

G-20

ローカル・ラグ制御機能を持った音響サーバの性能評価実験

Experimental Evaluation of Acoustic Server Having Local-lag Control

大島 義博† 片桐 滋† 大崎 美穂†

Yoshihiro Oshima Shigeru Katagiri Miho Ohsaki

1. はじめに

ネットワークを利用した遠隔コラボレーション支援システム“t-Room”の研究が行われている[1]。円滑な共同作業を実現する上で考慮すべき問題点に、ネットワーク上で発生する遅延が挙げられ、この遅延の存在は遠隔地点間における円滑な作業を行う上での障害になる恐れがある。我々は、この解決策として通信経路上で不可避の遅延（ラグ）の存在を前提としたローカル・ラグ法[2]に着目し、音メディアを中心とした音響サーバの開発と評価実験を行ってきた。ローカル・ラグ法とはリモート側の相手と通信を行う際に生じるラグをローカル側に同量のラグを加える事で遠隔地点間の同期性を保つ手法である。これまで進めてきた[3]から[6]における先行研究においては、主に次の2点、即ち①音響サーバの開発およびその性能評価と、②ローカル・ラグ制御の有用性の検証に焦点を合わせた研究を行ってきたが、実際の利用場面を想定した評価実験や検証はまだ十分ではなかった。本稿は、こうした点を背景として、①Wide Area Network (WAN) 環境下での性能評価実験と、②遠隔合奏や遠隔合唱（カラオケなども含む）などの利用シーンにおける音響サーバの有用性を示すために多地点間の例として3地点間での音響サーバの性能評価実験、③実際の利用状況における動的な音響サーバの処理遅延やネットワーク上の遅延対策としてリセット機能の追加と評価を行い、それらの結果を報告するものである。

2. ローカル・ラグの仕組み

本節ではローカル・ラグの仕組みを、遠隔合奏を例として紹介する。一般的な合奏では、演奏者は自身の演奏音と相手から届けられる演奏音の2種類の音を知覚する。本稿では自身の演奏音をフィードバック音、相手側からの演奏音をフィードスルー音と呼ぶこととする。離れた2地点間における遠隔合奏をネットワーク上で行った際の様子を図2.1に示す。図2.1の縦軸は時間の経過を表している。

図2.1において通信遅延量が十分に小さい場合、演奏者達は遅延を知覚しにくいと考えられる。そのため演奏者間の同期は保たれるため合奏が破綻する可能性は低いと思われる。しかしながら、通信遅延量が十分に大きい場合だと、演奏者達は遅延を知覚するため演奏者間の同期は損なわれ、合奏を続ける事が難しくなると考えられる。

この問題に対して我々は、ローカル・ラグ法を取り入れた音響サーバの開発・研究を行ってきた。ローカル・ラグ法とは遠隔地にいる利用者間で発生する遅延と同程度の遅延をローカル側に付加する事で相手との同期を保つ手法である。つまり図2.2より自身のフィードバック音に通信遅延量と同程度の遅延量を付加することで、音データの出力タイミングを整え演奏者間の同期を確保し合奏の破綻を防止するというものである。

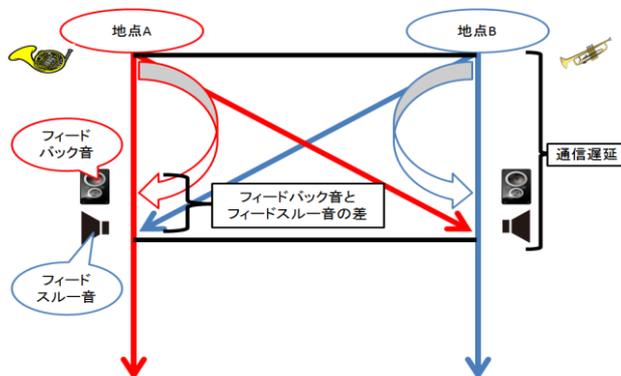


図2.1 ネットワーク経由の遠隔合奏における音データの流れ。

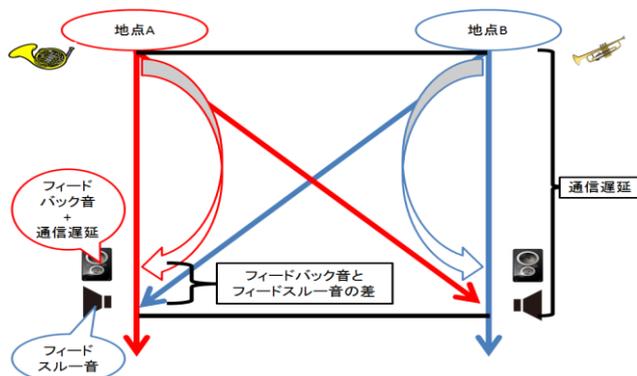


図2.2 ネットワーク経由のローカル・ラグを適用した遠隔合奏における音データの流れ。

3. ローカル・ラグ制御機能を持つ音響サーバ

3.1 同期制御サーバと音響サーバの機能概要

この節では同期制御サーバと音響サーバの機能について述べる。我々が開発してきた音響サーバを運用するには同期制御サーバが必要となる。同期制御サーバとは音響サーバ間の接続補助、音響サーバ間の通信遅延測定機能と音響サーバ間の最大遅延量の決定及び配信機能が実装されているサーバである。一方、音響サーバには音声出力機能とローカル・ラグ制御機能が実装されており、同期制御サーバの情報と音響サーバの機能によりローカル・ラグ制御を行う事が可能となる。

3.2 遅延測定の仕組み

音響サーバはローカル・ラグ制御を行うために、音データの通信を開始する前に通信遅延の計測を行っている。その測定方法について以下の図3.1に示す。図の縦軸は時間軸を表しており、音響サーバAと音響サーバB間の通信遅延を測定する場合を考える。音響サーバAが音データサイズ分のデータを送信した時刻を t_0 、音響サーバB

†同志社大学大学院 理工学研究科

が音響サーバ A から音データを受信した時刻を t_1 、音響サーバ B が音響サーバ A に音データを返信する時刻を t_2 、音響サーバ A が音響サーバ B から音データを受信した時刻を t_3 とする。往復遅延時間は $t_3 - t_0$ で表わされ、応答までの時間は $t_2 - t_1$ で表わされる。往復遅延量は通信遅延量と応答までの時間と双方向の通信遅延量と同量であることから

$$\text{通信遅延} = ((t_3 - t_0) - (t_2 - t_1)) / 2 \quad (1)$$

の計算により遅延計測を行っている。

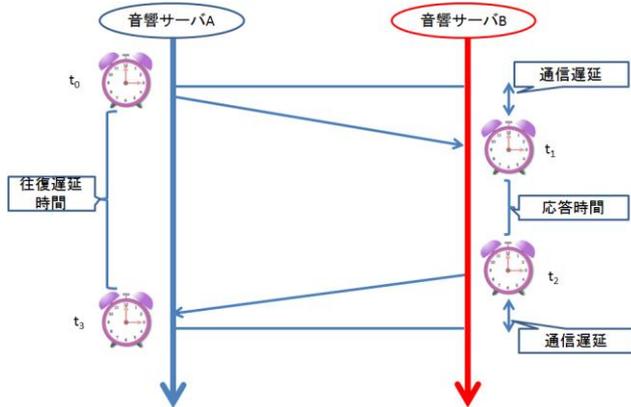


図 3.1 遅延の計測方法。

3.3 同期制御サーバの同期手順

全地点の同期を保つために同期制御サーバは地点間の遅延を測定し地点間の最大遅延量を導出、その後その遅延量を各地点に配信する。音響サーバは同期制御サーバから送信された最大遅延量を基準に補正遅延を計算し設定する。具体例として3地点の構成と各地点間の通信遅延の状態を図 3.2 にローカル・ラグを施した状態を図 3.3 に示す。図 3.2 では地点 A-地点 B 間の通信遅延が 20ms、地点 A-地点 C 間の通信遅延が 10ms、地点 B-地点 C 間の通信遅延が 40ms であり、最大遅延量は地点 B-地点 C 間の 40ms となる。図 3.3 より地点 A に着目した場合、地点 A で聞こえる音は、地点 A のフィードバック音と地点

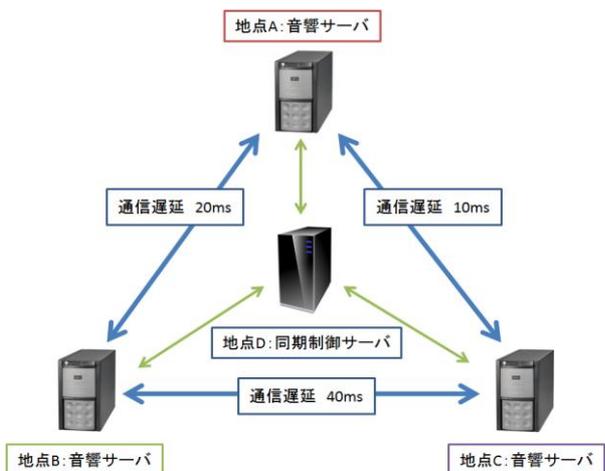


図 3.2 3地点の構成と各地点間の通信遅延の状態。

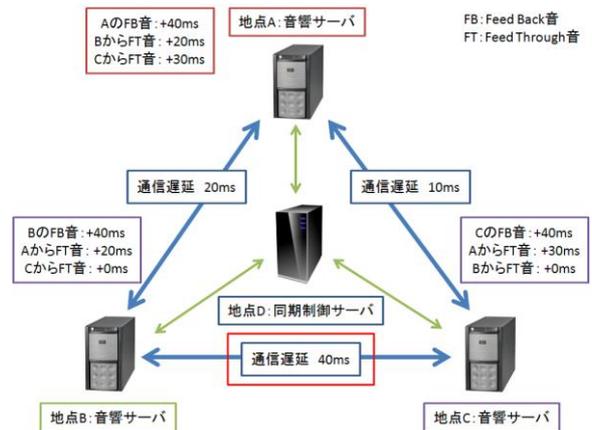


図 3.3 3地点間においてローカル・ラグ制御を行った状態。

Bからのフィードスルー音、そして地点 Cからのフィードスルー音の3つである。まず地点 Aのフィードバック音に関しては最大遅延量である 40ms に設定する。続いて地点 Cからのフィードスルー音に関しては、10msの通信遅延を伴って地点 Aに届くため、最大遅延量の 40ms に合わせるため 30ms 分の補正遅延を付加する。最後に地点 Cからのフィードスルー音に関しては、20msの通信遅延を伴って地点 Aに届くため、20ms 分の補正遅延を付加する。これにより地点 Aのフィードバック音と、地点 Bからのフィードスルー音、地点 Cからのフィードスルー音のそれぞれが 40ms のラグを伴って出力されるため同期が保たれた音が聞くことが可能となる。地点 B及び地点 Cについても同様である。データの送信には User Datagram Protocol (UDP) を採用している。送信後、受信スレッドは相手から送られてくる音データ(フィードスルー音)と自分自身から送られてくる音データ(フィードバック音)を受信し、その後、出力スレッドにより指定された遅延量に基づく時間調整を行い、ドライバに出力する。

各機能の詳細や処理の流れなどは下記を参照されたい。

オーディオバッファ：入力スレッドもしくはドライバに音データを渡すためのリングバッファ。音データは渡す際には PortAudio というライブラリのコールバック関数(ノンブロッキング処理)が使用されている。

入力スレッド：ドライバから音データを取得し、接続している全ての音響サーバ(自身も含む)に音データを送信するスレッド。通信プロトコルには UDP を採用している。

受信スレッド：対応する送信元(自身も含む)から送られてくる音データを受信し、受信時刻を記録してディレイバッファに格納するスレッド。

ディレイバッファ：受信スレッド毎に1つ用意されており、受信時刻が付加された音データを格納するリングバッファ(リングバッファ内で扱うデータサイズは各時点における波形データを表す short 型データの個数を基本としている)。単位補正遅延に基づいてこのバッファから音データを取り出すタイミングを調整することで、ローカル・ラグ機能を実現する。その時のタイミングは音響サーバ間の遅延量に依存する。

出力スレッド：受信スレッド毎に1つ用意されており、取りだし時刻を求めてディレイバッファから音データを取りだし、オーディオバッファに転送するスレッド。

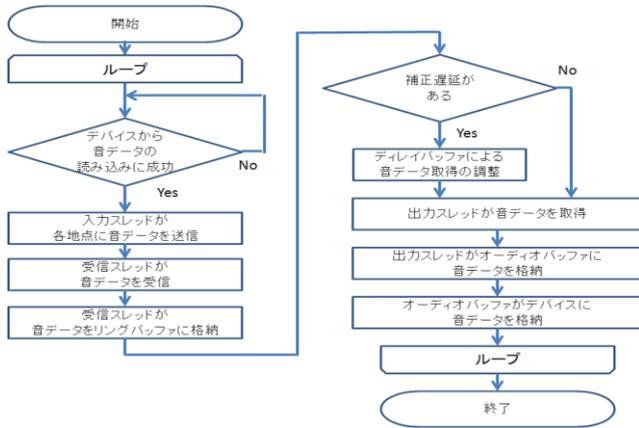


図 3.4 音響サーバのフローチャート

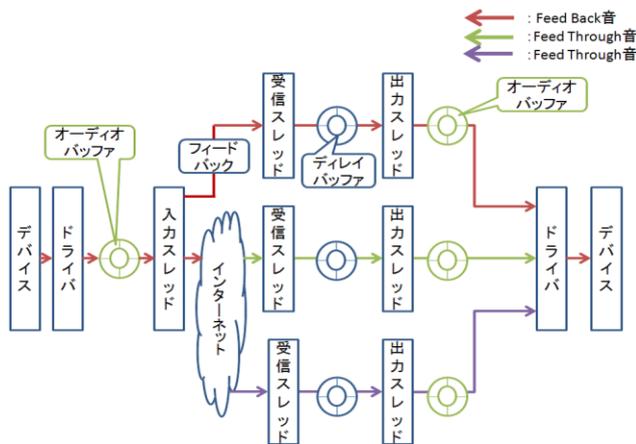


図 3.5 3地点接続時の処理の全体構成。

信遅延量を付加することで通信相手と同期を保つため、それらのずれがどのような挙動なのかを把握する必要がある。直接音とはレコーダーに入力される電子メトロノームの音のことである。フィードバック音は音響サーバに入力された電子メトロノームの音がシステムを通して出力される音のことである。フィードスルー音は相手から送信された音のことである。1回の計測を60分間とし各地点にある音響サーバをA, B, Cとしてレコーダーに録音された直接音とフィードバック音とのずれと、直接音と自身のフィードスルー音のずれを5分毎に、レコーディング・サウンド編集ソフトウェアであるAudacityを用いて計測した。取り扱う音データサイズは、128, 256, 512としている。これは先行研究の実験結果よりこれらのデータサイズが開発されてきた音響サーバ上で安定して動作することが示されているためである。以下に同期制御サーバと音響サーバの諸元を示す。

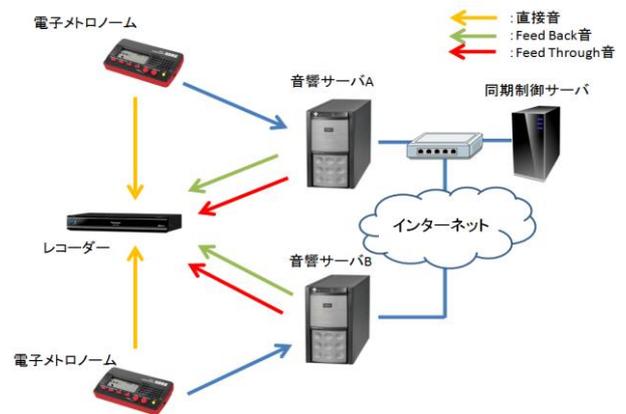


図 4.1 WAN環境下における2地点の実験環境。

4. 性能評価実験

4.1 実験概要

我々の先行研究では開発した音響サーバにおける処理時間の安定性とローカル・ラグ制御の有効性を示してきたものの、評価実験環境は全LAN環境下で行われてきた。また多地点接続機能がありながらもその評価実験を行ってこなかった。本稿の実験目的は開発した音響サーバの実用面の運用を考慮し、WAN環境下と多地点間の例として3地点間の処理時間の評価実験を行った。多地点接続において3地点間を取り上げた主な理由は以下の通りである。即ち、音響サーバは、その仕組み上、フィードバック音を扱う経路は常に1つである一方で、フィードスルー音を扱う経路は複数になり得る。しかしながら、これまでの評価実験では1経路のみの検証に留まっており、それを2経路にすることで複数経路としての動作検証を行うというものである。

次に実験環境と手順を述べる。図4.1はWAN環境下での実験環境、図4.2は3地点間の実験環境を示している。実験には120BPMの電子メトロノームとマルチトラックレコーダー（以下ではレコーダーと呼ぶ）を用いた。音響サーバの処理時間の安定性を評価するため、直接音とフィードバック音の時間的なずれと直接音とフィードスルー音の時間的なずれに注目した。それらのずれに注目した理由は、ローカル・ラグ制御はフィードバックに通

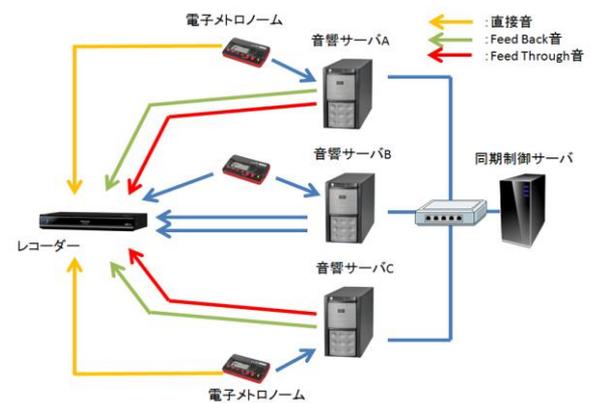


図 4.2 LAN環境下における3地点の実験環境。

同期制御サーバの諸元

- CPU : Inter(R) Core(TM) i7-3770 CPU
- OS : CentOS 6.4
- メモリ : 4GB

音響サーバ A, B, C の諸元

- CPU : Inter(R) Core(TM) i7-3770 CPU
- OS : CentOS 6.4
- メモリ : 4GB
- サウンドドライバ : ALSA
- サウンドカード : SoundBlaster X-Fi Fatal1ty FPS

4.2 WAN 環境下における 2 地点実験結果

実験結果を以下の図 4.3 から図 4.5 に示す。図はそれぞれ、データサイズを 128、256、512 に設定した場合に対応する。図の縦軸は、直接とフィードバック音のずれまたは直接音とフィードスルー音のずれを差という形で表している。図の横軸は 5 分ごとの計測時点を示している。各データサイズにおいて、3 回の計測 (1 時間/回) を行った。本稿ではページの関係上、1 回分のみを掲載している。

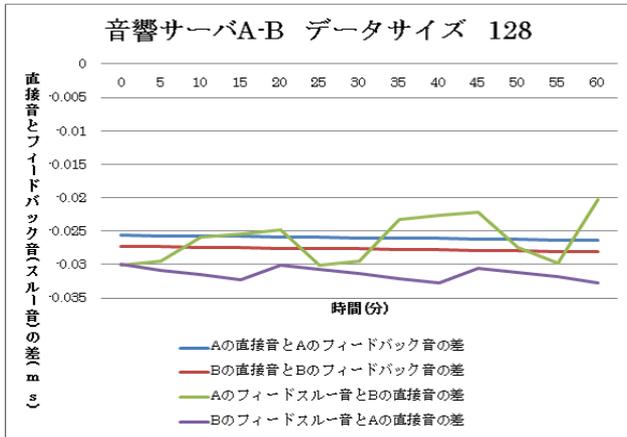


図 4.3 WAN 環境下での 2 地点間におけるデータサイズ 128 の実験結果。

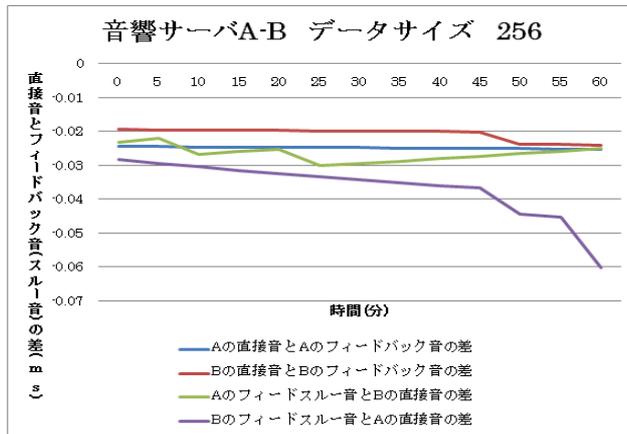


図 4.4 WAN 環境下での 2 地点間におけるデータサイズ 256 の実験結果。

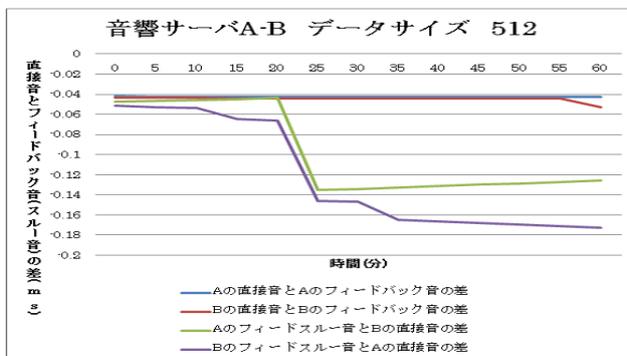


図 4.5 WAN 環境下での 2 地点間におけるデータサイズ 512 の実験結果。

4.3 WAN 環境下における実験結果の考察

4.2 節より各データサイズにおける直接音とフィードバック音のずれは 1 時間ではほぼ見られないため安定的に動作していることが示されている。この事はフィードバック音に通信遅延量を付加するローカル・ラグ制御を安定して行う事ができると考えられる。しかしながら、各データサイズにおける直接音とフィードスルー音のずれは 1 時間で時間とともに増減し、小さいずれでは 10ms、大きいずれでは 100ms 前後であるという事が明らかとなった。

この原因については、前者は、フィードスルー音がフィードバック音に対しサウンドカードを 2 回経由することから、直接音に対して波形が時間軸方向に伸縮するためであるという事が示されている[6]。また後者は、直接音とフィードスルー音の大きなずれは WAN を経由することから、LAN よりも動的で大きな遅延による影響、もしくは帯域制御の影響によるものであるという事が示されている[7]。

4.4 LAN 環境下における 3 地点実験結果

実験結果を以下の図 4.6 から図 4.14 に示す。図はそれぞれ、各地点におけるデータサイズを 128 とした場合からデータサイズを 512 に設定した場合に対応する。図の縦軸は、直接音とフィードバック音のずれまたは直接音とフィードスルー音のずれを差という形で表している。図の横軸は 5 分ごとの計測時点を示している。各データサイズにおいて、3 回の計測 (1 時間/回) を行った。本稿ではページの関係上 1 回分のみを掲載している。

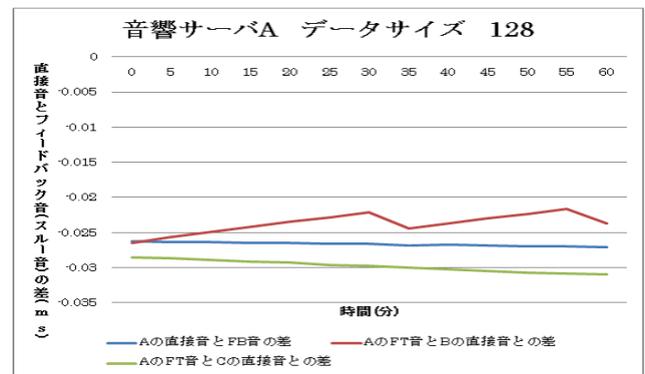


図 4.6 LAN 環境下での 3 地点間におけるデータサイズ 128 の音響サーバ A での実験結果。

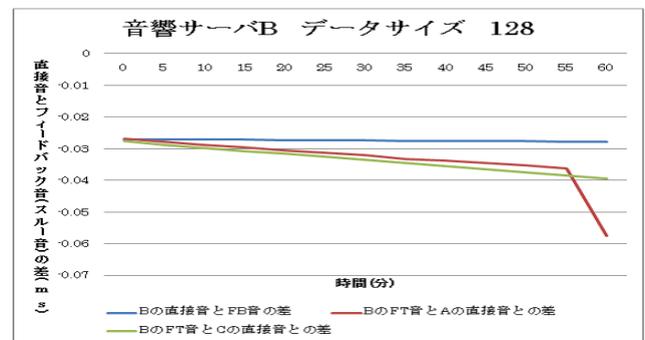


図 4.7 LAN 環境下での 3 地点間におけるデータサイズ 128 の音響サーバ B での実験結果。

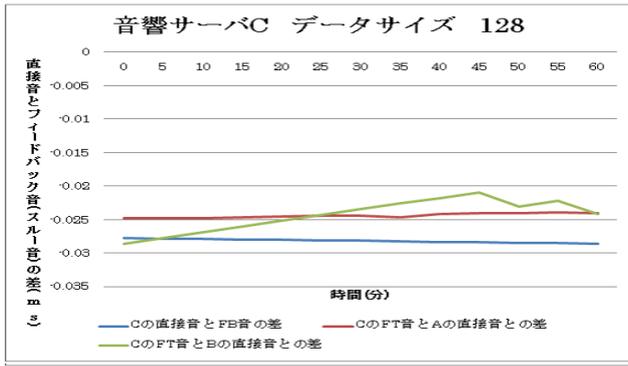


図 4.8 LAN 環境下での 3 地点間におけるデータサイズ 128 の音響サーバ C での実験結果。

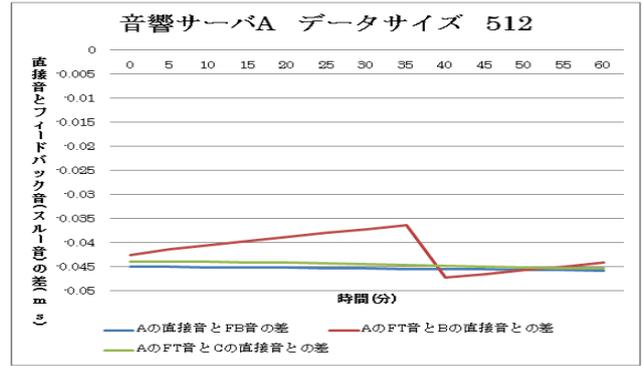


図 4.12 LAN 環境下での 3 地点間におけるデータサイズ 512 の音響サーバ A での実験結果。

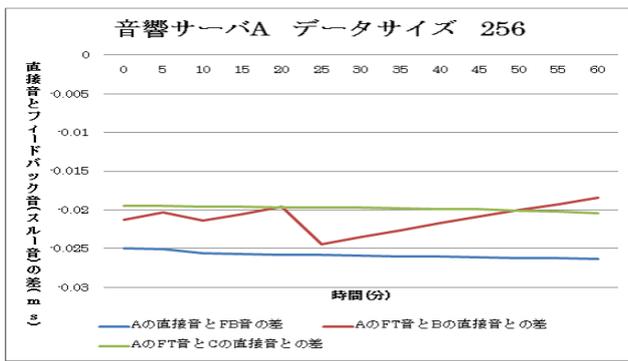


図 4.9 LAN 環境下での 3 地点間におけるデータサイズ 256 の音響サーバ A での実験結果。

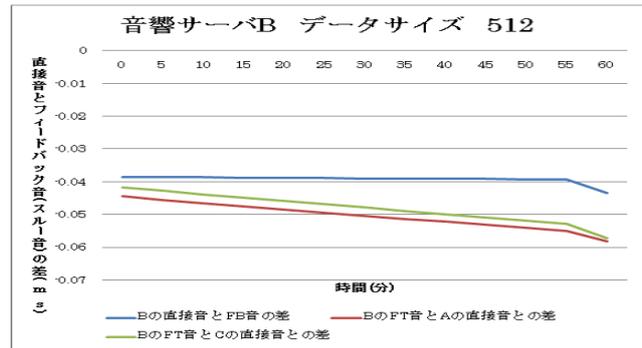


図 4.13 LAN 環境下での 3 地点間におけるデータサイズ 512 の音響サーバ B での実験結果。

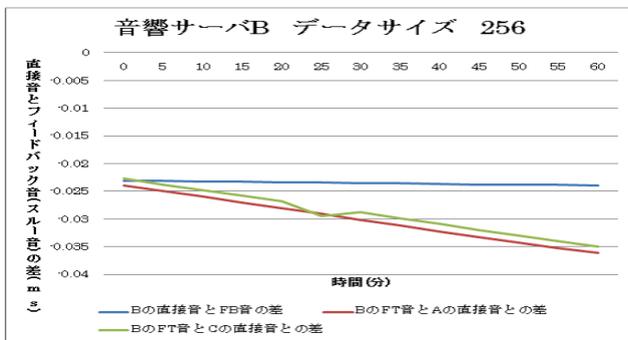


図 4.10 LAN 環境下での 3 地点間におけるデータサイズ 256 の音響サーバ B での実験結果。

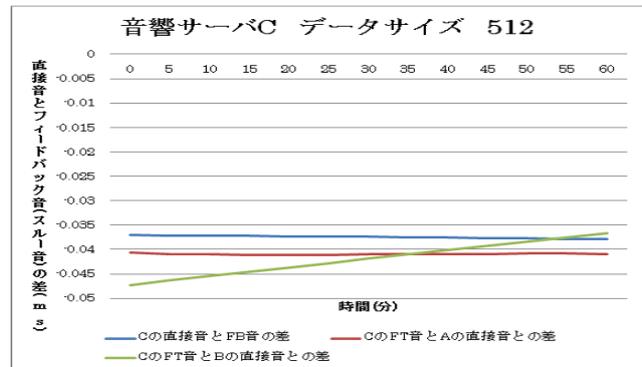


図 4.14 LAN 環境下での 3 地点間におけるデータサイズ 512 の音響サーバ C での実験結果。

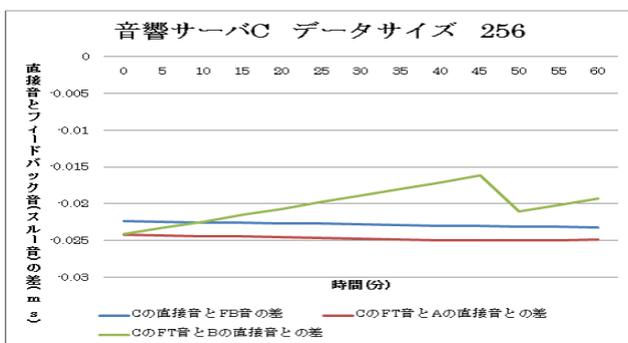


図 4.11 LAN 環境下での 3 地点間におけるデータサイズ 256 の音響サーバ C での実験結果。

4.5 3 地点接続実験結果に関する考察

4.4 節より各データサイズにおける直接音とフィードバック音のずれは 1 時間ではほぼ見られないため安定的に動作していることが示されている。この事はフィードバック音に通信遅延量を付加するローカル・ラグ制御を安定して行う事ができると考えられる。しかしながら、各データサイズにおける直接音とフィードスルー音のずれは 1 時間で時間とともに増減していることが明らかになった。また実験環境が 4.2 節と異なるため (WAN 環境下と LAN 環境下の違い)、各データサイズにおける直接音とフィードスルー音のずれは 1 時間ではおよそ 10ms から 20ms 前後と小さいものである事が示されている。

4.6 リセット機能の概要

4.3節、4.5節よりフィードバック音に関しては1時間の測定においても安定的に動作しているためローカル・ラグ制御を安定して行う事が可能であると考えられる。しかしながら、フィードスルー音においては直接音とのずれは安定しない事が明らかとなった。直接音とフィードスルー音のずれが安定しない場合、2地点間ないし3地点間においてもフィードバック音にローカル・ラグ制御を行っても各地点間の同期を保つ事は難しいと考えられる。そこで本稿では直接音とフィードバック音または直接音とフィードスルー音のずれに対してリセットする機能を実装した。

図 3.5 より開発した音響サーバは入力スレッドの左に位置するオーディオバッファから音データを取得していることが示されている。そのためこの箇所を初期化した後にスリープを行うことでリセット機能を実現できると考えた。4.7節と4.8節ではリセット機能を用いた上で4.2節および4.4節と同じ実験条件で同様の実験を行った。

4.7 WAN 環境下でのリセット機能を用いた 2 地点実験結果

実験結果を以下の図 4.19 から図 4.27 に示す。図はそれぞれ、データサイズを 128 とした場合からデータサイズを 512 に設定した場合に対応する。図の縦軸は、直接音とフィードバック音のずれまたは直接音とフィードスルー音のずれを差という形で表している。図の横軸は 5 分ごとの計測時点を示している。10 分おきにリセットを行い各データサイズにおいて、3 回の計測 (1 時間/回) を行った。本稿ではページの関係上 1 回分のみを掲載している。

4.8 LAN 環境下でのリセット機能を用いた 3 地点実験

実験結果を以下の図 4.16 から図 4.14 に示す。図はそれぞれ、データサイズを 128 とした場合からデータサイズを 512 に設定した場合に対応する。図の縦軸は、直接音とフィードバック音のずれまたは直接音とフィードスルー音のずれを差という形で表している。図の横軸は 5 分ごとの計測時点を示している。10 分おきにリセットを行い各データサイズにおいて、3 回の計測 (1 時間/回) を行った。本稿ではページ関係上一回分のみを掲載している。

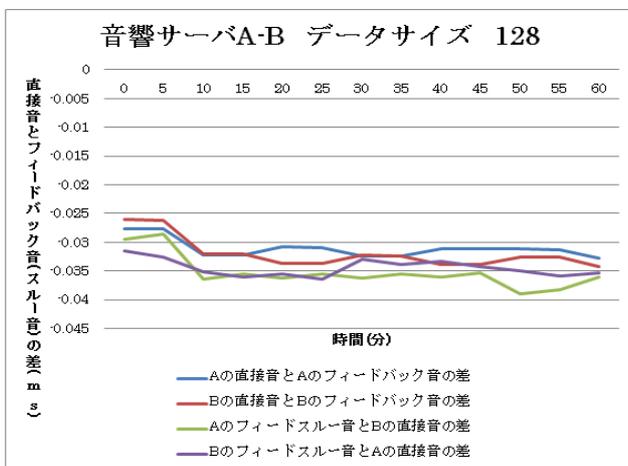


図 4.16 WAN 環境下での 2 地点間におけるリセット機能を用いたデータサイズ 128 の実験結果。

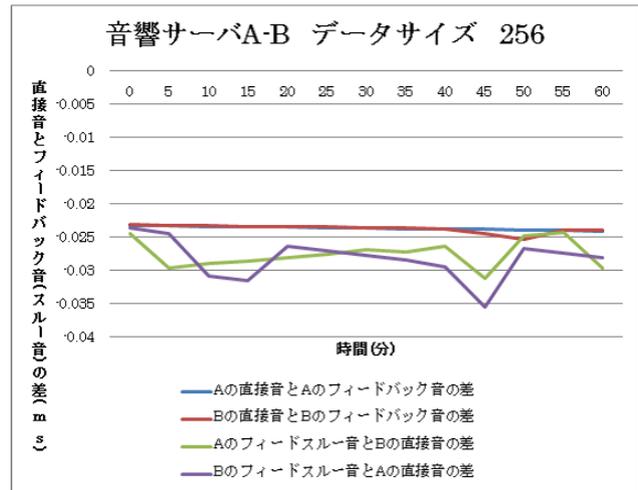


図 4.17 WAN 環境下での 2 地点間におけるリセット機能を用いたデータサイズ 256 の実験結果。

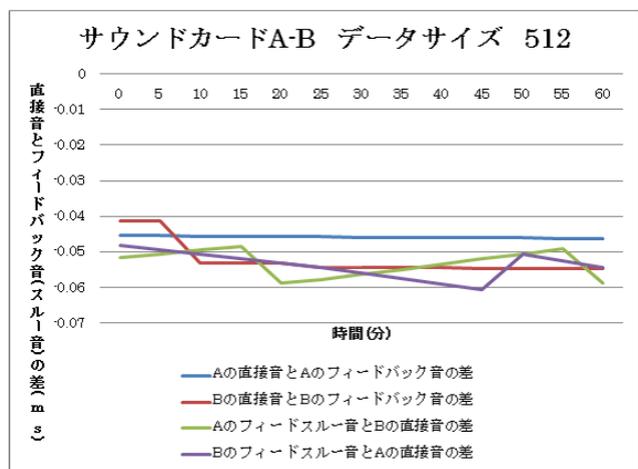


図 4.18 WAN 環境下での 2 地点間におけるリセット機能を用いたデータサイズ 512 の実験結果。

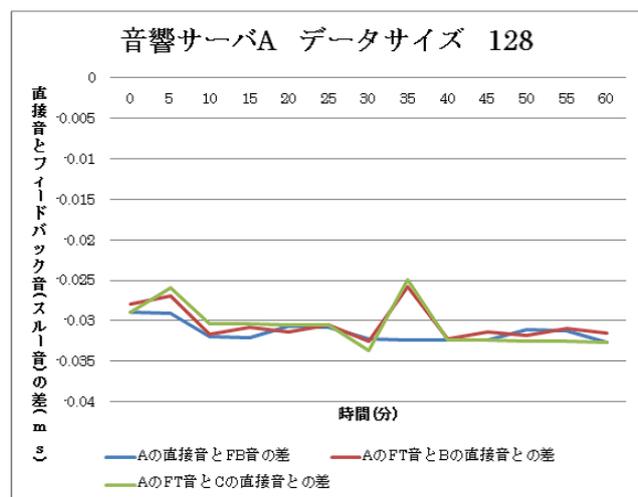


図 4.19 LAN 環境下での 3 地点間におけるリセット機能を用いたデータサイズ 128 の音響サーバ A での実験結果。

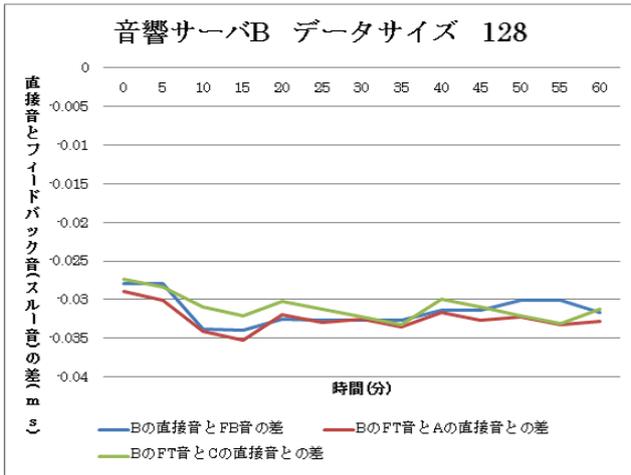


図 4.20 LAN 環境下での 3 地点間におけるリセット機能を用いたデータサイズ 128 の音響サーバ B での実験結果。

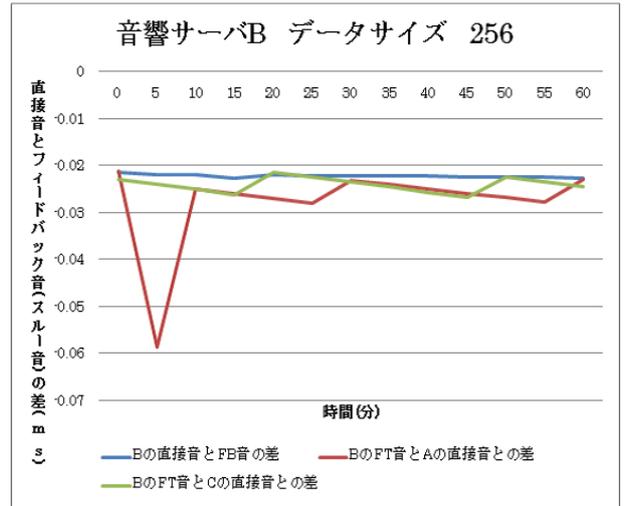


図 4.23 LAN 環境下での 3 地点間におけるリセット機能を用いたデータサイズ 256 の音響サーバ B での実験結果。

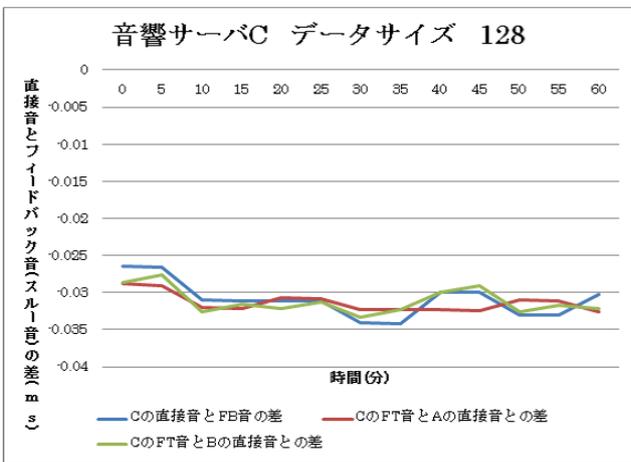


図 4.21 LAN 環境下での 3 地点間におけるリセット機能を用いたデータサイズ 128 の音響サーバ C での実験結果。

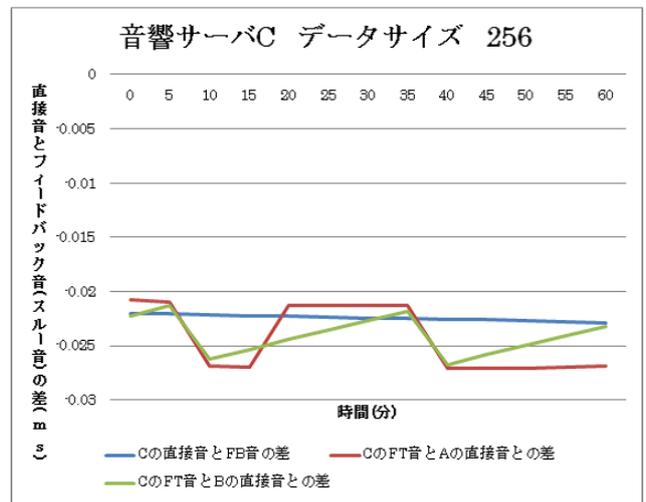


図 4.24 LAN 環境下での 3 地点間におけるリセット機能を用いたデータサイズ 256 の音響サーバ C での実験結果。

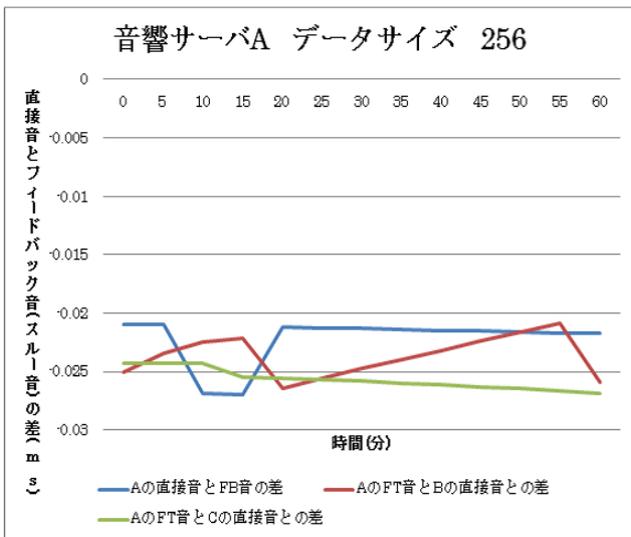


図 4.22 LAN 環境下での 3 地点間におけるリセット機能を用いたデータサイズ 256 の音響サーバ A での実験結果。

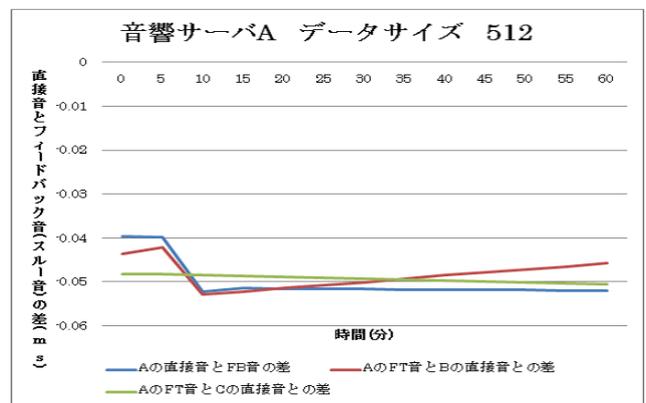


図 4.25 LAN 環境下での 3 地点間におけるリセット機能を用いたデータサイズ 512 の音響サーバ A での実験結果。

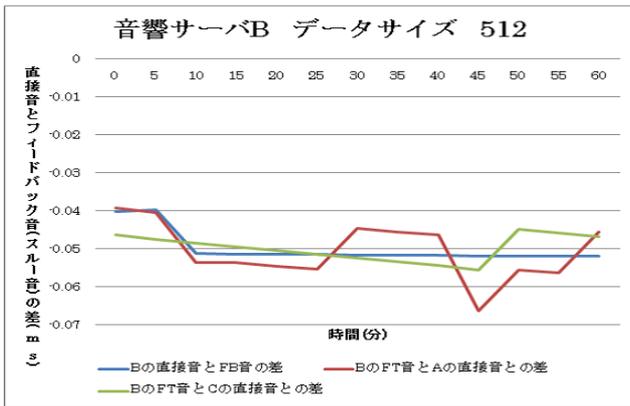


図 4.26 LAN 環境下での 3 地点間におけるリセット機能を用いたデータサイズ 512 の音響サーバ A での実験結果.

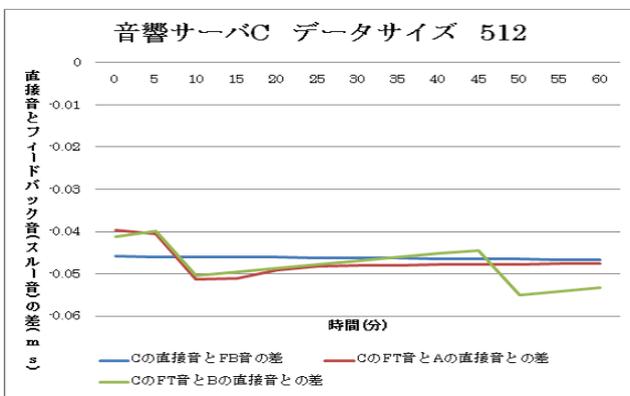


図 4.27 LAN 環境下での 3 地点間におけるリセット機能を用いたデータサイズ 512 の音響サーバ A での実験結果.

4.9 考察

4.6 節と 4.7 節より、フィードバック音においては 4.2 節と 4.4 節の結果と比較すると処理時間が不安定になることが示されたものの、一方で図 4.17 や図 4.23 より直接音とフィードスルー音のずれを補正できることが示された。実験結果より 5 ~ 10 分程度なら LAN と WAN との環境を問わずリセット機能を用いた状態でローカル・ラグ制御を行えば数値上では 2 地点または 3 地点において同期を確保することが可能であると考えられる。

5. まとめ

本稿では音響サーバの実用性を示すために、WAN 環境下での性能評価実験と 3 地点間での音響サーバの性能評価実験を行いその結果を明らかにした。その結果から 2 地点または 3 地点において直接音とフィードバック音のずれは一定かつ安定的であるもの、一方で直接音とフィードスルー音のずれは増減することが明らかとなった。また、LAN 環境下よりも WAN 環境下の方がよりフィードスルー音のずれが大きい事が明らかとなった。この問題に対してリセット機能を実装する事で直接音とフィードスルー音のずれを補正できることを示すことができた。ずれの補正とローカル・ラグ制御を行うことで 5 分から 10 分程度なら LAN 環境と WAN 環境を問わず、2 地点または 3

地点で音響サーバ間における同期を確保することが可能であると考えられる。

参考文献

- [1] Keiji Hirata, Yasunori Harada, Toshihiro Takada, Shigemi Aoyagi, Yoshinari Shira, Naomi Yamashita, and Junji Yamato; The t-Room: Toward the Future Phone, NTT Technical Review, Vol. 4, No. 12, pp. 26-33 (2006).
- [2] Dane Stuckel and Carl Gutwin: The Effects of locallag on Tightly-Coupled Interaction in Distributed Groupware: Computer Supported Cooperative Work, pp. 447-456 (2008).
- [3] 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, Vol. 2009-GN-73, No. 3 (2009. 11).
- [4] 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, 情報処理学会, オーディオビジュアル複合情報処理研究会 (2012. 2).
- [5] 前田佳奈, 竹森幸輝, 岩原正典, 片桐滋, 大崎美穂; ローカル・ラグ機能を持つ音響サーバを用いた遠隔合奏の評価, 情報処理学会, オーディオビジュアル複合情報処理研究会 (2012. 2).
- [6] 岩原正典; ローカル・ラグ制御機能と同期機能を持つ音響サーバの構築, 同志社大学大学院理工学研究科修士論文 (2013. 2).
- [7] 大島義博; 遠隔合奏支援システムが持つローカル・ラグ制御機能の精度評価実験, 同志社大学理工学部卒業論文 (2013. 2).