

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 6 月 12 日現在

機関番号：32503

研究種目：基盤研究（C）

研究期間：2010～2012

課題番号：22500166

研究課題名（和文） 注目領域を考慮した動画像符号化向け画面アスペクト比変換処理の研究

研究課題名（英文） Content-aware non-linear video resizing based on motion saliency

研究代表者

八島 由幸（YASHIMA YOSHIYUKI）

千葉工業大学・情報科学部・教授

研究者番号：60550689

研究成果の概要（和文）：テレビ，パソコン，タブレット端末，スマートフォンなど映像視聴端末の画面アスペクト比が多様化している．本研究では，様々な表示環境を持つ映像端末に対して効果的に画像を提供できる技術を開発した．本技術を使うことにより，画像の中から人間が注目する領域を自動検出し，顕著領域の大きさをできるだけ維持したまま画面の縦横比のみを変更することができる．本手法は，長時間動画像の効率的なブラウジングにも応用でき，定められた時間内にビデオ内の重要箇所だけを視聴することも可能になる．

研究成果の概要（英文）：Viewing aspect ratio of image/video display terminals such as television, PC, tablet and mobile phone has been more and more diversified. In this study, we have developed a new image processing technology which provides an effective display environment under various image/video terminals. Our proposed method can automatically detect salient regions from the input image/video, and enables non-linear spatial scaling while keeping their size or motion. Moreover, it is applicable to an effective browsing for long video sequences, and we can watch only important scenes of a movie within pre-determined viewing time.

交付決定額

（金額単位：3900000 円）

	直接経費	間接経費	合計
2010 年度	1200000	360000	1560000
2011 年度	1000000	300000	1300000
2012 年度	800000	240000	1040000
年度			
年度			
総計	3000000	900000	3900000

研究分野：情報通信

科研費の分科・細目：

キーワード：画像処理，画像符号化，非線形縮小，画像通信

## 1. 研究開始当初の背景

画像情報は非圧縮のままであると極めて情報量が多いため，種々の観点から情報圧縮技術が検討されてきており，国際標準画像圧縮符号化方式として JPEG や MPEG が定められ，放送・通信・家電にわたる産業界で広く利用されている．特に動画圧縮標準である MPEG-4

AVC/H.264（以下 AVC/H.264）は多くの圧縮パラメータを最適化することで，非圧縮 1Gbps の HDTV 信号を画質をある程度維持したまま 10Mbps 以下，すなわち 1/100 を超える圧縮率を達成することが可能になっている．近年の画像信号の高精細化や視聴環境の多様化に伴い，今後の画像圧縮への課題／要求条件と

して以下の3点が必要であるとされていた。

#### (1)さらなる高圧縮化

HDTVの16倍の画素数を持つスーパーハイビジョンに代表されるように、画像は超高精細化する一方で、AVC/H.264を上回る超高圧縮符号化が必要となる。

#### (2)新表現形態映像への対応

多視点カメラ映像や、毎秒60フレームを超える高フレームレート映像など新しい画像表現形態が登場してきており、これらに適した圧縮符号化技術が必要となる。

#### (3)画像視聴環境や個人嗜好の多様化

多様なネットワーク/多様な端末でもストレス無く視聴が可能なスケーラビリティと効率的な映像提示が必要となる。

このうち、(1)に関しては、国際標準化機関ITU-TおよびISO/IECにおいてHEVC(High Efficiency Video Coding)という形で検討が進んだ。(2)については自由視点テレビジョンFTVという形での国際標準化が進んでいる。一方、(3)の視聴環境適応型の映像圧縮符号化に関しては、国際標準の枠組みにおいては画像サイズのスケーラビリティを実現するSVC(スケーラブル符号化)が標準化されているものの、表示画面の縦横アスペクト比まで踏み込んだ、より柔軟な再生環境に適合する符号化技術の検討は十分行われていなかった。本研究は本課題に焦点を当て、種々の縦横アスペクト比を持つ再生環境に適した画像符号化を行う際の画像リサイズ技術を検討する。

## 2. 研究の目的

HDTV/SDTV/携帯のように画面サイズの異なる端末でも再生表示が可能な技術として、H.264/SVCに代表される空間スケーラブル符号化がある。従来は、入力となる原画像の縦横アスペクト比を保持したまま画面サイズを縮小する線形リサイズを基本としたスケーラビリティを実現することにどまっていた。これに対して本研究では、原画と異なる縦横アスペクト比での再生環境や、動画像に対して原画像と異なる時間再生環境まで意識し、時空間にわたる画像の非線形リサイズ技術を検討し、画像符号化への応用も図る。具体的には、画像処理技術の一つであるSeam carving(以下SC)やRibbon carving(以下RC)を応用する。これらは画像中の興味領域は保持しつつ、それ以外の領域を消去することで画像の空間方向および時間方向サイズを非線形変更するものである。

SCやRCを画像符号化に適用するためには、まずSC/RCそのものの処理性能を向上する必要がある。従来の検討では、移管問題が指摘されていた。

- ・画像によっては視覚的に重要な領域が不自然に削除されてしまう

- ・処理時間が膨大になり実用に供することが困難

- ・動画像への適用手法が明確化されていない。これは、SCがフレーム単位の処理を基本としていることから時間方向の連続性を考慮できないことに起因している。

本研究では、SC/RCを動画にも柔軟に適用できるような手法を検討し、時空間構造を持つ動画像から重要な被写体や領域のみを切り出して所望の縦横比画面(画面サイズのスケーラビリティ)/視聴時間(時間方向のスケーラビリティ)に違和感なく収めることを考える。さらに、非線形リサイズ手法を画像符号化に適用するアルゴリズムを提案する。本研究による成果により、JPEG, AVC/H.264, HEVCなど様々な画像圧縮技術との組み合わせで様々な視聴環境における柔軟な映像視聴が可能となり、産業分野における適用領域が拡大すると考えられる。

## 3. 研究の方法

研究の進め方としては、(1)静止画像に対するSeam carvingの特性の把握と視覚特性を利用した画質改善、(2)Ribbon carvingによる動画像の時間軸圧縮、(3)動画像に対する非線形リサイズにおける時間連続性保持技術の検討、(4)非線形リサイズ技術の画像符号化への適用、の4つのポイントから行った。基本的には、考案技術をCまたはC++言語にてプログラミングし、画像処理用高速ワークステーションにてシミュレーションを行い、処理画像を定量評価および主観評価する形で検討を進めた。

## 4. 研究成果

- (1)静止画像に対するSeam carvingの特性の把握と視覚特性を利用した画質改善

静止画像に対しては、画像の重要領域を推定し、重要領域を避ける形で非線形リサイズをするSeam carvingが提案されている。この手法は、重要でないと判断された画素を結んで出来る画像中の上端から下端まで(あるいは左端から右端まで)の一本の曲線(seam)を削除することで水平または垂直方向の1画素縮小を図り、これを所望の縦横アスペクト比になるまで繰り返すことで非線形リサイズを可能にする。

従来のSeam carvingでは、画素重要度として輝度信号のエッジ勾配を基本としたパラメータを利用するが、この手法だと、エッジ(勾配)の強さが必ずしも視覚的注視領域(興味領域)ではないといった理由により、非線形リサイズ画像が不自然になる場合があることが指摘されている。これを解決する

ために、本検討では、より直接的に視覚的注視領域を定量化できる顕著性マップ (Saliency map) を用いた。Saliency map は画像中で注意を引く領域を検出することが可能であるため、seam carving を行う際の画素重要度マップとして Saliency map を適用すれば、視覚的に良好なリサイズ処理が可能になる。今回、エッジ勾配と顕著性マップを適度な割合で重み付き平均することで新たな重要度マップとすれば、従来の問題を解決して自然な非線形リサイズ結果が得られることを明らかにした (図 1)。

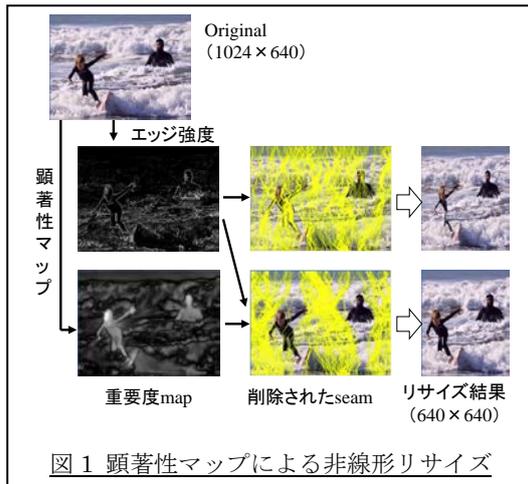


図 1 顕著性マップによる非線形リサイズ

## (2) Ribbon carving による動画の時間軸圧縮

監視ビデオや定点カメラの普及に伴い、長時間映像において重要な動オブジェクトを効率よくブラウザできる技術が重要である。これに対し、時間軸方向への content-aware な時間軸圧縮手法として ribbon carving が提案されている。Ribbon carving は空間方向の非線形リサイズ技術である seam carving を時間方向へ拡張したものと位置づけられる。Ribbon の算出には画素重要度最小パス検出に dynamic programming (DP) が求められるが、画像サイズが大きくと演算量が增大する欠点があった。本検討では、長時間動画映像を扱う際に不可欠となる処理時間の短縮化を実現する手法を提案・評価した。

まず、「線形フレーム縮小動画重要画素マップ (key-pixels map)」と「ribbon の空間拡大」を利用して時間軸非線形圧縮を行う手法を検討した。注目領域マップとしては背景差分を基本とする方式を用いた。さらに「時間軸線形縮小 (フレーム間引き) 画素重要度マップ」と「Ribbon の時空間拡大」を新たに導入して、さらなる高速化を図った。提案手法についてシミュレーション実験を行ったところ、最終的な処理画像の画質に影響を与えない縮小率としては、空間縮小方式では

縮小率 1/2, 時間縮小 (フレーム間引き) 法引きでは縮小率 1/6 が選択可能であり、その際の処理時間は、原画サイズで処理する従来方式に比べそれぞれ約 1/3, 約 1/6 にすることが可能であることがわかった (図 2(d))。

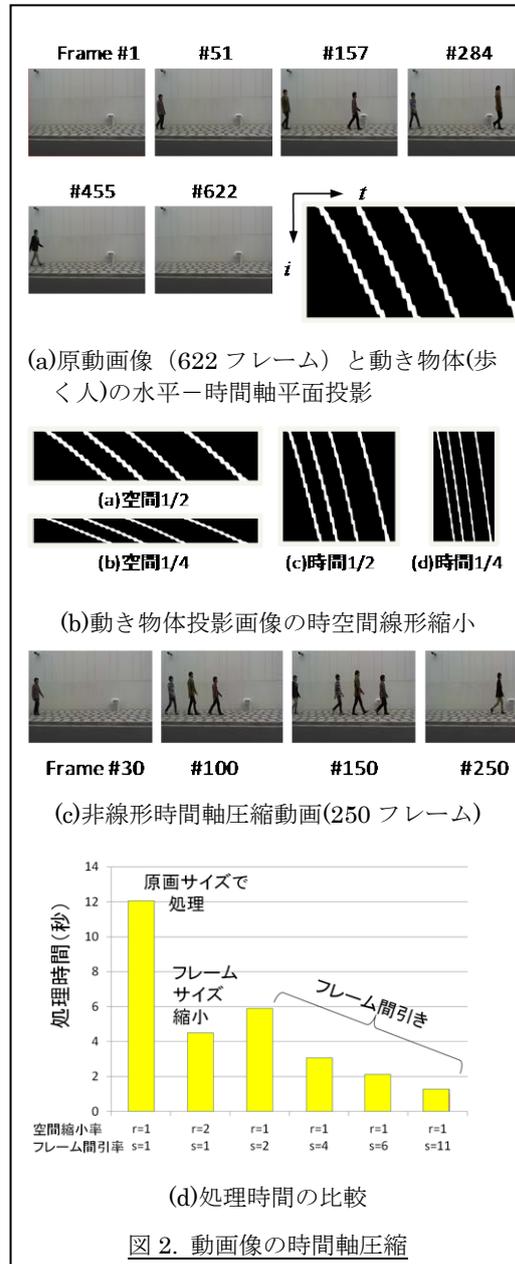


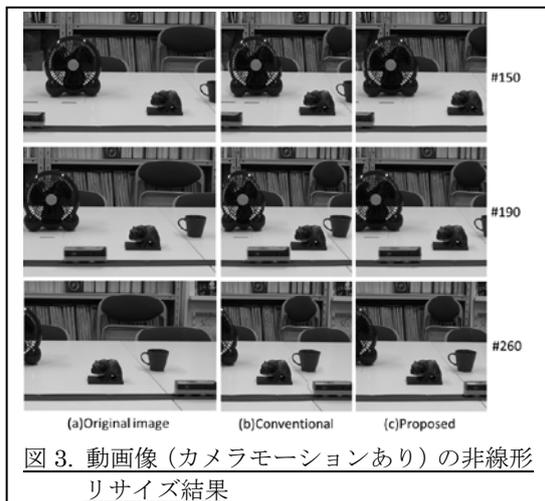
図 2. 動画の時間軸圧縮

## (3) 動画に対する非線形リサイズ

動画の画面アスペクトリサイズ向けの手法を検討した。動画の各フレームに対して、静止画で行われる Seam carving を適用すると、重要領域が時間とともに移動したり変形したりすることから、seam の時間的連続性が保たれず、視覚的に大きな違和感が発生する。

この問題に対し、まずカメラ静止の場合について、動的トリミング手法を導入した実験プログラムを作成し動作検証した。これは、ターゲットとする画面アスペクトサイズのトリミング枠を設け、そのトリミング枠の大きさを画面内の動き物体を追跡して切り出すものである。検討した動的トリミングでは、背景差分法による顕著性領域（動物体）の抽出、トリミング領域の自動決定、領域の適応的な拡大縮小、の3つを組み合わせることで良好な結果を得た。

次に、これらをカメラモーションが伴う動画像に発展させ検討を行った。カメラ操作が伴う動画像に対してもフレームごとに静止画対応のリサイズ処理を行う従来手法では、背景に大きなジャギー（画面のゆらぎ）が発生し視覚的に極めて大きな妨害となる。今回、カメラパラメータや被写体の動きを基にして、時空間動画像から一枚の大きなパノラマ静止画背景を再構成し、この背景画像を基に seam carving を適用する技術を新たに提案した。前景動オブジェクトは背景スプライト生成時のカメラパラメータとの整合性を取りながら最後に画面合成する。実際の動画像を用いた評価実験により、本手法は極めて自然な動画像非線形リサイズが実現できることが確認できた。図3は実験結果の一例であり、原動画像(a)に対して従来の非線形リサイズ結果(b)、本研究の提案手法による非線形リサイズ結果(c)を示している（図は上から順に150フレーム目、190フレーム目、260フレーム目を示す）。

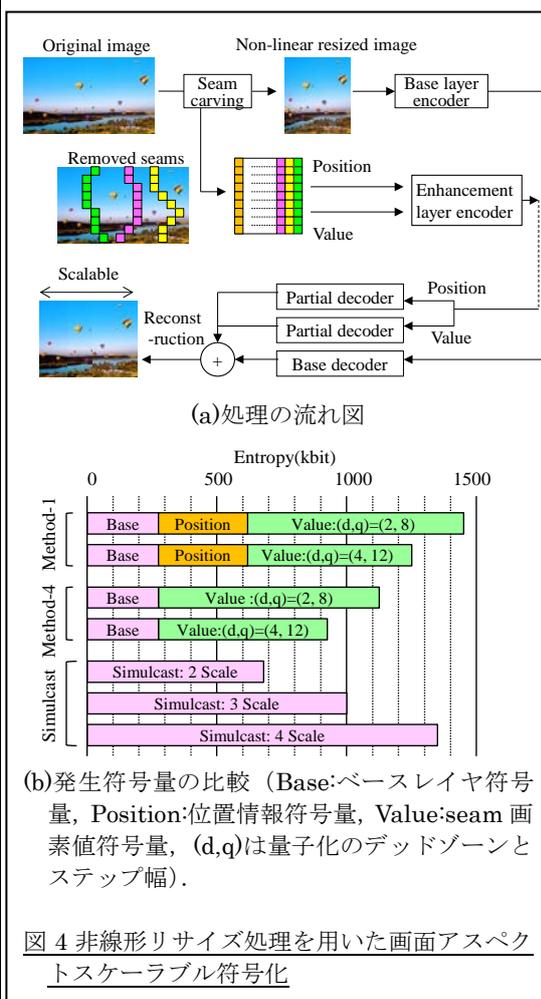


#### (4) 非線形リサイズ技術の画像符号化応用

非線形リサイズ技術の映像符号化への応用に関して検討を行った。非線形リサイズされた画像をベースレイヤ、リサイズの際に削除された seam を拡張レイヤとするスケーラブル符号化を実現した。ベースレイヤを通常

の JPEG または H.264/AVC で符号化し、拡張レイヤは seam の位置情報を算術符号化で、画素値情報を1次元 DPCM で符号化する。従来は、位置情報は必ず符号化伝送していたが、今回の検討では seam の画素値情報だけを符号化伝送し、復号側においてその位置情報を推定する手法を新たに考案した。

提案方式に対し、シミュレーション実験を行った（図4）。位置情報・画素値情報の両者を符号化する場合に比べて、画質を維持したまま符号量を約25%削減できることが検証できた。図4においてMethod-1は従来方式、Method-2は提案方式を示す。また、「Simulcast」はそれぞれ別々に符号化ストリームを生成する方式であり、提案方式は3種類以上の画面アスペクト比を実現する場合に Simulcast よりも有利であることを示している。さらに、本研究で提案した方式において、復号画像に細部の画素構造が崩れる（直線性の保持の不自然さ）場合があることがわかり、この問題点に対しても、seam ごとに位置情報と画素値情報を適応的に切り替えて符号化するハイブリッド方式を用いれば、符号量が若干増えるものの高画質を維持できることがわかった。



以上の検討により、様々な画面アスペクト比を有する端末に効率的に画像配信することが可能となり、今後の画像配信システムの設計やサービス普及に大きな役割を果たすと考えられる。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 3 件)

- ① 八島由幸, 注目領域を考慮した動画像符号化向け画面アスペクト比変換処理の研究, 2011 年度千葉工業大学プロジェクト研究年報, 査読無, Vol.8, 2011, pp. 145-146.
- ② 八島由幸, 注目領域を考慮した動画像符号化向け画面アスペクト比変換処理の研究, 2012 年度千葉工業大学プロジェクト研究年報, 査読無, Vol.9, 2012, pp. 141-142.
- ③ 八島由幸, Content-aware non-linear video resizing based on motion saliency, 2013 年度千葉工業大学プロジェクト研究年報, 査読無, Vol.10, 2013, pp. 145-146.

[学会発表] (計 4 件)

- ① N.Hirano, Y.Yashima, A Fast Ribbon Carving Method for Video Condensation, Workshop on Picture Coding and Image Processing (in IEEE PCS 2010), 2010 年 12 月 7 日, ウィンクあいち(愛知県).
- ② 平野那由他, 八島由幸, 高速 Ribbon Carving を用いた動画像時間軸圧縮に関する一検討, 電子情報通信学会画像工学研究会, 2011 年 2 月 22 日, 北海道大学(北海道).
- ③ N.Hirano, Y.Yashima, M.Takagi, H.Fujii and A.Shimizu, Viewing Aspect Ratio Scalable Image Coding using Adaptive Seam Transmission, 2013 International Workshop on Advanced Image Technology, 2013 年 1 月 07 日, 名古屋大学(愛知県).
- ④ 白鳥晃啓, 八島由幸, 背景スプライトを用いた動画像の非線形リサイズ手法の一検討, 電子情報通信学会画像工学研究会, 2013 年 3 月 11 日, 福岡工業大学(福岡県).

[その他]

非線形リサイズ結果については以下の Web ページに動画像として掲載している。

<http://www.vpc.net.it-chiba.ac.jp/research-v2.html>

#### 6. 研究組織

##### (1) 研究代表者

八島 由幸 (YASHIMA YOSHIYUKI)  
千葉工業大学・情報科学部・教授  
研究者番号: 60550689

##### (2) 研究分担者

( )

研究者番号:

##### (3) 連携研究者

( )

研究者番号: