

放送映像からの地名別風景ショットの抽出

Wang Jiani*, 出口 大輔(名大), 高橋 友和(岐阜聖徳大), 井手 一郎(名大/NII), 村瀬 洋(名大)

Extraction of scenery shots classified by place-name from broadcast video

Wang Jiani, D. Deguchi(Nagoya Univ.), T. Takahashi(Gifu Shotoku Univ.), I. Ide(Nagoya Univ./NII), H. Murase(Nagoya Univ.)

1. はじめに

近年、ウェブ上にある検索ツールである Google Maps などにおいて、風景が写った画像や映像を利用するサービスが盛んである。これらのサービスのさらなる利便性向上のために、風景ショットの特徴を分析し、映像中から抽出する技術に取り組んでいる。

本発表では、実際に放送された大量の映像の中に現れる風景ショットを抽出し、地名情報を索引付ける手法を提案する。放送映像から風景ショットを抽出することで、抽出された風景ショットを映像コンテンツ産業における環境映像や観光映像の素材としてユーザに提供することを目指す。例えば、旅行前に観光映像を見ることにより、目的地を選んだり、計画を立てたりする手助けになると考えられる。

2. 地名別風景ショットの抽出手法

本発表では、風景映像が多く含まれると考えられる特定の放送番組を対象とし、Fig. 1 に示すように、テキスト処理と画像処理の二つを組み合わせる処理を行う。

2.1 テキスト処理による地名抽出

まずテレビ番組の電子番組表 (Electronic Program Guide, 以降、EPG) から、地名を含む番組を抽出する。地名を抽出するには、番組表のテキスト情報を形態素解析する。形態素解析には、日本語形態素解析システム mecab を用いた。ただし、品詞が「名詞」で、かつ品詞細分類 2 が「地域」であるものを地名であるとする。次に抽出した番組に付随するクローズド・キャプション (Closed Caption, 以降、CC) から地名を含む箇所を検出し、地名別にまとめる。本発表ではこれらを「地名別 CC 群」と呼ぶ。

2.2 画像処理による風景ショットの抽出

前処理としてショット分割を行い、CCに付随する時刻情報をもとに地名別CC群に対応するショットを抽出する。次に、風景映像は緩やかなカメラワークを含むという特徴があることに注目し、オブティカルフローを用いたカメラワーク分類を行う。具体的には、カメラワークをfix, pan, zoom, motion に分類し、カメラワークがfix, pan, zoom のものを風景候補ショットとする。オブティカルフローの計算にはブロックマッチング法[1]を用いた。最後に、顔検出により人物が含まれるショットを除去することで、最終的な風景ショットを抽出する。

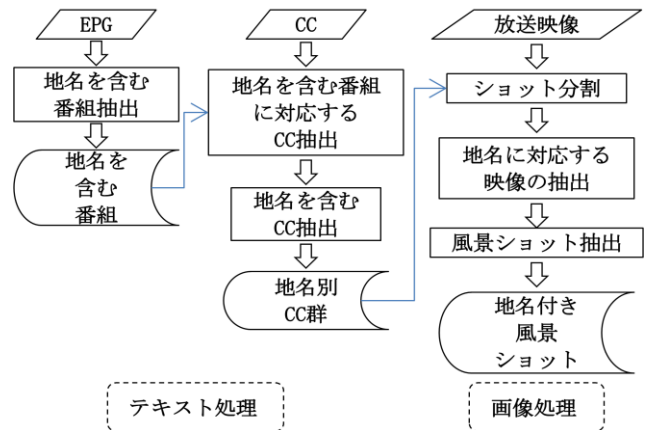


Fig. 1 Flow for the extraction of scenery shots classified by place-name.

3. 実験および考察

実験に用いた映像は2008年4月24日から2008年4月30日までの7日間に、東京地区において放送されたNHK教育テレビとテレビ朝日の映像である。これらの映像に提案手法を適用したところ、風景ショットは56件抽出された。Fig. 2 に抽出されたショットの例を示す。目視により確認したところ、その中で正しく抽出できたものは33件であった。なお、実際の風景ショットは44件存在していたため、提案手法の再現率は75%、適合率は59%であった。誤検出は24件で、未検出は12件であり、これらの主な原因は顔検出の失敗とカメラワークの誤分類であった。

4. むすび

本発表では、放送映像から地名別風景ショットを抽出する手法を提案した。実験により、放送映像から地名別風景ショットの抽出が可能であることを確認した。今後の課題として、カメラワークの分類精度の向上、人物ショットを除去する際の顔検出の精度向上が挙げられる。

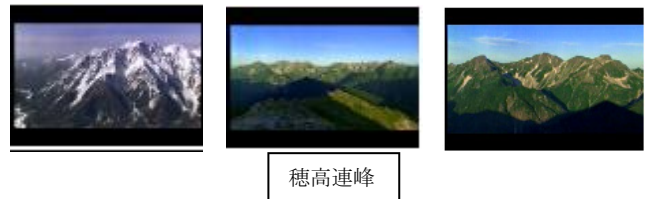


Fig. 2 Results of extracted scenery shots.

文 献

[1] 田村秀行: コンピュータ画像処理, オーム社, 2002.