Demo Abstract: Smart Camera Localization Using the Projection Matrix

John Kassebaum, Nirupama Bulusu, Wu-Chi Feng
Portland State University
{kassebaj, nbulusu, wuchi}@cs.pdx.edu

Categories and Subject Descriptors: C.2.1 **[Network Architecture and Design];** C.2.3 [**Network Operations**]: Computer-Communication Networks – *Network management, network monitoring, public networks.*

General Terms: Algorithms, Measurement, Design.

Keywords: Localization, Smart Camera Networks.

Smart camera network applications such as video surveillance, target tracking, traffic monitoring, and 3D scene reconstruction and analysis either require or are greatly improved by accurate 3D localization of the cameras in the network. Smart camera localization using the projection matrix provides an efficient and scalable method to determine each camera's (*x,y,z*) coordinates and three angles of orientation in a single, global 3D coordinate frame.

PROJECTION BASED LOCALIZATION

Projection based localization works by computing a projection matrix [1] for each camera in the network, then decomposing the matrix for the camera's *extrinsic* parameters, i.e., the camera's (*x,y,z*) position and three Euler rotation angles indicating the orientation of the camera in a 3D coordinate frame. Computing the projection matrix for a camera requires matching a set of 3D world points to the set of pixel coordinates of the world points found in an image taken by the camera. To this end, we propose the use of a feature rich localization target designed such that an efficient, robust algorithm can both detect the target and its feature points. Because the geometry of the target is known, the required set of 3D world points, i.e., the world coordinates of the feature points, is predetermined.

After decomposing a projection matrix, camera position and orientation are given in the target's 3D coordinate frame. Because the reference frame is external to the cameras, any two (or more) cameras simultaneously localizing themselves to the target automatically determine positions and orientations in the same 3D coordinate frame. Then, if one camera's global position and orientation is already known, a rotation and translation vector between its current relative frame and the global frame is computed and sent to the other camera which uses it to realign its relative frame and thus determine its own global position and orientation.

Projection based localization has key advantages over epipolar geometry based localization. Feature point extraction and matching are not done opportunistically which dramatically reduces costs incurred by image processing and message passing. Also, because the target's geometry provides the metric for the global coordinate frame, a centralized bundle adjustment to align all relative localizations is not required.

DEMONSTRATION OVERVIEW

This demonstration will localize a network of COTS cameras each connected to a Crossbow Stargate, and will show that projection based localization using a well-chosen target that fills only 5% of the total frame area determines camera positions accurate to within 1" and orientations to within 3° per orientation axis.

Because showing accuracy requires comparing actual, hand measured positions and orientations of the cameras with computed results, the cameras and target will be positioned on measured stands on a calibrated grid.

Localization of the network will progress by first localizing node 1 to the target. Then after moving the target into the field of view of nodes 1 and 2, nodes 1 and 2 will each localize to the new target position and then node 2 will be globally localized by realignment of its relative frame to the global frame as described above. The rest of the network will be localized in the same fashion.

Because no single localization target design will be suitable for all possible smart camera network deployments, we will use a target appropriate for indoor environments and of a sufficiently small size that the image space occupied by the feature points is less than 5% of the frame area. It is expected that larger deployments can be handled with scaled versions of the target.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. CNS-0722063. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

[1] Hartley, R. and Zisserman, A. Multiple View Geometry in Computer Vision. *Cambridge University Press*, 2000.