Geometry-Aware Point Cloud Learning for Robust and Efficient 3D Vision

Marianne Arriola¹, Xiangru Huang², Justin Solomon²
¹Department of Computer Science, University of California, Santa Barbara
²Department of Electrical Engineering & Computer Science, Massachusetts Institute of Technology

Can machines match human perception?
- We process the 3D physical world daily using human perception
- 3D vision models attempt to match human perception by making predictions using points picked up from the environment [Qi17]
- 3D vision models are used for autonomous driving, where accurately and efficiently detecting objects is crucial for safe and fast response

Input: scene | Human perception | Output: Is there a car? | Action: Where to go?
---|---|---|---
Human eyesight | Point representation | Machine perception | Yes/No

LiDAR sensor | Point representation | Machine perception | Yes/No

Point cloud learning is inefficient and not robust
- Point cloud learning is inefficient because it processes scenes that contain thousands/millions of points
- Point cloud learning is not robust because it must infer how noisy, unevenly sampled points compose coherent shapes

Can geometry information improve model efficiency and robustness?
- We hypothesize that summarizing points into the shapes that they constitute results in more efficient, robust predictive models
- Point summarization shrinks the size of the model input and provides added shape and object information to the model

Input: Point representation | Machine perception | Hybrid representation
---|---|---

Summarizing points into lines
- We test our approach by fitting lines to points [Fis81]
- We find the best line with the highest confidence by balancing two factors tuned by $\sigma$ and $\mathcal{T}$:
  - Distance of points from the line
  - Density of points along the line
  $$\text{dist}^2 \cdot \text{dens} \cdot c(l) = e^{-\left(\frac{\text{dist}(l)}{\bar{\text{dist}}}ight)^2 + \frac{\text{dens}(l)}{\bar{\text{dens}}}}$$

Input: Points | Output: Lines
---|---

Line representation definition
- Features at any point $\lambda$ along the line with endpoint features $f_0, f_1$ is:
  $$f_{\lambda}(l) = \lambda f_1 + (1 - \lambda)f_0, \lambda \in [0,1]$$
- Calculating features along the line can be visualized as:

Input: Line | Output: Line features
---|---

How do we learn from points and lines?
- We learn an informative hybrid representation for prediction by sampling the input and grouping information from nearby points
- Our model learns from lines by creating points on the line for sampling and preserving the line for the final predictions

Input: Hybrid representation | Create points on line | Sample | Group | Reduce to decision | Output: Prediction
---|---|---|---|---|---

Hybrid representations improve learning
- We test our model on a small-scale 2D example:
  - Given a drawing of a character, predict which character it is
  - Using only the point representation (PointNet++) [Qi17]
  - Using the hybrid representation

Input: Drawing | Point representation | Hybrid representation | Sample, group | New representation | Reduce to decision | Output: Which character?
---|---|---|---|---|---|---

Repeat sampling/grouping as necessary

How do we learn from points and lines?
- We learn an informative hybrid representation for prediction by sampling the input and grouping information from nearby points
- Our model learns from lines by creating points on the line for sampling and preserving the line for the final predictions

Input: Hybrid representation | Create points on line | Sample | Group | Reduce to decision | Output: Prediction
---|---|---|---|---|---

Hybrid representations improve learning
- We test our model on a small-scale 2D example:
  - Given a drawing of a character, predict which character it is
  - Using only the point representation (PointNet++) [Qi17]
  - Using the hybrid representation

Input: Drawing | Point representation | Hybrid representation | Sample, group | New representation | Reduce to decision | Output: Which character?
---|---|---|---|---|---|---

Repeat sampling/grouping as necessary

Geometry-aware learning shows promise
- Our positive results show that hybrid point cloud learning is feasible for larger-scale experiments
- This work can be extended to learn from hybrid representations with diverse shapes
  - Triangles, planes, etc.
- Future experiments needed for real-world 3D data
  - LiDAR sensor data for autonomous cars