|  |  |
| --- | --- |
| 研究所/中心 | 可视计算研究中心 |
| 成果（论文）名（英文） | LOGAN: Unpaired Shape Transform in Latent Overcomplete Space |
| 年份 | 2019 |
| 刊物名（英文） | ACM Transactions on Graphics (Proceedings of SIGGRAPH ASIA 2019) |
| 作者（多个作者按顺序用“,”隔开）（英文） | Kangxue Yin1, Zhiqin Chen1, Hui Huang2, Daniel Cohen-Or2,3, Hao Zhang1 |
| 全部作者单位（英文） | Simon Fraser University,Shenzhen University, Tel Aviv University |

### LOGAN: Unpaired Shape Transform in Latent Overcomplete Space

**ACM Transactions on Graphics (Proceedings of SIGGRAPH ASIA 2019)**

Kangxue Yin1    Zhiqin Chen1    [Hui Huang](https://vcc.tech/~huihuang" \t "_blank)2    [Daniel Cohen-Or](http://www.math.tau.ac.il/~dcor/" \t "_blank)2,3    [Hao Zhang](https://www2.cs.sfu.ca/~haoz/" \t "_blank)1

1Simon Fraser University    2Shenzhen University    3Tel Aviv University

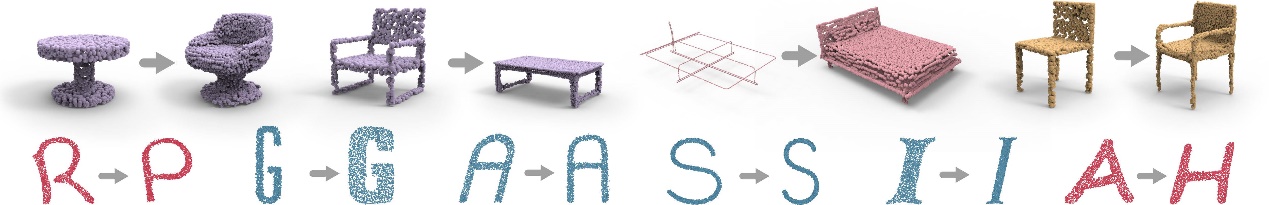


Fig. 1. We present LOGAN, a deep neural network which learns general-purpose shape transforms from unpaired domains. By altering only the two input data domains for training, without changing the network architecture or any hyper-parameters, LOGAN can transform between chairs and tables, from cross-sectional profiles to surfaces, as well as adding arms to chairs. It can also learn both style-preserving content transfer (letters R → P, A → H, in different font styles) and content-preserving style transfer (wide to narrow S, thick to thin I , thin to thick G, and italic to non-italic A.)

**Abstract**

We introduce LOGAN, a deep neural network aimed at learning *general-purpose* shape transforms from *unpaired* domains. The network is trained on two sets of shapes, e.g., tables and chairs, while there is neither a pairing between shapes from the domains as supervision nor any point-wise correspondence between any shapes. Once trained, LOGAN takes a shape from one domain and transforms it into the other. Our network consists of an autoencoder to encode shapes from the two input domains into a *common latent* space, where the latent codes *concatenate multi-scale* shape features, resulting in an *overcomplete* representation. The translator is based on a generative adversarial network (GAN), operating in the latent space, where an adversarial loss enforces cross-domain translation while a *feature preservation loss* ensures that the right shape features are preserved for a natural shape transform. We conduct ablation studies to validate each of our key network designs and demonstrate superior capabilities in unpaired shape transforms on a variety of examples over baselines and state-of-the-art approaches. We show that LOGAN is able to learn what shape features to preserve during shape translation, either local or non-local, whether content or style, depending solely on the input domains for training.

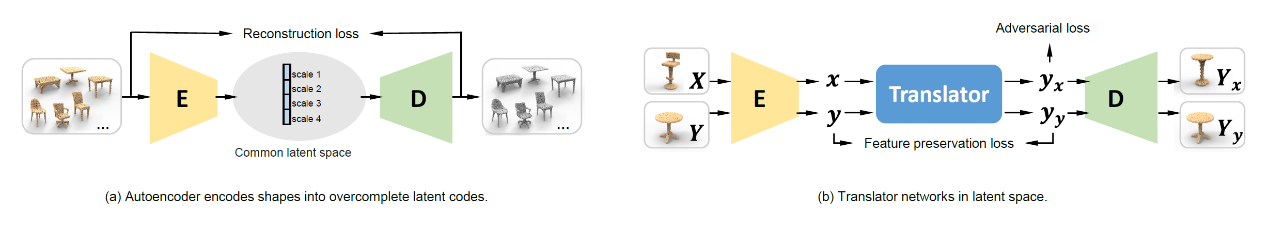


Fig. 2. Overview of our network architecture, which consists of an autoencoder (a) to encode shapes from two input domains into a common latent space which is overcomplete, and a GAN-based translator network (b) designed with an adversarial loss and a loss to enforce feature preservation.

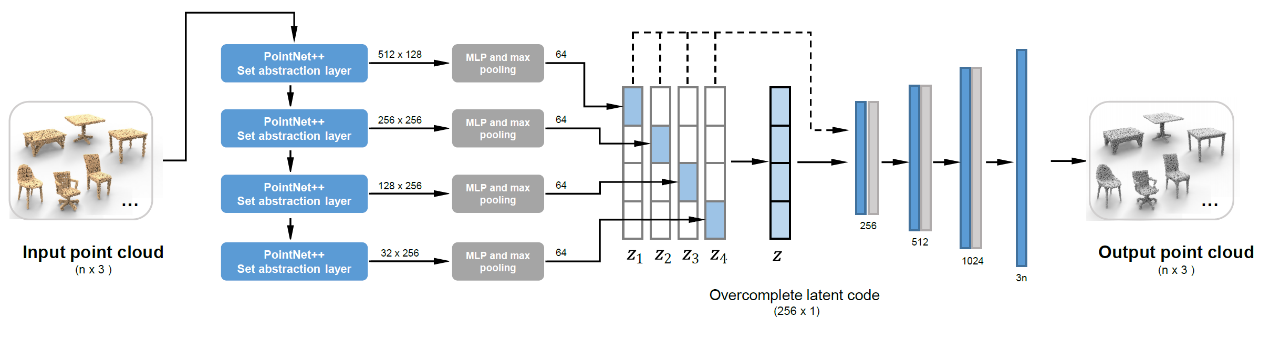


Fig. 4. Architecture of our multi-scale, overcomplete autoencoder. We use the set abstraction layers of PointNet++ [Qi et al. 2017b] to produce point features in different scales and aggregate them into four sub-vectors: z1, z2, z3, and z4. The four sub-vectors are padded with zeros and summed up into a single 256-dimensional latent vector z that is overcomplete; the z vector can also be seen as a concatenation of the other four sub-vectors. During training, we feed all the five 256-dimensional vectors to the decoder. In the decoder, the blue bars represent fully-connected layers; grey bars represent ReLU layers.

|  |  |
| --- | --- |
|  |  |
| Fig. 6. Architecture of our translator network. The blue bars represent fullyconnected layers; orange bars represent BN-ReLU layers. | Fig. 7. Architecture of the upsampling layer of our network after shape translation.We predict *m* local displacement vectors for each of the *n* points in the sparse point cloud, which results in a dense set of *mn* points. |



Fig. 8. Comparing chair-table translation results using different network configurations. Top four rows: chair → table. Rest: table → chair. (a) Test input. (b) LOGAN results with and without upsampling. (c) Retrieved training shapes from the target domain which are closest to the test input (left) and to our translator output (right). The retrieval was based on EMD between point clouds at 2,048 point resolution. Note that the chair dataset from ShapeNet has some benches mixed in, which are retrieved as “tables.” (d) Baseline AE 1 as autoencoder + our translator network. (e) Baseline AE 2 (λ1 = 0) + our translator network. (f) Our autoencoder (λ1 = 0.1) + WGAN & Cycle loss. (g) Our autoencoder (λ1 = 0.1) + WGAN & feature preservation (FP) loss.

|  |  |
| --- | --- |
|  |  |
| Fig. 10. Unpaired shape transforms between armchairs and armless chairs. The first two rows show results of armrest removal by LOGAN, while the last two rows show insertion. On the right, we show the mesh editing results guided by the learned point cloud transforms. | Fig. 11. Unpaired shape transforms between tall and short tables. Left: increasing height. Right: decreasing height. |

|  |  |
| --- | --- |
|  |  |
| Fig. 12. Comparisons on content-preserving style transfer, i.e., *regularA/HitalicA/ H, thinG/R-thickG/R,* and *wideM/N-narrowM/N* translations, by different methods. First two rows: regular-to-italic; middle two rows: thin-tothick; last two rows: wide-to-narrow. From left to right: input letter images; corresponding input point clouds; output point clouds from LOGAN; images reconstructed from our results; output images of CycleGAN; outputs from UNIT [Liu et al. 2017]; outputs from MUNIT [Huang et al. 2018]. For *wideM/N-narrowM/N* we align the letters by height for better visualization. | Fig. 13. Comparisons on style-preserving content transfer, i.e., *A-H, G-R,* and *M-N* translations, by different methods, including ground truth. |



Fig. 15. Comparisons between various network configurations, (supervised) P2P-NET, and ground truth targets, on shape transform examples from P2P-NET: skeleton→shape (rows 1-2), scan→surface (rows 3-4), and (cross-sectional) profiles→surface (rows 5-6). All point clouds have 2,048 points.

**Data & Code**

**Note that the DATA and CODE are free for Research and Education Use ONLY.**

**Please cite our paper (add the bibtex below) if you use any part of our ALGORITHM, CODE, DATA or RESULTS in any publication.**

**Link：[https://github.com/kangxue/LOGAN](https://github.com/kangxue/LOGAN" \t "_blank)**

**Acknowledgements**

The authors would like to thank the anonymous reviewers for their valuable comments. Thanks also go to Haipeng Li and Ali Mahdavi- Amiri for their discussions, and Akshay Gadi Patil for proofreading. This work was supported by NSERC Canada (611370), gift funds from Adobe, NSF China (61761146002,61861130365), GD Science and Technology Program (2015A030312015), LHTD (20170003), ISF (2366/16), and ISF-NSFC Joint Research Program (2472/17).

**Bibtex**  
@article{LOGAN19,  
title = {LOGAN: Unpaired Shape Transform in Latent Overcomplete Space},  
author = {Kangxue Yin and Zhiqin Chen and Hui Huang and Daniel Cohen-Or and Hao Zhang},  
journal = {ACM Transactions on Graphics (Proceedings of SIGGRAPH ASIA 2019)},  
volume = {38},  
number = {6},  
pages = {198:1--198:13},    
year = {2019},  
}

**Downloads**

|  |  |
| --- | --- |
| 研究所/中心 | 可视计算研究中心 |
| 成果（论文）名（英文） | LOGAN: Unpaired Shape Transform in Latent Overcomplete Space |
| 年份 | 2019 |
| 刊物名（英文） | ACM Transactions on Graphics (Proceedings of SIGGRAPH ASIA 2019) |
| 作者（多个作者按顺序用“,”隔开）（英文） | Kangxue Yin1, Zhiqin Chen1, Hui Huang2, Daniel Cohen-Or2,3, Hao Zhang1 |
| 全部作者单位（英文） | Simon Fraser University,Shenzhen University, Tel Aviv University |