

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 2 日現在

機関番号：12605
研究種目：若手研究(B)
研究期間：2012～2014
課題番号：24760288
研究課題名(和文) コンテンツ考慮型拡張を利用した静止・動画像符号化

研究課題名(英文) Image/video coding using content-aware resizing

研究代表者
田中 雄一 (Tanaka, Yuichi)
東京農工大学・工学(系)研究科(研究院)・准教授

研究者番号：10547029
交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：映像符号化とコンテンツ考慮型拡張を統合した ROI (region of interest) 符号化を実現した。コンテンツ考慮型拡張とはコンピュータグラフィックスで積極的に研究が行われている画像のリサイズ手法であり、自動で着目領域 (ROI) を抽出し、リサイズする手法である。本研究では研究代表者が進めているコンテンツ考慮型拡張を用いた画像符号化と、その動画像符号化への拡張を目的とした。結果として、従来のROIを考慮していない符号化と比較し、10%程度のビットレート増加でコンテンツ考慮型拡張を統合可能であることを示した。更に、本手法を他の画像処理へと応用し性能の向上を達成した。

研究成果の概要(英文)：We realized ROI (region of interest) image/video coding using content-aware resizing. Content-aware resizing is an image resizing method which has emerged in computer graphics. It automatically extracts ROI from an image and resizes it so that the resulting image preserves ROI as large as possible. We integrated it into image coding, and extended the method to video coding. In experimental results, content-aware image/video coding was realized with a small amount of overhead bitrates from image/video coding without considering ROI. We further extended our approach to other image processing applications.

研究分野：信号処理

キーワード：映像符号化 ウェーブレット・フィルタバンク リターゲティング

1. 研究開始当初の背景

静止・動画像符号化標準である JPEG2000 [1] や H.264/AVC [2] では、ROI (region of interest) 符号化を利用可能である。ROI 符号化では、ユーザが指定した領域の符号化ビットレートを上げることにより、content-aware、すなわち画像中の注目領域を考慮した符号化を実現している。しかしながら、個人が所有するデジタル画像の数は爆発的に増加しており、その一つ一つにユーザが注目領域を指定し ROI 符号化を行うのは現実的ではない。

他方、コンピュータグラフィックス (CG) において近年注目を集めているのが content-aware resizing (CAR) である。CAR とは、端的に言えば自動で ROI を抽出し、画像をリサイズする技術である。本分野が注目を集める嚆矢となった技術の一つが seam carving [3] である。Seam carving においては、seam と呼ばれる、画像の上端 (左端) から下端 (右端) へと繋がる 1 画素幅の画素群を繰り返し取り除くことにより、画像のリサイズが行われる。一般に seam はオブジェクトの輪郭やテクスチャ部などの画像の局所的なエネルギーが高い部分を避けるように配置されるため、結果として seam carving によって CAR が行われることとなる。図 1 に示すように、線形的にサイズ縮小を行うスケールリングと異なり、seam carving ではオブジェクト領域が原画像により近い構造及びサイズになっていることがわかる。

研究代表者 (以後代表者) らは、CAR を利用した画像符号化を提唱している。本手法においては、画像のエネルギー関数によって ROI が自動的に抽出され、それが ROI 画像として符号化される。また、ROI 外の領域 (non-ROI) は復号側において近隣の ROI から補間が行われる。

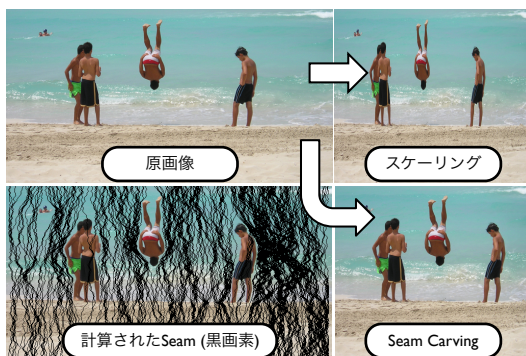


図 1 Seam carving の例

2. 研究の目的

CAR に基づく画像符号化では、従来の画像符号化と比較した場合、同一情報量で ROI の視覚的な画質向上が可能となる。もちろん、ユーザが特に指定すれば ROI を手動で決定することも可能である。そのため、本研究で提案する考え方は、ROI に対してより一般性がある画像符号化手法であると考えられる。

CAR を利用した画像符号化は、以下の点において更なる発展を遂げる必要がある。

- (1) 原解像度の画像で従来の符号化手法と比肩する性能が達成できるようになること
- (2) 静止画像符号化だけでなく、動画像においても CAR を用いた符号化を実現すること

研究開始当初時では、ROI においては優れた画質を実現できていた。しかしながら、ROI を考慮していない符号化と比較すると、non-ROI の画質劣化が、高ビットレートでの符号化の場合顕著であることがわかっていった。また、現在までに代表者らが提案している CAR 符号化は主に静止画像に対しての理論であったが、CAR 符号化は動画像にも同様に有効であると考えられる。しかしながら、動画像の性質は静止画像と異なるため、適切な動画像 CAR 手法及び符号化技術を新たに提案する必要があった。結局、本研究では以下を目的として研究を行った。

- (1) non-ROI の画素値を効率的に符号化する手法を確立すること

本研究では、復号側においてただ単に ROI からの補間処理で non-ROI を再構成するのではなく、non-ROI に適した符号化手法を実現し、原解像度において、ROI を考慮していない符号化と同等 (あるいはそれ以上) の画像符号化性能を達成することを目標とする。

- (2) 動画像の CAR 符号化における、効果的な CAR 手法及びそれに基づく符号化システムの実現

動画像は静止画像と異なり、フレーム間での動きが存在するため、静止画像の CAR をそのまま利用することは適切でない。また、CG 分野における動画像の CAR 手法 [4] は非常に計算量が多いため、高速な動画像の CAR 手法の提案と、それを用いた CAR 動画像符号化システムを実現する。

3. 研究の方法

- (1) コンテンツ考慮型拡張 (content-aware resizing: CAR) を用いた画像符号化において、ROI (Region-Of-Interest) 外の領域 (non-ROI) を効率的に符号化する手法に関して主に研究を行った。具体的には以下の 2 点のアプローチを用いた。

- ① CAR を行う際、我々がこれまで提案してきた rate-dependent seam carving を利用することで CAR の性能を保ちつつ、必要なサイド情報 (受信側へと伝送される) の符号量を抑えた。

- ② CAR によって取り除かれた領域 (すなわち、non-ROI) の画素を効率的に伝送するため、それらの画素を wavelet ハイパスフィルタにより変換した。更に、その高周波数成分と、ROI を wavelet 変換して得られた多重解像度画像を統合した親子関係を構築し、画像符号化を行った。その際、著名な wavelet に基づく画像符号化手法である SPIHT を利用し、その親子関係の一

般化として画像符号化手法を実現した。

(2) コンテンツ考慮型拡張 (content-aware resizing: CAR) を更に広範囲の画像処理へと応用するための理論とアルゴリズムに関して検討を行った。具体的には、画像の多重解像度表現手法の中で広く用いられている、ラプラシアンピラミッド (Laplacian pyramid: LP) における、低域通過フィルタとダウンサンプリング処理を CAR で置換した、retargeting pyramid (RP) を提案した。RP は LP と異なり、画像を均一に縮小するのではなく、CAR を用いて画像中の重要な部分は拡張し、重要でない部分は大きく縮小される。その結果、RP を利用したピラミッド表現では、上位の階層 (LP における低周波数成分) の画像においても、重要な部分のサイズと形状が保存される。



図 1 画像符号化性能の比較。(上) 通常符号化 (PSNR: 40.91 dB) (中) 従来の CAR 符号化 (PSNR: 38.40 dB) (下) 提案手法 (PSNR: 40.41 dB)

(3) 上記 2 種の研究の予想以上の進展から着想を得て、現在信号処理分野で大きな注目を集めているグラフ信号処理に関する研究も行った。グラフ信号処理は複雑・大規模なデータ (信号) に対する信号処理手法であり、特に画像に関しては、RP で用いられているようなメッシュの変形手法と深いつながりがある。本研究では、グラフ信号処理の基礎的検討として、グラフ信号をグラフ周波数帯域ごとへ分割するグラフフィルタバンクの研究を行った。更に、画像処理への応用として、強力な画像の平滑化手法である Trilateral filter をグラフ信号処理として定式化することに取り組んだ。

4. 研究成果

(1) コンテンツ考慮型拡張を用いた画像符

号化

従来符号化手法 (すなわち、ROI を考慮していない手法) と比較し、概ね 10% 程度のビットレート増加で CAR 機能が統合できることを示した。また、CAR に関しては、サイド情報の符号化を考慮していない CAR 手法である seam carving と遜色ない性能を示しつつ、必要なサイド情報の符号量を数十分の一〜数百分の一へ減少できることも明らかになった。また、動画像においても CAR 符号化を実現した。図 1 に、提案手法と従来手法とで画像符号化を行った画像を示す。

本成果は論文誌 APSIPA Transactions on Signal and Information Processing を始めとした複数の英文誌において採録されており、更に IEEE Signal Processing Society のフラッグシップ国際会議である ICASSP, ICIP においても複数の成果を発表した。

(2) Retargeting Pyramid

提案手法の性能の確認のため、方向性変換である contourlet transform (LP と方向性フィルタバンクを利用した 2 次元信号変換手法) へ RP を統合し、画像ノイズ除去シミュレーションを行った。結果として、ピーク信号対雑音比 (PSNR) において大幅な性能向上を果たした。更に、RP において、バイラテラルフィルタと Tikhonov 正則化を組み合わせることにより、ノイズ除去性能が上記 RP と比較して PSNR で 1dB 以上向上するという成果を得た。また、本手法を画像の強調・拡大に用いたところ、従来法と比較して大きな性能の向上を得た。成果は国際会議 ICASSP, ICIP に採録された。

(3) グラフ信号処理

従来のグラフ信号処理では実現されていなかった、冗長型グラフフィルタバンクの設計法を理論的に示した。本手法は冗長度をコントロール可能であり、グラフ周波数分解能と変換の低冗長度の両立が達成可能となる。

また、Trilateral filter においては、グラフ信号のノイズ除去に対し、図 2 のような大きな性能向上を達成した。

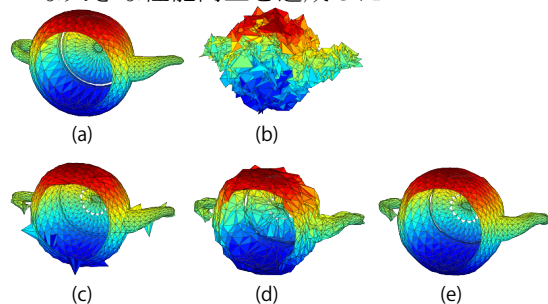


図 2 グラフ信号処理を利用した trilateral filter による 3 次元メッシュの平滑化。(a) 原信号 (b) 観測信号 (SNR: 17.7 dB) (c) Spectral graph wavelet による平滑化 (SNR: 26.42 dB) (d) グラフ信号処理を利用した bilateral filter による平滑化 (SNR:

26.07 dB) (e) 提案手法による平滑化 (SNR: 28.33 dB)

<引用文献>

- [1] B. E. Usevitch, "A tutorial on modern lossy wavelet image compression: foundations of JPEG 2000," *IEEE Signal Process. Mag.*, vol. 18, no. 5, pp. 22-35, 2001.
- [2] T. Wiegand, et al., "Overview of the H. 264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560-576, 2003.
- [3] S. Avidan and A. Shamir, "Seam carving for content-aware image resizing," *ACM Trans. Graph.*, vol. 26, no. 3, 2007.
- [4] L. Wolf, et al., "Non-homogeneous content-driven video-retargeting," in *Proc. ICCV' 07*, 2007.

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 4 件)

- ① Y. Tanaka and A. Sakiyama, "M-channel oversampled graph filter banks," *IEEE Trans. on Signal Processing*, vol. 62, no. 14, pp. 3578-3590, 2014.
DOI: 10.1109/TSP.2014.2328983
- ② V. H. Le, K. Shinoda, M. Hasegawa, S. Kato, and Y. Tanaka, "Modified high-order edge-directed interpolation for distributed video coding based on selective data pruning," *IEEE Transactions on Image Electronics and Visual Computing*, vol. 2, no. 1, pp. 26-36, 2014.
DOI:なし
- ③ Y. Tanaka, T. Yoshida, M. Hasegawa, S. Kato, and M. Ikehara, "Rate-dependent seam carving and its application to content-aware image coding," *APSIPA Trans. on Signal and Information Processing*, vol. 2, no. e1, 15 pages, 2013.
DOI: 10.1017/ATSIP.2013.3
- ④ Y. Tanaka, M. Hasegawa, and S. Kato, "Flexible combination of time-domain lapped transforms with various downsampling factors," *IEICE Trans. on Fundamentals*, vol. E95-A, no. 11, pp. 2049-2058, 2012.
DOI: 10.1587/transfun.E95.A.2049

[学会発表] (計 21 件)

- ① 森田亮佑, 田中雄一, "Performance analysis of retargeting pyramid and its applications," 電子情報通信学会 信号処理研究会, 2015年03月02日~2015年03月03日, ホテルミヤヒラ (石垣市・沖縄県)
- ② A. Hamamoto, M. Onuki, and Y. Tanaka, "Multidimensional linear-phase perfect reconstruction filter banks with higher order feasible building blocks," *ICIP 2014*, 2014年10月27日~2014年10月30日, パリ (フランス)
- ③ M. Onuki and Y. Tanaka, "Trilateral filter on graph spectral domain," *ICIP 2014*, 2014年10月27日~2014年10月30日, パリ (フランス)
- ④ A. Sakiyama and Y. Tanaka, "Oversampled graph Laplacian matrix for graph signals," *EUSIPCO 2014*, 2014年09月01日~2014年09月05日, リスボン (ポルトガル)
- ⑤ Y. Tanaka and A. Sakiyama, "M-channel oversampled perfect reconstruction filter banks for graph signals," *ICASSP 2014*, 2014年05月04日~2014年05月09日, フィレンツェ (イタリア)
- ⑥ R. Morita, K. Shirai, and Y. Tanaka, "Retargeting pyramid using direct decimation," *ICASSP 2014*, 2014年05月04日~2014年05月09日, フィレンツェ (イタリア)
- ⑦ 田中雄一, "グラフ信号処理のすゝめ", 第28回信号処理シンポジウム, 2013年11月19日~11月22日, 海峡メッセ下関 (山口県下関市) (招待講演)
- ⑧ M. Onuki and Y. Tanaka, "Design of optimized prefilters for time-domain lapped transforms with various downsampling factors," *ICIP 2013*, 2013年09月15日~2013年09月18日, メルボルン (オーストラリア)
- ⑨ H. V. Le, K. Shinoda, M. Hasegawa, S. Kato, and Y. Tanaka, "Selective data pruning based distributed video coding with modified high-order edge-directed interpolation," *ICIP 2013*, 2013年09月15日~2013年09月18日, メルボルン (オーストラリア)
- ⑩ Y. Tanaka and K. Shirai, "Scalable

image representation using improved retargeting pyramid,” ICASSP 2013, 2013年05月26日～2013年05月31日,バンクーバー (カナダ)

- ⑪ Y. Tanaka and K. Shirai, “Directional image decomposition using retargeting pyramid,” APSIPA Annual Summit and Conference 2012, 2012年12月03日～2012年12月06日, ハリウッド (米国)
- ⑫ T. Yoshida, Y. Tanaka, M. Hasegawa, S. Kato, and M. Ikehara, “Wavelet-based content-aware image coding with rate-dependent seam carving,” ICIP 2012, 2012年09月30日～2012年10月03日, オランダ (米国)

[図書] (計 0件)

[産業財産権]

○出願状況 (計 0件)

○取得状況 (計 0件)

[その他]

ホームページ等

<http://tanaka.msp-lab.org/>

6. 研究組織

(1) 研究代表者

田中 雄一 (TANAKA, Yuichi)

東京農工大学・大学院工学研究院・准教授

研究者番号：10547029

(2) 研究分担者

なし

(3) 連携研究者

なし