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

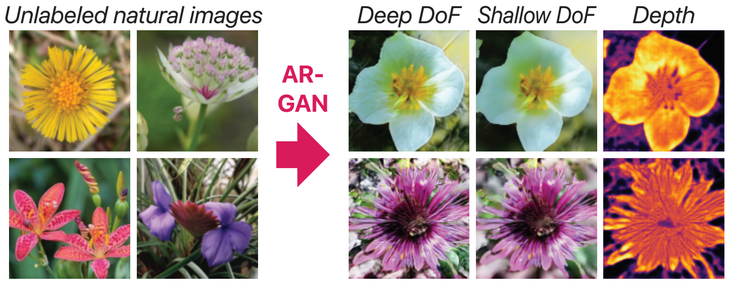
Takuhiro Kaneko
NTT Communication Science Laboratories, NTT Corporation

Project page

CVPR VIRTUAL JUNE 19-25

1 Introduction

Unsupervised learning of depth & DoF effect from natural images



Problem setting

- Only a collection of unlabeled natural images are available for training.
  - ***No*** ground-truth depth
  - ***No*** pairs data
  - ***No*** pretrained model

Objective

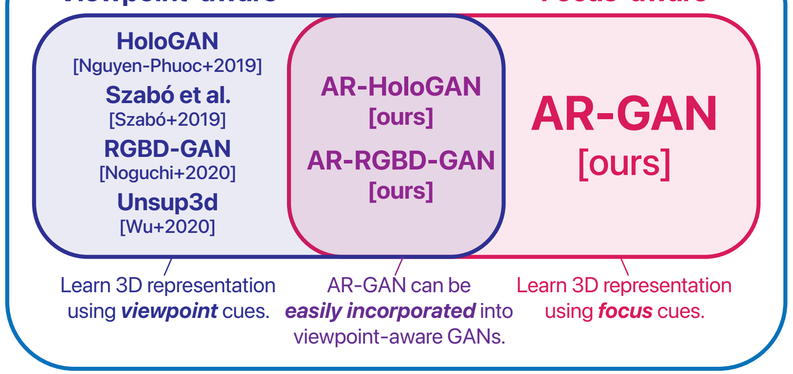
- Learn a generator that can synthesize tuples of
  - Deep DoF image
  - Shallow DoF image
  - Depth

  from random noise.

2 Positioning of research

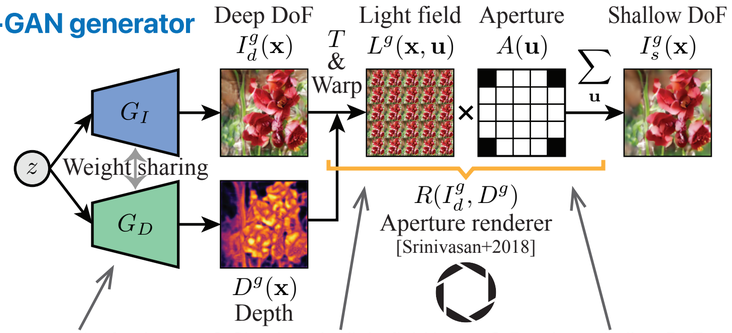
Fully unsupervised 3D representation learning

Only a collection of unlabeled natural images are available for training.



3 Overall pipeline

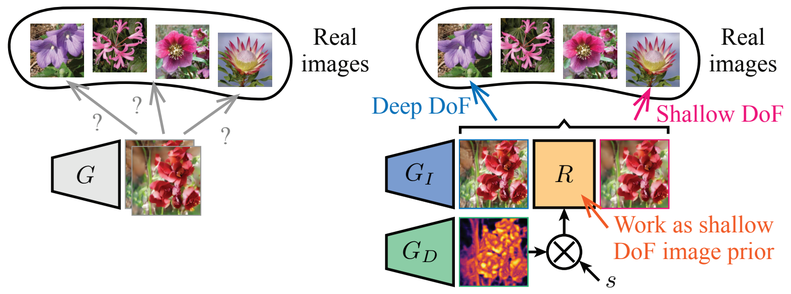
AR-GAN generator



1. Generate a pair of a deep DoF image and depth from a random noise.
2. Simulate the light field by warping the deep DoF image based on the depth.
3. Render a shallow DoF image by integrating the light field using the aperture.

4 DoF mixture learning

Learn real image distribution while generating diverse DoF images



Standard learning (baseline)

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

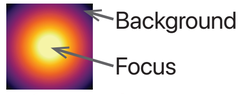
DoF mixture learning (proposed)

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

5 Center focus prior

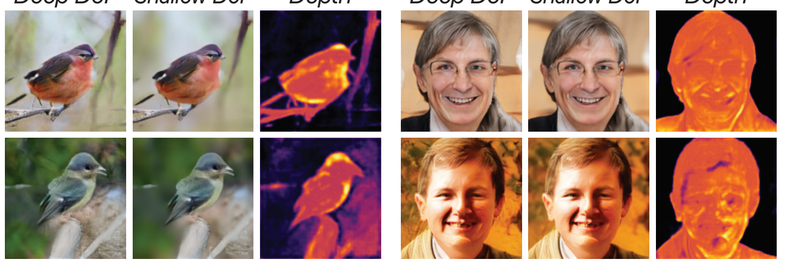
Solve ambiguity between fore/background blur by providing prior

Examples of center focused images



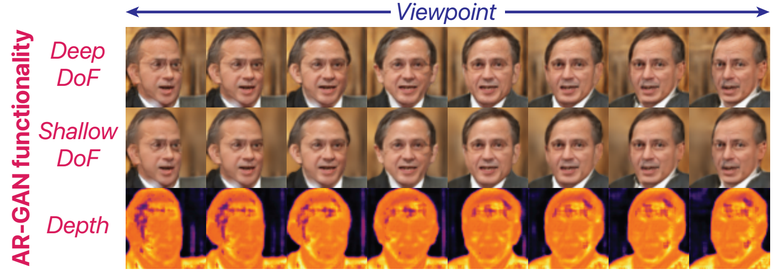
Center focus prior

6 Examples of generated data

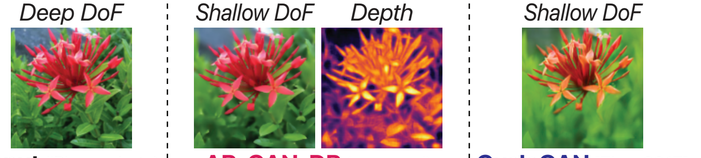


7 Portability analysis

AR-RGBD-GAN **RGBD-GAN** [Noguchi+2020] **functionality**



8 Application in shallow DoF rendering



Input (iPhone photo)

AR-GAN-DR (proposed)
W/ ***no*** supervision

CycleGAN [Zhu+2017]
W/ ***set-level*** supervision

9 Conclusion

- AR-GAN can learn depth & DoF effect from natural images in an unsupervised manner.
- AR-GAN has portability and can be easily incorporated into viewpoint-aware GANs.
- Generated data can be used for learning a shallow DoF renderer.
- We expect our findings will facilitate further studies on focus-aware image generation.