Generative Adversarial Image Synthesis with Decision Tree Latent Controller

Takuhiro Kaneko, Kaoru Hiramatsu, Kunio Kashino
NTT Communication Science Laboratories, NTT Corporation

1. Introduction

Motivation
- Create generative model that enables image generation to be controlled in coarse-to-fine manner
- Select & refine

Relationship to previous GANs

Contributions
- Derive novel functionality in deep generative model
- Propose new extension of GAN called DTLC-GAN
- Discover hierarchically interpretable representations with either unsupervised or weakly supervised settings

2. Related Work

Relationship to previous GANs

3. Proposed: DTLC-GAN (Decision Tree Latent Controller GAN)

Relationship to previous GANs

4. Experiments

1. Representation comparison

- Dataset: MNIST (Unsupervised)
- Categories: 20 (flat) vs. 10 x 2 (hierarchical)
- Evaluation metric: Inception score [Salimans+2016]

Hierarchical Conditional Mutual Information Regularization

- Discover hierarchically disentangled representations on basis of information gain
- Regularization for \( c \)
  - Unsupervised setting (\( c_{\text{flat}} \) latent)
  - Weakly supervised setting (\( c_{\text{sup}} \) supervised)
- Auxiliary classifier (AC-GAN)
- Full objective
  - Hierarchical conditional mutual information
  - Minimized for \( G, Q_1, \ldots, Q_L \)

2. Ablation study on curriculum learning

- Dataset: CIFAR-10 (Weakly supervised)
- Categories: 10 x 3 x 3 x 3
- Evaluation metric: For each layer, measure inter-category similarity on basis of SSIM

DTLC-GAN

3. Effect on image quality (w/ WGAN-GP)

- Dataset: CIFAR-10 (Unsupervised/supervised)
- Categories: 10 x 3 (L = 0, 1, 2) \( \Rightarrow \) 810 in L = 4
- Evaluation metric: Inception score [Salimans+2016]

4. Extension to continuous codes

- Dataset: 3D Faces (Unsupervised)
- Categories: 5 (discrete) \times 1 (continuous)

5. Application to image retrieval

- Dataset: CelebA (Weakly supervised)
- Categories: 1 (w/o attribute) \times 1 (w/ attribute) 3 x 3