

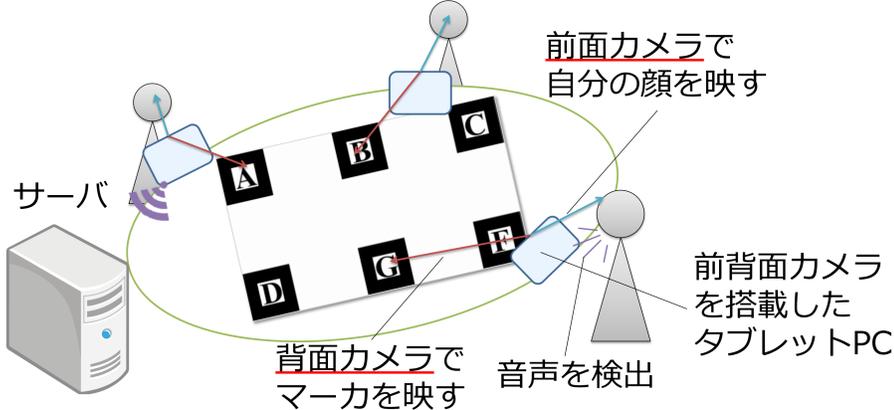
1. 研究背景・目的

対面での多人数会話において、人により発話の多寡があり、発話の少ない人からの情報が提供されないような場面が考えられる。そのような人から情報を引き出すよう働きかけたい。そのため、会話の活性化を目的とした、タブレットPCの前背面カメラを活用し、会話の様子を定量的に可視化するシステムを製作する。

2. 会話特徴量の可視化システム

システムの概要

- ユーザが囲むテーブル上にマーカを配置
- 前背面カメラを搭載したタブレットPCを所持した状態で会話
- 各ユーザの発話特徴量を取得
- サーバで発話特徴量を統合し、タブレットPCに可視化表示



個人の発話特徴量

- マーカを基準としたユーザの位置と顔の方向
- どのくらいの音量で発話しているか

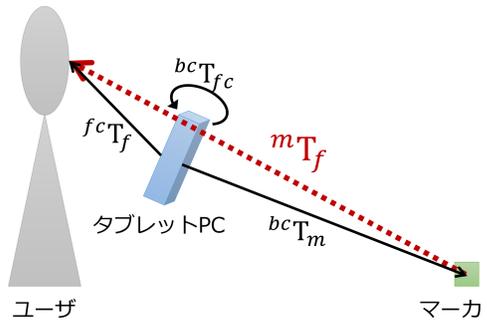
多者間の会話特徴量

- 個々人の発話特徴量を統合
- 誰が誰に話しているか、およびそのこれまでの累積量

3. 発話特徴量の計測

ユーザの位置・姿勢の推定

タブレットPCの前面・背面のカメラを同時に使用し、マーカを基準としたユーザの顔の位置と姿勢を求める。



- bcT_m : 背面カメラからマーカへの座標変換行列
- bcT_{fc} : 背面カメラから前面カメラへの座標変換行列
- fcT_f : 前面カメラから顔への座標変換行列

$$T = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} \text{回転行列} \\ + \\ \text{並進成分} \end{matrix}$$

bcT_m : ARToolKitを用いて計測



マーカに基づく背面カメラ画像からのタブレットPCの位置・姿勢推定

bcT_{fc} : 前背面カメラの位置関係が既知であることより定める

fcT_f : (株)オムロンのOKAO® Visionによる顔検出・顔姿勢推定, 物体の実サイズと投影サイズの関係のキャリブレーションに基き推定



前面カメラ画像からの顔検出・顔姿勢推定

3つの座標変換行列を掛け合わせる

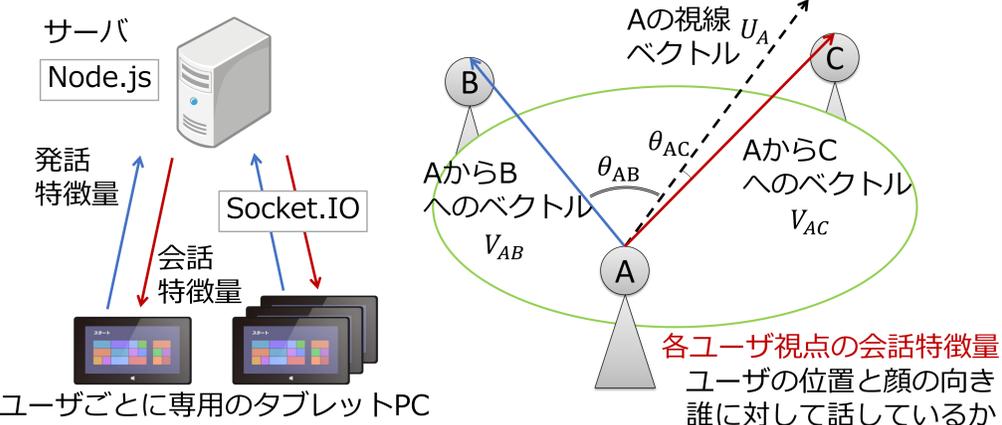
- マーカから顔への座標変換行列 $mT_f = (bcT_m)^{-1} bcT_{fc} fcT_f$
- テーブル座標系でユーザがどこにいてどちらを向いているか計測できる

発話音量の計測

- タブレットPCのマイク入力を利用し、閾値以上の音声が発出されたときを発話とする
- 発話しているとき/していないときの音声信号の二乗平均平方根 (RMS) の値から、あらかじめ閾値を定めておく

4. 会話特徴量の算出・可視化

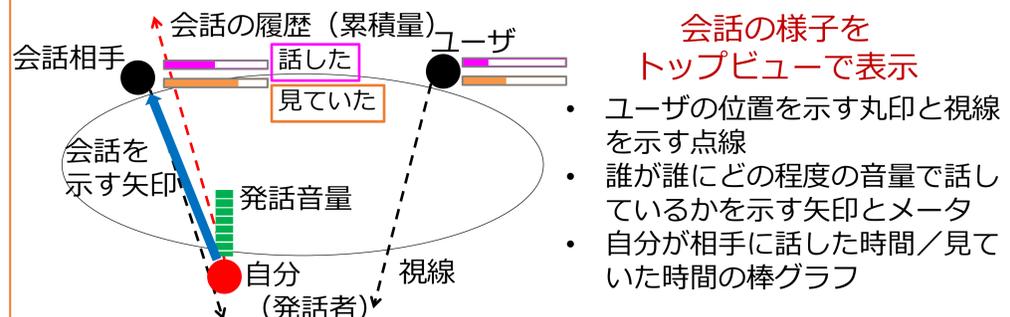
会話特徴量の算出



会話相手の算出

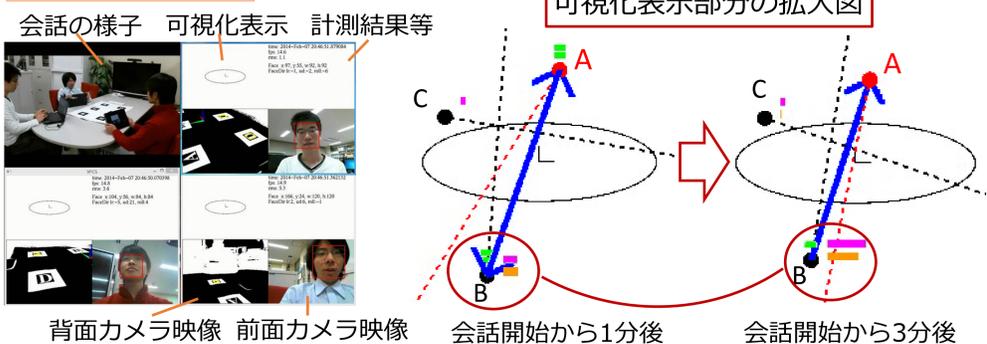
視線ベクトル U_i と他ユーザへのベクトル V_{ij} の類似度を比較
 類似度 $Sim_{ij} = \cos \theta_{ij} = \frac{U_i \cdot V_{ij}}{\|U_i\| \|V_{ij}\|}$
 最も類似している (Sim_{ij} が 1 に近い) 相手を算出する

会話特徴量の可視化



5. 会話特徴量の可視化の動作確認実験

動作確認実験



- 各ユーザの位置や顔の向きがおおむね正しく得られている
- 顔の向きの計測の誤差は水平方向に $\pm 2^\circ$ 程度
- 顔会話相手の算出ができており、会話履歴も記録できている

会話の活性化に向けた仕組みの例

会話の活性化に向け、会話へのゲーム性の導入例として、会話に得点を設け、それを可視化することを考える。ユーザの会話特徴量から、話したときの報酬 $R^{(S)}$, 話しかけられたときの報酬 $R^{(L)}$ を設定し、その合計の時系列積分値を得点 P とする。

話したとき $R^{(S)}$ 話しすぎると減点される

聞いたとき $R^{(L)}$ 話しを聞くと一定の報酬を得る

| ユーザ名 | 得点P |
|------------|-------------|
| UserA | [Bar chart] |
| UserB | [Bar chart] |
| UserC (自分) | [Bar chart] |

全員の累積発話量に対する自分の累積発話量の割合 $\frac{S_i}{\sum S_j}$

- うまく会話を進めるユーザほど得点が増加する
- 一方的に話すだけでなく、聞き手にも回る必要がある

6. 今後の課題

- 精度評価 (顔検出, マーカ検出, 音量検出におけるノイズ)
- 会話の活性化に向けた仕組みの妥当性・有効性の確認
- トップビューの可視化ではなく、背面カメラ映像への会話特徴量の重畳表示