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

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

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 textures in the images 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 occlusions for tracking objects consisting of multiple textured regions in video.

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

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 textures in the images 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 occlusions for tracking objects consisting of multiple textured regions in video.