

即時フィードバックを利用した VR プレゼンテーション練習システム

脇坂真広^{†1} 兼松祥央^{†2} 安原広和^{†2} 松吉俊^{†2} 三上浩司^{†2}

概要: プレゼンテーションを行う際の緊張感を想定した練習を行うための環境作りは容易ではない。そこで、本研究では仮想現実技術を用いて制作したプレゼンテーション環境に、動的な即時フィードバックシステムを実装した。仮想空間上に存在する聴衆の 3D モデルが利用者のプレゼンテーションに応じて、反応や態度の変化を行うことで、より臨場感のあるプレゼンテーション練習環境を構築した。

1. はじめに

社会のさまざまな場面でプレゼンテーションを行う機会は増えており、プレゼンテーション能力は社会を生きるために必要な能力となってきた。厚生労働省によるとプレゼンテーション能力は若者就職基礎力であり[1]、文部科学省によると当該の能力は学士力である[2]。また、コミュニケーションの得意度の調査[3]によると、いくつかあるコミュニケーション形式の中でも「複数の人の前で、発表すること」という項目が苦手な理由の多くの割合を占めている。

プレゼンテーションを上手く行うためには、練習を経て経験を積むことが必須である。一般的にプレゼンテーションを1人で練習する際、本番のプレゼンテーションとは異なる状況で行う場合が多い。そのため、こうして行った練習では実際のプレゼンテーションが持つ臨場感や緊張を感じることができない。このように、プレゼンテーション経験を積むことのできる本番に類似した練習環境を用意するのは難しいという問題がある。また、1人で練習を行ったとしてもそのプレゼンテーションの客観的な評価を得ることはできないため、練習をすることで発表内容の大きな改善に繋げることも難しい。

そこで、本研究の目的は仮想現実（以下「VR」）空間を用いて、より実践に近い擬似的なプレゼンテーション環境を制作することである。聴衆に見立てた 3D モデルにプレゼンテーションに応じたフィードバックを組み込むことで、本番さながらの練習を行うことのできる環境を提供する。

2. 関連研究

周ら[4]は Head Mounted Display（以下「HMD」）を利用してプレゼンテーション練習ができるシステムを提案した。このシステムでは 360 度映像を設置した VR 空間を用いてプレゼンテーションの会場を擬似的に再現している。また、視線認識を用いてプレゼンテーション中のユーザーのアイ

コンタクトを評価している。さらに、この研究ではビデオキャプチャーボードを利用して利用者が持ち込んだプレゼン資料を VR 空間内に映し出すことで、臨場感を向上させた。

樺山ら[5]も VR 空間でのプレゼンテーションを行うシステムを提案した。このシステムでは聴衆の視線を検出、再現することで、他の聴衆の存在を感じ辛いことを解消している。

栗原ら[6]はマイクと Web カメラを用いたプレゼンテーション練習システム『プレゼン先生』を開発した。このシステムではプレゼンテーション中のユーザーの良くない行動に対してリアルタイムに通知を行う。このシステムはリアルタイムの通知を行うため、顔の向きによってアイコンタクトを判定し、マイクから利用者の音声情報を取得して判定を行なっている。

高所や雷などの特別な環境に対して恐怖を抱く人に向けて暴露療法という医療手段が有効に利用されてきた。宮野ら[7]は VR を利用した暴露療法を用いて、高所恐怖症の人物に対して不安障害を緩和させる研究を行なった。この研究では 3DCG と VR を用いたコンテンツが高所恐怖症の緩和に一定の有効性があったことを示している。

周らの研究ではプレゼンテーション空間を作成するために 360 度映像を用いていた。360 度映像を用いたシステムでは常に同じ映像を用いていることや、相互的なリアクションがないことからリアルなプレゼンテーションと大きく環境が異なる。樺山らの研究ではオンラインで他の聴衆が存在する状態でのプレゼンテーションを前提としており、利用者1人での練習に最適化されているわけではない。

栗原らの研究ではリアルタイムなフィードバックを受けながら練習することで一定の向上効果が期待できる。しかし、アイコンタクトは視線を用いた表現であるため、頭の角度を視線として検出することは誤差が生じる可能性がある。また、聴衆役を実際に用意する必要があるため、利用者1人でこのシステムを用いて練習を行うことは難しい。

^{†1} 東京工科大学大学院

^{†2} 東京工科大学

本研究では利用者のプレゼンテーションに対するリアルタイムフィードバックを行うことができること、

即時フィードバックを行う 3D モデルを用いてプレゼンテーションシーンを作成することで、よりインタラクティブなプレゼンテーション練習環境を作成することを目標とする。

3. 提案手法

本研究では利用者が 1 人で利用できる、より本番に近いプレゼンテーションを疑似体験できる練習システムの提案を行う。そこで、本研究では 3D モデルを用いて、プレゼンテーションにリアルタイムに反応する聴衆アバターを配置し、利用者にインタラクションを感じられる環境を作成した。また、本システムを用いたプレゼンテーション後には、発表者の癖をデータ化した客観的な評価を提示する機能を搭載した。

3.1 臨場感を感じられる環境の作成

本システムは図 1 のような実際のプレゼンテーション環境を模した 3DCG 空間で、プレゼンテーションを練習するシステムである。

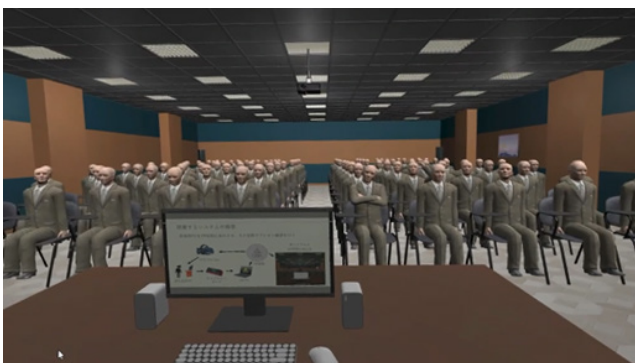


図 1：利用者のプレゼンテーション中の画面

利用者のプレゼンテーションに応じて聴衆として席に座っている 3DCG モデルのアバターがリアルタイムに評価を行い、その内容に応じてフィードバックとなるアニメーションを行う。また、本システムでは利用者の視線情報を用いてプレゼンテーション評価を行うため、VRHMD である VIVE Pro Eye を利用した。

本システムでは様々なシチュエーションを想定して、小規模、中規模、大規模の 3 つのステージを用意した。この 3 つのステージはそれぞれ部屋の広さや形状、聴衆アバターの数が異なる。図 2 に用意した 3 つのステージを示す。

本番のプレゼンテーションを行う際は直前までその会場を明確にイメージできない場合が多い。そのため、複数の環境を予め用意しておくことで、利用者の状況に合わせた練習を行うことができる。



図 2：システムで利用するプレゼンテーションステージ

3.2 プレゼンテーションに即時フィードバックする聴衆

一般的に良いプレゼンテーションは早口ではなく、言い淀みが少なく適切に聴衆に視線を向けて発表時間を守ることが重要だとされている。そこで、本システムで実装した聴衆アバターはそれぞれ、プレゼンテーション中の利用者の視線情報と音声情報に応じてリアルタイムなフィードバックを提示する。このフィードバックは、不機嫌そうな態度をとるなど、聴衆アバターのアニメーションとして提示される。図 3 は個々の聴衆アバターがフィードバックのアニメーションを再生するまでのフローである。

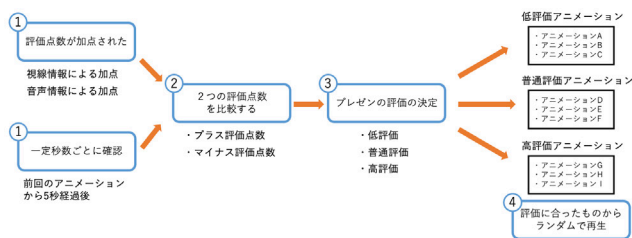


図 3：フィードバックが行われるまでのフロー

図 3 に記載された 2 つの①は、それぞれフローの開始条件を示す。これらの開始条件は後述するアバターがそれぞれ持つ 2 種類の評価点数が変動したとき、または評価点数が変動せずに 5 秒間経過したときに迎える。評価点数が変動しない場合でも 5 秒おきにアニメーションを変更するタイミングを設けることで、常に同じアニメーションを再生しつづける不自然な状態を緩和した。

また、図 3 の②のように、聴衆アバターはそれぞれ、利用者のプレゼンテーションの評価を記録するための 2 つの評価点数を持つ。これは、ポジティブな行為を行なった場合に加点するプラス評価点数と、ネガティブな行為を行った場合に加点するマイナス評価点数である。

次に、図 3 の②において視線情報を用いて評価点数を変動させる仕組みについて述べる。

本システムでは利用者が聴衆アバターに 1 秒以上視線を向けた場合、意識的にアイコンタクトをしていると判定してプラス評価点数が加点される。なお、このアイコンタクトし続けている時間のカウントは 2 秒間でリセットされる。したがって、例えば 4 秒同じアバターを見続けた場合は、2 サイクル分の 2 点が加点される。反対に、5 秒間 1 度もアイコンタクトの判定がなかったアバターはマイナス評価

点数が加点される。したがって、同じアバターのみを見続けてしまうと、他のアバターのマイナス評価が上昇しつづけることになる。また、聴衆アバターではなく周囲の壁やモニターを 10 秒以上見ていた場合は、向けるべき場所に視線を向けられていないと判定を行い、マイナス評価点数に加点される。

次に、図 3 の②において音声情報を元にした評価について述べる。音声情報は「言い淀みがないか」、「1 秒以上 2 秒以下の発言の間」が取れているか、「早口」となっていないか、「声の大きさ」の 4 つを基準に判定している。また、本システムにおける音声情報の検知には Windows Speech API と NMcCab を用いた。

まず、言い淀みの検知は Windows Speech API を用いてテキスト化された全ての発言内容のうち、『あー』や『えっと』といった、あらかじめ登録された内容を検知することで判定を行なった。言い淀みが検知されるごとに、マイナス評価点数が加点される。

発言の間は Windows Speech API が前回の検知を終えてから、次の検知を開始するまでの間を発言の間として判定を行なった。本システムでは文献[6]を参考に、1~2 秒の間が取れていた場合プラス評価点数に加点、間が 5 秒以上の場合はマイナス評価点数に加点、それ以外の秒数の場合はどちらにも加点されない。

次に、発言速度が速すぎる早口の判定には、1 センテンスごとの 1 秒間に発話した文字数で判定を行う。Windows Speech API と NMcCab を用いて文字数をカウントし、7.6 文字/秒を超えている場合は早口マイナス評価点数に加点される。反対に、7.6 文字/秒を超えていなかった場合は、プラス評価点数に加点をされる。

最後に、声の大きさの判定は 20db を閾値に判定している。20db 以下の発声が検知されるごとにマイナス評価点数が加点される。

図 3 の③は、上記で述べた視線情報や音声情報による評価点数の状態から、各聴衆アバターの状態を決定する工程である。各アバターは③の工程に入った時点で所持しているプラス評価点数、マイナス評価点数を次のルールで計算して、評価の状態を「低評価」「普通評価」「高評価」の中から決定する。

1. 現在のプラス評価点数とマイナス評価点数を比較する
2. 比較した際に「大きかった数値÷2 つの評価点数の合計値」を計算する
3. 手順 2 で求められた数値が 0.55 以下なら普通評価、それを超える数字の場合は手順 1 で大きかった数値がプラス評価点数なら高評価、マイナス評価点数なら低評価とする

最後に図 3 の④の手順では、③で決定した評価の状態に

応じたアニメーションリストから、アバターに適用するアニメーションがランダムに 1 つ選択される。

図 4 は左からそれぞれ聴衆アバターが行う低評価、普通評価、高評価状態でのアニメーション例である。



図 4：評価ごとのアニメーションの例

利用者の状況に合わせた練習を行うため、本システムでは難易度選択の機能を実装した。利用者は本システムを利用する際 Easy, Normal, Hard の 3 種類から自分に合った難易度を選択することができる。選択した難易度に応じて聴衆アバターが低評価と判断する係数が変化する。Easy と比べて Hard を選択した場合、聴衆アバターが利用者のプレゼンテーションに低評価を付けやすく、たくさんのフィードバックを行うようになる。

3.3 プレゼンテーション練習後の客観的評価機能

プレゼンテーションの準備を行う際、発表資料や発表の構成については書籍などを用いた準備が可能である。しかし、話し方や視線の向け方など、プレゼンテーションを行う際の所作については 1 人の練習で客観的に評価することが難しい。

そこで本システムでは、プレゼンテーション終了後、そのプレゼンテーションに関する情報を提示する機能を搭載した。ここでは、3.2 節のリアルタイムフィードバックに使用した視線・音声情報と利用者の発表時間をもとにしたデータが提示される。提示されるデータは次の 5 種類である。

- 利用者の視線のヒートマップ
- 利用者の発言の中で頻出の単語
- 利用者の言い淀み検知回数
- 利用者の発言の 1 秒あたり平均音節数
- 利用者のプレゼンテーション中に経過した時間

プレゼンテーションを行う際は、満遍なく様々なところに視線を向けてアイコンタクトをすることが望ましい。そこで本システムでは、視線を向けていた箇所のヒートマップを提示する。図 5 に示すように、利用者がプレゼンテーション中に視線を長時間向けたオブジェクトほど濃い赤色で表示される。このヒートマップ表示の対象はモニター、

壁、聴衆アバターである。

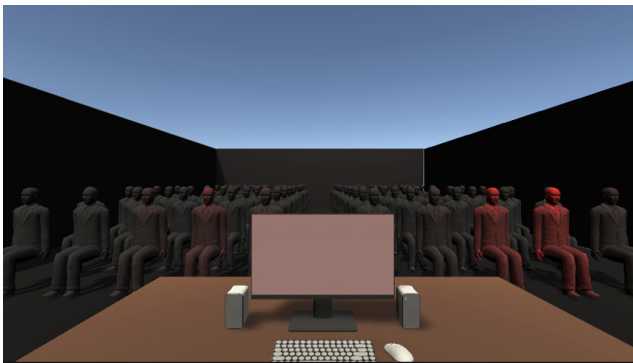


図5：利用者の視線が集まった場所の表示

次に、ヒートマップ以外の情報については、図6のように提示される。

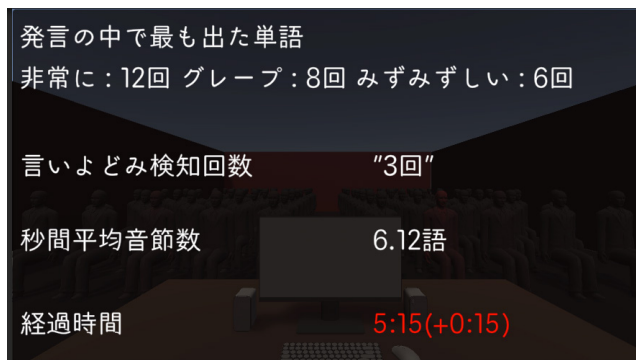


図6：終了後フィードバック画面

まず、頻出の単語の表示は Windows Speech API を利用して取得した利用者の発言内容を全てテキスト化したものを NMcCab によって形態素解析を行う。そうして取得されたデータリストから内容語のみをカウントし、カウント数が多かったものが提示される。利用者はこの情報をもとに、自身のプレゼンテーションではどのような言葉を用いて表現を行っていたか把握することができる。例えば、その言葉が直すべき口癖であった場合や、そのプレゼンテーションで適切ではない表現であった場合、次への改善に繋げることができる。

言い淀み検知回数および秒間平均音節数は、3.2 で述べた方法をもとに、プレゼンテーション開始から終了時点までの合計データが提示される。

最後に、経過時間の表示は本システムを利用する際に設定した制限時間を超えている場合、または制限時間の 80% に満たないような短い発表の場合は、警告として赤文字で表示される。

利用者はこれらの終了後フィードバックを認識することで、視線を全体に配っていたか？正しく伝えたいことを適

切な速度で伝えられていたか？といった要件を振り返ることができる。自身のプレゼンテーション能力向上に繋がる練習を行うためには、これらの項目を振り返ることが非常に重要である。

4. おわりに

本研究では、より臨場感のある環境 VR プレゼンテーション練習システムの提案を行った。また、利用者のプレゼンテーション内容に応じて客観的評価を提示することで、利用者が自身のプレゼンテーションの特徴を認識して改善に繋げることができるシステムを実装した。

今後の展望として、聴衆アバターの種類を性別、年齢などから利用者が選択することができる機能を検討したい。この機能を搭載することで、利用者がより苦手な空間を再現できるほか、より実際のプレゼンテーション環境に近い環境での体験が可能である。

また、利用者が目指すべきプレゼンテーションを分析したデータを用いて、3D アバターが行う評価の基準を変更することが挙げられる。利用者が自分の目指すべきプレゼンテーションをお手本として、話し方やパフォーマンスの練習ができるシステムへの発展が考えられる。

そして、利用者に対して評価実験を行い、本番のプレゼンテーションのような感覚を得ることができたかを評価検証したい。この結果をもとに、より本番のような臨場感を感じられるプレゼンテーション練習環境について検討を重ねていきたい。

参考文献

- [1] 厚生労働省若年者就職基礎能力修得のための目安委員会報告書. <https://www.mhlw.go.jp/houdou/2004/07/dl/h0723-4h.pdf>, (参照 2021-12-22).
- [2] 文部科学省中央教育審議会大学分科会(2008)学士過程教育の構築に向けて(審議のまとめ): 16. https://www.mext.go.jp/b_menu/shingi/chukyo/chukyo4/houkoku/080410.htm, (参照 2021-12-22)
- [3] JTB 広報室 コミュニケーションは苦手, 58%と過半数 主体的な発信は苦手, 受け身のコミュニケーションは得意. https://www.jtbcorp.jp/scripts_hd/image_view.asp?menu=news&id=00239&news_no=27, (参照 2021-12-22)
- [4] 周康, 三上浩司, 近藤邦雄, 鶴田直也. VR 技術を用いたプレゼンテーション練習システム. 映情学技報, Vol. 42, No. 12, AIT2018-79, pp. 135-136, 2018.
- [5] 樺山雄太, 横窪安奈, ロペズ ギョーム. 仮想現実を用いたオンラインプレゼンテーション支援システム. WISS 2020.
- [6] 栗原一貴, 後藤真孝, 緒方淳, 松坂要佐, 五十嵐健夫. "プレゼン先生: 音声情報処理と画像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, pp.59-64, 2006.
- [7] 宮野秀市. 全周囲パノラマ動画を利用した VR エクスパーチャー療法. VR 医学, Vol. 13, No. 1, pp. 1-10, 2015.