Robust Aleatoric Modeling for Future Vehicle Localization

Max Hudnell, True Price, Jan-Michael Frahm
{mhudnell, jtprice, jmf}@cs.unc.edu
THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

CVPR LONG BEACH CALIFORNIA June 16-20, 2019

Overview

Motivation: Future object localization is a new area of research with potentially useful application to collision prevention systems and tracking algorithms.

Goal: Accurately predict future object localizations (for any given future timepoint) while also modeling their uncertainty.

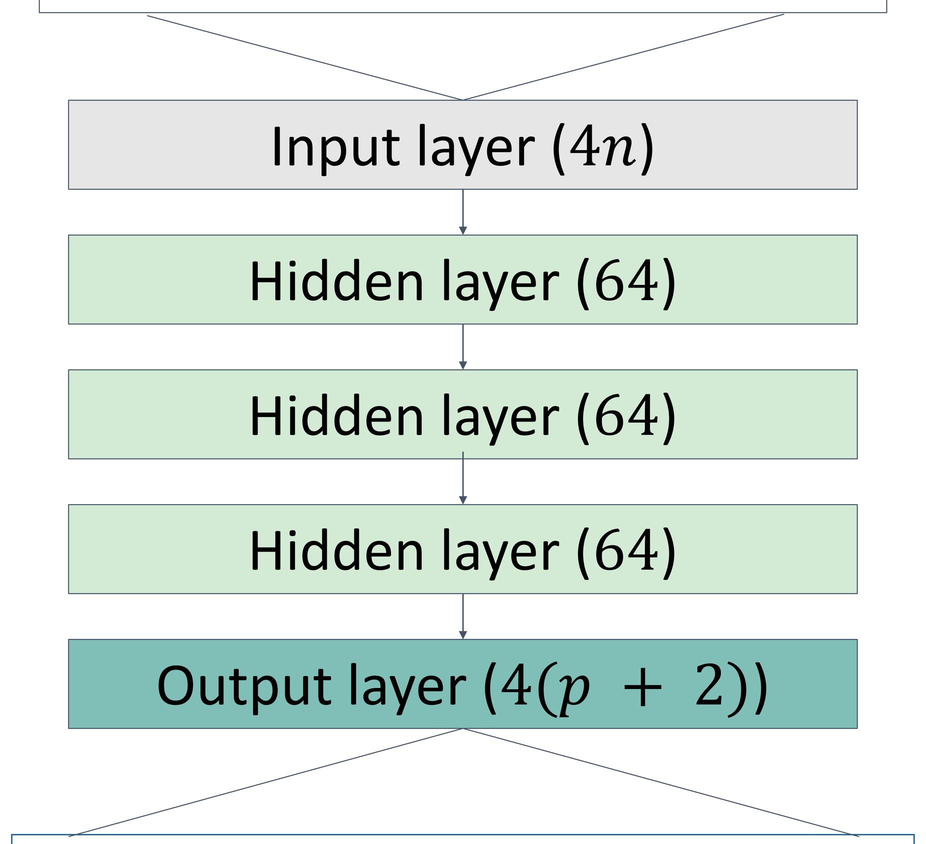
Challenges:
1. **Overfitting to rare instances** can occur when modeling Gaussian uncertainty.
2. **Current techniques fail to evaluate aleatoric uncertainty** by only evaluating prediction mean

Contributions:
- **We model predictions as a continuous function of time**: Less expensive without sacrificing accuracy
- **Robust distribution** formulation
- Introduce technique for **uncertainty evaluation**

Visualizing Estimations with Uncertainty

prior localizations, true +1s future localization (white), mean for +1s localization, location probability for +1s

Method



$$\mathbf{B} = \{B_{-n+1}, B_{-n+2}, \ldots, B_0\}$$

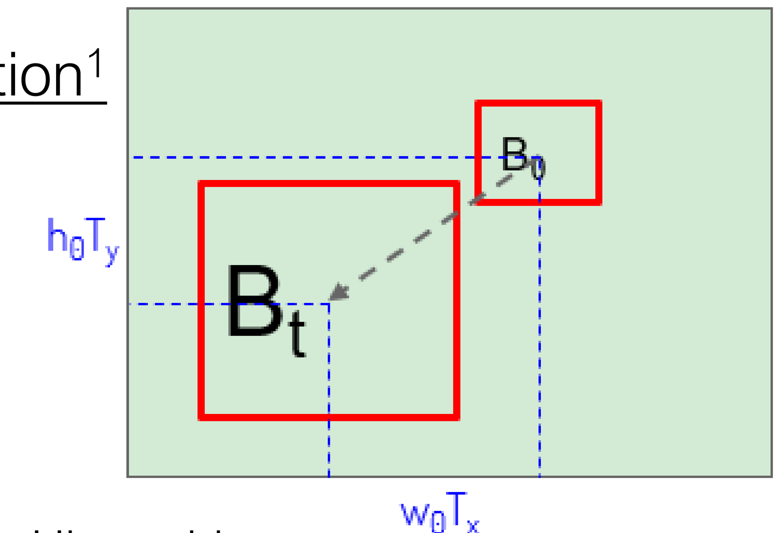
$$\theta_{\mathbf{B}} = \{\theta_B^x, \theta_B^y, \theta_B^w, \theta_B^h\}$$
$$\theta_\sigma = \{\theta_\sigma^x, \theta_\sigma^y, \theta_\sigma^w, \theta_\sigma^h\}$$

p: polynomial degree

Predicting Relative Transformations

Scale-invariant transformation[1]

$$B_x(t) = w_0 T_x + x_0$$
$$B_y(t) = h_0 T_y + y_0$$
$$B_w(t) = w_0 \exp(T w)$$
$$B_h(t) = h_0 \exp(T h)$$



[1]R. Girshick (2013), Rich Feature Hierarchies

Model Output: Functions of Time

- Output is a set of coefficients for the bounding box predictor T_d and uncertainty model σ_d
- Construct polynomial functions:

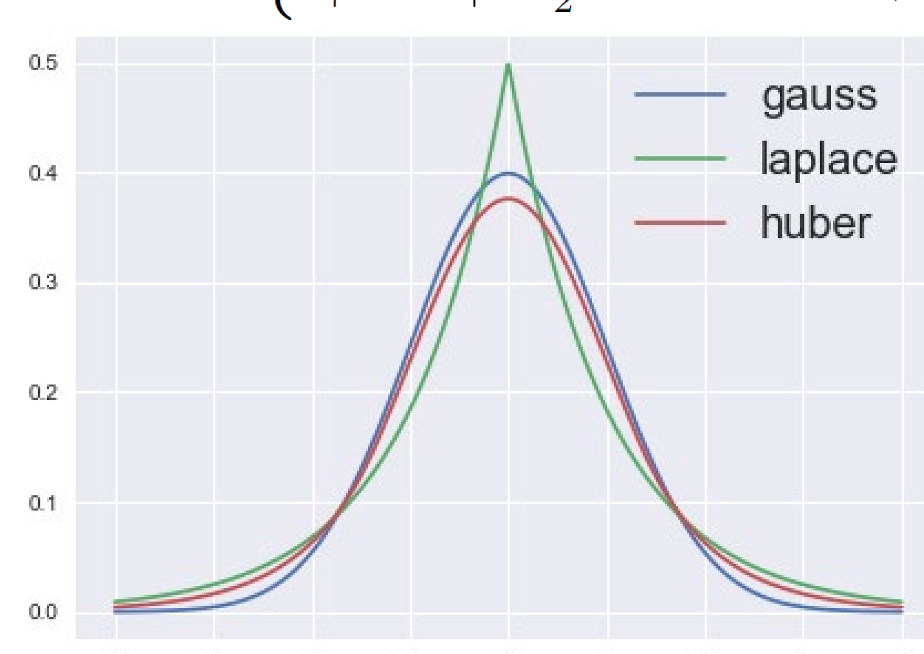
$$T_d(t; \theta_B^d) = \sum_{i=1}^{p} \theta_B^{d(i)} t^i.$$

$$\sigma_d(t; \theta_\sigma^d) = |\theta_\sigma^{d(1)} t| + |\theta_\sigma^{d(0)}|$$

Localization Loss w/o Uncertainty

- Evaluate at 10 future timesteps
- Minimize error from predicted $T^{(1)} \ldots T^{(9)}$ to ground truth $\hat{T}^{(1)} \ldots \hat{T}^{(9)}$ using Huber loss

$$H(\hat{x}, x) = \begin{cases} \frac{1}{2}(\hat{x} - x)^2 & \text{if } |\hat{x} - x| < \tau \\ \tau|\hat{x} - x| - \frac{1}{2}\tau^2 & \text{otherwise,} \end{cases}$$



Training for Confidence

- We maximize the likelihood of a distribution defined by the Huber loss
- Huber loss combines the *L1* and *L2* losses; the Huber distribution combines the *Gaussian* and *Laplace* distributions

$$p(\hat{x}|x, \sigma) = \begin{cases} \frac{1}{c}\exp\left(-\frac{(\hat{x}-x)^2}{2\sigma^2}\right) & \text{if } |\hat{x} - x| < \tau \\ \frac{1}{c}\exp\left(-\frac{\tau}{\sigma^2}|\hat{x} - x| + \frac{\tau^2}{2\sigma^2}\right) & \text{otherwise,} \end{cases}$$

Objective function:

$$\min_{\theta_B^d, \theta_\sigma^d} \sum_d H_d(\hat{T}_d(t), T_d(t; \theta_\sigma^d), \sigma_d(t; \theta_\sigma^d))$$

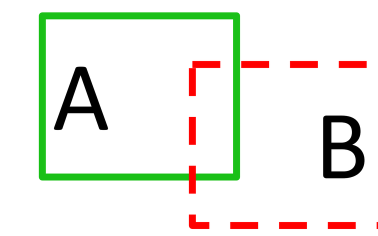
$$H_d(\hat{T}, T, \sigma) = \log c + \begin{cases} \frac{(\hat{T}-T)^2}{2\sigma^2} & \text{if } |\hat{T} - T| < \tau \\ \frac{\tau}{\sigma^2}|\hat{T} - T| - \frac{\tau^2}{2\sigma^2} & \text{otherwise,} \end{cases}$$

$$c = \sigma\sqrt{2\pi}\,\text{erf}\left(\frac{\tau}{\sigma\sqrt{2}}\right) + \frac{2\sigma^2}{\tau}\exp\left(-\frac{\tau^2}{2\sigma^2}\right) \qquad \tau = 1.345\sigma$$

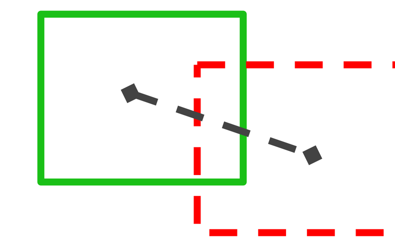
Results

Comparison of Loss / Model Type

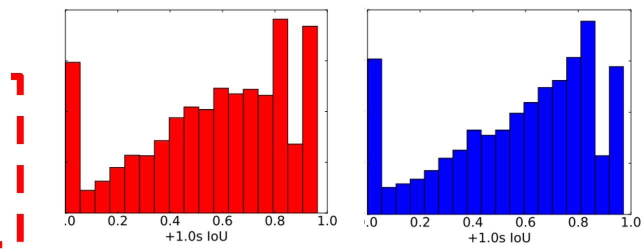
Intersection over Union (IoU)



Displacement Error



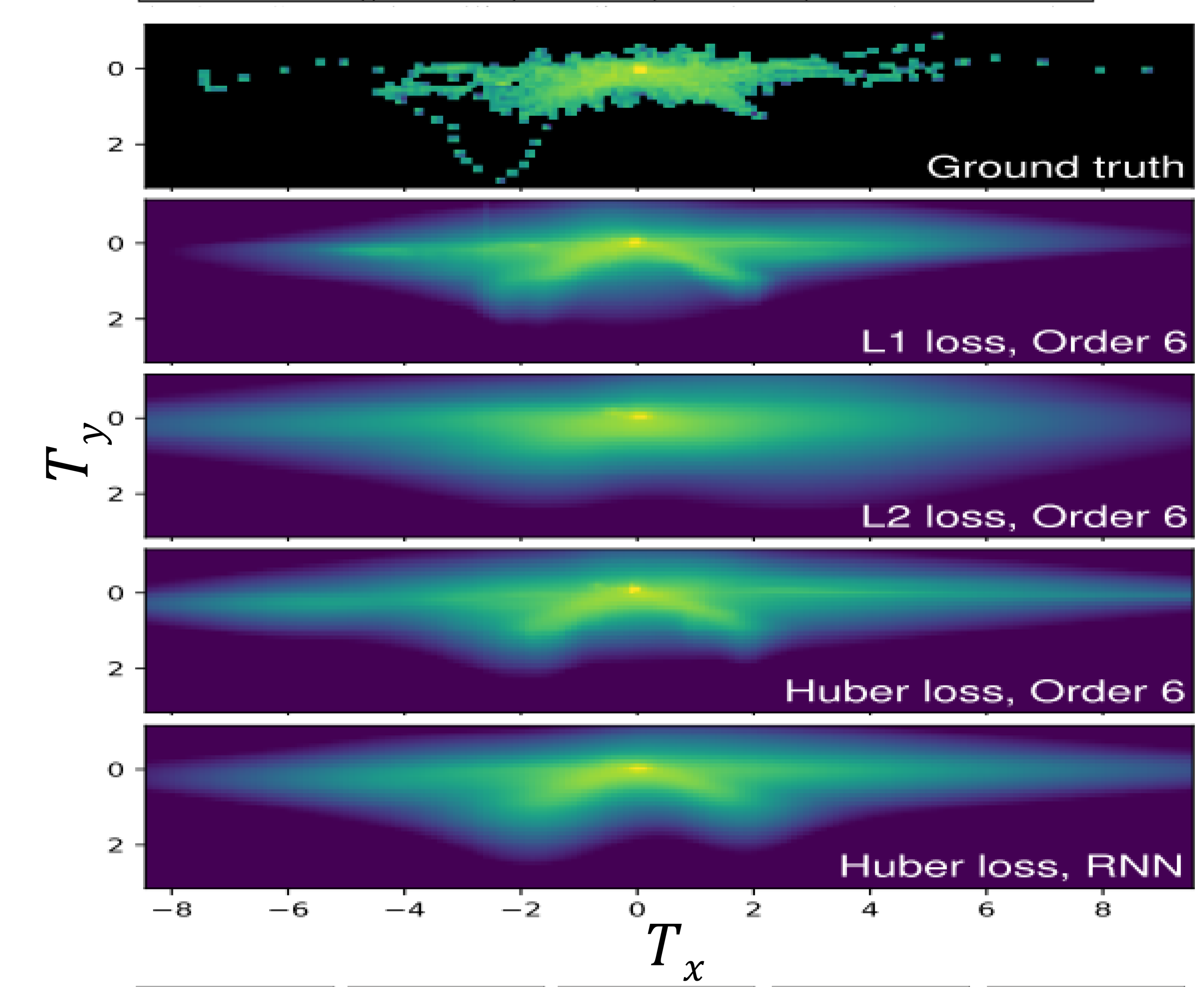
L1 vs. Huber IoU dist. comparison

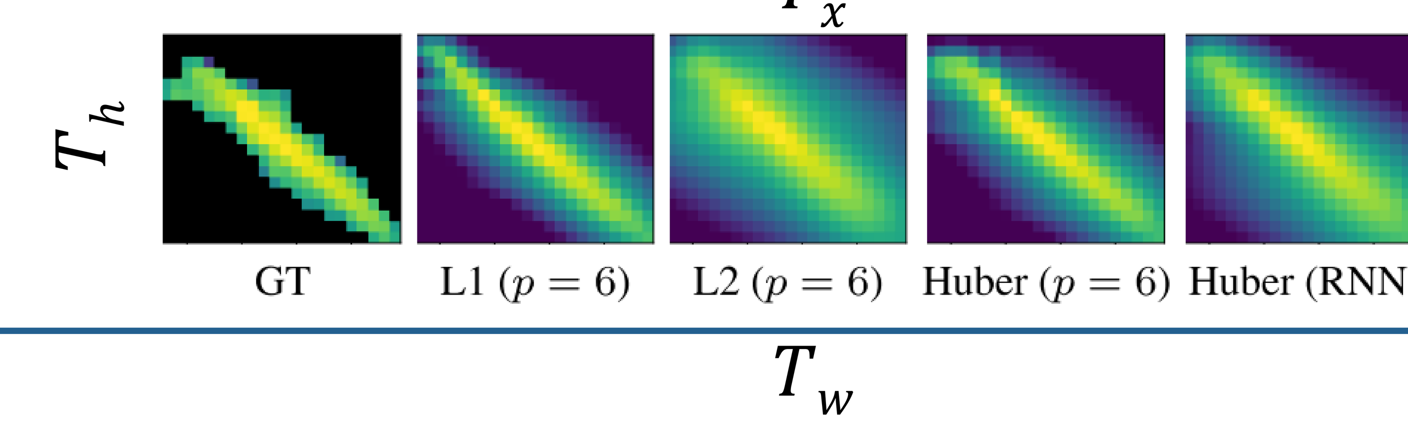


Loss	Func.	All DE +0.5s	All DE +1.0s	All ADE	All IoU +0.5s	All IoU +1.0s	Hard DE +0.5s	Hard DE +1.0s	Hard ADE	Hard IoU +0.5s	Hard IoU +1.0s
–	constant	32.06	72.15	36.98	0.498	0.339	44.23	102.40	51.62	0.326	0.128
–	linear	14.61	39.51	17.95	0.663	0.464	23.14	63.27	28.56	0.492	0.219
L1	$p = 6$	12.81	29.05	14.78	0.697	0.564	17.06	38.06	19.56	0.607	0.475
L2	$p = 6$	15.13	37.43	18.12	0.671	0.521	20.90	51.97	25.02	0.563	0.394
Huber	$p = 6$	12.58	29.18	14.72	0.708	0.584	16.18	36.76	18.76	0.622	0.488
Huber	RNN	14.01	31.37	16.12	0.686	0.570	19.62	44.63	22.71	0.577	0.442

Marginal log-probability distributions for +1s

Config.	L1	L2	Huber	Huber (RNN)
$\mathcal{H}^2(\mathcal{T}, \hat{\mathcal{T}})$	0.568	0.607	0.562	0.562





Supported by NSF grant No. CNS-1405847