MSc Internship proposal

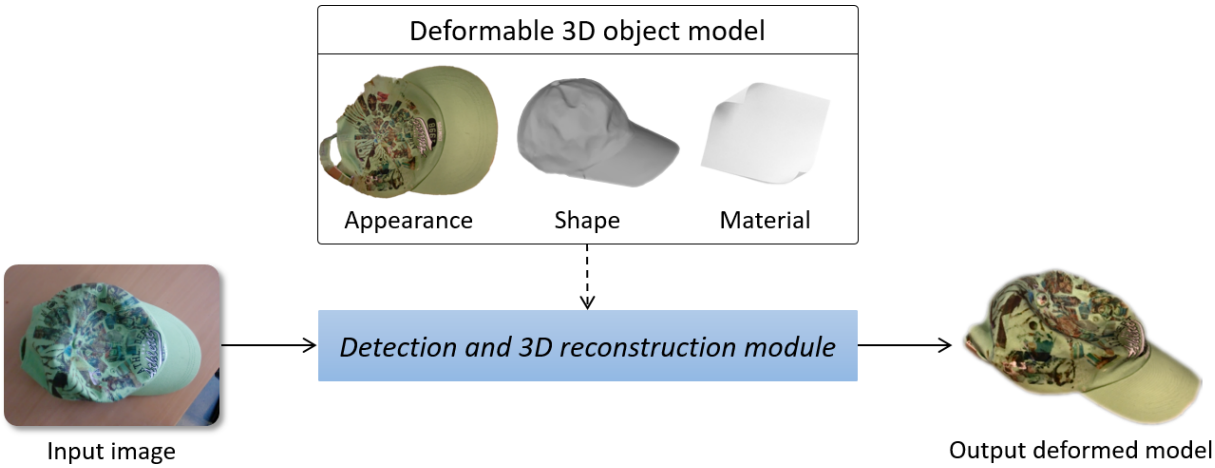
Deformable model-based object detection and 3D reconstruction

Many applications rely on the ability to **detect** objects and reconstruct their **depth** from a single image or a video stream. However, many applications also require that **registration** is computed with a given geometric object model. Here, registration means that any image pixel within the object is explicitly associated to a point on the 3D object model. Applications involving augmented reality for instance will be needing registration. Thinking about it, computing depth is not that difficult. First, it may be solved by hardware solutions such as the Kinect. Second, the state of the art has advanced remarkably with the deep learning approach over the recent years. However, computing the registration, in other words, matching the image to a 3D model at the pixel level, is still a largely open problem. To be clear, we are not just talking about pose computation here, which would involve computing only 6 degrees of freedom, and which has also remarkably progressed recently. Instead, we wish to compute a deformation field between the object model and the input image or video. This is because even if many objects are rigid, and thus pose suffices to explain their images, the standard is rather that objects be **deformable**.

The proposed subject is solving **deformable 3D model to image registration**. Any solution to this problem will also give 3D reconstruction. By 3D reconstruction, we mean that not just a depth map is computed, but also the object's hidden parts, within the camera coordinate frame. This is because, once registration is found, the given 3D model can be deformed to generate the observed state, corresponding to the desired 3D reconstructed. The proposed approach will rely on a deep neural network. The obvious difficulty is that collecting data with groundtruth is difficult if not impossible. Obviously, depth is simple to obtain using, again, the Kinect. Registration however cannot be obtained on real data. We will thus use a combination of supervised learning on simulated images obtained using Blender and self-supervised learning on real images.

The proposal is made by Adrien Bartoli and Nicolas Thome. They have prior expertise in 3D computer vision, object registration, detection, recognition and deep learning.

Apply by email to adrien.bartoli@gmail.com (http://igt.ip.uca.fr/~ab) and nicolas.thome@cnam.fr (cedric.cnam.fr/~thomen).



Proposed internship objective: to develop an object detection and 3D reconstruction module using a deep neural network, whose input is an RGB image and output is a 3D model of the observed object.

Bibliography

Shape-from-Template
A. Bartoli, Y. Gérard, F. Chadebecq, T. Collins and D. Pizarro
IEEE Transactions on Pattern Analysis and Machine Intelligence, 37(10):2099-2118, October 2015

Geometry-Aware Network for Non-Rigid Shape Prediction from a Single View
A. Pumarola, A. Agudo, L. Porzi, A. Sanfeliu, V. Lepetit, and F. Moreno-Noguer
IEEE Conference on Computer Vision and Pattern Recognition (CVPR), *2018*

Monocular 3D Reconstruction of Locally Textured Surfaces
A. Varol, A. Shaji, M. Salzmann and P. Fua
IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(6):1118-1130, June 2012