

## 力覚メディア・サウンド・ビデオを用いた遠隔合奏における ネットワーク遅延がQoEに及ぼす影響

曾 希† 岩田 海† 石橋 豊† 福嶋 慶繁† 菅原 真司†

†名古屋工業大学

### 1 まえがき

力覚メディアをサウンド、ビデオなどのメディアと組み合わせて用いることにより、遠隔地での様々な作業が可能となる。そのような作業を可能とするシステムの一例として、力覚メディア・サウンド・ビデオ伝送システム [1] がある。このシステムでは、利用者は一つの触覚インタフェース装置を用いて反力を感じながら、遠隔地にある別の触覚インタフェース装置を遠隔操作することができる。また、遠隔地で入力されたビデオとサウンドを視聴することもできる。本稿では、このシステムを使用して遠隔操作による合奏を扱う。

しかし、インターネットを介して力覚メディア、サウンド及びビデオを伝送すると、ネットワーク遅延やその揺らぎなどの影響により、メディアの時間関係に乱れが生じ、出力品質が大きく劣化してしまう恐れがある。この問題を解決するには、メディア同期制御を行うことが必要となる。高効率なメディア同期制御のためには、ネットワーク遅延がユーザ体感品質 (QoE: Quality of Experience) に及ぼす影響を調査する必要がある。

遠隔演奏や合奏を扱っている研究の例に、文献 [2], [3] がある。文献 [2] では、クラリネットとアルトサックスの合奏を行う場合に許容できるサウンドの遅延時間の調査と、ネットワークを介した実験を行っている。サウンドの許容遅延の調査では、二人の奏者の距離を離しながら合奏を行った結果、最大許容遅延は往復で約 60 ms であることを明らかにしている。また、ネットワークを介した実験では、二地点で合奏を行ったが、60ms より大きなネットワーク遅延が発生し、合奏が難しくなると報告している。しかし、この遠隔合奏ではサウンドとビデオのみが扱われており、力覚メディアは扱われていない。

文献 [3] では、力覚メディア・サウンド・ビデオ伝送システムを用いて、遠隔操作によってタンバリンを叩く作業を行い、受信バッファリング制御が QoE に与える影響を明らかにしている。しかし、この文献では合奏は行われていない。遠隔合奏では、サウンドを同期させなければならないため、ネットワーク遅延の影響によって一つの楽器を演奏するときよりも作業が難しくなる。そのため、遠隔合奏に対してネットワーク遅延が QoE に及ぼす影響を調査する必要がある。

そこで本稿では、文献 [3] のタンバリンの遠隔演奏に加えて、タンバリンを置いてある側の端末で鍵盤ハーモニカを演奏する遠隔合奏に対し、ネットワーク遅延が

QoE に及ぼす影響を調査する。

### 2 力覚メディア・サウンド・ビデオ伝送システム

図 1 に示すように、本システムは、ネットワークを介して接続されたマスタ端末とスレーブ端末から成る。各端末には触覚インタフェース装置として PHANToM Omni(以下、PHANToM と略す)、マイク付きビデオカメラ及びヘッドセットが備え付けられている。また、スレーブ端末側に楽器(タンバリン)が置かれている。力覚メディア、サウンド及びビデオは、両端末間を UDP によって双方向に伝送される。

マスタ端末の利用者は、ビデオを見ながらマスタ端末の PHANToM を用いてスレーブ端末の PHANToM を遠隔操作し、スレーブ端末側の楽器を叩く。このとき、スレーブ端末の PHANToM にはマスタ端末の PHANToM を操作した方向への反力が働き、マスタ端末の PHANToM にはスレーブ端末の PHANToM に働く反力とは逆方向の反力が提示されるため、利用者はあたかも自分のいる場所にある楽器を叩いているように感じる。この反力は、マスタ端末とスレーブ端末の PHANToM のカーソルの位置ベクトルの差により計算される。また、両端末の利用者は、ヘッドセットにより会話を行ったり、演奏するサウンドを聞くことができる。マスタ端末のビデオはマスタ端末の利用者の顔を映しており、スレーブ端末のビデオはタンバリンと PHANToM スタイルスの先端を映している(図 1 参照)。



図 1: システム構成

### 3 評価環境

#### 3.1 評価システム

評価システムは、図 1 のネットワークをネットワークエミュレータ (NIST Net) に置き換えて構成される。NIST Net は、通過する各パケットに対して固定遅延を付加するために用いられる。スレーブ端末側には、図 1

Influence of Network Delay in Networked Ensemble with Haptic Media, Sound, and Video

†Qi ZENG †Kai IWATA †Yutaka ISHIBASHI †Norishige FUKUSHIMA †Shinji SUGAWARA

†Nagoya Institute of Technology

に示すように、タンバリンの他に鍵盤ハーモニカが置かれる。

マスタ端末とスレーブ端末は、力覚メディアユニット（メディア同期のための処理単位。以下、MU と略す）、サウンド MU、ビデオ MU を互いに送信し合う。両端末はメディア同期アルゴリズムとして Skipping[4] を用いる。Skipping は、MU を受信すると直ちに出力する。ただし、順序制御を行い、最新の MU のみを出力する。サウンドの符号化方式にはリニア PCM を用いる。ビデオの符号化方式には MPEG-1 を採用し、インタラクティブ性を高く維持するために I ピクチャのみとする。ビデオの表示サイズは 320 × 240 ピクセルである。力覚 MU、サウンド MU、及びビデオ MU の転送レートはそれぞれ 1000MU/s、50MU/s、30MU/s であり、平均ビットレートはそれぞれ 320kbps、128kbps、1536kbps である。

### 3.2 評価方法

QoE 評価では、著者の一人がスレーブ端末にある鍵盤ハーモニカを演奏し、マスタ端末の各被験者が鍵盤ハーモニカのサウンドに合わせて遠隔操作によってタンバリンを叩き合奏を行った。演奏する曲は「メリーさんのひつじ」であり、タンバリンを叩く間隔は 2 拍に 1 回、曲の早さは 100bpm (beats per minute) とした。評価では、NIST Net によって固定遅延を双方向に付加し、0ms から 120ms まで 10ms または 20ms 間隔で変化させた。固定遅延は被験者毎にランダムに提示された。被験者には、評価の前に固定遅延のない状態で合奏を行い、それを基準として、PHANToM の操作性、インタラクティブ性、鍵盤ハーモニカとタンバリンの音の同期品質、および総合品質について、5 段階妨害尺度（5：劣化が感じられない、4：劣化が分かるが気にならない、3：劣化が気になるが邪魔にならない、2：劣化が邪魔になる、1：劣化が非常に邪魔になる）に従って評点をつけてもらい、MOS (Mean Opinion Score) を求めた。PHANToM の操作性、インタラクティブ性、鍵盤ハーモニカとタンバリンのサウンドの同期品質は、それぞれ PHANToM の動きし易さ、被験者が操作する PHANToM の動きとビデオの中のスレーブ端末の PHANToM の動きのずれ、鍵盤ハーモニカとタンバリンのサウンドがどれだけ合っていると感じられたかを表す。また、総合品質は他の三つの品質の加重和である。被験者は、21 歳から 30 歳の男女 15 名である。一回の評価時間は 20 秒であり、一人当たりの総評価時間は約 10 分であった。

## 4 評価結果と考察

固定遅延に対する PHANToM の操作性、インタラクティブ性、鍵盤ハーモニカとタンバリンのサウンドの同期品質、及び総合品質の MOS を図 2 に示す。この図には 95% 信頼区間も示されている。

図 2 より、固定遅延が増加するにつれて、全ての MOS は減少することが確認できる。この理由は次の通りである。まず、PHANToM の操作性に関しては、固定遅延が大きくなるにつれて、提示される反力が大きくなり、PHANToM の操作がし難くなるため、MOS は減少する。次に、インタラクティブ性については、往復の固定遅延

の分だけビデオが遅れて提示されるため、固定遅延が大きくなるに従い、MOS は減少している。そして、同期品質に関しては、固定遅延が大きくなるほど、PHANToM の操作がし難くなるため、鍵盤ハーモニカに合わせて叩こうとしても、サウンドをうまく合わせる事が難しくなり、MOS が低くなる。これらの理由から、三つの品質の加重和である総合品質の MOS も減少する。

また、図 2 では、固定遅延が約 40ms 以下のとき、総合品質の MOS は 3.5 以上の値をとっている。このため、文献 [5] によると、力覚メディア、サウンド及びビデオを用いた遠隔合奏において、ネットワーク遅延が約 40ms 以下であれば、許容可能であると言える。この最大許容遅延の値は、文献 [2] の結果と比べて大きくなっている。これは、文献 [2] では二地点で楽器が演奏されるのに対し、本稿では同一の地点で演奏されているためであると考えられる。

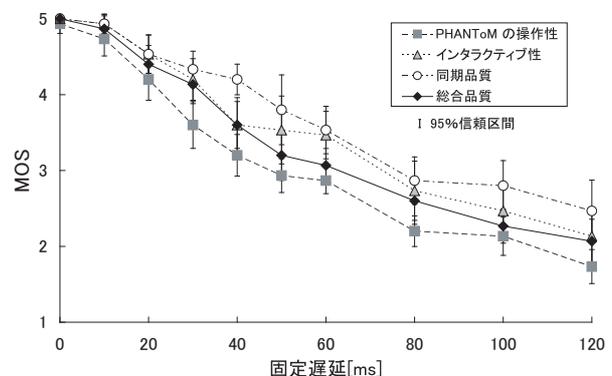


図 2: 固定遅延に対する MOS

## 5 むすび

本稿では、力覚メディア・サウンド・ビデオ伝送システムを用いて、鍵盤ハーモニカとタンバリンによる遠隔合奏を行い、ネットワーク遅延が QoE に及ぼす影響を調査した。その結果、ネットワーク遅延が約 40ms 以下であれば許容可能であるということが明らかとなった。

今後の課題として、本稿の結果を利用して、VTR (Virtual Time Rendering) アルゴリズム [4] において最大許容遅延を約 40ms に設定して、その効果を調査することが挙げられる。また、今回の評価ではタンバリンと鍵盤ハーモニカを同一の地点に置いたが、それらを別々の地点に配置する場合を扱うことも重要である。

謝辞 本研究は、平成 23 年度科学研究費補助金基盤研究 (C) (課題番号 22560368) によって行われた。

## 参考文献

- [1] 藤本 他, 映情学誌, vol. 61, no. 12, pp. 1756-1765, Dec. 2007.
- [2] 茂木 他, 情処研報, vol. 2002, no. 54, pp. 49-54, June 2002.
- [3] K. Iwata *et al.*, *ACM Computers in Entertainment*, vol. 8, issue 2, article 12, pp. 1-14, Dec. 2010.
- [4] Y. Ishibashi *et al.*, *Proc. CQR2002*, pp. 213-217, May 2002.
- [5] 浅谷, 信学会, ISBN978-4-88552-112-6, May 1993.