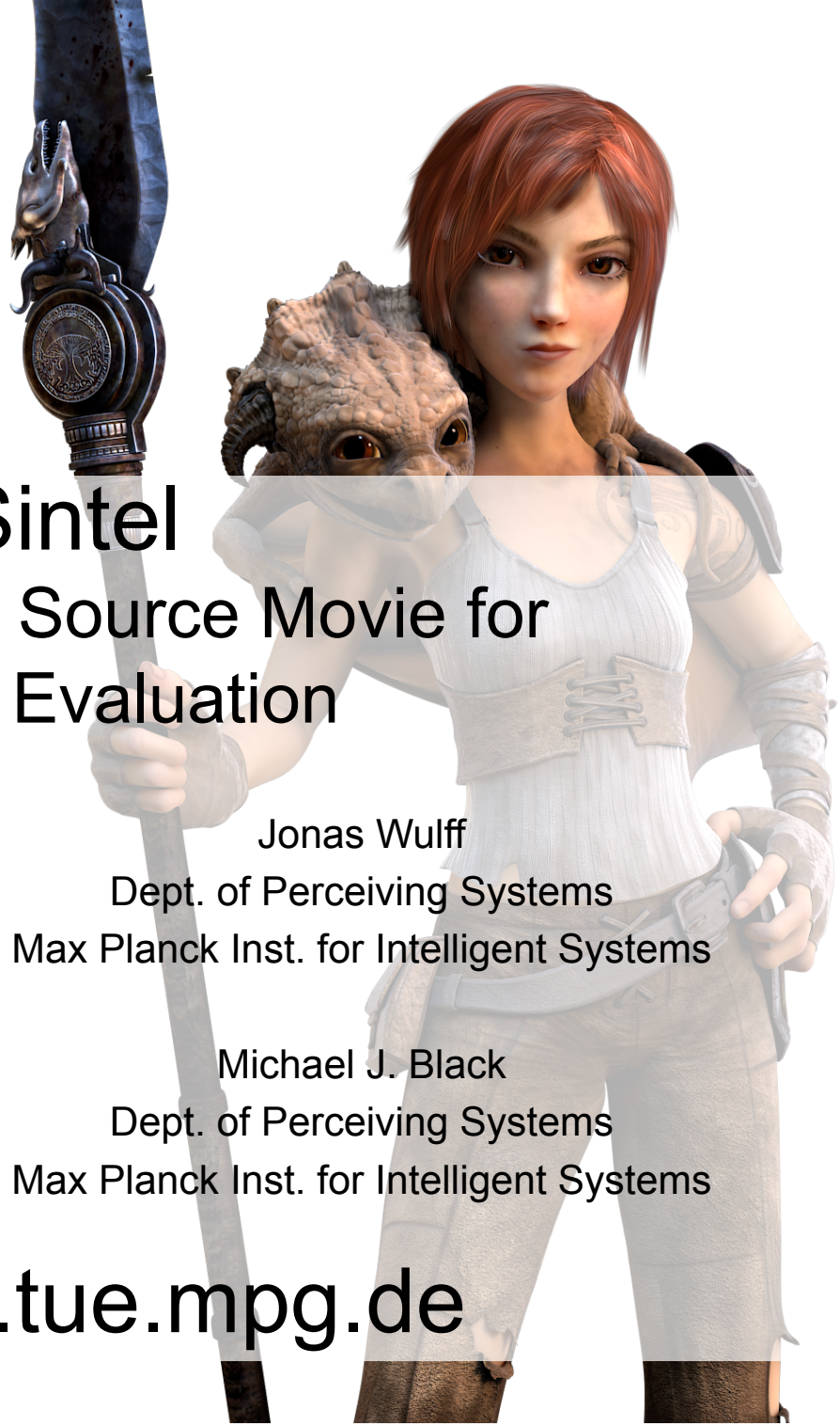


MPI-Sintel

A Naturalistic Open Source Movie for Optical Flow Evaluation

Daniel J. Butler
Dept. of Comp. Science & Engineering
University of Washington

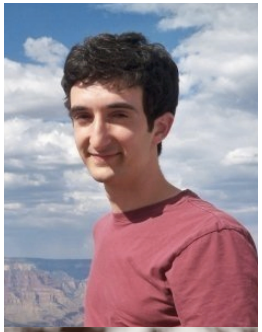
Jonas Wulff
Dept. of Perceiving Systems
Max Planck Inst. for Intelligent Systems

Prof. Garrett B. Stanley
Dept. of Biomedical Engineering
Georgia Tech

Michael J. Black
Dept. of Perceiving Systems
Max Planck Inst. for Intelligent Systems

http://sintel.is.tue.mpg.de

Collaborators



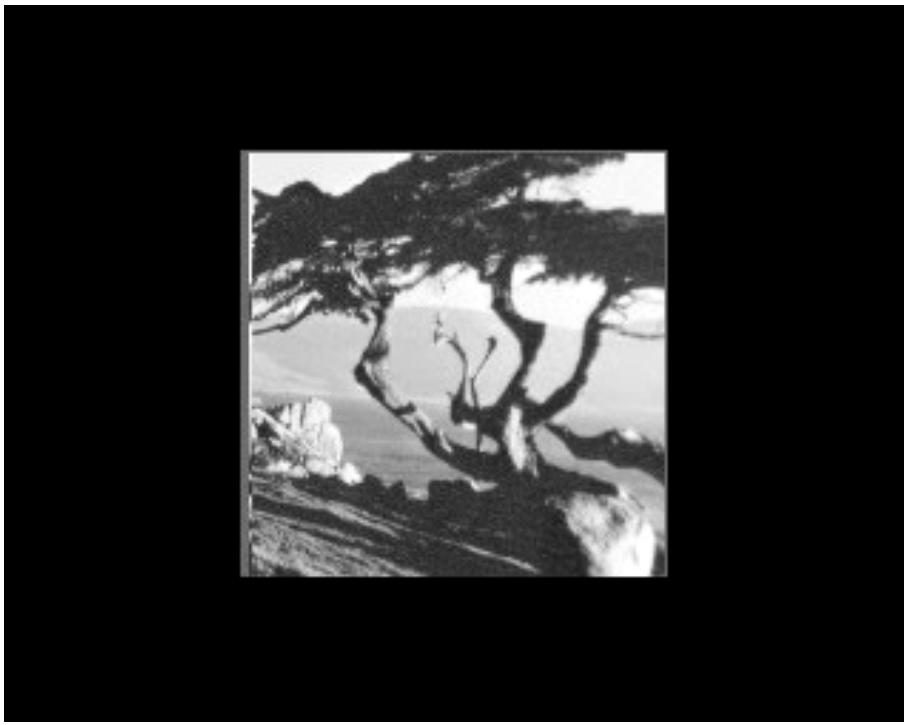
Dan Butler
University of Washington

Jonas Wulff
Max Planck Institute for Intelligent Systems

Garrett Stanley
Georgia Tech

Barron et al. Sequences (1994)

~10-60 frames/seq. 12 seqs ~300 frames total ~30 flow fields total

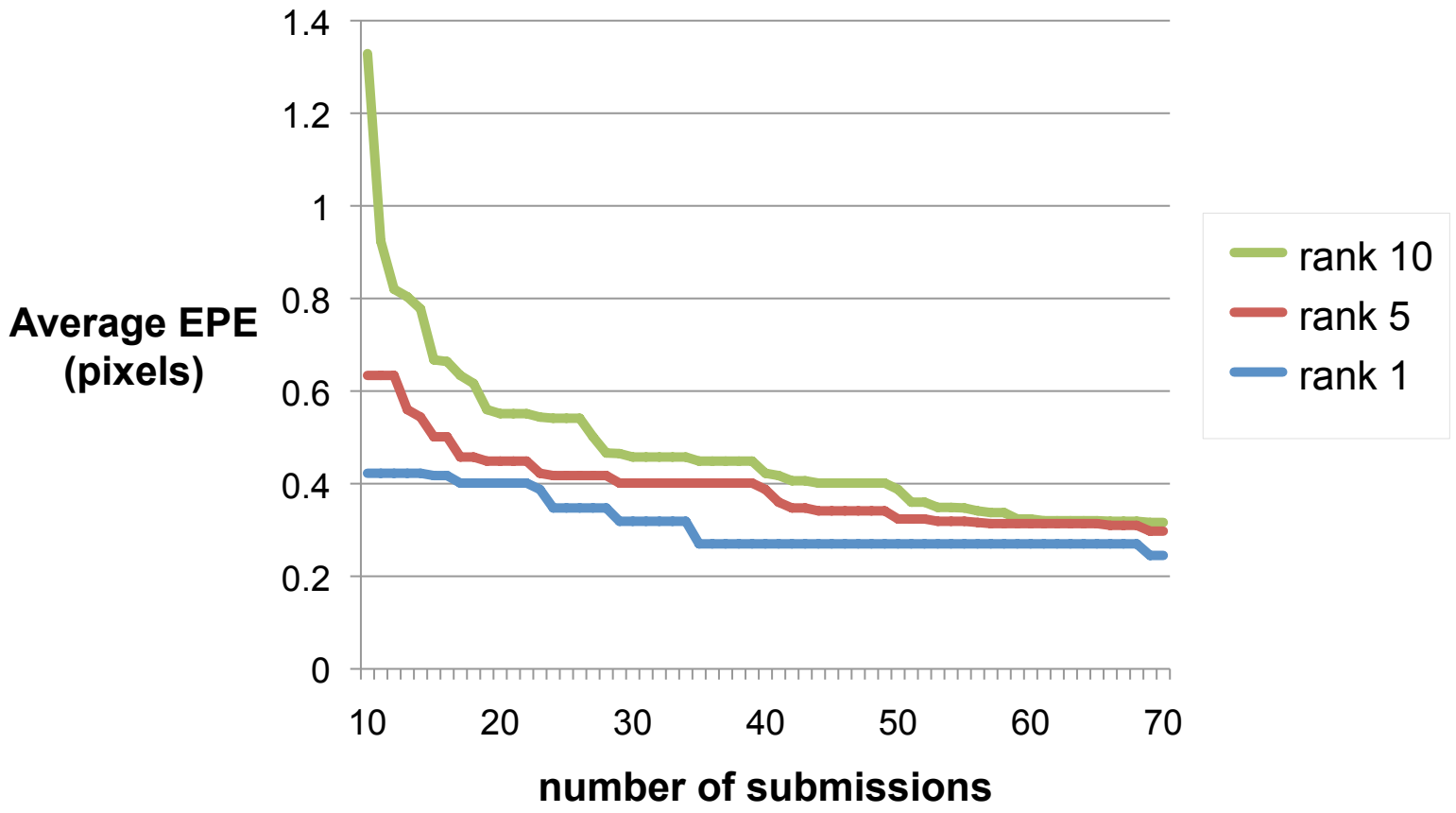


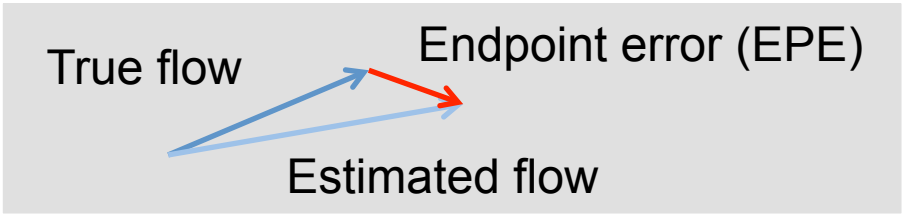
Barron, J. L., Fleet, D. J. and Beauchemin, S. S.,
Performance of optical flow techniques,
International Journal of Computer Vision, 12(1):43-77, 1994.

Middlebury Flow Dataset (2007)

Baker et al., *IJCV* 2011.

Error on Middlebury over time





We need a challenging new dataset

New datasets:

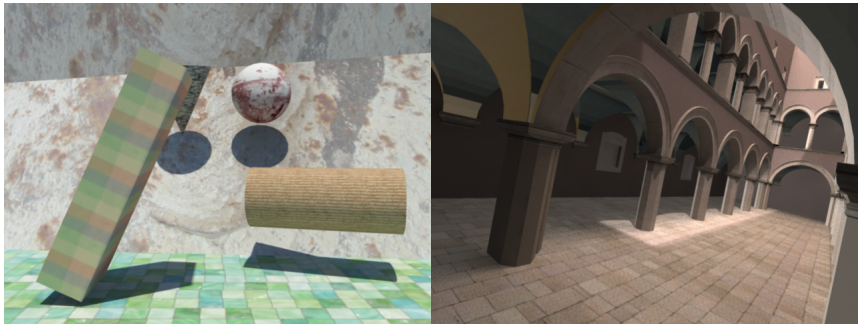
KITTI Vision Benchmark

Geiger et al., *CVPR* 2012.

HCI Robust Vision Challenge

Meister et al., *Optical Engineering,* 2012.

UCL Ground Truth Optical Flow Dataset



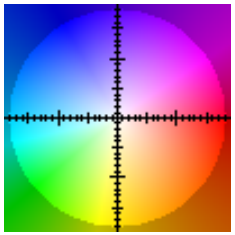
Mac Aodha et al., *PAMI,* 2012.

Human-Assisted Motion Annotation

Liu et al., *CVPR* 2008.

Introducing: MPI-Sintel

35 sequences, 1628 frames, 1593 flow fields



Sintel: a Blender Open Movie

Created in order to test and promote the Blender animation suite

Free and **Open**:

- All graphics data released under CC license
- Rendering software open source

Problem

- Can an animated movie teach us about optical flow in the real world?
- Will results generalize?
- Is it realistic enough?

- Solution: compare Sintel statistics with those of real scenes.
- How?

Idea: Lookalikes

- Lookalikes are “real” scenes that are semantically similar to Sintel scenes.

lookalike

Sintel

Idea: Lookalikes

- Lookalikes are “real” scenes that are semantically similar to Sintel scenes.

lookalike

Sintel

Idea: Lookalikes

- Lookalikes are “real” scenes that are semantically similar to Sintel scenes.

lookalike

Sintel

Idea: Lookalikes

- Lookalikes are “real” scenes that are semantically similar to Sintel scenes.

lookalike

Sintel

Idea: Lookalikes

- Lookalikes are “real” scenes that are semantically similar to Sintel scenes.

lookalike

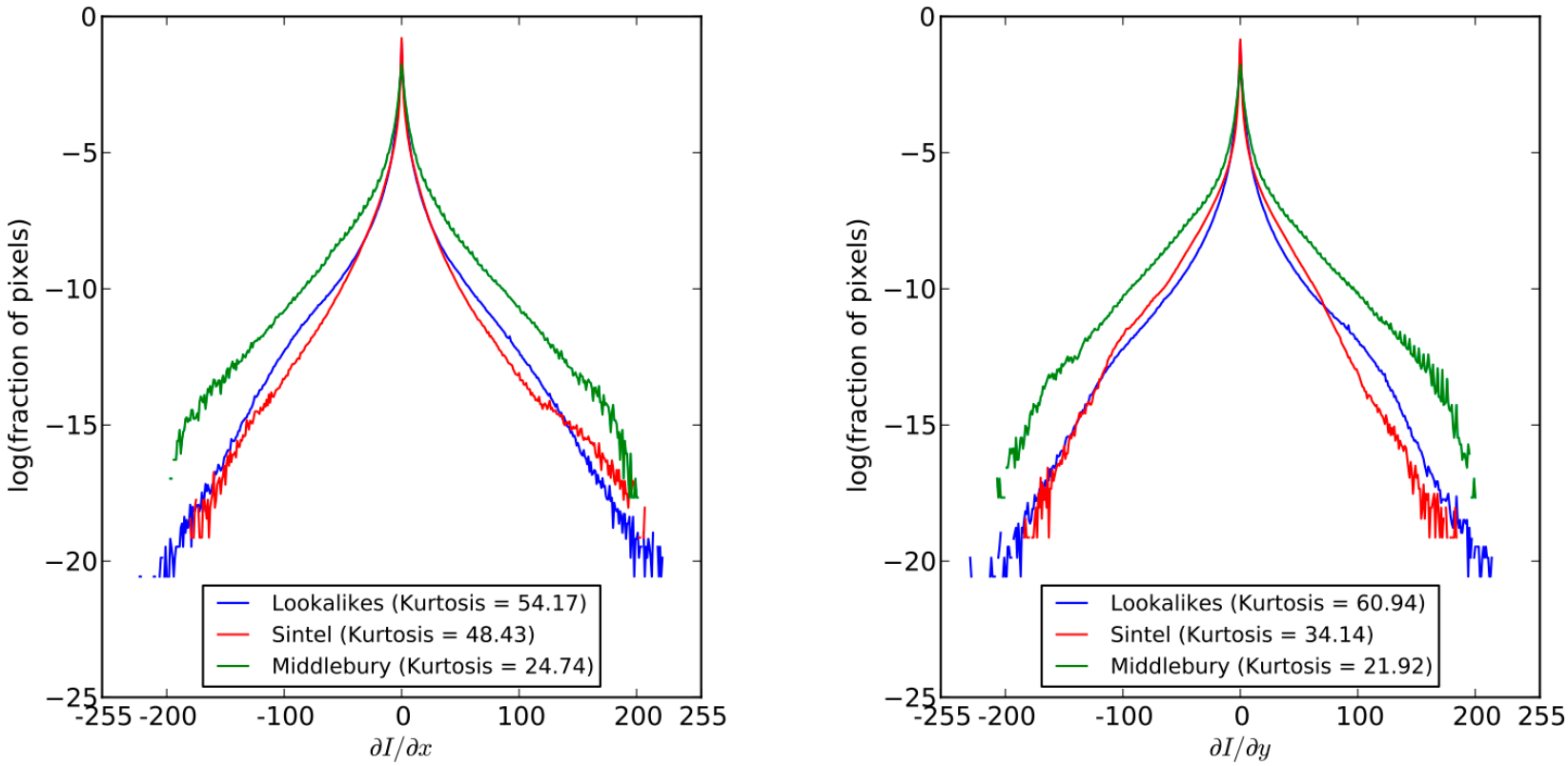
Sintel

Lookalikes

Image statistics:

- Luminance histograms
- Power spectra
- Derivative histograms

Image derivative log-histograms



Lookalikes
Sintel
Middlebury

What about motion statistics?

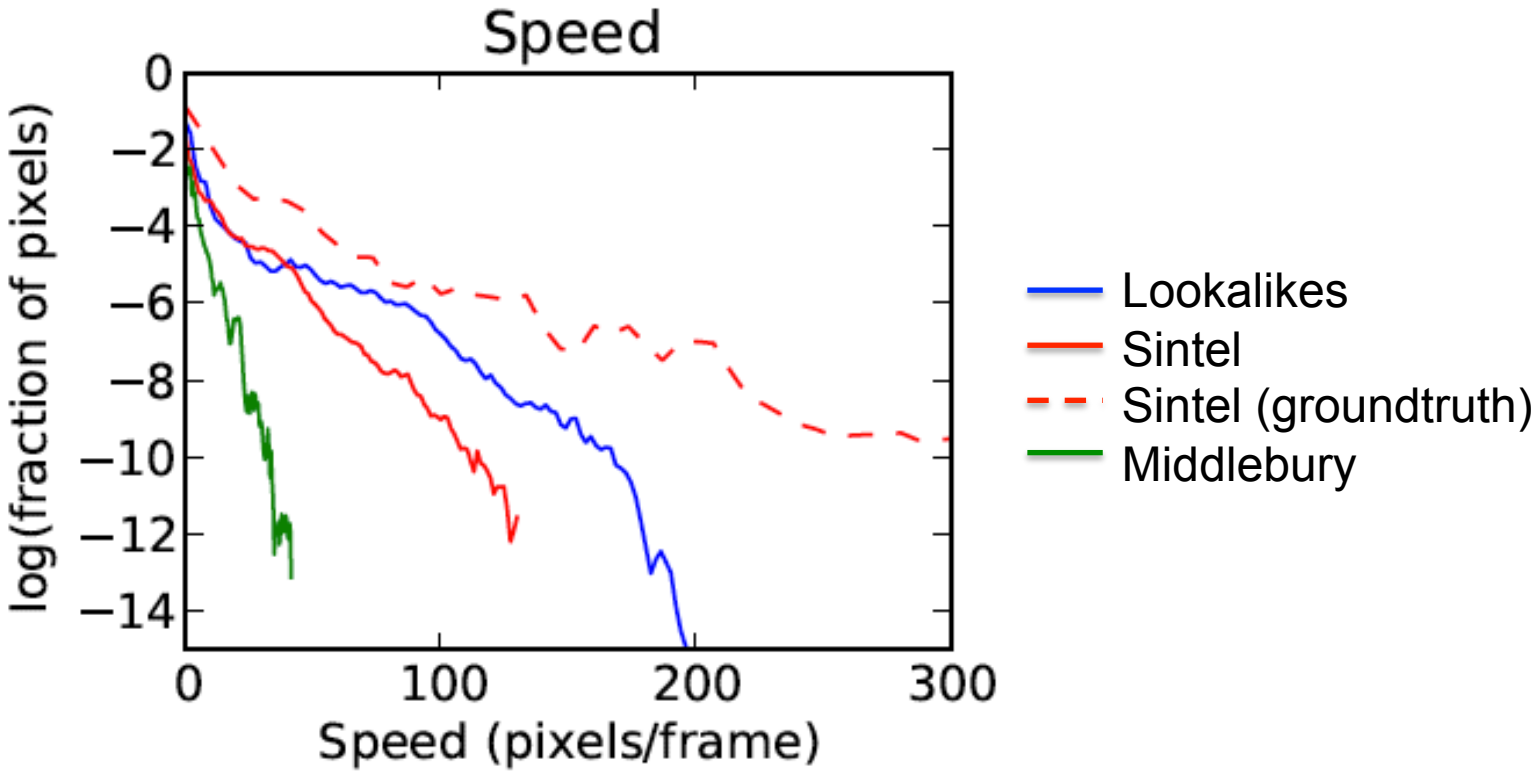
- Image statistics are only half the problem
- Do Sintel *motions* resemble *natural motions*?
  - Harder since we do not have ground truth flow for the lookalike sequences
- Approach: compare statistics of *estimated flow* on Sintel and lookalikes.

Flow statistics

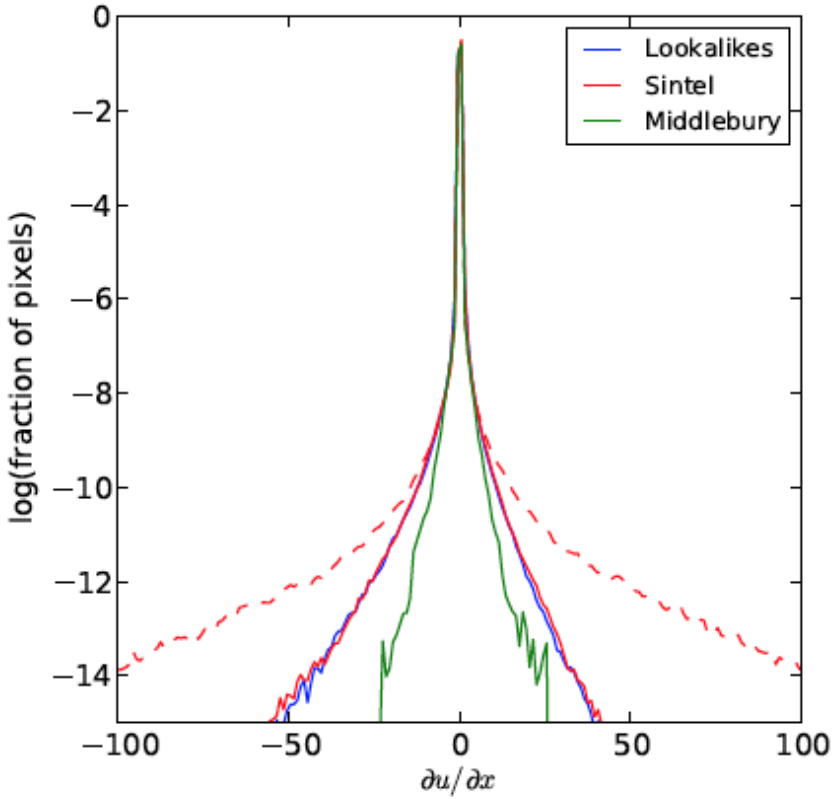
(*estimated flow*):

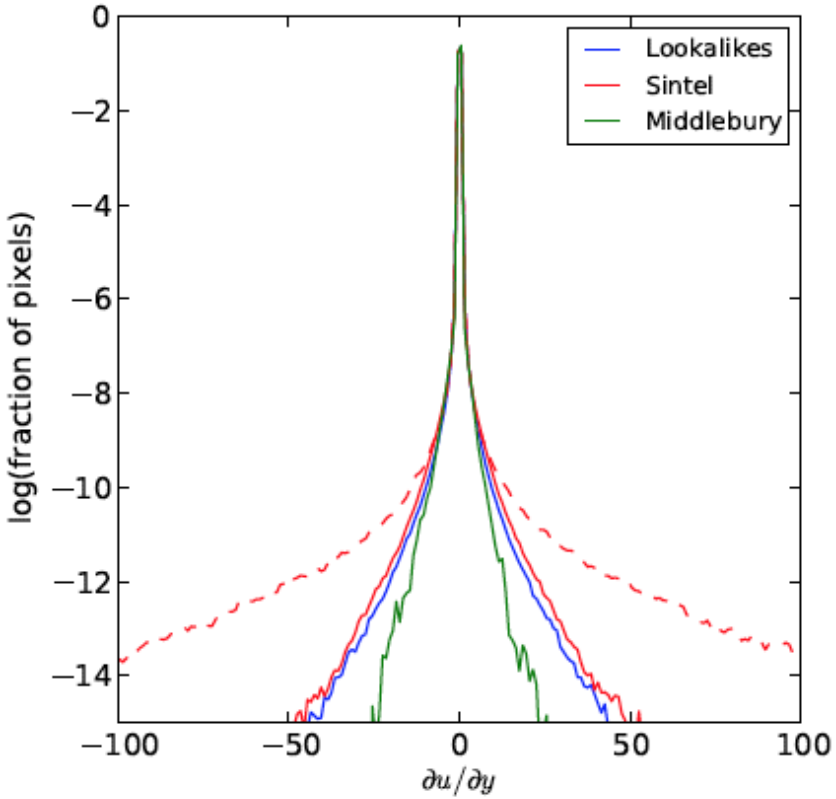
- Histograms of horiz. and vertical components
- Speed histograms
- Derivative histograms

Speed histograms



Flow derivatives





Lookalikes
Sintel
Sintel (groundtruth)
Middlebury

Realism story isn't over

- Obviously Sintel is not photorealistic
- However, it does pass some sanity checks

Future work:

1. Use photo-realistic graphics data
2. General problem of evaluating realism

Meister and Kondermann, *Conference on Electronic Media Technology (CEMT)*, 2011.

CG data is not just “good enough”…

… it has major **advantages**

Render passes

Motion boundaries

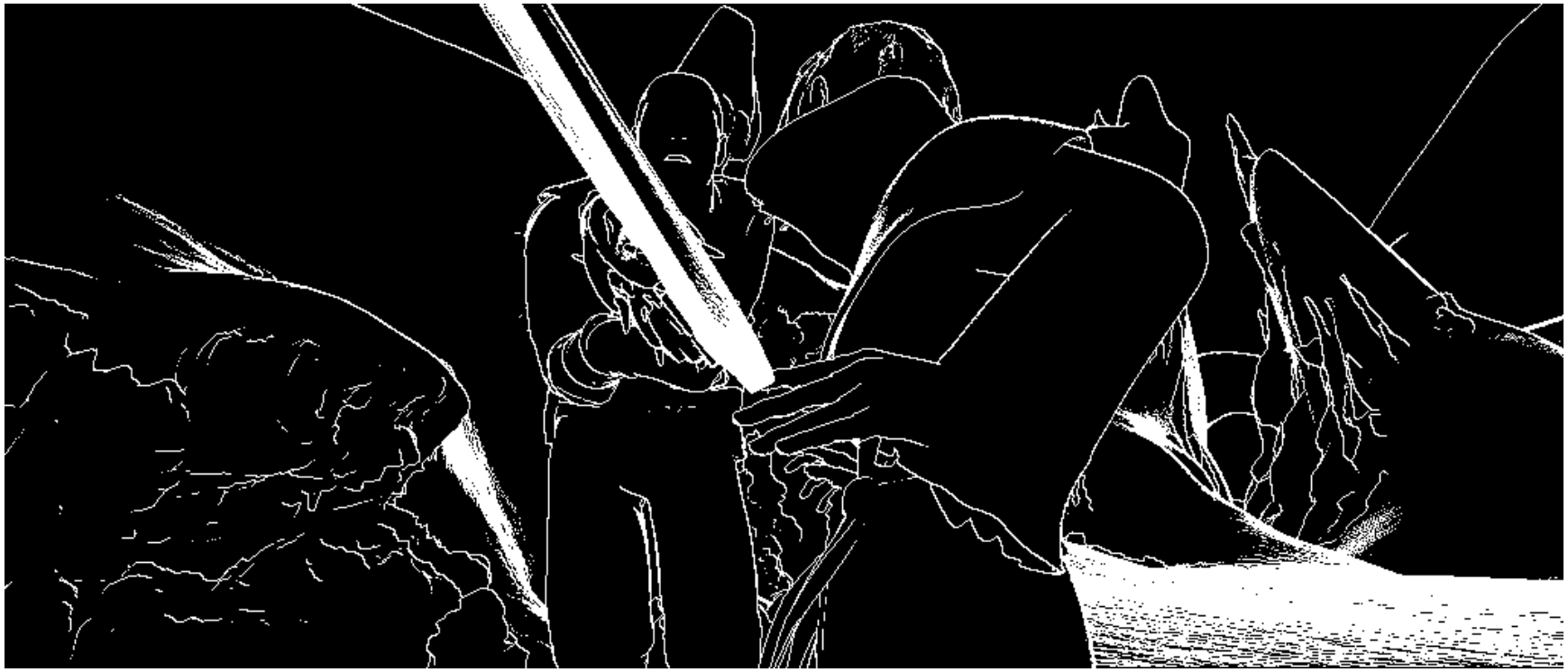
- Many definitions.
- Our goal: define places where current flow methods fail.

(Object boundary U Material boundary U thresholded depth gradient) ∩ thresholded flow gradient

Object boundaries

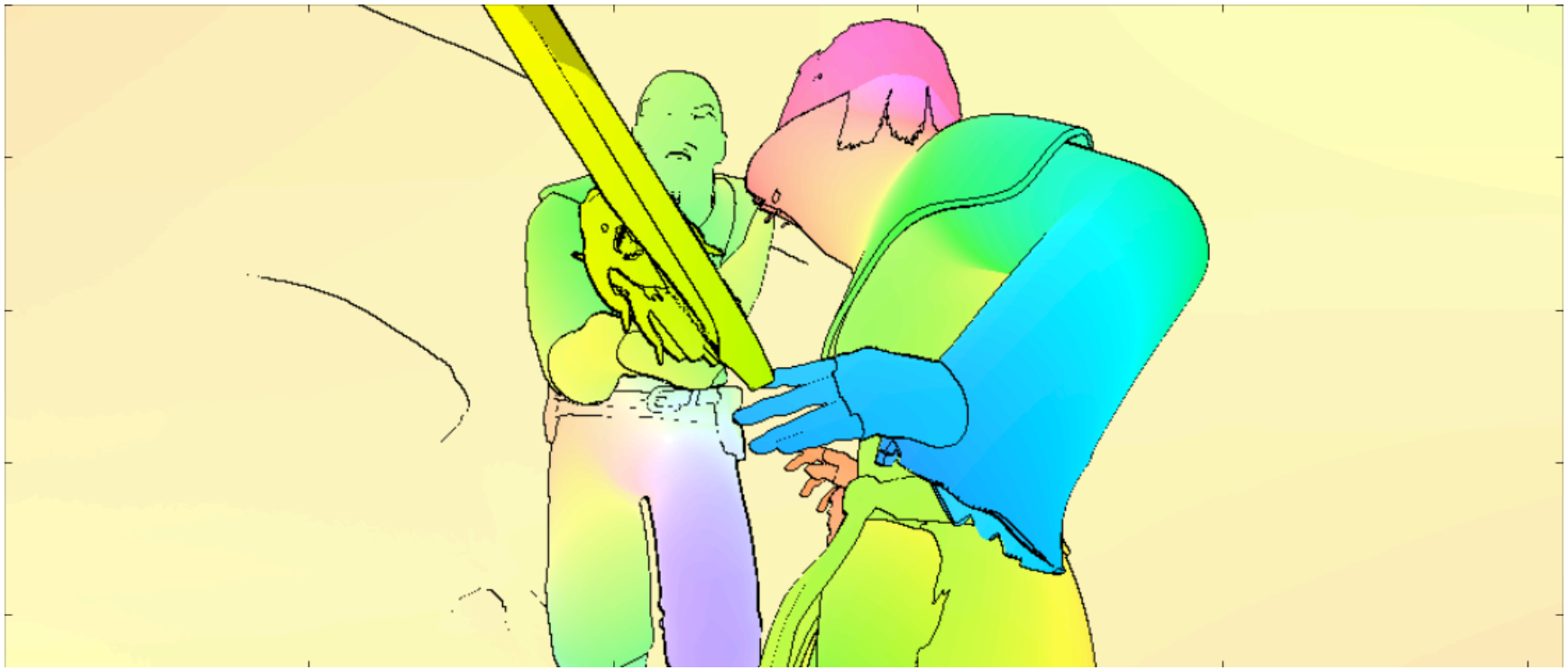
Material boundaries

Thresholded depth gradient

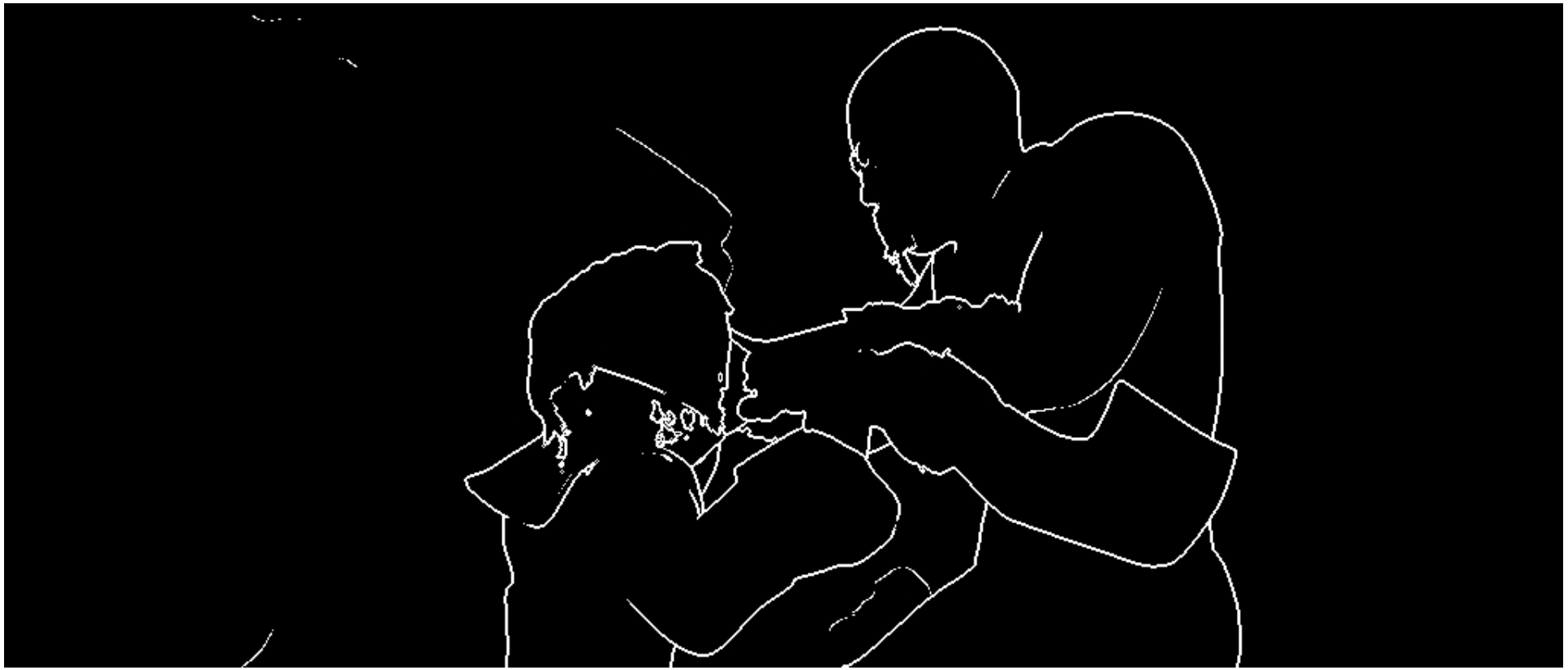


Thresholded flow gradient

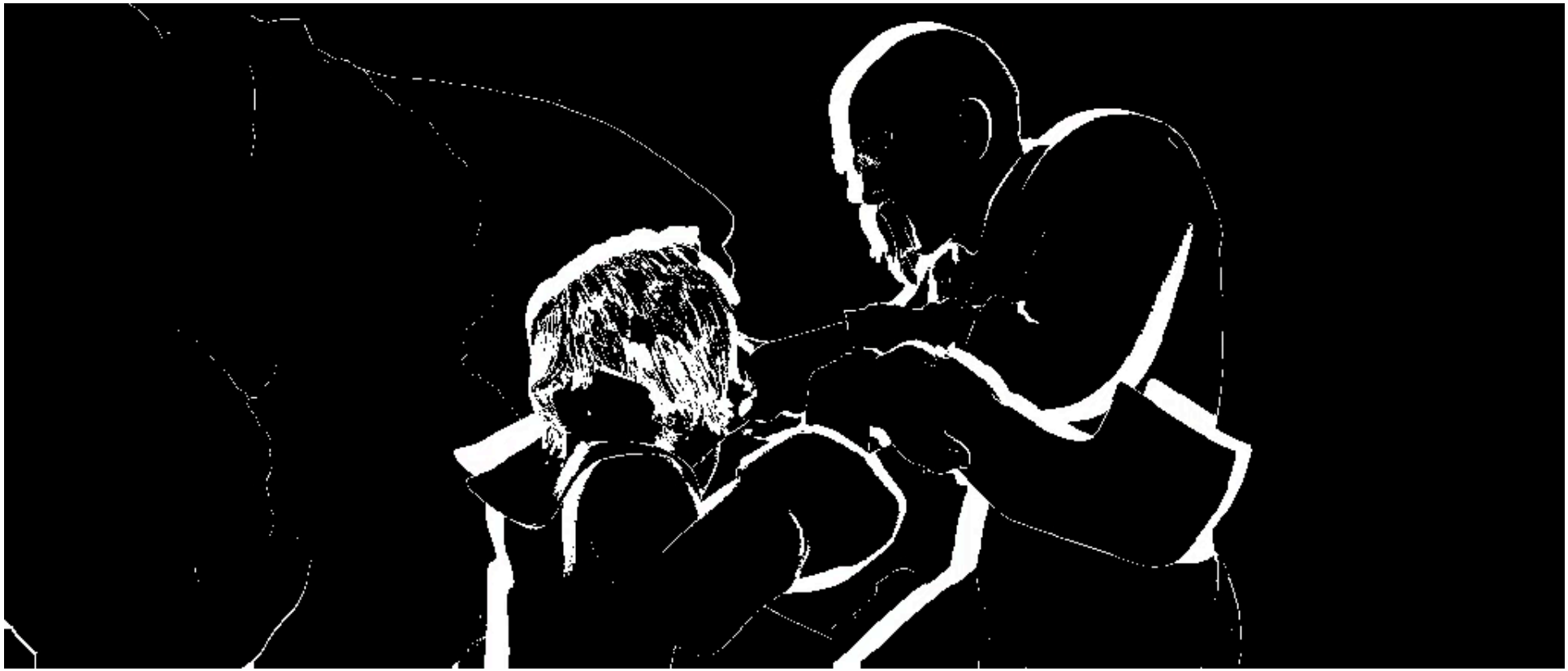
Final boundaries on flow



high flow gradient ∩ object boundaries



Unmatched regions



Beyond flow

Beyond flow

Material segmentation

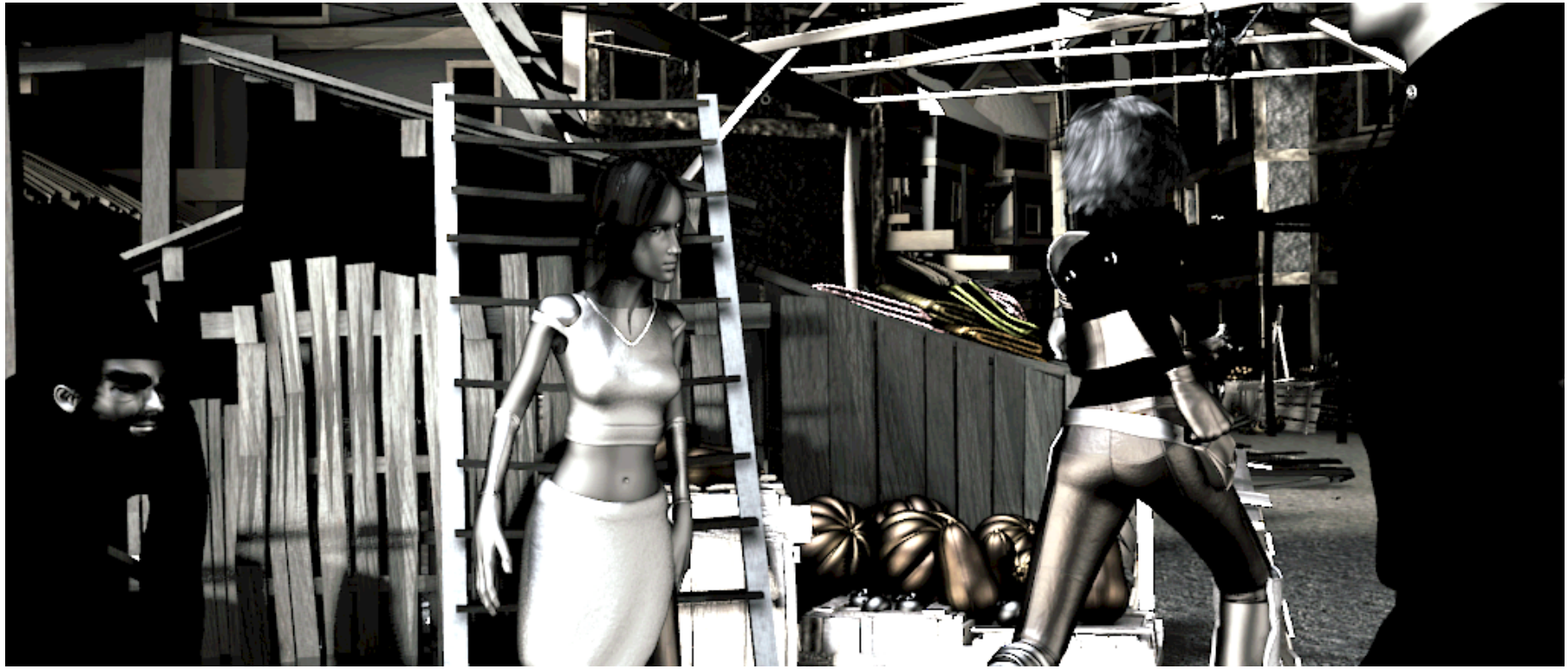
Beyond flow

Object segmentation

Beyond flow

Shading

Beyond flow



Specularities

Beyond flow

Depth

Problem: Cheating

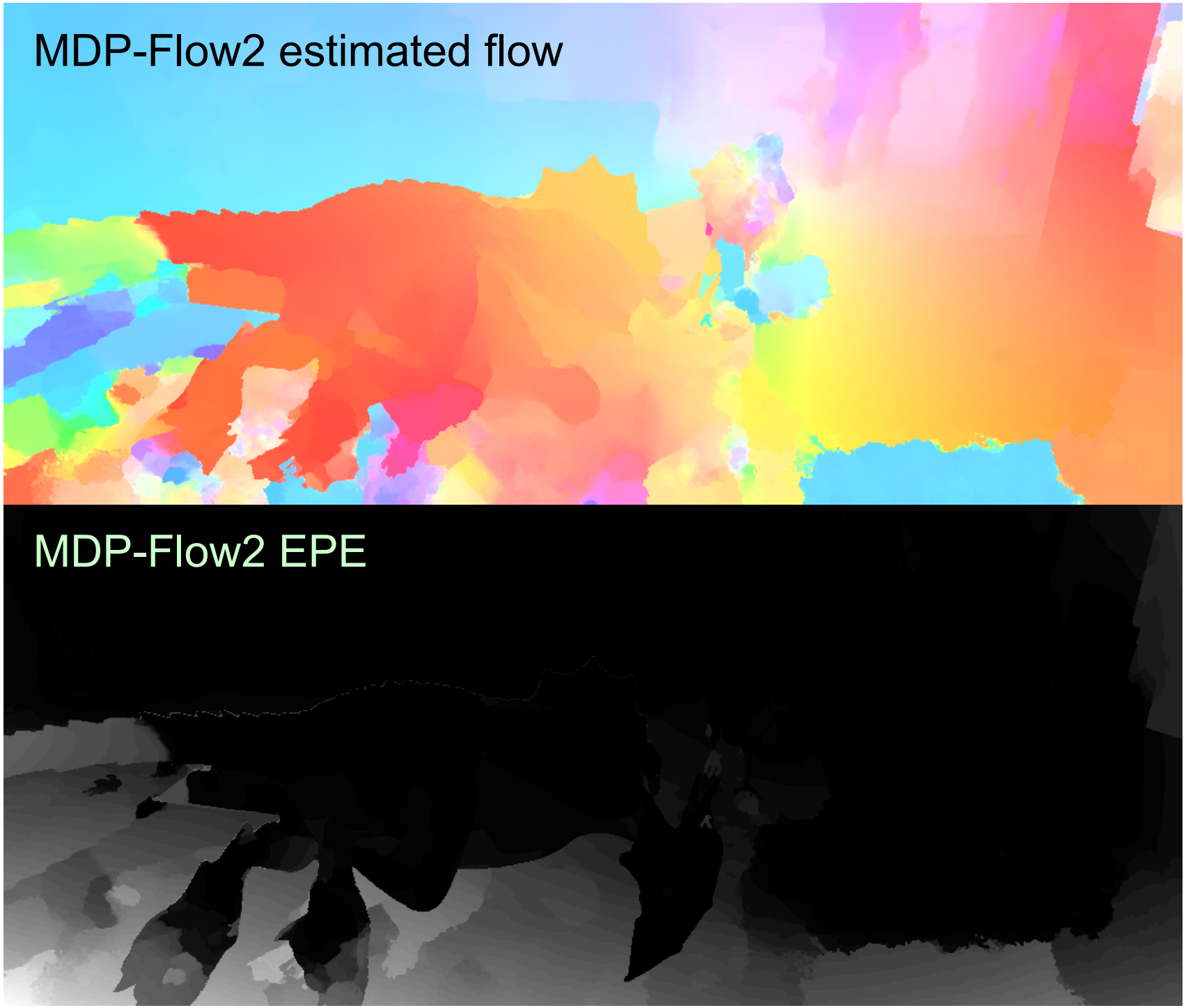
- Sintel is public domain
- People could compute the flow of the test sequences from the graphics data
- As a fraud check, we generate two perturbed sequences
  - One where the camera motion is different
  - One where the object motions are changed

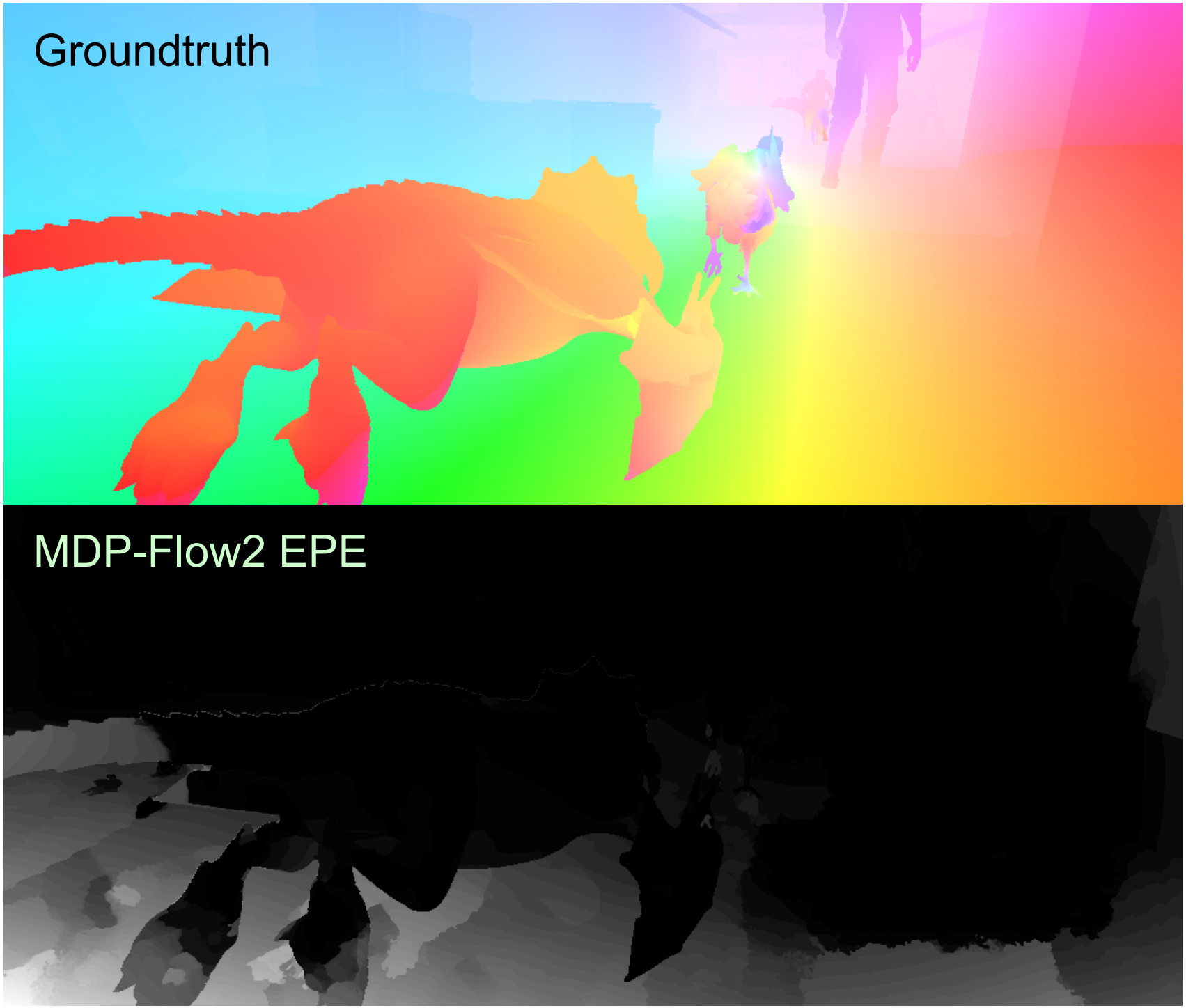
Evaluation

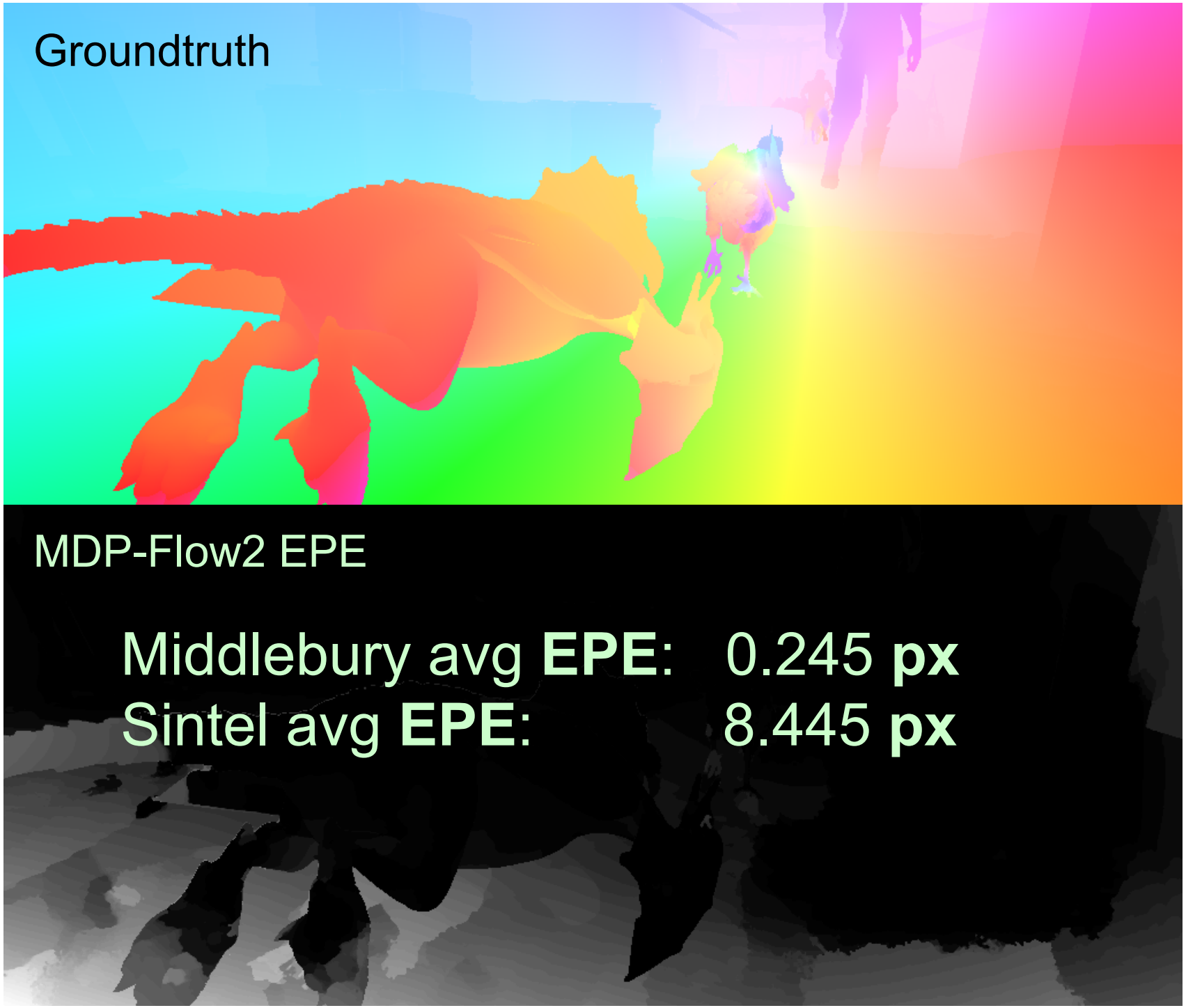
http://sintel.is.tue.mpg.de

Final | Clean

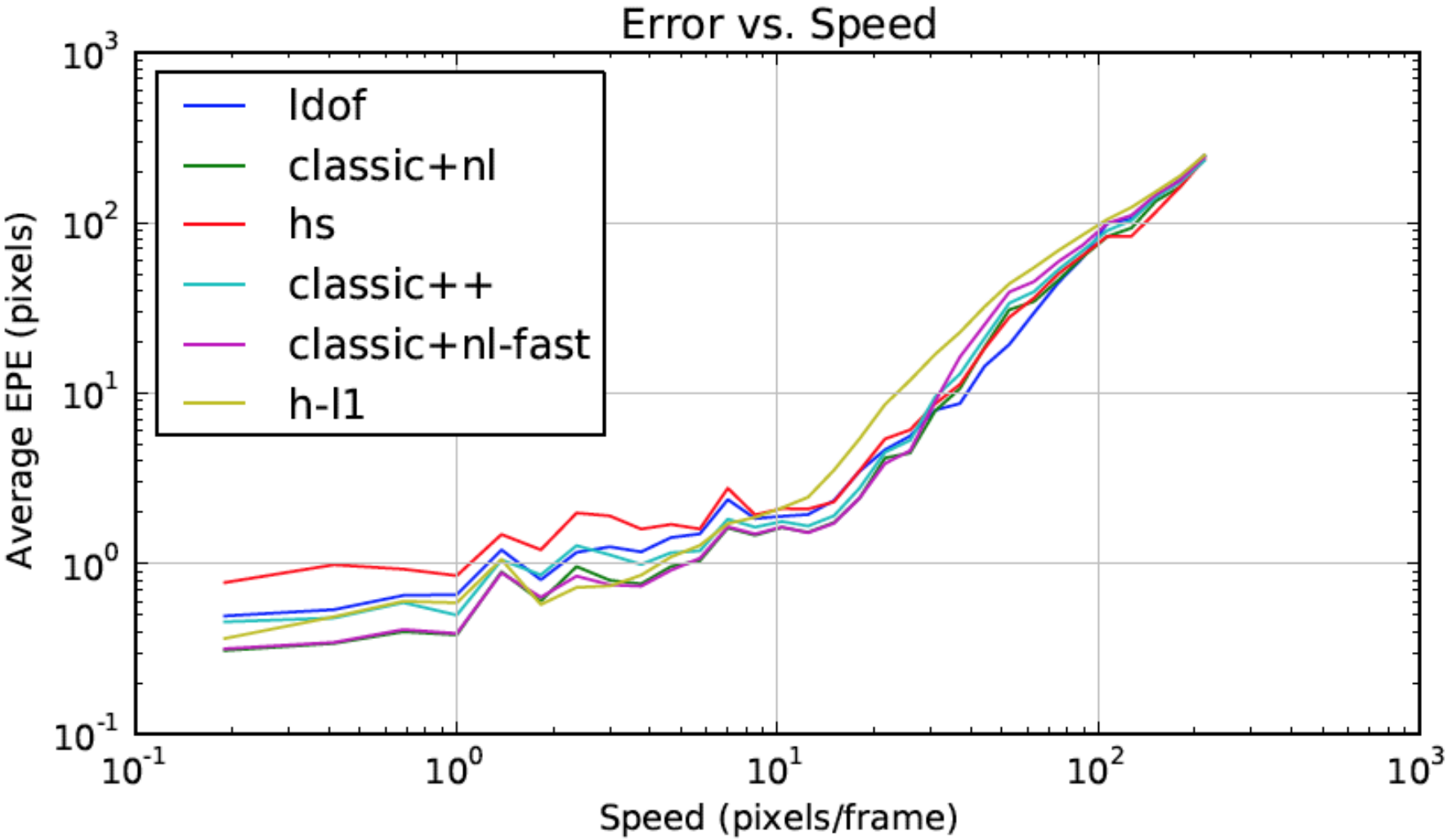
	EPE all	EPE matched	EPE unmatched	d0-10	d10-60	d60-140	s0-10	s10-40	s40+
GroundTruth [1]	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MDP-Flow2 [2]	8.445	4.150	43.430	5.703	3.925	3.406	1.420	5.449	50.507
LDOF [3]	9.116	5.037	42.344	6.849	4.928	4.003	1.485	4.839	57.296
Classic+NL [4]	9.153	4.814	44.509	7.215	4.822	3.427	1.113	4.496	60.291
Horn+Schunck [5]	9.610	5.419	43.734	7.950	5.658	3.976	1.882	5.335	58.274
Classic++ [6]	9.959	5.410	47.000	8.072	5.554	3.750	1.403	5.098	64.135
Classic+NL-fast [7]	10.088	5.659	46.145	8.010	5.738	4.160	1.092	4.666	67.801
AnisoHuber.L1 [8]	11.927	7.323	49.366	9.464	7.692	5.929	1.155	7.966	74.796


MDP-Flow2 estimated flow
MDP-Flow2 EPE

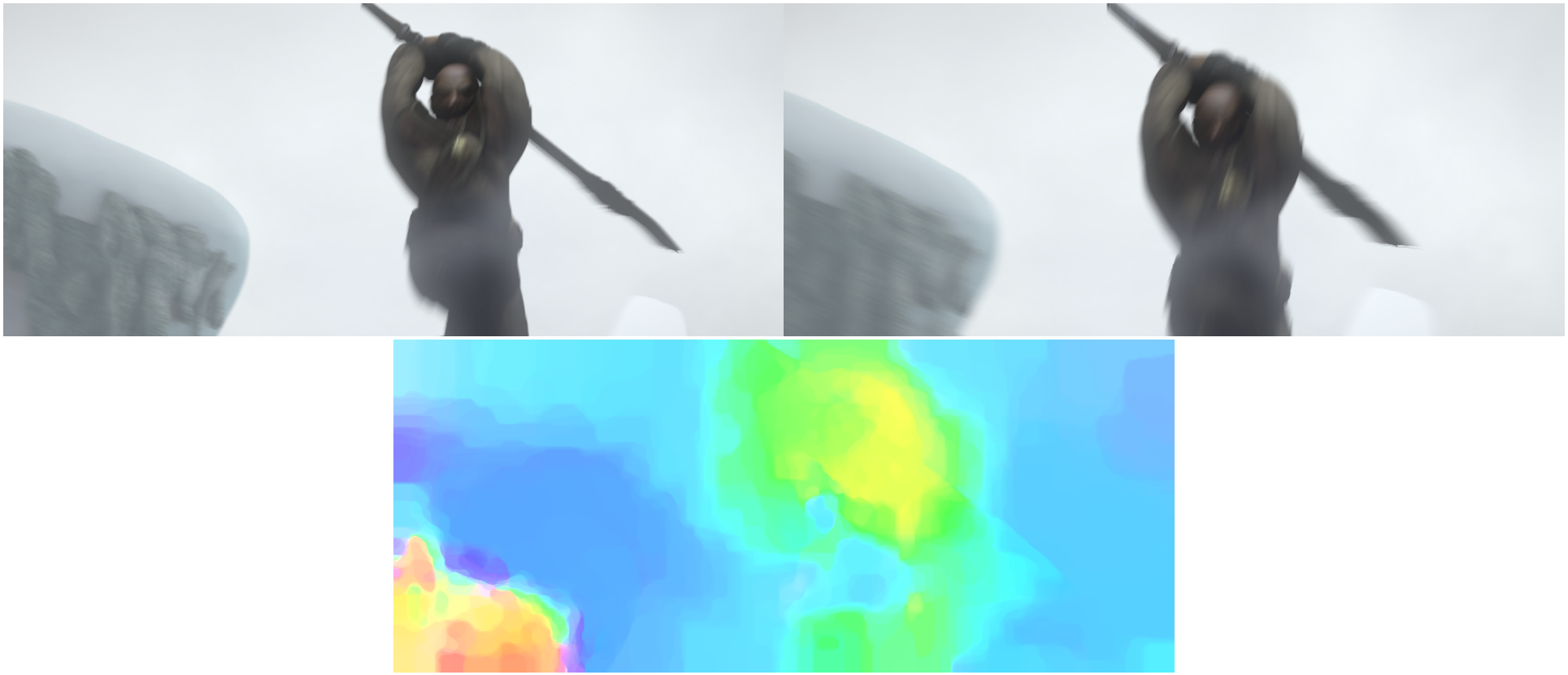

Groundtruth
MDP-Flow2 EPE


Groundtruth
MDP-Flow2 EPE
Middlebury avg **EPE**: 0.245 **px**
Sintel avg **EPE**: 8.445 **px**

Speed kills

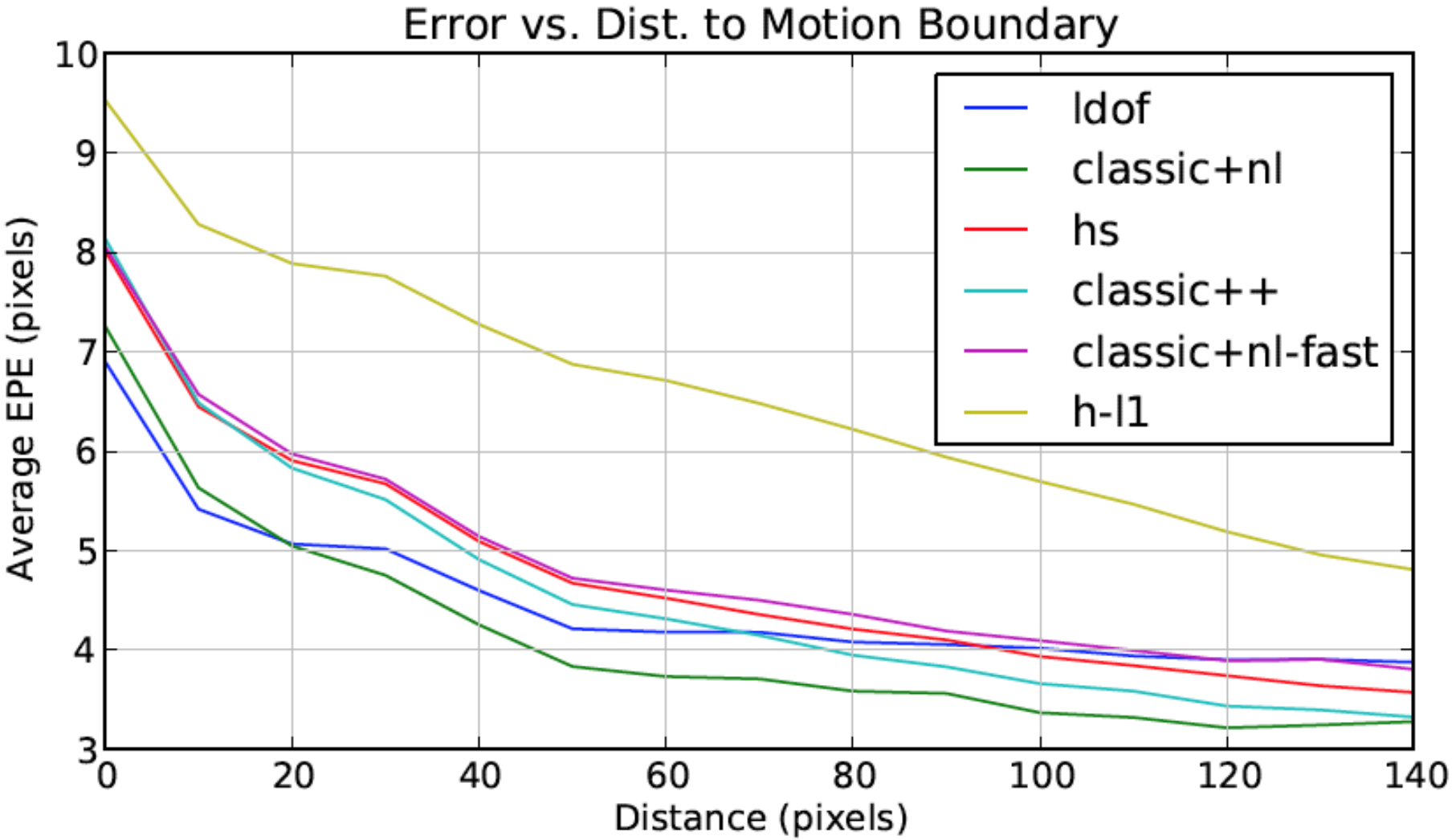

Error vs. Speed
ldof
classic+nl
hs
classic++
classic+nl-fast
h-l1
Average EPE (pixels)
Speed (pixels/frame)
10^{3}
10^{2}
10^{1}
10^{0}
10^{-1}
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}

Problems with pyramids



Small things, moving fast, get lost.
They disappear at the top of the pyramid and never come back.

Error and boundaries



Evaluation Take-aways

- Much larger errors than Middlebury (~35x)
- Unmatched regions are really hard
  ~45px error (vs. ~5px in matched regions)
- High speeds (>40 ppf) much worse than low speeds (<10 ppf)
  ~50px error vs. ~1.5px error
- Final pass harder than the Clean pass (15-40% greater error)

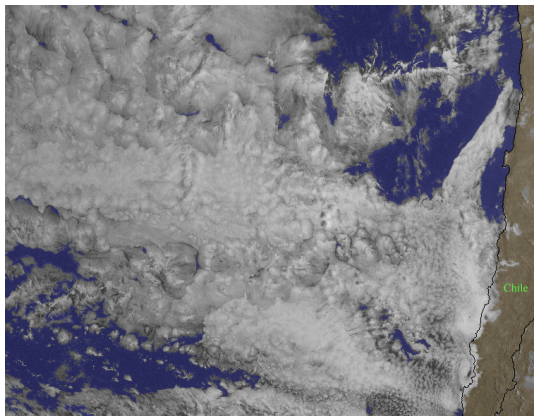
Lessons learned

- We thought this would be easy – it wasn't
- Movies just need to look good enough
- Full control of graphics data and rendering pipeline was necessary to create image sequences with accurate optical flow

See our poster today!

So what is missing?
What should future datasets address?

Snow, rain, smoke, fire….



Snowy scene flow – should the motion of the snow be in the ground truth?

Transparency

For flow evaluation, need an agreed standard for representing multiple motions at a pixel.

Small stuff

What is the optical flow of a single hair?

Motion and material

Apparent motion versus the motion field.

We went with the motion field.

What is the goal?

Accurate pixel motion?

Recognition of substances?

Recognition of motion “patterns”?

Goals may vary.

Challenges for optical flow supported by MPI-Sintel

1. Unmatched regions
   - Will encourage new methods that integrate information over time and incorporate layering

2. High speeds (>40px per frame)
   - Lookalikes exhibit these regions as well

3. Temporal processing of longer sequences
   – does it help?

4. Motion blur, defocus blur, atmospheric effects
   - Real world effects cause problems for current methods

http://sintel.is.tue.mpg.de

MJB and GBS were supported in part by NSF CRCNS Grant IIS-0904630.