

遠隔コミュニケーションにおける
発話衝突低減手法

平成24年度

玉木 秀和

主 論 文 要 旨

報告番号	甲 乙 第	号	氏 名	玉木 秀和
主 論 文 題 目： 遠隔コミュニケーションにおける発話衝突低減手法				
(内容の要旨) 環境負荷低減, 出張費削減などが求められるなか, 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なかでも Web 会議システムは, 複数の離れた拠点にいる会議の参加者が自席のデスクトップパソコンやノートパソコンから容易に参加できるシステムである. しかし, 複数の拠点を結んで行う Web 会議において, 参加者が互いの様子を正確に読み取るとは困難である. 個々の参加者の映像はディスプレイを分割した領域に表示されるため, 小さくなってしまいがちである. また, 帯域の保証されていないネットワーク上で Web 会議を実施する場合, 映像や音声は遅延して伝達される. そのため, Web 会議において話者交替が頻繁に起こる会議を実施した場合, 参加者同士の発話が衝突することが多く見受けられる. 新しいアイデアを考えだすための会議は創造会議といわれ, そのような会議では話者交替が頻繁に起こる. したがって, Web 会議において創造会議を行う場合, 発話衝突は大きな弊害になると考えられる. 本研究は, Web 会議での発話衝突を低減することにより, 自席のデスクトップパソコンやノートパソコンからでも, 創造会議を快適に行える遠隔会議環境を実現することを目的とする. 発話衝突を低減させる手法を模索するために, まず発話衝突の原因とその影響について調査した. Web 会議において, 音声遅延が発話衝突に大きな影響を与えるという仮説をたて, 実験を行った. 音声遅延量を変化させられる音声会議環境を用意し, 音声遅延量と発話衝突, またそれらに応じて参加者の受ける精神的ストレスの関係を調べた. 本実験環境では, 音声遅延量が 400msec を超える場合には発話衝突確率が増大し, 精神的ストレスが高まる結果となった. 本研究が対象とする Web 会議の環境では音声遅延量が 400msec を超えることが多いため, 発話衝突確率を低減させることが望まれる. 本論文で提案する発話衝突低減手法は, 次話者候補提示手法と発話欲求伝達手法から構成される. 次話者候補提示手法は, システムが会議参加者各々の予備動作を検知し, その種類と頻度から, 次に最も発話しそうな参加者を次話者候補として選定し, すべての参加者へ伝達する手法である. 発話欲求伝達手法は, システムによって推定された発話欲求の度合いを随時参加者に提示することにより, 参加者自身に予備動作の調整を促し, 発話欲求の推定精度を高める手法である. これらの手法の有効性を確認するため, 両方の機能を組み込んだプロトタイプシステムを実装し, 通常の Web 会議システムとの比較実験を行った. 実験結果より, 本提案手法が発話衝突確率を半分以下に低減できることを確認した.				

SUMMARY OF Ph.D. DISSERTATION

School School of Science for Open and Environmental Systems	Student Identification Number 80946801	SURNAME, First name TAMAKI, Hidekazu
Title A Method of Reducing Speech Contention in Distributed Communication		
Abstract <p>This research aims to realize an environment that facilitates creative, idea-generating meetings among remote participants. Currently, web conferencing is widely adopted because people can save on installation and travel costs. Web conferencing systems, however, provide inferior video and voice quality compared to specialized video conferencing systems, making it difficult for meeting participants to be aware of the other participants' movements and subtle changes in facial expression. Such deficiencies make it difficult to detect when others want to take the floor, and speech contentions become frequent.</p> <p>One of experiments this thesis describes investigates the influence of speech delay on speech contention and psychological stress. Results show that greater speech contention rates occur when audio delay is over 400 milliseconds, and longer audio delays cause higher stress.</p> <p>This thesis introduces a concept in which a system detects participants' nonverbal information and then presents that information to other participants. It is well known that face-to-face communication reduces speech contention and facilitates smooth turn-taking through the use of nonverbal information. However, participants cannot detect others' nonverbal information in Web conferences. The proposed method introduced in this thesis determines the participant who is most likely to speak based on several common body motions and vocal cues that precede floor-taking, and then introduces him/her to the other participants.</p> <p>Several improvements to the proposed method have also been made in order to increase the accuracy of communicating participants' desire to speak. These improvements allow participants to use an indicator to visually confirm how strong their own "desire to speak" is as detected by the system, and then adjust their behavior to ensure the indicator reflects their actual desire to speak. Evaluation results demonstrate that the proposed method has a potential to prevent speech contention and ensure smooth turn-taking in Web conferences.</p>		