

前景・背景分離に基づくカメラワーク特徴量の提案

加藤卓哉^{†1} 深山覚^{†2} 後藤真孝^{†2} 森島繁生^{†3}

概要: 動画コンテンツの視聴が普及により、動画コンテンツ検索の為の特徴量への注目が高まっている。中でもカメラワークは、動画を撮影する技術の特徴づける重要な動画特徴であり、その分析技術の研究は多くなされている。既存研究では、カメラワークは背景のみに依存すると定義し、カメラワーク特徴量が提案されている。しかしこうした手法では、トラッキングショットなどの前景の特徴に依存するカメラワークの違いを表現できず、分類できる動画の種類が限られている。

本研究では、動画中の前景と背景を分離し、それぞれを特徴量化することで、カメラワークを分類する手法を提案する。提案法では、背景をカメラの動きと定義し、前景をカメラと演者の位置関係や演者の行動と定義する新たな特徴量化手法を提案する。本手法を用いることで、従来手法では判別が困難であった様々なカメラワークの違いの分類に成功した。

キーワード: カメラワーク特徴量, 動画像検索, 前景・背景分離, 動画像解析

1. はじめに

近年、動画共有サイトやカメラ技術の発展に伴い、動画の撮影技法の分析が注目されている。例えば、映画などの映像作品のカメラワークを分析することで、カメラワークの似た動画を検索できたり、監督やカメラマンの育成に役立てたりできる。特に、撮影技術の特徴量として表現することで、機械学習などを用いた様々なアプリケーションが開発できる。例えば、人間によって撮影された動画を入力として、ドローンなどのロボットの動きを制御することができる。特に、映画や映像作品を入力とすることで、人間には撮影困難な環境においても自動でプロのカメラマンが撮影したような動画を撮影できる。さらに、3DCGアニメーションの作成においても、現実世界でのカメラワークと似た印象のカメラワークを再現可能とすることでより効率的に映像作成が可能となる。こうした背景のもと、撮影技法の分析のための様々な特徴量が考案されてきた。

一般に撮影技法は「カメラの動かし方」と「構図」の組み合わせによって定義される。[1] 例として、被写体(例: 人間・車など)を一定の位置に配置するように撮影するトラッキングショットと呼ばれる技法について考える。トラッキングショットにおいては、被写体の有無に加え、カメラがどのように動いているか、さらに被写体の数と位置によってシングルショットやツーショット、アップ、ワイドなどと細かく分類できる。こうした「カメラの動かし方」と「構図」の違いは、どのように被写体撮影するか特徴付けられるうえで重要であり、それぞれ独立に分類できる。そのため、動画の分析に両方の特徴の評価は欠かすことはできない。しかし、従来提案されてきた特徴量は、被写体の動作推定を行ったシーンの分類を目的としたものが多く、カメラが一定の動きをしていることが仮定されており、カメ

ラの動きが時間的に遷移する場合は評価することができなかった。さらに構図の違いを評価できる研究はほとんどない。このように、従来のカメラワーク特徴量は、プロの映像監督やカメラマンが用いるような撮影技法の分類できなかった。

そこで本研究では、こうした「カメラの動かし方」や「構図」の違いに分類するため、被写体の大きさや位置をカメラの動かし方と同時に定量化する新たな特徴量を提案する。被写体の位置の推定には、顕著性マップを用いた前景・背景分離手法を用いる。人間の視覚的特徴を考慮した顕著性マップを用

†1 早稲田大学

Waseda University

†2 産業総合技術研究所

National Institute of Advanced Industrial Science and Technology

†3 早稲田理工学術院総合研究所

Waseda Research Institute for Science and Engineering