Smart Libraries for Computer Graphics (2013)
Fabio Ganovelli and Paolo Pingi (Editors)

Scene analysis for automatic object segmentation and view suggestion in Assisted Multi-View Stereo Reconstruction

Emanuele Cavarretta[1], Matteo Dellepiane[1], Paolo Cignoni[1] and Roberto Scopigno[1]

[1]Visual Computing Lab, ISTI-CNR, Pisa

Abstract

Multi-view stereo reconstruction methods can provide impressive results in a number of applications. Nevertheless, when trying to apply the state-of-the-art methods in the case of a more structured 3D acquisition, the lack of feedback on the quality of the reconstruction during the photo shooting can be problematic.
In this poster we present a framework for the assisted reconstruction from images of real objects. In particular, the framework is able to separate the object of interest from the background and suggests missing points of view to the user, without any previous knowledge of the shape of the scene and the acquisition path. This is obtained by analyzing the sparse reconstruction and the connection between the reconstructed points and the input images.
The framework has been tested on a variety of practical cases, and it has proved to be effective not only to obtain more complete reconstructions, but also to reduce the number of images needed and the processing time for dense reconstruction.

1. Overview

In contrast with traditional blackbox approaches to MVS [GSC*07, FP10, FCS*10], a new method is proposed which allows the user to verify and control the completeness and quality of the reconstrucion as it occurs. This method is implemented in a fully working framework for Assisted-MVS, that is able to run on average-end laptops, thus allowing the user to perform the reconstruction directly on-site.
The core of the method is a novel scene analysis algorithm, that allows the system to suggest views to the user. This helps improving the coverage of the actual reconstruction. As a first step, the system understands which is the object that is being acquired, and segments it from the rest of the scene. In order to do this, a quality is computed for each point in the reconstruction, based on how much each feature that generates the point is near the center of the image. The quality value is then distributed among points by a smoothing operation, allowing for a better separation of the main subject from the background.

The actual segmentation occurs by finding a separation threshold among the quality values, organized in histograms, by dividing the groups of samples with higher quality from the background data. The defined threshold proved to be effective in both cases of a single object acquisition (tuttotondo) and a whole scene acquisition (panorama-like), see Figure 1. This breaks most of the assumptions made on similar approaches [WDAN07, DF09, TMDug].
Once the object of interest has been determined, view suggestion is supported in two steps. First, the object of interest is analyzed searching for points in areas with lower density, and a candidate point is selected. Then, a number of views are generated such that the selected point is in the center of the framing, and the view direction is perpendicular to that point. The direction is perturbated in order to generate views not too similar to existing ones (see Figure 2).

The proposed system gives the possibility to have a visula feedback about the coverage of a real objects when acquiring it using images. Moreover, it supports the user in completing the acquisition, preventing the shooting of unnecessary images and ensuring the highest completeness.

References

[DF09] DUNN E., FRAHM J.-M.: Next best view planning for active model improvement. In *BMVC* (2009). 1

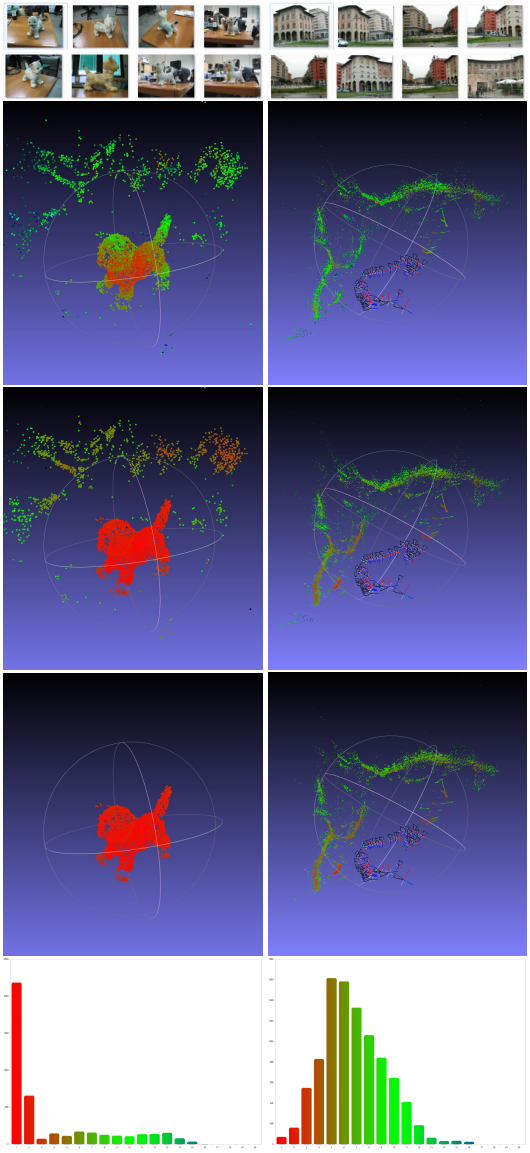

Figure 1: *Two examples of the quality value for background removal (red: highest quality, blue lowest quality). Red indicates high quality, blue low quality. First row: a subset of the images. Second row: the quality value based on the position of the generating feature in every image. Third row: the smoothed quality value. Fourth row: the automatic background removal. Fourth row: the histogram of quality values used for automatic segmentation.*

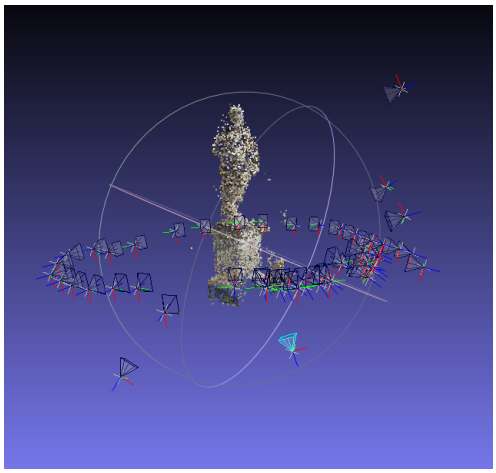

Figure 2: *An example of five suggested views (indicated by red circles) in the context of the acquisition of a statue. The views concentrated on the more detailed part of the object, and try to suggest different directions of view.*

[FCS*10] FURUKAWA Y., CURLESS B., SEITZ S. M., SZELISKI R., INC G.: R.: Towards internet-scale multiview stereo. In *In: Proceedings of IEEE CVPR* (2010). 1

[FP10] FURUKAWA Y., PONCE J.: Accurate, dense, and robust multiview stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence 32*, 8 (2010), 1362–1376. doi:http://doi.ieeecomputersociety.org/10.1109/TPAMI.2009.161. 1

[GSC*07] GOESELE M., SNAVELY N., CURLESS B., HOPPE H., SEITZ S. M.: Multi-view stereo for community photo collections. In *Proceedings of ICCV 2007* (Rio de Janeiro, Brazil, 2007), IEEE, pp. 265–270. 1

[TMDug] TRUMMER M., MUNKELT C., DENZLER J.: Online next-best-view planning for accuracy optimization using an extended e-criterion. In *Pattern Recognition (ICPR), 2010 20th International Conference on* (Aug.), pp. 1642–1645. doi:10.1109/ICPR.2010.406. 1

[WDAN07] WENHARDT S., DEUTSCH B., ANGELOPOULOU E., NIEMANN H.: Active visual object reconstruction using d-, e-, and t-optimal next best views. In *CVPR* (2007). 1