Spherical Regression:Learning Viewpoints, Surface Normals and 3D Rotations on$n$-Spheres

David Mohlin

KTH

2019

Introduction

- Try have networks model output on n-spheres. (S(n-1))
- $S(n) = \{x \in \mathbb{R}^{n+1} : ||x|| = 1\}$

Introduction

- Try have networks model output on n-spheres. (S(n-1))
- $S(n) = \{x \in \mathbb{R}^{n+1} : ||x|| = 1\}$
- Reason is that the S(n) set occurs naturally in several tasks, for example estimating directions

Introduction

- Try have networks model output on n-spheres. (S(n-1))
- $S(n) = \{x \in \mathbb{R}^{n+1} : ||x|| = 1\}$
- Reason is that the S(n) set occurs naturally in several tasks, for example estimating directions
- The space of rotations (SO(n)) has a topology similar to the one for S(m)
- $SO(n) = \{\mathbf{R} \in \mathbb{R}^{3\times 3} : \mathbf{R}^T\mathbf{R} = \mathbf{I}, |\mathbf{R}| = 1\}$

Motivation

- They want a mapping onto $S(N)$ which has a bounded jacobian

Motivation

- They want a mapping onto $S(N)$ which has a bounded jacobian
- To get this they use the mapping
- $|p_j| = \dfrac{exp(o_j)}{\sum\limits_{k=1}^{N} exp(o_n)} \ \ \forall j <= N$

Motivation

- $|p_j| = \dfrac{exp(o_j)}{\sum\limits_{k=1}^{N} exp(o_n)} \;\; \forall j <= N$
- This mapping will output N elements with l2 norm 1
- But all components are positive.

Motivation

- $|p_j| = \dfrac{exp(o_j)}{\sum\limits_{k=1}^{N} exp(o_n)} \;\; \forall j <= N$
- This mapping will output N elements with l2 norm 1
- But all components are positive.
- To fix this they have N 2 class classifiers determining $sign(p_j)$

Motivation

- $|p_j| = \frac{exp(o_j)}{\sum_{k=1}^{N} exp(o_n)} \;\; \forall j <= N$
- This mapping will output N elements with l2 norm 1
- But all components are positive.
- To fix this they have N 2 class classifiers determining $sign(p_j)$
- Final output is $p_j = sign(p_j) * |p_j|$

Datasets

- Pascal3D+: Real dataset with manually oriented orientations of objects
- 12 classes (bike, car, aeroplane etc) annotators have fitted cads onto images
- Also train on synthetic data

Datasets

- Pascal3D+: Real dataset with manually oriented orientations of objects
- 12 classes (bike, car, aeroplane etc) annotators have fitted cads onto images
- Also train on synthetic data
- NYU Depth v2: Normals estimated from kinnect, try to estimate those normals based on images.
- Also use synthetic data for training.

Datasets

- Pascal3D+: Real dataset with manually oriented orientations of objects
- 12 classes (bike, car, aeroplane etc) annotators have fitted cads onto images
- Also train on synthetic data
- NYU Depth v2: Normals estimated from kinnect, try to estimate those normals based on images.
- Also use synthetic data for training.
- Modelnet10-SO(3) synthetic dataset

Methods

- Pascal3D+: estimate 3 euler angles separately (3 elements from S(2)), similar to other method

Methods

- Pascal3D+: estimate 3 euler angles separately (3 elements from $S(2)$), similar to other method
- NYU Depth v2: Estimate point on $S(3)$, similar to other method.

Methods

- Pascal3D+: estimate 3 euler angles separately (3 elements from S(2)), similar to other method
- NYU Depth v2: Estimate point on S(3), similar to other method.
- Modelnet10-SO(3): Test several output representations, quaternions work the best

Results

- Pascal3d+ 11.6 $\rightarrow$ 9.2 degrees error (SOTA 10.1)

Results

- Pascal3d+ 11.6 $\rightarrow$ 9.2 degrees error (SOTA 10.1)
- NYU Depth v2: 21.7 $\rightarrow$ 19.7 degree error (SOTA 21.7 degrees)

Results

- Pascal3d+ 11.6 $\rightarrow$ 9.2 degrees error (SOTA 10.1)
- NYU Depth v2: 21.7 $\rightarrow$ 19.7 degree error (SOTA 21.7 degrees)
- Modelnet10-SO(3): 20.3 degree error

Comments

- For two of the datasets they rotate the dataset by 45 degrees around dataset varying axis.
- I assume the representation is not as general as they claim (it is hard to output some directions)

Comments

- For two of the datasets they rotate the dataset by 45 degrees around dataset varying axis.
- I assume the representation is not as general as they claim (it is hard to output some directions)
- The representation is not rotation invariant.

Comments

- For two of the datasets they rotate the dataset by 45 degrees around dataset varying axis.
- I assume the representation is not as general as they claim (it is hard to output some directions)
- The representation is not rotation invariant.
- The classification sign causes discontinuities.

Comments

- For two of the datasets they rotate the dataset by 45 degrees around dataset varying axis.
- I assume the representation is not as general as they claim (it is hard to output some directions)
- The representation is not rotation invariant.
- The classification sign causes discontinuities.
- They do not model Pascal3D+ and ModelNet in the same way despite trying to estimate a rotation matrix for both.

Comments

- For two of the datasets they rotate the dataset by 45 degrees around dataset varying axis.
- I assume the representation is not as general as they claim (it is hard to output some directions)
- The representation is not rotation invariant.
- The classification sign causes discontinuities.
- They do not model Pascal3D+ and ModelNet in the same way despite trying to estimate a rotation matrix for both.
- They have a larger error for a synthetic dataset than a real dataset.