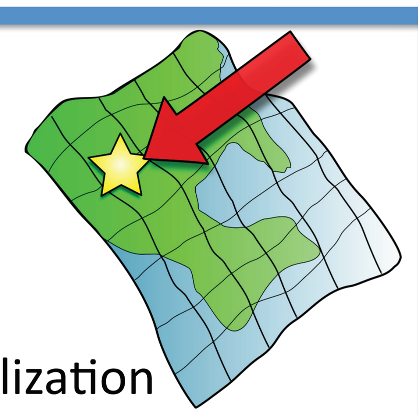
Lost! Leveraging the Crowd for Probabilistic Visual Self-Localization

Marcus A. Brubaker•, Andreas Geiger+ and Raquel Urtasun•

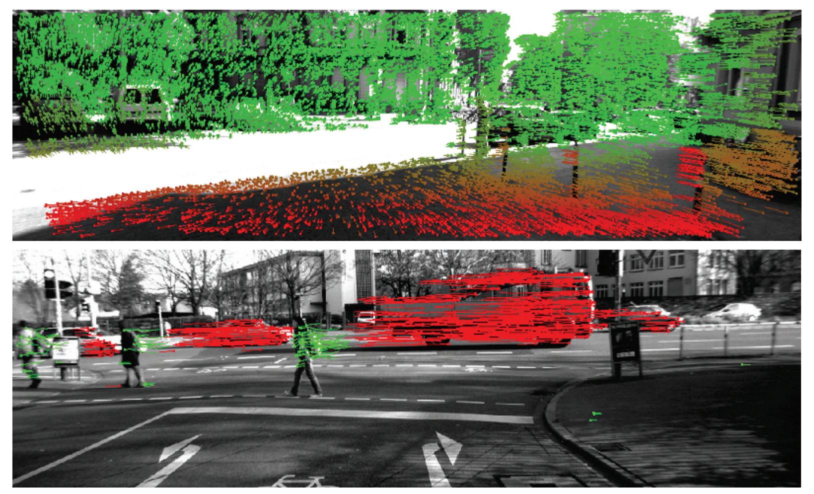
• Toyota Technological Institute at Chicago, + Karlsruhe Institute of Technology & MPI Tübingen

Introduction



- Localization is a critical part of any autonomous system
- GPS has limited availability; can be blocked or degraded
- Place recognition techniques rely on visiting locations before localization *[Dellaert et al, ICRA 1999; Thrun et al, AI 2001; Hays and Efros, CVPR 2008; Schindler et al, CVPR 2008; Crandall et al, WWW 2009; Kalogerakis et al, ICCV 2009]*
- Humans are able to localize given only a map of a region, can we do the same with a vision system?
- High-quality community developed maps are now freely available (OSM), making this a low-cost option
- We exploit the visual odometry to localize a vehicle in a given map to an accuracy of 3.1m on average
- ***Source code:*** *http://www.cs.toronto.edu/~mbrubake*

Localization using Visual Odometry

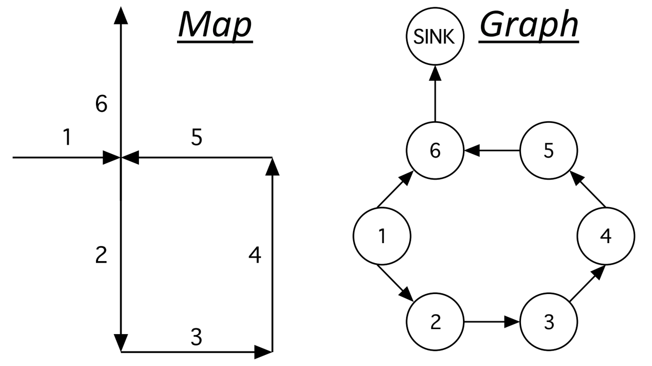


[Geiger et al, IV 2011]

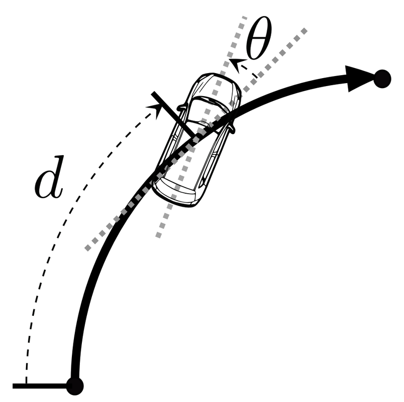
- Motion provides weak cues about location
  - ➢ Turns, curves and straight driving can limit possible locations in a region
  - ➢ Short sequences can be highly ambiguous
  - ➢ Visual odometry is noisy and suffers from drift over longer sequences
- Approach must be able to cope with high degree of uncertainty and ambiguity

Map-based Location Representation

- Map data is conveniently represented as a graph
  - ➢ Nodes u represent street segments
  - ➢ Edges represent connectivity between streets



- Given the street node, the vehicles position represented in terms of position and orientation on the street segment
  - ➢ d is the distance from the start of the street segment
  - ➢ θ Is the heading relative to the street segment



Probabilistic Localization with Visual Odometry

- The unknown state includes
  - ➢ u_t is the current street segment, and
  - ➢ $\mathbf{s}_t = (d_t, \theta_t, d_{t-1}, \theta_{t-1})$
- Odometry observations $\mathbf{y}_t$ are assumed to be corrupted with IID Gaussian noise

$$\mathbf{y}_t | u_t, \mathbf{s}_t \sim \mathcal{N}\left(\mathbf{M}_{u_t}\mathbf{s}_t, \Sigma^{\mathbf{y}}_{u_t}\right)$$

where $\mathbf{M}_{u_t}$ computes the change in position and orientation

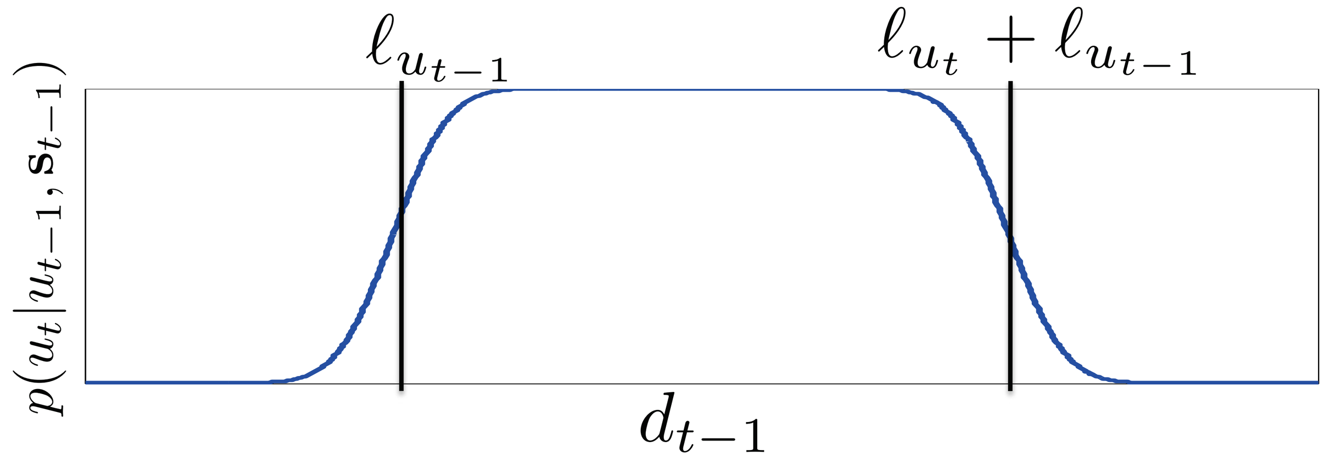
- A second order linear process, corrupted by Gaussian noise, is assumed for the continuous pose variables $\mathbf{s}_t$

$$\mathbf{s}_t | u_t, u_{t-1}, \mathbf{s}_{t-1} \sim \mathcal{N}\left(\mathbf{A}_{u_t,u_{t-1}}\mathbf{s}_{t-1} + \mathbf{b}_{u_t,u_{t-1}}, \Sigma^{\mathbf{s}}_{u_t}\right)$$

where $\mathbf{A}_{u_t,u_{t-1}}$ computes a constant velocity model

- Given the length of street segments ℓ_u and the connectivity defined by the street graph, one can derive the street transition probability to be:

$$u_t | u_{t-1}, \mathbf{s}_{t-1} \sim p(\mathbf{s}_t \text{ will be on } u_t | u_{t-1}, \mathbf{s}_{t-1})$$



- To represent the posterior

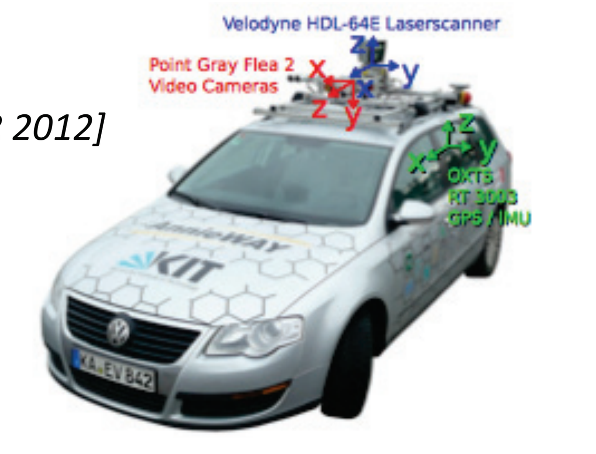
$$p(u_t, \mathbf{s}_t | \mathbf{y}_{1:t}) = p(\mathbf{s}_t | u_t, \mathbf{y}_{1:t}) p(u_t | \mathbf{y}_{1:t})$$

- Continuous portion represented with Mixture of Gaussians

$$p(\mathbf{s}_t | u_t, \mathbf{y}_{1:t}) = \sum_{i=1}^{N_{u_t}} \pi^{(i)}_{u_t} \mathcal{N}\left(\mathbf{s}_t | \mu^{(i)}_{u_t}, \Sigma^{(i)}_{u_t}\right)$$

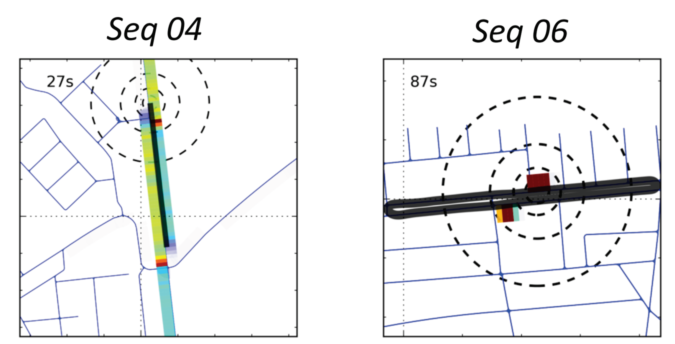
- Inference exploits Gauss-Linear structure of the model using a mix of Kalman filter-like updates and Monte Carlo approximations
- Derive a general algorithm to simplify mixture models to prevent the computational costs from growing

Experimental Results



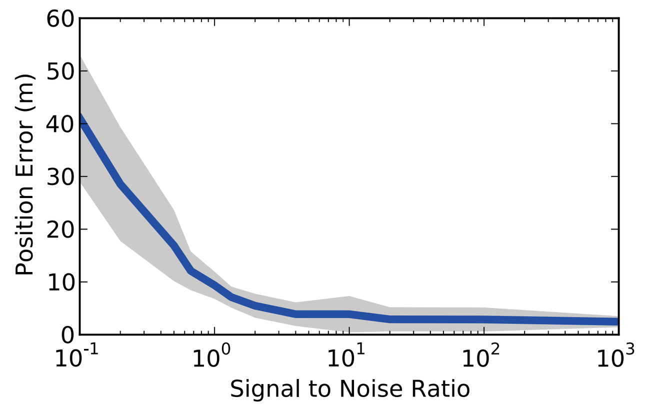
- Method validated on visual odometry sequences from the KITTI dataset *[Geiger et al, CVPR 2012]*
- Stereo and monocular odometry computed LIBVISO2 *[Geiger et al, IV 2011]*
- Error measure: heading angle and position
- GPS-based odometry and map projection error computed for comparison

Ambiguous Sequences

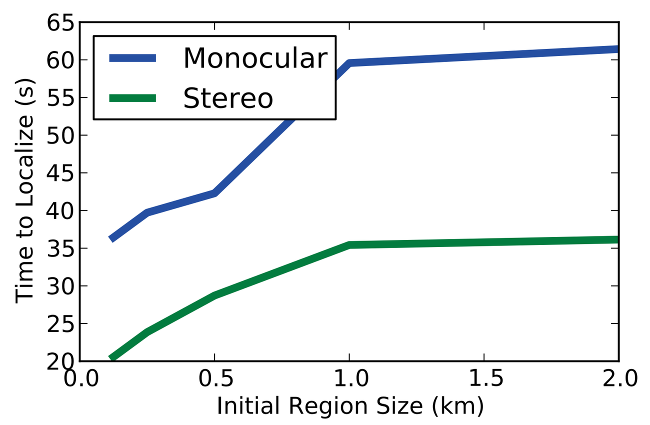


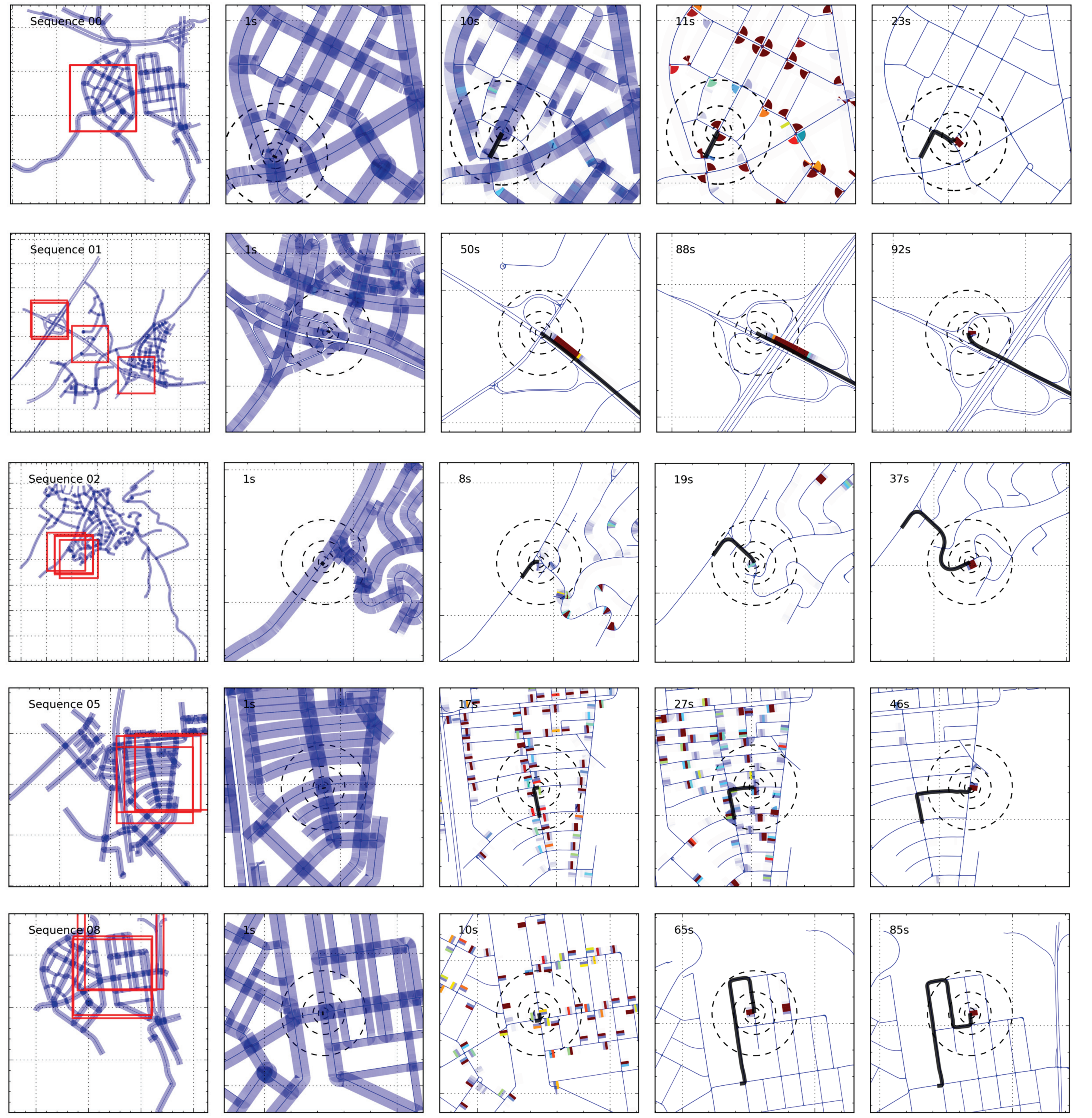
		00	01	02	03	04	05	06	07	08	09	10	Average
Position	Monocular	15.6m	*	8.1m	18.8m	*	5.6m	*	15.5m	45.2m	5.4m	*	18.4m
	Stereo	2.1m	3.8m	4.1m	4.8m	*	2.6m	*	1.8m	2.4m	4.2m	3.9m	3.1m
	GPS	1.8m	2.5m	2.2m	6.9m	*	2.7m	*	1.5m	2.0m	3.8m	2.5m	2.4m
	Map	0.8m	1.3m	1.0m	2.5m	3.9m	1.3m	1.0m	0.6m	1.1m	1.2m	1.1m	1.44m
Heading	Monocular	2.0°	*	1.5°	2.4°	*	2.0°	*	1.3°	10.3°	1.6°	*	3.6°
	Stereo	1.2°	2.7°	1.3°	1.6°	*	1.4°	*	1.9°	1.2°	1.3°	1.3°	1.3°
	GPS	1.0°	1.0°	0.8°	1.4°	*	1.2°	*	1.5°	1.0°	0.9°	1.0°	1.0°

Robustness to Noise



Influence of Region Size





Large Scale Maps

- 18km^2 with 2,150km of road
- See video for more results