

Model-Based Visual Tracking for Outdoor Augmented Reality Applications

Reinhold Behringer, Jun Park*, and Venkataraman Sundareswaran

Rockwell Scientific Company (RSC)

{rbehringer, vsundar}@rWSC.com, jpark@cs.hongik.ac.kr

* now working at Hongik University, Seoul, South Korea

Abstract

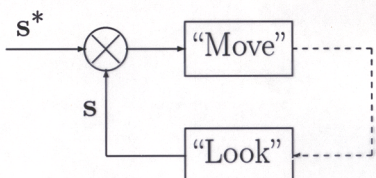
Outdoor Augmented Reality (AR) applications require robust tracking. RSC has developed a real-time visual tracking system that uses visual cues of buildings in an urban environment for improving the registration results of a conventional tracking system. This approach is based on the "Visual Servoing" method. We have demonstrated the feasibility of this approach on real buildings. This poster shows the approach and results. The concept can be generalized to any scenario where a CAD model of the object to be tracked is available.

Introduction

AR applications require a high registration precision [1]. In an outdoor scenario, GPS and digital compass provide position and orientation data, but their accuracy is not sufficient for a seamless and well-aligned overlay onto the user's view. Visual tracking (with a user-worn camera) can improve the tracking precision by using real-world (natural or man-made) features of the environment (e.g., [2]) and provide for example a correction of the drift of inertial trackers [3]. By using the knowledge of exact location of features in the environment, the complete 6 DOF registration can be obtained from visual tracking. RSC has implemented such a system, to be used in the NRL system BARS [4].

The Visual Servoing Approach

Assuming a known initial position/orientation of a camera, its motion in 6 DOF can be obtained by the "visual servoing" approach [5]. This algorithm is based on a minimization of the error in the forward projection of image features (known from a CAD model). The algorithm determines the optimal motion vector in 3D (translation and rotation) from the displacement of the 2D image features in order to provide a transition into the best match of measured 2D features and predicted features, based on the internal assumption of position/orientation. Its application for tracking in an AR system has been described in [6] and is patented (US 6,330,356 B1). It is independent on the type of employed image processing and generally can be used with any 2D feature tracking, as long as the CAD model includes these features.



Acknowledgements:

Part of this work was funded by ONR, contract N00014-99-3-0018.

Image Processing Techniques

We have developed a set of image processing routines for simple extraction of edges (gradient scan) and corners, the latter based on the Kanade-Lucas-Tomasi (KLT) tracker [7]. These routines are also suitable to be ported onto a low-power wearable platform. In the case of a specific demonstration environment, we chose the windows of a building as well trackable features.

Control Structure

Initialization: assumes that the visual tracking module is embedded in a hybrid tracking system providing approximate initial position and orientation. In stand-alone mode: user performs manually the initial feature correspondence by "dragging" features in the video image onto the overlaid (and not registered) CAD model.

Tracking: position and orientation of the camera (resp. the user) are calculated by the visual servoing algorithm.

camera moves → new features are predicted, based on the perspective projection of the internal CAD model.

Implementation

Microsoft Visual C++, on PC, running Windows 2000™.

Set of ActiveX controls for sharing of software components:

Video capture using virtually any commercially available camera (USB, PCI capture board, etc.), supporting VideoForWindows drivers.

Socket communication for platform-independent access of the tracking functionality as a server. A set of functions has been written for integration in a client application, using the GNU CommonC++ library for cross-platform compatibility.

Results:

For developing the algorithms, we have used video footage from our partners in this project at NRL from a campus building. We also built a 3D scale model for live and arbitrary capture. The software in this setup is running on a Pentium-4 CPU with 1.4 GHz with an Osprey video capture board, which allows video capture at 30 fps into a 640x480 pixel sized image from DV input or analog composite/S-video. The CPU load for this video capture procedure on our system is around 54%. The image processing algorithms do not add a significant load.

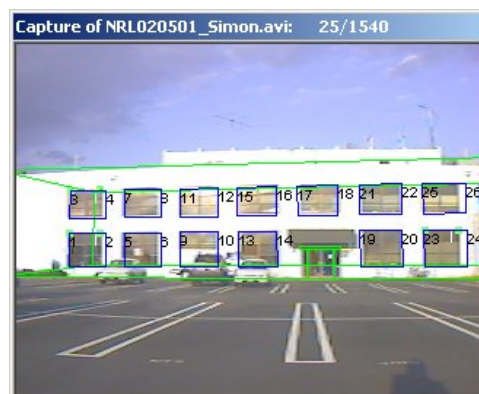
During June 3-7, 2002, we have successfully demonstrated real-time registration with this approach at the NRL campus in Washington DC. The setup for this demo was a Sony Vaio Laptop with a Pentium III CPU, running 750 MHz. The video capture format was chosen to 320x240 pixel at 15 fps, due to limitations of the connected USB camera. The CPU load here was 80% for the complete system, including video capture and image processing.

A fast capture rate proved to be important for reliable tracking, without the displacement getting too large. 30 fps in a 320 pixel wide video provided the capability to track during a rotation rate of approx. 40 deg/sec in stand-alone mode. In the anticipated hybrid mode (with tracking through compass / gyro), this rate describes the additional correction that the visual tracking can provide. We experimented with tracking of up to 28 features simultaneously, without any significant increase of the computational load (up to 60% CPU utilization).

Theoretical precision: 0.01-0.1 deg. Measured: 0.3-0.5 deg. The error sources: calibration, lens distortion, incorrect CAD data

References

- [1] R. Azuma, Y. Baillet, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre, "Recent advances in Augmented Reality." *IEEE Computer Graphics and Applications*, Vol. 21, No. 6, pp. 34-47. IEEE Computer Society, November 2001.
- [2] Ferrari, V., Tuytelaars, T., and van Gool, L., "Markerless AR with a real-time affine region tracker." *ISAR 2001*, New York, October 2001, pp. 87-96.
- [3] Satoh, K., Anabuki, M., Yamamoto, H., and Tamura, H., "A hybrid registration method for outdoor Augmented Reality." *ISAR 2001*, New York, October 2001, pp. 67-76.
- [4] Julier, S., Baillet, Y., Lanzagorta, M., Brown, D., Rosenblum, "BARS: Battlefield Augmented Reality System." NATO Symposium on Information Processing Techniques for Military Systems, 9-11 October 2000, Istanbul, Turkey.
- [5] B. Espiau, F. Chaumette, P. Rives, "A New Approach to Visual Servoing in Robotics." *IEEE Transactions on Robotics and Automation*, Vol. 8, No. 3, pp. 313-326, 1992.
- [6] V. Sundareswaran, and R. Behringer, "Visual Servoing-based Augmented Reality." *First IEEE Workshop on Augmented Reality (IWAR'98)*, San Francisco, November 1998.
- [7] Carlo Tomasi and Takeo Kanade, "Detection and Tracking of Point Features." Carnegie Mellon University Technical Report CMU-CS-91-132, April 1991.



Video courtesy Simon Julier, NRL