

パーソナル映像におけるカット編集点の傾向と自動推定

古川 拓弥[†] 藤吉 弘亘[†] 野村 亮之^{††}

[†] 中部大学工学部情報工学科

487-8501 愛知県春日井市松本町 1200

^{††} 株式会社 sus4

461-0004 愛知県名古屋市東区葵 1-15-18 オフィスサンナゴヤ 9F

E-mail: [†]takuya@vision.cs.chubu.ac.jp, hf@cs.chubu.ac.jp, ^{††}aki-nomura@sus4.co.jp

あらまし 近年、個人が撮影した映像を Web で配信し、多くの人々が楽しむことのできる動画共有コンテンツが普及している。配信時のデータ制限や誘目性の高い映像作成をするには映像編集が不可欠となるため、Web 上で簡易に映像編集を行う支援システムの需要が高まっている。本稿では、実際にユーザが Web 上で編集したパーソナル映像のカット点の傾向について調査し、その結果に基づいたカット点の自動推定法を提案する。パーソナル映像におけるカット編集点について ST-patch 特徴を用い相関を求めたところ、CRIM(Continues Rank-Increase Measure) の値が低いフレームにおいてカット編集を行っているという傾向が得られた。そこで、ST-patch 特徴から算出される CRIM に基づくカット点推定の結果、カット点の再現率 35.3%、適合率 45.8% 得た。

キーワード パーソナル映像, カット編集点, カメラワーク, ST-patch 特徴量

A Method for Estimating Cut Editing Points Based on Characteristics of Personal Video Contents

Takuya FURUKAWA[†], Hironobu FUJIYOSHI[†], and Akiyuki NOMURA^{††}

[†] Department of Computer Science, Chubu University

1200 Matsumoto-cho, Kasugai, Aichi, 487-8501 Japan

^{††} sus4 Co., Ltd. 1-15-18 Aoi, Higashi-ku, Nagoya, Aichi, 461-0004 Japan

E-mail: [†]takuya@vision.cs.chubu.ac.jp, hf@cs.chubu.ac.jp, ^{††}aki-nomura@sus4.co.jp

Abstract In recent years, there has been a significant increase in shared video contents; that is, video contents created by individuals, and distributed via the Internet for many people to enjoy. Because video must be edited to create eye-catching contents within the volume limitations placed on the video data to be distributed, there is a growing demand for support systems that enable users to edit video contents easily on the Internet. In this paper, we propose a method that surveys trends for cut points in personal video contents edited by users on the Internet, and automatically infers cut points based on the results of those surveys. When we investigated the relationship among cut editing points in personal videos using ST-patch characteristics, we determined that there was a tendency to make cut edits in frames with minimal value of CRIM(Continues Rank-Increase Measure)calculated from ST-patch. We evaluated the proposed method based on CRIM, and we obtained a recall ratio of 35.3% and a precision ratio of 45.8% for estimating cut editing point in personal video contents.

Keyword Personal Video, Cut Editing Point, Camerawork, ST-patch Feature

1. はじめに

近年では、パーソナルビデオカメラの低価格化や機能向上により、誰もが気軽に映像撮影が可能となっている。それに伴い、撮影した映像をインターネットで配信し、多くの人々に視

聴してもらうための動画共有コンテンツサイトが普及している。しかし、映像を Web 上で配信するには、映像のデータ容量の制限や、第三者に対して誘目性の高い映像作成を考慮しなければならない。そのために、撮影した映像を編集することが必要となる。しかし、映像を編集するには大変な労力と時間を

必要とするため、配信前に編集を行っているユーザは依然少ないという課題があり、映像編集を支援するシステムの需要が高まっている。

従来の自動編集手法に関する研究では、熊野らの野球中継のハイライトシーンを実時間配信を目的とした PC (Picher and Catcher) シーンの自動検出 [1] や、尾関らの料理や組み立て作業などの机上作業映像のためのイベント駆動型自動編集手法 [2] などが提案されている。しかし、これらの手法は限られた映像のみの自動編集であり、個人が撮影したパーソナル映像に対しては適用できないという問題がある。岩城らは、パーソナル映像を対象とした自動編集システム [3] を提案している。これは、シーン毎の特徴量を用いて重要度を設置し、指定した時間内の映像を選択するものである。しかし、重要度により選ばれたシーンが本来ユーザが求めるものであるのかどうかはわからない。また、プロのカメラマンは編集による効果を想定して、与えられた映像の意図を反映するショットを撮影を行うことが熊野ら [4] によって報告されている。一般ユーザにおいても意識的にもしくは無意識的に、編集を意識した映像撮影を行っていると考えられる。そこで、本研究ではカット編集に着目し、ユーザが Web 上で編集を行った映像からカメラワークにおける傾向と ST-patch 特徴によるカット編集点の傾向調査を行う。さらに、これらの傾向に基づいたカット編集点の自動推定を行う。

2. パーソナル映像のオンライン編集

本研究では、2006 年 8 月から 2008 年 8 月までに動画共有サイト ClipCast [7] に投稿されたパーソナル映像を対象とする。ClipCast では、Web 上で投稿動画の編集が可能となっている。投稿された動画は、動物や人、風景、スポーツなど、様々なシーンを撮影したものである。動画サイズは 320 × 240 で動画形式は flv である。2 年間に投稿された 6518 本の動画うち、サイト上でカット編集が行われた 1075 本の動画に対して調査を行う。以下に、ClipCast 上で編集を行った際に作成される編集データと各編集シーンにおけるカメラワークの傾向について述べる。

2.1 パーソナル映像の編集データ

ClipCast ではユーザが動画を投稿する際、図 1 のように簡易な操作で以下に示す 5 種類の編集を行うことができる。

- (1) シーンのカット
- (2) テキストの挿入
- (3) フレームの挿入
- (4) エフェクトの付加
- (5) トランジション (カット編集時における効果) の付加

ユーザが図 1 のようにサイト上で編集を行い動画作成をする際、編集を行う動画に対し xml 形式で編集データが出力される。編集データとして以下に示す 6 つの情報が記述される。

- (1) ユーザ ID
- (2) タイトル
- (3) カテゴリ、タグ情報
- (4) 使用した素材動画名



図 1 編集ステップ



図 2 カット編集例

(5) カット編集履歴

(6) テキスト、フレーム編集履歴

項目 (5) のカット編集履歴データより、ユーザがカット編集をした切り出しと切り終わりフレームを得ることができる。本研究では、図 2 のように切り出しと切り終わりのフレームをそれぞれ「IN 点」、「OUT 点」、切り出しされたシーンを「使用シーン」、それ以外のシーンを「未使用シーン」と定義する。なお、未使用シーンは使用シーンの前後で「前未使用シーン」と「後未使用シーン」と定義する。

2.2 カメラワークにおけるカット編集点の傾向

カット編集点とカメラワークについて相関があると考え、IN 点と OUT 点、使用シーン、未使用シーンにおいて目視でカメラワーク有無の傾向調査を行う。以下に示す 6 種類のカメラワークについて、フレーム単位でどのカメラワーク技法が使用されているかデータ収集を調査した。

- (1) 「フィックス」カメラを動かさずに撮影
- (2) 「パン」カメラを左右に振る
- (3) 「ティルト」カメラを上下に振る
- (4) 「ズーム」レンズを操作して被写体に寄るまたは引く
- (5) 「フォロー」動いている被写体に合わせカメラを動かす
- (6) 「ドリー」カメラマン自身が動いて撮影を行う

なお、本研究でのカメラワークの傾向調査において、カメラを動かさず撮影を行う「フィックス」についてはカメラワーク無しと判断する。

2.2.1 使用、未使用シーンでのカメラワーク発生割合

使用シーン、未使用シーンにおいて、フレーム単位でカメラ

表 1 使用シーンと未使用シーンにおけるカメラワークの割合と種類

	前未使用	使用	後未使用	All
総フレーム数	479117	1058145	421057	1958319
カメラワークフレーム数	36412	285699	34107	356218
割合	7.6%	27.0%	8.1%	18.7%
Follow	30.2%	24.1%	27.4%	24.4%
Left pan	12.8%	17.8%	14.5%	19.4%
Right pan	14.9%	19.0%	16.2%	18.8%
Dolly	13.8%	12.7%	14.0%	13.4%
Zoom in	8.0%	7.2%	7.8%	7.7%
Zoom out	9.1%	8.0%	8.9%	7.5%
Tilt up	5.0%	5.9%	5.4%	4.5%
Tilt down	6.2%	5.2%	5.8%	4.2%
その他	0.0%	0.1%	0.0%	0.0%

表 2 IN,OUT 点でのカメラワーク発生割合と種類

	IN 点	OUT 点
カメラワーク有	147	151
カメラワーク無	928	924
カメラワーク有の割合	13.7%	14.0%
Follow	32.4%	24.8%
Dolly	17.3%	23.4%
Left pan	13.9%	14.4%
Right pan	13.2%	13.0%
Zoom in	9.8%	8.9%
Zoom out	7.4%	6.7%
Tilt up	3.2%	3.1%
Tilt down	2.8%	5.7%

ワークの発生割合とその種類について調査した結果を表 1 に示す。表 1 より、全体を通して「フォロー」が多いことが判明した。一般ユーザは、対象物を追いかけながら撮影していると推測できる。前未使用シーン、後未使用シーンについて比較すると、カメラワークの割合について違いが見られなかった。カメラワークの種類については多少の変動があるものの、順位は変わらず顕著な傾向はないといえる。また、使用シーンと未使用シーンでのカメラワークの割合に注目すると、前未使用シーンで 7.6%、後未使用シーンで 8.1% という結果に対し、使用シーンでは 27.0% と 3 倍以上高いことがわかる。以上をまとめると、一般ユーザはカメラワークが発生しているフレームを使用シーンとしている確率が高いという傾向があるといえる。

2.2.2 IN 点, OUT 点のカメラワークの発生割合

IN 点, OUT 点のカメラワークの発生割合とその種類について調査した結果を表 2 に示す。表 2 の IN 点, OUT 点における各カメラワークの種類に注目すると、IN 点と OUT 点では大差がないことがわかる。次に、IN 点, OUT 点におけるカメラワーク有の割合に注目すると、IN 点で 13.7%、OUT 点で 14.0% とほぼ同じ値であるという結果が得られた。これより、IN 点, OUT 点では共にカメラワークの少ないフレームにおいてカット編集を行う傾向があるということがわかる。

以上の傾向より、カメラワークの有無を定量的に求めることでカット推定が可能と考えられる。

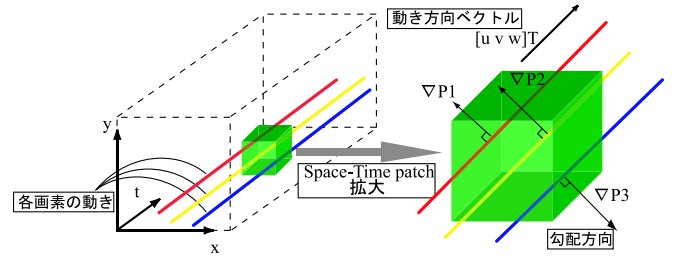


図 3 ST-patch の概要図

3. ST-patch によるカット編集点の傾向調査

映像中の動きを定量的に捉え、カット編集点の傾向調査を行う。画像上の動きの解析を行う手法として、物体の動きの移動量をベクトルデータとして表現したオプティカルフロー [5] が挙げられる。しかし、オプティカルフローではノイズの影響を受けやすく、動きの評価が困難となる。この問題に対して Shechtman らは、テンプレート映像と入力映像の時空間画像から抽出された 2 つの space-time patch(ST-patch) の相関関係を複数用いることにより、テンプレート映像の物体と同じ動きをしている物体を入力映像から識別する手法 [6] を提案している。この手法で用いられている ST-patch の相関関係は、テクスチャが異なる対象物体の動きや、非剛体の物体の様に複雑に動く物体に対しても動きの評価を行うことができる。そこで本研究では、カット点や各シーンにおける映像中の動きを捉えるために ST-Patch を用いる。以下に、ST-patch 特徴量の算出方法と動きベクトルの解析について述べ、ST-patch による編集点の調査結果について述べる。

3.1 ST-patch 特徴量

図 3 に ST-patch の概要図を示す。ここで、 x, y は画像の座標軸、 t は時間軸、3 本の線は個々の画像の動き、 $[u \ v \ w]^T$ は ST-Patch 内の動きベクトル、 ∇P_i は個々の画素の勾配方向ベクトルを表している。この ST-patch 特徴量を得るために、時空間画像においての x 軸、 y 軸の勾配を求める。画像中の動きが一定の場合、各軸に対するある画素 i の勾配 $\nabla P_i = (P_{xi}, P_{yi}, P_{ti})$ は、画素の動き方向ベクトル $[u \ v \ w]^T$ に対して垂直となる。よって式 (1) の関係が成り立つ。

$$\nabla P_i \begin{bmatrix} u \\ v \\ w \end{bmatrix} = 0 \quad (1)$$

画素数が n の場合、式 (1) は式 (2) となる。

$$\begin{bmatrix} P_{x1} & P_{y1} & P_{t1} \\ P_{x2} & P_{y2} & P_{t2} \\ \vdots & \vdots & \vdots \\ P_{xn} & P_{yn} & P_{tn} \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (2)$$

ST-patch 中の n 画素の ∇P_i からなる $n \times 3$ の行列を G とし、行列 G^T を掛けると式 (3) となる。

$$G^T G \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (3)$$

この様に、行列 $G^T G$ は 3×3 の行列となる。ここで、行列 $G^T G$ を M とすると式 (4) のように表すことができる。

$$M = G^T G = \begin{bmatrix} \sum P_x^2 & \sum P_x P_y & \sum P_x P_t \\ \sum P_y P_x & \sum P_y^2 & \sum P_y P_t \\ \sum P_t P_x & \sum P_t P_y & \sum P_t^2 \end{bmatrix} \quad (4)$$

式 (4) より求められる行列 M が 1 つの ST-patch から得られる特徴となる。行列 $M(3 \times 3)$ には、左上の 2×2 の行列にアピランスの情報、3 行目と 3 列目にはモーションの情報を持つ。

3.1.1 ランク制約による動きベクトル解析

式 (4) の 3×3 行列で表される ST-patch 特徴のランク制約を用いることにより、ST-patch 内の動きベクトルの解析を行うことができる。具体的には、動きベクトルの定常度、2 つの ST-patch 間の動きの類似度の 2 つを求めることができる。

3.1.2 ランク制約

ST-patch の時空間特性と画像空間特性を用いて、ランク制約による ST-patch 内の動きベクトルの定常度を求める。ここで、時空間特性とは式 (4) で示す行列 M を表し、画像空間特性とは以下に示す行列 M° を表す。

$$M^\circ = \begin{bmatrix} \sum P_x^2 & \sum P_x P_y \\ \sum P_y P_x & \sum P_y^2 \end{bmatrix} \quad (5)$$

ST-patch の時空間特性から、単一の動きの ST-patch は $\text{rank}(M) = \text{rank}(M^\circ)$ のようなランクとなる。そして、複数の動きを含む ST-patch では、1 つの行と列だけが行列 M から M° へ変換されているため、ランクの違いが 1 となる。よって式 (5) のような関係が成り立つ。

$$\Delta r = \text{rank}(M) - \text{rank}(M^\circ) = \begin{cases} 0: \text{単一の動き} \\ 1: \text{複数の動き} \end{cases} \quad (6)$$

3.1.3 CRIM (Continues Rank-Increase Measure)

ランクの増加量 Δr は、動きが単一か複数の 2 通りしか求めることができず、動きがどれだけ類似しているかを求めることはできない。そこで、行列 M と行列 M° より決められる固有値を用いて Continues Rank-Increase Measure (CRIM) Δr を算出する。 $\lambda_1 \geq \lambda_2 \geq \lambda_3$ を行列 M の固有値、 $\lambda_1^\circ \geq \lambda_2^\circ$ を行列 M° の固有値とする。行列 M 、行列 M° は共に対象行列であり、行列 M 、行列 M° の固有値はすべて実数となる。また、それぞれの固有値の関係は $\lambda_1 \geq \lambda_1^\circ \geq \lambda_2 \geq \lambda_2^\circ \geq \lambda_3$ となる。このような関係から、CRIM Δr を式 (6) のように定義する。

$$\Delta r = \frac{\lambda_2 \cdot \lambda_3}{\lambda_1^\circ \cdot \lambda_2^\circ} \quad (0 \leq \Delta r \leq 1) \quad (7)$$

ここで求められた Δr の値は、ST-patch 内の動きベクトルが一定の時に $\Delta r \approx 0$ となり、動きベクトルの変化が激しいときに $\Delta r \approx 1$ となる。これにより、ST-patch 内の動きベクトルの定常度を求めることができる。

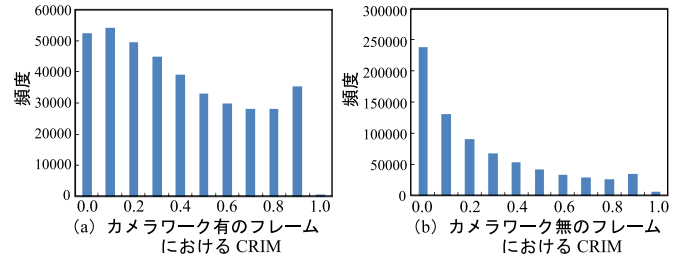


図 4 カメラワークの有無に対する CRIM

3.1.4 MC (Motion Correlation)

式 (7) から求められた CRIM Δr より、2 つの ST-patch P1 と P2 の動きの類似度を求めることができる。P1 の CRIM を Δr_1 、P2 の CRIM を Δr_2 、P1 と P2 をまとめた CRIM を Δr_{12} とすると、式 (8) により Motion Correlation (MC) m を求められる。

$$m = \frac{\min(\Delta r_1, \Delta r_2)}{\Delta r_{12}} \quad (8)$$

ここで求められた m は、 $m \approx 1$ の場合は動きの類似度が高く、 $m \approx 0$ の場合は動きの類似度が低いことを意味する。

3.2 カット編集点における傾向

ST-patch 特徴を用いてカット点における傾向調査を行う。ここで、ST-patch を動画サイズである 320×240 のサイズで作成すると、パッチサイズが大きくなるため計算コストが増大する。そこで、本手法では以下に示す 3 つのステップにより CRIM と MC の算出を行う。

(1) 入力画像のサイズを 1/3 にダウンサンプリング

(2) ダウンサンプリング画像のパッファから ST-patch 特徴量を算出

(3) 求められた ST-patch 特徴量を用いて CRIM と MC の算出

以下では、カット編集点における傾向を CRIM と MC を用いて調査した結果について述べる。

3.2.1 カメラワーク時における ST-patch 特徴

一般ユーザは使用シーンとしてカメラワークの多いシーンを使用している傾向が得られたため、カメラワークの発生しているフレームにおける CRIM を求めた。図 4 にカメラワークの有無に対する CRIM のヒストグラムを示す。図 4(a),(b) を比較するとカメラワークが発生しているフレームでは CRIM が高い値で分布し、カメラワークが発生していないフレームについては CRIM が低い値に分布しているのがわかる。これより、CRIM 値は、カメラワークの有無の判定に利用できるといえる。

3.2.2 IN 点、OUT 点での ST-patch 特徴

IN 点、OUT 点における CRIM のヒストグラムを図 5 に、MC のヒストグラムを図 6 示す。図 5 の (a),(b) と図 6 の (a),(b) をそれぞれ比較すると、IN、OUT 点間の分布に違いは見られない。従って、IN 点、OUT 点における傾向には差がないと判断できる。図 5 に着目すると、CRIM は低い値に分布をしていることがわかる。これは、式 (7) より動きベクトルの変化が少ないフレームにおいてカット編集を行っているとい

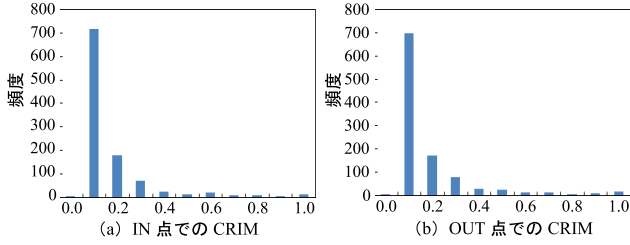


図 5 IN 点, OUT 点での CRIM(Continues Rank-Increase Measure)

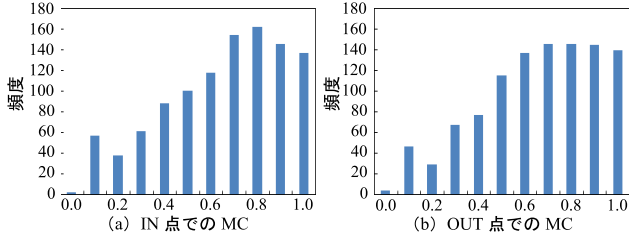


図 6 IN 点, OUT 点での MC (Motion Correlation)

える．図 6 の MC の分布では高い値に分布している．これは，式 (8) から前後フレームが類似しているところでカット編集を行っている傾向があるといえる．

以上より，IN 点，OUT 点では映像中における動きが少ないフレーム，すなわちカメラワークや映像中の被写体の動きが少ないフレームでカット編集を行っており，その傾向を ST-patch より捉えることができる．

4. カーネル密度推定を用いたカット編集点の推定

目視によるカメラワーク調査と ST-patch 特徴を用いた傾向調査により，使用シーンではカメラワークの発生割合が高く，IN 点，OUT 点は動きのベクトルの変化が少ないフレームにおいてカット編集が行われているという傾向が得られた．これらの傾向を用いたカット点の自動推定を行う．図 7 に自動推定の流れを示す．まず，3.2 節で述べたように，入力動画に対してダウンサンプリングを行い，ST-patch 特徴を算出し CRIM を求める．次に，CRIM に対してカーネル密度推定を行う．最後に，確率密度関数と閾値が交差するフレームをカット点候補とする．以下では，カーネル密度推定を用いたカット点の推定例と，推定したカット点の有用性を確認する評価実験について述べる．

4.1 カーネル密度推定を用いたカット編集点の推定

カーネル密度推定とは，パラメトリックな形式を仮定せずに確率密度関数を推定する手法の 1 つである．バンド幅が h で， N 個のデータ数の中にデータ点 x があるとき，カーネル密度推定で求められる確率密度関数は式 (9) となる．

$$p(x) = \frac{1}{N} \sum_{i=1}^N \frac{1}{(2\pi h^2)^{\frac{M}{2}}} \exp \left\{ -\frac{|x - x_i|^2}{2h^2} \right\} \quad (9)$$

カーネル密度推定を用いることで，どのような分布関数であっても，一応の近似が行える．カーネル密度推定を用いたカット

表 3 カット点の推定結果 [%]

閾値	再現率 R	適合率 P	F 値
0.1	31.0	35.2	30.0
0.2	31.6	38.3	31.1
0.3	31.3	40.2	31.2
0.4	35.3	45.8	34.8
0.5	32.3	46.9	32.8
0.6	29.8	48.4	31.5
0.7	27.2	48.9	29.3
0.8	24.0	50.4	27.0

編集点の推定例を図 8 に示す．

4.2 実 験

提案手法のカット点推定の有効性を確認するために，評価実験を行う．評価にはカット編集が行われた 1075 本のうち，ClipCast に投稿前に映像編集が行われていた映像データを除いた 666 本のデータを対象とした．推定精度の評価は，再現率 (Recall) と適合率 (Precision) により行う．さらに，再現率・適合率はともに高いことが望ましいため，両方を同時に評価するための F 値を用いる．再現率 R ，適合率 P ，F 値 F はそれぞれ以下の式により求める．

$$R[\%] = \frac{C}{T} \times 100 \quad (10)$$

$$P[\%] = \frac{C}{S} \times 100 \quad (11)$$

$$F[\%] = \frac{2PR}{P + R} \quad (12)$$

ここで， C は正解したフレーム数， T は正解のフレーム数， S は推定したフレーム数である．図 7 の閾値処理 (2) の閾値を 0.1 から 0.8 までを変化させたときの R ， P ， F 値にて評価を行う．カット点推定の実験結果を表 3 に示す．表 3 より，閾値が 0.4 のときに F 値が最も大きく 34.8% (再現率は 35.3%，適合率は 45.8%) を得た．

4.3 考 察

F 値が 90% を超えるカット編集点の推定に成功した動画を視聴した結果，これらの動画では，カメラワークが多く発生していたものが多いことを確認した．これは，カメラワークが発生しているシーンを使用シーンとして検出することができたからである．一方，F 値が 10% 未満のカット編集点の推定に失敗した動画を視聴した結果，図 8(b) のようにカメラを固定させてペットや赤ちゃんを撮影している映像が多くみられた．このようにカメラを固定し対象物の撮影を行っている映像については，提案手法は CRIM 値に変化がないため対応できない．固定カメラへの対応については，他のコンテキスト情報 (例えば顔など) を用いたカット点の推定等が考えられ，今後の課題としたい．

提案手法を用いたカット点推定では，F 値が最大で 34.8% であり，自動推定システムとして用いるには精度が低い．しかし，推定したカット点をカット候補点とし，ユーザが最終的なカット点を決定する支援システムとしての利用には有用であると考えられる．カット点をあらかじめ推定しておくことによって，ユーザの映像編集における負担を軽減することが期待で

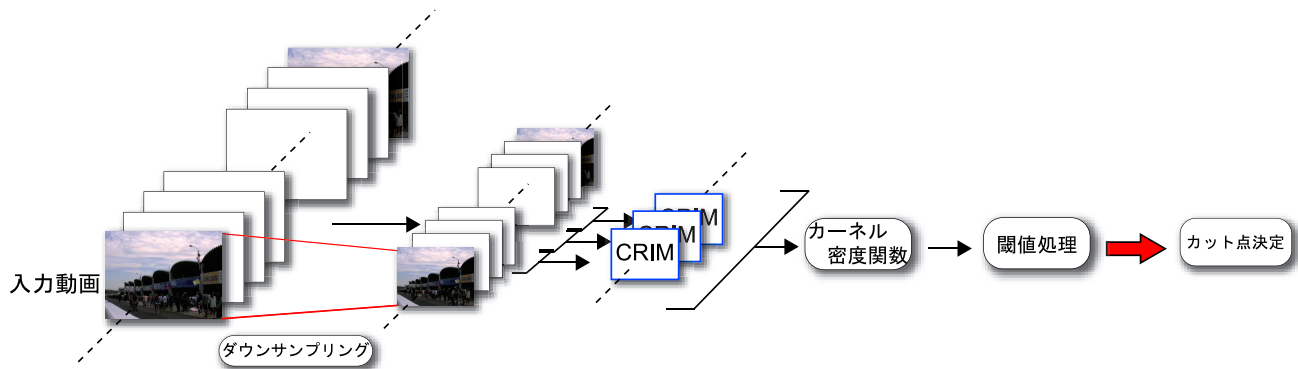


図 7 カット点推定の流れ

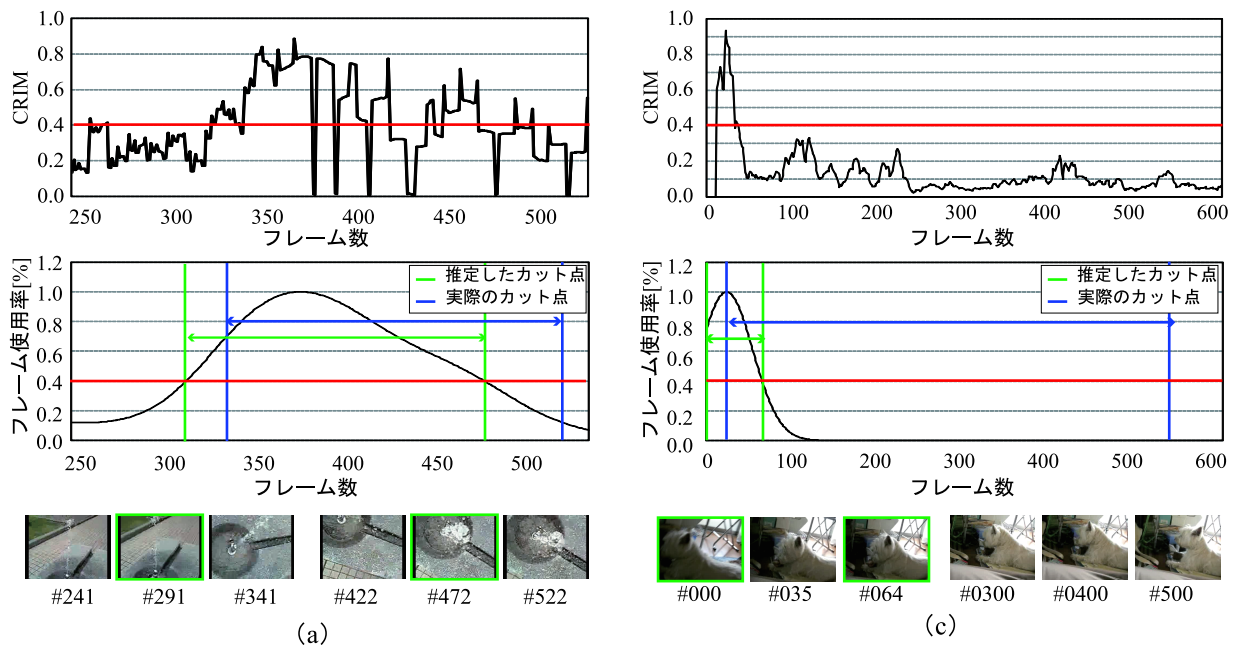


図 8 カット編集点の推定例

きる。

5. む す び

本稿では、一般ユーザが Web 上で映像編集を行った編集履歴を用いてカット点の傾向調査を行い、パーソナル映像におけるカット点の自動推定を行った。一般ユーザはカメラワークが発生しているシーンを使用シーンとし、映像中の動きの変化が少ないフレームでカット編集を行う傾向があることを確認した。この結果に基づき、ST-patch から得られた CRIM を用いてカット点の推定を行った結果、再現率、適合率は最大で 35.3% と 45.8% となった。この結果より、カット編集を行うユーザに対し、提案手法を用いてカット候補点の推定を行う編集支援システムとして用いることで、映像編集時にかかる負担軽減を行うことが可能であると考えられる。今後は、顔などの情報とカット編集点の傾向調査を行っていく予定である。

文 献

- [1] 熊野 雅仁, 神崎 伸夫, 藤本 雅清, 有木 康雄, “野球中継のハイライトシーン実時間配信を目的とした PC シーンの自動検出”, 電子情報通信学会技術研究報告, Vol.103, No.207, pp27-34 2003.
- [2] 尾関 基行, 中村 裕一, 大田 友一, “机上作業映像のためのイベント駆動型自動編集手法”, 電子情報通信学会技術研究報告, Vol.103, No.205 pp. 7-12 2003.
- [3] 岩城 澄, 中沢 実, 服部 進実, “パーソナル映像の自動編集システム”, 電子情報通信学会技術研究報告, Vol.102, No.737, pp.1-4 2003.
- [4] 熊野 雅仁, 天野 美紀, 有木 康雄, 上原 邦昭, “映像文法に基づいた実時間使用可能ショット識別による撮影ナビゲーションシステム”, 電子情報通信学会技術研究報告, Vol.104, No.369, pp.1-6 2004.
- [5] BKP.Horn and BG.Dchunck, “Determining ooptical flow”, Artificial Intelligence, vol.17, pp.185-203, 1981.
- [6] E. Shechtman and M. Irani, “Space-Time Behavior Based Correlation”, Computer Vision and Pattern Recognition, Vol.1, pp.405-412, 2005.
- [7] ClipCast, <http://clipcast.jp/>