

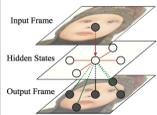
Motivation

Video Frame Synthesis

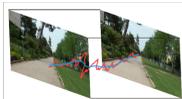


- * Complex motion (camera motion & scene motion)
- * High-res video generation (720p - 1080p)

Previous Attempts

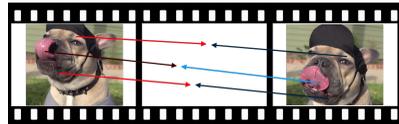


Hallucination-based Methods
[Ranzato et al.], [Srivastava et al.], [Mathieu et al.]

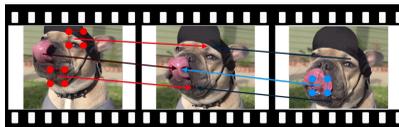


Flow-based Methods
[Baker et al.], [Mahajan et al.], [Walker et al.]

Our 3D Voxel Flow



- * Symmetric bi-directional flows
- * Selection mask between frames



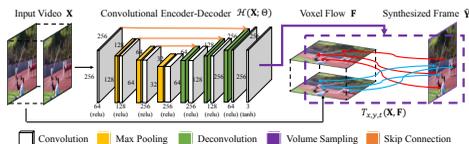
- * Differentiable bilinear sampling

Unsupervised Training

Any video can be used as training data by dropping, and then learning to predict, existing frames.

Approach

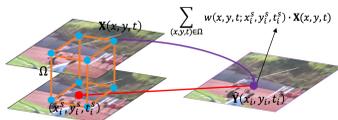
Deep Voxel Flow



□ Convolution □ Max Pooling □ Deconvolution □ Volume Sampling □ Skip Connection

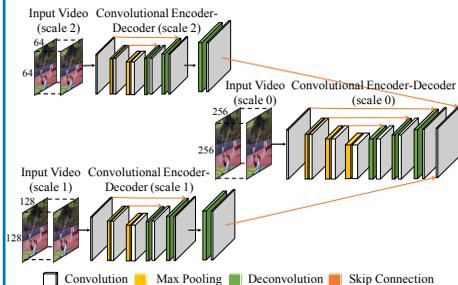
Conv. encoder-decoder with voxel flow layer

Voxel Flow Layer



$$\begin{aligned} \mathbf{W}^{000} &= (1 - [\mathbf{L}_x^0 - [\mathbf{L}_x^0]]) (1 - [\mathbf{L}_y^0 - [\mathbf{L}_y^0]]) (1 - \Delta t) \\ \mathbf{W}^{100} &= (\mathbf{L}_x^0 - [\mathbf{L}_x^0]) (1 - [\mathbf{L}_y^0 - [\mathbf{L}_y^0]]) (1 - \Delta t) \\ &\vdots \\ \mathbf{W}^{011} &= (1 - [\mathbf{L}_x^1 - [\mathbf{L}_x^1]]) (\mathbf{L}_y^1 - [\mathbf{L}_y^1]) \Delta t \\ \mathbf{W}^{111} &= (\mathbf{L}_x^1 - [\mathbf{L}_x^1]) (\mathbf{L}_y^1 - [\mathbf{L}_y^1]) \Delta t, \end{aligned}$$

Multi-scale Voxel Flow

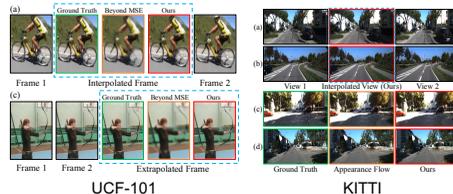


□ Convolution □ Max Pooling □ Deconvolution □ Skip Connection

Handle large motion via multi-scale guidance

Experiments

Video/View Interpolation & Extrapolation



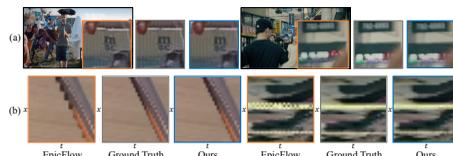
Self-supervised Learning

Method	EPE	Method	Acc.
L.D Flow [2]	12.4	Random	39.1
B. Basics [7]	9.9	Unsup. Video [33]	43.8
FlowNet [5]	9.1	ShuffNet&Learn [23]	50.2
EpicFlow [25]	3.8	ImageNet [15]	63.3
Ours (w/o fl.)	14.6	Ours (w/o fl.)	48.7
Ours	9.5	Ours	52.4

Left: flow estimation

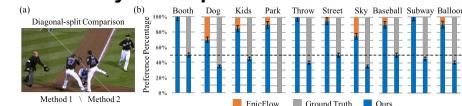
Right: action recognition

Slow-motion Effects from 15fps to 30fps



(a) close-ups of patches (b) temporal close-ups of x-t slices

User Study on 720p Videos



The result of voxel flow is indistinguishable from ground truth.

Project Page:

<https://liuziwei7.github.io/projects/VoxelFlow>

