

ネットワークプロトコルを用いた 遠隔対話における複数人インタラクションモデルの構築

Multi-person interaction model in remote discussion based on network protocol

末松圭史^{1*} 大坪正和¹ 金津達也¹ 吉田香¹
Keishi Suematsu¹ Masakazu Otsubo¹ Tatsuya Kanatsu¹ Kaori Yoshida¹

¹ 九州工業大学 大学院 生命体工学研究科

¹ Graduate School of Life Science and Systems Engineering, Kyushu Institute of technology

Abstract: During remote discussion using video conference system, utterance conflict is happened as a problem caused by lack of non-verbal communication. On the other hand, a network protocol for detecting / avoiding collision of data communication called CSMA (Carrier Sense Multiple Access). In our research, we try to replace human discussion with data communication on the network, and propose a multi-person interaction model. This paper shows the proposal of multi-person interaction model in remote discussion based on network protocol and the experimental results of network simulation.

1 はじめに

近年, Skype などのネットワークに繋がれたコミュニケーションツールを用いて, 遠隔地間で対話を行う遠隔対話システムの普及が進んでいる. 遠隔対話システムでは, 全参加者が同じ空間に居る必要はないため, 1対1の対話だけでなく, 3人以上での対話を行う手段として数多く利用されている. しかし, 映像と音声のみを用いた遠隔対話では, 誰がいつ話し始めるのかを判断するのが難しいため, 2人以上が同時に発話を開始してしまう発話衝突が多く発生することが問題となる. 先行研究によると, 遠隔対話では音声遅延が 300msec を超えると発話衝突の発生が顕著になり [1], 司会者を設けずアイデアを出し合う創造会議 [2] を行った場合は対面会議と比較して発話衝突が約 30 倍発生する [3] と報告されている. また, 発話衝突の発生により, 発話の中断や言い直しが多くなり [4], 遠隔対話システムの全参加者の精神的なストレスが増加する [5] と報告されている.

発話衝突の発生は, 対話を行う環境が統一されていない, 相手の存在感を感じにくい, 相手と視線を合わせることが難しいなどの制約が原因であると考えられる. 特に, 相槌などの言語情報 [6] に加えて, 身振り手振りを用いた非言語情報であるノンバーバルコミュニケーション [7] が困難であることが原因のひとつとして

考えられる. 本研究では, 腕組みをほどく, 身を前に乗り出す, 発話者の方へ身体や顔の向きを変える, 頭部動作 [8] などを発話予備情報と呼ぶ. 発話予備情報に関する先行研究として, 発話衝突問題の解決を目指して Kinect で取得した人の動作をヒューマノイド型ロボットに実装することで発話予備情報を補う研究 [9], Kinect で取得した人の動作を判別器にかけて発話者の選定を行う研究 [10], 人の注視の遷移パターンに基づいて発話の開始タイミングを予測する研究 [11] などが挙げられる.

一方, ネットワーク研究分野では, データ衝突を検知・回避するためのプロトコルが存在する [12]. 例えば, データ衝突を軽減する手法として, インターネットを用いられている通信プロトコル CSMA (Carrier Sense Multiple Access) [12] がある. CSMA では, データを送信してよいかどうか送信したいデータに先がけて送信要求を送り, 送信先からの受信許可を得た後にデータを送信する. そこで, 本研究では, 人の対話における発話衝突を通信プロトコルにおけるデータ衝突と捉え, ネットワークの考え方に基づいて人の対話をモデル化することを試みる. 具体的には, ネットワークにおけるデータの通信を人の対話に置き換え, 発話衝突の発生をネットワーク研究の視点から考察し, 遠隔対話を行う際の指標の一つとして用いることのできる複数人インタラクションモデルの構築を行う.

まず, 人の対話を通信プロトコルの上で定義し, ネットワークシミュレータを用いて人の対話の再現を試み

*連絡先: 九州工業大学 大学院 生命体工学研究科
〒 808-0196 福岡県北九州市若松区ひびきの 2 番 4 号
E-mail: suematsu-keishi@edu. brain. kyutech. ac. jp

る。その後、ネットワークシミュレータの結果から得られた知見を遠隔対話システムを用いた複数人の対話実験でも有効であることを検証することで、複数人インタラクションモデルを構築する。複数人インタラクションモデルの構築を行うことで、発話衝突の発生原因と発話衝突の回避手法を考察し、より快適に遠隔対話を行うことのできる遠隔対話システムの改良を目指す。本稿では、データ衝突を軽減する手法としてインターネットで用いられている通信プロトコルのひとつである CSMA について概略を解説し、CSMA において人の発話衝突の再現を試みたネットワークシミュレーション結果を報告する。

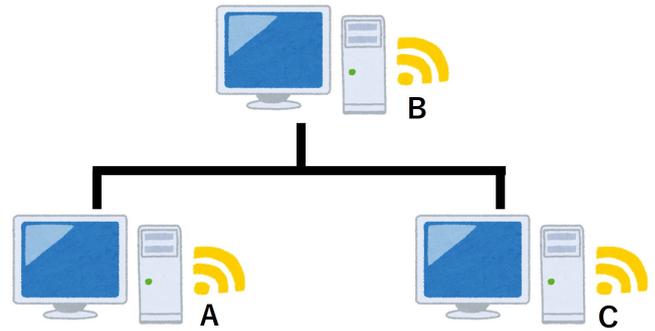


図 1: ネットワークトポロジ

2 通信プロトコルにおけるデータ衝突と人の対話における発話衝突

本章では、まずネットワーク研究分野で用いられている通信プロトコル CSMA (Carrier Sense Multiple Access) について説明する。その後、人の対話における発話衝突がどのように発生するのか CSMA を用いて仮説を述べる。

2.1 CSMA

ここでは、データ衝突を軽減する手法としてインターネットで用いられている通信プロトコル CSMA[12] について概説する。CSMA はデータリンク層の通信プロトコルとして使用されている。複数のノードが同じネットワークを共用し、データ通信を開始する前に一度受信を試みることで、現在通信を行っているノードが他に存在するかどうかを確認し、他ノードが通信を行っていないならば、自身のデータ通信を開始する手法である。有線ネットワーク通信と無線ネットワーク通信に 2 パターンがあり、有線通信の場合は CSMA/CD (Carrier Sense Multiple Access with Collision Detection) と呼ばれ、衝突 (Collision) を検知 (Detection) するプロトコルとなり、無線通信の場合は CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance) と呼ばれ、衝突を回避 (Avoidance) するプロトコルとなる。まずは、最もシンプルな状態である図 1 のようなネットワークトポロジ (ノードの配置) を想定する。

2.1.1 CSMA/CD

ここでは、有線ネットワークが採用している通信方式のひとつである CSMA/CD (Carrier Sense Multiple Access with Collision Detection) について概説する。CSMA/CD のフローチャートを図 2 に示す。各ノード

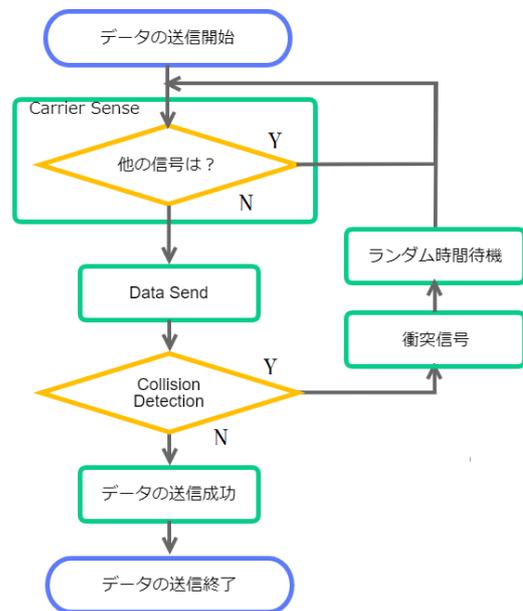


図 2: CSMA/CD

は Carrier Sense タームで、自身の通信可能範囲内に他にデータの送受信を行っているノードが存在しないことを確認する。次に、ノード A はノード B に向けてデータの送信を開始する。この時、ノード C もノード A と同タイミングで Carrier Sense タームを終えてしまった場合、ノード A と同じくノード C もノード B に向けてデータの送信を開始する。複数のノードが同タイミングでデータの送信を行ったことで、データ衝突が発生する。データ衝突が発生すると有線ケーブル内の電圧が上昇するため、各ノードは物理層レベルでデータ衝突の発生を検知することが可能である。衝突を検知したノード A とノード C はバックオフ時間 (ランダム時間) 待機した後、Carrier Sense タームを消化し、再度ノード B に向けデータの送信を開始する。

表 1: 通信プロトコルと人の対話の対応表

通信プロトコル	人の対話
ノード	遠隔対話システム参加者
Carrier Sense	発話してよいか探る (例) 周りを見渡す
Data Send	発話
Collision Detection	発話衝突の発生
送信要求 (RTS)	発話予備情報の表出
受信許可 (CTS)	発話許可を送る (例) (司会者がいる場合) ○○さんどうぞ

2.1.2 CSMA/CA

ここでは、無線ネットワークが採用している通信方式のひとつである CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance) について概説する。各ノードは Carrier Sense タームで、自身の通信可能範囲内に他にデータの送受信を行っているノードが存在しないことを確認する。次に、バックオフ時間 (ランダム時間) 待機する。バックオフ時間中に他のノードの通信を確認しなかった場合、(ノード A がノード C より先にバックオフ時間を消化したと仮定すると,) ノード A はノード B に向けてデータの送信を開始する。ノード B はデータの受信が完了した後、バックオフ時間待機した後、確認応答 (ACK) を送信する。ノード C はノード B の ACK を確認した後、Carrier Sense タームを消化し、残りのバックオフ時間を消化し、データの送信を行う。

無線通信において、ノード間に電波を通しにくい遮蔽物がある場合や、ノード間に距離があり互いの電波を検知できない場合は Carrier Sense により相手ノードの状態を把握することができない。そのため同タイミングでデータの送信を開始することが多くなり、データ衝突が発生しやすくなる。これを隠れ端末問題と呼ぶ。この問題を解決するために、CSMA/CA のオプション機能として、RTS (送信要求) /CTS (受信許可) を用いた制御を行うことができる。

2.2 人の会話における CSMA

ここでは、人の対話を通信プロトコル CSMA の上で定義することを試みる。人の対話と CSMA の用語の対応表を表 1 に示す。

人の対話は発話衝突の検知が容易であるため、CSMA/CD で行っていると考えることができる。諸研究 [9][10] では、RTS の送信は発話衝突の回避において大きな効果があると報告されている。

ネットワークプロトコルを用いると、発話衝突は以下の条件で発生すると考えられる。複数人が同時に発話を開始するほかに、2つの仮説を立てる。

- 仮説 1: 送信要求 (RTS) の送信を行わず発話を開始
- 仮説 2: 送信要求 (RTS) の送信を行うが、受信許可 (CTS) を受信したと勘違いし発話を開始

仮説 1: 電話など相手の身体情報を認識することのできない会議システムでは、送信要求 (RTS) の送信を行うことができない。そのような状況では発話衝突が数多く発生すると考えられる。この場合、送信要求 (RTS) /受信許可 (CTS) を用いた対話を行わないと想定される。

仮説 2: 創造会議 [2] では司会者を設けずに会議が進行されるため、受信許可 (CTS) の送信が司会者からの指名など明示的に行われず、各会議参加者の視線の移動や、身体の向きの変更、場の雰囲気などで暗示的に行われる。そのため、受信許可 (CTS) を受信したかどうか各会議参加者の判断に依存しているため、受信許可 (CTS) の受信の誤認識が発生し、発話してはいけないタイミングで発話してしまうことがある。この場合、送信要求 (RTS) の送信を行った後、数秒待機し、発話を行うと想定される。仮説を検証するために、まずシミュレーションモデルの構築を行い、複数人対話モデルの構築を行う。

3 シミュレーション

ここでは、前章で定義した人の対話を通信プロトコルにおけるデータの通信として行うネットワークシミュレーションについて説明する。図 1 のようなネットワークポロジ (ノードの配置) を想定し、各ノードは通信路で接続されているとする。シミュレーションにおける各パラメータは表 2 とする。前節で立てた仮説から、3 ノードは以下の 3 条件で通信を行い、それぞれの条件におけるデータ衝突数とデータ送信数の比較を行う。今回は遠隔対話環境を想定するため、全条件において Carrier Sense は行わないものとする。

1. 図 2 に示す CSMA/CD のアルゴリズム
2. 図 3 に示す CSMA/CDwithRTS のアルゴリズム
3. 図 4 に示す CSMA/CDwithRTSCTS のアルゴリズム

条件 1: 各ノードは 1~5sec のランダム時間で他の全ノードに向けてデータの送信を行う。各ノードのデータ送信のタイミングが被った場合、データ衝突が発生

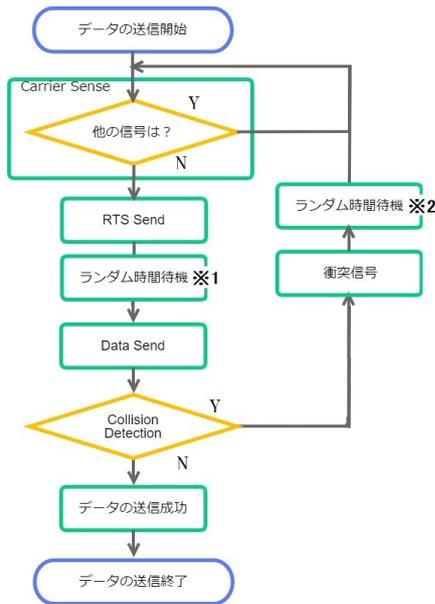


図 3: CSMA/CDwithRTS

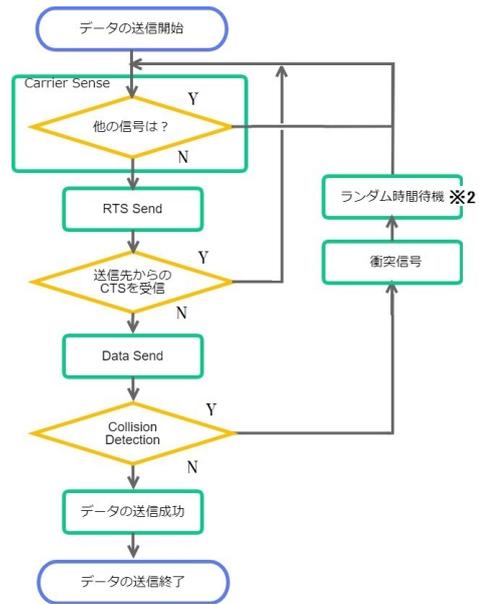


図 4: CSMA/CDwithRTSCTS

したと判断し、通信路から衝突信号の送信が行われる。衝突信号を受け取ったノードはデータの送信を停止し、1~5secの間でランダムに待機、データの送信を最初から再開する。

条件 2: 各ノードはデータ送信の前に RTS の送信を行い、1, 2, 3sec の間それぞれ待機する。その後、1~5sec のランダム時間で他の全ノードに向けてデータの送信を行う。各ノードのデータ送信のタイミングが被った場合、データ衝突が発生したと判断し、通信路から衝突信号の送信が行われる。衝突信号を受け取ったノードはデータの送信を停止し、1~5sec の間でランダムに待機、再度 RTS の送信を行い、待機した後、データの送信を最初から再開する。

条件 3: 各ノードはデータ送信の前に RTS の送信を行い、その後、通信路から自身のノード当りの CTS が送信されるのを待つ。1sec 以内に CTS を受信しない場合、再度 RTS の送信を行う。CTS を受信したノードは、1~5sec のランダム時間で他の全ノードに向けてデータの送信を行う。通信路は通信路中にデータが流れている場合、CTS の送信を行う事ができないため、ノードは自身のデータ送信を行う事ができず、データ衝突は発生しない。

4 結果

各条件で 3 分間のシミュレーションを 10 回行った後、ノード A, B, C のデータ衝突数とデータ送信数をそれぞれ合計し、平均と標準偏差を計算した結果を図 5

表 2: シミュレーションパラメータ

データの送信開始の発生回数	約 12 秒に 1 回
1 回のデータの送信にかかる時間	1~5sec
ランダム時間待機*1	1~3sec
ランダム時間待機*2	1~5sec

に示す。CD_RTS で 1sec 待機する場合、CD_RTS1 と表記し、2sec 待機は CD_RTS2, 3sec 待機は CD_RTS3 と表記する。CD は CD_RTS1, CD_RTS2, CD_RTS3 と比較し、データ衝突数は多いものの、データ送信数も多いという結果となった。また、CD_RTSCTS はデータ衝突数が 0 であり、データ送信数も 5 条件の中で最も多いという結果になった。

CD と CD_RTS1, CD_RTS2, CD_RTS3 の間において一元配置分散分析を行うため、まず等分散性の検定を行った結果、データ衝突数は $P = 0.799$ 、データ送信数は $P = 0.290$ となり、どちらも $P > 0.05$ であるため、等分散性があるといえる。次に一元配置分散分析を行った結果を表 3 と 4 に示す。表より、データ衝突数の P 値は、 0.141×10^{-11} 、データ送信数の P 値は 0.230×10^{-4} となり、どちらも平均値の間に有意差がみられた。

次に CD と CD_RTS1, CD_RTS2, CD_RTS3 の間において、どの要素の間に差があるのかを調べるため、Tukey の方法を用いて多重検定を行った結果を表 5 と 6 に示す。表より、データ衝突数においては、CD と CD_RTS 以外すべての組み合わせで有意差がみられ、

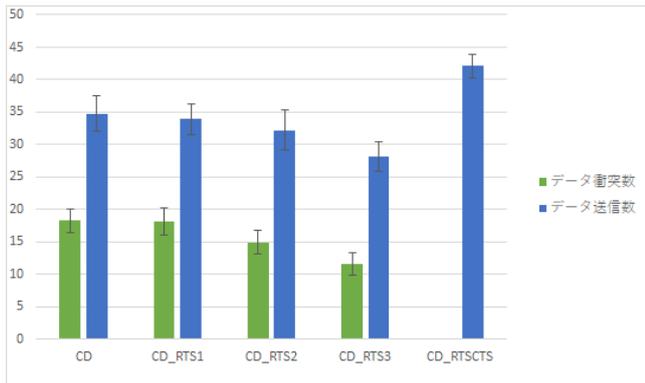


図 5: シミュレーション結果

表 3: データ衝突数の分散分析結果

変動要因	平方和	自由度	平均平方	F 値
群間	901.892	3	300.631	20.145
郡内	1731.10	116	14.923	
全体	2632.99	119	22.126	

データ送信数においては、CD と CD_RTS3、CD_RTS1 と CD_RTS3、CD_RTS2 と CD_RTS3 の間で有意差が見られた。

5 考察

CD は CD_RTS1、CD_RTS2、CD_RTS3 と比較し、データ衝突数は多いものの、データ送信数も多いという結果となった。また、データ衝突数においては、CD と CD_RTS2、CD と CD_RTS3 の間で、データ送信数においては CD と CD_RTS3 の間で、有意差が見られた。データ衝突数が減少することで、データ衝突後のランダム待機時間※1 (図 3) の発生回数が減少し、それによってデータ送信開始タイミングの発生回数が多くなったためと考えられる。このことから、データ送信前に RTS を送信し、2sec 待機した後、データ送信を開始することで、データ送信数は 0sec 待機と変わらず、データ衝突数を減少させることが可能であるといえる。実際の遠隔対話の場面においては、発話を行う前に、発話予備情報の発信を行った後、一呼吸おいて (2sec ほど) 発話を開始することで、発話衝突の発生を抑えつつも、発話数は減少しない対話が可能であると考えられる。

CD と CD_RTSCTS を比較すると、CD_RTSCTS のほうがデータ衝突数が少なく、データ送信数も多いという結果になった。ネットワーク通信において、RTSCTS を用いたデータの送受信は、スループットが半分以下

表 4: データ送信数の分散分析結果

変動要因	平方和	自由度	平均平方	F 値
群間	763.958	3	254.653	8.900
郡内	3319.03	116	28.612	
全体	4082.99	119	34.311	

表 5: データ衝突数の多重検定結果

比較	t 検定統計量	検定結果
CD, CD_RTS1	0.100	$P > 0.05$
CD, CD_RTS2	3.342	$P < 0.05$
CD, CD_RTS3	6.680	$P < 0.05$
CD_RTS1, CD_RTS2	3.242	$P > 0.05$
CD_RTS1, CD_RTS3	6.583	$P < 0.05$
CD_RTS2, CD_RTS3	3.342	$P < 0.05$

になるという課題があるが、人の対話モデルには影響を及ぼさないという結果となった。これは、CTS を用いてデータ送信を制御することで、データ衝突が発生することが無くなり、通信路からの衝突信号の送信とランダム待機時間※1 (図 3) の消費フェーズに遷移することが無くなるため、結果としてデータ送信数も増加したためと考えられる。また、ネットワーク通信において、データを送信する際、RTSCTS 情報 (ヘッダー) を送信データに付与するためその分送信データ量が減少する問題があるが、今回のシミュレーションでは送信データにヘッダーを付与しなかったため、この問題が発生しなかった点も理由として考えられる。実際の遠隔対話の場面においては、司会者を設けて、遠隔対話参加者の発言権をコントロールすることで発話衝突が発生することが無くなり、発話数も向上すると考えられるが、シミュレーション環境をより実際の遠隔対話環境に近づける必要があると言える。

一例として、CD_RTSCTS において、衝突が発生するようにアルゴリズムを改良する必要がある。RTS と CTS の送受信が、シミュレーション結果にどのような影響を及ぼすか調査するため、CD_RTSCTS 環境では衝突が発生しないアルゴリズムでデータ送信を行ったが、実際の対話では、発話予備情報の送信を何度も行うと、我慢できなくなり発話を開始してしまい、その結果発話が衝突してしまうことがある。この状況を再現するために、シミュレーションでは、RTS の送信を閾値 x 回以上行うと、CTS の受信を待たずにデータ送信を開始するようにアルゴリズムを改良することで、データ衝突が発生するようになり、より実際の対話環境に近付くと考えられる。

表 6: データ送信数の多重検定結果

比較	t 検定統計量	検定結果
CD, CD_RTS1	0.628	$P > 0.05$
CD, CD_RTS2	1.883	$P > 0.05$
CD, CD_RTS3	4.755	$P < 0.05$
CD_RTS1, CD_RTS2	1.255	$P > 0.05$
CD_RTS1, CD_RTS3	4.127	$P < 0.05$
CD_RTS2, CD_RTS3	2.872	$P < 0.05$

6 まとめ

Skype などのネットワークに繋がれたコミュニケーションツールを用いて遠隔地間で対話を行う遠隔対話システムの普及が進んでいるが、映像と音声のみを用いた遠隔対話では、誰がいつ話し始めるのかを判断するのが難しいため、2人以上が同時に発話を開始してしまう発話衝突が多く発生することが問題となる。

本研究では、データ衝突を軽減する手法としてインターネットで用いられている通信プロトコルのひとつである CSMA (Carrier Sense Multiple Access) に着目し、人の対話をネットワーク通信におけるデータ通信に置き換え、ネットワークシミュレータを用いて人の対話の再現を試みた。発話衝突発生の原因として、送信要求 (RTS) の送信を行わず発話を開始することと、送信要求 (RTS) の送信を行うが受信許可 (CTS) を受信したと勘違いし発話を開始することの 2 つを仮定し、5 つの条件でシミュレーションを行い、データ衝突数とデータ送信数の比較を行った。

シミュレーションの結果、RTS を送信した後、2sec 待機しデータ送信を行う事で、RTS を送信せず、待機も行わない場合と比較し、データ衝突数に有意差がみられるものの、データ送信数には有意差が見られないという結果になった。また、RTSCTS を用いたデータ送信を行う事で、データ衝突数は 0 であり、データ送信数は 5 つの条件の中で最も多いという結果になった。このことから、遠隔対話システムにおいては創造会議ではなく司会者を立てて対話をコントロールしたほうがよいと考えられる。しかし、人の対話においては衝突が起こるが今回のシミュレーションでは衝突が起こらない場合があるなど、シミュレーション環境の見直しが必要であることも分かった。

今後は、ネットワークシミュレーションにおいてアルゴリズムの改良を行い、より人の対話に近いシミュレーション環境を構築することが必要である。また、本稿で報告したシミュレーション結果をもとに対人実験を行い、シミュレーションと同等の効果が得られるかどうかの検証を行う予定である。将来的には、提案インタラクションモデルをロボットに実装し、ロボット

の自律対話システムを構築できる可能性がある。

参考文献

- [1] 鎧沢勇, 滝川啓, 大久保栄, 渡辺義郎. 衛星通信を利用した画像会議におけるエコー及び伝搬遅延の影響. 電子通信学会論文誌 B, Vol. 64, No. 11 (1981), pp. 1281-1288.
- [2] 高橋誠. 会議の進め方, 日本経済新聞出版社 (1987).
- [3] 玉木秀和, 中茂睦裕, 東野豪, 小林稔. 人のコミュニケーションリズムに着目した Web 会議円滑化手法. IEICE Technical Report MVE2009 (2009), pp. 101-106.
- [4] Sacks H, Schegloff AE, Jefferson G. A Simplest Systematics for the Organization of Turn-Taking for Conversation Language, Vol. 50, No. 4, Pt1, pp. 696-735(1974)
- [5] 玉木秀和, 東野豪, 小林稔, 井原雅行. 遠隔会議における発話の衝突と精神的ストレスの関係. グループウェアとネットワークサービス研究会, Vol. 2011-GN-79, No. 10, pp. 1-6(2011).
- [6] 堀口純子. コミュニケーションにおける聞き手の言語行動. 日本語教育, 64 号, pp. 13-25(1988).
- [7] Vargas, M. F., Louder than words: an introduction to nonverbal communication, Iowa State University Press, 1986
- [8] SAKAI, Kurima, et al. 発話者の音声に対応する動作生成と遠隔操作ロボットへの動作の付加効果. AI チャレンジ研究会 (第 39 回), 2014
- [9] 長谷川孔明, 中内靖. テレプレゼンスロボットによる無意識的身ぶりの表出が発話交替に与える影響. 日本機械学会 (2014)
- [10] 玉木秀和, et al. 発話がぶつからない Web 会議を実現するための発話欲求伝達手法. 情報処理学会 (2013)
- [11] 石井亮, et al. 複数人対話における視線交差のタイミング構造に基づく次話者と発話開始タイミングの予測. 人工知能学会全国大会 (2015)
- [12] Andrew S. Tanenbaum, David J. Wetherall. Computer Networks. Prentice Hall, 2010