

ローカル・ラグ制御機能を持つ 音響サーバを用いた遠隔合奏の評価

前田佳奈[†] 竹森幸輝[†] 岩原正典[†]
片桐滋[†] 大崎美穂[†]

高速デジタル・ネットワークの普及に伴い、様々なジャンルの芸術表現がネットワーク上で行われるようになってきた。中でも、遠隔合奏の研究が多くの関心を集めている。しかし、ネットワーク遅延が合奏に及ぼす影響は大きく、満足に演奏できる環境は必ずしも整っていない。本研究は、遠隔地に生じる遅延と同量の遅延をローカル・サイトの演奏者へのフィードバック音に加えるローカル・ラグ制御機能と、遠隔地サーバ間で数ミリ秒単位の同期をとることができる時刻同期機能を備えた音響サーバを開発し、それをを用いて遠隔合奏の実験を行い、開発したサーバの性能の分析とローカル・ラグ制御の有用性の確認を行うものである。

Evaluation of Remote Ensemble using Acoustic Server Having Local-lag Control Function

Kana Maeda[†] Koki Takemori[†] Masanori Iwahara[†]
Shigeru Katagiri[†] Miho Osaki[†]

Along with the popularization of high-speed digital networks, various sorts of artistic activities have been performed over networks. In particular, remote ensembles have attracted attention and been vigorously studied by many IT researchers and musicians. However, the network communication inevitably suffers from time lag, and such lag often degrades the quality of remote ensembles. To alleviate this lag-based problem, we develop in the present study a new server system that implements a local lag technique, which aims at

reducing the defect of lag by adding the same amount of lag as that for a remote-site player to the sound output of a local-site player, with millisecond-level synchronization between player sites. Through remote ensemble experiments, we report its performance in terms of synchronization/local-lag-control accuracy, and demonstrate its dependability in a real operation.

1. はじめに

デジタル・ネットワークの高速化が進み、音・映像配信の技術が向上する中で、ネットワークを介した遠隔合奏に関する様々なアプローチによる研究が行われている[1][2]。しかし、演奏者間で同期をとることが必要となる遠隔合奏における遅延（ネットワークに起因する伝送遅延やコンピュータ内の計算処理遅延）の影響は大きく、同室における合奏に近い演奏環境を作り出すことは未だに困難な課題となっている。

そうした中で、ネットワーク遅延対策の1つのアプローチとしてローカル・ラグ法と呼ばれる制御法を遠隔合奏に適用することが試みられた。ローカル・ラグとは、同時的コラボレーションを行う際に、作業者自身に対するフィードバックを一定時間だけ遅らせることで同期を図る方法である[3]。遠隔合奏では、遠隔地の演奏者同士が同時刻に同じ音環境を共有できるようになり、遅延が演奏に及ぼす影響を軽減できることが報告されている[4]。

しかし、文献[4]で用いられた遠隔合奏支援システムには、音データを処理する遠隔地のサーバどうしの同期がとられていない、さらには録音ファイルの開始時刻の同期がとられていないなど、幾つかの要改善点があった。録音ファイルが非同期であったことは、演奏結果を分析する際に手作業で録音データの同期をとる必要があるなど、分析をやや困難なものとする原因となっていた。従って、時刻同期などを自動的にとることができる実験用の遠隔合奏支援システムの開発が望まれるところである。

上記の背景に基づき、我々は、時刻同期機能や録音ファイルの開始時刻の同期機能を持つ新しい遠隔合奏支援システムの開発をするに至った[5]。本稿は、その新しいシステムを実際に遠隔合奏実験に用い、その時刻同期やローカル・ラグ制御の精度と遠隔合奏に対するシステムの効果の分析結果を報告するものである。

[†] 同志社大学大学院
Graduate School of Engineering, Doshisha University

2. ローカル・ラグ制御を用いる遠隔合奏支援

2.1 ローカル・ラグ制御法

遠隔地の演奏者間に横たわる遅延は主に、ネットワーク伝送に起因する遅延や音データを処理するコンピュータ内部 (AD/DA 処理なども含む) の計算処理遅延から成る。以下では、特に断らない限り、これらの種々の遅延の総体を遅延と呼ぶこととする。ローカル・ラグ制御法とは、遅延そのものを短縮するアプローチではなく、作業 (遠隔合奏の場合は演奏者) の行動に対する遅延の (悪) 影響を軽減するためのアプローチである。図 1 に、遠隔合奏への適用を例としたローカル・ラグ制御法の原理を図解する。簡単のため、ここでは各地点内における、計算機内部の処理に伴う遅延はないものとする。図中、地点 A の演奏者の演奏音はネットワークを経由して地点 B の演奏者の元に送られる。この時、地点 A の演奏者が聴く演奏音をフィードバック音と、地点 B の演奏者に届けられる演奏音とフィードスルー音とする。フィードバック音は、本来、演奏と同時に地点 A の演奏者の耳に届き、フィードスルー音はネットワーク遅延を伴って遅れて地点 B の演奏者の耳元に届く。ローカル・ラグ制御法とは、このネットワーク遅延と同量の遅延をフィードバック音に加え、演奏者がともに同時に同じ音を知覚できるようにするものである。演奏者は自身の演奏音をローカル・ラグ・フィードバック音として遅れて聴くことになるが、合奏者どうしの知覚に同期をとることで合奏の質の向上を目指すものである。

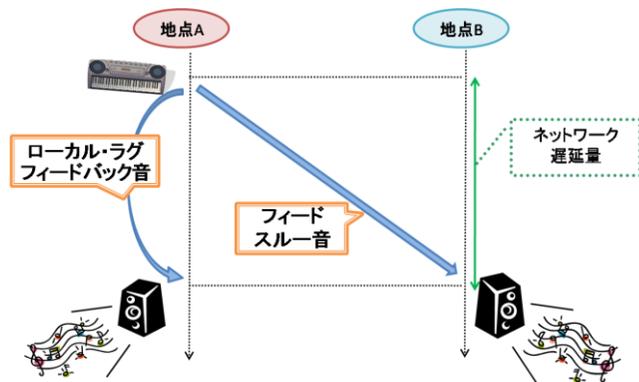


図 1 ローカル・ラグ制御法の原理。

2.2 遠隔合奏におけるローカル・ラグ制御の効果

著者らもかかわった文献[4]の実験においては、10msec から 100msec までのネットワーク遅延量を制御条件として、電子キーボードを用いた遠隔合奏実験を行い、ローカル・ラグ制御法が遠隔合奏の質の向上に一定の効果を持つことが示されている。し

かし、ネットワーク遅延あるいはそれに対応するローカル・ラグ量が一定値以上になると、キーボードの打鍵に対する音の再生があまりに遅れることにより発音の予測が困難となり、結果的に演奏も困難となることも明らかにされている。

3. 新しい音響サーバ

3.1 諸元及び遠隔合奏実験向けの実装

新しい音響サーバは、文献[5]に詳述されている新しい遠隔合奏支援システムを構成するものである。先行研究[4]で用いられたサーバと異なり、特に、遠隔地間の音響サーバどうして時刻を同期させる機能と録音ファイルのデータ開始時刻を合わせる機能を持っている。

遠隔合奏に本音響サーバを適用するに当たり、本稿では、そのローカル・ラグ制御能力の実験的な評価を行った。その評価にも影響する、本音響サーバの基本的な諸元を以下にまとめる。サーバは Linux を OS とするコンピュータ上に構築されている。

- CPU : Intel(R) Core(TM)2 Quad CPU (2.40 GHz)
- OS : Fedora14
- サウンドカード: Creative Sound Blaster 5.1
- サンプリングレート: 44.1 kHz
- 量子化ビット数: 16bit
- チャンネル数: 1 (モノラル)
- サウンドドライバ: ALSA (I/O ライブラリとして PortAudio を使用)
- 演奏用楽器とシステムの接続方法: PHONE 端子から音を出力

また、遠隔合奏実験に向けて、2 地点間の遠隔合奏を行う想定の下で遠隔合奏支援システムを実装した。この時、音響サーバを支援する同期サーバは、2 台のうちの 1 台の音響サーバと同じコンピュータ (ローカル・サイトに相当) 上に実装し、もう 1 台の音響サーバは別の (遠隔地に相当) コンピュータ上に実装する。いずれのコンピュータも上記の諸元を持つものである。また、時刻同期は文献[5]における、独自に導入したタイマに基づく手法で実現した。コンピュータのマルチコアの制御は特に行わなかった。

3.2 ローカル・ラグ制御精度の評価

遠隔合奏の実験に先立ち、上述のように実装された 2 式の音響サーバ間で遠隔合奏をする際に施すローカル・ラグ制御の、実際に達成される精度を計測した。

3.2.1 精度の計測方法

多チャンネル録音を可能とするオーディオインターフェースを用い、入力音源 (電子メトロノーム) の直接出力音データと音響サーバを介して出力された録音データの比較を行った。録音比較対象は下記の 3 つとし、それぞれの波形をサウンド編集ソフト

トを用いて視覚的に拡大表示し、目視で計測した。

- 入力音源からの直接出力音。
- 入力音源音を音響サーバに入力し、ローカル・ラグ制御をされたフィードバック音。
- ローカル地点で入力された後に遠隔地の音響サーバに送信されたフィードスルー音。

ネットワーク遅延量とそれに対応するローカル・ラグ量を 0ms~100ms の 10ms 刻みで設定した。電子メトロノームは $J=60$ に設定し、各制御遅延量で誤差（時間的ずれ）の計測を 10 試行を行った。実験は LAN 環境で行い、ネットワーク遅延はその環境中に別途構築した FreeBSD の Dummynet を用いて制御した。

3.2.1 計測結果

フィードバック音もフィードスルー音も、ネットワーク遅延とは別に、音響サーバ内の計算処理に伴う遅延を避けることはできない。そこでまず、ローカル・ラグ制御を伴わない条件（ネットワーク遅延がゼロ）下における音響サーバの計算処理に伴う遅延の影響に着目した。電子メトロノームからの直接出力音とフィードバック音及びフィードスルー音の誤差の平均値と標準偏差を 10 試行の結果から算出したものを表 1 に示す。

表 1 入力音源からの直接出力音とフィードバック音及びフィードスルー音の誤差。

	平均値(ms)	標準偏差(ms)
フィードバック音と直接出力音の誤差	51.19	0.011
フィードスルー音と直接出力音の誤差	55.42	0.060

表 1 より、音響サーバに音響信号が入力されてからフィードバック音が出力されるまでには 50ms 程度の時間がかかっていることが示唆された。この誤差は、サーバ入出力に伴う AD/DA 変換及び、サーバ内計算処理に起因すると考えられる。サーバ間で送信処理が行われるフィードスルー音では、フィードバック音に比べて数ミリ秒だけ誤差が大きくなっていた。これは送信に伴う処理時間に依るものと考えられる。また、フィードバック音もフィードスルー音も、その誤差計測に伴う標準偏差は小さく、計算処理は比較的安定して行われている様子が伺える。

次に、遅延を加えた場合についてまとめる。ここでは、表 1 の安定した結果を受けて、電子メトロノームからの直接出力音とフィードバック音の誤差に着目して計測を行った。ここでも試行数は 10 である。10 試行の誤差の平均値を図 2 に示す。なお、図 2 では、上述の AD/DA 変換及び、システム内処理に起因する遅れ（表 1 中の平均値）を取り除いて、ローカル・ラグ制御に伴う処理に起因する誤差に焦点を合わせた

表示を行った。

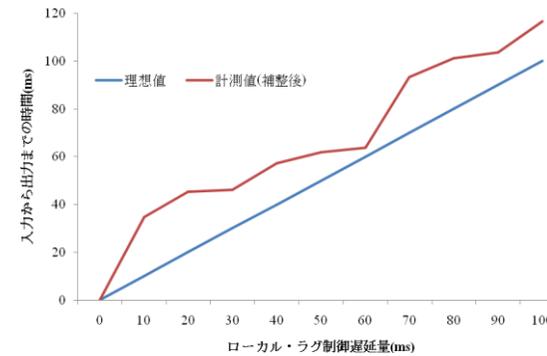


図 2 入力音源からの直接出力音とフィードバック音の誤差の平均値。

図 2 から、ローカル・ラグ制御には、理想値から最大 20ms 程度のばらつきを伴うことが示された。この誤差の原因として、時刻同期の基となるタイマの変動[5]などが考えられるが、より詳細の分析が必要と思われる。

計測の結果、遠隔合奏実験において、遅延量の差異が 20msec を下回るような制御条件どうしの結果の比較は望ましくないと判断される。従って、以下の遠隔合奏の実験では、ローカル・ラグ制御遅延量に 30ms 以上の差がある場合どうしの実験結果に着目して考察を行うことにする。

4. 遠隔合奏実験

4.1 概要と目的

実験では、次世代テレビ会議システムとされている遠隔コラボレーション支援システム「t-Room」を 2 部屋用い、各部屋に音響サーバと電子キーボードをそれぞれ 1 台ずつ設置することで、2 地点間での遠隔合奏実験を行った。なお、1 台の音響サーバには同期制御サーバも合わせて実装した。また、ネットワーク遅延の遠隔合奏への影響とローカル・ラグ制御適用効果を調査するため、ネットワーク遅延量及びローカル・ラグ遅延量の制御を行った。この実験は、新しい遠隔合奏支援システムの動作確認及び、より正確な演奏精度計測によるローカル・ラグ制御効果の再検討を目的としたものである。

4.2 実験環境

4.2.1 空間環境

実験は、t-Room を 2 部屋用いて行った。t-Room は 8 角形を形づくっており、6 面のディスプレイ壁面から構成されている。スピーカは演奏者の前面に設置し、演奏者のフィードバック音および演奏相手のフィードスルー音全てを 1 つのスピーカから出力した。実験時の全体構成図を図 3 に示す。

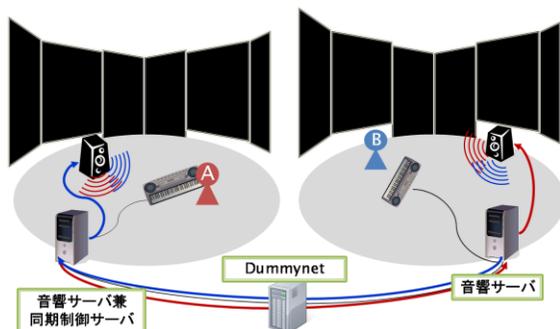


図 3 実験時の全体構成図。

実験中、t-Room で使用することのできる、遠隔地の映像情報は表示しなかった。これは、現在のシステム構成上、音と映像の同期がとれておらず、音と映像の遅延ずれによる演奏への影響が生じかねないためである。

4.2.2 音環境

実験に使用した電子キーボード設定の詳細を示す。

- 使用楽器： YAMAHA EOS B700, YAMAHA PSR-330 各 1 台
- 使用音： グランドピアノ音
- 電子キーボード設定： タッチレスポンス機能 ON
- 音圧： 70~75dB SPL

電子キーボードは、打鍵速度によって音圧が変化する、タッチレスポンス機能を ON にして用いた。音圧は騒音計を用い、実験に用いた課題を演奏して計測した。

4.3 実験参加者

プライベートでの鍵盤楽器演奏経験歴が 10 年以上の大学生 8 名が実験に参加し、2 名 1 組で演奏を行った。以後、便宜的に各 2 名を演奏者 A・演奏者 B と記す。経験豊富な実験参加者を用いた理由は 2 つある。1 つめは、鍵盤楽器に不慣れなために起こる演奏破綻を回避するためであり、2 つめは、熟達者ほど運動学習におけるフィードバック制御への依存度が低いためである[6]。

4.4 実験内容

4.4.1 呈示課題

先行研究に倣い、「かえるのうた」の演奏を課した。図 4 に呈示課題を示す。



図 4 呈示課題。

課題は演奏者それぞれが、片手で演奏するものとし、演奏者 A は、図 4 の青枠部分を、演奏者 B は赤枠部分を演奏した。

テンポ指示として、演奏者 A には演奏前に $J=100$ で電子メトロノームが呈示され、大凡のテンポをつかんだ上で演奏を行った。演奏中はメトロノームは呈示されなかった。課題は 4 分音符と 8 分音符のみで構成され、打鍵間隔は、4 分音符で平均 600ms、8 分音符で平均 300ms である。

4.4.2 遅延量制御パターン

演奏は 2 パターンの遅延量制御下で行い、ネットワーク遅延量のみを制御したパターンを NORMAL モード、ネットワーク遅延量とローカル・ラグ遅延量を同値に設定したパターンを LOCALLAG モードとした。各モードでは、遅延量が 0ms から 100ms までの 10ms 刻みで設定された。課題の演奏は各 11 試行とし、計 22 試行であった。なお、遅延量の設定は昇降順等での慣れの影響を避けるため、ランダム呈示にされた。

4.4.3 実験手順

実験参加者はまず、事前練習として、実験本番用の課題を各々で練習した後、ネットワーク遅延 0ms、ローカル・ラグ遅延 0ms の環境下で 3 回程度合奏練習を行った。その後、本実験へと移った。

なお、実験参加者に対し、ローカル・ラグ制御技術についての詳しい説明は行われなかった。これは、将来的にローカル・ラグ制御機能を持つ遠隔合奏支援システムが実用化される場合を考慮し、実験結果からその実現可能性について検討するためである。

実験参加者にはアンケートが課され、各試行を終える毎に回答した。NORMAL モードの 11 試行演奏後は休憩を取り、その後 LOCALLAG モードの演奏へと移った。

4.5 評価項目

4.5.1 定性評価

演奏者の自身の演奏に対する評価を調査するためアンケートを行った。アンケートには「相手との同期」を評価する5段階評価項目と自由記述欄を設けた。

4.5.2 定量評価

計測地点で、演奏のフィードバック音と遠隔地から届くフィードスルー音を別ファイルに録音した。これらの2つのファイルは、自動的に開始タイミングの同期が取れており、時間軸方向の調整操作なしに比較できるものである。波形解析には、フリーのサウンド編集ソフト「audacity」を用い、フィードバック音とフィードスルー音の立ち上がり時間を比較して、演奏者間での音のずれを算出した。計測には、課題中で演奏者Aと演奏者Bの演奏音が同時に聞こえることが望まれる表拍、計16音を用い、そのずれは、演奏者Aが先行した場合を+(プラス)、演奏者Bが先行した場合を-(マイナス)で示した。

4.6 実験結果

4.6.1 定性評価

演奏相手との同期の主観評価結果を図5に示す。横軸はNORMALモードでのネットワーク遅延量、LOCALLAGモードでのネットワーク遅延量及びローカル・ラグ遅延量であり、縦軸は実験参加者8名による5段階評価の平均値と標準偏差を表す。どの遅延量においても、NORMALモードでの演奏に比べてLOCALLAGモードでの相手との同期の感覚があったことがわかる。特に、遅延量100msでは、NORMALモードでは評価平均1.3であったが、LOCALLAGモードでは平均3.8と高い値を示した。それに比べ、NORMALモードではネットワーク遅延30msから、評価の平均が下がり続けた。

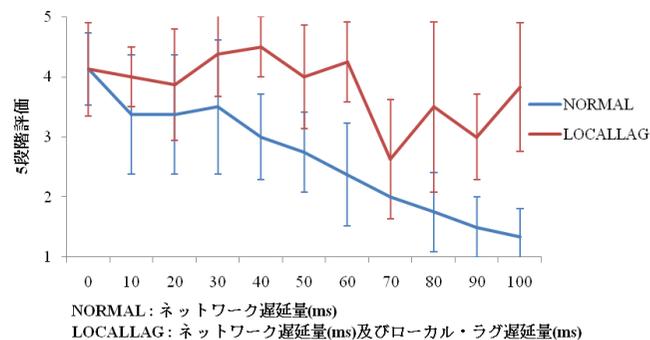


図5 演奏相手との同期の主観評価。

4.6.2 定量評価

録音ファイルから解析した演奏者間での音のずれを評価した。まず、ずれの絶対値の平均を実験参加者8名分で平均した結果について簡単に述べる。ネットワーク遅延量30ms以上では、NORMALモードでの音のずれがLOCALLAGモードでの音のずれの2倍近くになった。ネットワーク遅延量70ms以上では、ローカル・ラグ制御を行わない場合、音のずれが100msを超えた。

次に、例として実験参加者1組分の時間軸に沿った演奏音ごとの演奏者間の音のずれを示す。図6はNORMALモードでの演奏者間の音のずれであり、図7はLOCALLAGモードでのずれを示す。

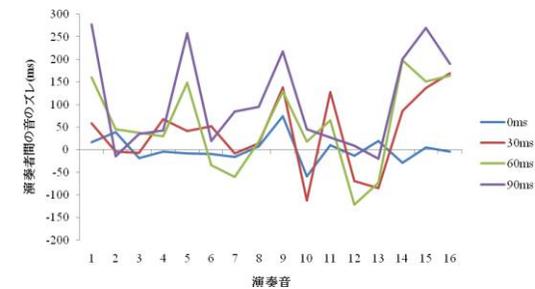


図6 NORMALモードでの演奏者間の音のずれ。

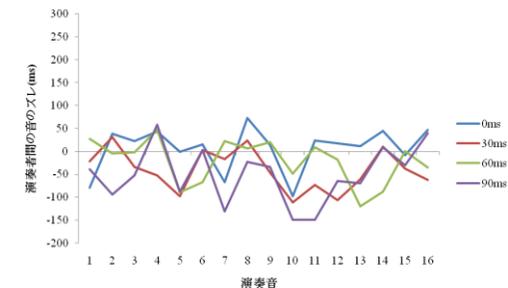


図7 LOCALLAGモードでの演奏者間の音のずれ。

横軸は音のずれの計測に用いた16音であり、縦軸はその音ごとの演奏者間での音のずれの大きさである。NORMALモードのネットワーク遅延0msの場合は、演奏音は0ms付近で前後した。ネットワーク遅延30msでは、課題の前半部分では安定していたが、後半部で不安定になった。なお、後半部での演奏者Aの演奏パートは旋律パー

トであり、ネットワーク遅延を考慮しても、最終小節で大幅に先行したことが分かる。ネットワーク遅延 60ms と 90ms ではずれの大きな揺れが見られ、演奏者間でテンポを維持できなくなったこと、ずれを修正しようとしていることが分かる。LOCALLAG モードでは、どの試行においても、-100ms から +50ms 付近での前後の揺れが見られた。遅延量の増加による演奏者間のずれは、ネットワーク遅延を考慮しても、NORMAL モードに比べて全体的に小さい。

4.7 考察

演奏相手との同期の主観評価においては、NORMAL モードに比べ、LOCALLAG モードでの評価が大幅に高かった。これは、ローカル・ラグ制御によって、ネットワーク遅延の演奏精度低下が改善されたためであると言えるであろう。ローカル・ラグ制御時の良好な演奏状態としては 2 状態が考えられる。1 つは、ローカル・ラグ制御に気がつかずに演奏できる状態であり、もう 1 つはローカル・ラグ制御によるフィードバック音の遅れに気がつきながらも上手く演奏できるという状態である。アンケートの自由記述部の結果によると、遅延が 70ms 程度になると多くの演奏者から発音の遅れに気がつき、同期の評価も下がった。しかし、その後は大幅に評価を下げることはなく、ローカル・ラグ制御が効果を発揮したことが図 5 から見てとれる。

録音ファイルの波形解析結果にもネットワーク遅延による演奏破綻改善が見られた。演奏者間での音のずれの絶対値の平均値は、NORMAL モードではネットワーク遅延が 70ms 以上になると音のずれが 100ms 以上になっていた。演奏音が 8 分音符の場合は打鍵間隔が平均 300ms であるのでその 3 分の 1 以上に相当し、大幅なずれであることが分かる。一方、LOCALLAG モードでは、ずれが 50ms 以下と、ローカル・ラグ制御が大いに貢献した。

実験参加者 1 組分の時間軸に沿った演奏者間での音のずれからは、ある音では相手より先行し、次の音では相手の音の後に演奏するという、演奏者間での特徴的なずれが見られた。特に NORMAL モードでネットワーク遅延量が大きくなるほどその傾向は顕著であった。LOCALLAG モードでは遅延量に関わらず、前後 50ms 程度のずれであった。これらのことから考察すると、NORMAL モードで 100ms 近く演奏者間のずれが前後している場合は、互いの演奏者のネットワーク遅延によって遅れた相手の音に合わせるための心掛けが、前後のずれの連鎖を生じさせたと考えられる。

最後に新たなシステムの動作確認について述べる。遠隔合奏実験は、予備実験も含め 5 組 10 名で行った。実験中は、どの課題においても問題なくシステムを運用することができた。このことより、新たなシステムが、遠隔合奏支援システムとして十分に動作すると言えるであろう。また、この安定した動作は、実験を監督した筆者も実際に確認した。

5. まとめと今後の展望

新しい音響サーバを用いて実際の遠隔合奏実験を行ったことにより、構築されたシステムはネットワーク遅延対策法を備えた遠隔合奏支援システムとして十分に効果を発揮することが確認された。また、システムに搭載されたローカル・ラグ制御機能は、精度計測により音の送信状況が明らかにされ、その信頼できる制御遅延の精度も明らかにになった。しかし、達成できた遅延制御量に約 20msec の誤差を伴っており、今後、この軽減をする必要がある。

ネットワーク遅延及びローカル・ラグ制御遅延の遠隔合奏への影響は、様々な要因によるものであると言える。そこで今後、ローカル・ラグ制御のさらなる可能性を追究するため、種々の呈示課題や演奏楽器を用いた演奏実験を行うことが望まれる。

また、音響情報と同期のとれた映像情報を提供することで、より一体感の高い遠隔合奏を実現することも期待される。

謝辞 音響サーバの開発に際して多くのご助力を賜った山口毅氏と、実験に御協力頂きました皆様に感謝致します。

参考文献

- 1) Juan-Pablo Caceres, Chris Chafe; JACKTRIP/SOUNDWIRE MEETS SERVER FARM: Proceedings of the SMC 2009 - 6th Sound and Music Computing Conference, 23-25 (2009. 07).
- 2) 岸田崇志, 前田香織, 河野英太郎, 近堂徹, 相原玲二; 多様な遠隔コラボレーションを実現する音声伝送システム(ネットワークサービス)(<特集>ブロードバンドネットワークサービス): 情報処理学会論文誌 45 (2), 517-525, 2004-02-15 (2004).
- 3) Dane Stuckel and Carl Gutwin; The Effects of locallag on Tightly-Coupled Groupware : Computer Supported Cooperative Work, 447-456 (2008).
- 4) 入江洋介, 青柳滋己, 高田敏弘, 平田圭二, 梶克彦, 片桐滋, 大崎美穂; t-Room のための遠隔合奏支援システムの構築: 情報処理学会グループウェアとネットワークサービス研究会研究報告 Vol.2009-GN-73 No.23 (2009).
- 5) 竹森幸輝, 前田佳奈, 岩原正典, 片桐滋, 大崎美穂; ローカル・ラグ制御機能とログ同期機能を持つ音響サーバの開発, オーディオビジュアル複合情報処理研究会 (2012. 02).
- 6) 高橋範行; ピアノ演奏における熟達 -演奏解釈と聴覚フィードバック利用に関して-:平成 18 年度 京都市立芸術大学博士論文 (2007).