

平成 25 年度 修士論文概要

主査	舟橋 健司	副査	本谷 秀堅, 山本 大介	研究室	舟橋研究室
入学年度	平成 24 年度	学籍番号	24417563	氏名	高柳 亮太

論文題目 撮影時におけるモバイル映像アノテーションとそれに基づく映像編集への応用
Mobile Video Annotation at Photographing and its Application for Video Editing

1 はじめに

近年、動画投稿サイトの発展・普及により一般ユーザが作成した映像コンテンツの数は爆発的に増加している。そのような背景から映像コンテンツの検索・要約に関する研究が数多く行われている。代表的な手法として、映像解析手法とアノテーション手法 [1] がある。映像解析手法は自動的に映像を検索・要約できるのでユーザの負担は小さいが、ユーザの嗜好を反映しにくい点や映像の質が影響するなどの課題がある。一方、メタデータを映像に付与することで多様な検索・要約が可能なアノテーション手法は、映像の質に依存しないという点で一般ユーザの映像を対象とした検索・要約に適していると考えられるが、アノテーションコストが課題である。このようにコストや信頼性などの点において一長一短ではあるが、共通して言えることはすでに存在している映像データを対象としている点である。そこで本研究では、アノテーションの利点を生かしつつ、アノテーションコストを最小化するため撮影中のユーザの自然な行動からメタデータを抽出することを検討し、撮影位置・方向を視線メタデータとして活用する。さらに、各撮影者の視線メタデータをオンライン上で共有し、撮影者の視線が集中した注目地点の提示や素材映像の共有といったメタデータ活用法を提案し、映像コンテンツ作成を支援する。

2 関連研究

本研究の先行研究 [2] では、動画編集作業において素材動画から欲しいシーンを抜き出すトリミングという工程に着目し、トリミング作業の効率化を目的に動画撮影中にメタデータを付与する手法を提案した。具体的には、撮影者が撮影中に印象的だと感じたときにボタンを押すことで映像シーンに対してメタデータを付与する。メタデータ付与によるトリミング作業の効率化に関しては一定の評価を得られたが、撮影時のボタン押下によるメタデータ付与は主観的性質が強くメタデータの信頼性の点や、撮影時間に比例してアノテーションコストが増加するという点に課題があった。

3 提案システム概要

本研究の目的は、アノテーションコストを抑えつつ、メタデータを利用して映像コンテンツ作成を支援することである。特に映像コンテンツ作成に必要な素材集めに焦点を当て撮影を支援するシステムを提案する。提案システムは撮影システムとメタデータ解析サーバの2つのサブシステムから構成される。撮影システムは携帯情報端末上で実装し、デバイス内蔵のGPSと電子コンパスから撮影者の撮影位置・方向を時刻とともに視線メタデータとして取得する。メタデータ解析サーバでは、オンライン上で不特定多数の撮影者から収集した視線メタデータから他の撮影者の位置や視線が集中している地点を導出・提示することで、撮影を支援する。つまり、人間の自然な行動から撮影を支援するためのメタデータを取得し、さらにオンライン上でメタデータを共有することで撮影者一人当たりのアノテーションコストを最小化するとともに、メタデータを利用して撮影者に有用な情報を提示することを目指す。本システムの構成図を図1に示す。

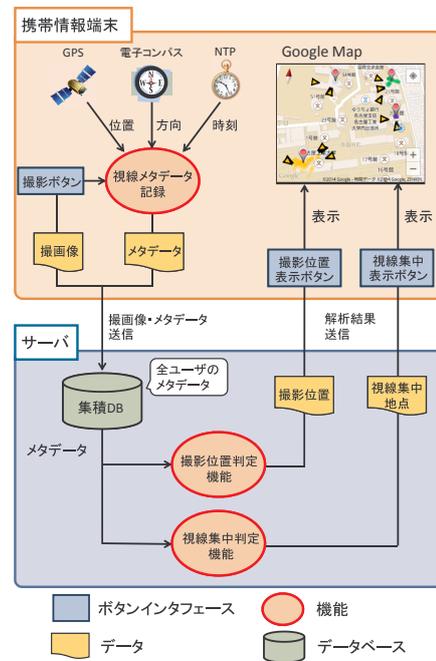


図 1: システム構成図

4 視線集中判定モデル

本章では、視線メタデータから不特定多数の撮影者の視線が集中した地点を導出する視線集中判定モデルについて説明する。視線メタデータは撮影システムを使用して静止画を撮影したときの撮影位置・方向を時刻とともに記録する。さらに、撮影位置の緯度経度を平面直角座標系の xy 座標に変換し、さらに撮影位置の xy 座標と撮影方向の値から向きと大きさをもつ視線ベクトルを導出する。撮画像とともに視線メタデータをメタデータ解析サーバへ送信する。不特定多数の撮影者から収集した視線メタデータに記録されている視線ベクトルの交差判定を行い、撮影者の視線の交点を求める。次に、導出された交点座標のデータセットに対して x -means 法によるクラスタリングを適用することで、交点座標をクラスタ分類する。 x -means 法はクラスター分析手法の k -means 法の逐次繰り返しと BIC による分割停止基準を用いることで最適なクラスター数を自動決定するアルゴリズムである。類似度の高い交点座標は同じクラスターに分類される。最後に各クラスターの交点座標に対してマハラノビス距離による外れ値除去法を適用することで、クラスター重心から極端に離れた交点を除去し、交点座標が密集した視線集中地点をユーザに提示する。

5 プロトタイプシステム

前述した提案モデルに基づいたプロトタイプシステムを開発した。撮影システムは Android OS 搭載の携帯情報端末 (CPU: NVIDIA TegraTM 2 mobile processor with Dual Core ARM CortexTM - A9 CPU) 上で Java 言語により実装した。また、メタデータ解析サーバは CPU: Intel Core i5-3337U, 1.80GHz 上で C++ 言語により実装した。

6 実験と考察

前述したプロトタイプシステムを使用して実験を行った。本実験では、名古屋工業大学キャンパス内の指定した撮影対象物 4 か所 (2 号館, 20 号館, 22 号館, 52 号館) を被験者 10 人 \times 5 枚 (計 50 枚) 撮影してもらい、実際に撮影した撮影対象物の位置と本提案手法により導出した視線集中地点が一致しているか確認した。その結果、図 2 に示すように撮影対象物付近に交点座標が集中していることが確認できる。表 1 は各クラスターの交点のうち撮影対象物 2 か所 (2 号館, 22 号館) の領域内に存在する交点数から適合率を求めた結果である。実験対象とした 2 地点において 84.00 % (2 号館), 81.33 % (22 号館) という高い値を示した。また、外れ値除去

を適用することで適合率が上昇することが確認できる。さらに、被験者へのインタフェースに関するアンケート調査を実施したところ、メタデータ付与が撮影ボタンを押すだけでいいので、アノテーションに対して特に負担を感じなかったという肯定的な意見が得られた。



図 2: 視線集中地点

マハラノビス距離	適合率	
	2 号館	22 号館
閾値なし	84.00 %	81.33 %
6.0	85.13 %	91.04 %
5.0	85.13 %	95.08 %
4.0	88.73 %	100 %
3.0	87.88 %	100 %
2.0	87.27 %	100 %
1.0	100 %	100 %

表 1: 適合率

7 むすび

本研究が提案したシステムにより、撮影時のアノテーションコストを抑えつつ、視線メタデータから撮影支援情報を提示することが可能となった。今後、ユーザ同士で素材映像を共有できる機能を実装し、更なる映像作成支援機能を実現していきたい。

参考文献

- [1] 大平茂輝, 長尾確, “アノテーションに基づくスポーツ映像要約とナレーション生成”, 第 67 回情報処理学会全国大会講演論文集, p55-56, 2005.3
- [2] 高柳亮太, 舟橋健司, 山本大介, “動画撮影時におけるメタデータ付与と動画編集への応用”, 平成 24 年度電気関係学会東海支部連合大会, D1-1, 2012-9-24.