

QoE ベースビデオ出力方式 SCS による音声・ビデオインタラクティブ IP 通信の QoE 向上

横井 友洋†

田坂 修二‡

布目 敏郎‡

名古屋工業大学大学院 工学研究科 情報工学専攻

1 まえがき

音声・ビデオ IP 伝送における QoE (Quality of Experience) 向上のためのビデオ出力方式として、SCS (Switching between error Concealment and frame Skipping) が提案されている [1].

SCS は、誤り補償とフレームスキップを切り替えることで QoE 向上を図る方式である。誤り補償では、ビデオ伝送中にパケットの欠落が生じ、ビデオフレーム内に欠落したスライスがある場合、そのスライスに対して他の情報を利用した補完を行う。しかし、完全な補完は困難であり、多くの場合ビデオの空間品質は劣化する。一方、フレームスキップは、パケット欠落によって空間品質が低下したビデオフレームを出力しない方式である。この方式を適用すると、出力されるビデオフレームの空間品質は元のままに維持されるが、時間品質は劣化してしまう。このように、空間品質と時間品質にはトレードオフの関係がある。SCS は、このトレードオフ関係を利用することで QoE の向上を図る。

SCS では、1 ビデオフレーム中の誤り補償されたスライスの割合 (誤り補償率) [%] を算出する。誤り補償率が設定した閾値 [%] を超えた場合、ビデオ出力方式がフレームスキップに切り替わる。誤り補償率が閾値を下回る 1 フレームが出現すると、ビデオ出力方式は誤り補償に切り替わる。

従来研究では、単方向マルチメディアストリーミングにおける SCS の有効性が示されている [1]。しかし、テレビ電話に代表されるインタラクティブ通信において、SCS が有効であるか否かは検証されていない。そこで本稿では、音声・ビデオインタラクティブ IP 通信に SCS を適用する。SCS 適用の効果を検証するために、二種類の対等インタラクティブ作業について主観評価実験を行い、QoE を多次元的に評価する。

2 実験方法

2.1 実験環境

実験のネットワーク構成を図 1 に、伝送する音声・ビデオの仕様を表 1 に示す。表 1 の MU (Media Unit) はメディア同期の処理単位であり、1 ビデオ MU は 1 フレームに、1 音声 MU は 40[ms] 分のサンプリングデータに相当する。

実験時には、負荷送信端末から負荷受信端末へ TCP と UDP の混合トラフィックを伝送し、メディアの品質を意図的に劣化させる。TCP トラフィックは、負荷受信端末上の複数のクライアントプロセスが HTML ファイルを要求することで生成される。1 クライアントにつき、平均 320[KB] の対数正規分布に従う大きさの HTML ファイルを、平均 1 秒の指数分布に従う時間間隔で要求する。この TCP トラフィックは、インターネット上に存在するファイルの大きさが対数正規分布に従うこと [2] に加え、Web ページの大きさの平均が 320.24[KB] であること [3] に基づき生成している。また、UDP トラフィックとして、ペイロード長 1472 バイトのパケットを指数分布乱数間隔で送信する。この際、パケット送信間隔の平均値は 500/クライアント数 [ms] とし、TCP クライアントの増加に伴い、UDP トラフィックも増加するように設定した。本実験では、負荷クライアントプロセス数を 50 (低

表 1 ビデオ・音声の仕様

	ビデオの仕様	音声の仕様
符号化方式	H.264 (x264) 640×480 画素	G.711 μ -law
平均ビットレート	2400[kb/s]	64[kb/s]
平均 MU レート	30[MU/s]	25[MU/s]
ピクチャパターン	IPPP, I+14P's	-
スライス分割数	20	-
作業時間	30[s]	-

表 2 SD 法で用いる形容詞対

評価項目	形容詞対
ビデオ空間品質	(ビデオが) 崩れた～まとまった
ビデオ時間品質	(ビデオが) ぎこちない～なめらかな
ビデオ総合品質	(相手の状態が) 把握しにくい～把握しやすい
音声品質	(音声が) きたない～きれいな
メディア同期	(音声とビデオが) ずれた～一致した
相互作用性	(相手からの応答が) 遅い～速い
相互作用性	(やり取りが) 不自然な～自然な
相互作用性	(作業が) 難しい～簡単な
相互作用性	(作業をしていて) 不快な～快適な
総合評価	(総合的な評価は) 非常に悪い～非常に良い

負荷) または 70 (高負荷) に設定している。

プレイアウトバッファリング時間は、120[ms]、140[ms]、300[ms] の 3 種類を採用する。120[ms] は、低負荷時でも出力時刻に間に合わないビデオ MU が発生する値である。140[ms] は、高負荷時に出力時刻に間に合わないビデオ MU が発生するが、低負荷時にはほぼ発生しない値である。300[ms] は、高負荷時でも出力時刻に間に合わないビデオ MU がほぼ発生しない値である。

SCS の閾値は、100%、50%、20%、0% の 4 種類を採用した。これらの値は、ビデオ MU 欠落率がほぼ等間隔となるように設定している。

実験における作業は、ビデオ重視の作業として計算作業を、音声重視の作業として数え上げ作業を採用した。作業内容の詳細は 2.2 節に記す。

今回の実験における刺激は各作業につき、ピクチャパターン 2 種類、負荷クライアント数 2 種類、プレイアウトバッファリング時間 3 種類、SCS の閾値 4 種類の計 48 個である。刺激の評価には SD (Semantic Differential) 法を採用する。SD 法で用いる形容詞対は表 2 の通りである。被験者は各刺激に対し、形容詞対毎に 1 から 5 までの得点を用いた系列カテゴリ法 [4] で評価を行う。得点は、表 2 の形容詞対列の左側に示すネガティブな形容詞が 1、右側に示すポジティブな形容詞が 5 に対応する。3 は“どちらでもない”とし、2 及び 4 は、それぞれ 1 と 3 の中間及び 3 と 5 の中間の評価とする。被験者は 10 代・20 代学生 32 名 (男性 16 名、女性 16 名) である。

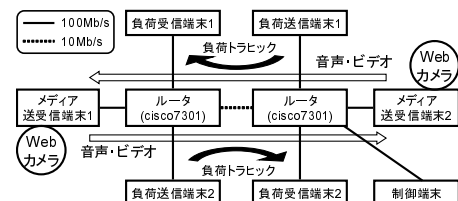


図 1 ネットワーク構成

2.2 作業内容

本実験では、二種類の対等インタラクティブ作業を行う。以下では、各メディア送受信端末の前に座った被験者を被験

Enhancement of QoE in Audiovisual Interactive IP Communications with the QoE-Based Video Output Scheme SCS

Tomohiro Yokoi † Shuji Tasaka ‡ Toshiro Nunome ‡
Department of Computer Science and Engineering, Graduate
School of Engineering, Nagoya Institute of Technology
†yokoi@inl.nitech.ac.jp ‡{tasaka, nunome}@nitech.ac.jp

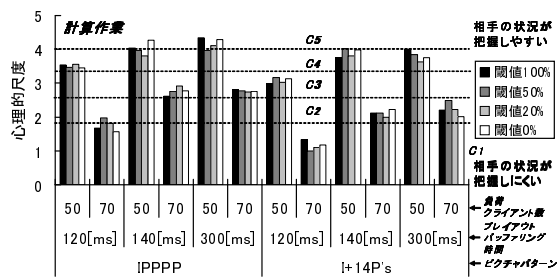


図2 相手の状態が把握しにくい～把握しやすい（計算）

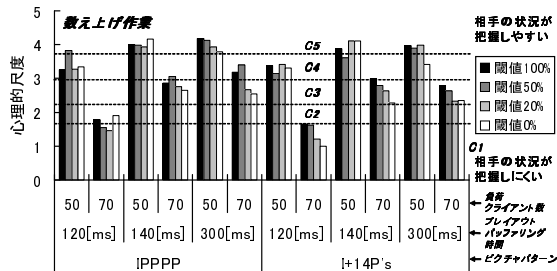


図3 相手の状態が把握しにくい～把握しやすい（数え上げ）

者A, 被験者Bと呼ぶ。

計算作業

被験者Aは任意のトランプを2枚選び、被験者Bに提示する。被験者Bは被験者Aの提示したトランプを確認次第、「+」、「-」、「×」のいずれかの演算子カードを被験者Aに提示する。被験者Aは被験者Bの提示した演算子カードを確認し、自分の選択したトランプの数字と相手の選択した演算子から計算式を作り、その式と答えを被験者Bに音声で伝える。この流れを役割を交代しながら30秒間続ける。

数え上げ作業

被験者Aは1から5までの任意の数字の一つを選び、1からその数字まで昇順にカウントを行う。被験者Bは被験者Aのカウントを確認次第、それを真似して伝え返す。この流れを役割を交代しながら30秒間続ける。

3 実験結果と考察

本実験では、QoE評価尺度として心理的尺度[5]を用いた。系列カテゴリー法によって得られた尺度値に対し、Mostellerの適合度検定[6]を行った。有意水準0.05において、系列カテゴリー法により得られた尺度値が測定結果に適合しなかった形容詞対に関しては、適合するまで推定値と実測値の誤差が大きい刺激を除外した。

以下では、2形容詞対に関して考察を行う。

相手の状態が把握しにくい～把握しやすい（図2, 3）

どちらの作業においても、プレイアウトバッファリング時間140[ms]以上で高いQoEを示している。計算作業では、SCSの閾値の違いによるQoEの差はさほど見られないため、プレイアウトバッファリング時間の設定が支配的であるといえる。一方、数え上げ作業では、多くの条件において閾値0%が悪い結果を示している。数え上げ作業は、ビデオの動きの大部分が口の動きであるため、フレームスキップにより音声と口の動きがずれることで、相手の状況を把握しにくいと感じる傾向にある。特に、ピクチャパターンI+14P'sかつ高負荷の場合には、フレームスキップによる時間品質の劣化が激しくなるため、この傾向が顕著である。

以上の結果より、計算作業では、実装の簡単なフレームスキップのみを適用しても、ある程度のQoEを維持出来ることが分かる。一方、数え上げ作業においては、フレームスキップのみの適用ではQoEを維持出来ないため、閾値0%以外を設定することが望ましいといえる。

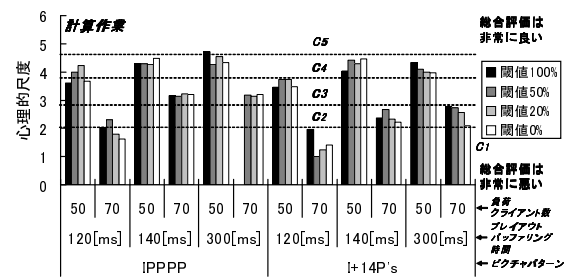


図4 総合的な評価は非常に悪い～非常に良い（計算）

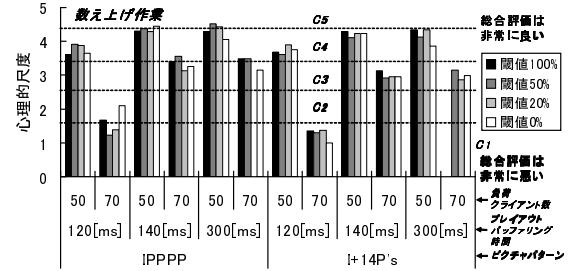


図5 総合的な評価は非常に悪い～非常に良い（数え上げ）

総合的な評価は非常に悪い～非常に良い（図4, 5）

計算作業における総合評価の結果は、図2に示した「相手の状態が把握しにくい～把握しやすい」の結果と同様であった。計算作業において、SCSの閾値が支配的にQoEに影響を及ぼさないという結果は、文献[1]における単方向ストリーミングの場合とは異なる結果となった。この原因として、計算作業では、トランプや演算子カードを選択すること、また、式を作り計算を行うこと等、ビデオから意識が離れる要素が多いことが挙げられる。一方、数え上げ作業は音声主体の単純な作業であり、作業中でもビデオに集中することが出来る。従って、被験者がフレームスキップによる時間品質の劣化に気づき、QoEの低下を招いたと考えられる。しかしながら、数え上げ作業においては、高負荷で図3の場合ほど閾値0%のQoEが低くなっていないことが分かる。これは、数え上げ作業は音声さえ聞こえていれば継続可能であり、ビデオ時間品質の劣化が総合評価に及ぼす影響が限定されたためである。

4 むすび

本実験の結果より、SCSの閾値を適切に変更することにより、インタラクティブ通信におけるQoEは改善されることが示された。また、単方向ストリーミングの場合とは異なり、プレイアウトバッファリング時間の違いが支配的にQoEに効くことも判明した。しかしながら、簡単な実装で手軽にQoE向上が狙えるという点で、インタラクティブ通信においてもSCSは有効であるといえる。SCSの適切な閾値は作業依存であるため、異なる作業の実験を行い、作業毎に適切な閾値を求めることが今後の課題として挙げられる。

謝辞

本研究は、平成23年度学術研究助成基金助成金・挑戦的萌芽研究（課題番号23656253）によって行われた。

参考文献

- [1] S.Tasaka et al., *Proc. ACM Multimedia 2008*, pp.259-268, Oct. 2008.
- [2] A. B. Downey and W. College, *Proc. MASCOTS 2001*, pp.361-370, Aug. 2001.
- [3] Google, <http://code.google.com/intl/ja/speed/articles/web-metrics.html>, 26 May 2010.
- [4] 田中 良久, “心理学的測定法 第2版”, 東京大学出版会, 1977.
- [5] S. Tasaka and Y. Ito, *Conf. Rec. IEEE ICC'2003*, pp.1880-1886, May 2003.
- [6] F. Mosteller, *Psychometrika*, vol.16, no.2, pp.207-218, 1951.