Expert Sample Consensus Applied to Camera Re-Localization

Eric Brachmann and Carsten Rother - Heidelberg University (HCI/IWR)

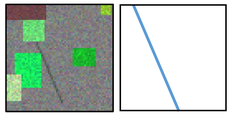
Visual Learning Lab Heidelberg

UNIVERSITÄT HEIDELBERG ZUKUNFT SEIT 1386

Code and trained models:

Problem Statement

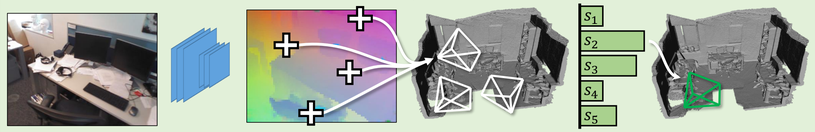
Estimate a parametric model from an observation. Observations are **complex** and **ambiguous**. There is no analytical solution to go from observation to model.



Observation Fitted Model

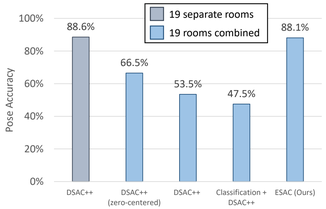
Our application: Estimate the **6D camera pose** (position + orientation) relative to a known scene **from a single RGB image**.

State-of-the-art: DSAC++ [Bra18] (our previous work)



observation scene coordinate regression hypothesis sampling scoring selection fitted model

Scalability issues in large and ambiguous environments:



Contributions

- Expert Sample Consensus (**ESAC**): a combination of **Differentiable RANSAC** (DSAC) and **Mixture of Experts** (MoE) [Jac91]
- ensemble of networks **trained jointly and end-to-end**
- **large-scale camera re-localization** in ambiguous environments

[Bra18] Learning Less is More - 6D Camera Localization via 3D Surface Regression, Brachmann and Rother, CVPR18
[Bra17] DSAC - Differentiable RANSAC for camera localization, Brachmann et al., CVPR17
[Jac91] Adaptive Mixtures of Local Experts, Jacobs et al., Neural Computation 1991
[Sho13] Scene coordinate regression forests for camera relocalization in RGB-D images, Shotton et al., CVPR13
[Val16] Learning to navigate the energy landscape, Valentin et al., 3DV16
[Sat18] Benchmarking 6dof outdoor visual localization in changing conditions, Sattler et al., CVPR18
[Sat16] Efficient & effective prioritized matching for large-scale image-based localization, Sattler et al., TPAMI16

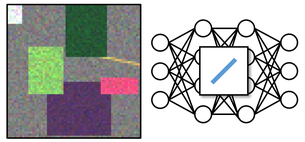
erc European Research Council Established by the European Commission

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 647769)

European Commission Horizon 2020 European Union funding for Research & Innovation

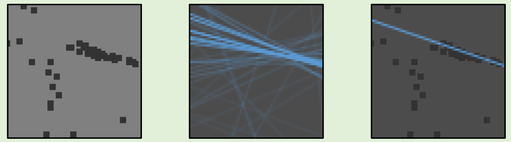
DSAC [Bra17] for a Toy Problem

Which line?



observation I network f_e

RANSAC / DSAC



prediction $\mathcal{Y}_e$ hypotheses $\mathbf{h}_j$ fitted Model $\hat{\mathbf{h}}$

RANSAC: $\hat{\mathbf{h}} = \mathbf{h}_j$ with $j = \underset{j}{\operatorname{argmax}}\, s(\mathbf{h}_j, \mathcal{Y}_e)$ with $\mathcal{Y}_e = f_e(I; \mathbf{w})$

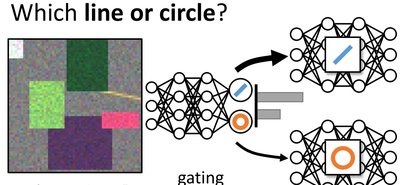
DSAC: $\hat{\mathbf{h}} = \mathbf{h}_j$ with $j \sim p(j) = \frac{\exp(s(\mathbf{h}_j, \mathcal{Y}_e))}{\sum_{j'} \exp(s(\mathbf{h}_{j'}, \mathcal{Y}_e))}$

soft inlier count: $s(\mathbf{h}, \mathcal{Y}) = \alpha \sum_{\mathbf{y} \in \mathcal{Y}} \text{sig}[\tau - \beta d(\mathbf{y}, \mathbf{h})]$

DSAC Training Objective: $\mathcal{L}(\mathbf{w}) = \mathbb{E}_{j \sim p(j)}\left[\ell(\mathbf{h}_j)\right]$

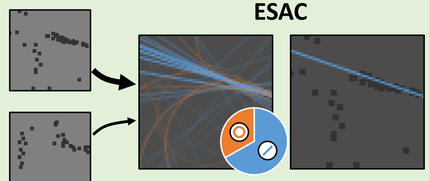
ESAC for a more Complex Toy Problem

Which **line or circle**?



observation I gating network g experts f_e

ESAC



predictions $\mathcal{Y}_e$ hypothesis pool $\mathcal{H}$ (line **and** circle hypothesis) fitted model $\hat{\mathbf{h}}$

The gating network **distributes model hypotheses** among experts: $p(\mathcal{H}) = \frac{N!}{\prod_e n_e!} \prod_e g(e, I; \mathbf{w})^{n_e}$

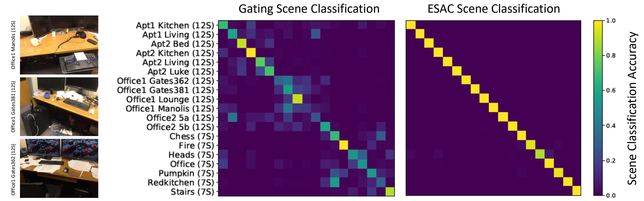
ESAC: $\hat{\mathbf{h}} = \mathbf{h}_{(e,j)}$ with $(e, j) \sim p(e, j|\mathcal{H}) = \frac{\exp(s(\mathbf{h}_{(e,j)}, \mathcal{Y}_e))}{\sum_{e'} \sum_{j'} \exp(s(\mathbf{h}_{(e',j')}, \mathcal{Y}_{e'}))}$

ESAC Training Objective: $\mathcal{L}(\mathbf{w}) = \mathbb{E}_{\mathcal{H} \sim p(\mathcal{H})} \mathbb{E}_{(e,j) \sim p(e,j|\mathcal{H})}\left[\ell(\mathbf{h}_{(e,j)})\right]$

ESAC for Indoor Camera Re-Localization

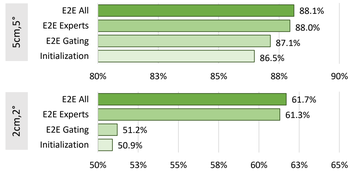
We combine the **7Scenes** [Sho13] and **12Scenes** [Val16] re-loc. datasets.

Both datasets contain multiple ambiguous offices, kitchens, etc.



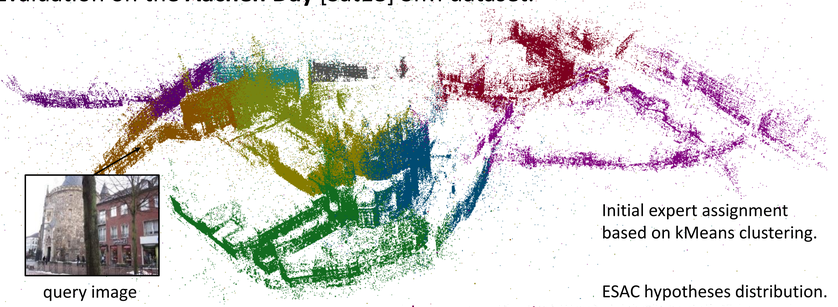
	Max. Experts	Avg. Experts	Accuracy	Avg. Time (ms)
DSAC++	-	-	53.3%	940
Expert Selection	1	1	47.5%	307
ESAC (Ours)	1	1	49.9%	276
	2	2	67.2%	343
	3	2.9	75.3%	398
	19	6.1	88.1%	555
Uniform Gating	19	19	87.8%	1,377
Oracle Gating	1	1	89.0%	120

Effect of end-to-end training.



ESAC for Outdoor Camera Re-Localization

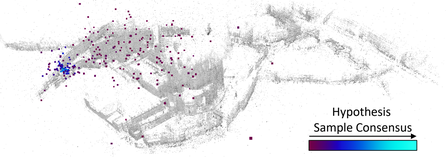
Evaluation on the **Aachen Day** [Sat18] SfM dataset.



query image

Initial expert assignment based on kMeans clustering.

ESAC hypotheses distribution.



Method	0.25m, 2° / 0.5m, 5° / 5m, 10°
DSAC++	0.4% / 2.4% / 34.0%
ESAC (10 Experts)	30.3% / 49.3% / 73.7%
ESAC (20 Experts)	39.7% / 55.9% / 77.8%
ESAC (50 Experts)	42.6% / 59.6% / 75.5%
Active Search [Sat16]	**57.3% / 83.7% / 96.6%**