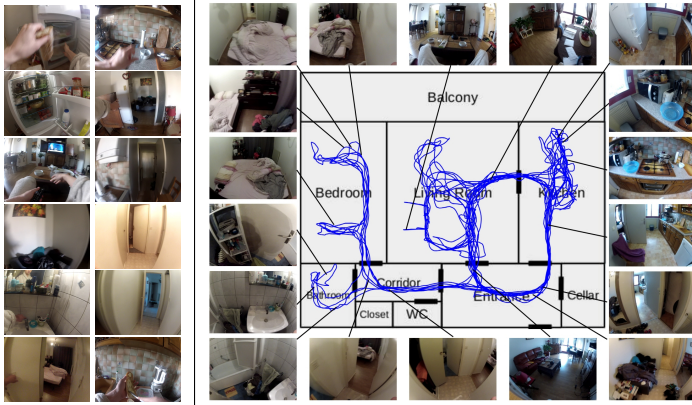


1 Introduction

In this paper, we are interested in Visual Indoor Localization (VIL) for challenging video sequences coming from a single monocular camera where the person wearing the camera performs daily living activities (see Fig.1(a)). The difficulty of this problem resides in the fact that: i) hand-held objects are frequently interposed between the camera and the environment; ii) strong motion blur and differences in illumination occur; iii) the environment changes between the images of the database and the video frames to localize, and the viewpoints can be significantly different.

We wish to develop a method that:

- relies only on the images coming from the wearable camera, i.e no other sensor such as Inertial Measurement Units should be used
- estimates the camera position with a sub-meter level accuracy as well as its orientation
- is consistent with the topology of the environment, i.e the camera trajectory should not cross walls
- is able to detect when the data is not sufficient to disambiguate the situation, i.e when the posterior distribution of the camera trajectory is multi-modal and/or too dispersed.



(a) Video frame examples of a person performing daily living activities. Hand-held objects are frequently interposed between the camera and the environment. (b) Illustration of a database of images of an apartment annotated with map coordinates. This database of 6000 images was generated automatically from a training video sequence. The blue line corresponds to the 2D positions of all the images of the database. Note that each image is actually annotated with a 6-dof pose (3D position and 3D orientation) but only the 3D position projected onto the plan of the apartment is presented in this figure.

Figure 1: (Best seen in colors) Left: Examples of video frames to localize. Right: Illustration of the database

2 Contribution

In this context, we propose a novel VIL framework which is able to satisfy the previous technical specifications. The contribution of this paper is threefold:

1. We formulate the localization problem as a target tracking problem on the Lie group of camera motions SE(3), where the measurements are map coordinates obtained by applying a Content Based Image Retrieval (CBIR) algorithm to the video frames.
2. In order to solve this problem, we derive a novel Rao-Blackwellized Particle Smoother on Lie Groups (LG-RBPS), which builds upon the recently proposed Extended Kalman Filter on Lie groups [1] and the Rauch-Tung-Striebel Smoother on Lie groups that we also derive in this paper.
3. To take into account the topology of the environment, we propose to introduce Virtual Measurements (VM) that guide the particles and prevent them from crossing walls.

3 Results

The proposed VIL framework is evaluated experimentally on several challenging video sequences where the person wearing the camera performs daily living activities.

	GO80	GO81	GO82	GO83	GO84	GO85
CBIR only (similar to [2])	1.7	1.3	2.4	2.2	1.3	2.2
CBIR + LG-RBPS No VM	0.5	0.7	1.7	0.9	<0.5	1.3
CBIR + LG-RBPS With VM	0.5	<0.5	0.8	0.7	<0.5	0.9

Table 1: Results on 6 challenging video sequences (GO80,..., GO85). Examples of video frames from these videos are presented Fig.1(a). The figures represent the RMSE in meter of the estimated trajectories w.r.t the ground truth which has an accuracy of 0.5m.

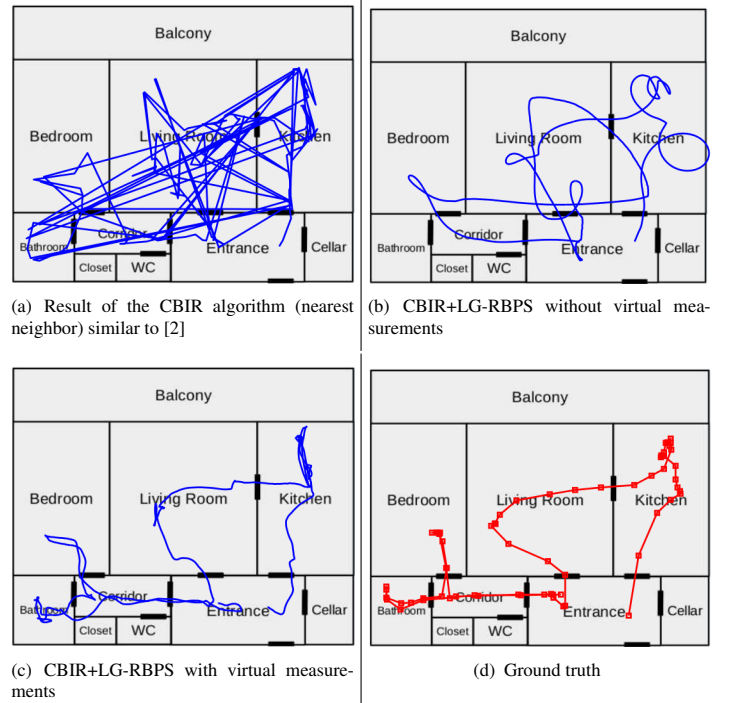


Figure 2: Illustration of estimated trajectories on the video sequence GO82. Only the projection of the 3D position onto the apartment plan is presented.

- [1] G. Bourmaud, R. M  gret, A. Giremus, and Y. Berthoumieu. Discrete extended Kalman filter on Lie groups. In *Signal Processing Conference (EUSIPCO), 2013 Proceedings of the 21st European*, 2013.
- [2] C.   . Conaire, M. Blighe, and N. E. O'connor. Sensecam image localisation using hierarchical SURF trees. In *Advances in Multimedia Modeling*, pages 15–26. Springer, 2009.

Acknowledgments The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007- 2013) under grant agreement 288199 - Dem@Care.