Texture-Aware Superpixel Segmentation

Rémi Giraud
Bordeaux INP
IMS

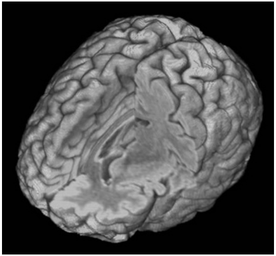
Vinh-Thong Ta
Bordeaux INP
LaBRI

Nicolas Papadakis
CNRS
IMB

Yannick Berthoumieu
Bordeaux INP
IMS

Large data $\rightarrow$ high computational times

Image HD



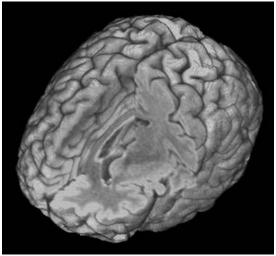
Volume 3D

Video

Introduction

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

Image HD

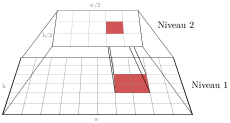


Volume 3D

Video

Large data → high computational times → Dimension reduction

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



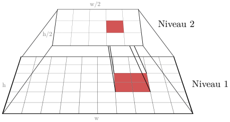

Image


Decomposition into blocks


Average colors

Large data → high computational times → Dimension reduction

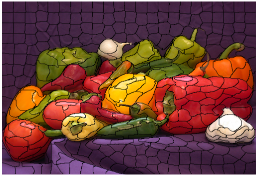
- Regular multi-resolution:
  Decompose the image into regular blocks
- Superpixels (since [Ren and Malik, 2003]):
  Local grouping of pixels with homogeneous colors




Image


Decomposition into blocks


Average colors


Decomposition into superpixels

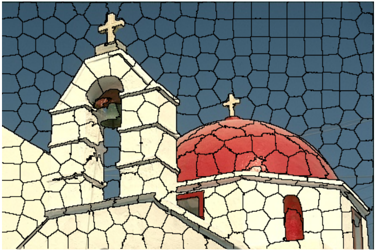

Average colors

Desired properties of superpixel methods:

- Relatively fast to compute ✓
- Limited parameter settings ✓

Desired properties of superpixel methods:

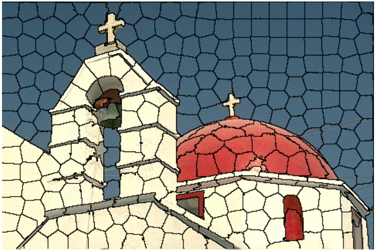
- Relatively fast to compute ✓
- Limited parameter settings ✓
- Both accurate and regular superpixels ∼



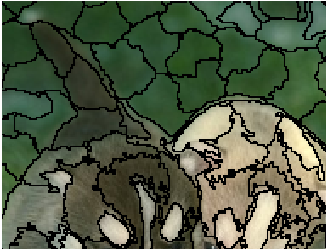
[Chen et al., 2017]

Desired properties of superpixel methods:

- Relatively fast to compute ✓
- Limited parameter settings ✓
- Both accurate and regular superpixels ∼



[Chen et al., 2017]

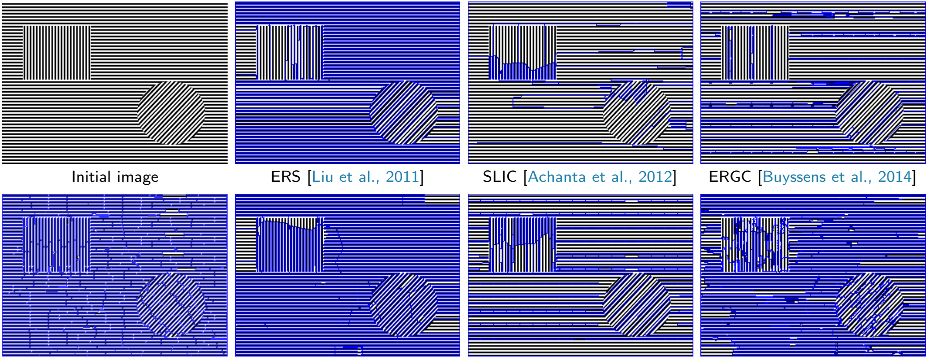


[Chen et al., 2017]

→ Irregular borders on textured regions

Robustness of state-of-the-art methods

What about textured images?



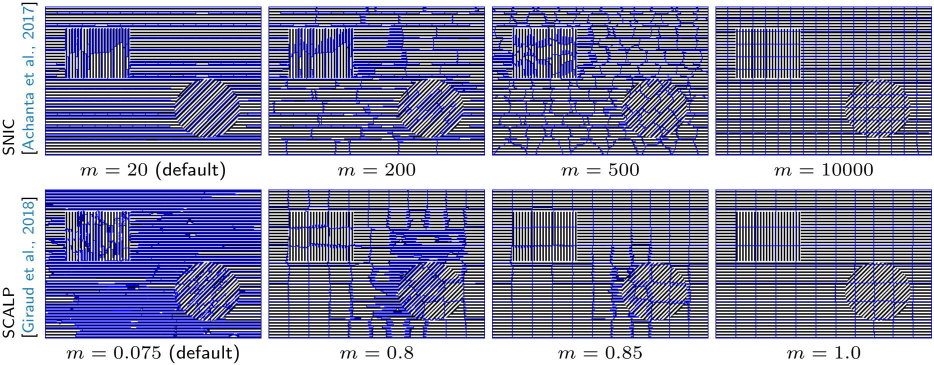
Initial image | ERS [Liu et al., 2011] | SLIC [Achanta et al., 2012] | ERGC [Buyssens et al., 2014]

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

Robustness of state-of-the-art methods

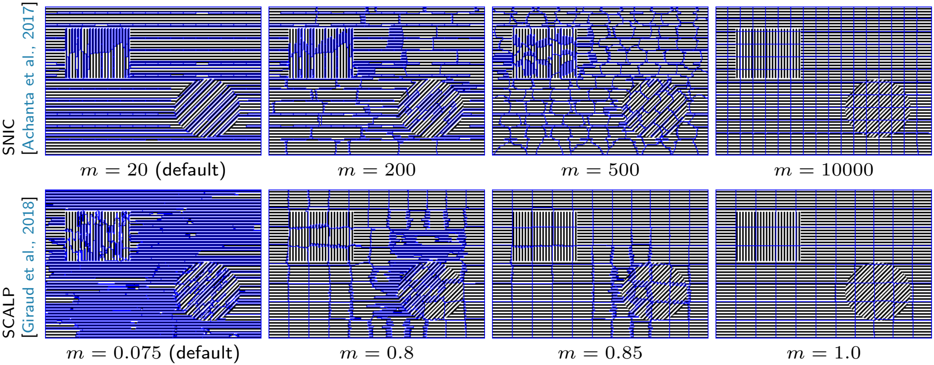
What about textured images?



$\rightarrow$ Even with manual regularity tuning, no explicit consideration of texture information

Robustness of state-of-the-art methods

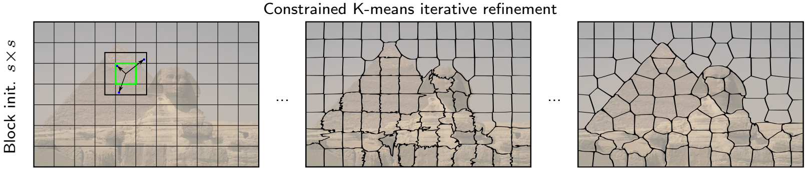
What about textured images?



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

→ TASP: Texture-Aware SuperPixel segmentation method

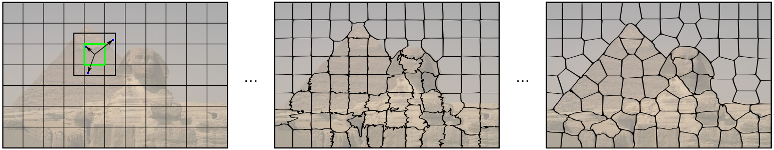
Simple Linear Iterative Clustering (SLIC) [Achanta et al., 2012]



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

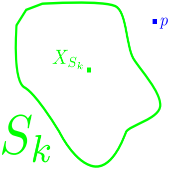
Constrained K-means iterative refinement

Block init. $s \times s$



Distance between a pixel p and a superpixel S_k:

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



$F_p = [l_p, a_p, b_p]$ 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

The SLIC method

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.

$m = 10$

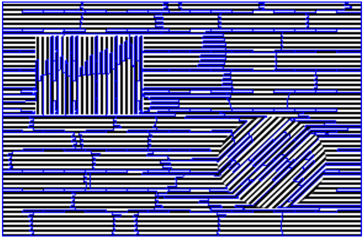
$m = 60$

Distance between a pixel p and a superpixel S_k:

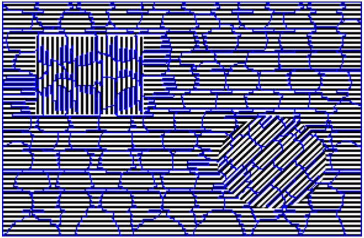
$$D(p, S_k) = d_{\text{color}}(F_p, F_{S_k}) + d_{\text{spatial}}(X_p, X_{S_k})m$$

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.

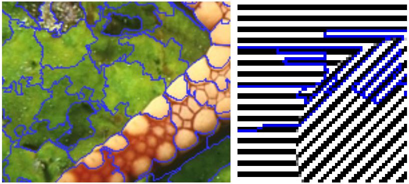


$m = 200$



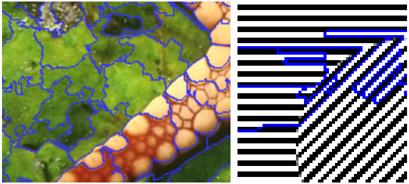
$m = 500$

- Automatic adaptation of the regularity parameter:



SLIC [Achanta et al., 2012]

- Automatic adaptation of the regularity parameter:

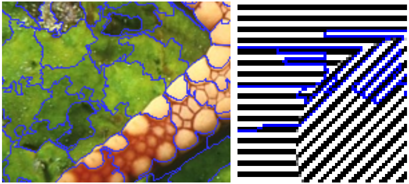


SLIC [Achanta et al., 2012]

Ponderation with feature variance within superpixels:

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

- Automatic adaptation of the regularity parameter:



SLIC [Achanta et al., 2012]

Ponderation with feature variance within superpixels:

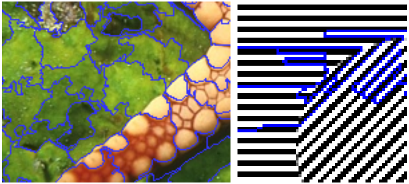
$$m_k = m \exp\left(\frac{\sigma(F_{p\in S_k})}{\beta}\right)$$

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

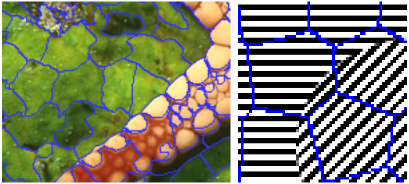
$$D(p, S_k) = d_{\text{color}}(F_p, F_{S_k}) + d_{\text{spatial}}(X_p, X_{S_k})m$$

The TASP method

- 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:

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

- Pixel to superpixel texture homogeneity term:

  $\rightarrow$ *Bench of filters?*

  Prior definition of filters
  Cannot be precisely averaged over a superpixel

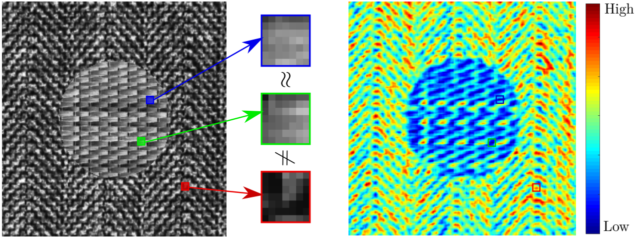
- Pixel to superpixel texture homogeneity term:

  → *Bench of filters?*

  Prior definition of filters
  Cannot be precisely averaged over a superpixel

  → *Patch-based distance?*

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



- Pixel to superpixel texture homogeneity term:

  Which patches to compare?

  → Patch on the superpixel barycenter?

  Not representative of the texture content

- Pixel to superpixel texture homogeneity term:

  Which patches to compare?

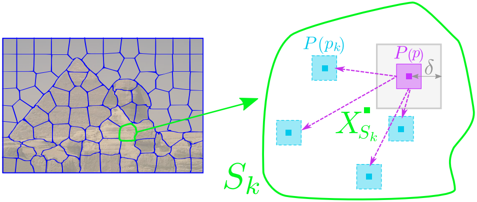
  → Patch on the superpixel barycenter?

  Not representative of the texture content

  → 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:

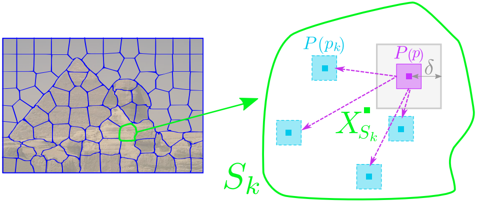
  Which patches to compare?

  → Patch on the superpixel barycenter?

  Not representative of the texture content

  → Nearest neighbor (NN) matching within the superpixel?

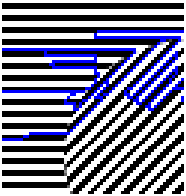
  Ability to find only similar texture patterns
  Fast selection of N similar patches with PatchMatch [Barnes et al., 2009]



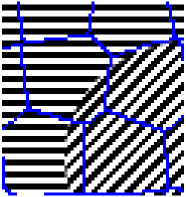
Texture homogeneity term:

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

- Pixel to superpixel texture homogeneity term:



SLIC [Achanta et al., 2012]

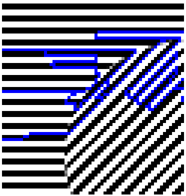


TASP w/ d_{texture}

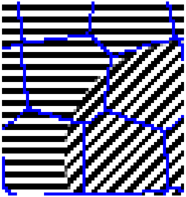
d_{texture} does not guarantee texture unicity within superpixels

The TASP method

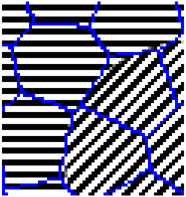
- Pixel to superpixel texture homogeneity term:



SLIC [Achanta et al., 2012]



TASP w/ d_{texture}



TASP w/ $d_{\text{texture}} + d_{\text{unicity}}$

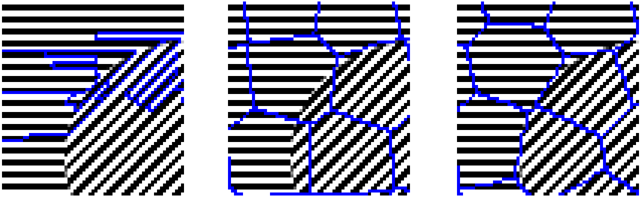
d_{texture} does not guarantee texture unicity within superpixels

$\rightarrow$ d_{unicity} forces the selection of patches p_k close to the superpixel barycenter:

Spatial distance on selected patches:

$$d_{\text{unicity}}(p, S_k) = 2.\frac{1}{N}\sum_{p_k \in \mathcal{K}_p}\left(1 - \exp\left(-\frac{\|X_{p_k} - X_{S_k}\|_2^2}{s^2}\right)\right)$$

- Pixel to superpixel texture homogeneity term:



SLIC [Achanta et al., 2012] — TASP w/ d_{texture} — TASP w/ $d_{\text{texture}} + d_{\text{unicity}}$

d_{texture} does not guarantee texture unicity within superpixels

$\rightarrow$ d_{unicity} forces the selection of patches p_k close to the superpixel barycenter:

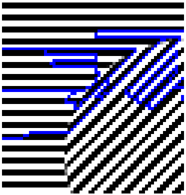
Spatial distance on selected patches:

$$d_{\text{unicity}}(p, S_k) = 2.\frac{1}{N}\sum_{p_k \in \mathcal{K}_p}\left(1 - \exp\left(-\frac{\|X_{p_k} - X_{S_k}\|_2^2}{s^2}\right)\right)$$

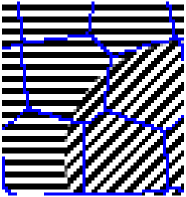
SLIC clustering distance [Achanta et al., 2012]:

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

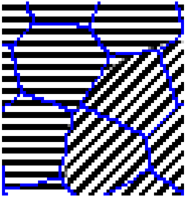
- Pixel to superpixel texture homogeneity term:



SLIC [Achanta et al., 2012]



TASP w/ d_{texture}



TASP w/ $d_{\text{texture}} + d_{\text{unicity}}$

d_{texture} does not guarantee texture unicity within superpixels

$\rightarrow$ d_{unicity} forces the selection of patches p_k close to the superpixel barycenter:

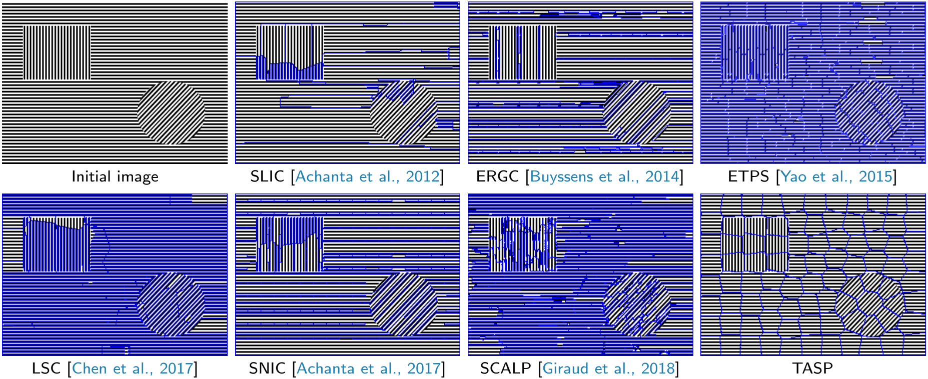
Spatial distance on selected patches:

$$d_{\text{unicity}}(p, S_k) = 2.\frac{1}{N}\sum_{p_k \in \mathcal{K}_p}\left(1 - \exp\left(-\frac{\|X_{p_k} - X_{S_k}\|_2^2}{s^2}\right)\right)$$

Final TASP clustering distance:

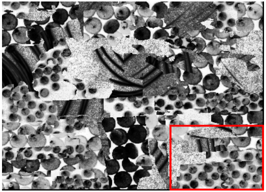
$$D(p, S_k) = d_{\text{color}}(F_p, F_{S_k}) + d_{\text{spatial}}(X_p, X_{S_k})m_k + d_{\text{texture}}(p, S_k) + d_{\text{unicity}}(p, S_k)m_k$$

On a very textured synthetic image:

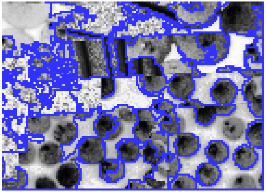


mix-Stripes: dataset of 10 images of size 300×400 with synthetic stripe textures

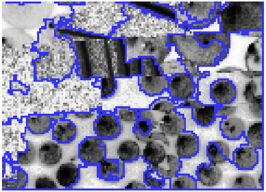
On a composite natural texture image:



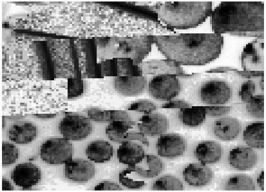
Initial image



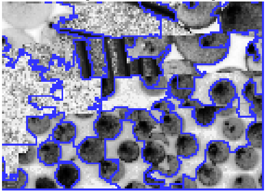
LSC [Chen et al., 2017]



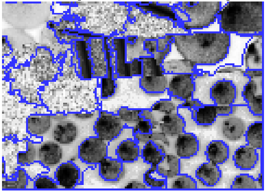
SNIC [Achanta et al., 2017]



Initial image



SCALP [Giraud et al., 2018]



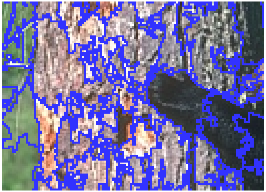
TASP

mix-Brodatz: dataset of 100 images of size 300×400 with natural textures [Brodatz, 1966]

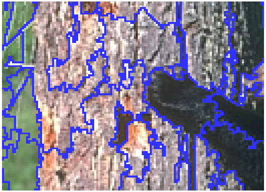
Results - Qualitative comparison to state-of-the-art

On a natural color image:

Initial image

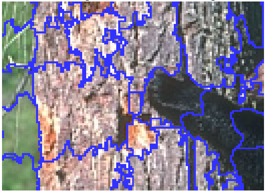


LSC [Chen et al., 2017]

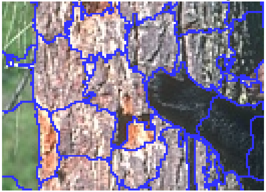


SNIC [Achanta et al., 2017]

Initial image



SCALP [Giraud et al., 2018]



TASP

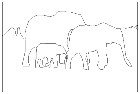
BSD: dataset of 200 natural color images of size 321×481 [Martin et al., 2001]

Results - Quantitative comparison to state-of-the-art

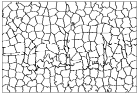
Standard metrics:

- Superposition with several objects: ASA
- Contour detection: F-measure

Image



Manual segmentation



Superpixels

	mix-Stripes (synthetic textures)		*mix-Brodatz* (natural textures)		*BSD* (natural color)	
Method	ASA	F	ASA	F	ASA	F
SLIC [Achanta et al., 2012]	0.7256	0.4048	0.7784	0.4607	0.9445	0.4706
ERGC [Buyssens et al., 2014]	0.6107	0.3717	0.7796	0.4677	0.9477	0.4571
ETPS [Yao et al., 2015]	0.7769	0.2953	0.7568	0.4354	0.9433	0.4710
LSC [Chen et al., 2017]	0.6979	0.3156	0.7908	0.4552	0.9503	0.4421
SNIC [Achanta et al., 2017]	0.6659	0.3529	0.7662	0.4815	0.9410	0.4617
SCALP [Giraud et al., 2018]	0.7307	0.3290	0.7977	0.4759	0.9499	0.4914
TASP	**0.8706**	**0.4232**	**0.8139**	**0.4824**	**0.9503**	**0.4992**

$\rightarrow$ Best performances on the three data types with the same parameters

Summary of contributions

- Superpixel method robust to texture
- Generic patch-based texture homogeneity term
- No need for manual regularity setting
- Accurate results on both texture and natural color datasets

Work in progress / Research perspectives

- Improvement of computational time (EUSIPCO 2019)
- Use of advanced texture descriptors
- Application to real data (3D medical, satellite, etc.)

Thank you for your attention

Check for source codes at

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

Bordeaux INP
ENSEIRB
MATMECA

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).

Brodatz, P. (1966). Textures: A photographic album for artists and designers. 1966. *Dover Publications*.

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.

References II

Liu, .-Y., Tuzel, O., Ramalingam, S., and Chellappa, R. (2011). Entropy rate superpixel segmentation. In *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 2097–2104.

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.

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.