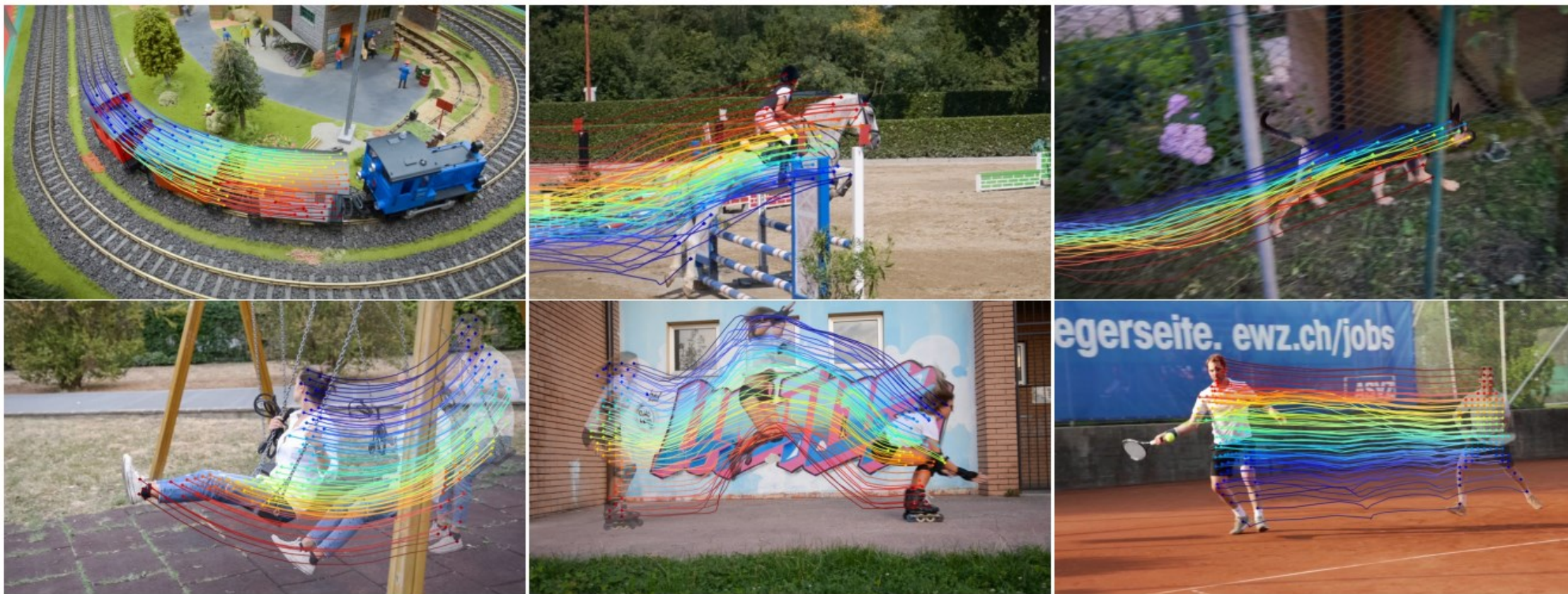


# Tracking Everything Everywhere All at Once

Qianqian Wang<sup>1,2</sup> Yen-Yu Chang<sup>1</sup> Ruojin Cai<sup>1</sup> Zhengqi Li<sup>2</sup>  
Bharath Hariharan<sup>1</sup> Aleksander Holynski<sup>2,3</sup> Noah Snavely<sup>1,2</sup>

<sup>1</sup>Cornell University   <sup>2</sup>Google Research   <sup>3</sup>UC Berkeley



# 背景

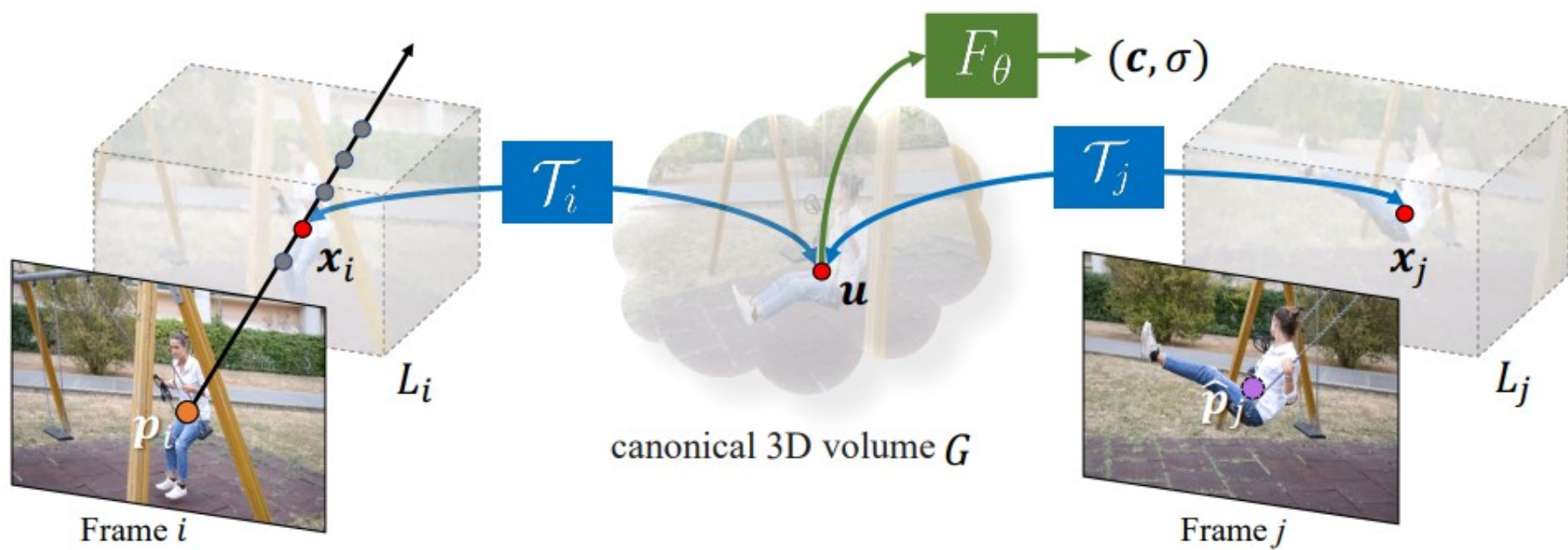
- 运动估计的重要性
- 目前的技术挑战:长序列跟踪、遮挡、时间一致性

# 本文的方法

- 表示视频为3D体积,定义网络F映射坐标到密度和颜色
- 密度反映规范空间信息,颜色用于损失约束

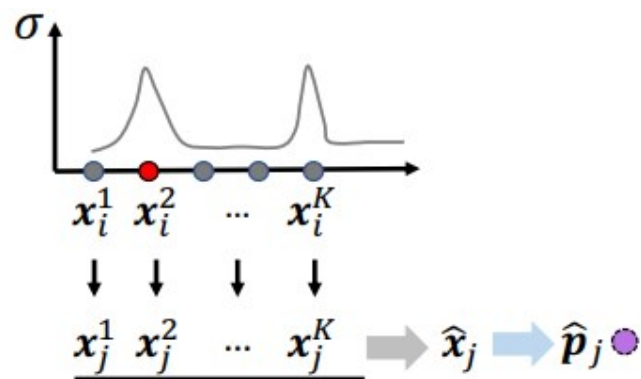


# 本文的方法



(a) OmniMotion representation

For a given  $p_i$  (orange dot):



→ Alpha composite    → Project

(b) Computing 2D motion

# 关键技术

- 3D双射
  - 定义双射映射,帧间3D点映射到规范3D坐标
  - 使用可逆网络实现
- 帧间运动
  - 通过采样提升2D点到3D,映射到目标帧
  - 获得颜色和密度,聚合目标帧点

# 优化

- 优化1

收集运动数据，使用光流方法,应用一致性检查过滤误匹配

- 优化2

应用损失函数，最小化流、光度、加速度损失

- 优化3

硬挖掘监督，计算误差图指导重要数据采样

# 实验结果

- 最左侧的图像显示了第一帧中的查询点，而右边的三个图像显示了随时间变化的跟踪结果。
- 本文的方法成功地跟踪了 swing 和 india 中的遮挡事件，而基线方法则失败了。
- 本文的方法还可以检测到遮挡并在点被遮挡时提供合理的位置估计。



# 实验结果

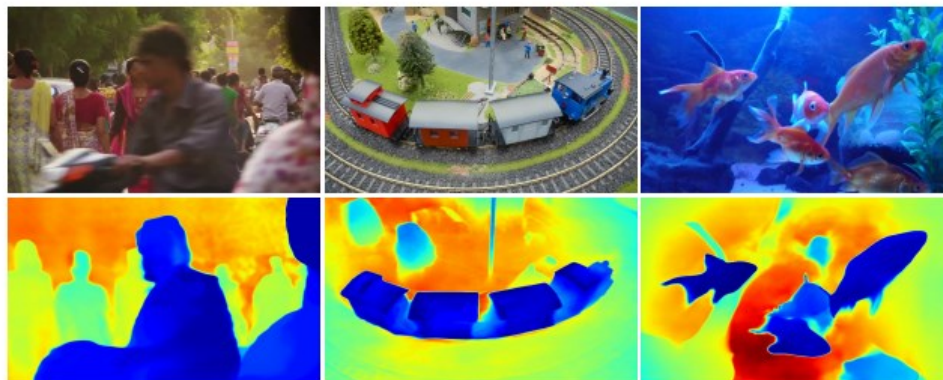
Method	Kinetics				DAVIS				RGB-Stacking			
	AJ $\uparrow$	$< \delta_{\text{avg}}^x \uparrow$	OA $\uparrow$	TC $\downarrow$	AJ $\uparrow$	$< \delta_{\text{avg}}^x \uparrow$	OA $\uparrow$	TC $\downarrow$	AJ $\uparrow$	$< \delta_{\text{avg}}^x \uparrow$	OA $\uparrow$	TC $\downarrow$
RAFT-C [66]	31.7	51.7	84.3	0.82	30.7	46.6	80.2	0.93	42.0	56.4	91.5	0.18
RAFT-D [66]	50.6	66.9	85.5	3.00	34.1	48.9	76.1	9.83	72.1	<u>85.1</u>	92.1	1.04
TAP-Net [15]	48.5	61.7	86.6	6.65	38.4	53.4	81.4	10.82	61.3	73.7	91.5	1.52
PIPs [23]	39.1	55.3	82.9	1.30	39.9	56.0	81.3	1.78	37.3	50.6	89.7	0.84
Flow-Walk-C [5]	40.9	55.5	84.5	<u>0.77</u>	35.2	51.4	80.6	0.90	41.3	55.7	<u>92.2</u>	<u>0.13</u>
Flow-Walk-D [5]	46.9	65.9	81.8	3.04	24.4	40.9	76.5	10.41	66.3	82.7	91.2	0.47
Deformable-Sprites [81]	25.6	39.5	71.4	1.70	20.6	32.9	69.7	2.07	45.0	58.3	84.0	0.99
Ours (TAP-Net)	<u>53.8</u>	<u>68.3</u>	<u>88.8</u>	<u>0.77</u>	<u>50.9</u>	<u>66.7</u>	85.7	<u>0.86</u>	<u>73.4</u>	84.1	<u>92.2</u>	0.11
Ours (RAFT)	55.1	69.6	89.6	0.76	51.7	67.5	<u>85.3</u>	0.74	77.5	87.0	93.5	<u>0.13</u>



# 消融实验

Method	AJ $\uparrow$	$< \delta_{\text{avg}}^x \uparrow$	OA $\uparrow$	TC $\downarrow$
No invertible	12.5	21.4	76.5	0.97
No photometric	42.3	58.3	84.1	0.83
Uniform sampling	47.8	61.8	83.6	0.88
Full	51.7	67.5	85.3	0.74

DAVIS 数据集的消融研究 [50]。



我们的表示法提取的伪深度图，其中蓝色表示较近的物体，红色表示较远的物体。