

映像通信の品質推定法に関する研究

Quality Estimation Methodology
for Video Communication

2013年7月

山岸 和久

Kazuhisa YAMAGISHI

映像通信の品質推定法に関する研究

Quality Estimation Methodology
for Video Communication

2013年7月

早稲田大学大学院 基幹理工学研究科

山岸 和久

Kazuhisa YAMAGISHI

目次

第1章 序論	1
1.1 背景	1
1.2 主観品質評価法	1
1.3 客観品質推定技術	3
1.3.1 客観品質推定技術の入出力	3
1.3.2 客観品質推定技術の適用シナリオ	4
1.4 客観品質推定技術と主な課題	7
1.4.1 品質設計技術構築に向けた課題	7
1.4.2 ヘッドエンド品質監視技術構築に向けた課題	7
1.4.3 エンドユーザ品質監視技術構築に向けた課題	9
1.5 本論文の構成および概要	9
第2章 テレビ電話サービスに対する品質設計法	11
2.1 まえがき	11
2.2 テレビ電話サービスに対する品質設計モデルのフレームワーク	12
2.3 主観品質評価	14
2.4 主観品質評価特性	15
2.4.1 符号化に対する映像品質評価特性	15
2.4.2 パケット損失に対する映像品質主観評価特性	15
2.5 映像品質設計モデル	16
2.5.1 符号化品質推定モデル	16
2.5.2 パケット損失品質推定部	18
2.6 映像品質設計モデルの数式の検証	20
2.7 映像品質設計モデルの検証	20
2.8 提案モデルを用いた品質設計	22
2.9 まとめ	23
第3章 3D映像サービスに対するヘッドエンド品質監視法	26
3.1 まえがき	26
3.2 主観品質評価	29
3.3 主観品質評価特性	32
3.4 提案客観品質推定モデル	35
3.5 提案モデルの品質推定精度評価	35
3.5.1 品質推定精度に対する要求条件	35
3.5.2 提案モデルのパフォーマンス	36
3.5.3 考察	36

3.6	まとめ	37
第4章	パケット損失パターンを考慮したIPTVエンドユーザ品質監視法	40
4.1	まえがき	40
4.2	提案フレームワーク	42
4.2.1	コンセプト	42
4.2.2	機能	44
4.2.3	モデルの動作	44
4.3	主観品質評価	45
4.4	主観品質評価特性	47
4.4.1	符号化に対する品質評価特性	47
4.4.2	パケット損失に対する品質評価特性	47
4.5	提案モデル	48
4.5.1	符号化品質推定モジュール	48
4.5.2	パケット損失品質推定モジュール	49
4.6	提案モデルのパフォーマンス評価	50
4.6.1	品質推定モジュールに対するパフォーマンス要件	50
4.6.2	符号化品質推定モジュールの品質推定精度	50
4.6.3	パケット損失品質推定モジュールの品質推定精度	51
4.6.4	提案モデルの係数の有効性評価	51
4.6.5	損失映像フレーム種別の影響	53
4.7	まとめ	54
第5章	映像フレームを用いたIPTVエンドユーザ品質監視法	55
5.1	まえがき	55
5.2	従来検討	57
5.3	提案モデルのコンセプト	59
5.3.1	提案モデルの定義	59
5.3.2	提案モデルの機能	61
5.4	主観品質評価	62
5.4.1	実験条件	62
5.4.2	主観品質評価特性	64
5.5	提案モデル	70
5.5.1	符号化に対する品質推定モジュール	70
5.5.2	パケット損失に対する品質推定モジュール	71
5.6	提案モデルのパフォーマンス評価	72
5.6.1	パフォーマンス要件	72
5.6.2	提案モデルの品質推定精度	73
5.6.3	提案モデルの品質推定精度に対する考察	76

5.7 まとめ	78
第6章 結論	82
謝辞	84
参考文献	85
研究業績	93

第1章 序論

1.1 背景

ネットワーク回線の広帯域化や映像コーデックの高圧縮化が進み、IP 網を介した映像通信サービス（例：テレビ電話、IPTV¹）が普及してきた。映像通信サービスの品質は映像の撮影・編集、符号化方式（例：MPEG²-2[1]、H.264[2]）、コーデックの実装、ネットワークの状態、受信端末に実装されるパケット損失隠ぺい処理（PLC³）やディスプレイ表示技術に影響を受ける。この映像処理のフローを映像処理チェーンと呼ぶ。映像処理チェーンにより発生する映像品質劣化は空間的歪と時間的歪に大別される。空間的歪は、モザイク状に発生する歪（ブロック歪）、エッジや色の変化が急峻なエリアで起こるモスキートノイズ、解像度低下による細かい模様や輪郭のぼけ、明るさや色が緩やかに変化するエリアに発生する擬似輪郭が代表的なものである。また、時間的歪は映像の動きの滑らかさが失われ、ギクシャクして見えるジャーキネス、輝度レベルが変化し、ちらつきが生じるフリッカ、再生停止によるフリーズが代表的なものである。ユーザに高い品質の映像通信サービスを提供するためには、これら映像品質劣化を考慮し、サービス提供前にサービス品質を設計することが重要である。加えて、サービス提供中の通信サービスの品質が高いことを監視したり、サービス品質低下時に品質劣化原因を早期に究明することが重要となる。

1.2 主観品質評価法

品質設計および品質監視はユーザが体感する品質に基づき実施されることが理想的であるため、品質を評価する手法が必要となる。映像通信サービスの品質評価は人間が品質を判断する主観品質評価が基本である。主観品質評価法はITU⁴により、ACR⁵[3]、ACR-HR⁶[3]、DCR⁷[3]、PC⁸[3]、DSCQS⁹[4]、SAMVIQ¹⁰[5] 法等が標準化されている。

ITU-T 勧告 P.910 で規定される ACR 法は単一刺激法（Single Stimulus Method）とも呼ばれ、通信サービスの品質評価に最も広く用いられている。評価者は 10 秒程度の評価映像を観視した後、続く 10 秒以内に 5 段階品質尺度（5: 非常に良い、4: 良い、3: 普通、2: 悪い、1: 非常に悪い）により映像品質の評価を行う。評価結果は評価者の各カテゴリへの投票率を評点で重み付けした「平均オピニオン評点（MOS¹¹）」で表す。ACR 法では、非常に悪い品質の映像を観視した後に軽微な品質劣化の映像を観視すると良いと評価されたり、

¹Internet Protocol TeleVision

²Moving Picture Experts Group

³Packet-Loss Concealment

⁴International Telecommunication Union

⁵Absolute Category Rating

⁶Absolute Category Rating with Hidden Reference

⁷Degradation Category Rating

⁸Pair Comparison

⁹Degradation Category Rating

¹⁰Subjective Assessment Methodology for Video Quality

¹¹Mean Opinion Score

逆に非常に良い品質の映像を観視した後に軽微な品質劣化の映像を観視すると悪いと評価されたりする場合がある。このように、映像品質評価値がその前に観視した映像の品質に影響を受ける順序効果があるため、評価映像の提示順序をランダム化して評価者毎に変化させたり、異なる提示順序で繰り返し評価するなどの対処が施される。

ITU-T 勧告 P.910 で規定される ACR-HR 法は、評価に用いる映像コンテンツの違いが評価値に与える影響を取り除くため、ACR 法で得られた評価結果に対して評価映像と基準映像（原映像）の評点の差に「5」を加算し算出（ $DMOS = \{ \text{評価映像の評点} \} - \{ \text{基準映像の評点} \} + 5$ ）する手法である。評価値は DMOS¹²と表現する。ただし、基準映像の品質は、映像品質関連業務に従事する専門家が「5: 非常に良い」あるいは「4: 良い」と判断する品質である必要がある。

ITU-T 勧告 P.910 で規定される DCR 法では、品質評価の基準となる基準映像と評価映像を 10 秒程度ずつ対にして評価者に提示し、続く 10 秒程度の時間内に 5 段階妨害尺度（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）により映像品質の評価を行う。評価者は品質評価の最初に必ず基準映像を観視し、この映像品質を基準として続く評価映像の品質を評価することから、ACR 法で見られる順序効果がある程度抑えられる。評価結果は ACR 法と同様に平均オピニオン評点で表すが、ACR 法で導出した MOS と区別するため、DMOS¹³と呼ばれる。DMOS 評価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。そのため、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。ただし、DMOS 評価では 1 つの評価対象の品質を評価するのに 2 つの映像の観視が必要なため、同じ条件数の評価値を得るためには ACR 法の約 2 倍の時間を要する。DCR 法では、DMOS = 4.5 を「検知限」、DMOS = 3.5 を「許容限」、DMOS = 2.5 を「我慢限」として品質の基準を与えることがある。なお、DCR 法は、二重刺激妨害尺度法（DSIS 法¹⁴）、もしくは EBU¹⁵ 法とも呼ばれることがある。

ITU-T 勧告 P.910 で規定される PC 法では、10 秒程度の評価映像を“対”にして評価者に提示し、後者の評価映像の品質を前者のそれと比較し、続く 10 秒以内にどちらの評価映像の品質が良かったかを判断し、7 段階比較尺度（-3: 非常に悪い、-2: 悪い、-1: やや悪い、0: 同じ、1: やや良い、2: 良い、3: 非常に良い）により評価する。どちらの映像を先に観視し、評価するかにより評価結果に順序効果が現れるため、同じ評価映像の対を用いた評価でも、提示順序を変えて評価を行うことが望まれる。PC 法では評価映像同士を直接比較して評価するため、細かい品質の差を検出することが可能となるが、品質評価結果が相対評価値として得られることや、ACR 法や DCR 法よりも多くの評価時間を要する（評価映像の組合せ数が多い）といった特徴がある。

¹²Differential Quality Score

¹³Degradation MOS

¹⁴Double Stimulus Impairment Scale Method

¹⁵European Broadcasting Union

ITU-R 勧告 BT.500-13 で規定される DSCQS 法は二重刺激連続品質尺度法と呼ばれ、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多く用いられる。DSCQS 法は品質の範囲が幅広く揃わないときに特に有効であり、基準映像と比べた評価映像の品質差と、評価映像の絶対品質を同時に評価できることが特徴である。DSCQS 法では、基準映像と評価映像を対にして2回提示し、2回目の提示時に両映像に対する評価を行う。一对で提示する映像のどちらが基準画像であるかは評価者に教えずにランダムに提示される。評価者は、5つに分類されたカテゴリ（ACR法の5段階品質尺度のカテゴリと同じ）に基づき、両映像に対する品質評価値を連続尺度上にマークする。評価尺度を0から100とし、条件毎に得られた一对の映像（基準映像と評価映像）に対し、両映像の評価値の差（基準映像の評価値から評価映像の評価値を引いた値）を計算する。この「画質差」を評価者で平均した値がDSCQS値である。DSCQS値は画質の差により導出された値であるため、値が小さいほど品質が良く、値が大きいほど品質が悪いことを表す。

ITU-R 勧告 BT.1788 で規定される SAMVIQ 法は、パソコン画面上で基準映像あるいは評価映像を再生できる評価環境を整え、評価者は自らの意思で映像を再生しながら評点を入力していく手法である。従来の映像品質評価はテレビジョン放送の映像を主要評価対象としてきたが、近年、パソコンやモバイル端末向けの映像配信サービスの普及に伴い、様々な映像フォーマットや観視環境に対応可能な品質評価手法の開発が求められ標準化された。映像の評価順序や観視回数は評価者に委ねられており、各評価映像を比較・確認しながら評価することができるため、再現性が高く、安定した評価値が得られる。

このような主観品質評価法は、評価時間、評価値の安定性や弁別性の観点を考慮し、用途に応じて使い分けられている。

1.3 客観品質推定技術

1.3.1 客観品質推定技術の入出力

前述の主観品質評価の実施には多大な時間・労力、専用の評価設備が必要となるため、サービスの品質評価・設計の効率化や品質監視・管理への適用が困難である。このような背景から、映像通信サービスに関する物理的な特徴量から品質を推定する客観品質推定技術の確立が求められてきた。

客観品質推定技術は入力情報の観点で、プランニングモデル（品質設計技術）[6, 7, 8, 9, 10, 11, 12, 13, 14]、メディアレイヤモデル[15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37]、パケットレイヤモデル[38, 39, 40, 41, 42, 43, 44, 45, 46]、ビットストリームモデル[47, 48, 49, 50, 51, 52]、ハイブリッドモデル[53, 54, 55, 56, 57]に分類される。

プランニングモデル（品質設計技術）（図1.1）は品質設計パラメータ（アプリケーション品質パラメータ（例：解像度、フレームレート、ビットレート）およびネットワーク品質パラメータ（例：パケット損失、遅延））を入力とし、メディア品質を推定する。ただし、プランニングモデルは品質設計パラメータからメディア品質を推定するため、システムに

関する品質推定上の仮定（例：コーデックや実装の情報）を必要とする。

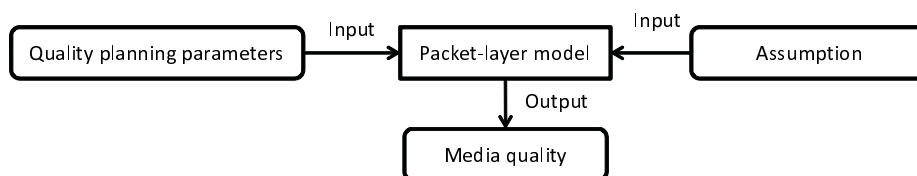


図 1.1 プランニングモデル

メディアレイヤモデル（図 1.2）はメディア信号を入力とし、メディア品質を推定する。更に、メディアレイヤモデルは無圧縮の原映像（SRC¹⁶）と劣化映像を入力とする FR¹⁷型メディアレイヤモデル、原映像から導出された特徴量と劣化映像を入力とする RR¹⁸型メディアレイヤモデル、劣化映像のみを入力とする NR¹⁹型メディアレイヤモデルに大別される。

パケットレイヤモデル（図 1.3）はパケットヘッダ情報を入力とし、メディア品質を推定する。ただし、パケットヘッダにはコーデック情報などが含まれていないため、システムに関する品質推定上の仮定（例：コーデックや実装の情報）を必要とする。

ビットストリームモデル（図 1.4）はビットストリーム情報からメディア品質を推定する。ただし、ビットストリームモデルはデコーダに実装される PLC の情報を得ることができない場合が多く、品質推定上の仮定として PLC の情報必要とする。

ハイブリッドモデル（図 1.5）は品質パラメータ、パケットヘッダ情報、ビットストリーム情報や、メディア信号からメディア品質を推定する。

1.3.2 客観品質推定技術の適用シナリオ

プランニングモデルは品質パラメータと品質の関係をモデル化した技術であるため、通信サービスの品質設計に適用可能である。これに対し、メディアレイヤモデル、パケットレイヤモデル、ビットストリームモデル、ハイブリッドモデルはサービス提供中の通信サービスの品質監視（インサービス品質監視）に適用可能である。

インサービス品質監視は、サービス提供側（IPTVにおけるヘッドエンド）の品質を監視するヘッドエンド品質監視、ネットワーク内の品質を監視するネットワーク品質監視、エンドユーザの体感する品質を監視するエンドユーザ品質監視に大別される。

ヘッドエンドにおける品質低下はサービスを享受するユーザすべてに影響を与えるため、そのインパクトも甚大である。ヘッドエンドにおける品質劣化は主に、符号化劣化と熱雑音などによるノイズであり、画素レベルで劣化を検知し、品質を監視することが適している。ヘッドエンドで利用できる画素信号は符号化前の無圧縮の原映像と符号化直後の劣化

¹⁶SoRCe

¹⁷Full Reference

¹⁸Reduced Reference

¹⁹No Reference

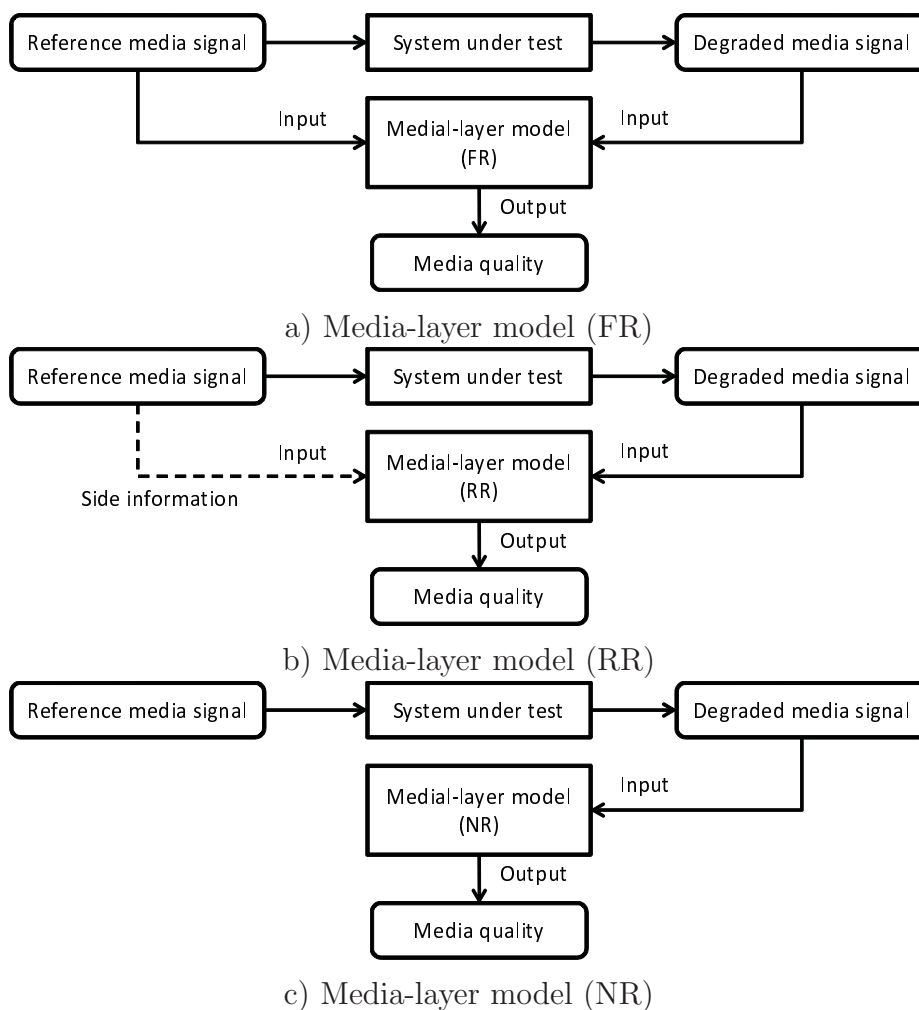


図 1.2 メディアレイヤモデル

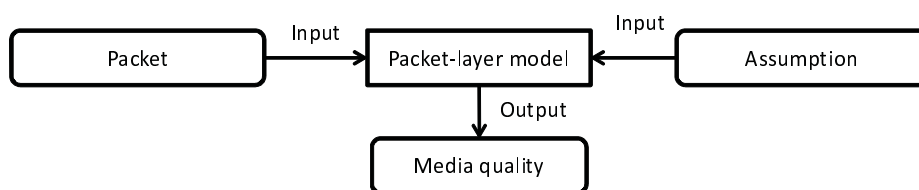


図 1.3 パケットレイヤモデル

映像である。無圧縮の原映像が利用できる場合は、FR型メディアレイヤモデルに基づき符号化直後の品質を監視することが適している。これに対し、過去に符号化を実施し原映像が存在しない場合や、他の場所で符号化したファイルのみが存在する場合には、原映像を用いた品質推定は行えない。その場合は、符号化直後の劣化映像のみを用いるNR型メ

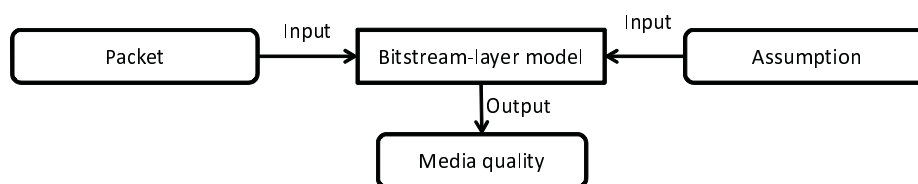


図 1.4 ビットストリームモデル

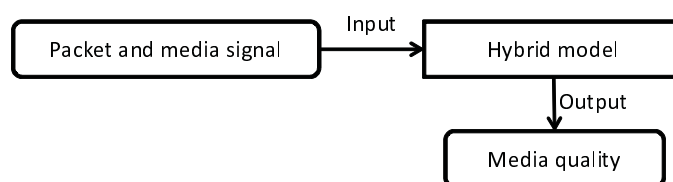


図 1.5 ハイブリッドモデル

ディアレイヤモデルに基づき品質を監視することが適している。

ネットワーク品質監視においては、ネットワークに多数の映像通信ストリームが流れるため、一つ一つのストリームを詳細に監視していくことはネットワーク機器の計算コストの観点で非常に困難である。そのため、ネットワーク品質パラメータを監視し、ネットワークの正常性を監視することが適している。

エンドユーザ品質監視は最終的にユーザが体感している品質（ユーザ体感品質）を監視するという観点で非常に重要であり、劣化検知の観点からはすべてのユーザの品質を常時監視することが理想的である。エンドユーザ端末では、劣化映像、パケット、ビットストリームが取得可能であるため、RR/NR型メディアレイヤモデル、パケットレイヤモデル、ビットストリームモデル、ハイブリッドモデルをエンドユーザ品質監視に適用可能である。しかしながら、すべてのユーザの品質を常時監視するためには、ユーザ端末に品質監視技術を組み込み監視することが適しており、必要以上の計算コストをかけることは困難である。そのため、計算量の観点からメディア信号を用いたRR/NR型メディアレイヤモデルやハイブリッドモデルによる品質推定は困難である。このような背景からエンドユーザ端末が受信するパケット情報のみを用いて品質を推定する必要がある。パケット情報を用いて品質を推定する技術は、パケットレイヤモデルとビットストリームモデルに大別される。ビットストリーム情報はしばしば、著作権保護の観点で暗号化されるため、品質推定に用いることができない。つまり、エンドユーザ品質監視を実現するためには、パケットヘッダ情報を用いるパケットレイヤモデルが適している。

1.4 客観品質推定技術と主な課題

1.4.1 品質設計技術構築に向けた課題

1.3節で述べたように、品質設計技術は品質設計に用いるアプリケーションおよびネットワーク品質パラメータを入力とする技術である。従来、音声通信サービスの品質を設計するITU-T²⁰勧告G.107[6] (E-model) およびG.107.1[7] (Wideband E-model) が標準化されている。E-modelは自由会話に対する品質(R値)を端末、ネットワーク、環境要因を含んだ21の品質パラメータを用い推定する技術である。E-modelは国内外の音声通信サービスの品質設計に幅広く用いられている。

映像通信サービスの品質設計を簡便に実施するためには、E-model同様に、アプリケーションおよびネットワーク品質パラメータから品質を推定可能なモデルを構築することが必要である。文献[14]ではアプリケーション品質パラメータを考慮し、マルチメディアサービスに対する品質を設計するモデルを提案している。しかしながら、文献[14]のモデルはコーデック種別(例:MPEG-4、H.264)、フレームレート低下、パケット損失が映像品質に与える影響を加味していない。映像通信サービスで用いられる映像コーデックの実装方法は多岐に渡り、映像コーデックの実装が映像品質に与える影響も大きい。そのため、映像コーデックの実装が映像品質へ与える影響を加味し映像品質を推定することが必要である。マルチメディア品質に関する推定技術に関しては、文献[60]で音声および映像の主観品質評価値、絶対遅延、メディア同期を用いマルチメディア品質を推定するモデルを提案している。このように、音声品質設計技術、マルチメディア品質設計技術は確立されているのに対し、映像品質設計技術は確立されていない。

1.4.2 ヘッドエンド品質監視技術構築に向けた課題

1.3節で述べたように、ヘッドエンド品質監視を実現するためには、メディア信号に基づきメディア品質を推定する技術の構築が必要である。従来、静止画像や動画像の品質を簡易に評価するため、PSNR²¹が用いられてきた。しかしながら、PSNRと主観品質の相関は必ずしも高くはない。そのため、静止画像の品質推定を目的に客観品質推定技術(PQS²²[18]、SSIM²³[19]、MS-SSIM²⁴[20])が構築されてきた。

文献[18]のPQSでは、基礎歪要因として5つの特徴量(ランダムノイズ(加重平均二乗誤差)、視覚の弁別閾を考慮した平均二乗誤差、符号化サブブロック間の誤差変化量、誤差の自己相関係数、輪郭近傍の可視性に対する平均二乗誤差)を計算する。次に、これら特徴量の相関性を削除するために、主成分分析を行い、得られた主成分と主観品質評価値に対し重回帰分析を施し、PQSを算出する。文献[19]のSSIMでは、基準画像と劣化画像の輝度値に対する平均値、分散、共分散から輝度、コントラスト、構造比較に関する特徴量

²⁰International Telecommunication Union Telecommunication Standardization Sector

²¹Peak-Signal-to-Noise Ratio

²²Picture Quality Scale

²³Structural SIMilarity

²⁴Multi-Scale SSIM

を導出し、SSIMを算出する。文献[20]のMS-SSIMは、まず、基準画像および劣化画像を縦横方向にローパスフィルタを施し、画像サイズを1/4にダウンサンプリングする処理を繰り返し、元の画像サイズを含み5つのサイズの画像を生成する。次に、SSIMと同様に、各レイヤの画像に対する輝度、コントラスト、構造比較に関する特徴量を導出し、これら特徴量を統合し、MS-SSIMを算出する。

これら静止画像に対する客観品質推定技術を動画像に対し拡張した客観品質推定技術の品質推定精度は必ずしも高くないため、VQEG²⁵にて、SD²⁶サイズの映像の符号化品質を推定するFR型メディアレイヤモデルの検討がされた。複数の提案方式の品質推定精度の性能評価の結果、4方式がITU-T勧告J.144[24]として標準化された。その中で、最も品質推定精度が高く、幅広く利用されてきたVQM²⁷は、輝度信号のエッジに関する特徴量2つ、輝度信号の水平/垂直エッジに関する特徴量2つ、輝度信号のフレーム内標準偏差と単純フレーム間差分標準偏差の積に関する特徴量、色差信号の差異に関する特徴量2つの計7つの特徴量を重回帰式により重みづけ加算し、VQM値を導出する技術である。

その後、パソコンや携帯端末を用いた映像配信サービスを想定し、QCIF²⁸、CIF、VGA²⁹サイズの映像の品質を推定するFR型メディアレイヤモデルの検討がされ、ITU-T勧告J.144同様に、提案方式の選定をし、ほぼ同等の品質推定精度を持つ4方式をITU-T勧告J.247[23]として標準化した。その中で、もっと品質推定精度が高いPEVQ³⁰は時空間整合処理や特徴量算出処理を行い、メディア品質を導出する。まず、基準映像と劣化映像を用い、フレームおよび画素位置の整合処理を行う。次に、輝度および色差信号に対しエッジに関する特徴量、時間的変動量に関する特徴量、時間的歪に関する特徴量を算出し、最終的にこれら特徴量を統合し、メディア品質を推定する。

更に、HD³¹サイズの映像の品質を推定するITU-T勧告J.341[25]が標準化された。ITU-T勧告J.341は入力映像をサブサンプリングし縮小した映像を用いて、時空間整合処理や特徴量算出処理を行い、メディア品質を導出する。まず、128×96画素に縮小された基準映像と劣化映像を用いて、両映像の対応するフレームを同期させる時間的な整合処理を行う。次に、960×540画素に縮小された映像を用い、基準映像と劣化映像の水平・垂直方向に対応する画素を整合させる。その後、映像の特徴量として、基準映像と劣化映像の画素値の相関、差分電力に基づく特徴量、ブロック歪、動きの滑らかさに関する特徴量を算出し、最終的にこれら特徴量を統合し、メディア品質を推定する。

これに対し、NR型メディアレイヤモデルについては、VQEGにて技術検証が実施されてきたが、品質推定精度の低さから技術の標準化には至っていない。つまり、2D映像に対するFR型メディアレイヤモデルの検討は多岐に渡り実施されており、その品質推定精度

²⁵Video Quality Experts Group

²⁶Standard Definition

²⁷Video Quality Metric

²⁸Quarter Common Intermediate Format

²⁹Video Graphics Array

³⁰Perceptual Evaluation of Video Quality

³¹High Definition

も高いレベルに達しているが、NR型メディアレイヤモデルは品質推定精度の高い技術が確立されていない。

3D映像に対するメディアレイヤモデルの検討は少なく、2012年からVQEGの3DTVプロジェクトにてFR型メディアレイヤモデルの技術検討が開始されたが確立された技術はない。3D映像の品質は左右眼映像の品質に依存するが、左右眼映像の品質が3D映像品質にどのように影響を与えるかは解明されていない。従来、3D映像品質を左右眼の2D映像品質の平均値から導出する技術が提案されている。しかしながら、H.264/MVC³²のように左右眼映像に割り当てるビットレートが異なることで、左右眼の映像品質に差が付く場合に従来技術を適用し高い品質推定精度を保てるかは解明されていない。

1.4.3 エンドユーザ品質監視技術構築に向けた課題

1.3節で述べたように、エンドユーザ品質監視を実現するためには、パケットヘッダ情報に基づき品質を推定する技術の構築が必要である。IPTVのパケットは、パケットヘッダに映像フレームの区切りを示すフラグが存在しない場合とフラグが存在する場合に大別できる。前者の場合、どの映像フレームが損失したかを考慮し品質を推定することは非常に困難である。パケットに映像フレームの区切りを示すフラグが存在しない場合、従来、パケットヘッダ情報を用いる技術の多くは、パケットヘッダからビットレートおよびパケット損失率を導出し品質を推定する技術が主流であった。しかしながら、ネットワークで発生するパケット損失の多くは、連続的にパケットが損失するバーストパケット損失である。そのため、ビットレートおよびパケット損失パターンに基づき映像品質を推定する必要がある。次に、後者の場合、各映像フレームに割り当てられたビット量や、パケット損失により損失した映像フレームを計算できる。映像品質は映像コンテンツに依存することが知られている。また、映像コーデックは映像コンテンツに応じて映像フレームに割り当てるビット量をコントロールしている。このような背景から、エンドユーザ体感品質を精度良く推定するためには、ビットレートに加え、映像フレームに割り当てられたビット量を考慮する必要がある。また、パケット損失発生時の映像品質を推定するには損失した映像フレームを考慮する必要がある。

1.5 本論文の構成および概要

以上述べた課題に対し、本論文では、2章で1対1のテレビ電話サービスを対象とした品質設計モデル、3章で3D映像サービスを対象としたヘッドエンド品質監視モデル、4章でパケット損失パターンを考慮したIPTVエンドユーザ品質監視モデル、5章で映像フレームを用いたIPTVエンドユーザ品質監視モデルについて論じる。以下に、各章の概要を述べる。

第2章テレビ電話サービスに対する品質設計法では、1対1のテレビ電話サービスの映像品質設計モデルの構築を目的に、映像ビットレート、フレームレート、パケット損失率

³²Multi View Coding

が映像品質に与える影響を主観品質評価特性から導出し、それら品質評価特性をモデル化する。具体的には、提案モデルでは、各映像ビットレートに対し映像品質を最大化する最適フレームレートおよび最適フレームレート時の映像品質をビットレートによりモデル化する。また、パケット損失発生時の映像品質はパケット損失率だけではなく、映像ビットレートおよびフレームレートに影響されるため、本特性をモデル化する。本映像品質設計モデルで推定した映像品質と主観品質評価値の関係から品質推定誤差が主観品質評価値の統計的信頼幅と同等のレベルになることを示す。

第3章 3D 映像サービスに対するヘッドエンド品質監視法では、3D 映像サービスのヘッドエンド品質監視モデルの構築を目的に、左右眼の2D 映像品質が3D 映像品質に与える影響を主観品質評価実験により解明し、解明された品質評価特性をモデル化する。具体的には、提案モデルでは、左右眼の映像品質のうち高い映像品質と、左右眼映像の品質差を用い3D 映像品質を推定する。提案モデルの品質推定精度を検証するため、左右眼の2D 映像品質の平均値から3D 映像品質を推定する従来モデルと品質推定精度を比較し、提案モデルは左右眼映像の品質差が大きい場合でも品質推定精度を高く維持できることを示す。

第4章 パケット損失パターンを考慮したIPTV エンドユーザ品質監視法では、映像フレームの区切りに関する情報がパケットヘッダに示されていないIPTV サービスに対するエンドユーザ品質監視モデルを構築する。具体的には、パケットヘッダ情報から導出されるビットレートおよびパケット損失回数と映像品質の関係を主観品質評価実験より導出し、本品質評価特性をモデル化する。提案モデルの品質推定精度を検証し、実用上十分な品質推定精度に達していることを示す。

第5章 映像フレームを用いたIPTV エンドユーザ品質監視法では、映像フレームの区切りに関する情報がパケットヘッダに示されているIPTV サービスに対するエンドユーザ品質監視モデルを構築する。具体的には、パケットヘッダ情報から導出されるビットレート、Iフレームのフレーム平均ビット量、劣化映像フレーム数と映像品質の関係を主観品質評価実験より導出し、本品質評価特性をモデル化する。提案モデルの品質推定精度を検証するため、従来モデルを組合せ、ビットレートおよび劣化映像フレーム数から映像品質を推定する比較モデルを用意した。実験結果より、比較モデルより提案モデルの品質推定精度が高いことを示す。

最後に、第6章 まとめで本論文をまとめる。

第2章 テレビ電話サービスに対する品質設計法

2.1 まえがき

IP¹ネットワーク、アプリケーションや端末技術の発展によりテレビ電話が急速に普及してきた。テレビ電話の品質は、送受信されるメディア信号、メディア符号化、ネットワーク品質、受信端末に依存する。そのため、高い品質でテレビ電話サービスを提供するためには、ユーザ体感品質 (QoE² [58]) を推定可能な客観品質推定技術を構築し、サービス提供に先立ち、品質を適切に設計することが重要である。

サービス提供者がサービス品質を設計する際、送受信されるメディア信号や受信端末はわからない。そのため、メディア信号と受信端末に仮定を置き、品質設計を行う必要がある。このような背景から、アプリケーション (例: ビットレート) やネットワーク品質設計パラメータ (例: パケット損失率) と主観品質の関係をモデル化した客観品質推定技術を構築する必要がある。

客観品質推定技術 [59] は入力情報の観点から、メディアレイヤモデル [24, 15, 16, 21, 22]、パケットレイヤモデル [38, 39, 40, 41]、プランニングモデル [6, 8, 9, 10, 11, 12, 13, 14]、ハイブリッドモデルに分類できる。ユーザ体感品質を推定するために、メディアレイヤモデルはメディア信号、パケットレイヤモデルはパケットヘッダ情報、プランニングモデルは品質設計パラメータ (例: ビットレート、パケット損失率)、ハイブリッドモデルはメディア信号、パケットヘッダ情報や品質パラメータから品質を推定する。

メディアレイヤモデルは復号された映像や音響/音声コンテンツからメディア品質を推定することができるため、映像や音響/音声コンテンツが品質に与える影響や、パケット損失隠ぺい処理を含むデコーダ特性が品質に与える影響を捉えた品質推定が行える。そのため、メディアレイヤモデルにより推定された品質は主観品質と高い相関を持つ。しかしながら、メディアレイヤモデルはユーザが視聴するメディア信号を入力とし品質を推定するため、品質設計パラメータと主観品質の関係をモデル化できておらず、品質設計に容易に適用できない。メディアレイヤモデルは主に、ベンチマークや品質監視に用いられる。

パケットレイヤモデルはメディアに関するペイロード情報を除いたパケットヘッダ情報のみから品質を推定するため、計算量が非常に少ない。しかしながら、パケットレイヤモデルは映像や音響/音声コンテンツに関する情報を示すペイロード情報から品質を推定しないため、メディア信号に対し品質推定上の仮定を置く必要がある。このように、パケットレイヤモデルはキャプチャされたパケットから品質推定を行うため、品質パラメータと主観品質の関係をモデル化できておらず、品質設計に容易に適用できない。パケットレイヤモデルは主に、インサービス品質管理に用いられる。

プランニングモデルは主観品質と品質パラメータの関係を定式化しているため品質設計に適している。品質設計時には、ユーザが視聴するコンテンツの種類がわからないため、

¹Internet Protocol

²Quality of Experience

パケットレイヤモデル同様に、映像や音響/音声コンテンツを加味した品質推定は困難である。そのため、コンテンツに対し品質推定上の仮定を置く必要がある。

品質設計技術の代表例に、IP 電話を含む音声サービスを対象とした ITU-T³勧告 G.107 (E-model) [6] がある。E-model は端末、ネットワーク、環境要因を含んだ 21 の品質パラメータを用い、会話品質 (R 値) を推定する技術である。一方、映像品質に対する品質設計技術の検討は少ない。文献 [14] ではアプリケーション品質パラメータを考慮し、マルチメディアサービスに対する品質を設計するモデルを提案している。しかしながら、文献 [14] のモデルはコーデック種別 (例: MPEG-4、H.264)、フレームレート低下、パケット損失が映像品質に与える影響を加味していない。マルチメディア品質に関する推定技術に関しては、文献 [60] で音声および映像の主観品質評価値、絶対遅延、メディア同期を用いマルチメディア品質を推定するモデルを提案している。

本稿では、アプリケーションおよびネットワーク品質パラメータを用いテレビ電話サービスの映像品質を推定する映像品質設計モデルを提案する。まず、映像品質と品質パラメータの関係を示す映像品質評価特性を解明する。次に、符号化およびパケット損失から影響を受ける映像品質を推定する映像品質設計モデルを構築する。映像品質設計モデルが異なるコーデック種別、解像度、表示サイズに対する映像品質を推定できることを確認する。更に、モデル構築時に用いた映像とは異なる映像を用い、映像品質設計モデルの品質推定精度の有効性を確認する。

2.2 テレビ電話サービスに対する品質設計モデルのフレームワーク

図 2.1 に示すように、テレビ電話サービスに対する品質設計モデルのフレームワークは文献 [60] に提案されている。文献 [60] のモデルは入力を映像および音声品質パラメータに加え、テレビ電話で行われるタスクとし、音声および映像の品質パラメータが品質に及ぼす影響を加味し、テレビ電話の品質を推定するモデルであり、音声品質推定モデル、映像品質推定モデル、マルチメディア品質統合モデルから構成される。以下に、音声品質推定モデル、映像品質推定モデル、マルチメディア品質統合モデルの特徴を述べる。

1) 音声品質推定モデル

音声品質推定モデルには既存技術である ITU-T 勧告 G.107 (E-model) [6] を適用し、音声品質パラメータから片道受聴音声品質を推定する。

2) 映像品質推定モデル

映像品質推定モデルは図 2.2 に示すように、ビットレート、フレームレート、パケット損失率を入力とし、片道観視映像品質を推定する。ジッタは受信端末のジッタバッファの中で遅延やパケット損失を発生させるため、ジッタによる影響は遅延とパケット損失に置き換えることが可能である。そのため、映像品質推定モデルでは符号化およびパケット損失に影響される映像品質を推定する。

³International Telecommunication Union Telecommunication Standardization Sector

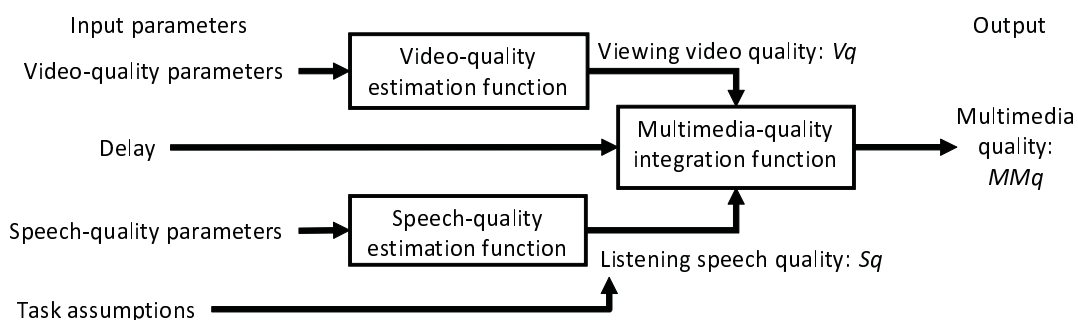


図 2.1 テレビ電話サービスに対する品質設計モデルのフレームワーク

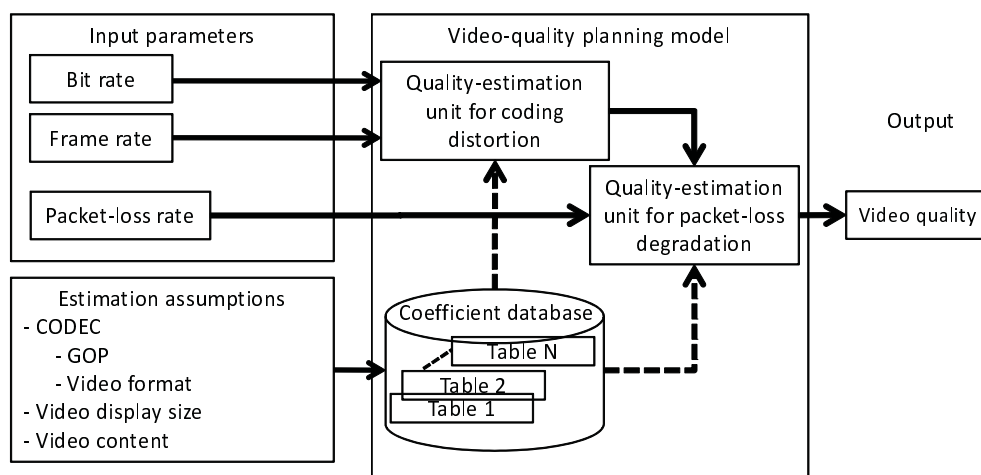


図 2.2 映像品質設計モデル

3) マルチメディア品質統合モデル

マルチメディア品質統合モデルは片道受聴音声品質、片道観視映像品質、伝送遅延を入力とし、個々のメディア品質、遅延（絶対遅延およびメディア間同期）および相互作用を考慮し総合品質を推定する。つまり、遅延による劣化はマルチメディア品質統合モデルでのみ考慮されることになる。本論文では、文献 [61] で提案されたモデルを採用することを想定している。

音声品質推定モデルおよびマルチメディア品質統合モデルは従来技術が確立されているため、本稿では、映像品質推定モデルの確立を行う。まず、映像品質主観評価特性の定性的傾向から映像品質推定モデルの数式を定義する。映像品質主観評価特性はコーデック種別、解像度、映像表示サイズに依存する。また、コーデックや端末の実装は多岐にわたるため、符号化や端末に関する情報のみからコーデックや端末に対する品質を推定することは困難である。そこで、本稿では、コーデック種別、解像度、映像表示サイズに対し、事

表 2.1 映像品質設計モデル構築のための実験条件（実験1）

Quality parameters	Experiment 1
Video codec	H.264
Video format	QVGA
Picture pattern	IPPPP
Key frame interval (s)	1
Video display size (inch)	4.2
Bit rate (kbps) Br	64, 128, 192, 256, 384, 512, 1024
Frame rate (fps) Fr	1, 2, 3, 5, 10, 15, 24, 30
Packet-loss rate (%) Ppl	0.0, 0.1, 0.2, 0.3, 0.5, 1.0, 2.0, 3.0, 5.0, 10.0
Packet-loss pattern	Random

前に導出された係数セットを変更することで映像品質を推定するモデルを構築する。加えて、品質を設計するという観点では、サービス提供前に、ユーザが視聴するメディア信号を事前に把握することはできないため、送受信されるメディア信号に対し仮定を置く必要がある。つまり、提案モデルによる映像品質推定は映像コンテンツが品質に与える影響を加味できないため、仮定された映像の品質を推定することとし、映像品質推定モデルの係数を仮定された映像セットごとに算出することとする。

2.3 主観品質評価

テレビ電話の品質を評価するため、各10秒の映像を4つ用意した。各評価映像には人がカラーブロックを組み立てる映像を用いた。実験に用いた映像は通常テレビ電話で映し出される顔と肩の映像と比較すると動きが多い。そのため、当該評価映像における符号化およびパケット損失に対する品質評価は厳しい評価となることが想定される。表示映像の解像度は、320×240 (QVGA⁴) を用い、表示映像サイズを4.2インチとした。映像は1280×1024の表示解像度を持つ17インチLCD⁵の中央に表示された。

実験に用いた品質パラメータは表2.1に示すように、ビットレート (Br)、フレームレート (Fr)、パケット損失率 (Ppl) とし、パケット損失はネットワークエミュレータにより発生させた。試験条件は全272条件とした。

主観品質評価実験では、映像品質を5段階ACR⁶法 [3] (5:非常に良い、4:良い、3:普通、2:悪い、1:非常に悪い) により評価した。マルチメディア品質の検討に関わらない非専門家である20から39歳の男女各32名の計64名の評価者が実験に参加した。評価者は視距離80cmで映像を評価した。映像品質 (Vq) は4つの映像に対する平均品質評価値

⁴Quarter Video Graphics Array

⁵Liquid Crystal Display

⁶Absolute Category Rating

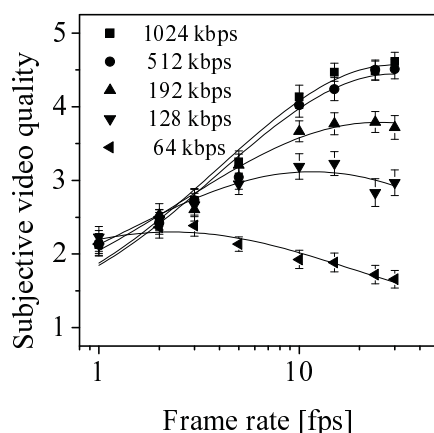


図 2.3 符号化に対する映像品質主観評価特性

とし、MOS⁷として表記する。

2.4 主観品質評価特性

本節では、映像品質および品質パラメータの関係を分析し、符号化およびパケット損失に対する映像品質評価特性を導出した。

2.4.1 符号化に対する映像品質評価特性

実験1において、パケット損失率が0%の場合の符号化に対する映像品質評価特性を図2.3に示した。また、主観品質評価値に対する95%信頼区間も図2.3に示した。図2.3より、各ビットレートに対し、映像品質を最大化する最適なフレームレート（最適フレームレート）が存在することがわかった。また、最適フレームレートの値はビットレートごとに異なるが、品質評価特性は凸型のガウス関数によりモデル化できることがわかった。この結果は、仮に、フレームレートが最適フレームレートより高い場合、映像の空間品質の劣化が時間品質の改善よりも大きくなることを意味し、逆に、フレームレートが最適フレームレートより低い場合、時間品質の劣化が空間品質の改善よりも大きくなることを意味している。

2.4.2 パケット損失に対する映像品質主観評価特性

次に、実験1において、パケット損失率(P_{pl})に対する映像品質評価特性を図2.4に示した。図2.3同様に、主観品質評価値に対する95%信頼区間を図2.4に示した。各ビットレートおよびフレームレートの組に対する曲線の傾きは異なるが、パケット損失率の増加とともに映像品質は指数的に低下することがわかった。つまり、品質評価特性は指数関数によりモデル化できることがわかった。

⁷Mean Opinion Score

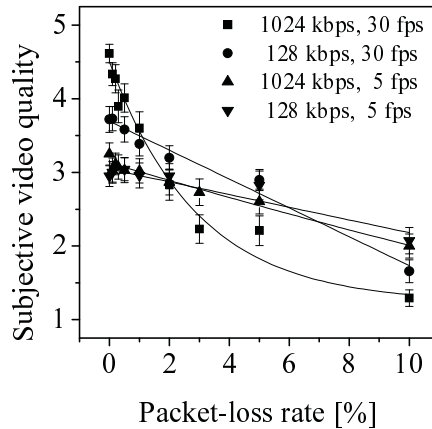


図 2.4 パケット損失に対する映像品質主観評価特性

2.5 映像品質設計モデル

本節では、まず、符号化に対する映像品質を推定するモデルを構築し、次に、符号化およびパケット損失に対する映像品質を推定するモデルを構築する。最後に、H.264で符号化した映像にパケット損失を発生させた映像に提案モデルを適用しモデルの有効性を検証した。

2.5.1 符号化品質推定モデル

符号化に対する映像品質を推定する符号化品質推定モデルを構築した。図 2.3 に示したように、符号化に対する映像品質評価特性はガウス関数によりモデル化でき、以下の数式(2-1)および(2-2)でモデル化した。

$$Vq(Fr, Br, Ppl) \Big|_{Ppl=0} = 1 + I_{coding} \quad (2-1)$$

$$I_{coding} = I_{Of_r} \exp\left[-\frac{\{\ln(Fr) - \ln(Of_r)\}^2}{2D_{Fr}^2}\right] \quad (2-2)$$

ただし、 $1 + I_{coding}$ は符号化に対する映像品質を、 Of_r は各映像ビットレートの映像品質を最大化する最適フレームレートを、 $1 + I_{Of_r}$ は各映像ビットレートの最大映像品質を、 D_{Fr} はフレームレートに対する映像品質の低下度合いを表す。

映像品質主観評価値を数式(2-1)および(2-2)に代入し、最小二乗法[62]を用い、各映像ビットレートに対する Of_r 、 I_{Of_r} 、 D_{Fr} を導出した。

その結果、図 2.5 (a) に示すように、最適フレームレート (Of_r) は映像ビットレート (Br) の増加とともに単調増加し、最大フレームレート (30fps) で飽和した。この特性を

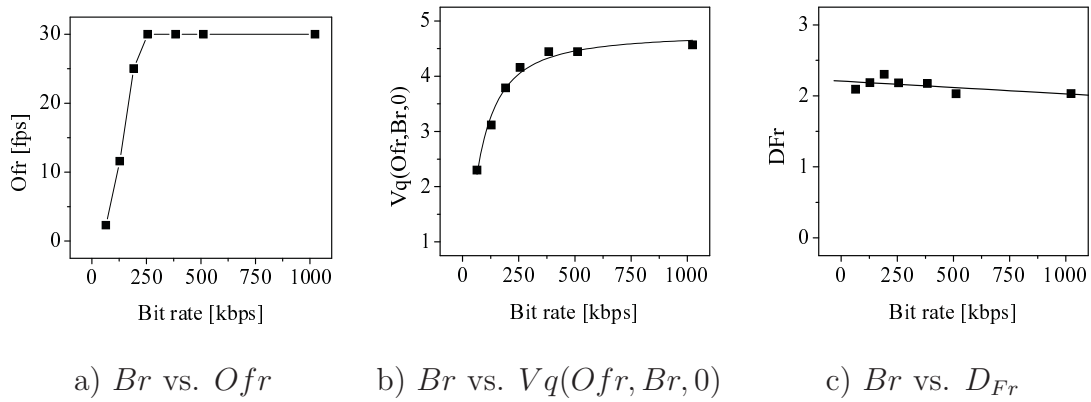


図 2.5 符号化に対する映像品質評価特性

以下の数式 (2-3) でモデル化した。

$$Ofr = v_1 + v_2 Br, \quad 1 \leq Ofr \leq 30 \quad (2-3)$$

ただし、 v_1 および v_2 は定数である。

次に、図 2.5 (b) に示すように、各映像ビットレート (Br) のフレームレート (Fr) を最適フレームレート (Ofr) とした場合、 $Vq(Ofr, Br, 0)$ は映像ビットレート (Br) の増加にともない増加し、一定値に飽和することがわかった。したがって、 I_{Ofr} は以下の数式 (2-4) で示すロジスティック関数でモデル化できた。

$$I_{Ofr} = v_3 - \frac{v_3}{1 + (Br/v_4)^{v_5}}, \quad 0 \leq I_{Ofr} \leq 4 \quad (2-4)$$

ただし、 v_3, v_4, v_5 は定数である。

最後に、図 2.5 (c) に示すように、 DFr は以下の数式 (2-5) で示す一次関数でモデル化した。

$$DFr = v_6 + v_7 Br, \quad 0 < DFr \quad (2-5)$$

ただし、 v_6 および v_7 は定数である。

実験 1 の主観品質評価データを用い、符号化に対する映像品質を推定する品質推定部の係数を最小二乗法に基づき導出した。次に、計算された係数セットを用い、主観品質を推定した。

実用上十分な品質推定精度を有するか判定するため、以下に示す二つの基準を設けた。

1. 相関係数 (c_V) ≥ 0.9 : $c_V \geq 0.9$ は、一般に、高い相関を示すため相関が 0.9 を上回ることを基準とした。

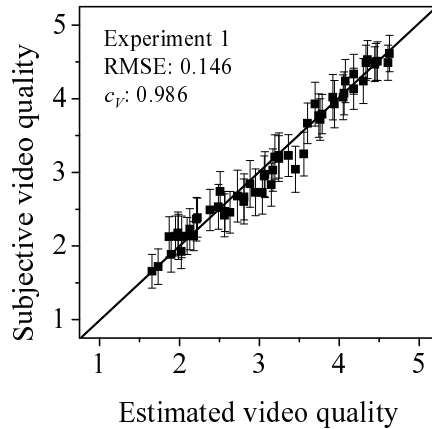


図 2.6 符号化に対する映像品質推定精度

2. $RMSE^8 \leq 0.3$: 映像品質主観評価値に対する 95%信頼区間の平均値 (MCI⁹) は主観品質値の統計的曖昧さを表すため、RMSEが実験1における MCI (0.3) を下回ることを基準とした (実験1の MCIは本章で用いた実験の中で最小値であり、文献 [60] の MCIの値とほぼ同等である)。

提案モデルを用いて推定した符号化に対する品質と主観品質の関係、相関係数 (c_V)、RMSEを図 2.6 に示した。その結果、 c_V は 0.9 以上、RMSE は 0.3 より小さかった。これより、符号化に対する映像品質は数式 (2-1) から (2-5) で定式化できることがわかった。

2.5.2 パケット損失品質推定部

次に、符号化およびパケット損失に対する映像品質を推定するパケット損失品質推定部を構築した。図 2.4 より $Vq(Fr, Br, Ppl)$ はパケット損失率の増加とともに指数的に低下した。そこで、この特性を以下の数式 (2-6) で表す指数関数で定式化した。

$$Vq(Fr, Br, Ppl) = 1 + I_{coding} \exp(-Ppl/D_{Ppl}) \quad (2-6)$$

ただし、 D_{Ppl} はパケット損失に対する映像品質の低下度合いを表す。

2.5.1 節で導出した係数セットを用い算出される I_{coding} および映像品質主観評価値を数式 (2-6) に代入し、パケット損失の低下度合い (D_{Ppl}) を各映像ビットレート (Br) およびフレームレート (Fr) に対し、最小二乗法に基づき算出した。ここで、最大の映像ビットレート (Br_{Max}) の時、 D_1 を以下の数式 (2-7) で定義する。

$$D_1 \stackrel{def}{\iff} D_{Ppl} \Big|_{Br=Br_{Max}} \quad (2-7)$$

⁸Root Mean Square Error

⁹Mean of the 95% Confidence Interval

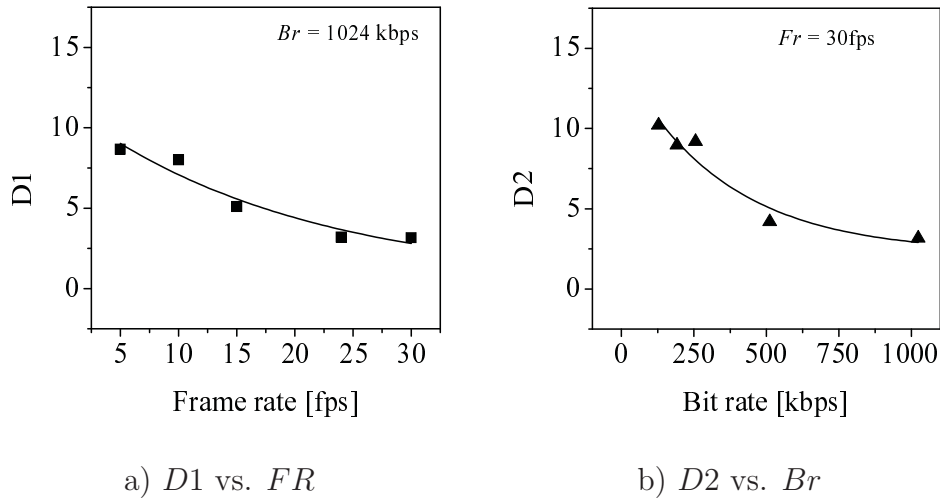


図 2.7 パケット損失に対する映像品質評価特性

同様に、最大フレームレート (Fr_{Max}) の時、 D_2 を以下の数式 (2-8) で定義する。

$$D_2 \stackrel{def}{\longleftrightarrow} D_{Ppl} \Big|_{Fr=Fr_{Max}} \quad (2-8)$$

図 2.7 (a) が示すように、 D_1 はフレームレート (Fr) の増加とともに単調に減少するため、以下の数式 (2-9) で定式化した。

$$D_1 = w + x \exp(-Fr/v_8) \quad (2-9)$$

同様に、図 2.7 (b) が示すように、 D_2 はビットレート (Br) の増加とともに単調に減少するため、以下の数式 (2-10) で定式化した。

$$D_2 = y + z \exp(-Br/v_9) \quad (2-10)$$

ただし、 w 、 x 、 v_8 、 y 、 z 、 v_9 は定数で、係数 w 、 x 、 y 、 z は暫定的に算出された値であり、品質推定に使用しない。

次に、数式 (2-9) および (2-10) を重回帰分析により統合し、以下の数式 (2-11) を定式化した。

$$D_{Ppl} = v_{10} + v_{11} \exp(-Fr/v_8) + v_{12} \exp(-Br/v_9), \quad 0 < D_{Ppl} \quad (2-11)$$

ただし、 v_8 、 v_9 、 \dots 、 v_{12} は定数である。

2.5.1 節同様に、実験 1 の主観品質評価データを用い、符号化およびパケット損失に対する映像品質を推定するパケット損失品質推定部の係数セットを最小二乗法により導出した。計算された係数セットを用い主観品質を推定した結果を図 2.8 に示した。図 2.8 に相関係数

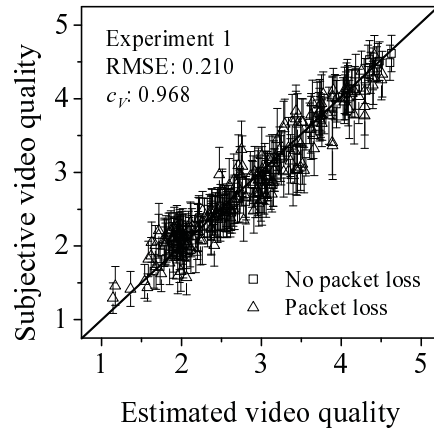


図 2.8 符号化およびパケット損失に対する映像品質推定精度

(c_v) および RMSE を示した。その結果、 c_v は 0.9 以上、RMSE は 0.3 (実験 1 の MCI) より小さかった。これより、映像ビットレート、フレームレート、パケット損失に影響される映像品質を数式 (2-1) から (2-11) で推定できることを示した。

2.6 映像品質設計モデルの数式の検証

映像品質主観評価値は 2.2 節で述べたように、コーデック種別、解像度、映像表示サイズに影響される。そこで、映像品質設計モデルの数式の形状が異なるコーデック種別、解像度、映像表示サイズの映像品質推定に適用可能か検証するため、4 つの実験 (実験 2 から 5) を実施した。実験 2 から 5 は以下の観点で実験 1 と異なる。実験 2 はコーデック種別と映像表示サイズ、実験 3 はコーデック種別、実験 4 はコーデック種別と映像表示サイズ、実験 5 はコーデック種別、解像度、映像表示サイズが実験 1 と異なる。検証試験の条件を表 2.2 および 2.3 に示す。実験 2 から 5 では、2.3 節で示した 4 つの映像コンテンツを用いた。映像品質 (V_q) は 2.3 節同様に、4 つの映像コンテンツに対する品質の平均値を用いた。

係数セット (v_1, v_2, \dots, v_{12}) は主観品質評価値を用い、実験ごとに最適化した。種々のコーデック種別、解像度、映像表示サイズに対する数式の形状の検証結果を表 2.4 に示す。この結果、 c_v は 0.9 以上、RMSE は 0.3 (実験 1 の MCI) より小さかったため、提案モデルの数式の形状は、コーデック種別、解像度、映像表示サイズに依存しないことがわかった。

2.7 映像品質設計モデルの検証

次に、最適化された係数セット (v_1, v_2, \dots, v_{12}) の有効性を検証するため、係数学習用および検証用として 6 つの実験を実施した。実験パラメータを表 2.5 から 2.7 に示した。映像品質設計モデルは実験 6-A、7-A、8-A の主観品質評価値を用い学習し、実験 6-B、7-B、8-B の主観品質評価値を用い係数セットの有効性を検証した。実験 6 および 7 では、2.3 節

表 2.2 実験条件 (実験 2 および 3)

(a) 実験パラメータ			(b) 実験条件		
Quality parameters	Experiment 2	Experiment 3		Experiment 2	Experiment 3
Video codec	H.264	MPEG4		256	49
Video format	QCIF	QVGA	Number of conditions		
Picture pattern	IPPPP		Number of sequences	4	
Key frame interval (s)	1		Number of subjects	64	32
Video display size (inch)	2	4.2	Viewing distance (cm)	30	80
Bit rate (kbps)	16, 32, 40, 64, 128, 192, 256, 384	512, 768, 1024, 1280, 1536, 2048			
Frame rate (fps)	1, 2, 3, 5, 10, 15, 24, 30	2, 5, 10, 15, 30			
Packet-loss rate (%)	0.0, 0.1, 0.2, 0.3, 0.5, 1.0, 2.0, 3.0, 5.0, 10.0	0.0, 0.1, 0.2, 0.5, 1.0, 2.0			
Packet-loss pattern	Random				

表 2.3 実験条件 (実験 4 および 5)

(a) 実験パラメータ			(b) 実験条件		
Quality parameters	Experiment 4	Experiment 5		Experiment 4	Experiment 5
Video codec	MPEG4			49	48
Video format	QVGA	VGA	Number of conditions		
Picture pattern	IPPPP		Number of sequences	4	
Key frame interval (s)	1		Number of subjects	32	
Video display size (inch)	8.5		Viewing distance (cm)	80	
Bit rate (kbps)	512, 768, 1024, 1280, 1536, 2048	1024, 1280, 1536, 2048, 3072, 4096			
Frame rate (fps)	2, 5, 10, 15, 30				
Packet-loss rate (%)	0.0, 0.1, 0.2, 0.5, 1.0, 2.0				
Packet-loss pattern	Random				

で示した4つの映像コンテンツとは異なる16の映像コンテンツを用い、それら映像コンテンツは人の上半身を映した映像コンテンツである。実験8では、人の上半身を映した映像コンテンツ4つと、カラーブロックを組み立てる人を映した映像コンテンツ4つを用いた。実験AおよびBで用いた映像コンテンツは異なるものとした。

実験6-A、7-A、8-Aの主観品質評価値を用い最適化された係数セットを表2.8に示した。学習データ(実験6-A、7-A、8-A)および検証データ(実験6-B、7-B、8-B)に対する相関係数(c_V)およびRMSEを表2.9に示した。この結果、相関係数(c_V)およびRMSEは2.5節で示した品質推定精度の基準を満たした。これより、各実験に対し最適化された係数を持つ提案モデルは非学習データに対し有効性を示した。

表 2.4 種々のコーデック種別、解像度、映像表示サイズに対する提案モデルの品質推定精度（実験2から5）

	c_V	RMSE		c_V	RMSE
Experiment 2	0.981	0.191	Experiment 4	0.973	0.141
Experiment 3	0.987	0.131	Experiment 5	0.972	0.181

表 2.5 実験条件（実験6）

(a) 実験パラメータ		(b) 実験条件	
Quality parameters	Experiment 6-A/B		Experiment 6-A/B
Video codec	MPEG4	Number of conditions	222
Video format	QVGA	Number of sequences	8 for Experiments A and B
Picture pattern	IPPPP	Number of subjects	48
Key frame interval (s)	1	Viewing distance (cm)	80
Video display size (inch)	4.2		
Bit rate (kbps)	320, 384, 512, 768, 1024, 1280, 1536, 2048		
Frame rate (fps)	1, 2, 5, 10, 15, 30		
Packet-loss rate (%)	0.0, 0.1, 0.2, 0.3, 0.5, 1.0, 2.0, 3.0, 5.0, 10.0		
Packet-loss pattern	Random		

2.8 提案モデルを用いた品質設計

ユーザが満足する映像品質を提供しかつ、過剰品質を避けるためには、アプリケーションおよびネットワーク設計が非常に重要である。本節では、実験 2.7 の条件に対し、以下に示すようにアプリケーションおよびネットワーク設計例を示す。

1. 各ビットレートに対する最適フレームレートは？
2. 品質要求条件 (例: $MOS \geq 3.5$ や $MOS \geq 2.5$) を満たす最小ビットレート (Br_{Min}) は？
3. 品質要求条件を満たす最大パケット損失率 Ppl_{Max} は？

例えば、ビットレートが 40、67、128、256 kbps の場合、数式 (2-3) より最適フレームレート (Ofr) は図 2.9(a) に示すように 15、30、30、30 fps と導出される。次に、品質設計者が $MOS \geq 3.5$ を品質要求条件とした場合、数式 (2-1)、(2-2)、(2-4) より提案モデルは図 2.9(a) に示すように最小ビットレート (Br_{Min}) = 67 kbps を導出する。品質設計者が $MOS \geq 2.5$ ($MOS = 2.5$ は 50% のユーザが品質を許容する閾値を意味する) を品質要求条件とした場合、提案モデルはフレームレート (Fr) > 4 fps 以上で最小ビットレート (Br_{Min}) = 40 kbps を導出する。

品質設計者が $MOS \geq 3.5$ を品質要求条件とし、ビットレート (Br) が 128 もしくは 256 kbps の場合、数式 (2-6) および (2-11) より提案モデルは図 2.9(b) に示すように最大パ

表 2.6 実験条件 (実験7)

(a) 実験パラメータ		(b) 実験条件	
Quality parameters	Experiment 7-A/B		Experiment 7-A/B
Video codec	MPEG4	Number of conditions	222
Video format	QQVGA	Number of sequences	8 for Experiments A and B
Picture pattern	IPPPP	Number of subjects	48
Key frame interval (s)	1	Viewing distance (cm)	30
Video display size (inch)	2.1		
Bit rate (kbps)	128, 160, 192, 256, 384, 512, 768, 1024		
Frame rate (fps)	1, 2, 5, 10, 15, 30		
Packet-loss rate (%)	0.0, 0.1, 0.2, 0.3, 0.5, 1.0, 2.0, 3.0, 5.0, 10.0		
Packet-loss pattern	Random		

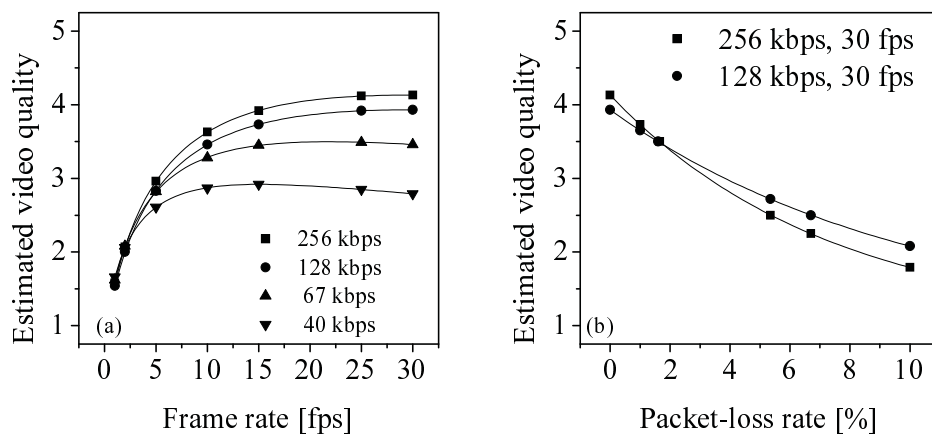
表 2.7 実験条件 (実験8)

(a) 実験パラメータ		(b) 実験条件	
Quality parameters	Experiment 8-A/B	Quality parameters	Experiment 8-A/B
Video codec	H.264	Number of conditions	158
Video format	QCIF	Number of sequences	4 for Experiments A and B
Picture pattern	IPPPP	Number of subjects	48
Key frame interval (s)	3	Viewing distance (cm)	30
Video display size (inch)	2		
Bit rate (kbps)	40, 64, 128, 160, 192, 256, 384		
Frame rate (fps)	1, 2, 5, 10, 15, 25, 30		
Packet-loss rate (%)	0.0, 0.5, 1.0, 2.0, 3.0, 5.0, 10.0		
Packet-loss pattern	Random		

ケット損失率 ($P_{pl_{Max}}$) を 1.7% と導出する。品質設計者が $MOS \geq 2.5$ を品質要求条件とし、ビットレート (Br) が 128 もしくは 256 kbps の場合、最大ケット損失率 ($P_{pl_{Max}}$) はそれぞれ 6.7% や 5.4% と導出される。すなわち、高いビットレートが必ずしも高い品質のサービスを提供できるわけではないことを意味する。このように、提案映像品質設計モデルはアプリケーションやネットワーク設計を簡便かつ効率的に実施できる有用なツールである。

2.9 まとめ

より良い品質設計を実現するため、本章では符号化およびパケット損失に対する映像品質を精度良く推定する映像品質設計モデルを提案した。まず、符号化およびパケット損失に影響される映像品質を推定する二つの映像品質推定部を定義した。前者の映像品質推定部は品質要求条件を満たす最小ビットレートの推定や各ビットレートに対する最適フレー



a) Frame rate vs. estimated video quality b) Packet-loss rate vs. estimated video quality

図 2.9 アプリケーションおよびネットワーク品質設計の例

第3章 3D映像サービスに対するヘッドエンド品質監視法

3.1 まえがき

両眼式3D映像に対するディスプレイおよび眼鏡技術の発展により、サービスプロバイダは3D映像を地上波、衛星、ケーブルやIPネットワーク [63] 経由で提供できるようになってきた。3D映像サービスの品質 (QoE¹) は時空間映像品質、奥行き感 [64, 65]、視覚的快適感 [66, 67] に依存すると一般的に言われている。加えて、クロストークやパペットシアター効果 [68] など新しい品質劣化要因が3D映像サービスのQoEに影響を与える。3D映像のQoEはカメラ、エンコーダ、伝送、デコーダ、ディスプレイや眼鏡で構成される3D映像処理チェーンに影響される [69, 70, 71]。カメラ、ディスプレイ、眼鏡は3D映像のQoEに影響を与えるが、仮に、カメラ、ディスプレイ、眼鏡が改善されても伝送帯域の制限により、3D映像品質は符号化に影響を受けるため、本章では、QoEの一要因として符号化に対する品質に焦点を当て検討する。

3D映像品質は符号化方式 (例: MPEG-2[1]、H.264/AVC²[2]、H.264/MVC³[2])、コーデックの実装 (例: プロファイル、レベル)、符号化フォーマット (例: フレームシーケンシャル、フレームコンパチブルフォーマット) に依存する。図3.1(a) (System A) に示すように、コーデックや伝送に対し既存のインフラを活用するため、サイドバイサイドフレームコンパチブル映像フォーマット [63, 72] では、左右眼映像の空間解像度を水平方向に半分にダウンコンバートし、フルHD⁴の2D映像の空間解像度を維持する。映像はMPEG-2やH.264/AVCで符号化され、セットトップボックス (STB⁵) のようなユーザ端末に伝送される。最終的に、サイドバイサイドフォーマットの映像はデコードされ、左右眼の映像は二つのフルHD映像信号にアップコンバートされる。このように、ユーザは左右眼の映像に対し空間解像度が減少することによる品質劣化を知覚する。ただし、左右眼の映像品質は基本的にはほぼ同等である [73]。

空間解像度の低下による品質劣化を防ぐため、左右眼映像に対し2つのフルHD映像信号を用いるフレームシーケンシャルフォーマット [73] の利用が理想的である。フレームシーケンシャルフォーマットに適用可能な2つの解決法が検討されている。一つは、図3.1(b) (System B) に示すように、H.264/MVCを利用する方法である。H.264/MVCは右眼映像のビットレートを低減させるため、左右眼の映像信号を用い右眼映像の符号化を実施するインタビュー予測技術を用いる。二つ目は、図3.1(c) (System C) に示すように、二つのエンコーダを用いる方法である。二つの符号化された映像ストリームは独立に伝送され、STBで左右眼映像の同期を取り再生される。一例として、左眼映像はMPEG-2やH.264/AVCで符号化され、地上波や衛星ネットワークにより伝送される。同時に、右眼映像はH.264/AVC

¹Quality of Experience

²Advanced Video Coding

³Multiview Video Coding

⁴High Definition

⁵Set-Top Box

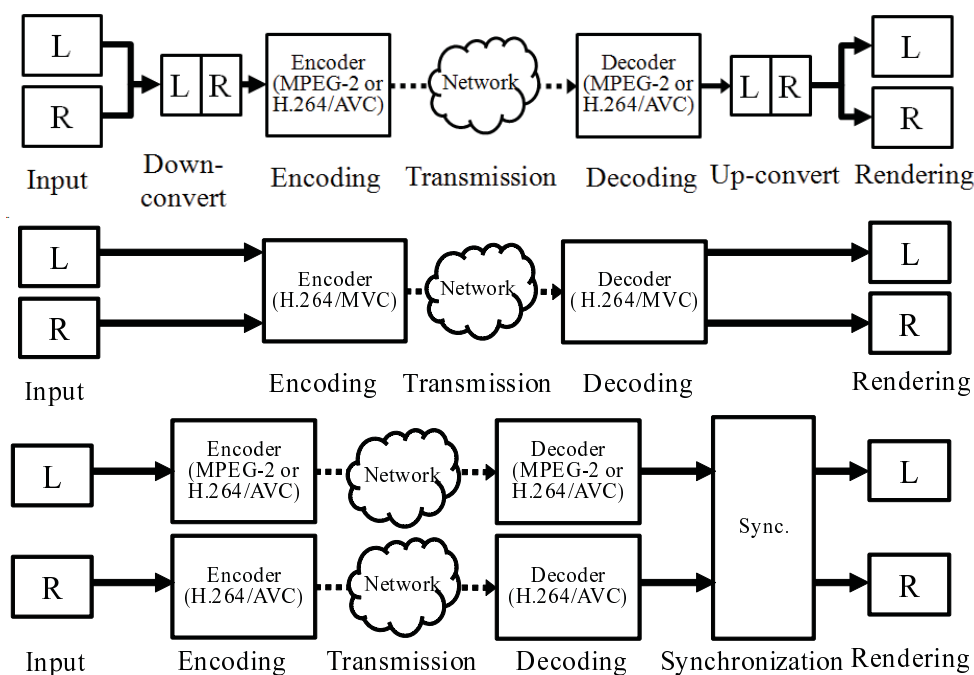


図 3.1 3D映像処理チェーン

で符号化され、IP ネットワークを介し伝送される。最終的に、STB で左右眼映像の同期を取り、3D 映像として表示する。System B および C に関しては、サービスプロバイダが両眼抑圧 (Binocular suppression) に基づき、しばしば右眼映像のビットレートを左眼映像のビットレートより低く符号化する。これらのシステムにおいて、左右眼映像に対する 2 つの 2D 映像はフル HD 解像度を持つが、左右眼映像の品質は一致しない。

前述のように、3D 映像品質は 3D 映像処理チェーンに依存する。そのため、3D 映像を高い品質でユーザーに提供するためには、サービスプロバイダが左右眼映像信号間の品質バランスを考慮し 3D 映像処理チェーンを最適化することに加え、3D 映像品質を監視することが重要である。映像処理チェーンを主観品質評価特性に基づき最適化するためには、左右眼映像品質の差が 3D 映像品質に与える影響を解明する必要がある。また、リアルタイムで 3D 映像品質を監視するためには、前述の品質評価特性をモデル化し、3D 映像サービスに対する客観品質推定技術を構築する必要がある。

3D 映像品質に関する従来検討について述べる。文献 [73] では、フル HD 解像度の左右眼映像の品質は水平方向に半分のサイズにした左右眼映像の品質と比べ、高いことを示している。加えて、文献 [74] では、左右眼映像の解像度が非対称であるときの主観品質評価特性について述べている。具体的には、右眼映像の解像度をフル HD とし、左眼映像の解像度をダウンサンプルした後にアップサンプルしたフル HD としている。その結果、左右眼映像に対し解像度が同じ場合と非対称の場合、3D 映像品質に有意な差が表れたことを示している。しかしながら、いずれの検討も空間解像度が 3D 映像品質にどの程度影響を与え

るかは述べていない。

文献 [75, 76, 77, 78, 79] では、左右眼映像の品質が非対称の場合の主観品質評価特性を示している。文献 [75] では、左右眼映像品質を非対称にするため、左眼映像に対する量子化パラメータ (QP⁶) と右眼映像に対する QP を異なる値に設定している。3D 映像品質は左右眼映像の品質のうち、高い品質に近づくよう二つの左右眼映像を融合する両眼闘争現象で示されるように、左右眼映像に対し非対称な品質は 3D 映像品質にあまり影響を与えないことを示している。しかしながら、文献 [75] の検討では、左右眼映像に対する QP の差が非常に小さいため、左右眼映像の品質差も小さい。文献 [76, 77] では、PSNR⁷ の観点で左右眼映像の PSNR を対称もしくは非対称に設定し、3D 映像品質への影響を評価している。その結果、左右眼映像の PSNR を非対称に符号化した映像の品質は左右眼映像の PSNR を対称に符号化した映像の品質と比較して低いことを示している。文献 [78] では、System B および C に対する 3D 映像品質を比較している。H.264/MVC において、左眼映像の右眼映像に対するビットレート比を 50% とした場合、非対称符号化に対する映像品質は対称符号化の映像品質とほぼ同等であることを示している。また、左眼映像の右眼映像に対するビットレート比を 35 – 5% とした場合、非対称符号化に対する映像品質は対称符号化の映像品質より低いことを示している。文献 [79] では、QP とガウシアンフィルタを変化させた際のブロックノイズとぼけに対する映像の非対称品質を検討している。その結果、評価者はぼけの非対称レベルと比較して、ブロックノイズの非対称レベルに基づき、非対称符号化のレベルを検知していることを示している。文献 [75, 76, 77, 78, 79] は左右眼映像の品質差が 3D 映像品質に影響を与えることを述べているが、左右眼映像の品質差がどの程度 3D 映像品質に影響を与えるかは示していない。また、当該品質評価特性をモデル化していない。

文献 [80] では、左右眼映像品質の平均値から 3D 映像品質を推定する客観品質推定モデルを提案している。文献 [80] の手法は左右眼映像品質差が小さい時に良い品質推定精度を持つことが想定される。しかしながら、左右眼の映像品質差が十分に大きいとき、モデルは 3D 映像品質を過大評価すると考えられる。System B や C では、左右眼の映像品質差が大きくなることが想定され、当該品質差がある場合の 3D 映像品質を評価可能なモデルが必要である。

これらの調査から、従来検討は、左右眼映像の 2D 映像品質差が 3D 映像品質にどの程度影響を与えるかを解明できておらず、その品質評価特性もモデル化できていない。そこで、本検討では、3D 映像品質評価特性を詳細に述べ、当該品質評価特性をモデル化し、提案モデルが従来モデルより良い品質推定精度を持つことを示す。具体的には、左右眼映像品質バランスを最適化することに加え、3D 映像品質を評価可能な客観品質推定モデルを提案する。まず、左右眼映像の 2D 映像品質の差が 3D 映像にどのように影響を与えるかを解明するため主観品質評価実験を実施する。次に、導出された品質評価特性を主観品質推定モデルとしてモデル化する。ただし、2D 映像品質を推定する客観品質推定モデルは標準化 [25]

⁶Quantization Parameter

⁷Peak-Signal-to-Noise Ratio

表 3.1 空間情報 (SI) および時間情報 (TI)

Group 1		Left		Right	
SRC No.	Title	SI	TI	SI	TI
1	Photo	65	5	71	5
2	Mirror	32	15	32	15
3	Bicycle (front face)	54	16	57	17
4	Basketball	67	47	63	47
5	Flower garden	71	50	72	50
6	Paint	67	57	68	57

Group 2		Left		Right	
SRC No.	Title	SI	TI	SI	TI
7	Well-dressed woman	29	10	32	10
8	Bicycle (side face)	63	27	72	28
9	Woman and maple leaves	42	8	47	8
10	Cheerleaders	55	60	63	60
11	Clown	122	76	118	76
12	Tropical fish	65	21	64	21

表 3.2 実験条件

(a) コーデック		(b) 符号化パラメータ	
Experiments A and C	H.264/AVC	Profile	High profile
Experiment B	H.264/MVC	Video format	1920 × 1080p
		Chroma format	4:2:0
		Frame rate	24 fps
		GoP	M=3, N=24 for H.264/AVC, M=3, N=22 for H.264/MVC

もされ、十分に検討されているため、本検討では、左右眼映像に対する主観品質評価値を用い、3D映像品質をモデル化する。つまり、提案モデルでは、2D映像の客観品質推定モデルの出力を提案モデルの入力として用いることを前提としており、2D映像の客観品質推定モデルの出力を用い提案モデルの品質推定精度を評価する点に関しては、本検討のスコップ外とする。

3.2 主観品質評価

主観品質評価特性を導出し、左右眼映像の2D映像品質の差が3D映像品質に与える影響をモデル化するため主観品質評価実験を実施した。実験では、12の原映像 (SRC⁸) と2つの映像コーデックを用いた。

映像コンテンツの選定は主観品質評価特性を適切に導出するための重要なポイントであ

⁸SouRCe

表 3.3 System A, B, C に対するビットレート

(a) System A および B		
System A	16, 12, 8, 4, 2, and 1 Mbps	
System B	32, 16, 8, 6, 4, and 2 Mbps	
(b) System C		
Overall bit rate (Mbps)	Bit rate for left view (Mbps)	Bit rate for right view (Mbps)
32	16	16
28	16	12
24	16	8
20	16	4
18	16	2
17	16	1
16	8	8
14	8	6
12	8	4
10	8	2
9	8	1
8	4	4
7	4	3
6	4	2
5	4	1
4.5	4	0.5
4	2	2
3	2	1
2.5	2	0.5
2	1	1
1.5	1	0.5

る。選択された映像コンテンツのセットは時空間情報の観点で幅広く選定されるべきである。10秒のフルHDの3D映像コンテンツを以下に示すように、12種類選定した。SRC1 (Photo) は女性が葉の写真を撮影する映像；SRC2 (Mirror) は女性が鏡の前で複数のドレスを試着する映像；SRC3 (Bicycle (front face)) は自転車をこぐ女性がカメラの方向へ直進する映像；SRC4 (Basketball) はバスケット選手がドリブルする映像；SRC5 (Flower garden) は男女が花畑で立ち、花を見ている映像；SRC6 (Paint) は女性が部屋で花の絵を描いている映像；SRC7 (Well-dressed woman) はドレスを着た女性がリビングを歩いている映像；SRC8 (Bicycle (side face)) は自転車をこぐ女性が公園を右方向に移動する映像；SRC9 (Woman and maple leaves) は女性がもみじの葉を眺めている映像；SRC10 (Cheerleaders) はゲームでチアリーダがダンスする映像；SRC11 (Clown) はピエロが風船を少女に与える映像；SRC12 (Tropical fish) は熱帯魚が水族館で泳いでいる映像である。SRC1、2、3、7、8、9は情報通信研究機構の映像、SRC4、5、6、10はデジタルコンテンツ協会の映像、SRC12

は日本電信電話研究所の映像である。ITU-T 勧告 P.910[3] で定義される SI⁹および TI¹⁰を表 3.1 に示した。提案モデルが非学習データに対し精度良く 3D 映像品質を推定できているか検証するため、SRC は映像グループ 1 および 2 として二つのグループに分類した。

実験 A1 および A2 (System A) において、サイドバイサイドフレームコンパチブルフォーマットの映像を H.264/AVC で符号化した。実験 B1 および B2 (System B) において、フレームシーケンシャルフォーマットの映像を H.264/MVC で符号化した。実験 C1 および C2 (System C) において、フレームシーケンシャルフォーマットの左右眼の映像を H.264/AVC で独立に符号化した。ただし、実験 C においては左右眼の映像信号に対し、同一の H.264/AVC を用いた。ここで、実験 XY を定義する。X はシステム (System A、B、C) を示し、Y は SRC の映像グループ (Group 1、2) を示す。符号化パラメータは表 3.2 に示した。

表 3.3 に示すように、System A では左右眼映像に対するビットレートは 1–16 Mbps、System B では 2–32 Mbps、System C では 1.5–32 Mbps とした。System B における左眼映像のビットレートは右眼映像のビットレートの約 2 倍である。System C においては、左右眼映像に対し幅広い映像品質をカバーするため、左右眼映像のビットレート比を 1:1 から 16:1 とした。

実験 A1、A2、B1、B2 の 3D 映像の PVS¹¹数は 36 (6 HRC¹² × 6 SRC)、実験 C1 および C2 の 3D 映像の PVS 数は 126 (21 HRC × 6 SRC) とし、実験 A1、A2、B1、B2 の左右眼の 2D 映像の PVS 数は 72 (6 HRC × 6 SRC × 2)、実験 C1 および C2 の 2D 映像の PVS 数は 84 (左眼映像の PVS 数：5 HRC × 6 SRC、右眼映像の PVS 数：9 HRC × 6 SRC) とした。主観品質評価実験において、3D および 2D の無圧縮の原映像をアンカーとして用いたが、無圧縮の原映像の主観品質評価値を統計分析には用いなかった。

主観品質評価特性はコーデックの実装に依存するため、System A、B、C に対する RD¹³カーブを示す。ただし、PSNR は左右眼映像のすべてのフレームを用いて計算した。また、ビットレートおよび PSNR は 12 映像すべてに対し平均をとった。System A においては水平方向にアップサンプリングした後に PSNR を計算した。図 3.2 より、System B の RD カーブは System A および C の RD カーブより高いことを示し、System A の RD カーブは System C の RD カーブとほぼ同等であることを示した。ここで、図 3.2(b) に示すように、System C の RD カーブは左右眼映像のビットレート比ごとに描いた。

主観品質評価において、3D 映像品質を 5 段階 ACR¹⁴法 (5：非常に良い、4：良い、3：普通、2：悪い、1：非常に悪い) により評価した。評価者は映像が表示された後、5 秒間で品質を評価するよう求められた。これら実験における PVS の表示順はランダムにした。主観品質評価値は MOS¹⁵で表した。ここで、MOS_OS、MOS_LS、MOS_RS はそれぞれ、

⁹Spatial Information

¹⁰Temporal Information

¹¹Processed Video Sequence

¹²Hypothetical Reference Circuit

¹³Rate Distortion

¹⁴Absolute Category Rating

¹⁵Mean Opinion Score

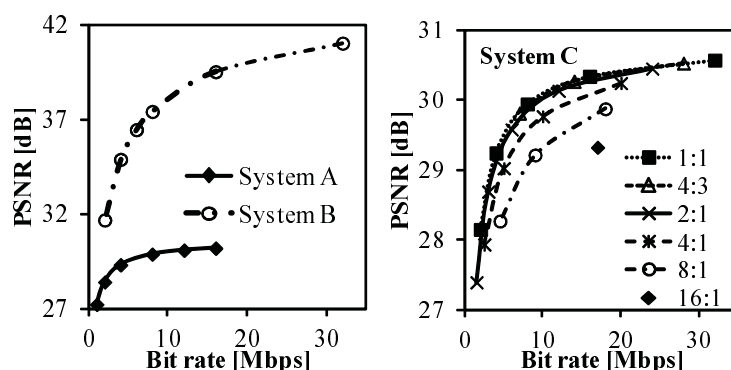


図 3.2 RD カーブ

3D映像、左眼映像、右眼映像の主観品質評価値を示す。主観品質評価実験を開始する前に、ITU-R 勧告 BT.1438[81] に示されている2つのスクリーニング試験 (Coarse stereopsis および Fine stereopsis 試験) を実施した。加えて、視力および色覚についてもスクリーニング試験を実施した。スクリーニング試験に合格した男女32名の評価者が主観品質評価実験に参加した。評価者は20から39歳であった。評価者は約150cm (3H: 画像高 (H)) の視距離で偏光眼鏡をかけ各3D映像を観視した。符号化された映像は40インチのLCDモニタ上に1920×1080のネイティブ解像度で表示された。実験環境として室内照度を20 luxとした。2D映像品質評価実験では、評価者は偏光眼鏡をはずして評価した。

3.3 主観品質評価特性

まず、主観品質評価の安定性について述べる。t-student分布を用い、32名に対する95%信頼区間 (CI¹⁶) を以下の数式で計算した。

$$CI = 2.040 \frac{\sigma}{\sqrt{32}},$$

σ は評価者ごとの評価値の標準偏差である。

図3.3はMOSとCIの関係を示している。3D映像品質評価実験における平均、最大、最小CIは0.261、0.377、0.000¹⁷であった。2D映像品質評価実験における平均、最大、最小CIは0.253、0.396、0.000であった。文献[82]において、二つの2D映像品質評価実験における平均CIが0.36および0.31であった。これより、3Dおよび2D映像品質評価実験の平均CIは文献[82]の実験の平均CIより小さかった。そのため、安定性については問題がなかったと考えられる。

実験では映像グループ1および2の映像を用いたが、これ以降では図面内のポイント数が多くなりすぎポイントの識別が困難なため、映像グループ1のデータのみを用いる。ま

¹⁶Confidence Interval

¹⁷CI=0.000 は全評価者が同じ評点を付けたことを意味する。

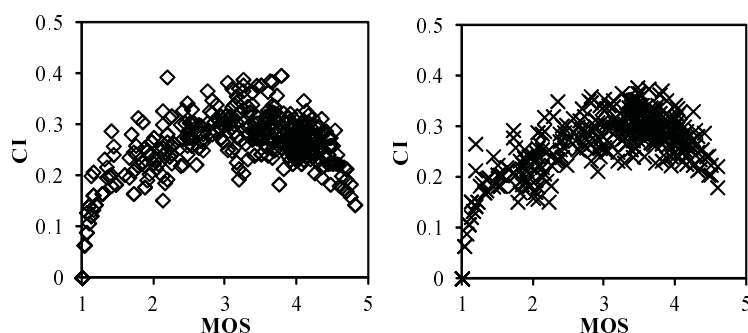


図 3.3 MOS vs. CI

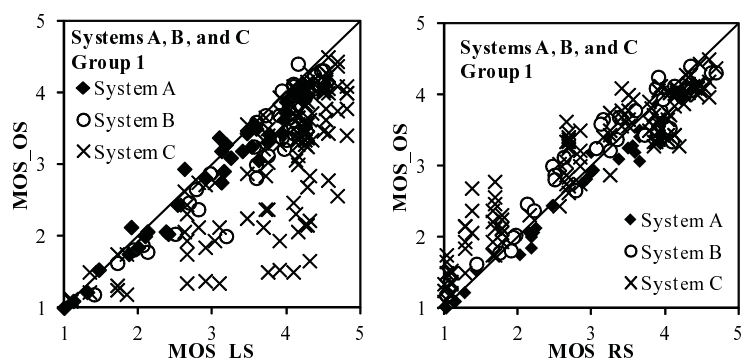


図 3.4 MOS_OS と MOS_LS もしくは MOS_RS の関係

た、映像グループ2の映像においてもほぼ同等の主観品質評価特性が得られていることは確認済みである。

まず、MOS_OS および MOS_LS、MOS_OS および MOS_RS の関係を調査した。System A において、MOS_OS は MOS_LS および MOS_RS のいずれとも高い相関を有した (図 3.4 a) および b))。System B および C においては、MOS_OS は MOS_LS および MOS_RS のいずれとも高い相関を有するポイントがあるが、MOS_LS および MOS_RS のいずれとも相関が低いポイントもあった。

次に、MOS_LS および MOS_RS の差分値の絶対値 $dMOS_{LR} (ABS(MOS_{LS}-MOS_{RS}))$ が MOS_OS に影響を与えるか調査した。図 3.5 a) および b) は System C の 32 および 17 Mbps のビットレートに対し、MOS_OS、MOS_LS および MOS_RS の関係を示している。図 3.5 a) に示されるように、 $dMOS_{LR}$ が小さい場合、MOS_OS は MOS_LS および MOS_RS とほぼ同じ値を示した。しかしながら、図 3.5 b) に示されるように、 $dMOS_{LR}$ が大きい場合、MOS_OS は左右眼の映像品質が非対称なため、MOS_LS および MOS_RS のうち高い方の品質から減少した。これら特性から $dMOS_{LR}$ は MOS_OS に影響を与えることがわかった。次に、図 3.5 b) に示されるように、SRC ごとの総ビットレートが同じ場合であっ

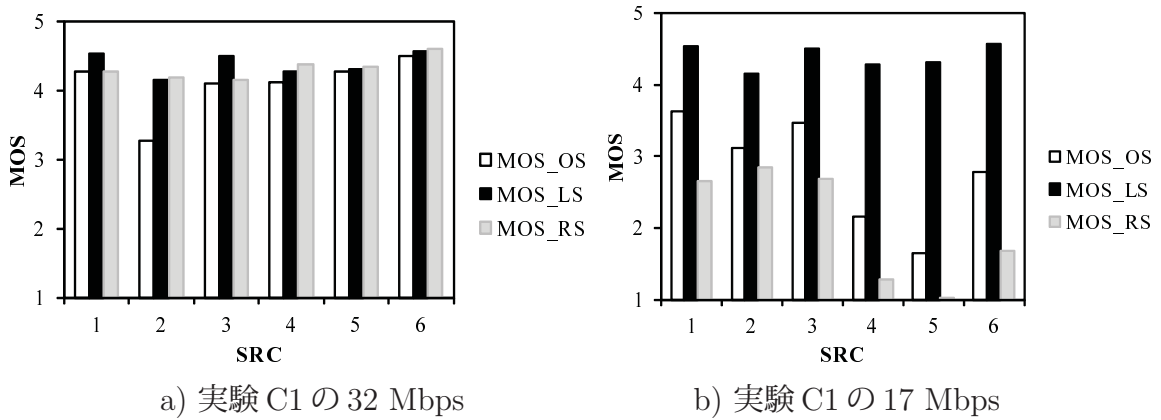


図 3.5 MOS_OS、MOS_LS、MOS_RS の関係

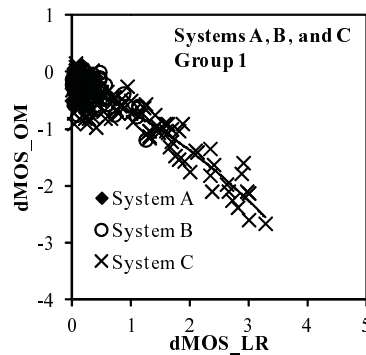


図 3.6 dMOS_LR vs. dMOS_OM

ても、SRC ごとの MOS_OS は大きく異なる。加えて、MOS_OS は dMOS_LR によって変化した。

前述のように MOS_OS は dMOS_LR により変化するため、dMOS_LR が MOS_OS へ与える影響を調査した。ここで、 $MAX(MOS_{LS}, MOS_{RS})$ および dMOS_OM を定義する。 $MAX(MOS_{LS}, MOS_{RS})$ は MOS_LS および MOS_RS の最大値とし、dMOS_OM は MOS_OS から $MAX(MOS_{LS}, MOS_{RS})$ を引いた差分値とする。図 3.6 は dMOS_LR および dMOS_OM の関係を示す。System A において、MOS_LS および MOS_RS はほぼ同等なため、プロットは図の左上に集中した。加えて、MOS_LS および MOS_RS が同じ場合、MOS_OS は MOS_LS および MOS_RS とほぼ同等なため、dMOS_OM はほぼゼロとなった。System B および C において、dMOS_OM は dMOS_LR に依存し変化した。加えて、図 3.6 に示される品質評価特性は二次関数でモデル化された。

前述の結果から主観品質評価特性を以下のようにまとめることができる。一つ目の特性は dMOS_LR が小さい時、両眼抑圧 [75, 76, 77, 78, 79] によって MOS_OS は MOS_LS および MOS_RS と高い相関を持った。二つ目の特性は dMOS_OM は dMOS_LR によって変化

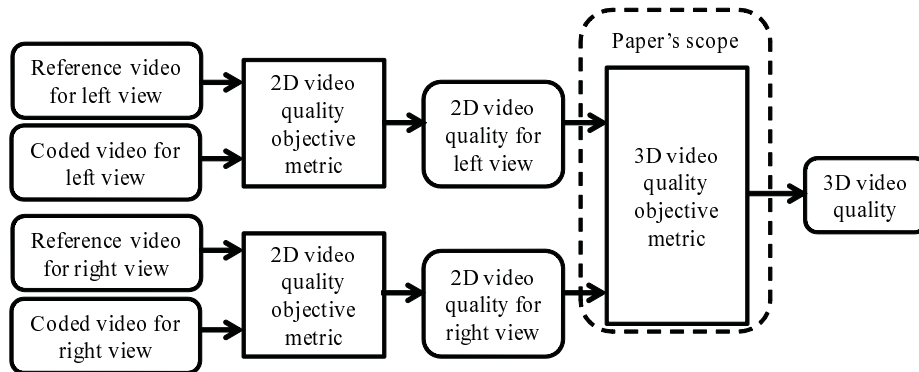


図 3.7 提案客観品質推定モデル

し、その特性は二次関数でモデル化された。

3.4 提案客観品質推定モデル

符号化された 3D 映像の品質の最適化および 3D 映像品質の監視に適用可能な客観品質推定モデルを提案する（図 3.7）。3.1 節で述べたように、左右眼映像の 2D 映像品質を推定可能な 2D 映像品質客観推定モデル（ITU-T 勧告 J.341[25]）は十分に検討されているため、本検討では、左右眼映像の主観品質評価値を入力とする 3D 映像主観品質推定モデルを構築した。

3.3 節で述べたように、3D 映像品質は左右眼映像の品質差に影響され、品質差が小さい場合、3D 映像品質は左右眼映像の 2D 映像品質のいずれかによって表される。そこで、上記品質評価特性を以下の数式（3-1）でモデル化した。

$$\begin{aligned}
 MOS_{OO} &= a + b \cdot \text{MAX}(MOS_{LS}, MOS_{RS}) + c \cdot \text{ABS}(MOS_{LS} - MOS_{RS}) \\
 &\quad + d \cdot (MOS_{LS} - MOS_{RS})^2, \\
 &\approx a + b \cdot \text{MAX}(MOS_{LO}, MOS_{RO}) + c \cdot \text{ABS}(MOS_{LO} - MOS_{RO}) \\
 &\quad + d \cdot (MOS_{LO} - MOS_{RO})^2
 \end{aligned} \tag{3-1}$$

ここで、 a 、 b 、 c 、 d は学習データに重回帰分析を適用し得られた係数、 MOS_{LS} および MOS_{RS} は左右眼映像に対する 2D 映像品質主観評価値、 MOS_{OO} は 3D 映像品質客観評価値、 MOS_{LO} および MOS_{RO} は 2D 映像品質客観推定モデルにより推定された左右眼映像に対する 2D 映像品質客観評価値である。

3.5 提案モデルの品質推定精度評価

3.5.1 品質推定精度に対する要求条件

本節では、提案モデルの品質推定精度の目標値について述べる。文献 [80] で提案された従来モデルと提案モデルの品質推定精度を比較した。具体的には、従来モデルは以下の数

表 3.4 提案および従来モデルの係数
(a) 提案モデル (b) 従来モデル

Coefficient	Value	Coefficient	Value
a	0.000	e	0.000
b	0.922	f	0.912
c	-0.329		
d	-0.104		

式 (3-2) で示されるように、左右眼の映像品質の平均値を用いて 3D 映像品質を推定する。

$$MOS_{OO} = e + f \cdot (MOS_{LS} + MOS_{RS}) / 2 \approx e + f \cdot (MOS_{LO} + MOS_{RO}) / 2 \quad (3-2)$$

ここで、 e および f は学習データに重回帰分析を適用し得られた係数、 MOS_{OO} 、 MOS_{LS} 、 MOS_{RS} 、 MOS_{LO} 、 MOS_{RO} の定義は数式 (3-1) と同じである。

学習データ (実験 C1) を用い最適化された提案モデルの品質推定精度が従来モデルの品質推定精度より高いことを本検討の要求条件とした。すなわち、非学習映像 (映像グループ 2) およびシステム (System A および B) に対し、提案モデルの品質推定精度が従来モデルの品質推定精度と比較して高いことを検証した。

品質推定精度の検証として、PCC¹⁸、RMSE¹⁹、OR²⁰を用いた。ただし、OR は 3.3 節で述べた CI より、 MOS_{OS} と MOS_{OO} の差分値の絶対値が大きいものの割合とした。

3.5.2 提案モデルのパフォーマンス

表 3.4 に示すように、学習データ (実験 C1) を用い、提案モデルの係数 a 、 b 、 c 、 d および従来モデルの係数 e 、 f を計算し、提案モデルと従来モデルの品質推定精度を比較した。提案および従来モデルの係数は 5% 有意水準を満たした。ただし、係数 a および e は 5% 有意水準を満たさなかったため、0 とした。実験 A1、B1、A2、B2、C2 を非学習データとした。

図 3.8 は MOS_{OO} および MOS_{OS} の関係を示す。表 3.5、3.6、3.7 に PCC、RMSE、OR を示した。ただし、品質推定精度の比較は、学習データを実験 C1 とした条件下で実施した。全実験において提案モデルの品質推定精度は従来モデルの品質推定精度より高かったため、提案モデルの品質推定精度は非学習システムおよび映像に対し十分な精度で品質を推定できることを示した。これらの結果より、提案モデルは 3D 映像処理チェーンの最適化および 3D 映像品質監視に適用可能であると結論付けた。

3.5.3 考察

実験結果を詳細に考察する。System B および C のプロットにおいて、従来モデルの品質推定精度が低かった。そこで、原因を分析するため、表 3.8 に SRC ごとの RMSE を示し

¹⁸Pearson's Correlation Coefficient

¹⁹Root Mean Square Error

²⁰Outlier Ratio

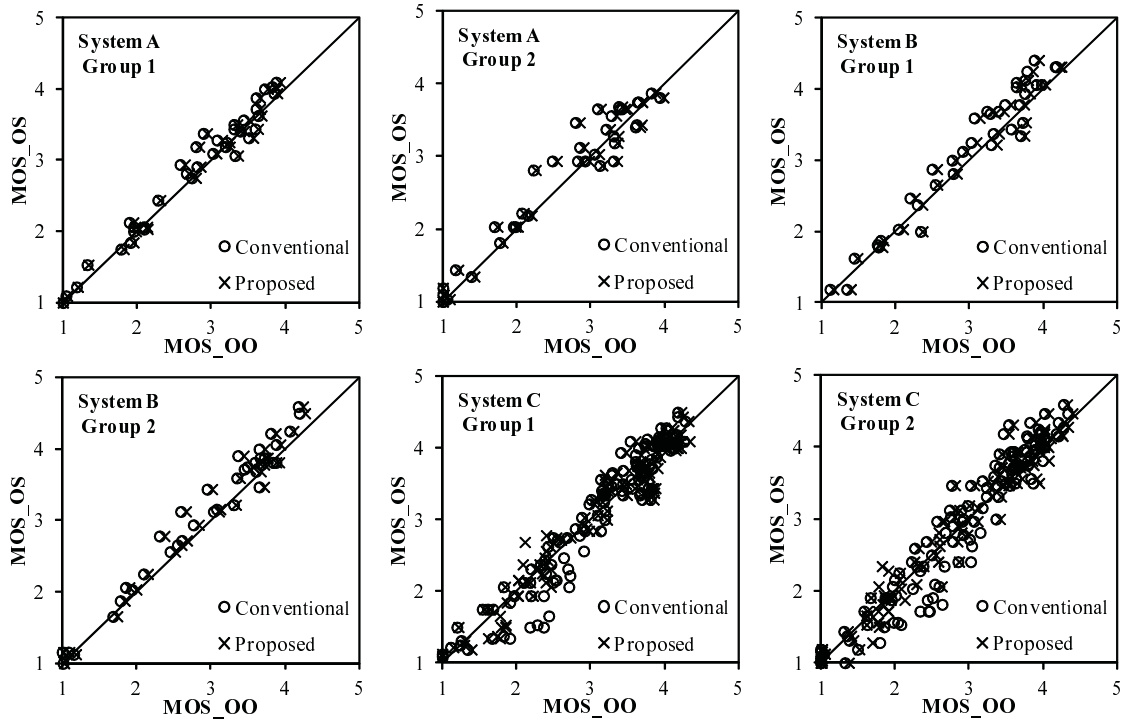


図 3.8 MOS_OO vs. MOS_OS

た。SRC2 および 3 を除き、提案モデルの RMSE は従来モデルの RMSE より低かった。全 HRC に対し dMOS_LR を平均した dMOS_LRavg を定義することにより、品質推定精度の改善が SRC に依存することを分析した。表 3.9 は SRC ごとの dMOS_LRavg を示す。表 3.8 および 3.9 に示されるように、dMOS_LRavg が大きいときに RMSE の改善も大きいため、品質推定精度の改善は SRC に依存することがわかった。

次に、従来モデルと比較することで、提案モデルが RMSE を改善した理由を分析した。表 3.10 に $0 \leq \text{dMOS_LR} \leq 1$ および $1 < \text{dMOS_LR}$ の範囲の提案および従来モデルの RMSE を示した。この結果、左右眼の映像品質差が小さい場合は従来モデルおよび提案モデルの RMSE はほぼ同等であるのに対し、左右眼の映像品質差が十分に大きい場合は提案モデルの RMSE が大きく改善されていることがわかった。つまり、提案モデルは左右眼の映像品質差によらず、精度良く 3D 映像品質を推定できることがわかった。

3.6 まとめ

左右眼の映像品質差を計算し、3D 映像品質を推定する新たな品質推定モデルを提案した。まず、左右眼の映像品質差が 3D 映像品質に与える影響に関し、従来検討の問題点を指摘した。次に、主観品質評価実験を実施することにより、左右眼の映像品質差が 3D 映像品質に与える影響を調査した。その結果、3D 映像品質は左右眼映像品質のいずれかに依存するのではなく、左右眼の 2D 映像品質差にも影響を受けることがわかった。また、dMOS_OM

表 3.5 提案および従来モデルの PCC

Experiment	Proposed	Conventional
A1	0.98	0.98
A2	0.97	0.97
B1	0.97	0.97
B2	0.99	0.98
C1	0.97	0.96
C2	0.98	0.95
A1, B1, and C1	0.97	0.96
A2, B2, and C2	0.98	0.96
All PVSs	0.97	0.96

表 3.6 提案および従来モデルの RMSE

Experiment	Proposed	Conventional
A1	0.16	0.18
A2	0.24	0.26
B1	0.24	0.27
B2	0.21	0.25
C1	0.23	0.28
C2	0.23	0.30
A1, B1, and C1	0.22	0.26
A2, B2, and C2	0.23	0.29
All PVSs	0.23	0.27

は dMOS_LR を変数とした二次関数でモデル化された。

左右眼の 2D 映像品質の平均値を用いる従来モデルと提案モデルを比較し、品質推定精度を確認した。その結果、提案モデルの品質推定精度は従来モデルの品質推定精度より高いことを示した。特に、左右眼の 2D 映像品質差が大きい場合、提案モデルは従来モデルと比較して、精度良く 3D 映像品質を推定することがわかった。

以下に今後の課題を示す。本検討では左右眼映像の主観品質評価値を提案モデルの入力とした。左右眼映像の主観品質評価値を用いることなく 3D 映像品質を客観的に推定するためには、2D 映像に対する客観品質推定モデル（例：ITU-T 勧告 J.341）を提案モデルに適用し、当該モデルの品質推定精度を検証する必要がある。

表 3.7 提案および従来モデルの OR

Experiment	Proposed	Conventional
A1	0.08	0.14
A2	0.22	0.25
B1	0.31	0.36
B2	0.25	0.28
C1	0.20	0.25
C2	0.22	0.33
A1, B1, and C1	0.20	0.25
A2, B2, and C2	0.23	0.30
All PVSs	0.21	0.28

表 3.8 提案および従来モデルに対する SRC ごとの RMSE

SRC	Proposed	Conventional	Improvement ratio
1	0.23	0.27	15%
2	0.27	0.26	-4%
3	0.21	0.19	-10%
4	0.22	0.29	24%
5	0.16	0.27	41%
6	0.20	0.27	25%
7	0.26	0.29	11%
8	0.21	0.24	14%
9	0.16	0.26	36%
10	0.20	0.28	30%
11	0.19	0.30	38%
12	0.33	0.34	5%

表 3.9 SRC ごとの dMOS_LRavg

SRC	1	2	3	4	5	6	7	8	9	10	11	12
dMOS_LRavg	0.59	0.50	0.66	0.74	0.78	0.72	0.44	0.53	0.92	0.75	0.86	0.71

表 3.10 品質推定精度の比較

	Proposed	Conventional
$0 \leq \text{dMOS_LR} \leq 1$	0.22	0.24
$1 < \text{dMOS_LR}$	0.23	0.35

第4章 パケット損失パターンを考慮したIPTVエンドユーザ品質監視法

4.1 まえがき

IP¹ ネットワークおよび映像コーデックの発展により、コンテンツ、ネットワーク、インターネットサービスプロバイダがTVコンテンツをIPネットワーク（IPTV²）を介して提供できるようになった。

IPTVサービスのユーザ体感品質（QoE³）[58]は多くの品質要因に影響を受ける。本検討では、IPTVサービスのQoEの一要因である映像品質に焦点を当て検討する。映像に関する品質要因は以下のように要約できる。まず、伝送前のソース映像コンテンツがどのように符号化されるか、次に、符号化されパケット化されたIPパケットはネットワークにどのように伝送されるか、最後に、符号化された映像がどのようにデコードされ、ユーザ端末（STB⁴、携帯端末）に表示されるかに映像品質は依存する。

ITU-T⁵勧告 G.1081 [83]は図4.1に示すように、サービスプロバイダやネットワークオペレータがエンドユーザに完全なIPTVサービスを提供できているか確認できるように、IPTVサービスの性能監視ポイントを定義している。図4.1に示すように、性能監視ポイントは5つある。ポイント1（PT1）において、ソースメディアおよびメタデータが適切な品質になっているか監視する。ポイント2（PT2）において、符号化され、パケット化されたソースメディアを監視する。ポイント3（PT3）において、パケット伝送特性を監視する。ポイント4（PT4）において、要求された品質を満たしているかを確認するため受信パケットを監視する。ポイント5（PT5）において、表示されるメディアをユーザ端末にて監視する。

PT1および2において品質を監視するいくつかのモデル[23, 24, 25]が提案され、ITUにて標準化されている。PT3においてはIPネットワークに多数のストリームが通過するため、計算負荷の観点からQoSパラメータ（例：ビットレート、パケット損失率）を監視することが適している。しかしながら、PT4および5において品質監視に適用可能なモデルの検討は少ない。

PT4および5のユーザ端末にて個々のユーザ体感品質を監視するためには、ユーザ端末から得られる情報のみで品質を推定するNR⁶客観品質推定モデルの構築が必要である。客観品質推定モデルは入力情報の観点で以下のように分類される[84]。具体的には、映像信号を入力とするメディアレイヤモデル[23, 24, 25, 26, 27, 28, 29, 30]、パケットヘッダを入力とするパケットレイヤモデル[42, 43, 44]、パケットのペイロード（ビットストリーム）

¹Internet Protocol

²IP TeleVision

³Quality of Experience

⁴Set-Top Box

⁵International Telecommunication Union - Telecommunication standardization sector

⁶No Reference

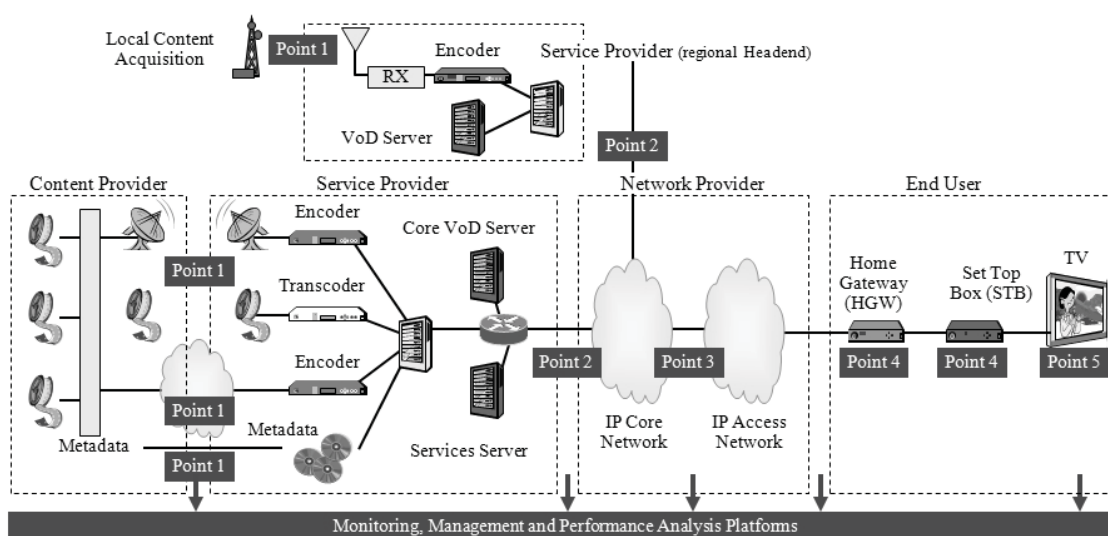


図 4.1 IPTV サービスの品質監視ポイント

を入力とするビットストリームモデル [47, 48]、映像信号、パケットヘッダ、ペイロードを入力とするハイブリッドモデル [53] に分類される。

一般に、ユーザ端末から映像信号、パケットのヘッダやペイロードが得られるため、メディアレイヤ、ビットストリーム、ハイブリッドモデルはユーザ端末における品質監視に適用可能である。しかしながら、HGW⁷やSTBのようなユーザ端末への組み込みを考えた場合、低い計算量で品質を推定することが求められる。このような場合、メディアレイヤモデルは映像信号をスキャンすることで品質を推定できるが計算量を必要とするため、伝送されたパケット情報から品質を推定するパケットレイヤやビットストリームモデルが適している。しかしながら、著作権を保護するため、ビットストリーム（例：動きベクトル、量子化パラメータ）はしばしば暗号化されるため、ビットストリームモデルを用いた品質推定は実施できない。仮に、パケットのペイロードが暗号化された場合でも、パケットレイヤモデルはペイロードではなくパケットヘッダ情報のみを用いるため、品質推定に適用可能である。このような理由から、HGW や STB で品質を監視するためには、パケットレイヤモデルが適している。

パケットレイヤモデルは映像に関するビットストリーム情報を除いたパケットヘッダ（例：IP、UDP⁸、RTP⁹、TS¹⁰、PES¹¹ヘッダ）を用いて典型的な映像コンテンツの平均映像品質を推定する。これらのモデルはビットストリームやコーデック情報（例：コーデック種

⁷Home GateWay

⁸User Datagram Protocol

⁹Realtime Transport Protocol

¹⁰Transport Stream

¹¹Packetized Elementary Stream

別、実装)にアクセスしないため、映像コンテンツやコーデックが映像品質に与える影響を加味することができない。そのため、映像コンテンツやコーデックに関して仮定を置き、品質を推定する必要がある。

従来のパケットレイヤモデルやビットストリームモデルは、文献[42, 43, 44]および[47, 48]にそれぞれ提案されている。文献[42, 43]で提案されるパケットレイヤモデルを用いた場合、ビットレートおよびランダムパケット損失率の関係が考慮されている。しかしながら、ネットワークでは連続パケット損失(バーストパケット損失)がしばしば発生する。そのような場合、ランダムパケット損失特性に基づくのみではパケット損失により低下する映像品質を精度良く推定できない[42]。これより、品質推定精度を改善するためには、パケットレイヤモデルを用いた場合、パケット損失パターンを考慮し映像品質を推定する必要がある。パケットレイヤモデルはビットストリームやメディア信号を用いないため、映像コンテンツが品質に与える影響を考慮できない。一方、ビットストリームモデルはビットストリーム(例:符号化パラメータ(フレームレート、動きベクトル、量子化パラメータ))を用いるため、映像コンテンツの品質への影響を考慮可能である。しかしながら、ビットストリームは著作権保護のためしばしば暗号化される。前述のように、パケットレイヤモデルはビットストリームではなくパケットヘッダのみを用いるためパケットペイロードが暗号化されていても品質推定を実施できる。

まず、IPTVの平均映像品質を推定するパケットレイヤモデルのフレームワークを提案する。次に、主観品質評価実験より符号化およびパケット損失に対する映像品質評価特性を導出する。実験より導出される品質評価特性に基づき、符号化およびパケット損失に対する平均映像品質を推定するパケットレイヤモデルを構築する。提案モデルの数式の形状は種々の映像コーデックや映像グループに対し統一的に扱えることを示す。次に、学習データと異なる非学習データに対し提案モデルの品質推定精度が高いことを示す。最後に、本検討をまとめ、今後の課題について述べる。

4.2 提案フレームワーク

4.2.1 コンセプト

パケットヘッダ(例:IP、UDP、RTP、TS、PESヘッダ)を用い品質を推定するパケットレイヤモデルはユーザ端末(PT4)でのインサービス品質監視に適している。IPTVの映像品質を監視する提案モデルを図4.2に示す。

IPTVサービスでは多くのプロトコルスタック(例:TCP/IP、UDP/IP、RTP/UDP/IP、TS/UDP/IP、TS/RTP/UDP/IP)を用いることができるが、放送、再送信、VoDにおけるIPTVサービスではTS/UDP/IPもしくはTS/RTP/UDP/IPが主に用いられている。そこで、提案モデルの入力はTS/UDP/IPもしくはTS/RTP/UDP/IPのヘッダとする。

提案モデルは映像コンテンツごとの映像品質評価値を出力するのではなく、想定された映像コンテンツに対し平均された平均映像品質(V_q)を出力する。品質を精度良く推定するためには、利用できる情報をすべて用いることが理想的である。しかしながら、コーデック

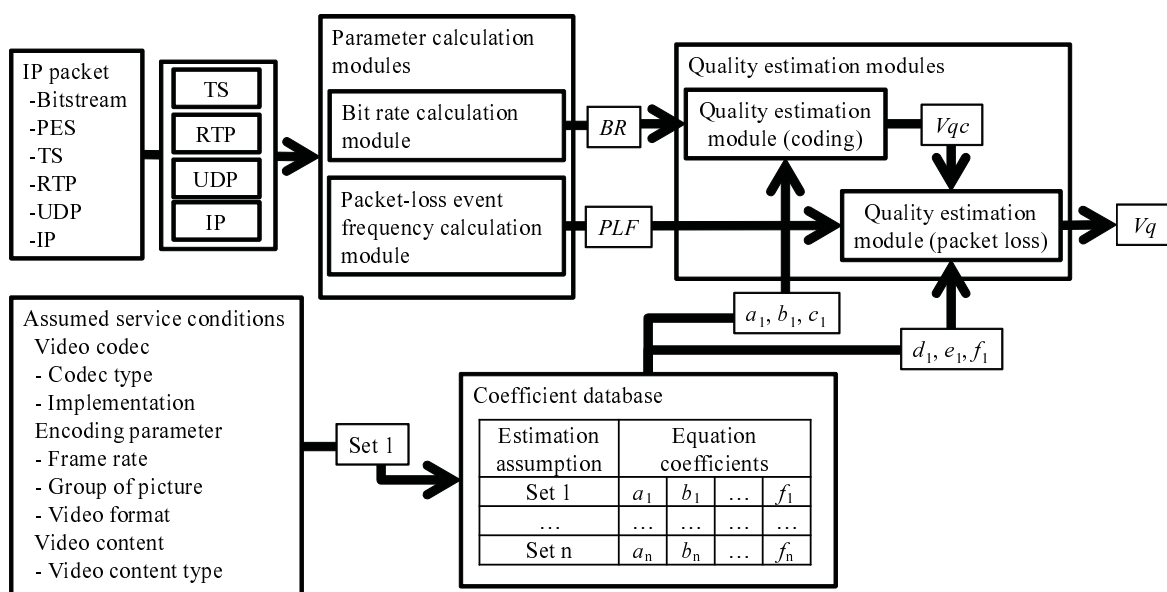


図 4.2 HD 映像の IPTV サービスの品質を監視するパケットレイヤモデル

ク種別 (例: MPEG-2、MPEG-4、H.264)、コーデックの実装 (例: 動き検索、レートディストーションアルゴリズム)、符号化パラメータ (例: 解像度、GoP¹²、フレームレート)、映像コンテンツの情報はパケットヘッダに含まれないため、これらの情報をパケットヘッダから直接得ることはできない。そのため、提案モデルはこれらの条件に対し事前情報を持つ必要がある。

IPTV サービスプロバイダは事前情報 (例: コーデック種別、実装、符号化パラメータ、映像コンテンツ) を把握しているため、事前情報は、例えば、IPTV サービスプロバイダによって提供される。

映像コンテンツセットの選定は、IPTV サービスプロバイダの品質監視方針に基づく。例えば、軽微な品質低下を検知したい場合、符号化効率の低い映像コンテンツセットを選択する。

提案モデルはビットレートおよびパケット損失に影響される映像品質を推定する。提案モデルをユーザ端末に組み込む場合、ジッタバッファにより廃棄されるパケットや、FEC¹³により回復されるパケットを加味し、提案モデルはパケット損失を計算できる。ただし、本検討において、ARQ¹⁴を検討対象外としており、提案モデルはリバッファリングにより影響される映像品質低下を考慮することができない。

一般に、映像コンテンツのビットレート (BR ¹⁵) やパケット損失率 (PLR ¹⁶) が同じ

¹²Group of Picture

¹³Forward Error Correction

¹⁴Automatic ReQuest

¹⁵Bit Rate

¹⁶Packet-Loss Rate

場合（例： $BR = 10$ Mbps、 $PLR = 0.1$ %）でも、映像品質低下は想定されるサービス条件に依存する [85]。しかしながら、品質パラメータ（例：ビットレート、パケット損失率）に対する映像品質低下の傾向は想定されるサービス条件に依存しない。例えば、映像品質はビットレートの増加とともに増加する。また、映像品質はパケット損失率の増加とともに低下する。このように、提案モデルの数式の形状は統一的に扱うことができるが、各想定されるサービス条件ごとにモデル係数を最適化する必要がある。

4.2.2 機能

符号化およびパケット損失による品質低下はビットレートやパケット損失に主に依存するため、パラメータ計算モジュールはビットレート (BR) およびパケット損失 (PL) を計算する。まず、映像パケットは TS ヘッダ内の PID (ITU-T 勧告 H.222 参照) により検出される。次に、ビットレート (BR) は検出された映像 TS パケット数に基づき計算される。最後に、パケット損失 (PL) は RTP ヘッダ内の RTP シーケンス番号もしくは、TS ヘッダ内の連続カウンタ (ITU-T 勧告 H.222 参照) に基づき計算される。

ビットレート (BR) およびパケット損失 (PL) を用い、品質推定モジュールは映像品質を推定する。前述のように、提案モデルの数式の形状（例：ロジスティック関数、指数関数）は想定されるサービス条件に対し同一とした。

提案モデルの数式の係数は想定されたサービス条件の下、実施された主観品質評価データを用い最適化される。これらの係数は係数データベースに保存される。係数データベースを構築するために、サービスプロバイダは保有するサービスシステムに対し主観品質評価実験を実施し、非線形回帰により係数を導出する必要がある。例えば、異なる GoP に対する映像品質は異なるため、各想定された符号化パラメータに対し、係数を選択する必要がある。加えて、パケットレイヤモデルは定義により映像コンテンツが与える品質への影響を加味することができないため、映像コンテンツ種別に対し仮定を置く必要がある。つまり、時空間情報の観点で、学習データセットに対する映像コンテンツの特性と、監視する映像コンテンツが異なる場合、モデル係数を変更する必要がある。

4.2.3 モデルの動作

提案モデルは以下のように動作する。まず、IPTV サービスプロバイダは想定されるサービス条件（例：図 4.2 の Set i （コーデック：H.264、フレームレート：30 fps、GoP： $M = 3$ 、 $N = 15$ 、解像度：HD、映像コンテンツ：映像品質に対し厳しいコンテンツ）を設定し、パケットヘッダ（例：IP、UDP、RTP、TS ヘッダ）を提案モデルに入力する。次に、パケットヘッダに基づきパラメータ (BR 、 PL) を計算し、想定されたサービス条件に基づいて係数データベースからモデル係数 (a_i 、 b_i 、..., f_i) を選定する。次に、パラメータおよび提案モデルの係数が品質推定モジュールに入力される。符号化に対する品質推定モジュールはビットレート (BR) を用い、符号化に対する平均映像品質 (V_{qc}) を推定する。最後に、パケット損失品質推定モジュールは符号化に対する平均映像品質 (V_{qc}) およびパケット損失 (PL) を用い、想定された映像セットに対する平均映像品質 (V_q) を出

表 4.1 各映像グループの映像コンテンツ
(a) 映像グループ A

No.	Title	Criticality [bit/pixel]
1	European market	0.3
2	Harbour scene	0.4
3	Whale show	0.7
4	Soccer action	0.8
5	Green leaves	1.1
6	Japanese room	0.2
7	Ice hockey	0.2
8	Weather report	0.1

(b) 映像グループ B

No.	Title	Criticality [bit/pixel]
1	Streetcar	0.2
2	Opening ceremony	1.0
3	Crowded crosswalk	0.3
4	Boy and toys	0.2
5	Buildings along the canal	0.3
6	Baseball	0.3
7	Summertime tanning	0.3
8	Flamingos	0.4

力する。

4.3 主観品質評価

パケットレイヤモデルの構築に必要な映像品質評価特性を導出するため観視システムを構築した。4.2節において、図4.2に示すように想定されるサービス条件ごとに最適化したモデル係数を変更することで、想定された映像コンテンツセットに対する平均映像品質を推定するモデルを提案した。そこで、提案フレームワークの有効性を検証するため、異なる2つのH.264コーデックおよび異なる2つの映像グループを用い、主観品質評価実験を3つ実施した。

ITU-T勧告P.910に示されるように、様々な時空間特性（例：空間的な精細さ、動き）を持つ映像コンテンツを選定する必要がある。そこで、表4.1に示す10秒の映像コンテンツを16種類用いた。クリティカリティ（ITU-R勧告BT.1210.3の図5参照）の範囲がほぼ同等となるように、映像コンテンツを2つのグループに分類した。映像グループAの映像コンテンツは実験1および2に、映像グループBの映像コンテンツは実験3に用いた。

符号化パラメータとして、表4.2に示す解像度、フレームレート、GoPを用いた。実験パラメータは表4.2に示すように、ビットレート（ BR ）、パケット損失回数（ PLF^{17} ）、平均バーストパケット損失長（ ABL^{18} ）とした。パケット損失を発生させるため、ネットワー

¹⁷Packet-Loss-event Frequency

¹⁸Averaged Burst packet-loss Length

表 4.2 実験条件 (実験 1、2、3)

(a) コーデック			(b) 符号化パラメータ		
Experiment 1	Product A without PLC		Video format	HD (1440 × 1080)	
Experiment 2	Product B with PLC		Frame rate	30 fps	
Experiment 3	Product A without PLC		GoP	M=3, N=15	

(c) ビットレート BR [Mbps]			(d) パケット損失回数 PLF		
Experiment 1	Experiment 2	Experiment 3	Experiment 1	Experiment 2	Experiment 3
20 p (Note 1)	20 p	20 p	0	0	0
18	18	17	1	1	1
16 p	16 p	15 p	2	2	2
14 p	14 p	13	3	3	3
12	12	11 p	5	4	5
10 p	10 p	9 p	10	5	10
8	8	7			
6 p	6 p	6 p			
4	4	4			
2	3	2			

(e) 平均バーストパケット損失 ABL [Packets]		
Experiment 1	Experiment 2	Experiment 3
1	1	1 (1/1) (Note 2)
2	2	2 (1/3)
3	3	8 (1/15)
5	5	-
10	10	-

注 1: p はパケット損失条件である。

注 2: 各 BL の範囲はカッコ内に示した。左の値は BL の最小値、右の値は BL の最大値とした。

クエミュレータを用いた。実験 1 および 2 ではバーストパケット損失長 (BL) は一定、実験 3 では一様分布に基づきバーストパケット損失長 (BL) を変化させた。表 4.2 に示すように、パケット損失は特定のビットレート (BR) にのみ発生させた。

実験において、2つの実装の異なる H.264 エンコーダ (製品 A、製品 B) を用いた。製品 A は実験 1 および 3 に、製品 B は実験 2 に用いた。パケット損失劣化隠ぺい処理 (PLC¹⁹) のないデコーダを実験 1 および 3 に用いた。実験 1 および 3 においては、PLC を用いていないため、パケット損失時にブロックノイズが発生した。実験 2 においては PLC 付きのデコーダを用いた。実験 2 においてはパケット損失時に映像フレームのフリーズが発生した。

1つのパケットは、188 バイトの TS パケット 7 つで構成した。実験 1、2、3 において、3 つの実験パラメータおよび映像コンテンツを用い、880 (110 条件 × 8 映像)、880 (110 条件 × 8 映像)、480 (60 条件 × 8 映像) 映像を作成した。

主観品質評価実験において、映像品質を DCR²⁰法にて評価した [3]。評価者は 1 対の評価

¹⁹Packet-Loss Concealment

²⁰Degradation Category Rating

映像を観視した：最初に提示する映像は常に無圧縮のソース映像とし、次に提示する映像は品質レベルの異なる同一映像シーケンスとした。評価者は5段階劣化尺度（5：劣化が全く認められない、4：劣化が認められるが気にならない、3：劣化がわずかに気になる、2：劣化が気になる、1：劣化が非常に気になる）を用いてソース映像と比較して2番目に提示した映像の劣化の程度を評価した。評価尺度は日本語で提示した。映像シーケンスおよび試験条件の提示順はランダムとした。

20から39歳の24名の評価者が各主観品質評価実験に参加した。評価者は映像品質に関する業務に関わらない非専門家とした。評価者は視距離約110 cm(3H、H：画面高)で各映像シーケンスを評価した。

映像品質主観評価値 (V_q) は表 4.1 に示される8つの映像コンテンツの平均値であるDMOS²¹で表した。

4.4 主観品質評価特性

各実験に対し、符号化およびパケット損失に対する映像品質評価特性を得た。

4.4.1 符号化に対する品質評価特性

パケット損失率 (PLR) が0%の場合、各実験に対し符号化により影響を受けた平均映像品質主観評価値 (V_q 、同一実験条件における8つの映像コンテンツの品質評価値の平均値) を図 4.3 に示した。図 4.3 には95%信頼区間も示した。ビットレート (BR) は伝送パケットから計算した。図 4.3 に示すように、映像品質のカーブは映像コーデックや映像グループに依存しているが、映像品質低下の定性的な傾向は映像コーデックや映像グループに依存しなかった。また、平均映像品質主観評価値 (V_q) はビットレート (BR) の増加とともに増加した。この結果はビットレート (BR) の低下が空間品質の低下を引き起こしたことを意味した。各映像コーデックおよび映像グループのビットレート (BR) に対する平均映像品質主観評価値 (V_q) はロジスティック関数で表せることがわかった。ただし、ロジスティック関数の係数は想定されるサービス条件ごとに異なった。

4.4.2 パケット損失に対する品質評価特性

各実験に対し符号化およびパケット損失により影響を受けた平均映像品質主観評価値 (V_q 、同一実験条件における8つの映像コンテンツの品質評価値の平均値) を図 4.4、4.5、4.6 に示した。図 4.4、4.5、4.6 には95%信頼区間も示した。パケット損失率 (PLR) およびパケット損失回数 (PLF) は伝送されたパケットから計算した。パケット損失率 (PLR) に対する映像品質低下の程度はそれぞれ異なるが、映像品質低下の定性的な傾向は映像コーデックの実装や映像グループに依存しなかった。また、図 4.4 に示すように、平均映像品質主観評価値 (V_q) はパケット損失率 (PLR) の増加とともに減少した。更に、平均映像品質主観評価値 (V_q) はバーストパケット損失長 (BL) を考慮することによって推定可能

²¹Degradation Mean Opinion Score

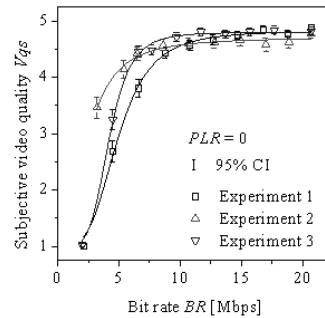


図 4.3 符号化に対する映像品質主観評価特性 (実験 1、2、3)

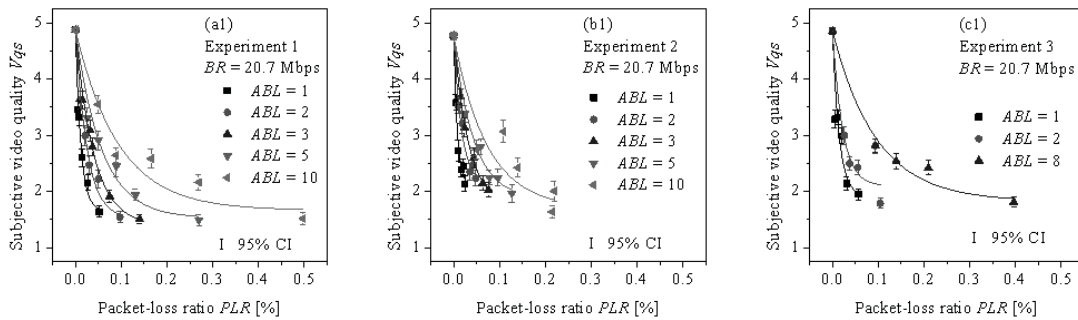


図 4.4 平均バースト損失長ごとの映像品質主観評価値とパケット損失率の関係

であることを示した。しかしながら、パケット損失回数 (PLF) に対する映像品質低下の程度は平均バースト損失長 (ABL) によらずほぼ同様な傾向であり、平均映像品質主観評価値 (V_q) は図 4.5 に示すように、パケット損失回数 (PLF) の増加とともに低下した。加えて、ビットレート (BR) が 6 Mbps 以上の場合、図 4.6 に示すように、ビットレート (BR) によらず、パケット損失回数 (PLF) に対する映像品質低下の程度はほぼ同様な傾向であった。これより、各映像コーデックおよび映像グループにおいて、パケット損失回数 (PLF) に対する平均映像品質主観評価値 (V_q) は指数関数で表せることがわかった。ただし、指数関数の係数は想定されるサービス条件ごとに異なる。

4.5 提案モデル

4.5.1 符号化品質推定モジュール

符号化に対する平均映像品質主観評価値 (V_{qs} 、同一実験条件において想定される映像コンテンツ群の主観品質評価値の平均値) を推定する品質推定モジュールを構築した。4.4.1 節に示すように、平均映像品質主観評価値 (V_{qs}) は以下の数式 (4-1) で示すロジスティック

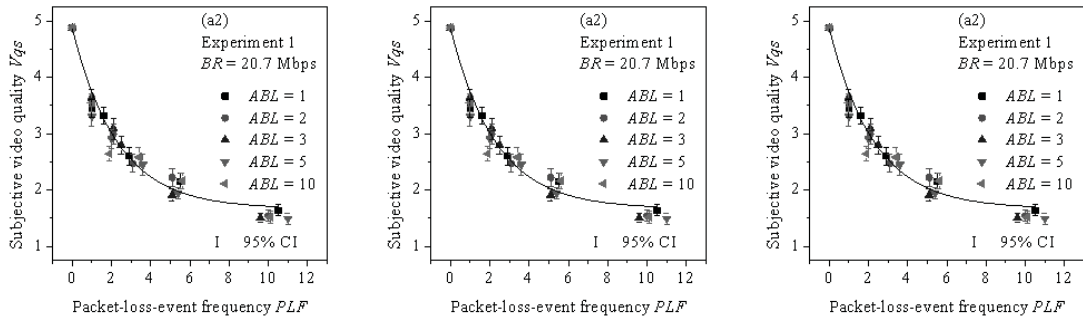


図 4.5 平均バースト損失長ごとの映像品質主観評価値とパケット損失回数との関係

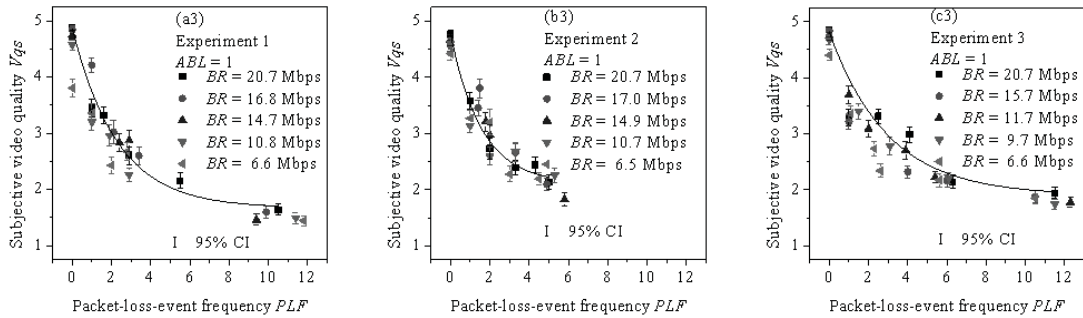


図 4.6 ビットレートごとの映像品質主観評価値とパケット損失回数との関係

ク関数で近似できた。

$$Vqs \approx Vqe \Big|_{PLF=0} = 1 + Ic = 1 + a - \frac{a}{1 + (\frac{BR}{b})^c} \quad (4-1)$$

ただし、 Vqe は平均映像品質推定値、 $1 + Ic$ はパケット損失回数 (PLF) が 0 の場合の平均映像品質評価値、 a 、 b 、 c は想定されたサービス条件に対する係数である。これら係数は学習データに非線形回帰を適用し最適化した。

4.5.2 パケット損失品質推定モジュール

符号化およびパケット損失に対する平均映像品質主観評価値 (Vqs 、同一実験条件において想定される映像コンテンツ群の主観品質評価値の平均値) を推定する品質推定モジュールを構築した。4.4.2 節に示すように、平均映像品質主観評価値 (Vqs) は以下の数式 (4-2) から (4-5) で示す指数関数で近似できた。

$$Vqs \approx Vqe = 1 + IcIp \quad (4-2)$$

$$I_p = (1 - d) \exp\left(-\frac{PLF}{e}\right) + d \exp\left(-\frac{PLF}{f}\right) \quad (4-3)$$

$$e \stackrel{def}{\iff} e \Big|_{BR=BR_{Max}, ABL=1} \quad (4-4)$$

$$f \stackrel{def}{\iff} f \Big|_{BR=BR_{Max}, ABL=1} \quad (4-5)$$

ただし、 I_p はパケット損失回数（ PLF ）に対する映像品質低下の度合い、 BR_{Max} は各実験で用いたビットレートの最大値、 d 、 e 、 f は想定されたサービス条件に対する係数である。これら係数は学習データに非線形回帰を適用し最適化した。

4.6 提案モデルのパフォーマンス評価

4.6.1 品質推定モジュールに対するパフォーマンス要件

提案モデルの品質推定精度の良し悪しを決定するため、相関係数（ r^{22} ）、平均二乗誤差（ $RMSE^{23}$ ）、外れ値率（ OR^{24} ）に対し、以下の要件を設けた。

1. $r \geq 0.94$
2. $RMSE \leq 0.27$
3. $OR \leq 0.66$

上記相関係数（ r ）、平均二乗誤差（ $RMSE$ ）、外れ値率（ OR ）はITU-T 勧告 J.247[23]における二次分析（全実験条件に対する平均値）で用いられた値であるため、品質推定精度の要件としては妥当である。

4.6.2 符号化品質推定モジュールの品質推定精度

符号化に対する品質推定モジュールの有効性を検証するため、非線形最小二乗法に基づき、実験1、2、3に対し、パケットレイヤモデルの係数（ a 、 b 、 c ）を計算した。各実験の係数を変更することによって、平均映像品質主観評価値（ Vq ）を推定した。提案モデルの品質推定精度を図4.7に示した。また、相関係数（ r ）、平均二乗誤差（ $RMSE$ ）、外れ値率（ OR ）も付した。

これらの実験において、相関係数（ r ）、平均二乗誤差（ $RMSE$ ）、外れ値率（ OR ）は4.6.1節で示した品質推定精度の要件を満たした。これより、同一実験条件において想定された映像コンテンツの品質評価値を平均した平均映像品質主観評価値（ Vq ）は数式（4-1）を用い、十分な品質推定精度で推定可能であると結論付けた。

²²Pearson's Correlation Coefficients

²³Root Mean Square Error

²⁴Outlier Ratio

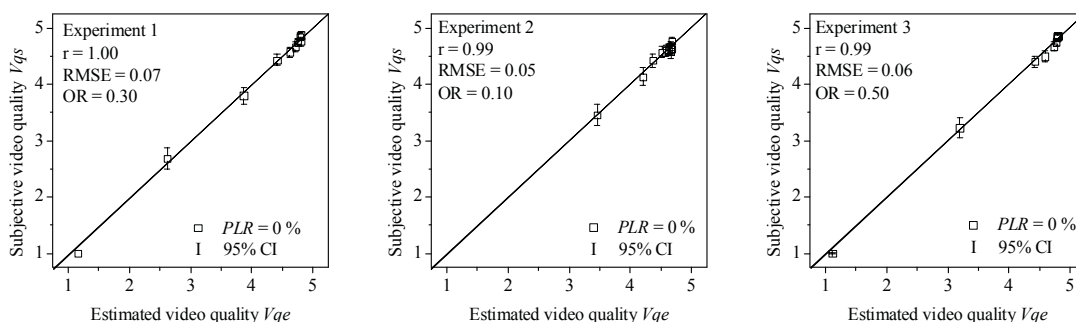


図 4.7 符号化に対する映像品質の推定精度 (実験 1、2、3)

表 4.3 パケットレイヤモデルの係数

Coefficient	Experiment 1	Experiment 2	Experiment 3
a	3.82	3.70	3.80
b	4.91	2.40	4.19
c	3.65	2.31	4.80
d	0.599	0.512	0.816
e	0.948	1.14	0.0305
f	8.04	10.0	6.56

4.6.3 パケット損失品質推定モジュールの品質推定精度

符号化およびパケット損失に対する品質推定モジュールの有効性を確認するため、非線形最小二乗法に基づき、実験 1、2、3 に対し、パケットレイヤモデルの係数 (d , e , f) を計算した。実験 1、2、3 に対する係数を表 4.3 に示した。各実験の係数を変更することによって、平均映像品質主観評価値 (V_q) を推定した。提案モデルの品質推定精度を図 4.8 に示した。また、相関係数 (r)、平均二乗誤差 (RMSE)、外れ値率 (OR) も付した。

これらの実験において、相関係数 (r)、平均二乗誤差 (RMSE)、外れ値率 (OR) は 4.6.1 節で示した品質推定精度の要件を満たした。これより、1) 符号化およびパケット損失により低下した平均映像品質評価値 (V_q) の推定に提案モデルを適用可能、2) パケット損失が発生した際に映像フレームをフリーズさせるパケット損失隠ぺい処理 (PLC) を持つコーデックの品質推定に提案モデルを適用可能、3) 想定されたサービス条件ごとに最適化されたモデル係数を用いることで提案モデルは平均映像品質評価値 (V_q) を推定可能であると結論付けた。

4.6.4 提案モデルの係数の有効性評価

最適化されたモデル係数 (a , ..., f) の有効性を検証するため、表 4.4 に示すように、映像グループ A とは異なる映像グループ B を用いて主観品質評価実験を実施した (映像グルー

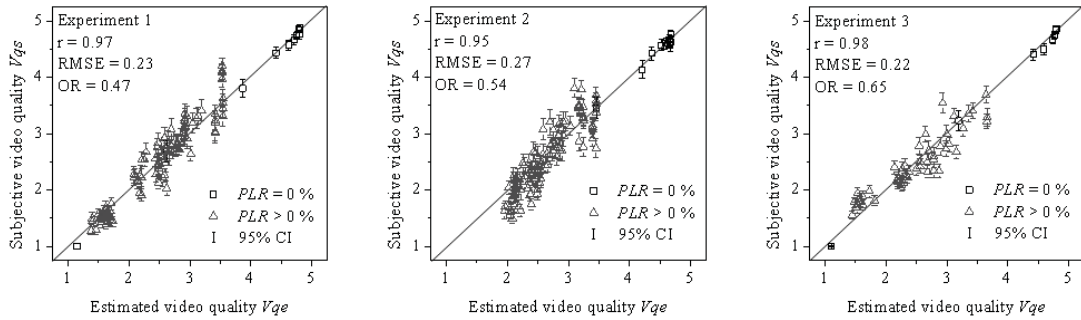


図 4.8 符号化およびパケット損失に対する映像品質の推定精度（実験 1、2、3）

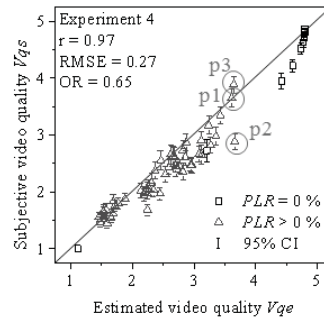


図 4.9 符号化およびパケット損失に対する映像品質の推定精度（実験 4）

プBは実験3と同様)。実験4では、実験3で用いた映像コーデックおよび符号化パラメータを用いた。バーストパケット損失長 (BL) は一様分布に基づき可変とした。表4.4に示すように、パケット損失は特定のビットレート (BR) のみに発生させた。条件数は3つの実験パラメータ (BR 、 PLF 、 ABL) と映像コンテンツ数の積である480映像 (60条件 × 8映像) とした。

表4.3の4列目の係数は実験3により学習したものである。実験4に示した非学習データを用い、これら係数の有効性を検証した。提案モデルの品質推定精度を図4.9に示した。また、相関係数 (r)、平均二乗誤差 (RMSE)、外れ値率 (OR) も付した。非学習データに対し、相関係数 (r)、平均二乗誤差 (RMSE)、外れ値率 (OR) が4.6.1節で示した品質推定精度の要件を満たした。これより、学習および非学習データに対する映像コンテンツの特徴がほぼ同等であれば、想定されたサービス条件に対し最適化された係数を持つ提案モデルは異なる映像コンテンツに対する品質を推定可能であると結論付けることができる。

表 4.4 実験条件 (実験4)

(a) コーデック

Experiment 4	Product A without PLC
--------------	-----------------------

(b) 符号化パラメータ

Video format	HD (1440 × 1080)
Frame rate	30 fps
GoP	M=3, N=15

(c) 実験パラメータ

BR [Mbps]	PLF	ABL [Packets]
20 p (Note 1)	0	1 (1/1) (Note 2)
17	1	2 (1/3)
15 p	2	8 (1/15)
13	3	
11 p	5	
9 p	10	
7		
6 p		
4		
2		

注1: pはパケット損失条件である。

注2: 各 BL の範囲はカッコ内に示した。左の値は BL の最小値、右の値は BL の最大値とした。

4.6.5 損失映像フレーム種別の影響

前述のように、提案モデルが品質推定精度の要件を満たすことを示した。しかしながら、図4.7から4.9に示すように、主観品質と推定品質の差を表す推定誤差が95%信頼区間より大きいものが存在する。本節では、提案モデルの品質推定精度に対し損失した映像フレーム種別の影響がどの程度あるか議論する。

一般に、映像品質は損失映像フレーム種別 (I、B、Pフレーム) に依存する [87]。文献 [87] によると、一つの損失パケットが複数の映像フレームにおいて大きな品質低下を招くことがある。劣化の伝搬時間長は損失したパケットを含む映像フレーム種別とGoP構造に依存する。例えば、図4.9に示すように、実験4において $PLF = 1$ および $ABL = 1$ の条件で、ビットレート (BR) が 20.7 Mbps の平均映像品質評価値 (Vq) は 3.7 (p1)、ビットレート (BR) が 15.7 Mbps の平均映像品質評価値 (Vq) は 2.9 (p2)、ビットレート (BR) が 11.7 Mbps の平均映像品質評価値 (Vq) は 3.9 (p3) であった。次に、一つの損失パケットによって劣化した映像フレーム数は、それぞれ、5.8、14.0、5.6 であった。これより、品質推定誤差は損失した映像フレームによって引き起こされる劣化の伝搬時間長に依存していることがわかる。しかしながら、映像フレーム種別はビットストリームに示される情報のため、定義によって、提案モデルは映像フレーム種別を用いて平均映像品質評価値 (Vq) を推定することができない。これより、これらの品質推定誤差は提案パケットレイヤモデルでは避けられないものと結論付けることができる。

4.7 まとめ

符号化およびパケット損失に対する平均映像品質評価値 (V_q 、同一実験条件において想定された映像コンテンツ群の品質評価値の平均値) を精度良く推定するパケットレイヤモデルを提案した。まず、2つの映像コーデックおよび2つの映像グループを用い、符号化およびパケット損失に対する2つの映像品質評価特性を導出するため、3つの主観品質評価実験を実施した。次に、パケット損失隠ぺい処理 (PLC) のない映像コーデックに対しパケットレイヤモデルを構築した。その後、モデル係数を変更することによって、パケット損失隠ぺい処理 (PLC) を持つ異なる映像コーデックおよび異なる映像グループに対する平均映像品質評価値 (V_q) を提案モデルが推定できることを検証した。最後に、非学習データに対し提案モデルの有効性を検証した。これより、符号化およびパケット損失に影響を受ける平均映像品質評価値 (V_q) を提案モデルにより推定できるため、IPTV サービスの平均映像品質評価値 (V_q) を推定する品質監視ツールとして有用であると結論付けた。

以下に今後の課題を示す。IPTV サービスを提供する際に STB を用いる場合は、ユーザ端末は十分なバッファサイズを有していると仮定できるが、モバイル端末に IPTV サービスを提供する場合は、リバッファリングがしばしば発生する [43]。このようなケースに対応するため、リバッファリングに対する主観品質評価実験から得られる品質評価特性に基づいたモデルを構築することが重要である。提案パケットレイヤモデルは映像に関する情報 (例: 映像信号、ビットストリーム) を用いないため、映像コンテンツごとの品質を推定することができない。そこで、映像に関する情報を用いることで映像コンテンツごとの品質を推定可能なモデルに拡張することが重要である。

第5章 映像フレームを用いたIPTVエンドユーザ品質監視法

5.1 まえがき

コーデック [1, 2] およびネットワーク技術は目覚ましく発展を遂げてきた。また、IPTV¹に関する国内および国際標準化が進められてきた [88, 89, 90, 91, 92, 93]。この結果、コンテンツ、ネットワーク、インターネットサービスプロバイダがHDTV²コンテンツをIPネットワーク経由で提供できるようになった。

一般にIPTVサービスはUDP³もしくはTCP⁴パケットにより提供される。伝送遅延を避けるために、放送およびVoD⁵においてはUDPベースのストリームが広く用いられている。一方、TCPやARQ⁶はVoDにおいて損失したパケットを回復するための再送方式として用いられている。IPTVサービスにおいてはUDPベースのストリームが幅広く用いられているため、本検討では、UDPベースのストリームに焦点を当て検討する。つまり、TCPやARQベースのストリームによる再送は本検討のスコープ外とする。

IPTVサービスのQoE⁷[58]はメディア品質、サービス料金やカスタマーサポート等の多くの要因に影響を受ける。IPTVサービスのQoEにおいては映像品質(VQ⁸)が支配要因であるため、本検討ではQoEの一部として映像品質に焦点を当て検討する。

コンテンツ、ネットワーク、インターネットサービスプロバイダが高い品質のIPTVサービスを提供するためには品質設計および監視を実施することが重要である。まず、映像品質主観評価特性に基づき、適切なコーデック(例:MPEG-2、H.264)および符号化パラメータ(例:解像度、フレームレート、GoP⁹、ビットレート)が決定され、ネットワークパフォーマンスが設計される。これをQoE設計と言う。次に、この設計されたサービスがエンドユーザに提供される間に、エンドユーザが高い品質のIPTVコンテンツを視聴しているかを確認するため、QoEはリアルタイムで監視される。これをインサービスQoE監視と言う。仮に、客観品質推定モデルが確立されていなくとも、主観品質評価試験の結果はQoE設計に適用可能である。しかしながら、プロバイダがQoEを監視する場合、QoE監視はリアルタイムで実施される必要があるため、客観品質推定モデルが必要である。

サービス品質管理はヘッドエンド、ネットワーク、エンドユーザ宅の3つのドメインで実施される。

ヘッドエンドにおいて、符号化品質を確認するためヘッドエンドQoE監視が実施される。

¹Internet Protocol TeleVision

²High Definition TeleVision

³User Datagram Protocol

⁴Transmission Control Protocol

⁵Video on Demand

⁶Automatic Repeat reQuest

⁷Quality of Experience

⁸Video Quality

⁹Group of Picture

ソース映像信号、すなわち、画素信号が利用可能なら、ソースおよび符号化映像信号を入力とするFR¹⁰メディアレイヤモデル [23, 24, 25, 30, 30] が適している。一方、ソース映像信号が利用できない場合（例：映像が異なる場所で符号化されるなどソース映像信号が存在しない場合）、NR¹¹メディアレイヤモデル [29, 32, 33] が適している。

ネットワークにおいて、ネットワークパフォーマンスを確認するためにネットワーク監視が実施される。IP ネットワークを多数のストリームが流れるため、計算量の観点から、スループット、パケット損失率や遅延のような QoS¹²パラメータを監視することが適している。

エンドユーザ宅において、ソース映像、エンコーダ、ネットワークパフォーマンス、デコーダおよびディスプレイに影響を受ける QoE を確認するため、エンドユーザ QoE 監視が実施される。RR¹³メディアレイヤモデル [34, 35]、NR メディアレイヤモデル、NR パケットレイヤモデル [43, 44, 45, 46]、NR ビットストリームモデル [47, 49, 50, 51, 52]、NR ハイブリッドモデル [54, 55, 56, 57] はエンドユーザ QoE 監視に適用可能である。QoE を推定するために、RR メディアレイヤモデルは劣化映像信号およびソース映像信号から導出される特徴量を入力とし、NR メディアレイヤモデルは劣化信号を入力とする。パケットレイヤモデルはビットストリームを除く伝送されたパケットヘッダ（例：IP、UDP、RTP¹⁴、TS¹⁵[93]、PES¹⁶[93]）を入力とする。ビットストリームモデルはビットストリーム（例：量子化パラメータ、動きベクトル）を入力とし、ハイブリッドモデルは映像信号、伝送されたパケットヘッダ、ビットストリームのいずれかの組を入力とする。

FR および RR メディアレイヤモデルは標準技術もあり、幅広く検討されているが、エンドユーザ QoE 監視に適用可能な NR メディアレイヤ、パケットレイヤ、ビットストリームおよびハイブリッドモデルは標準技術もなく、現在検討中である。そこで、本検討ではエンドユーザ QoE 監視に焦点を当て検討する。

エンドユーザ QoE 監視は一時監視と常時監視に分類できる。一時監視は、例えば、修理作業員が装置を修理する場合などに実施される。これをオンサイトエンドユーザ品質監視と言う。一方、常時監視は全ユーザもしくは一部のグループのユーザの QoE をリアルタイムで監視する際に実施される。これをリアルタイムエンドユーザ QoE 監視と言う。オンサイトエンドユーザ QoE 監視においては、修理作業員がラップトップコンピュータのような装置をエンドユーザ宅に持っていくことが可能なため、計算量はあまり重要な問題ではない。そのため、一般に、メディアレイヤやハイブリッドモデルは映像信号をスキャンし、QoE を推定するために多くの計算量を必要とするが、それらモデルはオンサイトエンドユーザ QoE 監視に適用可能である。しかしながら、リアルタイムエンドユーザ QoE 監視において

¹⁰Full Reference

¹¹No Reference

¹²Quality of Service

¹³Reduced Reference

¹⁴Real-time Transport Protocol

¹⁵Transport Stream

¹⁶Packetized Elementary Stream

は、文献 [43] で述べられるように、STB¹⁷やHGW¹⁸のようなユーザ端末に品質推定モデルを組み込むため、費用の観点から多くの計算量を必要とするモデルは致命的である。そのため、低計算量のNR客観品質推定技術をユーザ端末に組み込むことが望まれている。

伝送されるパケットが暗号化されている場合、ビットストリームモデルはビットストリームを入力として必要とするため、映像品質監視に適用することはできない。そのため、パケットレイヤモデルがリアルタイムエンドユーザQoE監視により適している。

そこで、我々はビットレート (BR¹⁹) およびパケット損失情報からIPTVサービスプロバイダが想定する典型的な映像コンテンツの平均映像品質評価値を推定するパケットレイヤモデルを提案した [45]。前記モデルに関しては、伝送されるパケットヘッダ情報のみを用いるため、映像コンテンツに仮定を置く必要がある。すなわち、ユーザが知覚する映像品質とモデルにより推定される映像品質に差が生じる。

このような背景から、映像コンテンツごとの映像品質 (Q) を推定可能なNR映像品質推定モデルの構築に焦点を当て検討をする。前述のように、従来のパケットレイヤモデルは映像コンテンツが映像品質に与える影響を加味することができない。そこで、推定対象の映像コンテンツの映像品質と文献 [45] のパケットレイヤモデルで推定された平均映像品質評価値の差分値を用いるモデルを提案する。提案モデルはビットレートBR (B) [43, 44, 45, 46] および従来モデルが用いるパケット損失により劣化した映像フレーム数DF²⁰ (D) [52] に加え、Iフレームのフレーム平均ビット量ABI²¹ (BI) を品質推定に用いる。一般に、映像フレーム種別 (I、B、Pフレーム) は伝送されるパケットヘッダに示されていない。そこで、文献 [94] の映像フレーム種別推定モデルのI、B、Pフレーム推定の誤差が非常に小さいため、文献 [94] で提案された映像フレーム種別推定モデルを提案モデルに組み込むことを前提とする。

5.2 従来検討

従来技術は映像に関するビットストリームを除いたパケットヘッダのみから典型的な映像コンテンツの平均映像品質評価値を推定するパケットレイヤモデル [43, 44, 45, 46] である。これらのモデルは映像信号、ビットストリーム、コーデック情報 (例: コーデック種別、実装) にアクセスしないため、映像コンテンツやコーデックの特徴に関し、仮定を置く必要がある。

ビットレートを用い符号化に対する平均映像品質評価値を推定する従来技術 [43, 44, 45] が提案されている。一般に、文献 [55] で述べられるように、符号化に対する映像品質は映像コンテンツに依存し、符号化に対する映像品質はビットレートを用いるのみでは推定できない。すなわち、モデルは映像コンテンツが映像品質に与える影響を考慮するために、

¹⁷Set-Top Box

¹⁸Home-GateWay

¹⁹Bit Rate

²⁰Damaged video Frame

²¹Average I-frame Bits

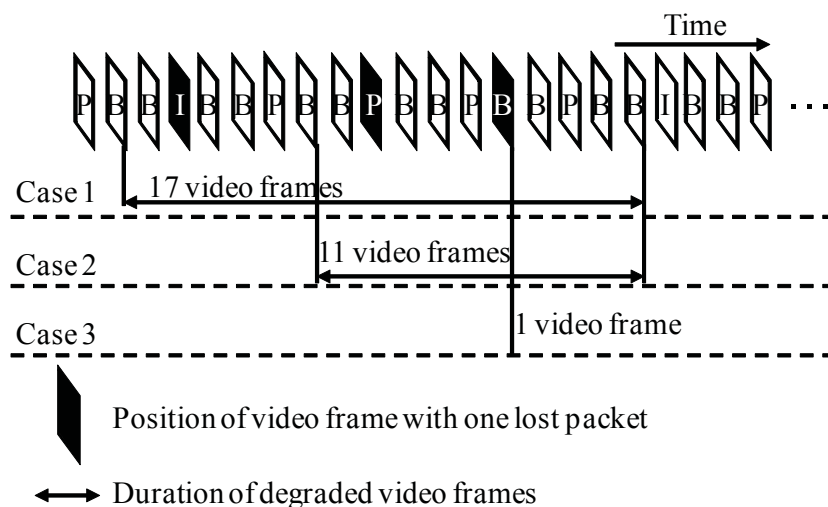


図 5.1 劣化映像フレーム数

その他の情報を用いて映像品質を推定する必要がある。

パケット損失率を用い平均映像品質評価値を推定するモデルが提案されている [43, 44]。しかしながら、連続パケット損失（バーストパケット損失）はしばしばネットワークで発生する。そのような場合、パケット損失に対する映像品質はランダムパケット損失特性のみに基づいて推定することができない [42, 48, 95]。このような技術的背景から、品質推定精度を改善するために、パケット損失回数（PLEF²²、例：一度に、5つのパケットが連続して損失した場合、パケット損失回数は1とする）に基づき平均映像品質評価値を推定する従来モデル [45, 46] が提案されている。しかしながら、文献 [52] に示されるように、パケット損失に対する映像品質は損失した映像フレーム種別（I、B、Pフレーム）の位置に依存するため、パケット損失回数を用いる従来モデル [45, 46] は映像コンテンツごとの映像品質を推定するという観点では十分な品質推定精度に達していない。加えて、図 5.1 に示すように、一つの損失パケットは複数の映像フレームを損失させる重大な品質劣化を引き起こす可能性がある。劣化映像フレームは損失したパケットを持つ映像フレームの種別（I、B、Pフレーム）およびGoP構造に依存する。例えば、図 5.1 の Case 1 において、Iフレームの一つのパケットが損失すると 17 映像フレームが劣化する。Pフレームのパケットが1つ損失すると、5から 14 映像フレームが劣化し、劣化映像フレーム数は損失したパケットを持つPフレームの位置に依存する（図 5.1 の Case 2 の場合、劣化映像フレーム数は 11 である）。図 5.1 の Case 3 において、Bフレームの一つのパケットが損失した場合、1 映像フレームが損失する。すなわち、パケット損失に対する映像品質を推定するためには劣化映像フレーム数を考慮する必要がある。

²²Packet-Loss-Event Frequency

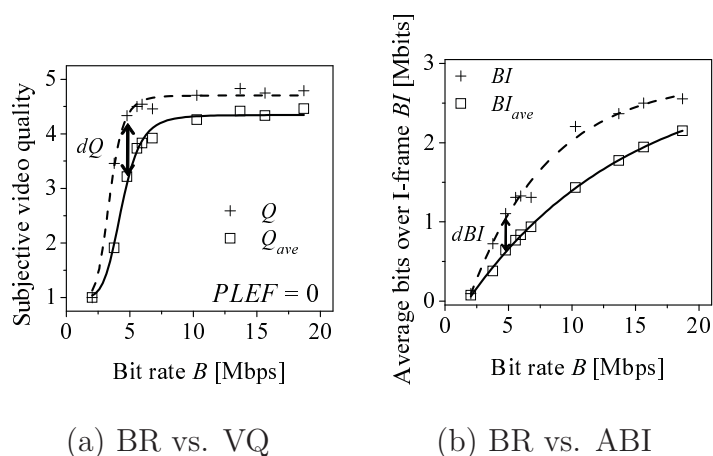


図 5.2 映像コンテンツが映像品質主観評価値およびIフレームのフレーム平均ビット量へ与える影響

5.3 提案モデルのコンセプト

5.2節で述べたように、ビットレートおよびパケット損失情報を入力とする従来のパケットレイヤモデルは映像コンテンツごとの映像品質を推定することはできないため、品質推定精度の改善の余地が十分ある。5.2節で述べた課題を解決するため、推定対象の映像の映像品質評価値と文献[45]で提案されたパケットレイヤモデルを用い推定した平均映像品質評価値 (Q_{ave}) の差分値 (dQ) に基づき映像品質評価値を推定するモデルを提案する。映像品質差分値 (dQ) は映像フレームごとのビット配分に依存し、特に、Iフレームのビット量に依存する。図5.2に示すように、例えば、ビットレート (B) が10 Mbpsの場合、Iフレームのフレーム平均ビット量 (BI) およびIフレームのフレーム平均ビット量の平均値 (BI_{ave} : 典型的な映像コンテンツのIフレームのフレーム平均ビット量の平均値) の差分値が映像品質差分値 (dQ) に影響を与えている。加えて、映像品質差分値 (dQ) はビットレート (B) にも依存している。これより、映像品質差分値 (dQ) はIフレームのフレーム平均ビット量 (BI) およびIフレームのフレーム平均ビット量の平均値 (BI_{ave}) の差分値 (dBI) およびビットレート (B) を用いて推定する。提案モデルのブロック図を図5.3に示す。提案モデルの数式は、主観品質評価実験より得られた主観品質評価特性を5.4節で述べた後、5.5節で示す。

5.3.1 提案モデルの定義

提案モデルはパケットヘッダから映像コンテンツごとの映像品質 (Q) を出力する。

5.1節で述べたように、提案モデルはPES/TS/UDP/IPもしくはPES/TS/RTP/UDP/IPで構成されたUDPベースのストリームを入力とする。しかしながら、提案モデルは映像フレームの区切りを示す情報がパケットヘッダに示されていない場合は適用できない。提案モ

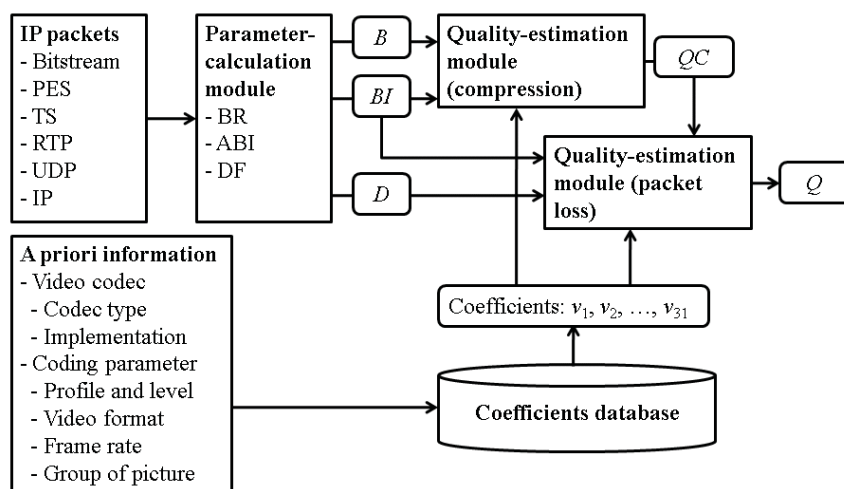


図 5.3 IPTV サービスに対する NR 映像品質推定モデル

モデルは以下の情報を映像フレームの区切りを識別するために用いる。映像フレームの区切りを示す情報は TS ヘッダ内の PUSI²³[89, 90, 93] もしくは PES ヘッダ内の PTS²⁴[89, 90, 93] に示されている。PUSI や PTS が映像フレームの区切りを示すか否かは、国内もしくは国際標準化機関が発行する勧告に示されている。このように映像フレームの区切りがパケットヘッダに示される場合、5.1 節で述べたように、映像フレームごとのパケット数に基づき映像フレーム種別 (I、B、P フレーム) を推定するモデルが文献 [94] で提案されているため、提案モデルでは、文献 [94] の映像フレーム種別推定モデルを用いることを前提とする。そのため、映像フレーム種別推定の検討は本検討のスコップ外とする。

映像品質を精度良く推定するためには、利用可能なすべての情報を用いることが理想的である。しかしながら、提案モデルはコーデック種別 (例: MPEG-2、MPEG-4、H.264) やコーデックの実装 (例: 動き検出アルゴリズム、レートディストーションアルゴリズム、パケット損失隠ぺい処理 (PLC²⁵)) に関する情報にアクセスしない。加えて、符号化パラメータ (例: プロファイル、レベル、解像度、フレームレート、GoP) は暗号化されているため、品質推定に用いることができない。そこで、提案モデルは、図 5.3 に示すように、事前情報を必要とする。IPTV サービスプロバイダは TV コンテンツを符号化するため、事前情報 (映像コーデック、符号化パラメータ) を知っているため、事前情報は IPTV サービスプロバイダによって提供することができる。

映像信号を用いることなしに、種々のコーデックに対し不変的なモデルを開発することは困難であるため、事前情報に対しモデル係数を最適化する文献 [45] のフレームワークを提

²³Payload-Unit-Start Indicator

²⁴Presentation Time Stamp

²⁵Packet-Loss Concealment

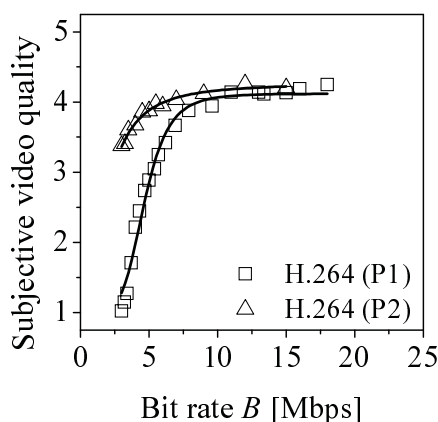


図 5.4 二つの異なるコーデックに対するビットレートおよび映像品質主観評価値の関係

案モデルに適用する。図 5.4 に示すように、各映像コンテンツのビットレート（例： $B = 5$ Mbps）が同じであっても、映像品質の劣化は事前情報に依存する [45]。しかしながら、ビットレートに対する映像品質の劣化の傾向は事前情報に依存しない。例えば、映像品質はビットレートの増加とともに増加 [45] し、映像品質は劣化映像フレーム数の増加とともに減少する [52]。つまり、モデル係数は事前情報（例：映像コーデック、符号化パラメータ）に対し最適化する必要があるが、提案モデルの数式の形状（例：ロジスティック関数、指数関数）は事前情報によらず、同一のものを用いることができる。モデル係数データベースを構築するために、サービスプロバイダは提供中のサービスシステムに対し主観品質評価実験を実施し、得られた主観品質評価値を用い非線形回帰分析を行い係数を最適化する。

5.3.2 提案モデルの機能

係数データベースは提案モデルの係数を保存し、IPTV サービスプロバイダによって提供される事前情報（例：エンコーダ：H.264、プロファイルおよびレベル：High profile level4、解像度：HD、フレームレート：30 fps、GoP：M=3,N=15、デコーダ：PLCなし）に対応する係数を映像品質推定モジュールに出力する。

パラメータ計算モジュールはパケット（例：PES/TS/RTP/UDP/IP）に基づき、パラメータ B （ビットレート）、 BI （Iフレームのフレーム平均ビット量）、 D （劣化映像フレーム数）を計算する。映像パケットはPMT²⁶に示されるPID²⁷[93]によって検出される。次に、検出された映像TSパケット数およびパケット到着時間に基づきビットレート（ B ）が計算される。Iフレームのフレーム平均ビット量はIフレームのTSパケット数およびIフレーム数に基づき計算される。次に、損失映像フレームの位置は、RTPヘッダのRTPシーケンス番号およびTSヘッダ内の連続カウンタ [93]に基づき検出される。最後に、劣化映

²⁶Program Map Table

²⁷Packet Identifier

像フレーム数 (D) は損失映像フレームの位置や映像フレーム種別に基づき算出される。

符号化に対する品質推定モジュールはビットレート (B)、Iフレームのフレーム平均ビット量 (BI)、係数 (v_1-v_{20}) を用い、符号化に対する映像品質評価値 (QC) を導出する。

符号化およびパケット損失に対する品質推定モジュールは符号化に対する映像品質評価値 (QC)、Iフレームのフレーム平均ビット量 (BI)、劣化映像フレーム数 (D)、係数 ($v_{21}-v_{31}$) を用い、映像品質評価値 (Q) を導出する。

5.4 主観品質評価

提案モデルを構築するために必要な主観品質評価特性を導出することに加え、提案モデルの有効性を検証するため主観品質評価実験を実施した。

5.4.1 実験条件

提案モデルを学習することに加え、非学習データに対する映像コンテンツごとの映像品質を精度良く推定できることを検証するため、16の無圧縮のソース映像コンテンツ (SRC²⁸) に対し、4つの主観品質評価実験 (実験1から4) を実施した。5.3節で述べたように、提案モデルはモデルの最適化のために事前情報を必要とする。そこで、文献 [45] で示されるフレームワークを検証するため、2つの異なる実装の H.264 エンコーダ (製品 P1、製品 P2) を用いた。表 5.1(a) で示すように、製品 P1 を実験 1 および 2、製品 P2 を実験 3 および 4 で用いた。これらの実験においては、既存の IPTV サービスで幅広く用いられているパケット損失隠ぺい処理 (PLC) のないデコーダを用いた。つまり、パケット損失が発生した場合、実験で用いたデコーダはブロックノイズを発生させる。これらの実験においては、表 5.1(b) で示す同一の符号化パラメータ (プロファイル、レベル、解像度、フレームレート、GoP) を用いた。ただし、符号化解像度は $1440 \times 1080i$ であるが、ディスプレイにおける表示サイズは 42 インチ LCD モニタのネイティブ解像度である 1920×1080 で表示した。提案モデルが非学習の実験条件に対し映像品質を精度良く推定できることを検証するため、実験パラメータとしてビットレートおよびパケット損失の異なる組を用いた。パケット損失はネットワークで実際に発生するランダムおよびバーストパケット損失の両者を用いた。表 5.1(c) から (e) に示すように、実験パラメータはビットレート (B)、パケット損失回数 ($PLEF$)、平均バーストパケット損失長 (ABL ²⁹、損失したパケット数をパケット損失回数で除算した値) とした。パケット損失を発生させるため、ネットワークエミュレータを用いた。バーストパケット損失長 (BL) は一様分布に基づき可変とし、表 5.1 に示すように、特定のビットレート (B) にのみパケット損失を発生させた。一つの IP パケットは 7 つの TS パケットにより構成され、TS パケットは 188 バイトとした。

ITU-T 勧告 P.910 に示されるように、主観品質評価実験を実施するためには、様々な時空間特性 (例: 精細度、動き) を持った映像コンテンツを選定する必要がある。提案モデルが映像コンテンツごとの映像品質を推定できるか検証するため、各 10 秒 (300 映像フレー

²⁸SouRCe

²⁹Average Burst Packet-loss Length

表 5.1 実験条件 (実験1から4)

(a) コーデック		(b) 符号化パラメータ	
Experiments 1 and 2	H.264 (Product P1*)	Profile and level	High profile level 4
Experiments 3 and 4	H.264 (Product P2*)	Video format	HD (1440 × 1080i)
		Frame rate	30 fps
		GoP	M=3, N=15

*: デコーダはPLCを持たない。

(c) ビットレート B [Mbps]		(d) パケット損失回数 $PLEF$	
Experiment 1	18 p**, 15, 13, 9.6, 6.2, 5.4, 5.0, 4.3 p, 3.4, 2.0	Experiment 1	1, 2, 4, 12
Experiment 2	16, 13.4, 11, 7.9, 6.9 p, 5.7 p, 4.7, 4.0, 3.7, 3.0	Experiment 2	1, 2, 4, 12
Experiment 3	15 p, 9 p, 7, 6, 5, 4.5 p, 3.5 p, 3.0 p	Experiment 3	1***, 2, 4, 7
Experiment 4	12 p, 5.5, 4.0, 3.25 p	Experiment 4	1, 2, 4, 7

**: p はパケット損失条件を表す。

***: $PLEF = 1$ の条件では損失箇所
の異なる三つの条件を用いた。

(e) 平均パケット損失長 ABL [Packets]	
Experiment 1	1 (1/1)****, 2 (1/3), 4 (1/7), 8 (1/15)
Experiment 2	1 (1/1), 2 (1/3), 4 (1/7), 8 (1/15)
Experiment 3	1 (1/1)
Experiment 4	8 (1/15)

****: カッコ内に BL の範囲を示し、左の値は BL の
最小値、右の値は BL の最大値である。

ム) の無圧縮のソース映像を 16 種類用いた [86]。表 5.2 に示すように SI³⁰ および TI³¹ の範囲がほぼ同等となるように、ソース映像を二つの映像グループ A および B に分類した。表 5.2 には、300 映像フレームのうち、最大の SI および TI を示した。

映像グループ A および B に対し、実験 1 および 2 では 336PVS³² (42 HRC³³ × 8 SRC)、実験 3 では 304PVS (38 HRC × 8 SRC)、実験 4 では 128PVS (16HRC × 8SRC) 用いた。PVS 数は前述の三つの実験パラメータとソース映像数の組に対応している。また、符号化に対する HRC 数は、実験 1 および 2 で 10HRC、実験 3 で 8HRC、実験 4 で 4HRC とし、パケット損失に対する HRC 数は実験 1 および 2 で 32HRC、実験 3 で 30HRC、実験 4 で 12HRC とした。

ここで各実験の目的をまとめる。映像グループ A に対する実験 1 および 3 (学習データ) は事前情報に対しモデル最適化のために必要となる学習データとして用いられる。映像グループ B に対する実験 1 および 3 は学習データと同じ HRC であるが異なる映像コンテンツを用いた実験であり、提案モデルが非学習映像に対し映像品質を精度良く推定できるかを確認するために用いられる。映像グループ A に対する実験 2 および 4 は学習データと同

³⁰Spatial Information³¹Temporal Information³²Processed Video Sequence³³Hypothetical Reference Circuit

表 5.2 各映像グループに対する映像シーケンス
(a) 映像グループ A (b) 映像グループ B

No.	Title	SI	TI	No.	Title	SI	TI
1	Soccer action	139	44	9	European market	93	68
2	Green leaves	115	39	10	Harbour scene	112	30
3	Baseball	65	34	11	Whale show	112	45
4	Weather report	81	11	12	Japanese room	66	28
5	Streetcar	77	25	13	Opening ceremony	129	15
6	Buildings along the canal	92	26	14	Crowded crosswalk	79	31
7	Summertime tanning	85	97	15	Boy and toys	53	27
8	Flamingos	39	18	16	Ice hockey	66	26
	Average	87	37		Average	89	34

一の映像コンテンツであるがビットレートおよびパケット損失の組が異なる HRC であり、提案モデルが非学習 HRC に対し映像品質を精度良く推定できるかを確認するために用いられる。映像グループ B に対する実験 2 および 4 は映像コンテンツ、ビットレートおよびパケット損失の組が学習データと異なる実験であり、提案モデルが非学習映像および HRC に対し映像品質を精度良く推定できるかを確認するために用いられる。

主観品質評価実験において、映像品質主観評価値は 5 段階 ACR³⁴法 [3] (5:非常に良い、4:良い、3:普通、2:悪い、1:非常に悪い) を用いて評価された。品質評価は日本語の品質尺度を用いた。主観品質評価実験では、PVS の提示順はランダムとした。

各実験に対し 20 から 39 歳の男女 24 名が参加した。評価者は映像品質評価に関する業務に関わらない非専門家とした。評価者は約 110cm (3H、H: 画面高) の視距離で各 PVS を観視した。

5.4.2 主観品質評価特性

本節では映像グループ A に対する実験 1 の結果を用い、I フレームのフレーム平均ビット量がどのように映像品質に影響するかを説明する。ただし、映像グループ A に対する実験 1 の結果が他の実験の結果とほぼ同等であることは確認済みである。

1) I フレームのフレーム平均ビット量に関する特性

映像グループ A に対する実験 1 のビットレート (B) および I フレームのフレーム平均ビット量 (BI) の関係を図 5.5 に示す。まず、平均 I フレームビット量 (BI_{ave})、最大 I フレームビット量 (BI_{max})、最小 I フレームビット量 (BI_{min}) を定義する。 BI_{ave} は各 HRC に対し 8 映像コンテンツの I フレームのフレーム平均ビット量を平均した値、 BI_{max} は各 HRC に対し 8 映像コンテンツの I フレームのフレーム平均ビット量の中で最大の値、 BI_{min} は各 HRC に対し 8 映像コンテンツの I フレームのフレーム平均ビット量の中で最小の値とする。ビットレート (B) はパケットから計算された。図 5.5 に示すように、ビットレート

³⁴Absolute Category Rating

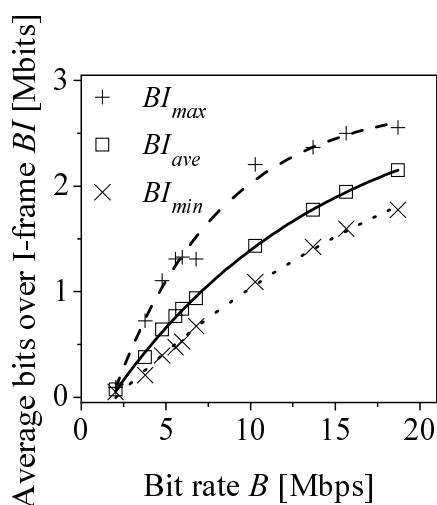


図 5.5 ビットレートおよびIフレームのフレーム平均ビット量の関係

(B) の増加がIフレームのフレーム平均ビット量 (BI) を増加させるため、ビットレート (B) に対する BI_{ave} 、 BI_{max} 、 BI_{min} のカーブは指数関数で表すことができる。しかしながら、ビットレート (B) の増加率は映像コンテンツやコーデックに依存するため、指数関数の係数は映像コンテンツおよびコーデックにより異なる。ビットレート (B) が一定の場合、動きの小さい映像 (映像グループ A において最小の TI を持つ weather report) は平均的な動きを持つ映像コンテンツと比較し、P や B フレームがビット量を必要としないため、I フレームのフレーム平均ビット量 (BI) が高い。動きの大きい映像では上記と逆の特性を持つ。

2) 符号化に対する品質評価特性

パケット損失回数 ($PLEF$) が 0 の場合の映像グループ A の実験 1 について、符号化に対する平均映像品質 (Q_{ave})、最大映像品質 (Q_{max})、最小映像品質 (Q_{min}) の特性を図 5.6 に示す。まず、平均映像品質 (Q_{ave})、最大映像品質 (Q_{max})、最小映像品質 (Q_{min}) を定義する。平均映像品質 (Q_{ave}) は各 HRC に対し 8 映像コンテンツの映像品質評価値 (Q) を平均した値、最大映像品質 (Q_{max}) は各 HRC に対し 8 映像コンテンツの映像品質評価値 (Q) の中の最大値、最小映像品質 (Q_{min}) は各 HRC に対し 8 映像コンテンツの映像品質評価値 (Q) の中の最小値とする。ビットレート (B) はパケットから計算された。図 5.6 の結果は、ビットレート低下が空間品質の低下を引き起こしたことおよび、これらのカーブは映像コンテンツに依存することを示した。図 5.6 が示すように、これらの映像品質のカーブは映像コンテンツに依存するが、映像品質低下の定性的傾向は映像コンテンツに依存しなかった (映像品質主観評価値はビットレートの増加とともに増加し、ある一定値に飽和する)。ビットレートに対する平均映像品質 (Q_{ave})、最大映像品質 (Q_{max})、最小映

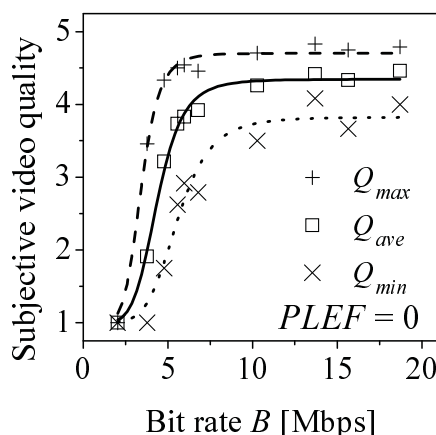


図 5.6 符号化に対する映像品質主観評価特性

像品質 (Q_{min}) はロジスティック関数で表すことができる。ただし、ロジスティック関数の係数はコーデックごとに異なる。

図 5.7 に示すように、ビットレート (B) と最大映像品質差分値 (dQ_{max}) もしくは最小映像品質差分値 (dQ_{min}) の関係を調査した。ここで、最大映像品質差分値 (dQ_{max}) および最小映像品質差分値 (dQ_{min}) を定義する。最大映像品質差分値 (dQ_{max}) は最大映像品質 (Q_{max}) から平均映像品質 (Q_{ave}) を減算した値、最小映像品質差分値 (dQ_{min}) は最小映像品質 (Q_{min}) から平均映像品質 (Q_{ave}) を減算した値である。図 5.7 の結果より、ビットレートに対する最大映像品質差分値 (dQ_{max}) および最小映像品質差分値 (dQ_{min}) は凸型の関数で表すことができること示した。つまり、映像品質差分値 ($dQ = Q - Q_{ave}$) はビットレートに依存した。

5.4.2 (1) 節で述べたように、ビットレート (B) が同一であっても、I フレームのフレーム平均ビット量 (BI) は映像コンテンツに依存する。そこで、I フレームのフレーム平均ビット量 (BI) と映像品質評価値 (Q) の関係を調査した。図 5.8 に示すように、映像コンテンツの I フレームのフレーム平均ビット量 (BI) が少ない場合、映像品質評価値は低かった。逆に、映像コンテンツの I フレームのフレーム平均ビット量 (BI) が多い場合、映像品質評価値は高かった。つまり、I フレームのフレーム平均ビット量 (BI) の変動が映像品質評価値 (Q) に影響し、各特性は一次関数で表せることがわかった。

次に、I フレームのフレーム平均ビット量の差分値 ($dBI = BI - BI_{ave}$) と映像品質差分値 (dQ) の関係を調査した。図 5.9 の線は図 5.8 の $B = 3.4$ Mbps の条件である。図 5.9 の点 X、Y、Z は $X = (BI_{max}, Q_{max})$ 、 $Y = (BI, Q)$ 、 $Z = (BI_{ave}, Q_{ave})$ を示す。図 5.9 に示すように、 $BI > BI_{ave}$ の場合、 dQ は dQ_{max} と $(BI - BI_{ave}) / (BI_{max} - BI_{ave})$ の積に比例した。一方、 $BI \leq BI_{ave}$ の場合、 dQ は dQ_{min} と $(BI - BI_{ave}) / (BI_{min} - BI_{ave})$ の積に比例した。加えて、図 5.8 で示すように、すべての線は線形であるため、 dBI および dQ に対

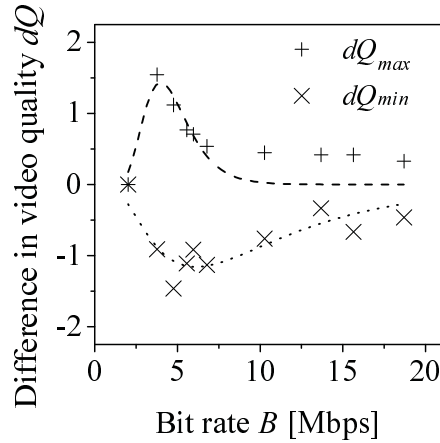


図 5.7 ビットレートおよび映像品質差分値の関係

する特性は同一である。そこで、映像品質差分値 (dQ) を以下の数式 (5-1) から (5-3) で示す一次関数で表した。

$$dQ = \begin{cases} a + b \cdot dQ_{max} \cdot F_{max}, & BI > BI_{ave} \\ a + b \cdot dQ_{min} \cdot F_{min}, & BI \leq BI_{ave} \end{cases} \quad (5-1)$$

$$F_{max} = \frac{BI - BI_{ave}}{BI_{max} - BI_{ave}} \quad (5-2)$$

$$F_{min} = \frac{BI - BI_{ave}}{BI_{min} - BI_{ave}} \quad (5-3)$$

ただし、上記不等式の条件によらず、係数 a および b は一定値とした。

3) パケット損失に対する品質評価特性

次に、パケット損失に対する映像品質差分値 (dQ) を調査した。パケット損失に対する映像品質評価値 (Q) は符号化にも影響される。そこで、映像品質評価値 (Q) に対する符号化の影響を取り除くため、正規化映像品質評価値 NVQ^{35} (N) を導入する。正規化映像品質評価値 (N) はパケット損失低下の程度を示す。各ビットレート (B) に対し、以下のように定義する。

$$N = \frac{Q(B) - 1}{Q(B)|_{PLEF=0} - 1}$$

³⁵Normalized Video Quality

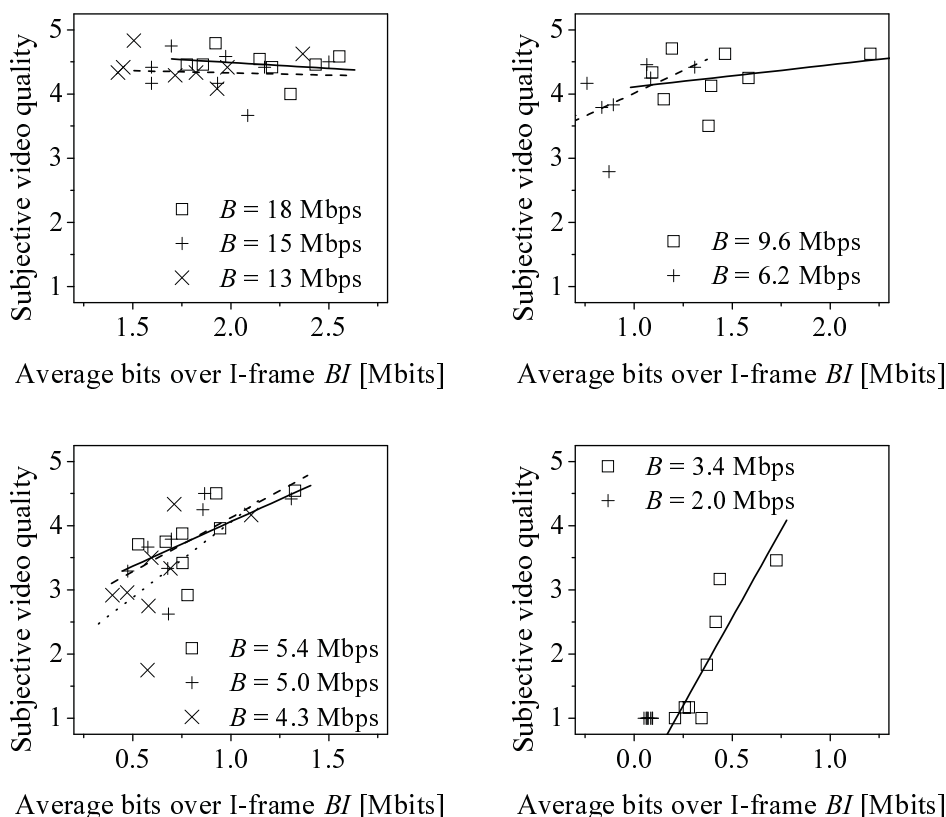


図 5.8 Iフレームのフレーム平均ビット量および映像品質主観評価値の関係

ここで、正規化平均映像品質 (N_{ave})、正規化最大映像品質 (N_{max})、正規化最小映像品質 (N_{min}) を以下のように定義する。正規化平均映像品質 (N_{ave}) は各 HRC に対し、8 映像コンテンツの正規化映像品質評価値 (N) を平均した値、正規化最大映像品質 (N_{max}) は各 HRC に対し、8 映像コンテンツの正規化映像品質評価値 (N) のうち最大の値、正規化最小映像品質 (N_{min}) は各 HRC に対し、8 映像コンテンツの正規化映像品質評価値 (N) のうち最小の値を示す。図 5.10 は、映像グループ A の実験 1 における劣化映像フレーム数 (D) に対する正規化平均映像品質 (N_{ave})、正規化最大映像品質 (N_{max})、正規化最小映像品質 (N_{min}) の特性を示す。

劣化映像フレーム数 (D) に対する正規化平均映像品質 (N_{ave})、正規化最大映像品質 (N_{max})、正規化最小映像品質 (N_{min}) のカーブはそれぞれ異なるが、カーブの定性的傾向は映像コンテンツに依存していない (正規化平均映像品質 (N_{ave})、正規化最大映像品質 (N_{max})、正規化最小映像品質 (N_{min}) は劣化映像フレーム数 (D) の増加とともに低下する)。劣化映像フレーム数 (D) に対する正規化平均映像品質 (N_{ave})、正規化最大映像品質 (N_{max})、正規化最小映像品質 (N_{min}) は指数関数で表すことができ、指数関数の係数はコーデックごとに異なることがわかった。図 5.10 に示すように、ビットレート (B) お

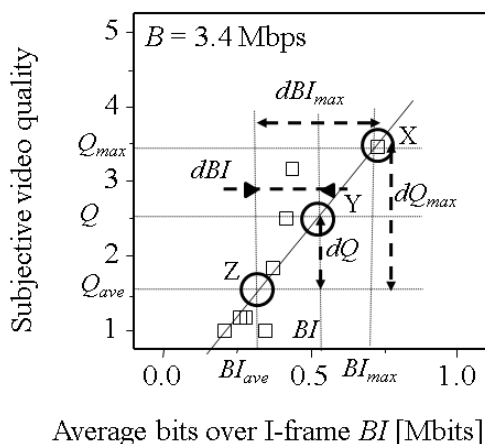


図 5.9 dBI および dQ の関係

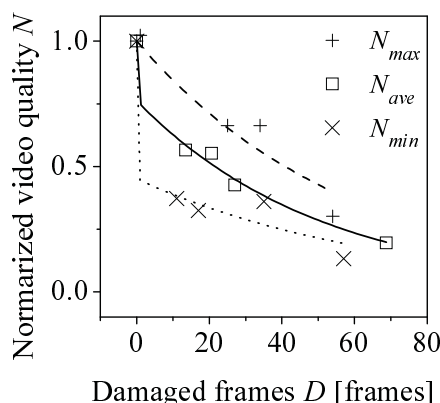


図 5.10 劣化映像フレーム数および正規化平均映像品質の関係

よび平均映像品質 (Q_{ave})、最大映像品質 (Q_{max})、最小映像品質 (Q_{min}) の関係と同様に、映像コンテンツにより特性の傾向が異なった。

そこで、劣化映像フレーム数 (D) と正規化映像品質の差分値 ($dN = N - N_{ave}$) の関係を調査した。ここで、正規化最大映像品質の差分値 (dN_{max}) は正規化最大映像品質 (N_{max}) から正規化平均映像品質 (N_{ave}) を減算した値、正規化最小映像品質の差分値 (dN_{min}) は正規化最小映像品質 (N_{min}) から正規化平均映像品質 (N_{ave}) を減算した値とする。劣化映像フレーム数 (D) および正規化最大映像品質の差分値 (dN_{max}) もしくは正規化最小映像品質の差分値 (dN_{min}) の関係を図 5.11 に示した。この品質評価特性は、ビットレート (B) および符号化に対する映像品質差分値 (dQC) の関係とほぼ同一である。

Iフレームのフレーム平均ビット量 (BI) の変動が符号化に対する映像品質差分値 (dQC) に影響を与えたように、Iフレームのフレーム平均ビット量 (BI) の変動が正規化映像品

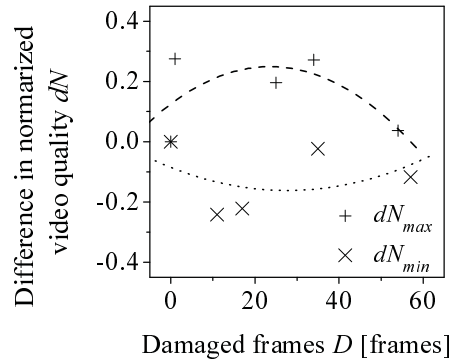


図 5.11 劣化映像フレーム数と正規化平均映像品質の差分値の関係

質の差分値 (dN) に影響を与えると想定される。 $BI > BI_{ave}$ の場合、 dN は dN_{max} と $(BI - BI_{ave}) / (BI_{max} - BI_{ave})$ の積に比例する。一方、 $BI \leq BI_{ave}$ の場合、 dN は dN_{min} と $(BI - BI_{ave}) / (BI_{min} - BI_{ave})$ の積に比例する。そこで、数式 (5-1) 同様に、正規化映像品質差分値 (dN) を以下の数式 (5-4) で示す一次関数で表した。

$$dN = \begin{cases} c + d \cdot dN_{max} \cdot F_{max}, & BI > BI_{ave} \\ c + d \cdot dN_{min} \cdot F_{min}, & BI \leq BI_{ave} \end{cases} \quad (5-4)$$

ただし、上記不等式の条件によらず、係数 c および d は一定値とした。

5.5 提案モデル

5.4.2 節で述べた主観品質評価特性から構築した提案モデルの数式を示す。本節において、 $v_i (i = 1, \dots, 31)$ は事前情報に対し最適化された定数であることに加え、非線形回帰分析により学習データを用いて算出された提案モデルの係数である。

5.5.1 符号化に対する品質推定モジュール

符号化に対する映像品質 (QC) を推定する品質推定モジュールを構築した。 $PLEF = 0$ の場合、パラメータ $Q, Q_{ave}, Q_{max}, Q_{min}, dQ$ は $QC, QC_{ave}, QC_{max}, QC_{min}, dQC$ と表す。

5.4.2(1) 節で述べたように、ビットレート (B) および $BI_{ave}, BI_{max}, BI_{min}$ の関係は以下の数式 (5-5) から (5-7) で表すことができる。

$$BI_{ave} = v_1 + v_2 \exp\left(-\frac{B}{v_3}\right) \quad (5-5)$$

$$BI_{max} = v_4 + v_5 \exp\left(-\frac{B}{v_6}\right) \quad (5-6)$$

$$BI_{min} = v_7 + v_8 \exp\left(-\frac{B}{v_9}\right) \quad (5-7)$$

5.4.2(2)節で述べたように、 $PLEF = 0$ の場合、ビットレート (B) および QC_{ave} 、 QC_{max} 、 QC_{min} の関係は以下の数式 (5-8) から (5-10) で表すことができる。

$$QC_{ave} = 1 + v_{10} - \frac{v_{10}}{1 + (B/v_{11})^{v_{12}}} \quad (5-8)$$

$$QC_{max} = 1 + v_{13} - \frac{v_{13}}{1 + (B/v_{14})^{v_{15}}} \quad (5-9)$$

$$QC_{min} = 1 + v_{16} - \frac{v_{16}}{1 + (B/v_{17})^{v_{18}}} \quad (5-10)$$

5.4.2(3)節で述べたように、映像品質差分値 (dQC) は以下の数式 (5-11) でモデル化できる。

$$dQC = \begin{cases} v_{19} + v_{20} \cdot dQ_{max} \cdot F_{max}, & BI > BI_{ave} \\ v_{19} + v_{20} \cdot dQ_{min} \cdot F_{min}, & BI \leq BI_{ave} \end{cases} \quad (5-11)$$

符号化に対する映像品質 (QC) は以下の数式 (5-12) でモデル化できる。

$$QC = QC_{ave} + dQC \quad (5-12)$$

5.5.2 パケット損失に対する品質推定モジュール

次に、符号化およびパケット損失に対する映像品質 (Q) を推定する品質推定モジュールを構築した。

5.4.2(3)節で述べたように、劣化映像フレーム数 (D) および正規化平均映像品質 (N_{ave})、正規化最大映像品質 (N_{max})、正規化最小映像品質 (N_{min}) の関係は以下の数式 (5-13) から (5-15) でモデル化できる。

$$N_{ave} = (1 - v_{21}) \exp\left(-\frac{D}{v_{22}}\right) + v_{21} \exp\left(-\frac{D}{v_{23}}\right) \quad (5-13)$$

$$N_{max} = (1 - v_{24}) \exp\left(-\frac{D}{v_{25}}\right) + v_{24} \exp\left(-\frac{D}{v_{26}}\right) \quad (5-14)$$

$$N_{min} = (1 - v_{27}) \exp\left(-\frac{D}{v_{28}}\right) + v_{27} \exp\left(-\frac{D}{v_{29}}\right) \quad (5-15)$$

符号化およびパケット損失に対する映像品質評価値 (Q) は符号化に対する映像品質評価値 (QC) および正規化映像品質 (N) によりモデル化できる。

$$\begin{aligned} Q &= 1 + (QC - 1) \cdot N \\ &= 1 + (QC_{ave} + dQC - 1) \cdot (N_{ave} + dN) \\ &= 1 + (QC_{ave} - 1) \cdot N_{ave} + dQC \cdot N_{ave} + (QC_{ave} + dQC - 1) \cdot dN \end{aligned} \quad (5-16)$$

ここで、正規化映像品質 (N) は以下の数式 (5-17) でモデル化した。

$$N = \begin{cases} N_{ave} + dN, & D > 0 \\ 1, & D = 0 \end{cases} \quad (5-17)$$

また、正規化映像品質差分値 (dN) を以下の数式 (5-18) でモデル化した。

$$dN = \begin{cases} v_{30} + v_{31} \cdot dN_{max} \cdot F_{max}, & BI > BI_{ave} \\ v_{30} + v_{31} \cdot dN_{min} \cdot F_{min}, & BI \leq BI_{ave} \end{cases} \quad (5-18)$$

結果、映像品質評価値 (Q) は以下のように変換できる。

$$Q = Q_{ave} + dQ \quad (5-19)$$

ここで、平均映像品質 (Q_{ave}) および差分映像品質 (dQ) は以下のようにモデル化できる。

$$Q_{ave} = 1 + (QC_{ave} - 1) \cdot N_{ave} \quad (5-20)$$

$$dQ = dQC \cdot N_{ave} + (QC - 1) \cdot dN \quad (5-21)$$

5.6 提案モデルのパフォーマンス評価

提案モデルの有効性を検証するため、モデル係数 (v_1-v_{31}) を計算し、非学習データにモデルを適用した。また、本節の最後に考察を述べる。

5.6.1 パフォーマンス要件

本節では、品質推定精度の目標値について述べる。エンドユーザ品質はサービスプロバイダにとって最も重要であるため、品質推定精度の高いFRメディアレイヤモデルの品質推定精度にパケットレイヤモデルの品質推定精度を近づけることは非常に重要である。しかしながら、5.3節で述べたように、パケットレイヤモデルはコーデック種別や実装に関する情報にアクセスしない。加えて、符号化パラメータは暗号化されているため、品質推定に用いることができない。そこで、より高い品質推定精度を達成するため、提案モデルは事前情報を用い学習する必要がある。

HDTV テストプラン [96] によると、VQEG³⁶はモデルの性能を検証するための基準としてRMSE³⁷を採用した。HDに対するFRメディアレイヤモデルであるITU-T 勧告 J.341[25] およびVGA、CIF、QCIFに対するFRメディアレイヤモデルであるITU-T 勧告 J.247[23]の試験結果によると、RMSEの最小値は約0.5³⁸であった。RMSEの値は、異なる解像度、映像コンテンツ、試験条件にも関わらず、ほぼ同等であった。PSNR³⁹がVQEGにおいて品質推定の最低要求条件として用いられたが、PSNRは一般に主観品質との対応が悪いため、PSNRは最低要求条件として適していない。そこで、モデルの品質推定精度が十分であることを検証するために、パフォーマンス要件として、“RMSE ≤ 0.5”を用いた。

映像品質差分値 (dQ) の有効性を検証するため、文献 [43, 44, 45, 52] で提案された従来モデルを統合したモデル、つまり、提案モデルに対し映像品質差分値 (dQ) を用いないモデルを比較モデルとして採用した。比較モデルの品質推定精度は、文献 [43, 44, 45, 52] でそれぞれ提案されたパケットレイヤモデルの品質推定精度より高い。比較モデルは以下の数式 (5-22) である。

$$Q = 1 + (QC_{ave} - 1) \cdot N_{ave}, \quad (5-22)$$

$$QC_{ave} = 1 + v_{10} - \frac{v_{10}}{1 + (B/v_{11})^{v_{12}}},$$

$$N_{ave} = (1 - v_{21}) \exp\left(-\frac{D}{v_{22}}\right) + v_{21} \exp\left(-\frac{D}{v_{23}}\right)$$

ここで、 v_{10} 、 v_{11} 、 v_{12} 、 v_{21} 、 v_{22} 、 v_{23} は提案モデルの係数値と同一である。

5.6.2 提案モデルの品質推定精度

5.3 節で述べたように、定義により、提案モデルはコーデックの実装にアクセスしないため、事前情報に対し最適化が必要である。映像グループ A の実験 1 (製品 P1) および実験 3 (製品 P2) に対し、提案モデルの係数 (v_1-v_{31}) を非線形最小二乗法に基づき計算した。同様に、映像グループ A の実験 1 および 3 に対し比較モデルの係数も計算した。表 5.3 に映像グループ A の実験 1 および 3 に対する係数を示した。以下、上記係数を用い映像品質主観評価値を推定した。

5.3 節で述べたように、映像フレーム種別推定モデルは本検討のスコープ外であるため、本検討では、映像フレーム種別の真値を用いた。結果として、映像フレーム種別の誤判定が映像品質へ与える影響が含まれた。

学習および非学習データに対する映像品質推定値および映像品質主観評価値の関係を図 5.12 および 5.13 に示した。RMSEを表 5.4 に示した。ここでPMは提案モデルを、CMは数式 (5-22) で表される比較モデルを表す。RMSEの改善率も示した。加えて、品質推定

³⁶Video Quality Expert Group

³⁷Root Mean Square Error

³⁸文献 [97] の表 7 および ITU-T 勧告 J.247 の表 1 から 3 に示される RMSE の最小値は、HD に対し 0.550、VGA に対し 0.571、CIF に対し 0.526、QCIF に対し 0.516 であった。

³⁹Peak-Signal-to-Noise Ratio

表 5.3 映像グループ A に対する提案モデルの係数

	Experiment 1 (Product P1)	Experiment 3 (Product P2)		Experiment 1 (Product P1)	Experiment 3 (Product P2)
v_1	2.921	3.024	v_{22}	2.960	4.163
v_2	-3.357	-3.021	v_{23}	52.053	63.376
v_3	12.693	12.323	v_{24}	0.760	0.721
v_4	2.799	2.669	v_{25}	3.979	0.018
v_5	-3.730	-3.643	v_{26}	71.838	58.996
v_6	6.345	3.769	v_{27}	0.750	0.462
v_7	3.400	2.566	v_{28}	0.995	7.031
v_8	-3.734	-2.698	v_{29}	37.740	51.452
v_9	21.894	12.439	v_{30}	-0.027	-0.009
v_{10}	3.346	3.327	v_{31}	0.362	-0.029
v_{11}	4.372	0.585			
v_{12}	5.817	1.188			
v_{13}	3.704	5.336			
v_{14}	3.417	0.013			
v_{15}	6.414	0.111			
v_{16}	2.825	2.779			
v_{17}	5.571	1.096			
v_{18}	5.726	1.795			
v_{19}	0.065	0.015			
v_{20}	0.540	0.144			
v_{21}	0.804	0.587			

精度を詳細に分析するため、相関係数 (R) および外れ値率 (OR⁴⁰) を表 5.5 および 5.6 に示した。非学習データ (映像グループ B の実験 1 および 3、映像グループ A および B の実験 2 および 4) に対する RMSE は学習データ (映像グループ A の実験 1 および 3) の RMSE より小さかった。これは非学習データに対する品質推定が学習データに対する品質推定よりわずかに容易であったと考えられる。

学習データ (映像グループ A の実験 1 および 3) および非学習データ (映像グループ B の実験 1 および 3、映像グループ A および B の実験 2 および 4) に対し、5.6.1 節で述べたパフォーマンス要件を提案モデルの RMSE が満たしたため、提案モデルの品質推定精度は FR メディアレイヤモデルの品質推定精度と同等であったと言える。加えて、比較モデルの相関 (R)、RMSE、外れ値率 (OR) と比較して、提案モデルの相関 (R)、RMSE、外れ値率 (OR) は、ほぼ同程度もしくは良い精度であった。これより、以下のようにまとめることができ、提案モデルはエンドユーザ QoE 監視に適用可能であると結論付けた。

1. 映像グループ A の実験 1 および 3 において高い品質推定精度で各コーデック (製品 P1 および P2) に対し提案モデルを最適化できた。
2. 映像グループ B の実験 1 および 3 において非学習映像に対する提案モデルの有効性を

⁴⁰Outlier Ratio

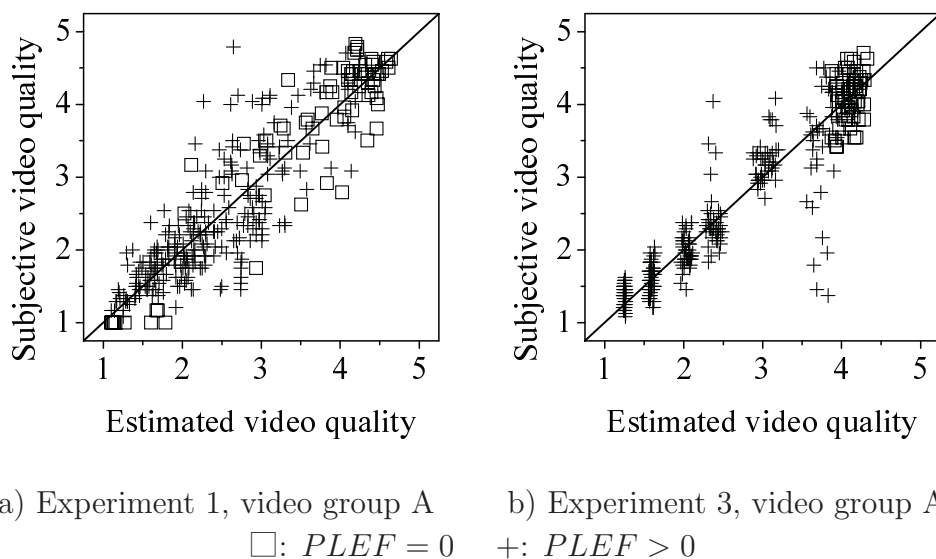


図 5.12 学習データに対する映像品質の推定精度

検証できた。

3. 映像グループ A の実験 2 および 4 において、ビットレート、ランダムおよびバーストパケット損失が異なる非学習 HRC に対し提案モデルは映像品質を精度良く推定することができた。
4. 映像グループ B の実験 2 および 4 において、映像コンテンツ、ビットレートおよびパケット損失が異なる非学習データに対し、提案モデルの有効性を検証した。
5. 提案モデルは 31 の係数を持つが学習および非学習データに対し、品質推定精度の低下はなかった。

表 5.4 学習および非学習データに対する RMSE
 a) 学習データ b) 非学習データ

Exp	VC	VG	CM	PM	IR	Exp	VC	VG	CM	PM	IR
1	Product P1	A	0.53	0.49	8%	1	Product P1	B	0.56	0.43	22%
3	Product P2	A	0.41	0.41	1%	3	Product P2	B	0.38	0.36	5%
Average			0.47	0.45	4%	2	Product P1	A	0.52	0.44	16%
						4	Product P2	A	0.47	0.45	3%
						2	Product P1	B	0.57	0.45	21%
						4	Product P2	B	0.43	0.40	7%
						Average			0.49	0.42	12%

注: Exp は実験、VC は映像コーデック、VG は映像グループ、CM は比較モデル、PM は提案モデル、IR は改善率を表す。

表 5.5 学習および非学習データに対する相関係数
a) 学習データ b) 非学習データ

Exp	VC	VG	CM	PM	Exp	VC	VG	CM	PM
1	Product P1	A	0.87	0.89	1	Product P1	B	0.87	0.91
3	Product P2	A	0.93	0.93	3	Product P2	B	0.94	0.94
Average			0.90	0.91	2	Product P1	A	0.88	0.91
					4	Product P2	A	0.90	0.90
					2	Product P1	B	0.86	0.90
					4	Product P2	B	0.92	0.93
					Average			0.89	0.91

注: 略語は表 5.4 と同じ。

表 5.6 学習および非学習データに対する外れ値率
a) 学習データ b) 非学習データ

Exp	VC	VG	CM	PM	Exp	VC	VG	CM	PM
1	Product P1	A	0.59	0.54	1	Product P1	B	0.57	0.52
3	Product P2	A	0.28	0.29	3	Product P2	B	0.32	0.30
Average			0.44	0.42	2	Product P1	A	0.58	0.54
					4	Product P2	A	0.40	0.38
					2	Product P1	B	0.60	0.51
					4	Product P2	B	0.41	0.41
					Average			0.48	0.44

注: 略語は表 5.4 と同じ。

5.6.3 提案モデルの品質推定精度に対する考察

学習データ（映像グループ A の実験 1 および 3）および非学習データ（映像グループ B の実験 1 および 3、映像グループ A および B の実験 2 および 4）に対し、考察を述べる。

学習データ（映像グループ A の実験 1 および 3）において、図 5.12 a) の中段より下側に品質推定精度の低い点があるが、仮に学習データとして映像グループ A および B の実験 1 および 2 を用いても品質推定精度は劇的に改善されないため、提案モデルは良く学習できていると考えられる。また、学習データに対する提案モデルの品質推定精度は十分なレベルであった。これより、学習データとして SRC および PVS 数は適当であったと考えられる。

学習および非学習データに対し、提案モデルの推定精度はパケット損失条件において品質推定精度が低い点があった。品質推定精度が低い原因は主に二つに分類できる。一つ目は、映像フレームの区切り（PUSI や PTS）が損失した際に発生する映像フレーム種別の誤検出である。二つ目は B フレームのパケットが損失した際に、パケット損失が映像品質に大きな影響を与える場合である。図 5.12 および 5.13 において 45 度の線より下の外れ値は、主に I もしくは P フレームの区切り（PUSI や PTS）が損失した場合か、損失した B フレームが映像品質に大きな影響を与えた場合である。一方、図 5.12 および 5.13 において 45 度の線より上の外れ値は、主に B フレームの区切りが損失した場合である。

まず、映像フレーム種別の誤検出が映像品質評価値に与える影響について述べる。Bフレームの区切りが損失した場合、多くの場合Bフレームの前はIもしくはPフレームであるため、BフレームはIもしくはPフレームとして検出される。IもしくはPフレームの区切りが損失した場合、IもしくはPフレームは多くの場合、Bフレームとして検出される。結果として、モデルは不正確な劣化映像フレーム数 (D) を導出する。このような場合、劣化映像フレーム数の正解値と不正解値の間に差が生じ、提案モデルは映像品質を精度良く推定することができない。しかしながら、映像フレームの区切りに関する情報がパケット損失により欠落した場合、定義によって、提案モデルはビットストリームを用いることができないため、劣化映像フレーム数の誤検出を避けられない。

次に、損失したBフレームが映像品質に大きな影響を与える場合について述べる。Bフレームが損失した際の劣化映像フレーム数 (D) が1であっても、パケット損失が映像品質に大きな影響を与えた。このような場合、提案モデルは映像品質を精度良く推定することができない。このような劣化の影響を考慮するためには、モデルは画素情報を入力として品質を推定する必要がある。ただし、劣化映像フレーム数の誤検出の発生頻度はBフレームが映像品質に大きな影響を与える頻度より高い。上記二つの品質推定精度が低い原因については、比較モデルにおいても同様の理由で映像品質を精度良く推定することができない。

図5.12b) および5.13b) において垂直方向に線上のプロットがある。この傾向は以下の理由がある。実験3においてはバーストパケット損失長が1であったが、実験1、2、4においてはバーストパケット損失長は多くの種類が用いられた。この結果、実験3ではGoPに対し一つのパケットが損失し、実験1、2、4では一つもしくは複数のパケットが複数のGoPに損失した。加えて、一般に、Iフレームのビット量はPやBフレームと比較して多いため、Iフレームが損失する確率は高い。このような場合、例えば、PLEFが1、2、4、7の時、Iフレームが1、2、4、7枚損失する (GoPサイズが15の場合、劣化映像フレーム数は17、34、68、119となる)。つまり、垂直方向の線上のプロットは同じ劣化映像フレーム数をもっているため、これらの点は垂直方向に線上となった。

実験1から4において二つのコーデックに対する提案モデルの品質推定精度の改善率について述べる。5.4節で述べたように、二つの実装の異なるH.264エンコーダ (製品P1、製品P2) を用いた。実験3および4で用いたH.264コーデック (製品P2) に対する提案モデルの品質推定精度の改善率は実験1および2で用いたH.264コーデック (製品P1) に対する提案モデルの品質推定精度の改善率より低かった。H.264コーデック (製品P2) に対する品質推定においてはビットレートおよび劣化映像フレーム数を用いることで十分であったが、H.264コーデック (製品P1) においてはIフレームのフレーム平均ビット量を用いることで品質推定精度を効率良く改善した。

提案モデルが映像コンテンツごとの映像品質を精度良く推定できたかを考察する。図5.14は製品P1およびP2に対し、提案モデルの品質推定精度が改善したことを示している。特に、実験2および3においてSRC12に対する提案モデルのRMSEは0.56および0.53であったが、比較モデルのRMSEは0.64および0.55であった。この傾向は他の映像コンテンツ、劣化映像フレーム数の誤検出、Bフレームの損失が映像品質に大きな影響を与える場合に

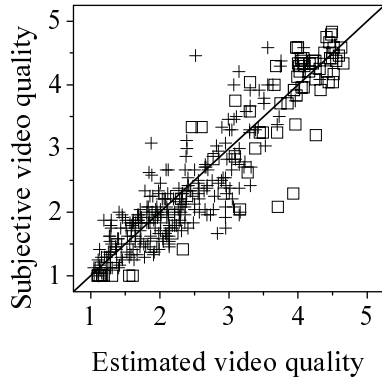
においても変わるものではない。これより、提案モデルは映像品質差分値を用いることで映像コンテンツごとの映像品質を精度良く推定できることがわかった。

5.7 まとめ

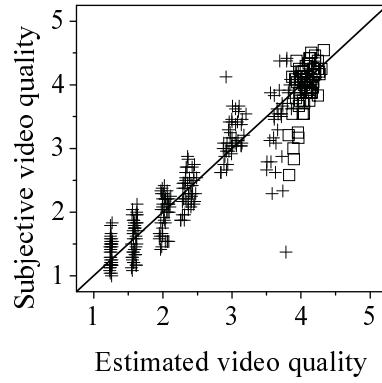
映像コンテンツが映像品質に与える影響を考慮するために、推定対象の映像コンテンツの品質評価値と平均映像品質評価値の差分値を計算する新しいモデルを提案した。まず初めに、映像コンテンツが映像品質に与える影響に関して従来モデルの問題点を指摘した。符号化およびパケット損失が映像品質に与える影響を調査するため、主観品質評価実験を実施した。この結果、平均映像品質 (Q_{ave}) はビットレート (B) および劣化映像フレーム数 (D) で表すことができ、映像コンテンツごとの映像品質 (Q) はIフレームのフレーム平均ビット量 (BI) に依存することがわかった。これらの特性を考慮するため、推定対象の映像品質評価値と平均映像品質評価値 (Q_{ave}) の差分値である映像品質差分値 (dQ) の概念を導入した。

非学習映像およびビットレートおよびパケット損失が異なる非学習HRCに対し提案モデルの有効性を検証するため、平均映像品質評価値 (Q_{ave}) を推定する比較モデルの品質推定精度と提案モデルの品質推定精度を比較した。実験結果より、提案モデルは映像コンテンツごとの映像品質を推定でき、モデル係数を変更することで実装の異なる映像コーデックの映像品質を推定できた。また、提案モデルの品質推定精度は比較モデルの品質推定精度より高いことを示した。しかしながら、映像フレームの区切り (PUSI、PTS) が損失した場合やBフレームがパケット損失を持ち映像品質に大きな影響を与える場合、提案モデルはビットストリームや画素情報を用いないため、品質推定精度が低下した。

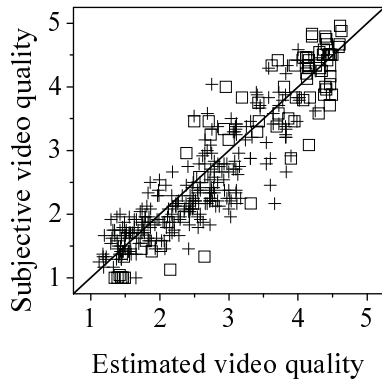
以下に今後の課題を示す。本検討では映像フレーム種別推定モデルを導入していないため、提案モデルに映像フレーム種別推定モデルを導入し、そのモデルの有効性を検証する必要がある。



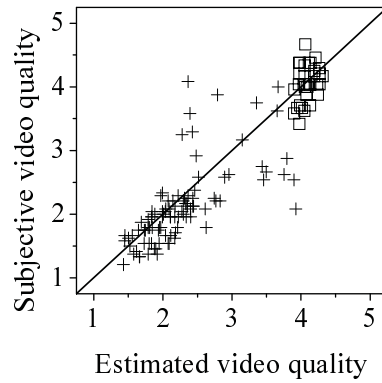
a) Experiment 1, video group B



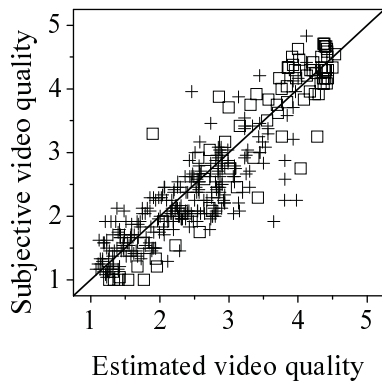
b) Experiment 3, video group B



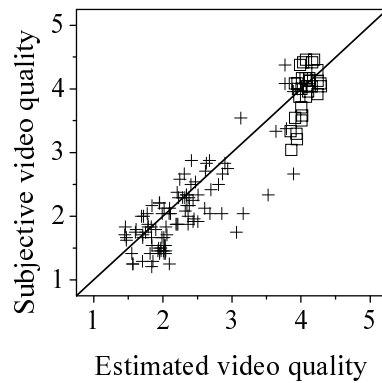
c) Experiment 2, video group A



d) Experiment 4, video group A



e) Experiment 2, video group B



f) Experiment 4, video group B

□: $PLEF = 0$ +: $PLEF > 0$

図 5.13 非学習データに対する映像品質の推定精度

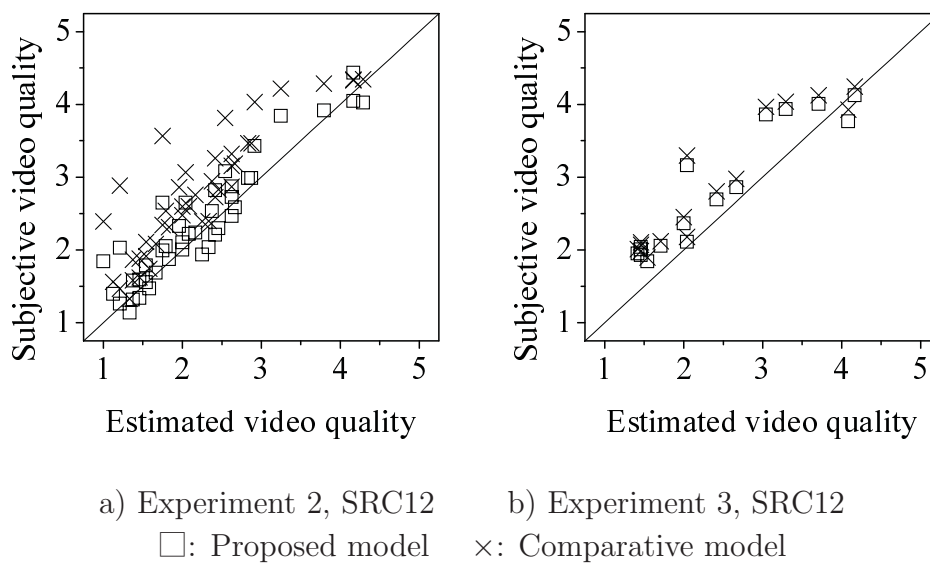


図 5.14 SRC12 の映像品質推定精度

表 5.7 略語

Abbreviation	Description
ABI	Average bits over I-frame
ACR	Absolute category rating
ARQ	Automatic repeat request
BR	Bit rate
CM	Comparative model
Codec	Encoder and decoder
DF	Number of video frames damaged by packet loss
FR	Full reference
GoP	Group-of-picture
H/E	Head-end
HDTV	High-definition television
IPTV	Internet protocol television
IR	Improvement rate
Mbps	Mbits/sec
MOS	Mean opinion score
MRFM	Multiple-reference-frames mode
NAMS	Non-intrusive parametric model for the assessment of the performance of multimedia streaming
NLSA	Nonlinear least-squares approximation
NR	No reference
NVQ	Normalized video quality
OR	Outlier ratio
PES	Packetized elementary stream
PID	Packet identifier
PLC	Packet-loss concealment
PLEF	Packet-loss-event frequency
PM	Proposed model
PMT	Program map table
PSNR	Peak-signal-to-noise ratio
PTS	Presentation time stamp
PUSI	Payload-unit-start indicator
PVS	Processed video sequence
QoE	Quality of experience
QoS	Quality of service
R	Correlation
RMSE	Root mean square error
RR	Reduced reference
RTP	Real-time transport protocol
SI	Spatial information
SRC	Source video sequence
TCP	Transmission control protocol
TI	Temporal information
TS	Transport stream
UDP	User datagram protocol
VC	Video codec
VG	Video group
VQ	Video quality
VQEG	Video Quality Experts Group

第6章 結論

本論文では、ユーザに高い品質の映像通信サービスを提供することを目的に、品質設計および品質監視モデルの研究を行った。

第1章で本研究の背景および課題をまとめた。

第2章テレビ電話サービスに対する品質設計法では、1対1のテレビ電話サービスの映像品質設計を目的に、映像ビットレート、フレームレート、パケット損失率が映像品質に与える影響を主観品質評価特性から導出し、それら品質評価特性をモデル化した映像品質設計モデルを提案した。まず、映像品質はビットレートに対し最適なフレームレートを持つことを主観品質評価特性より解明し、符号化に対する映像品質を推定するモデルを確立した。次に、パケット損失発生時の映像品質は、パケット損失の増加に伴い低下するのみだけではなく、ビットレートやフレームレートにも依存することを解明し、映像品質設計モデルを確立した。提案映像品質設計モデルにより推定された映像品質と主観品質評価値の関係から品質推定誤差が主観品質評価値の統計的信頼幅と同等のレベルになることを示した。これより、テレビ電話サービスの映像品質を適切に設計できることを示した。なお、提案映像品質設計モデルの有効性がITU-Tに認められ、ITU-T勧告G.1070の映像品質推定モデルとして標準化された。

第3章3D映像サービスに対するヘッドエンド品質監視法では、3D映像サービスのヘッドエンド品質監視を目的に、3D映像を構成する左右眼の2D映像品質から3D映像品質を推定するモデルを提案した。現行の3D映像サービスの方式の違いが3D映像品質に与える影響について述べ、その方式の違いによる3D映像品質の主観評価特性を導出した。次に、3D映像品質を左右眼の2D映像品質から推定可能なモデルを主観品質評価特性に基づき構築した。具体的には、左右眼の2D映像のうち高い映像品質と、左右眼の2D映像品質の差分値により3D映像品質を推定するモデルを構築した。左右眼の2D映像品質の平均値を用いて3D映像品質を推定する従来技術においては、左右眼の2D映像品質差が大きくなると品質推定精度が低下するが、提案技術は左右眼の2D映像品質差の大小にかかわらず、高い精度で3D映像品質を推定できることを示した。

第4章パケット損失パターンを考慮したIPTVエンドユーザ品質監視法では、パケットのヘッダ内に映像フレームの区切りを示すフラグが存在しない場合のIPTVサービスのエンドユーザ品質監視を目的に、パケットのヘッダ情報からビットレートおよびパケット損失回数を導出し、これらパラメータから映像品質を推定するモデルを提案した。まず、ビットレート、パケット損失率、バーストパケット損失長を変化させた映像に対し主観品質評価実験を実施した。その結果、ビットレートにより符号化に対する品質を推定可能であることを示し、次に、パケット損失率のみでは映像品質を推定できないことを示した。パケット損失率に対する映像品質はバーストパケット損失長に依存することを解明し、パケット損失回数により映像品質を推定できることを明らかにした。これら品質評価特性に基づき、映像品質をビットレートとパケット損失回数に基づき推定する映像品質監視モデルを構築し、主観品質評価値を実用上十分な精度で推定できることを示した。なお、提案映像品質

監視モデルの一部は、ITU-T 勧告 P.1201.1 の技術として標準化された。

第5章映像フレームを用いたIPTVエンドユーザ品質監視法では、映像フレームの区切りを示すフラグが存在する場合のIPTVサービスのエンドユーザ品質監視を目的に、パケットのヘッダ情報からビットレート、Iフレームのフレーム平均ビット量、劣化映像フレーム数を導出し、これらパラメータから映像品質を推定するモデルを提案した。まず、ビットレート、Iフレームのフレーム平均ビット量、劣化映像フレームに対する主観品質評価特性を導出し、映像コンテンツが映像品質に与える影響を加味するためには、Iフレームのフレーム平均ビット量を考慮することが効果的であることを示した。次に、ビットレートと劣化映像フレーム数のみに着目した従来モデルと提案モデルの品質推定精度を比較し、品質推定精度を大幅に改善することを示した。

本論文では、テレビ電話、3D映像、IPTVサービスを対象に品質設計モデル、ヘッドエンド品質監視モデル、エンドユーザ品質監視モデルを提案し、その有効性を示した。映像通信サービスは今後、HDから4k/8kへと高品質化が進む一方、モバイル端末向けの映像通信サービスが普及していくことが想定されている。このように、今後は、ネットワークやユーザ端末に適した品質の映像通信サービスを提供していくことが求められ、品質を可視化するだけでなく、品質を適切に制御できる技術が求められると予想される。ただし、主観品質が最大となるように品質を制御するためには、品質の可視化は必須であり、より品質推定技術が求められる時代になっていく可能性があることを述べ本論文を結ぶ。

謝辞

本研究を進めるにあたり、御多忙の中、貴重な時間を割いて熱心な御指導、御意見を賜りました方々に深く感謝致します。

NTT ネットワーク基盤技術研究所・林孝典グループリーダーには、第2から5章の研究をまとめる際に論文執筆に関し、貴重なご助言を頂いたことを感謝致します。

NTT ネットワーク基盤技術研究所・高橋玲グループリーダーのご指導により、本研究成果の一部がITU-Tの勧告として採択されたことに感謝致します。

第3章の研究を立ち上げる際に訪問研究員として受け入れてくれた Arizon State University に感謝するとともに、貴重なご助言を頂いた Prof. Lina Karam に感謝致します。

早稲田大学大学院基幹理工学研究科・甲藤二郎教授、後藤滋樹教授、同大大学院国際情報通信研究科・渡辺裕教授には、学位取得に関しご尽力を賜るとともに、本論文作成において多岐に渡りご指導を頂いたことに感謝致します。

最後に、学位取得を影で支えてくれた父、母、妻、二人の娘に感謝致します。

参考文献

- [1] ITU-T Recommendation H.262, “Information Technology – Generic Coding of Moving Pictures and Associated Audio Information: Video,” Feb. 2000.
- [2] ITU-T Recommendation H.264, “Advanced Video Coding for Generic Audiovisual Services,” Mar. 2010.
- [3] ITU-T Recommendation P.910, “Subjective Video Quality Assessment Methods for Multimedia Applications,” Apr. 2008.
- [4] ITU-R Recommendation BT.500–13, “Methodology for the Subjective Assessment of the Quality of Television Pictures,” Jan. 2012.
- [5] ITU-R Recommendation BT.1788, “Methodology for the Subjective Assessment of Video Quality in Multimedia Applications,” Jan. 2007.
- [6] ITU-T Recommendation G.107, “The E-model, A Computational Model for Use in Transmission Planning,” Nov. 2011.
- [7] ITU-T Recommendation G.107.1, “Wideband E-model,” Nov. 2011.
- [8] A. Takahashi, “Opinion Model for Estimating Conversational Quality of VoIP,” Proc. IEEE ICASSP’ 04, vol III, pp. 1072-1075, May 2004.
- [9] A. Takahashi, A. Kurashima, and H. Yoshino, “Objective Assessment Methodology for Estimating Conversational Quality in VoIP,” IEEE Transaction on Speech Audio and Language Processing, vol. 14, no. 6, pp. 1983-1993, Nov. 2006.
- [10] K. Yamagishi and T. Hayashi, “Analysis of Psychological Factors for Quality Assessment of Interactive Multimodal Service,” Electronic Imaging 2005, pp. 130-138, Jan. 2005
- [11] K. Yamagishi and T. Hayashi, “Opinion Model using Psychological Factors for Interactive Multimodal Services,” IEICE Transaction on Communications, vol. E89-B, no. 2, pp. 281-288, Feb. 2006.
- [12] K. Yamagishi and T. Hayashi, “Verification of Video Quality Opinion Model for Videophone Services,” 2nd ISCA/DEGA Tutorial & Research Workshop on Perceptual Quality of Systems, pp. 143 - 148, Sept. 2006.
- [13] K. Yamagishi and T. Hayashi, “Opinion Model for Estimating Video Quality of Videophone Services,” IEEE GLOBECOM Telecommunications Conference 2006, QRP-1, Nov. 2006.

- [14] N. Kitawaki, Y. Arayama, and T. Yamada, "Multimedia Opinion Model based on Media Interaction of Audio-visual Communications," MESAQIN2005, pp. 5-10, June 2005.
- [15] ITU-T Recommendation P.862, "Perceptual Evaluation of Speech Quality (PESQ): An Objective Method for End-to-end Speech Quality Assessment of Narrowband Telephone Networks and Speech Codecs," Feb. 2001.
- [16] ITU-R Recommendation BS.1387, "Method for Objective Measurements of Perceived Audio Quality," Nov. 2001.
- [17] ITU-T Recommendation P.863, "Perceptual Objective Listening Quality Assessment," Jan. 2011.
- [18] M. Miyahara, K. Kotani, V. R. Algazi, "Objective picture Quality Scale (PQS) for Image Quality," IEEE Transaction on Communications, vol. 46, no. 9, pp. 1215-1226, Sept. 1998.
- [19] Z. Wang, A. C. Bovik, H. R. Scheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," IEEE Transaction on Image Processing, vol. 13, no. 4, pp. 600-612, Apr. 2004.
- [20] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale Structural Similarity for Image Quality Assessment," Proc. IEEE Asilomar Conference on Signals, Systems and Computers, Nov. 2003.
- [21] VQEG. Video Quality Experts Group website, <http://www.vqeg.org/>.
- [22] J. Okamoto, T. Hayashi, A. Takahashi, and T. Kurita, "Verification of Objective Quality Assessment Method for Arbitrary Video Sequence," IMQA, Sept. 2005.
- [23] ITU-T Recommendation J.247, "Objective Perceptual Multimedia Video Quality Measurement in the Presence of a Full Reference," Aug. 2008.
- [24] ITU-T Rec. J.144, "Objective Perceptual Video Quality Measurement Techniques for Digital Cable Television in the Presence of a Full Reference," Mar. 2004.
- [25] ITU-T Recommendation J.341, "Objective Perceptual Multimedia Video Quality Measurement of HDTV for Digital Cable Television in the Presence of a Full Reference," Jan. 2011.
- [26] A. Ribman, V. Vaishampayan, and Y. Sermadevi, "Quality Monitoring of Video Quality Over a Packet Network," IEEE Transactions on Multimedia, vol. 6, no. 2, pp. 327-334, Apr. 2004.

-
- [27] H. Pinson and S. Wolf, "A New Standardized Method for Objectively Measuring Video Quality," *IEEE Transactions on Broadcasting*, vol. 50, no. 3, pp. 312–322, Sept. 2004.
- [28] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A Statistical Evaluation of Recent Full Reference Image Quality Assessment Algorithms," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3441–3452, Nov. 2006.
- [29] F. Yang, S. Wan, Y. Chang, and H.R. Wu, "A Novel Objective No-Reference Metric for Digital Video Quality Assessment," *IEEE Transactions on Signal Processing Lett.*, vol. 12, no. 10, pp. 685–688, Oct. 2005.
- [30] P.L. Callet, C. Viard-Gaudin, and D. Barba, "A Convolutional Neural Network Approach for Objective Video Quality Assessment," *IEEE Transactions on Neural Networks*, vol. 17, no. 5, pp. 1316–1327, Sept. 2006.
- [31] J. Okamoto, K. Watanabe, A. Honda, M. Uchida, and S. Hangai, "HDTV Objective Video Quality Assessment Method Applying Fuzzy Measure," *International Workshop on Quality of Multimedia Experience (QoMEX 2009)*, pp. 168–173, July 2009.
- [32] P. L. Callet, C. Viard-Gaudin, S. Pechard, and E. Caillault, "No Reference and Reduced Reference Video Quality Metrics for End to End QoS Monitoring," *IEICE Transactions on Communications*, vol. 89, no. 2, pp. 289–296, Feb. 2006.
- [33] M. Farias and S. Mitra, "No-Reference Video Quality Metric Based on Artifact Measurements," *IEEE ICIP 2005*, pp. 141–144, Sept. 2005.
- [34] ITU-T Recommendation J.246, "Perceptual Audiovisual Quality Measurement Techniques for Multimedia Services over Digital Cable Television Networks in the Presence of a Reduced Bandwidth Reference," Aug. 2008.
- [35] T. Yamada, Y. Miyamoto, Y. Senda, and M. Serizawa, "Video-quality Estimation Based on Reduced-reference Model Employing Activity-Difference," *IEICE Transactions on Fundamentals*, vol. E92-A, no. 12, pp. 3284–3290, Dec. 2009.
- [36] ITU-T Recommendation J.249, "Perceptual Video Quality Measurement Techniques for Digital Cable Television in the Presence of a Reduced Reference," Jan. 2010.
- [37] ITU-T Recommendation J.342, "Objective Multimedia Video Quality Measurement of HDTV for Digital Cable Television in the Presence of a Reduced Reference Signal," Apr. 2011.
- [38] A. Clark, "Modeling the Effects of Burst Packet Loss and Recency on Subjective Voice Quality," *IP Telephony Workshop*, 2001.

- [39] M. Masuda and T. Hayashi, "Non-intrusive Quality Monitoring Method of VoIP Speech Based on Network Performance Metrics," *IEICE Transaction on Communications*, vol. E89-B, no. 2, pp. 304-312, Feb. 2006.
- [40] ITU-T Recommendation P.564, "Conformance Testing for Narrowband Voice over IP Transmission Quality Assessment Models," July 2006.
- [41] IETF RFC 4445, "A Proposed Media Delivery Index (MDI)," Apr. 2006.
- [42] K. Yamagishi and T. Hayashi, "Parametric Packet-Layer Model for Monitoring Video Quality of IPTV Services," *IEEE ICC 2008*, pp. 110-114, May 2008.
- [43] J. Gustafsson, G. Heikkila, and M. Pettersson, "Measuring Multimedia Quality in Mobile Networks with An Objective Parametric Model," *IEEE ICIP 2008*, pp. 405-408, Oct. 2008.
- [44] A. Raake, M. Garcia, J. Berger, F. Kling, P. List, J. Johann, and C. Heidemann, "T-V-Model: Parameter-Based Prediction of IPTV Quality," *IEEE ICASSP 2008*, pp. 1149-1152, Mar. 2008.
- [45] K. Yamagishi and T. Hayashi, "Non-intrusive Packet-layer Model for Monitoring Video Quality of IPTV Services," *IEICE Transactions on Fundamentals*, vol. E92-A, no. 12, pp. 3297-3306, Dec. 2009.
- [46] F. You, W. Zhang, and J. Xiao, "Packet Loss Pattern and Parametric Video Quality Model for IPTV," *IEEE ICIS 2009*, pp. 824-828, June 2009.
- [47] O. Verscheurei and X. Garcia, "User-Oriented QoS in Packet Video Delivery," *IEEE Network*, vol. 12, no. 6, pp. 12-21, Nov. 1998.
- [48] S. Mohamed and G. Rubino, "A Study of Real-Time Packet Video Quality Using Random Neural Networks," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 12, pp. 1071-1083, Dec. 2002.
- [49] K. Watanabe, K. Yamagishi, J. Okamoto, and A. Takahashi, "Proposal of New QoE Assessment Approach for Quality Management of IPTV Services," *IEEE ICIP 2008*, pp. 2060-2063, Oct. 2008.
- [50] K. Jung, S. Lee, and D. Sim, "Perceptual Quality Assessment Method Based on the Parameter Extraction from H.264/AVC Bitstream," *Proc. of International Workshop on Image Media Quality and its Applications*, pp. 61-66, Sept. 2008.

- [51] A. Silva, P. Rodriguez-Bocca, and G. Rubino, "Optimal Quality-of-Experience Design for a P2P Multi-source Video Streaming," *IEEE ICC 2008*, pp. 22–26, May 2008.
- [52] M. Masuda, K. Ushiki, T. Tominaga, T. Hayashi, A. Takahashi, and K. Kawashima, "End-User QoE Estimation for Video Communication Services by Packet-Layer (in Japanese)," *IEICE Transaction on Communications*, vol. J94-B, no. 1, pp. 24–35, Jan. 2011.
- [53] D. Hands, "Quality Assurance for IPTV," *ITU-T Workshop on End-to-end QoE/QoS*, June 2006.
- [54] O. Sugimoto, S. Naito, S. Sakazawa, and A. Koike, "Objective Perceptual Video Quality Measurement Method Based on Hybrid No Reference Framework," *IEEE ICIP 2009*, pp. 2237–2240, Oct. 2009.
- [55] K. Yamagishi, T. Kawano, and T. Hayashi, "Hybrid Video-Quality-Estimation Model for IPTV Services," *IEEE Globecom 2009*, Nov. 2009.
- [56] A. Khah, L. Sun, and E. Ifeachor, "Content Clustering Based Video Quality Prediction Model for MPEG4 Video Streaming over Wireless Networks," *IEEE ICC 2009*, pp. 1–5, June 2009.
- [57] M. Ries, C. Crespi, O. Nemethovaand, and M. Rupp, "Content Based Video Quality Estimation for H.264/AVC Video Streaming," *IEEE WCNC 2007*, pp. 2668–2673, Mar. 2007.
- [58] ITU-T Recommendation G.100/P.10 Amendment 1, "New Appendix I – Definition of Quality of Experience (QoE), Jan. 2007.
- [59] A. Takahashi, H. Yoshino, and N. Kitawaki, "Perceptual QoS Assessment Technologies for VoIP," *IEEE Communication Magazine*, vol. 42, no. 7, pp. 28-34, July 2004.
- [60] T. Hayashi, K. Yamagishi, and H. Yoshino, "Perceptual QoS Evaluation Model for Audiovisual Communication Services," *World Telecommunications Congress 2006 (WTC' 06)*, May 2006.
- [61] T. Hayashi, K. Yamagishi, T. Tominaga, and A. Takahashi, "Multimedia Quality Integration Function for Videophone Services," *IEEE Global Telecommunications Conference 2007*, Nov. 2007.
- [62] D. C. Montgomery, E. A. Peck, G. G. Vining, "Introduction to Linear Regression Analysis," *Wiley-Interscience*, 2004.

- [63] DVB Document A154, “Digital Video Broadcasting (DVB); Frame Compatible Plano-Stereoscopic 3DTV (DVB-3DTV),” Feb. 2011.
- [64] W. J. Tam and L. B. Stelmach, “Psychovisual Aspects of Viewing Stereoscopic Video Sequences,” *Proceedings of the SPIE*, vol. 3295, pp. 226–235, Jan. 1998.
- [65] G. Leon, H. Kalva, and B. Furht, “3D Video Quality Evaluation with Depth Quality Variations,” *3DTV-CON*, pp. 301–304, May 2008.
- [66] M. T.M. Lambooi, W. A. IJsselsteijn, and I. Heynderickx, “Visual Discomfort in Stereoscopic Displays: A Review,” *Proceedings of the SPIE*, vol. 6490, pp. 64900I-1–13, Jan. 2007.
- [67] K. Ukai, “Human Factors for Stereoscopic Images,” *IEEE ICME*, pp. 1697–1700, July 2006.
- [68] H. Yamanoue, M. Okui, and F. Okano, “Geometrical Analysis of Puppet-theater and Cardboard Effects in Stereoscopic HDTV Images,” *IEEE Transaction on Circuits and Systems for Video Technology*, vol. 16, no. 6, pp. 744–752, June 2006.
- [69] B. F. Col and K. O’Connell, “3DTV at Home: Status, Challenges and Solutions for Delivering a High Quality Experience,” *International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM)*, Jan. 2010.
- [70] Q. Huynh-Thu, P. Le Callet, and M. Barkowsky, “Video Quality Assessment: From 2D to 3D - Challenges and Future Trends,” *International Conference on Image Processing (ICIP)*, pp. 4025–4028, Sept. 2010.
- [71] P. Merkle, K. Muller, and T. Wiegand, “3D Video: Acquisition, Coding, and Display,” *IEEE Transaction on Consumer Electronics*, vol. 56, no. 2, pp. 946–950, Jan. 2010.
- [72] K. Kim, J. Lee, D. Suh, and G. Park, “Efficient Stereoscopic Contents File Format on the Basis of ISO Base Media File Format,” *Proceedings of the SPIE*, vol. 7256, pp. 72560N-1–9, Jan. 2009.
- [73] K. Yamagishi, L. Karam, J. Okamoto, and T. Hayashi, “Subjective Characteristics for Stereoscopic High Definition Video,” *International Workshop on Quality of Multimedia Experience (QoMEX)*, pp. 37–42, Sept. 2011.
- [74] L. Stelmach, W. Tam, D. Meegan, and A. Vincent, “Stereo Image Quality: Effects of Mixed Spatio-temporal Resolution,” *IEEE Transaction on Circuits and Systems for Video Technology*, vol. 10, no. 2, pp. 188–193, Mar. 2000.

- [75] P. Aflaki, M. M. Hannuksela, J. Hakkinen, P. Lindroos, and M. Gabbouj, "Subjective Study on Compressed Asymmetric Stereoscopic Video," International Conference on Image Processing (ICIP), pp. 4021–4024, Sept. 2010.
- [76] G. Saygili, C. G. Gurler, and A. M. Tekalp, "Evaluation of Asymmetric Stereo Video Coding and Rate Scaling for Adaptive 3D Video Streaming," IEEE Transaction on Broadcasting, vol. 57, no. 2, pp. 593–601, June 2011.
- [77] G. Saygili, C. G. Gurler, and A. M. Tekalp, "Quality Assessment of Asymmetric Stereo Video Coding," Int. Conf. on Image Processing (ICIP), pp. 4009–4012, Sept. 2010.
- [78] A. Vetro, A. M. Tourapis, K. Müller, and T. Chen, "3D-TV Content Storage and Transmission," IEEE Transaction on Broadcasting, vol. 57, no. 2, pp. 384–394, June 2011.
- [79] V. D. Silva, H. K. Arachchi, E. Ekmekcioglu, A. Fernando, S. Dogan, A. Kondo, and S. Savas, "Psycho-physical Limits of Interocular Blur Suppression and its Application to Asymmetric Stereoscopic Video Delivery," IEEE 19th Int. Packet Video Workshop (PV2012), pp. 184–189, May 2012.
- [80] P. Campisi, P. Le Callet, and E. Marini, "Stereoscopic Images Quality Assessment," Proceedings of 15th European Signal Processing Conference (EURASIP), pp. 2110–2114, Sept., 2007.
- [81] ITU-R Recommendation BT.1438, "Subjective Assessment of Stereoscopic Television pictures," Mar. 2000.
- [82] M. Brotherton, Q. Huynh-thu, D. Hands, and K. Brunnström, "Subjective Multimedia Quality Assessment," IEICE Transaction on Fundamentals, vol. 89-A, no. 11, pp. 2920–2932, Nov. 2006.
- [83] ITU-T Recommendation G.1081, "Performance Monitoring Points for IPTV," Oct. 2008.
- [84] A. Takahashi, D. Hands, and V. Barriac, "Standardization Activities in ITU on QoE Assessment of IPTV," IEEE Communication Magazine, vol. 46, no. 2, pp. 78–84, Feb. 2008.
- [85] H. Yang and K. Rose, "Advances in Recursive Per-Pixel End-to-End Distortion Estimation for Robust Video Coding in H.264/AVC," IEEE Transactions on Circuits and Systems for Video Technology, vol. 17, no. 7, pp. 845–856, July 2007.

-
- [86] ITU-R Recommendation BT.1210-3, "Test Materials to be used in Subjective Assessment," Feb. 2004.
- [87] T. Hayashi, M. Masuda, T. Tominaga, and K. Yamagishi, "Non-intrusive QoS Monitoring Method for Realtime Telecommunication Services," NTT Technical Review, vol. 4, no. 4, pp. 35-40, Apr. 2006.
- [88] IPTV Forum Japan IPTVFJ STD0001, "Overview," Sept. 2009.
- [89] IPTV Forum Japan IPTVFJ STD0004, "IP Broadcasting Specifications," July 2010.
- [90] DVB Document A001, "Specification for the Use of Video and Audio Coding in Broadcasting Applications based on the MPEG-2 Transport Stream," Feb. 2007.
- [91] ITU-T Recommendation H.720, "Overview of IPTV Terminal Devices and End Systems," Oct. 2008.
- [92] ITU-T Recommendation H.721, "IPTV Terminal Devices: Basic Model," Mar. 2009.
- [93] ITU-T Recommendation H.222.0, "Information Technology – Generic Coding of Moving Pictures and Associated Audio Information: Systems," May 2006.
- [94] K. Ushiki, M. Masuda, T. Hayashi, and A. Takahashi, "Packet-layer Video-quality-estimation Method by using Frame-type Estimation (in Japanese)," IEICE Transaction on Communications, vol. J94-B, no. 1, pp. 36-48, Jan. 2011.
- [95] D. Hands and M. Wilkins, "A Study of the Impact of Network Loss and Burst Size on Video Streaming Quality and Acceptability," Proc. of the 6th International Workshop on Interactive Distributed Multimedia Systems and Telecommunication Services, vol. 1718, pp. 45-57, Oct. 1999.
- [96] VQEG, "Test Plan for Evaluation of Video Quality Models for Use with High Definition TV Content v3.0," 2009.
- [97] VQEG, "Report on the Validation of Video Quality Models for High Definition Video Content v2.0," 2010.

研究業績

種 類 別	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者（申請者含む）
○論文	1. Subjective Quality Metric for 3D Video Services、 IEICE Transactions on Communications、 Feb. 2013、 <u>K. Yamagishi</u> , T. Kawano, T. Hayashi, and J. Katto
○論文	2. No Reference Video-Quality-Assessment Model for Monitoring Video Quality of IPTV Services、 IEICE Transactions on Communications、 Feb. 2012、 <u>K. Yamagishi</u> , J. Okamoto, T. Hayashi, and A. Takahashi
○論文	3. Non-intrusive packet-layer model for monitoring video quality of IPTV services、 IEICE Transactions on Fundamentals、 Dec. 2009、 <u>K. Yamagishi</u> and T. Hayashi
○論文	4. Video-Quality Planning Model for Videophone Services、 ITE Journals、 July 2008、 <u>K. Yamagishi</u> and T. Hayashi
○国際学会	1. Effect of Difference in 2D Video Quality for Left and Right Views on Overall 3D Video Quality、 Proc. of IEEE ICIP2012、 Sept. 2012、 <u>K. Yamagishi</u> , T. Kawano, and T. Hayashi
○国際学会	2. Parametric Packet-Layer Model for Monitoring Video Quality of IPTV Services、 IEEE ICC 2008、 May 2008、 <u>K. Yamagishi</u> and T. Hayashi
○国際学会	3. Opinion Model for Estimating Video Quality of Videophone Services、 IEEE GLOBECOM 2006、 Nov. 2006、 <u>K. Yamagishi</u> and T. Hayashi
○国際学会	4. Verification of Video Quality Opinion Model for Videophone Services、 2nd ISCA Tutorial & Research Workshop on Perceptual Quality of Systems、 Sept. 2006、 <u>K. Yamagishi</u> and T. Hayashi
特許	1. IMAGE QUALITY ESTIMATION DEVICE, METHOD, AND PROGRAM、 EP02229001、 2012/08/29、 <u>K. Yamagishi</u> and T. Hayashi
特許	2. IMAGE QUALITY ESTIMATION DEVICE, METHOD, AND PROGRAM、 CN: ZL200880123599.0、 2012/08/29、 <u>K. Yamagishi</u> and T. Hayashi
特許	3. IMAGE QUALITY ESTIMATION DEVICE, METHOD, AND PROGRAM、 KR 01064220、 2011/09/05、 <u>K. Yamagishi</u> and T. Hayashi
特許	4. VIDEO QUALITY ESTIMATING DEVICE, METHOD, AND PROGRAM、 CA2604053、 2012/07/10、 <u>K. Yamagishi</u> and T. Hayashi
特許	5. VIDEO QUALITY ESTIMATING DEVICE, METHOD, AND PROGRAM、 US08154602、 2012/04/10、 <u>K. Yamagishi</u> and T. Hayashi
特許	6. VIDEO QUALITY ESTIMATING DEVICE, METHOD, AND PROGRAM、 US07965203、 2011/06/21、 <u>K. Yamagishi</u> and T. Hayashi
特許	7. VIDEO QUALITY ESTIMATING DEVICE, METHOD, AND PROGRAM、 CN: ZL200680009898.2、 2010/06/16、 <u>K. Yamagishi</u> and T. Hayashi
特許	8. 映像品質推定装置、方法、およびプログラム、特許第 4490483 号、2010/04/09、 <u>山岸和久</u> 、 <u>林孝典</u>
特許	9. Video Quality Estimation Apparatus, Method, and Computer-Readable Recording Medium for Recording Program、 KR0935650、 2009/12/29、 <u>K. Yamagishi</u> and T. Hayashi

種 類 別	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者 (申請者含む)
特許	1. VIDEO QUALITY ESTIMATING DEVICE, METHOD, AND PROGRAM, KR0933509、2009/12/15、 <u>K. Yamagishi</u> and T. Hayashi
特許	2. VIDEO QUALITY ESTIMATING DEVICE, METHOD, AND PROGRAM, CN: ZL200680009878.5、2009/12/09、 <u>K. Yamagishi</u> and T. Hayashi
	以下、本学位論文に直接関係のないもの
論文	1. 映像配信サービスを対象としたマルチメディア品質推定モデル、信学論(B)、Jan. 2011、富永聡子、 <u>山岸和久</u> 、林孝典、高橋玲
論文	2. Opinion Model using Psychological Factors for Interactive Multimodal Services、IEICE Transactions on Communications、Feb 2006、 <u>K. Yamagishi</u> and T. Hayashi
国際学会	1. Performance Comparison of Subjective Assessment Methods for 3D Video Quality、The fourth International Workshop on Quality of Multimedia Experience (QoMEX)、July 2012、T. Kawano、 <u>K. Yamagishi</u> 、and T. Hayashi
国際学会	2. Subjective Characteristics for Stereoscopic High Definition Video、The third International Workshop on Quality of Multimedia Experience (QoMEX)、Sept. 2011、 <u>K. Yamagishi</u> 、L. Karam、J. Okamoto、and T. Hayashi
国際学会	3. No Reference Video-Quality-Assessment Model for Video Streaming Services、The 18th International Packet Video Workshop (PV2010)、Dec. 2010、T. Kawano、 <u>K. Yamagishi</u> 、K. Watanabe、and J. Okamoto
国際学会	4. Hybrid video-quality-estimation model for IPTV services、IEEE GLOBECOM 2009、Nov. 2009、 <u>K. Yamagishi</u> 、T. Kawano、and T. Hayashi
国際学会	5. Multimedia Quality Assessment Method for IPTV Services、The Third International Workshop on Image Media Quality and its Applications 2008 (IMQA ' 08)、Sept. 2008、T. Tominaga、 <u>K. Yamagishi</u> 、and T. Hayashi
国際学会	6. Proposal of new QOE assessment approach for quality management of IPTV services、Proc. of IEEE ICIP2008、Oct. 2008、K. Watanabe、 <u>K. Yamagishi</u> 、J. Okamoto、and A. Takahashi
国際学会	7. Multimedia Quality Integration Function for Videophone Services、IEEE GLOBECOM 2007、Nov. 2007、T. Hayashi、 <u>K. Yamagishi</u> 、T. Tominaga and A. Takahashi
国際学会	8. Perceptual QoS Evaluation Model for Audiovisual Communication Services、World Telecommunications Congress 2006 (WTC' 06)、May 2006、T. Hayashi、 <u>K. Yamagishi</u> 、and Hideaki Yoshino
国際学会	9. State of the Art of Multimedia Quality Assessment Methods、ITU-T Workshop on Video and Image Coding and Applications (VICA)、July 2005、T. Hayashi and <u>K. Yamagishi</u>
国際学会	10. Analysis of psychological factors for quality assessment of interactive multimodal service、Electronic Imaging 2005、Jan. 2005、 <u>K. Yamagishi</u> and T. Hayashi
国際学会	11. A Study on Motion Compensated Prediction in DCT Domain with Multiple Mode Selection、Picture Coding Symposium、Apr. 2003、 <u>K. Yamagishi</u> 、J. Katto、and Y. Yasuda

種 類 別	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者 (申請者含む)
特許	1. 映像品質推定装置、方法、およびプログラム、特許第 4913893 号、2012/01/27、 <u>山岸和久</u> 、林孝典
特許	2. 品質推定方法、品質推定システム、ユーザ端末、品質管理端末およびプログラム、特許第 4914400 号、2012/01/27、 <u>山岸和久</u> 、林孝典、牛木一成
特許	3. 映像品質推定装置、方法、およびプログラム、特許第 4861371 号、2011/11/11、 <u>山岸和久</u> 、林孝典
特許	4. 映像品質推定装置、方法、およびプログラム、特許第 4802200 号、2011/08/12、 <u>山岸和久</u> 、林孝典
特許	5. 映像品質推定方法、装置およびプログラム、特許第 4802209 号、2011/08/12、 <u>山岸和久</u> 、林孝典
特許	6. 映像品質推定装置、方法、およびプログラム、特許第 4796019 号、2011/08/05、 <u>山岸和久</u> 、林孝典
特許	7. 映像品質推定装置、方法、およびプログラム、特許第 4787303 号、2011/07/22、 <u>山岸和久</u> 、林孝典
特許	8. 映像品質推定方法、装置、およびプログラム、特許第 4787210 号、2011/07/22、 <u>山岸和久</u> 、林孝典
特許	9. ネットワークコミュニケーションサービス満足度推定装置、方法、およびプログラム、特許第 4745721 号、2011/05/20、 <u>山岸和久</u> 、林孝典
特許	10. 映像品質推定装置、方法、およびプログラム、特許第 4740967 号、2011/05/13、 <u>山岸和久</u> 、林孝典
特許	11. 双方向マルチモーダルサービス品質推定装置および方法、特許第 4510583 号、2010/05/14、 <u>山岸和久</u> 、林孝典
特許	12. 映像品質推定装置、方法、およびプログラム、特許第 4451856 号、2010/02/05、 <u>山岸和久</u> 、林孝典
特許	13. 映像品質パラメータ推定装置、方法、およびプログラム、特許第 4451857 号、2010/02/05、 <u>山岸和久</u> 、林孝典
特許	14. 映像品質推定装置、方法、およびプログラム、特許第 4408120 号、2009/11/20、 <u>山岸和久</u> 、林孝典
特許	15. 双方向通信サービスシステムにおける品質推定方法および品質推定装置、特許第 4365291 号、2009/08/28、 <u>山岸和久</u> 、林孝典
特許	16. 映像品質推定装置、方法およびプログラム、特許第 4317540 号、2009/05/29、 <u>山岸和久</u> 、林孝典
特許	17. VIDEO COMMUNICATION QUALITY ESTIMATION DEVICE, METHOD, AND PROGRAM、CA 2617893、2011/05/03、T. Hayashi and <u>K. Yamagishi</u>
特許	18. VIDEO COMMUNICATION QUALITY ESTIMATION DEVICE, METHOD, AND PROGRAM、CN ZL200680030787.X、2010/06/16、T. Hayashi and <u>K. Yamagishi</u>
特許	19. VIDEO COMMUNICATION QUALITY ESTIMATION DEVICE, METHOD, AND PROGRAM、KR 0947275、2010/03/05、T. Hayashi and <u>K. Yamagishi</u>
	その他：国内大会 33 件、国内研究会 18 件、標準化寄書 38 件、NTT 紀行論文 11 件、ITU ジャーナル 10 件