Weakly-Supervised Sparse Coding with Geometric Prior for Interactive Texture Segmentation

Yuhui Quan, Huan Teng, Tao Liu and Yan Huang*

***Abstract*—Texture segmentation is about dividing a texture-dominant image into multiple homogeneous texture regions. The existing unsupervised approaches for texture segmentation are annotation-free but often yield unsatisfactory results. In contrast, supervised approaches such as deep learning may have better performance but require a large amount of annotated data. In this paper, we propose a user-interactive approach to win the trade-off between unsupervised approaches and supervised deep approaches. Our approach requires the user to mark one pixel in each texture region, whose label is directly propagated to its neighbor region. Such labeled data are of very small amount and even partially erroneous. To effectively exploit such weakly-labeled data, we construct a weakly-supervised sparse coding model that jointly conducts feature learning and segmentation. In addition, the geometric constraints are developed for the model to exploit the geometric prior on the local connectivity of region boundaries. The experiments on two benchmark datasets have validated the effectiveness of the proposed approach.**

***Index Terms*—Texture segmentation, sparse coding, weakly-supervised learning, geometric constraints**

I. Introduction

Texture regions of various types are prevalent in the images from daily life and real applications. Objects and backgrounds in natural scenes often consist of different textures [1]. An image about materials may encompass various texture regions. MRI (Magnetic Resonance Imaging) and CT (Computed Tomography), are about tissue cells modeled by mixtures of texture regions [2–4]. SAR (Synthetic Aperture Radar) images contain mountains, rivers and lands, which exhibit texture regions of different forms [5]. Separating such texture regions not only enables separate treatments to different image regions and objects, but also provides perceptual attributes and mid-level cues for visual recognition and understanding.

A closely-related topic in image processing is called texture segmentation [6, 7], whose task is about dividing a texture-dominant image into multiple disjoint homogeneous texture regions. Texture segmentation has its practical values in many fields, including the analysis on material, medical, biological and chemical images [3, 8–10], natural scene understanding [11], automatic navigation [12], remote sensing [5], etc.

Most traditional methods model texture segmentation as the patch-level clustering which involves two stages: *(i)* represent image patches by feature vectors (*e.g.* local spectral histogram [6, 13] and local regularity spectrum [14]); and *(ii)* conduct global segmentation via clustering in the feature space (*e.g.* k-means [6], graph cut [15] and mean-shift [16]).

Regarding the patch representation, traditional methods use hand-crafted features which are not adaptive to data. There are some approaches learning local features from input images for improvement. Khalilzadeh *et al.* [9] used sparse coding to extract patch features from MRI images. Rahmani *et al.* [5] used dictionary learning to obtain discriminative features for SAR images. Yuan *et al.* [7] applied low-rank factorization on the input to obtain the representative features of natural images. Regarding the feature clustering, most existing methods may ignore the spatial regularity of region boundaries. Thus, a few approaches [13, 17] use the Mumford-Shah models for enforcing global geometric properties on boundaries.

The performance of the two-stage approaches is extremely dependent on the effectiveness of the local features. For improvement, some approaches jointly conduct local representation and global clustering. Wang *et al.* [18] proposed a variational model that combines dictionary learning of image patches and Mumford-Shah active contours. Kiechle *et al.* [19] proposed to learn the filters from data such that the filtering responses tend to generate piece-wise constant label maps. All aforementioned approaches are unsupervised, whose results are often unsatisfactory, especially with unknown number of regions. Besides, since texture is a scale-related visual concept, users in different tasks may have different understandings on the texture regions being partitioned, even on the same images. Unsupervised methods are difficult to adapt to different cases.

Recently, a few supervised methods [3, 4, 20, 21] based on deep learning have shown improvement over the unsupervised one. Nevertheless, the deep learning needs an abundant of annotated data for supervision, the collection of which is expensive for texture segmentation. In particular, multi-time annotations are needed for an image, as subjective ambiguity probably exists on the boundaries of texture regions. In addition, the texture types may vary a lot over different segmentation tasks, making the transfer of annotated data infeasible. Furthermore, professional knowledge and specific devices are required in many fields [3]. To alleviate the requirement on annotations, Vincent *et al.* [20] used simple clustering and shallow segmentation to obtain the rough labels for supervision. Chen *et al.* [3] transferred the learned features from natural images for gland segmentation. Huang *et al.* [21] assumed all texture types are known and proposed an effective polygon-based training data generation scheme.

Unsupervised methods and supervised deep methods have their merits and weaknesses. To win their trade-off, we pro-

This research is supported by National Natural Science Foundation of China (61602184, 61872151, 61902130, 61672241, U1611461), Natural Science Foundation of Guangdong Province (2017A030313376, 2016A030308013), Fundamental Research Funds for the Central Universities (x2js-D2181690, x2js-D2190650), and China Postdoctoral Science Foundation (2019M652896).

All the authors are with School of Computer Science and Engineering, South China University of Technology, Guangzhou, 510006, China. Asterisk denotes the corresponding author (huangkaiyan@scut.edu.cn).