

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開発

趙 新博^{†1} 由井 蘭隆也^{†1} 宗森 純^{†2}

プレゼンテーション能力は社会スキルとして重要であり、その表現はシナリオ表現、スライド表現、身体表現、音声表現、質疑表現にまとめられる。その中、実演に必要な身体表現や音声表現の理解を支援するプレゼンテーション支援システムの開発を行った。そのシステムは身体表現の検出・記録機能、音声表現の検出・記録機能、スライド操作の記録機能とから成り立つ。本報告では、その開発について述べるとともに、開発システムの評価計画や試用について考察する。

Development of A Presentation Support System for Nonverbal Expression Skills

XINBO ZHAO^{†1} TAKAYA YUIZONO^{†1} JUN MUNEMORI^{†2}

Presentation skill becomes important as social skill, and its representation consists of scenario, slide, body, voice, and QA. We have developed a presentation support system to support understanding of body expression and voice expression in performance of presentation. The system has the function of body expression, the function of voice expression, the record function of slide control, and the real-time feedback function. And then, we described an evaluation plan of the system and discussed its preliminary usage.

1. はじめに

近年、インターネット上に情報公開するなど、コンピュータを用いた情報伝達は日常行為となっている。また、持ち運びができるノート PC を使用したプレゼンテーションは一般的であり、そのプレゼンテーションは知識や情報を伝達するための社会スキルとして重要となっている[1]。また 21 世紀の企業組織に必要な基礎スキルとして「クリティカルシンキングと問題解決」、「効果的なコミュニケーション」、「共同作業とチーム構築」、「創造性とイノベーション」の 4 つが取り上げられ、その中、効果的なコミュニケーションの支持率は 75.6% と高い[2]。つまり、知識や情報伝達を効果的にプレゼンテーションする技術は現代社会における基礎スキルとして重要である。

その中、プレゼンテーションの技術や経験を教えるための教科書[3]-[10]や数多くの取り組み[11]-[14]が知られている。これらはプレゼンテーションにおける準備から実演に必要な知識を網羅的に紹介している。その知識は経験的な知識が多く、プレゼンテーションを評価する技術は評価シートとしてまとめられている。その評価項目は、プレゼンテーション技術に必要な表現手段としてまとめると、シナリオ表現、スライド表現、身体表現、音声表現、質疑表現の 5 種類に分類できる。その中、シナリオ表現、スライド表現は事前準備が可能な表現であるが、身体表現、音声表

現、質疑表現はプレゼンテーションの実演に使用する表現である。よって、後者の三表現は、発表者が実演技能として身につける必要がある。

一方、マルチメディア処理技術が進み、人間の身体表現や感情表現に注目した研究が数多く行われている。例えば、アフェクティブ・コンピューティングと呼ばれる感情を計算機で処理する研究分野において人間の身体情報を処理するための様々な取り組みが行われている[15]。その技術は、対人コミュニケーションに必要なソーシャル技術の習得支援へと展開されている[16]。

我々は、マルチメディア処理技術を用いて、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てきた[17]。今回、プレゼンテーション支援システムにおいて身体表現と音声表現の理解を支援するシステムの開発について報告する。2.ではプレゼンテーション技術について述べ、3.では、システムの開発と実現機能について述べる。4.では開発システムの評価方針について述べ、5.ではシステムに対する試用をもとに考察する。

2. プレゼンテーション技術について

2.1 プレゼンテーション表現とプロセス

プレゼンテーション技術を理解するためにはプレゼンテーションに必要な表現とその評価方法を検討する必要がある。そこで、文献[3],[4]にあるプレゼンテーションのチェックシートと他の文献[5]-[14]の記載項目をもとに、プレゼンテーション表現を 5 つにまとめるとともに、それぞれの表現に応じた評価項目をまとめた。その結果を表 1 に示す。

^{†1} 北陸先端科学技術大学院大学
Japan Advanced Institute of Science and Technology

^{†2} 和歌山大学
Wakayama University

また図1に、プレゼンテーションのプロセスを示す。

最初の2つの表現はプレゼンテーションの準備段階に用意できる表現である。シナリオ表現はプレゼンテーション全体のストーリーや構成に関わる表現である。スライド表現はスライド1枚1枚の見やすさ等に関わる表現である。

次の3つの表現はプレゼンテーションの実演で使用される。身体表現は体の向きとアイコンタクト、ジェスチャ、顔の表情に関する表現である。音声表現は声の大きさ・強弱、話すスピード、聞き取りやすさに関する表現である。質疑表現はプレゼンテーション終了後に行われる質疑応答に必要な表現である。

これら5つの表現は実演後の反省を通して改善されることが期待される。

表1 プレゼンテーション表現と評価項目

シナリオ表現	アウトライン
	ストーリー構成
	発表目的
	発表時間
	聴衆考慮
スライド表現	タイトルやスライドタイトル
	簡潔明瞭な文章
	文字や図表のデザイン
身体表現	体の向きとアイコンタクト
	ジェスチャ
	顔の表情
音声表現	声の大きさ・強弱
	話すスピード
	聞き取りやすさ
質疑表現	質問への回答
	質問への準備

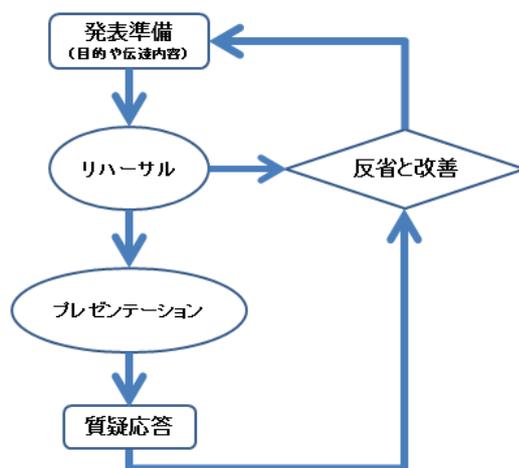


図1 プレゼンテーションのプロセス

2.2 関連システム

情報技術を用いたプレゼンテーション支援システムとして栗原による「プレゼン先生」がある[18]。その支援システムでは発表者の会話速度、声の抑揚、聴衆とのアイコンタクトの度合いを支援することを検討している。この研究はスライド表現と音声表現に関するものであるが、その支援機能の効果までは検証されていない。

プレゼンテーションのスライドデータを知識資源として再利用するために、スライドデータのレイアウトや視覚情報を構造抽出する研究がある[19]。この支援はスライド表現に関するものであり、本研究が対象とする実演技能の支援とは異なる。

一方、プレゼンテーションに関する聴衆のチャットからの重要発言を選択するシステムが開発されているのである[20]。これは質疑表現に関係ある支援である。

本研究で開発したシステムは身体表現と音声表現を記録し、その記録データをもとに発表者の実演スキルの理解・評価の支援を目標としている。

3. プレゼンテーション支援システムの開発

3.1 システムの開発概要

プレゼンテーションの実演において必要な身体表現と音声表現の理解を支援するためのシステム開発を行った。システム構成を図2に示す。システムは、発表者の実演データを取得するセンサ部分、その実演データから実演表現を検出・記録する部分、そして、フィードバック部分に分けられる。身体表現においてはイメージセンサを通して、音声表現においては音センサを通して処理が行われる。また、発表者へのフィードバックにはイヤフォンを通じた音声通知、画面を通じたイメージ通知の2通りある。

開発システムの利用風景を図3に示す。発表者は、ソフトウェアとしてPowerPoint (PPT) を用い、レクチャー卓にノートPCを置いて実演する。イメージセンサは図4に示すKinectを用いており、発表者の前0.8~4mで、高さは0.6~1.8mの範囲に置いて使用する必要がある。音セン

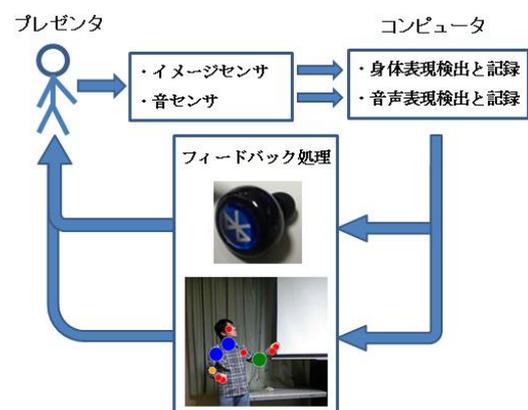


図2 システム構成



図3 システムの利用風景

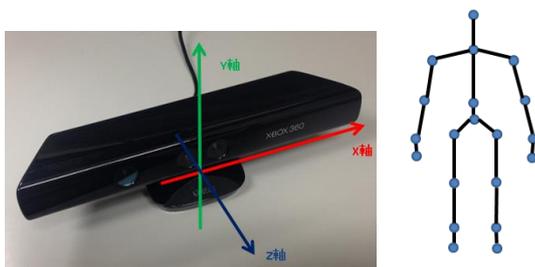


図4 Kinect その座標系および骨格モデル

サは本システムを動作させる PC の音声入力につながるマイクを使用する。

実現したシステム内容とプレゼンテーション表現との関係を表2に、開発したシステム画面を図5に示す。実現システムは身体表現の検出・記録機能、音声表現の検出・記録機能、スライド操作の記録機能とから成り立つ。システム画面における、左側のウィンドウは身体表現と音声表現についてセンサデータを介して取得された情報が表示されている。また右側のウィンドウはスライド操作を行うためのインターフェースを提供している。

表2 プレゼンテーション表現とシステムの実現内容

	評価項目	説明	実現内容
身体表現	体の向きとアイコンタクト	聴衆に向けて、アイコンタクトをする	○
	ジェスチャ	意味があるジェスチャ、姿勢をする	△
	顔の表情	笑顔など表情の表出	×
音声表現	声の大きさ・強弱	声の大きさと抑揚	○
	話すスピード	発話スピードの速さ	△
	聞き取り易さ	明瞭で、短く発話	△



図5 システム画面

身体表現の検出・記録機能では、イメージセンサ Kinect によって骨格モデルを抽出し、体の向きとアイコンタクト、ジェスチャに相当するデータ処理を実現している。特に、体の向きとアイコンタクトにおいては、図5の左側に示すようなフィードバック機能を備える。図5では認識された骨格モデルのノード部分を丸で表示すると共に、前を向かせる文字メッセージを表示している。

音声表現の検出・記録機能は音声センサから得られたデータをもとに、声の大きさ・強弱、聞き取り易さに関わるデータ処理を行う。

スライドの操作記録機能は発表者が使用しているスライドを記録するために作成したインターフェースである。

3.2 身体表現の検出・記録機能

・体の向きとアイコンタクト

発表者は聴衆に対する興味を示すために聴衆に体を向けるとともに、アイコンタクトを利用することが期待される。また体の姿勢はまっすぐ立つとよいとされる。そこで体の向き、頭の傾き、肩の傾きを検出する機能を実現した。

姿勢認識の様子を図6左側に示す。頭1点、肩2点、首1点を認識していることがわかる。特に首の点は、図6右側に示す首の角度が10度以上傾いているため、前向いていないと判断し、青い点で大きく表示されている。また、左肩は右肩より下がっていると判断し、左肩の点も青い点で大きく表示されている。

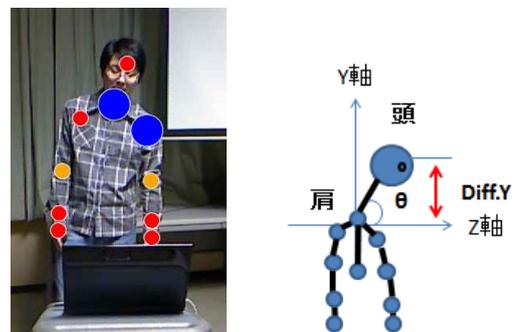


図6 肩の傾きと首の向きについて

この機能で得られた情報を用いて次のようなフィードバ

ック表示を実現している。首の傾きが 10 度以上の場合、パソコンを見続けている可能性が高いため、図 5 で示すように「前を見てください」という文字をシステム画面に表示している。また、肩の歪みを検出した場合、「肩を斜めにしないで下さい」という文字をシステム画面に表示する。なお発表時にはシステム画面を見ていない場合もあるため、音声によるフィードバックとしてピープ音を出す機能も備えている。

・ジェスチャ

ジェスチャ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。開発システムは発表者の腕の形を図 7 に示す 4 パターンに分けて、データ処理し、記録する。このデータは右腕と左腕、それぞれごとに検出して記録される。右腕左腕それぞれに 4 パターン識別するため、全部で 16 パターンのジェスチャを区別できる。例えば、手をよく動かしているかどうかの検出に用いることが期待できる。

実際の処理では、発表者の骨格モデルから腕ごとに肩、肘、手首の 3 点を抽出し、次の 4 パターンに分けて抽出している。

パターン 1：肩の縦軸座標は肘、腕首より高い、かつ肘の縦軸座標は腕首より高い。

パターン 2：肩の縦軸座標は肘、腕首より高い、かつ腕首の縦軸座標は肘より高い。

パターン 3：腕首の縦軸座標は肘、肩より高い、かつ肩の縦軸座標は肘より高い。

パターン 4：腕首の縦軸座標は肘、肩より高い、かつ肘の縦軸座標は腕首より高い。

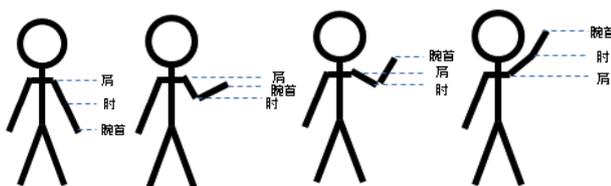


図 7 腕の形 (左腕の 4 パターン)

3.3 音声表現の検出・記録機能

・音声記録機能

プレゼンテーション中の声は大きいことが期待されるとともに、棒読みでない、抑揚をつけるとよいとされる。そこで、声の大きさ・抑揚を記録するための機能として、音声の大きさを時間とともに記録する機能を実現した。図 8(a)に示す音声波形を図 8(b)に示すように、日時と音声の大きさをデータセットとして記録できる。

音声の大きさの時間推移がわかるため、ある時間間隔における音声の大きさを平均値として求めることや、音声の時間変化を求めることができる。



(a)音声の波形表示

2014/11/05	16:04:27:	0.468352472488544
2014/11/05	16:04:27:	0.436751847309174
2014/11/05	16:04:27:	0.335963554357424
2014/11/05	16:04:27:	0.322722453278817
2014/11/05	16:04:28:	0.376293993889937
2014/11/05	16:04:28:	0.727589008635668
2014/11/05	16:04:28:	0.811288755702753
2014/11/05	16:04:28:	0.682524832773542
2014/11/05	16:04:28:	0.739566544599161
2014/11/05	16:04:28:	0.686345837265133
2014/11/05	16:04:28:	0.681609309151823
2014/11/05	16:04:28:	0.574429885854547
2014/11/05	16:04:29:	0.61706048225737
2014/11/05	16:04:29:	0.596928703279972
2014/11/05	16:04:29:	0.558586575688311
2014/11/05	16:04:29:	0.566828437098848

(b)音量の記録データ

図 8 音声の記録データ

・言い淀み言葉の検出機能

プレゼンテーション中に、「あの」、「え〜」、「その」、「まー」といった言葉が頻繁に出ると、プレゼンテーションに淀みを感じると言われる。そこで、これらを言い淀み単語と呼び、音声認識 API を用いて検出し、数える機能を実現した。例えば、リハーサル終了後、発表者は言い淀み単語の使用回数を知ることができる。

3.4 スライド操作の記録機能

プレゼンテーションの実演時に使用されているスライド番号を記録するために開発した機能である。システムと関連づけた PPT スライドに対して、前に戻る、次に進める、最初に移動する、最後に移動するという操作を行える。そして、スライドの移動操作を行うたびに、スライド番号と操作時間がログデータとして記録される。

操作インターフェースは、図 5 の右側に示すとおりであり、主にマウスで操作する。ただし、発表者の実演中の操作を簡単にするために、前に戻る、次に進める、という操作はキーボードの矢印キーを用いて操作できる。

その他の機能として、スライドのマーク機能、スライド番号による移動機能、時間表示機能をもつ。

4. システムの評価実験計画

プレゼンテーション支援システムの効果を評価するための実験計画について説明する。評価実験はシステム利用実験とビデオ利用実験の 2 通りである。この 2 つの実験を比較することによって、提案システムが従来のビデオ利用とどのように異なるか定量的に調査する予定である。

システム利用実験の手続きについて述べる。実験参加者は 1 回目のプレゼンテーションを行う。そのプレゼンテーション中のシステム画面はフィードバック情報として記録する。この記録には Cisco 社の WebEx Recorder を利用する予定である。実験参加者は 1 回目のプレゼンテーション終

了後、自身の発表に対して表1における身体表現と音声表現における評価項目について評価する。この評価を個人評価と呼ぶ。その後、フィードバック情報として記録されたシステムのビデオ画像を見て、反省する。これを個人反省と呼ぶ。再び、個人評価を行う。次に実験参加者は2回目のプレゼンテーションを行う。ただし、この際、リアルタイムフィードバックとして音声によるフィードバックを利用する。このプレゼンテーション終了後は1回目同様、個人評価、個人反省、個人評価の処理を行う。

対照実験として行うビデオ利用実験では、システム利用実験におけるフィードバック情報がビデオ記録データに入れ替えることと、2回目発表において音声によるフィードバックがない点のみ異なる。

この2つの実験を比較することによって、提案システムが従来のビデオ利用とどのように異なるかを定量的に調査する予定である。また1回目と2回目のプレゼンテーションでどのような違いが生じるかを考察すると共に、システムによる定量評価と個人評価との関係を検討する予定である。

5. システム試行と考察

開発したシステムの動作テストおよび収集データを検討するために大学院生1名Aに本システムを3回使用する試行調査を行った。なおシステム画面を記録するためにはCisco社のWebEx Recorderを利用した。

その試行結果を表3に示す。パターン変化回数は図7で説明した腕のパターンが変化した回数である。例えば、ある腕がパターンnである際、その腕が別のパターンになった場合、1回数え上げる。1回目と比べて、2回目、3回目は腕の動きのパターン変化回数が増えており、腕を活発に使ったことがうかがえる。言い淀み言葉は2回目において、「あの」という言葉が最多で2回使用されたのみであり、Aには言い淀みにつながる言葉の癖は少ない結果となった。

表3 試行結果

発表回数	1回目	2回目	3回目
発表時間	5分38秒	6分55秒	5分56秒
パターン変化回数	1	9	6
言い淀み言葉数	1	3	2

表4には、発表スライドの1枚目から3枚目を対象として、発表スライドごとに発表時間、体の向き、頭の傾き、肩の向き、音量を調べた結果を示す。発表時間はスライド2にかけていることがわかる。体の向きは、スライド1の場合、1回目と比べて3回目が前に向いている結果となっている。一方、スライド3の場合、体の向きはそれほど変化していない。これはAがノートPCを見て説明しているためと考えられる。また頭の傾きは20度から40度であり、ほとんど下を向いていたと考えられる。今後、プレゼンテ

ーションの記録データと同時に評価データを収集することによって、記録データの価値解釈を洗練させていく必要がある。

表4 スライドごとのデータ分析

		スライド1	スライド2	スライド3
第1回目	発表時間	11秒	34秒	10秒
	体の向き	右22.75度	右21.08度	右23.80度
	頭の傾き	23.78度	27.71度	19.26度
	肩の傾き	右1.00度	左1.69度	左0.70度
	音声の大きさ	0.54	0.59	0.53
第2回目	発表時間	16秒	42秒	15秒
	体の向き	右15.82度	右18.98度	右16.58度
	頭の傾き	26.25度	20.64度	20.40度
	肩の傾き	右1.09度	右0.45度	左0.47度
	音声の大きさ	0.47	0.52	0.51
第3回目	発表時間	9秒	35秒	15秒
	体の向き	右9.80度	右12.20度	右18.47度
	頭の傾き	29.62度	41.30度	26.36度
	肩の傾き	左1.20度	左4.00度	左2.80度
	音声の大きさ	0.56	0.58	0.56

6. おわりに

情報や知識を伝達するためのプレゼンテーション技術において、身体表現と音声表現の理解を支援するプレゼンテーション支援システムを開発した。身体表現の記録機能として、首や肩の角度検出、腕の4パターンを記録する機能を実現した。また、音声表現の記録機能として、音声の大きさを記録するとともに、言い淀み言葉を検出する機能を実現した。その実現したシステムを用いてスライド単位ごとに身体表現や音声表現のデータ解析をある程度、可能とした。

今後は、プレゼンテーション理解の評価実験を行う予定である。

参考文献

- 1) Knoblauch, H.: PowerPoint, communication, and the knowledge society, Cambridge press (2013).
- 2) American Management Association: AMA 2012 Critical Skills Survey: Executive Summary, <http://www.amanet.org/training/promotions/AMA-2012-Critical-Skills-Survey.aspx> (Access on 2015.2.15).
- 3) Wallwork, A.: 上出鴻子, 上出洋介 (訳): 実践で役立つ!! 英語プレゼンテクニック, 丸善出版株式会社(2013).

- 4) ロバート・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).
- 5) Steve Mandel： Effective Presentation Skills ， Crisp Publications(2000).
- 6) ジェレミードノバン著，中西真雄美（訳）：TED トーク世界最高のプレゼン術，株式会社新潮社(2014).
- 7) ジーン・セラズニー：数江良一，菅野誠二，大崎朋子（訳）：マッキンゼー流プレゼンテーションの技術，東洋経済新報社(2004).
- 8) 井庭崇：プレゼンテーション・パターン創造を誘発する表現ヒント，慶応義塾大学出版株式会社(2013).
- 9) 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).
- 10) 脇山真治：プレゼンテーションの教科書，日経 BP 社(2009).
- 11) 牧野由香里：プレゼンテーションにおける自律的学習のための学習環境デザイン，日本教育工学会論文誌，27(3)，pp.325-335(2003).
- 12) Ono, Y.A., Morimura, K.: Effective Methods for Teaching Technical Presentations in English to Japanese Engineering Students, IPCC, pp.1-6(2008).
- 13) 山下祐一郎，中島平：ビデオ映像とレスポンスアナライザを利用したプレゼンテーション能力の育成，日本教育工学会論文誌，pp.401-410(2010).
- 14) Hashimoto, K., Takeuchi, K.: Multimedia Learner Corpus of Foreigner's Basic Presentation in English with Evaluations, ICEIT, pp.469-473(2010).
- 15) Rosalind, W. P.: AFFECTIVE COMPUTING, MIT press (2000).
- 16) Hoque, M.E.: Picard, R.W.: Rich Nonverbal Sensing Technology for Automated Social Skills Training, IEEE Computer, Vol.47, Issue 4, pp.28-35(2014).
- 17) 趙新博，由井蘭隆也：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，情報処理学会研究報告．GN，pp.1-6（2014）.
- 18) 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).
- 19) 小林智也，西本一志：Chatplexer: チャットを併用する口頭発表における発表者のための重要発言選択支援の試み，情報処理学会論文誌，Vol.53, No.1, pp.12-21, 2012.
- 20) 羽山徹彩，難波英嗣，國藤進：プレゼンテーションスライド情報の構造抽出，電子情報通信学会論文誌，pp.1483-1494(2009).