

ノードの分散配置による会議空間構成手法に関する研究

高瀬宗祐[†] 大槻麻衣[†] 葛岡英明[†] 山下直美^{††}

遠隔会議において、会議が行われている空間での会話の内容を、遠隔からの参加者が理解することを支援する手法として、筆者らはノードの分散配置による会議空間構成手法を提案する。本手法では、会議室内に分散させた機器から映像と音声を異なるストリームとして配信することにより、遠隔からの参加者は会議の進行に応じて、適切な位置から映像と音声を取得することができる。本稿では、このノードの分散配置による会議空間構成を実現するシステムの開発及び、システムのプロトタイプによるコンセプトの評価について報告する。

Spatial Design for Teleconferences by Distributed Nodes

SOSUKE TAKASE[†] MAI OTSUKI[†] HIDEAKI KUZUOKA[†]
NAOMI YAMASHITA^{††}

To make participants in remote places comprehend what they discuss in a meeting, we propose a method that designing the meeting space with distributed nodes in the room. This method enables participants to select nodes for the image and the sound separately, therefore they can choose suitable media sources for understanding. In this paper, we report the progress of the implementation of the system and the assessment of its concept.

1. はじめに

CSCW (Computer Supported Cooperative Work) では、遠隔地とのコミュニケーションや協調作業を支援するシステムが多数提案されている。

昨今のネットワークの速度の向上と、電子情報機器の高機能化を受け、日常生活や企業の業務に導入することを目的とした、導入が比較的容易な遠隔会議システムの開発が盛んになっている。特にスマートフォンやタブレットコンピュータの普及に伴い、ビデオ会議や Web 会議システムが広がりを見せている[1]。

一般的な遠隔会議システムでは、会議を行う二地点にマイク及びカメラの機能を持つ機器を設置する必要があるが、単にこれらの機器を設置しただけではカメラの撮像範囲やマイクの集音範囲が制限されてしまい、コミュニケーションの障害となることが考えられる。特に、昨今普及しているスマートフォンやタブレットコンピュータを用いた遠隔会議の場合は、マイクやカメラが広範囲の撮像及び集音に対応しておらず、上記の制限が問題となる。

本稿ではこの問題に対し、映像と音声を別々に取得・配信するスマートフォンやタブレットコンピュータなどのモバイル端末をノードとして会議空間内に分散させることで、遠隔地からの会議参加者が任意の位置から映像と音声を取得できる、ノードの分散配置による会議空間構成の提案とその開発を行う。

2. 関連研究

遠隔会議を支援する研究は、主に映像・音声を送受信する装置が固定して設置されているものと、遠隔操作が可能なロボットを用いるものに分類できる。

複数の映像及び音声の送受信を行う端末を会議空間に設置する研究として、Buxton らの Hydra[2]がある。Hydra は会議の参加者一人に対し、一つの端末を割り当てることで、各参加者の視線や音声为谁に向けてのものであるかを認識できることを目指した。また、住谷らは多重ワークにおいて、会議が行われている空間内に複数のカメラを設置し、音声及び視線を監視することでカメラを切り替えるシステムを提案している[3]。この研究では多重ワーク中の会議への参加の支援を目的としており、遠隔地の参加者は発言等を行わず、またカメラの切り替えも自動的に行われるため、ユーザは能動的に制御できない。

遠隔地にいる人物がマイク及びカメラを任意の位置へと操作し、必要な情報を得る手段として、遠隔会議用のロボットを用いる手法がある。MeBot[4]は卓上に設置するロボットで、遠隔地からの会議の参加者は操縦桿を用いて、会議空間にいるロボットの操作を行う。Beam は Suitable Technology 社が販売する遠隔会議用ロボットで、床の上を移動することができ、会議空間内を自由に走行する[a]。これらのロボットを用いるシステムは、比較的自由的な操作が可能である反面、専用の装置を用意する必要がある。

本研究では、会議の参加者が所有するモバイル端末を利用することで導入を容易にし、且つ遠隔からの参加者に自由度の高い操作の提供を実現するシステムの開発を目指す。

[†] 筑波大学 システム情報工学研究科
^{††} NTT コミュニケーション科学基礎研究所

a Beam: <https://www.suitabletech.com/beam/>

3. ノードの分散配置による会議空間構成

モバイル端末を用いた遠隔会議では、使用する機器がカメラとマイクを内蔵していることから、カメラとマイクが同じ位置に存在しており、遠隔から会議に参加した人間は、同じ地点で取得された映像と音声を視聴することになる。ここで、一般的な会議を考えると、参加者の数が不特定であることや、各参加者の位置が分散していること、また、ある程度の広さを持つ空間でコミュニケーションが行われることが想定される。この際、音声の発生位置と参加者の注視対象が互いに遠く離れる場面が存在する。例えば、室内にあるホワイトボードに書かれた内容や、スクリーンにプロジェクタで投影された映像について、同じ室内の離れた位置から説明する場合などがある。こうした場合に、スマートフォンのようなカメラとマイクを内蔵する装置を用いると、カメラの撮像範囲と集音範囲のトレードオフが発生する(図1)。そのため、会議の進行に応じて、会議が行われている部屋にいる人間が、遠隔からの参加者に配慮し、その都度カメラやマイクの位置を移動させなければ遠隔からの参加者は必要な情報を得ることができない。

上記の問題に対し、2章で挙げたものをはじめとする既存の研究や遠隔会議を支援するシステムでは、遠隔から操作できるロボットを用いて映像と音声を取得する装置を適切な位置へと移動させる、広範囲からの音声や映像を取得できるマイクやカメラを接続するといった手法で解決を試みている。しかし、これらの手法は映像と音声のどちらもある程度の品質を期待できる中間地点から取得することとはできない。また、これらの遠隔会議支援のシステムは専用の装置を必要とするものが多いため、導入が難しく、設置できる環境が制限される場合があるという問題が存在する。

本研究ではこれらと異なるアプローチとして、スマートフォン、タブレットコンピュータ、ノートパソコンといった、移動が容易かつ映像及び音声の通信を行うことが可能なモバイル端末を、会議参加者が自由に会議空間内に分散して配置する、ノードの分散配置による会議空間構成手法を提案する(図2)。

この手法ではまず、ノードとなる複数のモバイル端末を、コミュニケーションが行われる空間内で、参加者の付近や資料を表示するスクリーンの正面など、会議進行において重要となる位置の映像又は音声を取得できるように配置する。映像及び音声を取得するノードを重要な位置に分散させる事により、専用の装置を必要とすることなく複数の位置からの映像や音声を取得でき、また、個々のノードはその対象に最適な位置から映像及び音声を取得することができる。

次に、各ノードは取得した映像と音声を、映像のみ、音

声のみの異なる二種類のストリームとして配信する。映像と音声を分離して配信することにより、遠隔からの参加者は映像と音声を異なるノードから取得することが可能となる。これにより遠隔からの参加者は、会議のある時点において映像及び音声をそれぞれ自由に選択することができる(図3)。

提案するノードの分散配置による会議空間構成では、上に挙げた2つの特徴を持つことで、遠隔からの会議参加者は適切な位置からの映像と音声を得ることができるため、会議内容の理解及び参加を支援する効果が期待できる。

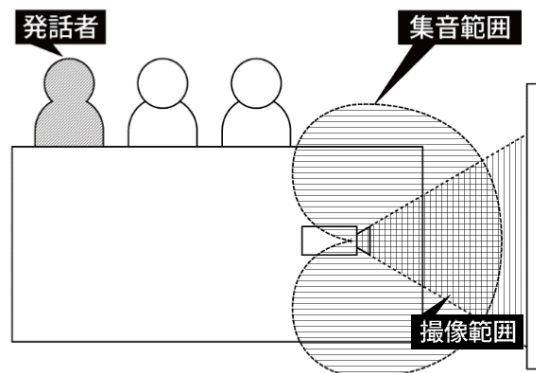


図1 撮像範囲と集音範囲のトレードオフの例
(撮像範囲をホワイトボード、スクリーン等に合わせると発話者が集音範囲外になってしまう)

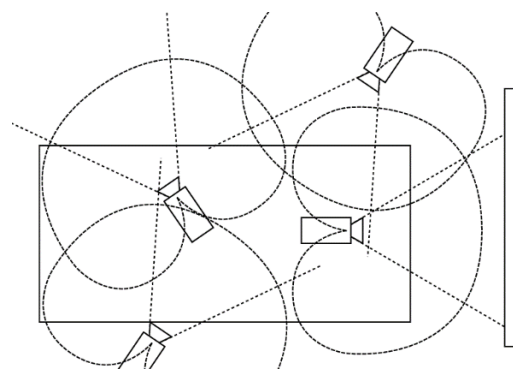


図2 ノードの分散配置による会議空間構成の例

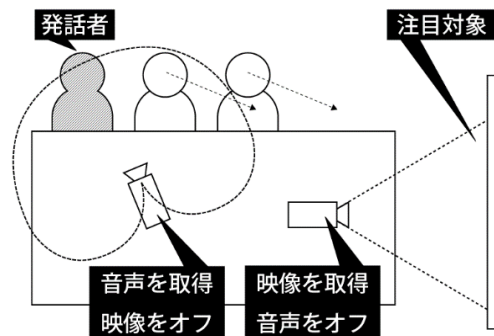


図3 映像と音声の分離

4. システム開発

4.1 システム構成

ノードの分散配置による会議空間構成を実現するための最低限の機能を持ったプロトタイプを実装した。システムはクライアントと、Web サーバ、DB サーバ、ICE サーバ、signaling サーバから構成される。

4.1.1 クライアント

会議空間内にノードとして分散させるものと、遠隔から会議に参加するために用いるカメラ及びマイクと通信機能を持つモバイル端末である。

4.1.2 Web サーバ

開発するシステムは Web ブラウザ上で動作する為、Web サーバを用意し、各端末上のブラウザからアクセスする。サーバアプリケーションとして今回は nginx を用いる。

4.1.3 DB サーバ

会議空間内に設置されているノード端末及び、遠隔からの会議参加者が使用している端末の状態や、各端末のカメラのキャプチャ画像及び取得した音量は、10 秒毎に送信され、DB サーバ上で更新される。システムのプロトタイプではデータベースに、ブラウザアプリケーションと親和性の高い MongoDB を使用する。

4.1.4 ICE サーバ, signaling サーバ

映像及び音声の遅延を防ぐため、これらのストリームは途中でサーバを介さずに P2P で通信できることが望ましい。ICE サーバは NAT を挟む P2P 通信を行うための STUN/TURN サーバ機能をもつサーバである。また、signaling サーバは端末間で映像及び音声通信の開始や終了などのメッセージを仲介する機能を持つ。

本システムでは、上記 2 種類の機能を提供する PeerJS[b] を利用する。

4.2 通信方式

システムの映像及び音声の通信手段として、WebRTC (Web Real-Time Communication) を用いる。

開発するシステムでは任意の端末をノードとして利用するが、ノード毎に Android や iOS 及び Windows といった複数の異なる動作環境に対応する必要がある。そのため、ノード上で動作するプログラムを、プラットフォーム依存の少ないブラウザ上で実行できる Web アプリケーションとして開発する。WebRTC はプラグイン無しで、Web ブラウザ間のビデオチャットやボイスチャット機能を提供する HTML5 の API である。

なお、開発するシステムは、使用する機器や環境に依存しない事を目指す。現時点で WebRTC が実験段階にあることから Web ブラウザごとに実装の対応が異なるため、今回は各端末上で使用する Web ブラウザを Google Chrome に統一する。

4.3 システムの動作

会議空間内に配置されたノードとなる端末は、一定時間毎にカメラからのサムネイル画像と、マイクからの音量の平均値を取得し DB サーバに送信する。

遠隔からの会議の参加者は、DB サーバからこれらの情報を参照し、どの端末の映像または音声を取得するかを指定する。

映像または音声を指定された会議空間内の端末は、リクエストと同時に指定されたデータの配信を開始する。このとき、遠隔からの参加者が要求したものと同一種類のデータを会議室側のノードも受信し、遠隔からの参加者が他のノードに対してリクエストを行うまで通信を継続する。

5. システムプロトタイプ

図 4 に開発したプロトタイプの画面を示す。クライアント上で動作する Web アプリケーションは会議空間内に配置された複数のノードも、遠隔からの参加者が使用する端末上で実行されるものも、共通の GUI を用いている。

画面左側の領域には、通信を行う相手のカメラからの映像が中央に表示され、右下には自身のカメラからの映像が表示される。

ユーザはまず、モバイル端末から、Web サーバにアクセスし、画面右上部のコンボボックスから、その端末が所属するドメインが Local (会議室) であるか Remote (遠隔側) であるかを選択する。

画面右側中央の領域には、Local 端末の場合は Remote 端末のカメラのサムネイルが表示され、Remote 端末の場合はすべての Local 端末のカメラのサムネイルが表示される (図 5)。表示されるサムネイルの周囲四辺には枠線が描画されており、その端末のマイクに入力された音量に応じて色と枠線の太さを変更する (図 6)。これにより、どの端末の付近に音声の発生源があるかを認識できる。

Remote 端末では、システムの画面に表示された Local 側のサムネイル画像をクリックすると、ダイアログが表示され、その端末から配信されるストリームのうち、映像か音声の一方、または両方のストリームを同時に取得するかを選択する (図 7)。

一方、Local 端末は、上で記したような Remote 端末やストリームの種類を選択する機能を提供しない。Local ノードは Remote 端末から映像又は音声を要求されると自動的に通信を開始する。

6. 評価実験

6.1 評価実験の実施

実際に、ノードの分散配置による会議空間構成を用いた会議を行い、提案手法を評価した。

b PeerJS <http://peerjs.com/>

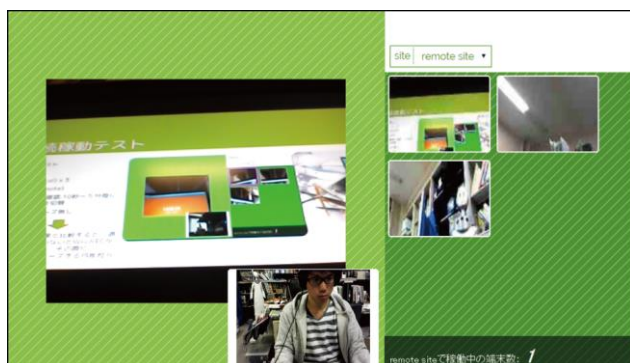


図 4 プロトタイプ画面

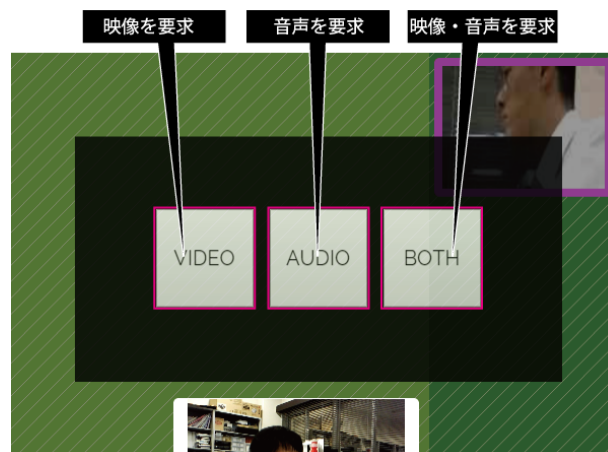


図 7 サムネイルクリック時のダイアログ



図 5 部屋の選択とサムネイル画像一覧 (Remote 端末)



図 6 音量の表示

ノードの分散配置による会議空間構成では、各ノードの情報の送受信や、配信するストリームの数が増えるため、通常の1対1のモバイル端末間の遠隔会議に比べ遅延が生じる可能性がある。複数のノードを利用することによって発生する遅延が遠隔からの参加者に対し影響を与えるかどうかを、通常の1対1の遠隔会議と比較し、アンケートによって評価した。また、会議への参加に関して、これもア

ンケートによる主観的な評価を行った。

6.2 実験内容

6.2.1 会議室

会議室には一定の距離を置いて机と椅子を2組配置し、実験者1名と被験者1名がそれぞれに座る。また、両者から見える位置に大型のプラズマディスプレイを設置し、実験者は机の上にあるコンピュータからプラズマディスプレイに表示される映像を指定できるようにする。実験者、被験者の前にはそれぞれの上半身がカメラに映るようにタブレット端末 Thinkpad 8 (OS: Windows 8.1 Pro, CPU: Intel® Atom™ Z3795 1.6GHz., RAM: 4GB) を2台配置し、プラズマディスプレイの前にも、画面が映るようにタブレット端末 (Nexus 7 2013, Android 4.4, CPU: Qualcomm Snapdragon™ S4 Pro 1.5GHz, RAM: 2GB) を1台、合計3台の端末をノードとして室内に配置する。

6.2.2 遠隔側の部屋

会議室とは別に、会議室内の光景や音声が直接取得できない場所に、遠隔からの会議参加者用の部屋を用意する。遠隔側の部屋には被験者1名が座る机と椅子が1組用意されている。また、机の上にはインターネットに接続したノートパソコン (hp Pavilion dm1, Windows 7, CPU: AMD E-350 Processor 1.6GHz, RAM: 4GB) が用意されており、被験者はノートパソコンから開発したシステムを用いて会議室で行われる会議に参加する。

6.2.3 会議内容

会議では国内旅行について、ランダムに47都道府県の中からランダムに1つの地域を選択し、その土地の観光について自由会話をを行う。被験者は、会議参加者である実験者に対し、会話に登場した事柄に対しインターネット検索を行い、画像などをプラズマディスプレイに表示させるよう指示する事が可能である。また、実験者は被験者からの指示が無い場合に、会話に登場した語句を検索し、関連する画像をプラズマディスプレイに表示する。

会議は以下の3つの条件別を実施する。

表 1 アンケート設問

質問項目	質問内容
質問 1	音声の遅延は感じられなかったか
質問 2	映像の遅延は感じられなかったか
質問 3	会話の音声は聞き取れたか
質問 4	会議室側で話者が何を見ているか分かったか
質問 5	会話の内容は理解できたか
質問 6	自分から発言できたか
質問 7	会議室で行われている会話に参加できていると感じたか

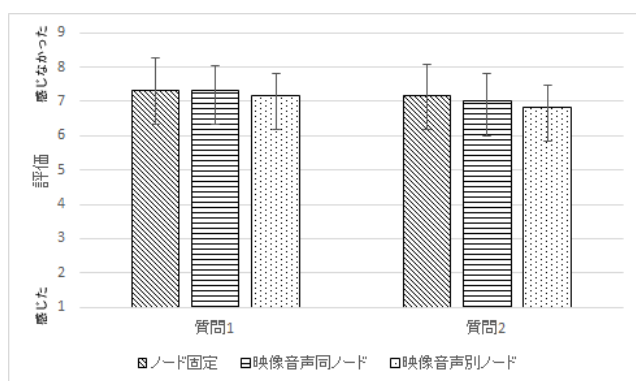


図 8 質問 1, 2 のアンケート集計結果

表 2 質問 1, 2 の一元配置分散分析結果

	p 値	有意差
質問 1	0.9845	有意差なし
質問 2	0.9576	有意差なし

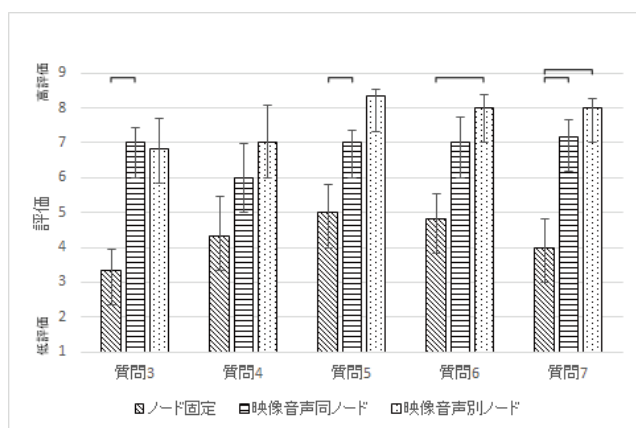


図 9 質問 3~7 のアンケート集計結果 (有意水準 $\alpha = 0.05$)

1. プラズマディスプレイ前のノードのみ選択可能
2. 全てのノードが選択可能. 音声と映像は分離不可
3. 全てのノードが選択可能で, 音声と映像を個別に選択可能

以降, 1 の条件を“ノード固定”, 2 の条件を“映像音声同ノード”, 3 の条件を“映像音声別ノード”とする.

3 種類の条件を 5 分ずつ, 被験者内配置で実施する. 被験者がシステムの操作や会議の内容に慣れる可能性があるため, それぞれの条件は被験者ごとに異なる順番で行う. 尚, 評価実験は大学生・大学院生の 6 名に対して行った.

6.2.4 アンケートの実施

実験では各条件で 3 回の会議を行った後, 遠隔から会議に参加した被験者はアンケートに回答する. アンケートには 9 段階のリッカート尺度を使った設問 (表 1) と自由記述を用意した.

6.3 実験結果と考察

質問 1 “音声の遅延を感じたか”, 質問 2 “映像の遅延を感じたか” のアンケート結果の算術平均と一元配置分散分析の結果を示す (図 8, 表 2).

質問 1, 質問 2 においてノード固定, 映像音声同ノード, 映像音声別ノード間における通信速度の遅延による有意差は見られなかった.

次に質問 3 “会話の音声は聞き取れたか”, 質問 4 “会議室側で話者が何を見ているか分かったか”, 質問 5 “会話の内容は理解できたか”, 質問 6 “自分から発言できたか”, 質問 7 “会議室で行われている会話に参加できていると感じたか” のアンケート結果の算術平均と, Bonferroni 法による各条件間の多重比較の結果を示す (図 9).

この結果に対し, 有意水準を 0.05 とする. まず, 質問 7 から, 映像音声同ノード, 映像音声別ノードともに, ノード固定に比べて会議室で行われている会話に参加できていると感じられることが分かった. 質問 3 と質問 5 より, 映像音声同ノードはノード固定よりも会話の内容が聞き取れる, 会話の内容が理解できると言える. また, 質問 6 より, 映像音声別ノードについてはノード固定よりも, 積極的に発言できることが分かった. 尚, 各質問において映像音声同ノードと映像音声別ノードの間に有意差は見られなかった. この原因については, 今後より詳細な分析を行う必要がある.

質問 4 “会議室側で話者が何を見ているか分かったか” では有意差は見られなかった. これに関し, 自由記述において「各ノードの位置関係が把握できなかった」「ディスプレイを見ていると, 他の人が何を見ているのかが分からなかった」などの意見があり, 会議空間全体の状態を把握できるような機能が必要であると考えられる.

7. 提案手法の定量的評価の検討

7.1 定量的な評価の必要性

6 章ではノードの分散配置による会議空間構成のためのシステムのプロトタイプを用いて, 評価実験を行った. その結果, 主観的評価において提案する手法は, 遠隔からの会議の参加を支援するという結果を得ることができた.

しかし, 参加者が会議の内容を理解できるかどうかは,

会議で扱われる話題に対する事前知識の有無や、コミュニケーションの中で自分自身が主導権を握るなど優位な位置に立つ事ができるかどうかなど、その時の条件や参加者の知識及び性格に影響される。そのため、より詳細な評価を行うには、システムによる遠隔からの会議参加者の支援の効果を定量的に評価する必要がある。

そこで著者らは現在、ノードの分散配置による会議空間構成を用いたシステムが、会議に置いて遠隔からの参加者の支援を定量的に評価する為の実験を検討している。以下にその概要を述べる。

7.2 実験の方針

実施する実験ではシステムが次の2点を支援するかについて評価する。

- 遠隔側の参加者の会議内容理解
- 遠隔側の参加者の積極的な発言

ノードの分散配置による会議空間構成手法を用いたシステムは、参加者の注視対象が会話の進行に応じて移動し、また発話者が交代するなど音声の発生源が移動するような場面において、特に有効であると考えられる。そのため、実施する実験においても、注視対象が移り変わり、また話者が定期的に切り替わるシナリオを用意する。

7.3 実験案

システム評価と同様、会議室で実験者と被験者の2名が会議を行い、遠隔側の部屋からもう1人の被験者が開発するシステムを用いて参加する。会議室内の配置については、6章で行った評価のものとは異なり、会話の進行に応じてメモや図を記すためのホワイトボードや黒板などを実験者の近くに設置する。これにより、会議参加者の注視の対象が増え、映像の切替えが行われるようにする。

遠隔会議における参加者の理解度や、会議中の発言などを分析する手法として、地図会話コーパス[5]やブックマーク X[6]では、地図を使い、提示された問題を解決するタスクを用いている。本研究で開発したシステムも、遠隔からの参加者が会議室側の参加者と視覚的に情報を共有し、かつ話者交代が発生するものとして、これらと同様、地図を用いた協調作業を実験タスクとして使用する。

会議は地図上の店を巡回するツアープランの計画について行う。被験者2名には最初に、ツアーを行う街の地図と、その街にある有名店の商品の写真と情報が与えられる。被験者の持つ店の情報は一部重複するが、相手が知らないものも存在しており、提示された時間と金額を考慮しながら効率的に店を巡回する計画を立てることを目指す。

定量的な評価を行うため、システム評価で用いたのと同じ条件を用意し、ノードの分散配置による会議空間構成を用いた場合と用いなかった場合、また映像と音声を異なるノードから選択できる場合とできない場合ではどちらの条件の際に、作成した計画が効率的であるか、遠隔からの会議参加者の発言数が多かったかを計測する。また、会議後

に会議内容に関する設問を用意し、会議の内容を理解しているかを評価する予定である。

8. まとめ

本研究では、ノードの分散配置による会議空間の構成手法を提案した。提案手法では、会議が行われる空間内に映像及び音声の送受信が可能なスマートフォンやタブレットコンピュータを分散配置し、遠隔からの会議参加者が任意の位置から映像と音声を取得することで、会議参加を支援する。また、Webブラウザ上で動作するプロトタイプを実装し、システムの動作評価及び、提案手法の有効性についてアンケートによる評価を実施した。

今後は、7章で述べたようにノードの分散配置による会議空間の構成の有効性を定量的に評価するための検証実験を実施し、その結果をもとにシステムの改善を行う予定である。

参考文献

- 1 (株)シードプランニング, テレビ会議/Web会議の最新市場とHD化動向—映像コミュニケーションの製品動向及び市場の将来性(2008)
- 2) A. Sellen, W. Buxton, and J. Arnott: Using spatial cues to improve videoconferencing, Proceedings of CHI '92, pp. 651-652 (1992).
- 3) 住谷哲夫, 高田格, 重野寛, 岡田謙一, 津村弘輔: 多重ワークにおける会議中継カメラの自動切替え手法, 情報処理学会論文誌 Vol.48, No.1, pp356-364(2007).
- 4) S.Adalgeirsson, and C.Breazeal.. MeBot: a robotic platform for socially embodied presence. Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction, 15-22(2010).
- 5) 青野元子, 市川薫, 他, 地図課題コーパス(中間報告), 情報処理学会研究報告音声言語情報処理 (SLP) Vol. 94, No. 90, pp.25-30 (1994).
- 6) 磯友輝子, 松田昌史, 大坊郁夫, 対面相互作用場面における課題達成と社会的スキルの関連, 電子情報通信学会技術研究報告, HCS, ヒューマンコミュニケーション基礎, Vol.107, No. 59, pp.7-12 (2007).