Weekly Update

Hanchen Wang

Ph.D Candidate in
Engineering, University of Cambridge

November 25, 2019

Overview

1. Stucture of First Year Report
   - More on PointNet Series
2. ideas on registration

What I have done on PointNet

- Ablation Studies(varying of {number of input points, transformers, batch size, pooling methods, dropout rates} VS {convergence speed, classification accuracy})
- **T-Net, analysis on the T-Net(spatial transformer) within the PointNet and some follow-up works**
- variations of LatentGNN, which is the structure designed to increase the spatial
- potential use of (different set functions other than poolings)

More on the PointNet Series- What does T-Net do?

usually in 3D we represent the homogeneous transformation of a rigid body by a (4,4) matrix T:

$$R_z = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} R_y = \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix} R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma & -\sin\gamma \\ 0 & \sin\gamma & \cos\gamma \end{pmatrix}$$

$$T = \begin{pmatrix} \cos\alpha\cos\beta & \cos\alpha\sin\beta\sin\gamma - \sin\alpha\cos\gamma & \cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma & x_t \\ \sin\alpha\cos\beta & \sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma & \sin\alpha\sin\beta\cos\gamma - \cos\alpha\sin\gamma & y_t \\ -\sin\beta & \cos\beta\sin\gamma & \cos\beta\cos\gamma & z_t \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

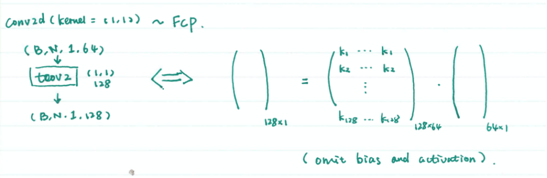
where α, β, γ is the yaw, pitch and roll of the camera/viewpoint respectively:

$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = T \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

in the PointNet, such transformation is done by the spatial transformer called **T-Net**(3, 3)

More on the PointNet Series: What does T-Net do?

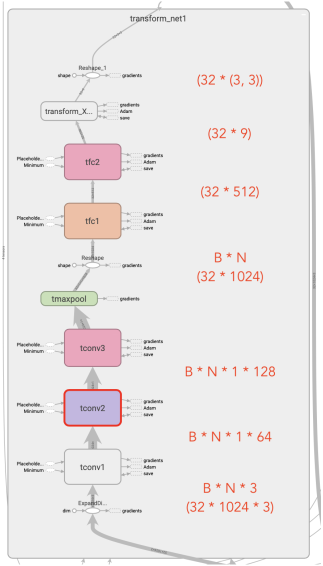
- structure of T-Net(visualized by TensorBoard)
- equivalence of conv2d(kernel size=(1,1)) and fcp:



- T-Net characterization 1: Pure Rotation:
- T-Net characterization 2: Rotation + Scaling:

$$\text{T-Net} = R + \begin{pmatrix} s_x & & \\ & s_y & \\ & & s_z \end{pmatrix}$$

- no translation in T-Net



More on the PointNet Series: What does T-Net do?

- T-Net characterization 1: Pure Rotation: to determine α, β, γ
- Loss $\mathcal{L}$ in based on the Frobenius Norm which is convex w.r.t. α, β, γ:

$$\mathcal{L} := ||\text{TNet} - R(\alpha, \beta, \gamma)||_F^2$$

$$= \left\| \begin{matrix} \cos\alpha\cos\beta - t_{11} & \cos\alpha\sin\beta\sin\gamma - \sin\alpha\cos\gamma - t_{12} & \cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma - t_{13} \\ \sin\alpha\cos\beta - t_{21} & \sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma - t_{22} & \sin\alpha\sin\beta\cos\gamma - \cos\alpha\sin\gamma - t_{23} \\ -\sin\beta - t_{31} & \cos\beta\sin\gamma - t_{32} & \cos\beta\cos\gamma - t_{33} \end{matrix} \right\|_F^2 = \sum_{i=1}^{3*3} (\cdot)^2$$

$$\begin{aligned} \Rightarrow \frac{1}{2}\frac{\partial\mathcal{L}}{\partial\alpha} &= (\cos\alpha\cos\beta - t_{11})(-\sin\alpha)\cos\beta + (\cos\alpha\sin\beta\sin\gamma - \sin\alpha\cos\gamma - t_{12})(-\sin\alpha\sin\beta\sin\gamma - \cos\alpha\cos\gamma) \\ &+ (\cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma - t_{13})(-\sin\alpha\sin\beta\cos\gamma + \cos\alpha\sin\gamma) + (\sin\alpha\cos\beta - t_{21})(\cos\alpha\cos\beta) \\ &+ (\sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma - t_{22})(\cos\alpha\sin\beta\sin\gamma - \sin\alpha\cos\gamma) \\ &+ (\sin\alpha\sin\beta\cos\gamma - \cos\alpha\sin\gamma - t_{23})(-\cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma) \end{aligned}$$

$$\begin{aligned} \Rightarrow \frac{1}{2}\frac{\partial\mathcal{L}}{\partial\beta} &= (\cos\alpha\cos\beta - t_{11})\cos\alpha(-\sin\beta) + (\cos\alpha\sin\beta\sin\gamma - \sin\alpha\cos\gamma - t_{12})(\cos\alpha\cos\beta\sin\gamma) \\ &+ (\cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma - t_{13})(\cos\alpha\cos\beta\cos\gamma) + (\sin\alpha\cos\beta - t_{21})(-\sin\alpha\sin\beta) \\ &+ (\sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma - t_{22})(\sin\alpha\cos\beta\sin\gamma) + (\sin\alpha\sin\beta\cos\gamma - \cos\alpha\sin\gamma - t_{23})(\sin\alpha\cos\beta\cos\gamma) \\ &+ (-\sin\beta - t_{31})(-\cos\beta) + (\cos\beta\sin\gamma - t_{32})(-\sin\beta)\sin\gamma + (\cos\beta\cos\gamma - t_{33})(-\sin\beta)\cos\gamma \end{aligned}$$

More on the PointNet Series: What does T-Net do?

- T-Net characterization 1: Pure Rotation: to determine α, β, γ
- Loss $\mathcal{L}$ in based on the Frobenius Norm which is convex w.r.t. α, β, γ:

$$
\begin{aligned}
\Rightarrow \frac{1}{2}\frac{\partial \mathcal{L}}{\partial \gamma} &= (\cos\alpha\sin\beta\sin\gamma - \sin\alpha\cos\gamma - t_{12})(\cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma) \\
&+ (\cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma - t_{13})(\cos\alpha\sin\beta(-\sin\gamma) + \sin\alpha\cos\gamma) \\
&+ (\sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma - t_{22})(\sin\alpha\sin\beta\cos\gamma + \cos\alpha(-\sin\gamma)) \\
&+ (\sin\alpha\sin\beta\cos\gamma - \cos\alpha\sin\gamma - t_{23})(\sin\alpha\sin\beta(-\sin\gamma) - \cos\alpha\cos\gamma) \\
&+ (\cos\beta\sin\gamma - t_{32})\cos\beta\cos\gamma + (\cos\beta\cos\gamma - t_{33})\cos\beta(-\sin\gamma) \\
\Rightarrow (\alpha', \beta', \gamma') &= (\alpha, \beta, \gamma) - \eta\nabla\mathcal{L} \quad \text{where } \eta \text{ is the (adaptive) learning rate}
\end{aligned}
$$

- we start search from the regions of the possiblt optimal values(it turns out that it is not convex):

$$
\alpha = \tan^{-1}(t_{21}/t_{11}) \quad \beta = \tan^{-1}\left(-t_{31}/\sqrt{t_{32}^2 + t_{33}^2}\right) \quad \gamma = \tan^{-1}(t_{32}/t_{33})
$$

Robustness of T-Net

- robustness of T-Net
- extended to five dimensions(Geometric Algebra)
- aggregated with point attributes such as RGB and normal

ideas on registration

- what traditional ICP do? (original and variations)
- how to incorporate deep learning into it
- **probabilistic models**
- **geometric algebra**

The End