

## 動きベクトルに基づく映像の短縮

## Video Skimming Based on Motion Vector

鈴木 史彦<sup>†</sup> 東海林 健二<sup>†</sup> 外山 史<sup>†</sup> 宮道 壽一<sup>†</sup>

Fumihiko Suzuki Kenji Shoji Fubito Toyama Juichi Miyamichi

## 1 はじめに

近年、監視カメラなどで撮影された長時間映像が増加していることによって、記録媒体の容量を有効活用することや映像分類の容易化などのために、重要な領域を保存したまま映像を圧縮する技術が注目されている。その技術の1つとして、文献[1]では、映像圧縮の手法の1つとして映像のフレームのアスペクト比を縮小するために、非線形リサイズ技術である Seam Carving を用いた手法が提案されている。しかし、この手法は時間軸方向の圧縮を行っていない。文献[2]では、グラフィック法を用いて、時間軸方向の短縮を行っているが、計算量が多く、処理時間がかかってしまう。

数多くある映像圧縮の手法の中で、時間軸方向の短縮のためには、動きの少ない領域を時間軸方向に削除する Ribbon Carving という手法が有効であると注目されている[3]。一方、動きの大きい領域を削除して時間軸方向の短縮を行うと、映像に対して一種の特殊効果を与えることができる。これは、固定カメラに対しては運動物体を削除した映像を作り出すことに相当する。

本研究では、映像の時間軸方向の短縮を行う際に、動きベクトルにさまざまな評価関数を適用して削除領域を決定することにより、映像要約や運動物体削除を含むさまざまな特殊効果を与える手法を提案する。また、本研究では Ribbon Carving を用いた処理を行っており、処理が高速な動的計画法を用いて削除位置を求めている。

## 2 Ribbon Carving

本研究では、時間軸方向の短縮に動的計画法を利用した Ribbon Carving を用いている。動きの少ない領域を削除するこの手法を従来手法とし、その概要を以下で述べる。まず、入力映像を  $F(x,y,t)$  とする。  $x,y$  は、その映像のフレーム画像の幅と高さを表し、  $x=1,\dots,W,y=1,\dots,H$  とし、  $t$  はフレーム数で  $t=1,\dots,N$  と表す。 Ribbon Carving は、映像から ribbon(3次元縫い目)を算出して、その ribbon の中で最小のコストを持つものを見つけ、最小となる位置の画素を削除する。映像から vertical ribbon と horizontal ribbon の2つの ribbon を算出する。 vertical ribbon は映像中の上下の動きベクトルを基に生成され、 horizontal ribbon は映像中の左右の動きベクトルを基に生成される。それら2つの ribbon を図1に示す。

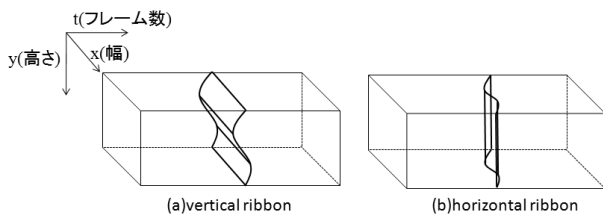


図1: Ribbon Carving

ここで、 horizontal ribbon  $R$  について述べる。  $y=1$  の  $x-t$  平面において  $R$  上の各画素は1つの  $x$  に対して1つの画素を持ち、  $t(x)$  は  $x$  に対して連続した値をとる。そして、  $(x,y,t(x))$  は、  $y=2,\dots,H$  において  $R$  上の画素となる。

図1から ribbon  $R$  のコストは式(1)により求めることが

<sup>†</sup>宇都宮大学大学院工学研究科

できる。  $C_R$  は  $R$  上の座標  $(x,y,t)$  のコスト  $C(x,y,t)$  の総和である。

$$C_R = \sum_{x,y,t \in R} C(x,y,t) \quad (1)$$

このときのコストの値が小さい位置を動きの小さい領域として検出する。 vertical ribbon, horizontal ribbon の最小累積コストを求めるために、  $C(x,y,t)$  をそれぞれ  $y-t$  平面、  $x-t$  平面に投影し、  $x$  軸、  $y$  軸方向にコストの値を足し合わせていく。このときの horizontal ribbon の場合のコスト算出式を式(2)に示す。  $x-t$  平面から求めた ribbon を  $R_h$  とする。 vertical ribbon の場合も同様にして  $x$  軸方向の総和から  $R_v$  を求める。

累積コストから生成した2次元画像を図2に示す。図2の2次元画像(a),(b)に対して Seam Carving 処理を施す。このとき、文献[4]で述べられているように動的計画法を用いることで Seam Carving を高速に処理することが可能である。

$$C_{R_h} = \sum_{y=1}^H C(t(x), x) \quad (2)$$

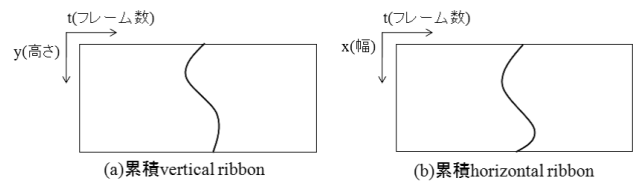


図2: 累積コストの2次元画像

求めた  $R_v$  と  $R_h$  の両者のコストを比較し、よりコストの小さい縫い目の画素を削除する。この処理1回で時間軸方向に1フレーム分の短縮を行うことができる。つまり、この処理の繰り返し回数は映像を短縮するフレーム数となり、映像中の重要な領域を残して映像の要約を行うことができる。

## 3 提案手法

本手法では、入力された映像中の運動物体の持つ動きベクトルに着目して、その動きベクトルに対して様々な評価関数を用いて削除領域を決定する。それにより、従来手法では動きの少ない領域に対してのみの処理であったが、提案手法では映像要約の他に、運動物体削除を含む様々な効果を与える。例えば、運動物体の速度に基づいて削除領域を変化させることや短時間の撮影で映像から背景画像を取得することが可能となると考える。以下で、提案手法について述べる。

## 3.1 評価関数

提案手法で ribbon を算出する際に用いた評価関数を以下に示す。本手法では、評価関数に1を加え、その逆数を計算することで運動物体の領域を求めている。処理するフレーム画像は、画素の明るさを取得しやすくするためにグレースケール画像に変換して明るさの差分を求め、フレーム画像から運動物体のエッジを検出する。式(3)は文献[3]で用いている評価関数で、各画素のグラディエントを求めている。式(4)は、映像中の画素  $(x,y,t)$  における明るさの差分を求め、フレーム画像からラプラシアンを計算して ribbon のコストを計算している。式(5)は、フレーム間差分によりフレーム間の変化のみが

らコストを求めている．以下の式で求めた値から式 (1) の和が最小となるものを運動物体の領域として検出する．

$$C_1(x, y, t) = \frac{1}{\sqrt{(I'_x(x, y, t))^2 + (I'_y(x, y, t))^2 + (I'_t(x, y, t))^2 + 1}} \quad (3)$$

$$C_2(x, y, t) = \frac{1}{|I''_x(x, y, t) + I''_y(x, y, t) + I''_t(x, y, t)| + 1} \quad (4)$$

$$C_3(x, y, t) = \frac{1}{|I'_t(x, y, t)| + 1} \quad (5)$$

### 3.2 映像に与える特殊効果

本手法では，上で述べた様々な評価関数を用いることで映像に対して特殊効果を与えることを期待する．例えば，入力映像中で複数の運動物体が同時に存在している場合，速度の速い運動体は，速度の遅い運動体より動きベクトルが大きくなる．そのため，動きベクトルの大きい方を削除するという処理が可能であると考えられる．

## 4 シミュレーション実験

実験で用いた映像はデジタルカメラを固定し，撮影した実映像である．実映像はフレームサイズを  $320 \times 240$ , 30fps のものを使用した．映像中に出現する運動物体は水平方向のみの運動を行うため，算出する ribbon は図 2 の (b) に示した horizontal ribbon に対して Seam Carving を行うだけとした．その累積 horizontal ribbon の 2 次元画像を基に画素の削除処理を行った．なお，両実験とも式 (4) を評価関数として用いた．

### 4.1 実験 1

フレーム数を 50，削除フレーム数を 40 として実験を行った．映像は開始時点の 0 フレーム目で人が右端から歩き始め，49 フレーム目で左端に到達する．映像の全フレームに運動物体が存在している．映像の一部フレーム画像を図 3 に示し，削除処理前の累積コスト 2 次元画像を見やすさのために  $t$  軸方向に 6 倍したものを図 4 に示す．そして，処理結果映像の一部のフレーム画像を図 5 に示す．

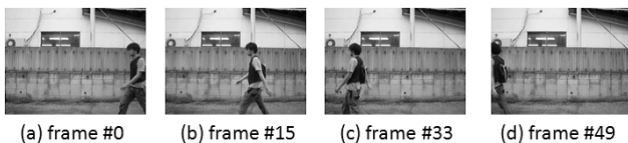


図 3: 入力映像のフレーム画像

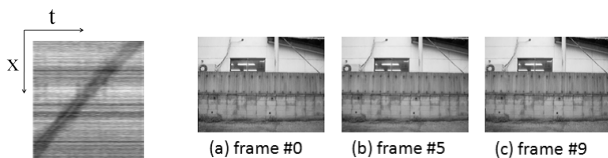


図 4: 累積コスト 図 5: 実験 1 の処理結果

実験の結果，フレーム中の歩行している人物は削除され，背景画像のみの映像が生成された．図 4 の累積コスト画像は，累積コストの値が最大となる位置を白画素で表している．図 4 から運動物体のエッジを検出できていることがわかり，運動物体の領域を正確に削除することができたと考えられる．本実験において，運動物体はすべてのフレームに存在しているため，単にフレームの切り出しを行うような編集では背景映像の作成は実現できない．そのため，本手法は背景画像を取得したい場合などに対しての有効であると考えられる．

### 4.2 実験 2

フレーム数を 70，削除フレーム数を 50 として実験を行った．映像は開始時点の 0 フレーム目に右端に存在している車が左方向に進行する．22 フレーム目で左端の車がフレームアウトすると同時に，左端からトラックがフレームインして右方向に進行する．37~57 フレーム目で 2 台の車はすれ違い，69 フレーム目でトラックがフレームアウトする．映像の全フレームに運動物体が存在している．映像の一部フレーム画像を図 6 に示し，削除処理前の累積コスト 2 次元画像を見やすさのために  $t$  軸方向に 6 倍したものを図 7 に示す．そして，処理結果映像の一部のフレーム画像を図 8 に示す．

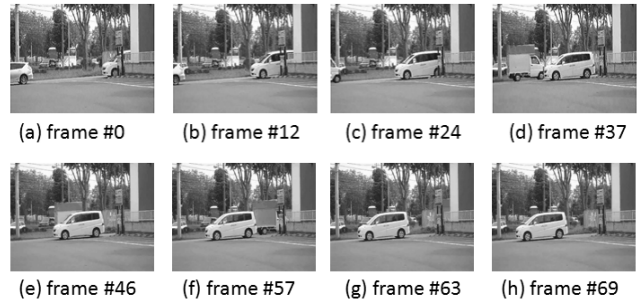


図 6: 入力映像のフレーム画像

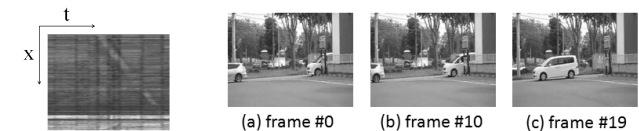


図 7: 累積コスト 図 8: 実験 2 の処理結果

実験の結果，運動物体のトラックが削除された．図 7 からトラックの運動領域を示すエッジを検出できたことがわかる．そのため，処理結果としてトラックのみが削除された映像を生成することができたと考えられる．入力映像中で運動物体が多くなると，エッジを正確に検出することが困難になり，処理結果に影響すると思われる．また，運動物体の速度も処理結果に関係する．

## 5 まとめ

本研究では，Ribbon Carving に様々な評価関数を用いて，運動物体領域を削除する手法を提案した．評価関数を変えて処理を行うことで，削除領域の異なる短縮映像を生成することが可能であるとわかった．処理に用いる評価関数をさらに検討していき，映像に与える特殊効果を調査する必要がある．今後の課題として，今回の実験では映像中の物体が水平方向のみに動くものを対象としたため horizontal ribbon のみを用いて短縮処理を行ったが，より複雑な動作をする物体映像の短縮や効果を与えるためにも様々な形式の映像について実験を行うことが挙げられる．

## 参考文献

- [1] M.Rubinstein, A.Shamir and S.Avidan, "Improved Seam Carving for Video Retargetting", ACM SIGGRAPH, vol.27, no.3, 2008.
- [2] B.Chen and P.Sen, "Video Carving", EUROGRAPHICS'08, 2008.
- [3] Z.Li, P.Ishwar, and J.Konrad, "Video Condensation by Ribbon Carving", IEEE Trans.on Image Processing, vol.18, no.11, pp.2572-2583, 2009.
- [4] S.Avidan and A.shamir, "Seam Carving for Content-Aware Image Resizing", ACM Trans.Graph, vol.26, no.3, 2007