

## Alpha Matting と境界最適化を用いた人物領域の抽出精度向上

藤原 恭雅<sup>†</sup> 小坂 知輝<sup>†</sup> 村木 祐太<sup>†</sup>大阪工業大学 情報科学部 情報メディア学科<sup>†</sup>

## 1. はじめに

近年、画像内の人物を切り抜き、人物や背景の加工に利用する技術が注目されている。人物領域を正確に切り抜くためには、編集ソフトを用いて手動で作業する必要があり、手間や専門的な技術を要する。この問題を解決するために、深層学習を用いた切り抜きの自動化手法が提案されている。しかし、深層学習を用いた手法では正確な切り抜きが困難であり、人物領域の一部が欠如、あるいは非人物領域と干渉してしまう問題がある。

本研究では、2種類のセグメンテーション手法と色情報、Alpha Matting を用いて、より高精度に人物領域を抽出する手法を提案する。提案手法では、2種類のセグメンテーション結果から人物領域を抽出する。次に、色情報と Alpha Matting を用いて抽出結果を補正する。以上の処理により、従来のセグメンテーション手法よりも高精度な人物領域の抽出を実現する。

## 2. 提案手法

提案手法の処理手順を図 1 に示す。また、提案手法では、RGB 画像を入力画像とする。入力画像の例を図 2 に示す。

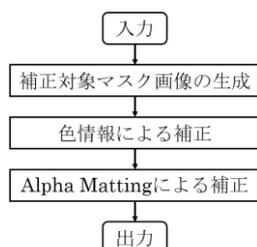


図 1 処理手順



図 2 入力画像

## 2.1 補正対象マスク画像の生成

補正対象マスク画像の生成では、セマンティックセグメンテーションとインスタンスセグメンテーションの実行結果から補正対象となるマスク画像を生成する。

入力画像に対してセマンティックセグメンテーションの一種である BEiT<sup>[1]</sup>とインスタンスセグメンテーションの一種である Mask2Former<sup>[2]</sup>を実行し、それぞれの実行結果から人物領域を抽出する。さらに、BEiT の抽出結果から人物領域内に存在する穴抜け領域を抽出する。そして、Mask2Former の人物領域と穴抜け領域の差分から、補正対象マスク画像を生成する。図 3 にセグメンテーション結果、図 4 に補正対象マスク画像を示す。

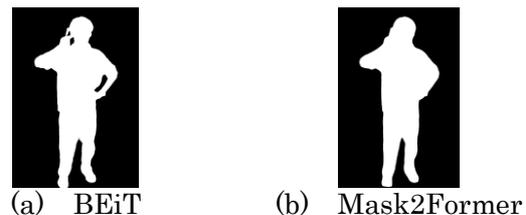


図 3 セグメンテーション結果



図 4 補正対象マスク画像

## 2.2 色情報による補正

色情報による補正では、同物体の隣接する画素同士が近い色情報を持つことを利用し、補正対象マスク画像の誤判定領域を補正する。

人物領域と非人物領域の境界にある人物領域の画素を取得し、色情報による補正の対象とする。そして、対象画素の近傍にある人物領域と非人物領域の輝度値の平均を算出する。その後、対象画素の輝度値と最も近い輝度値の平均を持つ領域を算出し、対象画素を算出した領域として再ラベリングする。ただし、算出した領域に近い平均値を持つ領域が存在する場合には再ラベリングしない。再ラベリング後、ラベルが変更された場合には、対象画素の 4 近傍画素かつ

異なるラベルの画素を補正対象に加える．色情報による補正結果を図5に示す．



図5 色情報による補正結果

### 2.3 Alpha Matting による補正

Alpha Matting による補正では，従来手法の Alpha Matting<sup>[3]</sup>を用いて色情報による補正では不十分であった箇所を補正する．

入力画像と色情報による補正結果からエッジを抽出し，2種類のエッジの差分から入力画像のエッジに沿っていない色情報による補正結果のエッジを補正対象として抽出する．次に，抽出したエッジを連結しているエッジごとに分解する．分解する際には，近傍に存在するエッジ同士も同一のエッジとして扱う．そして，色情報による補正結果と分解したエッジの1つから Alpha Matting の実行に必要な Trimap を生成し，Alpha Matting を実行する．実行後，実行結果のエッジと入力画像のエッジを比較し，実行結果のエッジが入力画像のエッジに沿う場合には実行結果を保持し，沿わない場合には，実行前の状態に戻す．最後に，実行結果と分解したエッジの1つを用いて再度 Trimap を生成し，同様の手順によって Alpha Matting を実行する．以上の処理を分解した全てのエッジを用いて実行し，最終的な Alpha Matting による補正結果を提案手法における出力結果とする．Alpha Matting による補正結果を図6に示す．



図6 Alpha Matting による補正結果

## 3. 実験と考察

### 3.1 実験

実験では，実験画像10枚に対して Mask2Former および提案手法を適用し，それぞれの実行結果と真値を比較した．そして，合致している画素と合致していない画素から合致率を算出し，提案手法の有用性を検証した．真値との比較結果を表1に示す．

表1 比較結果

実験画像	Mask2Former		提案手法	
	不一致 (px)	合致率 (%)	不一致 (px)	合致率 (%)
a	11300	99.48	6407	99.70
b	13733	99.36	7678	99.64
c	18042	99.16	9333	99.57
d	13255	99.39	10710	99.50
e	36861	98.29	30374	98.59
f	24619	98.86	14625	99.32
g	28081	98.70	6564	99.70
h	33754	98.44	11857	99.45
i	20276	99.06	7026	99.67
j	14531	99.33	7258	99.66
平均	21445	99.01	11183	99.48

### 3.2 考察

表1より，合致率は提案手法の方が全て上回り，不一致数が平均10262画素減少し，合致率は平均0.47%向上した．また，提案手法では，Mask2Former では取得できていない穴抜け領域が取得できており，真値に近い人物領域の抽出が可能となった．しかし，Mask2Former で取得できている人物領域を提案手法で取得できていない箇所が存在した．取得できていない原因は，Alpha Matting による補正の際に人物領域が非人物領域と推測されたためであると考えられる．改善方法として，補正対象であるエッジを抽出する際に，新たな条件を加える必要があると考えられる．

### 4. おわりに

本研究では，2種類のセグメンテーションモデルの実行結果に対して色情報と Alpha Matting を用いた補正を加えることで，人物領域の抽出精度向上を試みた．提案手法では，実験画像において従来のセグメンテーション手法よりも人物の形状に沿った抽出ができ，合致率が向上する結果となった．

今後の課題として，Alpha Matting による補正の対象となるエッジの抽出条件追加や色情報による補正が上手く機能しない箇所に対しての処理付与があげられる．

### 参考文献

- [1] Hangbo Bao, et al. “BEiT: BERT Pre-Training of Image Transformers”, Computer Vision and Pattern Recognition, (Sept 2021).
- [2] Bowen Cheng, et al. “Masked-attention Mask Transformer for Universal Image Segmentation”, Computer Vision and Pattern Recognition, (2022).
- [3] Thomas Germer, et al. “PyMatting: A Python Library for Alpha Matting”, Computer Vision and Pattern Recognition, (Mar 2020).