HAL
open science

Superpixels adaptés localement aux textures

Rémi Giraud, Vinh-Thong Ta, Nicolas Papadakis, Yannick Berthoumieu

To cite this version:

Rémi Giraud, Vinh-Thong Ta, Nicolas Papadakis, Yannick Berthoumieu. Superpixels adaptés localement aux textures. Colloque du Groupe de Recherche et d'Etudes de Traitement du Signal et des Images (GRETSI'19), Aug 2019, Lille, France. hal-02117558

HAL Id: hal-02117558

https://hal.science/hal-02117558

Submitted on 27 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Superpixels adaptés localement aux textures

Rémi GIRAUD[1], Vinh-Thong TA[2], Nicolas PAPADAKIS[3], Yannick BERTHOUMIEU[1]

[1]Bordeaux INP, Univ. Bordeaux, CNRS, IMS, UMR 5218, F-33400 Talence, France.
[2]Bordeaux INP, Univ. Bordeaux, CNRS, LaBRI, UMR 5800, F-33400 Talence, France.
[3]CNRS, Univ. Bordeaux, IMB, UMR 5251, F-33400 Talence, France.
remi.giraud@ims-bordeaux.fr, vinh-thong.ta@labri.fr,
nicolas.papadakis@math.u-bordeaux.fr, yannick.berthoumieu@ims-bordeaux.fr

Résumé – La plupart des méthodes superpixels calculent un compromis entre des descripteurs spatiaux et couleur à l'échelle pixellique. Elles nécessitent donc un réglage fin pour équilibrer ces mesures, et ne peuvent capturer une information de texture. Dans ce travail, nous répondons à ces problèmes avec une nouvelle méthode robuste aux textures. Pour capturer les zones texturées et homogènes, la contrainte spatiale est ajustée automatiquement en fonction de la variance locale. Ensuite, pour assurer l'homogénéité de texture des superpixels, une nouvelle distance basée patchs est introduite. La méthode proposée améliore la précision de celles de l'état de l'art sur des bases d'images texturées et naturelles couleur.

Abstract – Most superpixel methods compute a trade-off between spatial and color features at the pixel level. Hence, they need fine parameter tuning to balance the two measures, and fail to capture a texture information. In this work, we address these limitations with a new method that is robust to textures. To capture both textured and smooth areas, the spatial constraint is automatically adjusted according to the local feature variance. Then, to ensure the texture homogeneity within superpixels, a new patch-based distance is introduced. The proposed method outperforms the accuracy of the state-of-the-art ones on texture and natural color image datasets.

1 Introduction

L'objectif des superpixels est d'exploiter la redondance locale d'information pour réduire la charge de calcul et améliorer potentiellement les performances en réduisant le bruit d'un traitement à l'échelle pixellique. Les pixels sont localement groupés en régions, comme pré-traitement pour des applications telles que l'estimation de saillance visuelle [7], la classification [8], et la détection [17] ou le suivi d'objets [14].

La plupart des méthodes produisent des superpixels de tailles équivalentes contenant des pixels de couleurs homogènes, puisque les contours des objets d'une image sont souvent détectables par des changements de couleur. Les méthodes de l'état de l'art telles que [11, 16, 1, 5, 2], sont donc uniquement basées sur des espaces spatiaux et couleur à l'échelle pixellique. Dans [12, 6], des descripteurs couleur plus avancés sont définis. Plus récemment, [10] propose d'utiliser un a priori de contour dans la décomposition afin de garantir le respect des frontières des objets, au détriment d'une complexité supérieure.

Des approches basées apprentissage ont également été proposées, par ex., [15]. Malgré de potentielles bonnes performances, ces méthodes ne permettent pas d'influer sur la régularité, qui peut impacter les résultats d'applications basées superpixels [9], et présentent les inconvénients standards de ces approches en termes de ressources, de temps d'apprentissage, et d'applicabilité. Il est donc nécessaire d'augmenter la robustesse des méthodes superpixels sans apprentissage vis-à-vis des textures, tout en préservant leurs propriétés souhaitées : adaptabilité, faible complexité et paramètres limités.

Contributions. Dans ce travail, nous proposons une nouvelle méthode de Superpixels Adaptés Localement aux Textures (SALT) permettant de décomposer des images très texturées, comme naturelles couleur, avec les mêmes paramètres.

La méthode proposée, SALT, est comparée dans la Figure 1 aux approches de l'état de l'art sur une image texturée synthétique. SALT produit une segmentation pertinente, quand toutes les autres méthodes sont fortement induites en erreur par les textures, puisqu'elles se basent uniquement sur l'information du pixel ou ne capturent pas explicitement la texture dans leur modèle. À noter que nous avons utilisé les paramètres de régularité recommandés par les auteurs pour [1, 5, 18, 6, 10], mais aucun autre paramétrage n'aurait permis de capturer l'information de texture avec ces méthodes.

Pour pouvoir générer des superpixels pertinents sur des images texturées (cf. Figure 1), SALT ajuste sa contrainte spatiale en fonction de la variance des descripteurs dans le superpixel. De cette façon, SALT répond également au besoin d'un réglage manuel fin de la régularité. Ensuite, pour assurer la capture des textures, nous introduisons une nouvelle distance basée patchs permettant d'évaluer facilement la similarité d'un voisinage de pixels avec un superpixel.

Nous validons SALT sur la base d'images naturelles couleur standard [13], ainsi que sur deux nouvelles bases proposées pour évaluer les performances de segmentation de texture. SALT surpasse significativement les méthodes de l'état de l'art en termes de segmentation de texture, tout en obtenant des performances comparables voire supérieures sur les images naturelles, et ce, en utilisant les mêmes paramètres.

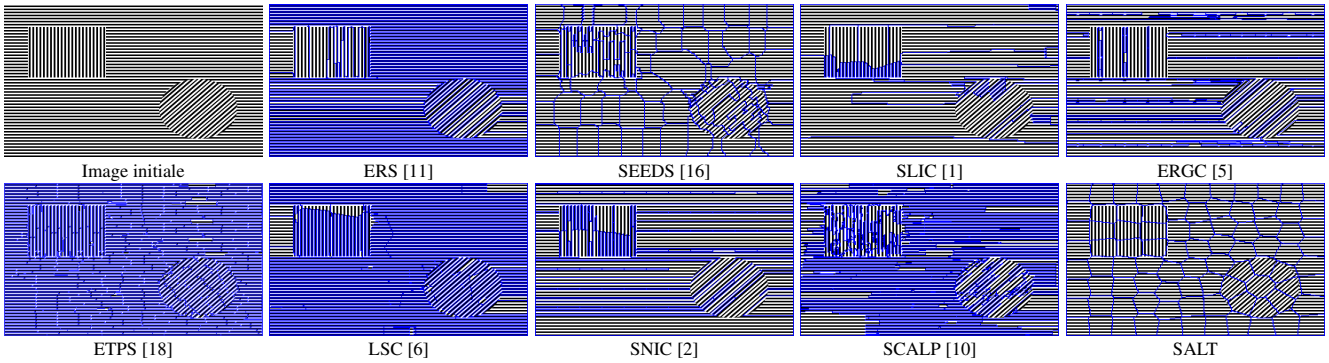


FIGURE 1 – Comparaison de SALT aux approches de l'état de l'art sur une image de texture synthétique pour 100 superpixels. SALT réussit à capturer les textures quand les autres méthodes sont extrêmement induites en erreur par les variations de contraste

2 Superpixels robustes aux textures

La méthode SALT améliore l'approche itérative utilisée dans [1, 6, 10], d'abord présentée dans cette section. Ensuite, nous proposons d'ajuster localement la régularité des superpixels, afin qu'ils s'adaptent automatiquement au contenu de l'image. Enfin, nous introduisons une distance basée patchs pixel-superpixel pour mesurer l'homogénéité de texture.

2.1 Algorithme de groupement itératif

La méthode [1] requiert uniquement le nombre voulu de superpixels et un paramètre de régularité. Elle est basée sur un groupement itératif de type K-moyennes. Les superpixels S_i sont d'abord répartis régulièrement en blocs de taille $s \times s$, et décrits par un descripteur d'intensité moyen F_{S_i} (couleurs CIELab pour [1]) et spatial avec le barycentre $X_{S_i} = [x_i, y_i]$ des pixels dans S_i. À chaque itération, chaque superpixel S_i est comparé à tous les pixels $p = [F_p, X_p]$, dans une zone de $(2s+1)\times(2s+1)$ pixels autour de son barycentre X_{S_i}. Un pixel p est associé au superpixel S_i qui minimise la distance D :

$$D(p, S_i) = d_F(F_p, F_{S_i}) + d_s(X_p, X_{S_i})\frac{m^2}{s^2}, \qquad (1)$$

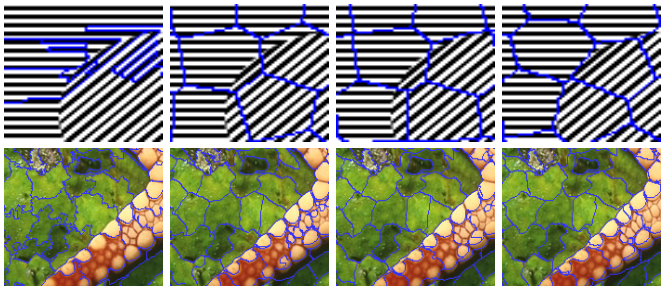
avec m le paramètre de contrainte spatiale de régularité de forme, $d_F(F_p, F_{S_i})$=$\|F_p - F_{S_i}\|_2^2$, la distance sur les descripteurs, et $d_s(X_p, X_{S_i})$=$\|X_p - X_{S_i}\|_2^2$ la distance spatiale. Enfin, un post-traitement assure la connectivité des superpixels.

Bien que cette méthode puisse grouper des pixels aux couleurs similaires, m est défini globalement et ne peut pas s'adapter localement à tous les contours de l'image. De plus, elle ne permet pas de capturer les motifs de texture, car elle ne prend en compte que l'information à l'échelle du pixel.

2.2 Adaptation locale de la régularité

La plupart des méthodes, dont [1, 6, 10], doivent régler manuellement le paramètre de régularité m selon la dynamique de la distance sur les descripteurs d_F. Ainsi, un paramètre choisi pour des images couleur naturelles peut mener (1) à décomposer très irrégulièrement des textures (voir Figure 1). Nous répondons à ce problème, en utilisant pour chaque superpixel S_i, un paramètre de régularité m_i défini selon la variance des descripteurs $\sigma_i(F_p)$ de tous les pixels $p \in S_i$:

$$m_i = m \exp\Big(\sigma_i(F_p)/\beta\Big), \qquad (2)$$



(a) SLIC [1] (b) SALT (2) (c) SALT (2), (3) (d) SALT (2), (3), (4)

FIGURE 2 – SLIC [1] à une régularité optimale pour des images couleur (a) vs SALT (b)-(d), qui décompose efficacement des images texturées et couleur avec les mêmes paramètres (d)

avec un paramètre d'échelle β. Le terme m_i permet d'augmenter la contrainte spatiale dans le modèle de SALT (5) pour des superpixels avec une forte variance, et de la réduire dans des zones lisses, pour que le superpixel capture les frontières des objets de l'image qui seraient difficilement perceptibles.

Ainsi, sans adapter manuellement m dans (2), SALT peut calculer des superpixels pertinents sur des images fortement texturées (Figure 2(b)-haut), et des images couleur naturelles (Figure 2(b)-bas), puisque m_i (2) ajuste automatiquement le compromis entre d_F et d_s dans (5). Néanmoins, la méthode doit toujours être améliorée pour capturer l'information de texture.

2.3 Mesure d'homogénéité de texture

Distance basée patchs pixel-superpixel. Dans cette section, nous proposons une méthode pour mesurer la similarité de texture entre un voisinage de pixels et le contenu d'un superpixel, donc entre deux régions de tailles différentes. Un descripteur de texture à l'échelle du patch et du superpixel augmenterait la complexité et entraînerait des paramètres supplémentaires dans le modèle. De plus, la texture ne peut pas être aussi bien préservée que pour l'information couleur et spatiale avec une moyenne sur le superpixel. La méthode doit conserver une complexité limitée et s'adapter à tout type d'image sans utiliser d'a priori. De telles contraintes empêchent également d'utiliser des approches coûteuses basées-dictionnaires ou apprentissage.

Pour répondre à ces contraintes, nous proposons une nouvelle méthode basée sur l'utilisation de patchs, qui capturent

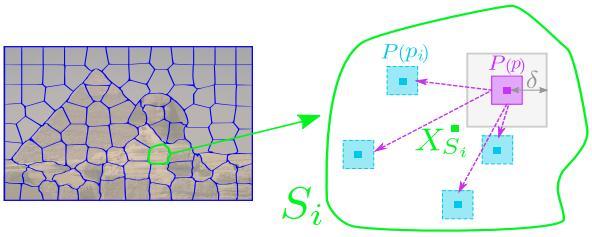


FIGURE 3 – Sélection de patchs similaires $P(p_i)$ dans un superpixel S_i en dehors d'un voisinage δ, pour calculer le terme d'homogénéité de texture (3) pour un patch $P(p)$

naturellement l'information de texture. Pour un pixel p, de patch $P(p)$, et un superpixel S_i, un algorithme de plus proches voisins est utilisé pour trouver N patchs similaires $P(p_i)$ tels que $p_i \in S_i$, en dehors d'un voisinage δ autour de p (voir Figure 3). Un nouveau terme d_P calcule la distance moyenne aux $p_i \in S_i$ sélectionnés :

$$d_P(p, S_i) = \frac{1}{N} \sum_{p_i \in \mathcal{K}_p} d_P(p, p_i), \tag{3}$$

avec $\mathcal{K}_p$ l'ensemble des N pixels $p_i \in S_i$, comparés par une distance basée patchs dans l'espace des descripteurs, telle que $d_P(p, p_i) = \frac{1}{n}\|F_{P(p)} - F_{P(p_i)})\|_2^2$, avec n la taille des patchs.

N'importe quel descripteur peut être utilisé dans (3). Nous proposons donc un modèle général évaluant facilement la conformité de texture d'un voisinage de pixel à un superpixel, sans recourir à une approche complexe de classification de texture.

Recherche de plus proches voisins basée patchs. La recherche de patchs similaires peut être effectuée par n'importe quelle méthode de plus proches voisins. Nous utilisons PatchMatch, un algorithme rapide de plus proches voisins approchés [3]. Le calcul de $d_P(p, S_i)$ peut être effectué directement pour tous les pixels p de la zone $(2s+1)\times(2s+1)$ autour de X_{S_i}. L'algorithme étant en partie aléatoire, N patchs dans S_i peuvent être cherchés en parallèle pour chaque pixel p, pour augmenter la robustesse du terme d'homogénéité de texture (3).

Unicité de texture des superpixels. Dans (3), la similarité est calculée sans information spatiale. Un patch $P(p)$ peut donc trouver des textures similaires dans des zones restreintes, menant un superpixel S_i à pouvoir grouper plusieurs textures (voir Figure 2(c)). Pour assurer l'unicité de texture, nous considérons dans d_P, la distance spatiale entre les patchs $P(p_i)$ en $X_{p_i} \in S_i$, et X_{S_i}, le barycentre spatial de S_i :

$$d_P(p, p_i) = \frac{1}{n}\|F_{P(p)} - F_{P(p_i)})\|_2^2 + \frac{m_i^2}{s^2}\,\Gamma\big(X_{p_i}, X_{S_i}\big), \tag{4}$$

avec $\Gamma(X_{p_i}, X_{S_i}) = 2s^2(1 - \exp(-\|X_{p_i} - X_{S_i}\|_2^2/s^2))$. Ce terme contribue itérativement à limiter la zone de recherche et la diversité des textures dans S_i en pénalisant fortement les patchs similaires trouvés loin de X_{S_i}. Par conséquent, le barycentre est encouragé à se déplacer vers une zone de texture homogène (voir Figure 2(d)).

SALT calcule le groupement des pixels selon la distance :

$$D(p, S_i) = d_F(F_p, F_{S_i}) + d_s(X_p, X_{S_i})\frac{m_i^2}{s^2} + d_P(p, S_i). \tag{5}$$

Ainsi, SALT est une méthode générale, robuste aux textures, et performante sur de nombreux types d'images, en utilisant les mêmes paramètres, comme démontré dans la section suivante.

3 Évaluation des performances

3.1 Validation

Bases d'images. Deux bases d'images composites sont proposées pour évaluer les performances de segmentation de texture [1]. Une base de 10 images de 300×400 pixels contenant différentes bandes noires et blanches, similaires à celles de la Figure 1 est proposée (mix-Stripes). Des images de textures naturelles, normalisées en intensité sont également prises de la base Brodatz [4], pour créer 100 images composites (mix-Brodatz), pouvant contenir jusqu'à 10 textures différentes. Chaque région contenant une texture ayant une taille d'au moins 1000 pixels. Enfin, la base BSD [13], contenant 200 images naturelles couleur de taille 321×481 pixels est également utilisée.

Métriques. Les méthodes sont évaluées par les métriques standards ASA (Achievable Segmentation Accuracy), et mesure F de détection de contours [10], et nous donnons les résultats quantitatifs pour un nombre moyen de 250 superpixels.

Paramètres. Les paramètres sont définis empiriquement une fois pour toutes. $N = 8$ patchs de taille 5×5 pixels sont sélectionnés en dehors d'un voisinage de $\delta = 3$ pixels. Dans (2), m est fixé à 0.1 et β à 25. Les descripteurs de couleur F sont calculés comme dans [10]. Enfin, 10 itérations sont effectuées pour calculer la décomposition, comme dans [1].

3.2 Résultats

Influence des contributions. L'impact visuel des contributions est montré dans la Figure 2, et les mesures ASA et F sont données dans le Tableau 1. Ce nouveau terme d'homogénéité de texture (3) et la contrainte d'unicité de texture (4) améliorent les performances sur chaque type de données.

TABLE 1 – Influence des contributions dans la méthode SALT

	mix-Stripes		mix-Brodatz		BSD	
Méthode	ASA	F	ASA	F	ASA	F
SALT sans (3),(4)	0.8303	0.3498	0.7969	0.4736	0.9484	0.4945
SALT sans (4)	0.8486	0.3882	0.8112	0.4812	0.9493	0.4961
SALT	**0.8706**	**0.4232**	**0.8139**	**0.4824**	**0.9503**	**0.4992**

Comparaison aux méthodes de l'état de l'art. SALT est comparé aux approches de l'état de l'art sur une image similaire à celles de la base d'images mix-Stripes dans la Figure 1, et aux méthodes les plus récentes sur des images tirées de mix-Brodatz et de la BSD dans la Figure 4.

Une évaluation quantitative est donnée dans le Tableau 2. SALT augmente significativement les performances sur les bases de textures synthétiques (mix-Stripes), et naturelles (mix-Brodatz), démontrant ainsi sa capacité à produire des superpixels robustes aux textures. SALT obtient également les meilleurs résultats sur des images naturelles couleur (BSD), en utilisant les mêmes paramètres, alors que les autres méthodes ne permettent pas de fournir des résultats précis sur les trois types de données en même temps.

Les méthodes comparées sont utilisées avec les paramètres recommandés par les auteurs. Néanmoins, aucune autre approche ne capture explicitement la texture. Par conséquent,

1. Bases de données disponibles à : `http://remigiraud.fr/salt/`

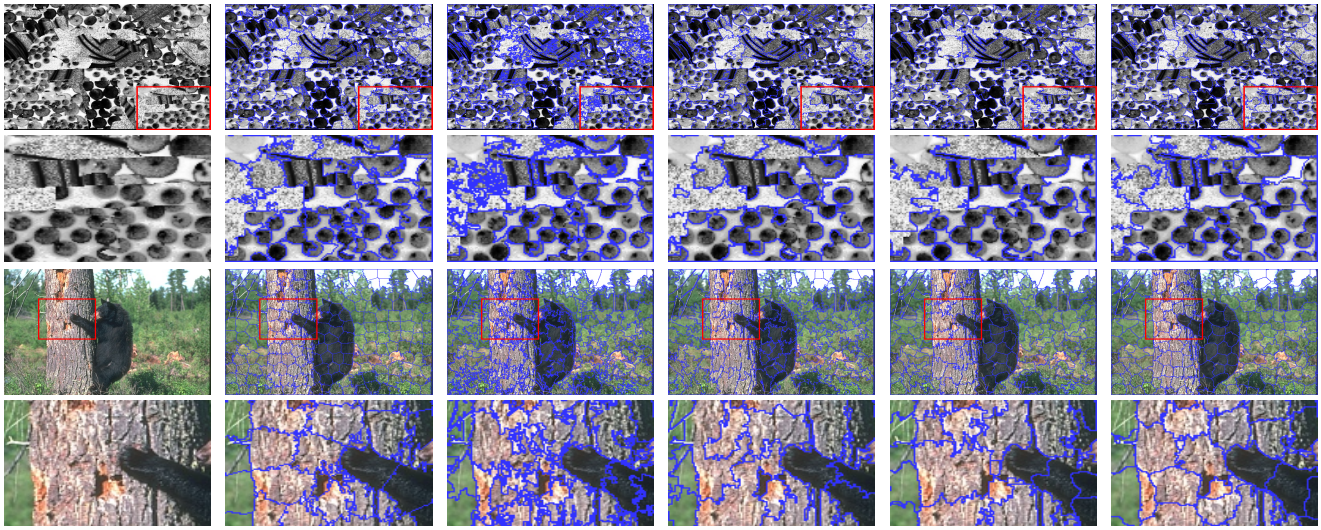


Image initiale | ETPS [18] | LSC [6] | SNIC [2] | SCALP [10] | SALT

FIGURE 4 – Comparaison entre SALT et les méthodes les plus récentes de l'état de l'art pour 200 superpixels. SALT produit un résultat plus pertinent sur une image composite de textures naturelles (haut) et sur une image couleur naturelle de la BSD (bas)

SALT surpasse toujours les méthodes de l'état de l'art optimisées manuellement pour chaque base d'images.

TABLE 2 – SALT vs les méthodes de l'état de l'art. Meilleurs et seconds résultats sont respectivement gras et soulignés

	mix-Stripes		mix-Brodatz		BSD	
Méthode	ASA	F	ASA	F	ASA	F
SLIC [1]	0.7256	<u>0.4048</u>	0.7784	0.4607	0.9445	0.4706
ERGC [5]	0.6107	0.3717	0.7796	0.4677	0.9477	0.4571
ETPS [18]	<u>0.7769</u>	0.2953	0.7568	0.4354	0.9433	0.4710
LSC [6]	0.6979	0.3156	0.7908	0.4552	<u>0.9503</u>	0.4421
SNIC [2]	0.6659	0.3529	0.7662	<u>0.4815</u>	0.9410	0.4617
SCALP [10]	0.7307	0.3290	<u>0.7977</u>	0.4759	0.9499	<u>0.4914</u>
SALT	**0.8706**	**0.4232**	**0.8139**	**0.4824**	**0.9503**	**0.4992**

4 Conclusion

Dans ce travail, nous répondons à la non-robustesse des approches superpixels en proposant une méthode robuste aux textures. Une nouvelle distance basée patchs est introduite pour grouper des pixels de même couleur et même texture locale. La méthode proposée est générale et supprime le besoin d'un réglage manuel de la contrainte spatiale. Nous améliorons la segmentation des méthodes de l'état de l'art sur des bases d'images texturées mais aussi couleur en utilisant les mêmes paramètres. Ce travail ouvre la voie à une plus large utilisation des superpixels, par ex., pour des applications complexes de segmentation d'images médicales ou de suivi d'objets vidéo.

Références

[1] R. Achanta, et al. *SLIC superpixels compared to state-of-the-art superpixel methods*, PAMI, 2012.

[2] R. Achanta, et al. *Superpixels and polygons using simple non-iterative clustering*, CVPR, 2017.

[3] C. Barnes, et al. *PatchMatch : A randomized correspondence algorithm for structural image editing*, ToG, 2009.

[4] P. Brodatz et al. *Textures : A photographic album for artists and designers*, Dover Publications, 1966.

[5] P. Buyssens et al. *Eikonal-based region growing for efficient clustering*, IVC, 2014.

[6] J. Chen, et al. *Linear spectral clustering superpixel*, TIP, 2017.

[7] S. He, et al. *SuperCNN : A superpixelwise convolutional neural network for salient object detection*, IJCV, 2015.

[8] R. Giraud, et al. *SuperPatchMatch : An algorithm for robust correspondences using superpixel patches*, TIP, 2017.

[9] R. Giraud, et al. *Evaluation framework of superpixel methods with a global regularity measure*, JEI, 2017.

[10] R. Giraud et al. *Robust superpixels using color and contour features along linear path*, CVIU, 2018.

[11] M. Y. Liu, et al. *Entropy rate superpixel segmentation*, CVPR, 2011.

[12] Y.-J. Liu, et al. *Manifold SLIC : A fast method to compute content-sensitive superpixels*, CVPR, 2016.

[13] D. Martin, et al. *A Database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics*, ICCV, 2001.

[14] M. Reso et al. *Temporally consistent superpixels*, ICCV, 2013.

[15] W.-C. Tu et al. *Learning superpixels with segmentation-aware affinity loss*, CVPR, 2018.

[16] M. Van den Bergh, et al. *SEEDS : Superpixels extracted via energy-driven sampling*, ECCV, 2012.

[17] J. Yan, et al. *Object detection by labeling superpixels*, CVPR, 2015.

[18] J. Yao, et al. *Real-time coarse-to-fine topologically preserving segmentation*, CVPR, 2015.