Texture Superpixel Clustering from Patch-based Nearest Neighbor Matching

Rémi Giraud  Yannick Berthoumieu
Large data $\rightarrow$ high computational times
Large data $\rightarrow$ high computational times $\rightarrow$ Dimension reduction
Large data $\rightarrow$ high computational times $\rightarrow$ Dimension reduction

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

![Image](image.png)

- **Image**
- **Decomposition into blocks**
- **Average colors**
Introduction

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

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

- Superpixels (since [Ren and Malik, 2003]):
  Local grouping of pixels with homogeneous colors
Robustness of state-of-the-art methods

What about the ability to handle texture?

Initial image  SLIC \cite{Achanta2012}  ERGC \cite{Buyssens2014}  ETPS \cite{Yao2015}

LSC \cite{Chen2017}  SNIC \cite{Achanta2017}  SCALP \cite{Giraud2018}

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

What about the ability to handle texture?

→ 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
Robustness of state-of-the-art methods

What about the ability to handle texture?

- 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)
1 Introduction

2 K-means-based Clustering Approach (TASP)

3 Proposed Nearest-Neighbor Superpixel Clustering (NNSC) approach

4 Results

5 Conclusion
K-means-based Clustering Approach (TASP)

- Simple Linear Iterative Clustering (SLIC) [Achanta et al., 2012]
K-means-based Clustering Approach (TASP)

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

Distance between a pixel \( p \) and a superpixel \( S_k \):

\[
D(p, S_k) = d_{\text{color}}(F_p, F_{S_k}) + d_{\text{spatial}}(X_p, X_{S_k})m_k
\]
K-means-based Clustering Approach (TASP)

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

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

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

→ Complexity $C_{\text{SLIC}} = \mathcal{O}((h \times w) \times 4 \times \text{Iter}_K\text{-means})$
K-means-based Clustering Approach (TASP)

- 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?
K-means-based Clustering Approach (TASP)

- 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:

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

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

Constrained K-means iterative refinement

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

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

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

→ Complexity \( C_{\text{SLIC}} = \mathcal{O}((h \times w) \times 4 \times \text{Iter}_{K\text{-means}}) \)
K-means Clustering Approach

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

**Constrained K-means iterative refinement**

Block init. $s \times s$

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

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

→ Complexity $C_{\text{TASP}} = \mathcal{O}((h \times w) \times 4 \times \text{Iter}_{\text{K-means}} \times \text{Iter}_{\text{NN}})$
The proposed NNSC approach

- NNSC: Nearest Neighbor-based Superpixel Clustering

Direct pixel label update using local NN search

Grid initialization

$\mathcal{L}_0$

$\mathcal{L}_{i-1}$

No update of $\mathcal{L}_i$

Update of $\mathcal{L}_{i+1}$

Update of $\mathcal{L}_i$

Update of $\mathcal{L}_i$...

No update of $\mathcal{L}_i$...

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

- **NNSC**: Nearest Neighbor-based Superpixel Clustering
  
  Direct pixel label update using local NN search

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

- **NNSC**: Nearest Neighbor-based Superpixel Clustering
  - Direct pixel label update using local NN search

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

Initialization | Propagation | Random search

**Iteration #1**

$V(p)$

$P(p)$
The proposed NNSC approach

- Aggregation of multiple clustering estimations from independent PM processes

Aggregation of $M$ label maps:

$$\mathcal{L}_{\text{final}}(p) = \arg\max_{l \in \{\text{labels}\}} \sum_{i=1}^{M} \delta \mathcal{L}_N^i(p),l$$

$\rightarrow$ Improve the robustness of the clustering
Results - Qualitative comparison to state-of-the-art

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]
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]

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

Standard ASA metric:
Superposition with image objects

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

→ Computational time from $\approx 60s$ for TASP $\rightarrow \approx 1.5s$ 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


K-means Clustering Framework

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.
K-means Clustering Framework

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?

SNIC [Achanta et al., 2017]

- $m = 20$ (default)
- $m = 200$
- $m = 500$
- $m = 10000$

SCALP [Giraud et al., 2018]

- $m = 0.075$ (default)
- $m = 0.8$
- $m = 0.85$
- $m = 1.0$

→ Even with manual regularity tuning, no explicit consideration of texture information
The proposed NNSC approach

- Automatic adaptation of the regularity parameter:

SLIC [Achanta et al., 2012]
The proposed NNSC approach

- Automatic adaptation of the regularity parameter:

\[ m_k = m \exp \left( \frac{\sigma(F_p \in S_k)}{\beta} \right) \]
The proposed NNSC approach

- Automatic adaptation of the regularity parameter:

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

SLIC [Achanta et al., 2012]

Ponderation with feature variance within superpixels:

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

SLIC clustering distance [Achanta et al., 2012]:
The proposed NNSC approach

- Automatic adaptation of the regularity parameter:

Ponderation with feature variance within superpixels:

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

TASP clustering distance:

\[ D(p, S_k) = d_{\text{color}}(F_p, F_{S_k}) + d_{\text{spatial}}(X_p, X_{S_k})m_k \]
The proposed NNSC approach - Impact of parameters

**Patch-size n**

**Number of label map estimations M**