Camera Matchmoving in Unprepared, Unknown Environments

Manolis I. A. Lourakis and Antonis A. Argyros
Institute of Computer Science
Foundation for Research and Technology - Hellas
Vassilika Vouton, P.O.Box 1385, GR 711 10,
Heraklion, Crete, Greece
Email: {lourakis | argyros}@ics.forth.gr

DESCRIPTION

Camera matchmoving is an application involving synthesis of real scenes and artificial objects, in which the goal is to insert computer-generated graphical 3D objects into live-action footage depicting unmodeled, arbitrary scenes. Graphical objects should be inserted in a way so that they appear to move as if they were a part of the real scene. Seamless, convincing insertion of graphical objects calls for accurate 3D camera motion tracking (i.e. pose estimation), stable enough over extended sequences so as to avoid the problems of jitter and drift in the location and appearance of objects with respect to the real scene. In addition to its theoretical interest, camera matchmoving finds several important applications in areas such as augmented reality, virtual studio shooting and the creation of special effects in the post-production/film-making industry.

This work addresses the problem of tracking the 3D motion of a camera in space, using only the images it acquires while moving freely in unmodeled, arbitrary environments. A novel feature-based method for camera tracking has been developed, intended to facilitate tracking in on-line, time-critical applications such as video see-through augmented reality and vision-based control. In contrast to several existing techniques which are designed to operate in a batch, off-line mode, assuming that the whole video sequence to be tracked is available before tracking commences, our method operates on images incrementally, as they are being acquired. Furthermore, it does not rely upon the presence in the environment of fiducial markers or special calibration objects. A brief overview of our approach is given in the next section. More detailed descriptions can be found in [1] and online at `http://www.ics.forth.gr/~lourakis/camtrack/`.

THE APPROACH

Our method is based on a feature-based 3D plane tracking technique, which permits the estimation of the homographies induced by a 3D plane between successive image pairs. At the core of plane tracking lies a homography "chaining" operation that is applied to triplets of consecutive images through a sliding time window and exploits the constraint that all images of a planar surface acquired by a rigidly moving observer depend upon the same 3D geometry. Since the tracked plane is not required to be physically present in the scene, a virtual one can be used instead. Plane tracking is achieved by matching between images the 2D projections of points from all over the scene. By doing so, all information conveyed by corresponding points is taken into account, without the need for continuously maintaining an image-based segmentation of the tracked plane.

Knowledge of the homographies induced by the selected virtual 3D plane between each pair of successive images allows the corresponding projection matrices encoding camera motion to be expressed in a common projective frame and, therefore, to be recovered directly, without estimating the 3D structure. Projective camera matrices are then upgraded to Euclidean and used for recovering the scene structure via triangulation. Following this, the scene structure is employed for refining the projection matrices through local resectioning. The proposed approach is causal (i.e. employs only past frames for deducing the camera pose corresponding to the current one), is tolerant to erroneous and missing feature matches, does not require modifications of the environment and has computational requirements that permit a near real-time implementation.

The accompanying video demonstrates the performance of the approach on several image sequences, which have been acquired using a variety of cameras undergoing different types of 3D motion. In order to achieve matchmoving, the estimated camera motions were exported to the 3DSMAX™[2] graphics package using MAXSCRIPT™and then the augmented sequences were generated with the aid of 3DSMAX's rendering engine that used the original sequence as a background.

ACKNOWLEDGEMENT

This work was partially supported by the EU IST - 2001 - 34545 project LifePlus.

REFERENCES

[1] M.I.A. Lourakis and A.A. Argyros. "Efficient, Causal Camera Tracking in Unprepared Environments," *Computer Vision and Image Understanding Journal*, 2005. Accepted for publication.

[2] Discreet (Autodesk Inc.). "3ds max," *[web page] http://www.discreet.com/3dsmax*, 2005. [Accessed on 20 Apr. 2005].