Deep Spherical Superpixels
— Supplementary Material —

Rémi Giraud[1] and Michaël Clément[2]

[1] Univ. Bordeaux, Bordeaux INP, IMS, CNRS UMR 5218, France.
remi.giraud@ims-bordeaux.fr
[2] Univ. Bordeaux, Bordeaux INP, LaBRI, CNRS UMR 5800, France.
michael.clement@labri.fr

1 Datasets

Ground-truth Examples of images with their ground-truth segmentation for the two datasets PSD [7] and WP [8] are respectively shown in Fig. 1 and 2. For PSD, the ground-truth annotation is fine-grained, with very small objects being segmented. For WP, we can note that the segmentation is more at semantic object level (car, truck, road, etc.), and also that only the lower half of images contain annotations.

List of images For PSD, we used 55 images for training (#4 to #58), 5 images for validation (#1,2,3 and #59,60), and 15 images for testing (#61-75). For WP, we used 300 images for training (#1-300), 100 for validation (#301-400) and 100 for testing (#401-500).

Fig. 1. Examples of spherical images and associated ground-truth segmentations considered for experiments from the Panorama Segmentation Dataset (PSD) [7]. Note the annotation precision with very small objects being segmented.

Fig. 2. Examples of spherical images and associated ground-truth segmentations considered for experiments from the WildPass (WP) dataset [8]. Only the bottom half of the image contains annotations.

2 Data Augmentation

In this Section, we illustrate the impact of the random parameters in our data augmentation techniques. In Fig. 3, random half-width crop are selected on the input images and mirrored to create new spherical image and ground-truth. In Fig. 4, the random settings of the k_x and k_y parameters in the panoramic stretching algorithm [6] enable to warp the image and ground-truth while preserving the spherical properties.

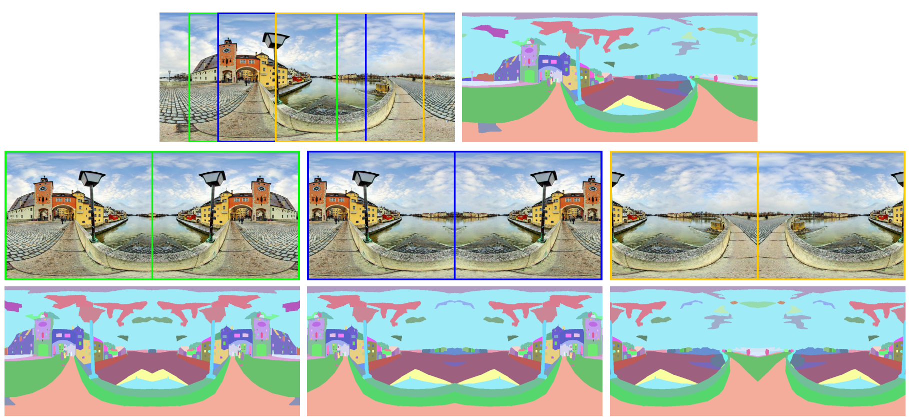

Fig. 3. Crop & mirror augmentation for different cropping positions. This method combines horizontal rolling, flipping and also creates information at the mirror border.

Fig. 4. Panoramic stretch augmentation for different settings. The parameters k_x, k_y when set > 1 or < 1 respectively correspond to an enlargement or a shrinking of the areas where 3D coordinates $|x| \approx 1$ and $|y| \approx 1$). The layout of the scene is represented by the green lines to more easily apprehend the distortion.

3 Parameter Settings and Training Details

The code of our method is based on a PyTorch implementation[3] of SSN. Our data augmentation is applied on-the-fly during training. These include (i) applying a random Gaussian blur with a variance σ ranging from 0 to 2, (ii) adding Gaussian noise of variance between 0 and 20, (iii) random horizontal rolling and flipping with half-width random crop and mirror with a 0.5 probability, and (iv) panoramic stretching with random parameters k_x and k_y between 0.5 and 2. Each channel of the input data (Lab F_c and 3D coordinates F_s) is normalized between -1 and 1. As in [4], color γ_c and position γ_s scale factors are respectively applied to input features F_c and F_s, and here set to and 0.6 and $10\times\sqrt{K}/h$ during training with $\lambda = 1$ in (2). Note that using different γ_c and γ_s for the inference can impact the regularity of the segmentation.

The training was conducted over 300k iterations, for $K = 200$ superpixels, with a batch size of 6 images, using Adam optimizer and a learning rate set at 1e-4, as in [4]. Training images were downsized to 256x512 pixels, so our model can understand the whole scene's geometry, contrary to the 201x201 crops used in [4]. Training was performed on a NVIDIA Titan V GPU with 12 GB of memory.

4 Qualitative Results

In Fig. 5, 6 and 7, we respectively show qualitative results on PSD images, noisy PSD images and WP images for the proposed DSS and state-of-the-art methods.

[3] `https://github.com/perrying/ssn-pytorch`

Fig. 5. Qualitative comparison on PSD images, for planar (left) and spherical methods (right) for two superpixel numbers $K = 1200$ (top-left) and $K = 400$ (bottom right).

References

1. Achanta, R., Süsstrunk, S.: Superpixels and polygons using simple non-iterative clustering. In: Conference on Computer Vision and Pattern Recognition (2017)
2. Ban, Z., Liu, J., Cao, L.: Superpixel segmentation using gaussian mixture model. IEEE Transactions on Image Processing **27**(8), 4105–4117 (2018)
3. Giraud, R., Pinheiro, R.B., Berthoumieu, Y.: Generalization of the shortest path approach for superpixel segmentation of omnidirectional images. Pattern Recognition **142**, 109673 (2023)

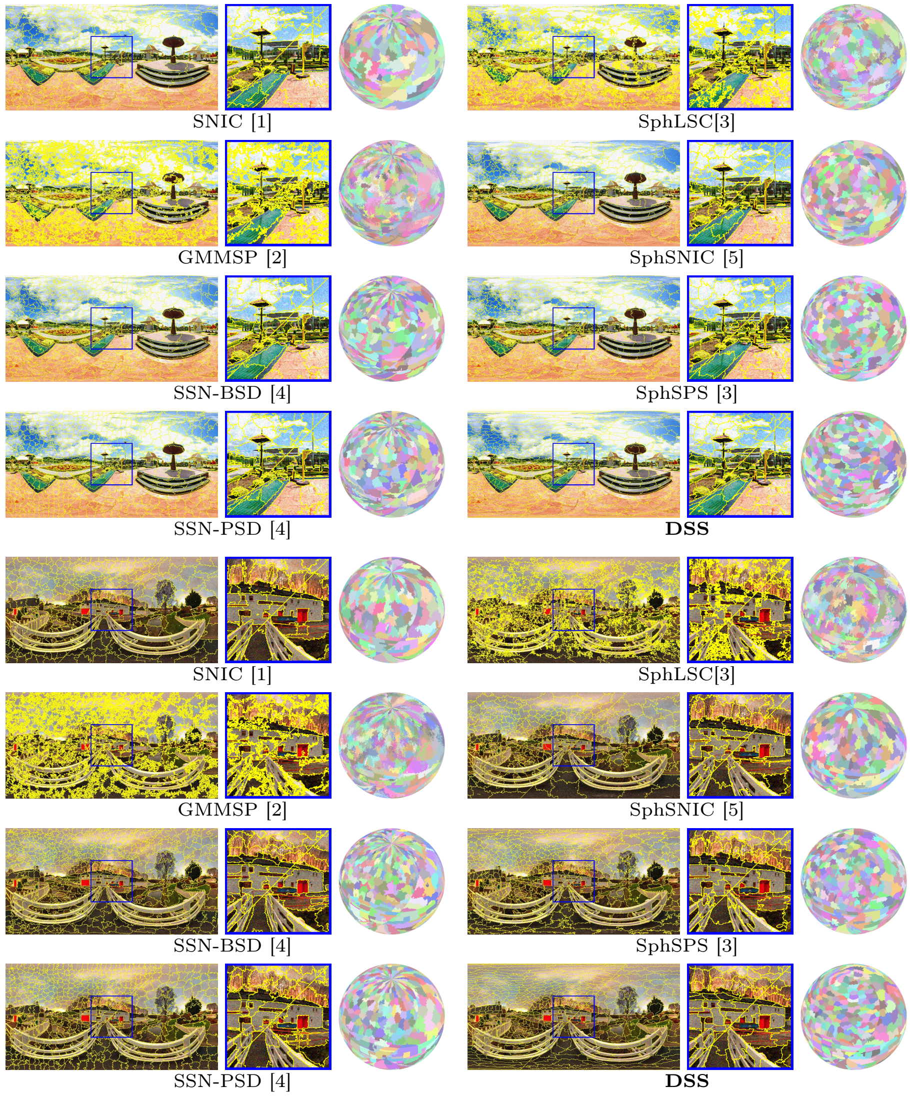

Fig. 6. Qualitative comparison on noisy PSD images, for planar (left) and spherical methods (right) for two superpixel numbers $K = 1200$ (top-left) and $K = 400$ (bottom right).

4. Jampani, V., Sun, D., Liu, M.Y., Yang, M.H., Kautz, J.: Superpixel sampling networks. In: European Conference on Computer Vision (2018)
5. da Silveira, T.L., de Oliveira, A.Q., Walter, M., Jung, C.R.: Fast and accurate superpixel algorithms for 360° images. Signal Processing **189** (2021)
6. Sun, C., Hsiao, C.W., Sun, M., Chen, H.T.: HorizonNet: Learning room layout with 1D representation and Pano Stretch data augmentation. In: Conference on Computer Vision and Pattern Recognition (2019)

Fig. 7. Qualitative comparison on WP images, for planar (left) and spherical methods (right) for two superpixel numbers $K = 1200$ (top-left) and $K = 400$ (bottom right).

7. Wan, L., Xu, X., Zhao, Q., Feng, W.: Spherical Superpixels: Benchmark and evaluation. In: Asian Conference on Computer Vision (2018)
8. Yang, K., Zhang, J., Reiß, S., Hu, X., Stiefelhagen, R.: Capturing omni-range context for omnidirectional segmentation. In: Conference on Computer Vision and Pattern Recognition (2021)