PhaseNet for Video Frame Interpolation

Simone Schaub-Meyer

Joint work with Abdelaziz Djelouah, Brian McWilliams, Alexander Sorkine-Hornung, Christopher Schroers

Work done at ETH Zurich and Disney Research

TECHNISCHE UNIVERSITÄT DARMSTADT

Motivation

Frame interpolation method without explicit correspondence handling

» Suitable for challenging scenarios (brightness change, motion blur)

Input frames

Interpolation result

Ours MDP-Flow2

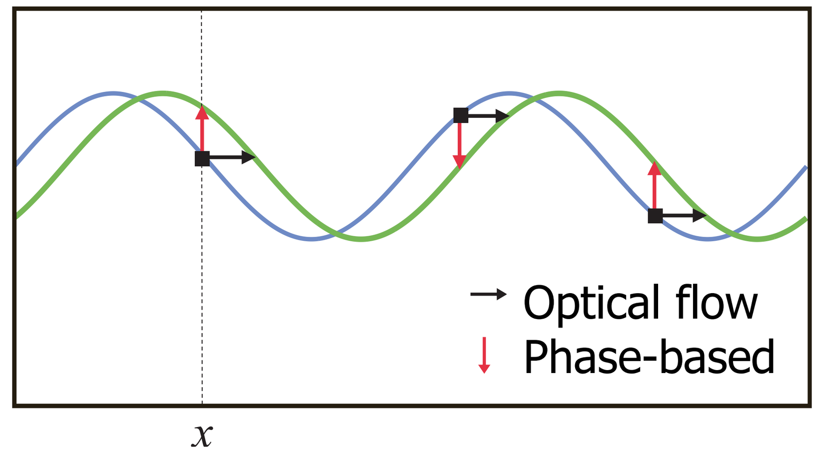
Contributions

» PhaseNet: Novel architecture which operates on steerable pyramid decomposition of images using phase-based motion representation

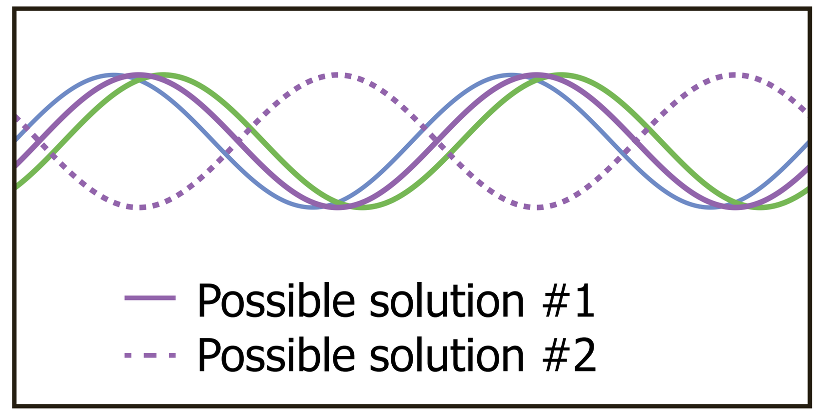
» Phase loss speeds up training and improves sharpness of result

Phase-based Motion Representation

Optical flow vs. phase-based approach

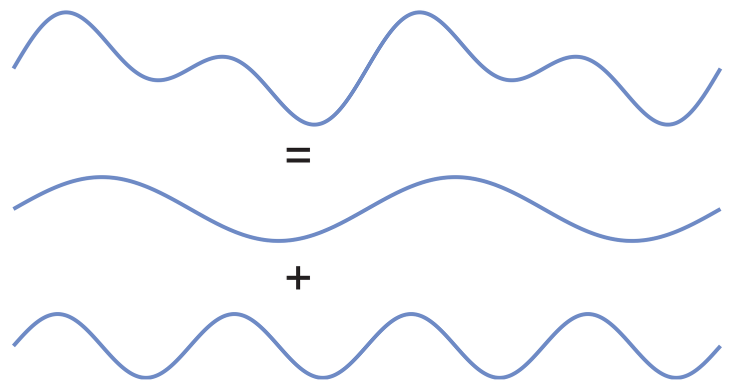


Problem: Periodicity of phase values



Decomposition in 1 dimension

$$f(x) = \sum_{\omega=-\infty}^{\omega=+\infty} A_\omega e^{i\phi_\omega}$$



Decomposition in 2 dimensions

» With complex steerable filters [3]

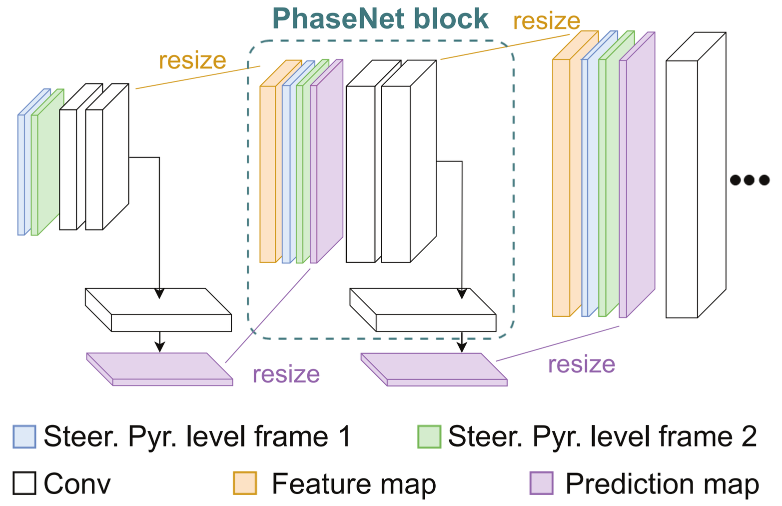
$$R_{\omega,\theta}(x,y) = (I * \Psi_{\omega,\theta})(x,y) = C_{\omega,\theta}(x,y) + i\,S_{\omega,\theta}(x,y) = A_{\omega,\theta}(x,y)\,e^{i\phi_{\omega,\theta}(x,y)}$$

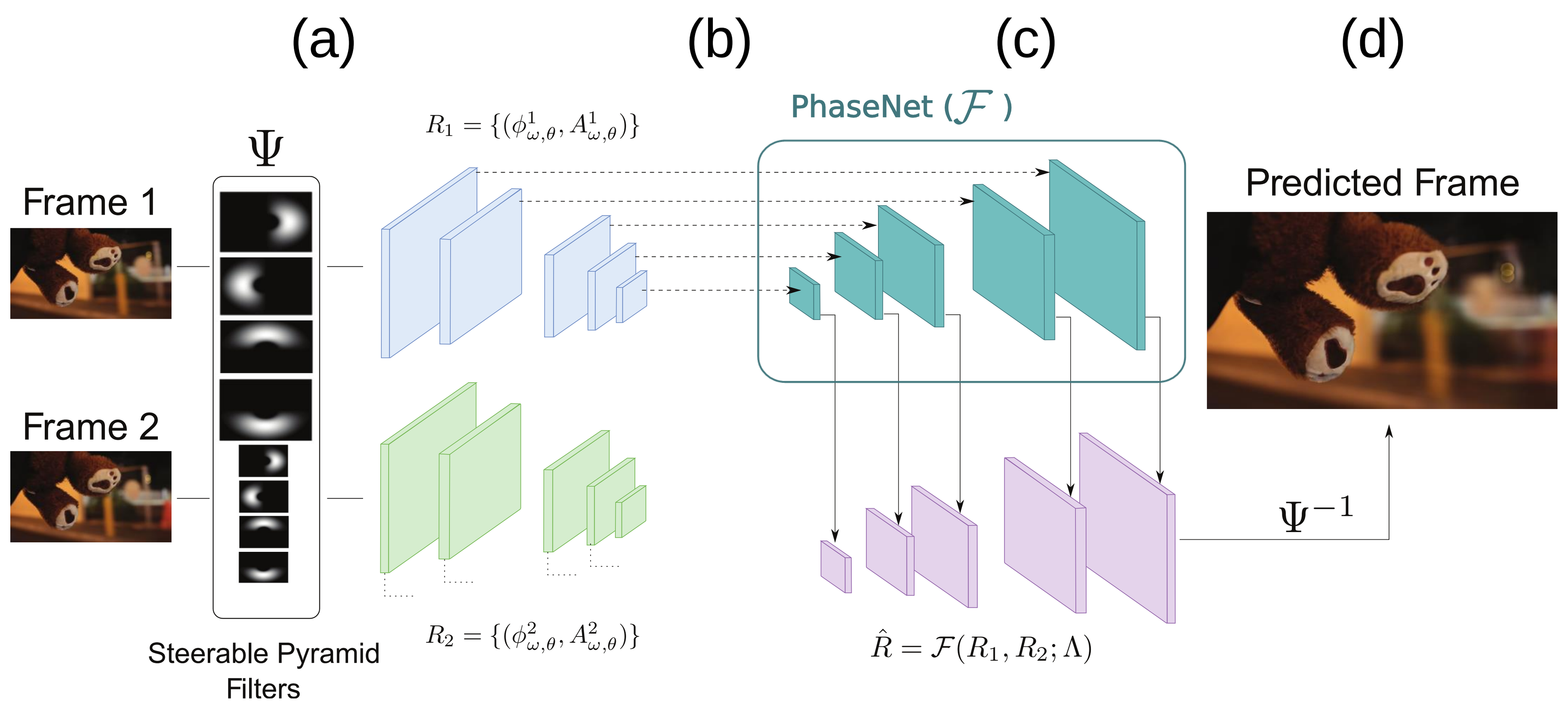
Amplitude $A_{\omega,\theta}(x,y) = |R_{\omega,\theta}(x,y)|$

Phase $\phi_{\omega,\theta}(x,y) = \mathrm{Im}(\log(R_{\omega,\theta}(x,y)))$

Algorithm

a) Decompose images
b) Input decomposition into PhaseNet
c) Predict pyramid decomposition
d) Reconstruct image from prediction



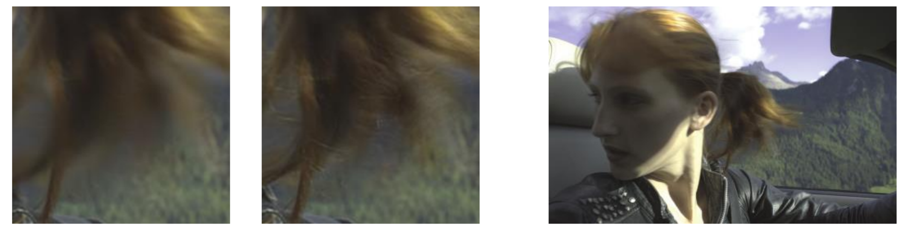


Loss Function

Total loss $\mathcal{L} = \mathcal{L}_1 + \nu\mathcal{L}_{\text{phase}}$

» Image loss $\mathcal{L}_1 = ||I - \hat{I}||_1$

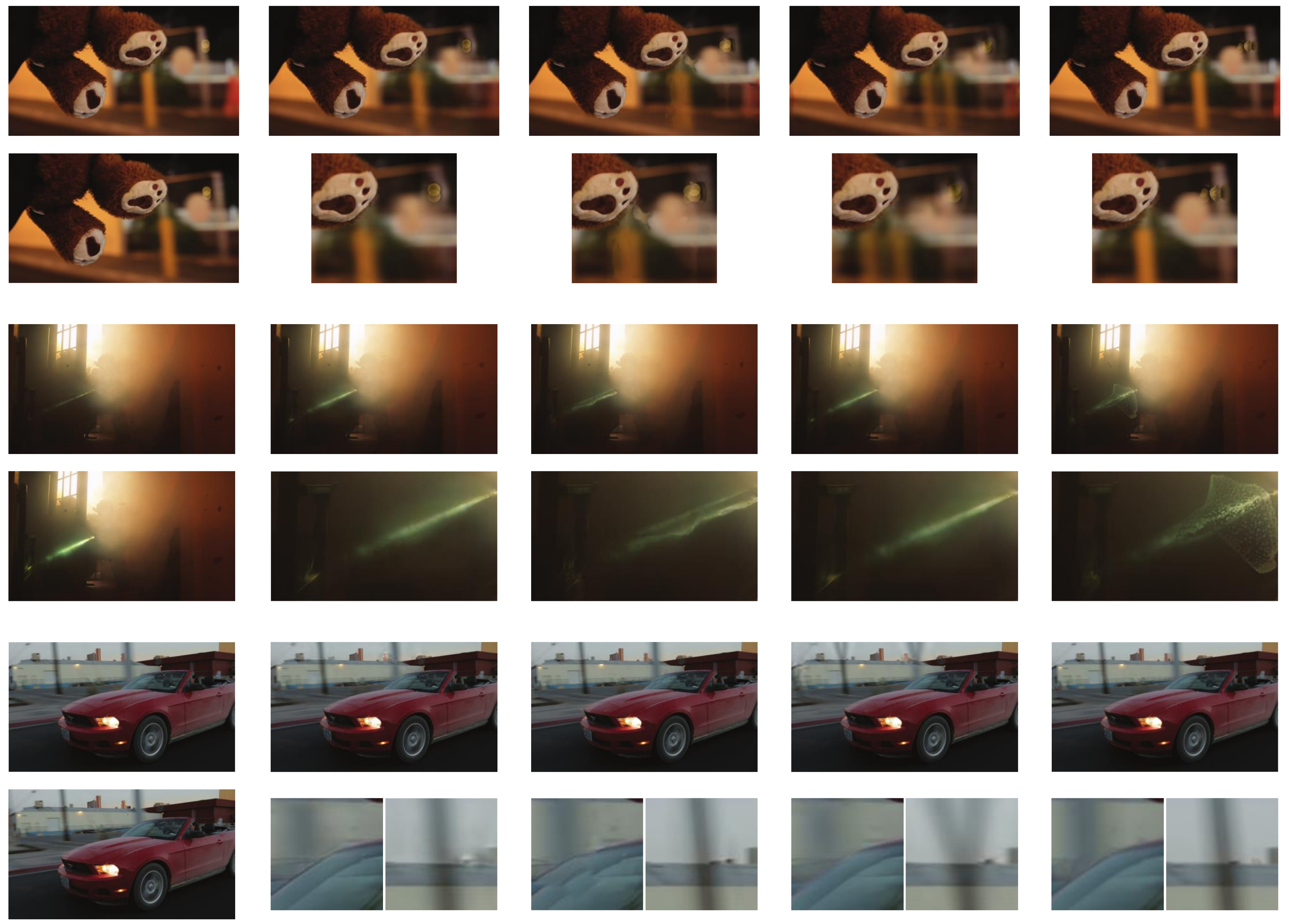
» Phase loss $\mathcal{L}_{\text{phase}} = \sum_{\omega,\theta} ||\Delta\phi_{\omega,\theta}||_1$



W/o $\mathcal{L}_{\text{phase}}$ Ours Ours

Measures phase difference $\Delta\phi_{\omega,\theta}$ to the ground truth pyramid decomposition

Results



Input Ours SepConv [2] Phase [1] MDP-Flow2 [4]

Conclusions

PhaseNet is well suited for interpolating challenging scenes with motion blur and brightness changes at the expense of preserving some fine detail.

References

[1] S. Meyer et al. Phase-based frame interpolation for video. CVPR, 2015.
[2] S. Niklaus et al. Video frame interpolation via adaptive separable convolution. ICCV, 2017.
[3] J. Portilla and E. P. Simoncelli. A parametric texture model based on joint statistics of complex wavelet coefficients. IJCV, 2000.
[4] L. Xu et al. Motion detail preserving optical flow estimation. PAMI, 2012.