



対面コミュニケーションにおける合意形成推定の一提案

藤本 雄樹* 大久保 雅史*

Proposal of Estimating Method for Agreement in Face-to-Face Communication

Yuki Fujimoto*, Masashi Okubo*

Abstract - Generally, in the meeting exchange of opinions, sharing of knowledge, and creation of new ideas are performed through the discussion. However it is difficult to make sure whether the participants agree with the opinion of the speaker. If there is a support system that estimates the agreement of participants with the opinion of the speaker, the meeting may advance more smoothly. In this research, we propose estimation method for agreement by using correlation between talker's motions in face-to-face communication.

Keywords: non-verbal information, agreement formation, communication motion, nod

1. はじめに

1.1 研究の背景

一般的にディスカッションなどの会議では、意見の交換や知識の共有化、新規アイデアの創出が行われる。しかし、会議の進行における問題点として、発言者の意見に対し、会議の参加者が同意か不同意かを見極めが困難ということがあり、合意形成の確認が不十分のまま会議が進行することがある。一方で、会議のような対面コミュニケーションにおいて、うなずきや表情などのノンバーバル情報が重要な役割を果たすと言われている^[1]。さらに、うなずきは同意を示す手段だけでなく、発話権の授受を制御し、会話の流れを円滑にするとされている^[2]。したがって、うなずきを推定することで、会議を円滑に支援できる可能性があると考えられる。

1.2 関連研究

うなずきの推定に関する研究として、多人数会話におけるうなずきの会話制御の分類を目的とした、斎賀らの研究がある^[3]。この研究では、前頭部に装着した小型無線加速度センサにより、首振り動作を取得している。うなずきの推定方法は、閾値処理を用いている。しかし、連続したうなずきは、ノイズ区間として誤検出される可能性があり、周波数解析により推定を行っている。また、頭部の動きの周波数から講師-受講者間のコミュニケーションの質を評価している木村らの研究がある^[4]。この研究では、首から下げたスマートフォンに搭載された加速度センサより X, Y, Z の 3 軸を計測し、各聴講者の加速度データの周波数分布を作成している。結果として、3~6 Hz は通常のうなずきや手の動き、6 Hz 以上は素早いうなずきや笑いに対応していることが確認されている。以上から、発表者や各聴講者の身体動作の特徴を周波数解析より推定できる可能性が示されている。しかし、周

波数解析によって抽出できるうなずきや笑いが会話にどのような影響を及ぼしているのか、明らかとなっていない。一方で、対面コミュニケーション中に生じるうなずきに着目した研究として前田らの研究がある^[5]。この研究は、対話によって生じるジェスチャー情報(うなずき)に、相互的關係が存在するかを分析することを目的としている。始めに、内部にビデオカメラを設置しているプロンプター・イヤホン・マイクで 2 者間の対面会話を収録している。会話後、収録データから対面会話中に表出しているジェスチャーをアノテートし、分析を行っている。結果として、うなずきはあいづちと同様に相手話者の発話に対する何らかの応答動作として生じるよりも、自己発話内で多く生じる傾向があり、さらに、うなずきが 2 人の話者で同時に発生する現象が多いことも示されている。

1.3 研究の目的

以上の関連研究では、周波数解析によりうなずきや笑いなどの特徴的な身体動作を推定できることを示している。しかし、それぞれの研究において、身体動作の計測に用いられているものは、頭部に装着した加速度センサや首から下げたスマートフォンであり、実際の対面会話を想定すると、より自然な状態での会話を計測することが望まれる。そこで本研究では、非接触で頭部の位置を取得する手段として Kinect を用いる。また、相手の発話に同意の意味を示すうなずきは、対話中に同時に起こりやすいとされていることから、対面会話におけるうなずき時の頭部動作の相関値が高いと推測される。すなわち、頭部動作の相関の強弱により、同意・不同意が推定できる可能性がある。発言者の意見に対する聞き手の同意・不同意の推定を支援することができれば、会議がより円滑に進む可能性があると考えられる。そこで、本研究では、話し手と聞き手の身体動作の相関の強弱を合意の指標とした合意形成推定の手法を提案する。

*: 同志社大学大学院 理工学研究科

*: Graduate School of Science and Engineering, Doshisha University

2. Kinect による頭部の位置の取得

より自然な状態で頭部の位置を推定するために、NUI(Natural User Interface)の1つで骨格情報を取得することができる Kinect を用いている。Kinect は原点を本体に持ち、人の頭部の位置データ(X 軸, Y 軸)を取得する。図 1 は、Kinect からの位置情報を取得する際の概略を示している。右の図は、1 回のうなずき時に得られる X 軸, Y 軸の時系列データの例である。うなずいた際には Y 軸の時系列データに図のような、下に凸の信号が得られる。

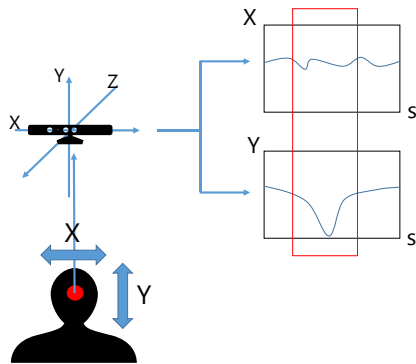


図 1 Kinect から得られる頭部の位置データ
Fig.1 Position data of head provided from Kinect

しかし、X 軸, Y 軸での頭部の位置データを取得する際、ノイズの影響で頭部の上下運動が小さい動作を取得することが困難であった。この問題を解決するために、顔の両目・鼻・両口角の 5 点から頭の回転運動を取得するクォータニオンを用いたアルゴリズムを開発している。このアルゴリズムでは、人の顔の動きロール・ピッチ・ヨーを取得する。本研究では、特に顔の X 軸周りに回転した角度(ピッチ)に着目する。図 2 は Kinect を正面として、作成したアルゴリズムで顔の X 軸周りに回転した角度を取得する概略を示している。この図の θ が顔の動いた角度であり、頭部の動きとして解析に利用している。

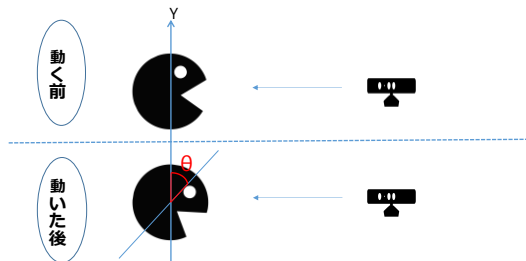


図 2 Kinect から得られる顔の縦方向の角度データ
Fig.2 Angle data of vertical direction of face provided from Kinect

3. 予備実験

3.1 実験目的・方法

1 回のうなずきの周波数帯域を求めるために予備実験を行っている。木村らの研究でうなずきと素早いうなず

きの周波数帯域が異なっていたため、深いうなずきと浅いうなずきの 2 種類の動作を対象に周波数帯域を求める。10 人の実験協力者に深いうなずきと浅いうなずきを 5 回ずつ行わせるのを 1 セットとし、5 セット行わせている。つぎに、得られたうなずきのデータに対して、深いうなずき・浅いうなずき共に、3s 間のデータでパワースペクトルを求め、最大の周波数の値を求める。図 3 は、1 人目の深いうなずきの 1 セット目のデータからパワースペクトルを求めた例を示している。他の 9 人の実験協力者に対しても同様な手法を行い、1 人ごとに深いうなずき・浅いうなずき共に 5 回ずつ、パワースペクトルの最大の周波数を求める。

○: 最大の周波数

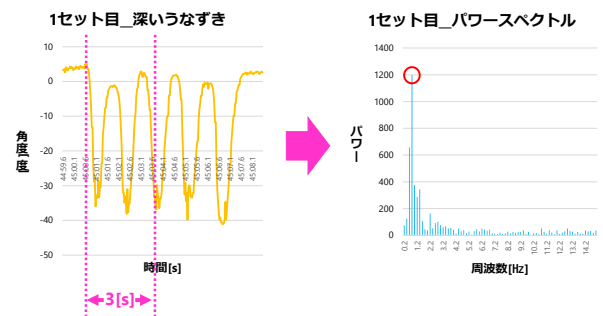


図 3 うなずきのデータとパワースペクトルの例
Fig.3 Example of nodding data and its power spectrum

3.2 予備実験の結果

図 4 に、深いうなずき・浅いうなずきの周波数の平均を示す。有意差の検定には t 検定を用いている。深いうなずきの平均が約 0.47Hz で浅いうなずきの平均が約 1.41 Hz で、有意な差が見られた。このことから、1 回の深いうなずきにかかる平均時間が 2.13 秒、浅いうなずきでは 0.71 秒と分かる。ここで、対面会話で行われるうなずきを考えた場合、深いうなずきより浅いうなずきの頻度が高く、会話中の同意によるうなずきの同期は浅いうなずきによるものだと考えた。そこで、本研究では浅いうなずきのみ着目し、1 回のうなずきの周波数帯域は、1~3 Hz と定めた。

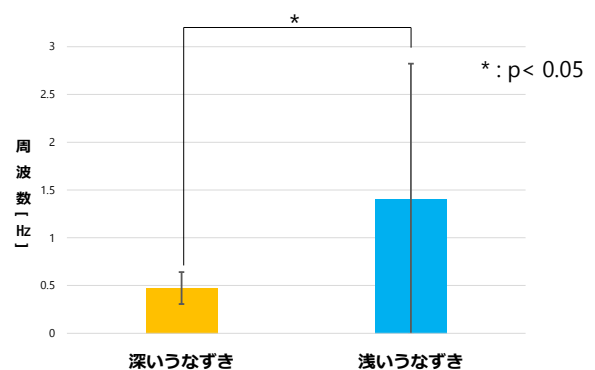


図 4 深いうなずき・浅いうなずきの周波数の平均
Fig.4. Average of frequency of deep nod and shallow nod

4. 本実験

4.1 実験目的・方法

話し手と聞き手の身体動作の相関の強弱から、聞き手が話し手の意見に対して合意しているか否かを推定することができるかを検証するために、同志社大学生2名(実験協力者A・実験協力者B)に会話をさせている。まず、会話の前に二者択一の質問が書かれたアンケートに回答させる。アンケートの回答が異なった質問をテーマとして実験協力者らに会話させる。1つのテーマごとに制限時間は設けていない。1つのテーマでお互いが話し合い、片方の回答に合意した場合、そのテーマでの会話を終了させ、実験協力者らで次のテーマを選択させ会話をさせる。会話中、各実験協力者の動作をKinectとビデオカメラで取得している。Kinectでは実験協力者の顔の縦方向に動いた角度を取得しており、ビデオカメラでは対話中の実験協力者の様子(表情・身体動作・音声)を収録している。図5に実験概要を示す。

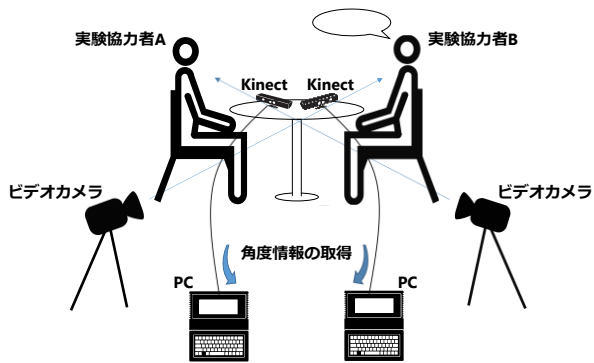


図5 実験概要

Fig.5 Experimental setup

実験後、Kinectから得られた各実験協力者の顔の縦方向に動いた角度データに対し、3sごとのデータでパワースペクトルを求める。求めた後、1sずらし再度3s間のパワースペクトルを求める。図6は、実験協力者Aの10分間の会話データにおいて、3sごとのデータのパワースペクトルを求める流れを示している。会話データの開始から3s間のパワースペクトルを T_1 とし、同様に T_{598} まで求める。

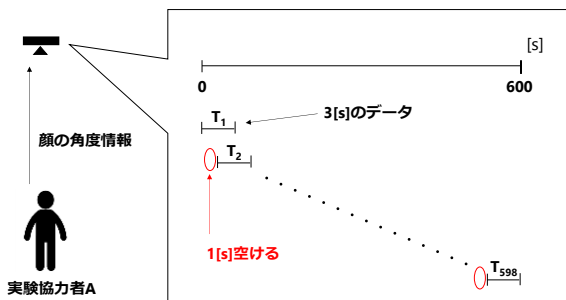


図6 解析領域を求める流れ

Fig.6 Flow of obtaining analyzing area

ここで、「3.2 予備実験の結果」から解析の対象となる周波数帯域は1~3 Hzと定めている。実験協力者A・実験協力者BのTごとのパワースペクトルの1~3 Hzでの相関値を求める。下の図は、実験協力者A・実験協力者BのTごとの相関値の時系列データを示している。この図で T_1 の相関値を求める式において、実験協力者Aの T_1 の周波数ごとのパワースペクトルの数値を A_i 、平均を A_{ave} としている(実験協力者Bも同様としている)。

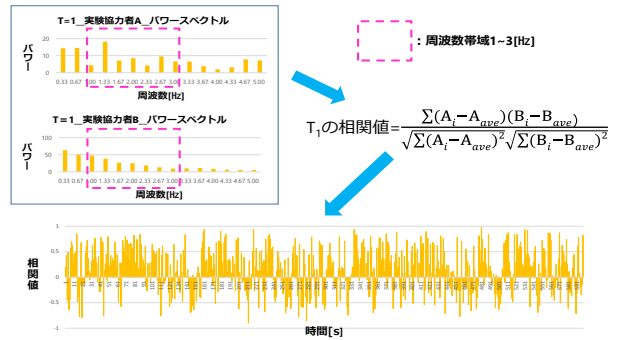


図7 相関値の時系列データ

Fig.7 Time series data of correlating coefficient

4.2 結果

作成した相関値の時系列において、相関の強弱を示すため、相関値が同じ傾向の区間が5つ以上、すなわち7s以上続いている箇所には色を付けていく。具体的には、相関値が[-1.0~-0.31]は青色、[-0.3~-0.3]は緑色、[0.3~1.0]は赤色としている。加えて、ビデオカメラの収録映像から、各実験協力者が発言・聴講中関係なく笑った箇所、会話が盛り上がっている箇所、話し手の意見に対して聞き手が同意している箇所(合意形成)を判断し、時系列のグラフに記入する。また、各実験協力者が発言した内容と聴講中に行っているうなずきなどの身体動作も、時系列のグラフに記入する。

本実験では、3つのテーマでの会話を収録しているが、3つ目のテーマでは、話し合いの折り合いがつかず、合意に至らなかった。そのため、1つ目・2つ目のテーマを対象に解析を行った。実験協力者A・実験協力者Bの2つのテーマでの解析結果を図8と図9に示す。

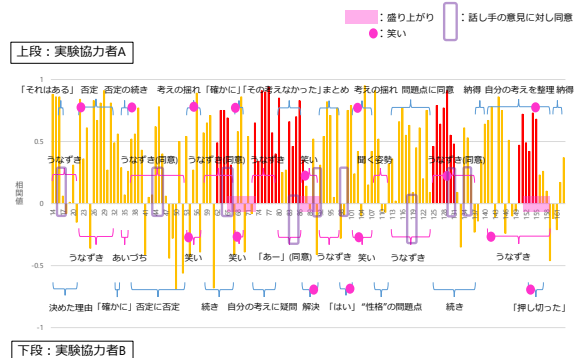


図8 テーマ1での解析結果

Fig.8 Analysis result with theme 1

図9 テーマ2での解析結果

Fig.9 Analysis result with theme 2

正の相関の連続区間である赤色で示した箇所では、笑いや会話が盛り上がっている箇所、話し手の意見に対して聞き手が同意している箇所が比較的によく見られ、また、視線がお互いに合うことや同時にうなずく動作も見られた。さらに、話し手の意見に対して聞き手が同意している箇所の周辺では、笑いや同時のうなずきや会話の盛り上がりが比較的よく見られた。

また、2つの会話を比較して、テーマ1よりテーマ2での会話の方が、正の相関の連続区間・笑い・話し手の意見に対し聞き手が同意している箇所・盛り上がりの箇所・発話の交替が多くみられた。また、テーマ2での各実験協力者が発言した内容において、相手の発言に対して否定する発言は行われなかった。

5. 考察

本研究では、対面コミュニケーションにおける合意形成推定の手法を提案することを目的とし、大学生2名に会話をさせ、会話中の身体動作の相関の強弱から、話し手の意見に対して聞き手が合意しているか否かを推定することができるかを検証している。結果として、正の相関の連続区間では、合意形成が行われていることが多いことや二者択一の質問の質の違いにより、聞き手が合意しているかの指標(同時にうなずいている箇所・会話の盛り上がりなど)を多くする必要性が示された。

まず、正の相関の連続区間では、笑いや会話が盛り上がっている箇所、聞き手が同意している箇所が多く見られ、また、視線がお互いに合うことや同時にうなずく動作も見られた。さらに、聞き手が同意している箇所の周辺では、笑いや同時にうなずくことや会話の盛り上がりが多く見られた。これらの解析結果より、正の相関の連続区間では、合意形成がされている可能性がある。また、お互いが同時にうなずくことや会話の盛り上がりも合意形成の指標になる可能性があることが示された。一方で、2つのテーマでの会話において、負の相関の連続した区間である[-1.0~-0.31]が見られなかった。

また、2つの会話を比較して、テーマ1よりテーマ2での会話の方が、正の相関の連続区間・笑い・話し手の意見に対し聞き手が同意している箇所・盛り上がりの箇所・発話の交替が多くみられた。また、テーマ2での各実験協力者が発言した内容において、相手の発言に対して否定する発言は行われなかった。これらは、二者択一の質問の質の違いによるものだと考えられる。テーマ1では、極端に二者択一の回答が分かれるが、テーマ2では、回答は分かれているものの、会話中に「どっちの回答でもよい」と述べられている。以上の解析結果より、二者択一の質問の質の違いにより、聞き手が合意しているかの指標(同時にうなずいている箇所・会話の盛り上がりなど)が多くする必要性がある。今後、さらに多くの会話を解析する必要がある。

6. おわりに

本研究では、対面コミュニケーションにおける合意形成推定の手法を提案することを目的とし、身体動作の周波数領域での相関の強弱から、聞き手が発話者の意見に対して合意しているか否かを推定することができるかを検証している。結果として、正の相関の連続区間では、合意形成が行われている可能性があることや二者択一の質問の質の違いにより、聞き手が合意しているかの指標(同時にうなづく箇所・会話の盛り上がりなど)を多く示す必要性が示された。今後、より詳細に検証するため、さらに多くの会話を解析する必要がある。

謝辭

本研究は科研費(26560016)の助成を受けたものである。

参考文献

- [1] 渡辺: 身体的コミュニケーションにおける引き込みと身体性 心が通う身体的コミュニケーションシステム E-COSMIC の開発を通して;ベビーサイエンス 2,4-12 (2003)
- [2] 黒川: ノンバーバルインタフェース, オーム社,(1994).
- [3] 斎賀,角,西田: 多人数会話におけるうなずきの会話制御としての機能分析;情報処理学会研究報告,Vol.2010-UBI-26 No.1 (2010)
- [4] 木村,加藤,萩沢,山本,三宅: 授業中の講師-受講者間のコミュニケーション計測システムの構築;ヒューマンインタフェース学会研究報告集,Vol.18 No.2 (2016)
- [5] 前田,堀内,市川: 自然対話におけるジェスチャーの相互関係の分析;情報処理学会研究報告,HI, ヒューマンインタフェース研究会報告,102 39-46 (2013)