

Title	メタバースコミュニケーションにおける生成AIを用いた概念具現化と相互理解支援システム
Sub Title	Concept sharing and mutual understanding support system using generative AI in metaverse communication
Author	リュウ, 徳ガイ (Liu, Dekai) 小木, 哲朗 (Ogi, Tetsuro)
Publisher	慶應義塾大学大学院システムデザイン・マネジメント研究科
Publication year	2024
Jtitle	
JaLC DOI	
Abstract	
Notes	修士学位論文. 2024年度システムエンジニアリング学 第382号
Genre	Thesis or Dissertation
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40002001-00002024-0066

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the KeiO Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

修士論文

2024 年度

メタバースコミュニケーション
における生成 AI を用いた
概念具現化と相互理解支援システム

LIU DEKAI

(学籍番号：82333627)

指導教員 小木 哲朗

2025 年 3 月

慶應義塾大学大学院システムデザイン・マネジメント研究科
システムデザイン・マネジメント専攻

Concept Sharing and Mutual
Understanding Support System Using
Generative AI in Metaverse
Communication

LIU DEKAI

(Student ID Number : 82333627)

Supervisor Tetsuro Ogi

March 2025

Graduate School of System Design and Management,
Keio University
Major in System Design and Management

論 文 要 旨

学籍番号	82333627	氏 名	劉徳凱
論文題目： メタバースコミュニケーションにおける生成 AI を用いた概念具現化と相互理解支援システム			
(内容の要旨) 近年のメタバース技術の発展に伴い、遠距離コミュニケーションの機会が増加している一方で、異なる文化や専門分野を持つ人々の間では、言語のみではニュアンスや抽象的な概念が十分に伝わらず、相互理解が困難となる課題が存在している。これにより、遠距離コミュニケーションの質が低下する可能性が指摘されている。 本研究では、こうした課題に対応するために、生成 AI を用いて言語表現を視覚化するリアルタイムシステムを開発し、メタバース環境におけるコミュニケーションの円滑化を目指した。本システムは、ユーザが言語で表現した内容をリアルタイムで画像として生成・共有することで、コミュニケーションの障壁を低減し、概念の共有と相互理解を支援するものである。 システム設計では、多空間デザインモデルを採用し、価値空間、意味空間、状態空間、属性空間の 4 つの視点から設計を行った。また、生成 AI モデルとメタバースプラットフォームを統合し、音声処理、キーワード抽出、プロンプト生成、画像生成、メタバース内表示までの一連のプロセスを構築した。 システムの有効性を評価するために、二つの実験を実施した。第一の実験では、システムがユーザの言語表現を正確に視覚化できる能力を検証した。結果として、ユーザのフィードバックを反映することで生成精度が向上することが確認された。第二の実験では、システムがユーザ間の相互理解を促進し、コミュニケーションの質を向上させる能力を評価した。結果として、視覚情報の共有が対話の補助ツールとして有効であることが示された。 実験結果から、本研究で開発したシステムは、生成 AI を活用した視覚化技術がメタバース内でのコミュニケーションを改善する可能性を示しており、異文化間の相互理解を深める新たなアプローチとして期待される。今後は、システムの多言語対応や初期生成精度の向上を目指し、さらなる研究を進める予定である。			
キーワード (5 語) メタバース、生成 AI、視覚化、相互理解、コミュニケーション支援			

SUMMARY OF MASTER' S DISSERTATION

Student Identification Number	82333627	Name	LIU DEKAI
Title Concept Sharing and Mutual Understanding Support System Using Generative AI in Metaverse Communication			
Abstract In this study, we aimed to address communication challenges in metaverse environments by developing a system that enhances mutual understanding and concept visualization using generative AI. While the metaverse offers opportunities for global interactions beyond physical boundaries, challenges remain in effectively communicating nuanced concepts and abstract ideas, especially among individuals with diverse cultural and professional backgrounds. These challenges often lead to misunderstandings and reduced communication quality. To solve these issues, we designed and implemented a real-time visualization support system that transforms language-based expressions into visual representations. The system integrates speech processing, keyword extraction, and image generation to assist users in conveying their ideas more effectively during metaverse-based communication. Utilizing the Multispace Design Model, we accounted for various design dimensions, including value space, meaning space, state space, and attribute space, to ensure comprehensive system functionality. We conducted two experiments to evaluate the system's effectiveness. The first experiment focused on the system's ability to accurately visualize users' linguistic expressions. Results demonstrated that user feedback improved the quality of generated visuals, making them more aligned with user intent. The second experiment assessed the system's role in enhancing mutual understanding and communication quality during interactions between users. The results confirmed the system's effectiveness in fostering mutual understanding and conceptual clarity. These findings suggest that generative AI-based visualization systems can play a critical role in overcoming communication barriers within the metaverse. Future research will focus on enhancing the system's initial visualization accuracy, expanding multilingual adaptability, and improving system response times to further support global communication needs.			
Key Word (5 words) Metaverse, Generative AI, Visualization, Mutual Understanding, Communication Support			

目次

第 1 章	序論	1
1.1	背景	1
1.1.1	メタバースの発展と可能性	1
1.1.2	遠距離コミュニケーションにおける課題	1
1.1.3	異文化間コミュニケーションの現状	2
1.2	問題提起	2
1.2.1	言語の限界と視覚化の必要性	2
1.2.2	コミュニケーションの重要性	3
1.2.3	メタバースにおけるコミュニケーションギャップ	4
1.2.3.1	非言語的要素の欠如	4
1.2.3.2	異文化間のギャップ	4
1.2.3.3	情報の過多と不足	4
1.3	研究目的と意義	5
1.3.1	本研究の目的	5
1.3.2	実社会への貢献	5
	(1) 教育分野への応用	5
	(2) グローバルビジネスでの活用	6
	(3) 障害者支援への応用	6
	(4) クリエイティブ分野での利用	6
第 2 章	先行研究	7
2.1	メタバースにおけるコミュニケーションの特性と課題	7
2.2	可視化ツールがコミュニケーションに与える影響	7
2.3	多空間デザインモデルの適用	8
2.3.1	多空間デザインモデルの概要	8
2.3.2	メタバースにおける適用可能性	8
2.4	差異の可視化による相互理解支援システム	9
2.5	先行研究の総括	9
2.6	本研究の独自性と貢献	9
2.6.1	従来研究との差異	9
2.6.2	本研究の新規性	10
第 3 章	システム設計	11
3.1	システム設計の決定方法と目的	11
3.1.1	設計における課題分析	11
	(1) 非言語的要素の欠如	11

(2)	言語的ギャップ	11
(3)	抽象的概念の共有の難しさ	11
3.1.2	解決方法の選定	12
3.1.2.1	非言語的要素の欠如に対して	12
3.1.2.2	言語的ギャップに対して	12
3.1.2.3	抽象的概念の共有の難しさに対して	12
3.1.2.4	解決方法を実現するためのアプローチ	13
(1)	多空間デザインモデルの適用	13
(2)	生成 AI 技術の統合	13
(3)	リアルタイム視覚化システムの構築	13
(4)	インタラクティブなユーザーインターフェースの開発	13
(5)	メタバース環境での統合と表示	13
(6)	フィードバックループの実装	14
3.2	多空間デザインモデルの適用	14
3.2.1	心理空間の設計	15
3.2.2	物理空間の設計	15
3.2.3	心理空間と物理空間の相互作用	16
3.2.4	知識空間の活用	16
3.3	システムの構造設計	16
3.3.1	音声処理モジュール	17
3.3.2	キーワード抽出アルゴリズム	17
3.3.3	プロンプト生成の仕組み	17
3.3.4	画像生成エンジンの選定	18
3.3.5	メタバース表示機能の設計	18
3.4	システムの実装環境	18
3.4.1	開発ツールとプラットフォーム	18
(1)	プログラミング環境	18
(2)	開発環境	19
(3)	メタバースプラットフォーム	19
①	Unity	19
②	Cluster	19
③	VRChat	19
(4)	クラウド環境	19
3.4.2	使用技術とフレームワーク	20
(1)	自然言語処理 (NLP)	20
(2)	画像生成技術	20
(3)	音声認識技術	20
3.4.3	システムの動作フロー	20

(1)	音声入力	21
(2)	キーワード抽出	21
(3)	プロンプト生成	21
(4)	画像生成	21
(5)	メタバース内での表示	21
(6)	フィードバックの反映	21
第 4 章	実験 1: システムの視覚化能力の検証	22
4.1	事前準備	23
4.1.1	被験者の選定	23
4.1.2	実験環境の準備	23
4.1.2.1	機材の準備	24
4.1.2.2	ソフトウェア環境	24
4.1.2.3	個室環境	25
4.2	実験内容の実施	25
4.2.1	ステップ 1 : テーマの提示	26
4.2.2	ステップ 2 : 説明と初期画像生成	26
4.2.3	ステップ 3 : フィードバックと反復的生成	26
4.2.4	ステップ 4 : 最終画像の評価	27
4.2.5	ステップ 5 : 使用体験の記録	27
4.3	アンケートの実施	27
4.3.1	質問 1 : 生成された最初の画像は、あなたがイメージしていた物や概念に近いものでしたか?	28
(1)	質問目的:	28
(2)	質問結果:	28
(3)	質問結果分析:	28
4.3.2	質問 2 : 最終的に生成された画像は、あなたがイメージしていた物や概念に非常に近いものでしたか?	29
(1)	質問目的:	29
(2)	質問結果:	29
(3)	質問結果分析:	29
4.3.3	質問 3 : 各回のフィードバックを経て、生成画像が改善されたと感じましたか?	29
(1)	質問目的:	30
(2)	質問結果:	30
(3)	質問結果分析:	30
4.3.4	質問 4 : システムが画像を生成するまでの時間に満足しましたか?	30
(1)	質問目的:	31
(2)	質問結果:	31

(3) 質問結果分析：	31
4.3.5 質問5：システムの使いやすさについてどう感じましたか？	32
(1) 質問目的：	32
(2) 質問結果：	32
(3) 質問結果分析：	32
第5章 実験2: システムによる概念具現化の支援能力の妥当性確認.....	34
5.1 事前準備.....	35
5.1.1 テーマの設定	35
(1) 身近で具体的な対象.....	35
(2) 多様な視点を引き出す可能性	35
(3) 具体性と抽象性の両立	35
5.1.2 被験者の選定	35
5.1.3 実験環境の準備	36
5.1.3.1 機材の準備	36
5.1.3.2 ソフトウェア環境	37
5.1.3.3 個室環境.....	37
5.2 実験内容の実施.....	37
5.2.1 ステップ1：対話プロセスの設定	38
5.2.2 ステップ2：視覚化とフィードバックのプロセス	38
5.2.3 ステップ3：生成プロセスの反復.....	38
5.2.4 ステップ3: 結果の収集	40
5.3 アンケートの実施.....	40
5.3.1 質問1：システムによって生成された画像は、相手の考えを理解する助けになりましたか？	40
(1) 質問目的：	40
(2) 質問結果：	40
(3) 質問結果分析：	40
5.3.2 質問2：システムを使用することで、自分の考えをより具体的に表現できましたか？	41
(1) 質問目的：	41
(2) 質問結果：	41
(3) 質問結果分析：	41
5.3.3 質問3：発言ごとに生成された画像が、相手のイメージと一致していくと感じましたか？	42
(1) 質問目的：	42
(2) 質問結果：	42
(3) 質問結果分析：	42
5.3.4 質問4：システムが画像を生成するまでの時間に満足しましたか？	43

(1) 質問目的：	43
(2) 質問結果：	43
(3) 質問結果分析：	43
5.3.5 質問 5：システムが対話の質を向上させるのに役立ったと感じましたか？	
44	
(1) 質問目的：	44
(2) 質問結果：	44
(3) 質問結果分析：	44
第 6 章 総合的な考察	46
6.1 実験 1 と実験 2 の結果統合	46
6.2 システムの全体的な有効性と限界	47
6.2.1 有効性の分析	47
(1) 抽象的概念の具体化	47
(2) コミュニケーションの効率化	47
(3) 動的プロセスの有効性	47
6.2.2 限界の分析	47
(1) 初期生成精度の課題	47
(2) 生成時間の問題	47
(3) ユーザ依存性の強さ	48
6.3 今後の課題	48
6.3.1 初期生成精度の向上	48
6.3.2 処理速度の向上	48
6.3.3 新システムの開発	49
第 7 章 結論	50
7.1 研究の概要	50
7.2 システム設計と実装	50
7.3 実験結果と考察	50
参考文献	52
謝辞	54
第 8 章 付録	55

図目次

図 1	思考空間	14
図 2	思考空間中価値空間、意味空間、状態空間、属性空間の定義	15
図 3	「実験 1: システムの視覚化能力の検証」の被験者各年齢層割合	23
図 4	「実験 1: システムの視覚化能力の検証」の被験者性別割合	23
図 5	実験中使用する HMD の例	24
図 6	「実験 1: システムの視覚化能力の検証」を行う個室環境	25
図 7	「実験 1: システムの視覚化能力の検証」を行う初期生成の様子	26
図 8	「実験 1: システムの視覚化能力の検証」を行う再度生成の様子	27
図 9	「実験 1: システムの視覚化能力の検証」の質問 1	28
図 10	「実験 1: システムの視覚化能力の検証」の質問 2	29
図 11	「実験 1: システムの視覚化能力の検証」の質問 3	30
図 12	「実験 1: システムの視覚化能力の検証」の質問 4	32
図 13	「実験 1: システムの視覚化能力の検証」の質問 5	33
図 14	「実験 2: システムによる概念具現化の支援能力の妥当性確認」の各被験者の国別割合	36
図 15	「実験 2: システムによる概念具現化の支援能力の妥当性確認」の各被験者学歴割合	36
図 16	「実験 2: システムによる概念具現化の支援能力の妥当性確認」実験環境	37
図 17	「実験 2: システムによる概念具現化の支援能力の妥当性確認」の初期画像生成の様子	39
図 18	「実験 2: システムによる概念具現化の支援能力の妥当性確認」の再度画像生成の様子	39
図 19	「実験 2: システムによる概念具現化の支援能力の妥当性確認」質問 1	41
図 20	「実験 2: システムによる概念具現化の支援能力の妥当性確認」質問 2	42
図 21	「実験 2: システムによる概念具現化の支援能力の妥当性確認」質問 3	43
図 22	「実験 2: システムによる概念具現化の支援能力の妥当性確認」質問 4	44
図 23	「実験 2: システムによる概念具現化の支援能力の妥当性確認」質問 5	45

表目次

表 4-1 実験 1: システムの視覚化能力の検証の概要	22
表 5-1 実験 2: システムによる概念具現化の支援能力の妥当性確認	34

第1章 序論

本章では、近年急速に発展しているメタバース技術の背景と、それが社会にもたらす可能性について議論する。また、メタバース環境におけるコミュニケーションの課題となる言語的および非言語的な要素の欠如による相互理解の難しさを指摘する。このような課題に対処するための生成 AI 技術の可能性についても触れる。さらに、本研究が取り組む具体的な課題と目的を明確にし、メタバースにおける相互理解の促進に向けたアプローチを示す。最後に、本論文の構成を簡潔に述べ、全体の流れを説明する。

1.1 背景

近年、社会的課題の複雑化やグローバル化に伴い、テクノロジーを活用した新しい形態のコミュニケーションが注目されている。その中でも、メタバース(Metaverse)技術は、現実世界と仮想空間を融合させることで、従来の物理的な制約を超えた革新的な体験や協働の場を提供している。この技術は、教育、医療、ビジネス、エンターテインメントなど、多岐にわたる分野での応用が進んでおり、次世代のデジタル基盤として注目されている[1]。

1.1.1 メタバースの発展と可能性

メタバースは VR(仮想現実)や AR(拡張現実)、3D モデリングといった先進技術を基盤に構築されており、インターネットの次世代形態として位置付けられている。その市場は急速に拡大しており、Grand View Research (2022)によると、メタバース経済の市場規模は 2021 年に約 388 億 5000 万ドルに達し、2022 年から 2030 年までの年平均成長率(CAGR)は 39.4%に達すると見込まれている。この成長率が維持されれば、2030 年までにメタバース市場は 7722 億 4000 万ドル規模に達する可能性がある。また、メタバースは単に仮想空間の体験だけでなく、金融活動や新たな経済圏の構築にも寄与すると予測されており、今後はブロックチェーン技術や暗号通貨を活用した経済インフラとしても期待されている [1] [2]。

1.1.2 遠距離コミュニケーションにおける課題

メタバース環境は、物理的な距離を超えた人々のつながりを促進する可能性を秘めている

が、現実空間とは異なる特有の課題が存在する。特に、非言語的要素の欠如が重要な問題として挙げられる。例えば、表情、声のトーン、ジェスチャーなどが伝わりにくい環境では、意思疎通が難しく、誤解が生じる可能性が高い[3]。さらに、Mehta の研究では、異なる文化背景を持つ人々が集まる環境では、言語や価値観の相違による理解不足が発生しやすいことが示されている。[4]このような課題を克服するためには、従来のコミュニケーションツールを補完する新たな技術の導入が必要である。

1.1.3 異文化間コミュニケーションの現状

異文化間での協働は、グローバル社会においてますます重要性を増しているが、同時に課題も多い。Knippenberg らの研究では、多様なバックグラウンドを持つメンバー間の情報統合が意思決定に大きな影響を与える一方で、文化的もしくは社会的な違いが対立や誤解の原因となる可能性も指摘されている。[5]これに加え、グループ内の結束が情報交換と相互理解を促進する要因であると述べている。メタバース環境では、これらの障壁をテクノロジーの力で緩和し、文化や言語の違いを超えたコミュニケーションが可能となる新しいアプローチが求められている[6]。

1.2 問題提起

メタバース技術の進展により、物理的制約を超えた新しい形態のコミュニケーションが可能になり、教育、ビジネス、エンターテインメントなど、多くの分野において新たな価値を創出している[7]。この技術革新は、地理的距離を超えた交流の場を提供し、多文化社会の統合や効率的な協働を促進する可能性を秘めている。しかしながら、その環境では、従来のコミュニケーション手法が適用しにくい課題が明らかになっている。その中でも重要なのが、「言語表現の限界」と「非言語的要素の欠如」によるコミュニケーションギャップである。

1.2.1 言語の限界と視覚化の必要性

デジタル技術の進展に伴い、オンラインコラボレーションはビジネス、教育、研究活動において不可欠な手段となっている。Osler らは、急速に普及したビデオ会議ツールが直面した課題として言語だけに頼ることの限界を指摘している。従来の対面コミュニケーションでは、表情やジェスチャー、姿勢といった非言語的要素が会話の理解を助ける役割を果たしていたが、オンライン環境ではこれらの要素が大幅に制限されることになる。文化的背景が異なる相手

との対話においては、言葉の選び方やニュアンスが伝わりにくくなるため、誤解が生じやすくなる。

Osler らの研究では、オンラインコラボレーションにおける対話の流動性の低下についても言及されている。ビデオ会議では、話者が一度に一人しか発言できないため、対話のテンポが遅くなり、自然なインタラクションが妨げられることが多い。また、音声の遅延や通信の途切れが、会話のリズムや相手の反応を正確に読み取ることを困難にし、意思疎通の質を低下させる。これにより、偶発的なコミュニケーションの減少が発生し、対面での雑談やちょっとした会話を通じたアイデアの共有が減ることが問題視されている。

さらに、オンライン環境では、視線のズレやカメラの角度の違いによって、相手の注意がどこに向いているのかを把握するのが難しくなる。対面でのコミュニケーションでは、アイコンタクトや身振りが会話の重要な手がかりとなるが、ビデオ会議ではこれらが制限されるため、相手が話を理解しているかどうかを判断するのが難しくなる。このような非言語的要素の欠如は、特に意思決定や議論を行う場面において、発言者の意図が正確に伝わらず、認識のずれを引き起こす可能性がある。

1.2.2 コミュニケーションの重要性

異なる文化的背景を持つ人々が協働する場面では、効果的なコミュニケーションが不可欠である。Mehta らは、チーム内の知識統合がチームのパフォーマンスに大きな影響を与えることを示し、特に異文化間のコミュニケーションでは言語的な違いや非言語的要素の欠如が誤解を生じさせる要因となることを指摘している[4]。言語のニュアンスの違いや文化特有の表現が相手に正しく伝わらない場合、意図とは異なる解釈がされることがあり、結果として情報伝達の障壁が生じる可能性がある。

また、Mehta らは、チームの知識統合が促進されるためには、メンバー間の相互理解が不可欠であり、特に多文化環境では、言語による意思疎通だけでなく、視覚的補助やフィードバックを活用した対話が求められると述べている。この視点は、メタバース環境におけるコミュニケーションにも適用できる。従来の対面での会話では、ジェスチャーや表情といった非言語的な手がかりがコミュニケーションの円滑化に寄与していたが、デジタル環境ではこれらの要素が不足し、意思疎通のギャップが生じやすい。

さらに、Mehta らの研究によると、チームの目標志向が知識統合の促進に重要な役割を果たすことが明らかになっている。学習志向のチームは、メンバーが積極的に知識を共有し、相互に補完し合うことで、より良い意思決定を行う傾向がある。一方で、競争的な目標を持つチームでは、情報が戦略的に保持されることがあり、知識の断片化が進み、誤解が増える可能性がある[4]。このことから、異文化間のコミュニケーションにおいては、開かれた対話と知識共

有を促進する環境を整えることが極めて重要である。

特に、メタバースのようなデジタル環境では、言語の壁を超えた効果的なコミュニケーション手段の導入が求められる。リアルタイムの翻訳技術や、生成 AI を活用した視覚的フィードバックシステムの導入により、異なる文化的背景を持つユーザ同士の相互理解を深めることができる。こうした技術を活用することで、異文化間の誤解を最小限に抑え、より円滑な知識統合が可能になる。

異文化間コミュニケーションにおいては、単に言語的な要素に依存するのではなく、相互理解を促進するための多様なアプローチが必要である。Mehta らの研究が示すように、知識統合の成功は、チームの目標志向や文化的多様性を考慮した適切なコミュニケーション戦略によって大きく左右される[4]。メタバース環境における効果的なコミュニケーションの実現には、技術的な補助と心理的な障壁を乗り越える工夫の両面が求められることを示す。

1.2.3 メタバースにおけるコミュニケーションギャップ

メタバース環境は、ユーザがバーチャル空間で他者と交流し、協力する場を提供する。しかし、現実空間と比較して、メタバース内のコミュニケーションには以下のようなギャップが存在する。

1.2.3.1 非言語的要素の欠如

現実空間では、表情、身振り、声のトーンといった非言語的要素が、情報の伝達を補完する重要な役割を果たしている。しかし、メタバース環境ではこれらが十分に再現されないため、相手の意図や感情を正確に理解することが難しい場合がある[3]

1.2.3.2 異文化間のギャップ

メタバースは多文化交流の場として機能するが、文化的な背景や価値観の違いによるコミュニケーション障害が発生することがある。たとえば、ある文化ではポジティブとされる表現が、別の文化ではネガティブに受け取られる場合がある。こうした文化的ギャップを埋めるためには、より普遍的で客観的なコミュニケーション手段が求められる[8]。

1.2.3.3 情報の過多と不足

メタバース環境では、大量の情報が同時に伝達される場面が多い一方で、必要な情報が不

足することもあり得る。このような情報の不均衡は、コミュニケーションの質や効率を低下させる要因となる[7]。

1.3 研究目的と意義

本研究では、メタバース環境におけるコミュニケーションの質的向上を目指し、生成 AI 技術を活用した視覚化システムを提案する。このシステムは、従来の言語中心のコミュニケーションでは克服できなかった課題を解決することを目的としており、その意義は多岐にわたる。本節では、研究の具体的な目的と、それが実社会においてどのように貢献するかを述べる。

1.3.1 本研究の目的

本研究の目的は、メタバース環境におけるコミュニケーションの質を向上させるために、生成 AI を活用した視覚化システムを開発することである。メタバースは、地理的制約を超えた新しい交流の場を提供する一方で、非言語的要素の欠如や言語表現の限界といった課題が存在する。これらの課題は、特に異文化間や異分野間でのコミュニケーションにおいて顕著であり、相互理解を妨げる要因となっている。

本研究の目的は、異なる文化的背景や専門知識を持つユーザ間での共通理解を促進することである。視覚化された情報を基に議論を進め、誤解を減らし、円滑なコミュニケーションを実現することを目指している。

1.3.2 実社会への貢献

本研究で提案する生成 AI 技術を活用した視覚化システム実社会において様々な分野で大きな可能性を秘めている。

(1) 教育分野への応用

本システムは、教育現場において複雑な概念を直感的に理解できるよう支援するツールとしての役割を果たす。特に、科学技術、医療、建築デザインなど、専門知識を伴う分野では、視覚的に補完された情報が学習者の理解を促進し、教育効果を向上させると考えられる[6]。遠隔教育やバーチャルクラスルームにおいても、学生が抽象的な概念を可視化することで学

習の質を高める可能性もある[3]。

(2) グローバルビジネスでの活用

グローバルビジネスでは、多国籍チームが直面する言語や文化の壁がコミュニケーションの障害となる場合が多い[4]。提案システムは、共通の視覚情報を基にした議論を可能にすることで、異文化間のギャップを埋め、意思決定プロセスを効率化する。これにより、多国籍チームの生産性を向上させ、より創造的かつ実践的な成果を促進する。

(3) 障害者支援への応用

本システムは、聴覚障害者や発話に困難を抱える人々のコミュニケーションを支援するツールとしての可能性を持つ。視覚化された情報は、音声や言葉に頼らず意思疎通を行うための代替手段となり得る。このような支援技術は、障害者が社会参加を拡大するための重要な一歩となる[8]

(4) クリエイティブ分野での利用

デザインやエンターテインメントの分野において、提案システムは創造プロセスの効率を大幅に向上させることが期待される。具体例として、製品デザインのプロトタイピングや映画制作におけるストーリーボードの作成が挙げられる。視覚化技術により、アイデアが即座に共有され、実行可能な形で進化することが可能となる[1]。

第2章 先行研究

2.1 メタバースにおけるコミュニケーションの特性と課題

メタバースは、仮想空間内でユーザが相互に交流し、協働することを可能にする次世代のデジタルプラットフォームとして注目されている。近年の研究では、メタバース内のコミュニケーションが現実世界のコミュニケーションと異なる特性を持つことが指摘されている。特に、非言語的要素の欠如や文化的背景の違いが、コミュニケーションギャップを生じさせる要因として挙げられている[3]。

Knippenberg らは、チームメンバー間の多様な視点を効果的に統合するプロセスが、意思決定や問題解決の質を向上させることを示した[5]。一方で、Mehta らは、多文化的なチームが持つ知識の多様性が新たなアイデアを生み出す可能性を秘めていることを指摘している[4]。しかし、これらの成果をメタバースに応用する際には、仮想空間特有の制約を克服する必要がある。

メタバース環境では、ユーザ間のコミュニケーションが文字情報や音声情報に依存する傾向があるため、非言語的要素の欠如が誤解や意思疎通の失敗を引き起こす可能性が高い。これを補完するために、生成 AI 技術を活用した視覚化システムが有効であると考えられる。本研究では、既存の研究を踏まえ、メタバース内での相互理解を支援するための新しいアプローチを提案する。

2.2 可視化ツールがコミュニケーションに与える影響

可視化ツールがコミュニケーションに与える影響およびその役割を明らかにすることを目的に着眼とした。コミュニケーション中知識の共有と収束を促進する可視化ツールの可能性を探るために、Fischer らの研究を参考にした。

Fischer らの研究では、可視化ツールが協調学習における知識の収束にどのような影響を及ぼすかを検討した。彼らは、コンテンツに依存しない汎用的な可視化ツールと、特定の内容に特化した可視化ツールを比較し、それぞれが学習者の知識共有にどのような影響を与えるかを分析した。その結果としてコンテンツ特化型の可視化ツールを用いることで、学習者がより適切な知識リソースを活用し、共通の知識体系へ収束しやすくなることが示された。また、ビデオ会議環境と対面環境の違いも調査され、ビデオ会議環境ではプロセスの収束が低くなる一方で、最終的な学習成果には大きな影響を与えないことが示された[13]。

Fischer らは、コンピュータ支援協調学習の文脈で、外部表象ツールの役割を分析した。彼

らの研究では、二人の学習者が共同作業を行う際の知識の共有と収束のプロセスを評価し、外部表象ツールがどのように学習の調整や意思疎通の円滑化を支援するかを検討した[13]。その結果、外部表象ツールを利用することで、学習者間の知識の相互影響が促進され、特に概念的な知識の適用において収束が高まることが示された。しかし、学習者が完全に同じ知識を共有するわけではなく、収束した知識と収束しなかった知識のバランスが学習効果に影響を与えることも指摘された[13]。

しかし、これらの研究では「マインドマップ」や「グラフ」といった静的な可視化ツールが主に検討され、リアルタイムに動的に生成される可視化ツールの可能性については十分に考慮されていない。本研究では、この知見を発展させ、生成 AI を活用したリアルタイム可視化システムを導入し、メタバース環境に適用する新しい手法を提案する。これにより、異なる背景を持つ人々の間での知識共有と収束を促進し、より直感的かつ効果的なコミュニケーションを実現することを目指す。

2.3 多空間デザインモデルの適用

2.3.1 多空間デザインモデルの概要

多空間デザインモデルは、価値空間、意味空間、状態空間、属性空間といった複数の異なる空間を統合的に設計し、複雑な課題に対する解決策を提供するためのフレームワークである。このモデルは、都市計画、製品設計、ソフトウェア開発などの分野で広く適用されている[9]。

2.3.2 メタバースにおける適用可能性

メタバース環境では、ユーザ間の価値共有や意味の統一がコミュニケーションの円滑化に寄与すると考えられる。多空間デザインモデルを活用することで、視覚化システムの設計が効率的に行えるだけでなく、異文化間や異分野間の相互理解を促進するための基盤を提供できる。本研究では、多空間デザインモデルを視覚化システムの設計に適用し、メタバース内でのコミュニケーションギャップを克服することを目指す。

2.4 差異の可視化による相互理解支援システム

本研究において、コミュニケーションにおける可視化ツールの影響を探ることが非常に重要となっている。特に、異なるバックグラウンドを持つ人々が概念の共有および相互理解を促進する方法として、可視化技術の有用性に着目した。大平の研究では、EVIDII (Environment for Visualizing Differences of Individual Impressions) を開発し、可視化技術が知識共有や意見の統合にどのような影響を与えるかを検討した[14]。EVIDII は、個人間の意見の違いを視覚的に表現することで、認識のズレを明確化し、合意形成を促進することを目的としている[14]。この研究では、議論の過程で生じる意見の相違や情報の非対称性を可視化することで、参加者がより深い理解に到達することができることを示した。しかし、EVIDII の可視化手法は主に静的なデータを基にしたものであり、リアルタイムでの動的な情報更新には対応していないという課題があった。

2.5 先行研究の総括

本研究では先行研究の知見を発展させ、生成 AI を活用したリアルタイム可視化システムを導入することで、メタバース環境における相互理解の促進を目指す。本システムは、会話の内容を解析し、リアルタイムで視覚的フィードバックを提供することで、非言語的な情報の欠如を補い、文化的背景の違いによる誤解を最小限に抑えることを目的とする。従来の静的な可視化ツールでは難しかった、リアルタイムでの概念のすり合わせを可能にし、異文化間や異分野間でのコミュニケーションをよりスムーズにすることが期待される。

2.6 本研究の独自性と貢献

2.6.1 従来研究との差異

従来の研究ではメタバース環境でのコミュニケーション課題に対し、主にテキストチャットや音声チャットといった言語的手段に依存していた。しかし、これらの手段では、非言語的要素の欠如や言語の限界による課題が十分に解決されていない。本研究は、生成 AI を用いた視覚化技術を導入することで、これらの課題を克服しようとする点で従来研究とは異なる。

2.6.2 本研究の新規性

本研究の新規性は、生成 AI 技術と多空間デザインモデルを統合し、リアルタイムでの視覚化を可能にするシステムを提案する点にある。このシステムは、言語情報を補完し、異文化間および異分野間での共通理解を促進する新しいアプローチを提供するものである。

第3章 システム設計

本章では、本研究で提案する視覚化システムの設計について述べる。まず、システム設計の目的と設計決定のプロセスを説明し、多空間デザインモデルを基盤とした各設計要素の内容を詳述する。さらに、システム構造の設計と実装環境について述べ、リアルタイム視覚化の実現に向けたアプローチを明確にする。本研究では、メタバース環境におけるコミュニケーションの課題を解決するために、生成 AI 技術を活用したシステムを構築することを目的としている。

3.1 システム設計の決定方法と目的

3.1.1 設計における課題分析

本研究が研究対象とするメタバース環境におけるコミュニケーションでは、三つの課題があることを確認した。

(1) 非言語的要素の欠如

メタバース環境では、表情や仕草などといった非言語的要素が伝わりにくく、これがニュアンスや感情の共有を妨げる要因となっている。

(2) 言語的ギャップ

異文化間では、言語表現が十分に理解されないことが多く、背景知識や解釈の違いが誤解を招く。

(3) 抽象的概念の共有の難しさ

技術的な内容や抽象的な概念を言語のみで正確に共有することは困難であり、意思疎通を妨げる要因となる。

これらの課題を解決するため、本研究では以下の 3 つのシステム要件を設計の目標として

掲げる。第一に、ユーザが入力した言語的な表現をリアルタイムで視覚化する機能を実現することを目指す。この機能により、ユーザは自分の意図やアイデアを即座に視覚的に確認できるようになる。第二に、生成 AI 技術を活用し、抽象的な概念を具体的かつ分かりやすい画像として表現する技術を導入する。これにより、複雑なアイデアやニュアンスを、より直感的に共有することが可能となる。第三に、ユーザ間のコミュニケーションをインタラクティブに支援する手段を提供し、相互理解を促進することを目標とする。このように、本研究のシステムは、視覚化と生成 AI 技術を組み合わせた新たなコミュニケーション手段を提案するものである。

3.1.2 解決方法の選定

3.1.1 中に挙げられたメタバース環境におけるコミュニケーション課題を解決するために、多空間デザインモデルを基盤として、生成 AI 技術を統合したシステムを設計する解決方法を通じて、ユーザ間の相互理解を促進することを目的としている。

3.1.2.1 非言語的要素の欠如に対して

視覚的な補完機能を導入する。生成 AI を活用して、ユーザの言語的表現を具体的な画像として視覚化することで、表情や仕草などの非言語的要素を再現し、ニュアンスや感情の共有を可能にする。この機能により、メタバース環境におけるコミュニケーションの精度が向上すると期待される。

3.1.2.2 言語的ギャップに対して

自然言語処理技術を基盤としたアルゴリズムを用い、異文化間での言語表現の意味を正確に解釈し補完する仕組みを提供する。この仕組みでは、言語モデルを活用して、ユーザが発する言葉の背景知識や文脈を分析し、誤解を最小限に抑えるプロンプトを生成する。

3.1.2.3 抽象的概念の共有の難しさに対して

リアルタイムでの視覚化機能とインタラクティブなフィードバック機能を組み合わせる。この機能では、ユーザが提示する抽象的なアイデアを具体的な画像として生成し、その画像を他のユーザと共有するプロセスを通じて相互理解を深める。また、ユーザからのフィードバックを基に画像を更新し、最適化することで、より正確な概念共有を実現する。

3.1.2.4 解決方法を実現するためのアプローチ

(1) 多空間デザインモデルの適用

システム設計を価値空間、意味空間、状態空間、属性空間の 4 つの側面から分析し、設計内容を構築する。ユーザの多様なニーズに応じた柔軟な設計が可能を実現することを目指し、非言語的要素の補完や言語的ギャップの解消が期待される。

(2) 生成 AI 技術の統合

生成エンジンとして、DALL-E や Stable Diffusion といった画像生成 AI モデルを採用し、内容生成においては ChatGPT を利用した。これにより、自然言語処理技術を組み合わせることで、ユーザの意図をより正確に反映したプロンプトを生成し、プロンプトを用いて抽象的な概念を分析し視覚的にわかりやすい画像として表現することを可能にする。

(3) リアルタイム視覚化システムの構築

音声処理モジュールやキーワード抽出アルゴリズムを導入し、ユーザが入力した言語的表現を即座に視覚化するシステムを構築する。このリアルタイム機能により、ユーザは生成結果を即座に確認し、必要に応じてフィードバックを提供することができる。

(4) インタラクティブなユーザーインターフェースの開発

ユーザが生成された画像をリアルタイムで確認し、修正や変更をリクエストできるインタラクティブなインターフェースを設計する。これにより、動的なコミュニケーションが可能となり、相互理解を促進する。

(5) メタバース環境での統合と表示

生成された視覚情報をメタバース環境内でリアルタイムに表示する仕組みを構築する。このために、cluster や unity などのプラットフォームを利用し、視覚的に洗練された表示方法を提供する。

(6) フィードバックループの実装

ユーザのフィードバックをリアルタイムで反映し、生成結果を反復的に改善するフィードバックループを実装する。これにより、生成画像の精度とユーザ満足度を向上させる。

以上の六つアプローチを組み合わせることで 3.1.1 に挙げた「非言語的要素の欠如、言語的ギャップ、および抽象的概念の共有」といった課題を解決し、ユーザ間の相互理解を支援するシステムの構築を目指す。

3.2 多空間デザインモデルの適用

多空間デザインモデル (Multispace Design Model) は、デザイン科学の理論に基づき、複雑な問題を体系的に解決するためのフレームワークである。このモデルは、思考空間と知識空間という 2 つの主要な構成要素から成り立つ。思考空間では、設計者の創造的な推論や構想が行われ、知識空間ではその推論を支える情報やデータが蓄積されている。本研究では、特に思考空間における心理空間と物理空間の相互作用を重視し、メタバース環境における相互理解を支援するシステムを設計した[11]。



図 1 思考空間

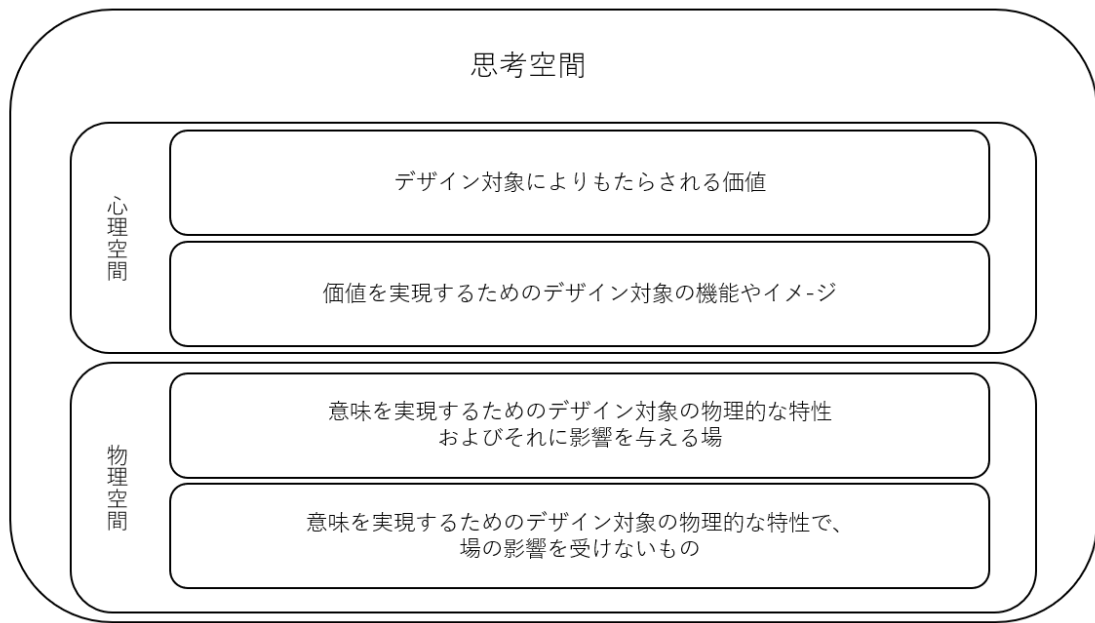


図 2 思考空間中価値空間、意味空間、状態空間、属性空間の定義

3.2.1 心理空間の設計

心理空間は、設計者の内面的な思考プロセスが展開される領域であり、抽象的な価値観や目標が形成される。本研究では、心理空間において、ユーザ間の相互理解を促進することを中心に設計を行った。この空間では、非言語的要素の補完や言語的ギャップの解消が重要な課題とされた。

生成 AI 技術を活用することで、抽象的な概念やニュアンスを視覚化し、ユーザ間で直感的に共有できる環境を構築した。また、自然言語処理 (NLP) 技術を用い、ユーザが発する言葉の背景や意図を解析し、それを基にした視覚情報の生成プロセスを心理空間の中核とした。これにより、ユーザが抱える認知的な負荷を軽減し、相互理解をよりスムーズに進めることが可能となった。

さらに、心理空間では、メタバース内での新たなユーザ体験の提供を目指し、感情や意図が正確に伝達される仕組みを構築した。この仕組みは、価値空間や意味空間に基づく具体的な目標設定を可能にし、システム全体の設計方向を決定づける役割を果たした。

3.2.2 物理空間の設計

物理空間は、心理空間で形成されたアイデアや目標を具現化し、現実世界で実装可能な形として提示する領域である。本研究では、物理空間において、生成された視覚情報をメタバ

ース環境内でリアルタイムに表示するシステムを設計した。

音声入力からテキストデータへの変換、キーワード抽出、プロンプト生成、画像生成、そしてメタバース内でのリアルタイム表示までのプロセスを統合的に設計した。また、ユーザが生成された画像にフィードバックを与え、そのフィードバックを基に画像を動的に更新する仕組みを導入した。この設計により、ユーザがインタラクティブに関与しながら、生成結果を逐次改善することが可能となった。

物理空間では、生成 AI の性能を最大限に活用するために、DALL-E や Stable Diffusion などの高精度な画像生成エンジンを採用した。これにより、ユーザの期待に応える高品質な視覚化を実現した。また、メタバース内での表示には **cluster** や **Unity** といった高度な開発プラットフォームを活用し、視覚的な洗練さを備えたユーザーインターフェースを構築した[11][12]。

3.2.3 心理空間と物理空間の相互作用

心理空間と物理空間の間には密接な相互作用が存在する。心理空間で形成された抽象的なアイデアや目標は、物理空間で具体化され、ユーザが視覚的に確認可能な形で提示される。一方で、物理空間で生成された結果やフィードバックは心理空間に戻され、次のデザインプロセスの改善に活用される。この循環的なプロセスにより、設計の創造性と実現可能性が向上し、システム全体の完成度を高めることができる。

3.2.4 知識空間の活用

本研究において、知識空間は、思考空間で行われる推論や設計活動を支える重要な基盤として機能した。知識空間には、デザインの理論的背景、ユーザデータ、生成 AI モデルのトレーニングデータなどが蓄積されており、これらの情報が設計プロセスに必要な知識を提供した。この知識空間は、思考空間における心理的および物理的プロセスをサポートし、設計の信頼性を高める役割を果たしている。

3.3 システムの構造設計

メタバース環境におけるコミュニケーション課題を解決するために、生成 AI 技術を活用した視覚化システムを構築した。本章では、このシステムの構造設計について述べる。システムは音声処理からキーワード抽出、プロンプト生成、画像生成、メタバース内でのリアルタイム表示

までを統合的に設計している。

3.3.1 音声処理モジュール

音声処理モジュールは、ユーザが発話した内容を高精度でテキストデータに変換する役割を担う。このモジュールでは、**whisper** や **Google Speech-to-Text API** などの最新の音声認識技術を採用し、多言語対応を実現した。また、背景ノイズの影響を最小化するために、ノイズキャンセリング技術を組み込み、認識精度を向上させている。

音声データがテキスト化された後、そのデータは次のキーワード抽出モジュールに送られ、さらなる処理が行われる。これにより、ユーザが自然な形で発話した内容が、システム全体で活用可能な形式に変換される。

3.3.2 キーワード抽出アルゴリズム

キーワード抽出アルゴリズムは、音声処理モジュールから得られたテキストデータを解析し、主要な情報を抽出するプロセスを担当する。このアルゴリズムには、自然言語処理 (NLP) 技術が使用されている。

入力されたテキストを形態素解析し、文脈を考慮した上で重要なキーワードを抽出する。このプロセスでは、ユーザの発話意図を正確に捉え、次のプロンプト生成モジュールで活用できるように構造化データを生成する。また、同義語の統一や曖昧表現の解消にも対応し、キーワード抽出の精度を向上させている。

3.3.3 プロンプト生成の仕組み

プロンプト生成の仕組みでは、抽出されたキーワードを基に、生成 AI エンジンで使用可能なプロンプトを作成する。このプロンプトは、ユーザの意図を反映しながら、抽象的な概念を具体的な視覚情報として表現するための指示文として機能する。

生成プロセスでは、キーワードを適切に組み合わせ、画像生成エンジンが要求するフォーマットに整形する。例えば、「椅子」や「快適さ」といったキーワードを「**A comfortable chair with ergonomic design**」といった具体的な表現に変換する。この仕組みにより、生成 AI が高品質な視覚化結果を提供できるようにしている。

3.3.4 画像生成エンジンの選定

画像生成エンジンには、最新の生成 AI モデルである DALL-E や Stable Diffusion を採用する。このエンジンは、プロンプトに基づいて高品質な画像を生成する能力を持ち、リアルタイム処理にも対応可能である。また、カスタマイズ可能な出力設定を備え、ユーザーのニーズに応じて画像のスタイルや詳細を調整できる。

3.3.5 メタバース表示機能の設計

メタバース環境におけるリアルタイムな視覚情報の共有を実現するために、本研究では軽量化を最優先事項としてメタバース表示機能を設計した。生成された画像を素早くユーザーに提示することを重視し、遅延の最小化とパフォーマンスの最適化を目指した。

3.4 システムの実装環境

本研究で提案する視覚化システムを構築するにあたり、実装環境として使用した開発ツールやプラットフォーム、採用した技術とフレームワーク、さらにシステム全体の動作フローについて詳述する。

3.4.1 開発ツールとプラットフォーム

システム開発には、以下の開発ツールとプラットフォームを使用した。

(1) プログラミング環境

開発言語としては、システム全体の柔軟性と拡張性を考慮し、**Python** を主に採用した。**Python** は、自然言語処理や画像生成において多くのライブラリが利用可能であり、AI 関連の開発に最適である。

(2) 開発環境

開発環境には **PyCharm** を使用し、コードの編集、デバッグ、およびプロジェクトの構成管理を効率的に行った。**PyCharm** は、**Python** 開発に特化した高度な機能を提供しており、効率的なコーディングをサポートする。

(3) メタバースプラットフォーム

本研究では、生成された画像をメタバース内でリアルタイムに表示する機能を実装するために、**Unity**、**VRChat**、**Cluster** の 3 つのプラットフォームを活用した。それぞれのプラットフォームは異なる特徴を持ち、本研究の目的に応じて適切に選定された。

① Unity

Unity は、3D 環境を構築するための高い柔軟性と豊富なプラグインを備えた開発環境である。本研究では、カスタマイズ性の高さとリアルタイムレンダリング機能を活かし、メタバース環境の基盤を構築するために採用した。

② Cluster

Cluster は、VR 空間内での画像共有機能が比較的容易に実装可能であるため、メタバース環境内でのコンテンツ共有を目的として採用した。**Cluster** は、クラウドベースのプラットフォームであり、PC、スマートフォン、VR デバイスなど、さまざまなデバイスでアクセス可能である。そのため、本研究のシステムを幅広いユーザに適用できる利点がある。

③ VRChat

VRChat は、ワールド内のカスタマイズ性が高く、動的コンテンツの配置が可能であることから、インタラクティブな要素を加えるために活用した。**VRChat** は、ユーザが自由にカスタマイズした 3D ワールドを構築できるため、本研究の目的に適していると考えられる。

(4) クラウド環境

計算リソースの最適化を図るため、**Google Cloud Platform (GCP)** を使用し、生成 AI モデルの実行とデータの保存を行った。特に、GPU を活用した高速処理が可能である点を重視した。

3.4.2 使用技術とフレームワーク

本システムの実装において、以下の技術とフレームワークを採用した。

(1) 自然言語処理 (NLP)

本システムでは、ユーザの発話内容を解析し、適切なキーワードを抽出するために GPT-4 を利用した。GPT-4 は、大規模な言語モデルを基盤とし、入力されたテキストの文脈と意図を正確に把握することが可能である。キーワードの抽出に加え、画像生成のためのプロンプトを適切に整形する役割も担っている。また、多言語対応が可能であり、異なる言語の入力に対して翻訳や文法補正を行う機能を備えている。

(2) 画像生成技術

本システムでは、視覚情報の生成に OpenAI の DALL-E および Stable Diffusion を使用した。DALL-E は、与えられたプロンプトから高品質な画像を生成し、複雑な概念を視覚的に表現することができる。Stable Diffusion は、オープンソースの画像生成技術であり、カスタマイズ性に優れ、柔軟な制御が可能である。

本研究では、DALL-E と Stable Diffusion それぞれの良さを考慮して、研究を進めた。DALL-E の方ではより発散的な場面で使用し、Stable Diffusion の方では SEED 値が固定できるため、文脈に沿った画像生成が実現できる。

(3) 音声認識技術

ユーザの発話内容をテキストに変換するため、音声認識技術として Whisper を採用した。Whisper は、高精度な音声認識機能を持ち、多言語対応が可能である。本システムでは、ノイズの影響を抑えるためにノイズキャンセリング技術を組み込み、認識精度を向上させている。さらに、音声データの自動翻訳機能を活用し、異なる言語間のコミュニケーションを円滑化することを目指している。

3.4.3 システムの動作フロー

本システムの動作フローは以下の手順で構成される。

(1) 音声入力

ユーザが発話した内容が音声処理モジュールでテキストデータに変換される。

(2) キーワード抽出

音声認識後のテキストデータが自然言語処理モジュールで解析され、重要なキーワードが抽出される。

(3) プロンプト生成

抽出されたキーワードを基に、生成 AI エンジン用のプロンプトが生成される。このプロンプトは、ユーザの意図を視覚化するための具体的な指示文となる。

(4) 画像生成

生成 AI エンジンがプロンプトに基づいて画像を生成する。生成された画像は、抽象的な概念やニュアンスを具体的に表現する。

(5) メタバース内での表示

生成された画像がリアルタイムで Unity を介してメタバース環境内に表示される。ユーザは結果を確認し、必要に応じてフィードバックを提供する。

(6) フィードバックの反映

ユーザのフィードバックがシステムに反映され、画像生成プロセスが最適化される。この循環的なプロセスにより、ユーザ体験が向上する。

第4章 実験 1: システムの視覚化能力の 検証

表 4-1 実験 1: システムの視覚化能力の検証の概要

目的	● 提案したシステムの性能を検証する。 ● メタバース環境におけるコミュニケーションの支援能力を評価する。 ● 改善に向けた意見を収集する。
対象	● 被験者 11 名
日時／場所	● 2024 年 8 月 ● 慶応義塾大学 日吉キャリアパス 協生館
実施内容	● 提案した視覚化システムがユーザの言語的表現を基に抽象的な概念を視覚化し、その生成結果がユーザの意図や期待に一致するかを評価することを目的に、実際のユーザ入力と反復的な生成プロセスを通じた検証を行った。
評価項目	● < 定量評価 > ● 質問 1:最初に生成された画像が被験者のイメージにどれだけ近い か(5 段階評価で測定し、平均スコアを算出) ● 質問 2:最終的に生成された画像が被験者のイメージにどれだけ近 づいたか(5 段階評価で測定し、改善度を分析) ● < 定性評価 > ● 被験者からの主観的な満足度の収集(アンケートの自由回答欄を 活用) ● 画像生成プロセス全体の使いやすさに関する意見の聴取
倫理審査承認 番号	

4.1 事前準備

4.1.1 被験者の選定

実験では、異なる年齢層を持つ計 11 名の被験者を選定した。

あなたの年齢を教えてください。

11 件の回答

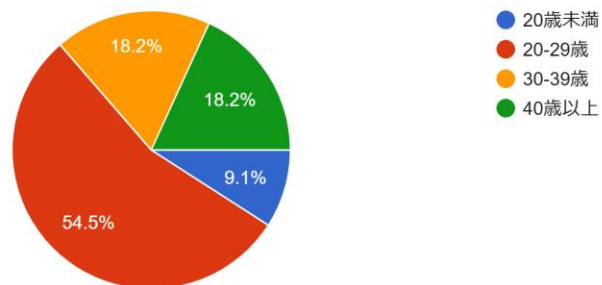


図 3 「実験 1: システムの視覚化能力の検証」の被験者各年齢層割合

性別を教えてください。

11 件の回答

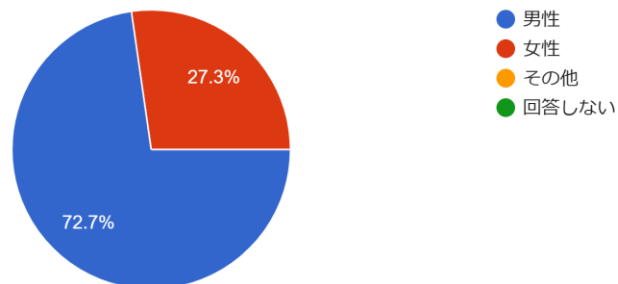


図 4 「実験 1: システムの視覚化能力の検証」の被験者性別割合

4.1.2 実験環境の準備

実験環境として個室を準備し、参加者が集中して実験に取り組めるよう配慮した。実験はオンラインメタバース環境で実施される機材と環境を整える。

4.1.2.1 機材の準備

- HMD(ヘッドマウントディスプレイ): Oculus Quest 2 を使用し、被験者がリアルタイムでメタバース内の視覚化結果を確認できるようにした。



図 5 実験中使用する HMD の例

- PC: 高性能 PC を準備し、生成 AI システムの動作やリアルタイム処理をスムーズに行える環境を構築する。

4.1.2.2 ソフトウェア環境

- 音声認識ツール: Google Speech-to-Text API の採用可能。
- 生成 AI エンジン: DALL-E モデルの使用可能。
- プロンプト生成: ChatGPT 4o の使用可能。

4.1.2.3 個室環境

- 被験者のプライバシーを保護し、集中を妨げない環境を提供するため、防音性の高い個室を使用。



図 6 「実験 1: システムの視覚化能力の検証」を行う個室環境

4.2 実験内容の実施

「実験 1: システムの視覚化能力の検証」は提案した視覚化システムを使用し、被験者が抽象的テーマについて説明を行い、その言語表現を基にシステムが画像を生成するプロセスを実施した。

4.2.1 ステップ1:テーマの提示

被験者に抽象的テーマ(例:「座り心地のいい椅子」)を提示し、そのテーマに基づく具体的なイメージを思い浮かべてもらう。

4.2.2 ステップ2:説明と初期画像生成

被験者は自分のイメージを言語で説明し、その音声データを基にシステムが初期の生成画像を提示する。



図 7 「実験 1: システムの視覚化能力の検証」を行う初期生成の様子

4.2.3 ステップ3:フィードバックと反復的生成

被験者は生成された初期画像を確認し、システムに対して改善点や修正内容をフィードバック。そのフィードバックに基づき、システムが画像を再生成するプロセスを複数回繰り返す。



図 8 「実験 1: システムの視覚化能力の検証」を行う再度生成の様子

4.2.4 ステップ4:最終画像の評価

最終的に生成された画像が被験者のイメージにどれだけ近づいたかをアンケートで評価する。また、初期画像から最終画像に至るまでのプロセスにおける改善度を測定する。

4.2.5 ステップ5:使用体験の記録

実験終了後、被験者に対してシステムの使いやすさ、生成結果の精度、改善プロセスの有効性についてアンケートを実施し、主観的な評価を記録した。

4.3 アンケートの実施

「実験 1: システムの視覚化能力の検証」終了後に、被験者に対してアンケートを実施し、提案システムの視覚化能力と使用体験に関する評価を収集した。

4.3.1 質問1:生成された最初の画像は、あなたがイメージ

していた物や概念に近いものでしたか？

(1) 質問目的:

システムが生成する初期の視覚化結果が、被験者の言語表現や概念にどれだけ近づいているかを評価することである。この評価により、システムの初期段階での視覚化精度を確認し、改善の必要性を把握することを目指す。

(2) 質問結果:

図 9 に示したアンケートの結果によると、「非常にそう思う」と回答した被験者が 9.1%となり、「そう思う」が 18.2%を占めた一方で、「どちらとも言えない」が 36.4%、「あまりそう思わない」が 27.3%、「全くそう思わない」が 9.1%であった。

(3) 質問結果分析:

この結果から、初期段階での視覚化は一部の被験者にとってイメージに近いものの、大半の被験者にとって完全に満足できる結果ではなかったことが示唆される。特に、「どちらとも言えない」や否定的な評価を示した回答が多いことから初期画像生成アルゴリズムの改善が必要であると考えられる。今後、プロンプト生成やモデル調整の最適化を通じて、初期段階での精度向上が求められる。

生成された最初の画像は、あなたがイメージしていた物や概念に近いものでしたか？

11 件の回答

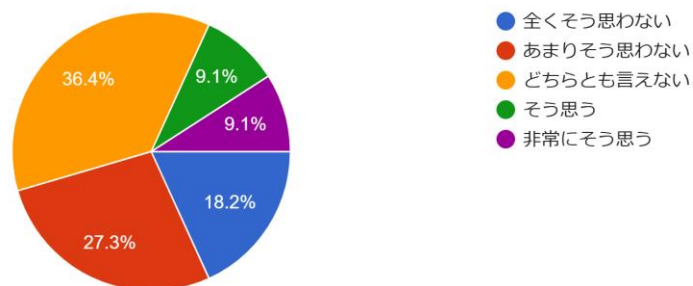


図 9 「実験 1: システムの視覚化能力の検証」の質問 1

4.3.2 質問2:最終的に生成された画像は、あなたがイメー

ジしていた物や概念に非常に近いものでしたか？

(1) 質問目的:

被験者が与えたフィードバックを通じて改善された最終生成画像が、被験者のイメージや概念にどれだけ一致しているかを評価することである。この評価を通じて、システムがユーザーフィードバックを反映して視覚化精度を向上させる能力を検証することを目指す。

(2) 質問結果:

図 10 に示したアンケートの結果によると、「非常にそう思う」と回答した被験者が 36.4%となり、「そう思う」が 36.4%を占めた。また、「どちらとも言えない」が 18.2%で、「あまりそう思わない」が 9.1%、「全くそう思わない」の回答はなかった。

(3) 質問結果分析:

この結果からは「非常にそう思う」と「そう思う」の合計として最終的に生成された画像が被験者のイメージや概念にかなり近いと感じた被験者が 72.8%であることが分かる。システムがフィードバックを通じて改善を行い、視覚化精度を向上させる能力を有していることを示している。一方で、「どちらとも言えない」や否定的な評価を示す回答も約 27.3%存在することから、一部のケースではフィードバックを十分に反映できていない可能性があると考ええる。

最終的に生成された画像は、あなたがイメージしていた物や概念に非常に近いものでしたか？
11 件の回答

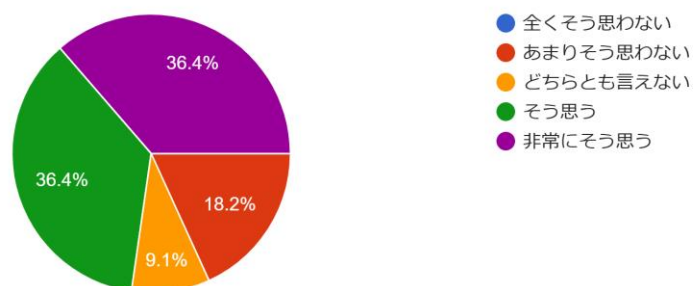


図 10 「実験 1: システムの視覚化能力の検証」の質問 2

4.3.3 質問3:各回のフィードバックを経て、生成画像が改

善されたと感じましたか？

(1) 質問目的:

システムが被験者からのフィードバックを効果的に活用し、生成された画像を段階的に改善する能力を検証することである。この評価を通じて、システムのインタラクティブ性および改善プロセスの有効性を確認することを目指す。

(2) 質問結果:

図 11 に示したアンケートの結果によると、「非常にそう思う」と回答した被験者が 45.5%、「そう思う」と回答した被験者が 27.3%を占めた。「どちらとも言えない」が 18.2%、「あまりそう思わない」が 9.1%であり、「全くそう思わない」の回答はなかった。

(3) 質問結果分析:

こうした結果からはプラス的に反映する「非常にそう思う」と「そう思う」の被験者が約 72.8%となり、フィードバックを経て生成画像が改善されたと感じていることが分かる。これは、システムがユーザのフィードバックを的確に反映し、画像生成プロセスを改善する能力を示している。一方で、「どちらとも言えない」や否定的な評価を示す回答が 27.3%存在することから、特定のケースではフィードバック処理が十分に機能していない可能性がある。

各回のフィードバックを経て、生成画像が改善されたと感じましたか？
11 件の回答

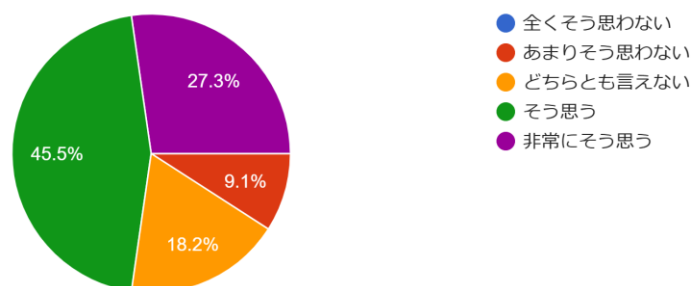


図 11 「実験 1: システムの視覚化能力の検証」の質問 3

4.3.4 質問4:システムが画像を生成するまでの時間に満

足しましたか？

(1) 質問目的:

提案システムの画像生成プロセスの時間的効率に対する被験者の満足度を評価することである。さらに、本システムが被験者の期待に応えられているかを確認することを目指す。

(2) 質問結果:

図 12 に示したアンケートの結果によると、「非常に満足している」と回答した被験者が 27.3%、「満足している」が 27.3%を占めたため、全体の 54.6%が肯定的な評価を示した。「どちらとも言えない」と回答した被験者は 27.3%であり、「あまり満足していない」が 18.2%、「全く満足していない」の回答はなかった。

(3) 質問結果分析:

こうした結果からは、54.6%の被験者が生成時間に満足していることが確認された。半数以上であることにより、システムがリアルタイムでの画像生成において一定の効果を発揮していると評価できる。一方で、「あまり満足していない」という否定的な回答が 18.2%あり、「どちらとも言えない」の回答も 27.3%に達していることから、生成時間に関する改善の余地があることも示唆される。

特に否定的な評価を示した被験者が存在する背景には、生成プロセスが遅延する場面が一部で発生した可能性が考えられる。そのため、システムの処理速度を向上させるとともに、生成中の状況をユーザに可視化するなど、待ち時間の心理的負担を軽減する工夫が求められると考えられる。

システムが画像を生成するまでの時間に満足しましたか？

11 件の回答

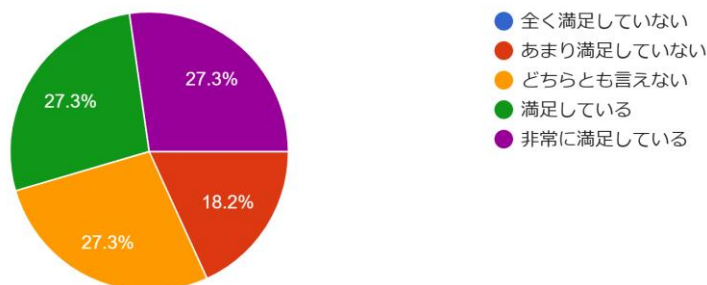


図 12 「実験 1: システムの視覚化能力の検証」の質問 4

4.3.5 質問5:システムの使いやすさについてどう感じましたか？

(1) 質問目的:

提案システムの操作性およびユーザビリティを評価し、システムが利用者にとってどの程度直感的かつ使いやすい設計となっているかを確認することである。被験者が快適に利用できたかどうかを測定することを目指す。

(2) 質問結果:

図 13 に示したアンケートの結果によると、「非常に使いやすい」と回答した被験者が 18.2%、「使いやすい」が 45.5%を占めし、全体の 63.7%が肯定的な評価を示した。その一方で、「どちらとも言えない」が 27.3%であり、「使いにくい」と回答した被験者が 9.1%であり、「非常に使いにくい」の回答はなかった。

(3) 質問結果分析:

こうした結果から 63.7%の被験者が提案システムの使いやすさを肯定的に評価していることが分かった。このことは、システムがユーザビリティを重視した設計に成功していることを示していると考えられる。「どちらとも言えない」と「使いにくい」との回答を合わせた 36.4%の被験者が、システムの操作性において一定の課題を感じた可能性がある。否定的および中立的な評価を示した背景として、システムのインターフェースが一部の被験者にとって直感的ではなかった可能性や、操作に必要な学習コストが想定よりも高かったことが挙げられる。

システムの使いやすさについてどう感じましたか？

11 件の回答

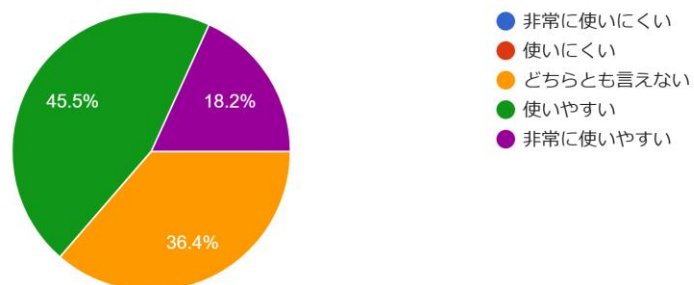


図 13 「実験 1: システムの視覚化能力の検証」の質問 5

第5章 実験 2: システムによる概念具現化の支援能力の妥当性確認

表 5-1 実験 2: システムによる概念具現化の支援能力の妥当性確認

目的	● 提案したシステムが異なる背景を持つユーザ間での相互理解をどの程度促進できるかを評価する。 ● 抽象的概念の共有と具現化を支援するシステムの有効性を検証する。 ● 改善に向けた意見を収集する。
対象	被験者 11 名
日時 / 場所	● 2024 年 8 月 ● 慶應義塾大学 日吉キャリアパス 協生館
実施内容	提案した視覚化システムを使用し、被験者と実験者がペアとなり、指定された抽象的テーマについて対話を行いながら実験を進めた。システムは対話中に得られた言語表現を基に視覚化を行い、生成された画像を被験者が参照しながら対話を進めた。被験者から提供されたフィードバックをもとに画像生成プロセスを複数回繰り返し、生成結果を改善することを目指した。また、システムの使用体験と、コミュニケーションの円滑化への効果についても評価を行った。
評価項目	< 定量評価 > ● 質問 1: 初期生成画像が被験者ペアまたは個人の共通イメージにどれだけ近いのか(5 段階評価で測定し、平均スコアを算出) ● 質問 2: 最終生成画像が被験者ペアまたは個人の共通イメージにどれだけ近づいたか(5 段階評価で測定し、改善度を分析) < 定性評価 > ● 被験者からの主観的な満足度の収集(アンケートの自由回答欄を活用) ● 画像生成プロセス全体の使いやすさに関する意見の聴取
倫理審査承認番号	

5.1 事前準備

5.1.1 テーマの設定

被験者と実験者の相互理解を評価する際に一定の安定性を保つため、テーマを「最近買いたい椅子」に設定した。このテーマの設定には三つの理由がある。

(1) 身近で具体的な対象

椅子という身近な物をテーマとすることで、短時間で被験者が容易に理解し、自身のイメージを形成できるよう配慮した。

(2) 多様な視点を引き出す可能性

異なる背景を持つ被験者それぞれが椅子に対する異なるニーズや価値観を持っているため、発散的な対話が期待できる。

(3) 具体性と抽象性の両立

椅子は具体的な対象である一方で、「デザイン」「快適性」「素材」など抽象的な要素についても議論が可能であり、システムの視覚化能力を十分に評価できる。

5.1.2 被験者の選定

実験では、異なるバックランドを持つ計 11 名の被験者を選定した。

出身地または国籍を教えてください。

11 件の回答

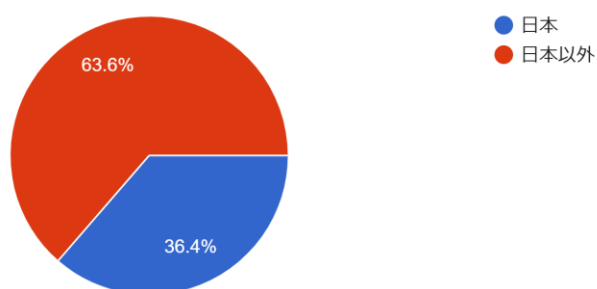


図 14 「実験 2: システムによる概念具現化の支援能力の妥当性確認」の各被験者の国別割合

最終学歴を選んでください。

11 件の回答

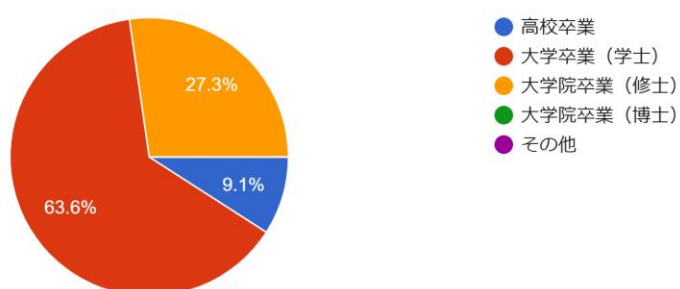


図 15 「実験 2: システムによる概念具現化の支援能力の妥当性確認」の各被験者学歴割合

5.1.3 実験環境の準備

実験は、個室環境およびオンラインメタバース環境で行うため、以下の環境を整えた。

5.1.3.1 機材の準備

HMD (Oculus Quest 2)

PC: 高性能な PC を用意し、生成 AI エンジンの円滑な動作を確保した。

5.1.3.2 ソフトウェア環境

音声認識ツール: Google Speech-to-Text API の採用可能。

生成 AI エンジン: DALL-E モデルの使用可能。

プロンプト生成: ChatGPT 4o の使用可能。

5.1.3.3 個室環境

実験用に防音個室(例;図)を準備し、被験者が集中して対話や視覚化を行える環境を提供した。



図 16 「実験 2: システムによる概念具現化の支援能力の妥当性確認」実験環境

5.2 実験内容の実施

「実験 2: システムによる概念具現化の支援能力の妥当性確認」では提案したシステムが被

験者と実験者のコミュニケーション中に実用し、概念具現化の支援能力をどの程度発揮できるかを評価する。

5.2.1 ステップ 1: 対話プロセスの設定

被験者と実験者が1組のペアを形成し、コミュニケーションを行う。対話のテーマは「最近買いたい椅子」として選定した。このテーマ設定により、短時間でイメージを具体化しやすくすると同時に、被験者ごとに異なるニーズや好みを引き出すことを目的とした。

5.2.2 ステップ 2: 視覚化とフィードバックのプロセス

被験者と実験者は対話を通じて、それぞれのイメージや考えを伝え合う。この過程で、システムが被験者の発話内容を処理し、生成された画像をリアルタイムで表示する。生成画像は対話に応じて次々と更新され、被験者と実験者は画像を活用して自身の考えを相手に伝えつつ、自分の脳内イメージを具体化していく。

5.2.3 ステップ 3: 生成プロセスの反復

対話のプロセスは1人につき3回繰り返し、発言内容に基づく画像生成の改善を図る。

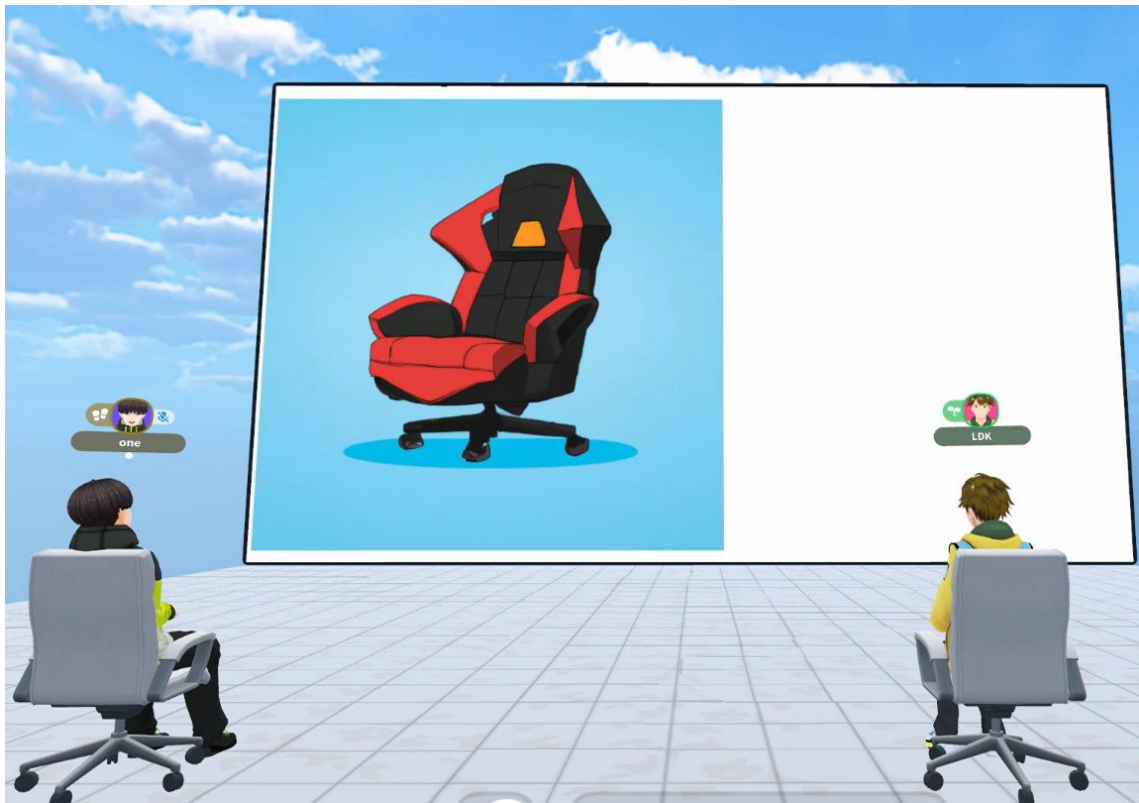


図 17 「実験 2: システムによる概念具現化の支援能力の妥当性確認」の初期画像生成の様子



図 18 「実験 2: システムによる概念具現化の支援能力の妥当性確認」の再度画像生成の様子

5.2.4 ステップ 3: 結果の収集

実験終了後、生成された画像のイメージ一致度、生成プロセスの改善効果、システムの生成時間に対する満足度、およびシステムの使いやすさに関する項目を含めて評価を行った。

5.3 アンケートの実施

「実験 2: システムによる概念具現化の支援能力の妥当性確認」終了後に、被験者に対してアンケートを実施して、提案システムの視覚化能力と使用体験に関する評価の収集を行った。

5.3.1 質問 1: システムによって生成された画像は、相手

の考えを理解する助けになりましたか？

(1) 質問目的:

生成された画像が相手の考えや意図を理解する上でどの程度有効であるかを測定する。提案システムがコミュニケーションを円滑化するためにどれだけ寄与しているかを明らかにすることを目指す。

(2) 質問結果:

図 19 に示したアンケートの結果によると、「非常にそう思う」が 27.3%で、「そう思う」が 54.5%を占め、全体の 81.8%が肯定的な評価を示した。一方、「どちらとも言えない」が 18.2%であり、さらに、「あまりそう思わない」や「全くそう思わない」との回答はなかった。

(3) 質問結果分析:

この回答結果から、生成された画像が相手の考えを理解する助けとなる点について、被験者の大多数が肯定的な評価を示していることが分かった。提案システムが相手の意図や考え

を視覚的に補完し、コミュニケーションを支援する能力を有していることを示している。まだ、一部の被験者が「どちらとも言えない」と回答しており、これは生成画像が完全に被験者の期待を満たしていないことも示唆している。

システムによって生成された画像は、相手の考えを理解する助けになりましたか？
11 件の回答

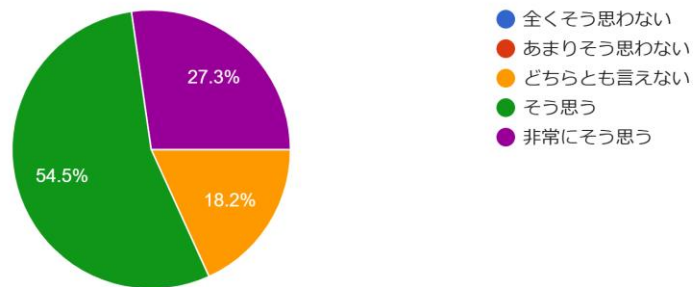


図 19 「実験 2: システムによる概念具現化の支援能力の妥当性確認」質問 1

5.3.2 質問 2: システムを使用することで、自分の考えをより具体的に表現できましたか？

(1) 質問目的:

システムを活用することで被験者が自分の考えやアイデアを具体的に表現する能力が向上したかどうかを評価することである。システムの支援機能がユーザの自己表現力にどの程度寄与するかを明確にすることを目指す。

(2) 質問結果:

図 20 に示したアンケートの結果によると、「非常にそう思う」と「そう思う」がそれぞれ 45.5% を占め、全体の 91% が肯定的な評価を示した。一方、「あまりそう思わない」が 9.1% であり、「全くそう思わない」や「どちらとも言えない」と回答した被験者はいなかった。

(3) 質問結果分析:

結果から、提案システムが被験者の考えを具体的に表現する支援機能として高い評価を得ていることが分かった。この結果は生成された視覚情報がユーザの抽象的なアイデアを具体

化し、他者に共有するプロセスを補完していることを示している。

システムを使用することで、自分の考えをより具体的に表現できましたか？
11 件の回答

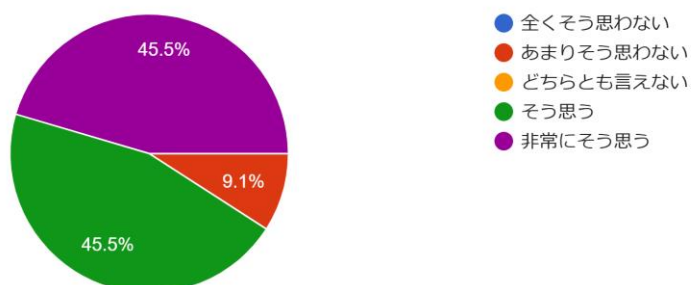


図 20 「実験 2: システムによる概念具現化の支援能力の妥当性確認」 質問 2

5.3.3 質問 3: 発言ごとに生成された画像が、相手のイメージと一致していくと感じましたか？

(1) 質問目的:

提案システムが対話中に生成する画像が、発言ごとに相手のイメージと徐々に一致していく効果をどの程度発揮したかを評価することである。生成画像が相互理解の促進にどれほど寄与したかを確認することが重要である。

(2) 質問結果:

図 21 に示したアンケートの結果によると、「非常にそう思う」が 36.4%、「そう思う」が 27.3% を占め、全体の 63.7% が肯定的な回答を示した。「どちらとも言えない」が 27.3%、「あまりそう思わない」が 9.1% となっている。

(3) 質問結果分析:

こうした結果から、提案システムが相互理解を促進する効果が一定程度評価されていることが分かる。肯定的な回答が全体の 6 割以上であり、多くの被験者により生成した画像が発言を重ねるごとにイメージと一致していくと感じていることを示唆している。一方で、「どちらとも言え

ない」と回答した割合も一定数存在しており、画像生成プロセスの精度向上や改善点の特定が今後の課題として挙げられる。

各発言ごとに生成された画像が、相手のイメージと一致していくと感じましたか？
11 件の回答

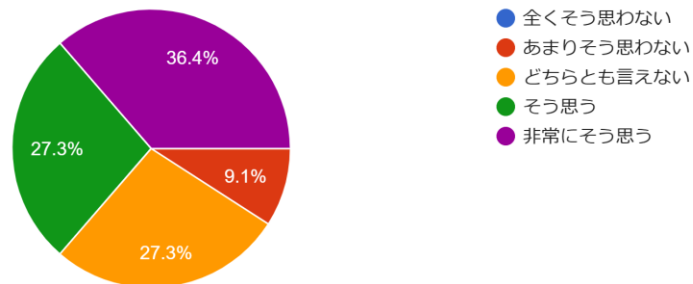


図 21 「実験 2: システムによる概念具現化の支援能力の妥当性確認」質問 3

5.3.4 質問 4: システムが画像を生成するまでの時間に満足しましたか？

(1) 質問目的:

提案システムの画像生成プロセスにおける待ち時間が、ユーザの満足度にどのように影響を与えているかを評価することである。

(2) 質問結果:

図 22 に示したアンケートの結果によると、「非常に満足している」が 27.3%、「満足している」が 36.4% で、全体の 63.7% が肯定的な回答を示した。「どちらとも言えない」は 27.3% であり、「あまり満足していない」は 9.1% だった。

(3) 質問結果分析:

多くの被験者が画像生成の速度に対して肯定的な評価を示していることがわかる。肯定的な回答が全体の 6 割を超えており、待ち時間が対話における大きな障害とはなっていないことが示唆される。一方で、「どちらとも言えない」や「満足していない」と回答した被験者も一定数

存在しており、画像生成時間が特定のシナリオにおいてユーザ体験を阻害している可能性が考えられる。

システムが画像を生成するまでの時間に満足しましたか？

11 件の回答

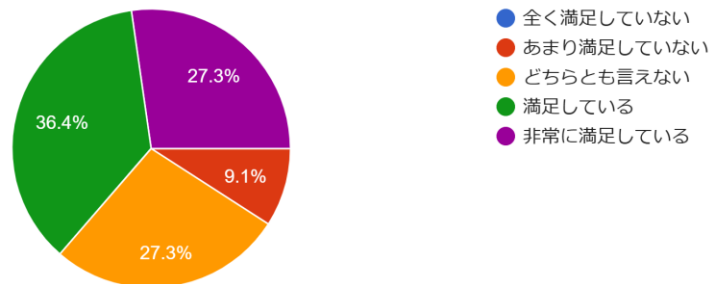


図 22 「実験 2: システムによる概念具現化の支援能力の妥当性確認」質問 4

5.3.5 質問 5: システムが対話の質を向上させるのに役立つ

ったと感じましたか？

(1) 質問目的:

提案システムが対話を通じて相互理解を深め、コミュニケーションの質を向上させる効果を持っているかどうかを評価することである。

(2) 質問結果:

図 23 に示したアンケートの結果によると、「非常にそう思う」が 45.5%、「そう思う」も 45.5% であり、合わせて 91% の被験者ではシステムが対話の質を向上させるのに役立ったと評価している。一方で、「どちらとも言えない」が 9.1% であり、「あまりそう思わない」や「全くそう思わない」という回答はなかった。

(3) 質問結果分析:

こうした結果から、多くの被験者がシステムの使用により対話の質が向上したと感じていることが明確となっている。視覚化による情報補完がコミュニケーションを支援し、より円滑な対話の実現に寄与していると考えられる。一方で、一部の被験者が中立的な評価を示した理由とし

て、テーマやシステムの特徴により支援効果が限定的だった可能性がある。

システムが対話の質を向上させるのに役立ったと感じましたか？

11 件の回答

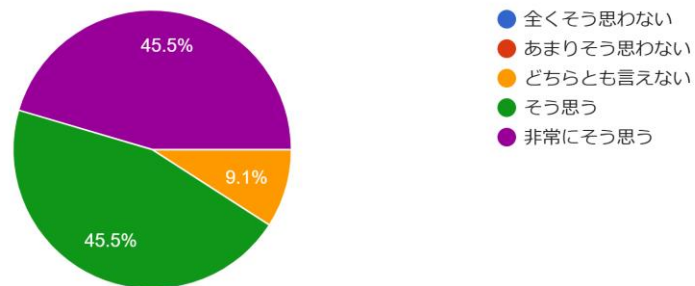


図 23 「実験 2: システムによる概念具現化の支援能力の妥当性確認」質問 5

第6章 総合的な考察

6.1 実験 1 と実験 2 の結果統合

本研究では提案したシステムの有効性を評価するために、「実験 1: 視覚化能力の検証」と「実験 2: 概念具現化支援能力の妥当性確認」の二つの実験を実施した。二つの実験から重要な知見を得ることができた。

実験 1 では、被験者が提示した抽象的なテーマをシステムがどの程度視覚化できるかを検証した結果、生成画像の初期精度には改善の余地がある一方、ユーザのフィードバックを反映することで視覚化精度が向上することが確認された。このような結果に導いていくのが、生成 AI モデルの特性と考えられる。生成 AI モデルの初期段階において、プロンプトが一般的に不完全であるため、生成がプロンプト内の曖昧な表現を解釈しきれず、ユーザの意図に完全に一致する画像を生成できなかった。生成 AI モデルは、学習データのバイアスや限界により、学習データ内に含まれる特定の概念を的確に視覚化する能力にもバラつきがあることと考えられる。その原因により、フィードバックループの導入により、ユーザが提供した追加情報が生成結果に直接反映されることで、視覚化精度が向上することになる。これにより、曖昧な表現が具体化され、生成 AI の内部モデルがユーザの意図に近い結果を出すことができた。さらに、プロンプトの影響を受けるため、ユーザが出したプロンプトが抽象的になるほど、生成 AI モデルが具体的なイメージを形成するのが困難となる。まだ、こうした抽象的なプロンプトでも、フィードバックを通じて具体性が追加されることで、システムの精度が改善する傾向が見られた。

実験 2 では、システムがユーザ間のコミュニケーションをどの程度促進するかを評価した。この実験で、視覚化による補助がユーザ間の相互理解を高めることが示された。このような結果になるのが、視覚情報の優位性だと考えられる。言語のみのコミュニケーションでは、文化的背景や専門知識の違いから誤解が生じる可能性が高い。視覚化された情報はより具体性と直感的な理解をもたらすため、こうしたギャップを埋める効果を発揮した。そのため、被験者と実験者がシステムを通じて、同時に同じ生成された画像を見られることで、対話の過程での相互補正が容易になり、最終的な合意形成がスムーズに行われる。被験者と実験者は異なる背景を持つために、文化や背景的影響を受けやすい。視覚情報が共通の参照として機能することで、言語による説明が不十分であった場合でも誤解が大幅に減少した。

6.2 システムの全体的な有効性と限界

6.2.1 有効性の分析

実験の結果により、以下の三つの点から有効性が確認された。

(1) 抽象的概念の具体化

抽象的な概念はシステムを通じて、具体的な画像に変換され、言語では困難となる理解が容易になった。

(2) コミュニケーションの効率化

コミュニケーション中に視覚情報を提示することで、概念や物体の理解が加速し、意思疎通にかかる時間が短縮された。

(3) 動的プロセスの有効性

フィードバックループを採用した生成プロセスにより、ユーザが逐次的にシステムの出力を調整することで、最終的な生成物の精度を向上させることができた。

6.2.2 限界の分析

(1) 初期生成精度の課題

画像生成の質はプロンプト内容や学習データにより大きく影響されているため、プロンプトの表現および学習データのカバレッジは生成画像の質を左右する。こうした原因により、初期生成の画像が制限される。

(2) 生成時間の問題

処理速度が一部のユーザにとって不十分であったのは、使用した生成モデルの計算負荷が大きいためである。この課題により、生成時間と生成画像は負の相関性がある。

(3) ユーザ依存性の強さ

システムの効果がユーザのプロンプトの質やフィードバックの内容に強く依存している点も課題として挙げられる。ユーザの熟練度が低い場合、システムの潜在能力が十分に発揮されない可能性がある。

6.3 今後の課題

最後に、本研究の今後の展望について説明します。

本研究では、生成 AI を活用した視覚化システムがメタバース環境におけるコミュニケーションの支援に有効であることを示しましたが、さらなる精度向上と実用化に向けて、以下の三つの改善点に取り組んでいきます。

6.3.1 初期生成精度の向上

現在、本システムでは一般的な学習データを活用することで、生成される画像の多様性を確保していますが、その一方で、被験者の意図に完全に合致しない画像が生成されることがある。

そこで、個人の学習データを活用し、ユーザごとに適応した学習モデルを構築することで、よりユーザの表現に適した初期画像の生成を実現し、精度の向上を図る。これにより、ユーザが求めるビジュアルイメージとの乖離を減らし、より直感的で効果的な視覚支援が可能となる。

6.3.2 処理速度の向上

現状のシステムでは、生成プロセスに一定の時間がかかるため、リアルタイム性に課題がある。今後は、モデルの軽量化や最適化を進めることで、生成プロセスの処理時間を短縮し、スムーズな対話支援を可能にする。特に、クラウドベースのインフラを活用し、分散処理を導入することで、複数のプロセスを並列化し、即時の視覚化を実現することを目指す。また、生成エンジンとしての効率的なキャッシュ機構を導入し、再生成時の負荷を低減することで、リアルタイム対話での快適な使用を可能にする。

6.3.3 新システムの開発

実験の結果、視覚化技術がコミュニケーションの補助として有効であることが確認されたが、従来のシステムでは画像の品質や安定性に課題が存在していた。これらの課題を解決するために、新たに **Stable Diffusion** を採用し、**SEED** 値の固定による安定的な表現を実現した。

Stable Diffusion は、従来の画像生成モデルと比較して、高解像度で一貫性のある画像を生成できるため、視覚的な伝達の正確性が向上した。特に、**SEED** 値の固定を導入することで、同一のプロンプトから再現性の高い画像を生成できるようになり、実験の一貫性を確保することが可能となった。これにより、異なるユーザ間で同じ概念を共有する際に、以前よりも安定した視覚情報の提供が実現できた。

さらに、新システムの導入に伴い、「発散(**Divergence**)」と「安定(**Stability**)」のバランスを比較する実験を開発した。発散とは、ユーザの入力に対して多様な視覚的解釈を提供することであり、安定とは、同じ概念に対して一貫した画像を生成することである。この新たな実験では、発散型のアプローチと安定型のアプローチが、それぞれコミュニケーションのどの側面で有効に機能するのかを検証することを目的としている。

本システムの導入により、コミュニケーションの場面に応じて、より柔軟で効果的な視覚情報の提供が可能になることが期待される。今後は、さらなる実験を通じて、新システムの最適な適用条件を明確化し、メタバース環境における実用化へ向けた発展を図る。

第7章 結論

本研究では、生成 AI 技術を活用してメタバース環境におけるコミュニケーションの質を向上させるシステムを提案し、その有効性を検証した。以下に、各章での成果をまとめる。

7.1 研究の概要

第 1 章では、メタバース技術の発展とその可能性、コミュニケーションの課題について述べた。メタバース環境では、非言語的要素の欠如や言語の限界によるコミュニケーションのギャップが問題として挙げられた。本研究の目的は、生成 AI を活用して視覚的な情報を提供することで、これらの課題を解決することである。

第 2 章では、先行研究として生成 AI やメタバースの活用事例を分析し、本研究が提案するシステムの独自性を明確にしました。多空間デザインモデルを採用し、既存の研究との差別化を図っている。

7.2 システム設計と実装

第 3 章では、システム設計の詳細について述べました。特に、音声処理モジュール、キーワード抽出、プロンプト生成、画像生成、およびメタバース表示機能の統合設計を実現しました。これにより、ユーザが発話した内容をリアルタイムで視覚化する一連のプロセスが可能になった。また、システム実装には最新の AI 技術を採用し、Python、Unity、DALL-E、Stable Diffusion などのツールを活用した。これにより、生成画像の精度とリアルタイム性を両立することが可能となり、メタバース環境内でのユーザ体験を向上させた。

7.3 実験結果と考察

第 4 章および第 5 章では、2 つの実験を通じてシステムの有効性を検証した。実験 1 では、ユーザの言語表現を視覚化する能力を評価し、生成画像がユーザの意図に近づくことが確認される。実験 2 では、視覚化が相互理解を促進し、コミュニケーションの質を向上させる効果があることを示す。さらに、フィードバックを通じた生成画像の改善プロセスが効果的であることが明らかになり、ユーザビリティに関する肯定的な評価が得られた。一方で、初期生成の精度や

生成時間の最適化といった課題も浮き彫りになった。

参考文献

- [1]. Aharon, D. Y., Demir, E., & Siev, S. (2022). Real returns from unreal world? Market reaction to Metaverse disclosures. *Research in International Business and Finance*, 63, 101778. <https://doi.org/10.1016/j.ribaf.2022.101778>
- [2]. Grand View Research. (2022). *Metaverse market size, share & trends analysis report by component (hardware, software), by technology (blockchain, VR & AR, mixed reality), by application, by region, and segment forecasts, 2022 - 2030*. Retrieved from <https://www.grandviewresearch.com/>
- [3]. Osler, C., & Peters, R. (2020). Virtual Collaboration and Communication: Enhancing Group Interaction in Digital Spaces. *Virtual Collaboration Research Journal*, 12(3),
- [4]. Mehta, A. and Mehta, N. (2018). Knowledge Integration and Team Effectiveness: A Team Goal Orientation Approach. *Decision Sciences*, 49(3),
- [5]. Knippenberg, D., De Dreu, C. K., & Homan, A. C., 2004. *Journal of Applied Psychology*
- [6]. Forsyth, D. R. (2021). *Group dynamics*. Belmont, CA: Cengage Learning.
- [7]. Ning, H., Wang, H., Lin, Y., Wang, W., Dhelim, S., Farha, F., Ding, J., & Daneshmand, M. (2021). A Survey on Metaverse: The State-of-the-art, Technologies, Applications, and Challenges. *Journal of Educational Evaluation for Health Professions*, 18(32).
- [8]. Schulz-Hardt, S., Brodbeck, F., Mojzisch, A., Kerschreiter, R., & Frey, D. (2006). Group decision making in hidden profile situations: Dissent as a facilitator for decision quality. *Journal of Personality and Social Psychology*,
- [9]. Liu, V., & Chilton, L. B. (2022). Design guidelines for prompt engineering text-to-image generative models. *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '22)*, April 29–May 5, 2022, New Orleans, LA, USA. ACM, New York, NY, USA. <https://doi.org/10.1145/3491102.3501825>
- [10]. Hirao, A., Arita, M., Kim, Y., & Matsuoka, Y. (2017). A systematic analysis of seating research knowledge based on Multispace design model. *Transactions of the JSME (in Japanese)*, 83(853). <https://doi.org/10.1299/transjsme.17-00090>
- [11]. Tsuchiya, H., Hashimoto, S., Zeng, J., Sakato, T., & Nakano, Y. (2022). Motivational interviewing conversational agents using LLM. *Seikei*

University Faculty of Science and Technology. Presented at the 2022 International Conference on Multimodal Interaction (ICMI '22).

- [12]. Matsuoka, Y. (2013). *Method design thinking on Multispace*. Kindai-Kagaku-Sha.
- [13]. Frank Fischer, Johannes Bruhn, Cornelia Gräsel, Heinz Mandl. Fostering collaborative knowledge construction with visualization tools. *Learning and Instruction*, 2002, 12(2), pp.213-232. 10.1016/S09594752(01)00005-6. hal-00190645
- [14]. Ohira, K., Takahashi, S., & Ueda, K. (2000). EVIDII: An environment for visualizing differences of individual impressions for knowledge sharing. *Information Processing Society of Japan Research Report*, 2000(3), 15-22.

謝辞

本研究の実施および論文の執筆にあたり、多くの方々からご支援をいただきました。

まず、慶應義塾大学大学院システムデザイン・マネジメント研究科の教授小木 哲朗先生には、主査として本論文の執筆に関する多大なご指導と貴重なご助言をいただき、心より感謝申し上げます。研究テーマの設定から実験の設計および分析に至るまで、多くの示唆をいただき、研究の方向性を定める上で大変貴重な機会を得ることができました。特に、研究が難航した際には親身にご指導いただき、研究を進める上での新たな視点を与えてくださいました。また、学会での発表を勧めていただいたこと、誠に感謝しております。発表の機会をいただいたことより、自分自身が研究に対する方向性を確認することができました。この確認を踏まえて、研究を進めることに欠かせない気付きを得られました。

また、副査を務めていただいた新妻 雅弘先生には、多忙な中にも関わらず、研究の細部に至るまで丁寧にご指導をいただきました。自身の知識を整理し、論理的に研究を深掘りし、進めることの重要性を学ぶことができたことは、本研究における大きな財産となりました。

本研究の実験に協力していただいたメディアシステム研究室の皆様にも、心より御礼申し上げます。研究テーマに興味を持ち、実験に参加していただいたことで、多様な視点から議論を深めることができました。また、意見交換を通じて新たなアイデアを得ることができ、本研究の発展に大きく寄与しました。

また、本研究を進めるにあたり、さまざまな形でご支援をいただいた慶應義塾大学大学院システムデザイン・マネジメント研究科の関係者の皆様にも深く感謝申し上げます。学内外のセミナーやディスカッションの機会を通じて、多くの助言をいただいたことで、研究の視野を広げることができました。

さらに、日々の支えとなってくれた家族や友人にも深く感謝いたします。困難に直面した際にも励ましてくれた皆様のおかげで、最後まで研究をやり遂げることができました。

本研究は、私一人の力では決して成し遂げることができなかったものです。この二年間、多くの方々との出会いがあり、それぞれの支えがあったからこそ、この研究を完成させることができました。この場を借りて、心より感謝申し上げます。

第8章 付録