Large Scale Vision-Based Navigation Without an Accurate Global Reconstruction

Sinisa Savic, Anthony Remazeilles, Albert Diosi and Francois Chaumette
http://www.irisa.fr/lagadic, IRISA/INRIA Rennes, France

Introduction

- We present a novel vision framework for scalable mapping and localization, enabling fault-tolerant feature-oriented appearance-based navigation in outdoor environments:
  - the mapping and navigation stages are considered separately, as an interesting and not completely solved problem;
  - the employed hierarchical environment representation has a graph of key-images at the top (scalability), and local 3D reconstructions at the bottom (feature prediction);
  - both mapping and localization framework components rely on a differential tracker featuring warp with isotropic scaling and affine contrast compensation [4].

- The topological representations have been used by the robotic community, in the context of range sensors: we introduce this idea to computer vision, together with [2-4].
  - Although related to [2,3], our work is novel since it better exploits the power of multi-view geometry techniques.
  - We share the problem with [3], but employ different techniques and do not require global consistency.
  - By posing weaker requirements we obtain scalable real-time mapping and can close loops regardless of the drift.
  - Experiments with 15000 landmarks have been performed without any performance degradation.

Mapping

- We consider the learning sequence loop-clouds, taken along a circular path of approximately 50m.
  - Fig. loop1 illustrates the mapping sensitivity w.r.t. parameters: (i) minimum count of features $n_i$, (ii) maximum allowed reprojection error $\epsilon$, and (iii) the tracking RMS threshold $R$. The presence of node $i$ indicates that the cycle at the topological level has been closed by wide-baseline matching.
  - The right-most map in Fig. loop1 was deliberately constructed using suboptimal parameters: our navigation approach works regardless of the accumulated drift.
  - Fig. loop2(left) illustrates the capability of the localization component to traverse a topological cycle: the navigation sequence was obtained during two rounds roughly along the same circular path.
  - The first round was used for mapping (loop-clouds, cf. Fig. loop1) while the localization is performed along the combined sequence, involving two complete rounds.
  - The localization was successful in both rounds.
  - Fig. loop(emp) shows results on sequence loop-clouds acquired along a similar circular path in bright sunlight.
    - Despite the different imaging conditions (cf. Fig. loop1), the localization was successful except in arc 10, 11 and 12.
    - The illumination difficulties were aggravated by a featureless tree and a considerable curvature of the learning path.
  - Fig. loop3 shows the localization after the reinitialization in arc 13, where the buildings behind the tree begin to appear.

Real-time navigation results

- The robot control has been implemented by a simple visual servoing scheme, the steering angle $\psi$ is determined as:
  - $\psi = \frac{1}{\lambda x} \sum_i \lambda_i$, where $\lambda_i \in \mathbb{R}^+$. We show an experiment on the map shown in Fig. nav1, offering a variety of driving conditions: narrow corners, slopes and driving under a building and trees for more experiments.
    - The speed was set to 30m/s in turns, otherwise 50 m/s.
    - Five reinitializations were required, as shown in Fig. nav1.
    - Between A and B, the robot drove above 740m despite the occlusions as shown in Fig. nav2.

Fig. 1

Localization

- Initialization: wide-baseline matching the current image with the two key-frames incident to the actual arc:
  - the correspondences are used to recover the three-view geometry (3Vg), and locate the mapped features.

- Feature prediction: the tracked features are used to estimate two-view geometries $W_{i+1}$ and $W_{i-1}$ for a more accurate loop-closure (i.e., $W_{i+1}=W_{i}$) by a decomposed approach.

Conclusions

- The poster presents a novel framework for scalable feature-oriented appearance-based navigation.
- The framework combines 3D prediction with 2D navigation:
  - the global topological map ensures unlimited scalability;
  - the local geometric representation enables recovery from tracking failures through feature prediction;
  - the desired feature positions in the next key-image allow simple and effective 3D navigation by visual servoing.
- The framework allows large-scale navigation without requiring a geometrically consistent global view of the environment.
  - in the experiment with a circular path, the navigation proceeds regardless of the extent of the drift.
- The framework is applicable in interconnected environments, where global consistency may be difficult to enforce.
- Very encouraging results in real experiments:
  - real-time navigation in public areas with other moving objects and moderately-to-large changes in imaging conditions;
  - difficulties: illumination variations, featureless areas, sharp turns, nearby vegetation.
- The main performance bottleneck is CPU power:
  - most of the processing time is spent within the multiresolution point feature tracker [6]; the mapping and localization throughput on a notebook with CPU equivalent to Pentium 450GHz is 5 Hz and T(X) (a 256x256 image) on 320x240 grayscale images;
  - a working system using only small images suggests that vision-based autonomous transportation is getting close.

Selected references


1. This work has been supported by the French national project Froliki sailing, for the project HATX project, ANR 08.
2. This work has been supported by Lagadic, IRISA/INRIA Rennes, France.
3. This poster was prepared with LATEX, with alternative color and grayscale versions.