Towards Geographical Referencing of Monocular SLAM Reconstruction Using 3D City Models
Application to Real-Time Accurate Vision-Based Localization

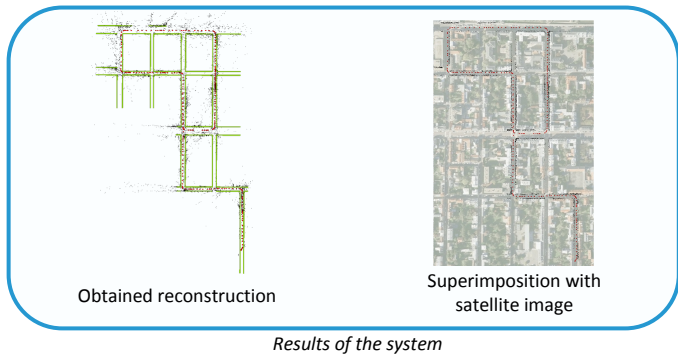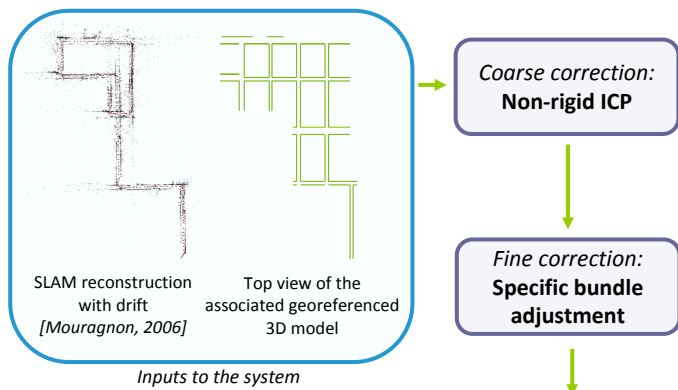P. Lothe[1]   S. Bourgeois[1]   F. Dekeyser[1]   E. Royer[2]   M. Dhome[2]

[1] CEA LIST, LSVE, Saclay, France   [2] LASMEA (CNRS / UBP), Clermont-Ferrand, France

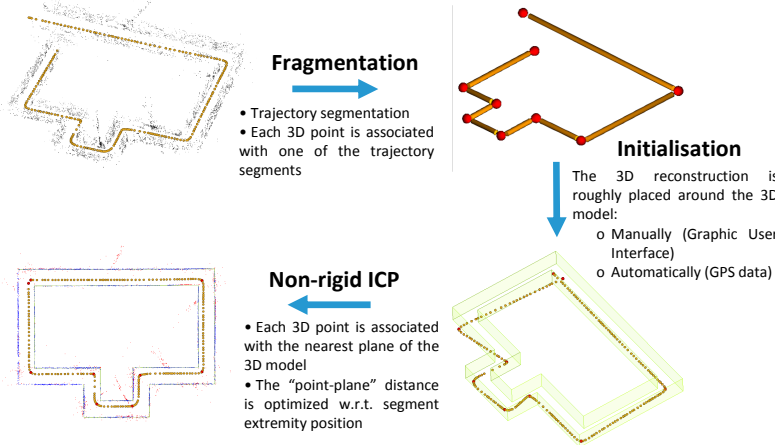**pierre.lothe@cea.fr**

LASMEA

1. Objective & Overview

▪ **Objective:** Feature landmark database construction for vision based localization.

▪ **Approach:** Fusion of monocular Structure-From-Motion and coarse 3D model of the environment **to correct SfM method drift**.

▪ **Contribution:** Image and CAD model errors are taken into account during the optimization.

Offline process

SLAM reconstruction with drift
[Mouragnon, 2006]

Top view of the associated georeferenced 3D model

Inputs to the system

Coarse correction:
Non-rigid ICP

Fine correction:
Specific bundle adjustment

Obtained reconstruction

Superimposition with satellite image

Results of the system

Current frame

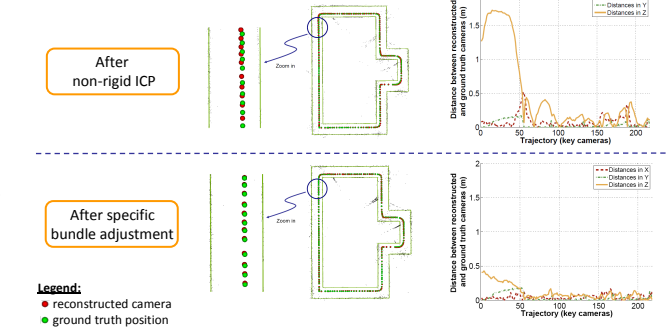**Online** Localization

Augmented frame

2. Non-Rigid ICP

▪ **Hypothesis**: drift negligible during turnings.

⇨ **Piecewise similarity transformation:** camera and 3D point positions are optimized w.r.t. the segment they belong to.
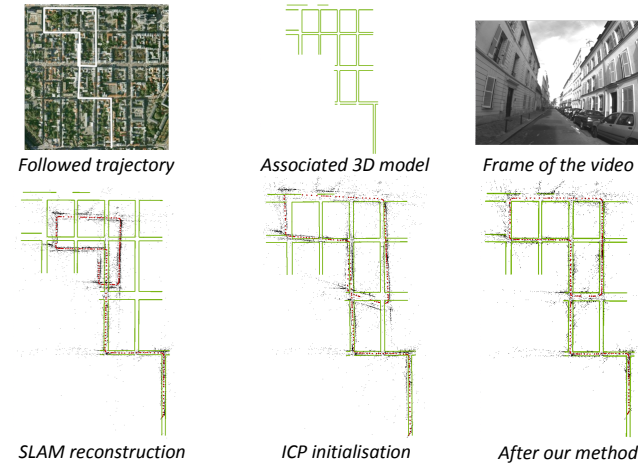

Fragmentation

• Trajectory segmentation
• Each 3D point is associated with one of the trajectory segments

Initialisation

The 3D reconstruction is roughly placed around the 3D model:
  o Manually (Graphic User Interface)
  o Automatically (GPS data)

Non-rigid ICP

• Each 3D point is associated with the nearest plane of the 3D model
• The "point-plane" distance is optimized w.r.t. segment extremity position

3. Specific Bundle Adjustment

1) **Bundle Adjustment**: Geometric and vision-based constraints are fused in a single term. Only camera parameters are optimized. The cost function is:

Legend:
- $\mathcal{C}_i^j$ Camera number « j » observing $\mathcal{Q}_i$
- $\mathcal{Q}_i$ 3D point
- q_i^j 2D observation of the 3D point « i » in the camera « j »

(a) 3D point-plane association.

(b) $\mathcal{Q}_i$ 2D observations backprojection onto its associated plane.

(c) Obtained 3D points barycentre computation.

(d) Barycentre projection into cameras images.

2) **3D point triangulation:** each 3D point position is then obtained from the new computed pose of the cameras.

3) **Loop:** Step 1 & 2 are iterated several times.

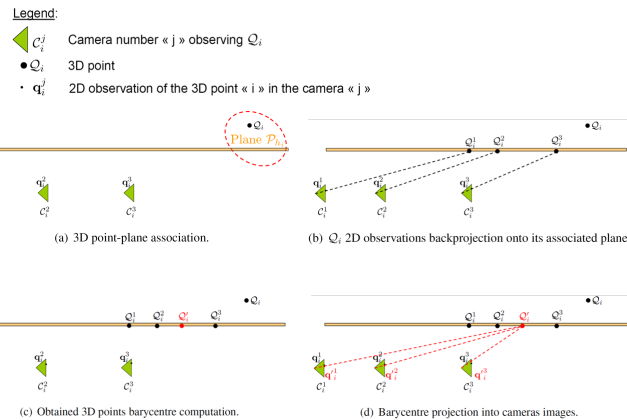▪ **Numerical comparison (on a synthetic sequence):**

After non-rigid ICP

After specific bundle adjustment

Legend:
● reconstructed camera
● ground truth position

4. Results

▪ **Real data (1.5 kms)**

Followed trajectory   *Associated 3D model*   *Frame of the video*

SLAM reconstruction   *ICP initialisation*   *After our method*

5. Conclusion

✓ Creation of georeferenced database for large scale sequences

✓ Interests:
  • 3D model improvement with visual features
  • SLAM drift correction