INTEGRATED MEDIA SYSTEMS CENTER
A National Science Foundation Engineering Research Center
at the UNIVERSITY OF SOUTHERN CALIFORNIA

PRINCIPAL INVESTIGATOR

Dr. Ulrich Neumann
uneumann@imsc.usc.edu

OTHER IMSC RESEARCHERS

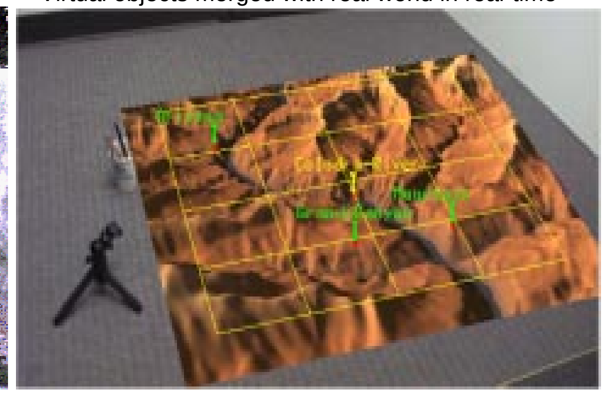
Dr. Suya You
suyay@imsc.usc.edu

TRACKING FOR AUGMENTED REALITY

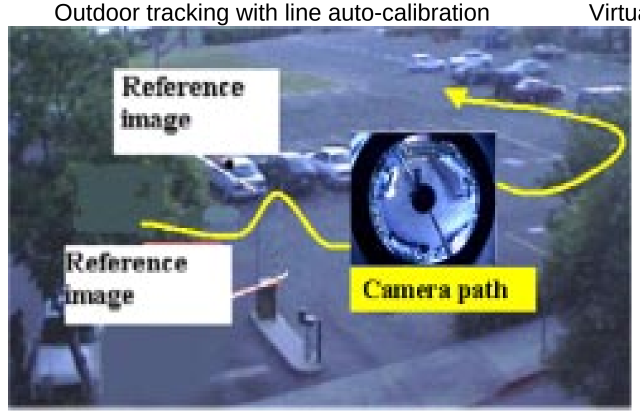
Robust 2D natural feature tracking

Virtual objects merged with real world in real-time

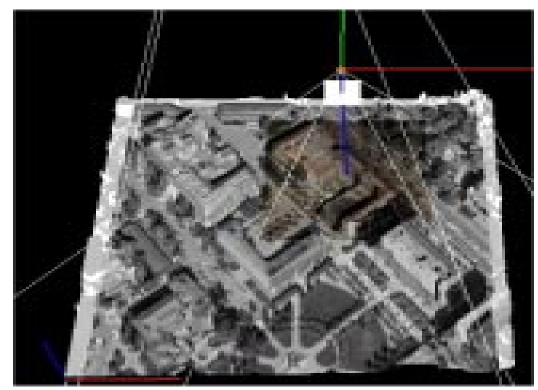
Outdoor tracking with line auto-calibration



Virtual Sand-table with 3D terrain model and text label



Wide area tracking with panoramic images



Fusion of multi-sensor projections

USC STUDENTS, DEGREES	OTHER RESEARCHERS, AFFILIATIONS
Bolan Jiang, Jong Weon Lee, JinHui Hu (Ph.D. students) Anand Srinivasan (MS students)	Dr. Anthony Majoros (Boeing), Avideh Zakhor (Berkeley) Dr. Bill Ribarsky (Georgia Tech)

BRIEF DESCRIPTION OF DEMONSTRATION

Tracking camera pose (position and orientation) enables merging 3D graphics (virtual objects) in real scenes. Tracking enables the creation of a real-time augmented reality (or special effects). Applications include AR training and maintenance instructions; AR entertainment; AR shopping.

DISTINGUISHING CHARACTERISTICS RELATIVE TO STATE-OF-THE-ART

In door tracking (based on fiducials) is scaleable to both near and far fields. Indoor and outdoor tracking makes use of natural visual features and multiple sensor technologies including inertial, GPS, and panoramic image sensors. Visualization merges images and models seamlessly

UNDERLYING TECHNOLOGIES

- Tracking from vision, GPS and inertial sensing
- Line and point feature auto-calibration
- Multi-sensor data fusion
- Real-time image analysis & computing
- Merging of 2D data using 3D model
- 3D graphics & visualization
- Cognitive perception

APPLICATIONS

Typical indoor applications include assembly, maintenance, and other tasks that require information about real objects and how to interact with them. Targets are industries with a need to visualize information in a spatial context to facilitate or speed up work, decrease errors, reduce training time. Military uses include navigation and iconic scene overlays to increase situational awareness of visualize battlefield environments. Entertainment applications include games or sports with real and virtual elements. Commerce and shopping applications include the visualization of virtual products in real environments and interactive browsing of video objects linked to web data.

RECENT HIGHLIGHTS, LEVEL OF DEVELOPMENT, UPCOMING MILESTONES

Offline systems allow tracking objects and camera motion in arbitrary environments. A real time system uses a fiducials (color ring or square landmark) and auto-calibration to allow extendible wide-area tracking for hand-held or head-mounted cameras. Panoramic imaging and inertial systems enable tracking in an uncalibrated outdoor environment. Research in multi-sensor dynamic AR visualization environment combines all manner of images, 3D models, video, and data in a coherent visualization that supports varied media types and layers of abstraction.

BRIEF DESCRIPTION OF UNDERLYING TECHNOLOGIES

1) Robust 2D natural feature tracking

A software algorithm performs two-dimensional Natural Feature (NF) tracking in a sequence of images. Features are detected and tracked as they move due to camera or scene motions. The algorithm is unique in its integration of several features:

a) It selects scene features in an image that are most appropriate for robust tracking.
b) It verifies that features are correctly tracked from frame to frame.
c) It attempts to locate and track a selected number of “best” features regardless of image conditions.

This algorithm executes at between 1-4 frames per second with 640x480 resolution images. The settings of various parameters, including the number of features to be tracked and the maximum rate of feature motion, may significantly affect the computing rate.

2) Real-time landmark detection and identification

A software algorithm detects and uniquely recognizes a set of novel landmarks or fiducials that are easily reproduced on home computer printers. The landmarks are high contrast patterns printed on paper. Their detection is robust over a wide range of lighting and viewing conditions. An alphanumeric or symbol region embedded in their design facilitates unique fiducial recognition from sets of 50-100 different symbols. The detection and recognition is robust to viewing direction and range variations. The detection and recognition of landmarks executes in approximate real time (20-25Hz), depending on the number of landmarks in the scene and the contrast complexity of the scene.

3) Augmented Video with Web Links as an Augmented Reality Modeling Language (ARML)

A software algorithm indexes Internet or other database information based on a use’s cursor motion over tracked points or regions in a video sequence. Video sequences are annotated during an authoring phase. The annotations are URLs or similar data that are meaningful data indices during the real time playback or interacting phase. During playback the user cursor selects objects in the video scenes by proximity or clicking, triggering sounds, speech, or one or more information or Internet browser windows displaying graphics, video, or text. This additional information may also overlay the video scene as an augmented reality. The authoring software application is interactive and includes real time and non-real-time functions. The playback software executes in real time while allowing user interaction.

4) Vision tracking and authoring system

A complete software system that integrates our latest tracking technologies provides users an integrated working environment to develop, author, test, and evaluate their own applications. The system allows users to easily acquire/track/edit/author the video stream as key frames along a timing-line. These included a user-friendly interactive interface, a variety of extendable functional buttons, frame timing-line edit functions, and different task control and information windows to clearly distinguish operational modes. The main system modules include:

a. *Video acquisition*: captures and digitizes video and saves on computer hard drive. Support most existing video capture systems.
b. *Video enhance processing*: performs necessary image processing functions for enhancements of the tracking and visualization.
c. *Database management*: interactive management of the annotation database that includes text label, 2D/3D object, and URL web-link.
d. *Video tracking*: performs feature selection and tracking that support both automatic and manual operations.
e. *Annotation authoring*: interactive authoring process that performs 2D or 3D tracking and annotation.
f. *Media playback*: user views and controls the playback of annotated video.
g. *Web link*: links annotations to URL web data appeared in neighboring browser windows.

The system supports both video and image sequence in several common used file formats, such as avi movie, ppm, and pgm. The authored video can be exported as avi movie file.

5) 3D/6D pose filter

A software algorithm estimates high-dimension parameters from multiple lower-dimension measurements. For example, the 3D position of a point can be estimated from a sequence of 2D images taken from cameras at known positions and orientations. This example describes the use of the filter for autocalibration, or modeling of the scene (see Autocalibration). Likewise, the position and orientation (6D pose) of a camera can be estimated from the positions of multiple 2D image features whose 3D positions are known. This latter example is exactly the calculation needed for tracking a camera's pose (from known fiducials) in order to overlay 3D annotations on a scene. The pose filter executes in real time (30Hz), given the necessary input measurement data.

6) Point and line feature autocalibration

A software algorithm simultaneously estimates the 6D pose of a camera and the 3D parameters of tracked features (points or lines) in the scene. An initial camera pose estimate is computed from a set of known calibrated features. Other features (intentional fiducials (IF) or natural features (NF)), at initially unknown positions, are tracked in the images produced as the camera moves. The IF or NF 3D positions are estimated (automatically calibrated) and their position estimates are used, in turn, to estimate the pose of the camera. This computation iterates and converges to produce both 6D camera pose and 3D IF or NF positions over a sequence of images. The iterative estimation executes in real time (30Hz), given tracked 2D feature positions in an image sequence.

7) Inertial/Vision fusion

A software algorithm integrates vision and gyroscope data from a rigidly connected camera and gyroscope assembly. The algorithm is a flexible framework with a two-channel complementary motion-filter structure that combines the low-frequency stability of vision sensors with the high-frequency tracking of gyroscope sensors, hence, achieving stable static and dynamic six-degree-of-freedom (6DOF) pose tracking. The complementary filter processes data independently, allowing for different sample rates of the sensor systems and reducing the end-to-end system delay. The fusion algorithm executes in real time (30Hz), given real time 3D gyro data and vision tracking data.

8) Wide area tracking with panoramic imaging

A software algorithm estimates 6DOF (six-degree-of-freedom) camera pose over wide areas (indoor or outdoor). By using a panoramic imaging sensor and our innovated tracking method, camera motion pose can be estimated robustly while requiring minimal 3D environment measurements or calibrations (normally 2 measurements are sufficient, in the form of reference images). The 6DOF camera pose is derived directly from a pair of 5DOF motion (orientation and translation direction) estimates measured between the two reference images and tracked images. The pose estimate algorithm executes in approximate real time (25Hz), given the 2D tracking data.

9) Multi-sensor data fusion and visualization

A software technique combines all manner of images, 3D models, video, and data in a coherent visualization that supports varied media types and layers of abstraction. The system uses sensor models and 3D models of the scene to integrate video or image data from different sources. Dynamic multi-texture projections enable real time update and "painting" of scenes to reflect the most recent visual scene data. The dynamic controls, including viewpoint as well as image inclusion, blending, and projection parameters, make for interactive real-time visualization of events occurring over wide areas such as a campus, airport, or battlefield.

10) AR applications

Demonstration applications illustrate the potential use of AR tracking technologies, such as Virtual Sand Table, Virtual Pictures, and Multi-sensor data fusion and visualization.

LIST OF PUBLICATIONS, REFERENCES, URLs

- Suya You and Ulrich Neumann, “Fusion of Vision and Gyro Tracking for Robust Augmented Reality Registration”, *", IEEE Virtual Reality*, March 2001, Japan.
- Bolan Jiang and Ulrich Neumann. “Extendible Tracking by Line Auto-Calibration”. ISAR 2001, USA.
- Majoros, A. E. and Ulrich Neumann. "Support of crew problem-solving and performance with augmented reality. *Proceedings of Bioastronautics Investigators' Workshop*, 2001.
- Suya You and Ulrich Neumann, “Automatic Mosaic Creation Based on Robust Image Motion Estimation”, *IASTED Signal and Image Processing*, Nov. 2000, USA.
- J. W. Lee, Suya You and U. Neumann, “Motion Estimation with Incomplete Information Using Omni-Directional Vision”, *IEEE Workshop on Omnidirectional Vision*, 2000, USA.
- Alok Govil, Suya You and Ulrich Neumann, "A Video-Based Augmented Reality Golf Simulator", Technical demonstration in *ACM Multimedia 2000*, March 2000, USA.
- J. W. Lee, S. You, and U. Neumann. “Large Motion Estimation for Omnidirectional Vision,” IEEE Workshop on Omnidirectional Vision, June 2000.
- B. Jiang, S. You, and U. Neumann. “Camera Tracking for Augmented Reality Media,” IEEE International Conference on Multimedia 2000, August 2000.

http://deimos.usc.edu/~suyay/demo (demo mpg clips)

http://deimos.usc.edu/~jonlee/Research/PanoTracking

http://deimos.usc.edu/~bjiang

http://www.usc.edu/dept/CGIT (CGIT lab web pages)

For additional information, please contact the Principal Investigator listed above via email, or contact

Isaac Maya, Ph.D., P.E. — 213-740-2592
Director, Industry and Technology Transfer Programs — imaya@imsc.usc.edu

Ann Spurgeon — 213-740-4877
Associate Director of Industry Programs — aspurgeo@imsc.usc.edu

Integrated Media Systems Center
3740 McClintock Avenue, Suite 131
Los Angeles, CA 90089-2561
213-740-8931 (fax)

For additional information on the Integrated Media Systems Center (IMSC), please visit our Web site at http://imsc.usc.edu