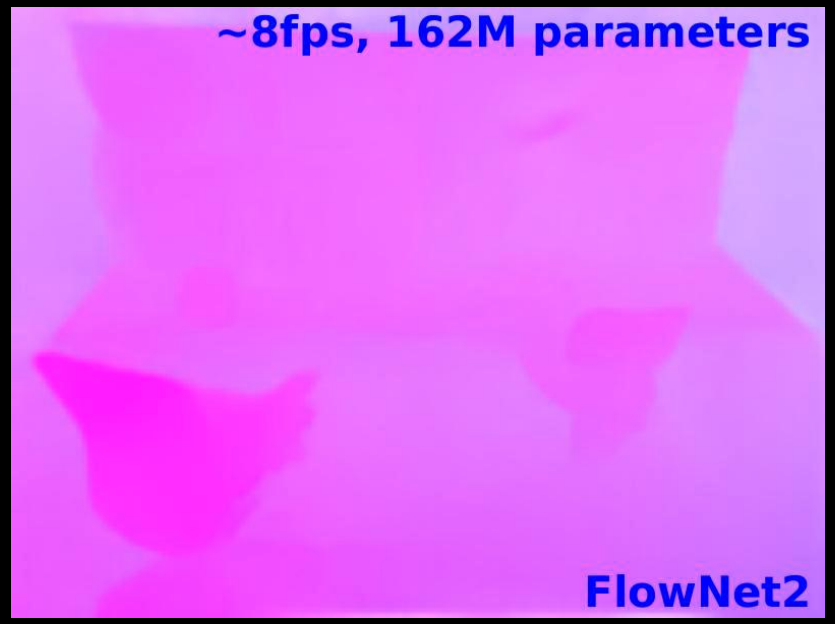
LiteFlowNet: A Lightweight Convolutional Neural Network for Optical Flow Estimation

T.-W. Hui, X. Tang, C. C. Loy
CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong

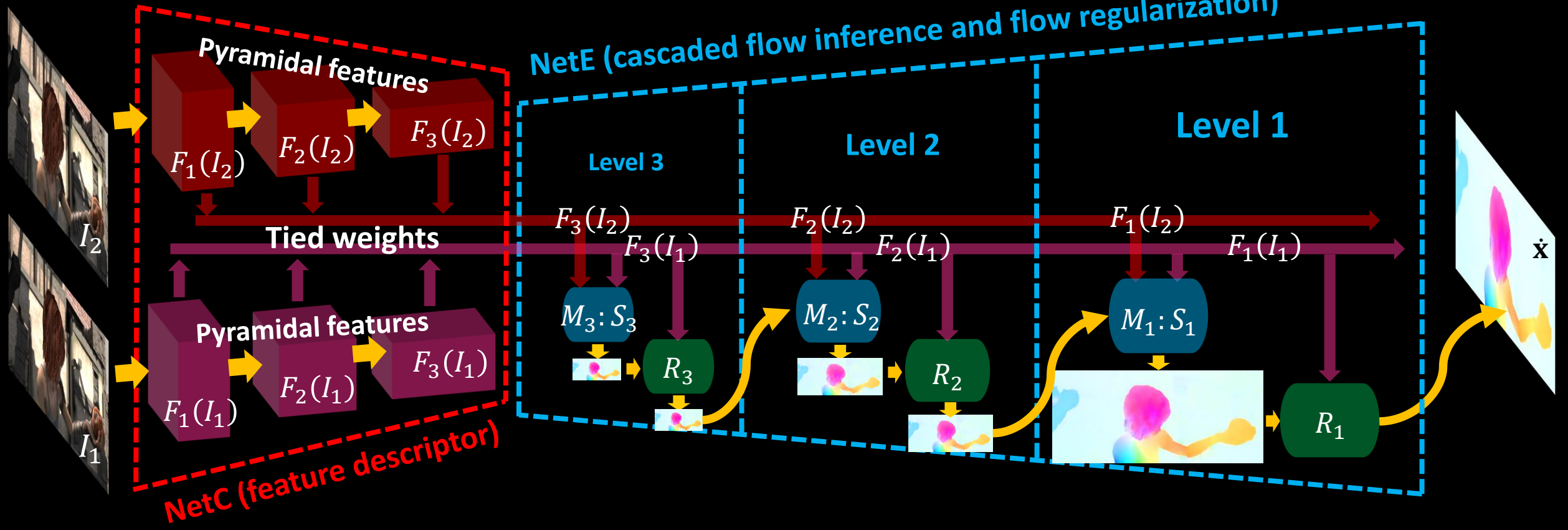
FlowNet2 (CVPR17)

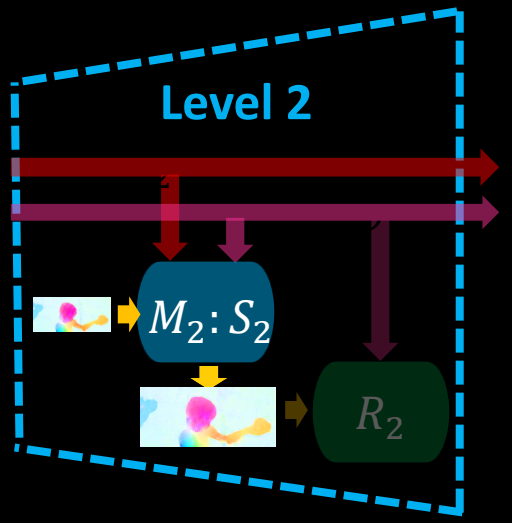
LiteFlowNet (CVPR18)



FlowNet2 (CVPR17)	LiteFlowNet (CVPR18)
U-Net architecture	Specialized architecture: - Data fidelity & regularization as variational methods
Image warping per cascade	Feature warping per pyramid level
Feature matching per cascade	Cascaded flow inference per level
	Flow regularization per level: - Feature-driven local convolution
Large network cascade (**~160M**)	Lightweight (**~5M**), faster (**1.36x**), and better performance on real data

LiteFlowNet





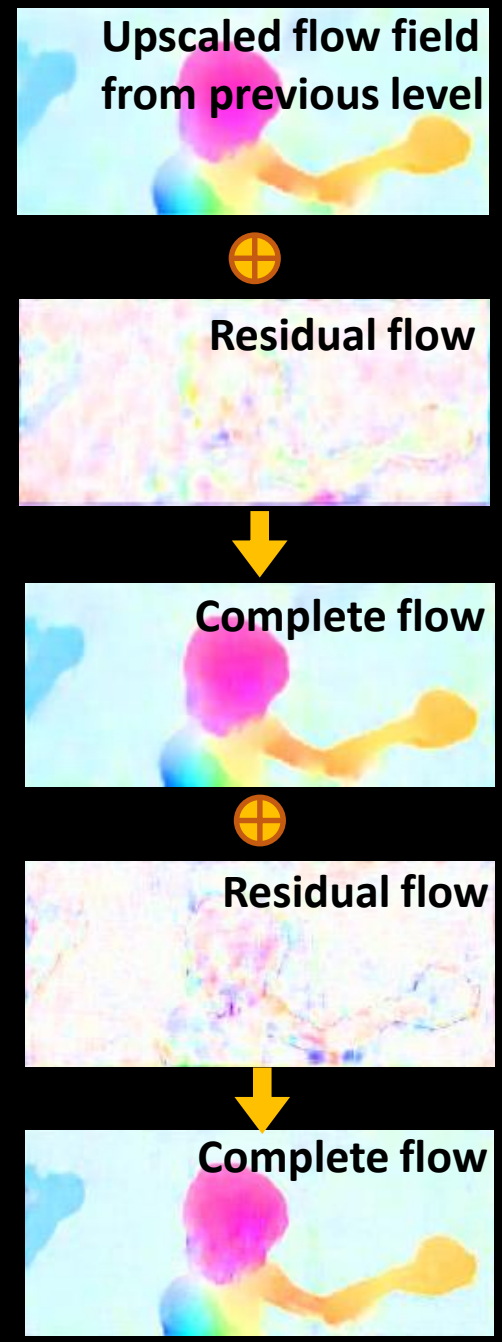
Cascaded Flow Inference

Stage 1:
Descriptor matching unit *M*

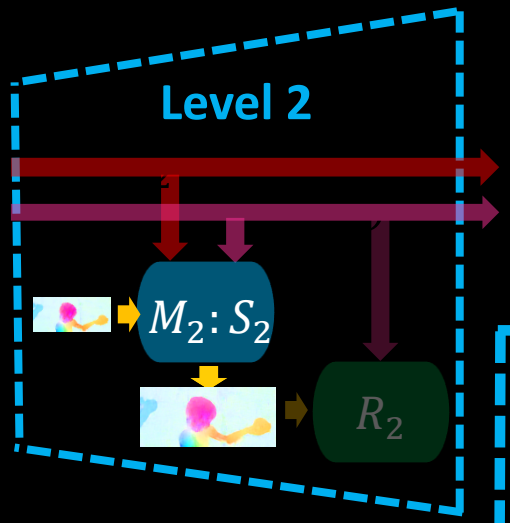
Generate a coarse flow estimate by computing correlation of high-level feature vectors

Stage 2:
Sub-pixel refinement unit *S*

Aim to refine the flow field to sub-pixel accuracy



Cascaded Flow Inference : Stage 1



Descriptor matching unit *M*

Upscaled flow field from previous level

Residual flow

Complete flow

Cost volume C

Upscaled in resolution and magnitude

f-warp

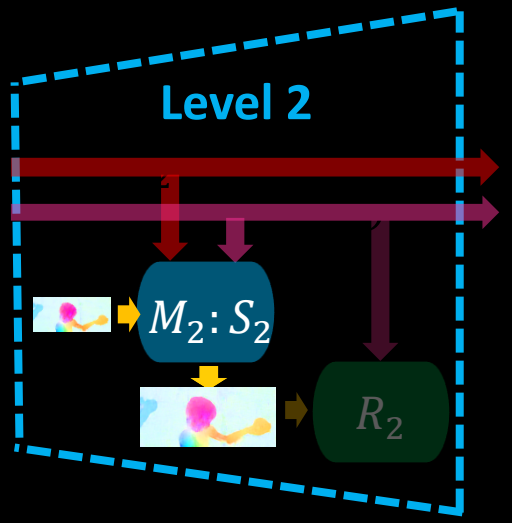
$\widetilde{F_2}(I_1)$

Pixel-to-pixel correlation

Flow field from previous level

$F_2(I_1)$

$F_1(I_2)$



Cascaded Flow Inference : Stage 2

Flow field from descriptor matching unit *M*

$F_2(I_1)$

$F_1(I_2)$

Sub-pixel refinement unit *S*

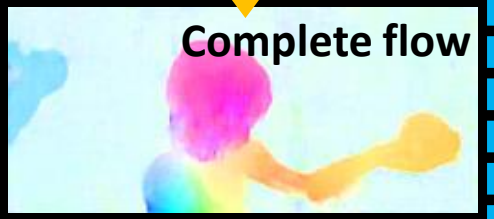
Flow field from descriptor matching unit *M*

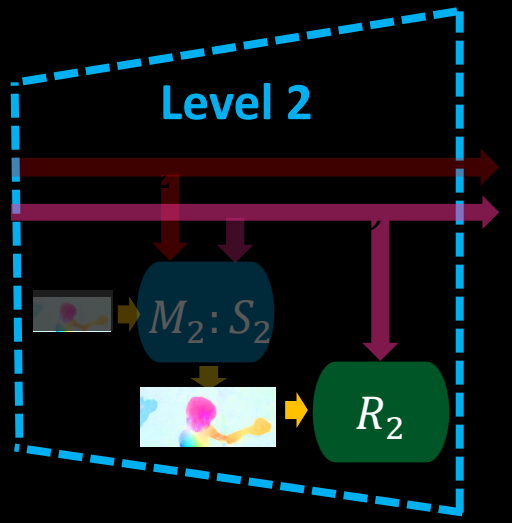
f-warp

$\widetilde{F}_2(I_1)$

Residual flow

Complete flow





Flow Regularization

Using data term alone, vague flow boundaries and undesired artifacts commonly exist in flow field

First in the literature: Perform flow field regularization by a feature driven local convolution (f-lcon)

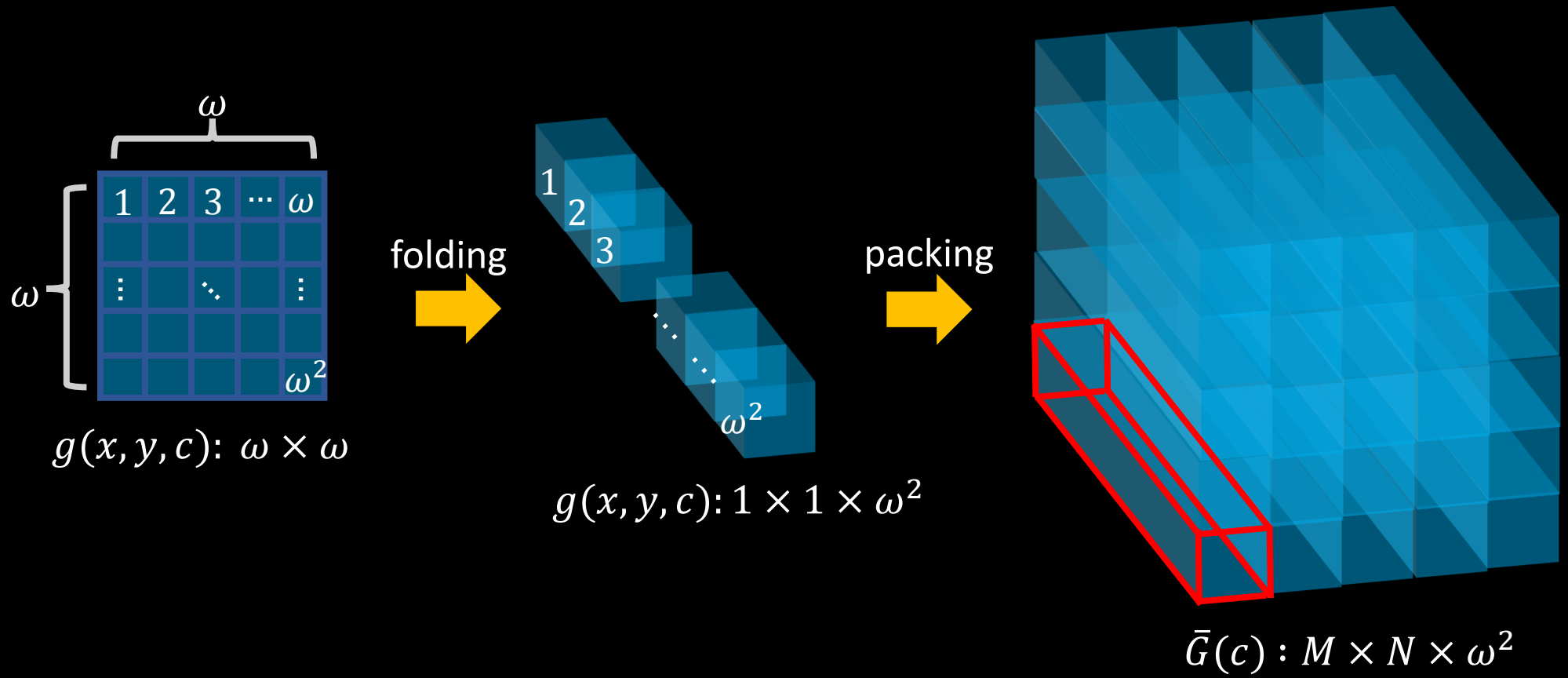
$$f_g(x, y, c) = g(x, y, c) * f(x, y, c)$$

filter flow patch

Each local filter g is *adaptive* and *both image- and flow-aware*

- *Pyramidal features*
- *Flow estimate*
- *Occlusion probability map*

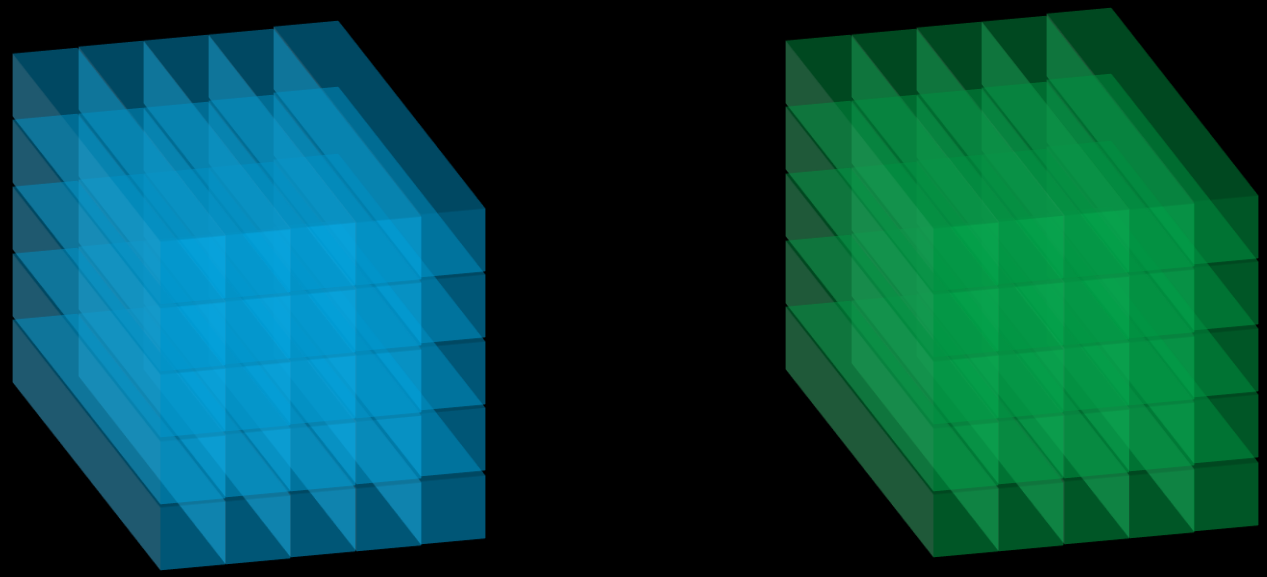
Efficient feature-driven local convolution



Folding and packaging each local filter g by GPU friendly `im2col`

Perform the same on the flow patch

Efficient feature-driven local convolution



3D tensor of a local filter

3D tensor of a flow patch

$$F_g(c) = \bar{G}(c) \odot \bar{F}(c)$$

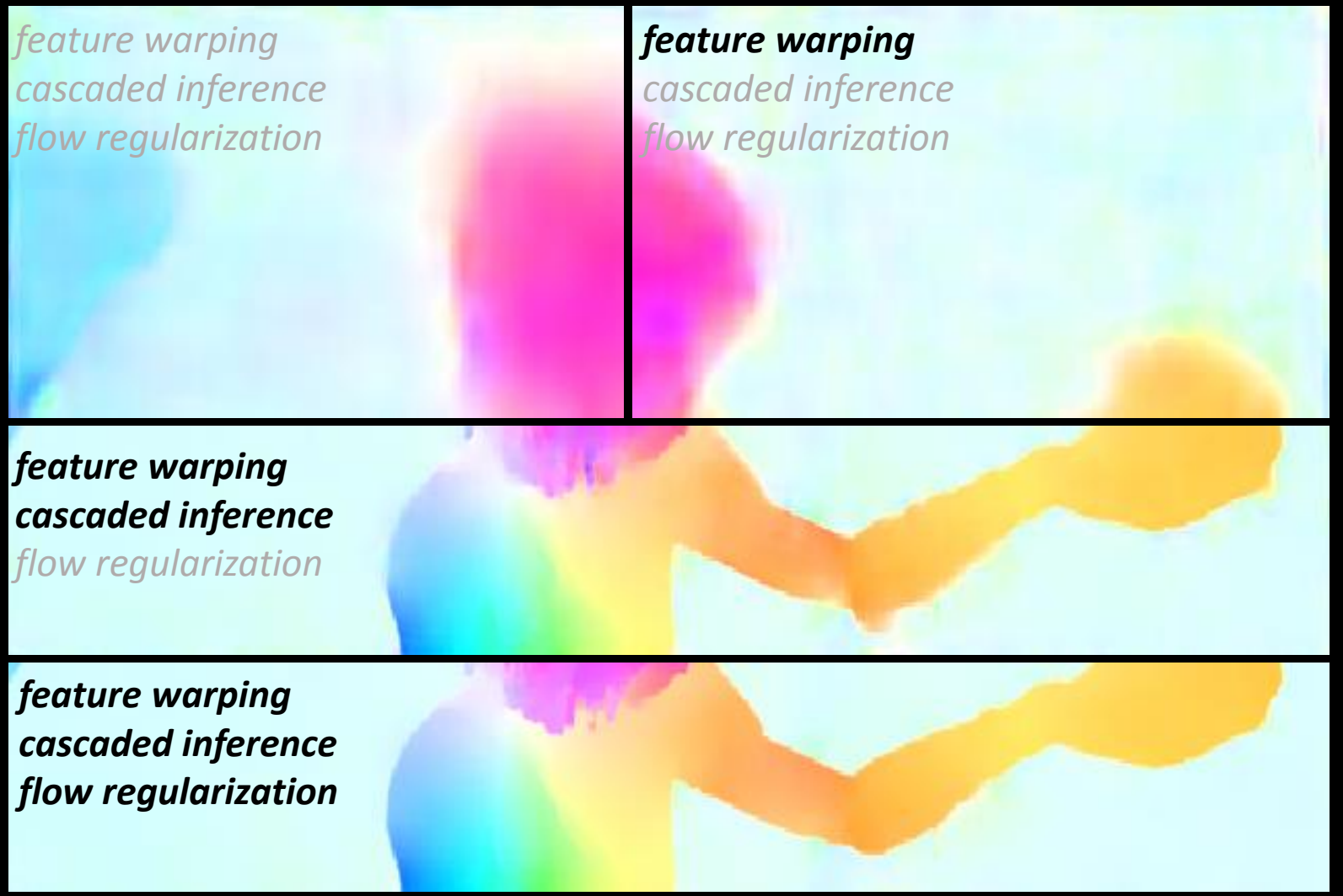
Local convolution can be computed efficiently by a tensor dot product

Effectiveness of different components

Feature warping improves flow sharpness

Cascaded inference generates sharper and more detailed flows

Regularization prevents flow bleeding and vague flow boundaries



Results on Sintel

~30 times smaller, 1.36 times faster

Project page (demo video, code, and more)

http://mmlab.ie.cuhk.edu.hk/projects/LiteFlowNet/