

静止画領域分割と動画像領域追跡を併用した アニメ動画像からのキャラクタ抽出

中嶋 亮太* 山藤 浩明 大倉 史生 松下 康之
大阪大学大学院情報科学研究科マルチメディア工学専攻松下研究室 OSAKA UNIVERSITY
* nakajima.ryota@ist.osaka-u.ac.jp

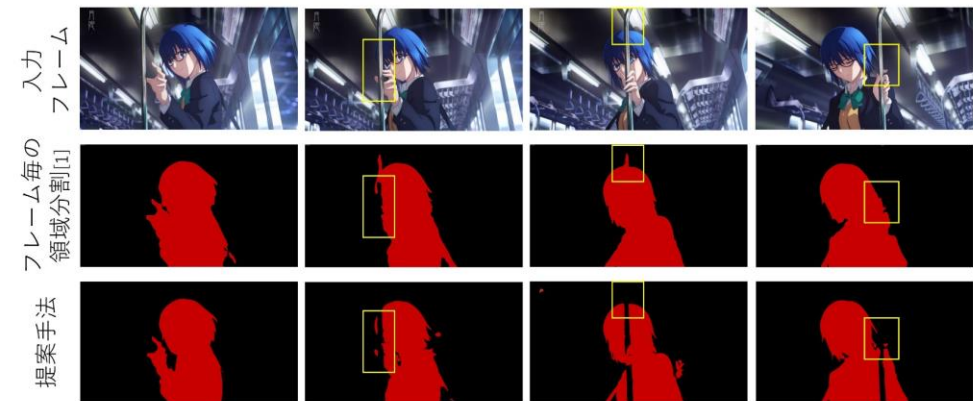
研究背景：アニメ動画像に適用できる領域分割手法がない

1. 既存手法の問題点
 - 動画像領域分割：実世界のデータのみによる学習済みモデル
 - 静止画キャラクタ抽出モデル：動画像での結果が不安定
 2. アニメ動画像領域分割への課題
 - アニメ動画像データセットがない→アニメ動画像に適したモデルの学習が困難
- 既存の学習済みモデルの併用によるキャラクタ抽出手法

提案手法：静止画のキャラクタ抽出を利用した領域追跡



実験結果：動画像領域追跡との併用によりキャラクター抽出が安定化



背景

提案手法

実験結果

補足

既存手法

- 動画像領域分割
 - ✓ 実世界のデータセットのみによるモデル
 - アニメ動画像には適用できない
- Yet-Another-Anime-Segmenter [1]
 - ✓ 静止画からのキャラクタ抽出モデル
 - ✓ 動画像ではシーンを通して結果が不安定

入力



出力



アニメ動画像領域分割への課題

- アニメ動画像データセットがない
- アニメ動画像に適したモデルを新たに学習することが困難

既存の学習済みモデルの併用によるキャラクタ抽出手法

[1] “Yet-Another-Anime-Segmenter,” <https://github.com/zymk9/Yet-Another-Anime-Segmenter>

[1] “Yet-Another-Anime-Segmenter,” <https://github.com/zymk9/Yet-Another-Anime-Segmenter>
[2] Ho Kei Cheng, et al., “XMMem: Long-Term Video Object Segmentation with an Atkinson-Shiffrin Memory Model,” ECCV2022.

静止画領域分割と動画像領域追跡を併用したアニメ動画像からのキャラクタ抽出

中嶋 亮太* 山藤 浩明 大倉 史生 松下 康之
大阪大学大学院情報科学研究科マルチメディア工学専攻松下研究室 OSAKA UNIVERSITY
* nakajima.ryota@ist.osaka-u.ac.jp

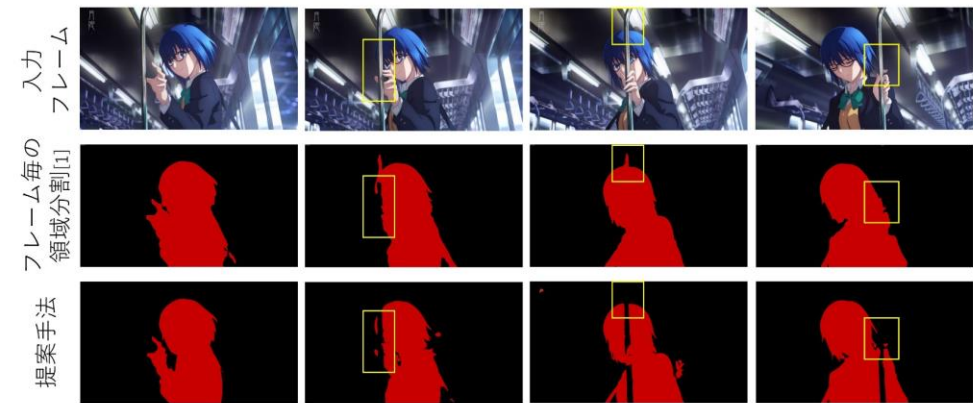
研究背景：アニメ動画像に適用できる領域分割手法がない

- 1. 既存手法の問題点
 - 動画像領域分割：実世界のデータのみによる学習済みモデル
 - 静止画キャラクタ抽出モデル：動画像での結果が不安定
 - 2. アニメ動画像領域分割への課題
 - アニメ動画像データセットがない→アニメ動画像に適したモデルの学習が困難
- 既存の学習済みモデルの併用によるキャラクタ抽出手法

提案手法：静止画のキャラクタ抽出を利用した領域追跡



実験結果：動画像領域追跡との併用によりキャラクター抽出が安定化



背景

提案手法

実験結果

補足

静止画領域分割と動画像領域追跡の併用

- 静止画領域分割 (Instance Segmentation)
 - ✓ Yet-Another-Anime-Segmenter [1]
 - ✓ 静止画からキャラクタ領域を獲得
- 動画像領域追跡 (Semi-supervised Video Object Segmentation)
 - ✓ XMem [2]
 - ✓ 入力：動画像フレーム， 1 番目のフレームのマスク
 - ✓ 出力：2 番目以降のフレームのマスク
 - ✓ 静止画領域分割の出力結果を入力マスクとして使用



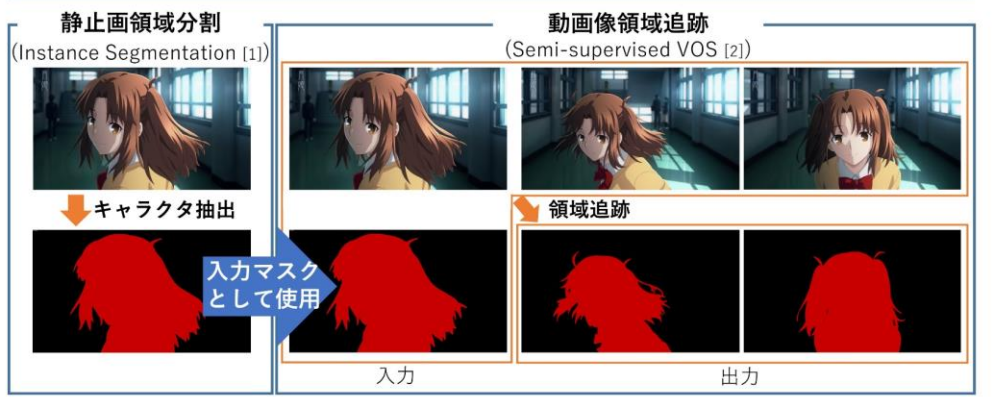
[1] “Yet-Another-Anime-Segmenter,” <https://github.com/zymk9/Yet-Another-Anime-Segmenter>

[2] Ho Kei Cheng, *et al.*, “XMem: Long-Term Video Object Segmentation with an Atkinson-Shiffrin Memory Model,” ECCV2022.

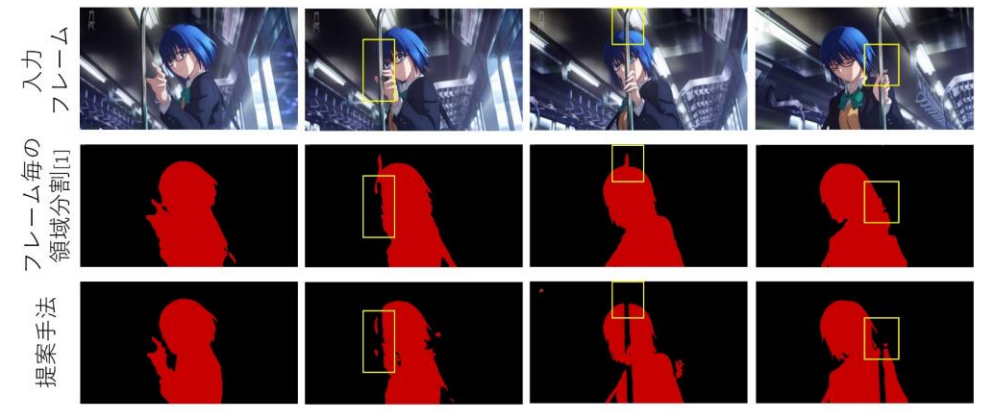
研究背景：アニメ動画像に適用できる領域分割手法がない

1. 既存手法の問題点
 - 動画像領域分割：実世界のデータのみによる学習済みモデル
 - 静止画キャラクタ抽出モデル：動画像での結果が不安定
2. アニメ動画像領域分割への課題
 - アニメ動画像データセットがない→アニメ動画像に適したモデルの学習が困難
- 既存の学習済みモデルの併用によるキャラクタ抽出手法

提案手法：静止画のキャラクタ抽出を利用した領域追跡



実験結果：動画像領域追跡との併用によりキャラクター抽出が安定化



背景	提案手法	実験結果	補足
----	------	------	----




入力マスクフレームの選択

- 静止画領域分割 [1] の出力結果は不安定
 - 動画像領域追跡 [2] の出力結果は入力マスクに依存
- 静止画領域分割 [1] の結果に応じた入力マスクの選択
- 静止画領域分割 [1] の信頼度スコアを利用




入力マスクに応じた動画の分割とフレーム順序の変更

- 入力マスクのフレーム前後で動画を分割
- 入力マスク前の部分は逆順にして推論

入力



[1]の出力



信頼度スコア

83%

79%

88%

入力マスクとして選択

[1] “Yet-Another-Anime-Segmenter,” <https://github.com/zymk9/Yet-Another-Anime-Segmenter>

[2] Ho Kei Cheng, *et al.*, “XMem: Long-Term Video Object Segmentation with an Atkinson-Shiffrin Memory Model,” ECCV2022.

[1] “Yet-Another-Anime-Segmenter,” <https://github.com/zymk9/Yet-Another-Anime-Segmenter>

[2] Ho Kei Cheng, *et al.*, “XMem: Long-Term Video Object Segmentation with an Atkinson-Shiffrin Memory Model,” ECCV2022.

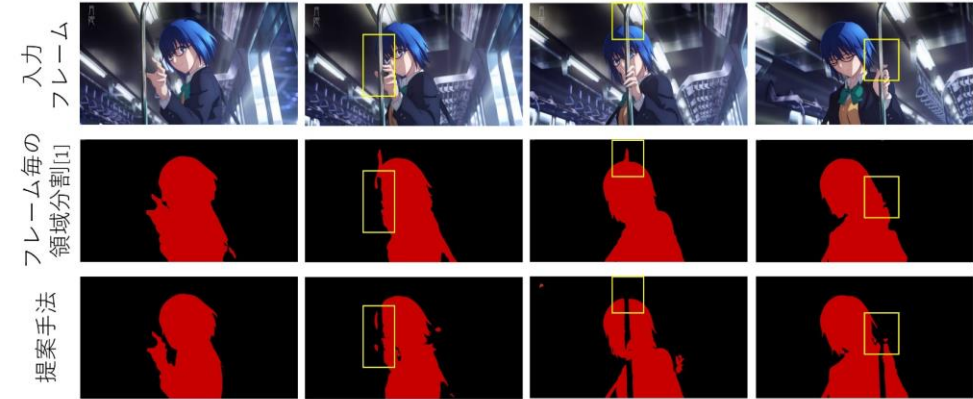
研究背景：アニメ動画像に適用できる領域分割手法がない

1. 既存手法の問題点
- 動画像領域分割：実世界のデータのみによる学習済みモデル
 - 静止画キャラクタ抽出モデル：動画像での結果が不安定
2. アニメ動画像領域分割への課題
- アニメ動画像データセットがない→アニメ動画像に適したモデルの学習が困難
- 既存の学習済みモデルの併用によるキャラクタ抽出手法

提案手法：静止画のキャラクタ抽出を利用した領域追跡



実験結果：動画像領域追跡との併用によりキャラクター抽出が安定化



[1] “Yet-Another-Anime-Segmenter,” <https://github.com/zymk9/Yet-Another-Anime-Segmenter>

[1] “Yet-Another-Anime-Segmenter,” <https://github.com/zymk9/Yet-Another-Anime-Segmenter>
[2] Ho Kei Cheng, et al., “XMem: Long-Term Video Object Segmentation with an Atkinson-Shiffrin Memory Model,” ECCV2022.

静止画領域分割と動画像領域追跡を併用したアニメ動画像からのキャラクタ抽出

中嶋 亮太* 山藤 浩明 大倉 史生 松下 康之
大阪大学大学院情報科学研究科マルチメディア工学専攻松下研究室 OSAKA UNIVERSITY
* nakajima.ryota@ist.osaka-u.ac.jp

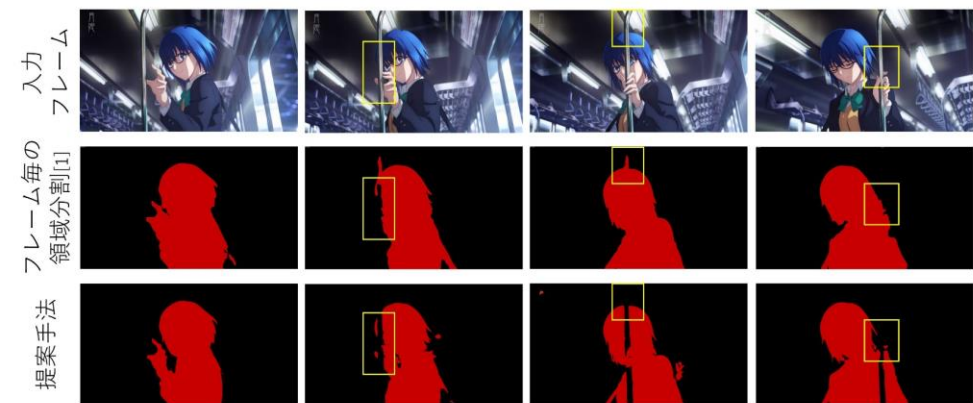
研究背景：アニメ動画像に適用できる領域分割手法がない

- 1. 既存手法の問題点
 - ・ 動画像領域分割：実世界のデータのみによる学習済みモデル
 - ・ 静止画キャラクタ抽出モデル：動画像での結果が不安定
 - 2. アニメ動画像領域分割への課題
 - ・ アニメ動画像データセットがない→アニメ動画像に適したモデルの学習が困難
- 既存の学習済みモデルの併用によるキャラクタ抽出手法

提案手法：静止画のキャラクタ抽出を利用した領域追跡



実験結果：動画像領域追跡との併用によりキャラクター抽出が安定化



[1] "Yet-Another-Anime-Segmenter," <https://github.com/zymk9/Yet-Another-Anime-Segmenter>
[2] Ho Kei Cheng, et al., "XMem: Long-Term Video Object Segmentation with an Atkinson-Shiffrin Memory Model," ECCV2022.

背景	提案手法	実験結果	補足
----	------	------	----

アニメ動画像セグメンテーションの応用例

- ・ データセット作成
 - ✓ キャラクタと背景が分離されたアニメ動画像データセットはない
- ・ キャラクタの行動推定
- ・ キャラクタのモーション補完

XMem [2] の選択理由

- ・ 実世界のデータセットのみによる学習済みモデル
- ・ アニメ動画像に対する結果も示されていた



[2] Ho Kei Cheng, et al., "XMem: Long-Term Video Object Segmentation with an Atkinson-Shiffrin Memory Model," ECCV2022.

静止画領域分割と動画像領域追跡を併用したアニメ動画像からのキャラクタ抽出

中嶋 亮太* 山藤 浩明 大倉 史生 松下 康之
大阪大学大学院情報科学研究科マルチメディア工学専攻松下研究室 OSAKA UNIVERSITY
* nakajima.ryota@ist.osaka-u.ac.jp

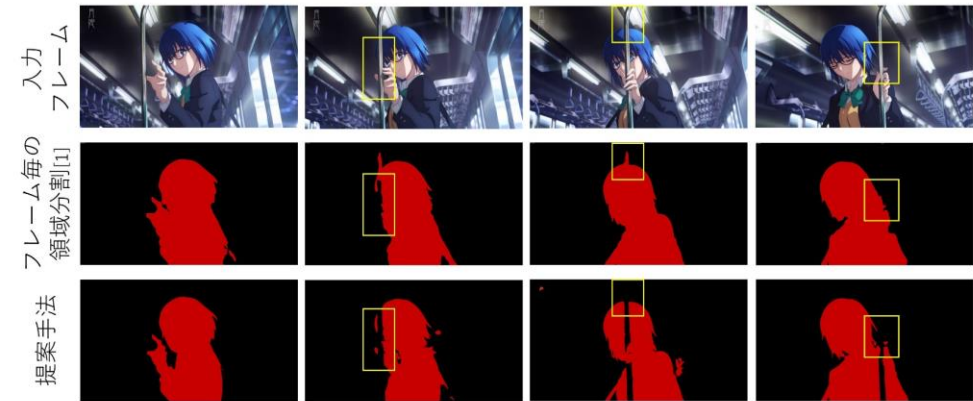
研究背景：アニメ動画像に適用できる領域分割手法がない

- 1. 既存手法の問題点
 - ・ 動画像領域分割：実世界のデータのみによる学習済みモデル
 - ・ 静止画キャラクタ抽出モデル：動画像での結果が不安定
 - 2. アニメ動画像領域分割への課題
 - ・ アニメ動画像データセットがない→アニメ動画像に適したモデルの学習が困難
- 既存の学習済みモデルの併用によるキャラクタ抽出手法

提案手法：静止画のキャラクタ抽出を利用した領域追跡



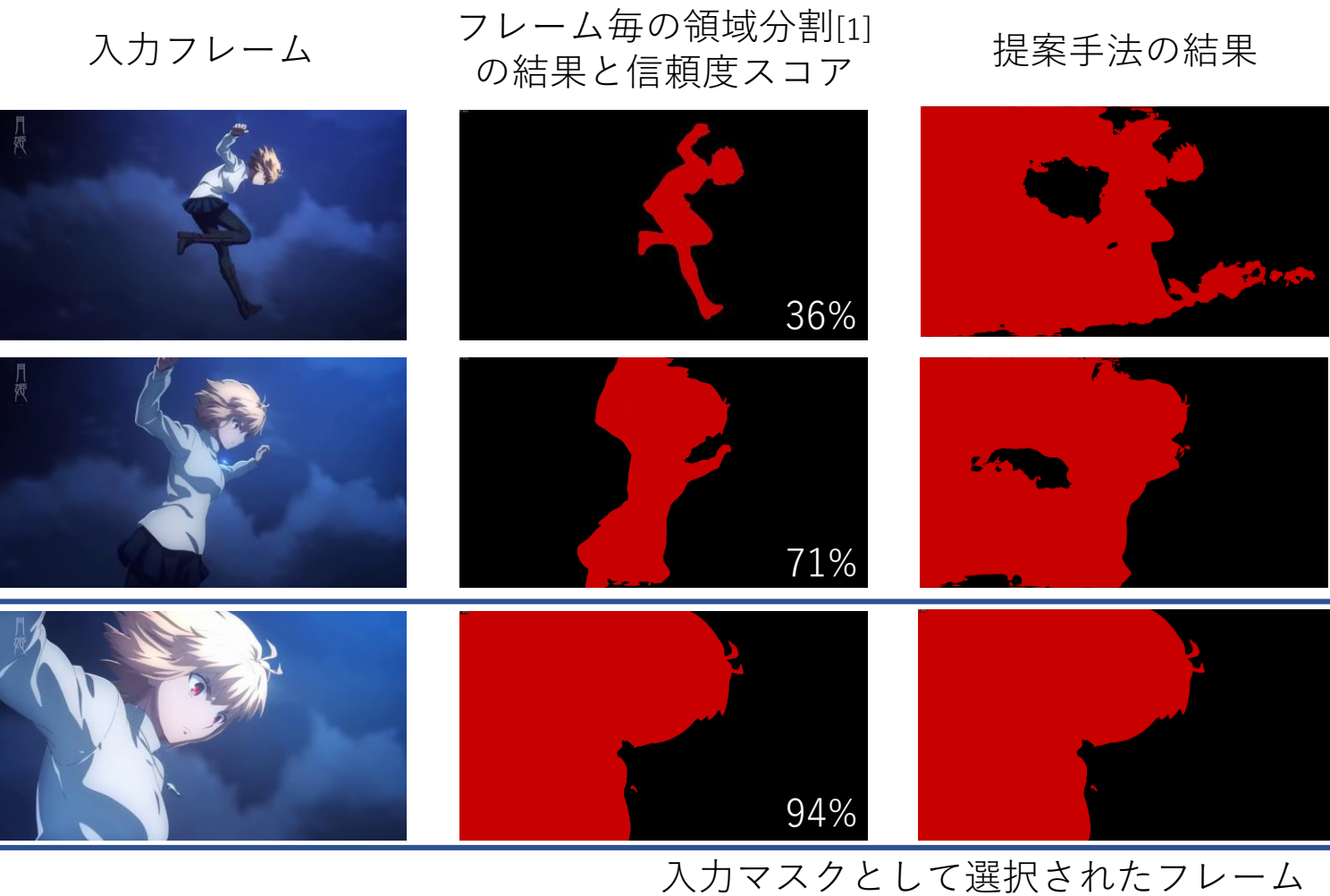
実験結果：動画像領域追跡との併用によりキャラクター抽出が安定化



背景	提案手法	実験結果	補足
----	------	------	----

信頼度スコアによる入力マスク選択の失敗例

- ・ 信頼度スコアが低いフレームの方が良いマスクが得られる場合
- 基準が信頼度スコアのみでは不十分



[1] “Yet-Another-Anime-Segmenter,” <https://github.com/zymk9/Yet-Another-Anime-Segmenter>

[1] “Yet-Another-Anime-Segmenter,” <https://github.com/zymk9/Yet-Another-Anime-Segmenter>
[2] Ho Kei Cheng, et al., “XMem: Long-Term Video Object Segmentation with an Atkinson-Shiffrin Memory Model,” ECCV2022.