

I-012

# フォロー対象とカメラの移動軌跡の類似性に基づくフォロー区間検出方式

## Follow-Scene Detection based on Similarities between Trajectories of Camera Motion and Followed Objects

石川真澄† 大網亮磨†  
Masumi Ishikawa Ryoma Oami

### 1. はじめに

蓄積機器の大容量化と低価格化により、放送局や家庭に大量の映像が蓄積されるようになった。これらの映像から編集する区間や視聴する区間を効率的に選択するために、映像内容の瞬時的な把握への要望がある。

素早く映像内容を把握するための技術として、代表画像検出技術が提案されている。代表画像として、ショットの先頭フレーム[1]や、フレーム間クラスタリングに基づく代表フレーム[2]、顔を含むフレーム[3]を選択する手法が提案されている。しかし、[1][2]は代表画像中に主要な対象が映っているとは限らない点が問題となる。また[3]は、対象が顔に限定されるため、人物以外を撮影した映像に対応するには一般対象への拡張が求められる。

主要対象を限定しない手法として、移動する対象を追ってカメラを動かす技法であるフォローによって撮影された区間(フォロー区間)から、フォローされた対象(フォロー対象)を含むフレームを選択する方式[4]が提案されている。しかし[4]では、2フレーム間での対象とカメラの移動ベクトルの類似性に基づくため、対象の動きの変化にカメラが遅れて追従する場合等、対象とカメラの動きが瞬間的には異なる場合に正しく判定できない。

そこで本稿では、カメラの移動軌跡とフォロー対象上の特徴点の移動軌跡の間の類似性に基づくフォロー区間検出方式を提案する。本方式では、移動軌跡を比較することで、瞬間的にカメラと対象の動きが異なるフォロー区間も検出可能とする。素材映像とホームビデオに本方式と従来法を適用し、フォロー区間検出精度を比較する。

### 2. 移動軌跡と移動ベクトル

まず、移動軌跡と移動ベクトルについて説明する。図1に、フォロー区間内のフレームと、特徴点とカメラの移動ベクトルおよび移動軌跡の一例を示す。特徴点の移動軌跡は区間内での特徴点の位置の遷移を、カメラの移動軌跡は区間内での撮影領域の位置の遷移を、あるフレーム(例えば  $F_i$ ) 上の座標系でそれぞれ表したものである。また、特徴点およびカメラの移動ベクトルは、それぞれ特徴点とカメラの2フレーム間での移動量を表す。

図1では、右に移動する対象に合わせて撮影領域を右に動かしつつ撮影している。そのため、フォロー対象上の特徴点の位置と撮影領域の位置は右に遷移し、特徴点の移動ベクトルとカメラの移動ベクトルは右向きとなる。

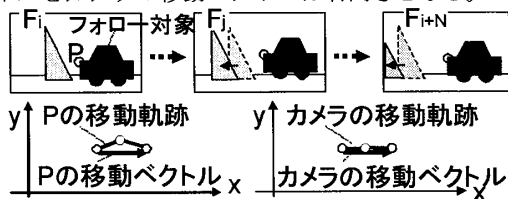


図1. 移動軌跡と移動ベクトル

### 3. 従来法[4]

#### 3.1 概要

2フレーム間(例えば  $F_i$  と  $F_{i+N}$ ) で求まる特徴点とカメラの移動ベクトルが類似することに着目し、カメラの移動ベクトルの大きさが閾値以上の場合に、カメラと類似する移動ベクトルを持つ特徴点をフォロー対象上の点として検出する。そして、検出特徴点数が閾値以上のときに  $F_i$  と  $F_{i+N}$  の間の区間をフォロー区間と判定する。

#### 3.2 問題点

フォロー対象の動きの変化にカメラが追従できない区間では、カメラと特徴点の移動ベクトルが異なる。このため、フォロー対象上の特徴点を検出できず、フォロー区間の検出漏れが生じる。

### 4. 提案法

#### 4.1 概要

カメラと特徴点の移動軌跡が類似することに着目する。 $F_i$  と  $F_{i+N}$  の間の複数の時刻で、カメラと特徴点の位置の差をもとにフォロー対象上の点か否かを仮判定する。複数の時刻での仮判定のうち、フォロー対象上の点とされた回数が否とされた回数以上の場合、特徴点をフォロー対象上の点と判定する。複数の仮判定の多数決処理によって、フォロー対象上の点の検出漏れを抑制する。

#### 4.2 処理の流れ

処理は以下の6ステップからなる。 $F_{i+N}$  を  $N$  間隔で選択し、(i)(ii)を  $F_i$  から  $F_{i+N}$  までの各隣接フレーム間で実行する。(iii)(iv)(v)で  $F_{i+N}$  内のフォロー対象の有無を判定し、(vi)で  $F_i$  から  $F_{i+N}$  の区間がフォロー区間か否かを判定する。

##### (i) 特徴点抽出・追跡

$F_i$  から特徴点を抽出し  $F_{i+N}$  まで追跡することで  $F_i$  から  $F_{i+N}$  までの各隣接2フレーム間での特徴点の動きベクトルを算出する。特徴点は、追跡に失敗しやすいフレーム枠付近の  $D_{rgn}$  画素の領域以外から抽出する。 $F_j$  ( $i < j < i+N$ ) でフレーム枠付近  $D_{rgn}$  画素の領域に移動した特徴点を消失点とし、 $F_{j+1}$  で消失点分の特徴点を追加する。

##### (ii) カメラワークパラメータ算出

各隣接2フレーム間でカメラワークパラメータを算出する。移動物体の動きによる影響を防ぐため、画面中央の  $C_{rgn}$  の領域以外から検出した動きベクトルを用いる。

##### (iii) 移動軌跡算出

動きベクトルとカメラワークパラメータをもとに、 $F_i$  と  $F_{i+N}$  の間のカメラと特徴点の移動軌跡を算出する。

##### (iv) 特徴点分類

カメラの移動軌跡と特徴点の移動軌跡から、 $F_{i+N}$  上の特徴点をフォロー対象上の点(フォロー点)かそれ以外の点(非フォロー点)かに分類する。

† 日本電気株式会社, NEC Corporation

(v) フォロー領域検出

フォロー点を位置に基づきグループ化し、グループ内フォロー点数が Pn 個以上のグループをフォロー領域とする。

(vi) フォロー区間判定

Fi+N からフォロー領域が検出された場合、Fi から Fi+N の区間をフォロー区間と判定する。

4.3 移動軌跡算出

特徴点 P の移動軌跡を {Xi, ..., Xi+N} とすると、P の Fj (i ≤ j < i+N) での座標 Xj : (xj, yj) と P の Fj+1 での座標 Xj+1 : (xj+1, yj+1) の関係は式(1)で表わされる。

$$X_{j+1} : (x_j - (x_{j+1}' - x_j) + V_x, y_j - (y_{j+1}' - y_j) + V_y) \quad (1)$$

なお、{z, θ, dx, dy} は Fj と Fj+1 の間でのカメラワークパラメータ、(Vx, Vy) は Fj と Fj+1 の間での点 P の動きベクトルである。Xj+1' (xj+1', yj+1') は、カメラの動きによって予測される Fj+1 での P の座標であり、以下で表される。

$$X_{j+1}' : ((x_j - \theta * y_j + dx)/z, (\theta * x_j + y_j + dy)/z)$$

また、P に対応するカメラの移動軌跡を {Xci, ..., Xci+N} とすると、Fj+1 での撮影領域の座標 Xcj+1 は式(2)で表される。

$$X_{cj+1} : (x_j - (x_{j+1}' - x_j), y_j - (y_{j+1}' - y_j)) \quad (2)$$

4.4 特徴点分類

式(3)を満たす特徴点 P をフォロー点と判定する。

$$\left( \sum_{j=0}^{N-1} |X_{j+1} - X_j| \right) / N < TH_1 \quad \text{かつ} \quad \left( \sum_{j=1}^N G(X_j - X_{c_j}) \right) \geq N/2 \quad (3)$$

$$G(D) = \begin{cases} 1 & \text{if } |D| < TH_2 \\ 0 & \text{if } |D| \geq TH_2 \end{cases}$$

5. 評価実験

5.1 実験条件

素材映像とホームビデオの各 10 本(1 分, 320x240, 29.97fps)を用い、手動付与したフォロー区間と抽出結果とを比較した。映像中、フォロー/非フォロー区間の比率はほぼ同じである。[4]と同様、特徴点追跡に Lucus-Kanade-Tomasi 法を、カメラワークパラメータ算出に点組み合わせフィルタリングを用いた。予備実験の結果から、Drng=8 画素、Crgn=50%、Pn=4 個、TH1=1 画素とした。手振れの多いホームビデオよりもプロの撮影した素材映像の方がフォロー対象とカメラの位置の差が小さい傾向があるため、素材映像で TH2=15 画素、ホームビデオで TH2=30 画素とした。N=10 フレームのときの提案法と従来法のフォロー区間検出精度を、適合率と再現率により比較した。

5.2 結果と考察

表 1 に全映像の総合精度を、図 2 と図 3 に各映像に関する精度を、素材映像とホームビデオとに分けて示す。

表 1 から、提案法により再現率が素材映像で 15%、ホームビデオで 11%向上し、適合率は素材映像およびホームビデオでともに数%劣化することがわかる。

再現率の向上は、人物の歩行速度の加速に対して遅れてカメラが追従した場合でも、フォロー区間と判定できたためだった。提案法で検出漏れになる区間は、画面端の対象が中央に移動するようにカメラを対象より遅く動かしたケースであった。また、適合率の低下の要因は、一様な模様の背景上の特徴点に対する追跡ミスによる誤検出であった。適合率、再現率の低下要因に対する対策は、今後の検討課題である。

図 2 と図 3 より、全映像について提案法は従来法と同等または高い再現率を得た。図 2 と 3 を比較すると、両手法でホームビデオの方が素材映像よりも映像間の適合率の差が大きい。これは、手振れを含む区間で、手振れによるカメラの移動軌跡と、追跡に失敗した特徴点の移動軌跡とが類似し、フォロー点が誤検出されたためだった。手振れを考慮した類似性の判定が必要と考えられる。

6. まとめ

本稿では、特徴点とカメラの移動軌跡の類似性に基づくフォロー区間検出について報告した。本方式では、カメラと特徴点の位置を複数の時刻で比較することでフォロー区間の検出漏れを抑制する。素材映像とホームビデオの各 10 本に適用し、従来法よりも再現率が素材映像で 15%、ホームビデオで 11%向上することを確認した。

参考文献

- [1] Y. Tonomura et al., "Structured video computing," IEEE Multimedia, vol.1, no.3, pp.34-43, 1994.
- [2] Uchihashi, S. et al., "Video manga: Generating semantically meaningful video summaries," In Proceedings of the ACM Multimedia Conference, pp.383-392, 1999.
- [3] D. Byrne et al., "A User-Centered Approach to Rushes Summarization Via Highlight-Detection Key Frames," TVS'07, Augsburg, Bavaria, Germany, pp.35-39, September 28, 2007.
- [4] 鳥井陽介他, "映像の動きを用いた動物体アップショット・フォローショット検出", MIRU2005 pp.24-31. 2005.07.

表 1 総合精度

	素材		ホームビデオ	
	再現率	適合率	再現率	適合率
提案法	67%	93%	62%	70%
従来法	52%	95%	51%	74%

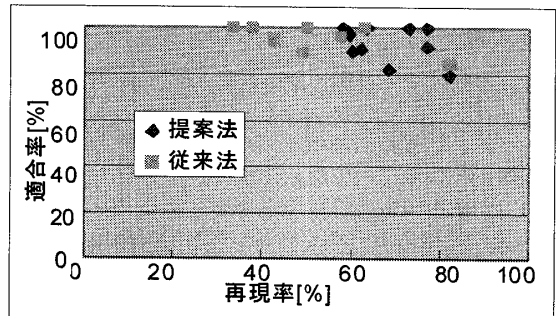


図 2 フォロー区間検出精度 (素材映像)

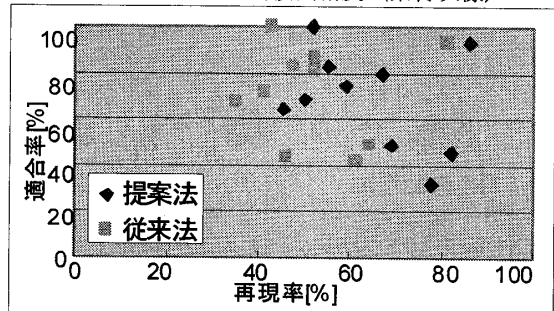


図 3 フォロー区間検出精度 (ホームビデオ)