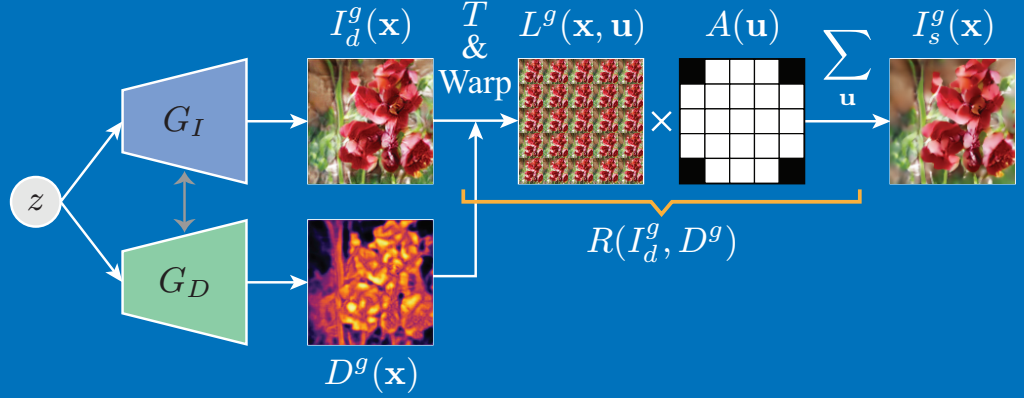
Project page

Unsupervised Learning of Depth and Depth-of-Field Effect from Natural Images with Aperture Rendering Generative Adversarial Networks

Takuhiro Kaneko



NTT Communication Science Laboratories, NTT corporation

Overview

Unsupervised learning of depth and DoF effect from natural images

Unlabeled natural images

Training data

AR-GAN

Training

Deep DoF *Shallow DoF* *Depth*

Generated data

- No ground-truth depth
- No paired data
- No pretrained model

$$G : z \mapsto (I_d^g, I_s^g, D^g)$$

Positioning of research

Fully unsupervised 3D representation learning

Only a collection of unlabeled natural images are available for training

Viewpoint-aware

Learn 3D representation using ***viewpoint*** cues

HoloGAN
[Nguyen-Phuoc+2019]

Szabó et al.
[Szabó+2019]

RGBD-GAN
[Noguchi+2020]

Unsup3d
[Wu+2020]

Viewpoint change

HoloGAN

Inapplicable to a viewpoint-biased dataset

(e.g., Oxford Flowers dataset)

Positioning of research

Fully unsupervised 3D representation learning

Only a collection of unlabeled natural images are available for training

Focus-aware

Learn 3D representation using ***focus*** cues

Applicable to a viewpoint-biased dataset if the dataset includes various DoF images

Training data

(e.g., Oxford Flowers dataset)

Training

Deep DoF *Shallow DoF* *Depth*

AR-GAN

[ours]

Positioning of research

Fully unsupervised 3D representation learning

Only a collection of unlabeled natural images are available for training

Viewpoint-aware

Learn 3D representation using ***viewpoint*** cues

HoloGAN
[Nguyen-Phuoc+2019]

Szabó et al.
[Szabó+2019]

RGBD-GAN
[Noguchi+2020]

Unsup3d
[Wu+2020]

AR-HoloGAN
[ours]

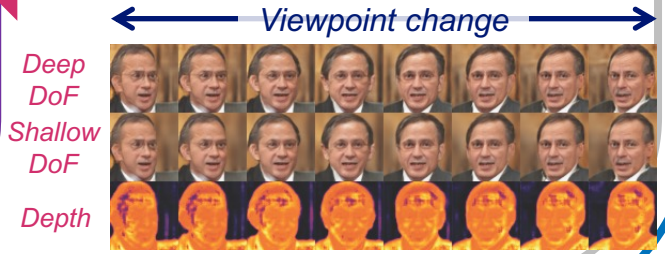
Compatible

AR-RGBD-GAN
[ours]

Focus-aware

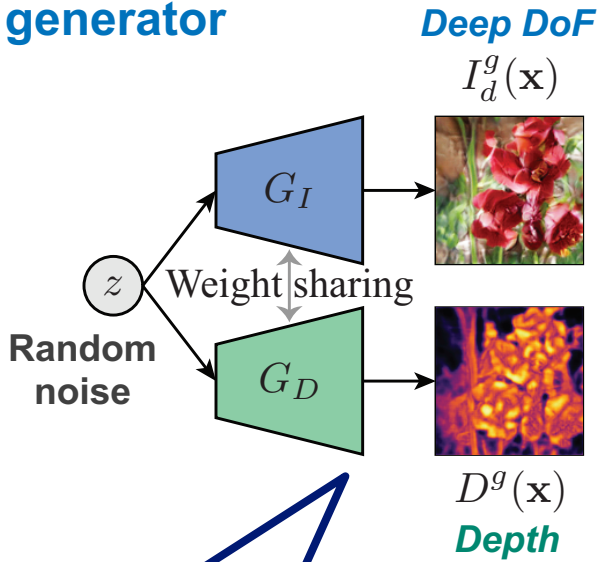
Learn 3D representation using ***focus*** cues

AR-GAN
[ours]



Overall pipeline 1/3

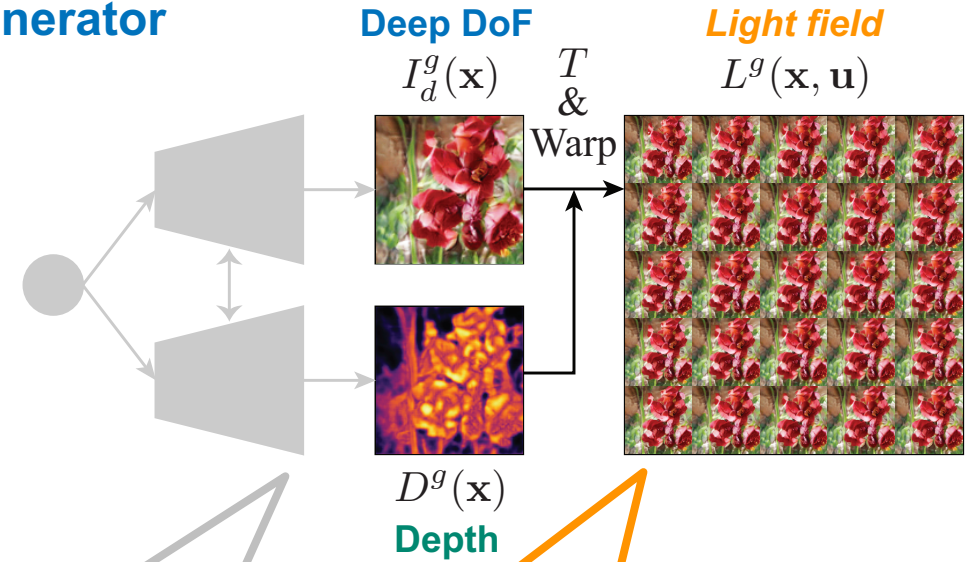
AR-GAN generator



1. Generate pair of ***deep DoF image*** and ***depth*** from **random noise**

Overall pipeline 2/3

AR-GAN generator

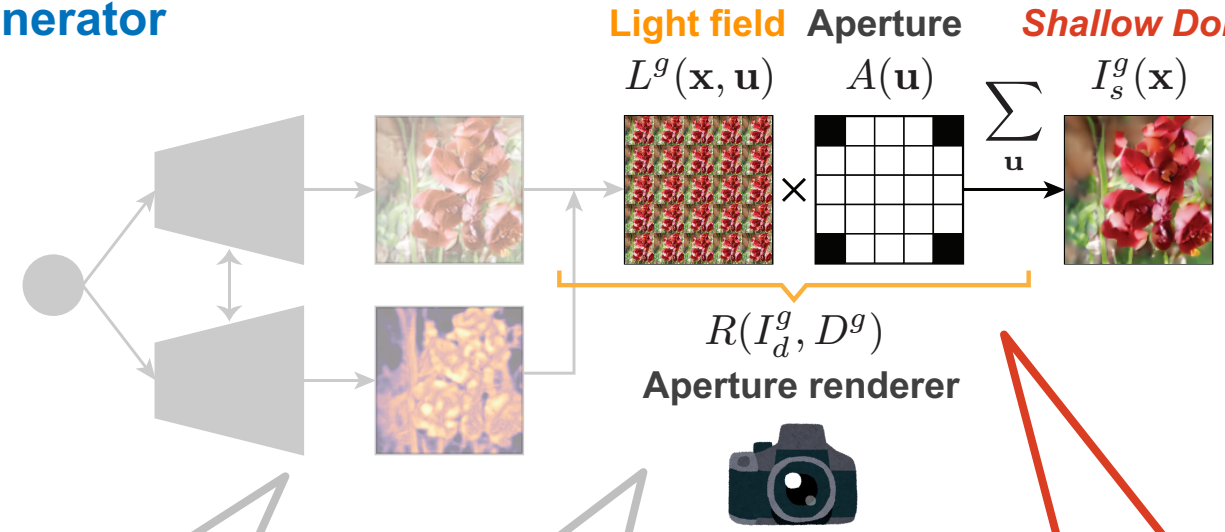


1. Generate pair of ***deep DoF image*** and ***depth*** from **random noise**

2. Simulate ***light field*** by warping **deep DoF image** based on **depth**

Overall pipeline 3/3

AR-GAN generator



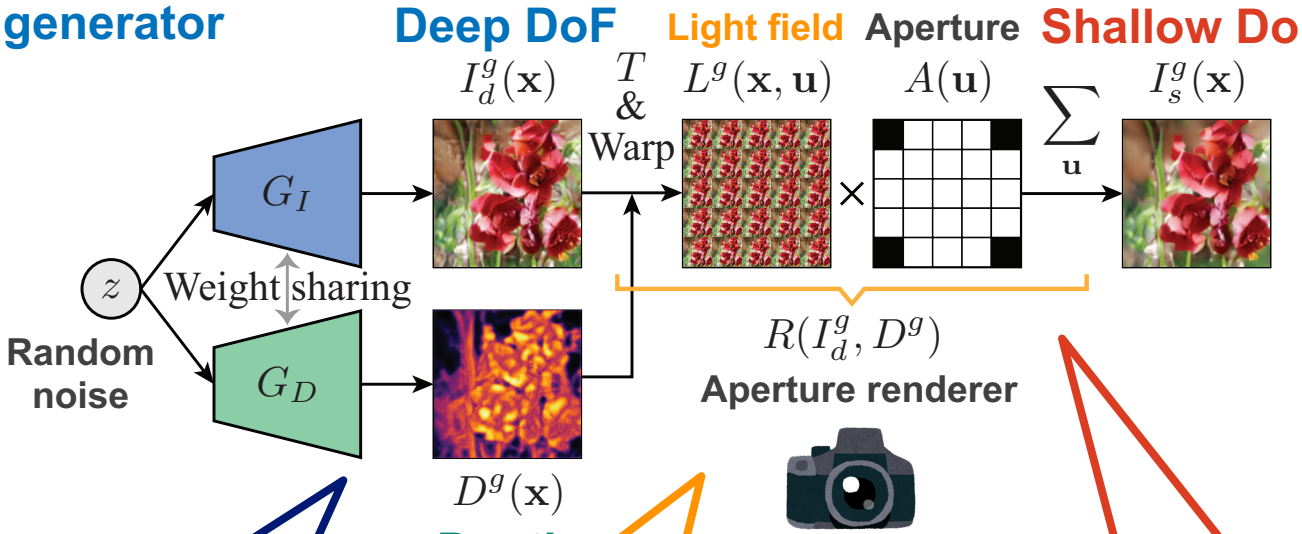
1. Generate pair of ***deep DoF image*** and ***depth*** from **random noise**

2. Simulate ***light field*** by warping **deep DoF image** based on **depth**

3. Render ***shallow DoF image*** by integrating **light field** using **aperture**

Aperture renderer [Srinivasan+2018]

Overall pipeline



1. Generate pair of ***deep DoF image*** and ***depth*** from **random noise**

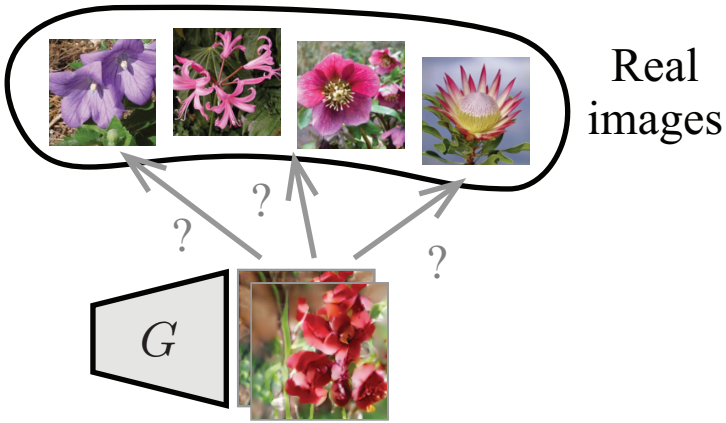
2. Simulate ***light field*** by warping **deep DoF image** based on **depth**

3. Render ***shallow DoF image*** by integrating **light field** using **aperture**

Aperture renderer [Srinivasan+2018]

Standard learning (baseline)

Learn real image distribution without any constraints

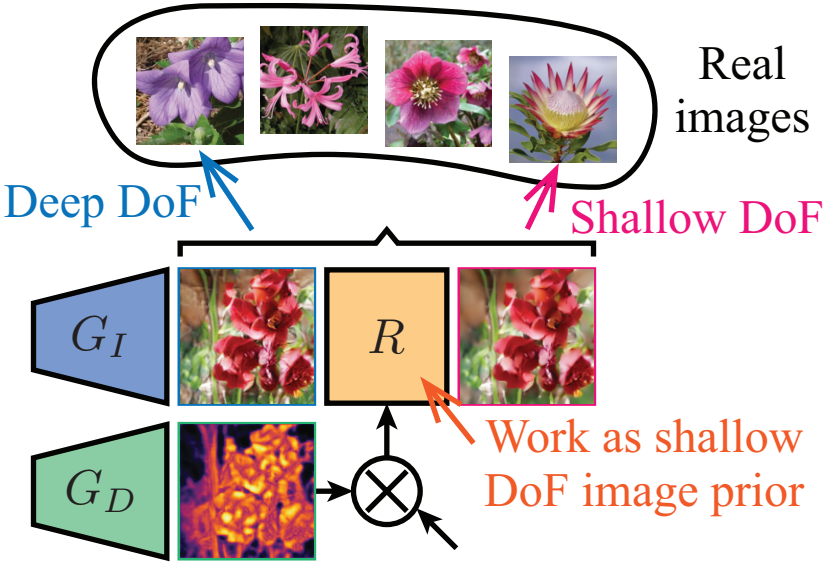


Standard learning (baseline)

$$\mathcal{L}_{\mathrm{GAN}} = \mathbb{E}_{I^r}[\log C(I^r)] + \mathbb{E}_z[\log(1 - C(G(z)))]$$

DoF mixture learning (proposed)

Learn real image distribution while generating diverse DoF images

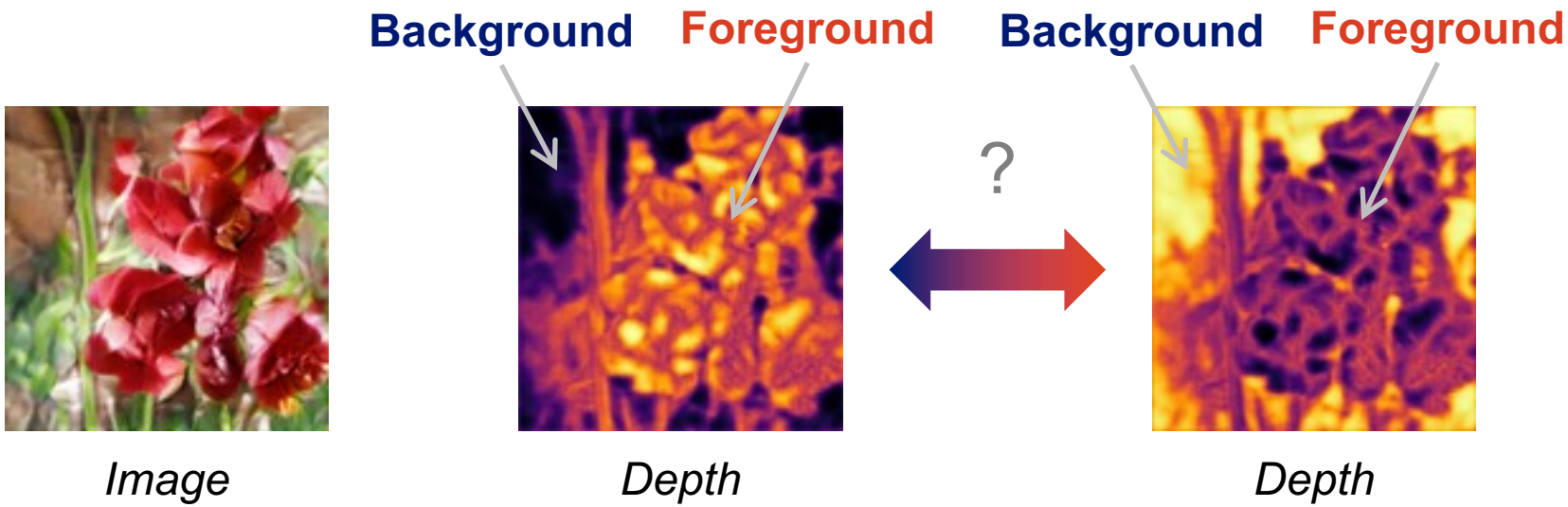


DoF mixture learning (proposed)

$$\mathcal{L}_{\text{AR-GAN}} = \mathbb{E}_{I^r}[\log C(I^r)] + \mathbb{E}_{z,s}[\log(1 - C(R(G_I(z), sG_D(z))))]$$

Difficulty in unsupervised learning

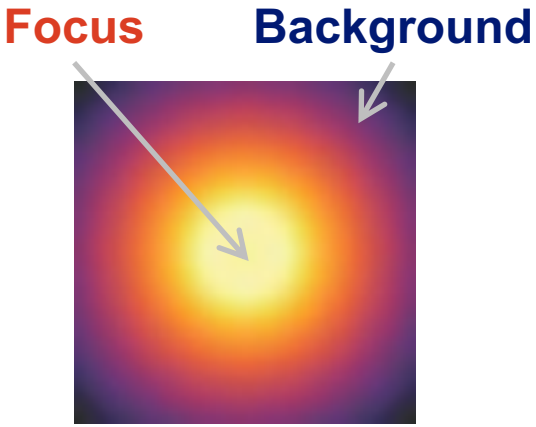
Ambiguity between fore/background blur



Center focus prior

Solve ambiguity between fore/background blur by providing prior

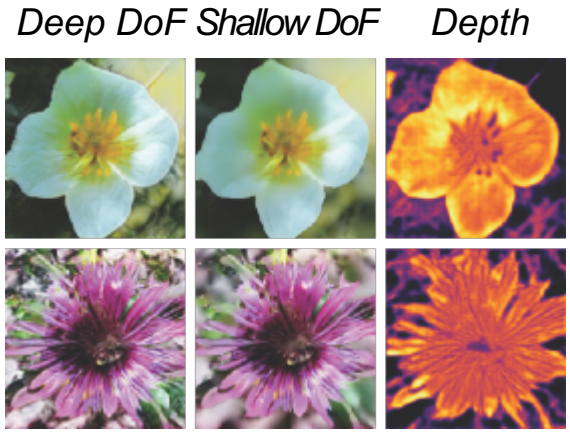
Examples of center focused images



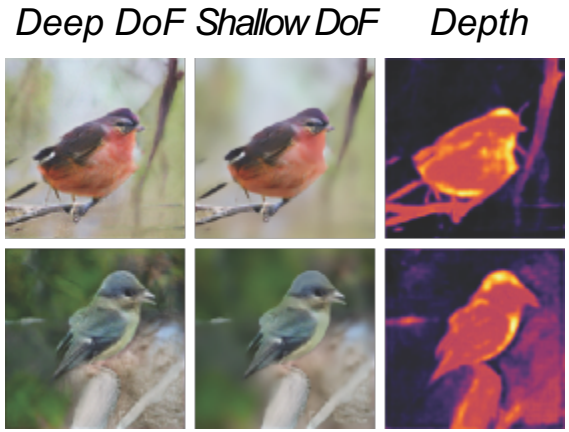
Center focus prior

Examples of generated data

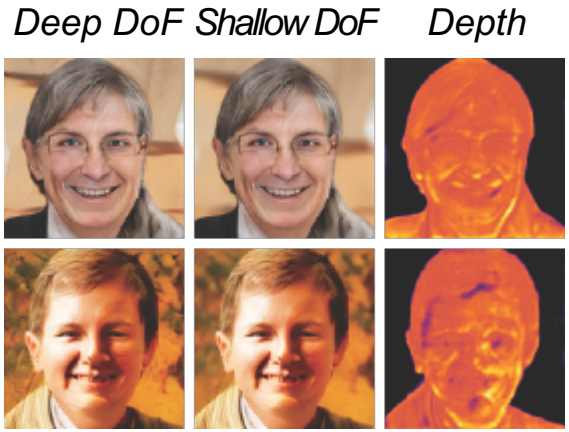
Examples of data generated from AR-GAN



Generated data
[Oxford Flowers]



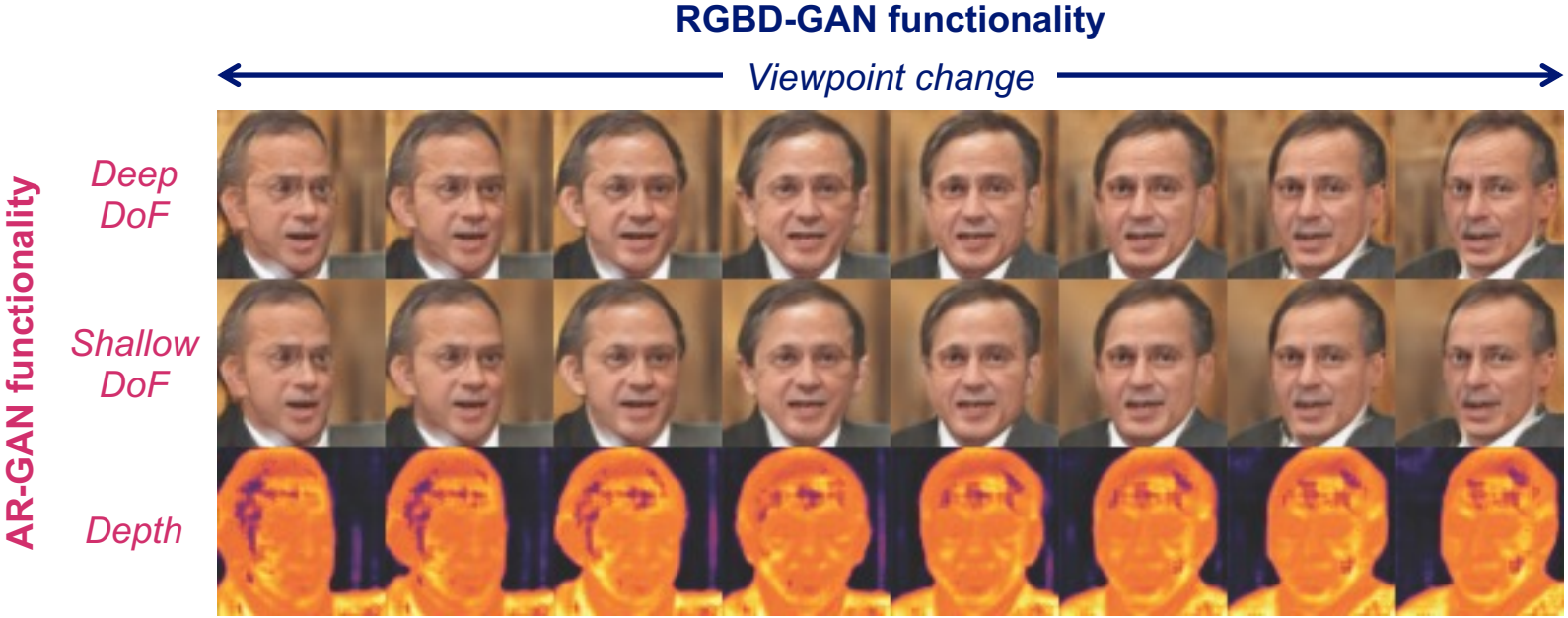
Generated data
[CUB-200-2011]



Generated data
[FFHQ]

Portability analysis

Examples of data generated from AR-RGBD-GAN (AR-GAN + RGBD-GAN)



Generated data
[FFHQ]

RGBD-GAN [Noguchi+2020]

Application in shallow DoF rendering

NTT

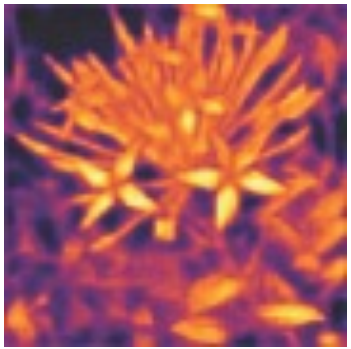
Learn shallow DoF renderer (Deep DoF → Shallow DoF) using generated data

Deep DoF

Input
(iPhone photo)

Shallow DoF

Depth



AR-GAN-DR
(proposed)

W/ ***no*** supervision

Shallow DoF

CycleGAN
(baseline)

W/ ***set-level*** supervision

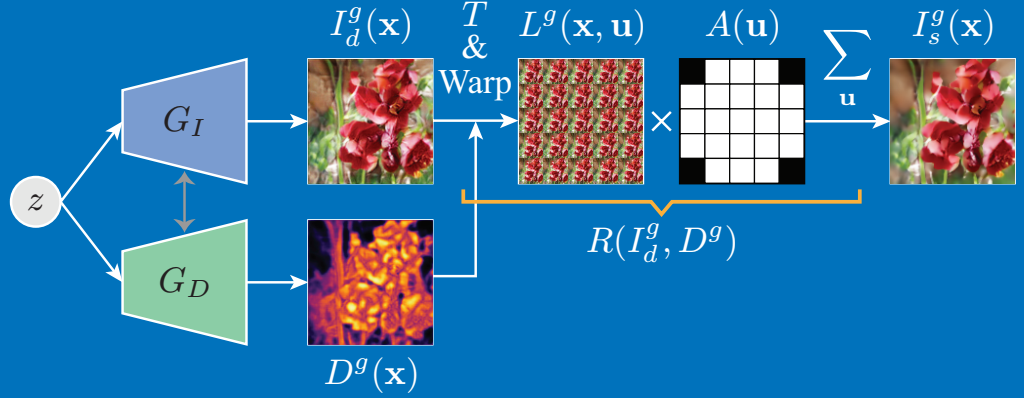
Deep or shallow DoF

iPhone photo and CycleGAN [Zhu+2017]

Project page

Unsupervised Learning of Depth and Depth-of-Field Effect from Natural Images with Aperture Rendering Generative Adversarial Networks

Takuhiro Kaneko



NTT Communication Science Laboratories, NTT corporation