

Progress Report

Bardia Mojra

December 23, 2020

Robotic Vision Lab

University of Texas at Arlington

1 Long-Term Goals

- Agile Manufacturing

2 Mid-Term Goals

- Bayesian Pose Estimation
- ARIAC
- Pose Estimation Survey

3 To Do

- Dissect Kendall's [?]: in progress.
- Write a literature review for [?]: in progress.
- Dissect BayesOD [?]: in progress.
- Write a literature review for [?]: Done.
- Implement PoseCNN, DOPE, and BayesOD.
- Look into transfer learning.
- Read and write a review for [?]: Done.
- Look into domain randomization and adaptation techniques.
- MoreFusion [?]: Wrote a literature review. Done.
- Reading list: [?] and [?].
- Read [?].
- Search for recent pose estimation survey papers.
- Generate new data set for PoseCNN and DOPE using UE4.
- Read more papers on deep pose estimation.
- Learn to use UE4.

4 Progress

Following items are listed in order of priority:

- Pose Estimation with BNN's: I am currently working on dissecting [?] and write a review for it simultaneously. Next, I will expand the review I have written for BayesOD and add more details as I look into the source code. My goal is to do some research every morning and to end that with a page of written material. I will focus on Bayesian vision as it offers probabilistic predictions by using per pixel Bayesian inference rather than deterministic output predictions based on MLE parameters as currently used by CNN's.
- Bayesian Uncertainty, [?]: Current pose estimation models with performance comparable with state of the art such as [?], [?], [?], and [?] all use CNN's.

Although use of CNN's remains a powerful tool as a mean to compute rich feature maps by taking advantage of color channels; it has its short comings, mainly that predictions are deterministic even though the model is probabilistic.

Classical neural networks and CNN's use maximum likelihood to calculate network weights and biases which derive network outputs. Such models are represented by conditional PDF $p(\mathbf{y}|\mathbf{x}, \theta)$ trained on data $D = \{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^N$ and converging to a set of learned parameters $\theta = \{\mathbf{W}, \mathbf{b}\}$ where

$$p(D|\theta) = \prod_{i=1}^N p(\mathbf{y}_i|\mathbf{x}_i, \theta)$$

defines that PDF. For training these models, parameters \mathbf{W} and \mathbf{b} are learned through back-propagation using Maximum Likelihood Estimate (MLE) method, as defined by

$$\hat{\theta}_{MLE} = \underset{\theta}{argMax} \sum_{i=1}^N \ln p(\mathbf{y}_i|\mathbf{x}_i, \theta)$$

but in bayesian —

We start with the two basic rules for Bayesian machine learning, sum rule and product rule.

Sum rule:

$$p(\mathbf{X}) = \sum_{\mathbf{y}_i} p(\mathbf{y}_i, \mathbf{x}_i)$$

Product rule:

$$p(\mathbf{X}_i, \mathbf{Y}_i) = p(x)p(\mathbf{y}_i|\mathbf{x}_i)$$

Bayesian Learning: Learning is done by applying Bayes' rule to current state of knowledge as new evidence becomes available throughout training process.

$$p(\theta|D, M) = \frac{p(D|\theta, M) p(\theta|M)}{p(D|M)}$$

Where $p(D|\theta, M)$ is the likelihood of parameters θ in model M. Prior belief or probability of parameters for given model M. And posterior belief of θ given model M and training data D is represented by $p(\theta|D, M)$.

Bayesian Prediction:

$$p(x|D, M) = \int p(x|\theta, D, M)p(\theta|D, M)d\theta$$

Bayesian Model Comparison:

$$p(M|D) = \frac{p(D|M)p(M)}{p(D)}$$

On the other hand, Bayesian Neural Networks or BNN's infer posterior probability distribution frame by frame by marginalizing new likelihood over the distribution of parameters, hence converging to steady state posterior distribution. Moreover, Bayesian deep learning tools presented in [?] and [?] allows for study and analysis of *aleatoric* and *epistemic* uncertainties together in one frame work.

Aleatoric:

Epistemic:

Learned attenuation:

- ARIAC: For now, I will focus on implementing pose estimation and BayesOD implementations.

- Pose Estimation Survey Paper Feedback:

Need a strong introduction to the problem

Problem statement needs to be both quantitatively and qualitatively (figures)

Make a to do list of this

The paper is in rough shape and far from complete. It needs a strong introduction to the problem. The paper should explain quantitatively (math) and qualitatively (figures) what exactly the pose estimation problem is. Related work should go back at least 20-30 years. That does not mean technology/methods from the 1990s have to be described in detail, but it has to at least be mentioned/referenced.

One way to possibly improve the organization/structure is to look at trends/paradigms since the start of the pose estimation problem. There was obviously a shift in paradigm with the advent of deep learning. What were the trends prior to that? Try to organize the paper in chronological order based on those trends.

It's a good start, but writing a survey paper requires reading nearly all of the (highly cited) prior work and having an understanding of what that work accomplished. This work all needs to be briefly described in the survey with references. Please look at other recent survey papers (e.g., robot learning) and talk to and read Joe's survey on movement primitives as well.

- Joint Probability:

Joint probability $p(D, W)$ relates prior belief to posterior belief given new evidence $p(D)p(W|D) = p(D, W)[\text{jointprob.}] = p(W)p(D|W)[\text{conditionalprob.}]$

Posterior probability for a particular weight values, W , for given training data D .

Bayesian approach to deep neural network learning allows for quantifying uncertainties related to model parameters as well as uncertainties rooted in model structure. Uncertainty

- TensorFlow [?]: I am still working through chapter 2.
- Project Alpe with Nolan: On pause for right now.

- UR5e:
- Fellowship:
- System Identification Presentation:

5 Plans

Following items are listed in order of priority:

- (On pause) Continue with ROS Industrial tutorials and documentation.
- (On pause) Resume Robotic Perception course as soon as possible.
- (On pause) Read Digital Image Processing by Gonzalez and Woods.