Range-nullspace Video Frame Interpolation with Focalized Motion Estimation

Zhiyang Yu[1], Yu Zhang[2], Dongqing Zou[2,3], Xijun Chen[1], Jimmy S. Ren[2,3] and Shunqing Ren[1]

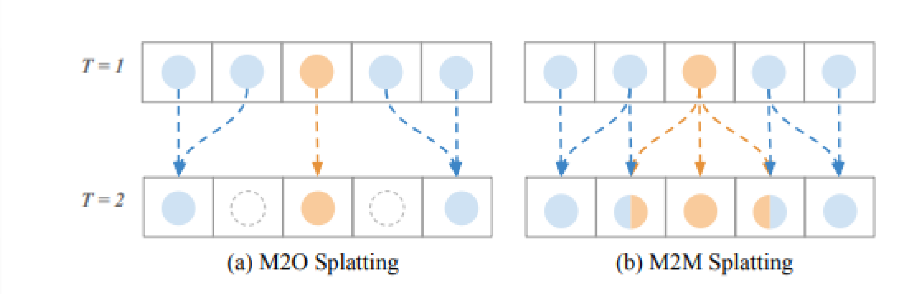
[1]Harbin Institute of Technology, [2]SenseTime Research and Tetras.AI
[3]Qing Yuan Research Institute, Shanghai Jiao Tong University

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

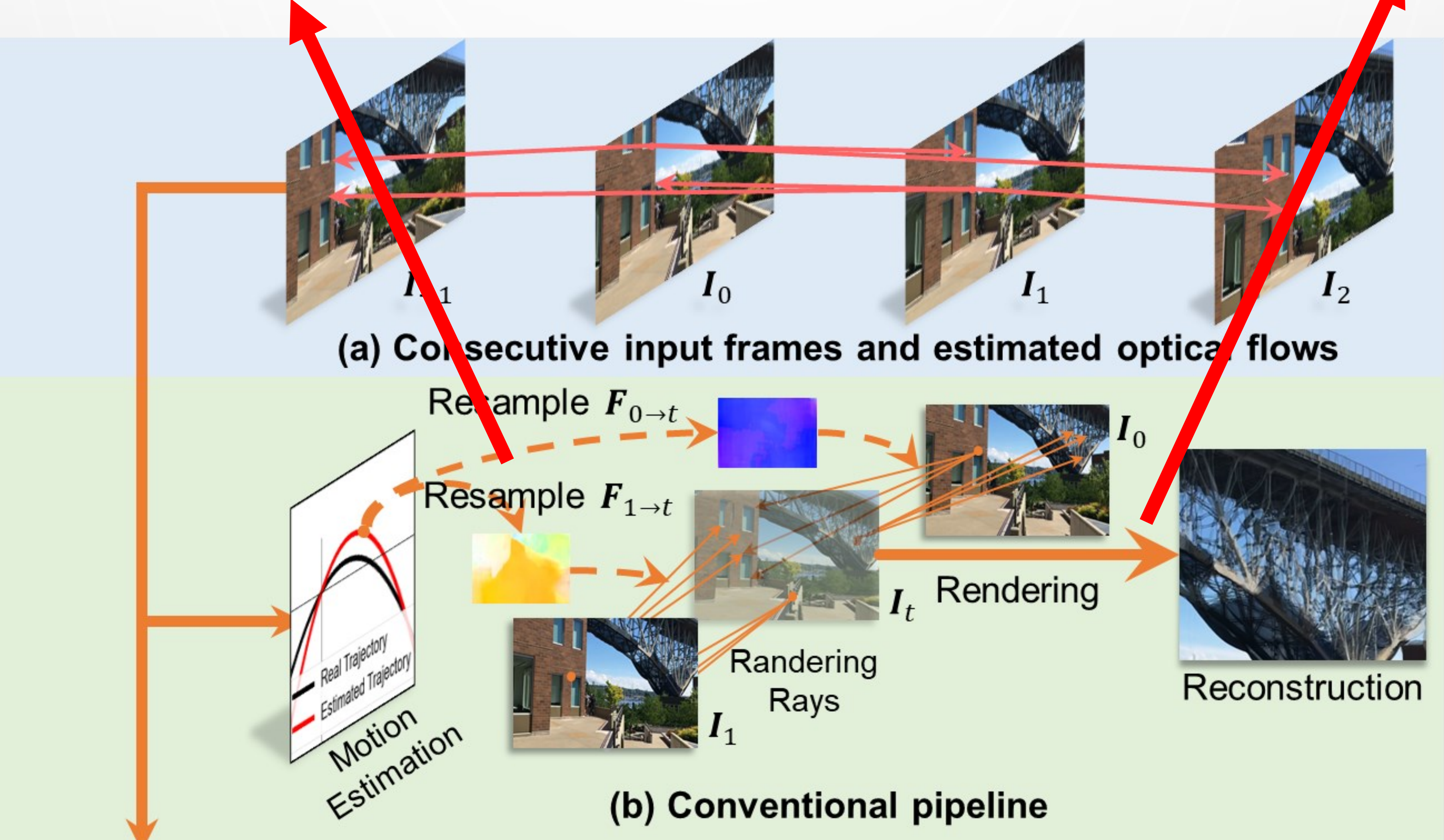
Continuous-time video frame interpolation

Motivation

frame -1
frame 0
frame 1
frame 2
ground truth
linear model
quadratic model

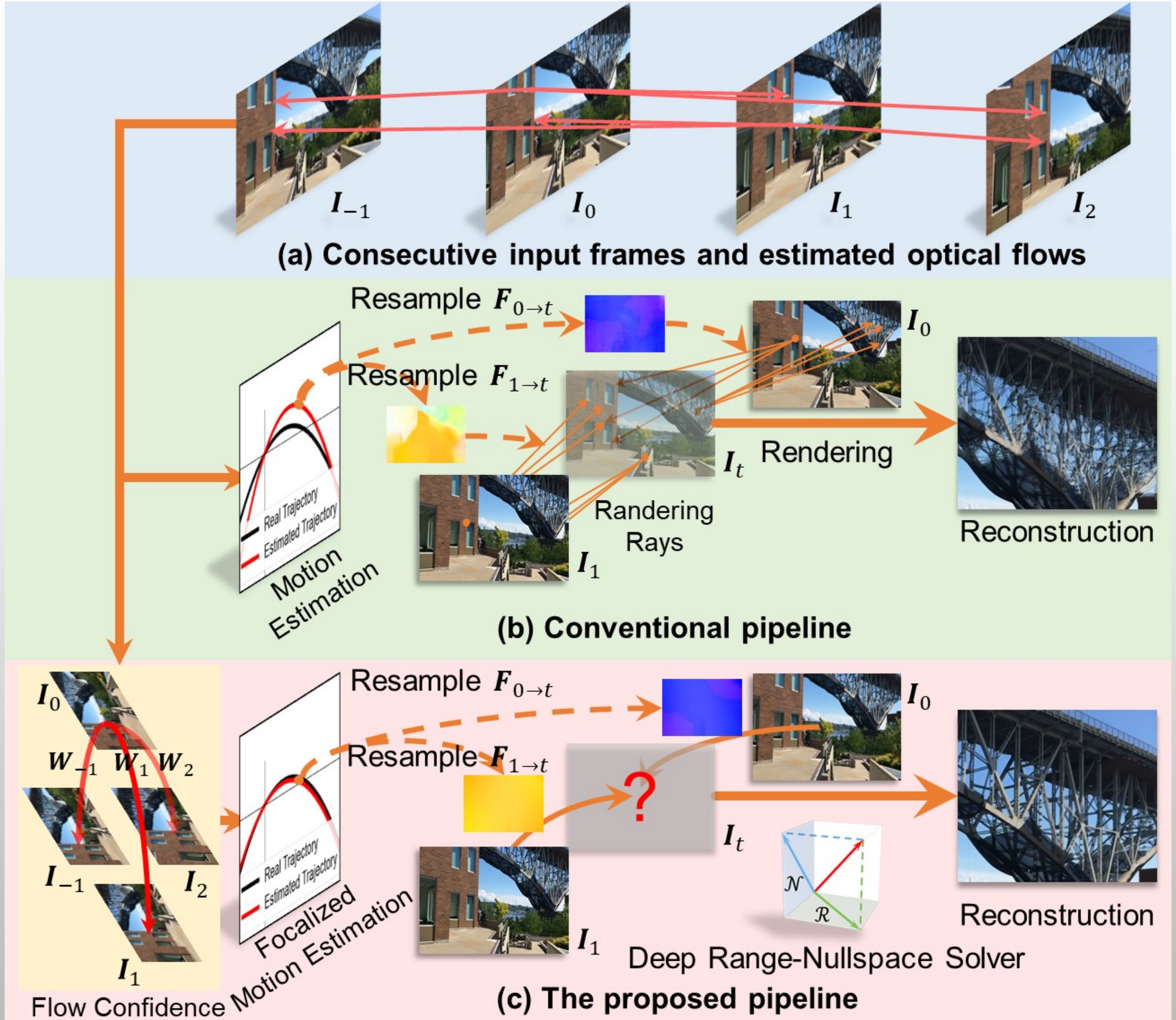


T = 1
T = 2
(a) M2O Splatting
(b) M2M Splatting



I_{-1}
I_0
I_1
I_2
(a) Consecutive input frames and estimated optical flows
Resample $F_{0\to t}$
Resample $F_{1\to t}$
Real Trajectory
Estimated Trajectory
Motion Estimation
I_0
I_t
Rendering
Randering Rays
I_1
Reconstruction
(b) Conventional pipeline

Motivation



I_{-1}
I_0
I_1
I_2
(a) Consecutive input frames and estimated optical flows
Resample $F_{0\to t}$
Resample $F_{1\to t}$
Real Trajectory
Estimated Trajectory
Motion Estimation
I_0
I_t
Rendering
I_1
Randering Rays
Reconstruction
(b) Conventional pipeline
I_0
W_{-1} W_1 W_2
I_{-1}
I_2
I_1
Flow Confidence
Resample $F_{0\to t}$
Resample $F_{1\to t}$
Real Trajectory
Estimated Trajectory
Focalized Motion Estimation
?
I_0
I_t
I_1
$\mathcal{N}$
$\mathcal{R}$
Deep Range-Nullspace Solver
Reconstruction
(c) The proposed pipeline

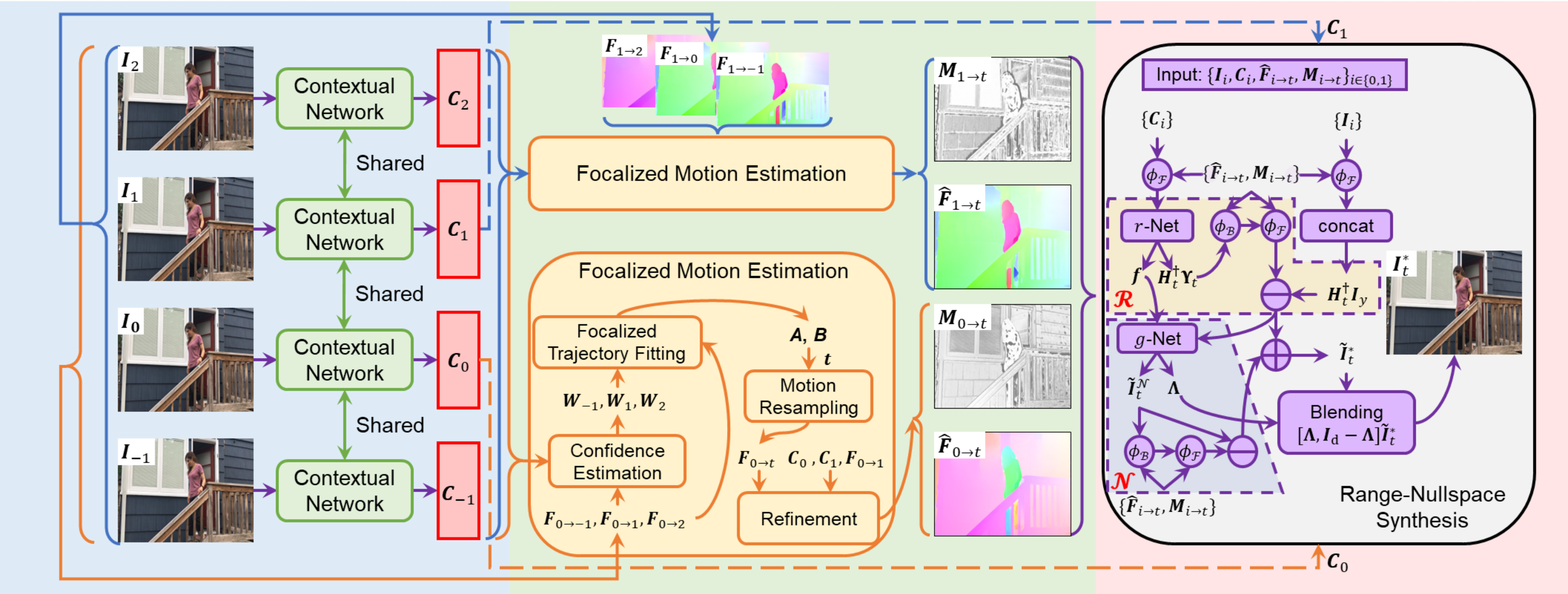
Contributions

1. A novel lightweight VFI framework is proposed, which refreshes the records of 7 out of 10 public VFI benchmarks.

2. The idea of focalized trajectory fitting, which improves parametric motion estimation in VFI and generates better resampling quality of intermediate flows.

3. A new perspective that treats intermediate frame synthesis as an ill-posed problem, solved with a deep range nullspace solver that decouples frame synthesis into several orthogonal tasks.

Pipeline



$$\hat{\mathbf{F}}_{0\to t}, \mathbf{M}_{0\to t} = \mathrm{FME}(t, \{\mathbf{F}_{0\to j}\}_{j\neq 0}, \{\mathbf{C}_i\}_{\forall i}),$$

$$\hat{\mathbf{F}}_{1\to t}, \mathbf{M}_{1\to t} = \mathrm{FME}(1-t, \{\mathbf{F}_{1\to j}\}_{j\neq 1}, \{\mathbf{C}_i\}_{\forall i})$$

$$\mathbf{I}_t^* = \mathrm{RNS}\left(\{\hat{\mathbf{F}}_{i\to t}, \mathbf{M}_{i\to t}, \mathbf{C}_i\}_{i\in\{0,1\}}\right)$$

Focalized Motion Estimation

Confidence Estimation

$$\mathbf{R}_i = \mathbf{C}_0^b - \phi_{\mathcal{B}}\left(\mathbf{C}_i^b, \mathbf{F}_{0\to i}\right), i \in \{-1, 1, 2\}$$

$$\mathbf{W}_{-1}, \mathbf{W}_1, \mathbf{W}_2 = \mathcal{W}\left(\mathbf{R}_{-1}, \mathbf{R}_1, \mathbf{R}_2\right)$$

Focalized Trajectory Fitting

$$f(\tau; \mathbf{A}, \mathbf{B}) = \mathbf{A}\tau^2 + \mathbf{B}\tau, \mathbf{A}, \mathbf{B} \in \mathbb{R}^{H\times W\times 2}$$

$$\min_{\mathbf{A},\mathbf{B}} \sum_{i\in\{-1,1,2\}} \left\| \sqrt{\mathbf{W}_i} \odot \left(\mathbf{F}_{0\to i} - f(i; \mathbf{A}, \mathbf{B})\right)\right\|_2^2$$

$$\hat{\Theta}_{x,y,z} = \left(\mathbf{X}\mathbf{W}_{x,y}\mathbf{X}^T\right)^{-1} \mathbf{X}\mathbf{W}_{x,y}\mathbf{F}_{x,y,z}$$

$$\mathbf{W}_{x,y} = \begin{bmatrix} \mathbf{W}_{-1}(x,y,0) & & \\ & \mathbf{W}_1(x,y,0) & \\ & & \mathbf{W}_2(x,y,0)(x,y) \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 4 \\ -1 & 1 & 2 \end{bmatrix}, \mathbf{F}_{x,y,z} = \begin{bmatrix} \mathbf{F}_{0\to -1}(x,y,z) \\ \mathbf{F}_{0\to 1}(x,y,z) \\ \mathbf{F}_{0\to 2}(x,y,z) \end{bmatrix}$$

Motion resampling and refinement

$$\boldsymbol{F}_{0\to t} = f(t; \hat{\mathbf{A}}, \hat{\mathbf{B}})$$

$$\Delta\mathbf{F}_{0\to t}, \mathbf{M}_{0\to t} = \mathcal{M}\left(\mathbf{F}_{0\to t}, \mathbf{C}_0, \overline{\mathbf{C}}_{1\to 0}\right)$$

$$\overline{\mathbf{C}}_{1\to 0} = \phi_{\mathcal{B}}(\mathbf{C}_1, \mathbf{F}_{0\to 1})$$

Range-nullspace Synthesis

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

$$\begin{bmatrix} \mathbf{I}_0 \\ \mathbf{I}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{0\to t} \\ \mathbf{H}_{1\to t} \end{bmatrix} \mathbf{I}_t + \begin{bmatrix} \boldsymbol{\Upsilon}_{0\to t} \\ \boldsymbol{\Upsilon}_{1\to t} \end{bmatrix}$$

$$\mathcal{R}(\boldsymbol{x}) \triangleq \mathbf{H}^{+}\mathbf{H}\boldsymbol{x}, \quad \mathcal{N}(\boldsymbol{x}) \triangleq (\mathbf{I}_d - \mathbf{H}^{+}\mathbf{H})\boldsymbol{x}$$

$$\underbrace{\begin{bmatrix} \mathbf{I}_0 \\ \mathbf{I}_1 \end{bmatrix}}_{\mathbf{I}_y} = \underbrace{\begin{bmatrix} \mathbf{H}_{0\to t} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{1\to t} \end{bmatrix}}_{\mathbf{H}_t} \tilde{\mathbf{I}}_t + \underbrace{\begin{bmatrix} \boldsymbol{\Upsilon}_{0\to t} \\ \boldsymbol{\Upsilon}_{1\to t} \end{bmatrix}}_{\boldsymbol{\Upsilon}_t}, \ \text{s.t. } \mathbf{C}\tilde{\mathbf{I}}_t = \mathbf{0}$$

$$\tilde{\mathbf{I}}_t = [\tilde{\mathbf{I}}_{t,1}^{\mathrm{T}}, \tilde{\mathbf{I}}_{t,2}^{\mathrm{T}}]^{\mathrm{T}}, \ \mathbf{C} = [\mathbf{I}_d, -\mathbf{I}_d]$$

$$\tilde{\mathbf{I}}_t^{*} = \mathbf{H}_t^{+}\mathbf{I}_y - \mathcal{R}(\mathbf{H}_t^{+}\boldsymbol{\Upsilon}_t) + \mathcal{N}(g(\cdot))$$

Range-nullspace Synthesis

$$\boldsymbol{f}, \mathbf{H}_t^{+}\boldsymbol{\Upsilon}_t = r\left(\left\{\phi_{\mathcal{F}}(\mathbf{C}_i; \hat{\mathbf{F}}_{i\to t}, \mathbf{M}_{i\to t})\right\}_{i\in\{0,1\}}\right)$$

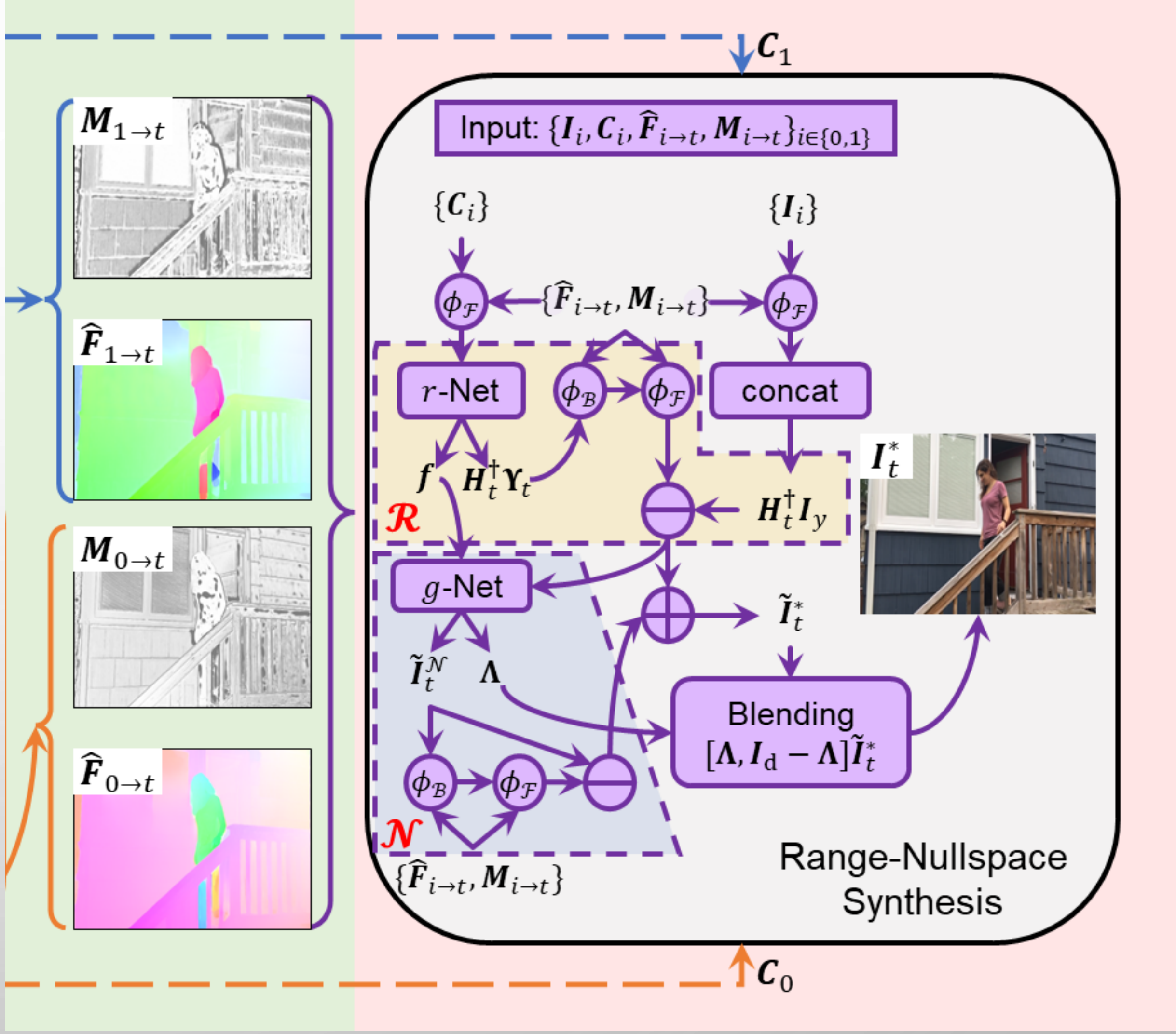
$$\boldsymbol{\Lambda}, \tilde{\mathbf{I}}_t^{\mathcal{N}} = g\left(\mathbf{H}_t^{+}\mathbf{I}_y - \mathcal{R}(\mathbf{H}_t^{+}\boldsymbol{\Upsilon}_t), \mathbf{H}_t^{+}\mathbf{I}_y, \boldsymbol{f}\right),$$

$$\tilde{\mathbf{I}}_t^{*} = \mathbf{H}_t^{+}\mathbf{I}_y - \mathcal{R}(\mathbf{H}_t^{+}\boldsymbol{\Upsilon}_t) + \mathcal{N}(\tilde{\mathbf{I}}_t^{\mathcal{N}}),$$

$$\mathbf{I}_t^{*} = [\boldsymbol{\Lambda}, \mathbf{I}_d - \boldsymbol{\Lambda}]\tilde{\mathbf{I}}_t^{*}.$$

$$\mathbf{H}_{i\to t}\boldsymbol{x} = \phi_{\mathcal{B}}(\boldsymbol{x}; \hat{\mathbf{F}}_{i\to t}),$$

$$\mathbf{H}_{i\to t}^{+}\boldsymbol{x} \approx \phi_{\mathcal{F}}(\boldsymbol{x}; \hat{\mathbf{F}}_{i\to t}, \mathbf{M}_{i\to t}).$$



Experiments

8× interpolation.

- **Training Set**: GoPro(Train Spilt) or X4K1000FPS (Train Spilt)
- **Evaluation**: GoPro(Test Split), Adobe240 and X4K1000FPS (Test Spilt)

2× interpolation.

- **Training Set**: Vimeo-90k(Septulets, Train Split)
- **Evaluation**: Vimeo-90k(Septulets, Test Split), UCF101, DAVIS and SNU-FILM (Easy Medium Hard Extreme)

Experiments

Table 1: Quantitative results of 8× VFI in terms of PSNR/SSIM and the number of parameters, evaluated on GoPro, Adobe240 and X4K1000FPS datasets. The best performed model is highlighted in red and the second best is colored in blue.

	GoPro	Adobe240	X4K1000FPS	Param(M)
SloMo	29.71/0.924	29.63/0.927	25.07/0.795	39.61
QVI	30.52/0.941	31.41/0.955	28.06/0.855	29.23
EQVI	30.81/0.942	32.13/0.959	26.96/0.843	28.07
DAIN	29.53/0.920	30.53/0.939	27.28/0.835	24.03
EDSC	29.20/0.916	29.87/0.931	25.30/0.811	8.95
FLAVR	31.10/0.942	30.92/0.938	24.50/0.791	42.06
XVFI	29.80/0.925	29.74/0.930	28.42/0.881	5.61
M2M	30.52/0.933	29.93/0.931	30.04/0.905	7.61
IFRNet	30.00/0.928	29.62/0.925	23.77/0.793	19.69
RIFE_m	29.79/0.925	29.81/0.930	28.70/0.880	10.71
DBVI	31.73/0.947	33.28/0.965	31.10/0.928	15.18
Ours	32.31/0.951	33.32/0.964	31.97/0.932	10.10

Table 2: Quantitative results of 8× VFI on X4K1000FPS dataset

	AdaCoF	FeFlow	SloMo	QVI	EQVI	DAIN	FLAVR	XVFI	M2M	IFRNet	RIFE_m	DBVI	Ours
PSNR	25.81	25.16	27.77	29.96	28.27	27.52	27.92	30.12	30.84	28.36	30.75	32.89	33.88
SSIM	0.772	0.783	0.849	0.892	0.860	0.821	0.853	0.870	0.916	0.870	0.914	0.939	0.946

Experiments

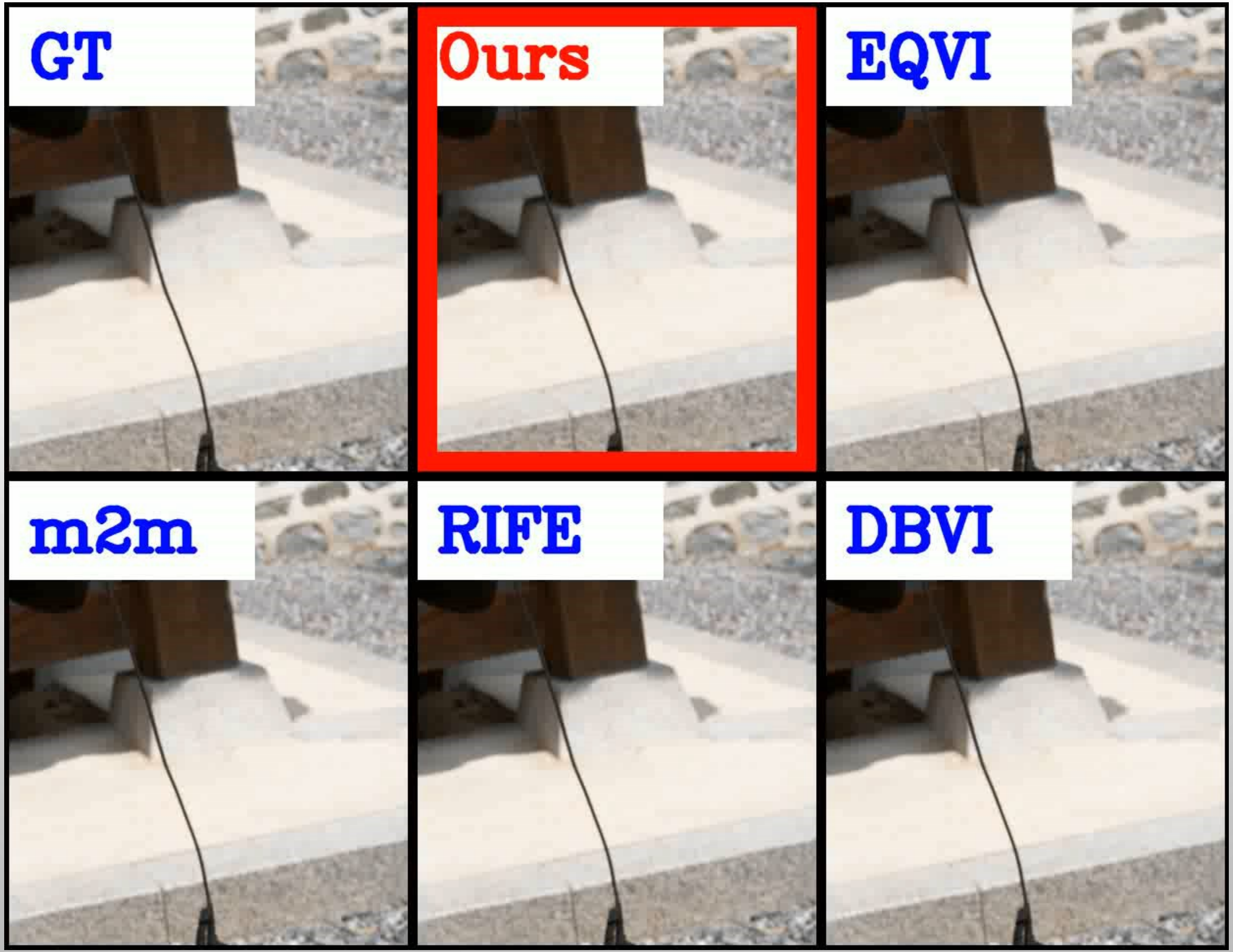
Table 1: Quantitative results on 2× interpolation in terms of PSNR/SSIM with number of parameters and flops. All the methods are trained on the training set of Vimeo-90K (septulets).

	Vimeo-90K (septulets)	UCF101	DAVIS	SNU-FILM				Param(M)
				Easy	Medium	Hard	Extreme	
SloMo	34.43/0.969	32.45/0.967	26.10/0.862	36.12/0.984	33.44/0.972	29.17/0.928	24.14/0.843	39.61
QVI	34.98/0.970	32.87/0.966	27.20/0.874	39.53/0.990	36.43/0.983	31.07/0.947	24.96/0.856	29.23
EQVI	35.16/0.973	32.99/0.970	27.51/0.891	37.44/0.978	35.19/0.981	30.72/0.946	25.42/0.868	28.07
XVFI	35.21/0.970	32.68/0.968	26.89/0.868	39.21/0.989	34.96/0.977	29.43/0.928	24.02/0.841	5.61
DAIN	33.57/0.964	31.65/0.963	26.61/0.867	38.53/0.988	34.34/0.974	29.50/0.930	24.54/0.851	24.03
BMBC	34.76/0.965	32.61/0.955	26.42/0.868	39.90/0.991	35.34/0.978	29.34/0.927	23.65/0.837	11.01
ABME	35.67/0.972	32.81/0.969	27.00/0.868	39.59/0.990	35.77/0.977	30.58/0.936	25.42/0.864	18.1
EDSC	34.52/0.967	32.67/0.968	26.28/0.849	40.01/0.990	35.37/0.978	29.59/0.926	24.39/0.843	8.95
GDConv	35.58/0.972	33.11/0.969	27.02/0.870	40.36/0.991	36.14/0.982	30.25/0.940	24.82/0.860	5.14
CAIN	34.69/0.969	32.40/0.966	27.12/0.872	39.33/0.989	35.34/0.977	30.15/0.933	24.88/0.855	42.78
FLAVR	36.30/0.975	33.33/0.971	27.44/0.873	40.44/0.991	36.37/0.981	30.87/0.942	25.18/0.862	42.06
M2M	35.56/0.973	32.70/0.969	27.57/0.887	39.35/0.989	35.23/0.977	29.99/0.934	24.83/0.857	7.60
IFRNet	36.37/0.976	32.87/0.969	27.94/0.890	39.68/0.990	35.57/0.979	30.21/0.938	24.71/0.855	19.69
Softsplat	35.76/0.972	32.89/0.970	27.42/0.878	-	-	-	-	12.46
$RIFE_m$	35.87/0.974	32.64/0.969	27.75/0.886	39.50/0.990	35.46/0.978	30.17/0.936	24.79/0.854	10.71
VFIT-B	36.96/0.978	33.44/0.971	28.17/0.889	40.57/0.991	36.54/0.982	31.04/0.945	25.50/0.867	29.1
ST-MFNet	36.46/0.976	33.46/0.971	28.30/0.896	40.78/0.992	37.12/0.984	31.61/0.951	25.78/0.874	21.03
DBVI	36.17/0.976	33.01/0.970	28.61/0.905	40.46/0.991	36.95/0.985	31.68/0.953	25.90/0.876	21.69
Ours	36.33/0.975	33.25/0.970	28.84/0.905	40.67/0.991	37.36/0.985	32.21/0.955	26.22/0.877	10.10

Visual Results


Input
JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA
sachtler

Output (8X)

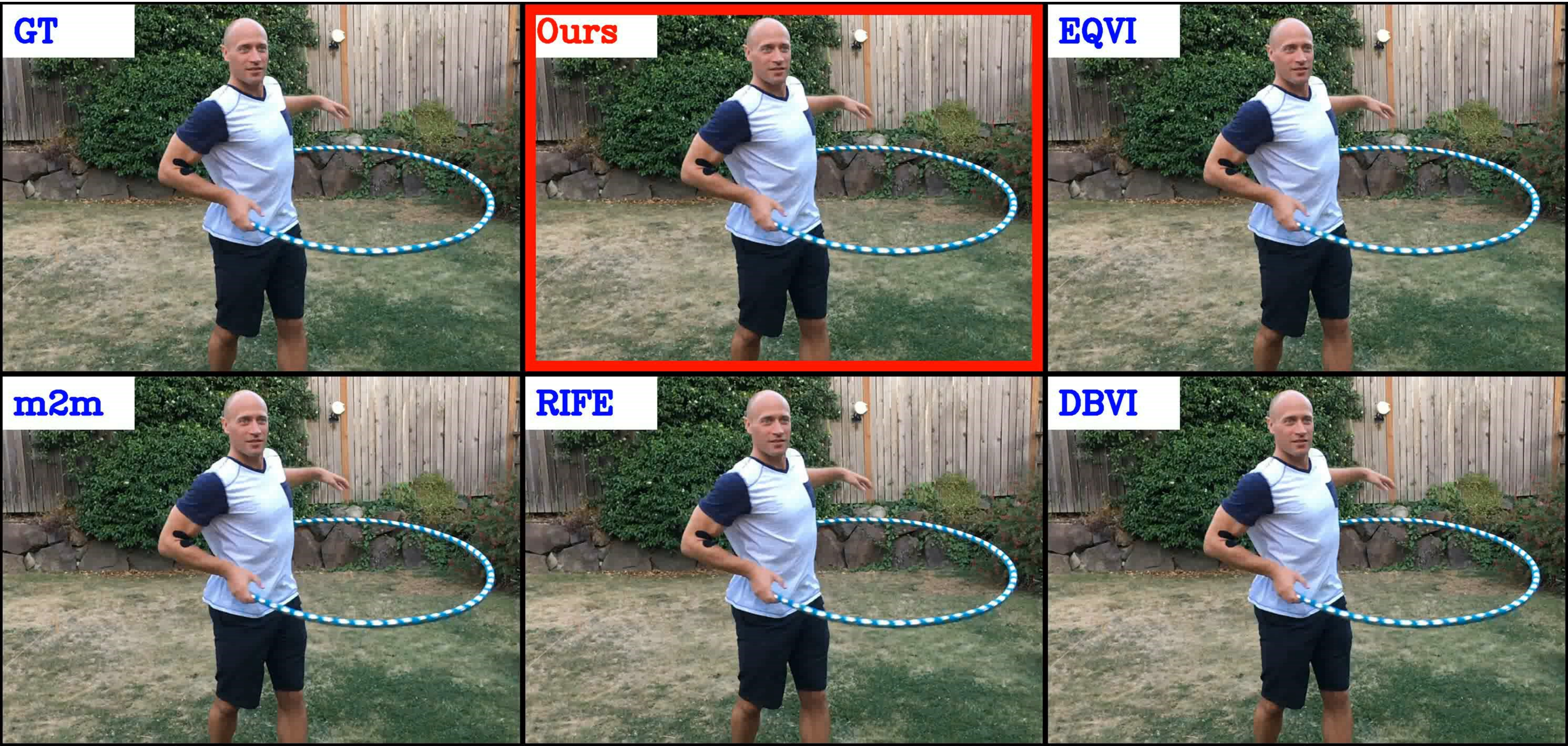



Input

Output (8X)


Input

Output (8X)



THANKS!