情報処理学会論文誌 Vol.59 No.7 1425–1434 (July 2018)

視覚特性に基づく高効率映像圧縮伝送システム

岡田 光弘[1,a)] 佐藤 拓杜[1] 稲田 圭介[1] 谷田部 祐介[1] 伊藤 浩朗[2] 小味 弘典[3]

受付日 2017年9月29日, 採録日 2018年4月4日

概要：人間の視野における解像度は視野中心で最も高く，周辺部ほど低下する．この視覚特性に基づき，視野周辺の映像の解像度を削減する Foveated Imaging 処理（FI 処理）と呼ばれる画像処理技術がある．本論文では，フル HD 映像を対象に主観的な解像感を維持したまま伝送ビットレートを削減するために，エンコード前の映像に FI 処理を適用した映像圧縮伝送システムを構築する．さらに，このシステムの有効性を評価するために，Degradation Category Rating（DCR）法と呼ばれる主観画質の評価手法を用いて，基準映像に対する FI 処理を適用した映像の画質劣化を 5 段階で評価した．その結果，Degradation Mean Opinion Score（DMOS）の平均値が 4.25 と画質劣化が気にならないレベルの画質を維持しつつ，9.2～44.7%の伝送ビットレートの削減効果を確認できた．

キーワード：Foveated Imaging，視覚特性，映像圧縮伝送システム，アイトラッキング

High Efficiency Video Coding Transmission System Based on Human Vision Properties

Mitsuhiro Okada[1,a)] Takuto Sato[1] Keisuke Inata[1] Yusuke Yatabe[1] Hiroaki Ito[2] Hironori Komi[3]

Received: September 29, 2017, Accepted: April 4, 2018

Abstract: Resolution of human visual field is the highest at the center of the view field, and it decreases in the visual periphery. There is an image processing technique called Foveated Imaging that reduce the image resolution around the visual periphery to take advantage of human vision properties. In this paper, we develop a video coding transmission system that utilizes Foveated Imaging in order to reduce the video transmission bandwidth while maintaining the image quality perceived by humans. The image quality degradation of proposed videos against reference videos was evaluated on a scale of 1 to 5 using a subjective image quality evaluation method called Degradation Category Rating (DCR) method. The evaluation results indicate the average value of Degradation Mean Opinion Score (DMOS) is 4.25 which suggests little to no effect on image quality, while reducing the video transmission bandwidth by 9.2–44.7%.

Keywords: Foveated Imaging, human vision properties, video coding transmission system, eye tracking

1. はじめに

近年，ドローンの長距離遠隔操縦や，災害時の危険な現場での機器の遠隔操作など遠隔地にある機器を遠隔操作したいという需要は高まっている．**図 1** に想定する遠隔操作システムの例を示す．図 1 のように遠隔地と操作室に分かれており，遠隔地では，カメラを取り付けた遠隔操作対象の機器があり，カメラで撮影した映像をエンコーダで圧縮して，Long Term Evolution（LTE）回線などの無線ネットワークを介して操作室に送る．一方，操作室では，デコードした映像を遠隔操作者が視聴しながら，遠隔操作対象の機器を操作する．遠隔地と操作室の間の距離は，遠隔操作システムに対するニーズ [16] より，最大 5 km 程度

1 株式会社日立製作所
Hitachi Ltd., Kokubunji, Tokyo 185–8601, Japan
2 日立オートモーティブシステムズ株式会社
Hitachi Automotive Systems, Ltd., Hitachinaka, Ibaraki 312–8503, Japan
3 株式会社日立産業制御ソリューションズ
Hitachi Industry & Control Solutions, Ltd., Taito, Tokyo 110–0006, Japan
a) mitsuhiro.okada.uf@hitachi.com

© 2018 Information Processing Society of Japan

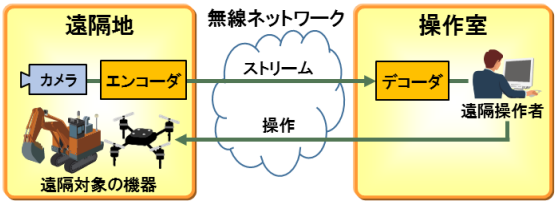


図 1　遠隔操作システム
Fig. 1　Remote control system.

を考えている．

このような遠隔操作システムでは，遠隔操作時の臨場感向上のために伝送する映像を高解像度化したいという要望がある．しかし，現行システムではネットワーク帯域を削減するため，画像を縮小する，フレームレートを削減する，エンコード時の量子化ステップ幅を大きくするなど，画像全体を一律で画質劣化させる方法で伝送ビットレートを削減している [1], [2]．

一方，伝送ビットレートを削減する手段として，視聴者の注視点情報に基づいて解像度を削減する Foveated Imaging 処理（FI 処理）と呼ばれる画像処理技術が提案されている [3], [4], [5]．この処理は，人間の視野の中心では最高解像度で物体をとらえることが可能であるが，視野の周辺になるほど解像度が低下するという視覚特性を利用している．この特性によれば，視聴者の注視点を逐次追跡し，各時刻の映像から知覚できない解像度を削減しても，視聴者は情報損失に気づかないということになる．この思想に基づく画像処理は，近年のアイトラッキングデバイスの高性能化や低価格化を受けて実用性が高まっている．

本論文では，フル HD 映像を対象に主観的な解像感を維持したまま伝送ビットレートを削減するために，エンコード前の映像に FI 処理を適用した映像圧縮伝送システムを構築する．さらに，構築したシステムの有効性を評価する．

2.　関連研究

人間の視覚特性に基づき，主観的な画質を維持したまま圧縮効率を改善する手法の関連研究について述べる．

文献 [3], [4], [5] は，エンコーダに入力する前の映像に注視点情報に基づいた FI 処理を施すことにより圧縮効率を改善する手法を提案している．また，文献 [6], [7] は，エンコード処理の中に注視点情報に基づいた処理を追加することで圧縮効率を改善する方法を提案している．これらの論文は，固定の注視点を用いて評価しており，リアルタイムで視聴者の注視点情報を取得して動作するシステムを構築して評価していない．

一方，文献 [8] は，ポインティングデバイスを使用してリアルタイムでエンコード処理に反映するシステムを構築している．この論文は，ポインティングデバイスで注視点情報を送信しており，アイトラッキングデバイスを用いた場合については評価されていない．

3.　解決すべき課題

遠隔操作システムに FI 処理を適用する場合，遠隔操作者の注視点情報取得，注視点情報のネットワーク伝送，注視点情報を利用した FI 処理，エンコード処理，ストリームのネットワーク伝送，デコード処理という順番で処理が実行される．そのため，遠隔操作者の注視点情報を取得してから遠隔操作者がデコード映像を見るまでタイムラグ（システムの遅延時間）が存在する．このシステムの遅延時間が大きい場合，FI 処理がデコード映像に反映されるのが遅くなるため主観画質に悪影響を与えるが，文献 [3], [4], [5], [6], [7], [8] では，注視点情報の取得を含めてリアルタイム動作する評価環境を構築しておらず，システムの遅延時間を考慮した評価ができていないという課題があった．

また，文献 [5] では，フル HD 映像を対象に CPU で FI 処理とエンコード処理をしているが，1 フレーム当り約 166 ms の処理時間がかかっており，CPU ではリアルタイム処理が難しいという課題があった．

本論文では，リアルタイム動作するシステムを実現するため，ネックとなっていた FI 処理とエンコード処理を Field-Programmable Gate Array（FPGA）に実装する．さらに，実際にリアルタイム動作する映像圧縮伝送システムを構築し，有効性を評価する．

GPU ではなく FPGA を選定した理由は，主観画質への影響を少なくするためには処理遅延を少なくすることが必要であり，パイプライン処理できる FPGA の方が低処理遅延で実装できると考えたからである．

4.　視覚特性に基づく映像圧縮伝送システム

図 2 に構築した視覚特性に基づく映像圧縮伝送システムの構成を示す．

遠隔操作対象の機器がある遠隔地は，FI 処理部と High Efficiency Video Coding（HEVC）[9] エンコーダを実装した FPGA ボード，任意の映像入力ができる再生装置およびビデオカメラで構成した．FPGA ボードへの映像入力は，High-Definition Multimedia Interface（HDMI）インタフェースを採用し，再生装置とビデオカメラを評価項目に応じて切り替えることができる構成とした．

一方，遠隔操作者がいる操作室は，PC，視線情報を取得するアイトラッキングデバイス，遠隔地の映像を表示するディスプレイで構成した．PC 上では，注視点情報を FPGA ボードに送信する注視点送信ソフトとネットワークから受信したストリームをデコード処理する HEVC デコーダが動作する．

また，PC と FPGA ボードの間は無線ではなく有線の Local Area Network（LAN）ケーブルで接続した．これは，ネットワーク上のパケットエラーや遅延ばらつきによる画

© 2018 Information Processing Society of Japan

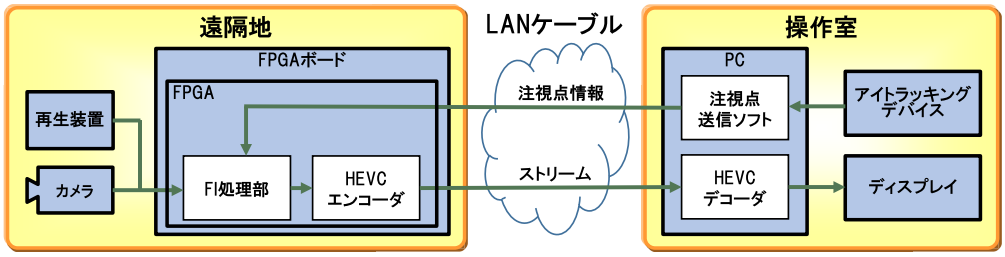


図 2　視覚特性に基づく映像圧縮伝送システム
Fig. 2　Video coding transmission system based on human vision properties.

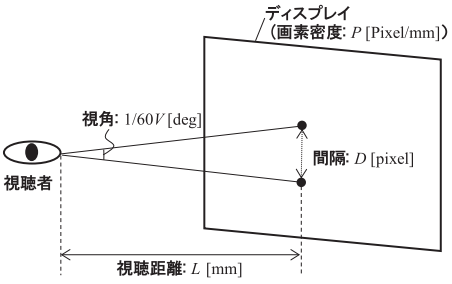


図 3　視力の定義
Fig. 3　Define of visual sensitivity.

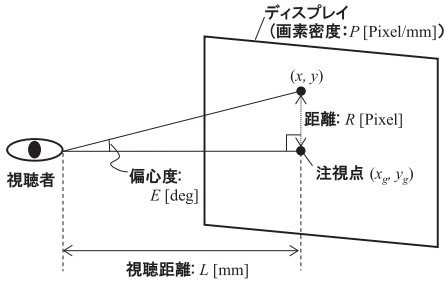


図 4　偏心度の定義
Fig. 4　Define of eccentricity.

質評価への影響を排除するためである．LAN ケーブル上は，PC から FPGA ボードに注視点情報が送られ，FPGA ボードから PC には，HEVC でエンコード処理したストリームが伝送される．

以降，5 章では FI 処理部の処理内容，6 章では FPGA 実装の詳細，7 章では PC で動作するソフトウェアについて述べる．

5.　Foveated Imaging 処理

5.1　視力と遮断周波数の関係

視力 V とは，2 つの物を分離して見ることができる最小の間隔のことである．実際には，この間隔の視角を分（度の 60 分の 1）で表し，その逆数が視力として定義される．すなわち，判別可能な最小の間隔が視角 1 分なら視力 1.0，視角 5 分なら視力 0.2 ということになる [10]．この定義をふまえ，**図 3** に示した視聴距離 L，ディスプレイ画素密度 P，視角 $1/60V$ における，視聴者が判別する必要のある間隔 D を求めると以下で表される．

$$D = 2L\tan\left(RADIANS\left(\frac{1}{120V}\right)\right)\cdot P \tag{1}$$

ここで，本論文で扱う最低視力は 0.1 程度であり，式 (1) の $1/120V$ は 1 より小さくなるため $\tan\theta \approx \theta$ で近似する．

$$D = \frac{LP}{3438V} \tag{2}$$

一方，視力の定義を言い換えると，間隔 D 以下の隙間は，判別できなくてよいと解釈できる．そこで，この間隔 D の周波数を遮断周波数 f_c と考えると，視力に対する遮断周波数の関係式が得られる．

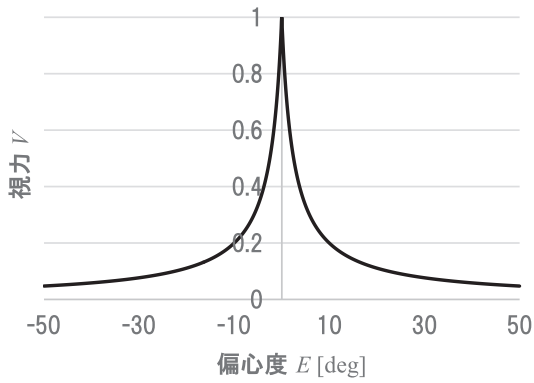


図 5　偏心度視力モデル
Fig. 5　Eccentricity-visual sensitivity model.

$$f_c = \frac{1}{2D} = \frac{1719V}{LP} \tag{3}$$

5.2　偏心度と視力の関係

偏心度 E とは，**図 4** に示されるように，注視点の向きを基準とした角度のことである．一般的に人間の目は注視点から離れる（偏心度が大きくなる）にしたがって視力は低下する．この特性は以下の式でモデル化されている [11]．

$$V(x,y) = \frac{V_f}{1 + E/2.5} \tag{4}$$

ここで，$V(x,y)$ はディスプレイ上の位置 (x,y) の視力，V_f は注視点 (x_g, y_g) の視力を示しており，本論文では一般的な視力として，$V_f = 1.0$ とした偏心度視力モデル（**図 5**）を用いて議論する．

© 2018 Information Processing Society of Japan

表 1　ガウシアンフィルタの遮断周波数
Table 1　Cut-off frequency of Gaussian filters.

タップ数	遮断周波数 f_c
3tap	0.181
5tap	0.129
7tap	0.107
9tap	0.093
11tap	0.084

表 2　遠隔操作環境の条件
Table 2　Condition of remote control environment.

項目	条件
画素密度 P [pixel/mm]	3.68
視聴距離 L [mm]	700

表 3　注視点からの距離
Table 3　Distance from gaze point.

タップ数	注視点からの距離 R[pixel]
3tap	303
5tap	474
7tap	599
9tap	712
11tap	805

5.3　偏心度視力モデルの近似

文献 [4], [5] は，図 5 の偏心度視力モデルを忠実に再現するために画素ごとにローパスフィルタの係数を変更している．FPGA で画素ごとに係数を変更する場合，巨大な係数テーブルを持ち，それを切り替えて使う必要があり現実的ではない．そこで，この係数テーブルの容量を削減するために，5 種類（3tap，5tap，7tap，9tap，11tap）の 2 項分布型ガウシアンフィルタを使用し，注視点から適切な距離でタップ数の多いガウシアンフィルタに徐々に切り替える方式を実装した．

具体的に各ガウシアンフィルタを切り替える距離を算出する．まず，図 4 より，距離 R と偏心度 E の関係式を求める．

$$R = L\tan(RADIANS(E)) \cdot P \tag{5}$$

次に，式 (3) と式 (4) を代入すると以下が得られる．

$$R = LP\tan\left(RADIANS\left(\frac{4298}{f_c LP} - 2.5\right)\right) \tag{6}$$

式 (6) に，あらかじめ計算した各ガウシアンフィルタの遮断周波数（**表 1**）と，遠隔操作環境の条件（**表 2**）を代入して，各ガウシアンフィルタの注視点からの距離 R を計算すると，**表 3** の結果が得られる．

本論文では，表 3 の距離 R をそのガウシアンフィルタへ切り替えるポイントとすることにより，図 5 の偏心度視力モデルを近似した．なお，注視点の座標を (x_g, y_g) としたときの距離 R は $\sqrt{(x_g - x)^2 + (y_g - y)^2}$ で求めており，**図 6** に示すように徐々にタップ数の多いガウシアンフィルタに切り替えている．

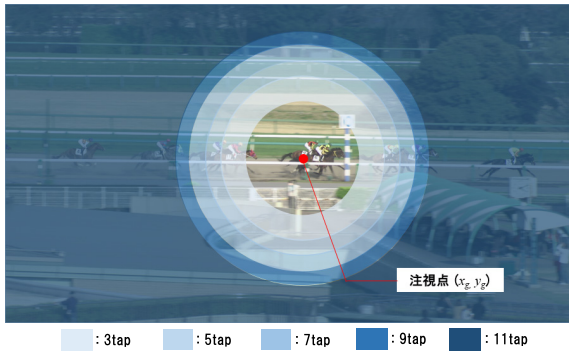


図 6　ガウシアンフィルタ切り替えの例
Fig. 6　Gaussian filters switching example.

6.　FPGA 実装

6.1　回路構成

FPGA の全体回路構成を**図 7** に示す．HDMI 入力部は，FPGA 外部の HDMI インタフェースからラスタスキャン順に入力される映像を FPGA の内部動作クロックに乗せかえる機能と，フレーム開始を示す垂直同期信号（Vsync）を生成する機能を持つ．FI 処理部は，HDMI 入力部経由で入力された映像に対して，5 章に記載した FI 処理を行う．HEVC エンコーダは，FI 処理後の映像をエンコード処理してストリームを生成する．User Datagram Protocol（UDP）パケット送信部は，ストリームを UDP パケット化して，LAN ケーブルに出力する．

ここで，HEVC エンコーダのスペックを**表 4** に示す．解像度 1,920 × 1,080 画素，YUV 4:2:0 8 bit，フレームレート 30 fps の映像をリアルタイムエンコードできる HEVC エンコーダを使用した．また，低処理遅延でエンコードするために画像を溜めてエンコードする必要がある B ピクチャは使用せず，初めのフレームのみ I ピクチャでエンコードし，その後は P ピクチャでエンコードした．ただし，ネットワーク上でパケットロスしたとき，デコード映像が破綻した状態から復帰するためにイントラブロックの帯を巡回させるイントラリフレッシュ [12] を採用している．

一方，UDP パケット受信部で受信した注視点情報のパケットは，デパッキングして UDP パケット受信部内にある注視点レジスタに上書き記憶する．FI 処理部は，フレーム先頭を示す Vsync で注視点レジスタの値をラッチして保存し，1 フレーム期間はこの保存した注視点情報を用いる．これにより，注視点情報を送信するタイミングと FI 処理部を開始するタイミングの同期をとる必要がないようにしている．

6.2　動作タイミングの設計

FPGA での処理遅延を極力短くするため，**図 8** に示すように各処理が最小限の処理遅延で動作するように論理を設計した．

© 2018 Information Processing Society of Japan

HDMI 入力部からラスタスキャン順に出力された映像は，FI 処理部で 10 ライン分バッファリングし，11 ライン目とバッファリングした 10 ライン分の画素を用いて最大 11tap のフィルタ処理を行う．ここで，FI 処理部の 10 ラインのバッファの内訳は，フィルタ処理前に画素を保持する 5 ラインと，フィルタ処理後に画素を保持する 5 ラインである．そのため，FI 処理部の遅延時間は 5 ラインとなる．

HEVC エンコーダは，16 ライン（1Coding Tree Unit (CTU)）分の映像が入力されたら処理を開始する．このエンコーダは，パイプライン処理化されており，1CTU ごとにストリームが出力される．

UDP パケット送信部は，ストリームが 64 byte 溜まったら，すぐに UDP パケットを生成してネットワーク上に出力する．このため，遅延時間は無視できる．

以上のように FPGA に映像が入力されてから，ネットワーク上にストリームが出力されるまでの遅延時間は，合計 21 ライン（約 0.7 ms）程度となるように設計した．

6.3　実装規模

本論文では，Xilinx 社の Virtex Ultrascale FPGA を使用した．実装規模は，179 KLUTs，592 Block RAMs，315 DSPs で実装できた．

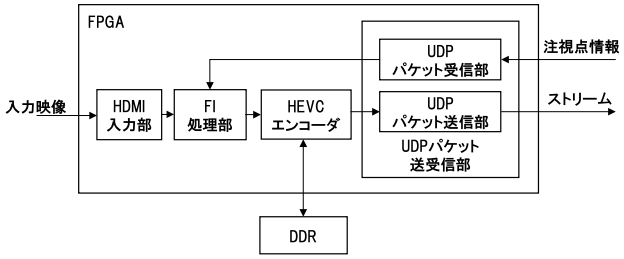


図 7　FPGA 全体回路構成
Fig. 7　Block diagram of the FPGA.

表 4　エンコーダのスペック
Table 4　Encoder specification.

項目	パラメータ
解像度	1920×1080 画素
フォーマット	YUV 4:2:0 8bit
フレームレート	30fps
ピクチャ構造	IPPPP…(イントラリフレッシュ)

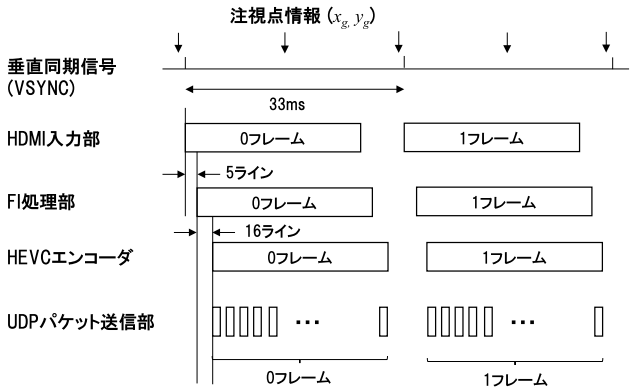


図 8　各処理の動作タイミング
Fig. 8　Operation timing of each process.

7.　PC で動作するソフトウェア

7.1　注視点送信ソフト

アイトラッキングデバイスに付属の SDK を用いて注視点情報を取得し，UDP パケットを生成し，FPGA へ UDP パケットを送信するプログラムを作成した．注視点情報の送信間隔は，エンコード開始と注視点情報の同期をとらなくてもタイムラグを小さくできるように 1 フレームのエンコード時間 33 ms の半分の約 16 ms 間隔で送信している．

7.2　HEVC デコーダ

オープンソースとして配布されている VLC メディアプレーヤ [13] のネットワークストリーム再生機能を用いて実現した．VLC メディアプレーヤのネットワークプロトコルの設定で，使用する UDP パケットのポート番号を指定することで再生可能となる．その際，デコーダ側の受信バッファのキャッシュ設定は，バッファアンダーフローが発生しない 200 ms を設定した．

8.　評価環境および評価条件

実機評価環境を図 9 に示す．この実機評価環境は，表 5 に記載した機材を用いて構築している．

また，評価条件として，表 2 に記載の遠隔操作環境の

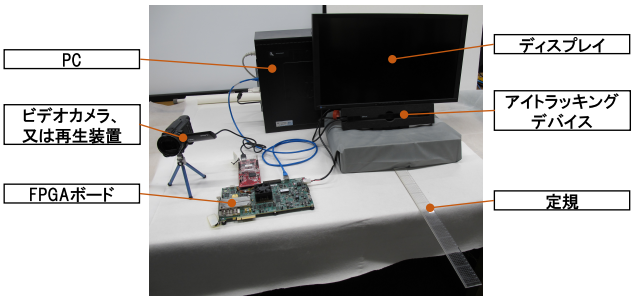


図 9　実機評価環境
Fig. 9　Evaluation environment.

表 5　評価環境の仕様
Table 5　Specification of evaluation environment.

機材名	仕様
PC	OS：windows 7 professional 64bit CPU：Intel core i7-3770K 3.5GHz メモリ：16GB
アイトラッキングデバイス	Tobii EyeX
ディスプレイ	EIZO FORIS FG2421 画面サイズ：23.5 インチ 画面解像度：1920×1080 画素 画素密度：3.68pixel/mm
再生装置	Blackmagicdesign Intensity Pro 4K
ビデオカメラ	SONY FDR-AX55
FPGA ボード	Xilinx VCU 108 評価ボード
LAN ケーブル	2m

© 2018 Information Processing Society of Japan

条件を用いる．ディスプレイは，画素密度 3.68 pixel/mm の 23.5 インチディスプレイ，視聴距離は，ディスプレイの高さの約 2 倍の 700 mm，LAN ケーブルは，2 m とした．さらに，ディスプレイと評価者の視聴距離を一定にするため，ディスプレイ前にある定規を腹部に当てた状態で評価した．評価内容は，システムの遅延時間，主観画質，および伝送ビットレートの削減効果の 3 項目を評価する．

9. 評価方法

9.1 システムの遅延時間

映像入力にビデオカメラを用いて End-to-End の遅延時間の測定を行う．具体的には，ストップウォッチとディスプレイに表示されるデコード処理後のストップウォッチの両方が写るようにビデオカメラで撮影し，デコード画像をキャプチャする．そして，キャプチャしたデコード画像に写った 2 つのストップウォッチの時刻の差分で遅延時間を測定する．

9.2 主観画質

主観画質の評価は，ITU-T 勧告 P.910 [14] に記載の Degradation Category Rating（DCR）法を用いて，基準映像に対する提案映像の劣化度合いを評価する．

図 10 に DCR 法の評価手順を示す．1 つの映像の評価は，基準映像 10 秒，インターバル 2 秒，提案映像 10 秒，評価時間 10 秒の順番で実施する．評価尺度は，**表 6** に記載の 5 段階で評価を実施する．また，評価点が 4 以下の場合は，映像ごとに劣化理由のヒアリングを実施した．

評価人数は，評価の安定性を高めるため映像関係の経験者 15 名とした．評価映像は，ハイビジョン・システム評価用標準画像第 2 版 [15] の 4 種類（S204，S208，S210，S218）の映像の先頭から 300 フレームを評価対象とした（**図 11**）．

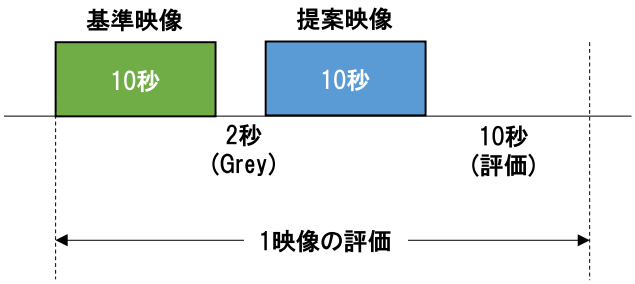


図 10　DCR 法の評価手順
Fig. 10　Evaluation procedure of DRC method.

表 6　評価尺度
Table 6　Evaluation score.

評価点	評定語
5	劣化が認められない
4	劣化が認められるが気にならない
3	劣化が認められ，わずかに気になる
2	劣化が認められ，気になる
1	劣化が認められ，非常に気になる

この 4 種類の評価映像は，遠隔操作で重要な 3 つの視点移動（連続的な遅い視点移動，連続的な早い視点移動，瞬間的な視点移動）を網羅できるように選択した（**表 8**）．

HEVC エンコーダの画質制御に関しては，注視点付近の画質を同等画質で評価するために，すべてのフレームで固定の Quantization Parameter（QP）値を使用する固定 QP 方式を採用した．QP 値は，システムの遅延時間を考慮した FI 処理の主観画質の評価を主目的とするため，一般的に圧縮ノイズが目立たないとされる Peak Signal-to-Noise Ratio（PSNR）が 35 dB 以上となる QP 値を評価映像ごとに使用した（**表 7**）．

また，適切に基準映像と提案映像を比較するためには注視点を同一にする必要がある．そこで，評価者には，評価前に表 8 の注視ポイントを伝えて，注視ポイントを追従してもらうことで，基準映像と提案映像ともに同一注視点での評価を可能にした．

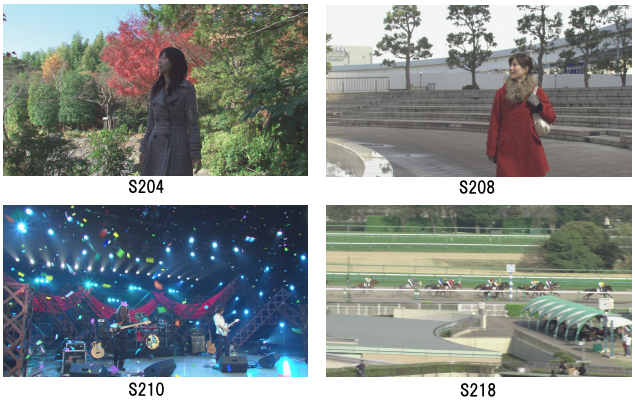


図 11　評価映像
Fig. 11　Evaluation videos.

表 7　エンコード QP 値
Table 7　Encoded QP value.

映像名	QP 値
S204	25
S208	32
S210	32
S218	36

表 8　評価映像の注視ポイント
Table 8　Gaze points of evaluation videos.

映像名	注視ポイント	視点移動の種類
S204	歩いている女性の顔を追従する．	連続的な遅い視点移動
S208	歩いている女性の顔を追従する．	連続的な遅い視点移動
S210	黒服ギタリスト，白服ギタリスト，ドラマーの順で約 3 秒毎に切り替えて見る．	瞬間的な視点移動
S218	画面右から左へ先頭の馬から順に注視点を移動させる．シーンチェンジ後は，場群中央にいる 10 番の馬を見る．	連続的な早い視点移動，瞬間的な視点移動

© 2018 Information Processing Society of Japan

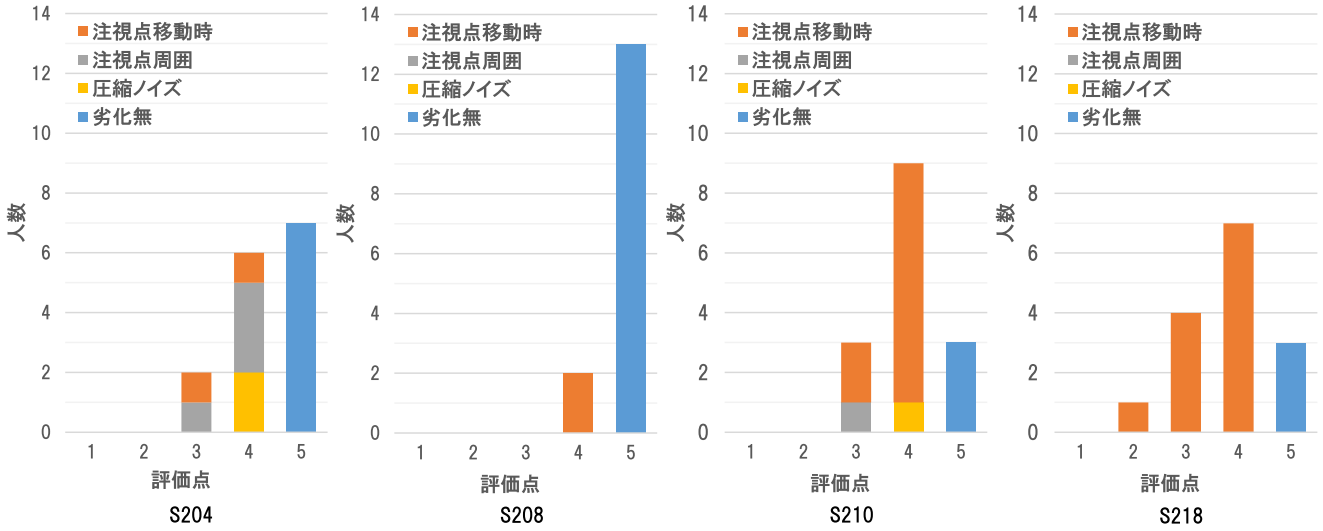


図 12　主観画質評価の人数分布と劣化理由
Fig. 12　Distribution of head count and degradation factors.

9.3　伝送ビットレートの削減効果

主観画質評価と同一の評価映像とエンコード QP 値を使用し，ビットレートの削減効果を評価する．この評価は，図 9 の実機評価環境ではなく，FPGA に実装した HEVC エンコーダと同一動作のソフトウェアを用いて実施した．注視点情報については，表 8 の注視ポイントに基づきフレームごとに注視点情報 (x_g, y_g) を手動で作成した．

10.　評価結果および考察

10.1　システムの遅延時間

9.1 節の方法でシステムの遅延時間を測定した結果，遅延時間は 0.37 秒であった．遠隔操作で操作可能な遅延時間の限界は 1.5 秒といわれており [1]，本論文のシステムは遠隔操作可能な遅延時間の要件は満たしたといえる．

遅延時間の内訳については，FPGA の処理時間が 0.7 ms 程度であるため，VLC メディアプレーヤの受信バッファ制御を含むデコード処理が大部分を占めている．システムの遅延時間を短くするためには，デコード処理の高速化が必要であると考える．

10.2　主観画質

10.2.1　DMOS の評価結果

表 9 に各映像の Degradation Mean Opinion Score (DMOS) を示す．4 映像の DMOS を平均すると 4.25 と劣化が認められるが気にならないレベルの主観画質を実現できた．また，4 映像を個別に見てもすべての映像で DMOS の「許容限」とされている 3.5 を超えており，有効性を確認できた．

10.2.2　考察

図 12 に主観評価結果の人数分布と劣化理由を示す．劣化理由は，主観評価時のヒアリング結果を基に「注視点移動時の解像度劣化」，「注視点周囲の解像度劣化」，「注視点付近の圧縮ノイズによる劣化」の 3 種類に分類している．それぞれについて考察する．

表 9　主観画質評価の結果
Table 9　Results of subject image quality evaluation.

映像名	DMOS
S204	4.33
S208	4.87
S210	4.0
S218	3.8
平均	4.25

(a)　注視点移動時の解像度劣化に関して

図 12 より，S210，S218 で注視点移動時に解像度劣化を感じた人数が S204，S208 に比べて多いことが分かる．S204，S208 は，ゆっくり歩いている女性を追っており視点の移動速度が遅いため解像度劣化を認識できなかったが，S210，S218 は，視点の急な切り替えや右から左への早い視点移動が発生しており，システムの遅延時間の影響が現れたからであると考える．

本論文で構築したシステムは，DMOS が 3.5 以上のため実用上は問題ないといえるが，無線ネットワークにした際，システムの遅延時間が大きくなるため問題となる可能性がある．現行の LTE-Advanced 回線の場合，ネットワーク遅延は 10 ms 程度 [17] といわれており，影響は少ないと考えるが，解像度劣化が生じた場合は，構築したシステムで遅延時間の大部分を占めているデコード処理の高速化を検討する必要があると考える．

(b)　注視点周囲の解像度劣化に関して

FPGA に実装した FI 処理の妥当性を評価することができる．その理由は，図 5 の偏心度視力モデルに基づいて，注視点周囲の解像度を削減しており，注視点周囲の解像度

© 2018 Information Processing Society of Japan

図 13　再生画像
Fig. 13　Reconstructed images.

劣化が確認できるということは，FPGA に実装した FI 処理に問題があるということを意味するからである．

図 12 より，注視点付近が比較的複雑な絵柄の映像（S204，S210）で注視点周囲の解像度劣化の指摘が 5 件あった．これは，ガウシアンフィルタを用いて高周波成分のカットしたことよる画質の変化が S208，S218 に比べて大きかったため，劣化を認識できたと考える．ただし，DMOS が 3 以上とわずかに気になるレベルに解像度劣化を抑えられており，実用上は問題ないと考える．

また，本論文では，FPGA 実装を容易にするために，5 種類のフィルタ係数を注視点からの距離で切り替える方式を採用したが，わずかに気になるレベルに解像度劣化を抑えられており，画素ごとにフィルタ係数を切り替える既存手法 [4], [5] と比べて遜色がないと考える．

© 2018 Information Processing Society of Japan

表 10 伝送ビットレート削減効果
Table 10 Reduction effect of video communication bandwidth.

映像名	基準映像 [Mbps]	提案映像 [Mbps]	削減率[%]
S204	21.27	11.76	44.7
S208	6.54	5.29	19.1
S210	10.29	8.51	17.3
S218	3.04	2.76	9.2

(c) 注視点付近の圧縮ノイズによる劣化に関して

注視点付近の圧縮ノイズによる劣化の指摘は，3 件あった．PSNR が 35 dB 以上となる QP 値を用いて，ガウシアンフィルタをかけていない注視点付近を同等画質で評価したのにもかかわらず，圧縮ノイズが認識された．これは，圧縮処理の原理上，QP 値を同一にしても画質が完全に同一になることはなく，局所的に見るとわずかな違いがあるからだと考える．

本論文では，PSNR が 35 dB 以上となる QP 値を用いて，圧縮ノイズによる劣化の影響を極力排除して評価した．しかし，実際のシステムでは，ネットワークの伝送帯域に合わせて指定の伝送ビットレートにしなければいけないため，レート制御（QP 値を動的に変更）を行う必要がある．実用化に向けて，圧縮ノイズが目立たないレート制御についても，検討する必要があると考える．

10.3 伝送ビットレートの削減効果

表 10 に各映像の伝送ビットレートの削減効果を示す．表 10 より，9.2〜44.7%の伝送ビットレートの削減効果が確認できた．

伝送ビットレートの削減効果が映像によって大きく異なっているが，その要因は 2 つあると考える．1 つ目は，各映像の絵柄の複雑さの違いである．**図 13** に基準映像と提案映像の 1 フレーム目の再生画像を示す．提案映像に示した赤色の点は注視点である．S204 は全体的に高周波成分の多い複雑な絵柄の映像なので，注視点周囲の解像度劣化が他の映像と比べて大きいことが分かる．このように，高周波成分が多い複雑な絵柄であるほど，ガウシアンフィルタにより大幅に高周波成分がカットされるので，伝送ビットレートの削減効果も大きくなる．2 つ目は，各映像で注視ポイントが異なる点である．たとえば，画面中央を注視しているときは，フィルタ処理をしない領域が最も大きくなり，伝送ビットレートの削減効果が小さくなる．一方，画面隅を注視するときは，画面中央に比べてフィルタ処理をしない領域が小さくなるため，伝送ビットレートの削減効果が大きくなる．

上記どちらの要因が，大きく影響をしているかを区別するため，すべての映像で画面中央に注視点を固定した場合の伝送ビットレートの削減効果を確認した．その結果，表 8 の注視ポイントを見る場合と比較して，最大で 3.2%の差異となった．この差異は，9.2〜44.7%と比べて小さいため，表 10 の伝送ビットレートの削減効果の違いは，映像の絵柄の複雑さの違いの方が大きく影響していると考える．

本論文では解像度劣化を感じさせずに伝送ビットレートを削減することを目的にして議論したが，伝送ビットレートの削減を重視して，解像度劣化を許容する方法もある．その場合でも，本方式を用いると視覚的に重要度の低い視野周辺のガウシアンフィルタの標準偏差 σ をより大きくする，ガウシアンフィルタを切り替える距離を短くするなどして，極力解像度劣化を感じさせない方法で伝送ビットレートの削減効果を高めることが可能となる．このように，本方式は視覚特性を利用して伝送ビットレートを削減できるので，注視点を考慮せずに画像全体を一律に画質劣化させることで伝送ビットレートを削減している既存の遠隔操作システム [1], [2] に比べて優位性があると考える．

11. まとめ

本論文では，フル HD 映像を対象に主観的な解像感を維持しつつ伝送ビットレートを削減するために，エンコード前の映像に FI 処理を適用した映像圧縮伝送システムを構築し，その有効性を評価した．その結果，DMOS の平均値が 4.25 と画質劣化が気にならないレベルの主観画質を維持しつつ，9.2〜44.7%の伝送ビットレートの削減効果を確認できた．

今後は，実用化に向けて，圧縮ノイズが目立たないレート制御，デコード処理の高速化，無線ネットワークでの評価を行う予定である．

商標について

HDMI は，HDMI Licensing LLC の商標または登録商標です．Xilinx，Virtex は，Xilinx Inc. の商標または登録商標です．windows は，Microsoft Corporation の商標または登録商標です．Intel Core i7 は，Intel Corporation の商標または登録商標です．EIZO，FORIS は，EIZO（株）の商標または登録商標です．Blackmagicdesign は，Blackmagic DesignPty, Ltd. の商標または登録商標です．SONY は，SONY（株）の商標または登録商標です．

参考文献

[1] 新田恭士，松尾 修，北原成郎，黒田 昇，田村圭司，下田孝徳：長遠距離無人化施工技術の適用性に関する考察—雲仙普賢岳における超長距離遠隔操作実証実験の概要，建設ロボットシンポジウム論文集，Vol.13, pp.41–51 (2012).

[2] 柏尾知明，松木剛志，出口幹雄，白井みゆき，占部弘治，栗原義武，今井伸明：拡張現実感を用いたロボットのリアルタイム遠隔操作システムの開発—あかがね工業博 2012 トレジャーハンターロボットへの応用例，計測自動制御学会論文誌，Vol.50, No.4, pp.383–385 (2014).

[3] Kortum, P. and Geisler, W.: Implementation of a foveated image coding system for image bandwidth reduction, *Proc. SPIE*, Vol.2657, pp.350–360 (1996).

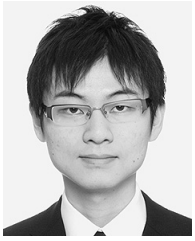
[4] da Costa, A.L.N.T. and Do, M.N.: A Retina-Based

© 2018 Information Processing Society of Japan

Perceptually Lossless Limit and a Gaussian Foveation Scheme with Loss Control, *Proc. IEEE Journal of Selected Topics in Signal*, Vol.8, No.3, pp.438–453 (2014).
[5] 佐藤拓杜，谷田部祐介，伊藤浩朗，中村克行，小味弘典：人間の知覚特性に基づく高効率映像圧縮技術の検討，第31回信号処理シンポジウム講演論文集，pp.93–94 (2016).
[6] Wang, Z. and Bovik, A.C.: Embedded Foveation Image Coding, *IEEE Trans. Image Processing*, Vol.10, No.10, pp.1397–1410 (2001).
[7] Lee, S. and Bovik, A.C.: Fast Algorithms for Foveated Video Processing, *IEEE Trans. Circuits and Systems for Video Technology*, Vol.13, No.2, pp.149–162 (2003).
[8] Geisler, W.S. and Perry, J.S.: A real-time foveated multiresolution system for low-bandwidth video communication, *Proc. SPIE*, Vol.3299, pp.294–305 (1998).
[9] ITU-T H.265: High Efficiency Video Coding (HEVC) (12/2016)
[10] 横澤一彦：視覚科学，勁草書房，p.9 (2010).
[11] J.M. フィンドレイ，I.D. ギルクリスト，本田仁視（監訳）：アクティヴ・ビジョン，北大路書房，pp.15–16 (2006).
[12] 溝添博樹，岡田光弘，小味弘典，佐々本学，羽鳥好律：コンシューマ用途向け超低遅延 H.264 符号化制御アルゴリズムおよびシステム，情報処理学会論文誌コンシューマ・デバイス&システム，Vol.3, No.2, pp.1–9 (2013).
[13] VideoLAN Organization: VLC media player, available from ⟨https://www.videolan.org/⟩ (accessed 2017-08-31).
[14] ITU-T P.910: Subjective video quality assessment methods for multimedia applications, pp.7–8 (2008).
[15] 映像情報メディア学会：ハイビジョン・システム評価用標準画像，第 2 版 (2009).
[16] 杉野　勲：IoT 時代に向けた移動通信政策の動向，総務省移動通信課，pp.25–29, 入手先 ⟨http://www.soumu.go.jp/main_content/000454145.pdf⟩（参照 2018-01-27）.
[17] Chen, W.: Vehicular Communications and Networks, Woodhead Publishing Series in Electronic and Optical Materials, p.199 (2015).

岡田 光弘　（正会員）

2004 年東京理科大学理工学部電気工学科卒業．2006 年同大学大学院修士課程修了．同年（株）日立製作所に入社．画像の高効率符号化および IT プラットフォームに関する研究に従事．



佐藤 拓杜　（正会員）

2013 年東北大学工学部情報知能システム総合学科卒業．2015 年同大学大学院修士課程修了．同年（株）日立製作所に入社．画像処理およびセンサデータからの行動解析に関する研究に従事．

稲田 圭介

1993 年大阪大学大学院工学部通信工学科卒業．1995 年同大学大学院工学研究科通信工学専攻修士課程修了，同年（株）日立製作所入社，現在に至る．映像信号および符号化処理技術の研究開発に従事．映像情報メディア学会会員．

谷田部 祐介　（正会員）

1999 年東京理科大学理工学部電気工学科卒業．2001 年同大学大学院理工学研究科修士課程修了．2018 年大阪府立大学大学院工学研究科電気・情報系専攻博士課程修了．現在，（株）日立製作所に勤務．画像圧縮伸張技術，映像認識技術の研究開発に従事．映像メディア学会会員．

伊藤 浩朗

1990 年横浜国立大学工学部電子情報工学科卒業．1992 年同大学大学院修士課程修了．（株）日立製作所を経て，2017 年に日立オートモーティブシステムズ（株）に入社．自動運転システムの開発に従事．映像情報メディア学会会員．

小味 弘典

1992 年大阪府立大学電気工学科卒業．1994 年同大学大学院修士課程修了．（株）日立製作所を経て，2018 年（株）日立産業制御ソリューションズに入社．映像圧縮伸張技術，映像メディア処理の開発に従事．1999～2000 年南カリフォルニア大学客員研究員．映像情報メディア学会会員．

© 2018 Information Processing Society of Japan