

令和 5 年 5 月 24 日現在

機関番号：11301

研究種目：基盤研究(B)（一般）

研究期間：2020～2022

課題番号：20H04201

研究課題名（和文）重要領域検出による超高圧縮映像符号化技術の開発

研究課題名（英文）Development of Technology for Realizing High-Compression of Video by Extracting Important Regions

研究代表者

大町 真一郎 (Omachi, Shinichiro)

東北大学・工学研究科・教授

研究者番号：30250856

交付決定額（研究期間全体）：（直接経費） 13,300,000円

研究成果の概要（和文）：映像から重要領域を検出し、映像としての価値や必要な情報を損なわずに高圧縮を実現する技術を開発することを目的として研究を行った。映像全体の品質を保つ映像符号化のこれまでの流れとは異なり、映像としての価値や必要な情報を損なわずに高圧縮を目指す点に大きな特徴がある。その結果、画像中の重要領域を検出する技術、画像の視認性を評価する技術、機械学習を用いた高圧縮画像符号化技術などを開発した。

研究成果の学術的意義や社会的意義

人が生成するデータの量は爆発的に増加している。その中でも映像データはインターネットトラフィックの8割を占めており、効率的な保存や伝送が求められている。本研究は画像中の各領域の重要度の推定や視認性を評価する指標を開発し、映像中から真に必要な情報を残しつつ高圧縮するための手法を開発したものである。映像通信の効率を向上させることが可能となり、様々な場面において映像データを有効活用するための基盤技術となり得る。

研究成果の概要（英文）：The purpose of this research was to develop a technology for detecting important regions in video and achieving high compression without losing the value of the video or necessary information. Unlike the conventional trend of video coding that preserves the quality of the entire image, the main characteristic of this technology is that it aims for high compression without losing the value and necessary information as a video. As a result, we developed a technology for detecting important regions in images, a technology for evaluating image visibility, and a high-compression image coding technology using machine learning.

研究分野：パターン認識

キーワード：映像符号化

## 1. 研究開始当初の背景

人が生成するデータの量は爆発的に増加している。2020年のデータ生成量は1人あたり毎秒1.7MBであり、換算すると全世界で1日あたり1.1ゼタバイト（ $1.1 \times 10^{21}$ バイト）ものデータが生成されると予測されている。これは使用できるストレージの量をはるかに上回っており、生成されるデータの多くはそのままの形では保存できずに捨てられてしまう。また、シスコ社の統計によると、インターネット上の映像トラフィックは2017年から2022年までの5年間に4倍に増加し、全トラフィックの82%を占めるようになると予測されている。すなわち、上記の問題は映像データにおいて最も顕著であると言える。例えば監視カメラは自然災害や犯罪など我々の生命を脅かすさまざまな場面になくはならない存在であるが、映像を限られたストレージに長時間記録したり狭帯域の通信回線を介して伝送したりするためには高い圧縮率で符号化する必要がある。カメラの解像度がいくら高くても、符号化による画質の劣化によって時には映像としての価値が大きく損なわれてしまうことがある。

上記の問題は、映像から時間的・空間的に真に重要な領域を検出し、映像としての価値や必要な情報を保存したまま高圧縮することができれば解決可能であると考えられる。映像はデータサイズが大きい一方で冗長性も高く、広い意味での冗長性を極限まで削減することでこのような技術は実現可能であると考えている。

## 2. 研究の目的

本研究では、映像から時空間上の重要領域を検出し、映像としての価値や必要な情報を損なわずに高圧縮を実現する技術を開発することを目的とする。映像全体の品質を保ち、PSNRやSSIMで性能を評価するのではなく、映像の価値や情報の必要性を考慮した符号化技術を開発する点に大きな特徴がある。

本研究の目的が達成できれば、あふれ続ける映像データの中から真に重要な情報を残していくことが可能となるほか、映像通信の効率も飛躍的に向上させることが可能となり、テレビ会議、IoTやセンサネットワークなど、さまざまな場面において映像データを有効活用するための基盤技術となり得る。さらに副次的な成果として、重要度によるランク付けや検索時のメタデータ生成など、映像にさまざまな付加価値を付与することも可能になる。

## 3. 研究の方法

映像から重要領域を検出する手法を開発する。また、重要な領域の視認性を保つための品質評価手法を開発する。さらに、機械学習を用いた画像・映像の高圧縮手法を開発する。これらを組み合わせ、重要領域の品質を保った高圧縮符号化技術を開発する。

## 4. 研究成果

### (1) 画像中の重要領域検出手法の開発

画像中から重要な領域を検出するための手法として、これまでは視覚的顕著性が用いられることが多かった。顕著性マップに関しては多くの研究がされている。しかし、顕著なオブジェクトは、必ずしも画像中の重要な被写体であるとは限らない。すなわち、視覚的に顕著であるにもかかわらず、ユーザーにとって重要ではないオブジェクトが存在する可能性がある。一般に、オブジェクトの重要性は画像のコンテキストに依存するため、画像内のどの領域が重要であるかを判断することは容易ではない。

本研究では、画像内の重要な領域を推定するために画像キャプションを用いる手法を提案した。キャプションを生成するためには画像のコンテキストを理解する必要があるため、生成されたキャプションは画像に関する重要な情報が含まれていることが期待される。画像中の主なオブジェクトは通常、キャプションにおける主語と目的語に相当することが多いため、キャプションから主語と目的語を抽出し、それらの単語を生成する際に注目した領域を特定することで重要な領域を推定する。

提案手法は、画像特徴抽出、キャプション生成、マスク生成、主語と目的語の検出、および重要度計算の5つの主要なモジュールで構成されている。画像の特徴抽出にはFaster R-CNNを用いた。キャプション生成には既存の手法を活用したが、提案手法では単にキャプションを生成するだけでなく重要な領域を推定することが主目的であるため、グリッドベースの特徴量ではなく領域ベースの特徴量を使用した。マスク生成とはオブジェクトの詳細な領域を抽出する処理であり、Mask R-CNNを用いて実現した。主語と目的語は、自然言語処理ツールStanzaによる係り受け解析を利用することにより検出した。最後に、キャプションの主語と目的語に対応する単語との関連度を領域ごとに計算することで重要度とした。

提案手法により重要領域を検出した例を図1に示す。図1(a)の画像から提案手法により各画素の重要度を推定したものが図1(b)である。画素の輝度値が重要度を表している。図1(c)は顕著性マップを求めた結果である。顕著性を用いた従来手法では画像の多くの部分が選択されてしまっているのに対し、提案手法では中央の傘を持った人物と傘が特に重要であると推定されていることが分かる。

また、提案手法を評価するための実験を行った。画像中の重要度を定義したデータセットを主観評価により作成し、F 値を用いて視覚的顕著性を用いた手法と提案手法の精度を定量的に比較した。その結果、視覚的顕著性を用いた最新の既存手法と比較して提案手法はより高精度で重要領域を推定することができることが示された。



図1 キャプションを用いた重要領域検出

### (2) 視認性を考慮した画像符号化手法の開発

まず情景画像中のテキストを対象とし、品質推定モデルを開発した。情景画像ではなく光学的スキャナでスキャンした文書画像内の文字の品質については推定する手法がいくつか提案されている。機械学習を利用する方法としては、文書画像を小領域に分割して特徴抽出する手法等が提案されている。しかし、情景画像中のテキストを対象とした研究はほとんどなかった。提案手法では情景画像中のテキスト領域の画像を入力とし、畳み込み層で画像特徴を抽出しつつ、文字認識モデルにより文字認識結果を得る。これらを機械学習モデルである Transformer に入力し、両特徴を統合して学習させる。

提案した品質モデルを用いた画像符号化手法も提案した。画素単位での画像の品質を定義した品質マップを与え、品質マップに応じて画像の品質を変える適応的符号化手法を活用した。これにより視認性を保った画像の符号化が可能になる。提案手法による処理結果の例を図2に示す。図2(a)の画像を、提案手法でテキストの視認性を保証しつつ圧縮した画像を図2(b)に示す。一般的な画像圧縮手法である JPEG2000 の結果 (図2(c)) と比較してテキストの視認性が優れていることが分かる。

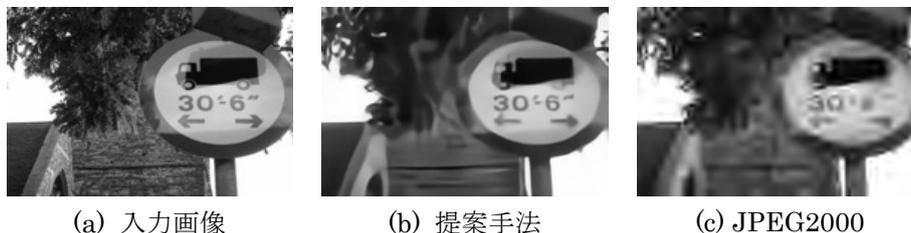


図2 視認性を考慮した画像符号化

### (3) 機械学習を用いた画像の高圧縮手法の開発

機械学習を用いて画像を高圧縮する手法を開発した。敵対的生成ネットワークを活用し、元の画像を忠実に再現しながらデータ量をできるだけ削減できるように学習を行う。図3に例を示す。図3(a)の画像を、1画素あたり0.031ビットにまで圧縮した画像を図3(b)に示す。細かい情報が失われているものの、元画像と大差ない画像が得られていることが分かる。



図3 機械学習を用いた画像の高圧縮

## 5. 主な発表論文等

〔雑誌論文〕 計14件（うち査読付論文 13件 / うち国際共著 1件 / うちオープンアクセス 7件）

1. 著者名 Shohei Uchigasaki, Tomo Miyazaki, Shinichiro Omachi	4. 巻 -
2. 論文標題 Deep Image Compression Using Scene Text Quality Assessment	5. 発行年 2023年
3. 雑誌名 Pattern Recognition	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Taku Suzuki, Daisuke Sato, Yoshihiro Sugaya, Tomo Miyazaki, Shinichiro Omachi	4. 巻 10
2. 論文標題 Important Region Estimation Using Image Captioning	5. 発行年 2022年
3. 雑誌名 IEEE Access	6. 最初と最後の頁 105546 ~ 105555
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/ACCESS.2022.3211260	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 大町真一郎, 宮崎 智	4. 巻 34
2. 論文標題 情景画像中の重要テキストの検出と伝送	5. 発行年 2023年
3. 雑誌名 画像ラボ	6. 最初と最後の頁 17 ~ 22
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Yongsong Huang, Qingzhong Wang, Shinichiro Omachi	4. 巻 13583
2. 論文標題 Rethinking Degradation: Radiograph Super-Resolution via?AID-SRGAN	5. 発行年 2022年
3. 雑誌名 Lecture Notes in Computer Science	6. 最初と最後の頁 43 ~ 52
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/978-3-031-21014-3_5	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Zhengni Tang, Tomo Miyazaki, Yoshihiro Sugaya, Shinichiro Omachi	4. 巻 30
2. 論文標題 Stroke-Based Scene Text Erasing Using Synthetic Data for Training	5. 発行年 2021年
3. 雑誌名 IEEE Transactions on Image Processing	6. 最初と最後の頁 9306 ~ 9320
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/TIP.2021.3125260	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Junpei Masuho, Tomo Miyazaki, Yoshihiro Sugaya, Masako Omachi, Shinichiro Omachi	4. 巻 70
2. 論文標題 A Framework for Estimating Gaze Point Information for Location-Based Services	5. 発行年 2021年
3. 雑誌名 IEEE Transactions on Vehicular Technology	6. 最初と最後の頁 8468 ~ 8477
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/TVT.2021.3101932	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Huy Manh Nguyen, Tomo Miyazaki, Yoshihiro Sugaya, Shinichiro Omachi	4. 巻 11
2. 論文標題 Multiple Visual-Semantic Embedding for Video Retrieval from Query Sentence	5. 発行年 2021年
3. 雑誌名 Applied Sciences	6. 最初と最後の頁 3214 ~ 3214
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/app11073214	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Shoma Iwai, Tomo Miyazaki, Yoshihiro Sugaya, Shinichiro Omachi	4. 巻 -
2. 論文標題 Self Texture Transfer Networks for Low Bitrate Image Compression	5. 発行年 2021年
3. 雑誌名 Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops	6. 最初と最後の頁 1901 ~ 1905
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/CVPRW53098.2021.00214	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Shuya Sano, Tomo Miyazaki, Yoshihiro Sugaya, Naohiro Sekiguchi, Shinichiro Omachi	4. 巻 9
2. 論文標題 Mackerel Fat Content Estimation Using RGB and Depth Images	5. 発行年 2021年
3. 雑誌名 IEEE Access	6. 最初と最後の頁 164060 ~ 164069
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ACCESS.2021.3134260	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Shinichiro Omachi and Masako Omachi	4. 巻 -
2. 論文標題 Correlation-Based Data Augmentation for Machine Learning and its Application to Road Environment Recognition	5. 発行年 2022年
3. 雑誌名 IEEE Transactions on Vehicular Technology	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/TVT.2022.3167048	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Kota Odaira, Tomo Miyazaki, Yoshihiro Sugaya, Shinichiro Omachi	4. 巻 -
2. 論文標題 Importance Estimation for Scene Texts Using Visual Features	5. 発行年 2022年
3. 雑誌名 Interdisciplinary Information Sciences	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.4036/iis.2022.A.06	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Huy Manh Nguyen, Tomo Miyazaki, Yoshihiro Sugaya, Shinichiro Omachi	4. 巻 11
2. 論文標題 Multiple Visual-Semantic Embedding for Video Retrieval from Query Sentence	5. 発行年 2021年
3. 雑誌名 Applied Sciences	6. 最初と最後の頁 3214 ~ 3214
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/app11073214	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Yoshito Nagaoka, Tomo Miyazaki, Yoshihiro Sugaya, Shinichiro Omachi	4. 巻 21
2. 論文標題 Text Detection Using Multi-Stage Region Proposal Network Sensitive to Text Scale	5. 発行年 2021年
3. 雑誌名 Sensors	6. 最初と最後の頁 1232 ~ 1232
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/s21041232	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Shoma Iwai, Tomo Miyazaki, Yoshihiro Sugaya, Shinichiro Omachi	4. 巻 -
2. 論文標題 Fidelity-Controllable Extreme Image Compression with Generative Adversarial Networks	5. 発行年 2021年
3. 雑誌名 Proceedings of the 25th International Conference on Pattern Recognition	6. 最初と最後の頁 8235 ~ 8242
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICPR48806.2021.9412185	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計17件 (うち招待講演 3件 / うち国際学会 7件)

1. 発表者名 岩井翔真, 宮崎智, 大町真一郎
2. 発表標題 GANを使った画像符号化における識別器に関する検討
3. 学会等名 画像の認識・理解シンポジウム
4. 発表年 2022年

1. 発表者名 宇川慧, 宮崎智, 大町真一郎
2. 発表標題 物体のテクスチャに注目したGANによる熱赤外線画像着色手法の検討
3. 学会等名 画像の認識・理解シンポジウム
4. 発表年 2022年

1. 発表者名 梶原颯希, 宮崎智, 大町真一郎
2. 発表標題 編集距離に基づいた距離学習による情景画像文字認識結果の修正に関する検討
3. 学会等名 画像の認識・理解シンポジウム
4. 発表年 2022年

1. 発表者名 谷本衡哉, 宮崎智, 大町真一郎
2. 発表標題 U-Netを用いた電子部品画像からの配線検出
3. 学会等名 画像の認識・理解シンポジウム
4. 発表年 2022年

1. 発表者名 三ツ井悠翔, 宮崎智, 大町真一郎
2. 発表標題 Masked Image Modelingを利用した情景画像中のテキスト認識
3. 学会等名 言語処理学会第29回年次大会
4. 発表年 2023年

1. 発表者名 Shinichiro Omachi
2. 発表標題 Detecting Important Information from Scene Images via Texts
3. 学会等名 The 6th Tohoku Uni-NTU Symposium on Interdisciplinary AI and Human Studies (招待講演) (国際学会)
4. 発表年 2023年

1. 発表者名 Shinichiro Omachi
2. 発表標題 Image and Video Compression Preserving Important Information for Humans
3. 学会等名 2021 6th International Conference on Signal and Image Processing (ICSIP 2021) (招待講演) (国際学会)
4. 発表年 2021年

1. 発表者名 Shoma Iwai, Tomo Miyazaki, Yoshihiro Sugaya, Shinichiro Omachi
2. 発表標題 Semantic Guided Extreme Image Compression with Generative Adversarial Networks
3. 学会等名 電気関係学会東北支部連合大会
4. 発表年 2021年

1. 発表者名 Zhengmi Tang, Tomo Miyazaki, Yoshihiro Sugaya, Shinichiro Omachi
2. 発表標題 Erasing Scene Text Based on Stroke Prediction
3. 学会等名 電気関係学会東北支部連合大会
4. 発表年 2021年

1. 発表者名 内ヶ崎翔平, 菅谷至寛, 宮崎智, 大町真一郎
2. 発表標題 文字の可読性を考慮した画像の高効率符号化手法の実験的検証
3. 学会等名 画像の認識・理解シンポジウム
4. 発表年 2021年

1. 発表者名 岩井翔真, 宮崎智, 菅谷至寛, 大町真一郎
2. 発表標題 参照型超解像による特徴転移を利用した高品質な画像符号化
3. 学会等名 画像の認識・理解シンポジウム
4. 発表年 2021年

1. 発表者名 Shinichiro Omachi
2. 発表標題 Extraction of important text information from images
3. 学会等名 Symposium of Yotta Informatics Research Platform for Yotta-Scale Data Science 2021 (招待講演) (国際学会)
4. 発表年 2021年

1. 発表者名 Shoma Iwai, Tomo Miyazaki, Yoshihiro Sugaya, Shinichiro Omachi
2. 発表標題 Fidelity-Controllable Extreme Image Compression with Generative Adversarial Networks
3. 学会等名 25th International Conference on Pattern Recognition (国際学会)
4. 発表年 2021年

1. 発表者名 Taku Suzuki, Daisuke Sato, Tomo Miyazaki, Yoshihiro Sugaya, Shinichiro Omachi
2. 発表標題 Study of Detecting Important Regions in Natural Images with Image Captioning
3. 学会等名 The 17th International Workshop on Emerging ICT (国際学会)
4. 発表年 2020年

1. 発表者名 Takuto Toida, Tomo Miyazaki, Yoshihiro Sugaya, Shinichiro Omachi
2. 発表標題 Survey of Automatic Video Colorization by DeepNeural Network
3. 学会等名 The 17th International Workshop on Emerging ICT (国際学会)
4. 発表年 2020年

1. 発表者名 Shoma Iwai, Tomo Miyazaki, Yoshihiro Sugaya, Shinichiro Omachi
2. 発表標題 Two-Stage Training for High-Fidelity Image Compression with Generative Adversarial Networks
3. 学会等名 The 17th International Workshop on Emerging ICT (国際学会)
4. 発表年 2020年

1. 発表者名 鈴木 拓, 佐藤大亮, 宮崎 智, 菅谷至寛, 大町真一郎
2. 発表標題 キャプションを用いた画像中の重要領域判別手法の検討
3. 学会等名 画像の認識・理解シンポジウム
4. 発表年 2020年

〔図書〕 計1件

1. 著者名 計測自動制御学会、浮田 浩行、濱上 知樹、藤吉 弘亘、大町 真一郎、戸田 智基、岩崎 敦、小林 泰介、鈴木 亮太、木村 雄喜、橋本 大樹、玉垣 勇樹、水谷 麻紀子、永田 毅、木村 光成、李 晃伸、川嶋 宏彰	4. 発行年 2022年
2. 出版社 コロナ社	5. 総ページ数 240
3. 書名 機械学習の可能性	

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	宮崎 智  (Miyazaki Tomo)  (10755101)	東北大学・工学研究科・助教    (11301)	
研究分担者	大町 方子  (Omachi Masako)  (90316448)	仙台高等専門学校・総合工学科・教授    (51303)	
研究分担者	菅谷 至寛  (Sugaya Yoshihiro)  (80323062)	東北大学・工学研究科・准教授    (11301)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関