

Adding virtual texture to real objects

Thierry Molinier, David Fofi, Patrick Gorria, and Joaquim Salvi

Structured light provides an automatic, scalable means of adequately rendering ‘look and feel’ without having to use physical markers.

Augmented reality is the name given to the process of seamlessly adding computer-generated information to the real world. A common example is projecting texture onto 3D objects such as archeological artifacts or archival materials that are too fragile to touch. Texture mapping of virtual objects is relatively straightforward because all relevant information—e.g., color and geometry—is perfectly known. In contrast, texturing real, but unfamiliar 3D objects poses a number of challenges. A successful solution to the problem would have a wide range of applications, including sensing, industrial inspection of manufactured parts, reverse engineering, object recognition, as well as in clothing design, virtual museums, and the film industry.

Texturing an object requires estimating its pose (i.e., position and orientation) vis-à-vis the projector used for patterning. The task is particularly difficult because there are no direct relationships between projector and scene. Points of correspondence must be found to ensure acceptable results. This can be accomplished manually, as shown, for example, by a project to illuminate a model of the Taj Mahal, where moving a cross-hair projected onto the physical object registers (aligns) points of interest.¹ An alternative system, called DOME, employs a back-projection screen shaped into a curved surface² that makes it easy to establish the relationships between projector and object. Other approaches use physical markers³ such as pins. Still others manage to project texture by ‘confounding’ the silhouette of the real object with that of the virtually textured object. The problem with all of these techniques is that they either require human intervention or are limited to planar scenes.

We propose an automatic method of adequately projecting texture onto real objects that requires no prior knowledge of the exact pose or any use of physical markers. Our approach, which uses two cameras and one projector, can be generalized to any number of cameras and projectors. We estimate the pose

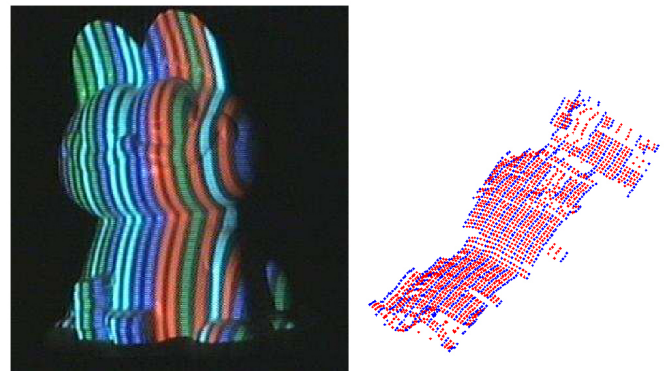


Figure 1. Structured light and 3D reconstruction. Coded structured light is projected onto a textureless model. Cameras then record images from which points of correspondence are extracted, and the 3D position of the points is estimated. Finally, a scan of the model is aligned with the 3D reconstruction.

of the object relative to the projector using ‘structured light,’ a technique based on a coded pattern that can be identified in the image captured by a camera.⁴

The method comprises three steps. The two cameras are first stereocalibrated using Bouguet’s Camera Calibration Toolbox.⁵ A pattern is then projected and identified by the cameras, and its 3D position estimated. Once the correspondence between 3D and 2D points is known, the projector can be calibrated using the method of Roger Tsai.⁶ In the second step, we estimate the pose of the real object with respect to the projector using virtual markers enabled by structured light.⁷ The coded patterns generated by this technique can either be colored or not. We use color, which is labor-intensive because it requires identifying the intensity and hue of each line of light. We estimate the points of correspondence between the cameras and the projector, and reconstruct 3D points from the object’s surface using the calibration from the first step. Figure 1 shows a model illuminated with color-structured light based on a de Bruijn sequence (i.e., each triplet of color is unique⁸) and a 3D reconstruction of

Continued on next page



Figure 2. Blank models.

the model with interpolation to increase the number of points for improved registration at a later stage. We use only one-shot structured-light techniques for reconstruction because they are faster than multiple-shot methods, although not as accurate or as well resolved.

The third step consists of registering our reconstruction with a scan of the object. The accuracy of the registration depends on knowing the movement between these two sets of points. The pose of the scanned model with respect to the projector can be guessed or, alternatively, estimated using a rough registration technique. Subsequently applying a finer scheme—which requires that the two sets of points be close to one another—gives a better result. Once the pose of the scan has been determined, all that remains is to synthesize the view, add texture to the image, and reproject the image onto the object.

We have tested our method and system on several textureless objects with different poses (see Figure 2). Figure 3 (left) shows a structured-light projection of an object (turtle), the synthesized projector view (center), and the real, virtually textured object. Figure 4 shows a number of differently sized and shaped textured models. In almost all cases, the result is satisfactory. For the 3D points, the registration error is less than 0.1mm. We synthesized ten views of the model to simulate rotational movement. Only one registration failed, owing to the curved back of the object: the 3D reconstruction had assumed a different surface.

Conclusion

We have described a complete, automatic method for projecting texture onto real objects without using physical markers. The three-step technique consists of calibration, structured light, and registration. The final step is the most critical, as any error could



Figure 3. (left) Structured-light view, (center) synthesized view, and (right) virtually textured model.



Figure 4. Additional virtually textured models.

significantly affect the object's visual appearance. In future, we intend to replace the last two steps by fusing a silhouette of the object and the synthetic view instead of using 3D models, to make the system more easily adaptable to changes of equipment. We also plan to work with mobile objects in real time.

Author Information

Thierry Molinier, David Fofi, and Patrick Gorria
Laboratoire d'Electronique, Informatique et Image
UMR CNRS 5158
Le Creusot, France

Joaquim Salvi
Computer Vision and Robotics Group (Vicorob)
University of Girona
Girona, Spain

References

1. O. Bimber and R. Raskar, **Spatial Augmented Reality: Merging Real and Virtual Worlds**, A. K. Peters, Natick, MA, 2005.
2. S. Webb and C. Jaynes, *The DOME: a portable multi-projector visualization system for digital artifacts*, **Proc. IEEE Virt. Real.** **2005**, pp. 245–248, 2005.
3. F. Bérard, *The magic table: computer-vision based augmentation of a whiteboard for creative meetings*, **IEEE Worksh. Projector-Camera Syst.**, 2003.
4. T. Molinier, D. Fofi, P. Gorria, J. Salvi, and F. Meriaudeau, *Projector view synthesis and virtual texturing*, **2nd Int'l Topical Mtg. Opt. Sens. Artific. Vision**, 2008.
5. J.-Y. Bouguet, *Camera Calibration Toolbox for Matlab*.
<http://www.vision.caltech.edu/bouguetj/>. Accessed 7 July 2009.
6. J. Salvi, X. Armangué, and J. Batlle, *A comparative review of camera calibrating methods with accuracy evaluation*, **Pattern Recognit.** **35** (7), pp. 1617–1635, 2002.
7. J. Salvi, J. Pagès, and J. Batlle, *Pattern codification strategies in structured light systems*, **Pattern Recognit.** **37** (4), pp. 827–849, 2004.
8. J. Pagès, J. Salvi, and J. Forest, *Optimized de Bruijn patterns for one-shot shape acquisition*, **Image Vision Comput.** **23** (8), pp. 707–720, 2005.