



AIを利用した映像圧縮と無線伝送への応用

AI-based Video Processing and Its Application to Wireless Transmission

岳 麗*
Li Yue

前田 直樹
Naoki Maeda

豊田 重治
Shigeharu Toyoda

小河 昇平
Shouhei Ogawa

半田 和也
Kazuya Handa

藤田 康仁
Yasuhito Fujita

近年AIによるヒト・モノ・コトなどの映像解析が幅広い分野で進んでいる。同時に、より広範囲に、より高精細に見たいニーズから、4Kや8Kなど、カメラ映像の高解像度化が進んでいる。それに伴い、映像データの伝送量や保存コスト、AI分析の処理負荷の増大が課題となっている。今回、データ量を大幅に削減でき、且つ圧縮伝送後のAI処理負荷を低減できるAI応用映像処理技術（AVP: AI-based Video Processing）を試作した。その有効性を検証するために行った工場実証実験では、従来圧縮技術と比較して平均ビットレートを92.2%削減できたと同時に、クラウド側AI処理負荷の低減により限られた計算リソースでより高精細映像の解析ができ業務改善に繋がった。

In recent years, AI-based video analysis has been widely used in various fields. Demand for higher definition and wider area viewing has led to the development of cameras with higher resolutions such as 4K and 8K. As a result, cost for video data transmission, storage, and analysis has increased significantly. In response to this, we have developed a prototype AI-based video processing technology (AVP) that can significantly reduce the amount of video compression data and cloud AI computation costs after transmission. In a field test conducted at a factory, AVP reduced the average bit rate by 92.2% compared to conventional video compression technology, while reducing the AI computation cost on the cloud side, enabling highly accurate analysis of workers' movement flow using high-resolution video.

キーワード：AVP（AI-based Video Processing）、高効率映像圧縮、AI、工場IoT、5G

1. 緒 言

従来、映像は人が目で見て楽しむものであったが、近年その用途が広がり、AIによるヒト・モノ・コトなどの映像解析が幅広い分野で進んでいる。同時により広範囲に、より高精細に見たいニーズから、4Kや8Kなどカメラ映像の高解像度化が進んでいる。この結果、車、ロボット、工場、道路など、さまざまなデバイスや場所に設置された多数の高解像度カメラから大量の映像データが伝送され、今後、映像データ量が爆発的に増加すると予測されている。

AIによる映像解析には深層学習が多用されるが、多くの計算パワーを使用する解析はリソースに限りがあるエッジ（カメラ）側だけでは処理しきれず、クラウド側と連携した処理が有効となる。しかし、大量の映像データのクラウド側への伝送はネットワーク帯域を圧迫してしまう問題がある。近年、高速／大容量な5G（第5世代移動通信システム）が普及しつつあるが、大量の映像データを伝送するためには、5G回線を用いる場合であっても伝送の効率化が大きな課題となっている。そのため、ネットワーク側技術の進化だけでなく、エッジ側技術の進化も必要である^{(1)~(3)}。

それに対して、当社はAVP（AI-based Video Processing）という高効率映像圧縮伝送技術を開発した。AVPによる圧縮伝送後の映像はAIによる解析又は人の目視による状況確認に用いられることを想定している。そのため、従来の映像圧縮技術が指標とする画素単位での映像の忠実な再現ではなく、圧縮伝送後のAI解析精度と意味レベルでの再現を重視している。また、従来の圧縮技術に比べて10分の1以

下で映像を圧縮することが可能である。本論文では、AVPの特徴、およびその技術を活用した工場でのIoT応用事例を報告する。

2. 高効率映像圧縮伝送技術 AVP

AVPは見たい部分を高精細に、それ以外の部分を不鮮明にすることによって消費するエネルギーを低く抑える人間の眼の特徴を再現し、圧縮後のデータ量を最小限に抑えながら必要な情報を高画質でリアルタイムに伝送することが特徴である。図1で示すように、AVPは注目領域抽出AIと領域別圧縮の2つのコア処理からなる。注目領域抽出AIは入力映像画面の中に圧縮伝送後のAI解析又は人間の状況理解に影響する領域（以降、注目領域という）を画像から抽出し、領域別圧縮では、注目領域だけを高画質に圧縮して

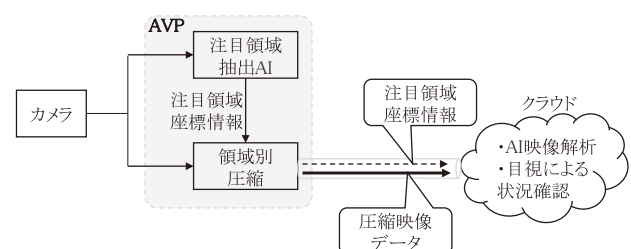


図1 AVPブロック図

それ以外の領域を低画質に圧縮する。こうすることによって、画面全体を均一画質で圧縮する従来技術と比較して、必要な情報を失わずに圧縮率を大幅に向上することが可能である。

AVPの特長は下記の4点である。

- ① 領域別画質制御により、従来均一圧縮対比1/10以下のビットレート
 - ② 標準準拠圧縮で、汎用デコーダや映像再生ソフトで再生可能
 - ③ 注目領域座標情報を映像圧縮データのベンダー拡張エリアに格納し伝送し、クラウド側AIの処理負荷を軽減可能
 - ④ 通信スループット変動に応じたビットレート制御
- 上記特長を活かす適用先として下記3つの例を挙げる。
- (1) 高精細映像による高精度AI解析のニーズが高まる工場IoT
 - (2) モバイル回線を経由した車内外の遠隔監視が必要だが映像伝送コストが高い自動運転バス/タクシー
 - (3) モバイル回線を経由した遠隔監視・制御が必要とされる自動配送ロボット、警備ロボットなどの屋外自律型ロボット

以下、AVP各機能の詳細を説明する。

2-1 注目領域抽出AI

注目領域抽出AIは深層学習をベースとしており、クラウド側のAI解析対象となる物体、あるいは人間が状況理解の対象となる物体（以降、注目物体という）を含む領域を抽出する。

注目領域抽出AIは図2で示すように、画面を複数のブロックに分割し、各ブロックで注目物体（図2では車両）の有無を判断する。注目物体の位置特定に物体検出AIがよく使われるが、注目領域抽出AIは物体検出AIが行っている物体の種類と正確な位置情報まで判断することをせずより処理がシンプルであるため、より軽量のニューラルネットワークで実現できる。注目領域抽出AIはブロック内における注目物体の有無だけではなく、小さい注目物体が存在するか、大きい注目物体が存在するかまで判断できる。以降、小

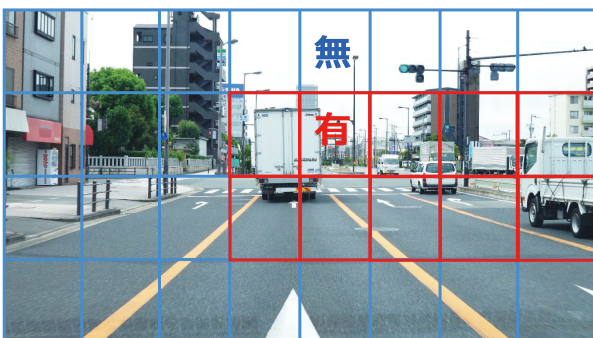


図2 注目領域抽出AIの出力例

い注目物体が存在するブロックを「細注目領域」、大きい注目物体が存在するブロックを「粗注目領域」、注目物体が入っていないブロックを「非注目領域」という。

2-2 領域別圧縮

領域別圧縮は、細注目領域を高画質（低圧縮率）、粗注目領域を中画質（中圧縮率）、非注目領域を低画質（高圧縮率）に圧縮する（図3参照）。画質（圧縮率）の制御は量子化パラメータQP（Quantization Parameter）^{*1}の変更により行う。各領域のQP値は、圧縮伝送後クラウド側のAI解析精度を保持できるように最適化されている。標準圧縮技術を使用しており、汎用デコーダや映像再生ソフトで再生できるため、特殊なデコーダを用意する必要がない。

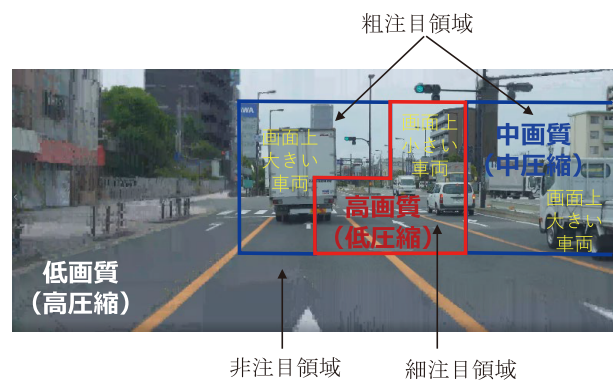


図3 注目領域と領域別圧縮例

2-3 注目領域座標情報の利用によるクラウド側AI処理負荷低減

注目領域抽出AIの出力である注目領域座標情報は、クラウド側AIの処理負荷低減に利用可能である。AI処理の計算量は一般的に入力画像の画素数に比例することが知られている。よって、クラウド側では、デコードした画像全画面をAIに入力する代わりに、図4で示すように、注目領域座標情報をもとに注目領域だけを切り出してAIに入力することによって、AIの計算量を大幅に低減できる。

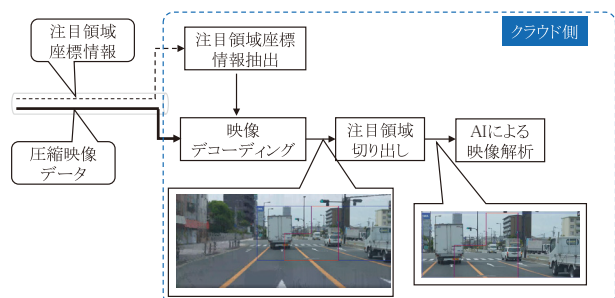


図4 注目領域座標情報利用ダイアグラム

2-4 ネットワーク状況監視に基づいたビットレート制御

モバイル回線を経由した遠隔監視・制御では、通信可用帯域^{*2}の変動が深刻な課題として知られている。それによって映像パケットの到着間隔が大きくなり、パケット欠損して正しく復号化できず画面が大きく乱れたり、という問題が起きる。

AVP では、不安定な伝送路でも安定的に通信品質を確保し、秘匿性の高い配信を可能にする映像伝送プロトコルを採用している。映像伝送プロトコルからさまざまなネットワーク統計情報をリアルタイムに取得できるため、これらの統計情報を使用する送信ビットレート制御技術の開発に取り組んでいる。具体的には、複数の統計情報を組み合わせるリアルタイムに通信帯域の混雑状況を検知し、**図5**で示すように、可用帯域を超えるまで送信ビットレートを増やし、可用帯域を超えたことを検知したら送信ビットレートを減らす仕組みを実装している。

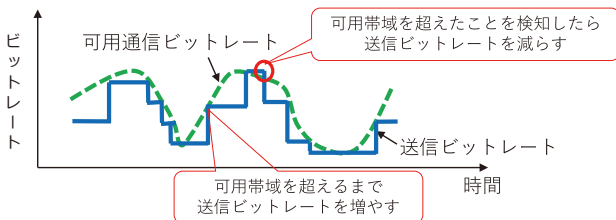


図5 可用帯域検知に応じたビットレート制御

3. AVPの工場IoT応用事例

当社はソフトバンク(株)と、5GやAI、IoTを活用したDX(デジタルトランスフォーメーション)による工場での生産性向上の取り組みを進めている。工場内の様々なセンサ、設備、カメラからのデータを、当社産業用5G端末を使ってクラウドサーバに伝送し分析することで、簡易なネットワーク機器構成で、簡単、安定的に生産現場を見える化できる。

これまでの実証実験^{(4),(5)}では、工場にFull HD(1920×1080)カメラを設置し、映像から骨格認識AI^{*3}を用いて作業者を検出してその結果に基づいて作業者の動線解析を実施し、作業時間をリアルタイムに見える化することで予実差の大きい作業等を迅速に現場にフィードバックできる仕組みを構築してきた。しかし、3つの制約(1)解析時間、(2)GPUメモリ使用量、(3)伝送量)からFull HD映像を640×640にリサイズしてから骨格認識AI処理をしていたため、AI入力の映像解像度が低く、画面奥に存在する小さく映った作業者は輪郭がつぶれ検出ができず、骨格認識AIの検知率が58.2%と低く、改善に必要な検知率である95%を満たしていなかった。対策として、映像をリサイズせずに元の解像度でAI処理したり、さらに4K、8Kなど、

より高解像度な映像でAI解析を行うことが考えられる。しかしその場合、クラウドサーバへ映像を送信する際の伝送量の大幅な増加や、AIの処理負荷や処理時間の大幅な増加につながる。伝送量については、高速/大容量な5G回線を用いる場合であっても伝送の効率化が大きな課題となる。これらの課題に対して、今回、**図6**で示すように4Kカメラと試作したAVP装置を利用することで、5G回線のトラフィックとクラウドサーバの処理負荷を増加させることなく、骨格認識AIの検出率を大幅に向上することができた。

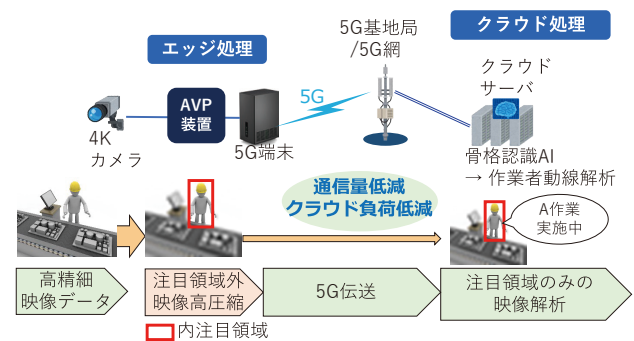


図6 AVP導入後のシステム構成図

4K高精密映像をそのまま伝送した場合は5G回線でも帯域不足でパケットロス(データ欠損)が大量に発生したが、AVP処理を付加することで平均ビットレートが92.2%削減され問題なく伝送できるようになった。クラウド側の映像解析では、4K映像から注目領域の座標情報に基づいて作業者の領域を切り出す。続いて、切り出した画像を640×640にリサイズして骨格認識AIに入力する。ここで、切り出された作業者の画像は640×640の画像サイズを大幅に上回ることには少ないため、ほとんどの場合画像を縮小することなくAI処理を行うことができ検知精度を向上させることができた。結果、作業者の検知率を99.1%まで改善できた。また、骨格認識AI処理は、カメラの撮影解像度(4K)に依存せず、切り出した小さな画像に対して処理するため、クラウドサーバの計算パワーを追加することなく解析することができた。

AVP導入前後の比較を表1に示す。AVPを導入することによって高精密カメラを使用でき、骨格認識AIの検知率が大幅改善されたため、精度高い作業時間に見える化が実現

表1 AVP導入前後の比較

	AVP導入前	AVP導入後
カメラ解像度	1920×1080	3840×2160
平均ビットレート	1.6Mbps	0.5Mbps
骨格認識AI検知率	58.2%	99.1%

でき業務改善につながった。さらに映像のデータ量を低く抑えられたため、映像の保存コスト削減の効果もある。

4. 結 言

高効率映像圧縮技術 AVP を開発し、工場実証実験を通じて、伝送ビットレート的大幅削減とクラウド側 AI 処理負荷低減を確認した。当技術はモバイル網経由の映像による遠隔監視を必要とする自動運転バス/タクシーや、自動配送ロボットと警備ロボットを代表とする屋外自律型ロボットにも有効だと考えており、今後実証実験を通じて効果を確認していく予定である。

5. 謝 辞

本研究開発にあたっては名古屋大学大学院情報学研究科・村瀬洋特任教授、出口大輔准教授、川西康友講師に有益なご指導ご助言をいただきました。心より感謝申し上げます。

用語集

※1 量子化パラメータ QP

量子化ステップ幅を設定するパラメータであり、QP が 6 増える毎に量子化ステップ幅は 2 倍になる。

※2 可用帯域

通信路中のボトルネックリンクの物理帯域から他者が利用しており使えない帯域を除いた、実際に利用できる帯域。

※3 骨格認識 AI

人物の関節点を検出する深層学習アルゴリズム。

参 考 文 献

- (1) Y. Shinohara, H. Itsumi, B. Florian, T. Iwai, "Video compression estimation recognition accuracy for remote site object detection," IWCMC (2020)
- (2) L. Galteri, M. Bertini, L. Seidenari, A. Del Bimbo, "Video Compression for Object Detection Algorithms," ICPR (2018)
- (3) 久保田智規、中尾鷹詔、吉田英司、「深層学習を用いた映像解析のための高圧縮な動画画像符号化手法」、信学技報、vol. 119、no. 456、IE2019-121、pp. 121-126 (2020年3月)」
- (4) プレスリリース (2019年11月12日)、「住友電工とソフトバンク、工場での5G活用に向けた実証実験を実施」
<https://sei.co.jp/company/press/2019/11/prs089.html>
- (5) プレスリリース (2021年6月9日)、「5GとAIを活用した映像伝送・解析で、工場の作業を自動的かつリアルタイムに見える化」
<https://sei.co.jp/company/press/2021/06/prs047.html>

執 筆 者

岳 麗* : 情報ネットワーク研究開発センター
グループ長
博士 (工学)



前田 直樹 : 情報ネットワーク研究開発センター
主幹



豊田 重治 : 情報ネットワーク研究開発センター
主席



小河 昇平 : 情報ネットワーク研究開発センター
主席



半田 和也 : IoT 研究開発センター



藤田 康仁 : IoT 研究開発センター 部長



*主執筆者