

# 単一の全方位画像における空の動き生成

角方 元紀<sup>†</sup> 池林 ハキーム<sup>†</sup> 河合 紀彦<sup>†</sup>

大阪工業大学 情報科学部<sup>†</sup>

## 1. はじめに

現在 Google Street View<sup>[1]</sup>といったサービスがあるが、静止画像で臨場感が高くないという問題がある。この解決策として、定点での動画撮影を行うアプローチもあるが、広域な範囲で動画を撮影する必要がある場合には、撮影に膨大な時間がかかるという問題がある。そこで本研究では画像中の空領域に注目し、一枚の全方位画像から空の動きを再現した全方位動画を生成することで臨場感を向上させる手法を提案する。

動画生成を行う従来手法として、一枚の透視投影画像からニューラルネットワークを用いて空や川の動きを再現した動画を生成する Animating Landscape<sup>[2]</sup>が提案されている。しかし、この研究では透視投影画像を学習データとしているため、全方位画像に適用した場合にも、透視投影画像であることを前提とした動きが再現され、空方向を透視投影画像として見た場合には、違和感のある空の動きになる問題がある。また、パラメータに大きく依存し、通常動きがない領域でも動きが発生する問題がある。

これに対して、提案手法では、従来のニューラルネットワークを用いる手法と異なり、3次元空間で動きをモデル化して、それを2次元画像上での動きに変換する。また、セマンティックセグメンテーション<sup>[3]</sup>による領域分割を行い対象領域以外は動かないようにすることで、全方位画像上の空領域のみに動きを再現した全方位動画を生成する。

## 2. 提案手法

### 2.1 概要

提案手法では、一枚の空を含む風景画像を入力し、図1に示すような全方位画像にセマンティックセグメンテーションを用いて領域分割を行う。その後、地平線より上部の空領域でない領域に対して Inpainting<sup>[4]</sup>を行い、地平線より上部が全て空となる画像を生成する。次に、その画像の上半分の領域に対して動きを加え、空領域以外の部分を静止画の入力画像に置き換える



図1 入力画像例

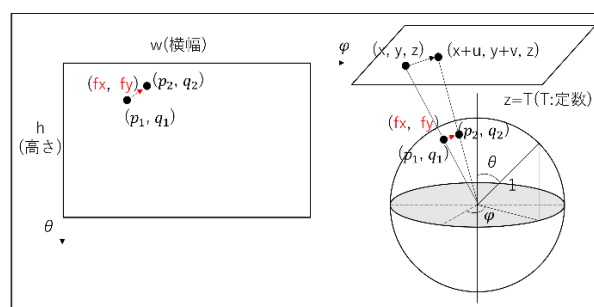


図2 球と平面の3次元位置の関係

ことで、空領域のみに動きがある動画を生成する。ただし、セマンティックセグメンテーションにおける境界領域の精度の不十分さを補償するため、空以外の領域の膨張を行い、さらに境界領域に対して Alpha blending を施すことで合成箇所の違和感を低減する。次節では、空の動き生成について詳述する。

### 2.2 空の動きの算出

空の動きについては、空の雲が三次元空間での上空での平面上を真っすぐ動くという仮定を用いて、その3次元の動きを全方位画像上での2次元オプティカルフローとして表現する。

具体的には、図2に示すように、全方位画像の画素 $(p_1, q_1)$ とそれに対応する上空の平面上の位置 $(x, y, z)$ の関係は、球の中心を通る直線と球と平面との交点から、以下の式により算出される。

$$x = T \tan \frac{\pi q_1}{h} \cos \frac{2\pi p_1}{w} \quad (1)$$

$$y = T \tan \frac{\pi q_1}{h} \sin \frac{2\pi p_1}{w} \quad (2)$$

$$z = T \quad (3)$$

ただし、変数の $w$ ,  $h$ は全方位画像の幅と高さ、 $T$ は定数である。

次に、求めた平面上の点の $x, y$ 座標を、図2に示すようにそれぞれある一定距離 $u, v$ 移動させ、それを球の表面に投影することで、上空の平面

の座標 $(x + u, y + v, z)$ に対応する全方位画像の画素 $(p_2, q_2)$ を以下のように算出する.

$$p_2 = \frac{w}{2\pi} \tan^{-1} \frac{y+v}{x+u} \quad (4)$$

$$q_2 = \frac{h}{\pi} \cos^{-1} \frac{z}{\sqrt{(x+u)^2 + (y+v)^2 + z^2}} \quad (5)$$

最後に変換された画素 $(p_2, q_2)$ と変換前の画素 $(p_1, q_1)$ の差を計算し、その画素におけるオプティカルフロー $(f_x, f_y)$ を算出する.

以上の処理を画像の全画素に対して行うことで全画素のオプティカルフローを算出し、これに基づき画素値をコピーすることで雲が移動した画像を生成する.  $u, v$ を線形に増やしながらかこれを繰り返すことで、空が動く動画像を生成する.

### 3. 実験

#### 3.1 概要

空の動きを生成する実験を行った. 実験では全方位カメラ (RICOH THETA Z1) を用いて撮影し、1600×800の解像度にリサイズした正距円筒図法で展開された全方位画像である図1の画像を使用した. なお本実験では、球の半径を1、定数Tを2、空を表す平面上での移動量は1フレーム当たりy方向に0.0015として、180フレームの動画像を生成した. また、Alpha blendingを行う境界領域の幅は20画素とした. また、これらの結果を、透視投影画像に対してニューラルネットワークで動きを生成する手法<sup>[2]</sup>を用いて生成した動画像と比較した.

#### 3.2 実験結果

生成した全方位動画の各フレームを特定方向の透視投影画像として表示したものを図3に示す. 実験結果から空領域が左の一定方向に動いていることが確認できる. また、上側にある雲が遠くにある雲よりも速く動いていることから遠近感が出ていることが分かる. また、空領域以外の領域が止まっており、空と建物などの境界領域に対しても違和感なく動きが付与されていることが確認できる.

次に従来手法との比較を行う. 図4はそれぞれの生成された動画の特定のフレームについて透視投影画像として同方向を表示したものになる. この結果から従来手法では建物の輪郭が崩れて雲の形が不自然に歪んでいるのに対して提案手法では輪郭が保たれ雲も違和感なく動いていることが分かる. また、従来手法では画像の両端の部分で動きが異なり不自然な歪みが形成されているのに対して提案手法では、境界部分に線

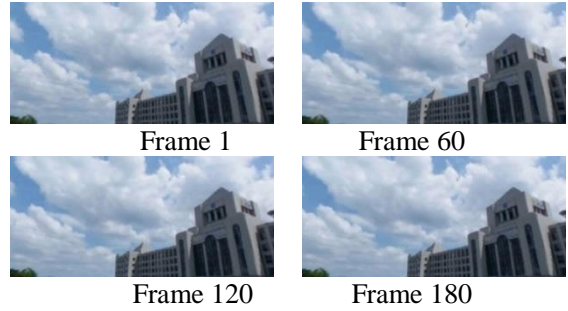


図3 y方向に動かした各フレームの結果

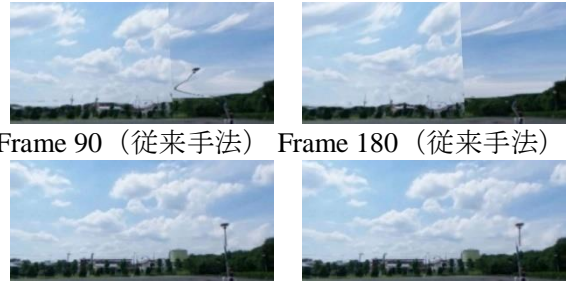


図4 両手法の特定のフレームの結果

や歪みが形成されることなく動画が生成されていることが確認できる. これらのことから、通常動きがない領域が動くという問題と空が違和感のある動きになるという問題が解消されていることが分かる.

### 4. まとめ

本研究では、単一の全方位画像から空の動きのある動画を生成する手法を提案した. 実験を通して提案手法の有効性を確認した. 現在は空領域にとどまっているが、今後は水や木などにも動きを加えることでさらに臨場感の高い動画の生成を行う.

### 謝辞

本研究の一部は、JSPS 科研費 JP21H03483 の助成を受けて実施した.

### 参考文献

- [1] Google Street View, <https://www.google.co.jp/maps> (参照9月30日)
- [2] Y. Endo, Y. Kanamori, S. Kuriyama, "Animating Landscape: Self-Supervised Learning of Decoupled Motion and Appearance for Single-Image Video Synthesis", ACM Transactions on Graphics, Vol.38, No.6, pp.175:1-175:19, 2019.
- [3] J. Lambert, Z. Lie, O. Sener, J. Hays, V. Koltun, "MSeg: A Composite Dataset for Multi-domain Semantic Segmentation", Proc. CVPR, 2020.
- [4] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, T. S. Huang, "Free-Form Image Inpainting With Gated Convolution", Proc. ICCV, pp. 4471-4480, 2019.