

2024年度 前期雑誌会

情報認識学研究室 M1 吉川 大貴

目次

- 研究背景
- 関連研究
 - 直交座標系を学習する手法
 - 線形ベクトル演算を定義する手法
 - ベクトル場を定義する手法
 - 曲線座標系を定義する手法
- 修論に向けて

研究背景

研究背景

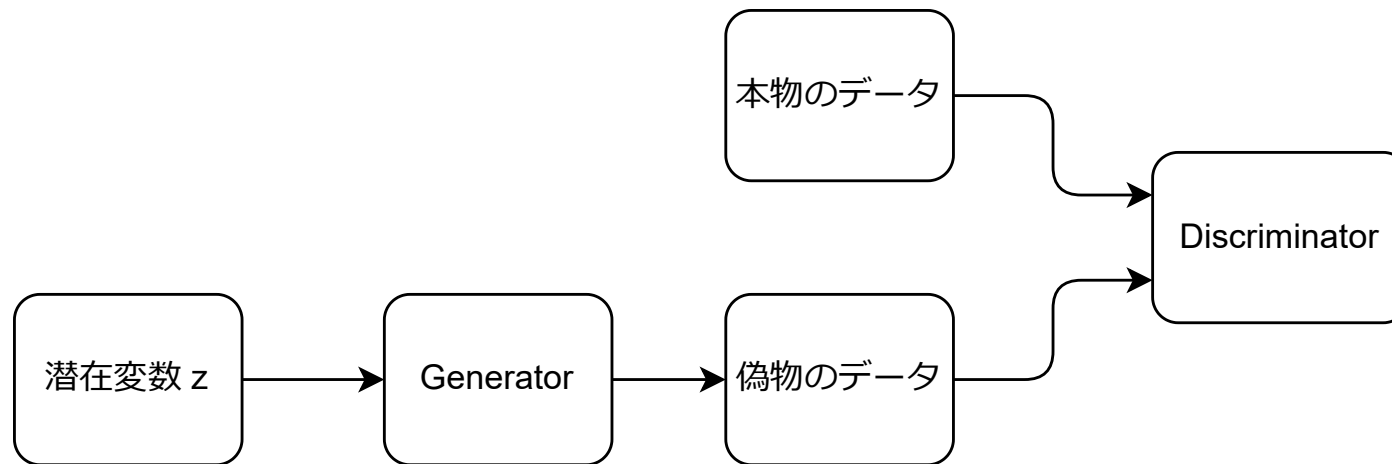
- 画像を自由に作り出すことはコンピュータビジョンの研究における究極のゴールの一つ [1]
- 深層生成モデルを用いた画像生成手法が提案されてきた
 - 例：GAN, VAE, 拡散モデル

[1] T. Aoshima, T. Matsubara (2023). Deep Curvilinear Editing: Commutative and Nonlinear Image Manipulation for Pretrained Deep Generative Model. CVPR

研究背景

◆ GAN (Generative Adversarial Networks) [2]

- 2つの敵対的なネットワークを競わせることで学習
- 生成器は識別機を騙すように、識別器は偽物を見破るように学習
- 学習後は潜在変数から生成器を通してデータを生成



[2] Goodfellow, I. J., et al. (2014). Generative adversarial nets NIPS

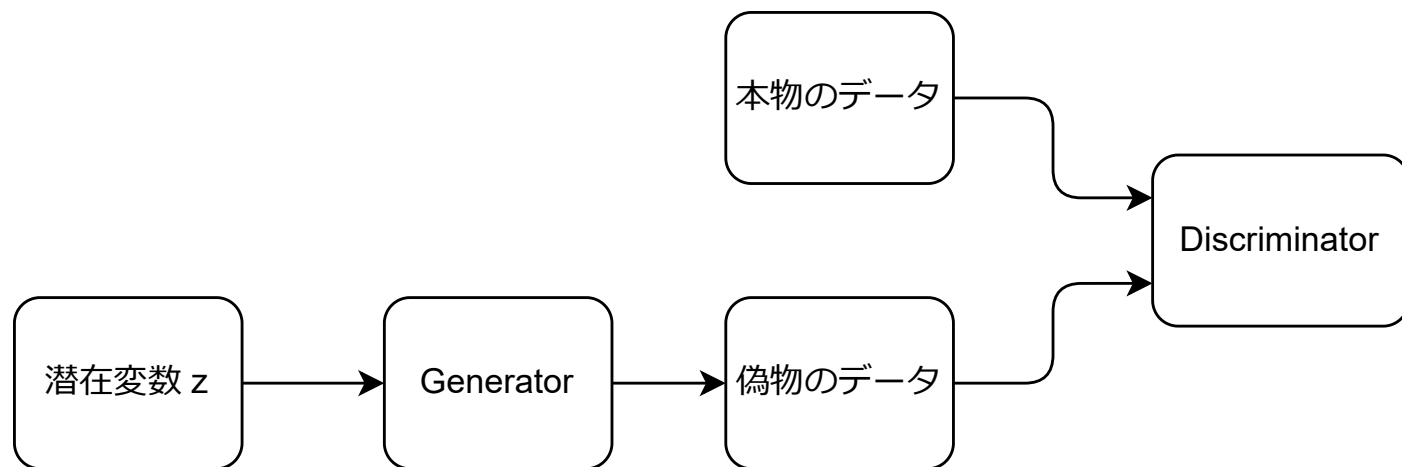
2024年度前期雑誌会

研究背景

◆ GAN (Generative Adversarial Networks) [2]

- 生成器 G と識別器 D の最適化問題

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$

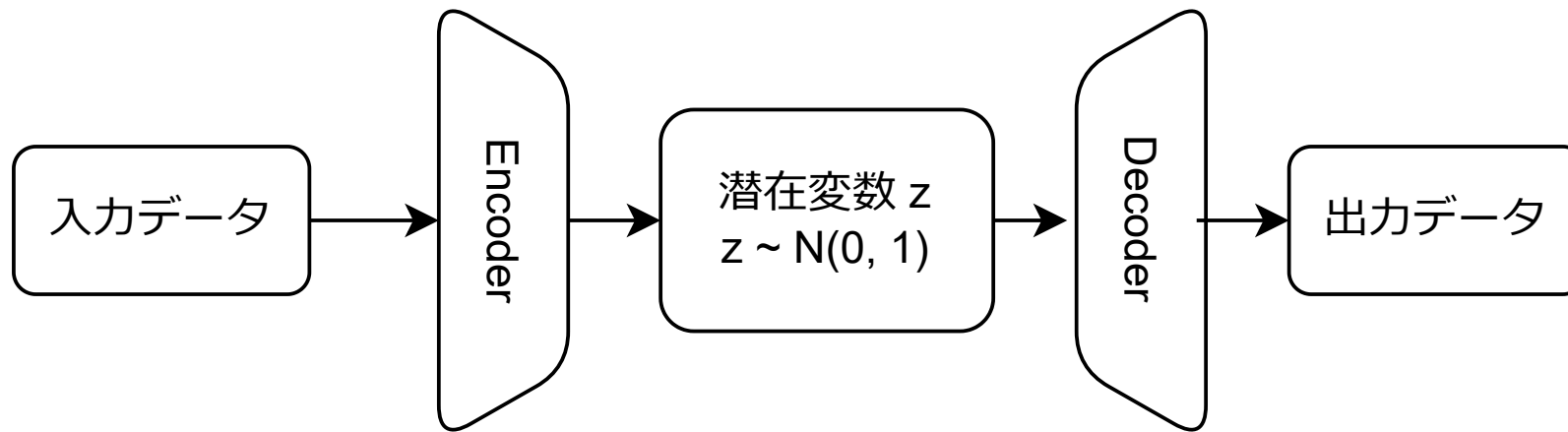


[2] Goodfellow, I. J., et al. (2014). Generative adversarial nets NIPS

研究背景

◆ VAE (Variational Autoencoder) [3]

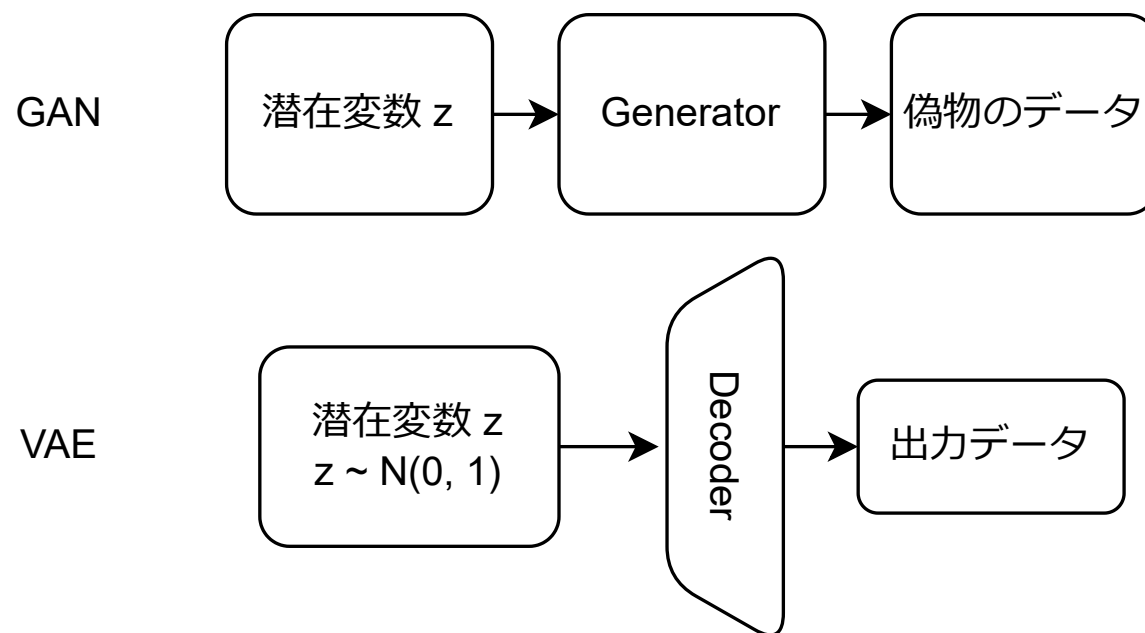
- EMアルゴリズムと変分ベイズを融合させた手法
- 潜在空間におけるデータの分布を学習
- 学習後は潜在変数からデコーダーを用いてデータを生成



[3] Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes ICLR

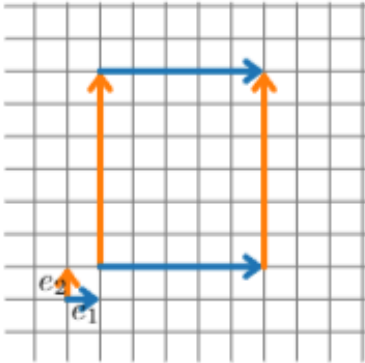
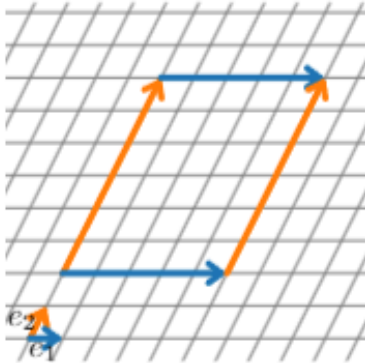
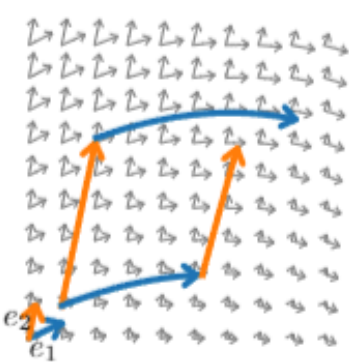
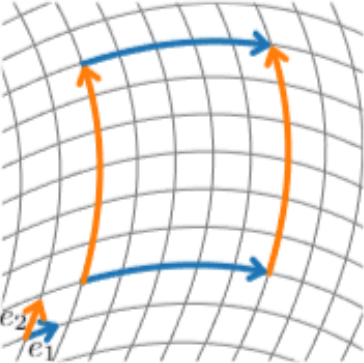
研究背景

- 深層生成モデルにおいて、生成される画像は潜在空間上の点の座標で決まる
→ 潜在変数の座標を操作することで画像編集が可能



研究背景

- 自然で高精度な編集をするための手法が複数提案されている [1]

	Training under constraints [10, 45, 46]	Linear arithmetic [27, 55, 62]	Vector fields/Local basis [12, 52, 60]	DeCurvEd (proposed)
Global coordinate	Cartesian	oblique	(only local)	curvilinear
No retraining	✗	✓	✓	✓
Nonlinear edit	—	✗	✓	✓
Commutative edit	✓	✓	✗	✓
Conceptual diagram				

[1] T. Aoshima, T. Matsubara (2023). Deep Curvilinear Editing: Commutative and Nonlinear Image Manipulation for Pretrained Deep Generative Model. CVPR

関連研究

関連研究

直交座標系を学習する手法

- 表現ベクトルの各要素に特定の属性が紐づけられるように制約
 - GANにおいて潜在変数と画像の相互情報量を最大化 [4]
 - VAEにおいて目的関数のKL項の影響を大きくする [5]

問題点

- 再学習が必要となり、学習コストが高い
- 学習が不安定化し、生成画像の品質が低下する
→ 学習済みモデルに対して編集を行う手法が必要

[4] Xi Chen. et al.(2016) InfoGAN: Interpretable Representation Learning by Information Maximizing Generative Adversarial Nets

[5] I. Higgins et al. (2017) beta-VAE: Learning Basic Visual Concepts with a Constrained Variational Framework

関連研究

線形ベクトル演算を定義する手法

- 学習済みモデルのパラメータを解析 [6]
- PCAを用いて潜在空間において属性に対応する主成分を抽出 [7]
- 生成された画像を属性によってグループ分けし、その差を属性ベクトルとみなす [8]
- 表現ベクトルを変化させ、それによって引き起こされる画像の変化をグループ分けする [9]

[6] Y. Shen, et al. (2021) Closed-Form Factorization of Latent Semantics in GANs

[7] E. Härkönen, et al. (2020) GANSpace: Discovering Interpretable GAN Controls

[8] A. Radford, et al. (2016) Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks

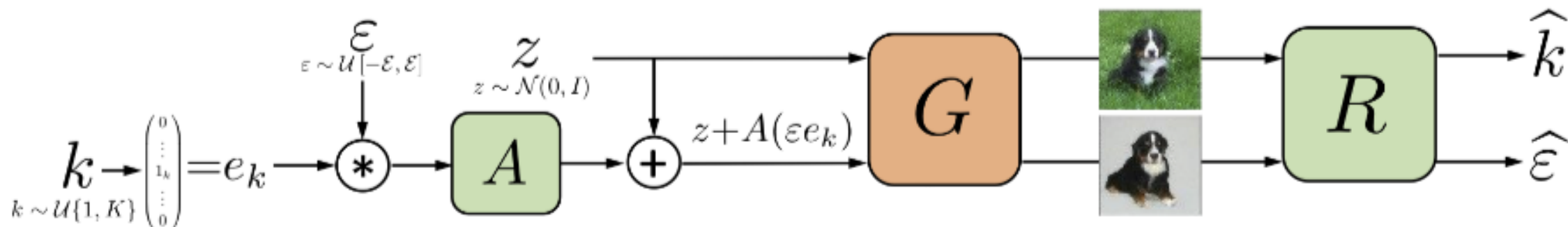
[9] A. Voynov, A. Babenko (2020) Unsupervised Discovery of Interpretable Directions in the GAN Latent Space

関連研究

線形ベクトル演算を定義する手法 [9]

生成器 G は固定したまま、次の2つを学習

1. 行列 $A \in \mathbb{R}^{d \times K}$: 属性インデックスから潜在変数の変化量を決定
2. リンコンストラクタ R : 編集前後の画像から属性ベクトルと変化量を回帰



[9] A. Voynov, A. Babenko (2020) Unsupervised Discovery of Interpretable Directions in the GAN Latent Space

関連研究

線形ベクトル演算を定義する手法

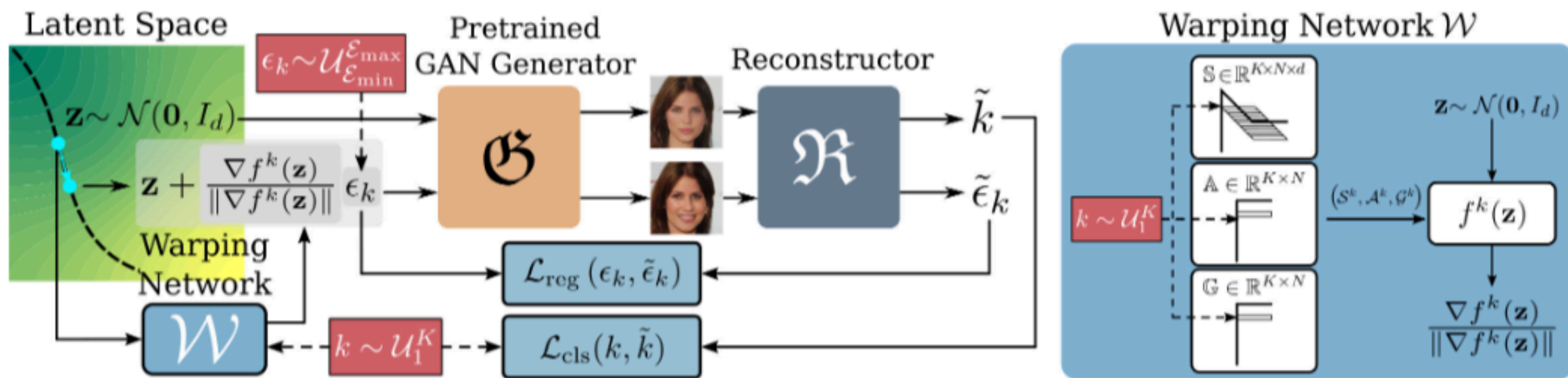
- 現実存在するデータには偏りやゆがみ、属性間の相関がある
- 潜在空間中の基準点によって属性ベクトルの向きが異なる [10]
→属性ベクトルの向きを潜在空間の座標に依存する形で定義

[10] V. Khruikov, et al. (2021) Latent Transformations via NeuralODEs for GAN-based Image Editing

関連研究

ベクトル場を定義する手法

- RBFカーネルの重み付き和で属性ごとのベクトル場を定義 [11]



[11] C. Tzelepis, et al. (2021) WarpedGANSpace: Discovering and Interpolating Interpretable GAN Controls

関連研究

ベクトル場を定義する手法

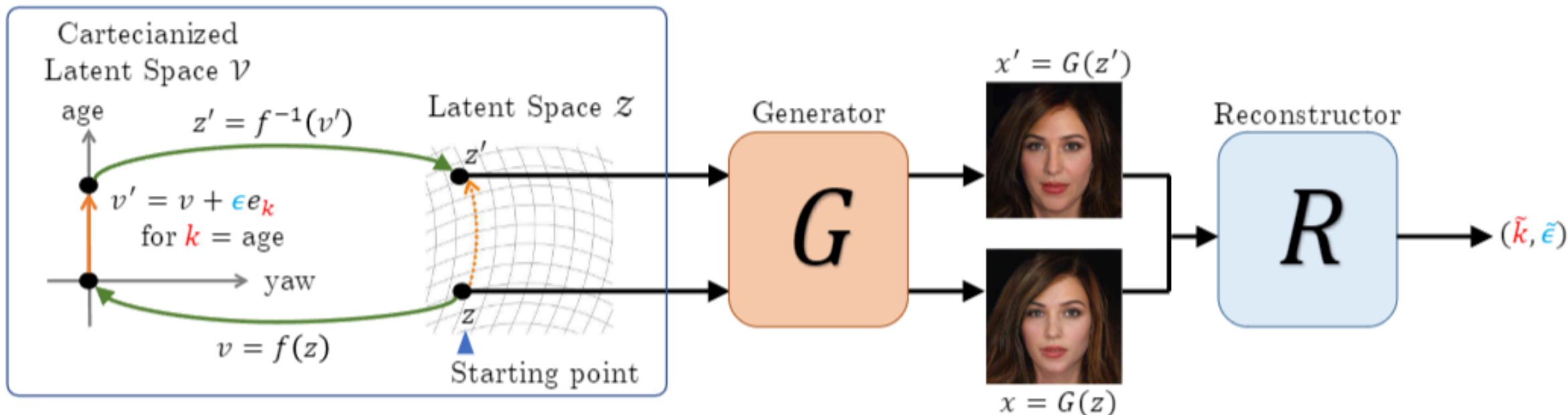
- 座標が局所的にしか定義されていないため、大域的には不整合が起ころ可能性がある [1]
- ベクトル場は一般に非可換であり、編集が非可換になる
→ 可換なベクトル場を定義する手法が必要

[1] T. Aoshima, T. Matsubara (2023). Deep Curvilinear Editing: Commutative and Nonlinear Image Manipulation for Pretrained Deep Generative Model. CVPR

関連研究

曲線座標系を定義する手法 [1]

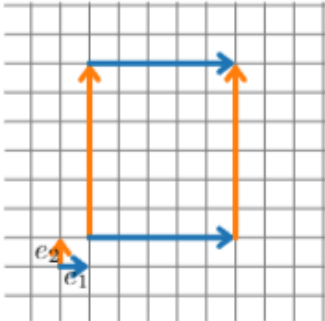
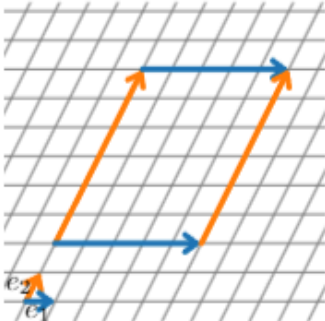
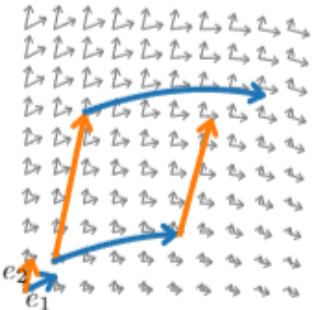
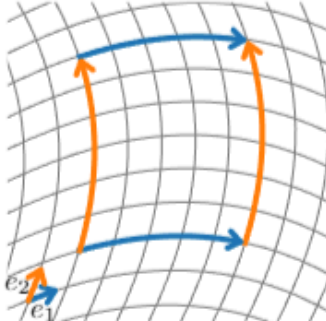
- 潜在空間に曲線座標系を仮定し、直交座標系への写像 $f: \mathcal{Z} \rightarrow \mathcal{V}$ を学習



[1] T. Aoshima, T. Matsubara (2023). Deep Curvilinear Editing: Commutative and Nonlinear Image Manipulation for Pretrained Deep Generative Model. CVPR

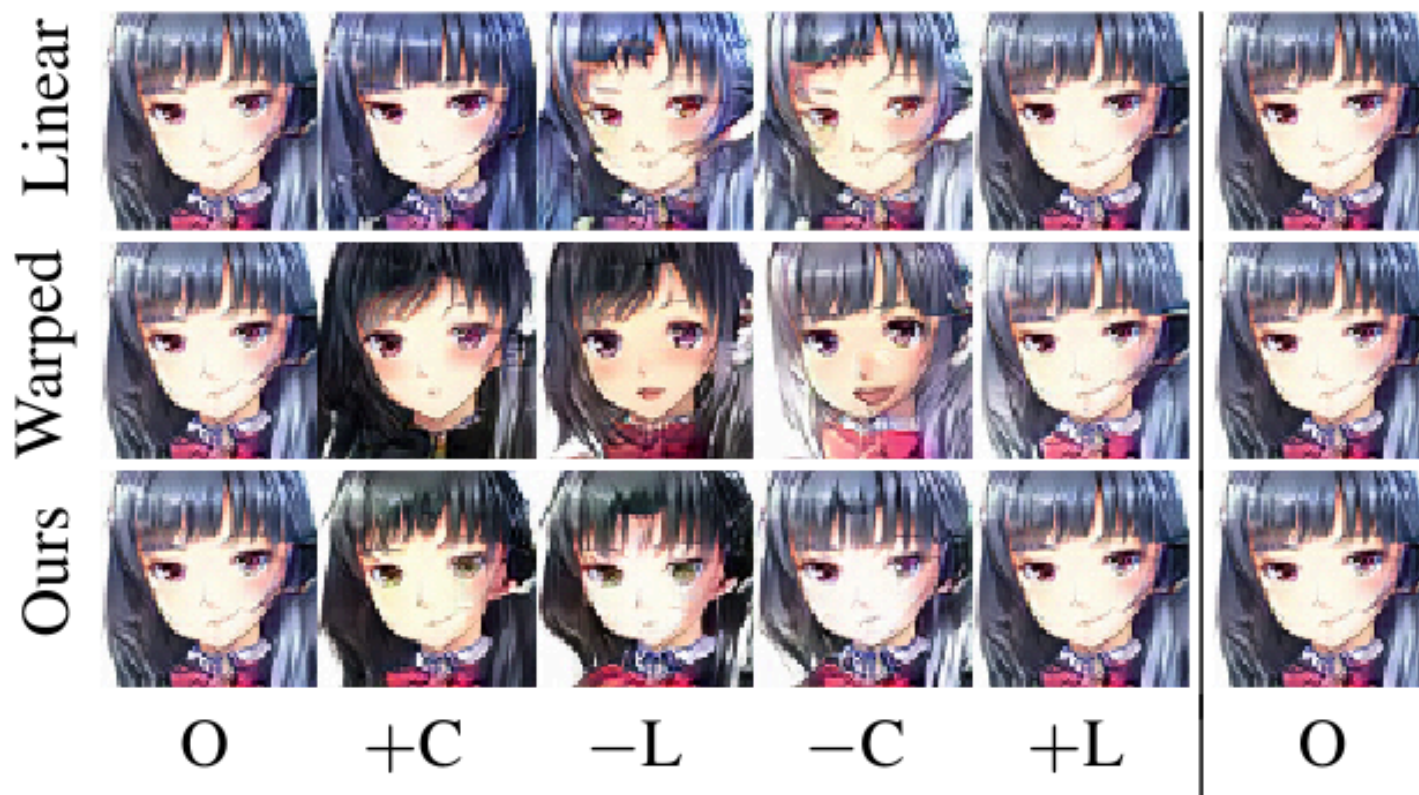
曲線座標系を定義する手法 [1]

- 線形ベクトル演算、ベクトル場の両方の利点を持つ

	Training under constraints [10, 45, 46]	Linear arithmetic [27, 55, 62]	Vector fields/Local basis [12, 52, 60]	DeCurvEd (proposed)
Global coordinate	Cartesian	oblique	(only local)	curvilinear
No retraining	✗	✓	✓	✓
Nonlinear edit	—	✗	✓	✓
Commutative edit	✓	✓	✗	✓
Conceptual diagram				

関連研究

曲線座標系を定義する手法 [1]



[1] T. Aoshima, T. Matsubara (2023). Deep Curvilinear Editing: Commutative and Nonlinear Image Manipulation for Pretrained Deep Generative Model. CVPR

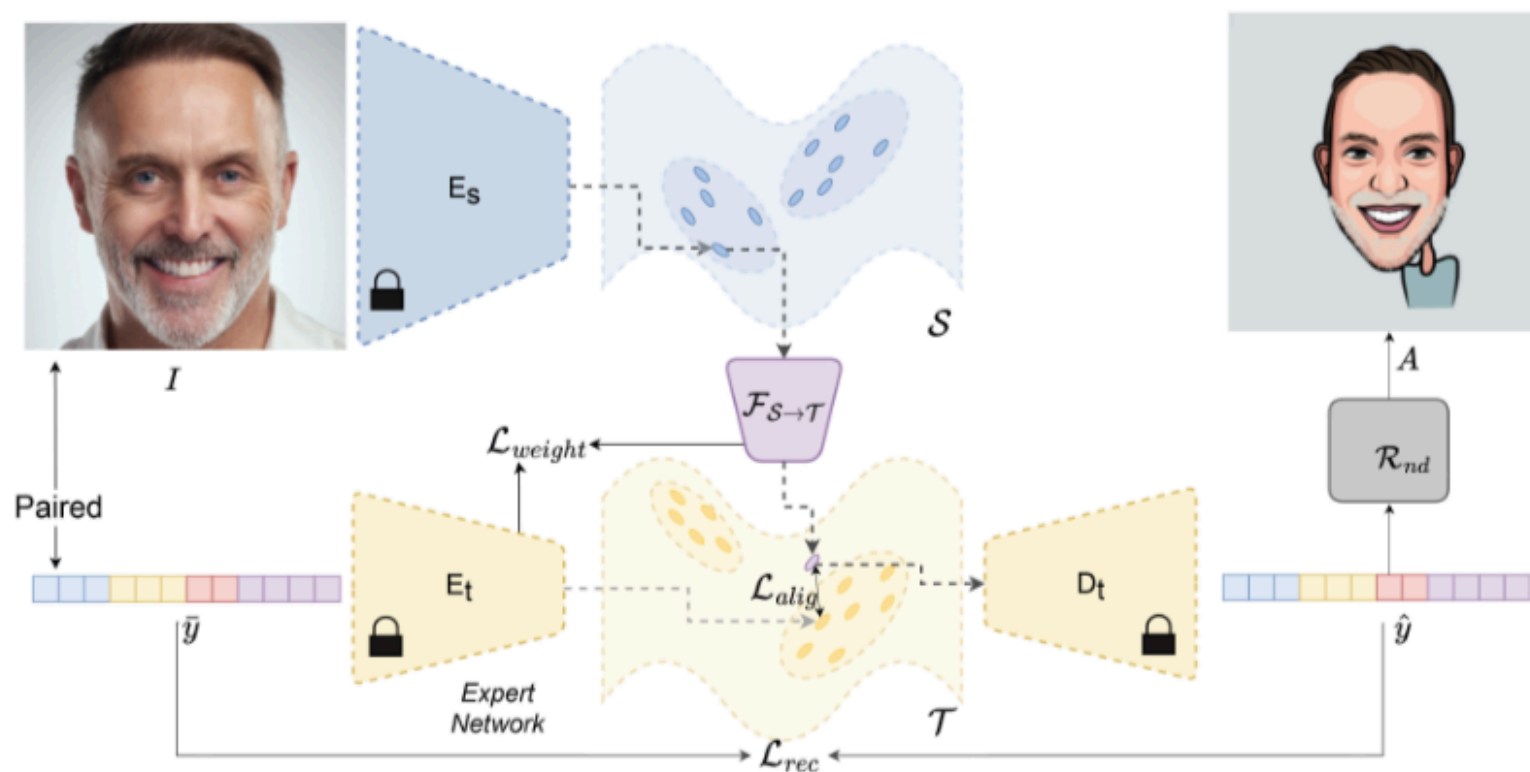
修論に向けて

修論に向けて

- DeCurvEdは潜在変数に条件づけられた任意のモデルに適用可能
 1. VAE, Conditional PixelCNNといったモデルでも有効性を示す
 2. 生成モデルのアラインメント
 - 入力画像と異なるドメインの画像を生成
例：人の顔写真にそっくりなPixar風の画像を生成

修論に向けて

生成モデルのアライメント



M. Ladron de Guevara, et al. (2023) Cross-modal Latent Space Alignment for Image to Avatar Translation.