

早稲田大学大学院情報生産システム研究科

# 博士論文概要

## 論文題目

**Research on visual features based Content Adaptive  
Strategies for Video Coding**

申請者

Minghui WANG

情報生産システム工学専攻  
マルチメディアシステム研究

2012年9月

動画像の非可逆符号化においては、人間が知覚しにくい画質の歪をある程度許容することにより、冗長な情報データを見つけ削除することにより、データを圧縮する処理が行われる。従来の MPEG2 また H.264 等の標準化方式では、画質の評価に PSNR (Peak Signal-to-Noise Ratio) という客観的な評価方法を採用している。しかし、客観的と主観的な品質は必ずしも合致しないため、客観的評価のみならず、人間の視覚特性に基づいた主観的評価が必要である。国際標準方式として H.264/AVC は、高い符号化効率を追求するために、可変ブロックサイズ動き予測 (VBSME)、複数参照フレーム (MRF)、フレーム内予測 (INTRA) 等の従来の MPEG 1、2 では複雑度の観点から採用されなかった技術が多く使用されている。その結果、H.264 は MPEG-2 の 10 倍以上の演算量を必要としている。客観的な品質を落とさずに演算量を削減し、主観的にも良い品質を目指すことが重要である。

人間の網膜の中心部の視覚解像度は網膜周辺部の解像度より高い。このために、人間に興味ある物体は網膜の中心部に位置すると考えられ、画面内に人間の関心領域 (RoI: Region of Interest) が導入されている (JPEG2000 標準)。また、人間にとって静止画や動画像内の物体の縁 (Edge) は平坦部より注目度が高く、動画像においては、動きのある部分は静止部分より注目度が高い (Waloszek, 2005)

これらの考察に基づき、本論文では人間の視覚にとって画像品質を保ちながら、RoI に基づく圧縮法に取り組み、計算量が少ない手法を提案している。一方、マルチビュー動画像の奥行き画像 (Depth Map) に対して、輪郭線に基づく圧縮法を取り組み、高画質を達成するために、新しい方法を提案している。

第 1 章 [Introduction] では、動画像圧縮の基本原理と標準方式を紹介している。国際標準である H.264/AVC、マルチビューでの奥行き画像を使った画像生成 (DIBR: Depth Image Based Rendering)、及び、画像符号化に関わる人間の視覚特性を紹介し、本研究において取り組む研究課題を示している。

第 2 章 [Region-of-interest (ROI) based video coding strategy] では、低消費電力が要求される携帯電話や監視カメラなどの応用に適用することを目指して、ROI に基づく画像圧縮手法を提案している。携帯電話ではカメラの焦点が合った人物の顔が、監視カメラでは、カメラの焦点の合った動きある前景が ROI としてあげられる。既存の H.264 エンコーダでの高速化手法は、マクロブロック (MB: Macro Block) の予測モードと動きベクトル (MV: Motion Vector) の値を統計的処理により、生起する確率

をもとに、高速な演算を図っている。本手法は各 MB を ROI か非 ROI かに分類し、非 ROI の MB の予測モードの決定と MV の探索を厳密に行わないことで、演算量の大幅な削減により高速化を図ることを提案している。

ROI 検出について、Liu の手法 (Y. Liu, IEEE TCSVT2008) は、フレーム間の差分と肌の色の特徴を利用し、ROI の MB を検出している。Zhang の手法 (T. Zhang, IEICE 2011) は肌の色と顔の輪郭特徴を利用し、ROI を検出している。本論文では、先ず、空間的特徴を抽出し、次に、時間的特徴と肌の色の特徴を組み合わせ、ROI を検出している。空間的特徴として、特定の周波数や鮮明度が挙げられ、その抽出方法を Itti (IEEE TPAMI2008) が提案しているが、検出性能が良いが、演算量が大きい。本論文では演算量の少ない空間特徴抽出方法を考案したことに特徴がある。また、Zhang の提案は ROI 検出をエンコーダ自身に組み込んで高速化を図っているが、本論文では ROI 検出をエンコーダの前処理器として行うために、計算時間はかかるが、様々なエンコーダとの統合化が容易である。

H.264 のリファレンスソフトウェアである JM14 と比較し、計算時間は約 50% を削減できた。Zhang と Liu の提案と画質を比べたところ、平均で 0.36 dB または 0.26 dB 向上し、符号量は 8.9% また 6.2% を削減できた。主観的評価では、JM14 と比較し、ビデオ携帯電話向けの ROI の場合での観察者が 27% が “より良い” と 57% が “同じ” と答えた。監視カメラ向けの ROI の場合で 20% の観察者が “より良い、” 63% が “同じ” と回答した。

また、動きのある前景の領域を ROI として定義し、ROI を用いた奥行き画像符号化法を提案している。フレーム間の動き差分と空間的特徴の差分を取ることで、ROI を検出し、エンコーダの高速化を図っている。ROI に基づく方法は、JM14 と比較し、PSNR で 0.27dB の画質が悪化するが、計算時間は平均で約 50% 削減でき、画質の悪化がない場合では 24.6% の計算時間の削減が可能となっている。

第 3 章 [Framework of contour based depth map coding (CBDC) strategy] では、DIBR システムでのマルチビュー動画の奥行き画像 (Depth Map) に対して、新しい符号化法を提案している。奥行き画像での物体の輪郭の質を高めること、および合成された画像を高品質に保つために、輪郭線に基づく圧縮法 (CBDC) を提案している。ここで提案している CBDC 法とは、符号化の基本単位として、従来のマクロブロックの代わりに領域 (Region) を採用し、領域を輪郭線と内部領域に分けて、符号化を行う方法である。輪郭線の符号化は可逆で、内部領域は不可逆な方式を用いている。DIBR システムにおいて、従来のブロックに基づく方法と異なり、奥行き画像に対して任意形状の領域を符号化する方式であることが特徴

である。

奥行き画像は、各シーンでの画素点の奥行きの深度を表し、DIBR システムで新たに採用されたものである。奥行き画像内の鮮明度の高い輪郭は、合成画面上での ROI と見なされる。従来のテクスチャマップと比べ、奥行き画像は空間的な冗長性が高く、区分的 (piece-wise) 特徴を有している。この区分的特徴を利用して、ブロックに基づく符号化に適用した方法が提案されている (P. Merkel, SPIC2009)。ブロックを基本単位にすることは、分割しやすく、統一的に符号化しやすいが、空間的な冗長度を大幅に圧縮することはできなく、H.264 符号化よりも圧縮度は低かった。本論文ではフレームを任意形状の領域に分割し、符号化の基本単位はこの分割された領域であることが特徴である。

領域に対して符号化を行う手法として、F. Jager (VCIP2011) 及び J. Gautier (PCS2012) の方法がある。彼らの輪郭線の符号化手法は、JBIG と呼ばれる Bi-level 画像圧縮方法を使って輪郭線を符号化しているが、本論文ではベクトルを用いた方法を提案している。輪郭線内に大量の線分が存在すると、本提案方法は JBIG より効率良く符号化できる。内部領域の符号化手法として Jager は線形モデリング法を提案したが、本論文は 2 次元モデリングと線形モデリングを統合した方法を提案している。合成された画像で評価したところ、PSNR の値が Jager の方法より 3 dB 向上した。JM14 と比較したところ、本論文は輪郭線上点の対する輝度差の総和 (SAD: Sum of the absolute difference) が毎点平均的 1.803 を少なくなり、物体の縁はより高い品質となっている。SSIM (Structural Similarity Image Measurement) では 0.0002 の向上が見られ、ほとんど同じ画質である。主観的評価では、JM14 と比較し、28% が “より良い”、58% の観察者が “同じ” と答えている。

第 4 章 [Parallelization design for contour based depth map coding] では、高速化のために、CBDC の並列化手法を提案している。本論文で提案した CBDC 法は奥行きのレイヤーごと、あるいは領域ごとに独立な処理が可能であり、データの従属性が低い。各レイヤーや各領域にプロセッサを割り当て並列処理が可能となり、高速化が図ることができる。PC (Intel Core i7 4 core @ 2.93GHz, 16GB Memory) 上で並列化を行い実験したところ、2 プロセッサで 1.8 倍、4 プロセッサで 2.3 倍の速度向上が図られた。

第 5 章 [Conclusion] では、全篇の総括または将来の研究について検討している。