

Web 会議における発話衝突回避のための 発話欲求伝達手法

玉木秀和[†] 東野豪[†] 小林稔[†] 井原雅行[†]

本稿では、手軽に使用できる Web 会議において、発話の衝突を低減し、円滑な話者交替を可能にすることを目指し、参加者各々に発話欲求インジケータを表示させ、参加者の動作から推定される発話欲求度をリアルタイムに表示し、その発話欲求度が、設定された閾値を超えた場合、その参加者の映像に黄色い枠を表示し、他の参加者へと知らせる手法を提案する。小規模な評価実験を実施した結果、提案手法により発話欲求を他の参加者へ伝えられ、発話衝突を低減できる可能性が示唆された。

A Method Communicating Desire of Speaking to Avoid Speech Contention in Web Conferences

Hidekazu Tamaki[†] Suguru Higashino[†]
Minoru Kobayashi[†] and Masayuki Ihara[†]

Our aim is to avoid speech contention in Web conferences. We propose a method communicating desires of speaking. Participants can check how strong their desires are detected with an indicator. And if desire level increases over a threshold, frame of his/her picture is set color to yellow to indicate his/her desire to other participants. We took a brief test and investigated that our concept has a potential to avoid speech contention and realize a smooth turn taking in Web conferences.

1. はじめに

CO₂ 排出量低減、節電施策、世界不況、パンデミック対策などが叫ばれ、遠隔会議システムの需要が増加してきている。遠隔会議システムの中でも Web 会議システムは、自席で利用できる上、ソフトウェアのみで導入できるなどの手軽さがあり、市場も成長傾向にある[1]。

しかし、Web 会議は TV 会議専用システムに比べて、デスクトップ上で利用し、転送効率が悪いネットワークを通じて行うという特性上、個々の参加者の映像が小さい、フレームレートが低い、画質が悪い、リップシンクがとれない、映像、音声遅延が生じるなどの制限がある。このため他の参加者の様子を読み取りづらく[2]、誰がいつ発話し始めるのかを判断しにくい。そして、いざ発話を開始すると、他の参加者の発話と衝突してしまう、ということが多い。音声遅延が 300ms を超えると、特にこの発話の衝突が顕著になると報告されている[3]。また、Sacks は発話が発話すると発話を諦めて中断する傾向が高いことを指摘している[4]。このような状態では、素早い話者交替を頻繁に行う会議はしづらく[3]、消極的な会議になりかねない。著者らの分析結果では、盛んに話者交替の起こる、司会者のいない創造会議[5]（アイデアを考え出す会議）を行う場合に、Web 会議では対面会議と比較して発話の衝突が 30 倍近く起こることが分かった[6]。

人は普段、対面したコミュニケーションの場面では顔色や姿勢の変化など、非言語情報をうまく利用して発話の衝突を避け、円滑に話者交替している[7][8]。しかし、Web 会議では先に述べた制約のため非言語情報を利用しにくく、発話の衝突が起こりやすい。

本研究の目的は、Web 会議において発話の衝突を低減し、素早く頻繁に話者交替する会議を実施可能にすることである。著者らはこれまで、Web 会議において、人が発話の前に行う特徴的な動作（以降「予備動作」と呼ぶ）を検知して、最も次に発話しそうな参加者を選定し、全参加者へ示すことで話者交替を円滑化する手法を提案してきた[9]。検討を重ねた結果、参加者の意図と、システムが提示する情報とが近い程、参加者はその情報を受け入れ、滞りなく話者交替を行うことが可能になることが分かってきた。そこで本稿では、システムに推定されている発話欲求を参加者ごとにインジケータで確認でき、そのインジケータがある閾値を超えて振り切れた時に、それを前参加者へと通知する手法（図 1）を提案する。

[†] 日本電信電話株式会社 サイバーソリューション研究所
NIPPON TELEGRAPH AND TELEPHONE CORPORATION NTT Cyber Solutions Laboratories

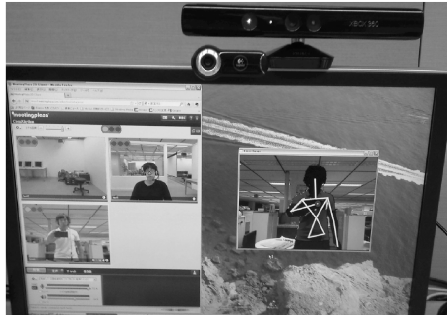


図1 発話欲求を伝達可能な遠隔会議システム

2. ビデオ会議と話者交替

様々な映像コミュニケーションシステムが開発され、伝達できる情報量が増加している。しかし、Sellen はこれらにはまだコミュニケーションを行う際に弊害があると指摘した[10]。1対1の遠隔コミュニケーションであっても、遅延のある環境では発話の衝突が多く起こる[3]。1対1のコミュニケーションであれば、発話権が交互に遷移するが、多人数のコミュニケーションではそれは複雑になる。商用化されている会議システムにも発話権の遷移をサポートする機能が実装されているものがある。ボタンを押してマイクのオン・オフを切り替えることにより発話権を獲得、放棄する機能や、挙手アイコンを表示させるためのボタンを備えたものもある[11]。人と機械の会話を自然に感じさせるために発話前の非言語情報を表現する機能をエージェントシステムに組み込んだ試みも存在している[12][13]。石井らはアバタを介した遠隔コミュニケーションシステムにおいて、次に発話することを促す視線を表現する手法を提案した[14]。しかしこの手法では、音声に基き動作を疑似的に作り出しており、実際の参加者の動作は反映されない。川島らは2者間の遠隔コミュニケーションにおいて、フィラー（間を埋める音声）を、映像によって表現した[15]。しかし参加者は2人に限定され、これも実際の参加者の動作は反映されない。

3. Turn-take Smoother の概要

3.1 研究目標

我々の研究目標は、Web 会議が持つ手軽さ・使用場所の制約の低さはそのままに、発話衝突せず円滑に話者交替し、気軽に質の高い遠隔会議ができるという、新しいコミ

ュニケーションメディアを“近い将来に実用化できるレベル”で実現することである。より具体的には、人が話者交替をするときに使用する予備動作の伝達を支援し、円滑な話者交替を可能にする仕組みを想定している。これを実現するには、下記3つの要件を満たす必要がある。

要件1) 普及している Web 会議システムを基とする
実現性を高めるため、実験用のための大きな機材を用いたシステムではなく、実際に動き、普及している商用システムを基とした実装を行う。

要件2) ユーザに特別なデバイスを装着させない
脈波、脳波、その他をセンシングするデバイスを増やし、ユーザへの装着を要求することは、導入しやすさ、利用しやすさを損ない、さらに使用時のストレスを増加させる恐れがある。

要件3) 近い将来に普及する可能性のあるレベルのデバイスを用いる
カメラ、ヘッドセットなどの各種デバイスは、数年のうちにデスクトップ PC に組み込まれる可能性のある程度の性能を備えたものを用いる。

3.2 Turn-take Smoother のコンセプト

上記要件を満たす方法を検討する。(1)まず、要件1のように既に広く普及している商用 Web 会議システムとして、MeetingPlaza[11]を用いる。国内 Web 会議システムにおいてトップシェアを獲得しているため、考案したアイデアを組み込み、効果を検証するために最適であると考えた。(2)次に、要件2を勘案し、センシングデバイスとして、装着型の脈拍センサ、脳波計、視線計測機などは使用せず、現在 Web 会議システムを使用する際に装着する最低限のデバイスであるヘッドセットと、映像を撮影するためのカメラのみを使用することとした。(3)センシングデバイスであるカメラのうち、要件3を満たす中で最も精度の高いデバイスとして、Kinect センサ[16]を使用する。Kinect センサは、比較的小型のデバイスながら高精度に映像とその奥行き情報を取得できる。さらに価格も低く、数年後にはデスクトップ PC、ひいてはノート PC にさえ、この程度の精度のカメラが搭載されることは不可能ではないと考える。

(1)~(3)を踏まえると、映像の表示領域は、従来の Web 会議システムと同等に、限られた大きさとなる。すなわち、できる限り高解像度のディスプレイを用いても、複数人の参加者へ割り当てるために分割された画面の表示サイズは小さく、各々の参加者の非言語情報を取りこぼさず読み取ることは困難である。そこで著者らは、発話する前に出現する予備動作をシステムがセンシングし、それをシグナルへと変換して表示し、他の参加者へと伝えることで、それを読み取った参加者が、それを活用し、円滑な話者交替ができる手法の実現を目指す。

4. 発話欲求フィードバックインジケータ表示手法の提案

4.1 研究課題

前章で述べたコンセプトを実現するには、予備動作を検知して、シグナルへと変換しなくてはならない。対面会話における話者交替の場面で、ある参加者に発話欲求が生じてから実際に行動するまで、図2の上段のようなステップを踏むと考えた。発話欲求が生じたときに、著者がこれまでに明らかにした予備動作[9]、すなわち「頷き」「相槌」「手を顔周辺へ動かす」「体を前に乗り出す」といった動作を、随意あるいは不随意に行う。そして他の参加者の予備動作を認知し、自身が発話するか、発話権を他の参加者へ譲るかを判断し、実際に発話する、もしくは発話権を譲る。このステップの流れを基に、Web会議における話者交替を考えると、これまでも述べている通り、「予備動作」と「認知」ステップの間が上手く流れず、誤った判断で行動するため、同時に複数の参加者が発話開始する状況が増えると考えられる。

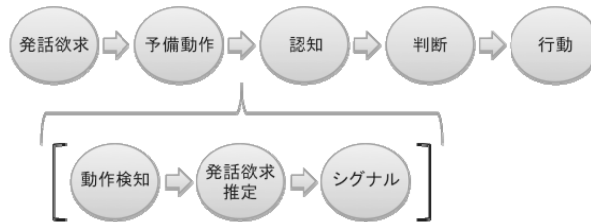


図2 発話欲求が生じてから行動するまでのステップ

そこで著者が行うアプローチは、「予備動作」と「認知」ステップの間に、図2下段にあるような3つのステップを行う仕組みを挿入することである。すなわち、参加者の行った予備動作を検知し、それを基に、その参加者がどれだけ発話欲求を持っているかを推定する。そして推定した結果を基に、その参加者が発話欲求を持っていることを、他の参加者へと伝えるためのシグナルを提示する。このシグナルを適切に提示することができれば、それを認知することにより、各々の参加者が、他の参加者の発話欲求を知り、正しい判断の基に行動できると考える。

我々はこれまで、検知した予備動作を基に発話欲求を推定し、そのうち最も発話しそうな参加者一人を次話者候補として選定し、その参加者の映像に黄色い枠をシグナルとして表示し、強調する方法を取ってきた[6]。しかし、Webカメラの画像のみを用いた予備動作の検知精度は完璧ではなく、発話欲求を正確に推定するモデルも確率されていないため、次話者の選定結果が実際の参加者の意図と異なっていた場合に、参加者を躊躇させてしまうことがあった。つまり、提示されるシグナルが、元々の発話欲

求と乖離した場合に、参加者の判断を狂わせてしまい、さらにはその乖離度合いが高まると、シグナルへの信用度合いも下がることが分かってきた。

そこで本稿では、発話欲求を正しく反映したシグナルの表示手法の確率を研究課題として設定する。

4.2 提案手法

前節で述べた研究課題を解決するためには2つの条件がある。1つ目は、予備動作自体を検知する精度を高めることである。2つ目は、検知された予備動作を基に、正しく発話欲求を推定することである。そこで著者は、まず1つ目の条件をクリアするために、2次元画像とその奥行き情報を高い精度で検出できる Kinect センサを採用する。参加者の予備動作を正確に検知できたとしても、そこから発話欲求を正確に推定することは困難である。そこで著者は発想を転換し、参加者へ、推定されている自身の発話欲求度合をリアルタイムにフィードバックし、参加者自身が動作調整することにより、発話欲求度合と、表示されるシグナルの整合率を高めることのできる仕組みを提案する。すなわち、参加者各々に発話欲求インジケータを表示させ、参加者の動作から推定される発話欲求度合をリアルタイムに表示し、その発話欲求度合が、設定された閾値を超えた場合、その参加者の映像にシグナルを表示し、他の参加者へと知らせる。このときのシグナルは、映像の枠に黄色い枠を表示させるものとする。この方法を選択した理由は、使用した MeetingPlaza では、発話しているときに、参加者の映像の枠に赤色の枠が表示されるよう設定されているため、発話につながる、前の状態ということが分かりやすいと考えた。

これにより、表示されているインジケータが、実際の発話欲求よりも高い場合、随機的に動作を控え、インジケータを下げるのが可能となる。自身で調整可能である。逆に、発話欲求が高い場合には、インジケータを確認しながら、予備動作を行うことで、インジケータを上げて、他の参加者の注意を引くことが可能となる。この予備動作は、ボタンを押す動作などとは異なり、普段対面会話において話者交替する際に使用しているものなので、Web会議中の話者交替においても自然に取り入れられると考ええる。

5. 発話欲求フィードバックインジケータ生成モジュールの実装

会議システム全体の系としては、図3のようになる。各クライアントマシンにおいて、Kinect センサから得られた情報を基に各参加者の予備動作を検知し、そこから発話欲求フィードバックインジケータや、参加者映像へ黄色い枠を表示させる要求を MeetingPlaza Client システムへ送る。そこからインターネットを経由し、サーバマシンを介し、他のすべてのクライアントマシンへと情報が共有され、発話欲求フィードバックインジケータや黄色い枠の状態が同期される。MeetingPlaza はブラウザ上で動作

する会議システムだが、その画面上に表示される映像は Web カメラから取得し、マイクから取得した音声とともに、MeetingPlaza Client システムへ取り込まれ、後は前述と同様の流れで、各クライアントマシンへ共有される。今回、Kinect センサの他に、MeetingPlaza 画面上に共有される映像の取得用として Web カメラを使用した。Kinect センサと同等レベルのデバイスが PC に組み込まれた際にはそれに一本化し、そのセンサから取得した映像を予備動作の検知用と、MeetingPlaza 画面用に用いることを想定している。1 人の参加者が使用するシステムの環境は図 4 の通りである。

発話欲求フィードバックインジケータ生成モジュールは、予備動作検知パートと発話欲求フィードバックインジケータ可視化パートから成る。

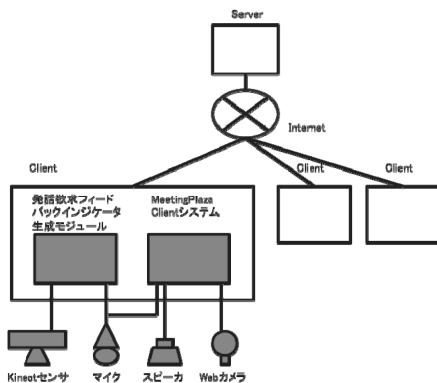


図3 会議システム全体の構成図

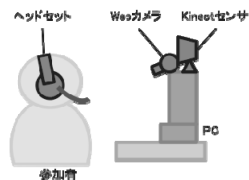


図4 1人の参加者のシステム使用環境

5.1 予備動作検知パート

Kinect センサを用い、人を認識し、その頭部、首、肩関節、肘関節、手首などの三次元座標をリアルタイムに取得する。それらの情報から、著者らがこれまでに、人が

発話する前に行う予備動作であると明らかにした、下記4つの予備動作[9]を検知した。

- ・ 手を挙げる
- ・ 手を顔周辺へ持っていく
- ・ 頷く
- ・ 身体を前に乗り出す

<手を挙げる>

手首の位置が、肘関節の位置よりも、天地方向で、高くなった場合、手を挙げていると判定した。

<手を顔周辺へ持っていく>

手首と頭部の三次元座標の距離が、ある一定距離よりも短くなった場合、手を加尾周辺へ持っていく予備動作であると判定した。

<頷く><身体を前に乗り出す>

今回、これら2つの動作に関しては、同じ判定基準を適用した。すなわち、奥行き方向の軸を z 軸とし、Kinect センサから離れる方向を正とすると、頭部の z 座標値を h_z 、首の z 座標値を n_z 、ある時間経過した後の z 座標値をそれぞれ h_z' 、 n_z' とすると、

$$h_z - n_z > h_z' - n_z' \quad \dots (1)$$

が満たされる場合、頷く、もしくは身体を前に乗り出している、と判定した。

5.2 発話欲求フィードバックインジケータ可視化方法

前節で検知した予備動作を基に、参加者毎に発話欲求フィードバックインジケータを表示した(図5)。予備動作ごとに一定の値を決め、検知された予備動作に応じた値を発話欲求度として加算した。発話欲求度は0から100までの値を取り、90を超えたときに、その参加者映像の枠を黄色く強調した。発話欲求フィードバックインジケータは、参加者映像下部左側に緑色で表示させ、同右側にはマイク入力音声レベルインジケータを赤色で表示させた。発話欲求度が100のとき、丁度左半分を緑色の発話欲求フィードバックインジケータが埋めるようにした。マイクから入力音声がある場合、発話欲求度は常に100となるようにした。また、発話欲求度は時間経過とともに一定の速度で減少させるようにした。



図5 発話欲求フィードバックインジケータ

6. 評価実験

6.1 実験概要

この実験の目的は、Turn-take Smoother の中核技術である、発話欲求フィードバックインジゲータ表示手法の長所・短所を洗い出すことである。具体的には、3 人の被験者に Web 会議システムを用いて研究議論をするタスクを行い、提案手法の(1)発話欲求を他の参加者へ伝えることができるか(2)発話衝突確率は低減できるかを検証する。比較条件として、通常の Web 会議システムを使用した条件と、発話欲求を伝えるために「挙手ボタン」を使用した条件とを用意した。今回はまず小規模な実験を行ったが、今後大規模な実験を行い、有効性を確かめる必要がある。

6.2 実験方法

20 代男女 3 人の被験者を会議の参加者とした。被験者はすべて、ヒューマンインタフェースの研究者であった。会議には MeetingPlaza を使用した。会議参加者の映像は、2 人分を横に並べて配置し、もう 1 人を、下に配置した。すべての参加者は、マイクとヘッドホンの組み込まれたヘッドセットを装着した。いずれかの参加者のマイクに音声入力があった場合、その参加者映像に赤色の枠が表示されることで、その参加者が発話中であることが、全ての参加者へと伝えられた。

4 つの条件を用意した。1 つ目は「通常の Web 会議」条件とし、上記記者を知らせる機能のみを使用し、参加者は特別な操作をしないものとした。2 つ目は「挙手ボタン」条件とし、通常の Web 会議条件に加えて、各参加者が、自身の映像の右下に表示された挙手ボタンをクリックすることにより、その参加者映像の左上に挙手アイコンが表示される機能を使用することとした。次に発話したいことを他の参加者へ知らせるため、挙手ボタンをクリックしてアイコンを表示させ、他の参加者へと知らせるよう指示した。3 つ目は「黄色枠」条件とし、Turn-take Smoother を使用し、発話欲求フィードバックインジゲータと、発話欲求を他の参加者へ伝えるための黄色枠表示機能を ON にした。被験者には、喚知される予備動作の種類を教示し、黄色枠を使用して発話欲求を伝えるよう指示した。図 1 のように、画面右側に、自身の映像と、予備動作が検知されているか否かをモニタリングできるウィンドウを表示した。4 つ目は「黄色枠・静止画」条件とし、上記「黄色枠」条件と基本的に同じ条件とし、MeetingPlaza 画面上の、それぞれの参加者映像を静止画とした。「通常 Web 会議」「挙手ボタン」「黄色枠」3 つの条件にて、7 分間ずつ使用し、操作に慣れてから実験に臨んだ。

会議の内容は、ヒューマンインタフェース研究について議論するというものであった。4 つの条件を、「通常の Web 会議」「挙手ボタン」「黄色枠」「黄色枠・静止画」の順に、それぞれ 7 分ずつ議論を行った。議題は、実験を通して同じものを用いた。実験中の参加者のモニタをビデオカメラで撮影した。後にビデオを分析し、各条件とも 5 分間の発話回数と発話衝突回数を記録した。条件を切り替えた直後は、参加者の意識が、

変化した環境へ適応することへ向くため、各条件とも、開始 1 分から 6 分までの 5 分間を分析対象とした。

また、実験後に、各条件の比較について、自由記述形式でアンケートを行った。

6.3 実験結果

条件ごとに、ターン数、発話回数、発話衝突回数、発話衝突確率（発話衝突回数／発話回数）を以下の表 1 にまとめる。また、挙手ボタン／黄色枠の使用状況の比較を表 2 にまとめる。

表 1 発話／衝突状況の比較

条件	ターン数	発話回数	発話衝突回数	発話衝突確率
通常 Web 会議	33 ターン	41 回	4 回	9.8%
挙手ボタン	13 ターン	15 回	2 回	12.5%
黄色枠	33 ターン	35 回	1 回	2.9%
黄色枠・静止画	44 ターン	49 回	3 回	6.1%

表 2 挙手ボタン／黄色枠使用状況の比較

条件	挙手／黄色枠表示回数	挙手／黄色枠表示後発話回数	挙手／黄色枠表示せず発話した回数	挙手／黄色枠表示後発話した確率	全発話中、挙手／黄色枠表示後に発話した確率
挙手ボタン	12 回	11 回	5 回	91.7%	68.8%
黄色枠	24 回	21 回	14 回	87.5%	60.0%
黄色枠・静止画	37 回	32 回	16 回	86.5%	65.3%

試行回数が少ないため、傾向を調べることはできないが、挙手ボタン条件で発話回数が少なかった。また、黄色枠を表示した条件では、発話衝突確率が少なかった。挙手ボタン／黄色枠使用状況においては、各条件とも挙手／黄色枠表示後に発話した確率が高かった。

6.4 考察

挙手ボタン条件では、挙手後の発話確率が高いことから、発話欲求を明確に伝えられていることが分かる。しかし、発話が長くなり、ターンテイキングの頻度が減少した。さらに、挙手ボタンを下げ忘れる場面も散見された。2 人の参加者が同時に挙手し、片方の参加者が挙手状態を解除することでもう一方の参加者へターンを譲る場面が 1 回見られた。発話衝突が起こった場面では、誰も挙手状態になっている参加者がいないときに、2 人の参加者が同時に発話開始してしまった場面が 1 回と、1 人の参加者が挙手していたが、他の 1 人の参加者も、それに気づかず、同時に発話してしまった場

面が1回観察された。

黄色枠条件で発話衝突の起こった場面では、2人の参加者映像に同時に黄色枠が表示され、そのまま2人が同時に発話してしまった。

実験後に行ったアンケートによると、最も会話しやすかったのは「黄色枠」条件だったと答えた被験者が多く、最も会話し辛かったのは「通常 Web 会議」条件だったと答えた被験者が多かった。

アンケートから、「意識しなくても、発話欲求の高まったときに黄色枠が表示されることが多かった」「黄色枠で発話欲求を伝えることができた」「黄色枠により注意を引きつけられた」「発話欲求がないときにインジケータが高まった場合は、自身で動きを調整し、インジケータを下げることができた」といった意見が得られた。

しかし、挙手ボタンと比較し、「挙手ボタンは確実な発話欲求を伝える。挙手ボタンの方が、話者の意図（話したいのか、相手に話す権限をゆずったのか）が明確なので、遠慮せずに話せた。挙手ボタンの方が、話の始めと終わりが明確にわかるので、会話の衝突が少ない気がした。黄色枠は誰が話しているのかわからないことが多かった」という意見も得られた。挙手ボタンは、発話する前に挙手状態にし、発話終了したときにそれを解除する、という使い方をされ、発話開始点、発話終了点が明確だった点が、次の話者が安心して発話開始できることに貢献したと考えられる。

想定外のコメントとしては「黄色枠は発話欲求が生まれる前の、興味を持って聞いている段階でも表示された。興味を持って聞いている段階、発話欲求の生まれた段階と、多段階にレベルを設定してほしい」「発話欲求以外に、興味を持って聞いているときにも表示されることが多く、興味を持っている、盛り上がっているというのが分かり、音声の入っていないときに安心感を与えた」「黄色枠・静止画条件でも、静止画でないような感じがした」といったコメントが得られた。

総じて、黄色い枠は、挙手ボタンに比べ、発話欲求を伝える手段として確実性は下がるが、自然に表示させられ、自身で調整ができ、他の参加者の注意を引くことができ、発話衝突確率を低減させられる可能性がありそうということが分かった。本当に発話欲求があるレベルと、会話に興味を持っているレベルなどの個人差を吸収することで、より円滑な話者交替が実現可能になると考える。

7. おわりに

本稿では、手軽に使用できる Web 会議において、発話の衝突を低減し、円滑な話者交替を可能にすることを目指し、参加者各々に発話欲求インジケータを表示させ、参加者の動作から推定される発話欲求度合をリアルタイムに表示し、その発話欲求度合が、設定された閾値を超えた場合、その参加者の映像に黄色い枠を表示し、他の参加者へと知らせる手法を提案した。

小規模な評価実験を実施した結果、提案手法により発話欲求を他の参加者へ伝えられ、発話衝突を低減できる可能性が示唆された。また、当初想定していなかった効果として、黄色枠が表示されることで、発話欲求の他に、興味がある、盛り上がっているという情報を伝えられると感じ、音声途切れたときや、映像が静止しているときにも安心できる、といった意見が得られた。

今後は使用する予備動作の頻度や種類の個人差を埋めるため、予備動作が検知されてから実際に発話するか否かを変数に、リアルタイムな学習を行い、発話欲求度へ加算する値を変化させていく、などの方法を検討している。

参考文献

- 1) 2008年版テレビ会議/Web会議の最新市場とHD化動向；シードブランニング(2008)
- 2) 徳麿, 友保康成, 渋谷雄, 田村博: テレビ会議技術の課題と利用法についての考察, 8th Symposium on Human Interface, pp.207-212, 1992.
- 3) 鑑沢勇, 滝川啓, 大久保榮, 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.J64-B, No.11, pp.1281-1288, 1981.
- 4) Sacks H., Schegloff A. E., Jefferson G.: A Simplest Systematics for the Organization of Turn-Taking for Conversation, Language; 50, 4 Pt 1, 696-735, 1974.
- 5) 高橋誠: 会議の進め方, 日本経済新聞出版社, 1987.
- 6) 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, pp.101-106, 2009.
- 7) 松尾隆: コミュニケーションの心理学, ナカニシヤ出版, 1999.
- 8) マジョリー・F・ヴォーガズ: 非言語コミュニケーション, 新潮社, 1987.
- 9) 玉木秀和, 中茂睦裕, 東野豪, 小林稔: 遠隔会議における話者交替円滑化手法, マルチメディア, 分散, 協調とモバイルシンポジウム DICOMO2010, pp.108-116, 2010.
- 10) Abigail J. Sellen: SPEECH PATTERNS IN VIDEO-MEDIATED CONVERSATIONS, CHI'92, pp.49-59, 1992.
- 11) <http://www.meetingplaza.com/index-j.html> (2011年10月現在)
- 12) Cassell J. and Vilhja'lmsson H.: Fully Embodied Conversational Avatars: Making Communicative Behaviors Autonomous; Autonomous Agents and Multi-Agent Systems, 2, pp.45-64 (1999).
- 13) López B., Hernández Á., Díaz D., Fernández R., Hernández L.: Design and validation of ECA gestures to improve dialogue system robustness, Proceedings of the Workshop on Embodied Language Processing, pp. 67-74, 2007.
- 14) 石井亮, 宮島俊光, 藤田欣也: アバタ音声チャットシステムにおける会話促進のための注視制御, ヒューマンインタフェース学会論文誌 Vol.10, No.1, 2008.
- 15) Kawashima H., Nishikawa T., Matsuyama T.: Visual Filler: Facilitating Smooth Turn-Taking in Video Conferencing with Transmission Delay, CHI 2008 Extended Abstract, pp. 3585-3590, 2008.
- 16) <http://www.xbox.com/ja-JP/kinect> (2011年10月現在)