

会議における発言の音響的特徴と言語的特徴に基づく自動評価

岡本 康佑[†] 松原 茂樹[‡] 長尾 確[‡]

名古屋大学工学部電気電子・情報工学科[†] 名古屋大学大学院情報科学研究科[‡]

1 はじめに

大学のセミナーなど様々な場面で会議が行われている。充実した会議を行うためには参加者の議論能力を高めることが重要である。

会議中に参加者が互いの発言を評価することによって議論能力や参加者の会議への参加意欲の向上を促す研究が行われている[1]。しかし、会議中に他者の発言を評価することは参加者の負担が大きく、また、客観的な基準で他者を評価することは難しい。

そこで本稿では、会議中に取得した発言の音声と書記が作成した議事録を用いて、発言をその音響的特徴と言語的特徴に基づき自動評価するシステムを提案する。

2 会議における発言と議論の構造化

著者らの研究室では、ディスカッションマイニングという会議記録システムを提案・運用しており[2]、図1に示すような環境で会議を行っている。参加者はリモコンとタブレットを所持しており、発言する際はリモコンを用いて発言を登録し、それが話題を提起する発言(導入発言)であるか話題を展開する発言(継続発言)であるかを表明する。発言が終わったら発言の終了をリモコンで表明する。

システムは発言の音声を記録している。また、導入発言から次の導入発言の直前までの継続発言を議論セグメントとして分割し、発言間の関係に基づき議論セグメントを木構造として表現している[2]。



図 1: 会議の様子

Automatic Evaluation of Statements in Meetings Based on Acoustic and Linguistic Features

[†]OKAMOTO, Kosuke(okamoto@nagao.nuie.nagoya-u.ac.jp)

[‡]MATSUBARA, Shigeki(matubara@nagoya-u.jp)

[‡]NAGAO, Katashi(nagao@nuie.nagoya-u.ac.jp)

Department of Information Engineering, Nagoya University[†]

Graduate School of Information Science, Nagoya University[‡]

3 発言の自動評価

会議では参加者は同じ事柄について話し合い、相手の話を分析し自分の主張をわかりやすい話し方で伝える必要がある。文献[3][4]を参考に、自動評価の評価指標を定めた。

自動評価は発言ごとに行う。また、音声認識システム Julius[5]により発話区間を検出し、自動評価に用いる。

評価指標は音響的特徴に基づくものと言語的特徴に基づくものがある。

3.1 音響的特徴に基づく評価指標

以下は音響的特徴のみで評価する指標である。

- 声の大きさ**:参加者が十分に聞き取れるくらい声が大きく、その一方で、感情的になり大き過ぎないのが良い。発言中の各発話の音量[dB]を計測し、評価する。
- 声の抑揚**:抑揚のない発話は聞き手を退屈させる要因となる。発話における声の高さ(F0)[Hz]を測定し、その標準偏差の値が高いものを良い評価とする。
- 話の速さ**:発話は速すぎても遅すぎても聞き取りづらい。話速(時間当たり音節数)が適切な範囲内であれば良い評価とする。
- 流暢さ**:言い淀みや沈黙の多い発言は聞き取りづらく理解しづらい。発言中の「えー」などの有声休止(母音の引き延ばし)による言い淀みの回数、及び、発話間の2秒以上の無音区間を沈黙としたときの沈黙回数が少ない発言を良い評価とする。
- 非単調性**:強調したい部分が明確であると話を理解しやすい。強調したい部分をゆっくり話したり、その前に間を設けるなど、話が非単調であることが効果的である。話速の標準偏差、及び、発話の間にある2秒以下の無音区間を「間」としたときの「間」の回数に基づいて発言の非単調性を評価する。

3.2 言語的特徴に基づく評価指標

言語的特徴による評価指標を以下に示す。

- 簡潔さ**:発言は簡潔であると理解しやすい。評価のために音声認識による会議中の発言の音節数と議事録のテキストの音節数を比較する。書記は発言内容を要約して

記述するため、発言の音節数と議事録のテキストの音節数が近ければ発言が簡潔であるとみなせる。

- g. **主題との合致性**: 発言が主題に合致している、すなわち、議論の主題に沿っている必要がある。議論中の発言が話題提起の発言の内容と共通していれば主題に沿っているといえる。導入発言者の発言との関連度(後述)を求めることで評価する。
- h. **結束性**: 発言が結束している、すなわち、議論の流れに沿っている必要がある。発言が議論構造上の親の発言の内容と共通していれば流れに沿っているといえる。継続発言に対してその親の発言との関連度(後述)を求めることで評価する。

発言間の関連度は以下の方法で算出する。1回の会議の各発言を1つの文書としたTF-IDF法を用いて単語を以下の式

$$f(t, d_1, d_2) = \left(\frac{tf(t, d_1)}{\sum_{s \in d_1} tf(s, d_1)} + \frac{tf(t, d_2)}{\sum_{s \in d_2} tf(s, d_2)} \right) * idf(t)$$

で重みづけし、関連度を求める2つの文書 d_1, d_2 に共通して現れる単語についてはその値を加算し、片方にしか現れない単語についてはその値を減算し、すべての単語について合計する。この値を発言間の関連度とする。

4 自動評価結果のフィードバック

評価した値をもとに、各評価指標について「良い」「普通」「悪い」の3段階で発言をランク付けし、参加者にフィードバックする。

4.1 会議中の評価の提示

音響的特徴のみを用いた評価指標 a. ~ e. は会議中に評価をフィードバックできる。声の大きさ、声の抑揚、話の速さについてリアルタイムに評価し、「悪い」評価となったとき、即座にスクリーン上に警告をポップアップする(図2)。

また会議中、参加者のタブレットに対して、発言が終わるごとに声の大きさ、声の抑揚、話の速さ、発言の流暢さ、発言の非単調性の評価

をアイコンにより提示する。アイコンはランクに応じて色分けされ、一目で自分の評価を知ることができる(図3)。

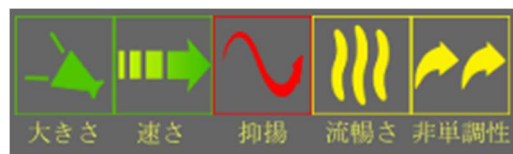


図3: 評価アイコン

4.2 会議終了後の評価の提示

会議終了後には、議事録を用いることにより、言語的特徴に基づき評価できる。参加者がWeb上で会議を振り返るための既存のインタフェース[2]上に評価結果を示すアイコンを提示する。アイコンはランクに応じて色分けされ、各評価についての改善の方針を示すコメントを参照することができる(図4)。

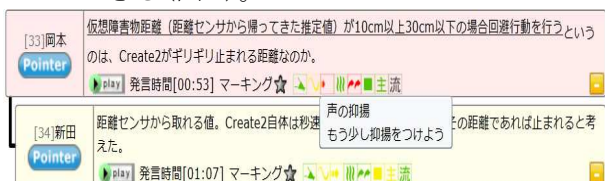


図4: Web上の議事録での提示

また、参加者に対して、その会議を通した各評価指標ごとの平均、参加者全員の評価指標ごとの平均、前回から向上した評価指標、及び、次回への改善点をメールで通知する。

5 おわりに

本稿では、会議中に取得した発言の音声と議事録を用いて、発言を音響的特徴と言語的特徴に基づき自動評価し、評価結果をフィードバックする手法の提案を行った。

今後の課題として、本システムによる評価の信頼性の向上が挙げられる。また、評価に基づく発言の分析や総評などの議論能力向上のための効果的な提示方法を検討する必要がある。

参考文献

- [1] 川西, 大平, 長尾, 議論能力向上のためのゲーミフィケーションシステムとその評価, 情報処理学会論文誌, Vol. 1, pp. 277-278, 2014.
- [2] 長尾, ディスカッションマイニング: 対面式会議での議論からの知識発見, 信学技報, TL, Vol. 112, No. 339, pp. 59-64, 2012.
- [3] 松本, 日本語ディベートの技法, 七寶出版, 2001.
- [4] 足立, 議論の論理: 民主主義と議論, 木鐸社, 1984.
- [5] 李: 大語彙連続音声認識エンジン Julius, <http://julius.osdn.jp/>, (2015年1月参照).



図2: スクリーン上のメッセージ表示