



Aoki  
Media  
Sensing  
Lab.

# 慶應義塾大学 青木研究室紹介

---

理工学部 電気情報工学科

教授

青木 義満

[aoki@elec.keio.ac.jp](mailto:aoki@elec.keio.ac.jp)

<https://aoki-medialab.jp/>



# 画像・音・言語・・・ マルチモーダルAI研究と実社会応用



# 次世代のAI実現へ向けて

-大規模基盤モデルを前提とした-

- 画像/動画/3Dの認識・生成
- マルチモーダルAI
- 学習効率化
- イベントカメラ・センシング
- 実世界センシング



# 音で姿勢を聴く？ Sound To Pose (CVPR2023)

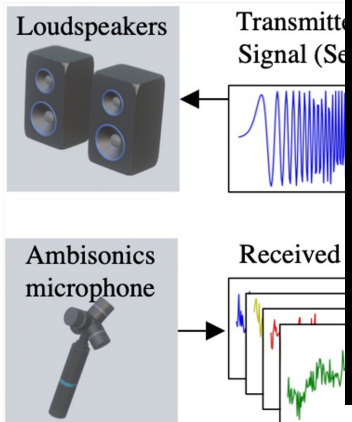
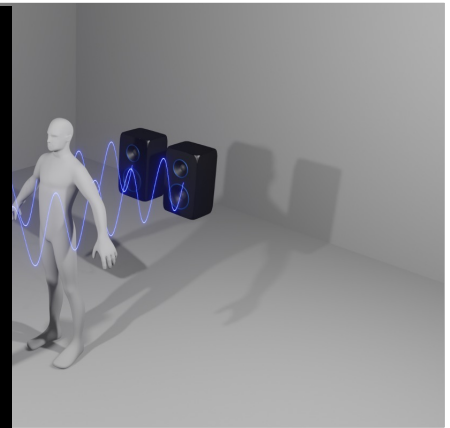
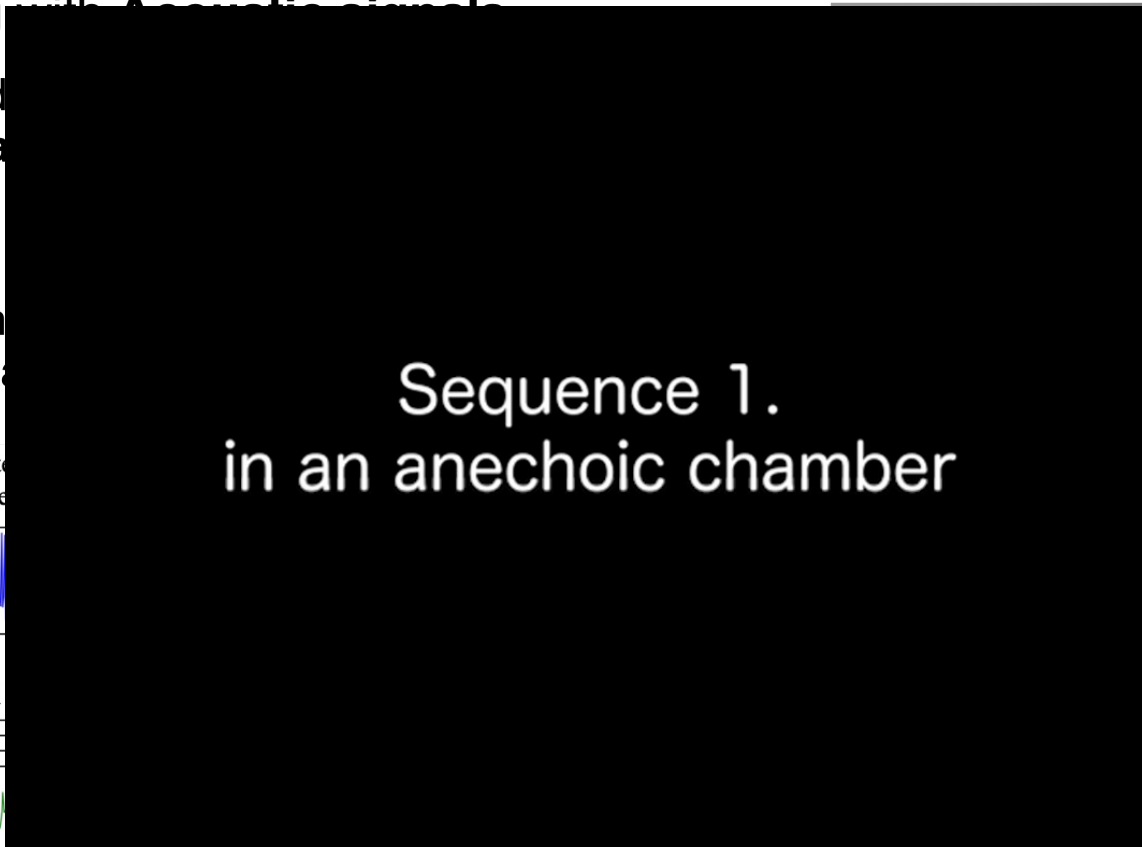
画像 + 音

Active remote sensing with Acoustic signals

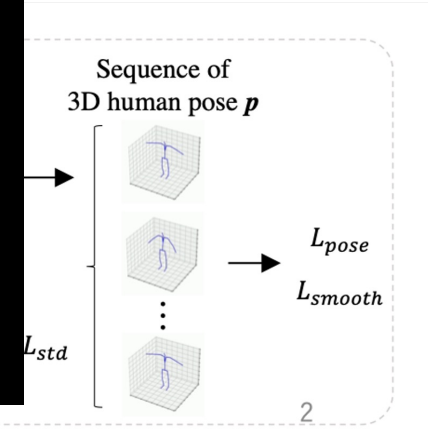
- ↳ Not using sound
- Prevents **Privacy**

Why sound ???

- ↳ We can use sound anywhere 📶 🏠 & 🏥

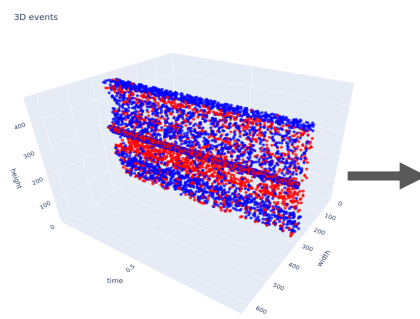
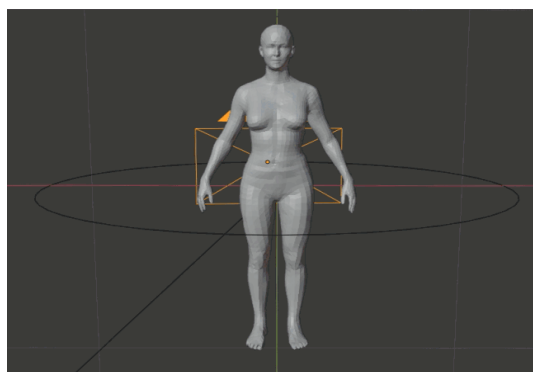


Sequence 1.  
in an anechoic chamber

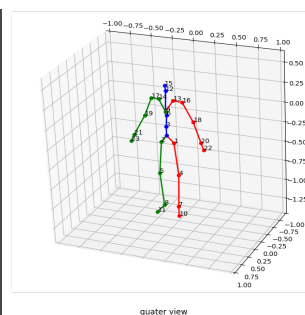
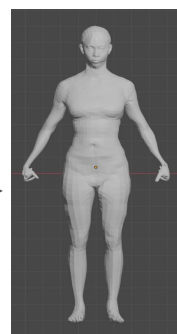


# イベントカメラを用いた高精度な人体スキャン [CV4MR] in CVPR2024

- 動くイベントカメラによる、イベントのみを用いた静的な人体メッシュ復元 + 姿勢推定手法を提案
- フレームベースの既存手法を上回る精度を達成
- 従来手法が苦手とする暗所やブラーを含むシーンに対する頑健性を示した



イベント



メッシュ+三次元姿勢

|                    | PEL-MPJPE ↓  |              | CD ↓         |              |
|--------------------|--------------|--------------|--------------|--------------|
|                    | Pose1        | Pose2        | Pose1        | Pose2        |
| Ours (Event-based) | <b>58.11</b> | <b>65.64</b> | <b>7.589</b> | <b>13.36</b> |
| EasyMocap [32]     | 59.54        | 97.29        | 36.18        | 30.09        |
| PyMAF [13]         | single       | 78.40        | 234.1        | 49.83        |
|                    | multi        | 63.90        | 185.4        | 30.19        |
| PyMAF-X [14]       | all          | 64.09        | 195.2        | 29.82        |
|                    | single       | 92.10        | 182.9        | 48.68        |
| PyMAF-X [14]       | multi        | 82.49        | 109.1        | 22.54        |
|                    | all          | 81.95        | 108.6        | 22.16        |



# スポーツ x コンピュータビジョン x AI

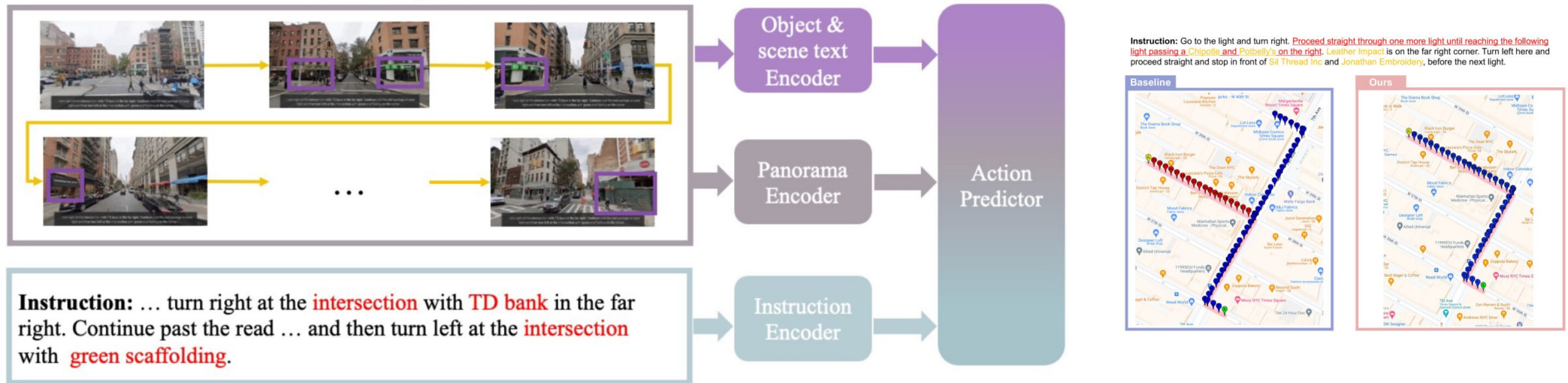


楽天GORAのブースでは ゴルファーのために  
特別に開発された新技術が先行公開されました



# Vision-and-Language Navigation (ICRA2024)

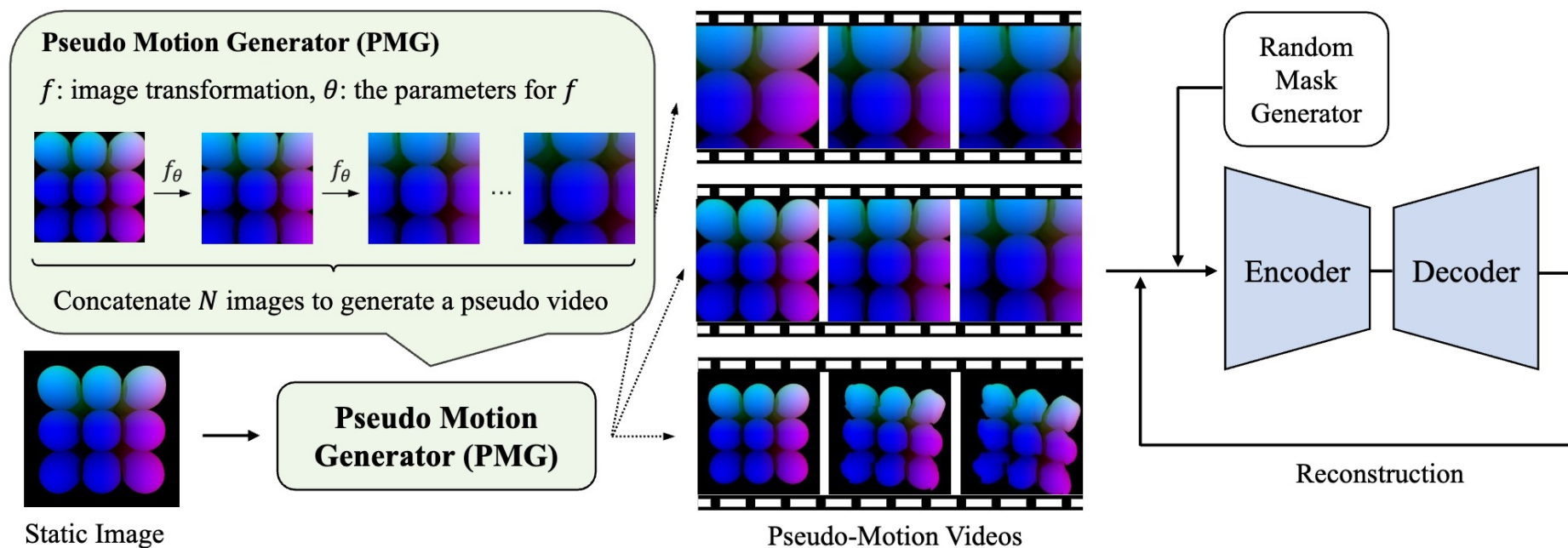
画像 + 自然言語



- エージェントが視覚情報と自然言語の指示をもとに未知の環境を移動し、特定のゴールに到達するタスク
- 視覚と言語を統合して行動を決定する能力が必要

ロボット自律移動, ナビゲーション支援,  
視覚と言語のマルチモーダルインターフェース

# 疑似動画で動画認識モデルを事前学習 (ECCV2024)

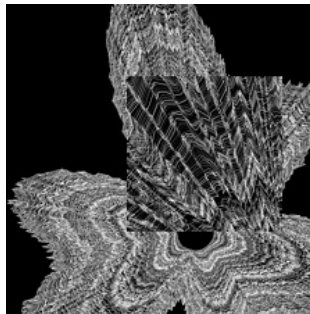
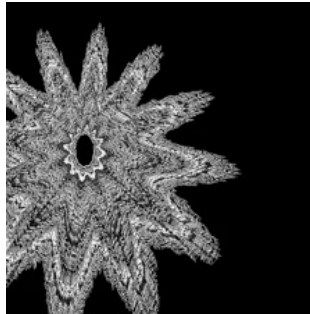
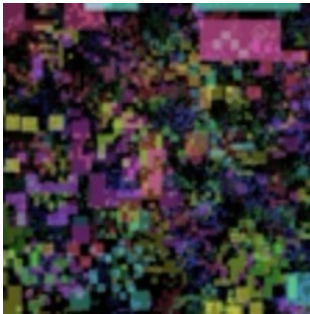
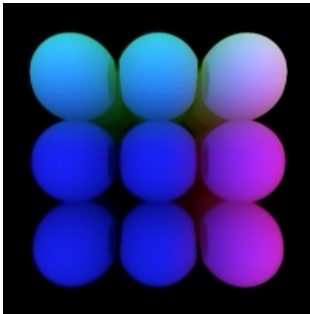
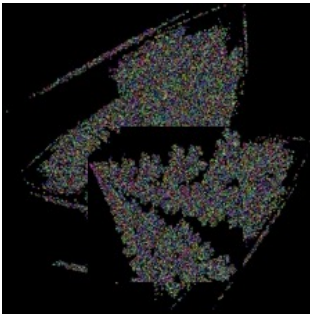
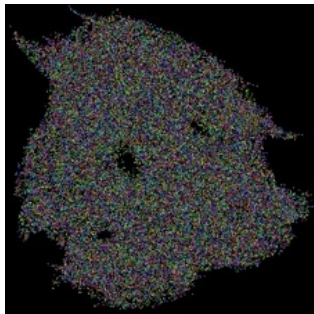
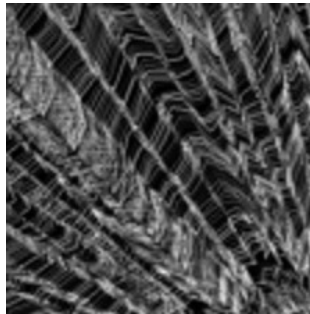
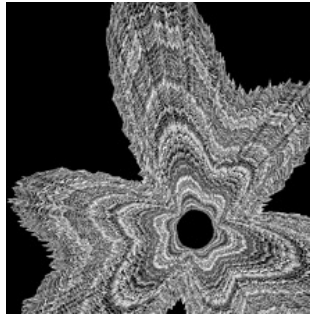
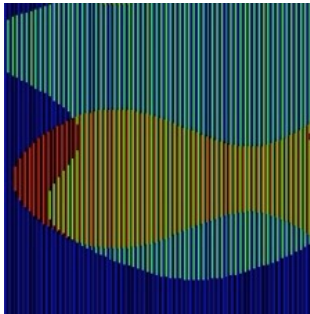
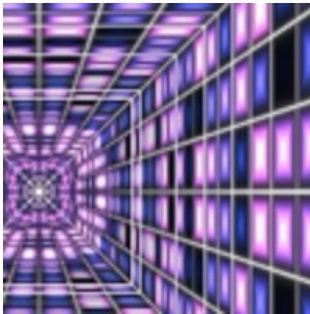
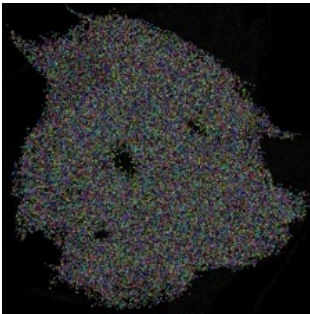
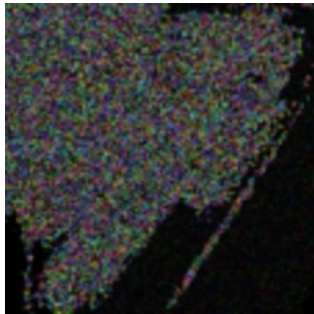


- ・ 人工動画を直接作るのではなく、より簡単な生成画像から、擬似的な動きを持つ動画を生成
- ・ 擬似的な動きを持つ動画を用いてVideoMAEを学習
- ・ VideoMAEは低次元な特徴を学習しているため、人物や物体が映っていない画像でも学習可能





# 生成される擬似動画



FractalDB

Shaders1k

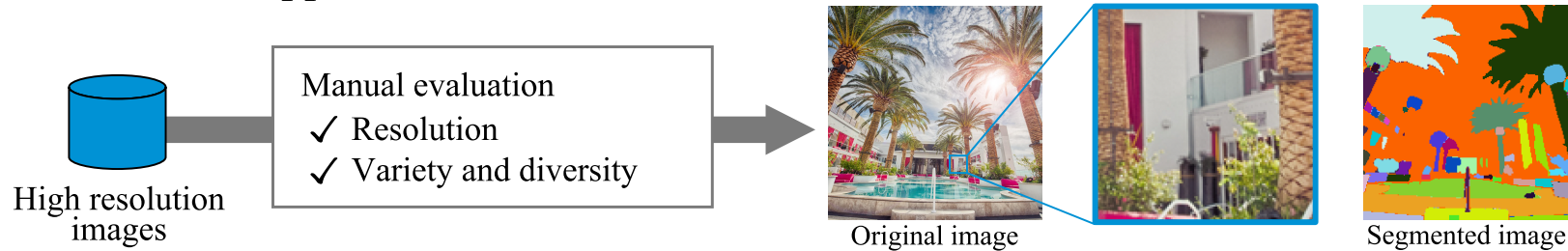
Visual Atom



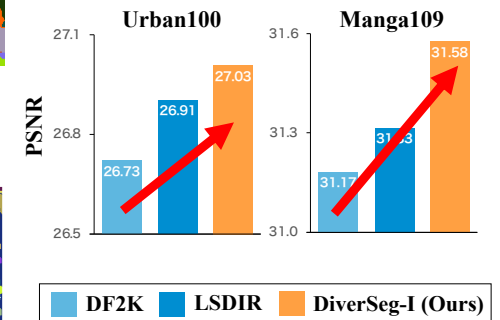
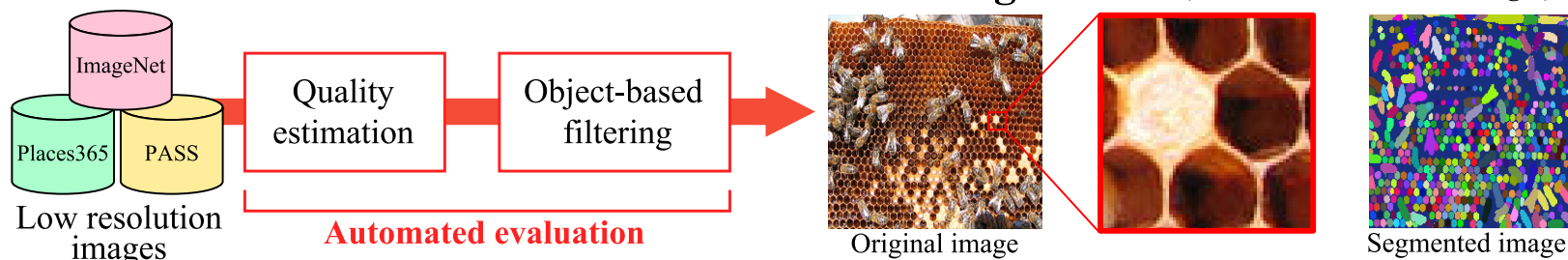
# Rethinking Image Super-Resolution from Training Data Perspectives (ECCV2024)

**品質**と**多様性**の2段階フィルタリングを行うことでWeb画像で構築された低解像度画像にも関わらず, 従来の高解像度データセットより性能向上

## Conventional approach



## Our approach



# 画像生成

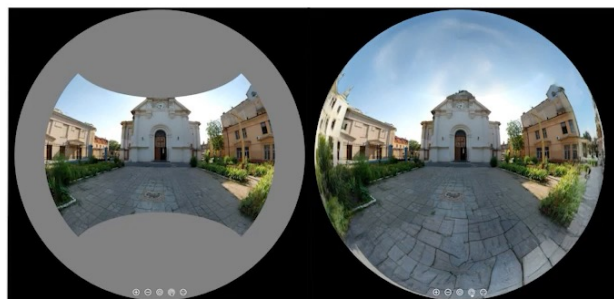
## Diverse Plausible 360-Degree Image Outpainting for Efficient 3DCG Background Creation

Naofumi Akimoto, Yuhi Matsuo, Yoshimitsu Aoki  
Keio University

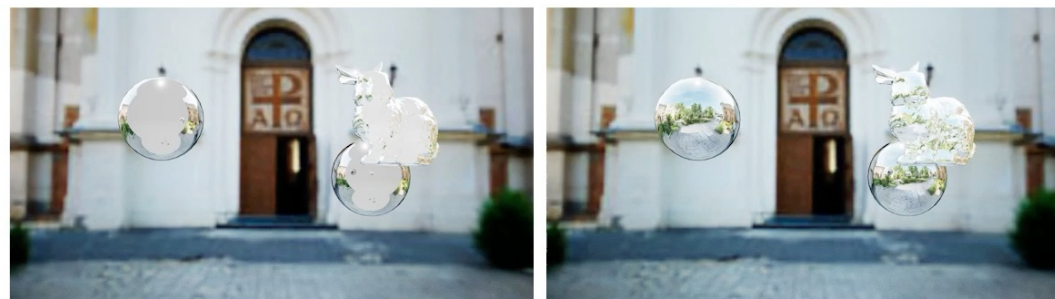
**Input Image**



**Diverse Outpainted Images**



**Spherical Visualization**



**Background Creation Demo**



# 青木研メンバー (2024年度)



**博士課程 : 12名**  
**修士課程 : 13名**  
**学部生 : 5名**  
**合計30名!!**



# 慶應AIセンター, Keio-CMU連携プロジェクト参画

## 一緒に慶應のAI研究を世界へ!

