Shape-Tailored Local Descriptors and their Application to Segmentation and Tracking

Naeemullah Khan[1], Marei Algarni[1], Anthony Yezzi[2], Ganesh Sundaramoorthi[1]
[1]King Abdullah University of Science and Technology (KAUST), Saudi Arabia
[2]School of Electrical and Computer Engineering, Georgia Institute of Technology, USA

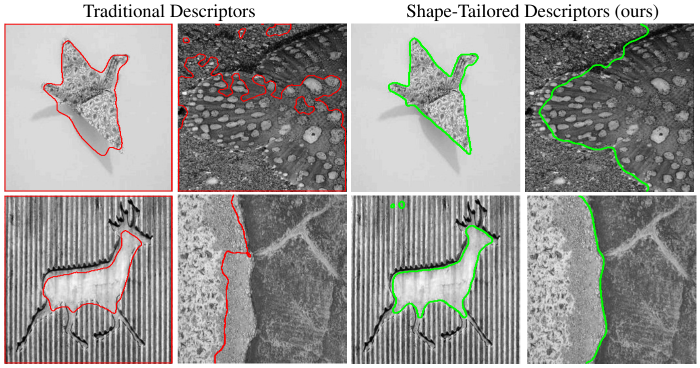

Figure 1: [Left]: Descriptors that aggregate local image data across boundaries of textured regions lead to segmentation errors. The problem is exacerbated as the texton size increases. [Right]: Segmentation by Shape-Tailored Descriptors (our method).

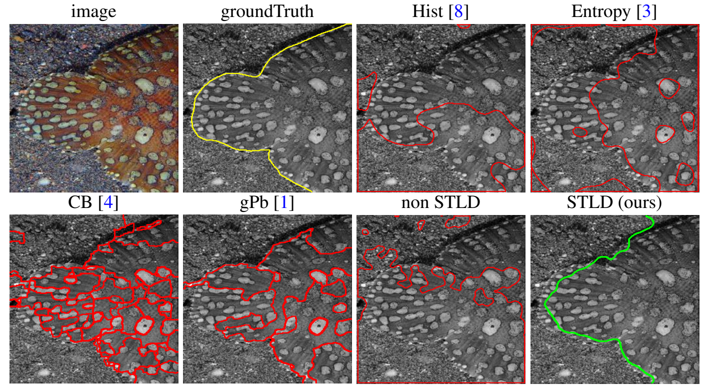

Figure 2: A sample result from our texture segmentation dataset, and comparison to state-of-the-art in texture segmentation.

In this paper, we propose new dense local invariant descriptors and show their effectiveness in texture segmentation. Local invariant descriptors are image statistics (typically oriented gradients) at each pixel that describe neighborhoods in a way that is invariant to geometric and photometric nuisances. These descriptors play an important role in characterizing local textural properties. Existing local invariant descriptors aggregate oriented gradients in predefined pixel neighborhoods that could contain image data from different textured regions. This leads to ambiguity in grouping descriptors, especially for descriptors near the boundary. This could lead to segmentation errors if descriptors are grouped to form a segmentation. The problem is exacerbated when the textons in the textures are large (see Fig. 1).

To solve this problem, one would need to compute oriented gradients only from within textured regions. However, the segmentation is not known a-priori. Thus, it is necessary to solve for the local descriptors and the region of the segmentation in a joint problem. We solve this in two steps. First, we construct novel dense local invariant descriptors, called *Shape-Tailored Local Descriptors (STLD)*. These descriptors are formed from shape-dependent scale spaces of oriented gradients. The shape dependent scale spaces are the solution of Poisson-like partial differential equations (PDE). Of particular importance is the fact that these scale-spaces are defined within a region of arbitrary shape and do not aggregate data outside the region of interest. Second, we incorporate Shape-Tailored Descriptors into the Mumford-Shah energy [7] as an example energy based on these descriptors. Optimization jointly estimates Shape-Tailored Descriptors and their support region, which forms the segmentation.

We have evaluated STLD on the problem of discriminating real-world textures at various scales, and other under geometric and photometric nuisances. We have collected a dataset of 256 images that contain two textures. An example result is shown in Fig. 2. A summary of results on this dataset is given in Fig. 3. We also show the application of STLD in detecting disocclusions for tracking objects consisting of multiple textured regions in video.

	Contour		Region metrics					
	F-meas.		GT-cov.		Rand. Index		Var. Info.	
	ODS	OIS	ODS	OIS	ODS	OIS	ODS	OIS
STLD	**0.58**	**0.58**	**0.87**	**0.87**	**0.87**	**0.87**	**0.59**	**0.59**
non-STLD	0.17	0.17	**0.81**	0.81	**0.82**	0.82	**0.77**	0.77
gPb [1]	**0.50**	**0.54**	0.74	**0.84**	0.78	**0.86**	0.80	**0.65**
CB [4]	0.48	0.52	0.64	0.70	0.66	0.75	0.89	0.78
SIFT	0.10	0.10	0.55	0.55	0.59	0.59	1.44	1.44
Entropy [3]	0.08	0.08	0.74	0.74	0.75	0.75	0.95	0.95
Hist-5 [8]	0.14	0.14	0.66	0.66	0.70	0.70	1.18	1.18
Hist-10 [8]	0.13	0.13	0.66	0.66	0.70	0.70	1.19	1.19
Chan-Vese [2]	0.14	0.14	0.71	0.71	0.73	0.73	1.04	1.04
LAC [5]	0.09	0.09	0.55	0.55	0.58	0.58	1.41	1.41
Global Hist [6]	0.12	0.12	0.65	0.65	0.67	0.67	1.12	1.12

Figure 3: Quantitative Results on a texture dataset of 256 images. Algorithms are evaluated using contour and region metrics. Higher F-measure for the contour metric, ground truth covering (GT-cov), and rand index indicate better fit to the ground truth, and lower variation of information (Var. Info) indicates a better fit to ground truth. Bold red indicate best results and bold black indicates second-best results.

[1] Pablo Arbelaez, Michael Maire, Charless Fowlkes, and Jitendra Malik. Contour detection and hierarchical image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(5):898–916, 2011.

[2] Tony F Chan and Luminita A Vese. Active contours without edges. *Image processing, IEEE transactions on*, 10(2):266–277, 2001.

[3] Byung-Woo Hong, Stefano Soatto, Kangyu Ni, and Tony Chan. The scale of a texture and its application to segmentation. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.

[4] Phillip Isola, Daniel Zoran, Dilip Krishnan, and Edward H Adelson. Crisp boundary detection using pointwise mutual information. In *Computer Vision–ECCV 2014*, pages 799–814. Springer, 2014.

[5] Shawn Lankton and Allen Tannenbaum. Localizing region-based active contours. *Image Processing, IEEE Transactions on*, 17(11):2029–2039, 2008.

[6] Oleg Michailovich, Yogesh Rathi, and Allen Tannenbaum. Image segmentation using active contours driven by the bhattacharyya gradient flow. *Image Processing, IEEE Transactions on*, 16(11):2787–2801, 2007.

[7] David Mumford and Jayant Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on pure and applied mathematics*, 42(5):577–685, 1989.

[8] Kangyu Ni, Xavier Bresson, Tony Chan, and Selim Esedoglu. Local histogram based segmentation using the wasserstein distance. *International Journal of Computer Vision*, 84(1):97–111, 2009.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.