

## 多地点ビデオ会議コーパスに基づく会議途中参加支援機能の研究

中山 彰<sup>†</sup> 細田 真道<sup>†</sup> 犬童 拓也<sup>†</sup> 小林 稔<sup>†</sup> 岩城 敏<sup>†</sup>

あらまし 多地点ビデオ会議システムにおいて、途中参加支援機能（「おっかけ会議」）の実現を狙いに、会議中に発生するマルチモダリティな時系列データから会議の構造、要所に関する情報をマイニングし、それを表現するための研究を進めている。本報告では、その研究の構想および研究基盤としての収録・再生システムの概要、ビデオ会議コーパス、収録状況、解析結果に述べる。

### Video Conference Grasper Based on a Video Conference Corpus and its Multimedia Data Mining

Akira NAKAYAMA Masamichi HOSODA Takuya INDO Minoru KOBAYASHI  
Satoshi IWAKI

**Abstract** For the purpose of realization of a video conference participating support function, we are studying to extract and express the information about the structure of meetings and their key points through the data mining of multi-modality data generated during meetings. This report describes the outline of the research approach, a meeting recording / reproduction system, specifications video meeting corpus, the current state of the corpus recording, and its preliminary analysis result.

## 1 はじめに

ブロードバンドインターネット接続環境の普及、高性能な計算機の普及、そして、社会的な合意（裁量労働制、在宅勤務）が進んだことにより、個人でもマルチメディア会議システムを活用できる環境は整ってきた。またインターネットビデオ会議システムが必須の就労形態も出現してきている [1]。インターネットビデオ会議システムはユーザの移動時間や会議スペースなどの時間的・空間的コストの削減を大きな利点であるが、一般家庭ではまだ爆発的な普及にいたっているとは言いがたい。その理由を追求するために、われわれは、複数の家庭と会社を商用インターネットビデオ会議システムで結んだ在宅勤務実験を約一年間実施してきた [2]。その実験においては、会議での途中参加支援機能（例えば、いままでの議論過程をみながら会議に参加するなど）や会議議事録作成機能、会議概要（雰囲気・議事）把握機能などの要望が数多くあげられた。併せて行った会議のビデオ録画の解析の結果、専用の部屋で行われるオフィスでの実会議とは大きく異なり、在席での会議においては、実環境からのインタラプト（話しかけ、電話、緊急の用事）や、PC 関連のトラブルや接続環境のトラブル等が原因で、すべてのユーザが会議の最初から最後まで参加することはむしろ稀であることがわかってきた。また実験当日に行わ

れた会議に参加できなかったユーザのために、会議アーカイブ（画面のビデオキャプチャ）も公開していたが、利用者はわずかであり、より利便性の高い会議アーカイブ閲覧ツールの必要性も指摘された。

われわれはこれらの考察をすすめた結果、自分の時間が許す範囲で参加したり、必要話題における議論だけ参加する、会議の要所・要点だけ閲覧する、雰囲気の変移だけ味わう、また会議の要所・要点などを閲覧しながら会議に参加するなど幅広い意味でのコミュニケーション場への途中参加を支援する機能の重要性を認識するに至った。そこで本研究ではこのように途中参加を支援することで、「時間の束縛」から人間を解放する未来型ビデオ会議システムの実現を目指す。そしてその象徴的な名称として会議途中参加者が本会議の流れに追いつくという意味を込めて「おっかけ会議」と呼んでいる。以下ではこの「おっかけ会議」に関連する技術分野を (1) グループウェアと CSCW (Computer Supported Cooperative Work) (2) 「メディア処理」の観点から鳥瞰し、克服すべき技術課題を整理する。

従来の途中参加支援のためのグループウェアと CSCW という視点での研究としては、ソフトウェア開発における会議を対象に、マルチメディア議事録の会議録の研究 [3] がある。しかしながらこの方式では、会議での発言を構造づけるため、また発言の書き起こしのために、議事録専任者が必要である。また同期・非同期統合型会議システム ASSIST [4] では、会議参加者に発言の意図（賛成・反対・提案など）について入力させたり、またあらかじめ議事を整理して登録しておかなければならないという手間

<sup>†</sup> 日本電信電話株式会社 NTT サイバースペース研究所  
〒 239-0847 神奈川県横浜須賀市光の丘 1-1  
nakayama.akira@lab.ntt.co.jp  
NTT Cyber Space Laboratories, NTT Corporation  
1-1 Hikarinooka Yokosuka-Shi Kanagawa JAPAN

が発生する。パワーポイントのスライドの切れ目により会議の構造化を試みた研究 [5] もあるが、報告型の会議を対象にしたものであり、スライドの区切れより小さい会議の構造を捉えることができないという限界がある。さらに学会講演発表を対象に会議の構造付けを行った例 [6] や、ボイスメールを対象に音声認識を試み、各メールにキーワードレベルの要約 [7] をつけた例もある。この二つの例は、会議などの碎けた発話に対しての現状の音声認識技術の限界をあらわしたものと見える。限界をまた意図的に入力される「拍手」や「会議発言者の指名」によって、導かれた会議のダイジェスト音声を提示することにより会議への途中参加支援を実現した例 [8] もあるが、音声・ホワイトボードのみの会議を対象にしていることなど適用範囲が限定される。

過去の研究を総括すると、会議などの途中参加支援、閲覧のために、会議参加者・議長になんらかの構造化のための作業を強いること [3, 4]、自然発話の音声認識が困難であるため自然言語処理をベースとした会議要約・分析技術が未成熟 [6, 7]、音声以外のモダリティを使用した例としては、スライドの切れ目、拍手の入力 [5, 8] などの例があるが、ほかのビデオ会議で活用可能なモダリティの存在を考慮に入れると検討が不十分である。それゆえ「おっかけ会議」実現に向けた問題は解決されていない状況と言える。また会議中の雰囲気再現が不十分であったり、またそれらをもとに途中参加・検索・閲覧できる仕組みを備えていなかった。

一方で、メディア処理（計算機による対話・コミュニケーション場の処理）という視点では、顔表情、身振りやうなずき、姿勢、パラ言語情報（声の抑揚・大きさ・沈黙・いい間違い・発話量）などの非言語的情報 [9, 10] がコミュニケーションにおける重要な要素であるとして、大規模な音声対話コーパスとそれに含まれる非言語的情報を用いてコミュニケーションの構造推定・雰囲気推定を信号処理で目指した研究例もある。例えば、松坂は二人話者での合意形成型対話を対象に、話者の頭部に現れる「うなずき」「かしげ」「首振り」などのジェスチャをオプティカルフローなどの特徴量を元に認識し談話（主張、合意、否定、確認）の認識を試みている [11]、日高らは音声からパワー、高さ、発話速度を抽出し音声の強調部分をもとめビデオメールの要約をしている [12]。また山下らはホテル予約会話において、話題の区切れの際に発話のピッチが高くなる、パワーが大きくなることを報告している [13]。また小椋らは、電話での二人話者での発話を対象に話題境界の特定に貢献する言語特徴を検討し、話題境界後の発話単位の単語数をもっともそれに貢献していることを見出した [14]。綿貫らは自由対話を分析し、気分の高まりとともに、二人の話者の発話量が等しくなること、

頭部の動き、ピッチの変化から話者の感情の度合いが同期していることを見出した [15]。以上のこれらの研究は単一のモダリティそして一人もしくは二人の情報発生源（話者）の信号から会話の構造・雰囲気推定を試みていると考えることができ、一定の成果を収めている。しかしながら 3 人以上の発話者が存在する会議のような場での話者間の相互作用やまた各モダリティ間の相互作用、また司会者の存在が想定される場で検討した例は少ない、またビデオ会議特有のモダリティを活用した例も皆無である。その理由のひとつとしては、いままでのコーパスが、独話や、1 対 1 の対話収録に重きをおいており、多人数対話（会議）コーパスも少なく、テレビ会議コーパスの存在は皆無である [16]。

上記の研究例と問題点を整理すると、われわれの目指す「おっかけ会議システム」実現に対して不足している技術は下記に集約できる。

1. 会議参加者に手間を強いることなく、マルチモダリティな信号処理による多人数マルチメディア会議の構造化・キーポイント・雰囲気抽出する技術
2. (1) に基づき、会議内容を可視化・ブラウズ・検索できる技術

一方、本研究を客観的・体系的・効率的に進める上での環境・基盤という視点に立つと、上記 (1) の課題解決の前提となるデータ、すなわちマルチメディアビデオ会議コーパスがこれまでまったく存在していないことが課題として浮かび上がってくる。

そこで本研究ではこのコーパス構築を最優先課題と設定し、これまでにコーパス構築のための会議収録システムを開発し、それを用いた会議収録実験を進めてきている。

本稿ではまず本題に入る前に、本研究全体の枠組み、アプローチ、適用対象とする会議のモデルを明示することで、我々の研究領域と方向性を整理する。次に、コーパス構築用に開発した多人数マルチメディア会議収録システムの特徴、コーパス設計指針、会議収録実験の進捗状況について述べる。最後にこれまでに取得したコーパスを用いて統計的解析を実施した結果、話題の区切れに関する知見が得られたので併せて紹介する。

## 2 研究全体の枠組み

### 2.1 おっかけ会議システムの目指すもの

「おっかけ会議システム」の最終的な目標は会議の内容を要約した文章を提示するだけでなく、検索・ブラウジング（ばらばら見ること）などにより

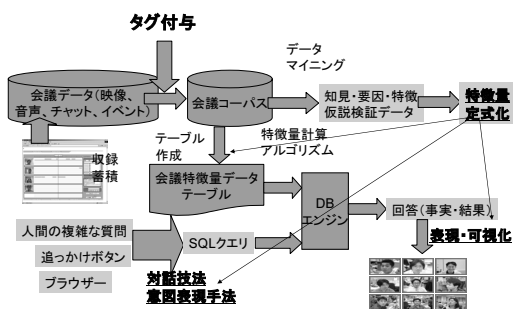


図 1: 研究全体の枠組み

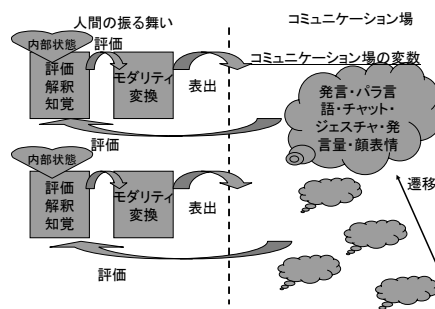


図 2: 会議と会議参加者のモデル

会議途中参加者・会議閲覧者に対して、記録された会議情報に「効率的」「効果的」にアクセスすることで、あたかも最初から議論に参加していたのと同等の以上の効果をユーザに与え得るメディアを構築することである。

この目標のためには、(1) マルチモダリティの情報を活用し、会議の構造・雰囲気把握すること (2) 利用できるメディアを活用し会議の雰囲気・構造を伝えること (現在なされている会議を阻害することなく同時に提示できること) (3) 参加者の要望・興味に応じた場所を伝えること (4) 興味有る場所の指定方法を創出すること、といった、複数の重要な課題があると筆者らは考えており、下記のアプローチにより今後解決を図っていく。

## 2.2 研究全体の枠組みとアプローチ

研究全体の枠組みを、図 1 に示す。研究全体としては、まず会議を行いマルチモダリティな情報を収録・蓄積する必要がある、そして、会議に対して人手でタグ (感情・態度・雰囲気・発言の価値評価・構造など) を付与し会議コーパスを作成する。そのタグのデータ、会議の観察・比較評価を経て会議特徴量のアルゴリズムを考案・実装しタグデータとの比較を行うことで、前節の課題 1 の実現を目指す。またそれらの会議特徴量を利用して課題 2 の実現を図る。またそれらの物理量を検索可能な形で登録し、そして、興味あるポイントの指定方法を翻訳するモジュールを考案・実装することで、前節の 3、4 の課題にアプローチしていく。前節 3、4 の課題についての議論は別稿に譲る [17]。

## 2.3 本研究で取り扱うビデオ会議のモデル

本研究で仮定する会議および会議参加者の振る舞いモデルを図 2 に示す。このモデルでは、おのおの独立した会議の参加者が、会議での参加者発する発言、非言語情報 (パラ言語、ジェスチャなど) を知

覚・解釈・評価しそれが心の内部状態を変更する。またその評価結果は各会議の参加者が持つであろうモダリティ変換機構を通じて発言や非言語情報として現れると考えている。表出・評価の繰り返し議論の発散・収束を生みだし、時間の経過とともに全体として意思統一が図られていくような会議を対象とする。

そしてそれらの発言や非言語情報の全体集合、そして時間的な変化が会議の要所・構造・雰囲気の手がかり情報になるとの仮説を設定し、われわれはその仮説の解明のために、発言・非言語情報を生み出す会議参加者の「心の内部状態」に注目する。そのためコーパスではアンケート、議事録などを通じて「心の内部状態」も観測し、タグとして付与していく。

## 3 コーパス取得システム概要

### 3.1 収録システム概要

[要件] 本研究を進める上では、プラットフォームとしての、会議収録システムを作成する必要がある。下記の 3 点を主眼に目的に設計・インプリメンテーションを行った。

1. ビデオ会議に初めて参加する人でも自然に会話に参加できる品質・空間を確保する
2. 画像解析・音声解析に耐えうる十分な品質の動画像・音声を収録する
3. マルチメディアデータを時刻同期した形で収録し、センサの拡張に柔軟に対応可能とする

[ハードウェア構成] 図 3 に概要を示す。テキスト・イベント系データ収録サーバ、多地点 A/V 通信・配信・収録サーバ、および USB カメラ (Logicool QCAM Pro4000)、ヘッドセットマイクロホン (SONY DR-140)、2 台のディスプレイ (顔画像表示用、共有書類表示用) を備えたクライアント 4 台から構成されている。サーバクライアントはギガビットプライベートネットワークで接続されている。クライアント 4

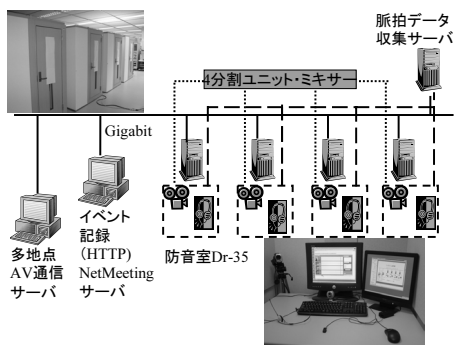


図 3: 会議収録・閲覧システムブロックダイアグラム

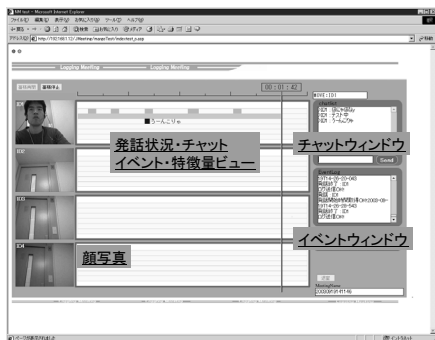


図 4: 収録ブラウザ画面

台の入出力デバイスは遮音等級 Dr-35 の 1.5 畳の防音室に格納されており、4 人の話者間の聴覚的・視覚的な直接の結合が生じないように、かつ外来のノイズを防ぐようになっている。そのほか、ユーザの脈拍を収集するための PC、各防音室内に設置されたデジタルビデオカメラからの映像・音声を集約・記録するための、4 分割ユニット、4 チャンネルミキサー、デジタルビデオデッキが設置されている。

[ソフトウェア構成]サーバソフトウェアは、カスタマイズ性に優れ、品質に定評のある、MacroMedia 社の Flash Communication Server MX 1.5 を使用した。その機能に覆いかぶさる形で FlashMX を用いてクライアントプログラムを新たに作成し、4 拠点での動画音声配信、記録・チャット通信を実現した。また NetMeeting3.0 の Window 共有機能と PowerPoint を組み合わせ、書類共有機能を実現した。また HTTP サーバに対するポストという形でイベント記録（チャット入力時刻、発話開始・終了タイミング、共有資料のページ切り替えタイミング、資料提示タイミングなど）もミリ秒の精度で実現した。ポストメソッドを採用したため新たなイベント記録などに柔軟に対応可能な設計となっている。

[性能]表 1 に本システムの性能概要を示す。音声の遅延量 220 ミリ秒は、音声遅延が通話の妨害感に与える研究 [18] によれば、十分許容できる範囲だ

表 1: 研究プラットフォーム性能概要

| 項目        | 性能                                  |
|-----------|-------------------------------------|
| フレームレート   | 15 フレーム/秒                           |
| フレームサイズ   | 160*120                             |
| サンプリングレート | 22050Hz                             |
| 音声遅延量     | 約 220 ミリ秒程度                         |
| 同時参加数     | 4 人                                 |
| 記録可能イベント  | 発話開始・終了時刻<br>チャット送信時刻<br>共有資料への操作時刻 |
| イベント記録精度  | ミリ秒単位                               |

といえる。音声と映像のずれはほぼ無視できる程度であった。

### 3.2 タグ付与・再生システム

収録した会議を再生するためのビューワーも新たに開発した。音声・映像ももちろんのこと、発話状況、会議中におきたイベントについても時刻同期した形で再生されるように構成されている。また会議の任意の時刻にラベルを付与するアノテーション機能、また会議中の任意の特徴量をグラフとして同期して表示するトラックも設け、各自の音声・動画像、イベント、会議特徴量、タグなどを一覧しつつ、会議のダイナミクスを検討できるようになっている。

## 4 会議コーパスの構築

### 4.1 ねらい

マルチメディアデータマイニングのための客観的で豊富な素材収集を目的にコーパス取得をおこなった。これは、途中参加者・閲覧者に対して、有益な情報（話題の切れ目、重要箇所、雰囲気、各参加者の感情・態度）をマルチメディアデータ会議データの信号処理により推定するための「手がかり」およびその信号処理アルゴリズム、ルールなどの妥当性の検証のためである。その目的のためビデオ会議の様子を収録するのはもちろんのこと、重要箇所、雰囲気、各参加者の感情・態度を収集するための被験者向けの時系列的なアンケート、そしてマクロな状況をどのように捕捉しているのか、記憶しているのかを調査するための議事録、そしてどのような発言をメモしているのか調査するためのメモ用紙などの配布も併せて実施・収集している。以下に収録方法やデータファイルの仕様、設計仕様について述べる。

### 4.2 形態と議題の設計

前節のねらいを達成するためにまず最初に収録する課題・形態は、「会議が活性化し」、「要所が存在し」、「全員の意思統一が可能な」、合議タイプの企

画・意思決定会議とした。「会議の活性化のため」また「社会的立場」の影響を少なくするため、参加する被験者はよく知り合った親しい友人・知人同士で、社会的な上下関係のないものとした。

また議題についても上記の観点から、無人島に遭難した時をもっていく道具をリストから8つ選択し優先度をつける課題、また、飲料メーカーの企画会議を想定し、新商品の詳しい特徴・コンセプトが書かれたシートを参照しながらその名前をつける名前付け話題とした。

### 4.3 アンケートの設計

前章での会議および会議参加者の振り舞いモデルの仮定のもと、できるだけ細かい時系列でかつ、判断しやすい単位とするため各発言ごとにアンケートの記入をするようにした。アンケートの項目は大きく4つに分け、それぞれ他人の発言に対する態度、価値判断、会議中になされた「主張」への態度、会議中の進行状況（構造の把握状況）について調査することとし設計を行った。以下にアンケートの項目を詳細に示す。

**態度** 各発言について抱いた感情について回答させるために、心の2次元感情モデル [19, 20] を参考に二つの項目（活性度合い、快・不快度合い）を設けた。各項目は5段階で記入を行うようにした。  
 [活性度合い] 興奮している、やや興奮している、平静、やや沈んでいる、沈んでいる  
 [快・不快度] 心地良い、やや心地良い、どちらでもない、少し不快、不快

**価値判断** 各発言についての価値判断について回答させるために、二つの項目（重要度合い（その発言についてどの程度の重要性（会議への寄与、アイディアの重要性、発言の重みなどを感じたか？））、展開度合い（発言がどの程度会議の流れを変更したのか（話題の展開、新しい視点・見方の導入）））を設けた。各項目は3段階で記入を行うようにした。自分自身の発言についても、自己評価を行い、記入させるようにした。

[重要度合い] かなり重要、少し重要、なんでもない  
 [展開度合い] 流れを変更した、やや変更した、追従

**主張への態度** 事実の交換以上に、本人の意思・意図・推論などが含まれた他人の考えを変更する目的でなされた発言を「主張」と定義し、主張の強さおよび、賛成度合いを記入をさせるようにした。

[主張度合い] 強い主張、弱い主張、独り言  
 [賛成度合い] 賛成、やや賛成、どちらでもない、やや否定的、否定的

**構造の把握状況** 被験者の話題の構造の把握状況調査するために、話題の切れ目状態について、話題の

表 2: 動作書き起こしタグの一覧

| 項目   | タグ       | 項目   | タグ         |
|------|----------|------|------------|
| 言語   | うなずき     | 上体移動 | 移動（前後）     |
|      | 笑う       |      | 移動（左右）     |
| 頭部移動 | 微笑む      | 特殊動作 | 特殊な動作（サイン） |
|      | 移動（前後）   |      | 特殊な動作（指示）  |
|      | 移動（左右）   |      | 特殊な動作（例示）  |
| 首の振り | 移動（上下）   | その他  | 顔の前で手を合わせる |
|      | 首を振る（右）  |      | 手振り        |
|      | 首を振る（左）  |      | ヘッドセット位置直し |
| 視線移動 | 首を振る（左右） |      | ほおづえ       |
|      | 移動（上下）   |      | 座りなおし      |
|      | 移動（左右）   |      | 後ろに体をそる    |
|      | 方向（見回す）  |      | 前のめりになる    |
|      |          |      | 顔を手で触る     |
|      |          |      | 頭を触る       |

切れ目度合いを記入を行わせるようにした。また、各話題にふさわしい題名をつけて記入させるようにした。

[切れ目度合い] 大きな切れ目、小さな切れ目

### 4.4 議事録フォーマットの設計

会議のマクロな特徴・印象・流れを会議参加者がどのように把握しているか、そして記憶しているのかを調査するために会議議事録を被験者に作成させた。会議の流れ、会議の要点、議事要点、重要発言、盛り上がりを感じた場所などを、「だれの」「いつごろの」「どういった」「何に関する」などの項目を詳細に列挙しできるだけ細かく記述させるようにした。また議事録を書くにあたってメモしておけばよかったと思ったことなども同様のフォーマットで記述させた。

### 4.5 音声書き起こしと動作タグの設計

RWCP 会議コーパス [16] の仕様にならない音声の書き起こしを実施することとした。またこれに加えて TEI [21] の基準にならない「笑い、咳、ノック音、キーボード音」などの音声イベント、音声のテンション、声質も書き起こしている。また会議開始時間からのミリ秒単位の精度で、会議参加者ごとに、顔表情（Ekman の基本 6 表情）、動作、姿勢についても書き起こした。動作の書き起こしタグについては表 2 にまとめた。

### 4.6 取得実験の流れ

まず最初に被験者に対して実験の目的（研究用であり、通常の会議と同様の状況を収録したいこと）を詳細に説明し、理解させた。ビデオ会議システムに習熟させるため、しりとり、ゲーム、また被験者にとって日常的な話題などで練習させた。その後、収録する話題について説明し、「会議収録後、議事

表 3: 取得データの概要

| 回 | 話題 | 男女構成                  | 職業      | 時間<br>[分] | chat<br>[回数] | 発言回数<br>[ /10 分] |
|---|----|-----------------------|---------|-----------|--------------|------------------|
| 1 | 選択 | 女 4                   | 電話オペレータ | 29        | 164          | 210              |
| 2 | 選択 | 女 4 <sup>C1</sup>     | ファイリング  | 37        | 32           | 193              |
| 3 | 名付 | 女 4 <sup>C2</sup>     | 建築学科卒   | 61        | 54           | 146              |
| 4 | 名付 | 男 2 <sup>C2</sup> 女 2 | 専門学校    | 15        | 35           | 322              |

録を作成してもらう、メモ用紙は自由に使用していい」旨を伝えた。各個人が具体的な考えを持つための時間も設定した後、本収録を実施した。会議収録後、すぐさま被験者には、会議中使用を許したメモ用紙のみを参照して議事録フォーマットに沿って議事録を作成させた。その作業と並行して会議が活性化した場所を含む任意の箇所を 10 分間取り出しその部分のみ書き起こし、アンケート用紙を作成した。そのアンケート用紙をもとに会議参加者にはブラウザを使用し会議の様態をみながらアンケートを記入してもらった。

#### 4.7 収録概要

現在までに、4 回分の会議の収録が終わり、練習課題も含めると 8 話題分のデータが蓄積されている。取得状況を表 3 に示す。チャット経験がある被験者には“C”のマークを挿入している。ビデオ会議を経験者はいなかった。10 分間あたりの発言数は 150 から 320、チャットの回数も 35 回から 164 回となりビデオ会議の使用が初めてであるにもかかわらず、通常時とかわらない活発な議論が行われていることが示唆される。収録後のインタビューでも多少遅れを感じるがあったが、まったく普段とかわらない様子でリラックスして話ができたという回答が全員から寄せられ、自然な会議を実施できたといえる。

### 5 収録会議の解析の一例

第 3 回収録アンケートの構造タグと発言モーラ数の関連の解析を行った。3 人以上の被験者の区切れが一致する場所の前後の発言単位の平均モーラ数を比較した結果、切れ目直前は約 4 モーラ、切れ目が明確に意識されていないところでは平均的には約 7 モーラ、切れ目直後には約 32 モーラと話題の切れ目の直後の発話量が著しく増大することがわかった。これは話者二人の会話を対象に話題の区切れの解析を試みた結果 [14] を拡張し、信号処理による多人数ビデオ会議の構造推定の可能性を示す例である。

### 6 おわりに

「おっかけ会議」研究の構想・現状の進捗状況について報告し、信号処理による会話構造解析の可能性を示した。今後は収録したタグ、発話・動画、動作書き起し、議事録などの解析またそれらの関連の解析などを進め会議の構造・雰囲気解析技術の確立につとめていきたい。

### 参考文献

- [1] ascii24.com, <http://ascii24.com/news/i/net/article/2002/03/11/print/634296.html>
- [2] 小林稔, 岩城敏, “多地点接続インターネット会議システムを用いた在宅勤務実験,” 信学技報, OIS2002-11, pp.25-29, 2002.
- [3] 海谷治彦, “ハイパー議事録システムに関する研究,” 東京工業大学大学院理工学研究科博士論文, 1994.
- [4] 田中充, 福宿光徳, 西堀良久, 勅使河原可海, “同期・非同期型統合型マルチメディア会議システム ASSIST におけるマルチメディア議事録の開発と評価,” DICOMO99 予稿集, pp.79-84, 1999.
- [5] 横森正利, 上野和彦, “操作情報を利用した会議進行の記録・再生システム,” 情処学会, GN39-12, pp.65-70, 2001.
- [6] 長谷川将宏, 秋田祐哉, 河原達也, “談話標識の抽出に基づいた講演音声の自動インデキシング,” 情報処理学会研究報告, SLP-36-6, 2001.
- [7] Julia Hirschberg ら, “SCANMail : Browsing and Searching Speech Data by Content,” Eurospeech98, 1998.
- [8] 川口明彦ら, “同期型電子会議へのスムーズな途中参加支援のための一方式,” 情報処理学会論文誌, Vol.42, No.12, pp.3031-3040, 2001.
- [9] 大坊郁夫, “しぐさのコミュニケーション—人は親しみをどう伝え合うか—,” サイエンス社, 1998.
- [10] 泉子 K メイナード, “会話分析への招待,” 1996.
- [11] 松坂ら, “合意形成型対話における声・顔の機能的表情の分析と認識,” 信学技報, PRMU2002-118, pp.31-36, 2002.
- [12] 日高ら, “音声強調に着目したマルチメディアコンテンツ要約技術,” FIT2002 予稿集, K-36, 2002.
- [13] Yoichi Yamashita and Michiyo Murai, “An Annotation Scheme of Spoken Dialogues with Topic Break Indexes,” ICSLP2000, 2000.
- [14] 小椋敦子, “音声対話における意図や話題移行の表出に関わる言語特徴の分析,” 奈良先端修士論文, NAIST-IS-MT9951025, 1999.
- [15] 綿貫啓子, 外川文雄, “対話における感情の変化の解析,” 情報処理学会研究会報告, 音声言語処理 7-13, pp.79-84, 1995.
- [16] 田中和世, 伊藤克巨, 伊原正典, 岡隆一, “会議音声データの収録とデータファイル化,” 情報処理学会研究会, 音声言語情報処理 37-15, pp.85-90, 2001.
- [17] 犬童拓也ら, “会議途中参加支援を目的とした会議情報検索インタフェース—GijiLook の提案—,” 情報処理学会研究会, GN-49-21, 2003.
- [18] 栗田孝昭, 井合知, 北脇信彦, “オーディオビジュアル通信における伝搬遅延の影響,” 電子情報通信学会論文誌, Vol.J76-B-I, No.4, pp.331-339, 1993.
- [19] バイロン・リーブス, クリフォード・ナス, “人はなぜコンピュータを人間として扱うか,” pp.183-196, 翔泳社, 2001.
- [20] 波多野誼余夫ら, “認知科学の新展開 2 コミュニケーションと思考,” pp.118-133, 岩波書店, 2001.
- [21] Text Encoding Initiative, “The XML Version of the TEI Guidelines 11 Transcriptions of Speech,” <http://www.tei-c.org/P4X/TS.html>, 2001.