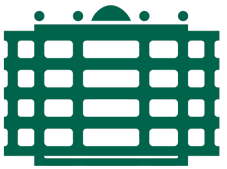
Evaluating Superpixels in Video: Metrics Beyond Figure-Ground Segmentation

Peer Neubert
Peter Protzel

TU Chemnitz



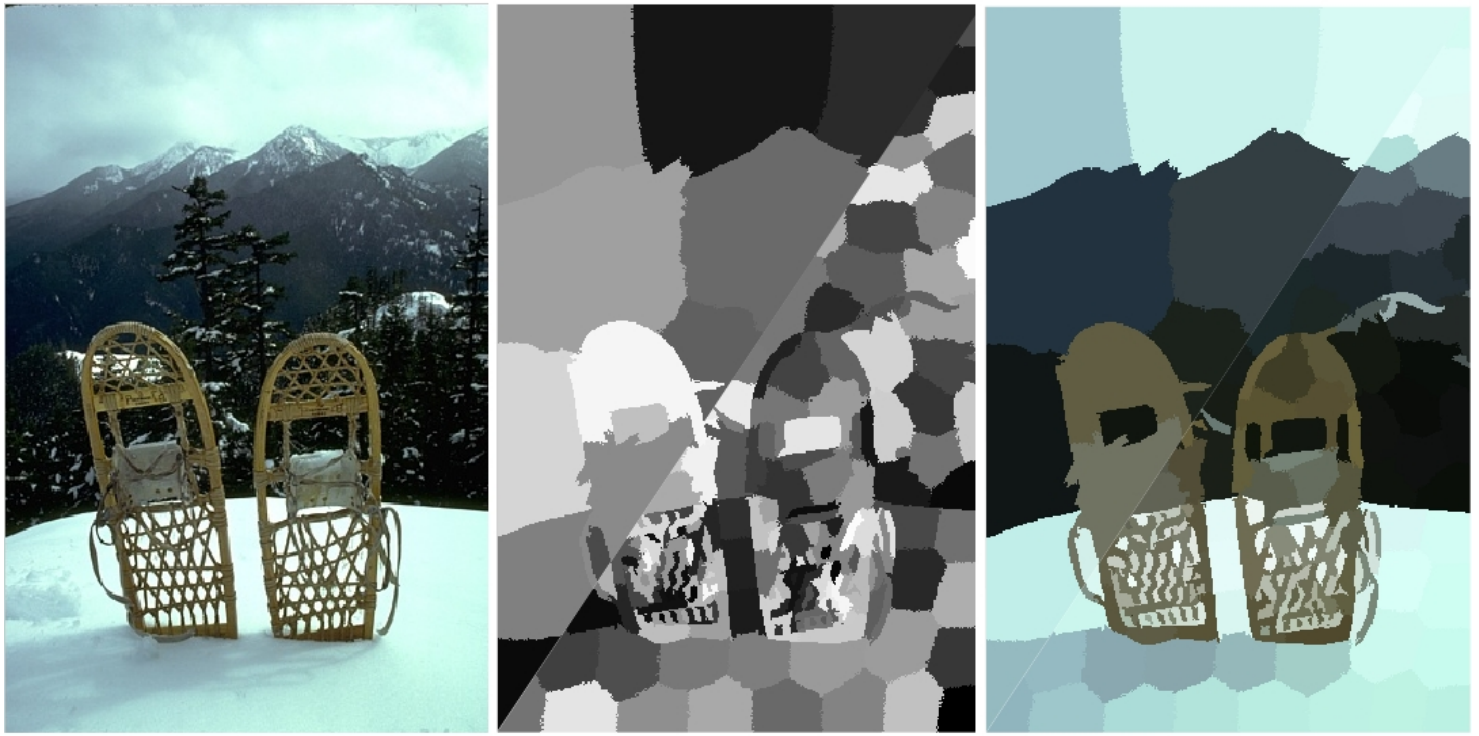
CHEMNITZ UNIVERSITY OF TECHNOLOGY

process automation

Outline

- What are **superpixels** and what are they good for?
- Introduce two **new criteria** for evaluation based on ground truth **optical flow**
- Results of various open source **algorithms**

What are Superpixels?

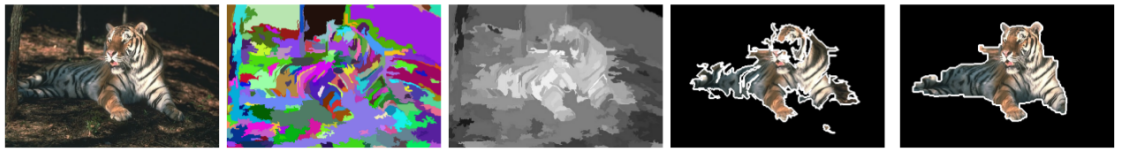


Pixel ⟷ Superpixel ⟷ Segments / Objects

- Increasing size
- Increasing meaning

What are Superpixels good for?

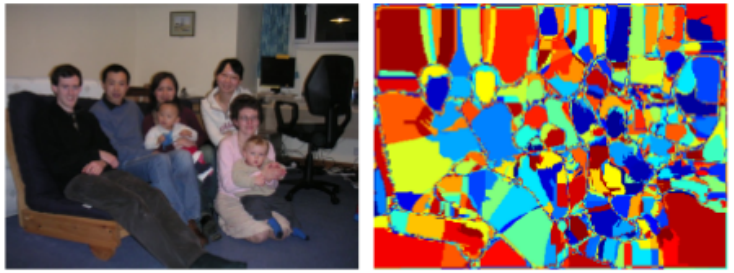
Segmentation



(a) original image (b) superpixels (c) confidences (d) classification (e) GC refinement

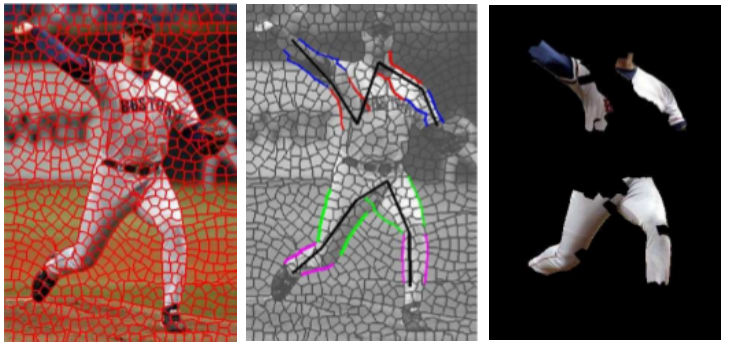
Mehrani and Veksler, 2010

Object Recognition



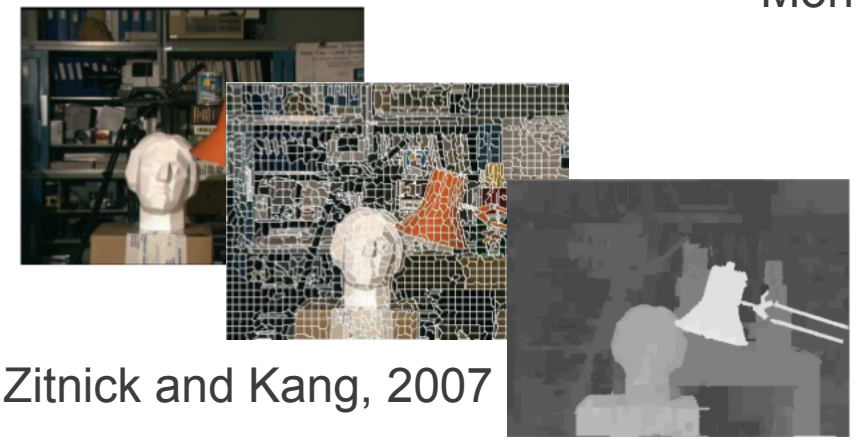
Pantofaru et al., 2008

Body model estimation



Mori, 2005

Depth estimation

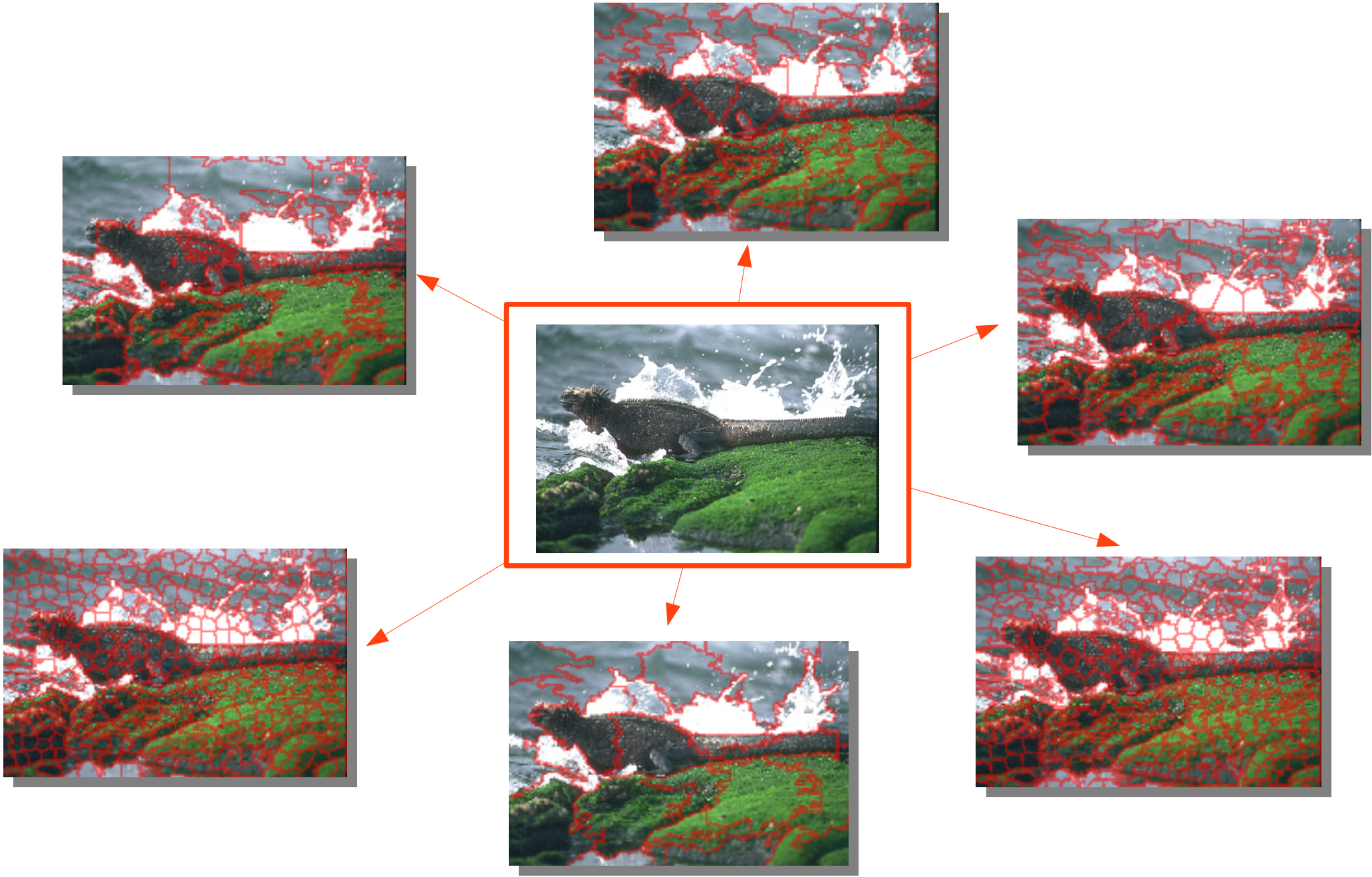


Zitnick and Kang, 2007

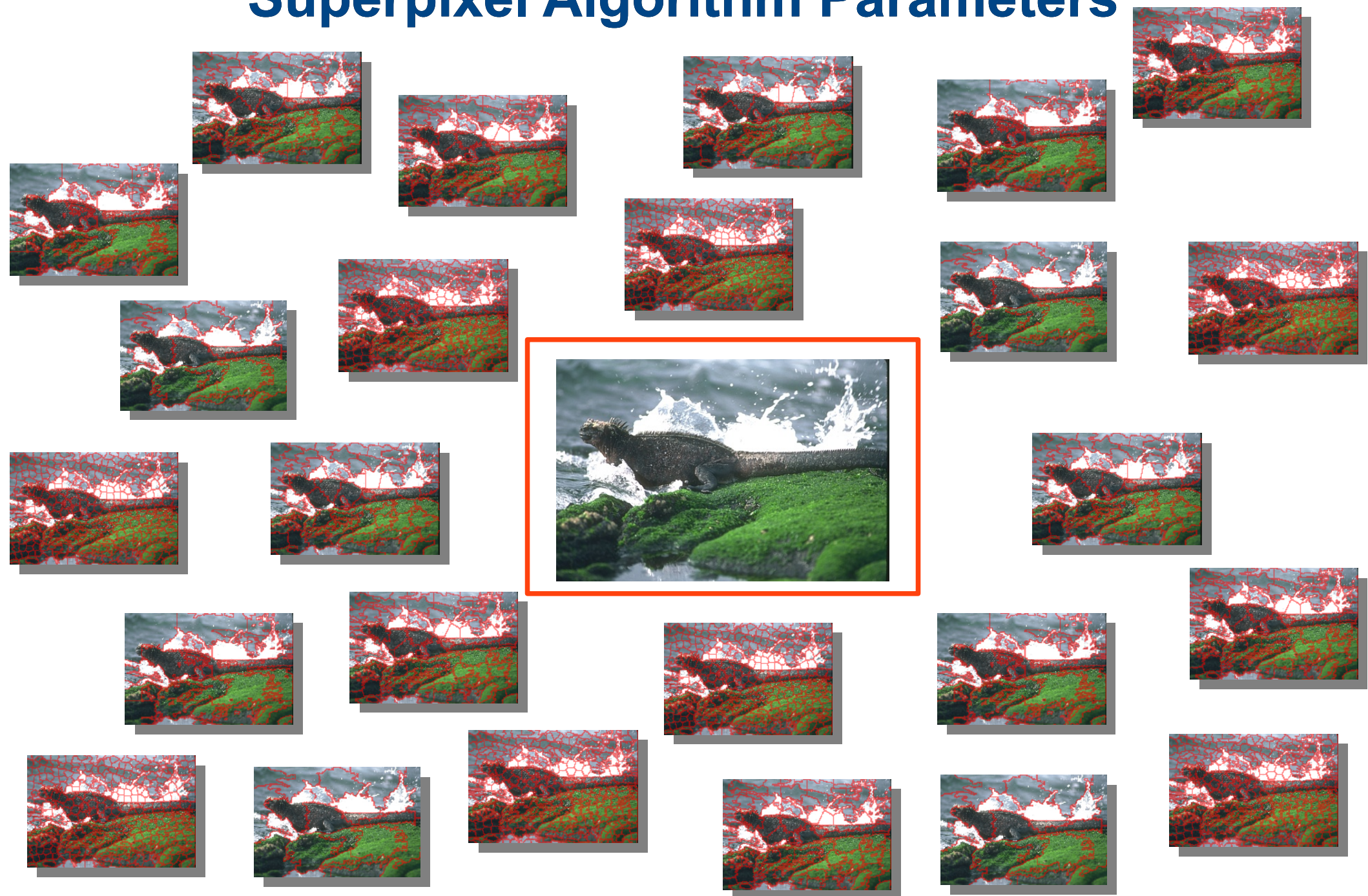
Multi-class object segmentation

Yang et al., 2010

Superpixel Algorithms



Superpixel Algorithm Parameters



What are desired properties of superpixel algorithms?

- Segmentation quality
  - Boundary recall
  - Undersegmentation error

  → Existing metrics based on figure ground segmentations
- Runtime
- Availability (implementation)
- So … **Why am I standing here?**

Stability-Criteria

Does the segmentation algorithm find the same regions or object boundaries robust to changes in the image?

Unstable Segmentation

VS

Reasonably Stable Segmentation

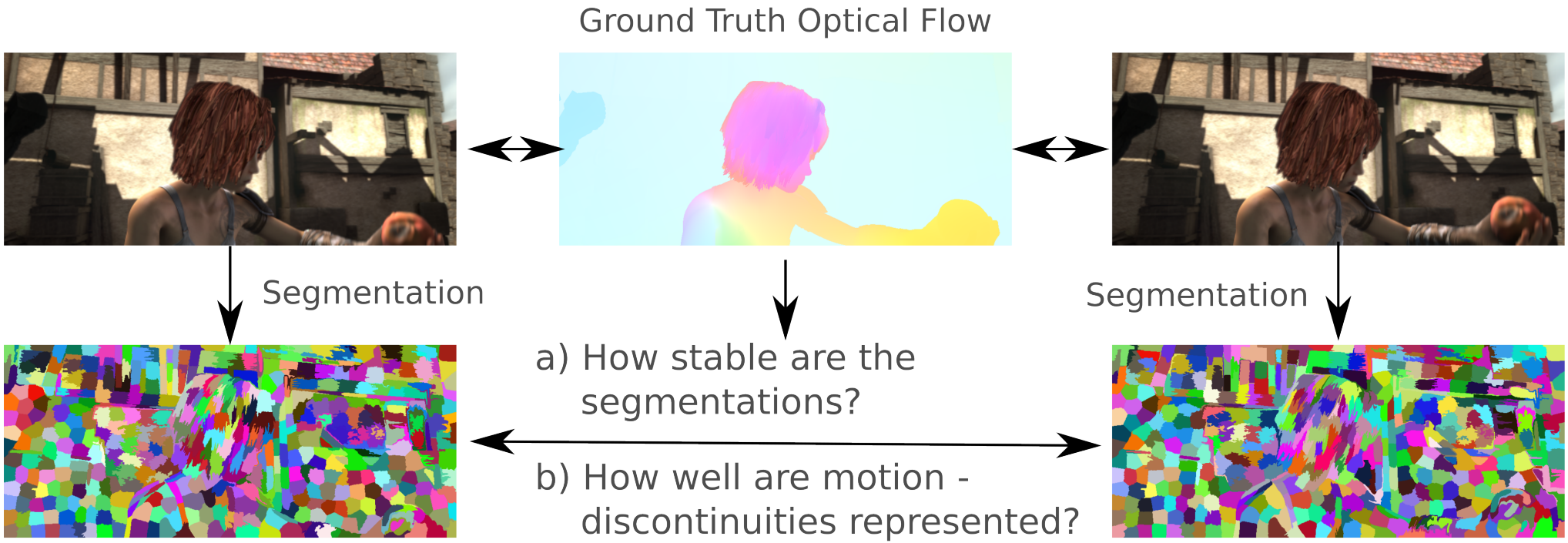
Discontinuity-Criteria

How well are motion discontinuities in the image sequence represented by the algorithms segment boundaries?

E.g.:

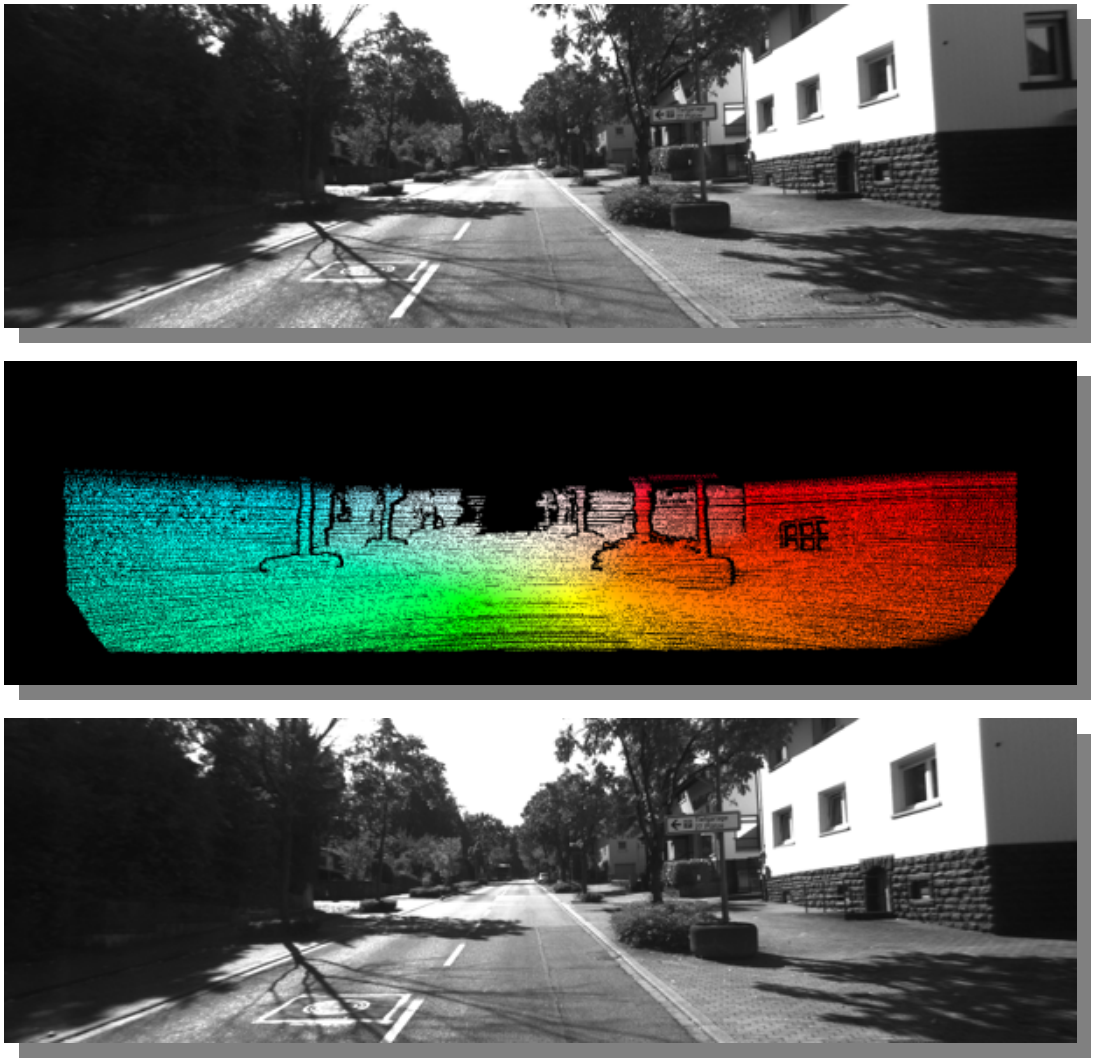
- Motion gradient between a moving foreground object and the background
- The interior of a non-rigid object

Idea: Exploit Ground Truth Optical Flow



Optical Flow Datasets: KITTI

- 194 image pairs of street scenes
- Ground truth optical flow from 3d laser data
- Geiger et al., 2012



Optical Flow Datasets: SINTEL

- Based on animated short movie Sintel
- 23 sequences with 20-50 frames
- Almost perfect ground truth optical flow computed from rendering information
- Butler et al., 2012

So far:

- Introduced two new criteria for a good superpixel algorithm
  - Stability
  - Representation of motion discontinuities
- Proposed to use ground truth optical flow for evaluation of superpixel algorithms
- Introduced the two used datasets (KITTI & Sintel)

→ Now the more technical part: Metrics

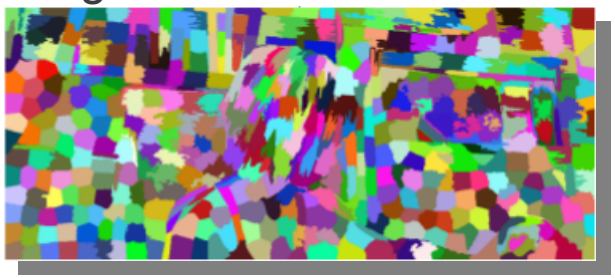
→ Followed by results on several available algorithms

Measuring the Segmentation Stability

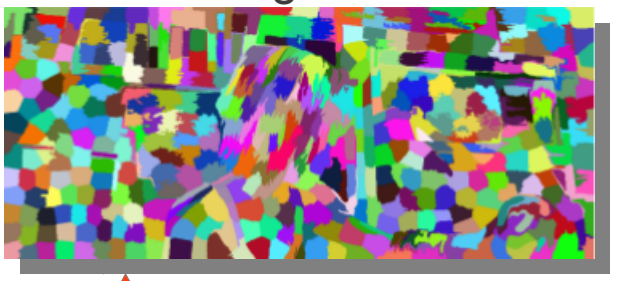
Image 1

Image 2

Segmentation L1

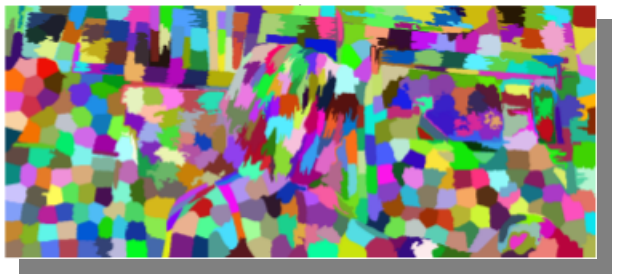


Segmentation L2



Apply ground truth optical flow

Undersegmentation Error



Measuring the Segmentation Stability: MUSE

How well is the segmentation L2 reconstructed by the transformed L1?

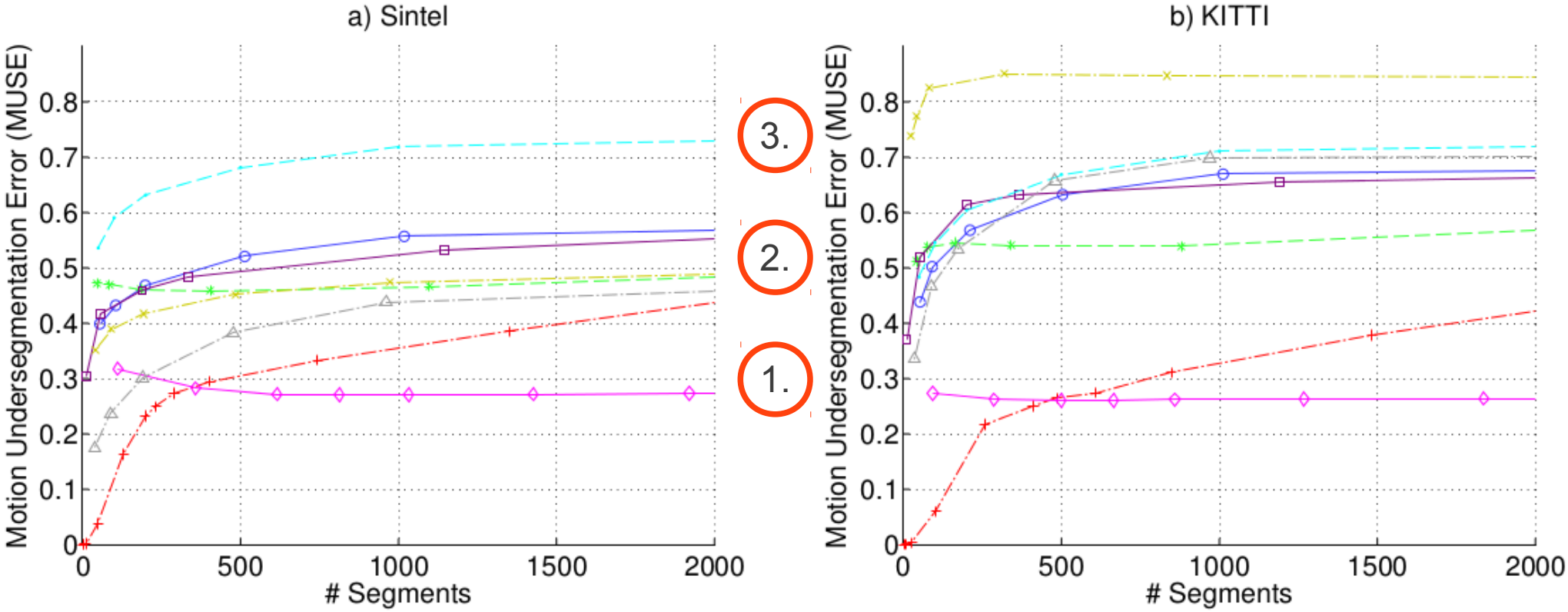
$$MUSE = \frac{1}{N}\left[\sum_{a \in L_1^F}\left(\sum_{b \in L_2 : a \cap b \neq \emptyset} min(b_{in}, b_{out})\right)\right]$$

- Motion undersegmentation error: MUSE
- Reconstruct each segment of L2 with segments of L1
- Pixels without valid flow information are ignored
- Since undersegmentation error is not a symmetric metric we compute the average of both cases

Results: Compared Algorithms

- Requirements:
  - Available open source implementation
  - Reasonable runtime on images of size 1024 × 436 for application on videos
- Compared Algorithms:
  - Felzenszwalb-Huttenlocher Segmentation (FH)
  - Edge Augmented Mean Shift (EAMS)
  - Quickshift (vlQS)
  - Marker-Controlled Watershed (WS)
  - Entropy Rate Superpixel Segmentation (ERS)
  - Two implementations of Simple Linear Iterative Clustering
    - Original implementation (oriSLIC)
    - Implementation of VlFeat (vlSLIC)
  - Baseline: Regular grid (BOX)

Results: MUSE

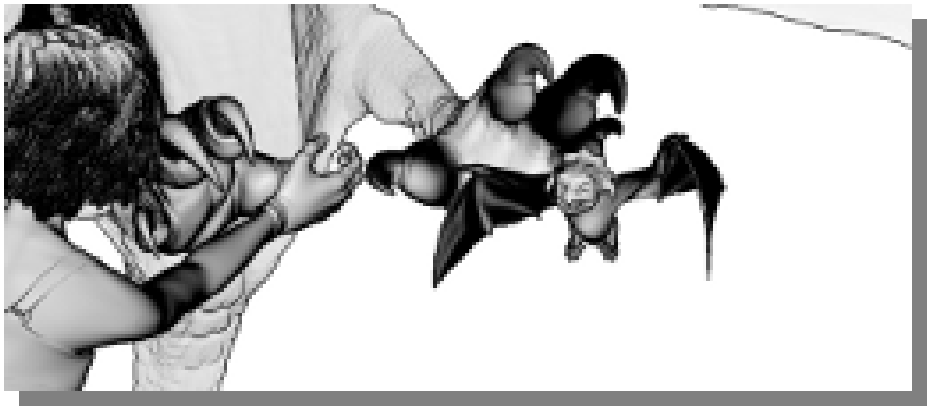


We can distinguish three groups of algorithms:

1. Gradient affine algorithms without regulation of size or distribution
2. Algorithms with additional stable boundaries (compactness constraints)
3. Unstable Algorithms (parameter sets)

Measuring the Accordance with Motion Discontinuities

- Motion discontinuities result in high gradients in the optical flow field



- High motion gradient indicates objects or object parts that can move differently and thus should probably be handled individually in the application.
- Motion Discontinuity Error: MDE

Measuring the Accordance with Motion Discontinuities: MDE

Summarize over all image pixels (i,j)

Strength of motion discontinuity at this pixel

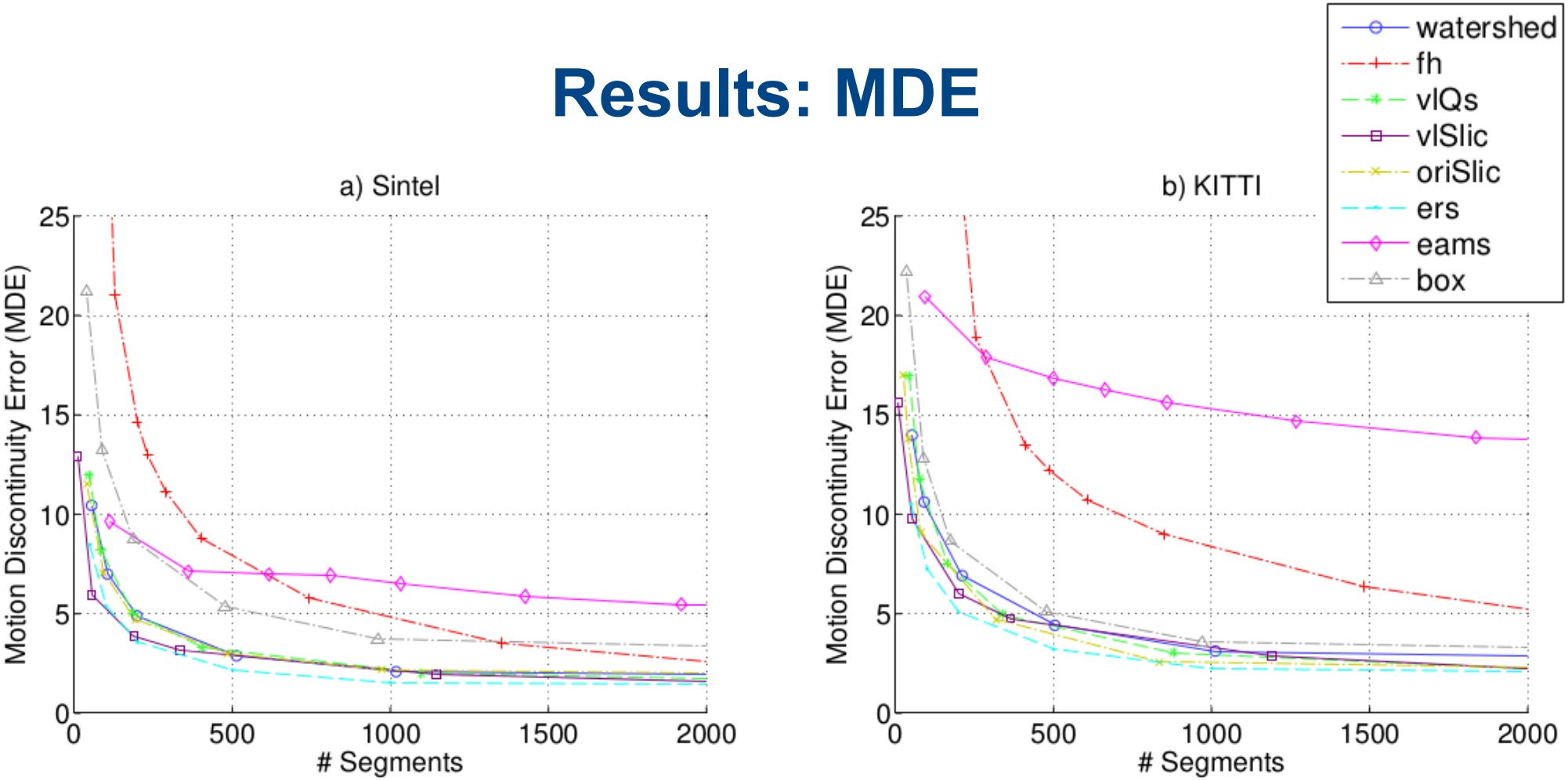
$$MDE = \frac{1}{\sum_i \sum_j \|\nabla F(i,j)\|_2} \sum_i \sum_j \|\nabla F(i,j)\|_2 \cdot D(B(i,j))$$

Total amount of motion

How far is the pixel from a segment border?

F: A ground truth optical flow field from an image I to another image
B: The boundary image of a segmentation of image I
D(B): The distance transform of B containing for each pixel the distance to the nearest segment boundary

Results: MDE



Results are inverted:

- Compactness constraints help a lot
  - ➔ They introduce additional borders that decrease the average values of the distance transform of the boundary image

- Exclusively relying upon image gradients is not enough

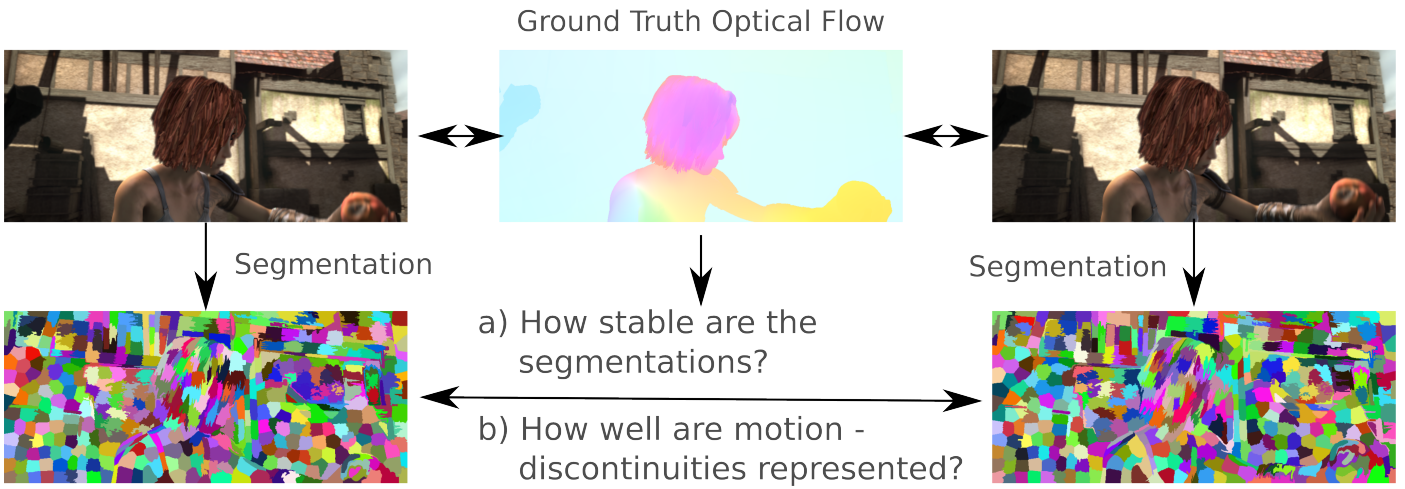
Discussion of the metrics

- MDE favors segmentations with many boundary pixels
  - Similar to measures like boundary recall
  - Should be used together with complementing measurements
- MDE and MUSE are somehow complementary

  → Look for well balanced algorithms/parameters
- MDE and MUSE do not replace figure-ground metrics!

Conclusions

- Proposed two novel criteria for evaluating superpixel segmentations
  - Segmentation stability
  - Representation of motion discontinuities
- Exploit ground truth optical flow data to measure them
  - Motion Undersegmentation Error (MUSE) evaluates the stability of segmentations
  - Motion Discontinuity Error (MDE) evaluates how well differently moving image parts are separated by the segmentation algorithms.
  - Matlab Toolbox: http://www.tu-chemnitz.de/etit/proaut/forschung/superpixel.html
- (vlSLIC and Quickshift algorithms showed best balanced results)



Questions?

- E.g.:
  - How are the parameters of the algorithms chosen?
  - Stable superpixels for video? Why not supervoxels!?
  - What should the perfect superpixel algorithm look like?
  - What about other algorithms like Normalized Cuts, Turbopixels, Superpixel Lattice, gPb-OWT-UCM, …?

Or let's discuss at the coffee break