香港中文大學
The Chinese University of Hong Kong

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

CVPR LONG BEACH CALIFORNIA June 16-20, 2019

Self-Supervised Learning via Conditional Motion Propagation

Xiaohang Zhan[1], Xingang Pan[1], Ziwei Liu[1], Dahua Lin[1], Chen Change Loy[2]

[1]The Chinese University of Hong Kong [2]Nanyang Technological University

CVPR 2019

Learn from Motion

- Motion: easy to obtain, without manual annotations.
- Goal: learn image representations and object properties from motion.

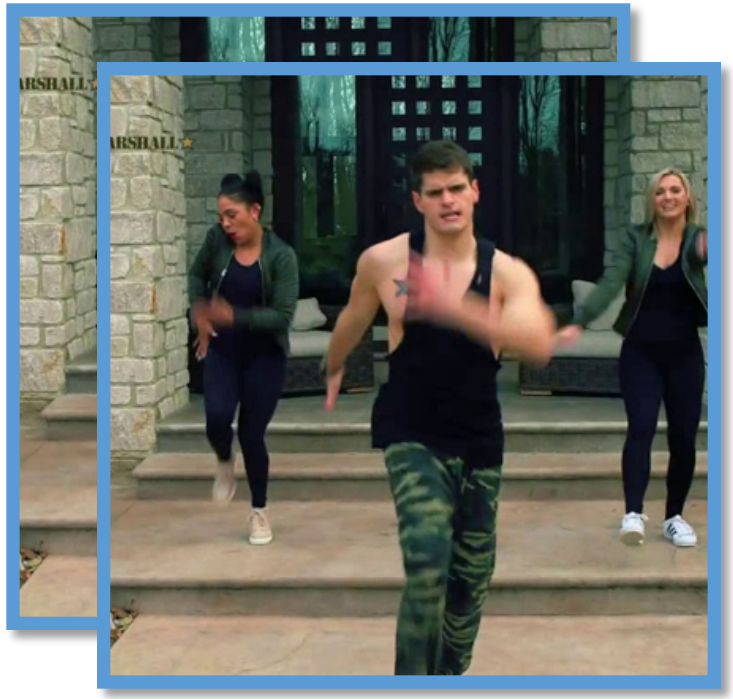


Image pair

Optical flow estimation

Motion (optical flow)

Previous Works

- **Motion prediction** ☹ **Motion is ambiguous.**

(a) Input Image (b) Prediction (c) Ground Truth

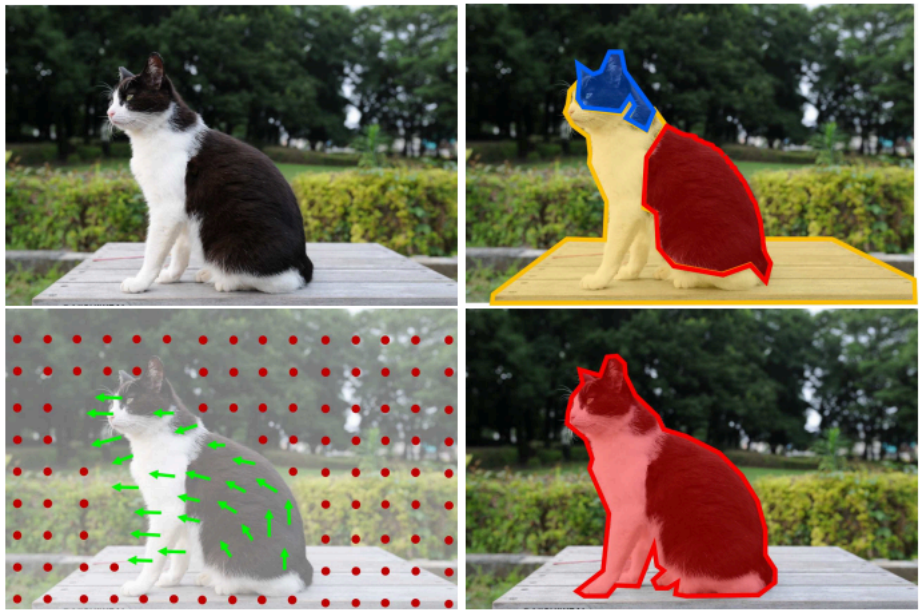
Motion prediction from static images. [1]

[1] Walker J, Gupta A, Hebert M. "Dense optical flow prediction from a static image." *In CVPR*, 2015.

Previous Works

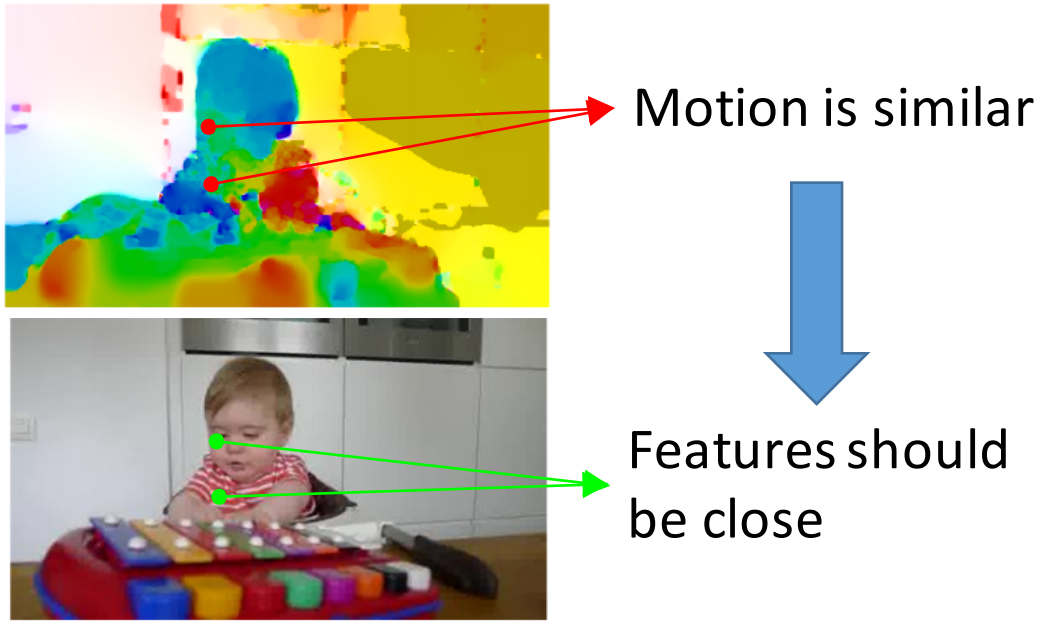
- **Motion as regularization**

  Assume pixels on the same object should have similar motion.



Learn to segment moving foreground [2]

☹ **Motion is complicated.**

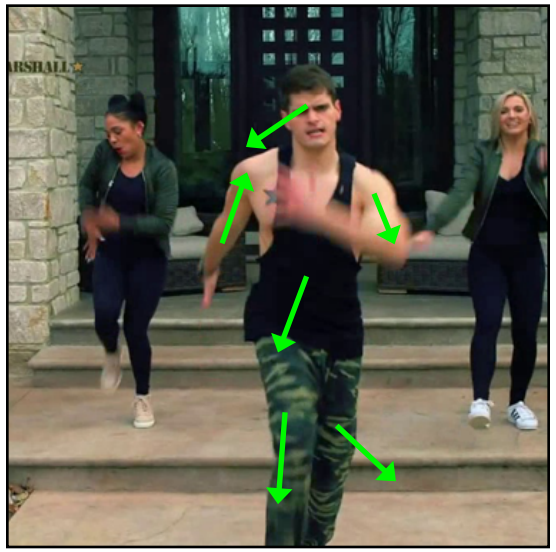


Learn motion consistency [3]

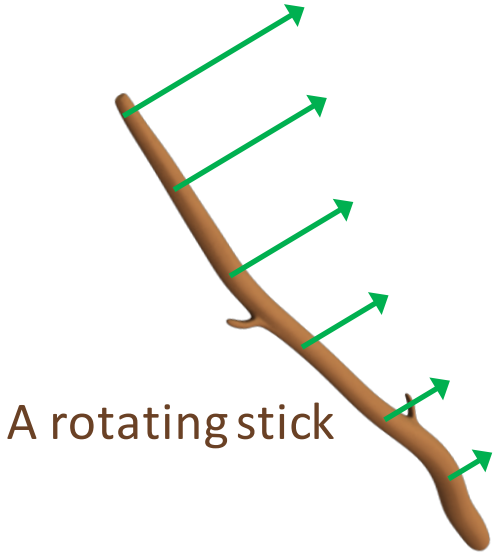
[2] Pathak D, Girshick R, Dollár P, Darrell T, Hariharan B. Learning features by watching objects move. In CVPR, 2017.
[3] Mahendran A, Thewlis J, Vedaldi A. Cross pixel optical-flow similarity for self-supervised learning. In ACCV, 2018.

Motion is complicated

Do pixels on the same object have similar motion?



Objects with high degrees of freedom

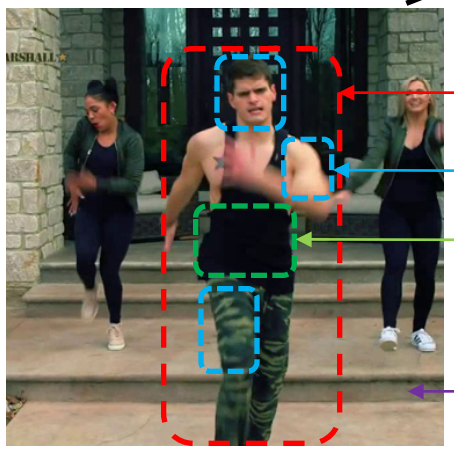


Rigid objects

We need a better paradigm to learn from motion.

How does human imagine motion?

Human brain



Static image

articulated

rigid

deformable

background

Kinematic properties

+

- keep left arm still
- raise right arm
- put right leg down
- lower the head
- …

Abstract instructions

=

Full motion

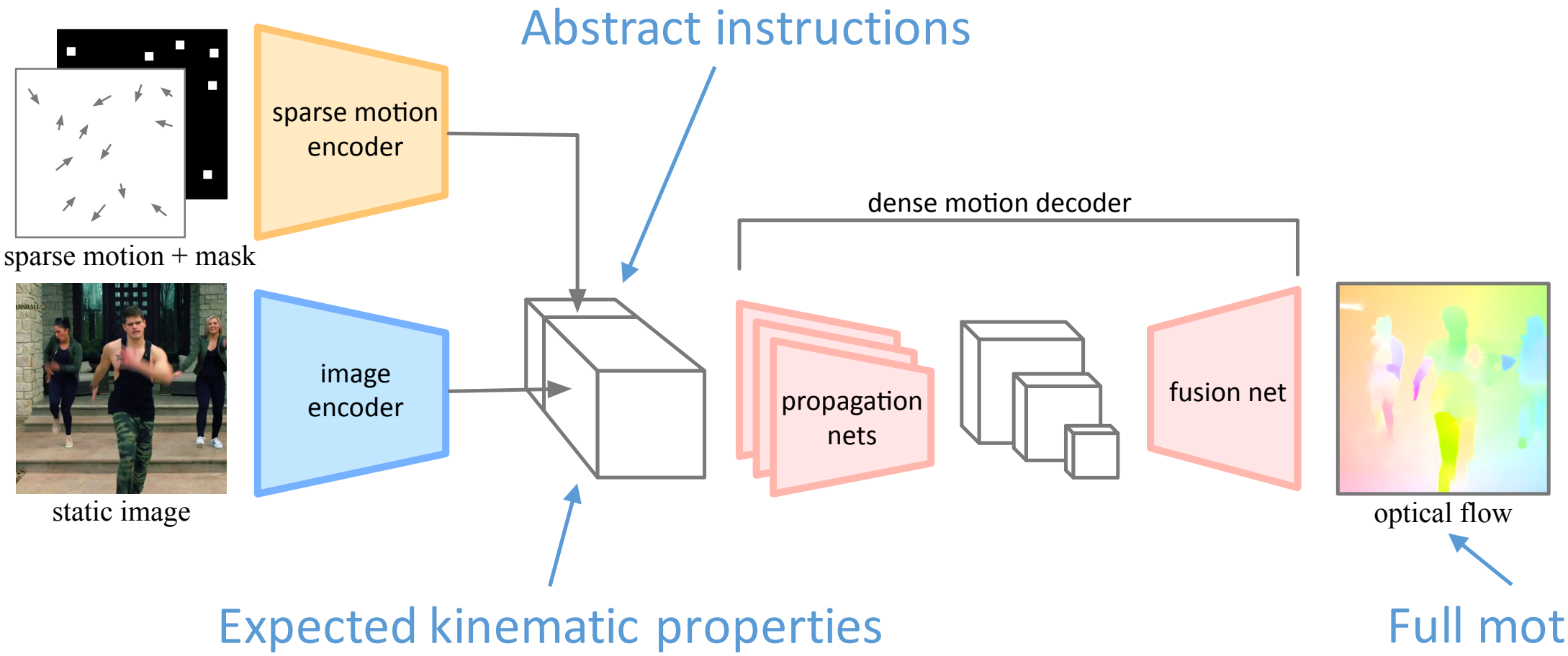
Observed

What we want to learn

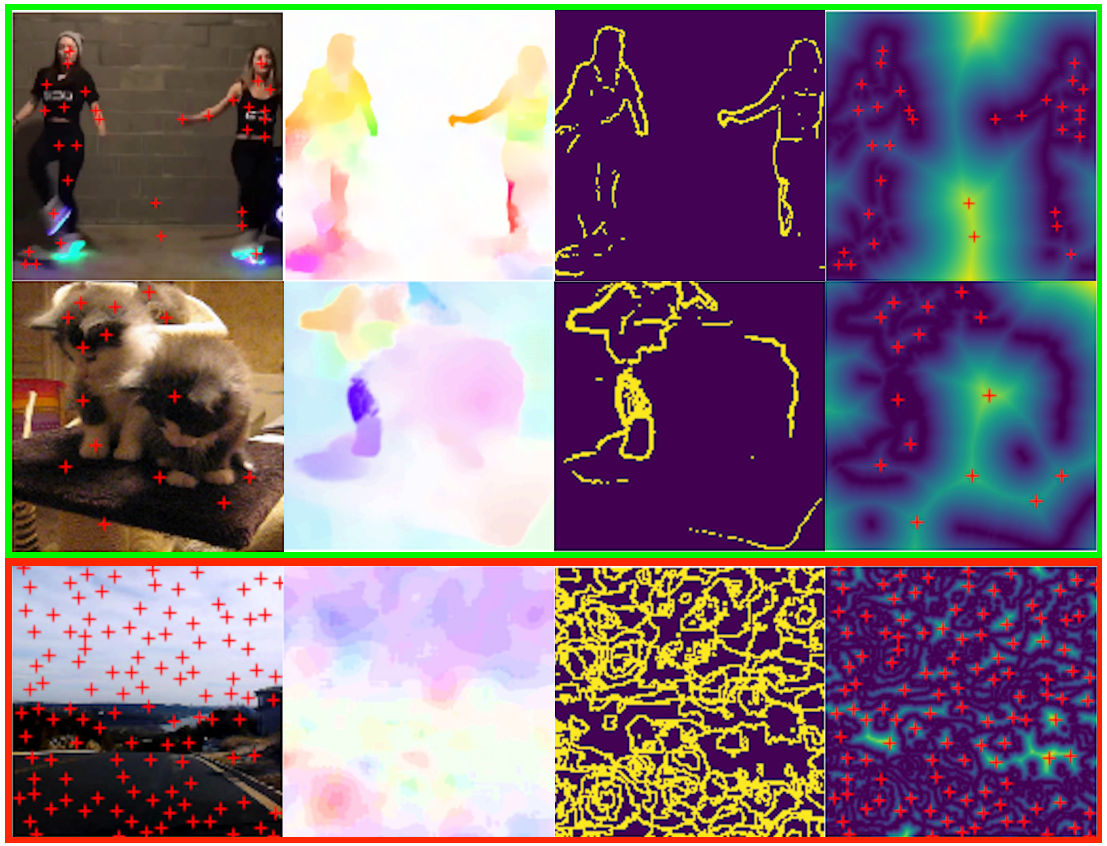
Sparse motion

Observed

Conditional Motion Propagation



Details



Sparse motion sampling

Flow quantization:
regression as classification.

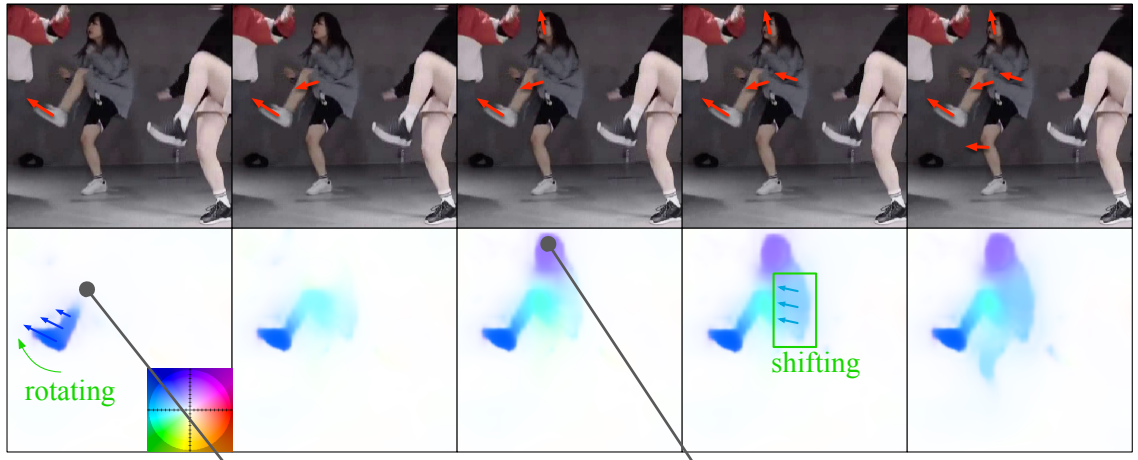
$$L_x = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C}\left(\mathbb{1}(Q_i^x = c)\log P_{i_c}^x\right),$$

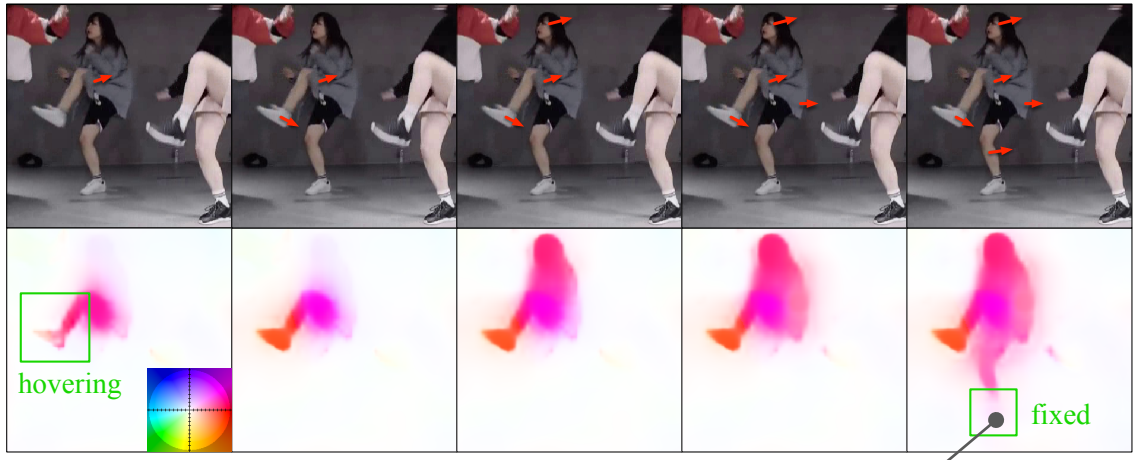
$$L_y = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C}\left(\mathbb{1}(Q_i^y = c)\log P_{i_c}^y\right)$$

L_x, L_y: loss in x, y direction.
Q: quantized flow. P: probability.

Codes and demos

- https://github.com/XiaohangZhan/conditional-motion-propagation





Kinematic property

Rigidity

Physical feasibility