

研究速報

多地点間の合唱におけるメディア同期のユーザ体感品質評価

Quality of Experience Assessment of Media Synchronization in Chorus among Multi-Points

宮下 裕 司[†], 正会員 石 橋 豊[†], 正会員 福 嶋 慶 繁[†],
菅 原 真 司[†], コスタス プサニス^{††}

Yuji Miyashita[†], Yutaka Ishibashi[†], Norishige Fukushima[†], Shinji Sugawara[†] and Kostas E. Psannis^{††}

Abstract In this paper, we apply inter-destination (or group) synchronization control, which adjusts the output timing of media among multiple terminals, to chorus with voice and video among multi-points to synchronize singing voices for the listener. By quality of experience (QoE) assessment, we show that the inter-destination synchronization control can maintain the media synchronization quality in chorus to a greater degree than when the control is not used.

キーワード：多地点間通信, 音声, ビデオ, ユーザ体感品質, メディア同期, 端末間同期制御

1. ま え が き

複数の利用者が音声やビデオ等を用いて、協調作業や競合作業等を行う多地点間通信が注目されている¹⁾²⁾。多地点間通信では、複数の端末間でメディアの送受信が行われる。しかし、ネットワーク遅延やその揺らぎ等により、メディアの時間関係が乱される恐れがある。これはメディアの出力品質を大きく劣化させる要因となる。また、各受信端末で複数の送信端末から発生したメディアの出力タイミングが異なると、協調作業や競合作業において、作業が成立しなくなったり、公平性が損なわれたりする可能性も生じる³⁾。そこで、複数の端末で発生したメディアの出力タイミングを複数端末間でそろえる端末間同期制御⁴⁾が必要となる⁵⁾。

文献3)では、音声とビデオを用いた競合作業として、早押しクイズ形式の名前当てゲームを対象とし、ネットワーク遅延が公平性に及ぼす影響を調べている。その結果、端末間でネットワーク遅延に差が生じると、端末によって出力タイミングが異なり、公平性が劣化することを示している。文献5)では、文献3)で扱われた競合作業である早押しクイズ形式の名前当てゲームと、協調作業としてネットワーク上で

行うじゃんけんおよび数字を読み上げる作業を対象として、端末間同期制御が公平性のユーザ体感品質 (QoE: Quality of Experience)⁶⁾に及ぼす影響を調査している。そして、端末間同期制御により、早押しクイズ形式の名前当てゲームの公平性を高く保てることを示している。

この他、複数端末で発生したメディア間の出力タイミングをそろえる必要のある作業として、合唱や合奏など音楽を扱う協調作業^{7)~10)}が考えられる。合唱では、音声・ビデオが用いられ、タイミングの基準を示す指揮者、指揮に合わせて歌う歌い手、歌を聴く視聴者などが存在する。これらの人々が離れた地点に別れ、ネットワークを介して合唱を行う場合、各地点間のネットワーク遅延の違いにより、視聴者のいる地点で出力される歌声の間にタイミングのずれが生じる。したがって、視聴者のいる地点では、制御によって、複数の地点からの音声の出力タイミングをそろえることが考えられる。また、歌い手がいる地点では、指揮者のいる地点の指揮あるいは歌声に合わせて歌うので、指揮者のいる地点から受信した音声・ビデオの出力品質の劣化を防ぐため、メディア内およびメディア間同期制御¹¹⁾を行う必要がある。しかし、従来の研究では、ユーザ体感品質評価は行われておらず、端末間同期制御の効果は十分に調査されていない。

ネットワークを介して大規模な合唱を行った代表例として、長野オリンピックの開会式が挙げられる⁷⁾。長野オリンピックの開会式では、五都市(ニューヨーク、北京、シドニー、ベルリン、ケープタウン)を衛星回線で結んでの「第九」の遠隔合唱が行われている。この例では、指揮者とオーケストラおよび合唱隊のいる長野県民文化会館、合唱隊の

2011年9月7日受付, 2011年12月27日再受付, 2012年2月14日採録

[†]名古屋工業大学 大学院 工学研究科 創成シミュレーション工学専攻
(〒466-8555 名古屋市昭和区御器所町, TEL 052-732-2111)

^{††}マケドニア大学 技術マネジメント学科
(〒54006 ギリシャ テッサロニキ)

[†]Department of Scientific and Engineering Simulation, Graduate School of Engineering, Nagoya Institute of Technology
(Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan)

^{††}Department of Technology Management, University of Macedonia
(Thessaloniki 54006, Greece)

いる五都市の会場、観客のいる長野オリンピック会場を拠点としている。指揮者のいる地点から五都市の会場に指揮者のビデオを送信し、五都市の会場で合唱隊が指揮者のビデオに合わせて歌う。そして、五都市の会場の音声・ビデオを長野オリンピック会場で出力する。ここで、長野オリンピックと五都市の各会場間の音声・ビデオのネットワーク遅延は予め計測されており、そのうち、ネットワーク遅延が最も大きい会場の音声・ビデオに合わせて、他の会場の音声・ビデオを出力している。このように、予め計測したネットワーク遅延の内最も大きい値を選択して基準値とし、すべての地点からの音声・ビデオのネットワーク遅延を基準値にそろえることで、合唱のメディア同期を実現している。しかし、ネットワーク遅延の基準値を予め設定する方法では、ネットワーク遅延の変化に対する柔軟性に欠ける。つまり、ネットワーク遅延が基準値を超えて大きくなる場合、メディア同期が実現できなくなる。また、基準値を大きく設定しすぎると、バッファリング時間が必要以上に大きくなってしまい、リアルタイム性が損われる。

そこで本論文では、長野オリンピックの遠隔合唱と同じように、指揮者、歌手、視聴者を別々に配置する状況を考え、音声・ビデオを用いた合唱に対して端末間同期制御を適用する。そして、ユーザ体感品質評価により、ネットワーク遅延が視聴者の感じるメディア同期品質に及ぼす影響や端末間同期制御の効果を調査する。

以下では、まず、2章で端末間同期制御について述べる。次に、3章で評価システムと評価方法を述べ、4章では評価結果について考察する。

2. 端末間同期制御

本論文では、基準となる指揮者（指揮をしながら歌う）がいる地点（地点数: 1）、指揮者の指揮と歌に合わせて歌う歌手がいる地点（地点数: $M(\geq 1)$ ）、各地点からの歌を視聴する視聴者がいる地点（地点数: 1）を想定する*。そして、視聴者のいる地点において歌声を同期させることを目的とする。以下では、指揮者がいる地点の端末を端末 1、歌手がいる地点の端末を端末 2, 3, ..., $M+1$ と呼び、視聴者のいる端末における端末間同期制御について説明する。また、歌手のいる端末における制御についても述べる。

本論文では、音声をマスタメディア、ビデオをスレーブメディアとする¹¹⁾。そして、音声に対してはメディア内同期制御を行い、視聴者の端末においては端末間同期制御も行う。ビデオに対しては、メディア内同期制御に加え、メディア間同期制御を行う。ここで、メディア内およびメディア間同期制御には、VTR(Virtual-Time Rendering) アルゴリズム¹¹⁾を用いる。VTR アルゴリズムでは、メディアユニット（メディア同期制御のための処理単位。以下、MU と略す。例えば、ビデオフレーム）のスキップ、ポーズなどを

行い、ネットワーク遅延の揺らぎに応じて動的にバッファリング時間を変更することによって、メディア内およびメディア間同期を達成しようとする。バッファリング時間の変更は、仮想時間軸を伸縮することによって行われ、各 MU はこの仮想時間軸に沿って出力される。ただし、仮想時間軸の伸長は最大許容遅延 $\Delta_{al}(> 0)$ ms 以下の範囲で行われる。

本論文で用いる端末間同期制御は、指揮者及び歌手の端末から視聴者の端末に送信された MU の出力タイミングを視聴者の端末においてそろえる**。本論文では、同期マエストロ方式⁴⁾による端末間同期制御を用いる。同期マエストロ方式では、ネットワーク上に一つの同期マエストロを設ける。同期マエストロは、受信したすべての情報を比較し、基準とすべき出力タイミング⁴⁾を決定する。簡単のため、本論文では同期マエストロを視聴者の端末に置く。視聴者の端末における端末 $i(i = 1, 2, 3, \dots, M+1)$ からの音声 MU の出力タイミングを、指揮者の端末 1 における音声 MU の発生時刻に $\Delta_S^{(i)}(> 0)$ ms を加えたものとする。以下では簡単のため、出力タイミングを $\Delta_S^{(i)}$ で表す。基準とすべき出力タイミング Δ_G は、 $\Delta_S^{(1)}, \Delta_S^{(2)}, \dots, \Delta_S^{(M+1)}$ のうち最も大きい値とする。そして、各端末 i からの音声 MU の出力タイミング $\Delta_S^{(i)}$ を基準とすべき出力タイミング Δ_G に近づけることで、すべての歌声の同期を達成しようとする。

合唱では、端末 $i(i = 2, 3, \dots, M+1)$ の歌手は端末 1 から受信した指揮者の音声・ビデオに合わせて歌う。視聴者の端末において歌手の端末 i からの音声 MU の出力タイミングを計算するため、まず、指揮者の端末 1 で音声 MU が発生してから、歌手の端末 i でそれが出力されるまでの時間 $\Delta_1^{(i)}$ を知る必要がある。そこで、歌手の端末 i では、最後に出力された MU の MU 遅延（一つの MU が発生してから、その MU が出力されるまでの時間）の値を $\Delta_1^{(i)}$ とし、自端末で発生する最新の音声 MU に付加して視聴者の端末に送信する。これは、出力した音声 MU と送信する音声 MU を対応付け、視聴者の端末において $\Delta_S^{(i)}$ の計算に用いるためである。なお、歌手の端末は、メディア内同期制御によって指揮者の端末からの最初の音声 MU を出力し、それと同時に自端末の MU の入力を開始する。次に、視聴者の端末では、歌手の端末 i で音声 MU が発生してから、出力されるまでの時間 $\Delta_2^{(i)}$ を求め、 $\Delta_S^{(i)} = \Delta_1^{(i)} + \Delta_2^{(i)}$ とする。

視聴者の端末では、各端末からの音声 MU の出力タイミングを Δ_G にそろえるため、次のように出力タイミングの変更を行う。指揮者の端末 1 からの音声 MU に対しては、 $\Delta_S^{(1)}$ の値を変更することで出力タイミングを Δ_G に徐々に近づけていく。歌手の端末 $i(i = 2, 3, \dots, M+1)$ からの音声 MU に対しては、 $\Delta_2^{(i)}$ の値を変更することで、出力タイミング $\Delta_S^{(i)} = \Delta_1^{(i)} + \Delta_2^{(i)}$ を Δ_G に徐々に近づける（歌

* 本論文では、説明の簡単のため、視聴者のいる地点数を 1 としている。この地点数を複数にしても同じ制御が適用可能である。

** 各 MU にはタイムスタンプを付加する。各端末はタイムスタンプを用いて受信した MU を破棄するか否かを判断する。また、NTP (Network Time Protocol)¹²⁾等を用いることにより、すべての端末で時刻同期がとれているものとする。

い手の端末 i においては端末間同期制御による $\Delta_1^{(i)}$ の変更を行わないことに注意されたい). 具体的には, 音声 MU が Δ_G よりも $N_a(\geq 1)$ 回連続で早く出力された場合, 出力タイミングを $\Delta_I(> 0)$ だけ遅らせる. 音声 MU が Δ_G よりも $N_b(\geq 1)$ 回連続で遅く出力された場合, 出力タイミングを $\Delta_D(> 0)$ だけ早める. そして, すべての端末からの音声 MU の出力タイミングが Δ_G と等しくなったとき, 同期がとれたと判断する. なお, $\Delta_S^{(i)}$ および $\Delta_2^{(i)}$ の値は $\Delta_L(\geq 0)$ ms 以上 $\Delta_H(\geq \Delta_L)$ ms 以下の範囲内で変化させる.

3. 評価環境

本評価では, 歌手のいる地点数を $M = 2$ とする. そして, 端末間同期制御を行う場合と行わない場合のユーザ体感品質を比較する. 本章では, 評価システムと評価方法について述べる.

3.1 評価システム

評価システムは, 図 1 に示すように, 4 台の端末 (端末 1 ~ 4) とネットワークエミュレータ (NIST Net¹³⁾) およびスイッチングハブから構成される. このように接続することで, すべての端末間の通信にネットワークエミュレータを介することができる. なお, 4 台の端末は別々の部屋に置かれる. 端末 1 にはタイミングの基準を示す指揮者, 端末 2, 3 には基準に合わせて歌う歌手, 端末 4 には視聴者を配置する. NIST Net は, 端末間で送受信されるパケットに対して, パレート正規分布¹³⁾ に従う付加遅延を発生させるために用いられる. また, MU は UDP により転送される.

端末 4 を除く各端末は, マイクから 20ms 毎に音声サンプル (符号化方式: PCM, ビットレート: 64kbps) をキャプチャリングし, 他端末に音声 MU をユニキャストする. また, 33ms 毎にビデオカメラからビデオを入力し, MPEG-1(GOP: I ピクチャのみ, 表示サイズ: 253 × 189 ピクセル, 平均ビットレート: 922kbps) に符号化した後, ビデオ MU を他端末にユニキャストする. なお, 自端末で発生した音声は, その端末においてエコーバックされない. また, ビデオは, 受信する端末だけでなく, 発生した端末においても出力される. 端末 1 は他の三つの端末に音声・ビデオを送信する. 端末 1 では他の地点の音声・ビデオは出力されず, 自端末のビデオのみが出力される. また, テンポの基準を設けるため, 端末 1 の地点にメトロノームを配置する. 端末 2, 3 は端末 1 から送信された音声・ビデオを受信および出力し, 自端末の音声・ビデオを端末 4 に送信する. また, 端末 2, 3 では自端末のビデオも出力される. 端末 4 は他の三端末から送信された音声・ビデオを受信および出力する. 図 2, 3 にそれぞれ歌手の端末と視聴者の端末におけるビデオの表示例を示す.

3.2 評価方法

ユーザ体感品質評価において, 被験者は端末 4 の視聴者を担当し, 合唱のメディア同期品質を評価する. 楽曲は 4 分

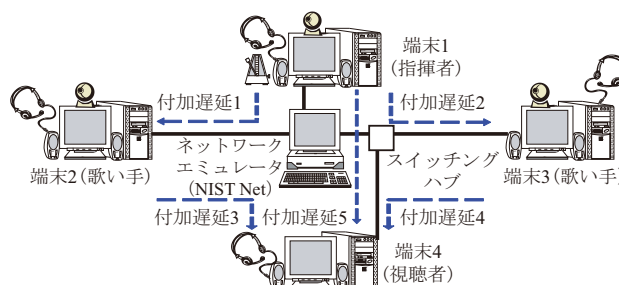


図 1 評価システムの構成
Configuration of assessment system.



図 2 歌手の端末におけるビデオの表示例
Displayed images at singer's terminal.



図 3 視聴者の端末におけるビデオの表示例
Displayed images at listener's terminal.

の 4 拍子の「かえるの合唱」であり, テンポはメトロノームにより 130bpm (beats per minute: 1 分間における拍の数) に設定する. 合唱の作業は以下のように行われる.

指揮者は「1, 2, 3, はい」と掛け声を掛け, 指揮 (手で拍子を取る) と歌唱を始める. そして, メトロノームのテンポに合わせて 20 秒程歌を歌う. 歌手は指揮者の掛け声に合わせて歌唱を開始し, 指揮者の歌に合わせて歌う. 視聴者は他の三地点から送信されてくる音声・ビデオを視聴し, 合唱の作業を観測する.

作業終了後, 視聴者は合唱のメディア同期品質を評価する. ここで, 合唱のメディア同期とは, 音声およびビデオによって観測される各歌声や指揮の動作が総合的に同期していることとする. ユーザ体感品質評価には ITU-R BT.500-11 の単一刺激法¹⁴⁾ を用いる. 被験者は評価の前に, 基準となる状態で作業を行う. 基準となる状態は, 付加遅延 1 ~ 5 の値を 0ms とし, メディア内同期制御, メディア間同期制御および端末間同期制御を行わない状態とする. その後, 付加遅延の値や端末間同期制御の有無を, 被験者毎にランダムに変えて, 合唱の作業と評価を繰り返す. なお, 端末間同期制御を行わない場合では, 端末 2, 3, 4 において VTR アルゴリズムによるメディア内およびメディア間同期制御を行う. 被験者は表 1 に従って基準からの劣化を評価する. 被験者は 21 歳から 24 歳までの男女 20 名である. 作業は被験者 1 人につき 24 回行い, 評価時間は 1 人あたり 30 分程度である.

評価では, 図 1 における付加遅延 2, 3, 4 の平均を 50ms とする. そして, 付加遅延 1 の平均を 50ms から 300ms

表 1 5 段階妨害尺度
Five-grade impairment scale.

評価	評価基準
5	遅延による劣化がわからない
4	遅延による劣化がわかるが気にならない
3	遅延による劣化が気になるが邪魔にならない
2	遅延による劣化が邪魔になる
1	遅延による劣化が非常に邪魔になる

まで 50ms 間隔で変化させる。また、付加遅延 5 の平均は 50ms と 300ms に変化させる。なお、各付加遅延の標準偏差を 10ms とする。端末間同期制御で用いる Δ_I および Δ_D は 20ms⁴⁾ とし、 N_a および N_b を 5 に設定する。さらに、 $\Delta_H = 500\text{ms}$, $\Delta_L = 40\text{ms}$ とする。VTR アルゴリズムのパラメータについては、最大許容遅延 Δ_{al} を Δ_H と同じく 500ms に、初期バッファリング時間 $J_{\max}^{(1)}$ を Δ_L と同じく 40ms に設定する。 Δ_{al} と J_{\max} は、音声およびビデオの出力品質の劣化を抑えるため、付加遅延の揺らぎを充分吸収できる値に設定されている。

4. 評価結果と考察

図 4 に付加遅延 1 の平均に対する合唱のメディア同期品質の MOS(平均オピニオン評点)¹⁴⁾ を示す。図 4 の考察のため、各端末間の音声およびビデオの平均 MU 遅延を計測したところ、ビデオの平均 MU 遅延は、メディア間同期制御により音声の平均 MU 遅延と同程度になっていた。そのため、以下では、音声の平均 MU 遅延に着目して考察を行う。ここで、指揮者の音声の平均 MU 遅延は、指揮者の端末 1 で発生した MU が視聴者の端末 4 で出力されるまでの平均時間であり、 $\overline{\Delta}_S^{(1)}$ で表される。端末 2, 3 の音声の平均 MU 遅延は、端末 1 の指揮者が発声してから、それに合わせた端末 2, 3 の歌手の音声は視聴者の端末で出力されるまでの時間の平均であり、それぞれ $\overline{\Delta}_S^{(2)}$, $\overline{\Delta}_S^{(3)}$ と表す。

図 4 より、端末間同期制御を行う場合では、付加遅延 1 および 5 の値に関らず、合唱のメディア同期品質の MOS が高く保たれていることがわかる。この理由は、端末間同期制御により、 $\overline{\Delta}_S^{(1)}$, $\overline{\Delta}_S^{(2)}$, $\overline{\Delta}_S^{(3)}$ が同程度の値にそろえられ、歌声が同期して聴こえるためである。

また、図 4 では、付加遅延 5 の平均が 50ms で端末間同期制御を行わない場合、付加遅延 1 の平均が大きくなる程、MOS が小さくなる傾向が見られる。この理由は、付加遅延 1 の平均が大きくなる程、 $\overline{\Delta}_S^{(2)}$ が大きくなり、端末 2 の歌手の歌声がより遅れて聴こえ、歌声の同期のずれが大きくなるためである。

さらに、図 4 から、付加遅延 5 の平均が 300ms で端末間同期制御を行わない場合では、付加遅延 1 の平均が 50ms から 200ms 程度までの範囲において、MOS は端末間同期制御を行う場合と同程度の高い値となっていることがわかる。この範囲において、 $\overline{\Delta}_S^{(1)}$, $\overline{\Delta}_S^{(2)}$, $\overline{\Delta}_S^{(3)}$ を測定したところ、 $\overline{\Delta}_S^{(1)} \geq \overline{\Delta}_S^{(2)} \geq \overline{\Delta}_S^{(3)}$ となり、 $\overline{\Delta}_S^{(1)}$ と $\overline{\Delta}_S^{(3)}$ の差が約 160ms となっていた。つまり、この範囲においては、端末 3 の歌

い手の歌声が最も早く聴こえ、指揮者の歌声が最も遅れて聴こえている。特に、付加遅延 1 の平均が 50ms の場合、付加遅延 2 の平均も 50ms であるため、 $\overline{\Delta}_S^{(2)}$ と $\overline{\Delta}_S^{(3)}$ がほぼ同じ値になる。この場合では、視聴者の端末において、先に 2 人の歌手の歌声が揃って出力され、その約 160ms 後に指揮者の歌声が出力される。そして、付加遅延 1 の平均が 200ms の場合、 $\overline{\Delta}_S^{(1)}$ と $\overline{\Delta}_S^{(2)}$ を測定したところ、両者はほぼ同じ値になっていた。この場合では、先に端末 3 の歌手の歌声のみが出力され、その約 160ms 後に指揮者と端末 2 の歌手の歌声がほぼ揃って出力される。また、付加遅延 1 の平均が 100ms, 150ms の場合では、 $\overline{\Delta}_S^{(1)} > \overline{\Delta}_S^{(2)} > \overline{\Delta}_S^{(3)}$ となっていた。すなわち、最も早く聴こえる端末 3 の歌手の歌声と最も遅く聴こえる指揮者の歌声の間に端末 2 の歌手の歌声が聴こえる。このとき、 $\overline{\Delta}_S^{(1)}$ と $\overline{\Delta}_S^{(2)}$ の差と $\overline{\Delta}_S^{(2)}$ と $\overline{\Delta}_S^{(3)}$ の差は、それぞれ 60ms~100ms 程度となっていた。これらのことより、今回の評価環境における合唱では、指揮者のビデオおよび音声は歌手のそれらより遅いか同時に出力される場合、最も早く出力される歌声と最も遅く出力される歌声が約 160ms ずれてもほとんど気にならないことがわかる。

一方、付加遅延 1 と 5 の平均が 50ms で端末間同期制御を行わない場合、付加遅延 2 の平均も 50ms であるため、 $\overline{\Delta}_S^{(2)}$ と $\overline{\Delta}_S^{(3)}$ がほぼ同じ値になる。この場合では、 $\overline{\Delta}_S^{(1)}$ の測定結果は約 90ms であり、 $\overline{\Delta}_S^{(2)}$ および $\overline{\Delta}_S^{(3)}$ は約 180ms であった。つまり、視聴者の端末では、先に指揮者のビデオと歌声が出力され、約 90ms 後に 2 人の歌手の歌声が同時に出力される。前述のように、図 4 の付加遅延 1 の平均が 50ms, 付加遅延 5 の平均が 300ms で端末間同期制御を行わない場合にも、2 人の歌手の歌声が同時に出力される共通点があるにも関わらず、二つの場合で MOS は大きく異なっている。この違いについて調査するため、二つの場合でのユーザ体感品質の比較を行った。比較では、各被験者は二つの場合の合唱をランダムな順序で聴いた。そして、より同期して感じたのはどちらかを選択し、指揮者のビデオのずれによる印象の違いなど、気付いたことを書いた。20 人の被験者による評価の結果、18 人の被験者が付加遅延 5 の平均が 300ms の場合の方が同期して感じると答えた。また、付加遅延 5 の平均が 300ms の場合の方が良いと答えた被験者の主な印象の違いなどをまとめると、付加遅延 5 の平均が 50ms の場合では、指揮者の動きと歌手の歌声が揃っていないと感じた (7 人が回答)、付加遅延 5 の平均が 300ms の場合では、ずれていると感じなかった (5 人)、付加遅延 5 の平均が 50ms の場合では、指揮者のみが先行しているため、違和感を感じた (4 人) となった。よって、指揮者のビデオと歌声が歌手の歌声よりも先行する場合の方が、歌手の歌声が先行する場合よりもずれが気になることがわかった。

この他、図 4 において、付加遅延 5 の平均が 300ms, 付加遅延 1 の平均が約 250ms 以上で端末間同期制御を行わない場合に MOS 値が大きく減少する傾向が見られる。この

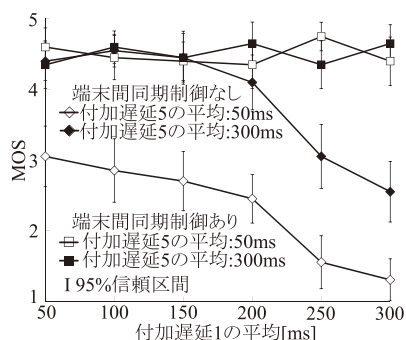


図 4 付加遅延 1 の平均に対する合唱のメディア同期品質の MOS
MOS of media synchronization quality in chorus versus average of additional delay 1.

理由は次の通りである。付加遅延 1 の平均が約 250ms 以上の場合では、 $\Delta_S^{(1)}$ より $\Delta_S^{(2)}$ の方が大きくなる。そのため、最も早い端末 3 の歌い手の歌声が聴こえてから、最も遅い端末 2 の歌い手の歌声が聴こえるまでの時間がより長くなり、メディア同期品質が損われている。

今回の結果から、指揮者のビデオと歌声が歌い手の歌声より先行するか否かなどで、ユーザ体感品質が異なることがわかった。よって、メディア同期品質には、コンテンツ依存性があると考えられる。そのことを詳しく調査するため、本論文と異なるテンポの合唱¹⁵⁾、異なる楽曲の合唱、楽器による合奏に対しても本論文と同様の評価を行った。その結果、端末間同期制御により、ユーザ体感品質を高く維持できるという点で同様の結果が得られた。しかし、端末間同期制御を行わない場合、テンポや楽曲、音の種類等、コンテンツの違いにより、メディア同期品質のユーザ体感品質は異なることがわかった。よって、ユーザ体感品質評価を行う上で、コンテンツ依存性は重要な観点である。

5. む す び

本論文では、多地点間において音声とビデオを用いて合唱を行う場合を扱い、視聴者のいる地点で多地点からの歌を同期させるために端末間同期制御を適用した。そして、ユーザ体感品質評価により、端末間同期制御の効果を調査した。その結果、端末間同期制御を行うことで、その制御を行わない場合に比べ、合唱のメディア同期品質を高く維持できることを示した。

今後の課題としては、多様なネットワーク環境で端末間同期制御の効果を調べることが挙げられる。また、メディア同期品質のコンテンツ依存性を調べるため、多様な楽曲やテンポで本論文と同様の調査を行うことも重要である。

【文 献】

- 1) 栗田 孝昭, “多地点間映像音声通信サービスにおける会話立場の違いが総合品質へ及ぼす影響”, 信学技報, CQ2005-32 (July 2005)
- 2) F. Gong, “Multipoint audio and video control for packet-based multimedia conferencing”, in *Proc. ACM Multimedia'96*, pp. 308-314 (April 1996)
- 3) Y. Ishibashi, M. Nagasaka, and N. Fujiyoshi, “Subjective assessment of fairness among users in multipoint communications”, in *Proc. ACM SIGCHI ACE'06* (June 2006)

- 4) Y. Ishibashi and S. Tasaka, “A group synchronization mechanism for live media in multicast communications”, in *Conf. Rec. IEEE GLOBECOM'97*, pp. 746-752 (Apr. 1997)
- 5) 細谷 和輝, 石橋 豊, 菅原 真司, “音声・ビデオ通信における端末間同期制御の効果”, 映像学誌, **62**, 12, pp. 2053-2058 (Dec. 2008)
- 6) 阿部 威郎, 石橋 豊, 吉野 秀明, “次世代のサービス品質技術動向”, 信学誌, **91**, 2, pp. 82-86 (Feb. 2008)
- 7) 山北 淳, 古川 賢信, 須賀川 豊, 渡辺 義典, “TLA(タイムラグアジャスター) が結んだ 5 大陸 長野オリンピック開会式「五大大陸 第九合唱」”, 映像学技報, **22**, 25, pp. 43-48 (May 1998)
- 8) M. Yoshida, Y. Obu, and T. Yonekura, “A protocol for real-time remote musical session”, *IEICE Trans. Inf.&Syst.*, **88**, 5, pp. 919-925 (May 2005)
- 9) O. Nakamura, K. Sugiura, S. Yamamoto, N. Shigechika, A. Kato, K. Hasebe, and J. Murai, “Internet Metronome: An experimental remote jazz session with uncompressed HDTV transmission over lightpaths”, *IEICE Trans. Commun.*, **89**, 4, pp. 1052-1058 (Apr. 2006)
- 10) J. P. Caceres, R. Hamilton, D. Iyer, C. Chafe, and G. Wang, “To the edge with China: Explorations in network performance”, in *Proc. the 4th International Conference on Digital Arts*, pp. 61-66 (Nov. 2008)
- 11) Y. Ishibashi and S. Tasaka, “A synchronization mechanism for continuous media in multimedia communications”, in *Proc. IEEE INFOCOM'95*, pp. 1010-1019 (Apr. 1995)
- 12) D. L. Mills, “Internet time synchronization: The network time protocol”, *IEEE Trans. Commun.*, **39**, 10, pp. 1482-1493 (Oct. 1991)
- 13) M. Carson and D. Santay, “NIST Net - A Linux-based network emulation tool”, *ACM SIGCOMM Computer Commun. Review*, **33**, 3, pp. 111-126 (July 2003)
- 14) ITU-R BT.500-11, “Methodology for the subjective assessment of the quality of television picture”, International Telecommunication Union (June 2002)
- 15) 宮下裕司, 石橋豊, 福嶋慶繁, 菅原真司, K. E. Psannis, “多地点間における音声・ビデオを用いた合唱の QoE 評価”, 信学技報, CS2010-51 (Dec. 2010)



みやした ゆうじ
宮下 裕司 2010 年, 名古屋工業大学工学部情報工学科卒業。2012 年, 同大学院工学研究科創成シミュレーション工学専攻博士前期課程修了。在学中, 音声・ビデオを用いた多地点間通信における QoS 制御の研究に従事。



いしばし ゆたか
石橋 豊 1981 年, 名古屋工業大学工学部情報工学科卒業。1983 年, 同大学大学院修士課程修了。同年, 日本電信電話公社入社。NTT ヒューマンインタフェース研究所主任研究員を経て, 1993 年より名古屋工業大学工学部電気情報工学科助教授。現在, 同大学大学院教授。分散マルチメディアの研究に従事。工学博士。正会員。



ふくしま のりしげ
福嶋 慶繁 2004 年, 名古屋大学工学部電気電子情報工学科卒業。2009 年, 同大学大学院工学研究科電子情報システム専攻博士課程後期課程修了。同年, 名古屋工業大学大学院助教授。多視点映像処理の研究に従事。博士(工学)。正会員。



すがわら しんじ
菅原 真司 1994 年, 東京工業大学工学部電気電子工学科卒業。1996 年, 同大学大学院修士課程修了。1999 年, 同大学大学院博士課程修了。同年, 電気通信大学電気通信学部情報通信工学科助手。2005 年, 名古屋工業大学大学院助教授。現在, 同大学大学院准教授。分散データベース, 情報検索型通信の研究に従事。博士(工学)。



コスタス プサニス From 2004, Kostas E. Psannis served as a Visiting Assistant Professor in the Department of Technology Management, University of Macedonia, Greece, where currently he is a Lecturer. He has extensive research, development and consulting experience in the area of telecommunications technologies. He received a degree in Physics from Aristotle University of Thessaloniki (Greece), and the Ph.D. degree from the Department of Electronic and Computer Engineering of Brunel University (UK).