LMI-based 2D-3D Registration: from Uncalibrated Images to Euclidean Scene

Danda Pani Paudel[1], Adlane Habed[2], Cédric Demonceaux[1], Pascal Vasseur[3],
[1]Le2i laboratory, CNRS, University of Bourgogne. [2]Icube laboratory, CNRS, University of Strasbourg. [2]LITIS EA laboratory, University of Rouen.

Using information from both 2D and 3D sensors provides several advantages ranging from texture mapping to scene understanding [1, 2]. When the 3D sensor and 2D cameras are free, many systems require the 2D cameras to be internally calibrated and registered with the 3D sensor at all time. We investigate the problem of registering a scanned scene, represented by Euclidean 3D point coordinates, and two or more uncalibrated cameras. The proposed approach assumes camera matrices to be calculated in some arbitrarily chosen projective frame and no calibration or autocalibration is required. When cameras are uncalibrated, the transformation relating the cameras to the scene is projective. Our registration solution is based on a Linear Matrix Inequality (LMI) framework that allows simultaneously estimating this unknown projective transformation and establishing 2D-3D correspondences without triangulating image points. The proposed LMI framework allows both deriving triangulation-free LMI cheirality conditions and establishing correspondences between 3D volumes (boxes) and 2D pixel coordinates. Two registration algorithms, one exploiting the scene's structure and the other concerned with robustness, are presented. Both algorithms employ the Branch-and-Prune (BnP) paradigm and guarantee convergence to a global solution under some mild initial bounding conditions. Our algorithms require initial "box"-2D correspondences with 5 non-overlapping boxes around scene points and/or camera centers. Finding initial bounds on camera positions is relatively easy as far as hand-held or GPS-equipped cameras are concerned. Our 2D-3D registration approach is mainly based on the following propositions:

Proposition 0.1 *Let $\mathcal{S}x = \{(X_j, \mathcal{B}_j)\}_{j=1}^{m}$ be a set of putative point-to-box correspondences: i.e. each point X_j, projecting onto image points $\{x_j^i\}_{i=1}^{n}$, is assigned to its associated box $\mathcal{B}_j$. If all correspondences in $\mathcal{S}x$ are correct, then LMIs*

$$\mathsf{B}_j(\mathsf{H}) \succeq \mathtt{I}, \; j = 1 \ldots m \tag{1}$$

must be simultaneously feasible for the true unknown transformation matrix H *registering the cameras system and scene. Each LMI in (1) is constructed from known image points, projective camera matrices and bounding boxes.*

Proposition 0.2 *Let $\mathcal{S}c = \{(C_i, \mathcal{C}_i)\}_{i=1}^{n}$ be a set of camera-to-box correspondences: i.e. each camera center C_i is assigned to its associated box $\mathcal{C}_i$). If $\mathcal{S}c$ correspondences are correct, then LMIs*

$$\mathsf{D}_i(\mathsf{H}) \succeq \mathtt{I}, \; i = 1 \ldots n \tag{2}$$

must be simultaneously feasible for the true registration matrix H. *Each LMI in (2) is constructed from known projective camera matrices and bounding boxes.*

Key idea: When a set of points and/or camera centers are putatively assigned to bounding boxes $\mathcal{B}_j$ and/or $\mathcal{C}_i$, LMIs (1) and (2) can be simultaneously tested for feasibility. Should they be infeasible, one is guaranteed that at least one point or one camera center has wrongly been assigned to a box.

Algorithms:

- Two registration algorithms, **SSR** and its robust version **RR**, are proposed. Both algorithms exploit the Branch-and-Prune (BnP) paradigm and explore a dynamically-built search tree.
- **SSR** explores the 3D scene's space directly. It processes and subdivides non-empty bounding boxes to which points are assigned in order to iteratively obtain tighter boxes. This algorithm exploits the fact that scanned scenes consist of surface points and much of the explored space is void. **SSR** requires that matched 2D features have their corresponding 3D points scanned.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.

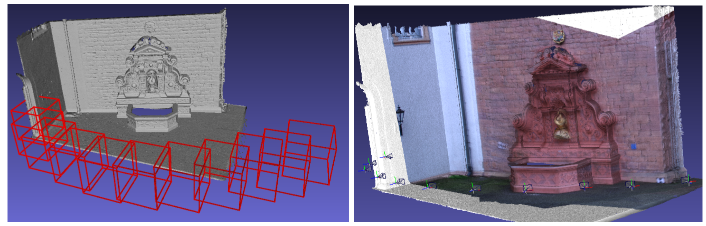

Figure 1: Fountain sequence [3]: (left) 11 camera 2m Bbx and scene, (right) estimated cameras in textured scene using SSR.

Figure 2: Herz-Jesu sequence [3]: (left) matched 2D features with outliers in red, (right) texture-mapped scene using RR.

- **RR** explores the space of parameters defined by the 15 bounded entries of the sought registration matrix. It allows to recursively obtain tighter bounds on this matrix while guaranteeing that at least a predefined number of points are assigned to non-empty boxes.
- Pruning in both algorithms is based on LMI feasibility tests.
- Active nodes correspond to viable point-box correspondences. They are locally boosted by a projective ICP-like 3D-2D refinement.
- The algorithms terminate when the cost of the projective ICP-like refinement reaches a predefined objective or when all branches have been processed (up to bound gap in the branching parameters). In the latter case the best solution is returned.

Results: We tested the proposed methods using synthetic and real images. The results of our experiments in terms of running-time and various registration errors are reported in the paper. Both **SSR** and **RR** have provided very satisfactory results. We also compare the results we obtained with our methods against those obtained using other registration methods from the literature: RISAG [1] and Go-ICP [4]. For qualitative analysis, the results of our registration method were used for texture mapping using real images. Some such results are shown in Figures 1-2. Note that a small error in pose can significantly affect the texture mapping.

[1] Corsini, M. and Dellepiane, M. and Ganovelli, F. and Gherardi, R. and Fusiello, A. and Scopigno, R., Fully Automatic Registration of Image Sets on Approximate Geometry, Int. J. Comput. Vision, 2013.

[2] D. P. Paudel, C. Demonceaux, A. Habed, and P. Vasseur, Localization of 2D Cameras in a Known Environment using Direct 2D-3D Registration, International Conference on Pattern Recognition (ICPR), 2014.

[3] Strecha, C. and von Hansen, W. and Van Gool, L. and Fua, P. and Thoennessen, U.On benchmarking camera calibration and multi-view stereo for high resolution imagery, CVPR, 2008.

[4] J. Yang, H. Li, and Y. Jia., Go-icp: Solving 3d registration efficiently and globally optimally, ICCV, 2013,