Using Shading and a 3D Template to Reconstruct Complex Surface Deformations

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

ISIT, UMR 6284 CNRS
Université d'Auvergne
Clermont-Ferrand, France

Motivations The goal of Shape-from-Template (SfT) is to register and reconstruct the 3D shape of a deforming surface from a single image and a known deformable 3D template. Most SfT methods use only motion information and require well-textured surfaces which deform smoothly. Consequently they are unsuccessful for poorly-textured surfaces with complex deformations such as creases. However, Shape-from-Shading methods permit to reconstruct textureless surfaces and complex deformations since it uses all image pixels and the photometric relationship. We overcome the shortcomings of previous attempts by proposing a novel, *(i)* fully-integrated approach to combine shading constraints with SfT in order to *(ii)* reconstruct complex deformations on all visible regions, both textured and textureless, *(iii)* without any *a priori* photometric calibration.

Template, illumination and camera modeling We define the template as a texture-mapped thin shell 3D mesh in a known reference pose with M vertices. At each time t, each vertex is deformed into the unknowns 3D camera coordinates $\mathbf{x}_t \in \mathbb{R}^{3\times M}$. We upgrade the template with an photometric texture map which defines how each point of the template's surface reflects light. We assume Lambertian model and compute this map using an intensity-based segmentation of the texture-map. It gives constant albedo regions with $\alpha = \{\alpha_1, ..., \alpha_K\}$, the K unknown albedo values. The scene is illuminated by an unknown illumination $\mathbf{l}$ which is constant over time, fixed in the camera coordinates and modeled by spherical harmonics (4 and 9 coefficients). The camera has a linear response, $\beta_t \in \mathbb{R}^+$, which is unknown and time-varying.

Integrated cost function The deformation $\mathbf{x}_t$ is constrained by image data and deformation priors (*isometry* and *smoothing* constraints), and $\mathbf{l}$, β_t and α are constrained by the shading term and the batch of images. We use the *shading* relationship to enforce similarity between the modeled and the measured pixel intensities. As it uses all image pixels, mis-alignement may induce errors. Thus, we use *motion* and *boundary* constraints to align the projected 3D surface with its input image. We also use a robust *smoothing* based on an *M-estimator*, which permits piecewise constant 3D reconstructions, such as creases.

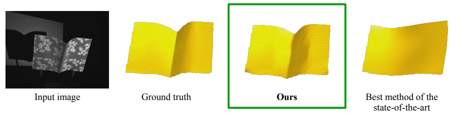

Figure 1: 3D renderings for the input image $n°6$ of the *floral plane* dataset.

Strategy solution The integrated cost function is large scale and highly non-linear, but all constraints are sparse with respect to $\mathbf{x}_t$. We use a cascaded initialization for the four types of unknowns: first $\mathbf{x}_t$, then using a batch of input images $\mathbf{l}$, β_t and finally α. Using Gauss-Newton iterations with line-search, a refinement process minimizes the whole integrated cost function for the batch of images. We found that a dense mesh with vertices of order $\mathcal{O}(10^4)$ is sufficient to capture the creases.

Experimental results We compare our approach on three datasets with four SfT methods and we see that our method is capable of capturing non-smooth deformations, better than others, as figure 1 shows, using shading without any *a priori* photometric calibration, which was not possible with previous methods in SfT or SfS.