WEAKLY SUPERVISED POINT CLOUD UPSAMPLING VIA OPTIMAL TRANSPORT

Zezeng Li[1] Weimin Wang [1] Na Lei [1] Rui Wang [1]

[1]School of Software, Dalian University of Technology, Dalian, 116024, People's Republic of China

Overall

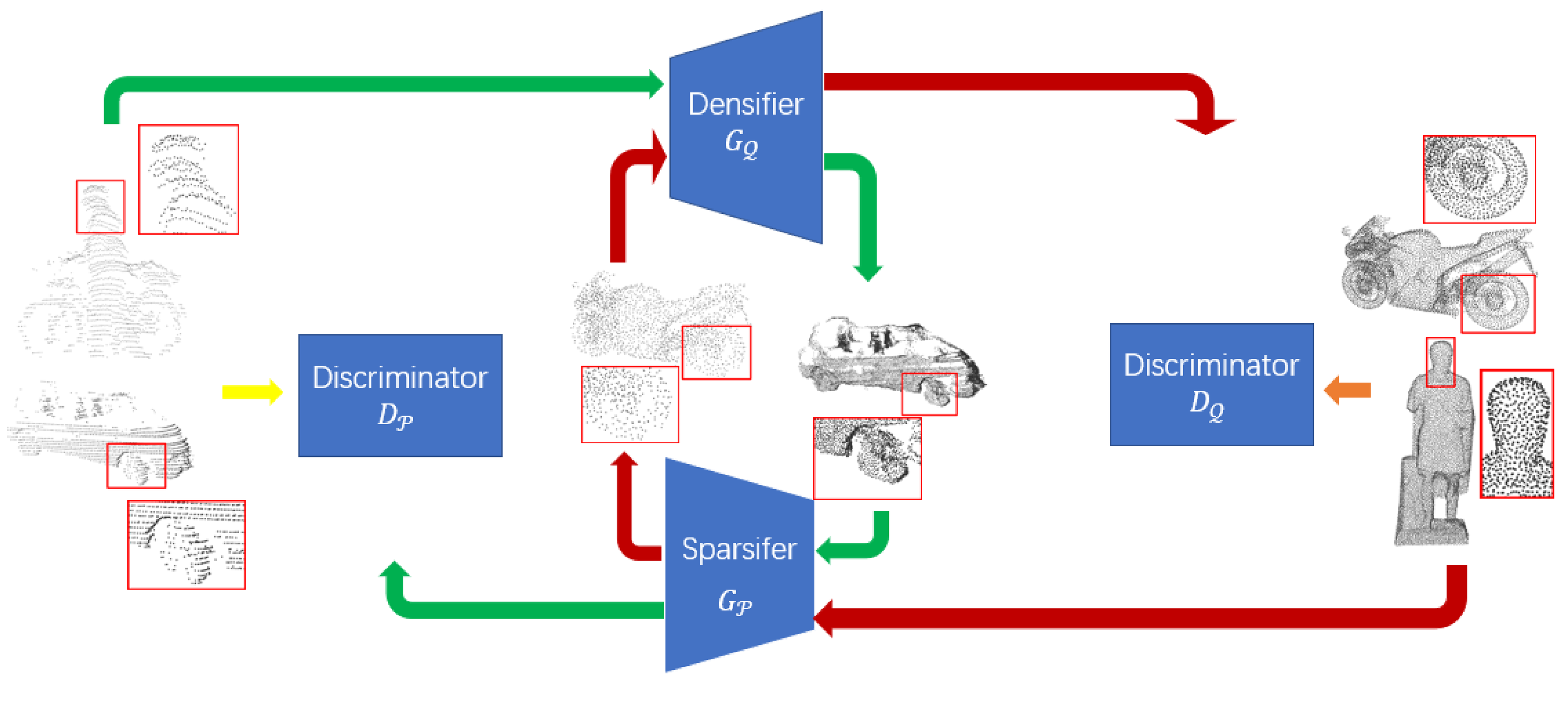

Figure 1. The diagram of the proposed PU-CycGAN.

Abstract

Existing learning-based methods usually train a point cloud upsampling model with synthesized, paired sparse-dense point clouds. However, the distribution gap between synthesized and real data limits the performance and generality. To solve this problem, we innovatively **regard the upsamplig task as an optimal transport (OT) problem** from sparse to dense point cloud. Further we propose PU-CycGAN, **it can be directly trained for upsampling with unpaired real sparse point clouds**, so that the distribution gap can be filled via the learning. Especially, quadratic Wasserstein distance is introduced for the stable training.

Introduction

In real world, the raw point clouds produced from depth cameras and LiDAR sensors are often sparse, noisy, and non-uniform. The point cloud upsampling technology that aims at generating dense, uniform and complete point clouds. Various learning-based methods have been proposed to continuously improve the upsampling performance. However, these methods require paired sparse-dense data in the network training, these supervised methods cannot be trained with real-scanned datasets such as ScanNet and KITTI where paired dense point clouds are unavailable. To resolve the above problems, We cast point cloud sampling as the OT problem and propose PU-CycGAN which can be trained with unpaired point sets.

Main Contributions

1. We propose a weakly supervised point cloud upsampling framework that trains the model with unpaired point clouds.
2. We notably regard point cloud upsampling as an OT problem, and design a quadratic Wasserstein distance to stabilize GAN's training.
3. We introduce consistency loss and self-restraint loss to improve the performance of the model in underlying surface representation.

Network Structure

Given M sparse point sets $\mathcal{P} = \{\{p_i^j\}_{j=1}^N\}_{i=1}^M$ and unpaired dense point sets $\mathcal{Q} = \{\{q_i^k\}_{k=1}^{rN}\}_{i=1}^M$, we aim to learn a map which transports the sparse point sets to dense and uniformly distributed point set. Where N denotes the number of points in each sparse point set, r is the upsampling rate.
Fig. 1 shows the overall network architecture of PU-CycGAN. Herein,

- Densifier $G_{\mathcal{Q}}$ and Sparsifier $G_{\mathcal{P}}$ are generators which are used to fit the map $\mathcal{P} \to \mathcal{Q}$ and $\mathcal{Q} \to \mathcal{P}$.
- In addition, we introduce Sparse discriminator $D_{\mathcal{P}}$ and Dense discriminator $D_{\mathcal{Q}}$.

where $D_{\mathcal{P}}$ aims to distinguish between $\mathcal{P}$ and generated sparse point sets $G_{\mathcal{P}}(\mathcal{Q})$, $D_{\mathcal{Q}}$ aims to discriminate between $\mathcal{Q}$ and $G_{\mathcal{Q}}(\mathcal{P})$.
Through the sparse-dense-sparse (shown in green arrows) and dense-sparse-dense data (shown in red arrows) cycles, our model is expected to capture the inherent upsampling patterns and generate dense patches that are uniformly distributed on the target surface.

Loss Function

Quadratic Wasserstein Loss The objective function of discriminator DP and DQ is:

$$\begin{aligned}\min_w \frac{1}{2}\left(\frac{1}{B}\sum_{i=1}^{B} D_w(y_i) - \frac{1}{B}\sum_{i=1}^{B}\phi(y_i)\right)^2 \\ + \frac{1}{2}\left(\frac{1}{B}\sum_{i=1}^{B}(D_w(y_i) - \psi(y_i))^2\right) \\ + \frac{\gamma}{B}\sum_{i=1}^{B}(\|\nabla_x D_w(x_i)\| - d_{EM}(x_i, y_i))^2\end{aligned}$$

Further, to stabilize the optimization of the generator, we set the adversarial loss of the generators $G_{\mathcal{Q}}$ and $G_{\mathcal{P}}$ as a quadratic function, which is

$$\mathcal{L}_{adv}(x_i, y_i) = \left(\frac{1}{B}\sum_{i=1}^{B} D_w(y_i) - \frac{1}{B}\sum_{i=1}^{B} D_w(x_i)\right)^2$$

Cycle Consistency Loss To eliminates the need of paired data, we proposed an point cloud upsampling consistency loss which is defined as follows:

$$\mathcal{L}_{cyc} = d_{EM}(p_i, G_P(G_Q(p_i))) + d_{EM}(q_i, G_Q(G_P(q_i)))$$

Self Restraint Loss Without paired point sets as the supervision, we define a self-restraint loss to ensure that the generated points are distributed on the underlying surface. Herein, self-restraint loss uses the Chamfer distance to measure the loss between sparsified or densified point set and the original one, which is

$$\mathcal{L}_{sel}(x_i, z_i) = \frac{1}{N}\sum_{j=1}^{N}\min_{x_{ik}\in x_i}\|z_{ij} - x_{ik}\|_2^2 + \frac{1}{rN}\sum_{k=1}^{rN}\min_{z_{ij}\in z_i}\|z_{ij} - x_{ik}\|_2^2$$

Experimental Results

To compare with baseline methods, we first train models on PU1K and PU-GAN's datasets which are cropped into dense and downsampled sparse point clouds patches respectively. Then we train a model with upaired real-scanned sparse KITTI and dense SEMANTIC3D data to demonstrate the capability and advantages of our method in the real applications with unpaired data.
We qualitatively and quantitatively compare our method with several baselines.

Quantitative Analysis

Method	P2F(10^{-3})	CD(10^{-3})	HD(10^{-3})
PU-Net	4.834	1.155	15.170
MPU.	3.551	0.935	13.327
PU-GAN	**1.590**	**0.420**	5.390
PU-GCN	2.499	0.585	7.577
Dis-PU	3.143	1.151	14.680
Ours	2.080	0.551	**2.919**

Table 1. Comparisons on PU1K against supervised methods.

As shown in Table 1, although PU-CycGAN is weakly supervised, it clearly outperforms the strong supervision method PU-Net, MPU, PU-GCN and Dis-PU in all three metrics.

Qualitative Analysis

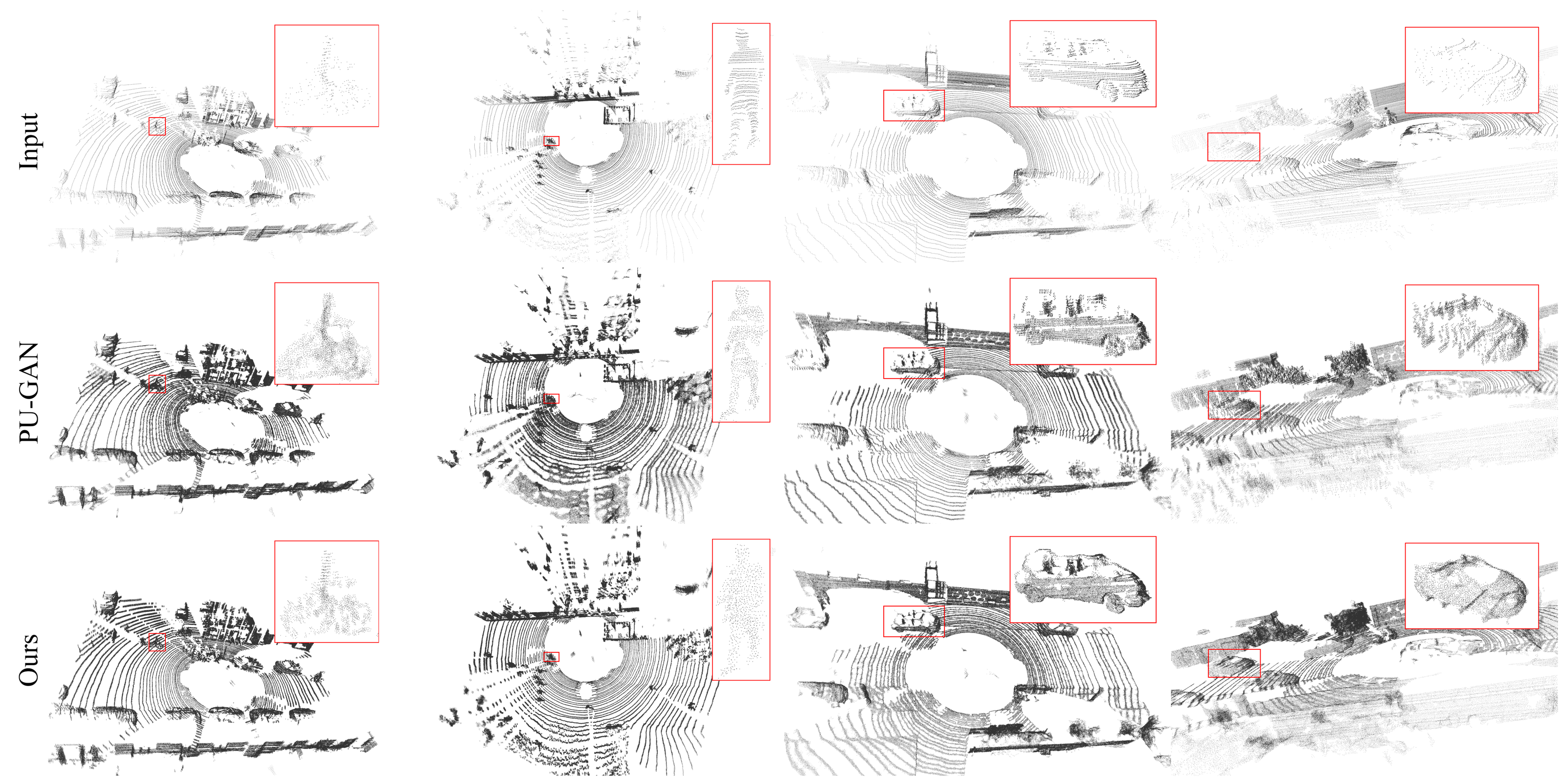

Figure 2. Qualitative comparisons on KITTI.

Fig. 2 compares PU-CycGAN against the most competitive upsampling methods PU-GAN on KITTI. We observe that PU-GAN tends to produce more outliers and overfill holes (e.g. the first close-up on bicycle wheels and the second close-up), while PU-CycGAN preserves better underlying surface and repair fine-grained details.