

## 研究課題別中間評価結果

1. 研究課題名： マルチモーダルな場の認識に基づくセミナー・会議の多層的支援環境

2. 研究代表者： 河原 達也（京都大学 学術情報メディアセンター 教授）

### 3. 研究概要

人間の知的活動の源泉ともいえる音声コミュニケーションを、マルチモーダルな観点で分析・モデル化した上で、リアルタイムに支援したり、効果的なアーカイブ化を行うための情報環境を構築する。まず、学術的なイベント等で一般的になっているポスター形式の発表を対象に、多様なセンサを備えた大型液晶ディスプレイによる「スマートポスターボード」の設計・実装を行っている。ポスター発表は今でも紙を用いる場合が多く、センサを備えたインタラクション環境は世界的にも例がない。本研究では、視線やあいづち・うなずきなどの聴衆の反応に着目したアプローチを提案し、学会等でデモ展示が行えるレベルに達しつつある。今後、実装・改良とポータブル化を進めて、様々な場面に適用・実証実験できるようにする。

次に、講演形式のセミナーを対象に、音声認識による字幕付与と、言語解析・知識処理による用語解説を目指して研究を行っている。音声認識を用いた字幕付与については、京都大学 OCW(OpenCourseWare)への適用や、一般向けのシンポジウムで実演を行うレベルに達している。今後、専門用語の解説の機能や聴衆の反応に基づく処理を導入するとともに、幅広い適用を目指す。

さらに、衆議院の会議録作成のための音声認識を 2010 年度から継続的に運用・改善し、音声認識と話し言葉処理に関する技術の洗練を行っている。これは、議会審議の可視化という観点で、政治学の研究者との共同研究にも展開されている。

### 4. 中間評価結果

#### 4-1. 研究の進捗状況及び研究成果の現状

##### 進捗状況及び研究成果

衆議院での会議録作成支援システムは、音声認識で 90%の正解率を達成し、本格的に運用されたことは評価する。この成果は、世界的に最先端の音声技術を持つチーム成果として会議録作成システムとしての先進性は高く、さらに、現場で実システムとして運用されたインパクトは大きい。

さらに、字幕付与の技術については、対象を OCW に変更したことも妥当であり、以上の二つの音声認識技術によって蓄積された経験は、今後の音声認識研究の一層の進展に寄与するものと確信する。その一方で、採択時に設定された「マルチモーダルな場の認識に基づくセミナー・会議の多層的支援環境」という研究課題の柱となるべき研究項目 1 については、国際会議で基調講演やデモを行うなどの成果はあるものの十分な進展がない。むしろ、既存技術の応用という範囲にとどまっており、それほど高い先進性や独創性が見られない。

##### 研究体制

自動音声認識による衆議院の速記代替技術を着実に向上させていること、字幕技術の対象を OCW に変更していることなど、要素技術が活用できる分野を明確に見据えた研究方針は、音声認識技術を背景とする分野では、一貫性を持った運営が行われてきた。

このため、成果として音声認識技術分野が強調された形になっており、「マルチモーダル」を着眼点とする研究項目 1 についての推進体制が弱い。今後の研究では、あくまで、研究課題の優先度が「マルチモーダルな場の認識に基づく」視点であることを再認識し、これに応じた最終的なゴールの提示方法について明確な方向性を打ち出すための、より強いリーダーシップを期待したい。

#### 4-2. 今後の研究に向けて

今後の研究に向けて、以下の3項目に注力した研究が求められる。

- 研究構想の軸に沿った研究展開を強化すべきである。本研究課題は、「マルチモーダルな場の認識に基づく」視点の提案である研究項目1に魅力を感じて採択されたものであることの再認識を促したい。研究項目2（字幕）、研究項目3（議事録）に関わる音声認識技術は、1の基礎となるべき重要な技術項目であるが、ある程度成熟した技術レベルに達している。
- スマートポスターボードの重要性について賛否両論あるが、一つの選択肢である。このスマートポスターボードを研究期間内の具体的なマイルストーンとし、さらに、例えば「自動プレゼンテーション」といった応用イメージの手がかりを明示するところまでの進展を期待する。
- 「マルチモーダルな場の認識に基づく」視点の研究は歴史が長く、関連の情報処理は日進月歩であるため、世界的な研究開発・技術動向に充分注意を払いながら研究を推進することが求められる。

#### 4-3. 総合的評価

- 衆議院での会議録作成支援システムは、音声認識で充分高い正解率を達成し、本格的に運用されたこと、字幕付与の技術については、専門性の高い講義である京都大学 OCW への音声認識技術の活用など、先進性や独創性は高く、現場での実システムとして運用された世界的なインパクトは大きい。このように、研究成果として音声認識技術に基づいたものが強調されており、これらは、研究代表者の専門性にバイアスされたものと推測する。
- スマートポスターボードに導入するマルチモーダル技術については現時点で具体的な姿が明確でない。国際会議で基調講演やデモを行うなどの成果はあるものの、現時点のシステムは既存の技術の応用にとどまっており、世界的に最先端の音声認識技術を活かした独創性と先進性を持つマルチモーダル情報処理技術が生み出されていない。今後は、全グループを上げて、この部分の研究に傾注し、勢いよく加速することを期待する。
- 当初提案された目標を達成できるかどうかは、今後のマルチモーダル情報処理技術の進展にかかっているといっても過言ではない。