3D Object Classification via Spherical Projections

Zhangjie Cao[1], Qixing Huang[2], and Ramani Karthik[3]

[1]School of Software
Tsinghua University, China

[2]Department of Computer Science
University of Texas at Austin, USA

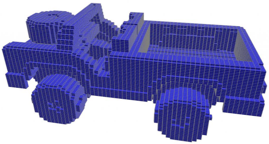
[3]School of Mechanical Engineering
Purdue University, USA

International Conference on 3DVision, 2017

Main-stream Methods

- Two main-stream 3D classification methods: image-based and 3D-based.

(a) Image-based



(b) 3D-based

- Spherical projections combine key advantages of these two main-stream 3D classification methods.

Depth-based Projection

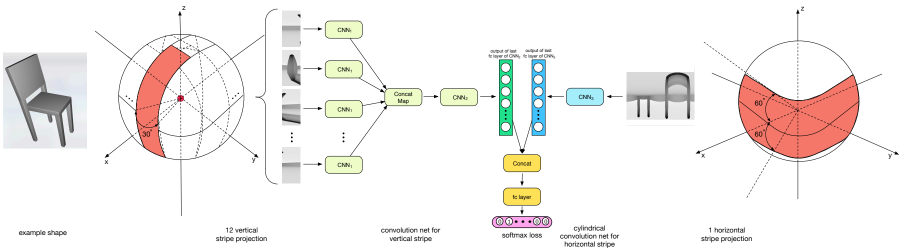
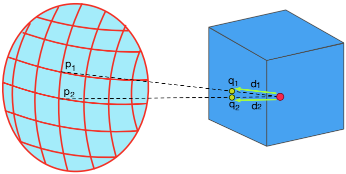
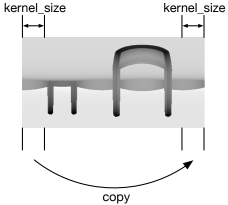


Figure: Depth-based Projections and Networks

Depth-based Projection



(a) Depth-based Projection Method



(b) Cylindrical Depth-based Projection

Figure: Details on Depth-based Projection

- Depth values are recorded as the distance to the first hitting point
- First compute depth values for vertices of a semi-regular quad-mesh
- Then generate the depth value of other points by linear interpolation.

Contour-based Projection

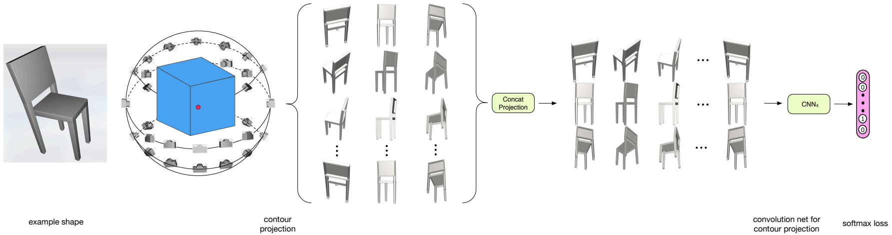


Figure: Contour-based Projections and Networks

Experiments Setup

- **Datasets:** ModelNet40, ShapeNetCore
- **Parameter selection:** cross-validation by jointly assessing
- **Methods to compare with:** Image-based methods: MVCNN, MVCNN-MultiRes; 3D-based methods: 3D ShapeNets, Voxnet, Volumetric CNN, OctNet; combined methods: FusionNet. All of these methods use the upright orientation but do not use the front orientation.

Results

- Accuracy of our approaches and the various baseline methods on ModelNet40 and ShapeNetCore and two curated subsets.

Method	ModelNet40	ShapeNetCore	ModelNet40-SubI	ShapeNetCore-SubI
3D Shapenets	85.9	na	83.33	na
Voxnet	87.8	na	85.99	na
FusionNet	90.80	na	89.54	na
Volumetric CNN	89.9	na	88.65	na
MVCNN	92.31	88.93	91.22	88.64
MVCNN-MultiRes	93.8	90.01	92.60	90.00
OctNet	87.83	88.03	86.45	87.85
depth-base pattern	91.36	89.45	90.25	89.13
contour-based pattern	93.31	90.49	92.20	90.80
overall pattern	94.24	91.00	93.09	91.22

Results

- Accuracy Before and After Pre-training on ModelNet40

Method	Before Pre-training		After Pre-training	
	Accuracy (class)	Accuracy (instance)	Accuracy (class)	Accuracy (instance)
MVCNN	82.15	87.15	90.35	92.31
MVCNN-MultiRes	88.12	91.20	91.40	93.80
depth-base pattern	80.44	86.09	87.32	91.36
contour-based pattern	88.33	91.48	91.16	93.31
overall pattern	88.53	91.77	91.56	94.24

- Accuracy Before and After Pre-training on ShapeNetCore

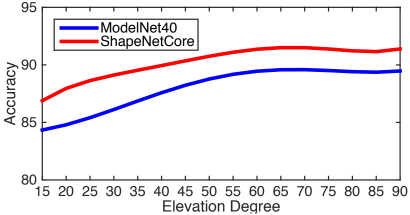
Method	Before Pre-training		After Pre-training	
	Accuracy (class)	Accuracy (instance)	Accuracy (class)	Accuracy (instance)
MVCNN	67.84	84.55	78.79	88.93
MVCNN-MultiRes	75.34	88.4	79.01	90.01
depth-base pattern	70.55	85.15	78.84	89.45
contour-based pattern	74.52	88.54	79.38	90.49
overall pattern	75.60	88.87	80.38	91.00

Analysis

- Accuracy w.r.t Number of Views for Depth and Contour Pattern on ModelNet40 and ShapeNetCore

Pattern	Number of Views	ModelNet40	ShapeNetCore
depth-based	6	90.87	88.95
	24	91.02	89.23
	12	91.36	89.45
contour-based	6	92.26	89.23
	24	93.12	90.36
	12	93.31	90.49

- Accuracy w.r.t Elevation degree of the strip parallel to the latitude



Summary

- We introduce a spherical representation exploiting both depth variation and contour information which can capture geometric details and data dependencies across the entire object.
- We develop deep neural networks incorporating large-scale labeled images for training to classify spherical representations of 3D objects.
- In the future, we plan to define convolutional kernels directly on spherical domains.