Fast Point Cloud Sampling Network

Tianxin Huang[a], **Yong Liu** *[a], **Jie Liang**[a, b]

[a]*Laboratory of Advanced Perception on Robotics and Intelligent Learning, College of Control Science and Engineering, Zhejiang University, Hangzhou, China*
[b]*Beijing Institute of Mechanical and Electrical Engineering, Beijing, China*

A B S T R A C T

The increasing number of points in 3D point clouds has brought great challenges for subsequent algorithm efficiencies. Down-sampling algorithms are adopted to simplify the data and accelerate the computation.

1. Introduction

Existing works [1–3] often use random sampling and the farthest point sampling (FPS) to down-sample the point clouds.

The differences between our work and former learning-based works are presented in Fig. 1.

The discrepancy between progress-net and our method is presented in Fig. 1-(b) and (c).

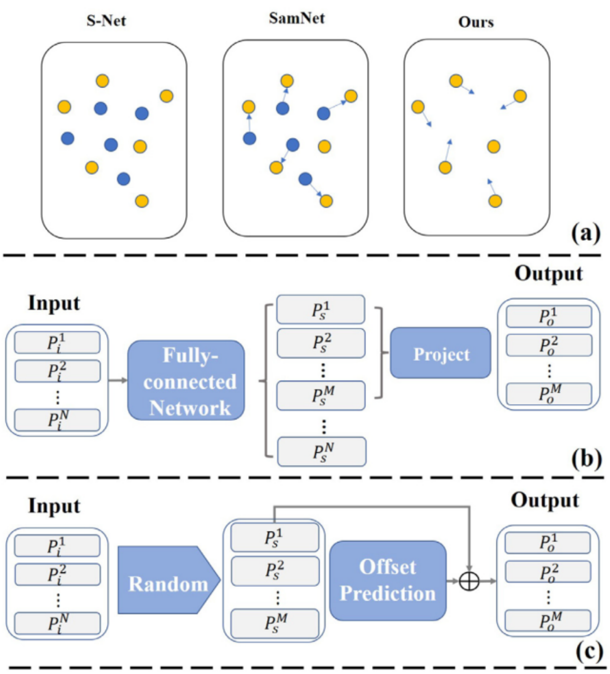

Fig. 1. (a) shows the differences between learning-based sampling strategies, while (b) and (c) present the discrepancy between progress-net and our method in multi-resolution sampling.

Our contributions can be summarized as:

- We propose a novel learning-based point cloud sampling framework named fast sampling network (FPN) by driving existing randomly sampled points to better positions;
- We introduce a hybrid training strategy to help FPN adapt to different sampling resolutions by randomly introducing selecting the resolution of initial points during training.

2. Methodology

2.1. Basic Pipeline

The basic pipeline of FPN is presented in Fig. 2. We aggregate global features from the input points with a set of multilevel perceptions (MLPs) and Max Pooling following PointNet [4].

2.2. Hybrid Training Strategy

The achievement of HTS is presented as Algorithm 1.

2.3. Loss Function

The range constraint can be presented as

$$\mathcal{L}_{rc} = \frac{1}{N}\sum \|S_0 - S_i\|_2, \tag{1}$$

For reconstruction-related tasks, it may be Chamfer Distance or Earth Mover Distance [5] defined as

$$\begin{aligned}\mathcal{L}_{task} &= \mathcal{L}_{CD}(S_1, S_2)\\ &= \frac{1}{2}\left(\frac{1}{|S_1|}\sum_{x\epsilon S_1}\min_{y\epsilon S_2}\|x-y\|_2 + \frac{1}{|S_2|}\sum_{x\epsilon S_2}\min_{y\epsilon S_1}\|x-y\|_2\right),\end{aligned} \tag{2}$$

or

$$\mathcal{L}_{task} = \mathcal{L}_{EMD}(S_1, S_2) = \min_{\phi:S_1\to S_2}\frac{1}{|S_1|}\sum_{x\epsilon S_1}\|x-\phi(x)\|_2. \tag{3}$$

where S_1 and S_2 are input and output. ϕ is a bijection from S_1 to S_2.

*Corresponding author.

3. Experiments

3.1. Dataset and implementation details

Table 1
The number of neurons in networks. f_1 , f_2 , f_3 are modules in Fig. 2.

	f_1	f_2	f_3
MLPs	(128,256,256)	(128,256,256)	(128,128,3)

Table 2

The comparison on optimal clustering.

Center	Iterations	1	10	100
16	FPS	2.43	2.00	1.98
	Ours	**2.16**	**1.98**	**1.96**
32	FPS	1.20	1.02	1.00
	Ours	**1.11**	**1.00**	**1.00**

The hyper-parameter λ is tuned on the validation split of ShapeNet. Detailed network structures are shown in Table 1.

3.2. Discussion about clustering

Except down-stream tasks such as reconstruction or recognition, down-sampled points can also be adopted as the initial clustering centers.

The results are presented in Table 2.

3.3. Ablation Study

The influence of range constraint. Note that this is only conducted to observe the influence of range constraint weight λ on sampling performances instead of the tuning of λ, which is chosen according to the val set introduced in subsection 3.1.

Algorithm 1 Training with Hybrid Training Strategy

Input: data X, the number of iterations $iter$, the number of resolutions m;
$prob_1, prob_2, \ldots, prob_m = \frac{1}{m}, \frac{1}{m}, \ldots, \frac{1}{m}$;
for $i = 1$ **to** $iter$ **do**
  Select the resolution r according to $prob_1, \ldots, prob_m$;
  Train FPN by descending gradient:$\Delta_{\theta_{FPN}} \mathcal{L}_{loss}\left(Y_{X,r}\right)$
end

Data availability

Data will be made available on request.

Acknowledgement

We thank all reviewers and the editor for excellent contributions. This work is supported by the Key Research and Development Project of Zhejiang Province under Grant 2021C01035.

References

[1] Charles Ruizhongtai Qi et al. "Pointnet++: Deep hierarchical feature learning on point sets in a metric space". In: *Advances in neural information processing systems* 30 (2017).

[2] Qingyong Hu et al. "Randla-net: Efficient semantic segmentation of large-scale point clouds". In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020, pp. 11108–11117.

[3] Charles R Qi et al. "Deep hough voting for 3d object detection in point clouds". In: *proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019, pp. 9277–9286.

[4] Charles R Qi et al. "Pointnet: Deep learning on point sets for 3d classification and segmentation". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 652–660.

[5] Haoqiang Fan, Hao Su, and Leonidas J Guibas. "A point set generation network for 3d object reconstruction from a single image". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 605–613.

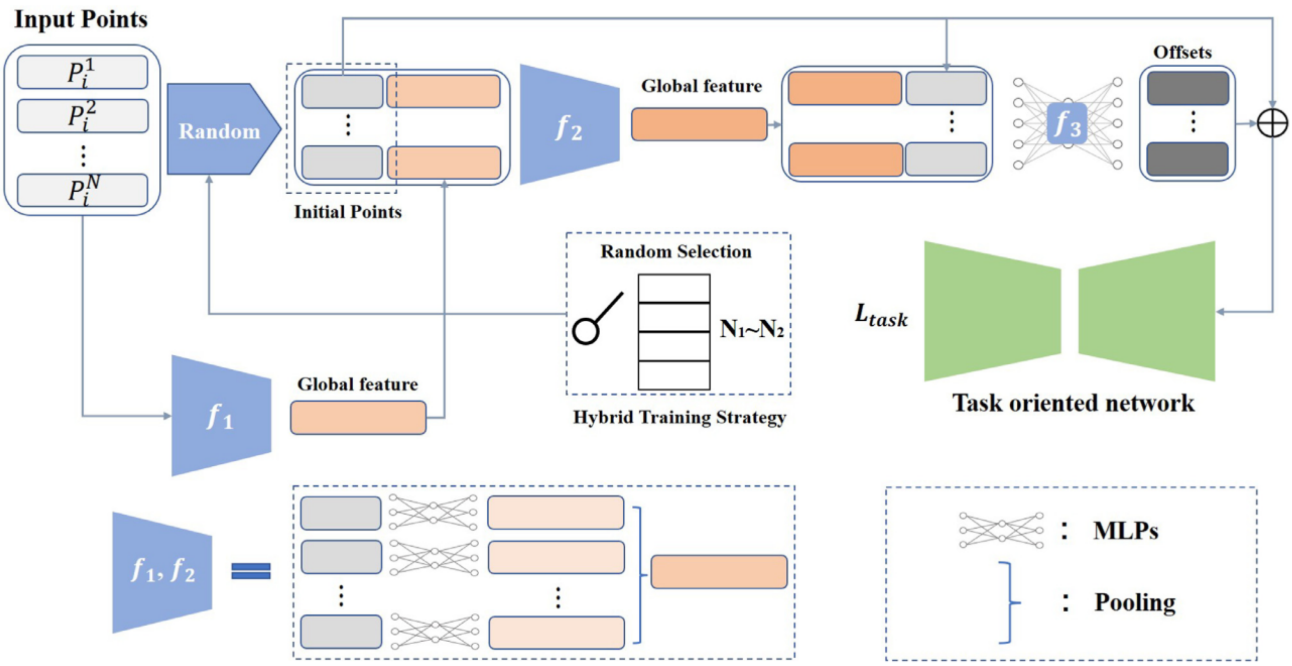

Fig. 2. The whole pipeline of FPN. The + denotes element-wised addition. f_1 and f_2 aggregate features by MultiLayer Perceptrons(MLPs) and pooling, while f_3 is a group of MLPs to predict offsets from coordinates and features. The task network is corresponding to the specific task, such as point cloud recognition and reconstruction. L_{task} is the loss constrained the task network.