

# ノンバーバル表現に注目したプレゼンテーション 支援システムの提案

趙新博<sup>†1</sup> 由井 蘭隆也<sup>†1</sup>

プレゼンテーション能力は授業、学会発表、ビジネスなど多くの場面で求められている。本研究は効果的なプレゼンテーションに必要なノンバーバル表現に注目した。そして、プレゼンテーション支援システムを提案する。その支援システムはプレゼンターの姿勢検出、言い淀み検出、イントネーション検出、ジェスチャー検出、ビデオ録画機能を提供する。

## The proposal of a presentation support system which took notice of nonverbal expression

Zhao Xinbo<sup>†1</sup> Yuizono Takaya<sup>†1</sup>

The ability of presentation can be applied in many areas, such as teaching, demonstration and commerce. This study discusses the necessary nonverbal expression required by an effective presentation. And a presentation support system is proposed. This system has presenter's posture detection, hesitation detection, intonation detection, gesture detection, and a video recording function.

### 1. はじめに

現在、プレゼンテーション能力は情報や知識を多くの相手に伝えるコミュニケーション手段として注目されている。例えば、日本政府が主導しているグローバル人材育成のプログラムにおいて、プレゼンテーションの能力は必要な項目とされている。またプレゼンテーション能力はコミュニケーション能力の一つとして、授業、学会発表、国際会議、ビジネスなど多くの場面で求められてきている[1][2][3][4]。

しかしながら「Yahoo! リサーチ」の調査によると、プレゼンテーションに対して苦手な人が多く、ジェスチャー、音声の抑揚、視線などのノンバーバル表現をうまく運用できていないことが知られている[5]。またプレゼンテーションにおいては、態度、熱意などの感性的な表現は聴衆に影響を及ぼすことが知られている [6][7]。それら表現要素はプレゼンテーションに対して重要とされる[8]。つまり、ノンバーバル表現や感情表現がプレゼンテーションには重要である。

一方、ノンバーバル表現には国による違いがあるとされている[9]。近年、国際会議におけるプレゼンテーションでは英語圏の習慣に従う傾向がある。そのため、ジェスチャーや音声の抑揚が控えめである東洋人（日本人や中国人）のプレゼンテーションは魅力的でないという指摘もある [1]。よって、ノンバーバル表現を身に付けることは、英語を用いた国際会議において魅力的なプレゼンテーションを行う

ために必要と考える。

その中、人間のノンバーバル表現を理解するために研究が進められてきている。例えば、アフェクティブコンピューティング (Affective Computing) [10] と呼ばれる計算機による感情処理技術が発展している。人間は相手と対話する場合、五感を通して表情、ジェスチャー、視線や体の動きなどの要素を相互作用として処理し、相手の感情を理解する。これらを計算機で処理・理解する方法が検討されている[11][12]。また、心理学では、怒り、恐れ、悲しみという感情が心拍、指先体温、脳波、呼吸、血圧などの生体情報と相関があるという結果が報告されている[13]。また、感情を込めた朗読は感情を込めない時と比べて、呼吸が深く、音量や音量変化が大きくなるという実験報告もある [14]。このように人間の感情表現と生体情報との関係に関する知見が蓄積されつつある。これらの情報はプレゼンテーションにおけるノンバーバル表現の測定に応用できる可能性がある。

本研究では効果的なプレゼンテーションのためのノンバーバル表現に注目する。そのために、ノンバーバル表現の理解を支援するシステムを提案する。そして、提案機能の一部実現したプロタイプとその試用評価について報告する。

### 2. 関連知識

プレゼンテーションの教育方法やその支援技術は、プレゼンテーション中のリアルタイムに使用するもの、その前に行うリハーサル支援、または、その後に行う振り返りに使うものに分けることができる。また、プレゼンテーション支援の対象要素はバーバル要素、パラ言語、ノンバー

<sup>†1</sup> 北陸先端科学技術大学院大学  
Japan Advanced Institute of Science and Technology

バル要素とに分けることができる(図1)。バーバル要素は言葉そのものである。パラ言語にはイントネーション、ポーズ(発話の一時中止)、リズムがある。ノンバーバル要素には顔認識、呼吸、心拍、ジェスチャーとがある。

その中、心理学的観点から、発表スライドの構成、聴衆とのインタラクションなどプレゼンテーションに役に立つ支援手段が多く述べられている[15]。また、情報技術を用いたプレゼンテーション支援システム「プレゼン先生」が開発されている[16]。その支援システムではWebカメラとマイクから得た情報をもとにプレゼンターの振る舞いを分析し、話速度、声の抑揚、聴衆とのアイコンタクトの度合いなどを支援対象としている。よって、その支援は図1におけるノンバーバル要素の顔認識とパラ言語に当てはまる。一方、顔の表情認識やジェスチャーなどの理解を用いた支援は今後の課題としている。

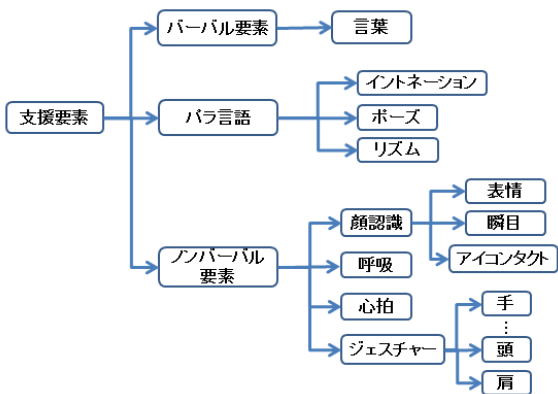


図1 プレゼンテーション支援の対象要素

### 3. プレゼンテーション支援システムの設計

プレゼンテーション支援システムとして、図1に示した対象要素において、ノンバーバル要素を中心に支援するシステムを検討した。現在、検討している支援システムの機能を表1に示す。それら機能は姿勢検出機能、言い淀み検出機能、イントネーション検出機能、ジェスチャー検出機能、そして、プレゼンテーションの評価入力機能である。

システム利用の様子を図2に示すとともに、システム処理の流れについて図3に示す。プレゼンテーション支援システムはカメラやマイクを通してプレゼンターの情報を入手する。その情報からプレゼンテーションの表現要素を切り出し、それを処理したアドバイスをプレゼンターにリアルタイムにフィードバックする。この処理によって、プレゼンテーション中の表現について注意を喚起する。また、システムはプレゼンテーションのリハーサルに使用できる。ショープレゼンテーションの全体を動画として記録することに加えて、プレゼンテーション全体を通しての癖などを検出してプレゼンターに表示する。これにより、振り返り

型の学習を支援する。

表1 提案システムの機能一覧

機能	説明	実装状況
姿勢検出機能	頭と肩の姿勢を検出する	○
言い淀み検出機能	「あの」、「その」など単語を検出する	○
イントネーション検出機能	発話のスピードや区切りなどの要素を検出する	×
ジェスチャー検出機能	情報伝達に意味があるジェスチャーを検出する	×
評価入力機能	自己評価と第三者評価によりプレゼンテーションを改善する	×

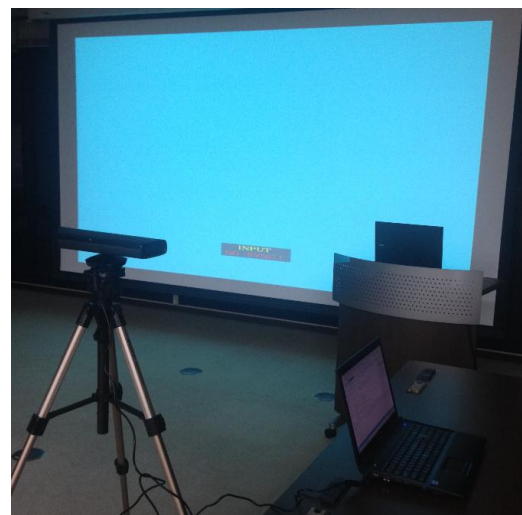


図2 システム利用像

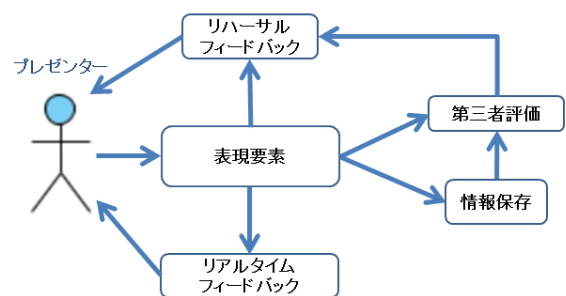


図3 システム処理の流れ

## 4. システムの開発状況

### 4.1 基本認識機能

ノンバーバル表現を中心としたプレゼンテーション支援システムを実現するためには画像処理を用いた認識処理が必要となる。今回、提案システムのの prototypes 開発には Kinect を用いた。Kinect SDK には人間の骨格を認識する機能が基本的に備わっている。Kinect は各種センサを使って、人間の全身を 20 個の参照点を抽出する機能を提供している (図 4)。この人間骨格認識を中心として、姿勢検出機能やジェスチャー機能を実現する。Kinect は X,Y,Z 軸の座標 (図 5) を用いて、図 4 の参照点を 3 次元データとして入手できる。

一方、マイクを通して得られる音声データをもとにした音声認識機能によって言い淀み検出機能を実現する。また、イントネーション検出機能は音声の波形処理によって実現する。

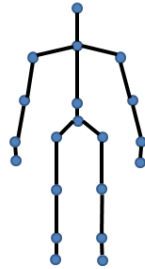


図 4 Kinect 骨格認識イメージ

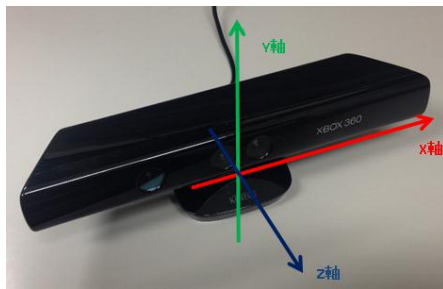


図 5 Kinect 座標軸

### 4.2 姿勢検出機能

プレゼンターは聴衆たちの様子を注意しながら発表する必要があると指摘されている[8]。聴衆とのアイコンタクトが必要であり、少なくとも顔を聴衆に向けることが期待される。

今回、対象とするプレゼンテーションはプレゼンターがレクチャー卓にノート PC を用いて行うものとしている。そのため、プレゼンターの上半身のみを取り扱う。そのため、図 6 に示す上半身部分をモデルとした姿勢の角度抽出を行う。

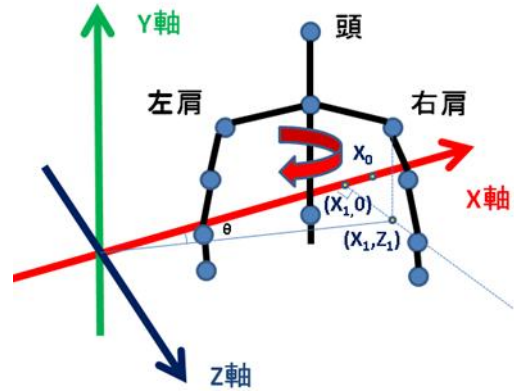


図 6 姿勢の角度検出

図 6 をもとに、プレゼンターの水平の向きを式(1)を用いて計算できる。また同様な式を用いて垂直方向の頭の向きを計算できる。これら角度をもとにプレゼンターの頭が前の方向を向いているかどうか、つまり、聴衆の方向を向いているかどうかの判定に使用できる。

$$\theta = \text{ATAN}(Z_1 / X_1) \quad (1)$$

現在実装中の姿勢認識機能の画面を図 7 に示す。下の 2 点は肩、上の点は頭を認識しており、それら 3 点の中央にある青い点は首の位置を認識しているが、通常は他の 3 点と同じように赤い点で表示される。しかし、図 7 では通常的位置より下がっており、下を向いているとシステムが判定した結果、青い点が大きく表示している。

この機能を通して次のような支援ができる。例えば、ずっと頭を下げて、パソコンを見続けているプレゼンターの場合、聴衆たちに頭を向けるようにシステムがプレゼンターにフィードバックする。同様に、肩の歪みもフィードバックできる。



図 7 うつむき状態を検出している様子

### 4.3 言い淀み検出機能

プレゼンテーション支援において、ノンバーバル表現だけでなくバーバル表現における個人の癖に注目した機能である。例えば、プレゼンターによっては「あの」、「なー」、

「その」、「まー」という表現を頻繁に使うものがある。これらは言い淀み表現と呼ばれ、プレゼンに淀みをもたらすため悪い印象をもたらすとされている。そこで、言い淀みを抽出して、カウントする機能を作成した(図 8)。現在は、「あの」、「えー」、「その」、「まー」が検出される場合にカウントしている。

現在、リハーサル終了後、プレゼンターは言い淀みの回数を知ることができる。今後、言い淀みの癖があるプレゼンターには、リアルタイムに癖のある言い淀みを診断し、自己認識を高める支援が考えられる。

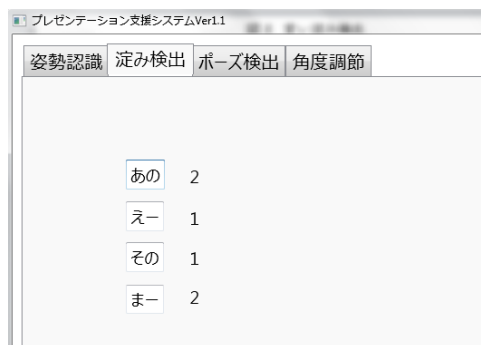


図 8 言い淀み検出

#### 4.4 ジェスチャー検出機能

ジェスチャーは他人に何かを伝えるために使われる身振り手振りの表現である。例えば、英語表現のボディランゲージとして、強調する単語において、頭を下に動かすことや片方の手の肘から下を上げて下げる、という表現がある。このように、プレゼンターは自分の観点をうまく伝えるためにジェスチャーを他の表現手段と一緒に使用できる。

本システムでプレゼンテーション向けに整理されたジェスチャー表現を支援する予定である。特に、4.2 で説明した姿勢認識機能を発展させて、手の位置を座標で検出し、その座標の動きをもとに手振りを検出予定である。

#### 4.5 イントネーション検出機能

プレゼンテーション要素としてパラ言語がある。例えば、イントネーションなど声に伴って様々な意味やニュアンスを付加できる[16][17]。また、発話中における不自然に感じさせる長い発話停止としてポーズ (Pause) 表現が知られている[17]。話し言葉では区切りがなければ、話を理解することは難しい。さらに、日本語と英語では、区切り長短の印象は違うことが指摘されている[14]。

以上より、プレゼンテーション要素として言語にイントネーション与えるポーズやトーンは重要である。よって、これらイントネーション支援機能を開発する予定である。

#### 4.6 評価機能

プレゼンテーションを評価するための項目が示されている。これら項目を用いた評価は自己評価や第三者評価で行われる。

提案システムでは、聴衆がリアルタイムにプレゼンテーションの評価する機能を検討している。この評価情報とプレゼンテーションのビデオを同期して記録することができる。これにより、プレゼンテーション中に悪いと評価された時点のビデオ情報を後からみることによって、何が問題であるか検討することも支援可能とする予定である。

### 5. プロトタイプの試用評価

開発中のプロトタイプの支援機能を評価するため、JAIST 知識科学研究科修士 2 年の学生 3 人 (中国人留学生 A と B, 日本人学生 C) に対してシステムの試用実験を行った。その発表時間は 15 分である。マルチメディアセンサーとして Kinect を使用した。そのセンサーとプレゼンターとの距離は 0.8~4m であり、高さは 0.6~1.8m の間に置いた。また Cisco 社の Webex Recorder を利用し、提案システムのウィンドウ画面を記録した。図 9 にシステムの利用画面を示す。

実験参加者はプレゼンテーションを終えた後、記録したビデオ画面をみる。そして、図 10 に示す個人評価シートに記入を行う。個人評価シートの評価項目は頭 (下がるか)、視線 (アイコンタクト)、肩 (歪みがあるか) などの 10 項目である。とても悪いからとても良いの範囲で 5 段階評価を行った。また、気づいた点を自由記述させた。



図 9 試用実験中における姿勢検出の様子

評価項目	表現
頭 (下がるか)	(とても悪い) 1 -- 2 -- 3 -- 4 -- 5 (とても良い)
視線 (アイコンタクト)	(とても悪い) 1 -- 2 -- 3 -- 4 -- 5 (とても良い)
肩 (歪みがあるか)	(とても悪い) 1 -- 2 -- 3 -- 4 -- 5 (とても良い)
手振り	(とても悪い) 1 -- 2 -- 3 -- 4 -- 5 (とても良い)
ポーズの頻度	(とても悪い) 1 -- 2 -- 3 -- 4 -- 5 (とても良い)
音量	(とても悪い) 1 -- 2 -- 3 -- 4 -- 5 (とても良い)
音声の抑揚	(とても悪い) 1 -- 2 -- 3 -- 4 -- 5 (とても良い)
発話のスピード	(とても悪い) 1 -- 2 -- 3 -- 4 -- 5 (とても良い)
言い淀みの頻度	(とても悪い) 1 -- 2 -- 3 -- 4 -- 5 (とても良い)
態度	(とても悪い) 1 -- 2 -- 3 -- 4 -- 5 (とても良い)
気づいた点を記述する	

図 10 個人評価シート

### 6. 評価結果と考察

#### 6.1 評価結果

参加者は記録されたシステム画面をプレゼンテーション

終了後みた。そして、その記録を見るまでは自分の姿勢や、言い淀みに関して気づいてなかったという感想を述べていた。

表2に言い淀みの検出結果を示す。留学生AとBは「あの」の検出頻度が、日本人学生に比べると多いことがわかる。また、留学生Bと日本人学生は「えー」という癖があることがわかる。一方「まー」の認識は実験者の認識と比べて多い回数となった。これは誤認識であるかどうか、今後、記録データと比較してみる必要がある。

表2 言い淀みの検出結果

検出内容	留学生 A	留学生 B	日本人 C	平均値
あの	6	8	1	5.0
えー	2	6	5	5.0
その	1	1	0	0.7
まー	5	4	3	4.0

実験参加者はプレゼンテーションを記録したビデオを見て、図11のように気づいた点を記述した。三人のデータをまとめた結果を表3に示す。

プレゼンテーション個人評価シート

項目	評価	コメント
項目1	頭	(とても悪い) 1-①-3-4-5 (とても良い)
項目2	視線	(とても悪い) ①-2-3-4-5 (とても良い)
項目3	肩	(とても悪い) ①-2-3-4-5 (とても良い)
項目4	手	(とても悪い) 1-②-3-4-5 (とても良い)
項目5	ポーズ	(とても悪い) 1-2-③-4-5 (とても良い)
項目6	音量	(とても悪い) 1-2-④-4-5 (とても良い)
項目7	抑揚	(とても悪い) 1-2-3-⑤-5 (とても良い)
項目8	スピード	(とても悪い) 1-⑥-3-4-5 (とても良い)
項目9	言い淀み	(とても悪い) 1-⑦-3-4-5 (とても良い)
項目10	態度	(とても悪い) ①-2-3-4-5 (とても良い)
気づいた点を記述する	体がフラフラ 慣れて、態度が悪い 視線を相手に送る回数と時間が少ない フラフラ見るくらいなら見ないほうがいい(ふ、めい)	

図11 個人評価シートの記入例

表3 検証テスト参加者の得点状況

評価項目	留学生 A	留学生 B	日本人 C	平均値
頭	4	5	2	3.7
視線	4	4	1	3.0
肩	5	4	1	3.3
手	5	3	2	3.3
ポーズ	2	3	3	2.7
音量	2	5	3	3.3
抑揚	4	3	4	3.7
スピード	4	2	2	2.7
言い淀み	2	1	2	1.7
態度	5	5	1	3.7

## 6.2 各機能に関する検討

試用実験を通して気づいた点を下記に記す。

### ・姿勢検出機能について

実験から得られた録画データより、本システムはプレゼンターの姿勢情報を概ね正しく検出できたことがわかった。ただし、プレゼンターの動きが Kinect の検出境界や検出範囲を超えた場合、一時的に認識できない場合が見られた。今後、2, 3 台のセンサーを同時に利用することによって検出範囲を広げるという方法も考えられる。また、頭部のみ動かす場合は、検出できないということがわかった。これについては、画像処理を用いた角度推定などを検討する必要がある。

### ・言い淀み検出機能について

音声検出機能そのものはプレゼンターが連続に早く話す場合、認識できない場合があったが、言い淀み検出の場合は、区切りなどが発生することが多いため、誤認識することは少ないことが観察された。

### ・評価機能について

試用実験で実施したプレゼンターによる自己評価は紙シートであった。今後、携帯やタブレット端末を利用できるネットワークアプリケーション的な評価機能を実装する予定である。

## 6.3 今後の展望

現在のプロトタイプシステムでは、姿勢検出機能と言い淀み検出機能を実現した。

今後、グローバルコミュニケーションの視点から留学生と日本人における第二外国語を用いたプレゼンテーションに対する支援も検討して行く予定である。また、アフェクティブ・コンピューティングの成果を取り入れ[18]、感情を込めたプレゼンテーション支援も行っていきたい。

そのためには、ジェスチャー検出機能やイントネーション検出機能の実装を進める予定である。

## 7. おわりに

効果的なプレゼンテーションは自分の考えを多くの人びとに伝達するための重要な技能である。そこで、ノンバーバル表現に注目したプレゼンテーション支援システムを提案した。そのシステムでは、プレゼンターの姿勢検出(上半身)機能や言い淀み検出機能を実現した。そして、実験参加者3人によるシステムの試用実験を実施できた。

今後は、ジェスチャー検出機能、イントネーション検出機能、評価入力機能の実装を進める予定である。

**謝辞** プレゼンテーション練習にご協力頂いた北陸先端科学技術大学院大学知識科学研究科 M2 の皆様に、謹んで感謝の意を表す。

## 参考文献

- 1) ロバード・R・H・アンホルト: 理系のための口頭発表術, 株式会社講談社(2013).
- 2) 八幡ひろし: 自分の考えをしっかりと伝える技術, PHP 研究所 (2008).

- 3) Giving Presentations: Expert Solutions to Everyday Challenges, Harvard Business Review Press (2007).
- 4) Edward, Z.: Presentation Skills for Scientists, Cambridge University Press (2010).
- 5) 即興！恐怖の3分プレゼン, 日経BPネット(2007).
- 6) 望月理香, 永徳真一郎, 茂木学[他]: ライフログを活用した感性コミュニケーション実現方式の提案-実感をともなう情報伝達効果に関する検証, 情報処理学会論文誌 53(1), pp.30-38 (2012).
- 7) 市野順子, 西野哲生, 八木佳子: Ambient facilitator:姿勢から推定される参加態度情報を用いて発言の促進や抑制を図るテーブルトップ環境の提案, GN ワークショップ 論文(2013).
- 8) 末武国弘: より良いプレゼンテーションの仕方, 電子情報通信学会誌 83(7), pp.581-588 (2000).
- 9) Adrian, W.: 実践で役立つ!! 英語プレゼンテーション, 丸善出版株式会社(2013).
- 10) Rosalind, W. P.: AFFECTIVE COMPUTING, MIT press (2000).
- 11) Andrea, K. and Nadia, B.B.: Affective Body Expression Perception and Recognition: A Survey, IEEE TRANSACTIONS ON AFFECTIVE COMPUTING, VOL. 4, NO.1, pp.15-33 (2013).
- 12) Sander, K. et al.: DEAP: A Database for Emotion Analysis Using Physiological Signals, IEEE TRANSACTIONS ON AFFECTIVE COMPUTING, VOL3, NO.1, pp.18-31 (2012).
- 13) Susan, N.H. et al.: ヒルガードの心理学 (第15版), 金剛出版 (2012).
- 14) 中村敏枝: コミュニケーションにおける「間」の感性情報心理学, 音声研究 13(1), pp.40-52 (2009).
- 15) 海保博之: わかりやすさとコミュニケーションの心理学, 朝倉書店(2010).
- 16) 栗原一貴, 後藤真孝, 緒方淳 (他): プレゼン先生:画像情報処理と音声情報処理を用いたプレゼンテーションのトレーニングシステム, WISS 第14回インタラクティブシステムとソフトウェアに関するワークショップ論文集, pp.59-64 (2006).
- 17) 近藤富英: ノンバーバル・コミュニケーション行動としてのポーズの機能と役割への一考察, 人文科学論集. 文化コミュニケーション学科編 40, pp.129-136 (2006).
- 18) Michelle, K. et al.: Body Movements for Affective Expression: A Survey of Automatic Recognition and Generation, Affective Computing, IEEE Transactions on, Volume: PP, Issue: 99, pp.1-20 (2013).