ERCIM “Alain Bensoussan”
Fellowship Scientific Report

Fellow: Frédéric Larue
Visited Location : CNR, Istituto di Scienza e Tecnologie dell'Informazione (ISTI), Pisa, Italy
Duration of Visit: 9 months

I - Scientific activity

The work achieved during this fellowship period is divided into two different activities.

The first one took place into the frame of the European 3D-COFORM Consortium, which aims at establishing 3D documentation as an affordable, practical and effective mechanism for long term documentation of tangible cultural heritage. The work I realized aims at synthesizing diffuse textures (namely, color information deprived of illumination effects) from the data provided by a 3D scanning system built by a partner of the consortium, enabling to perform real-time acquisition by moving the object by hand in front of the device until the object surface be completely acquired. These data consist, for each scanning session, in a 3D model and a video acquired by a color camera embedded in the scanner, and for which the motion with respect to the object is known.

In such a case, the texture reconstruction process consists, for each surface point of the model, in gathering color samples from the pictures into which the point is visible and to combine them to produce a final color. Due to uncertainty in the camera calibration and color inconsistencies from one picture to the other, naive combining approaches (like a simple averaging or like taking the median value among all samples) generally fail to generate a qualitatively good texture, often leading to visible artifacts like cracks, or misalignment/blurring of fine pictorial features. Moreover, in the general case, the removal of shadows and specular highlights is quite hard, since the lighting environment is not necessarily known or cannot be controlled.

To avoid artifacts in the final texture, we extended some previous work developed in the host research group, consisting in computing the final color at each surface point as a linear combination of all input samples. The weights used are determined from various criteria: the distance to the camera (the closer the viewpoint, the more precisely chromatic details are acquired), the surface orientation with respect to the camera (when the viewing direction becomes too grazing, the distinctness of the observed chromatic information decreases drastically), the proximity of step-discontinuities, namely, significant depth steps perceived by the camera (which correspond to regions where calibration inaccuracy may cause the major problems).

Moreover, the specific hardware configuration of this scanner has made possible the definition of other criteria accounting for lighting. Indeed, during a scanning session, the only light source influencing the appearance of the object comes from the scanner itself, for which the position with respect to the object is perfectly known. Thus, during the color combination, are also introduced in the weight computation criteria enabling to ignore shadowed regions, to reduce influence of highlight in the final reconstructed texture, and to avoid the loss of luminosity that often occurs while applying the aforementioned naive approaches.

While observing 3D models provided by our partner, we also noticed that the geometrical accuracy of the scanner is relatively poor. For this reason, in addition to a color texture, we also implemented the fitting of a normal map, thanks to *shape-from-shading* techniques, in order to capture directly from the input pictures the finest geometric features that cannot be properly recovered by the scanning device. The final result is a 3D mesh with two textures synthesized over its surface, one for the diffuse color and the other containing normal vectors for high quality relighting.

The second activity I worked on concerns the alignment of video sequences onto 3D models, once again for the purpose of color texture reconstruction, but this time (and conversely to the work previously presented) acquired with a standard video camera and in arbitrary conditions. My main contribution to this project until now has consisted in improving the video-to-mesh alignment stage.

The current framework uses *mutual information*, which is a tool that computes a degree of similarity between two images. The principle of the alignment process is then to perform a rendering of the 3D model from the point of view of a virtual camera that must be aligned with the real one, and to compare it to the acquired picture using *mutual information* as a quality estimator. This process is coupled to a minimization algorithm that adjusts iteratively the position of the virtual camera until the rendering of the 3D model matches as best as possible the acquired picture. Like any minimization technique, this approach obviously needs an initial alignment relatively close to the wished solution to be able to achieve good results without falling into local minima. It is then asked to the user to provide the alignment for the first frame of the video, and the alignment for all successive frames is automatically computed by mutual information using as starting point the position of the camera obtained for the previous frame.

The way the rendering of the 3D mesh is performed while computing mutual information is very important, and may impact the quality of the final alignment in a drastic way. Until now, we mostly tried and compared different heuristics for this particular purpose.

Only recently, we have effectively started to explore some smart ways to synthesized purely diffuse textures onto the 3D mesh from the aligned video flow, which includes among other the difficult task of specular highlights detection in the case of an unknown lighting environment. I am currently still involved in this project, in as much as it is already planned to keep on collaborating on it after the end of the fellowship.

II- Publication(s) during your fellowship

A paper on the enhanced scanner (adding color processing to a geometry-only 3D scanner) will be submitted soon to an international conference.

III -Attended Seminars, Workshops, and Conferences

None.

IV – Research Exchange Program (12 month scheme)

Not concerned.