



1. はじめに

映像符号化技術の圧縮率の向上，通信ネットワークの伝送速度の高速化，録画／再生デバイスの性能向上により，高画質映像を配信・記録できるようになってきた．このような映像では，高解像度（例：4K，8K），高フレームレート（例：120 fps，240 fps），高輝度（HDR：High Dynamic Range）が用いられるため，従来の HD（High Definition）映像よりもデータ量が多い．そのため，映像を符号化によりデータ圧縮して伝送・記録することは不可欠であり，符号化後の映像品質を適切に設計・管理することは非常に重要である．

映像品質を設計・管理するためには，主観品質に基づき実施することが基本であり，主観品質評価法が必要となる．ITU-R（International Telecommunication Union Radiocommunication Sector）は，テレビジョン映像を対象とした主観品質評価法を BT シリーズとして勧告を制定し，ITU-T（Telecommunication Standardization Sector）は，IP ネットワークで配信される映像に対する主観品質評価法を P シリーズとして勧告を制定している．映像配信サービスを提供する事業者は，これら主観品質評価法を用い，映像品質を定量化し，テレビジョン映像に対するビットレートを設計してきた．

一方で，昨今利用されているアダプティブビットレート映像配信では，固定のビットレートを用いるのではなく，シーンごとに画質レベルが一定になるようにビット量を割り当てることが多くなってきている．このような場合，主観品質評価を実施し，ビットレートを決定することはできないため，映像信号を用いた客観品質評価技術を用いることが一般的である．客観品質評価技術については，ITU-R が BT シリーズ，ITU-T が J 若しくは P シリーズの勧告として標準化を進めている．

以降，高画質映像を対象とした主観品質評価法及び客観品質評価技術について概説し，今後の展望を述べる．

2. 主観品質評価法

映像の主観品質評価法では，評価対象となる映像を評価者が観視し，体感した品質を評価する．評価結果は，評価に用いる映像ソース，ブース等の観視環境等の影響を受けることから，再現性の高い評価結果を得るために ITU ではこれらの条件を規定した勧告を標準化している．テレビジョン映像を対象とした主観品質評価法は ITU-R 勧告 BT.500⁽¹⁾，IP ネットワークを経由する映像配信等を対象とした主観品質評価法は ITU-T 勧告 P.910⁽²⁾にそれぞれ規定されている．

2.1 観視条件・環境

勧告 BT.500 や勧告 P.910 には，主観評価を安定的に実施するために観視条件・観視環境が規定されている．観視条件として室内照度やモニタの輝度等が規定されており，勧告 BT.500 では観視環境としてラボ環境（laboratory viewing environment）と，実際にユーザが視聴する環境を想定した家庭環境（home viewing environment）が規定されている．ラボ環境では軽微な劣化を検知できるよう，観視条件が規定されている（例：室内

恵木則次 正員 日本電信電話株式会社 NTT ネットワークサービスシステム研究所
E-mail noritsugu.egi.bn@hco.ntt.co.jp
山岸和久 正員 日本電信電話株式会社 NTT ネットワークサービスシステム研究所
E-mail kazuhisa.yamagishi.vf@hco.ntt.co.jp
増田征貴 正員 日本電信電話株式会社 NTT ネットワークサービスシステム研究所
E-mail masataka.masuda.at@hco.ntt.co.jp
Noritsugu EGI, Kazuhisa YAMAGISHI, and Masataka MASUDA, Members (NTT Network Service Systems Laboratories, NIPPON TELEGRAPH AND TELEPHONE CORPORATION, Musashino-shi, 180-8585 Japan).
電子情報通信学会誌 Vol.106 No.1 pp.14-19 2023 年 1 月
©電子情報通信学会 2023

照度はラボ環境の方が家庭環境より低い等)。なお、HDR 映像の観視条件・環境については勧告 BT.2100⁽³⁾ に規定されており、ディスプレイの最小輝度は 0.005 cd/m^2 以下、ピーク輝度は $1,000 \text{ cd/m}^2$ 以上を用いる必要がある。

観視距離については、評価者が好ましいと感じる距離 (PVD: Preferred Viewing Distance) とスクリーンサイズと解像度に基づき設定される距離 (DVD: Design Viewing Distance) が勧告 BT.500 では規定されている。このうち DVD は、隣り合う二つのピクセルに対する視野角が $1/60$ 度となる距離としている。これは、これよりも長い視距離では視力 1.0 の評価者が画素構造を弁別できないためである。そのため、同一のスクリーンサイズであれば解像度が大きいほど、画面の大きさに対して短い視距離、広い視野角が、それぞれ規定されている (表 1, 図 1)。図 1(b) に示すように、8K 解像度の映像評価においては水平方向の視野角が 96° と非常に広くなっており、これにより 8K 映像を観視する評価者は高い臨場感を体感することができる。

2.2 映像ソース

符号化された映像に対する品質は選択した映像ソースに強い影響を受けることから、テレビジョン映像の評価に用いる映像ソースが規定されている。ITU-R 勧告 BT.2245⁽⁴⁾ はテレビジョンシステムにおける画質評価用の映像ソースをまとめたレポートであり、HDTV ($1,920 \times 1,080$) 以上の解像度に対応した映像が含まれている。勧告 BT.2245-1 (2014) は 8K 映像が、勧告 BT.2245-5 (2018) は HDR 映像がそれぞれ規定されている。2022 年 7 月時点の最新版として勧告 BT.2245-9 が制定されており、約 300 種類の映像ソースに関する情報が記述されている。

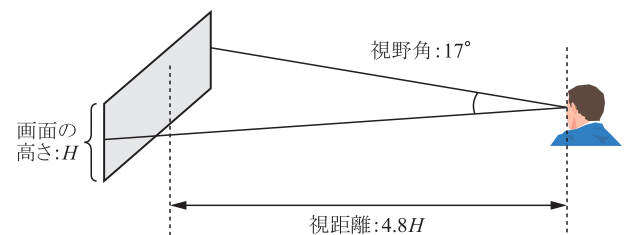
勧告 P.910 には、評価映像の適切な選定を行うための映像の時空間特徴量を示す SI (Spatial Information), TI (Temporal Information) が規定されており、これらの値が広範囲に分布するように映像を選定することが推奨されている。現在は SDR 映像及び 8 bit 深度の映像に適用範囲が限定されていることから、映像品質専門家グループ (VQEG: Video Quality Expert Group) から HDR や 10 bit 深度の映像に対応した SI, TI の算出方法についての提案がなされ、勧告 P.910 の改訂が進んでいる。

2.3 評価者及び評価法

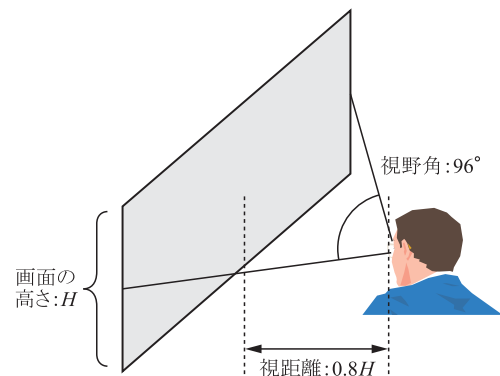
主観品質評価の結果は、映像品質に関する専門家と非専門家では大きく異なることが知られている。そのため、システム上で発生する映像劣化に関する知識を有さない非専門家で、かつ一定の視力、正常な色覚を有する人物に評価を依頼することが一般的である。また、安定

表 1 最適な視聴角, 視距離

解像度	アスペクト比	最適な視野角	最適な視距離 (H : 画面の高さ)
640×480	4 : 3	11°	$7H$
$1,024 \times 768$		17°	$4.5H$
$1,400 \times 1,050$		23°	$3.3H$
$1,280 \times 720$	16 : 9	17°	$4.8H$
$1,920 \times 1,080$		31°	$3.2H$
$3,840 \times 2,160$		58°	$1.6H$
$7,680 \times 4,320$		96°	$0.8H$



(a) 解像度 $1,280 \times 720$ の場合



(b) 解像度 $7,680 \times 4,320$ の場合

図 1 視聴角と視距離の関係 高解像度映像の観視時は相対的に短い視距離、及び広い視野角となる。

した評価結果を得るため、15 名以上の評価者で実験を実施することが推奨されている。

勧告 BT.500 には、評価法として 8 種類の評価法が定義されている。以降では代表的な三つの手法である SS (Single-Stimulus) 法, DSIS (Double-Stimulus Impairment Scale) 法, DSCQS (Double-Stimulus Continuous Quality-Scale) 法について述べる。

(i) SS 法

SS 法による映像サービスの品質評価として最も用いられている評価法は ACR (Absolute Category Rating) 法である。ACR 法では、評価映像を観視し、5 段階品質尺度 (「非常に良い (Excellent), 良い (Good), 普通 (Fair), 悪い (Poor), 非常に悪い (Bad)」) を用い、

評価を行う。ACR 法では品質評価基準が存在しないため、評価結果は実験で取り扱う評価映像の品質の変化範囲（評価実験の枠組み）に強く影響を受ける。そのため、評価実験の枠組みをどの程度に設定するかが重要となる。

(ii) DSIS 法

DSIS 法はテレビジョン映像における典型的な評価方法の一つであり、最初に基準映像を観視した後に評価映像を観視し、5 段階劣化尺度を用いた相対評価を行う (DCR (Degradation Category Rating) 法)。通常は基準映像、劣化映像の順に 1 回映像を観視し評価を行うが、基準映像との軽微な品質差を対象とした評価実験を行う場合には、基準映像、劣化映像の順にそれぞれを 2 回観視した上で評価を行うことが推奨されている。DCR 法では評価の比較対象が存在するため、ACR 法に比べて劣化がより敏感に評価できるため、比較的劣化が小さい評価対象系の評価では ACR 法よりも DCR 法の方が適している。

(iii) DSCQS 法

DSCQS 法はテレビジョン放送に関わるシステムや伝送路の品質評価に用いられる評価法である。基準映像、評価映像のいずれかを映像 A、もう一方を映像 B とし、評価者は映像 A、映像 B の順に映像を観視し、再度映像 A、映像 B の順に映像を観視する中でそれぞれについて図 2 に記載する連続尺度を用いて評価する。このとき、基準映像が映像 A、映像 B のどちらであるかはラ

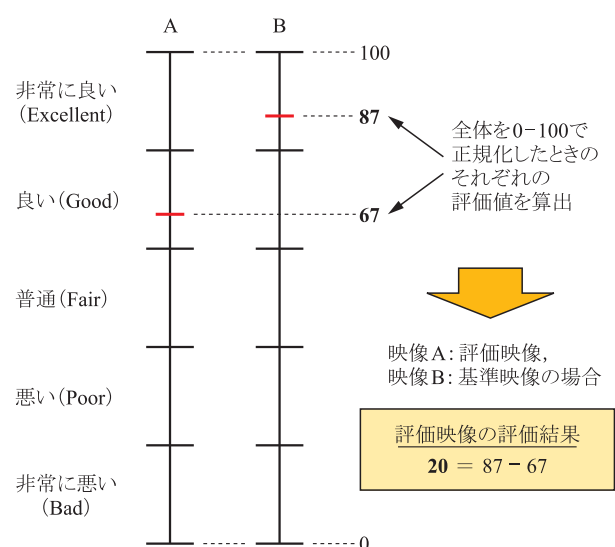


図 2 DSCQS を用いた評価 映像 A、映像 B に対する評価を行い、基準映像（ここでは映像 B）の評価値から評価映像（ここでは映像 A）の評価値を減算した値を、評価映像に対する評価結果とする。

ンダムとし、評価者にはその情報を伝えない。各評価映像に対する評価結果の差分から評価映像の評価結果を算出する。

これらの三つの方法により取得した各評価者の評点については、一貫性のない評価を行っている評価者のスクリーニング、評価傾向が異なる評価者のバイアス補正を実施した上で、評価映像ごとに算出される全評価者の評点に対する平均値を用いる。

評価法をどのように使い分けるかは評価映像のどのような劣化要因を含むかで大別される。符号化劣化の軽微な品質差を識別したい場合には原映像と劣化映像の比較をする DSCQS 法や DCR 法を用いることが一般的である。一方で、ネットワークを経由することで発生するパケット紛失に伴う劣化や再生停止が含まれる映像では ACR 法が用いられることが一般的である。8K や 4K のような高画質映像においても上記観点を考慮し評価法を使い分ける必要がある。

3. 客観品質評価技術

主観品質評価は、観視環境構築のための専用設備の準備、評価者の確保などの観点から多大な時間、労力を必要とする。そのため、評価映像の画素情報から主観品質評価値を推定する客観品質評価技術が多く検討されてきた。本技術は評価映像と原映像を用いる Full-Reference (FR) 法、評価映像と原映像に関する特徴量を用いる Reduced-Reference (RR) 法、評価映像のみを用いる No-Reference (NR) 法の 3 種類に分けられる。

ITU-T SG12 では 4K 映像に対応した客観品質評価技術を勧告 P.1204.4⁽⁵⁾として制定している。以降、本技術について概説する。

3.1 勧告 P.1204.4 の概要

勧告 P.1204.4 は RR 法に分類される技術であり、最大 4K の解像度、60 fps のフレームレートの映像における、プログレッシブダウンロードやアダプティブトリミング等の映像ストリーミングにより伝送された映像の品質推定を行う。図 3 に示すように評価映像の画素情報 (①) のほかに、基準映像の画素情報から抽出した 256 kbit/s 以下のサイド情報 (②) を使用する。ただし、同じ評価映像であっても、TV やスマートフォン等の観視環境により評価結果は異なるため、デバイス関連情報 (③) も使用する。

コアアルゴリズムでは、符号化圧縮等の処理による評価映像の品質劣化量を算出する。時間情報の欠落による品質劣化量 (例：フレームレート低下) は、評価映像情報から抽出したフレームレートや動きの大きさに関する特徴量を用いて算出する。空間情報の欠落による品質劣化量 (例：ブロック雑音やぼけ) は、基準映像と評価映

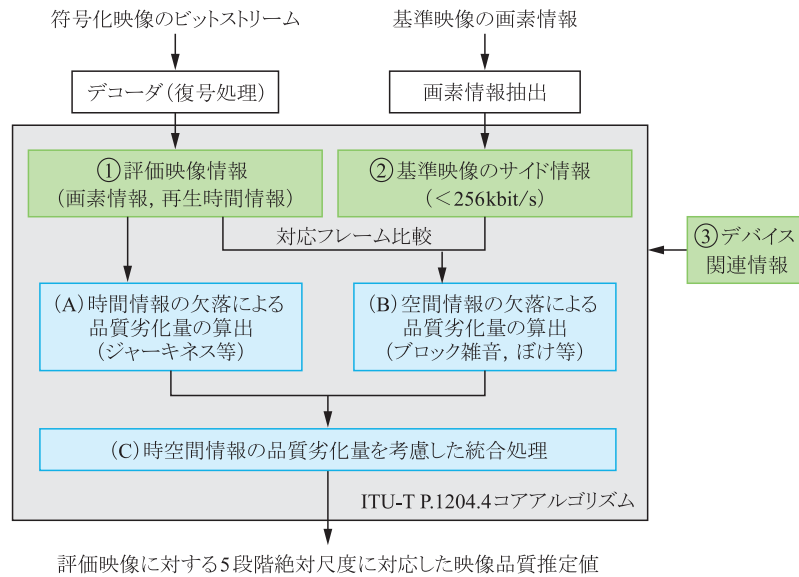


図3 P.1204.4の映像品質推定ブロック図 評価映像と基準映像の情報から算出した時空間情報の欠落による品質劣化量を統合することで、評価映像の品質を推定する。

像との対応フレーム画像の特徴量を比較することで算出する。そして、時空間劣化量を統合することにより、評価映像の品質を推定する。スマートフォンやタブレットに比べ、ディスプレイサイズの大きいPCモニターやTVモニターの方が同一の評価映像に対して厳しく評価されることを考慮し、コアアルゴリズム内の算出式の係数はデバイスにより異なる値が設定されている。

3.2 勧告 P.1204.4 における劣化量算出、統合処理

本節では、コアアルゴリズム内での各品質劣化量算出方法及び前記劣化量の統合処理について概説する。

(A) 時間情報の欠落に関する品質劣化量算出

フレームレートが低いほど、更には評価映像の動きが大きいほど、前後のフレーム間の画像の差分が大きくなる。このような映像を観視した場合、ジャーキネス（フレームレートの低下により動きの滑らかさが失われる現象）により品質が低下する。この品質低下を考慮するために、図4に示すフローにより評価映像情報からフレームレート低下による品質劣化量を算出する。

フレームレートは評価映像情報に含まれる各フレームの再生時間に基づき、単位時間当りのフレーム数を算出する。評価映像の動きに関する特徴量は、各フレームに対応する画像のエッジ情報を用いて算出する。エッジとは、画像内において隣接する画素の輝度が急激に変化する箇所であり、画像解析における物体の輪郭抽出に用いられる。動きが大きい映像では輪郭情報がより変化することから、連続するフレーム画像間でのエッジ情報を比

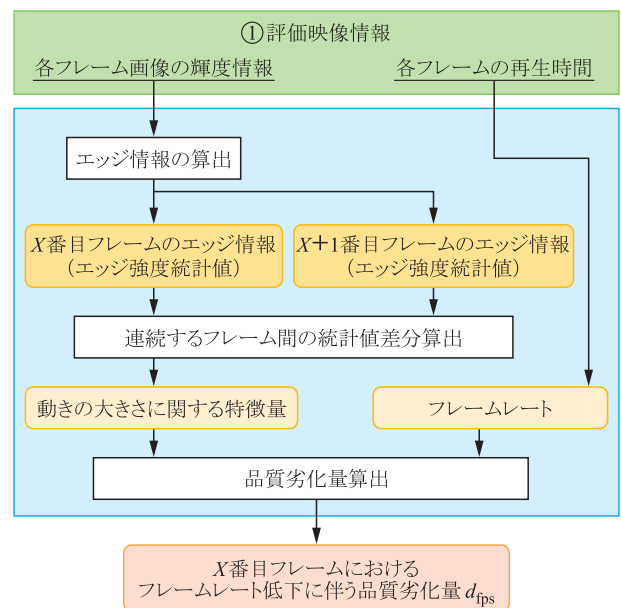


図4 フレームレート低下に伴う品質劣化量の算出 フレームレート、及び連続するフレーム画像間のエッジ強度に関する統計量差分から算出した動きの大きさに関する特徴量から、品質劣化量を算出する。

較し、この平均差分を動きの大きさに関する特徴量として扱う。なお、本アルゴリズムにおけるエッジ情報は、中解像度にリスケージングした評価映像の各フレーム画像に対して、エッジの強度（隣接する画素の輝度差分の大きさ）に基づき算出される統計値としている。

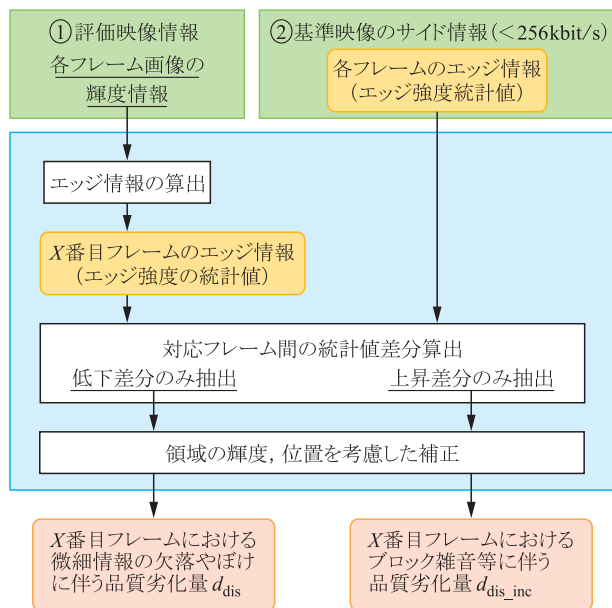


図5 画質低下に伴う品質劣化量の算出 評価映像と基準映像の対応するフレーム間のエッジ強度に関する統計値差分に対して、低下差分／上昇差分における劣化のタイプを考慮し、それぞれの品質劣化量を算出する。

最後に、算出したフレームレートと動きの大きさに関する特徴量を用いて、フレームレート低下による品質劣化量 d_{fps} を 0～1 の範囲で算出する。

(B) 空間情報に関する品質劣化量算出

符号化圧縮により発生するブロック雑音やぼけ、微細な情報の欠落に起因した画質低下に伴う品質劣化量を算出する。空間情報の欠落の検出では、評価映像と基準映像の対応するフレーム画像の画素情報を比較することが望ましいが、RR 型である本アルゴリズムでは利用可能な基準映像のサイド情報の容量は制限されている。この点を踏まえ、1 画像に対してデータ量を抑制した特徴量を用いており、具体的には (A) で記述したエッジ情報を用いている。図5に画質低下に伴う品質劣化量の算出フローを示す。

事前に測定した基準映像のエッジ情報がサイド情報に含まれている。そこで、基準映像と評価映像の対応フレームを抽出し、両フレームのエッジ情報を比較することで、各空間領域の画質低下を検出する。エッジに関する統計値の低下は、本来あるべき輪郭情報が失われていることを表しているため、微細情報の欠落やぼけではこの傾向となりやすい。一方、エッジに関する統計値の上昇は、基準映像にはない輪郭情報が発生していることを表しているため、ブロック雑音発生に伴う新たな輪郭の発生ではこの傾向となりやすい。劣化のタイプにより観視者に与える影響は異なることから、エッジに関する統

計値の差分は低下と上昇とで分けて処理され、二つの品質劣化量 d_{dis} , d_{dis_inc} を 0～1 の範囲で算出する。このとき、輝度が大きい領域や画面中央に近い領域は、より観視者が注視する傾向にあることを考慮し、各品質劣化量を算出している。

上記劣化量の算出では中解像度にリスケールした映像を用いるのに対して、これとは別に、高解像度の映像を用いてエッジ強度に関する一つの統計値をフレームごとに算出している。この統計値は画像全体の輪郭の強度を示すシャープネスに近い値であり、例えば評価映像を解像度にダウンスケールした後にアップスケールを行った場合において、画像全体に影響する処理による画質の低下を検出する。基準映像に対する本統計値もサイド情報に含まれており、基準映像と評価映像とでこの統計値を比較する。ここでも、エッジ強度に関する統計値の変動が低下と上昇で劣化のタイプが異なることから、低下と上昇とで分けて処理され、それぞれ 0～1 の範囲で二つの品質劣化量 d_{sharp} , d_{sharp_inc} が算出される。

(C) 各品質劣化量の統合

(A)と(B)で述べた合計五つの品質劣化量について、それぞれ 1 から減算した値を乗算することにより、1 フレームに対する総合品質値が 0～1 の範囲で算出される。最後に評価対象とする時間に対応する全てのフレームの統合品質値に基づき、最終的な映像品質として 5 段階絶対尺度に対応した映像品質推定値を算出する。

4. 今後の展望

本稿では、高画質映像を対象とした主観品質評価法及び客観品質評価技術について、代表的な標準化技術について解説した。

映像品質評価に関するワークアイテムは ITU-T SG12 で検討されている。ワークアイテム名 J.noref は前述した客観品質評価技術の NR タイプに該当する技術、ワークアイテム名 J.src-vq は基準映像そのものが問題ないかを確認するための評価技術をそれぞれ検討しており、いずれも参照する比較対象映像の情報がない中で映像品質推定を実現するモデルの標準化を目指している。ワークアイテム名 J.q-uhd は、UHD (Ultra-High Definition) 映像や HD から UHD にアップスケールした映像に対して、高解像度、ディスプレイサイズの拡大、相対的視距離の短縮、視野角の拡大による知覚品質への影響を考慮した主観／客観品質評価技術が検討されている。いずれも 2023 年の勧告化を目指している。

文 献

- (1) ITU-R Recommendation BT.500-14, "Methodologies for the subjective assessment of the quality of television images," Oct. 2019.
- (2) ITU-T Recommendation P.910, "Subjective video quality assessment methods for multimedia applications," Nov. 2021.
- (3) ITU-R Recommendation BT.2100-2, "Image parameter values for high dynamic range television for use in production and international programme exchange," July 2018.
- (4) ITU-R Recommendation BT.2245-9, "HDTV and UHDTV including HDR-TV test materials for assessment of picture quality," Nov. 2021.
- (5) ITU-T Recommendation P.1204.4, "Video quality assessment of streaming services over reliable transport for resolutions up to 4K with access to full and reduced reference pixel information," Jan. 2020.

(2022 年 7 月 25 日受付 2022 年 9 月 1 日最終受付)



えぎ のりつぐ
恵木 則次 (正員)

平 15 東北大・工・通信卒。平 17 同大学院博士課程前期了。平 17 日本電信電話株式会社入社。以来、オーディオや映像の品質評価技術の研究に従事。現在、同社ネットワークサービスシステム研究所主任研究員。平 20 年度本会学術奨励賞受賞など。



やまぎし かずひさ
山岸 和久 (正員)

平 13 東京理科大・理工・電気卒。平 15 早大大学院博士課程前期了。平 25 同大学院で博士号を取得。平 15 日本電信電話株式会社入社。以来、映像通信サービスの品質評価技術の研究に従事。現在、同社ネットワークサービスシステム研究所通信品質研究グループリーダー。工博。平 18 年度本会学術奨励賞受賞など。



ますだ まさたか
増田 征貴 (正員)

平 9 芝浦工大・工・工業経営卒。平 11 同大学院修士課程了。平 23 東京農工大大学院で博士号を取得。平 11 日本電信電話株式会社入社。以来、IP 電話や IPTV、Web サービス等を対象とした通信品質の評価・設計・管理法の研究に従事。現在、同社ネットワークサービスシステム研究所トラヒックエンジニアリング研究グループリーダー。工博。平 14 年度本会学術奨励賞受賞など。

