

# テレワーク2.0を支えるマルチモーダル会話分析技術

富士通九州ネットワークテクノロジーズ株式会社

# テレワーク2.0を支えるマルチモーダル会話分析技術！

## ■ 目指す価値

COVID-19に伴い、ワークスタイルが対面からリモートへと急速に変化。

- ・ 仕事とプライベート空間が両立可能なテレワーク環境の提供。
- ・ 会話録画データを付加価値の高いデータへ変換して提供。

**「誰が・いつ・何を」話したかを可視化、構造化データへ変換。**

会話内容の整理・閲覧が容易となり、顧客満足度分析、マーケティング分析、従業員教育、FAQ蓄積、業務記録自動化などに多種多様な用途で活用可能！

## ■ 技術の特徴

- ・ 従来の音声技術単体では困難であった会話データの構造化（誰が・いつ・何を・どんな意図で）を**音声・画像・AI技術の融合**により実現。
  - 発話時の声と顔の動きに基づいて、誰が何を話したかを可視化
  - 上記とパラ言語情報(\*1)に基づいて、発話文言とその意図などを構造化

(\*1)言語情報のうち  
イントネーション、リズム、ポーズ、声質といった周辺言語

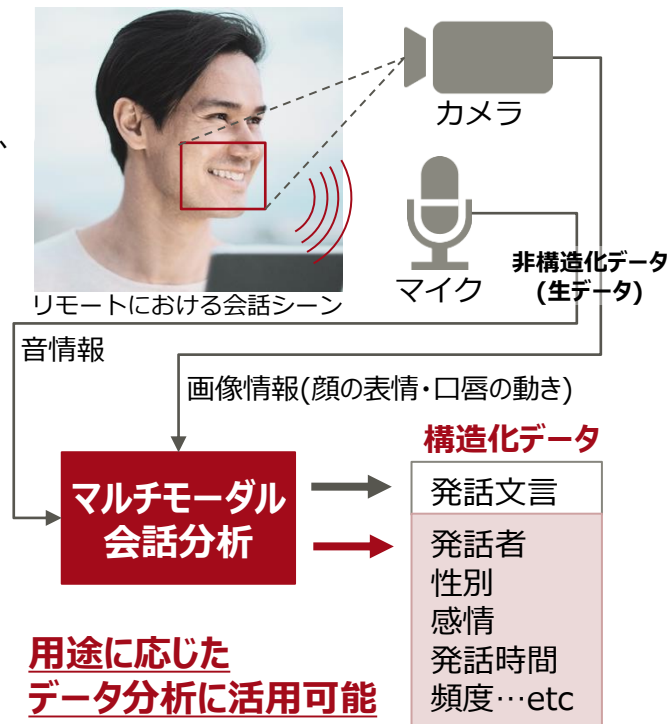
## ■ 技術開発の状況

- ・ 発話動画から発話者とその発話内容の可視化・保存を実現。
- ・ 発話動画から発話内容のテキスト化とその意図分析を検討中。

## ■ 活用計画

- ・ 多地点でのリモート会議やコミュニケーション・カウンセリングなど
- ・ 小売り業や教育現場などのサービス業にも領域を拡大していく予定。

開発戦略本部 第二統括部 第一開発部



# 世の中はテレワーク2.0へ

- COVID-19（新型コロナウイルス）の感染拡大により、コミュニケーションや仕事など生活様式が対面からリモートへと変化。
- その中で、ワークスタイルもオフィスとリモートの併用から、完全リモート(テレワーク2.0)の実現に向けた動きが加速。
  - 教育(オンライン授業)、医療(遠隔診療)、旅行業、小売業など様々な業界に広がると予想



対面会議

テレワーク2.0

完全移行



リモート会議

## ■ 課題 1 : 自宅での、仕事とプライベート空間の両立

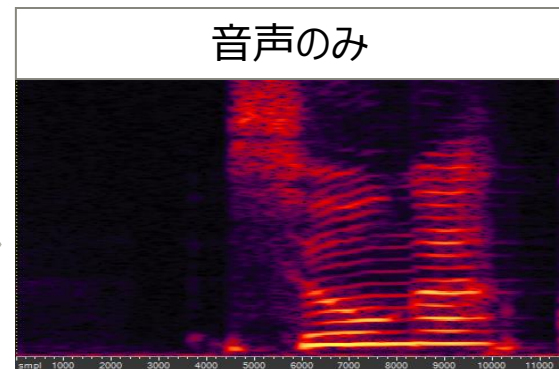
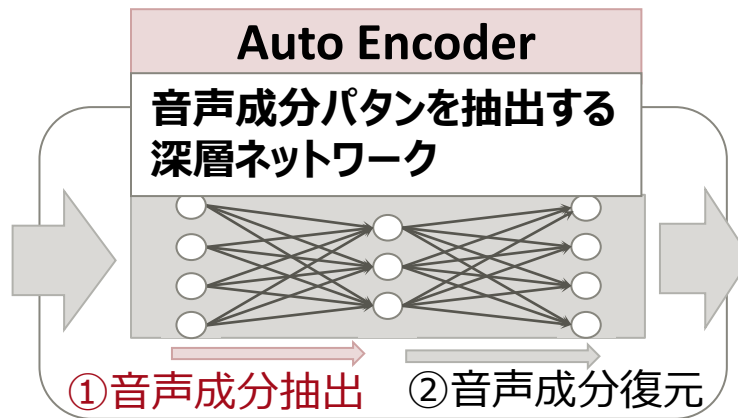
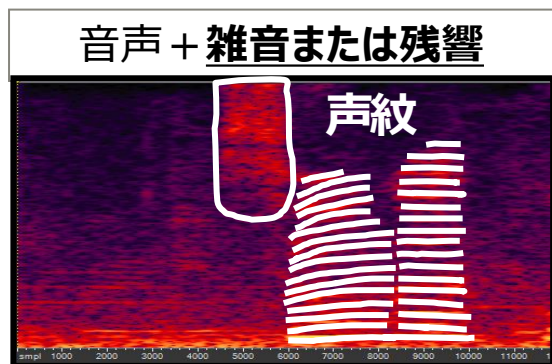
- ワークスタイルの変化に伴い、自宅スペースの確保が必要だが、物理的に困難。
- このため、電話会議において**生活音などの周囲騒音**や発話の声漏れなど会話の聞き取りにくさや、**同居家族への配慮**などが問題となる。

## ■ 課題 2 : 増え続ける会議音声データの整理・閲覧や利活用

- 電話会議に変化したことで、会議音声の録画が容易になりデータ活用のチャンス到来だが、
  - 録画データは増えるが、**データ整理や所望の情報を探し出すのが困難**。
  - 大量の**会話データを利活用した新たな価値やビジネス創出**をしたい。
    - 共通：会話音声の要約議事録、業務記録自動化
    - 医療：カウンセリング結果の傾向分析
    - サービス業：顧客満足度分析、マーケティング分析、従業員教育、業務ノウハウ蓄積

# 課題 1 : 仕事とプライベート空間の両立(1/2)

- 周囲騒音を除去し、話者音声だけを相手に届ける雑音除去技術。
  - 2018年度開発
- DeepLearning技術と音声信号処理技術の融合により、**多種多様な雑音の中から音声を直接推定し、音声以外を除去。**



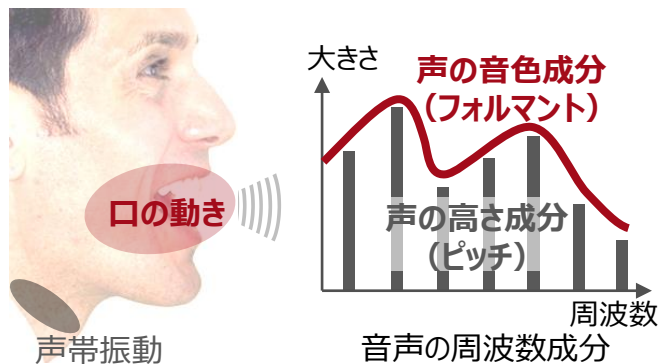
# 課題 1 : 仕事とプライベート空間の両立(2/2)

## ■ マルチモーダル音声抽出技術

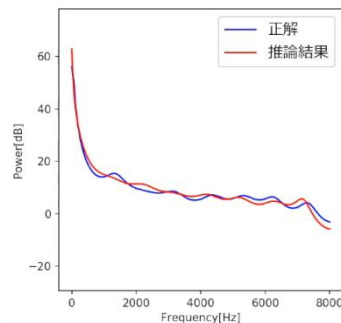
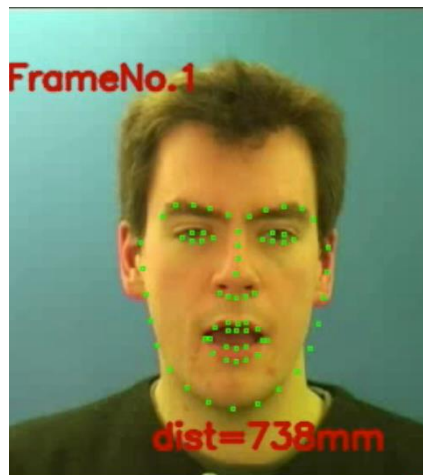
■ 2019年度開発

## ■ 口の動きから音声成分を推定する深層学習モデルを音声信号分析と融合 **雑音が大きい・声が小さい状況**でも音声成分のみを抽出

### 発声の仕組み



口の動き ⇒ カメラ



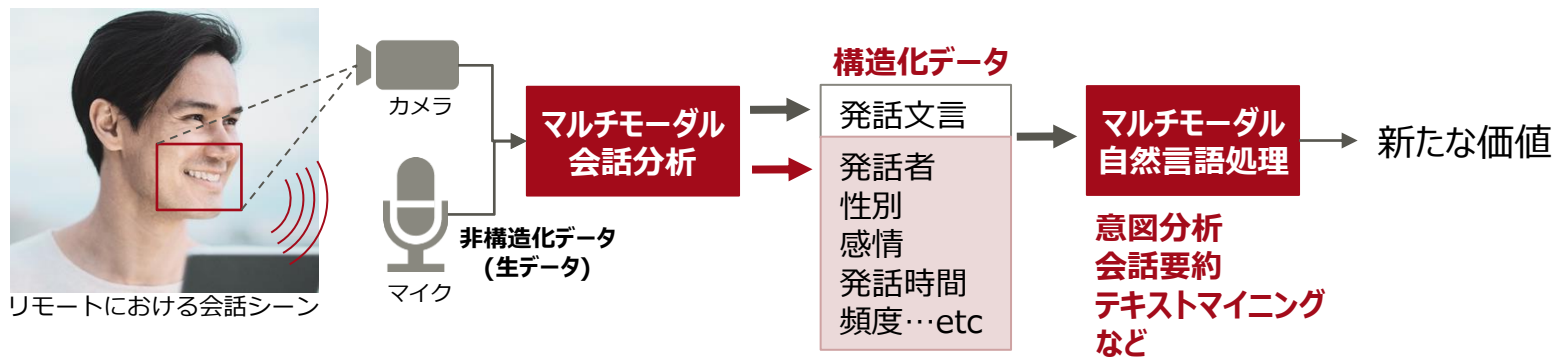
# 課題 2 : 増え続ける会議音声データの整理・利活用

## ■ 検索・分析しやすい形に構造化する、マルチモーダル会話分析技術。

- 音声や顔の情報に基づいて、「いつ・だれ」が話したかを構造化する  
発話検知・話者識別 **It's New!!**

- 更に、音声に基づくパラ言語情報の構造化データを用いたマルチモーダル自然言語処理  
(2020年度開発中) **It's New!!**

⇒ 発話の裏に含まれる話者の意図を分析に利用



# デモ：マルチモーダル会話分析

## 分析動画



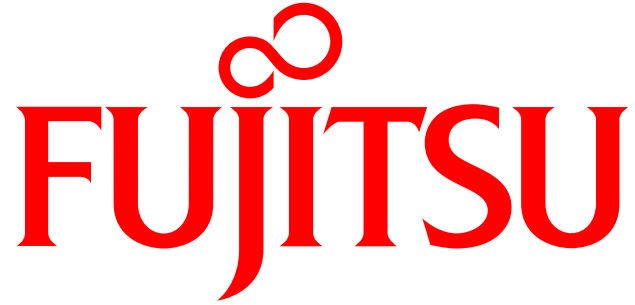
## 会話ログ

誰が	いつ	何を
C	00:05:53	おはようございます
B	00:10:44	おはようございます
A	00:14:78	おはようございます
C	00:18:32	こんにちは

※オフラインシミュレーション結果を用いて作成したアプリイメージです。



- 多地点でのリモートによる、コミュニケーション・カウンセリングなどで雑音除去により、どんな場所でも快適にリモート会議が可能。
- サービス業などは、本来テレワーク担当と顧客の中で閉じてしまい非参加者からは会話内容が見えないが、雰囲気も含めて見える化することで、テレワークでも対面と同じサービス品質を実現できる。



shaping tomorrow with you