

顕著性と運動量ベクトル解析に基づく 主被写体検出と縦型ビデオクロッピング

大田 海生[†] 村木 祐太[†]

大阪工業大学大学院 情報科学研究科 情報科学専攻[†]

1. はじめに

近年、スマートフォンや SNS の普及により、アスペクト比が 9:16 の縦型動画に対する需要が高まっている。そのため、横型動画を縦型にクロッピングする手法が検討されているが、クロッピング位置が固定であると移動する被写体を捉えることができない。また、顕著性が高い領域を含むようにクロッピング位置を変動させる SalCrop^[1]が提案されているが、被写体が複数の場合と同じ被写体を検出し続けることや、人物が被写体として検出されやすい問題がある。

そこで本研究では、顕著性と被写体の運動量ベクトル解析に基づいて主被写体を検出し、適切な位置でクロッピングする手法を提案する。これにより、より印象的な部分を残したクロッピングを可能にする。

2. 提案手法

2.1 概要

提案手法の処理手順を図 1 に示す。

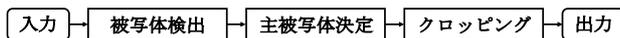


図 1 提案手法の処理手順

提案手法では、被写体を検出した後、各被写体の運動量ベクトルを基に主被写体を決定する。そして、主被写体の顕著性座標を基にクロッピングを施す。以上の処理を各フレームに対して施し、クロッピング動画を生成する。

2.2 被写体検出

被写体検出の処理手順を図 2 に示す。

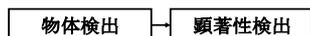


図 2 被写体検出の処理手順

2.2.1 物体検出

入力動画内の物体を検出するため、物体検出を施す。物体検出結果を図 3 に示す。

2.2.2 顕著性検出

入力動画内の顕著な領域を検出するため、顕著性検出を施す。顕著性検出結果を図 4 に示す。



図 3 物体検出結果 図 4 顕著性検出結果

物体検出で得られたバウンディングボックス (以下、BB) 内に顕著な領域が存在した場合は、その物体を被写体とする。

2.3 主被写体決定

検出された被写体から、主被写体を決定する。被写体の BB 内の運動量ベクトルの平均が最大の被写体を主被写体とする。

2.4 クロッピング

クロッピングの処理手順を図 5 に示す。

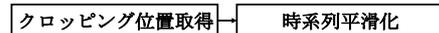


図 5 クロッピングの処理手順

2.4.1 クロッピング位置取得

主被写体の動きに合わせて、クロッピング位置を取得する。クロッピング位置は主被写体の BB 内に存在する顕著性の中心 X 座標とする。

2.4.2 時系列平滑化

クロッピング位置の動きを動画全体に渡って滑らかにするため、時系列の平滑化を施す。

Main Subject Detection and Vertical Video Cropping Based on Saliency and Momentum Vector Analysis
Kaiki Ota[†] Yuta Muraki[†]
Osaka Institute of Technology Graduate School of Information Science and Technology[†]

3. 実験と考察

3.1 実験

実験では、提案手法の優位性を検証するため、男女 30 名に対してアンケートを実施した。アンケートは 3 つの手法で生成した 18 種類のクロッピング動画が適切であるかを、5 を最も良い評価とする 5 段階で評価する。手法は提案手法、従来手法である SalCrop と Adobe Auto Reframe^[2]の 3 つである。アンケート結果を図 6、表 1 に示す。

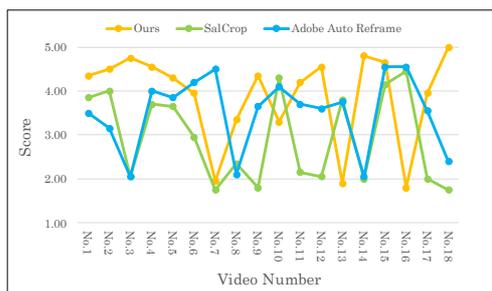


図 6 アンケート結果

表 1 アンケート結果

	Proposed Method	SalCrop	Adobe Auto Reframe
Average	3.90	2.93	3.51

3.2 考察

図 6 より、従来手法よりもクロッピングが適切であることが確認できた。No.3の入力を図 7、各手法の出力結果を図 8~10 に示す。各図(a)に 10 フレーム、各図(b)に 30 フレーム、各図(c)に 50 フレームを示す。



図 7 No.3 の入力

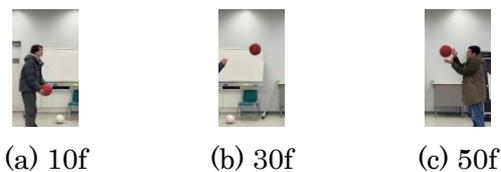


図 8 提案手法

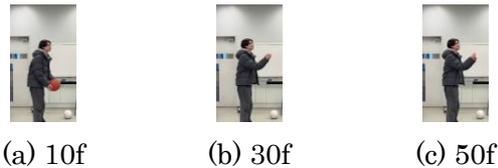


図 9 SalCrop

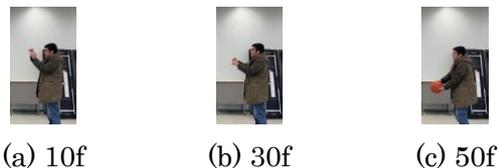


図 10 Adobe Auto Reframe

図 7 は主被写体が変化する動画であり、提案手法では主被写体の切り替えがスムーズに実現できている。

一方、カメラを大きく動かしながら撮影している場合や、ピントが合っていない被写体が存在する場合は、従来手法と比べて精度が劣る結果となった。解決策として、カメラの動きを検出することや、主被写体を決定する際に鮮明度を考慮することが挙げられる。

4. おわりに

本研究では、横型動画から縦型にクロッピングする手法を提案した。提案手法では、顕著性や運動量ベクトル解析に基づいて主被写体を決定し、クロッピング位置を取得した。

今後の課題として、カメラを大きく動かしながら撮影している場合や、ピントが合っていない被写体が存在する場合の対応が考えられる。

<参考文献>

- [1] Kao Zhang, Yan Shang, Songnan Li, Shan Liu, Zhenzhong Chen, “SalCrop: Spatio-temporal Saliency Based Video Cropping”, 2022 IEEE International Conference on Visual Communications and Image Processing (VCIP), 2022.
- [2] Adobe, “Automatically reframe video for social media channels”, <<https://helpx.adobe.com/premiere-pro/using/auto-reframe.html>>, (最終閲覧日 2023/12/23).