

VoiStrap : ID カードネックストラップ型の音声センシングデバイスを用いたヒューマンコミュニケーション可視化システム

原田陽雄 米山博人 下谷啓 藤居徹 西野洋平 飯田靖 西原義雄

富士ゼロックス株式会社 システム技術研究所

1. はじめに

近年、グローバルマーケットにおける日本企業のプレゼンスの低下が危惧されている。その要因としてコミュニケーショントラブルによる組織力の低下を指摘する声は強い。かつての競争力を取り戻すため、組織内コミュニケーションの活性化が必要との機運は年々高まっている[1]。しかし、コミュニケーショントラブルが組織の文化や風土に起因するケースは多く、スローガンのような情緒的なアプローチだけで改善を進めることは容易ではない。そこでわれわれは、データに基づく科学的アプローチがコミュニケーション活性化の切り札になると考え、発話のセンシングによってヒューマンコミュニケーションを可視化する技術“VoiStrap”を開発している。本稿では VoiStrap システムの概要と、会議を対象とした可視化例を報告する。

2. VoiStrap システムの構成

図 1 に VoiStrap システムの基本構成を示す。設計コンセプトとして、①対話状況をリアルタイムに可視化できる、②特別な器具の装着等でユーザーに肉体的負担を与えない、③ユーザーにプライバシーやセキュリティの不安を抱かせない、④シンプルな構成で簡単に現場導入できる、を重視した。

多くのオフィスワーカーが日常的に装着しているネックストラップ型の ID カードに着目し、センシングエンジンを組み込んだ。ストラップ内に埋め込んだ小型マイクで発話状況をセンシ

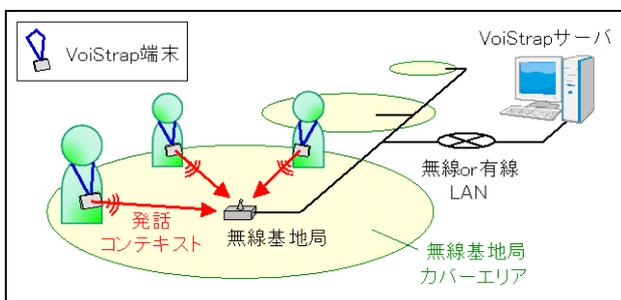


図 1 VoiStrap システムの基本構成

VoiStrap: Human communication visualizing system using ID-card with neck strap type voice sensing device
Haruo Harada, Hirohito Yoneyama, Kei Shimotani, Akira Fujii, Yohei Nishino, Kiyoshi Iida and Yoshio Nishihara
Fuji Xerox Co., Ltd. System Technology Laboratory



図 2 VoiStrap 端末の概観

ングし、点在する無線基地局にデータをリアルタイム送信する。発話内容(コンテンツ)ではなく、発話タイミングや韻律(コンテキスト)だけをデータ化し、プライバシーやセキュリティへの不安を回避する。無線基地局が受信したデータを有線または無線 LAN を介してサーバーに集め、信号解析/データマイニングを施してコミュニケーション状況を可視化する。無線基地局はタバコケース程度のサイズで、20~30m おきに机上設置される。サーバーには一般的な PC を利用でき、大がかりな設置工事を必要とせず簡単にシステムを構築できる。

図 2 に VoiStrap 端末の概観を示す。バッジ部のサイズは 91mm×58mm×12mm で、表側にクレジットカード大の ID カードを装着できる。端末全体の重量は約 60g である。これまでに行った検証実験で重量や装着感に対するクレームは受けていない。バッジ内部に、マイク信号の増幅回路、マイコンを用いた信号解析回路、ZigBEE 規格の無線回路、電源回路、リチウムイオンバッテリーを格納している。無音時には回路をスリープさせるためバッテリーライフは発話量に依存する。50%duty で発話を行った場合でも 8 時間以上の動作が可能である。

3. 基本アルゴリズム

図 3 に装着者本人の発言か他者の発言かを分離するアルゴリズムを示す。ストラップ内の首元およびバッジ近傍にマイクを装着することがポイントである。例に示すように発声点と 2 つのマイクとの距離は自分では大きく、他者では小さくなる。一方、音圧はほぼ距離の二乗に反

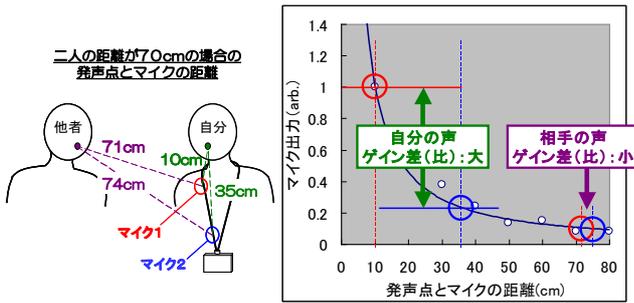


図3 自他発言分離アルゴリズム

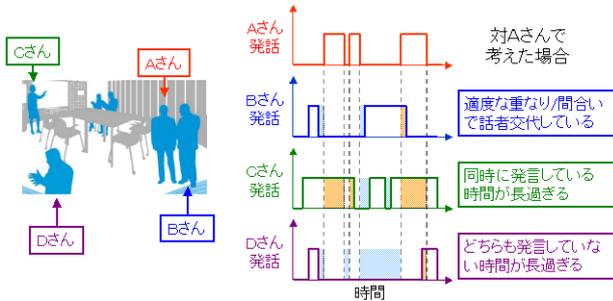


図4 会話ペア判定アルゴリズム

比例して減衰するため、自分発話のように音源に近い領域では急峻に低下する。これらから2つのマイク信号の差（または比）を観測すれば、自分では大きな値、他者では小さな値として発話者を分離できる。音圧の相対値を用いているため、声の大きさに依存せず安定した判断が可能である。

さらに自他分離アルゴリズムで求めた発話タイミングから、誰と誰が会話しているかを推定する。図4の概念図に示すように会話はキャッチボールという前提で、交互にタイミング良く発話しているかどうかを評価する。この例では、発話信号の重なり、間合いのいずれも小さいAさんとBさんが会話していると判断する。

4. 会議におけるコミュニケーション可視化例

5人の参加者で行った会議での測定例を紹介する。図5(A)に自他分離アルゴリズムで抽出した各参加者の発話タイミングを示す。実際のシステムでは1秒おきにデータが更新され、外部からでもリアルタイムに発話状況を確認できる。同図(B)に5分ごとの話者交代回数を示す。会議全体の流れの中で会話の密度がどのように変化していったかを把握できる。場の盛り上がり推定への応用などを想定している。

特徴的な結果が得られた2つの会議について、各参加者の発話割合および各参加者間での会話回数を図6に示す。(A)では進行役の発話量が圧倒的に多く、会話も進行役と各参加者間を中心

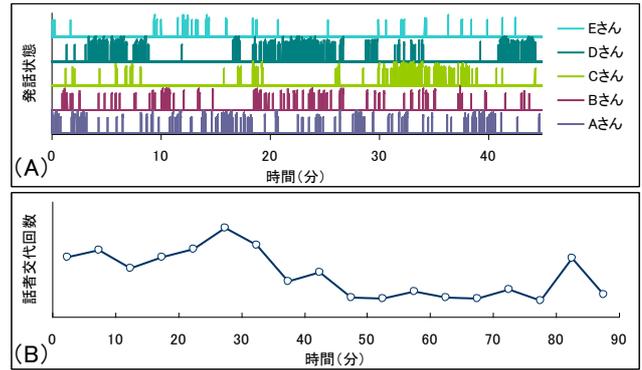


図5 会議可視化例：(A)発話状態 (B)話者交代回数

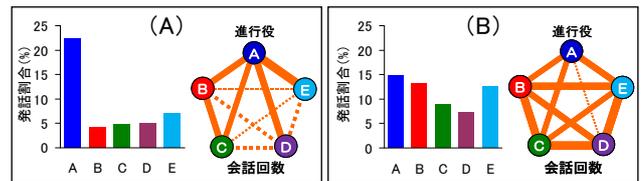


図6 会議可視化例：発話割合と会話回数

に行われている。一方(B)では各参加者が均等に発言しており、参加者同士の会話も活発に行われていることが分かる。

現在、このような VoiStrap データと、会議目的の達成度や参加者の満足度の相関を調査している。これまで抽象的に語られていたベストプラクティスを定量化し、コミュニケーション活性化を自律サポートするシステムを実現したいと考えている。

5. おわりに

今後、会議に限らずあらゆるビジネスシーンへの応用や、教育現場など他分野への展開も視野に入れて研究を進めていく。より可視化レベルを高めるため、センサデバイスとしての精度アップと、発話コンテキストから多くの情報を抽出できるアルゴリズムの探索を続けていく。一方、システムコンセプトを崩すことなく、位置や感情などの付加情報をセンシングする技術にも挑戦していく予定である。

長年の歴史の中で形成された組織の文化や風土を変革するには大きなエネルギーを必要とする。VoiStrap の提供する客観性が、組織改革や自己変革に伴う痛みを許容できるだけの“納得感のあるところへの働きかけ”に繋がることを期待している。

参考文献

[1] たたとえばビジネスコミュニケーション白書 2010, 一般社団法人 日本経営協会