Deep Decoupling of Defocus and Motion Blur for Dynamic Segmentation

Abhijith Punnappurath, Yogesh Balaji, Mahesh Mohan, A. N. Rajagopalan

Indian Institute of Technology Madras

www.ee.iitm.ac.in/~ee10d038/d3m

ECCV'16

EUROPEAN CONFERENCE ON COMPUTER VISION

Goal

Segment dynamic objects given a single space-variantly blurred image of a 3D scene captured using a hand-held camera

Challenges

Single image

Camera/ object motion ⇒ motion blur

3D scene ⇒ defocus blur

General camera motion/ 3D scene ⇒ space-varying blur

Depth-motion ambiguity

Static camera
Dynamic object
Only object pixels blurred

Moving camera
Dynamic object
Only background pixels blurred

Moving camera
Dynamic/ stationary objects
All pixels blurred

Our approach

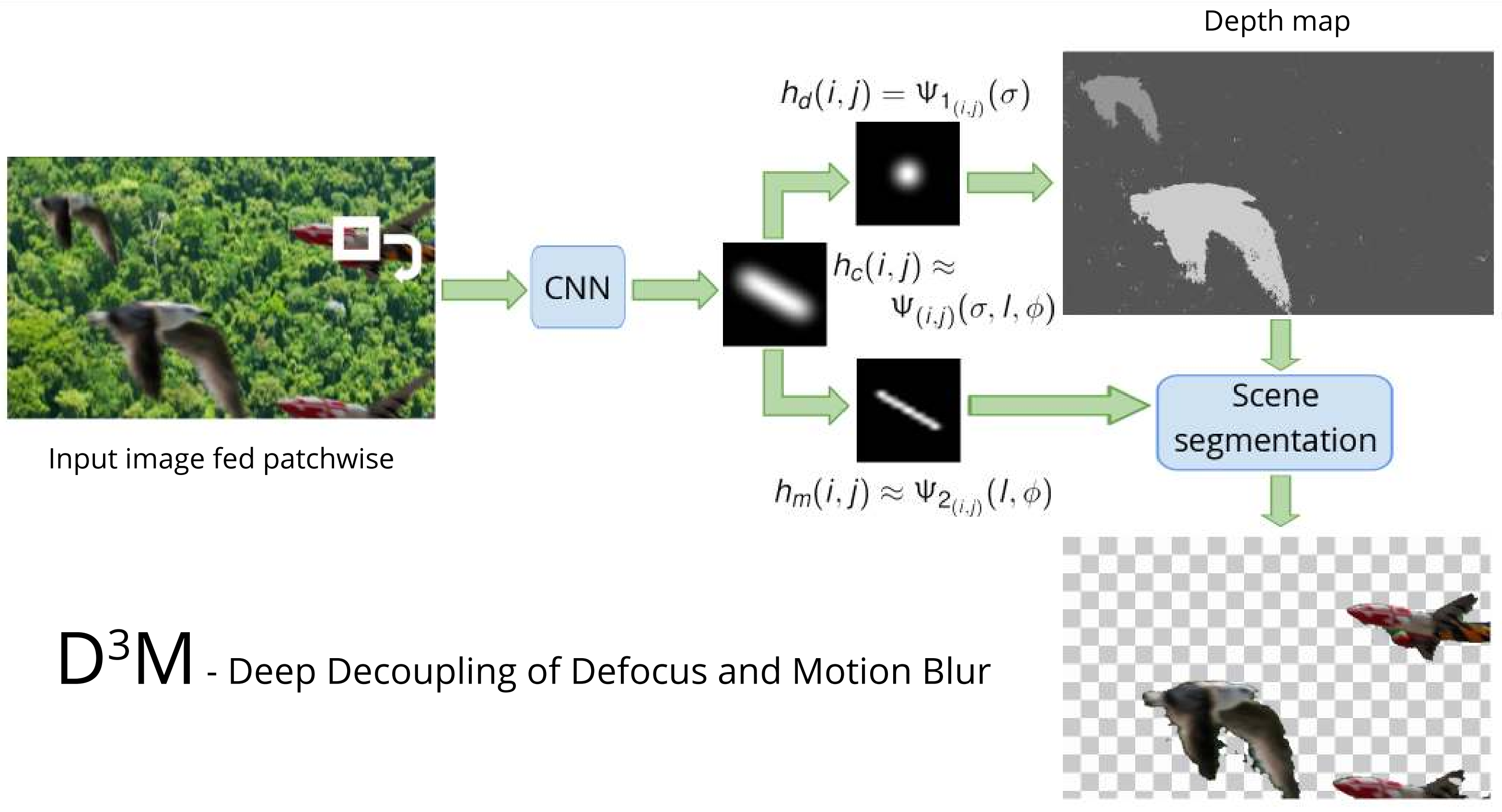
Train a CNN to predict the composite kernel h_c at each pixel

Composite kernel is convolution of defocus h_d and motion h_m kernels

Use defocus cue to recover the depth map

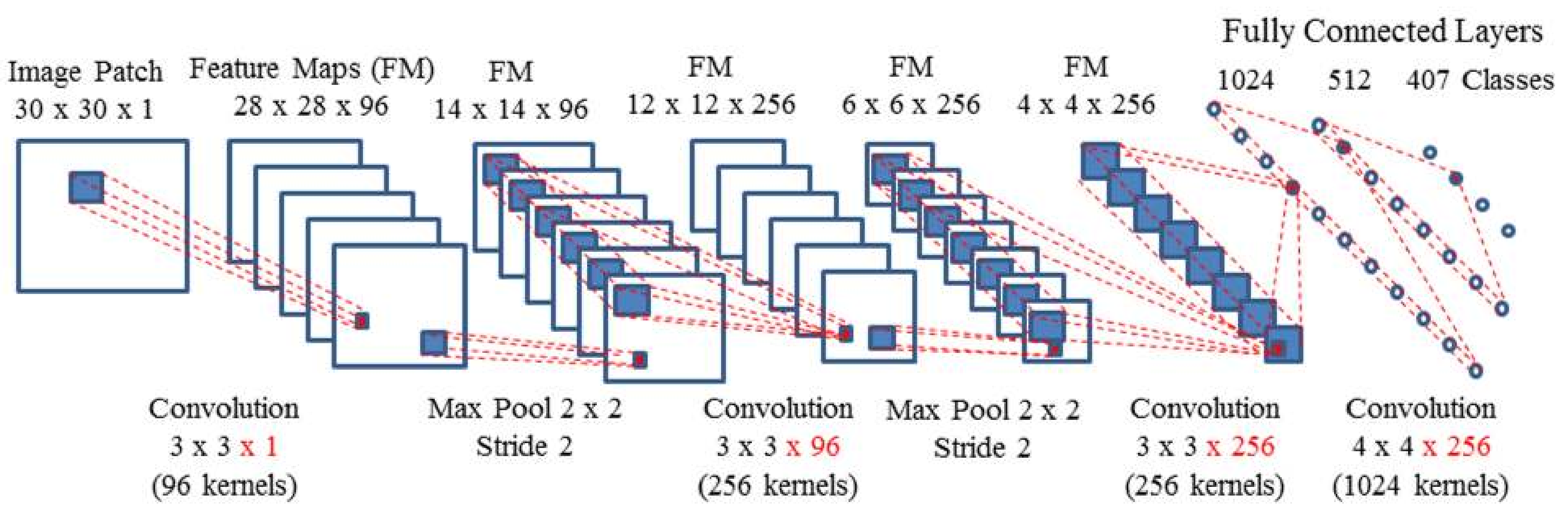
Use motion kernels to segregate the dynamic objects at each depth layer

Joint model for defocus and motion helps resolve depth-motion ambiguity


Depth map
$h_d(i,j) = \Psi_{1_{(i,j)}}(\sigma)$
CNN
$h_c(i,j) \approx \Psi_{(i,j)}(\sigma, l, \phi)$
$h_m(i,j) \approx \Psi_{2_{(i,j)}}(l, \phi)$
Scene segmentation
Input image fed patchwise
Segmented dynamic objects

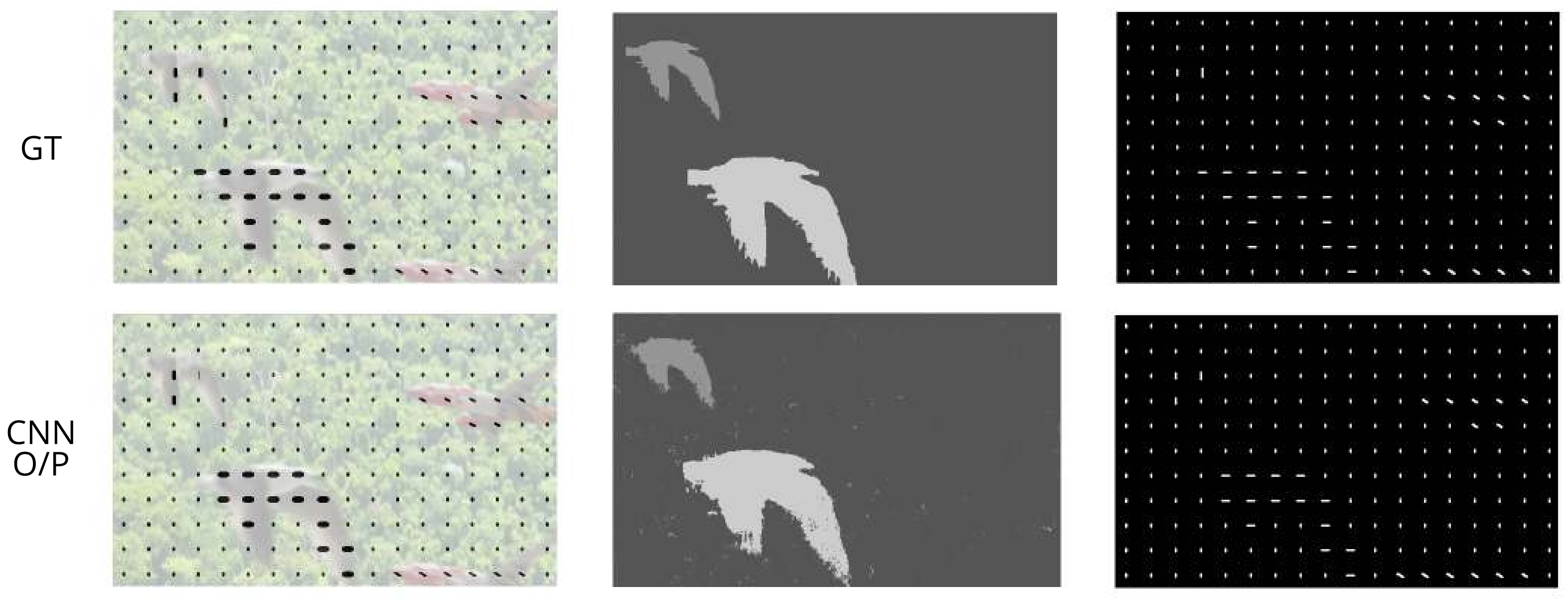
D^{3}M - Deep Decoupling of Defocus and Motion Blur

Kernel classification using CNN



Our CNN architecture for predicting the composite kernels

A synthetic example



Composite kernels

Depth map

Motion kernels

Scene segmentation

Layer with maximum area in depth map = Reference depth layer d_0

Segmenting moving objects in the reference depth layer d_0

Blur on dynamic object pixel ≠ Blur on pixel affected only by camera motion

Non-uniform camera motion blur model for a static fronto-parallel planar scene

$$g = \sum_{k=1}^{|\mathbf{T}|} \omega_0(k) f_k$$

g : blurred image, f : latent image, $\mathbf{T}$: discrete camera pose space,

ω_0 : camera motion parameter, f_k : f warped by the homography $\mathbf{H_k}(t_{x_k}, t_{y_k}, \theta_{z_k})$

$$h(i,j;m,n) = \sum_{k=1}^{|\mathbf{T}|} \omega_0(k) \times \delta(m-(i_k-i), n-(j_k-j)) \qquad (1)$$

h : space-varying motion kernel, (i,j) : image pixel coordinates,

(i_k, j_k) : transformed coordinates when $\mathbf{H_k}^{-1}$ is applied on (i,j)

Blur compatibility test [2]

- Select two pixels with motion kernels h_{m_1} and h_{m_2}
- Find $\tau_\upsilon = \{k : h_{m_\upsilon}(i,j; i_k - i, j_k - j) > 0\}$,where $\upsilon = 1,2$
- Calculate $\boldsymbol{\tau_{12}} = \boldsymbol{\tau_1} \cap \boldsymbol{\tau_2}$
- Regenerate $\hat{h}_{m_1}$ and $\hat{h}_{m_2}$ using $\boldsymbol{\tau_{12}}$
- The two pixels are NOT "blur compatible" if h_{m_1} and h_{m_2} have positive entries at locations other than those in $\hat{h}_{m_1}$ and $\hat{h}_{m_2}$

Apply test on all pixels at d_0

Segmenting moving objects at other depths d_p

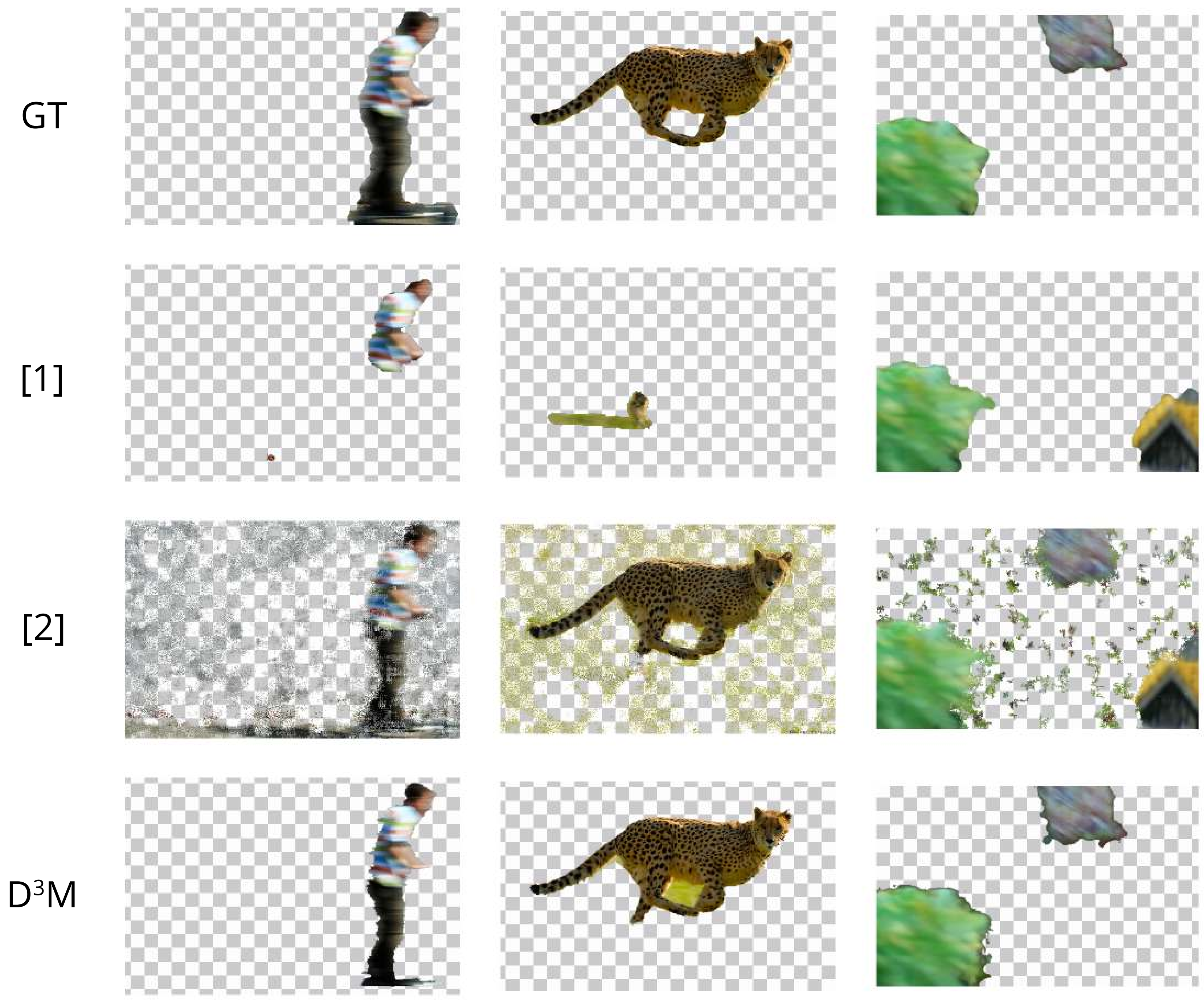
Depth map and motion experienced by reference layer are known ⇒

Kernel at a pixel lying on any other depth layer can be determined

- Compute relative depth $s_p = \frac{d_p}{d_0}$ from $\frac{\sigma_0}{\sigma_p} = \frac{\left(\frac{1}{u} - \frac{1}{d_0}\right)}{\left(\frac{1}{u} - \frac{1}{d_p}\right)}$, where $u =$ aperture radius
- Estimate ω_0 using the method in [2]
- Calculate $\mathbf{H}_{k_p}(t_{x_{k_p}}, t_{y_{k_p}}, \theta_{z_k})$, where $t_{x_{k_p}} = \frac{t_{x_k}}{s_p}$, $t_{y_{k_p}} = \frac{t_{y_k}}{s_p}$
- The motion kernel $\hat{h}_{m_p}$ at any other depth d_p can be estimated from equation (1) with (i_k, j_k) replaced by (i_{k_p}, j_{k_p}), where (i_{k_p}, j_{k_p}) is obtained by applying $\mathbf{H}_{k_p}^{-1}$ on (i,j)
- Let h_{m_p} = motion kernel predicted by our CNN.
Cross-correlation$(\hat{h}_{m_p}, h_{m_p}) <$ threshold ⇒ dynamic pixel

At all other depth layers d_p

Results



References

[1] A. Chakrabarti, T. Zickler, W. T. Freeman, "Analyzing spatially-varying blur," In Proc. CVPR 2010.

[2] C. Paramanand, A. N. Rajagopalan, "Motion blur for motion segmentation," In Proc. ICIP 2013.