# Texture Superpixel Clustering from Patch-based Nearest Neighbor Matching

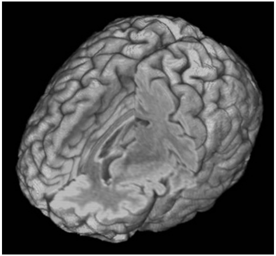
Rémi Giraud Yannick Berthoumieu

Bordeaux INP ENSEIRB MATMECA

ims

Large data $\rightarrow$ high computational times

Image HD

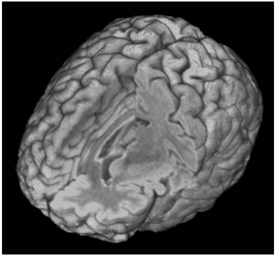


Volume 3D


Video

Large data $\rightarrow$ high computational times $\rightarrow$ Dimension reduction

Image HD

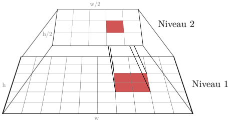


Volume 3D


Video

Large data → high computational times → Dimension reduction

- Regular multi-resolution:
  Decompose the image into regular blocks





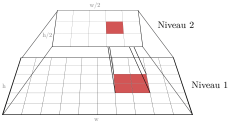
Image


Decomposition into blocks


Average colors

Large data → high computational times → Dimension reduction

- Regular multi-resolution:
  Decompose the image into regular blocks
- Superpixels (since [Ren and Malik, 2003]):
  Local grouping of pixels with homogeneous colors

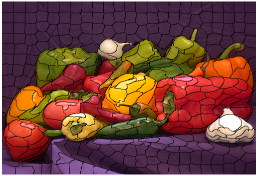




Image


Decomposition into blocks

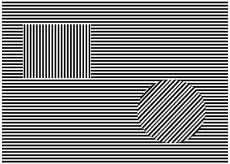

Average colors

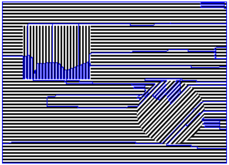

Decomposition into superpixels

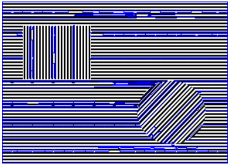

Average colors

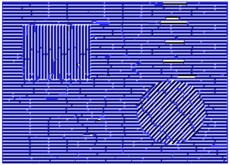
# Robustness of state-of-the-art methods

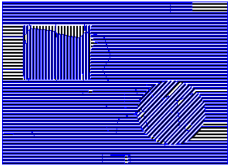
What about the ability to handle texture?

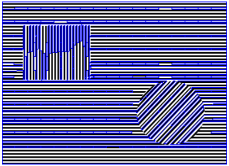

Initial image

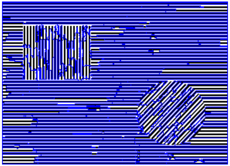

SLIC [Achanta et al., 2012]


ERGC [Buyssens et al., 2014]


ETPS [Yao et al., 2015]

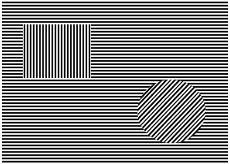

LSC [Chen et al., 2017]

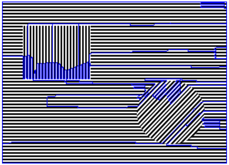

SNIC [Achanta et al., 2017]

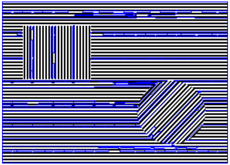

SCALP [Giraud et al., 2018]

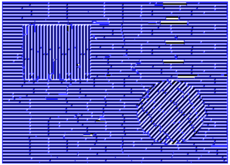
→ All state-of-the-art methods severely fail at clustering textures

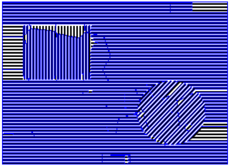
What about the ability to handle texture?

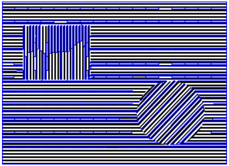

Initial image

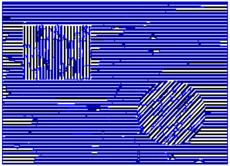

SLIC [Achanta et al., 2012]

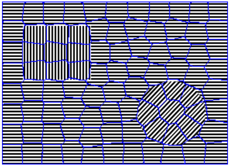

ERGC [Buyssens et al., 2014]


ETPS [Yao et al., 2015]


LSC [Chen et al., 2017]


SNIC [Achanta et al., 2017]


SCALP [Giraud et al., 2018]

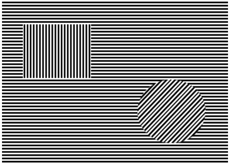

TASP [Giraud et al., 2019]

→ All state-of-the-art methods severely fail at clustering textures

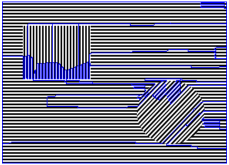
→ Introduce a texture homogeneity term using patch-based distances

→ K-means-based clustering approach (TASP) → high complexity

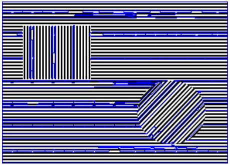
What about the ability to handle texture?



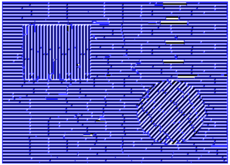
Initial image



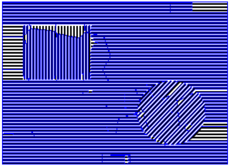
SLIC [Achanta et al., 2012]



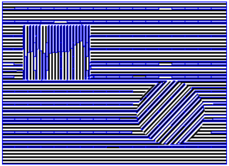
ERGC [Buyssens et al., 2014]



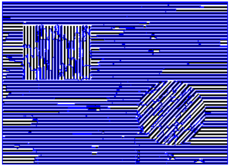
ETPS [Yao et al., 2015]



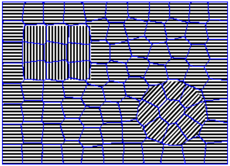
LSC [Chen et al., 2017]



SNIC [Achanta et al., 2017]



SCALP [Giraud et al., 2018]



TASP [Giraud et al., 2019]

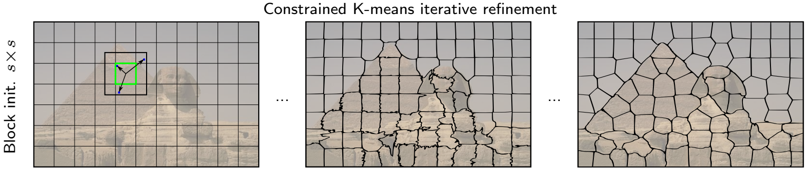
→ All state-of-the-art methods severely fail at clustering textures

→ Introduce a texture homogeneity term using patch-based distances

- → K-means-based clustering approach (TASP) → high complexity
- → Nearest Neighbor-based Superpixel Clustering (NNSC)

# K-means-based Clustering Approach (TASP)

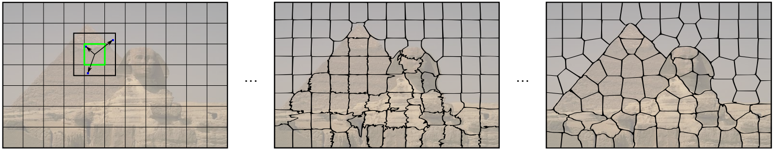
- Simple Linear Iterative Clustering (SLIC) [Achanta et al., 2012]

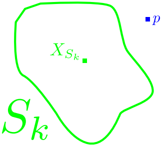


- Simple Linear Iterative Clustering (SLIC) [Achanta et al., 2012]

Block init. $s \times s$

Constrained K-means iterative refinement



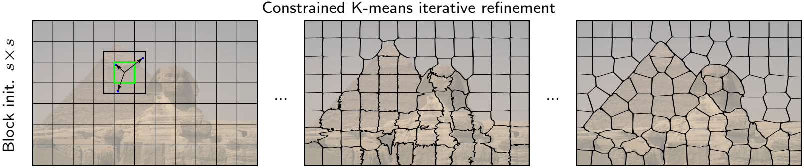



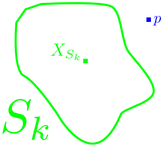

$F_p = [l_p, a_p, b_p]$ color in the CIELab space
$X_p = [x_p, y_p]$ position
$F_{S_k}, X_{S_k}$ average on pixels $\in S_k$
$m_k$ regularity parameter

Distance between a pixel $p$ and a superpixel $S_k$:

$$D(p, S_k) = d_{\text{color}}(F_p, F_{S_k}) + d_{\text{spatial}}(X_p, X_{S_k})m_k$$

- Simple Linear Iterative Clustering (SLIC) [Achanta et al., 2012]




$F_p = [l_p, a_p, b_p]$ color in the CIELab space
$X_p = [x_p, y_p]$ position
$F_{S_k}, X_{S_k}$ average on pixels $\in S_k$
$m_k$ regularity parameter

**Distance between a pixel $p$ and a superpixel $S_k$:**

$$D(p, S_k) = d_{\text{color}}(F_p, F_{S_k}) + d_{\text{spatial}}(X_p, X_{S_k}) m_k$$

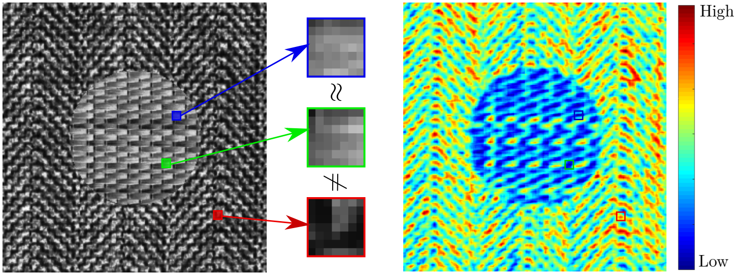
$\rightarrow$ Complexity $\quad C_{\text{SLIC}} = \mathcal{O}((h \times w) \times 4 \times \text{Iter}_{\text{K-means}})$

- Pixel to superpixel texture homogeneity term:

  → *Using patch-based distances*

  *No complex texture classification approach*

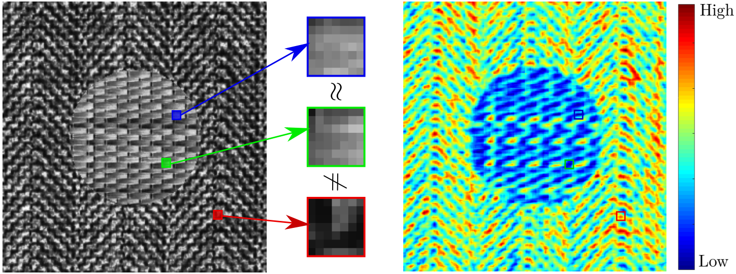
  *Remains in the same feature space than pixel to superpixel distances*



- Pixel to superpixel texture homogeneity term:

  → *Using patch-based distances*

  *No complex texture classification approach*
  *Remains in the same feature space than pixel to superpixel distances*



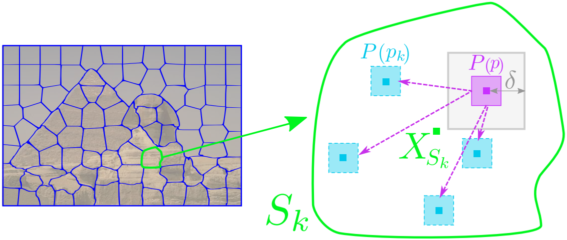

Which patches to compare?

- Pixel to superpixel texture homogeneity term:

  → Nearest neighbor (NN) matching within the superpixel

  *Ability to find only similar texture patterns*

  *Fast selection of $N$ similar patches with PatchMatch [Barnes et al., 2009]*

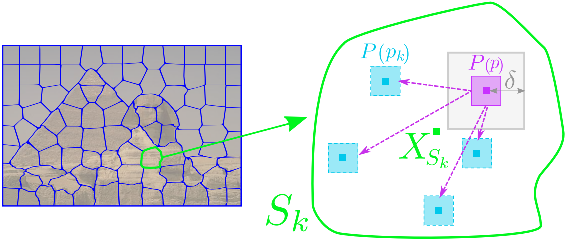


- Pixel to superpixel texture homogeneity term:

  → Nearest neighbor (NN) matching within the superpixel

  *Ability to find only similar texture patterns*

  *Fast selection of $N$ similar patches with PatchMatch [Barnes et al., 2009]*



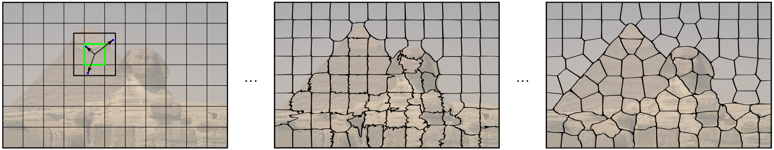

Texture homogeneity distance:

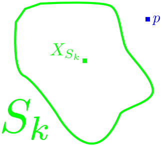
$$d_{\text{texture}}(p, S_k) = \frac{1}{N} \sum_{p_k \in \mathcal{K}_p} \frac{1}{n} \|F_{P(p)} - F_{P(p_k)})\|_2$$

- Simple Linear Iterative Clustering (SLIC) [Achanta et al., 2012]

Constrained K-means iterative refinement

Block init. $s \times s$




$F_p = [l_p, a_p, b_p]$ color in the CIELab space
$X_p = [x_p, y_p]$ position
$F_{S_k}, X_{S_k}$ average on pixels $\in S_k$
$m$ regularity parameter

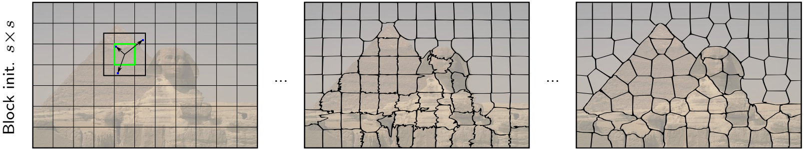
**Distance between a pixel $p$ and a superpixel $S_k$ (SLIC):**

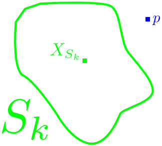
$$D(p, S_k) = d_{\text{color}}(F_p, F_{S_k}) + d_{\text{spatial}}(X_p, X_{S_k}) m_k$$

$\rightarrow$ Complexity $\quad C_{\text{SLIC}} = \mathcal{O}((h \times w) \times 4 \times \text{Iter}_{\text{K-means}})$

- Texture-Aware SuperPixels (TASP) [Giraud et al., 2019]

Constrained K-means iterative refinement




$F_p = [l_p, a_p, b_p]$ color in the CIELab space
$X_p = [x_p, y_p]$ position
$F_{S_k}, X_{S_k}$ average on pixels $\in S_k$
$m$ regularity parameter

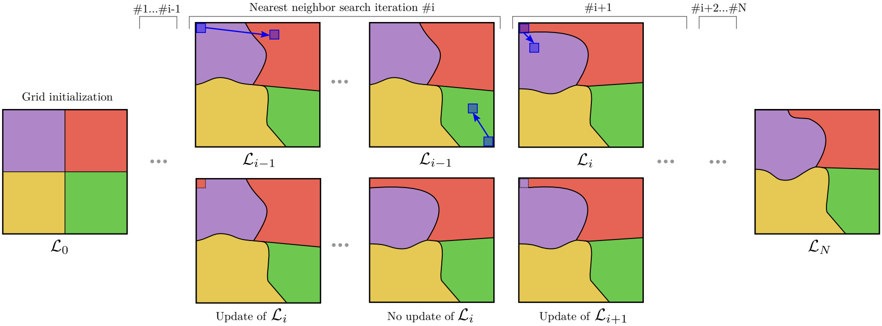
**Distance between a pixel $p$ and a superpixel $S_k$ (TASP):**

$$D(p, S_k) = d_{\text{color}}(F_p, F_{S_k}) + d_{\text{spatial}}(X_p, X_{S_k}) m_k + d_{\text{texture}}(p, S_k)$$

$\rightarrow$ Complexity $C_{\text{TASP}} = \mathcal{O}((h \times w) \times 4 \times \text{Iter}_{\text{K-means}} \times \text{Iter}_{\text{NN}})$

- NNSC: Nearest Neighbor-based Superpixel Clustering

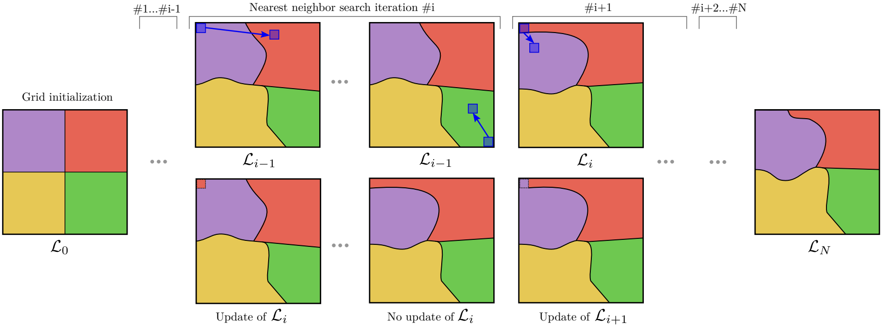
  Direct pixel label update using local NN search



# The proposed NNSC approach

- NNSC: Nearest Neighbor-based Superpixel Clustering

  Direct pixel label update using local NN search

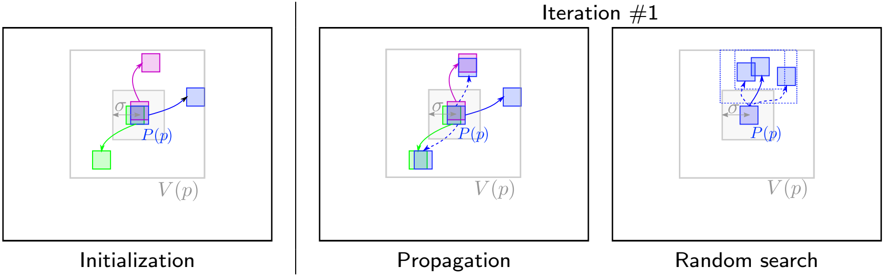



→ Complexity reduced to $C_{\text{NNSC}} = \mathcal{O}((h \times w) \times \text{Iter}_{\text{NN}})$

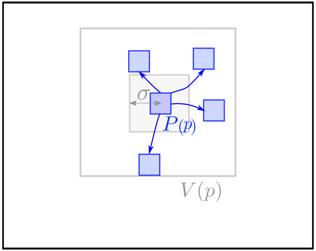
- NNSC: Nearest Neighbor-based Superpixel Clustering

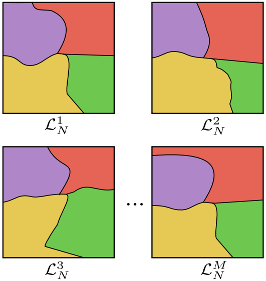
  Direct pixel label update using local NN search

  Constrained PatchMatch (PM) [Barnes et al., 2009] algorithm:



- Aggregation of multiple clustering estimations from independent PM processes



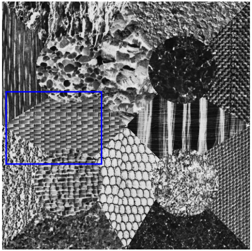



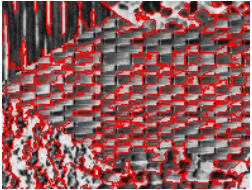

**Aggregation of $M$ label maps:**

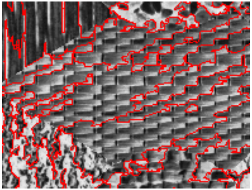
$$\mathcal{L}_{\text{final}}(p) = \underset{l \in \{labels\}}{\text{argmax}} \sum_{i=1}^{M} \delta_{\mathcal{L}_N^i(p),l}$$

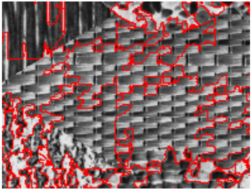
→ Improve the robustness of the clustering

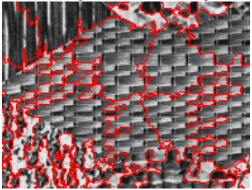
On a composite natural texture image:

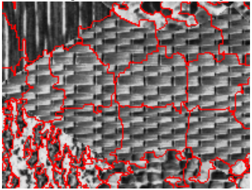

Initial image


LSC [Chen et al., 2017]


SNIC [Achanta et al., 2017]

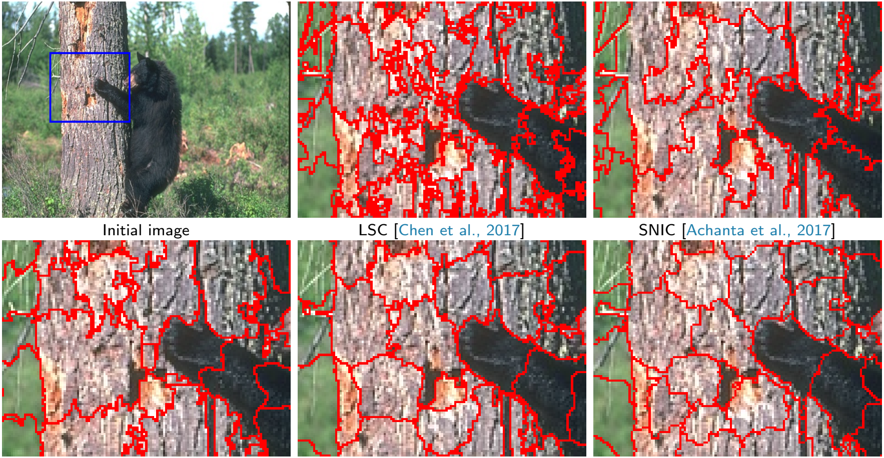

SCALP [Giraud et al., 2018]


TASP [Giraud et al., 2019]


NNSC

CTI99: dataset of $10$ images containing up to 16 different textures [Randen and Husoy, 1999]

On a natural color image:



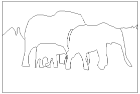
Initial image | LSC [Chen et al., 2017] | SNIC [Achanta et al., 2017]

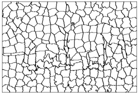
SCALP [Giraud et al., 2018] | TASP [Giraud et al., 2019] | NNSC

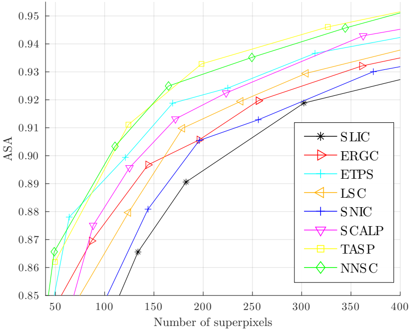
BSD: dataset of $200$ natural color images of size $321 \times 481$ [Martin et al., 2001]

Standard ASA metric:
Superposition with image objects

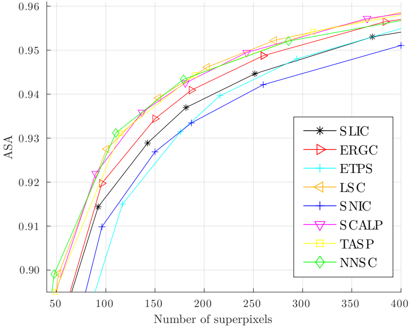

Image


Manual segmentation


Superpixels



Texture dataset (CTI99)



Natural color dataset (BSD)

→ Best performances on the two data types with the same parameters

→ Computational time from ≈ 60s for TASP → ≈ 1.5s for proposed NNSC

## Summary of contributions

- New superpixel method robust to texture
- Faster direct patch-based nearest neighbor framework
- Accurate results on both texture and natural color datasets

## Work in progress / Research perspectives

- Use of advanced texture descriptors
- Application to real data (3D medical, satellite, etc.)

Texture Superpixel Clustering from Patch-based Nearest Neighbor Matching

# Thank you for your attention

Check for source codes at

http://rgiraud.vvv.enseirb-matmeca.fr

Bordeaux INP ENSEIRB MATMECA

ims

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S. (2012). SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)*, 34(11):2274–2282.

Achanta et al., R. (2017). Superpixels and polygons using simple non-iterative clustering. In *Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on*, pages 4895–4904.

Barnes, C., Shechtman, E., Finkelstein, A., and Goldman, D. B. (2009). PatchMatch: A randomized correspondence algorithm for structural image editing. *ACM Trans. on Graphics (ToG)*, 28(3).

Buyssens, P., Gardin, I., Ruan, S., and Elmoataz, A. (2014). Eikonal-based region growing for efficient clustering. *Image and Vision Computing*, 32(12):1045–1054.

Chen, J., Li, Z., and Huang, B. (2017). Linear spectral clustering superpixel. *IEEE Trans. on Image Processing (TIP)*, 26(7):3317–3330.

Giraud, R. and Berthoumieu, Y. (2019). Texture Superpixel Clustering from Patch-based Nearest Neighbor Matching. In *European Signal Processing Conference (EUSIPCO 2019)*.

Giraud, R., Ta, V.-T., and Papadakis, N. (2018). Robust superpixels using color and contour features along linear path. *Computer Vision and Image Understanding (CVIU)*, 170:1–13.

Giraud, R., Ta, V.-T., Papadakis, N., and Berthoumieu, Y. (2019). Texture-Aware Superpixel Segmentation. In *IEEE International Conference on Image Processing*.

Martin, D., Fowlkes, C., Tal, D., and Malik, J. (2001). A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proc. of IEEE International Conference on Computer Vision (ICCV)*, volume 2, pages 416–423.

Randen, T. and Husoy, J. H. (1999). Hierarchical image segmentation via recursive superpixel with adaptive regularity. *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)*, 21:291–310.

Ren, X. and Malik, J. (2003). Learning a classification model for segmentation. In *Proc. of IEEE International Conference on Computer Vision (ICCV)*, pages 10–17.

Yao, J., Boben, M., Fidler, S., and Urtasun, R. (2015). Real-time coarse-to-fine topologically preserving segmentation. In *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 2947–2955.

Distance between a pixel $p$ and a superpixel $S_k$:

$$D(p, S_k) = d_{\text{color}}(F_p, F_{S_k}) + d_{\text{spatial}}(X_p, X_{S_k})m$$

Limitations:

- Global regularity parameter $\rightarrow$ irregular borders with low $m$ / inaccurate borders with high $m$.
- Only local pixel color considered $\rightarrow$ not robust to texture.

$m = 10$

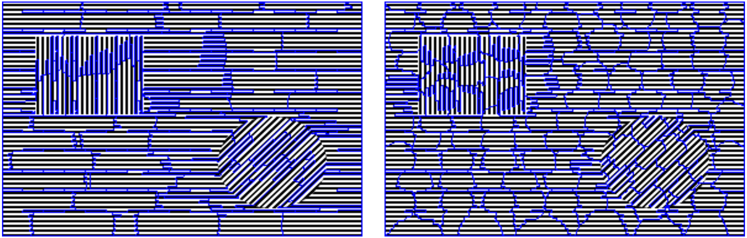
$m = 60$

Distance between a pixel $p$ and a superpixel $S_k$:

$$D(p, S_k) = d_{\text{color}}(F_p, F_{S_k}) + d_{\text{spatial}}(X_p, X_{S_k})m$$

Limitations:

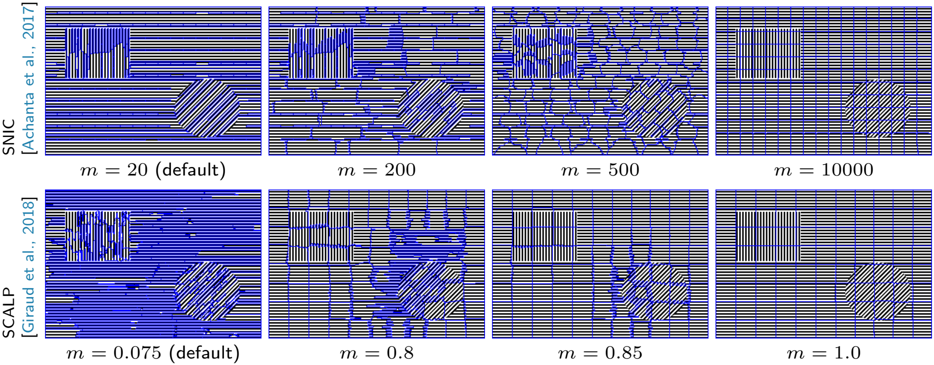
- Global regularity parameter $\rightarrow$ irregular borders with low $m$ / inaccurate borders with high $m$.
- Only local pixel color considered $\rightarrow$ not robust to texture.



$m = 200$ $m = 500$

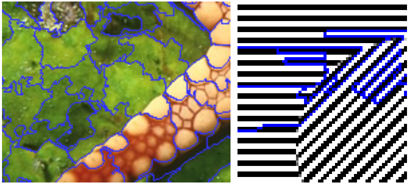
# Robustness of state-of-the-art methods

What about textured images?



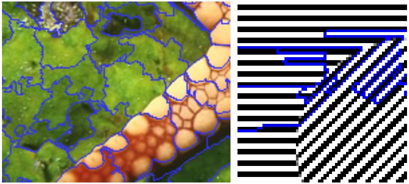

→ Even with manual regularity tuning, no explicit consideration of texture information

- Automatic adaptation of the regularity parameter:



SLIC [Achanta et al., 2012]

- Automatic adaptation of the regularity parameter:



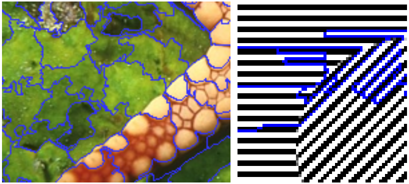
SLIC [Achanta et al., 2012]

**Ponderation with feature variance within superpixels:**

$$m_k = m \exp\left(\frac{\sigma(F_{p \in S_k})}{\beta}\right)$$

# The proposed NNSC approach

- Automatic adaptation of the regularity parameter:



SLIC [Achanta et al., 2012]

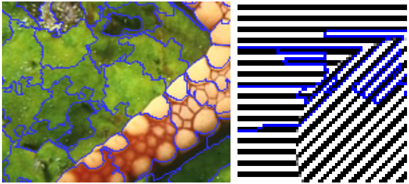
**Ponderation with feature variance within superpixels:**

$$m_k = m \exp\left(\frac{\sigma(F_{p\in S_k})}{\beta}\right)$$

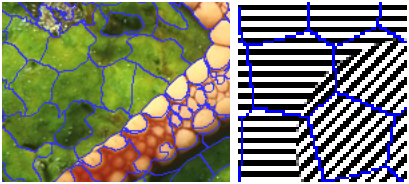
**SLIC clustering distance [Achanta et al., 2012]:**

$$D(p, S_k) = d_{\text{color}}(F_p, F_{S_k}) + d_{\text{spatial}}(X_p, X_{S_k})m$$

- Automatic adaptation of the regularity parameter:



SLIC [Achanta et al., 2012]



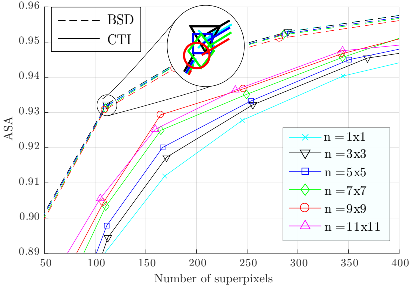
TASP

**Ponderation with feature variance within superpixels:**

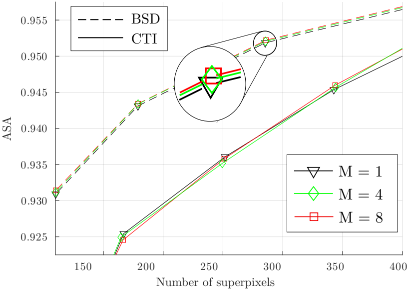
$$m_k = m \exp\left(\frac{\sigma(F_{p \in S_k})}{\beta}\right)$$

**TASP clustering distance:**

$$D(p, S_k) = d_{\text{color}}(F_p, F_{S_k}) + d_{\text{spatial}}(X_p, X_{S_k}) m_k$$




Patch-size n




Number of label map estimations M