Leibniz Universität Hannover

Temporally Consistent Superpixels

ICCV 2013 Paper

Matthias Reso
Jörn Jachalsky
Bodo Rosenhahn
Jörn Ostermann

Zürich, 13. September 2014

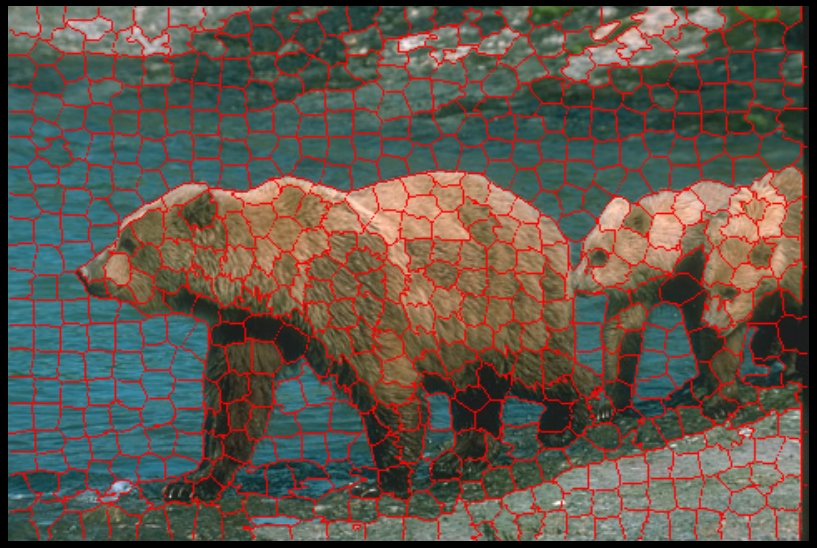
Institut für Informationsverarbeitung
Leibniz Universität Hannover

technicolor

Introduction

Original

Boundary image



Label map

Average map

Matthias Reso
reso@tnt.uni-hannover.de

SLIC Superpixels

- Introduced by Achanta et al. [1]
- Pixels are viewed as data points in a 5D feature space [labxy]
- Superpixels form clusters in this 5D feature space and are represented by cluster centers
- Cluster analysis by minimizing the energy term [2]

$$E_{total}(\Delta) = \sum_{n \in \mathcal{N}} (1-\alpha) E_{color}(n, \delta_n) + \alpha E_{spatial}(n, \delta_n)$$

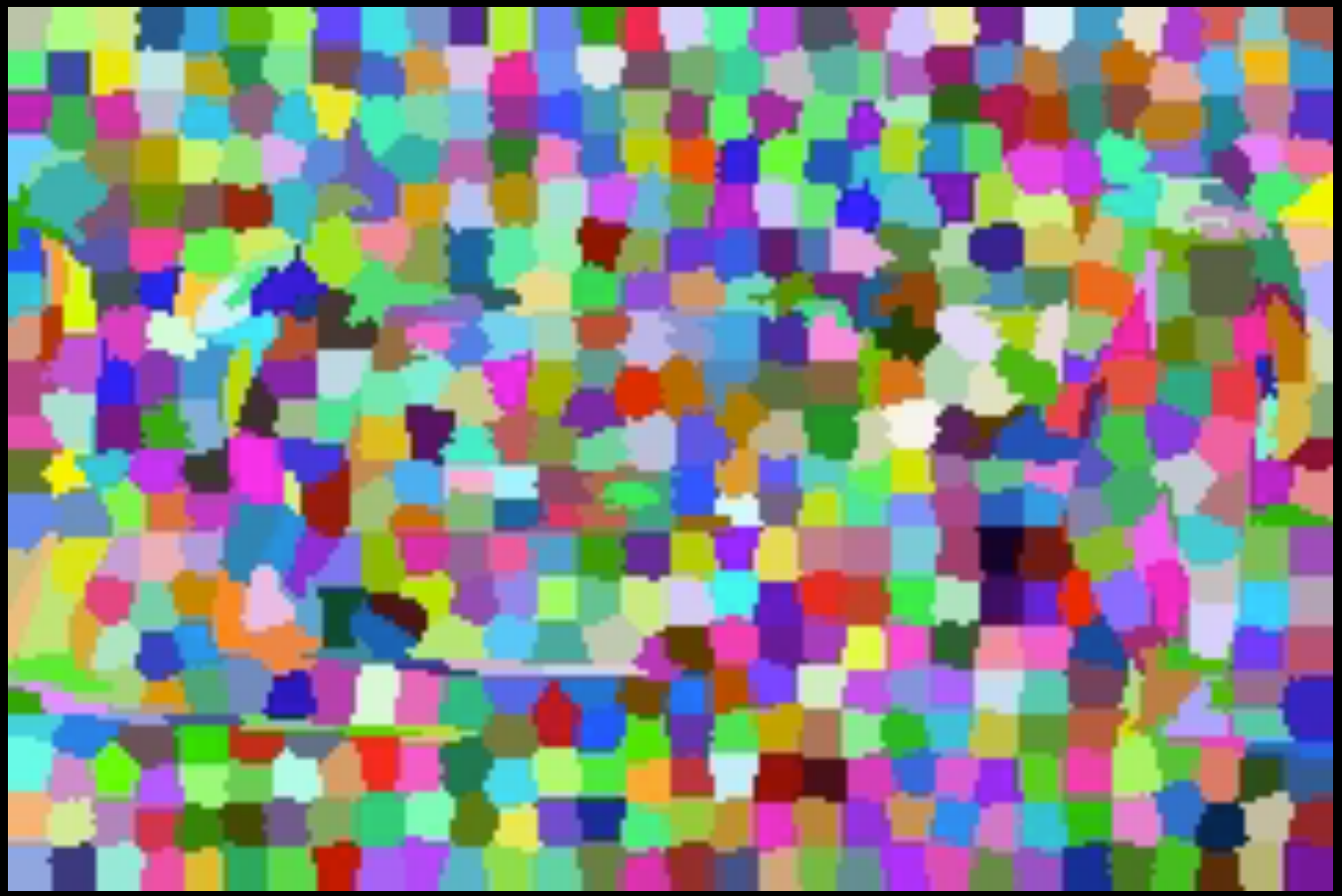
- Energies E_{color} and $E_{spatial}$ are proportional to Euclidean distances
- Optimization by employing an Expectation-Maximization-scheme

[1] Achanta, Shaji, Smith, Lucchi, Fua, and Susstrunk. SLIC superpixels compared to state-of-the-art superpixel methods. TPAMI, 34(11):2274–2282, 2012

[2] Schick, Fischer, and Stiefelhagen. Measuring and evaluating the compactness of superpixels. In ICPR, 930–934, 2012

Introduction

First Challenge: Temporal Consistency

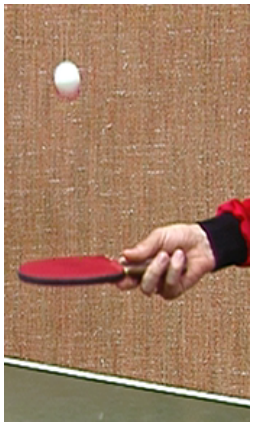


Matthias Reso
reso@tnt.uni-hannover.de

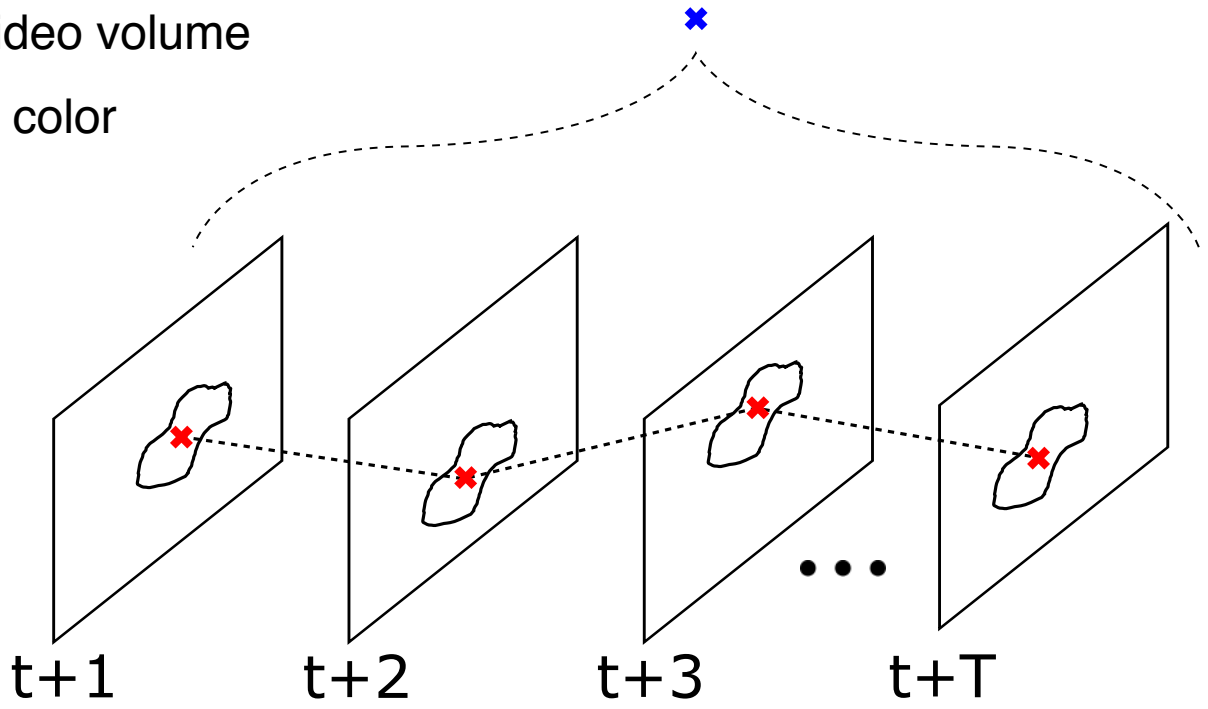
Basic Idea

- Extend SLIC to the video volume
- We assume constant color
- But we can't assume constant position

Frame t



Frame t + 1



1 Global Color Center

T Local Spatial Centers

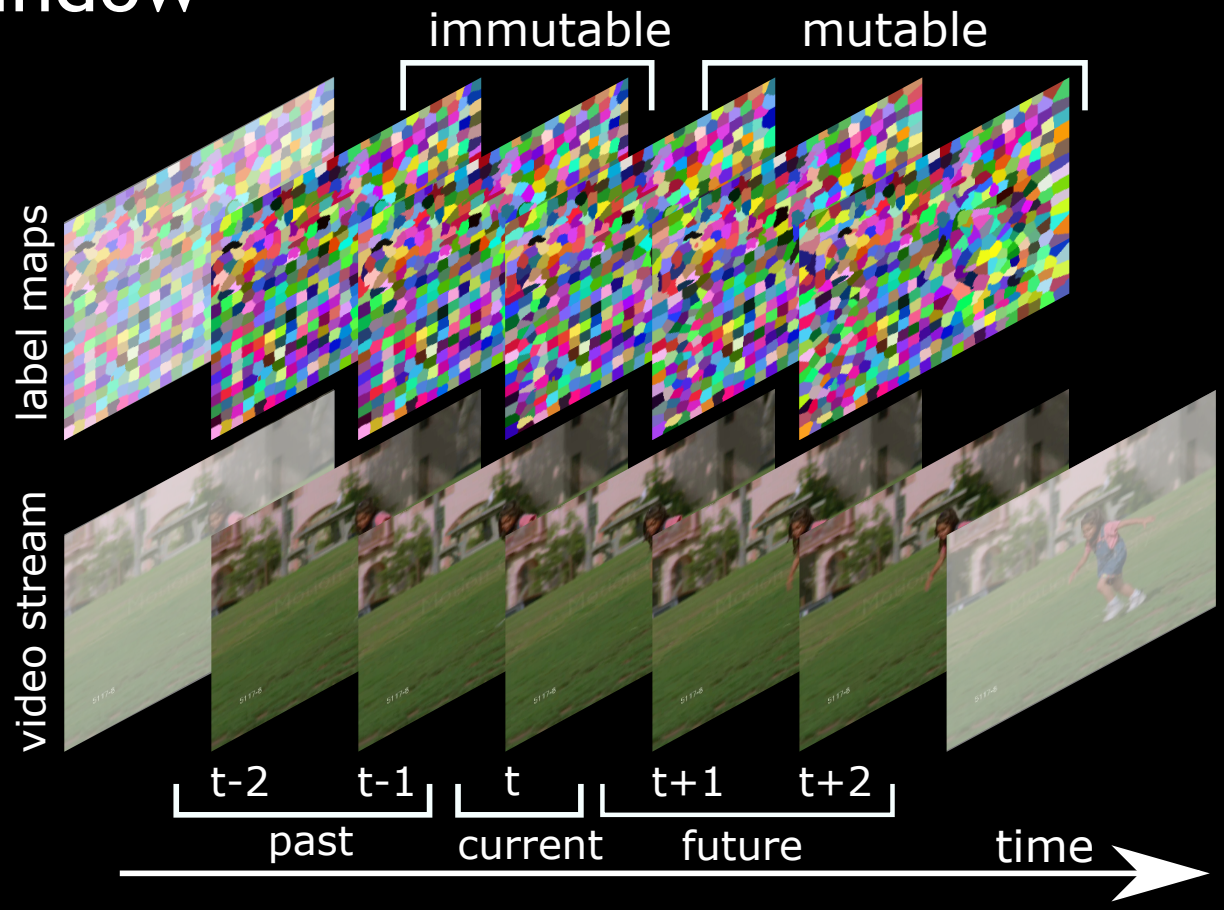
Extended Energy Term

- Minimization over multiple-frames

$$E_{New}(\Delta) = \sum_{\tau \in \mathcal{T}} \sum_{n \in \mathcal{N}(\tau)} (1-\alpha) E_{color}(n, \delta_n) + \alpha E_{spatial}(n, \delta_n, \tau)$$

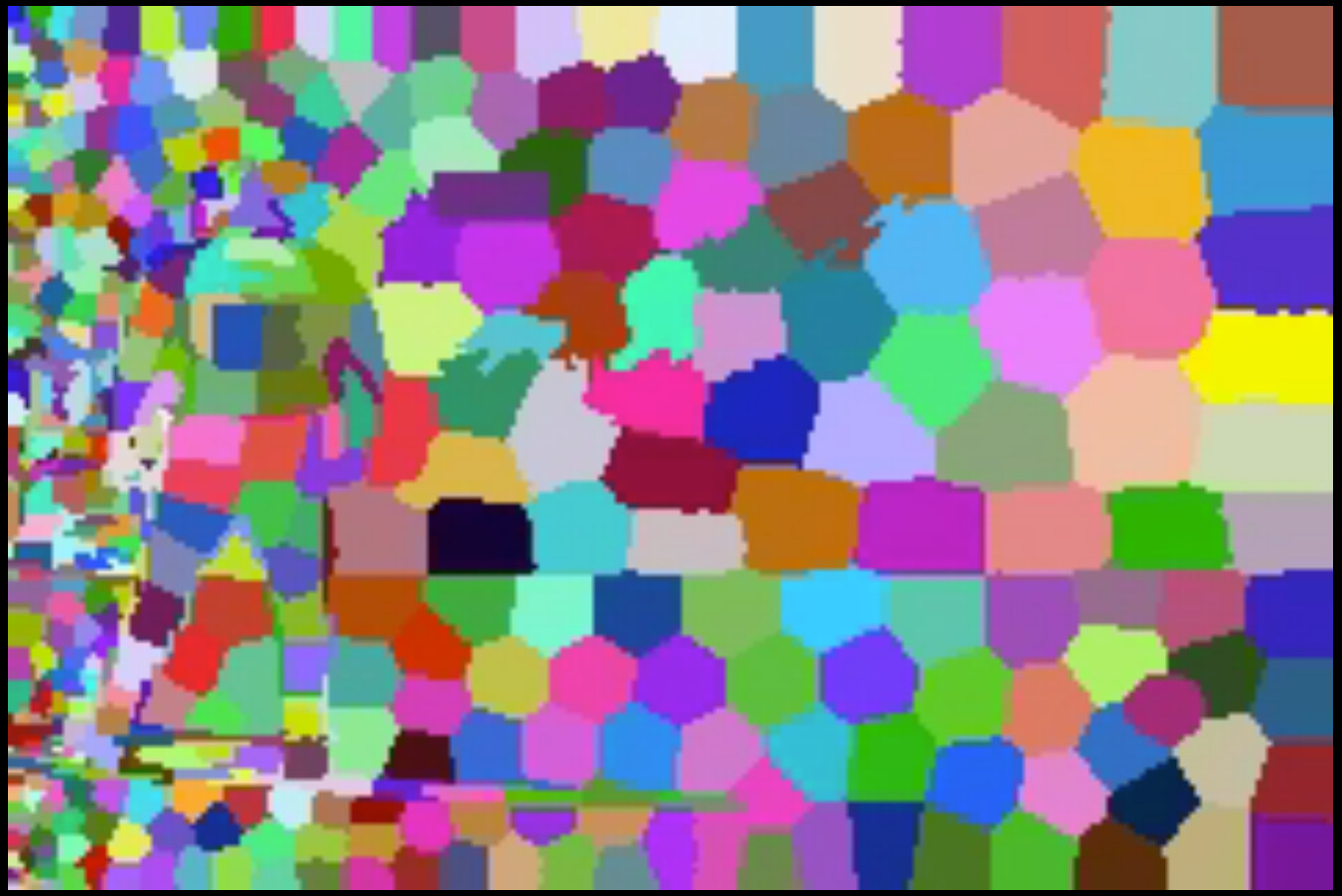
- Energy $E_{spatial}$ now depends on frame τ
- A sliding window is introduces to allow
  - Streaming capabilities
  - Gradual illumination changes

Sliding Window



– New frames are initialized using the optical flow

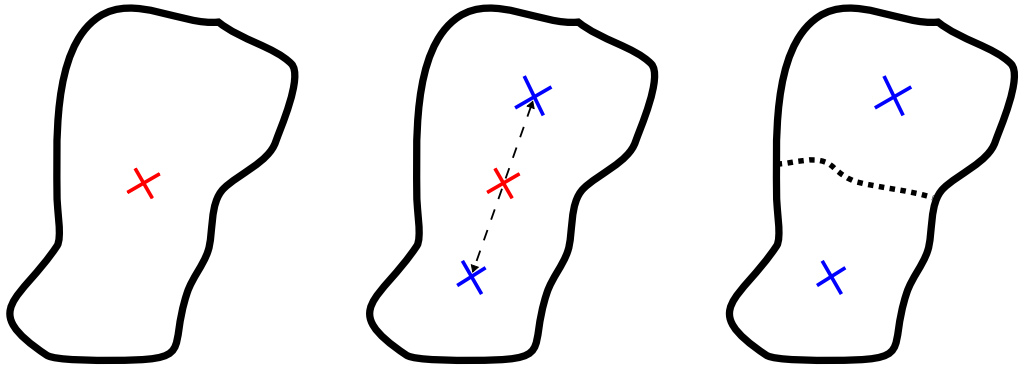
Further Challenge: Handling of Structural Changes



Matthias Reso
reso@tnt.uni-hannover.de

Structural Changes and Spatial Coherency

- Monitor superpixel size inside sliding window
- Extrapolation based on linear growth assumption
- Balanced splitting and termination of superpixels



- Ensuring spatial coherency by a new contour evolution based post-processing step

Benchmarks

- Comparison to
  - SLIC supervoxel approach of Achanta et al. [1]
  - Streaming Hierarchical Video Segmentation of Xu et al. [3]
- Quantitative results on Chen's dataset [4]
  - 3D Undersegmentation Error [5]
  - Mean Temporal Length [5]

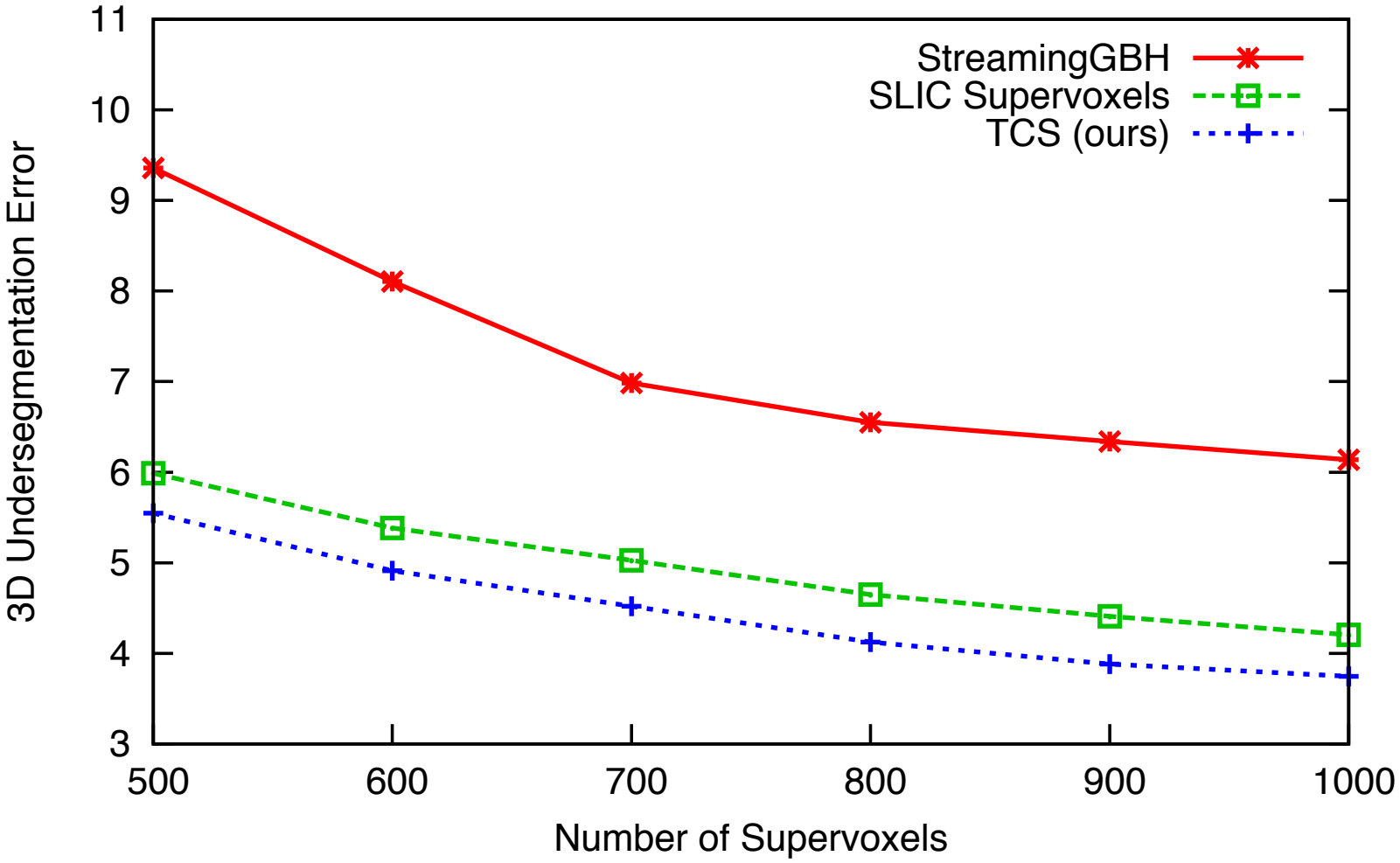
[1] Achanta, Shaji, Smith, Lucchi, Fua, and Susstrunk. SLIC superpixels compared to state-of-the-art superpixel methods. TPAMI, 34(11):2274–2282, 2012

[3] Xu, Xiong, and Corso. Streaming hierarchical video segmentation. In ECCV, pages 626–639, 2012

[4] Chen and Corso. Propagating multi-class pixel labels throughout video frames. In WNYIPW, pages 14–17, 2010

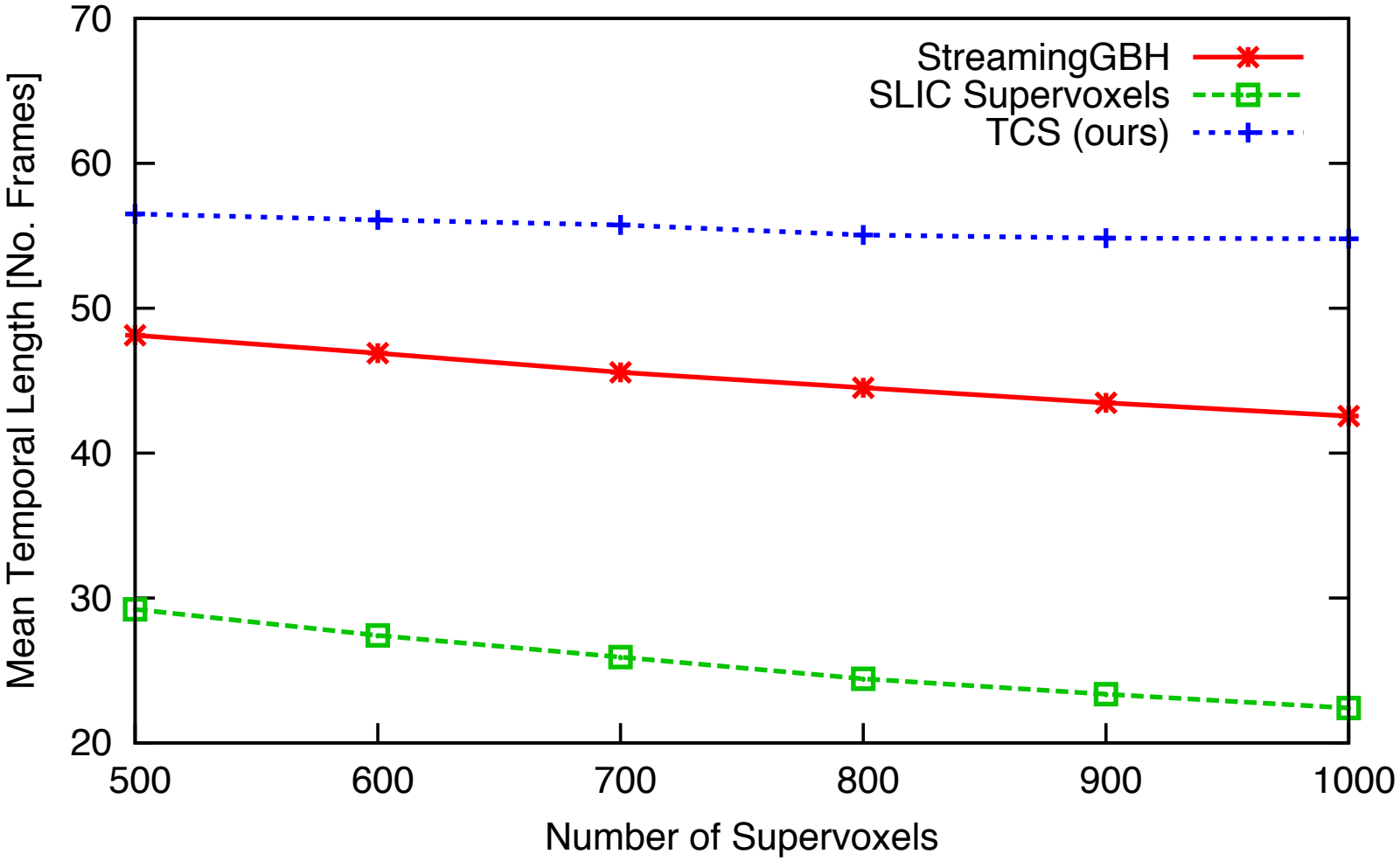
[5] Calculated using LIBSVX http://www.cse.buffalo.edu/~jcorso/r/supervoxels/

3D Undersegmentation Error



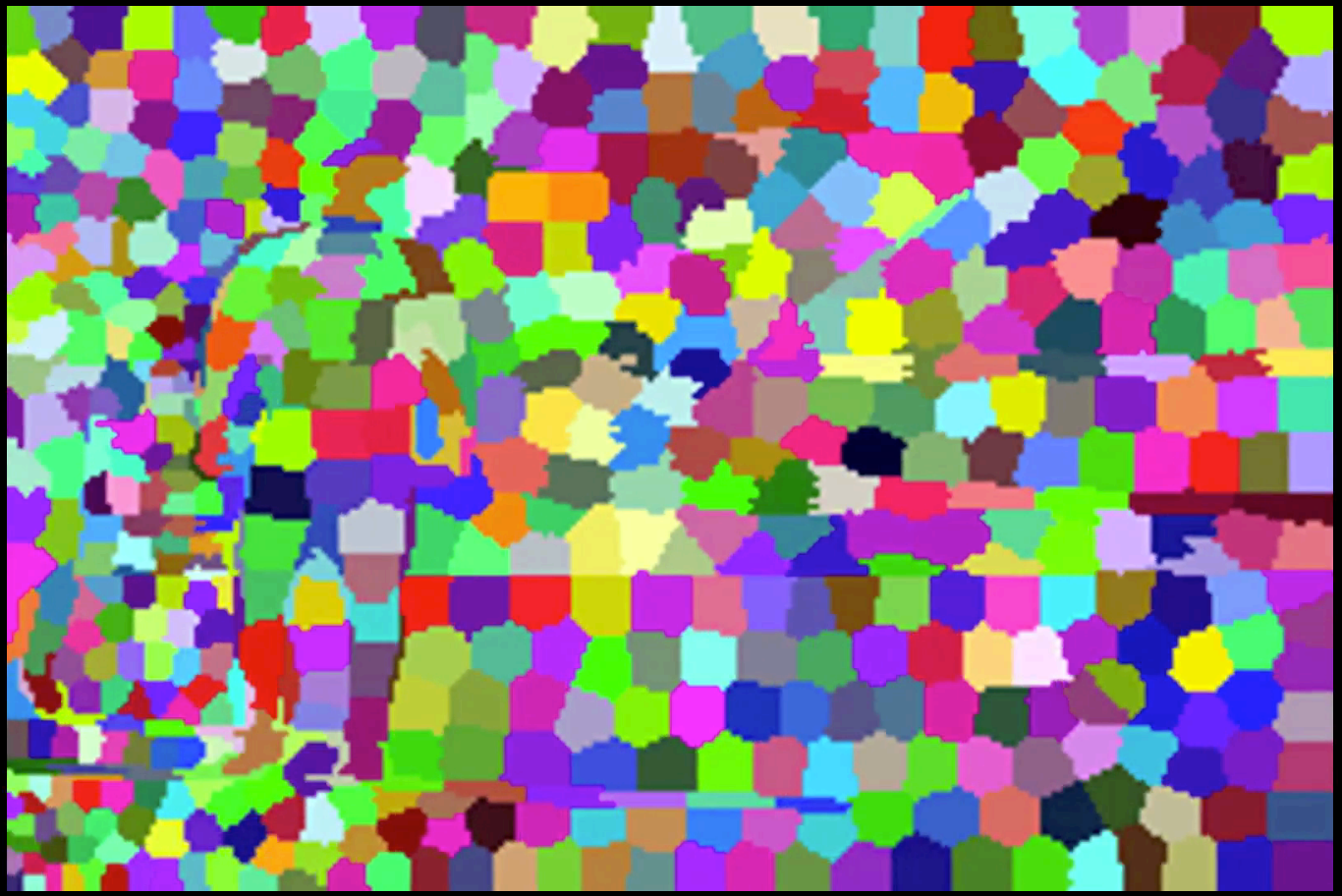
Matthias Reso
reso@tnt.uni-hannover.de

Mean Temporal Length



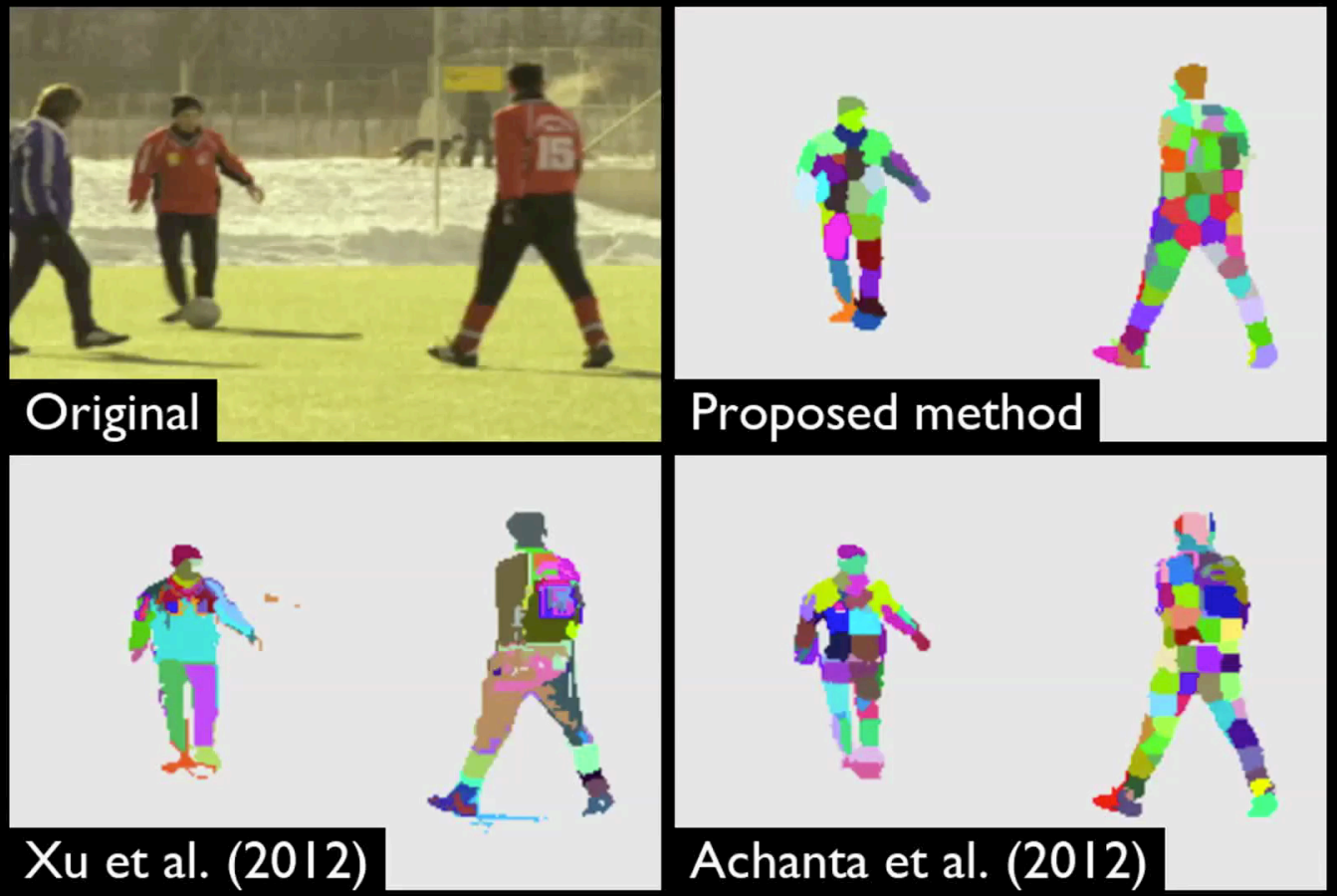
Matthias Reso
reso@tnt.uni-hannover.de

Qualitative Results



Matthias Reso
reso@tnt.uni-hannover.de

Qualitative Results



Matthias Reso
reso@tnt.uni-hannover.de

Summary

- Hybrid clustering approach working on the video volume
  - Global color cluster centers
  - Local spatial cluster centers
- Sliding window to enable streaming capabilities
- Handling of structural changes
- Improved 3D undersegmentation error and mean temporal length