

レート依存シームカービングを用いた content-aware 画像符号化

Content-aware image coding using rate-dependent seam carving



田中 雄一 (Yuichi TANAKA, Ph. D.)

東京農工大学大学院 生物システム応用科学府 准教授
(Associate Professor at Graduate School of BASE, Tokyo University of Agriculture and Technology)

IEEE 電子情報通信学会 画像電子学会

受賞：丹羽保次郎記念論文賞(2010) 電気通信普及財団賞 テレコムシステム技術賞 奨励賞(2011) APSIPAASC 2014 Best Paper Award(2014)

研究専門分野：多次元マルチメディア信号処理：特にグラフ信号処理 ウェーブレット・フィルタバンク 映像符号化

あらまし 本研究では、ビットレート依存シームカービングを用いた content-aware 画像符号化を提案する。Content-aware 画像符号化とは、コンテンツ考慮型画像拡張、もしくはリターゲティングと呼ばれる画像サイズ変更手法を画像符号化（圧縮）と統合する試みである。本研究では特に、シームカービングと呼ばれるコンテンツ考慮型画像拡張手法に着目し、復号側へ伝送するサイド情報の削減およびウェーブレット変換に基づく画像符号化手法との融合を図った。結果として、従来符号化手法と比較し、ビットレート増加を 10%程度に抑制しながら、コンテンツ考慮型画像拡張が実現可能となった。

1. 研究の目的、狙い

コンピュータビジョン・コンピュータグラフィックス分野では、コンテンツ考慮型画像拡張 (content-aware image resizing : CAIR)、あるいは画像のリターゲティング (image retargeting) が注目を集めている [1]-[8]。これは「洗練された」画像のリサイズ手法の一種であると言えることができる。例えば

画像のアスペクト比を変更したり、複雑な構造を持つ画像をサイズ変更したりする場合を考えたとき、伝統的な手法である、均一な画像拡大・縮小（スケーリング）や、画像の切り抜き（クロッピング）などでは、ユーザが満足する画像が得られない場合が多い。一方、CAIR では画像内の特徴に応じ、適応的に画素のダウンサンプリングを行うことができる。画像処理や信号処理の観点から言えば、CAIR は注目領域 (region-of-interest : ROI) を抜き出すことを目的とした非均一な信号のダウンサンプリング手法の一種である。一方、CAIR は局所的・全体的な画像の構造を解析する必要があるため、一般に非常に計算量が多いことが知られている。そのため、モバイルデバイス等のソフトウェア上で CAIR 処理を行うためには、(CAIR 自体は特にモバイルデバイスで望まれているにも関わらず) 課題が多い。

シームカービング (seam carving) [1][2]は CAIR の嚆矢となった技術の一つである。シームカービングでは、画像を縦断（横断）する画素の集合としてシーム（縫い目）を定義し、シームを繰り返し取り除く（挿入する）ことで画像を縮小（拡大）している。また、他の手法と比較して、単純なアルゴリズムで CAIR が実現できることが知られている。図 1 にシームカービングによる画像縮小の例を示す。図から分かるように、シームカービングでは背景等の変動の少ない部分を（自動的に）抽出し、取り除いていることが分かる。

CAIR を画像符号化（圧縮）と統合する、すなわちコンテンツ考慮型画像符号化 (content-aware image coding : CAIC) [9][10][11]を実現することは、注目領域 (region-of-interest : ROI) を考慮した符号化 (ROI 符号化) の一種を実現することである。ROI 符号化自体は、JPEG2000 [12]や H.264/AVC [13]等の映像符号化標準で利用可能である。これらの手法では、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。また、通常 ROI は矩形で与える必要がある。そのため、CAIC の

レート依存シームカービングを用いた content-aware 画像符号化

Content-aware image coding using rate-dependent seam carving



図1 シームカービングの例 (Mt. Evans)

ように自動で非矩形の ROI を計算し、それを考慮に入れた符号化を実現することは、ROI に対してより一般性のある画像符号化手法の実現と言う観点から価値があると言える。

2. 研究の背景、同じ分野の研究に関する国内外の研究動向

CAIC には、現在までいくつかの手法が提案されている。例えば、原画像から変動の少ない列・行を取り除き、復号側で補間によって復元する、selective data-pruning (SDP) と呼ばれる手法が存在する[14]。SDP で取り除かれる部分は、補間した際に誤差が小さくなるような領域が選択される。SDP を利用した画像符号化手法は、低ビットレートにおいて通常の符号化を用いた手法と比較し、良好な画質が実現できる。しかしながら、画像から多くの列が取り除かれた場合には、補間処理で大きな誤差が生じる場合がある。また、補間処理で取り除いた部分を再構成しているため、高ビットレートでの符号化性能(圧縮率)は従来手法より劣ることが知られている。

Nguyen らは、シームカービング様の技術を利用した CAIC を提案した[9]。本手法では、まず原画像に

CAIR を適用し、縮小された「核」となる画像を JPEG 等の画像符号化標準で圧縮する。同時に、シーム経路情報(seam path information : SPI)、及びシーム上の画素値が符号化され、サイド情報として伝送される。復号側では、まず「核」画像が再構成される。もしユーザが「核」画像と異なる解像度の画像を所望している場合には、サイド情報が非 ROI 領域の再構成に用いられる。しかしながら、SPI 及びシーム上の画素の符号化には非常に多くのビットレートが必要となるため、原画像に対する符号化性能は通常の画像符号化と比較し、非常に低い。

筆者らは以前、柱状のシームからなる CAIC を提案した[10][11]。これは SDP が画像の 1 列(1 行)全体を取り除くのと異なり、短い柱状の画素を連結させた画素集合(シームの一種として考えられる)を取り除く手法である。復号側では、SDP と同様に補間によって画像を再構成する。本手法は、SDP と比較し良好な性能を示した。一方、シームの形状にそれほど自由度がないため、画像のサイズ変更手法として考えると、シームカービングほど良好な結果が得られなかった。動画像符号化においてもシームカービングを利用した手法が存在する[15]が、これも SDP と同様、補間によ

レート依存シームカービングを用いた content-aware 画像符号化

Content-aware image coding using rate-dependent seam carving

って画素を再構成しているため、原理的に、復号側で原画像が完全には復元できないという問題点があった。

また、ウェーブレット領域におけるシームカービングを利用した画像符号化手法も報告されている [16][17]。SPI を削減するため、本手法では縮小された画像（ウェーブレット領域の最低周波数成分）に対しシームカービングを行っている。しかしながら、空間領域では、本手法で提案しているシームはブロック状となるため、ウェーブレット変換のレベル数を多くすることができない（普通、レベル数が多くなれば符号化性能が向上する）。そのため、従来のウェーブレット変換と比較し、符号化性能がかなり劣化している。

すなわち、CAIC では以下の 2 点の問題点を解決する必要はある。

1. CAIR の性能を保ったまま、サイド情報の符号量を削減すること
2. 原サイズでの符号化性能を維持すること

3. 研究の方法、研究の結果

本研究では、上記 2 点の問題点を解決するため、ビットレート依存シームカービングと新しい親子関係*1を利用したウェーブレット画像符号化を提案する。簡単に言えば、本研究におけるシームカービングは、通常のシームを圧縮、すなわち非可逆符号化したものと捉えることができる。提案手法はシームの区分的近似によって実現され、ビットレートに応じて近似の粗さを調整することが可能である。また、非 ROI 画素値を効率的に符号化するため、ウェーブレット変換符号化で利用されている SPIHT[18]に対し、新しい親子関係を提案する。

3.1 区分的近似を利用したビットレート依存シームカービング

SPI の伝送に必要なデータ量を削減するために、本研究では、従来手法で計算されたシームを図 2 のようないくつかのテンプレートで近似する。復号側へ伝送するのは（符号化された）テンプレートのインデックスだけで良いため、結果として SPI の圧縮が実現できる。シームの選択は、原シームと近似シームの二乗誤差が小さくなるようなものを選択する。また、誤差が

ある閾値より大きくなる場合には、シームを再帰的に分割して近似することで、大きな誤差が発生しないようなアルゴリズムを提案した。

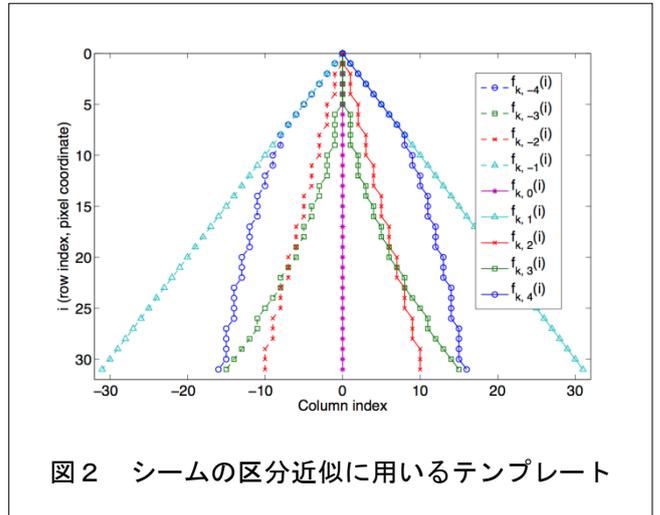
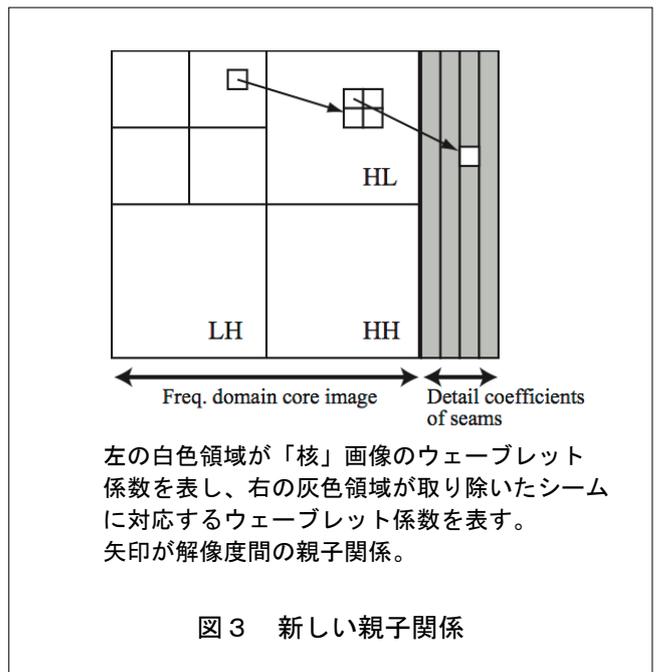


図 2 シームの区分近似に用いるテンプレート

3.2 ウェーブレット変換に基づく CAIC

ビットレート依存シームカービングを利用することで、SPI は効率的に圧縮可能となった。さらに取り除いたシームに対応する画素を符号化するため、「核」画像とシームに対しそれぞれウェーブレット変換を適用し、図 3 に示すような親子関係を構築する。



左の白色領域が「核」画像のウェーブレット係数を表し、右の灰色領域が取り除いたシームに対応するウェーブレット係数を表す。矢印が解像度間の親子関係。

図 3 新しい親子関係

レート依存シームカービングを用いた content-aware 画像符号化

Content-aware image coding using rate-dependent seam carving

さらに、親子関係を利用した画像符号化手法である SPIHT を CAIC のために改変した手法を提案することで、全体の CAIC システムを実現した。本符号化により、取り除いたシームであっても、ウェーブレット係数が大きい信号値に関しては符号化順位が高くなる。そのため、「核」画像だけでなく、任意の解像度における画質を改善することが可能となる。

3.3 結果

本稿では、テスト画像として図 4 に示すような Park Joy, Pisa Tower, Beach, Crew の 4 種類のグレースケール画像を用いた。

図 5 に SPI の伝送に必要なビットレートの比較を示す。提案手法はパラメータによって近似精度が変化するため、必要なビットレートがパラメータに応じて変化しているが、どのパラメータにおいても従来手法と

比較して数分の一から数十分の一程度のビットレートで SPI が符号化できていることが分かる。

また、図 6 に近似パラメータを変化させ、ビットレート依存シームカービングを行った結果を示す。見て分かる通り、パラメータが変化するに従って近似が粗くなっている（それに伴い必要ビットレートは減少している）。

図 7 に画像の縮小結果を示す。上述したように、画像のアスペクト比を変更した場合、スケーリングだと特に人物や建物に違和感が生じる。一方、シームカービングではアルゴリズムによって不要と判断された部分が取り除かれているため、それほど違和感がなく画像縮小が可能となっている。また、シームカービングと提案手法を比較すると、シーム経路を近似しているにも関わらず、視覚的にはほとんど同一の縮小画像が生成されていることが分かる。



Park Joy (704 × 1024 画素)



Pisa Tower (768 × 1024 画素)



Beach (512 × 640 画素)

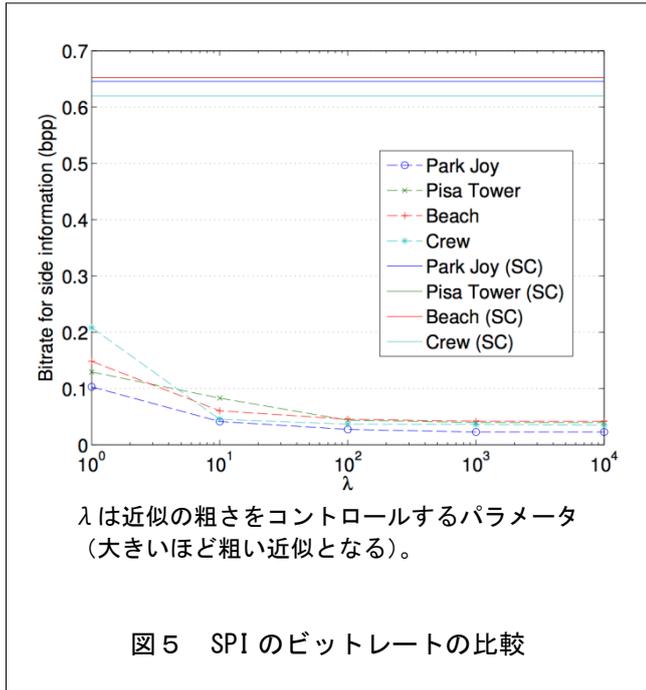


Crew (576 × 704 画素)

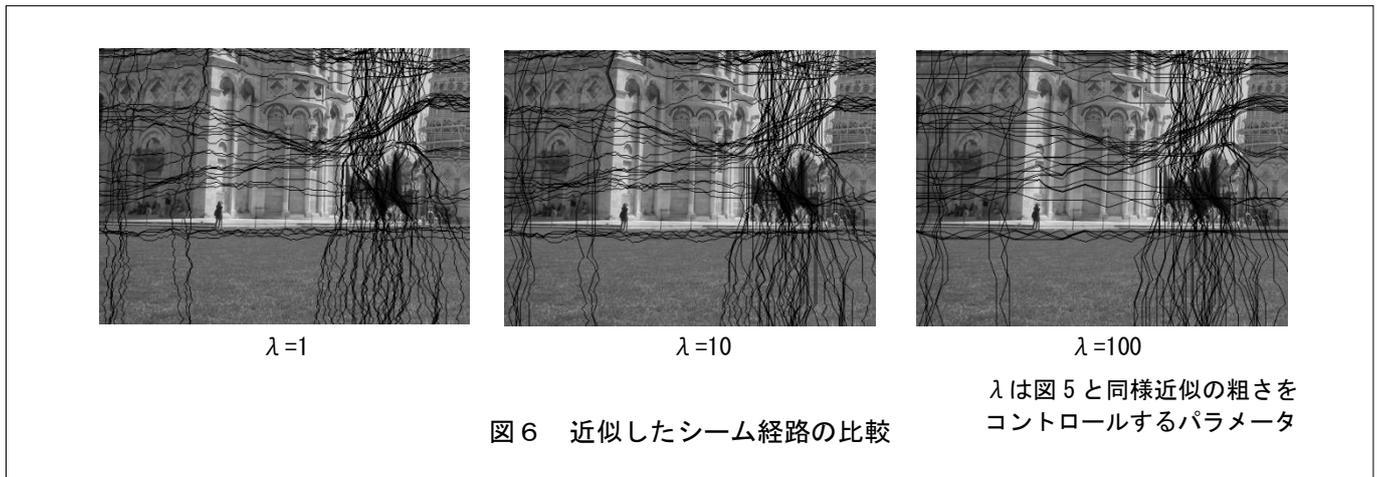
図 4 テスト画像

レート依存シームカービングを用いた content-aware 画像符号化

Content-aware image coding using rate-dependent seam carving



符号化した際の、原画像に対するレート歪曲線*2を 図 8 に示す。図において、SPIHT[18]は ROI を考慮していないウェーブレット画像符号化に、CD-SPIHT は文献[10][11]に、SC-SPIHT は文献[9]に、CA-SPIHT は提案手法に対応している。ROI を考慮していない従来符号化と比較し、提案符号化は同画質で 10%程度のビットレート増加で抑えられていることが分かる。一方、従来 CAIC では、特に高ビットレート (低圧縮率) で符号化した場合に、大きな画質の差が現れている。原サイズの復号画像を図 9 に示す。レート歪曲線と同様、視覚的にも提案 CAIC は通常のウェーブレット変換に基づく手法と遜色ない性能を発揮している。従来 CAIC は、補間によって非 ROI の画素を再構成しているため、不自然なアーティファクトが生成されていることが見て取れる。



レート依存シームカービングを用いた content-aware 画像符号化

Content-aware image coding using rate-dependent seam carving

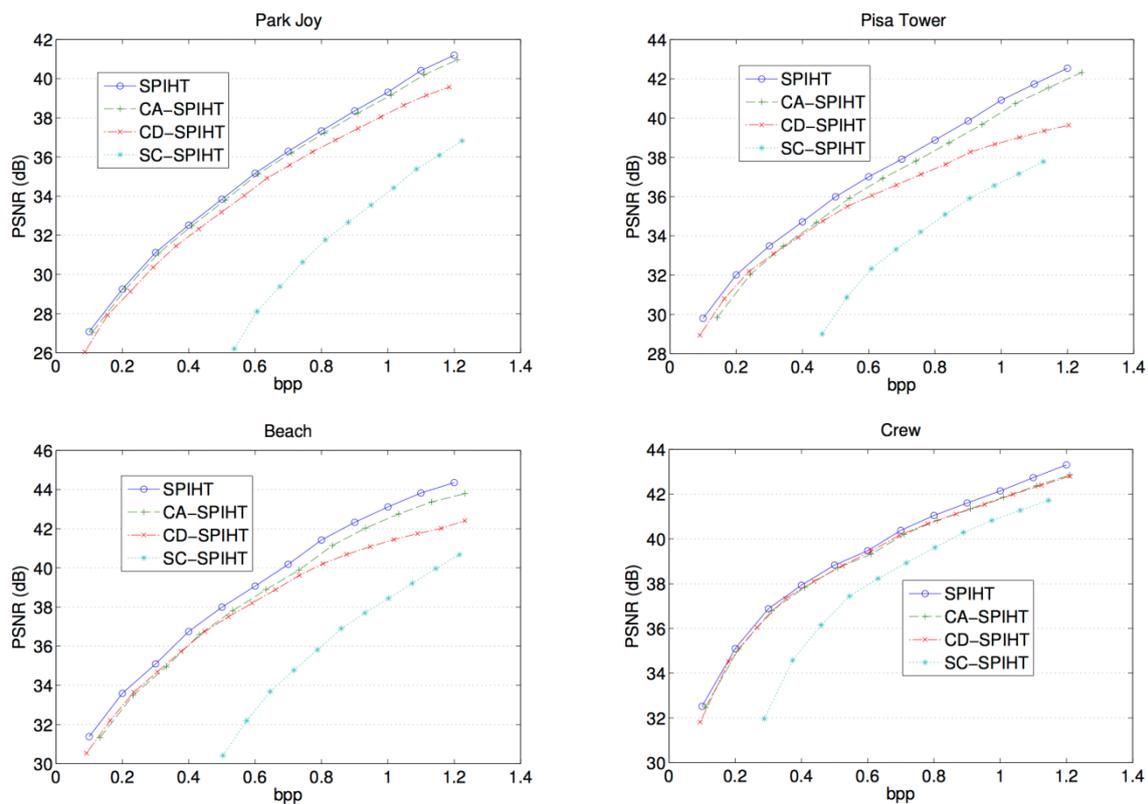


図 8 原サイズ画像の符号化性能に対するレート歪曲線

SPIHT



SC-SPIHT



CD-SPIHT



CA-SPIHT
(提案法)



図 9 復号画像の比較 (原サイズ、全ての画像を 1.0 bit/pixel で符号化)

レート依存シームカービングを用いた content-aware 画像符号化

Content-aware image coding using rate-dependent seam carving

4. 将来展望、将来における応用分野

本研究では、ビットレート依存シームカービングと、それを利用したコンテンツ考慮型画像符号化を提案した。本手法は符号化側（計算資源が豊富なサーバサイド）でコンテンツ考慮型拡張のための前処理を行うため、モバイルデバイスでの利用に有用であると考えられる。今後は本手法を動画像のサイズ変更・符号化へ拡張することで、映像を用いる、より広範囲の分野へと適用可能であると思われる。一方、動画像符号化は H/26x を始めとする符号化標準が実デバイスにおけるスタンダードな手法である。そのため、符号化標準を改変せずに直接利用可能なコンテンツ考慮型符号化を実現することも必要であると考えている。

また、本稿では flickr (www.flickr.com) よりダウンロードした以下の画像を creative commons license に基づき使用した（括弧内はユーザ名）：Mt. Evans (krossbow), Pisa Tower (wsuph001), Beach (weno)。

用語解説

- *1 親子関係を利用した画像符号化：ウェーブレット変換の低周波数成分と高周波数成分の係数間には相関があることを利用した画像符号化。
- *2 レート歪曲線：映像符号化で良く用いられる図で、横軸をビットレート (bit/pixel)、縦軸を定量的画質評価尺度（ピーク S/N 比が多い）としてプロットした図。曲線が上にあればあるほど良好な符号化手法だと考えて良い。

参考文献

- [1] S. Avidan and A. Shamir, “Seam carving for content-aware image resizing,” *ACM Trans. Graph.*, vol. 26, no. 3, 2007.
- [2] M. Rubinstein, A. Shamir, and S. Avidan, “Improved seam carving for video retargeting,” *ACM Trans. Graph.*, vol. 27, no. 3, 2008.

- [3] L. Wolf, M. Guttman, and D. Cohen-Or, “Non-homogeneous content-driven video-retargeting,” in *Proc. ICCV'07*, 2007.
- [4] Y. Wang, C.-L. Tai, O. Sorkine, and T.-Y. Lee, “Optimized scale-and-stretch for image resizing,” *ACM Trans. Graph.*, vol. 27, no. 5, 2008.
- [5] Y. Guo, F. Liu, J. Shi, Z. H. Zhou, and M. Gleicher, “Image retargeting using mesh parametrization,” *IEEE Trans. Multimedia*, vol. 11, no. 5, pp. 856–867, 2009.
- [6] D. Domingues, A. Alahi, and P. Vandergheynst, “Stream carving: An adaptive seam carving algorithm,” in *Proc. ICIP'10*, 2010.
- [7] W. Dong, N. Zhou, J. C. Paul, and X. Zhang, “Optimized image resizing using seam carving and scaling,” *ACM Trans. Graph.*, vol. 28, no. 5, 2009.
- [8] M. Rubinstein, A. Shamir, and S. Avidan, “Multi-operator media retargeting,” *ACM Trans. Graph.*, vol. 28, no. 3, 2009.
- [9] A. Nguyen, W. Yang, and J. Cai, “Seam carving extension: a compression perspective,” in *Proc. 17th ACM International Conference on Multimedia*, 2009, pp. 825–828.
- [10] Y. Tanaka, M. Hasegawa, and S. Kato, “Image coding using concentration and dilution based on seam carving with hierarchical search,” in *Proc. ICASSP'10*, 2010, pp. 1322–1325.
- [11] Y. Tanaka, M. Hasegawa, and S. Kato, “Improved image concentration for artifact-free image dilution and its application to image coding,” in *Proc. ICIP'10*, 2010, pp. 1225–1228.
- [12] 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.
- [13] 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.

レート依存シームカービングを用いた content-aware 画像符号化

Content-aware image coding using rate-dependent seam carving

- [14] M. Decombas, F. Capman, E. Renan, F. Dufaux, and B. Pesquet-Popescu, "Seam carving for semantic video coding," in Proc. SPIE, Applications of Digital Image Processing XXXIV, 2011.
- [15] C. Deng, W. Lin, and J. Cai, "Content-based image compression for arbitrary-resolution display devices," in Proc. ICC 2011, 2011, pp. 1–5.
- [16] C. Deng, W. Lin, and J. Cai, "Content-based image compression for arbitrary-resolution display devices," IEEE Trans. Multimedia, vol. 14, no. 4, pp. 1127–1139, 2012.
- [17] D. T. Vo, J. Sole, P. Yin, C. Gomila, and T. Q. Nguyen, "Selective data pruning-based compression using high-order edge-directed interpolation," IEEE Trans. Image Process., vol. 19, no. 2, pp. 399–409, 2010.
- [18] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," IEEE Trans. Circuits Syst. Video Technol., vol. 6, no. 3, pp. 243–250, 1996.

この研究は、平成22年度SCAT研究助成の対象として採用され、平成23～25年度に実施されたものです。