Supplementary Material for "Dense Non-Rigid Structure-from-Motion and Shading with Unknown Albedos"

Mathias Gallardo[1], Toby Collins[2] and Adrien Bartoli[1]

[1] EnCoV, IP, UMR 6602 CNRS, Université Clermont Auvergne, SIGMA, France

[2] IRCAD, Strasbourg, France

Mathias.Gallardo@gmail.com, Toby.Collins@gmail.com, Adrien.Bartoli@gmail.com

In this document, we provide supplementary information about the 'boundariness map' E_t computation for the boundary contour term. This term encourages the boundary of the surface to project to strong intensity edges in the image and it is a very useful constraint for poorly-textured surfaces. However, boundary may be attracted by false surface edges, which corresponds to background clutter or surface texture. The segmentation cue to remove such false positives from the 'boundariness map' depends on the particular dataset. According to the dataset, we use two segmentation cues: the image projection of the estimated surface and the color information of the image.

We define an intermediate 'boundariness map' e_t where strong image gradients are represented by a low value. This map acts such as a potential well. We compute it in two steps. First, we compute a blurred grayscale version of the image I_t using a Gaussian filter (h, σ). We denote it $I_t^{(h,\sigma)}$. Then, we use the following formula:

$$e_t = \exp\left(-\frac{|\nabla I_t^{(h,\sigma)}|}{s}\right), \qquad (1)$$

where $\nabla I_t^{(h,\sigma)}$ is the gradient of the blurred image $I_t^{(h,\sigma)}$ and s is the bandwidth of the potential well.

Projection-based segmentation. This type of segmentation handles datasets where the object surface has the same color distribution as the background. This is the case of the *floral paper*, *paper fortune teller* and *Kinect paper* datasets. For this case, we use simple morphological image processing techniques to create a boundary mask and propose a three-levels image pyramid to improve the convergence of the boundary constraint.

Firstly, we use the projection of the current estimated surface on the image plane. We erode the region defined by the convex hull of this projection using a square of 40×40 pixels, which gives a rough mask of the foreground R_{fg}. Then, we dilate the region defined by the convex hull of this projection using a square of 90×90 pixels and compute its complementary, which gives a rough mask of the background R_{bg}. Then, we define as R_E the complementary of the union of R_{fg} and R_{bg}. R_E corresponds to the boundary mask, where each pixel belonging to a region close to the surface edges is set to one.

Secondly, we use a three-levels image pyramid which gives coarse-to-fine versions of the 'boundariness' map and which increases thus the convergence basin. We define them as following:

- for the first level, we use
- for the second level
- for the finest level, we use the finest

Color-based segmentation. Color information can be used for datasets where the color distribution of the object surface is different from the one of the background. Such segmentation cue is used for the *creased paper* dataset and for the two new datasets proposed in the supplementary video, the *pillow cover* and the *hand bag* datasets. For this, we create a color-based foreground and background detectors which are trained respectively on the image projection of the surface and on the

This works by applying a color-based foreground detector, trained on the target surface to each input image pixel, and setting I B = 1 for any pixel which has a detection score below a threshold T d . We train the detector using the foreground of input image (in our experiments we use an RGB Gaussian Mixture Model of 4 components) and use a default threshold of T d = 50. In figure 3(b) and (d) we show the difference between the naive boundariness map and the boundariness map using the color-based statistical filter. Here we see that many false boundary edges in the background have been removed.

One future work is to improve the way of computing the 'boundariness map' by making it automatic and generic.

The boundary contour term. This constraint works for surfaces with disc topology. It encourages the surface's boundary contour to lie close to image edges, and was shown to significantly help register surfaces with weak texture [1, 2]. We discretize the boundary of Ω to obtain a set of boundary pixels $\mathcal{B} \triangleq \{\mathbf{u}_{k\in[1,B]}\}$. We then compute a 'boundariness map' for each image $E_t : \mathbb{R}^2 \rightarrow \mathbb{R}^+$ where high values of $E_t(\mathbf{p})$ correspond to a high likelihood of pixel $\mathbf{p}$ being on the boundary contour. The term is evaluated as:

$$C_{bound}(\mathcal{V}_t; E_t) \triangleq \frac{1}{|B|} \sum_{\mathbf{u}_k \in \mathcal{B}} \rho\left(E_t(\pi_t \circ f(\mathbf{u}_k; \mathcal{V}_t)\right). \quad (2)$$

We found that E_t cannot be built naively using for instance an edge response filter, because of many false positives, particularly with background clutter and strong object texture. Instead we build it using an edge response filter that is modulated to suppress false positives according to one or more segmentation cues. The right cue depends on the particular dataset, for example if the background is constant over the image set, or if the object has a distinct color distribution to the background. We give the exact formula for computing E_t for each tested dataset in the supplementary material.

References

[1] M. Gallardo, T. Collins, and A. Bartoli. Can we Jointly Register and Reconstruct Creased Surfaces by Shape-from-Template Accurately? In *ECCV*, 2016.

[2] Q. Liu-Yin, R. Yu, L. Agapito, A. Fitzgibbon, and C. Russell. Better Together: Joint Reasoning for Non-rigid 3D Reconstruction with Specularities and Shading. In *BMVC*, 2016.