in \mathcal{N}}\left(\sum_{\mathbf{c}_l \in \mathcal{C}} \omega_{\mathbf{c}_l}^k \mathbf{H}_{(\delta_{r_{\underline{\mathbf{y}}}},\mathbf{c}_l)}^k \mathbf{f}^t\right)(\underline{\mathbf{y}})\right)\right)$$

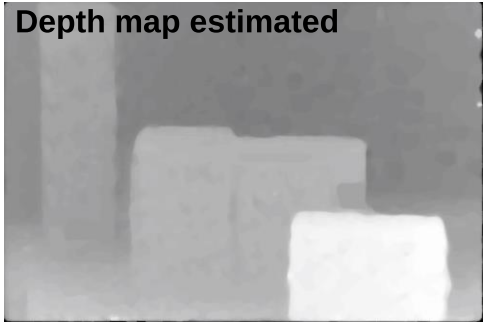
$$+ \sum_{\underline{\mathbf{y}} \in \mathcal{N}} \mu \min(|\delta_{r_\mathbf{y}} - \delta_{r_{\underline{\mathbf{y}}}}|, \beta)$$

Smoothness of depth map

Results

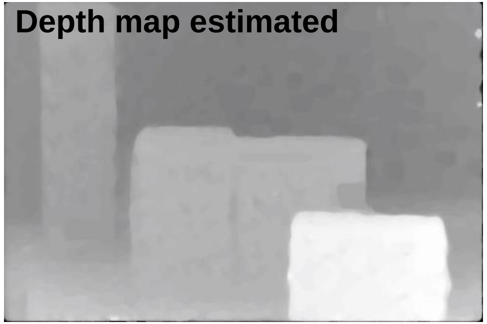
Blurred low resolution input images

Results



Input LR

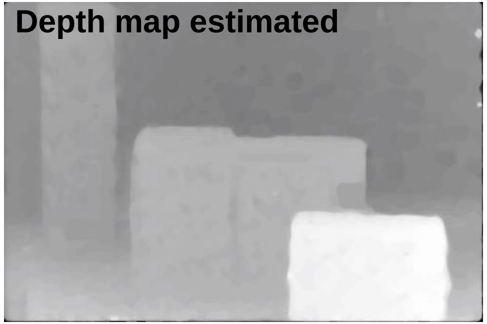
Results



Input LR

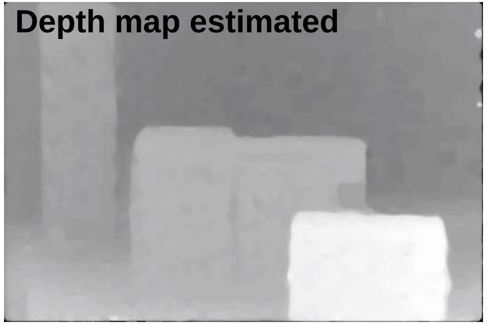
Sroubek et al. TIP 2007

Results



Input LR | Ma et al. CVPR 2015 | Sroubek et al. TIP 2007

Results



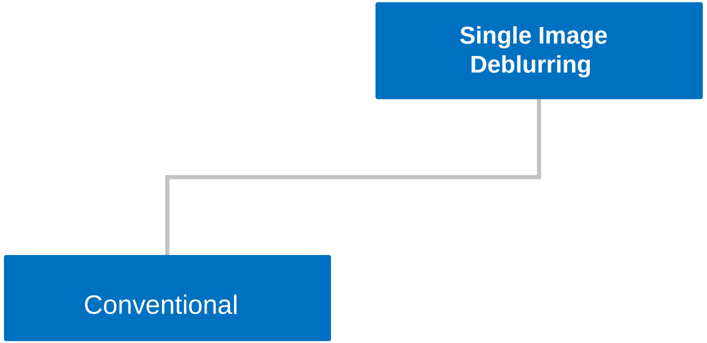
Input LR | Ma et al. CVPR 2015 | Sroubek et al. TIP 2007 | Ours

Abhijith Punnappurath, T. M. Nimisha, and A.N. Rajagopalan "Multi-image blind super-resolution of 3D scenes," IEEE Transactions on Image Processing, vol. 26, No. 11, pp. 5337-5352, November 2017.

(3) Dictionary Replacement for Single Image Deblurring

Problem Statement: Deblurring and depth estimation from a single blurred image

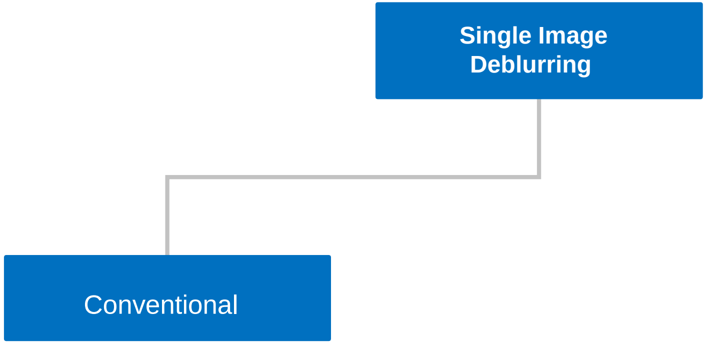
Related Works



Pan et al. CVPR 2016

Pan, J., Sun, D., Pfister, H., Yang, M.H.: Blind image deblurring using dark channel prior. CVPR 2016

Related Works



Challenges:

- Highly ill-posed problem
- Too many unknowns to evaluate
- Requires prior

Related Works

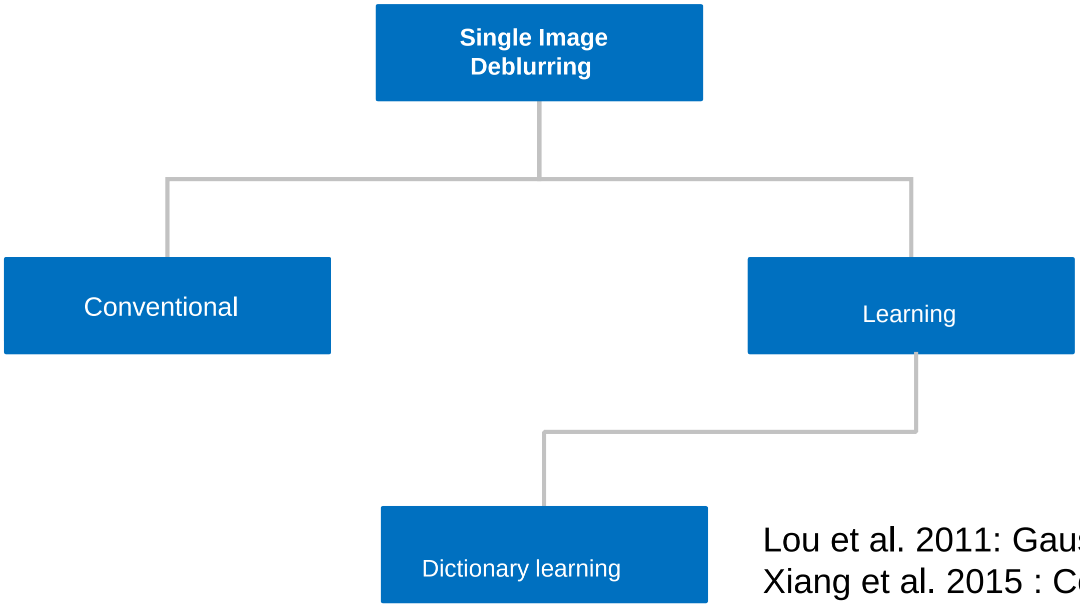
Single Image Deblurring

Conventional

Drawbacks:

- Selection of prior
- Selection of prior weightage
- Slow optimization
- Does not estimate depth maps

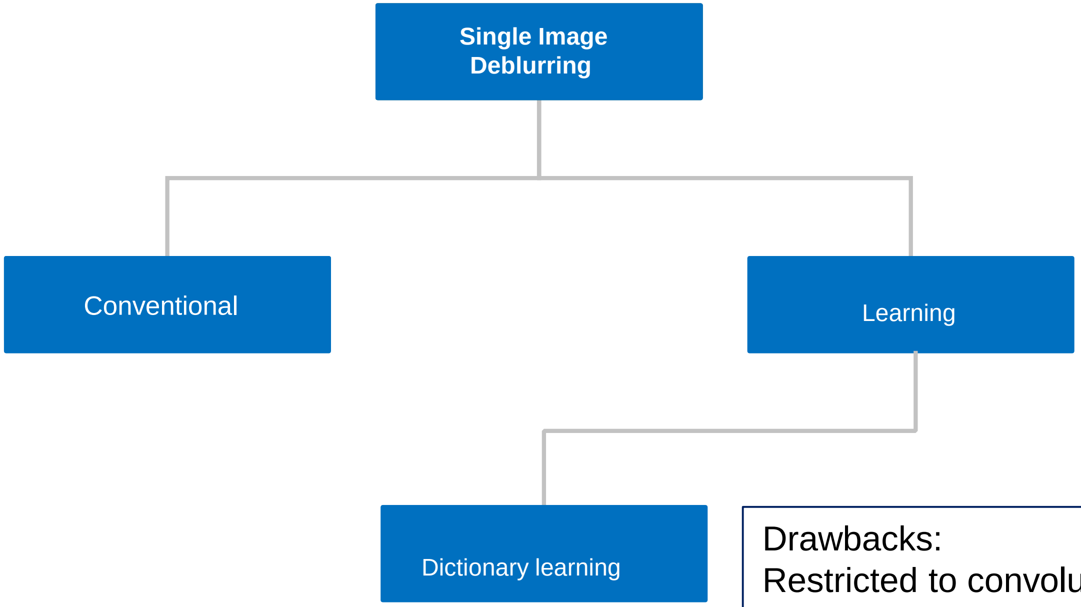
Related Works



Lou et al. 2011: Gaussian deblurring
Xiang et al. 2015 : Coupled dictionary based deblurring

Lou, Y., A. L. Bertozzi, and S. Soatto (2011). Direct sparse deblurring. Journal of Mathematical Imaging and Vision, 39(1), 1–12.
Xiang, S., G. Meng, Y. Wang, C. Pan, and C. Zhang . Image deblurring with coupled dictionary learning. IJCV 2015

Related Works



Drawbacks:
Restricted to convolutional blur
Fails for 3D scenes
Learns dictionary for each blur type

Related Works

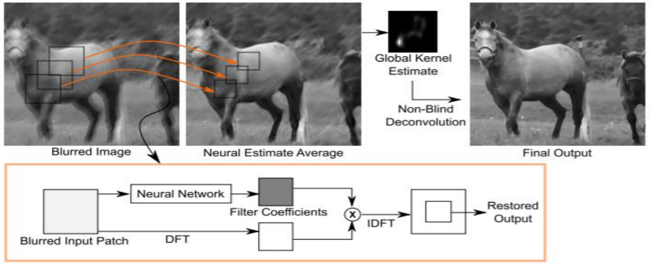
Single Image Deblurring

Conventional

Learning

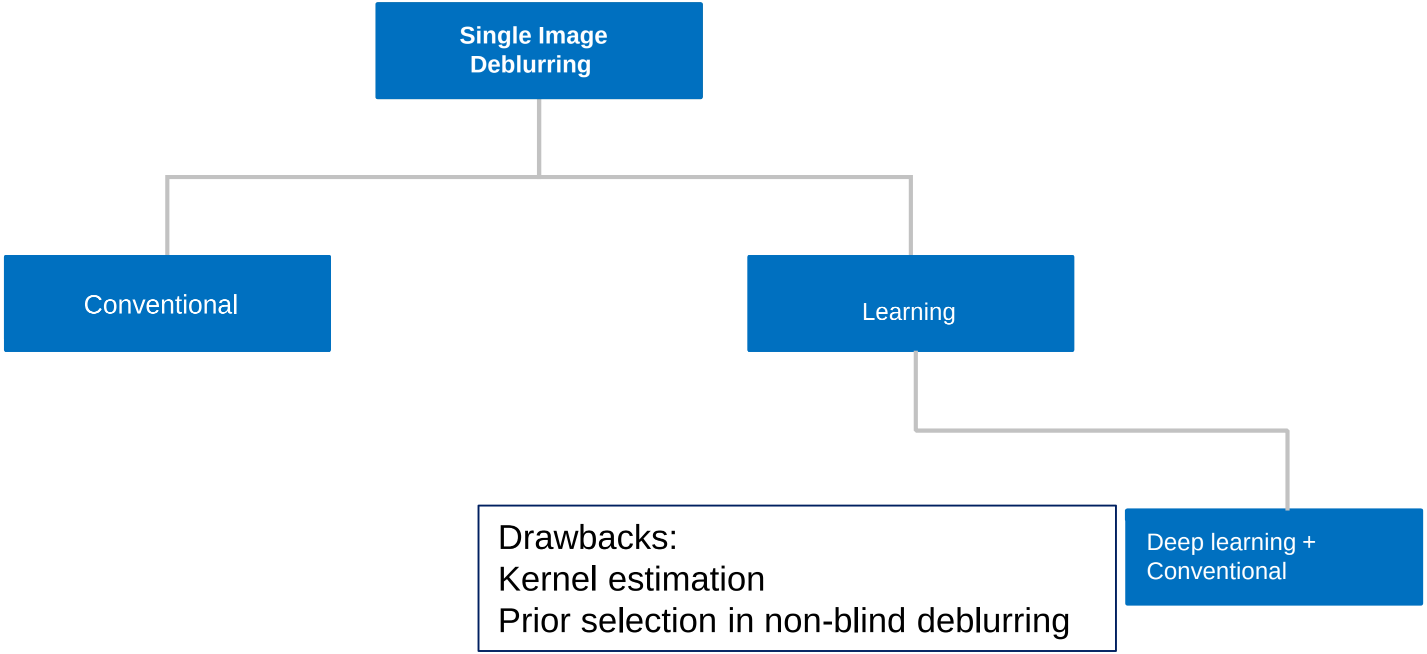
Deep learning + Conventional

Chakrabarti. 2016



Chakrabarti, A., A neural approach to blind motion deblurring. ECCV 2016

Related Works



Related Works

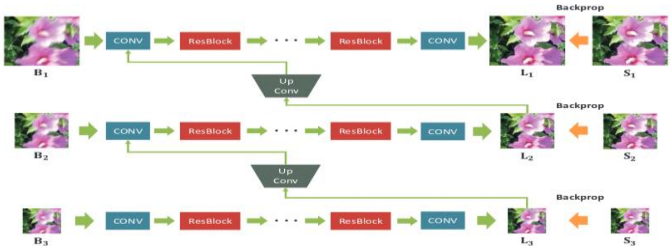
Single Image Deblurring

Conventional

Learning

Deep learning

Nah et al. 2017



Nah, S., Kim, T.H., Lee, K.M.: Deep multi-scale convolutional neural network for dynamic scene deblurring. In: CVPR 2017

Related Works

Single image deblurring + depth:

Hu et al. 2014 : Camera motion blur only
Solve for segment-wise depth with user-assisted segmentation

Hu, Z., L. Xu, and M.-H. Yang, Joint depth estimation and camera shake removal from single blurry image. In: CVPR 2014

Related Works

Single image deblurring + depth:

Hu et al. 2014 : Camera motion blur only
Solve for segment-wise depth with user-assisted segmentation

Our method:

- Depth and deblurring from single blurred frame

Related Works

Single image deblurring + depth:

Hu et al. 2014 : Camera motion blur only
Solve for segment-wise depth with user-assisted segmentation

Our method:

- Depth and deblurring from single blurred frame
- No user input

Related Works

Single image deblurring + depth:

Hu et al. 2014 : Camera motion blur only
Solve for segment-wise depth with user-assisted segmentation

Our method:

- Depth and deblurring from single blurred frame
- No user input
- Works irrespective of the blur type (for which blur-depth assumption holds)

Dictionary Replacement for Single Image Deblurring

- **Contributions**:
  - First attempt in deblurring and depth estimation from space-variant blur using dictionary replacement

Dictionary Replacement for Single Image Deblurring

- **Contributions**:
  - First attempt in deblurring and depth estimation from space-variant blur using dictionary replacement
  - Works irrespective of blur type (motion/defocus)

Dictionary Replacement for Single Image Deblurring

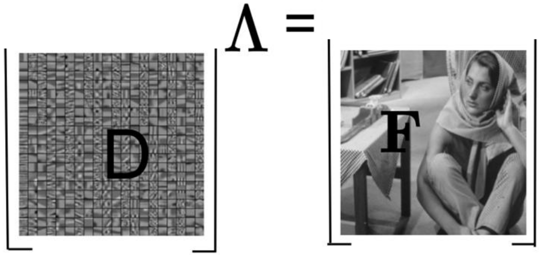
- **Contributions**:
  - First attempt in deblurring and depth estimation from space-variant blur using dictionary replacement
  - Works irrespective of blur type (motion/defocus)
  - Depth varying cases also handled

Dictionary Replacement for Single Image Deblurring

- **Contributions**:
  - First attempt in deblurring and depth estimation from space-variant blur using dictionary replacement
  - Works irrespective of blur type (motion/defocus)
  - Depth varying cases also handled

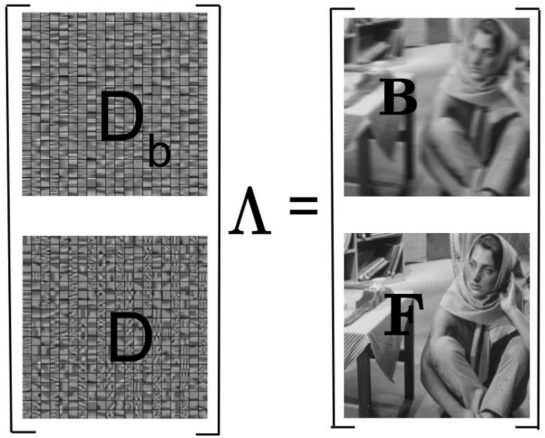
- **Assumptions**: Camera motion-- in plane translations
  Blur-depth relation holds

Blur-Invariant Representation



$$\mathbf{F} = \mathbf{D} \circ \Lambda$$

Blur-Invariant Representation

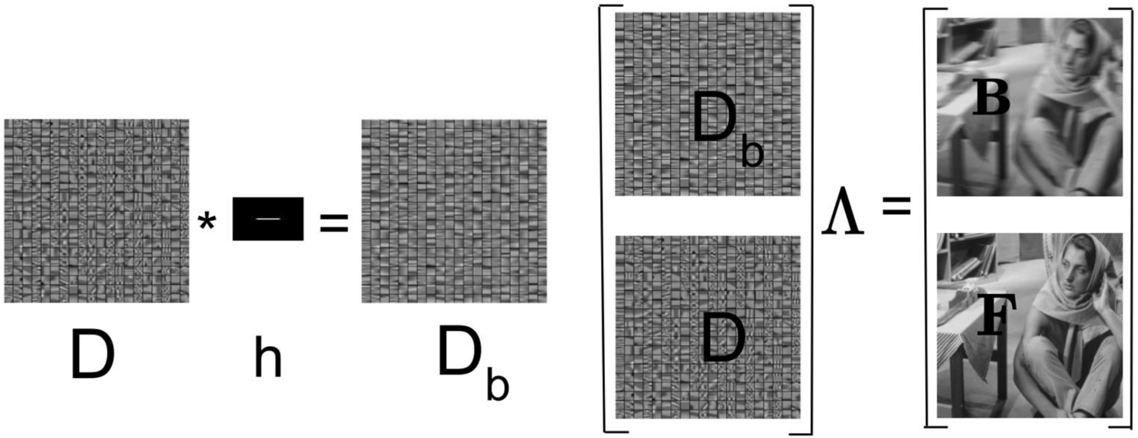


Space-invariant blur

$$\mathbf{B} = \mathbf{h} \otimes \mathbf{F}$$

$$\mathbf{F} = \mathbf{D} \circ \Lambda$$

Blur-Invariant Representation



Space-invariant blur

$$\mathbf{B} = \mathbf{h} \otimes \mathbf{F} = \mathbf{D}_b \circ \Lambda$$

$$\mathbf{F} = \mathbf{D} \circ \Lambda$$

Proposed Approach

Aim : Given a single blurred image $\mathbf{B}$ estimate latent image $\mathbf{F}$ and depth map

Blur at different depths are scaled versions of each other

Estimate blur $\mathbf{h}^{init}$ at a layer

Input blurred frame

Proposed Approach

Aim : Given a single blurred image $\mathbf{B}$ estimate latent image $\mathbf{F}$ and depth map

Blur at different depths are scaled versions of each other

Estimate blur $\mathbf{h}^{init}$ at a layer → Scale this kernel to fit at other locations

$$DC_s(\mathbf{x}) = ||\mathbf{B}(\mathbf{x}) - \bar{\mathbf{B}}^s(\mathbf{x})||_2^2 \quad \text{where} \quad \bar{\mathbf{B}}^s = \mathbf{D}_b^s \circ \Lambda^s$$

$$\mathbf{D}_b^s = \mathbf{h}^s * \mathbf{D}$$

Proposed Approach

Aim : Given a single blurred image $\mathbf{B}$ estimate latent image $\mathbf{F}$ and depth map

Blur at different depths are scaled versions of each other

Graph cut

Estimate blur $\mathbf{h}^{init}$ at a layer

Scale this kernel to fit at other locations

$$DC_s(\mathbf{x}) = ||\mathbf{B}(\mathbf{x}) - \bar{\mathbf{B}}^s(\mathbf{x})||_2^2 \quad \text{where} \quad \bar{\mathbf{B}}^s = \mathbf{D}_b^s \circ \Lambda^s$$

$$\mathbf{D}_b^s = \mathbf{h}^s * \mathbf{D}$$

Estimated depth map

Proposed Approach

Aim : Given a single blurred image $\mathbf{B}$ estimate latent image $\mathbf{F}$ and depth map

Blur at different depths are scaled versions of each other

Estimate blur $\mathbf{h}^{init}$ at a layer → Scale this kernel to fit at other locations → Stack deblurred result at each scale $\hat{\mathbf{F}}^s = \mathbf{D} \circ \Lambda^s$

Deblurred output at two scales

Proposed Approach

Aim : Given a single blurred image $\mathbf{B}$ estimate latent image $\mathbf{F}$ and depth map

Blur at different depths are scaled versions of each other

Estimate blur $\mathbf{h}^{init}$ at a layer → Scale this kernel to fit at other locations → Stack deblurred result at each scale $\hat{\mathbf{F}}^s = \mathbf{D} \circ \Lambda^s$

Input blurred frame

Deblurred output

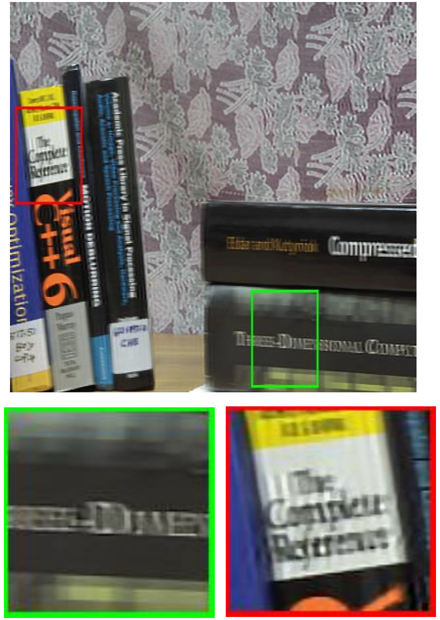
Results

Input

Results

Input

Xu et al. CVPR 2013



Xu, L., S. Zheng, and J. Jia, Unnatural l0 sparse representation for natural image deblurring. CVPR 2013

Results

Input | Estimated Depth | Our output | Xu et al. CVPR 2013

T.M Nimisha, M. Arun, and A.N. Rajagopalan, “Dictionary Replacement for Single Image Restoration of 3D Scenes,” in British Machine Vision Conference (BMVC), York, UK. September 2016.

(4) Blur-Invariant Feature Learning for Single Image Deblurring

Problem Statement: End-to-end learning for deblurring for general camera motion

(4) Blur-Invariant Feature Learning for Single Image Deblurring

Problem Statement: End-to-end learning for deblurring for general camera motion

- **Contributions:**
  - End-to-end network, skips need for kernel estimation and prior weight selection

(4) Blur-Invariant Feature Learning for Single Image Deblurring

Problem Statement: End-to-end learning for deblurring for general camera motion

- **Contributions:**
  - End-to-end network, skips need for kernel estimation and prior weight selection
  - New architecture with Autoencoders (AE) and Generative Adversarial Networks (GAN)

(4) Blur-Invariant Feature Learning for Single Image Deblurring

Problem Statement: End-to-end learning for deblurring for general camera motion

- **Contributions:**
    - End-to-end network, skips need for kernel estimation and prior weight selection
    - New architecture with Autoencoders (AE) and Generative Adversarial Networks (GAN)
    - Can handle both space-varying and invariant blur scenarios

(4) Blur-Invariant Feature Learning for Single Image Deblurring

Problem Statement: End-to-end learning for deblurring for general camera motion

- **Contributions:**
  - End-to-end network, skips need for kernel estimation and prior weight selection
  - New architecture with Autoencoders (AE) and Generative Adversarial Networks (GAN)
  - Can handle both space-varying and invariant blur scenarios
  - Works in dynamic blur cases as well

(4) Blur-Invariant Feature Learning for Single Image Deblurring

Problem Statement: End-to-end learning for deblurring for general camera motion

- **Contributions:**
  - End-to-end network, skips need for kernel estimation and prior weight selection
  - New architecture with Autoencoders (AE) and Generative Adversarial Networks (GAN)
  - Can handle both space-varying and invariant blur scenarios
  - Works in dynamic blur cases as well

- **Approach:** Learn clean feature domain and map blurred images to clean features
  Two stage network

(4) Blur-Invariant Feature Learning for Single Image Deblurring

Problem Statement: End-to-end learning for deblurring for general camera motion

- **Contributions:**
  - End-to-end network, skips need for kernel estimation and prior weight selection
  - New architecture with Autoencoders (AE) and Generative Adversarial Networks (GAN)
  - Can handle both space-varying and invariant blur scenarios
  - Works in dynamic blur cases as well

- **Approach:** Learn clean feature domain and map blurred images to clean features

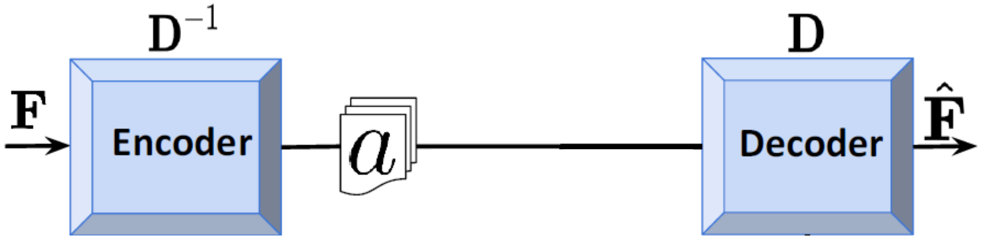
  Two stage network

  Stage I: Learns clean image feature representation using AE

  Stage II: Map blurred images to clean representations using GAN

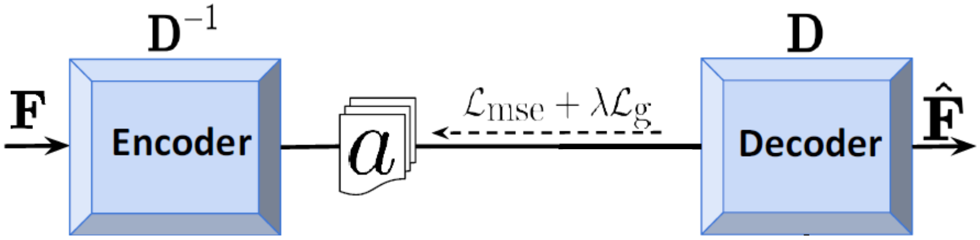
Proposed Network and Training Losses

Stage I: Learns clean image feature representation using AE



Proposed Network and Training Losses

Stage I: Learns clean image feature representation using AE



Stage I

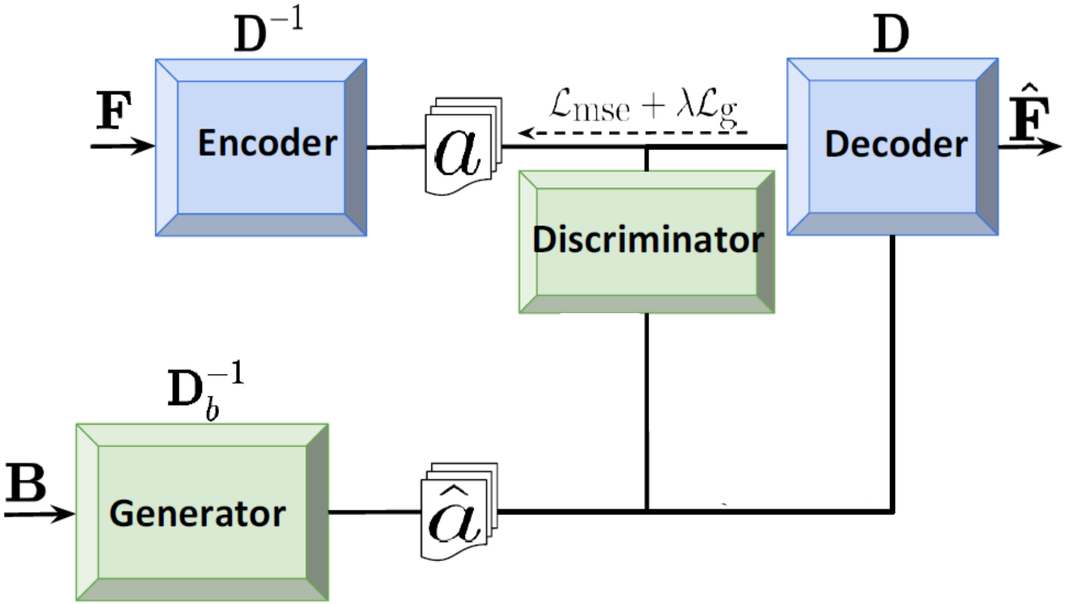
$$\mathcal{L}_{\mathrm{mse}} + \lambda \mathcal{L}_{\mathrm{g}}$$

$$\mathcal{L}_{\mathrm{mse}} = \|\mathcal{D}e(\mathcal{E}n(\mathbf{F} + \mathbf{n})) - \mathbf{F}\|_2^2$$

$$\mathcal{L}_{\mathrm{g}} = \|\nabla \mathcal{D}e(\mathcal{E}n(\mathbf{F} + \mathbf{n})) - \nabla \mathbf{F}\|_2^2$$

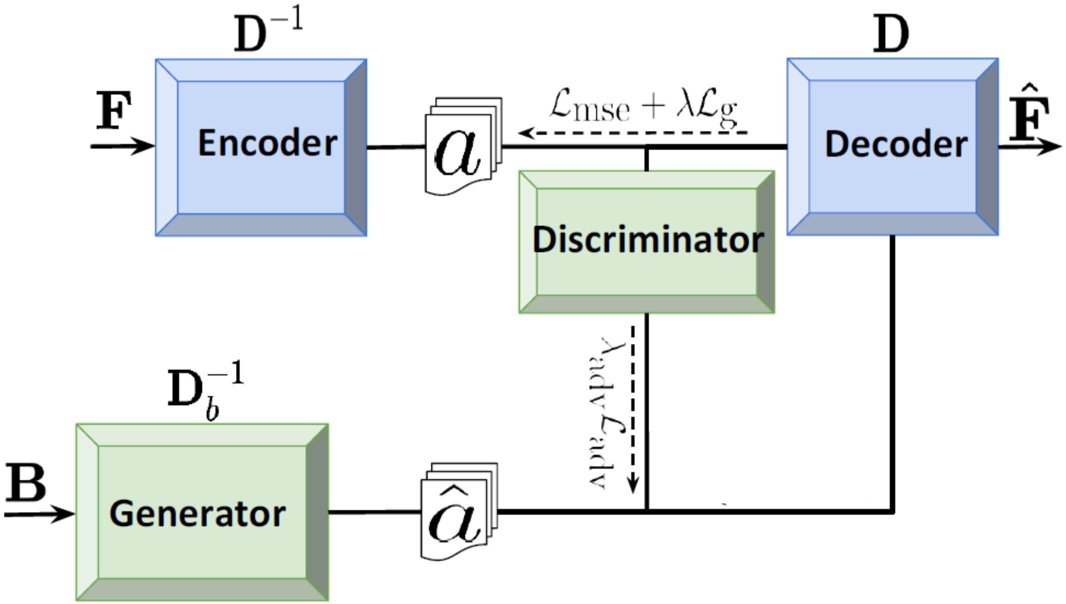
Proposed Network and Training Losses

Stage II: Map blurred images to clean representations using GAN



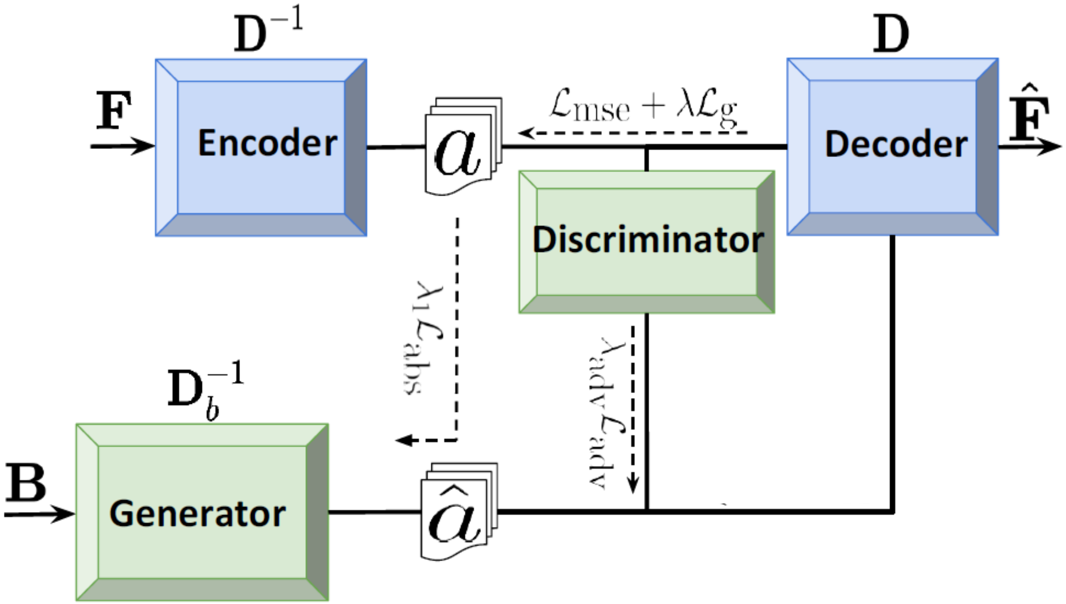
Proposed Network and Training Losses

Stage II: Map blurred images to clean representations using GAN



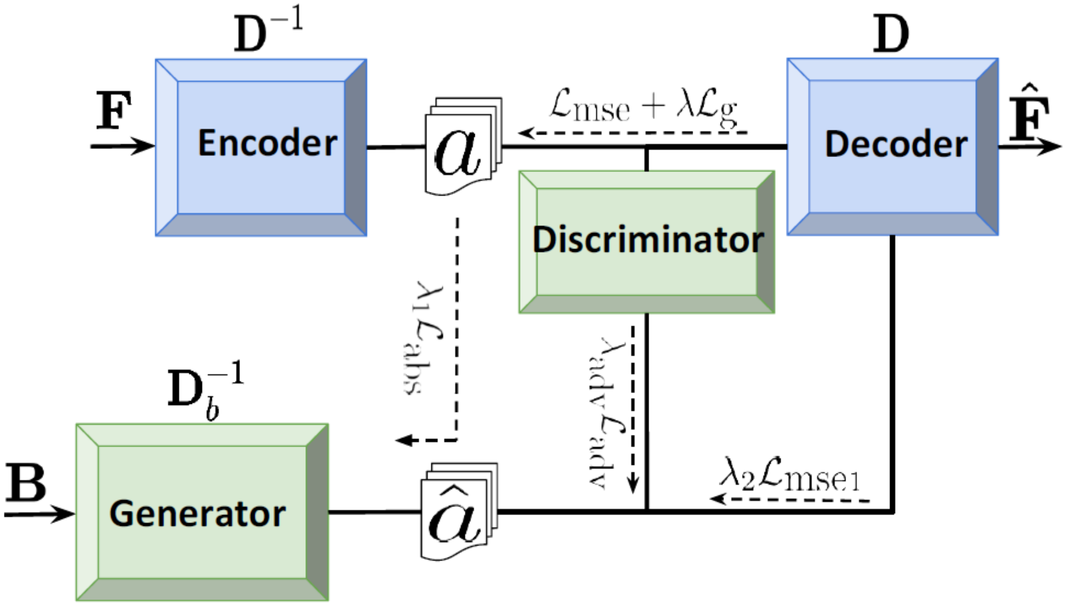
Proposed Network and Training Losses

Stage II: Map blurred images to clean representations using GAN



Proposed Network and Training Losses

Stage II: Map blurred images to clean representations using GAN



Stage II

$$\lambda_{\text{adv}}\mathcal{L}_{\text{adv}} + \lambda_1\mathcal{L}_{\text{abs}} + \lambda_2\mathcal{L}_{\text{mse1}}$$

$$\boxed{\begin{aligned}\mathcal{L}_{\text{abs}} &= \|\mathcal{G}(\mathbf{B}) - \mathcal{E}n(\mathbf{F})\|_1 \\ \mathcal{L}_{\text{mse1}} &= ||\mathcal{D}e(\mathcal{G}(\mathbf{B})) - \mathbf{F}||_2^2\end{aligned}}$$

Quantitative Results

Dataset Sun et al. 2013	Xu et al. CVPR 2013	Pan et al. CVPR 2016	Whyte et al. IJCV 2012	Ours
PSNR				
MSSIM				
Run time Image size: 1024 X 700				

Quantitative Results

Dataset Sun et al. 2013	Xu et al. CVPR 2013	Pan et al. CVPR 2016	Whyte et al. IJCV 2012	Ours
PSNR	28.11			
MSSIM	0.9177			
Run time Image size: 1024 X 700	(Matlab, CPU) 34 sec			

Quantitative Results

Dataset Sun et al. 2013	Xu et al. CVPR 2013	Pan et al. CVPR 2016	Whyte et al. IJCV 2012	Ours
PSNR	28.11	31.16		
MSSIM	0.9177	0.9623		
Run time Image size: 1024 X 700	(Matlab, CPU) 34 sec	(Matlab, CPU) 40 min		

Quantitative Results

Dataset Sun et al. 2013	Xu et al. CVPR 2013	Pan et al. CVPR 2016	Whyte et al. IJCV 2012	Ours
PSNR	28.11	31.16	26.335	
MSSIM	0.9177	0.9623	0.8528	
Run time Image size: 1024 X 700	(Matlab, CPU) 34 sec	(Matlab, CPU) 40 min	(Matlab, CPU) 4 min	

Quantitative Results

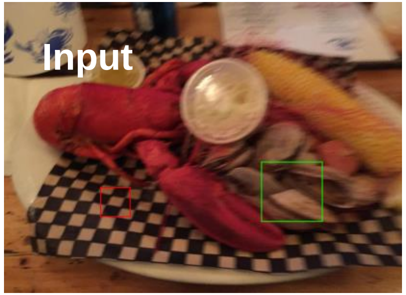
Dataset Sun et al. 2013	Xu et al. CVPR 2013	Pan et al. CVPR 2016	Whyte et al. IJCV 2012	Ours
PSNR	28.11	31.16	26.335	30.54
MSSIM	0.9177	0.9623	0.8528	0.9553
Run time Image size: 1024 X 700	(Matlab, CPU) 34 sec	(Matlab, CPU) 40 min	(Matlab, CPU) 4 min	(Torch, GPU/CPU) 3.4 sec/2 min

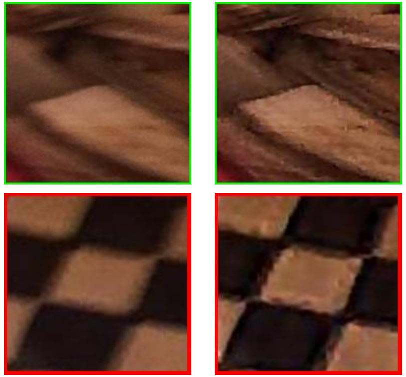
Quantitative Results

Dataset Sun et al. 2013	Xu et al. CVPR 2013	Pan et al. CVPR 2016	Whyte et al. IJCV 2012	Ours
PSNR	28.11	31.16	26.335	30.54
MSSIM	0.9177	0.9623	0.8528	0.9553
Run time Image size: 1024 X 700	(Matlab, CPU) 34 sec	(Matlab, CPU) 40 min	(Matlab, CPU) 4 min	(Torch, GPU/CPU) 3.4 sec/2 min

Best but run time too high

Qualitative Results

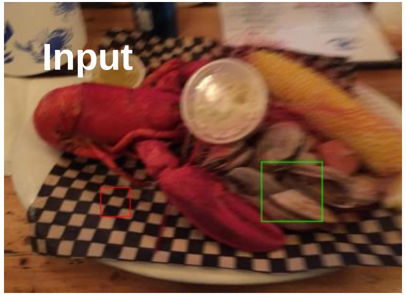


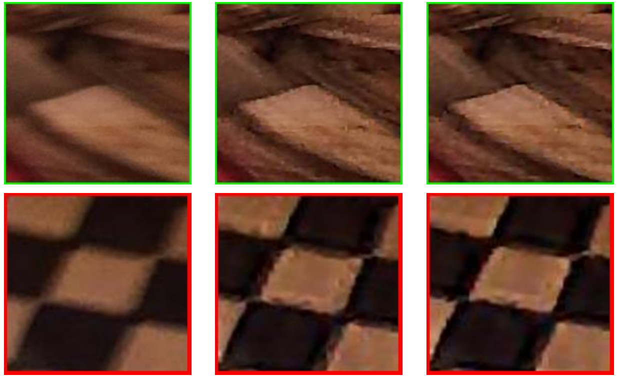


Input

Xu et al.
CVPR 2013

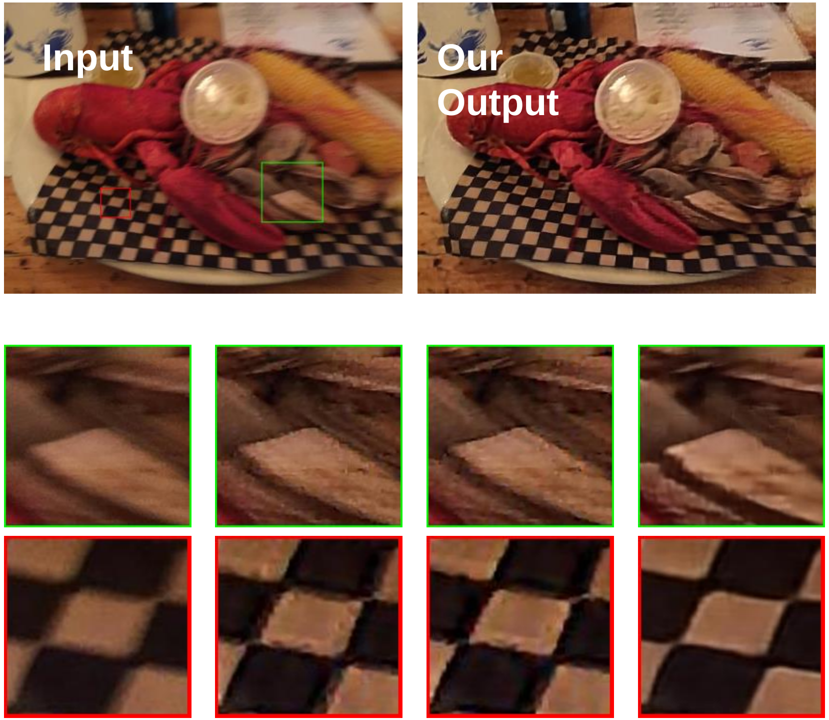
Qualitative Results





Input | Xu et al. CVPR 2013 | Pan et al. CVPR 2016

Qualitative Results



Input

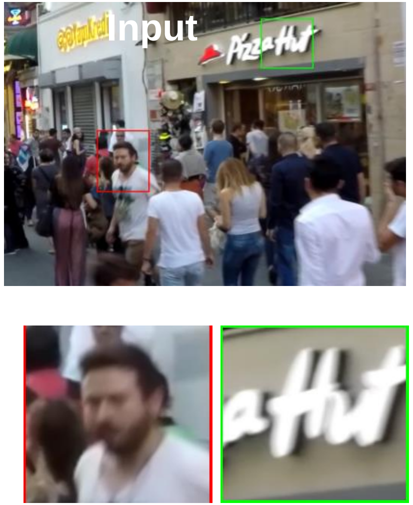
Xu et al.
CVPR 2013

Pan et al.
CVPR
2016

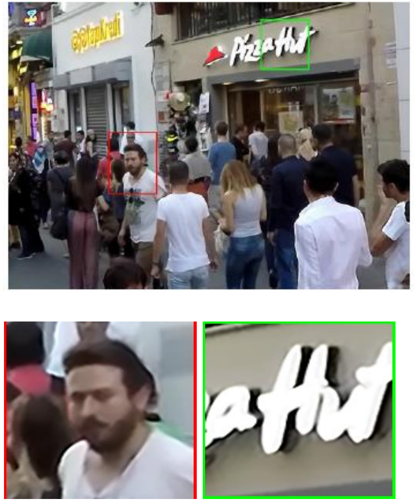
Ours

Qualitative Results

Dynamic Scene



Input



Kim and Mu CVPR 2015

Ours

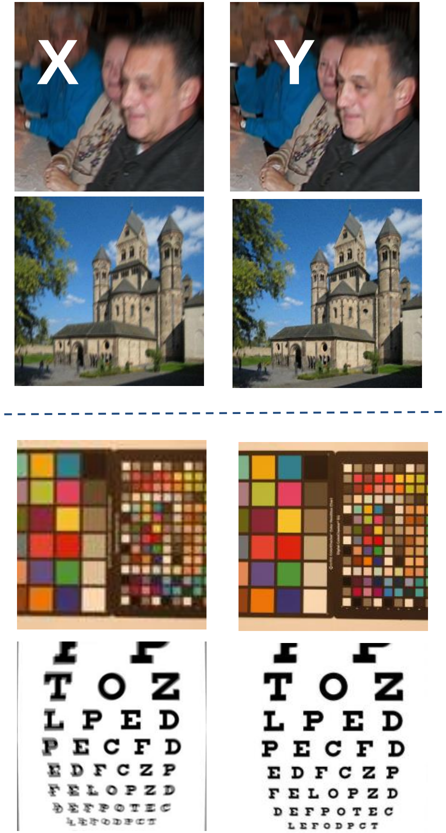
T.M Nimisha, Akash Kumar Singh, and A.N.Rajagopalan, "Blur-Invariant Deep Learning for Blind Deblurring," IEEE International Conference on Computer Vision (ICCV), Venice, Italy, October 2017.

(5) Unsupervised Class-Specific Single Image Deblurring

Problem Statement: End-to-end deblurring from single image without paired dataset for learning

Need for Unsupervised Methods

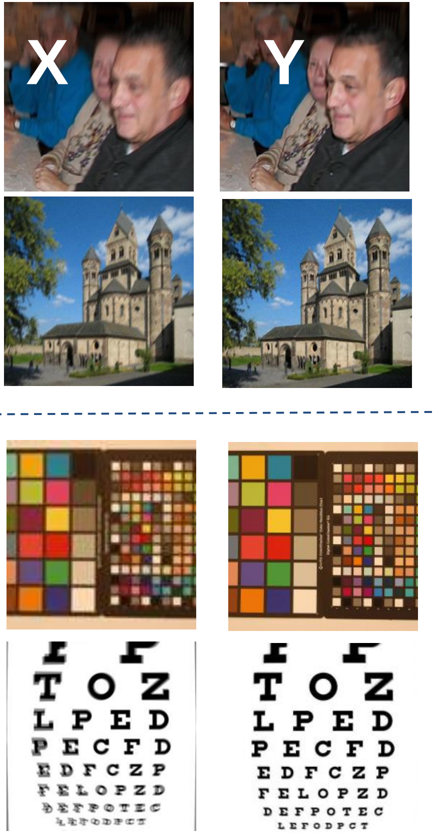
Paired dataset



Drawbacks:
- Capturing is difficult and expensive
- Problems with aligning data

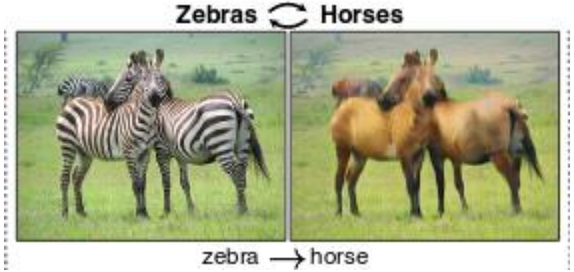
Need for Unsupervised Methods

Paired dataset



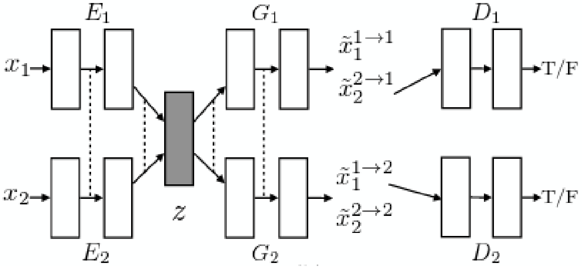
Unpaired dataset

Related Works



Zhu et al. [ICCV 2017]

- Uses cyclic consistency loss

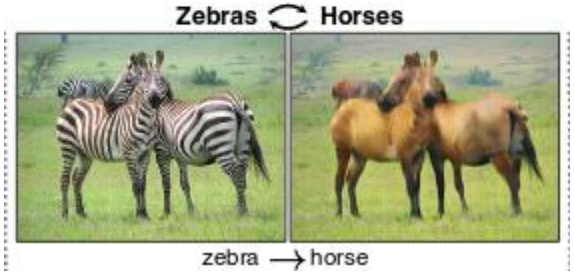


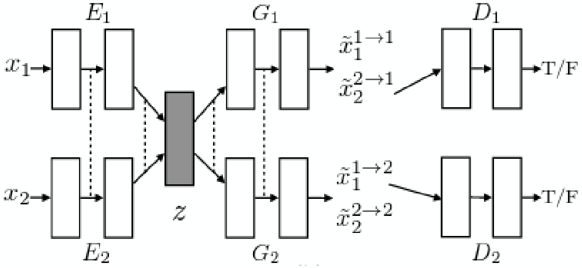
Liu et al. [NIPS 2017]

- Shared latent codes using weight sharing

Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. ICCV 2017
Liu, M.Y., Breuel, T., Kautz, J.: Unsupervised image-to-image translation networks NIPS 2017

Related Works





Drawbacks:

- Deterministic one-to-one mapping
- Sensitive to initialization and requires repeated attempts to converge to a satisfactory mapping

Unsupervised Class-Specific Single Image Deblurring

- **Contributions:**
  - First ever data-driven approach for deblurring from unpaired data

Unsupervised Class-Specific Single Image Deblurring

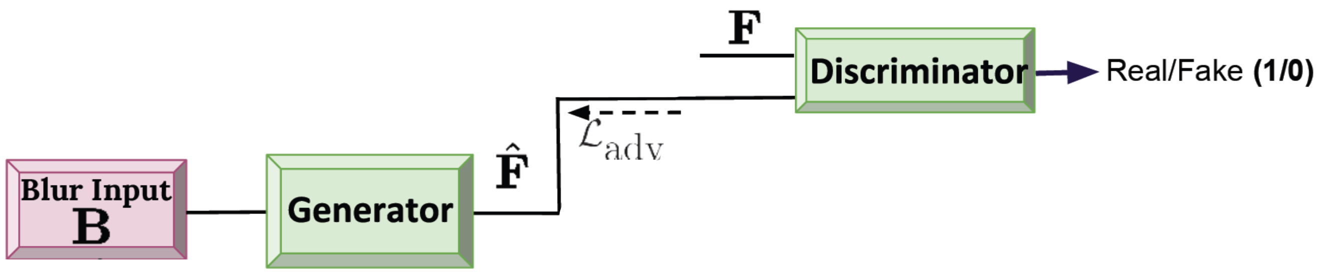
- **Contributions:**
  - First ever data-driven approach for deblurring from unpaired data
  - Proposed self guidance modules (reblurring and gradient) for convergence

Unsupervised Class-Specific Single Image Deblurring

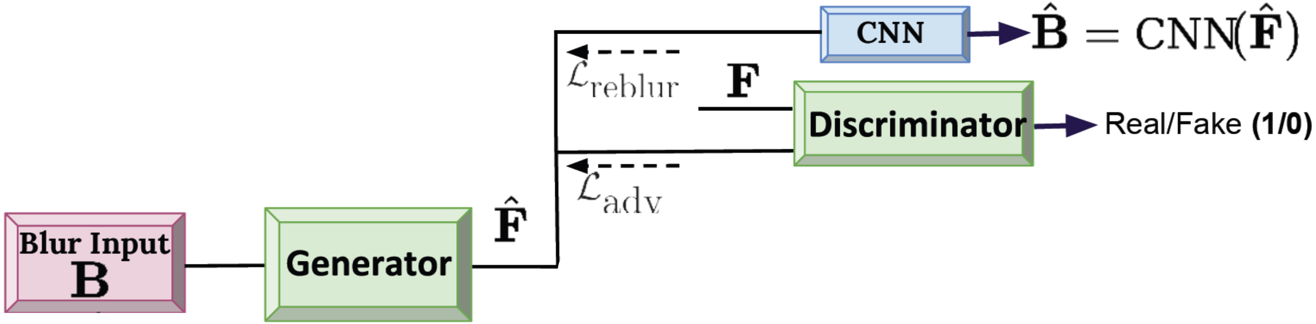
- **Contributions:**
  - First ever data-driven approach for deblurring from unpaired data
  - Proposed self guidance modules (reblurring and gradient) for convergence

- **Approach:** Learn the mapping from blur to clean using Generative networks
  - Class-specific approach
  - Add additional costs to constrain the solution space

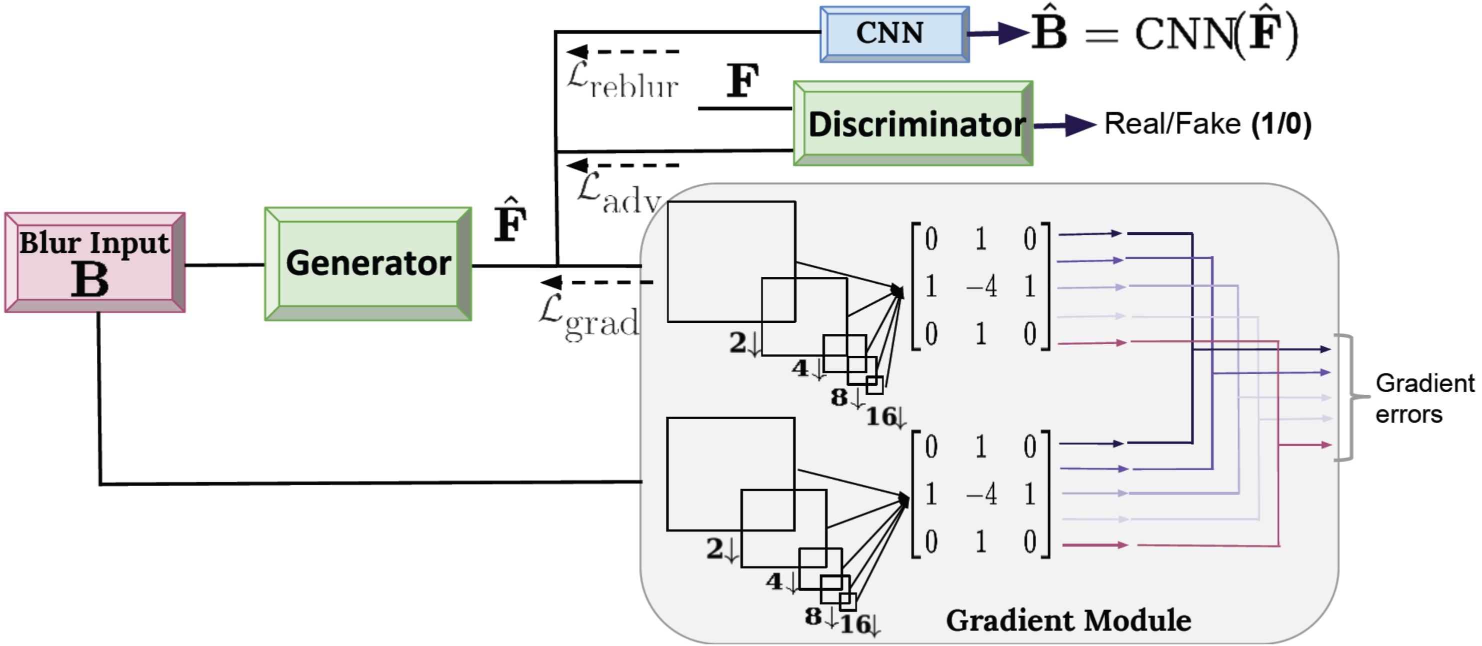
Proposed Network and Training Losses



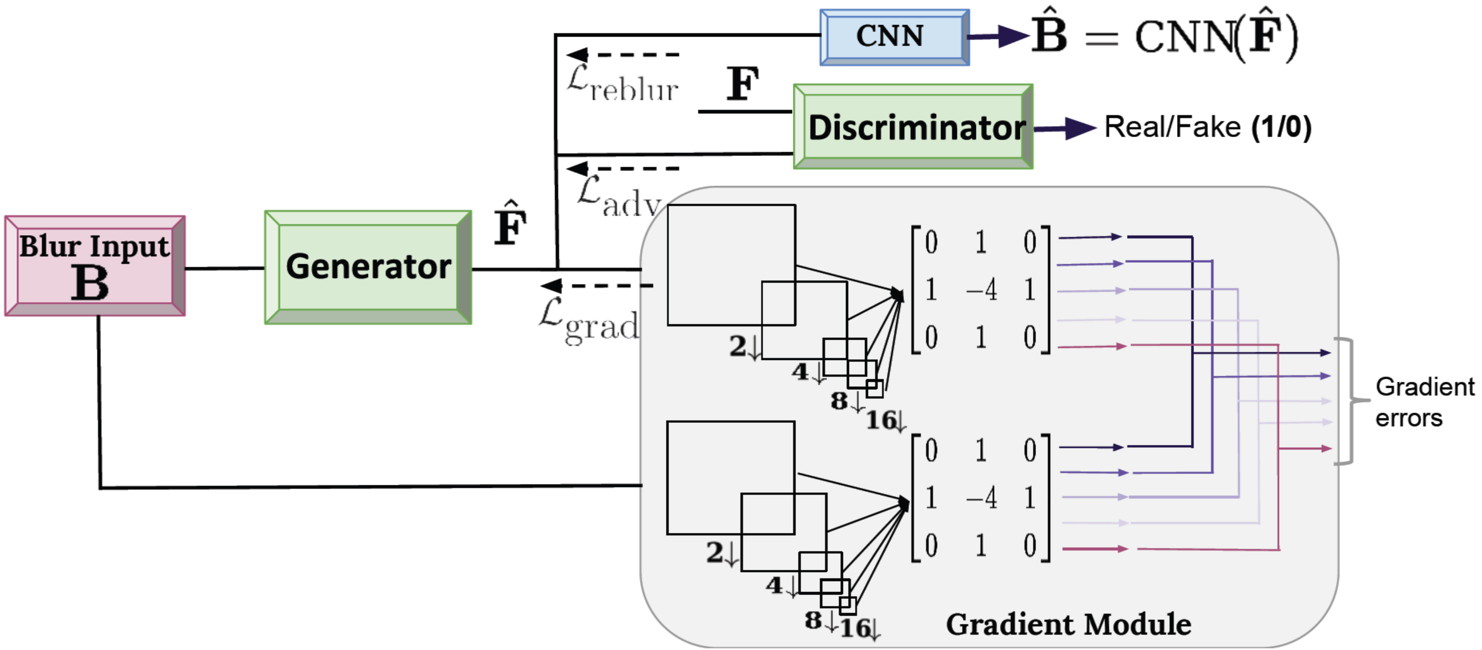
Proposed Network and Training Losses



Proposed Network and Training Losses



Proposed Network and Training Losses



Loss Functions:

$$\mathcal{L}_{\mathcal{G}} = \lambda_{\text{adv}}\mathcal{L}_{\text{adv}} + \lambda_{\text{reblur}}\mathcal{L}_{\text{reblur}} + \lambda_{\text{grad}}\mathcal{L}_{\text{grad}}$$

$$\mathcal{L}_{\text{adv}} = \min_{\theta} \frac{1}{N} \sum_{i} \log(1 - \mathcal{D}(\mathcal{G}_{\theta}(\mathbf{B}_i)))$$

$$\mathcal{L}_{\text{reblur}} = ||\mathbf{B} - \mathrm{CNN}(\hat{\mathbf{F}})||_2^2$$

$$\mathcal{L}_{\text{grad}} = \sum_{s \in \{1,2,4,8,16\}} \lambda_s |\nabla \mathbf{B}_{s\downarrow} - \nabla \hat{\mathbf{F}}_{s\downarrow}|$$

Effect of Loss Functions

Input

Target

Effect of Loss Functions

Input

GAN only

Target

Effect of Loss Functions

Input

GAN only

GAN+CNN

Target

Effect of Loss Functions

Input GAN only GAN+CNN GAN+CNN+GRAD Target

Results

Quantitative Evaluation: Conventional method

Face Dataset

Method	PSNR	SSIM	KSM
Pan et al. CVPR 2016	19.38	0.7764	0.7436
Ours			

KSM: Kernel Similarity Measure

Input

Pan et al. CVPR 2016

Results

Quantitative Evaluation: Deep learning method

Face Dataset

Method	PSNR	SSIM	KSM
Pan et al. CVPR 2016	19.38	0.7764	0.7436
Nah et al. CVPR 2017	24.12	0.8755	0.6229
Ours			

KSM: Kernel Similarity Measure

Input

Nah et al. CVPR 2017

Results

Quantitative Evaluation: Unsupervised method

Face Dataset

Method	PSNR	SSIM	KSM
Pan et al. CVPR 2016	19.38	0.7764	0.7436
Nah et al. CVPR 2017	24.12	0.8755	0.6229
Zhu et al. 2017	8.93	0.4406	0.2932
Ours			

KSM: Kernel Similarity Measure

Input

Zhu et al. 2017

Results

Quantitative Evaluation:

Face Dataset

Method	PSNR	SSIM	KSM
Pan et al. CVPR 2016	19.38	0.7764	0.7436
Nah et al. CVPR 2017	24.12	0.8755	0.6229
Zhu et al. 2017	8.93	0.4406	0.2932
Ours	22.80	0.8631	0.7536

KSM: Kernel Similarity Measure

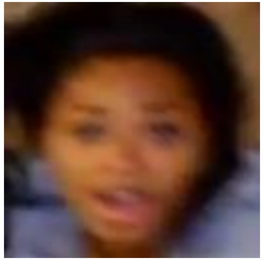
Input

Ours

T. M. Nimisha, Sunil Kumar, and A N Rajagopalan, "Unsupervised Class-Specific Deblurring," European Conference on Computer Vision (ECCV), Munich, Germany, September 2018.

Comparison with Face Deblurring Work

Input



Comparison with Face Deblurring Work

Chrysos, G., Zafeiriou, S.: Deep face deblurring. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (2017)

Comparison with Face Deblurring Work

Input | Chrysos 2017 | Ours

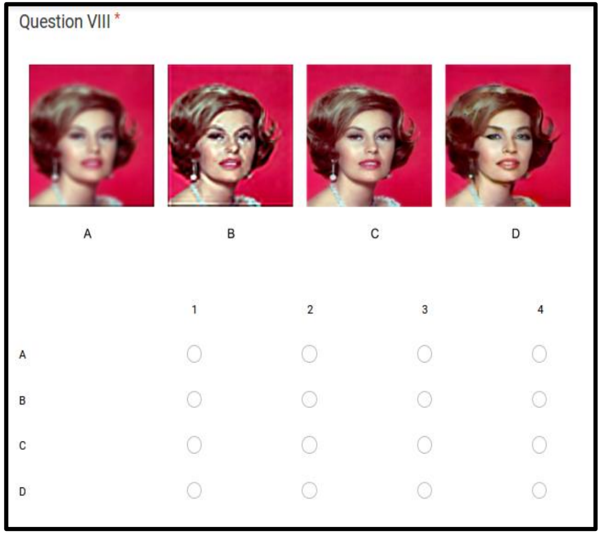
Human Perception Ranking

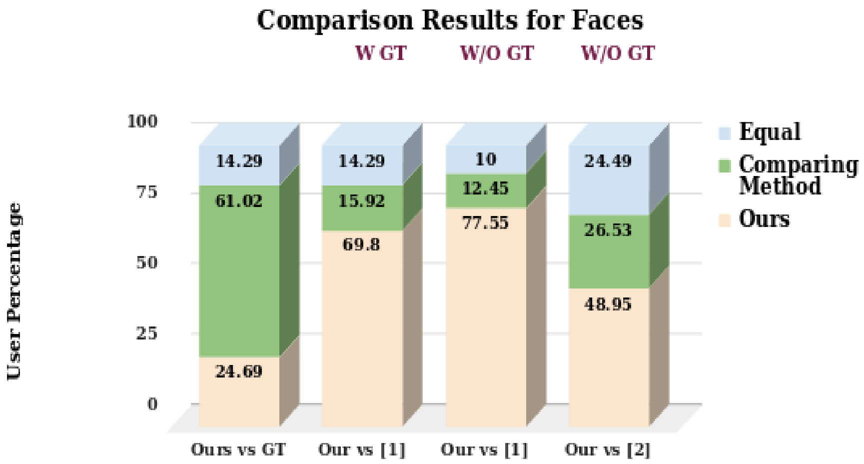
Question VIII *

A B C D

	1	2	3	4
A	○	○	○	○
B	○	○	○	○
C	○	○	○	○
D	○	○	○	○

Human Perception Ranking





Publications Related to Thesis

Journal papers:

- T. M. Nimisha, A. N. Rajagopalan, and R. Aravind, “Generating High Quality Pan-Shots from Motion Blurred Videos,” Elsevier Journal: Computer Vision and Image Understanding (CVIU), Vol. 171, pp.20-33, June 2018.
- Abhijith Punnappurath, T. M. Nimisha, and A.N. Rajagopalan “Multi-image blind super-resolution of 3D scenes,” IEEE Transactions on Image Processing, Vol. 26, No. 11, pp. 5337-5352, November 2017.

Conference papers:

- T. M. Nimisha, Sunil Kumar, and A N Rajagopalan, "Unsupervised Class-Specific Deblurring," European Conference on Computer Vision (ECCV), Munich, Germany, September 2018.
- T.M Nimisha, Akash Kumar Singh, and A.N.Rajagopalan, "Blur-Invariant Deep Learning for Blind Deblurring," IEEE International Conference on Computer Vision (ICCV), Venice, Italy, October 2017.
- T.M Nimisha, M. Arun, and A.N. Rajagopalan, “Dictionary Replacement for Single Image Restoration of 3D Scenes,” in British Machine Vision Conference (BMVC), York, UK. September 2016.