A DATA-CENTRIC APPROACH TO UNSUPERVISED TEXTURE SEGMENTATION USING PRINCIPLE REPRESENTATIVE PATTERNS

Kaitai Zhang[1*], *Hong-Shuo Chen*[2*], *Xinfeng Zhang*[1], *Ye Wang*[1] *and C.-C. Jay Kuo*[1]

Ming Heish Department of Electrical Engineering, University of Southern California, Los Angeles, USA[1]
College of Electrical and Computer Engineering, National Chiao Tung University, Hsinchu, Taiwan[2]

ABSTRACT

Features that capture textural patterns of a certain class of images are crucial for texture segmentation tasks. This paper introduces a data-centric approach to efficiently extract and represent textural information, which adapts to a wide variety of textures. Based on the strong self-similarities and quasi-periodicity in texture images, the proposed method first constructs a representative texture pattern set for the given image by leveraging the patch clustering strategy. Then, pixel-wise texture features are designed according to the similarities between local patches and the representative textural patterns. Moreover, the proposed feature is generic and flexible, and can perform segmentation task by integrating it into various segmentation approaches easily. Extensive experimental results on both textural and natural image segmentation show that the segmentation method using the proposed features achieves very competitive or even better performance compared with the stat-of-the-art methods.

***Index Terms*—** Unsupervised texture segmentation, Data-centric feature extraction, self-similarity

1. INTRODUCTION

Image segmentation is one of the most fundamental tasks in image processing and computer vision research, which has a wide range of applications like autonomous driving, remote sensing and medical image diagnosis. Herein, texture segmentation is a more frequently occurring problem appearing in various circumstances, which partitions an image into multiple regions with similar textural patterns. However, due to the complexity of the textures, their segmentation is more challenging compared with that on natural images, where the structures are more regular.

In the past decades, texture analysis and segmentation[1, 2, 3, 4] have been well studied and numerous methods have been proposed. For these methods, the local features representing image structural information play a crucial role in segmentation task by integrating into machine learning frameworks. Especially, various handcrafted features have been designed to characterize textural patterns. Herein, filter banks are very popular feature extraction schemes, such as Gabor filters[1], gradients filters, Laplacian filters, and Gaussian filters. The responses to such filter banks, as well as statistics based on them, are used to obtain local texture features[5]. Another type of textural features include Local Binary Patterns[6], co-occurrence matrices [7], and wavelet transforms [8, 3].

For unsupervised texture segmentation, those features should be further processed with some well-defined algorithms, *e.g.*, graph cut, clustering[1, 9], and region merging. More recent works employing matrix factorization[10] and energy function minimization [11, 12] have achieved excellent performance in segmentation. However, these works mainly focus on the segmentation part without specifically local feature design, and the further improvement of their performance is prohibited by the limitations of the existing local feature extraction strategies.

The first limitation for most of the existing local feature extractors, *e.g.* filter banks, is the inefficient adaptivity to various texture contents due to the pre-defined filter parameters. Therefore, the effectiveness of these local features varies significantly from different classes of texture or even depends on a single texture image. To overcome this problem, a straightforward method [10, 11] is to manually select a subset of filters from the aforementioned large set of filter banks, with a belief that there are filters suitable for the textural patterns in the processed images. However, the manual selection requires human experience and insights, which typically is an expensive and time-consuming task. In such situation, it would be desirable to have a feature extraction method which can automatically adapt to various kinds of texture image and learn effective features in a data-driven manner.

Besides the adaptivity, the absent of the global contrast information in existing local features, the computation of which is constrained inside small image patches, makes them are blind to image context. The contrast information between texture and its context in a image is never taken into account during the whole feature extraction process, and only utilized in subsequent segmentation part, which limits the generalization of the local features. In addition, the high dimensionality of existing local features increases the computational bur-

*Equally contributed to this paper and listed as co-first author.

dens since the segmentation algorithms need to calculate the feature distance. FSEG method[10] introduces a matrix factorization model to bypass iterative high-dimension distance calculations and achieves fast segmentation. The PCA-MS method[11] employs principal component analysis(PCA) to reduce dimension explicitly, which shows the inherent redundancy of those handcrafted local features. Therefore, a good solution is to design a compact texture features to deal with various texture patterns adaptively by leveraging both the image local and global information.

In this paper, we propose a data-centric textural feature extraction method using Principle Representative Patterns. The proposed method clusters image patches, and selects those cluster centroids as major patterns, denoted as Principle Representative Patterns(PRP). Texture features are constructed by measuring the similarities between local patch and those Principle Representative Patterns. Herein, the contrast information is utilized the proposed feature representation, which further improves their discriminative power. The proposed PRP features can perform image segmentation task by further integrating them into existing segmentation algorithms. Since the proposed method is a data-driven strategy in an unsupervised manner, it is adaptive to various textural structures. Moreover, it offers compact feature representation with a tunable hyperparameter to control the feature dimension. Extensive experimental results on both textural and natural image segmentation tasks show the superiority of the proposed method.

2. TEXTURAL FEATURE BASED ON PRINCIPLE REPRESENTATIVE PATTERNS

2.1. Pixel-wised Texture Representation

To tackle the unsupervised texture segmentation problem, we first build a pixel-wised texture representation. Since the textural information is based on local structures instead of a single pixel, we model our texture representation in basis of image patches. Given an image I, the texture information of each pixel i, denoted as $T(i)$, is represented as below,

$$T(i) = \{w(P_i, P_j) |\, j \in I\}, \forall i \in I. \tag{1}$$

Herein, we use the function $w(P_i, P_j)$ to describe the relationship between two patches, P_i and P_j. P_i denotes of a square patch of the fixed size and centered at a pixel i.

It is obvious that $T(i) \in R^N$ is a high dimension vector, where N is the number of pixels in image I. It will be computationally unacceptable to utilize $T(i)$ directly. Inspired by the strong self-similarity and quasi-periodicity of texture images, we assume that there exist a set of Principle Representative Patterns (PRPs) in image I, and they contain all the major textural patterns of the image. Then, we can only use these PRPs to describe $T(i)$ to reduce the dimension of the feature vector in Eq.1, with a negligible information loss. Thus, the

(a) (b) (c)

Fig. 1. Segmentation results step by step. (a) A texture mosaic with 5 different components from Brodatz texture dataset. (b) Raw segmentation results with proposed PRP features. (c) Final segmentation mask after post-processing.

feature vector can be rewritten as,

$$T(i) = \{w(P_i, P_j) |\, P_j \in R\}, \forall i \in I, \tag{2}$$

where R denotes the set of all Principle Representative Patterns selected for the image I.

Our texture feature $T(i)$ model above considers not only the local patch P_i, but also its contrast with patches among the whole image, including local and non-local ones. Moreover, the whole process is driven by the processed data in an unsupervised manner, and the usage of PRPs provides us more freedoms in dimension control of $T(i)$ by manipulating the cardinality of R. Another thing need to be noticed is that we only use image patch P_i here to illustrate our idea, and P_i could actually be extended to generalized feature representation $g(P_i)$ extracted from P_i.

2.2. Principle Representative Pattern (PRP)

To select those representative patterns among all possible candidates in image I, an effective method is to select one patch from a collection of very similar patches as the representative one. We preform the K-means clustering method to get cluster centroids $\{u_i\}_{i=1}^{K}$ on all the image patches of I according to the following distance metric,

$$\arg\max \sum_{i=1}^{K} \sum_{j \in I} Dist(P_j, u_i). \tag{3}$$

Those cluster centroids are utilized to form the Principle Representative Pattern set R. The number of centroids, K, could be tuned as hyperparameter in practice according to the performance requirement and computation source limitation.

The idea of Principle Representative Pattern tries to directly take advantage of the unique properties of texture images, *i.e.,* quasi-periodicity and self-similarity. In other words, it uses the fact of high degree of redundancy in texture images. In fact, natural images also contain different levels of redundancy, and the proposed feature representation can also be extended to natural images.

2.3. PRP Features and Segmentation

With the Principle Representative Patches, we could get $T(i)$ by computing the similarities between each image patch P_i and PRPs. Inspired by the famous non-local means denoising algorithm[13], we employ Gaussian weighting function to measure the similarity between two patches.

$$w(P_i, P_j) = \frac{1}{Z(i)} e^{\frac{\|P_i - P_j\|_{2,\sigma}^2}{h^2}}, \tag{4}$$

where $\sigma > 0$ is the standard deviation of the Gaussian kernel, h is a smooth factor and $Z(i)$ is a normalization constant.

$$Z(i) = \sum_j e^{\frac{\|P_i - P_j\|_{2,\sigma}^2}{h^2}}. \tag{5}$$

Based on the above formulation, the proposed textural feature, $T(i)$, can be interpreted as a probability mass function, or an energy spectrum. Each entrance of $T(i)$ describes the probability of P_i to be this major textural pattern, or energy that P_i projects to the corresponding representative textural pattern. Higher value means higher probability and intense energy concentration, while lower value offers complementary side evidences to better describe P_i.

The proposed method can be further extended to a general feature encoding technique. Replacing local patch P_i with any local feature, our method could encode those local features into a compact and context-aware form. The generated feature map $\{T(i)\}$ is also completely flexible to be combined with any cutting-edge segmentation algorithms, such as Mumford-Shah functional [11] and matrix factorization method [10]. The above facts make our method have full potential to evolve with emerging local features and segmentation algorithms in the future.

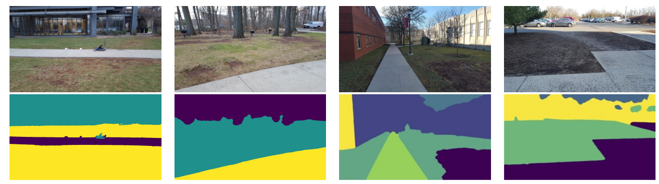
3. EXPERIMENTAL RESULTS

We show segmentation results on various natural and textural images in the real world and, then, provide quantitative evaluations by comparing the proposed method with other stat-of-the-art methods on the Prague Unsupervised Texture Segmentation Benchmark [14, 15].

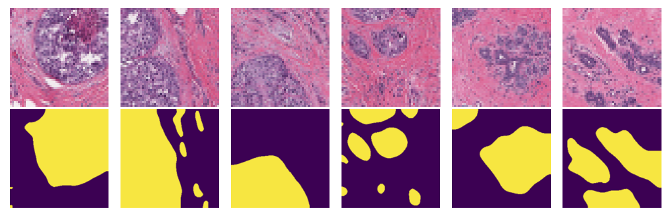
3.1. Qualitative Segmentation Results

We first test it on several natural scene images and animals images. Fig.2(a) illustrates some examples of the segmentation results using the proposed features, and those images are from Berkeley Segmentation Dataset(BSDS500) [16]. We can find that the proposed approach obtains all of main regions segmented correctly, even without any usage of object-specific knowledge. Although there are still some flaws, *e.g.*, the missing of goose's beak and the merging for trees and

(a) Natural scenes and animals



(b) Ground Terrain Textures.



(c) Histology images.

Fig. 2. Segmentation results of various natural texture images in the real world.

their shadow, those flaws are reasonable since no semantic meaning is included in our model.

We also test the proposed method in the ground terrain textures, which can be seen everywhere and have many variations under different weather and lighting conditions. In Fig.2(b), some examples from Ground Terrain Outdoor Scenes Dataset(GTOS) [17] are showed with the corresponding segmentation results. The boundaries between two different terrain are shape and fine in segmentation map, which proves that the proposed method has an excellent ability to handle different terrain textures and even on the situation where texture is nonuniform and noisy.

Histology image segmentation which provides precise location information of different tissues is always desirable by surgeons and researchers. Fig.2(c) shows several histology images with segmentation results using the proposed method. We can see that the irregular boundaries between two tissues are well-located, and both main regions and secondary regions can be well distinguished and separated.

3.2. Comparison Results on Texture Mosaics

We further compare our method with three state-of-the-art methods on the Prague segmentation benchmark, which contains 80 color texture mosaics of size 512×512. For each of the 80 texture mosaics, we learn a separate set of Principle Representative Patterns(PRP) and compute the segmen-

Table 1. Comparison Results of proposed approach with various segmentation methods on the Prague Unsupervised Texture Segmentation Benchmark. Up arrows indicate better results correspond to large values, and down arrows the opposite. Boldface highlights the best, and a star denotes the second-best value in each column.

Method	CS↑	OS↓	US↓	O↓	C↓	CO↑	CC↑	I.↓	II.↓	RM↓
PMCFA[18][1]	**75.32**	11.95*	9.65*	**4.51**	8.87	**88.16**	90.73*	**11.84**	**1.47**	**3.76**
Ours	72.75*	**8.11**	9.80	6.76*	**6.18**	87.20*	87.65	12.80*	2.28	3.72*
PCA-MS [11]	72.27	18.33	**9.41**	7.25	6.44*	85.96	**91.24**	14.40	1.59*	4.45
FSEG [10]	69.18	14.69	13.64	9.25	12.55	84.44	87.38	15.89	2.60	4.51

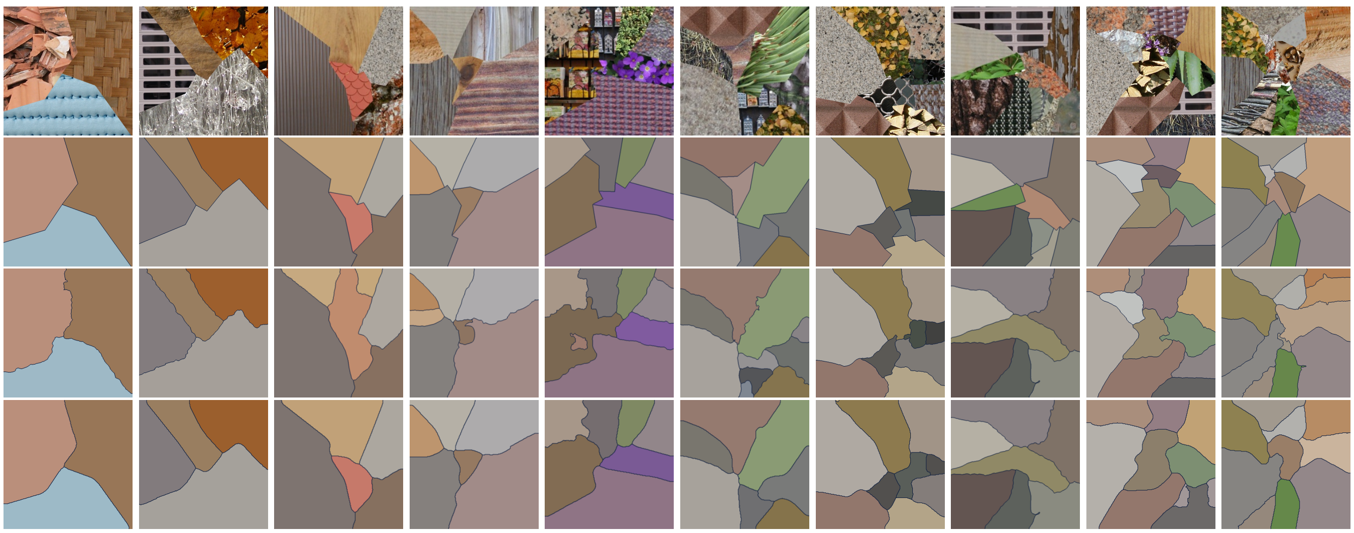


Fig. 3. Example results on Prague dataset. The first row shows original images, the second row shows ground truth, the third shows segmentation results from FSEG[10] and the last row shows results from the proposed method.

tation results subsequently. The parameters for learning the features are set empirically and remain fixed for all instances in the dataset. The images are converted from RGB color space to L*a*b* color space, which is more perceptually uniform. The patch size is 7×7, the number of PRPs is 50, and the scale parameter of Gaussian kernel is 0.1. To avoid usage of prior knowledge about the number of different textures components, we determine the number of clusters automatically by search it according to overall intra-cluster distance. Furthermore, the Conditional Random Field(CRF)[19] with a local voting scheme is utilized to refine the segmentation boundary as post-processing.

Table 1 shows numerical results for various segmentation schemes, where the best and the second best results are highlighted by boldface fonts and asterisks, respectively. We see that the proposed method offers the top 2 results on most evaluation metrics. It achieves the best scores in over segmentation (OS) and commission error (C) metrics. In total, our method achieves 2 best scores and 5 second-best scores among those 10 indicators. Fig. 3 shows the visualization results of segmentation on texture mosaics where the amount of texture components changes from 3 to 12 corresponding the images from left to right. We can see that the segmentation results by our method are very approach to the ground truth in most cases, and obviously outperforms FSEG. We also notice that the errors mainly occur in mosaics with higher number of components, *e.g.*, the first and second images from the right side. This is because the optimal hyperparameter setting for an image mosaic with a different number of components should be quite different, but we have to make a trade off to get an unified hyperparameter setting for the whole dataset.

[1]Detailed introduction of PMCFA algorithm could be found online at https://sites.google.com/site/costaspanagiotakis/research/imagesegmentation

4. CONCLUSION

An effective textural feature extraction method for unsupervised texture segmentation was presented. Features are learned from data in an unsupervised manner. They encode local features as well as contrast information. It was shown by extensive experimental results that the proposed method offers the state-of-the-art performance.

5. REFERENCES

[1] Anil K Jain and Farshid Farrokhnia, "Unsupervised texture segmentation using gabor filters," *Pattern recognition*, vol. 24, no. 12, pp. 1167–1186, 1991.

[2] Guodong Wang, Jingge Lu, Zhenkuan Pan, and Qiguang Miao, "Color texture segmentation based on active contour model with multichannel nonlocal and tikhonov regularization," *Multimedia Tools and Applications*, vol. 76, no. 22, pp. 24515–24526, 2017.

[3] Tianhorng Chang and C.-C. Jay Kuo, "Texture analysis and classification with tree-structured wavelet transform," *IEEE Transactions on image processing*, vol. 2, no. 4, pp. 429–441, 1993.

[4] Jingge Lu, Guodong Wang, and Zhenkuan Pan, "Non-local active contour model for texture segmentation," *Multimedia Tools and Applications*, vol. 76, no. 8, pp. 10991–11001, 2017.

[5] Xiuwen Liu and DeLiang Wang, "Image and texture segmentation using local spectral histograms," *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 3066–3077, 2006.

[6] Timo Ojala and Matti Pietikäinen, "Unsupervised texture segmentation using feature distributions," *Pattern recognition*, vol. 32, no. 3, pp. 477–486, 1999.

[7] Patrick C Chen and Theodosios Pavlidis, "Segmentation by texture using a co-occurrence matrix and a split-and-merge algorithm," *Computer graphics and image processing*, vol. 10, no. 2, pp. 172–182, 1979.

[8] S Arivazhagan and L Ganesan, "Texture segmentation using wavelet transform," *Pattern Recognition Letters*, vol. 24, no. 16, pp. 3197–3203, 2003.

[9] Michael Unser, "Texture classification and segmentation using wavelet frames," *IEEE Transactions on image processing*, vol. 4, no. 11, pp. 1549–1560, 1995.

[10] Jiangye Yuan, Deliang Wang, and Anil M Cheriyadat, "Factorization-based texture segmentation," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3488–3497, 2015.

[11] Niklas Mevenkamp and Benjamin Berkels, "Variational multi-phase segmentation using high-dimensional local features," in *Applications of Computer Vision (WACV), 2016 IEEE Winter Conference on*. IEEE, 2016, pp. 1–9.

[12] Martin Storath, Andreas Weinmann, and Michael Unser, "Unsupervised texture segmentation using monogenic curvelets and the potts model," in *Image Processing (ICIP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 4348–4352.

[13] Antoni Buades, Bartomeu Coll, and J-M Morel, "A non-local algorithm for image denoising," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. IEEE, 2005, vol. 2, pp. 60–65.

[14] M. Haindl and S. Mikes, "Texture segmentation benchmark," in *Proceedings of the 19th International Conference on Pattern Recognition, ICPR 2008*, Los Alamitos, December 2008, pp. 1–4, IEEE Computer Society.

[15] S. Mikes, M. Haindl, G. Scarpa, and R. Gaetano, "Benchmarking of remote sensing segmentation methods," *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, vol. 8, no. 5, pp. 2240–2248, May 2015.

[16] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. 8th Int'l Conf. Computer Vision*, July 2001, vol. 2, pp. 416–423.

[17] Jia Xue, Hang Zhang, Kristin J Dana, and Ko Nishino, "Differential angular imaging for material recognition.," in *CVPR*, 2017, pp. 6940–6949.

[18] Costas Panagiotakis, Ilias Grinias, and Georgios Tziritas, "Natural image segmentation based on tree equipartition, bayesian flooding and region merging," *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2276–2287, 2011.

[19] John Lafferty, Andrew McCallum, and Fernando CN Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," 2001.