VR/AR空間における既知攻撃の技術的実現性の検証とリスク要因の整理

鈴木 裕友[1]　大塚 航世[1]　藤田 真由[1]　倉﨑 翔大[1]　森 亮太[3]　金岡 晃[1]　大東 俊博[3]　大木 哲史[2]

概要：VR/AR 技術の普及に伴い新たなセキュリティリスクが指摘されてきたが、それらの多くは概念的な指摘にとどまり実際に攻撃が実現可能かどうかについての技術的検証は十分に行われていない。本研究では、これまでの研究で示されてきたリスクと我々が新たに設計した攻撃を対象に、Meta Quest 3 と Unity 環境を用いてその実現可能性を検証した。また、各攻撃を成立させるために必要な機能を整理し、それらが実際にどのような形で開発者に提供されているかを分析した。さらに、複数の攻撃に共通する要因を明らかにし、対策可能な設計ポイントについて議論を行う。最後に、他の VR/AR 開発プラットフォーム（Unreal Engine、Apple ARKit、WebXR）における同様の攻撃の成立可能性についても考察する。

キーワード：VR/AR セキュリティ

Proof-of-Concept-Based Feasibility Analysis of Known Attacks in VR/AR Spaces and Functional Risk Mapping

Yusuke Suzuki[1]　Kosei Otsuka[1]　Mayu Fujita[1]　Shodai Kurasaki[1]　Ryota Mori[3]
Akira Kanaoka[1]　Toshihiro Ohigashi[3]　Tetsushi Ohki[2]

Abstract: With the widespread adoption of VR/AR technologies, various new security risks have been identified; however, many of these have remained conceptual, with limited technical verification of their actual feasibility. This study investigates the feasibility of previously proposed risks and newly designed attacks through implementation on the Unity platform with Meta Quest 3. The functions required for each attack are identified, and their availability to developers within the platform is analyzed. Common enabling factors across multiple attacks are also extracted, and potential design-level mitigation points are discussed. Finally, the feasibility of similar attacks is considered for other VR/AR development platforms, including Unreal Engine, Apple ARKit, and WebXR.

Keywords: VR/AR Security

1. はじめに

VR（Virtual Reality）/AR（Augmented Reality）技術は、HMD（Head-Mounted Display）をはじめスマートフォンや PC など多様なデバイス上で実用段階に達し、エンターテインメントや教育のみならず、製造や保守、医療、建設といった産業用途へ急速に広がっている。高度化したセンシングとレンダリングが人の知覚とより緊密に結び付くという特性は高い価値を生む一方で、セキュリティとプライバシの新たなリスクを顕在化させる。しかし既存の情報機器と同じ前提では捉えきれない論点が多く、VR/AR 技術に関するセキュリティとプライバシの研究はこの 2〜3 年で関連研究は増加しているものの、対象デバイスやプラットフォーム、前提条件の違いにより知見は断片化して

[1] 東邦大学
Toho University
[2] 静岡大学
Shizuoka University
[3] 東海大学
Tokai University

おり、俯瞰的な比較と体系化は十分とは言い難い。

従来研究は、センサ悪用、空間オブジェクトの偽装、UIの誘導など多様なリスクの存在を指摘してきたが、具体的にどの攻撃面 （Attack Surface）を経由し、どの機能・API・権限をどう組み合わせれば、どの機種で攻撃が現実に成立するのかという PoC（Proof of Concept）に基づく検証は著者の知る限りほとんど存在しない。脅威モデルの明示や、研究間で比較可能な評価の前提も十分に整っていないため、実効性や優先度付けの議論が難しい。

本研究では、攻撃者モデルを「拡張リソースのマーケットを経由する供給者」として定義する。すなわち、ブラウザの Google Chrome やエディタの Microsoft Visual Studio Code（VSCode）の拡張機能と同様に、公式マーケットで無害に見えるリソースとして配布されながら実際には有害な機能をユーザに意図させず適用させる攻撃を想定する。VR/AR アプリ開発の主要基盤である Unity の Asset Store は、この拡張モデルと構造的に類似している。Asset には 3D モデルだけでなくスクリプトが含まれ、3D データとスクリプトを組み合わせた Prefab をロードするだけで多機能な 3D オブジェクトが利用可能になる。この供給形態は利便性に大きく貢献する一方で、スクリプトによる悪意の混入が流通段階で潜在化しうるという攻撃面を生む。

本研究の中心的な着眼点は、個別の攻撃テクニックの多様さよりも、それらに共通する「プリミティブな行為」に置いた。前述の攻撃者モデルを前提に、プラットフォームが提供する許可済み機能の範囲内で、どのプリミティブが実現可能か、可能ならばどの程度の規模、精度、影響で実現できるかを PoC で検証する。たとえば、(i) 空間アンカやトラッキング情報への依存度を利用した位置合わせ、(ii) レンダリング順序や深度の操作、(iii) 入力・通知チャネルの乗っ取りに類する UI 介入、(iv) ネットワーク同期やアセット更新の挙動を利用した状態注入、といった行為を独立に切り出し、成立条件と制約を明らかにする。

実装は Unity 上で行い、実機として Meta Quest 3S を用いる。別の開発環境やデバイスに関する一般化可能性については、Quest 3S で観測された挙動とプラットフォーム仕様の対応から考察する。なお本稿では、ユーザ知覚への影響の測定や包括的な評価指標の設計は扱わずにユーザ影響は各リスク研究に委ね、プリミティブの成立性と実装上の条件に焦点を絞る。測定指標や比較枠組みは考察章で整理し、今後の統一的評価へ向けた論点として提示する。

本研究の貢献は次の通りである。第一に、拡張リソースのマーケット経由という供給形態に着目した攻撃者モデルを定式化し、Unity の Asset/Prefab/スクリプト構成がもたらす具体的な攻撃面を明確化する。第二に、VR/AR における攻撃実現の共通要素をプリミティブとして抽出し、その成立条件・必要権限・依存機能・失敗モードを体系化する。第三に、複数のプリミティブを PoC 実装し、Quest 3S 上で再現性のある手順とともに規模・精度・影響の範囲を示す。第四に、既往のリスク研究に対して、各プリミティブがどのリスクをどの条件で実現しうるかをマッピングし、リスクの実装到達可能性を構造的に位置付ける。第五に、得られた知見に基づき、プラットフォーム設計・Asset 審査・実行時制御の各層における実装可能な防御点（権限最小化、動的検査の挿入、レンダリング境界の分離、透明性インジケータの付与、更新・同期経路の検証強化など）を提示する。

2. 関連研究

近年、XR 環境における攻撃実証や攻撃可能性の研究が多くなっている。ここでは、これらを整理し、それぞれの研究の特徴を示していく。

認知や行動操作の領域では、Human Joystick 攻撃が行動誘導の代表例として提示されている [1]。より一般化された枠組みとして、ユーザの認知特性を手掛かりに行動や判断を誘導する PMA が提案され、視覚誘導を用いた行動操作の具体化が示されている [2], [3]。一方、聴覚モダリティを正面から扱い、音響刺激により視点を制御する可能性と攻撃ベクトルを整理した研究も現れている [4]。さらに、反射的や不可避的な知覚反応を突く誘導や、視覚配置の工夫によるダイアログ読取り妨害、VR 酔いの誘発など、可用性低下と誘導が交差する事例が報告されている [5], [6], [7]。

プライバシ推論や識別については、モーション特徴からの個人識別、点群やマップ情報に基づく空間や室内位置推論、動画や音などのサイドチャネルにより入力や行動を間接的に復元する研究がある [8], [9], [10], [11], [12], [13], [14], [15], [16]。生体指標の推定に踏み込む試みもあり、センシングの多重化が推論精度とリスクを同時に高めつつあることがわかる [17]。また、外向きカメラに写り込む周辺者のプライバシは、ユーザ本人以外の主体に対するリスクとして独自の論点を形成している [18], [19]。

コンテンツ偽装と完全性の観点では、3D オブジェクトの外観や属性を改変して利用者の理解を誤らせる Spoofing が中心的な話題であり、その概要と具体的手法が示されている [20], [21]。UI 上のクリック阻止やリダイレクション、重畳表示による誘導、偽カーソルや広告による視覚的欺きなど、表示層の操作を通じて意図された入力をゆがめるアプローチも広がっている [22], [23], [24]。フレームや音の挿入・削除による速度知覚の操作は、知覚の連続性を用いた偽装の一形態と捉えられる [25]。

プラットフォーム、ネットワーク、マルチユーザ環境にまたがる研究では、攻撃者が仮想空間へ実質的に侵入する事例、遠隔からの読み書きや GPS スプーフィングなど、権限や配置に依存する攻撃面が報告されている [26], [27]。ま

た、GPU 描画の改変やネットワーク負荷の付与といったシステム資源への介入は、表示品質や応答性を通じてユーザ体験と安全性に影響を与える [28]。

WebXR（WebVR）、拡張機能、開発エコシステムに基づく攻撃は、配信経路の多様化とともに存在感を高めている。WebXR（WebVR）の欠陥悪用により、広告や偽 UI を介して入力や視線を操作する手口が提示されており、同種の重畳・誘導は悪性広告や拡張機能を通じても実現しうる [22], [23], [24]。また、開発エコシステムにおける拡張やパッケージの汚染は、サプライチェーン経由でアプリ実行環境に侵入する経路として Google Chrome Extension や Microsoft VSCode Extension のマーケットでの事例が知られている [29], [30], [31], [32]。

防御や倫理、概念整理の系譜として、現実世界のスミッシングを AR で検知する試みや、周辺者プライバシに関する設計上の配慮、記憶操作をはじめとする人間中心の現象理解が挙げられる [18], [19], [33], [34]。

3. 攻撃者のモデル

想定する攻撃者は、プラットフォーム脆弱性の悪用者ではなく、Unity Asset Store といった拡張リソースのマーケットを経由して無害に見えるアセットやスクリプトを配布する供給者とした。現時点ではそういった攻撃の報告は著者らの知る限りでは無いものの、Google Chrome や VSCode で類似事例が報告されている [29], [30], [31], [32]。開発者や開発組織が当該リソースをプロジェクトに組み込むことで、攻撃者のコードがアプリ権限の範囲で実行される。Unity では、スクリプトを含む Prefab やアセットパッケージの導入だけで多機能オブジェクトが動作し得るという供給形態が、サプライチェーン由来の攻撃面を生む。

このような攻撃者の能力は、公開 API の利用とデータや表示状態の操作に関して自由度が高い。実行時にシーン内オブジェクトの属性や階層、レンダリング順序の変更、新規オブジェクトの生成や削除、入力情報（頭部姿勢・ハンドジェスチャ等）の並行参照、外部サーバへの送信、他スクリプトの停止、既存ノード配下への潜伏といった、プラットフォームが開発者に許容している機能を組み合わせて振る舞うことが可能になる。これらは PoC で検証可能な「攻撃プリミティブ」に分解でき、実行可能な攻撃は原則として許可済み機能の組み合わせとして成立する。

一方で、この攻撃者モデルには制約もある。本モデルは、未公開/内部 API の知識や利用を仮定しない。また、ドライバ層やハードウェア信号への直接介入、ファームウェア改変、カーネル権限の取得といった低層操作も想定しない。ネットワークの中間者化等の脆弱性依存の攻撃は、Man-in-the-Room 攻撃 [26] など既報が限定的であることから本モデルでは前提に置かないこととした。

前提条件として、対象アプリは標準的なビルド設定と権限で実行され、マーケットからのアセットや拡張の採用があり、プラットフォーム規定のサンドボックス内でのネットワーク送信など一般に許容された動作が可能であることを置く。

4. 攻撃実現のための調査項目

第 3 章で定義したサプライチェーン起点の攻撃者モデルを前提に、個別攻撃を 1 つずつ再現するのではなく、複数の攻撃に共通する行為を抽出し調査項目を設定する。通常利用可能な API と通常のアプリ権限のみを用い、どの行為がどの程度まで実装可能かを PoC で確認することで、既存研究に現れる多様な攻撃を一般化して評価可能になる。

4.1 調査項目（プリミティブ）の概要

攻撃者が取り得る技術的機能は次の 3 つの方向性に分けることができる。1 つ目は、シーン内リソースのライフサイクルと属性の制御である。2 つ目は、入力と状態ストリームの読取りである。そして 3 つ目は、プロセス境界外との情報のやり取りと、既存ロジックへの干渉である。その方向性を踏まえ、VR/AR に共通的に利用可能な情報と機能を考慮し、下記の 7 つの項目を設定した。項目 A〜G は、この 3 つの方向性を基本的な API の操作単位に分解した最小構成であり、いずれも脆弱性の存在を要さず、様々な形で利用可能な粒度で定義されている。

A. 既存オブジェクトの情報変更が可能か
- 目的: 位置、情報（名称、タグ、レイヤ等）の改変可否を確認する
- 方法: 悪意のあるスクリプトから他の既存オブジェクトのプロパティ変更を適用し、変更の影響を確認する

B. 新規オブジェクトが生成可能か
- 目的: スクリプト付きオブジェクト（Prefab）設置によるアプリ起動時オブジェクト生成の可否を確認する
- 方法: アプリ起動時の新規オブジェクト生成と配置を試験し、持続性を確認する

C. 既存オブジェクトの削除が可能か
- 目的: 実行時に他オブジェクトの破棄や無効化の可否を確認する
- 方法: アプリ起動後の既存オブジェクトの Destroy、無効化、不可視化、サイズ変更、遠方移動の各手段を試験し、影響を確認する。

D. ハンドジェスチャや HMD の向きと位置の重複参照が可能か
- 目的: 他スクリプトが入力を Listen 中でも並行参照できるかを確認する
- 方法: アプリ起動後に悪意のあるスクリプトによる各種情報の取得を試験し、取得情報を確認する。

E. 情報を外部サーバに送信可能か
- 目的: 実行時のネットワーク送信チャネルと送信可能性を把握する
- 方法: 悪意のあるスクリプト実行により既存オブジェクトの各種情報を収集し、それらをテキストファイル群とし zip ファイル化したのちに、固定 IP アドレスのサーバに HTTP POST 通信での送信を行い、送信可否を検証する

F. 他の実行中スクリプトを停止可能か
- 目的: 他スクリプトの停止・フリーズの可否を確認する
- 方法: アプリ実行中に、特定オブジェクトにアタッチされたスクリプトを Disable することを試験し、停止状態を確認する。

G. 既存オブジェクト下に悪意のあるオブジェクトを移動可能か
- 目的: 親子関係変更や深度順序調整により悪意のあるオブジェクトから既存オブジェクト下への移動・コピーの可否を確認する
- 方法: 新規生成オブジェクトや自分自身を既存ノード配下へのコピーを試験し、その結果を確認する

4.2 調査項目の妥当性考察

攻撃者モデルとこの 7 つの項目をもとに、2 章で挙げたいくつかの攻撃はこれらプリミティブの組合せで示すことができる。たとえば Casey らの Human Joystick Attack[1] や Cheng らの PMA[2] は、表示されているシーンやオブジェクトの変更や追加により認知を誤認させるため、項目 A と項目 B の実施可能性検証により実現が可能となる。また、攻撃モデルを越えるものについては、攻撃が実施不可能であることも同時に示すことができる。たとえば Chandio らの Visual Attack や Inertial Attack[25] はビジュアルフレームの追加削除や慣性データなどの直接的な操作が必要とされるが、これらは API の標準機能としては実施できないため、上記 7 項目ではその攻撃可能性は示すことができず、攻撃の実施が難しいことを示すことができる。

検証可能性の観点からも、この 7 つは適切であると考えられる。いずれも PoC で手順化でき、成功可否、必要権限、依存設定、安定性といった評価軸で客観的に評価可能である。また、倫理面においても、PoC 開発とその検証に脆弱性悪用や低層の直接操作を含まないため、一版ユーザやデバイスに過度のリスクを与えずに再現性の高い評価が可能である。

5. 攻撃実装結果

本章では，調査項目（A〜G）について，Unity を用いた Meta Quest 3 環境で実装と検証した結果を示す。

5.1 調査項目 A：既存オブジェクトの情報変更

name、Active/Inactive、Transform、Renderer の各属性は実行時に変更可能で、変更はシーン進行に追従して持続する。layer の変更はレイキャストや描画経路に直接影響し、Renderer、enabled の切替は可視性制御の基本操作として機能する。AudioSource は事前のリソース配置に依存するが、実行時付与とパラメタ更新が可能である。

tag と layer の変更には制限がある。いずれもプロジェクトに登録済みの値にのみ変更でき、存在しない値には設定できない。このため、攻撃者は未登録の分類へ変更することはできず、開発者が設定した既存の分類の範囲で付け替えや再配置を行うことになる。一方で、既存 layer のうちカメラの culling mask やレイキャストの対象外に設定されているものがあれば、そこへ移すだけで不可視化や非選択化が実現可能になる。

5.2 調査項目 B：新規オブジェクト生成

起動 f 直後およびイベント発生に応じたオブジェクト生成と配置は問題なく成立した。生成位置、親子関係、描画順序を A と合わせて調整することで、重畳表示や注意誘導 UI の出現など、表示層の振る舞いを攻撃者が構成可能である。

5.3 調査項目 C：既存オブジェクト削除

Destroy による削除と非アクティブ化はいずれも機能し、前者では OnDestroy、後者では OnDisable の呼出しを観測した。これらはイベントとして発生するため、検知が可能である。一方で、スケールの極小化や遠方への座標移動といった、データとしては存在するもののユーザの視覚的な認識には上がらないような準削除的手段でも不可視化が達成できる。これらは表示妨害やクリック阻止、コンテンツ隠蔽の実用的な手段となりうる。

5.4 調査項目 D：入力情報の重複参照

頭部姿勢やハンドジェスチャ等の入力は、専有は不要で他スクリプトが Listen 中であっても参照可能であることが確認された。この性質は、注意誘導やユーザ状態に適応する誘導 UI、さらにはプライバシ推論に利用可能である。

5.5 調査項目 E：外部送信

収集した情報をローカルで加工し、HTTP POST により外部サーバへ送信できた。パケット送出はプラットフォームのサンドボックスが許容する範囲で成立し、追加権限を要しない構成であっても実用的な送信経路が確立できることを確認した。これにより、D で取得した入力・状態の継続的な漏えいが現実的となる。

5.6 調査項目 F：他スクリプト停止

対象オブジェクトにアタッチされたスクリプトの停止が成立したことを確認した。更新処理の停止は保護的ハンドラや監視処理を無効化し得るため、他の項目と組み合わせることで検知回避や妨害に繋がる可能性がある。

5.7 調査項目 G：オブジェクト移動（潜伏）

既存オブジェクト配下への移動は問題なく成立し、親子関係と深度順序の変更により、生成物を自然にシーンの配下に含ませられることを確認した。これは検知回避や持続化において、A や B と併用することで表示偽装の可能性がより高まると言える。

6. 既存研究の攻撃と調査プリミティブの対応

本章では、第 5 章で確認したプリミティブ A〜G の成立性に基づき、既存研究で提示された各攻撃が攻撃者モデルでどこまで到達し得るかを、対応関係として示す。検討した対応関係表を表 1 に示す。

6.1 対応の考え方

表 1 の対応付けは、各攻撃の成立条件を、A〜G いずれの操作が不可欠かという観点で判断している。表中の「A, B」といった記載は、両者が同時に必要となるものではなく、どちらかが成立すれば成立することを示している。

6.2 構成されるプリミティブのパターン

表 1 から、いくつかの傾向が読み取れる。

複数の研究にまたがって同一のプリミティブに依存する場合、そのプリミティブは共通の攻撃面として機能していると言える。B 単独（8/37 件）、または B と A の組合せ（11/37 件）で成立する攻撃が多く存在することから、共通の攻撃面であると言えよう。重畳表示、カーソル偽装といった表示層の介入は、多くの場合、生成（B）と属性調整（A）の組合せで達成可能であり、サプライチェーン起点のコードでも実装容易性が高いことがわかる。

次に、D の並行参照が加わると、ユーザ状態に適応する誘導や識別系の攻撃へ到達しやすくなる。D は他コンポーネントの処理を阻害せずに状態を取得できるため、A や B と合成した適応的表示が技術的に実現しやすい。

E と F は支援的役割を持つことがわかる。E は継続的送出や遠隔設定の導入を可能にし、F は保護的ハンドラや監視ロジックの停止を通じて攻撃を安定化させる側面があると言える。

C は不可視化や遮蔽と親和性が高く、A や B、G と併用されると表示妨害の一貫性が向上してしまう。

なお、この整理は「共通面に含まれない機能が安全である」ことを意味してはいないことに注意が必要である。むしろ、現時点では攻撃者や研究者の関心が集まりにくい領域が残存しており、それらを利用した新たな攻撃、または本研究で扱うプリミティブとの合成による攻撃が今後増える可能性があると言える。したがって、防御設計においては、共通面の強化に加えて、周辺的に見える機能についても監査と抑止の方策を広範に検討する必要がある。

6.3 実施可能性が高い領域

本研究の調査項目には含めていないものの、攻撃者モデルに照らして、既存研究のうち、本攻撃者モデルにおいて技術的成立性が高いと推定される領域を挙げる。

Haptic grabbing[6] では、左右いずれかの手に選択的な振動を与え、注意や移動方向の偏りを生じさせ、通知の見落としや視線誘導を誘発する。プラットフォームが公開するハプティクス API の利用範囲次第では実現可能性が高いと考えられ、A（属性変更）および D（入力参照）と組み合わせることで、状態適応的な誘導として具体化し得る。

Remote Read/Write[27] では、Remote Read/Write (Limited User)、Remote Write (Global) が該当する。ユーザ生成コンテンツやセッション同期に公開 API 経由でフックできる場合、B（生成）と G（階層操作）により成立可能性が高い。

6.4 プラットフォーム依存で横断性の低い領域

特定の実行形態やプラットフォームに依存し横断的な成立性は低いが、環境次第で実現し得る事例を整理する。

Abuse of an Auxiliary Displa[24] では、HMD 装着中にユーザが注視しない PC モニタ等の補助ディスプレイで広告や動画を再生する。スタンドアロン HMD では成立しにくい一方、デスクトップ連携やミラーリングを介する環境では成立し得る。

GUI Switch [22] では、悪性広告からのスクリプト混入により画面キャプチャを取得し、性能劣化を誘発して没入モードを離脱させ、見た目が同一の別ページへ差し替える。Quest 3 のスタンドアロン環境では前提が整わない場合が多いが、WebXR を軸とするプラットフォームでは成立可能性がある。

6.5 OS/ソフトウェアの脆弱性を前提とする領域

Man-in-the-Room[26] は、脆弱性を突いてプライベートな VR ルームに不可視状態で侵入する。侵入後の移動・観察といった振る舞い自体は A（属性変更）等のプリミティブで表現可能だが、初期侵入は脆弱性前提であり、本研究の攻撃者モデルからは外れる。

6.6 物理・通信・低層信号の操作を前提とする領域

Odeleye@VR4Sec2021[28] の GPU 資源占有による描画

表 1　既存研究の攻撃と調査プリミティブの対応

文献	攻撃名	対応プリミティブ
Casey @ IEEE TDSC 2019[1]	Human Joystick Attack	A, B
Cheng @ USENIX Security 2023 [2]	PMA	A, B
藤田@ IoR-WS 2023[21], HCII 2025[20]	Spoofing (Superimposition)	B
	Spoofing (Transparent)	A, C
倉﨑@ IEEE MetaCom 2025 [3]	VPMA	A, B
大塚@ IEEE MetaCom 2025 (CSS2025) [4]	AVMA	A, B
森@ IEEE VRW 2025[5]	Misoperation-Inducing Attack	B
Lee @ USENIX Security 2021 [24]	Gaze Cursor-Jacking	A, B, C
	Controller Cursor-Jacking	A, B
	Abuse of an Auxiliary Display	
Wang @ IJHCI 2023 [6]	Lighting Interference	B
	Object Interference	B
	Haptic Grabbing	
Mukherjee @ USENIX Security 2025 [22]	Malvertising	B, D
	DoS through Overriding	B
	Visual Overlapping	B
	GUI Switch	
	Sequential Rendering	A, B
de Haas @ ISMARWS 2022 [35]	Auditory Guidance (Movement/Viewpoint)	B, D
	Auditory Disorientation (Cybersickness)	A, D
Vallasciani @ ISMARWS 2024 [36]	Shared State Manipulation / Leakage	A, E
Abraham @ NordiCHI 2022 [37]	False Information Display	A, B
	Perceptual Manipulation	A, B
	Impersonation	A, B
Sajid @ IEEE VR 2025 [23]	Click Redirection Attack	A, B
	Object Occlusion Attack	B
	Spatial Occlusion Attack	B
	Latency Attack	
Ye @ IEEE S&P 2025 [17]	BPSniff	D
de Guzman @ ESORICS 2019 [9]	Spatial Inference Attack	D
Slocum @ USENIX Security 2024 [27]	Remote Read/Write (Limited User)	
	Remote Write (Global)	
	Remote Read (Global)	
Vondráček @ Computers & Security 2023 [26]	Man-in-the-Room	
Odeleye@VR4Sec2021[28]		
Chandio@IEEE AIxVR2024 [25]	Visual Attack	
	Inertial Attack	

遅延と VR 酔い誘発、Chandio@IEEE AIxVR2024[25] のビジュアルフレームに対する中間者攻撃や慣性センサの信号の操作、Remote Read (Global)[27] に必要な GPS 偽装、Latency Attack [23]、といった攻撃群は物理層やドライバ層、ネットワーク制御層の介入を前提としている。これらは本研究の攻撃者モデルでは想定外であり、公開 API と通常権限のみでは再現しない。

7. 議論

7.1 攻撃者モデル

本研究では、サプライチェーン由来の通常権限と公開 API のみを用いる攻撃者を前提とした。これは、現状の実装可能性から有効である一方、将来の攻撃面を十分に見積もるには拡張の検討をすべきである。とくに、非公開低レイヤ API や、デバイス/OS 提供会社のアプリで用いられる内部 API の存在は知られており [38], [39]、これらが悪用されるシナリオは理論上排除できない。

7.2 攻撃フェーズに基づく対策ポイント

攻撃は概ね「配置→探索→取得→追加/変更」という 4 つのフェーズに分解でき、各フェーズで対策の可能性を議論することが可能になる。

7.2.1 悪意オブジェクトの配置

この段階では B と G が対象となる。防止としては、アセット導入時の署名検証と出所検証、実行時のオブジェクト生成ポリシ（生成可能な型や親子関係の制限、生成レー

トの上限）が有効であると考えられる。検知としては、生成イベントと階層再親化のログ、生成先レイヤやカメラ可視範囲との不一致検出が考えられる。

7.2.2 攻撃対象オブジェクトの探索

ここではAとDが対象となる。防止は、重要オブジェクトのtag/layer変更不可化や参照ガード、公開インタフェースの最小化で達成できる。検知は、特定メソッド/クラス経由の探索呼出しの集計と異常な探索頻度の検出、実行時リフレクションの使用監査が考えられる。

7.2.3 情報の取得

中心となるのはDとEである。防止では、入力/状態ストリームのアクセス範囲を機能単位で明示し、不要な並行参照を禁止することが考えられる。検知は、参照のサンプリング頻度や外部送出のエンドポイントを監査し、予期しない高頻度参照や未知の送信先を警告ことなどが考えうる。

7.2.4 情報の追加/変更

A、C、Fが関与する。防止は、重要属性の動的変更禁止、描画順序やculling設定の固定、クリティカルコンポーネントの停止不可化が考えられる。検知は、RendererやCollider等の有効/無効切替、削除/不可視化、停止要求のイベント化と監査が考えられる。

8. まとめ

本研究はサプライチェーン起点、通常権限、公開APIのみを前提に攻撃を7つプリミティブへ分解し、Unityを開発環境としMeta Quest3を用いて実装可能性を検証した。そして、既存の多くの攻撃を可能にすることをこれまでのVR/ARに関する攻撃に言及した論文との対応表により明らかになった。一方、ドライバやセンサ、ネットワークへの低レイヤ介入や未公開API依存の手法は本モデルの外側に位置づきことを示し、より強い攻撃者モデルを要することを議論した。本研究により、実装可能性に基づくこれまでの研究で示された脅威の現実性の概観と、設計・審査・運用で優先すべき対策点が明確になった。

謝辞 本研究は、JST、CREST、JPMJCR22M4の支援を受けたものである

参考文献

[1] Casey, P., Baggili, I. and Yarramreddy, A.: Immersive Virtual Reality Attacks and the Human Joystick, *IEEE Transactions on Dependable and Secure Computing*, Vol. 18, No. 2, pp. 550–562 (online), DOI: 10.1109/TDSC.2019.2907942 (2021).

[2] Cheng, K., Tian, J. F., Kohno, T. and Roesner, F.: Exploring User Reactions and Mental Models Towards Perceptual Manipulation Attacks in Mixed Reality, *32nd USENIX Security Symposium (USENIX Security 23)*, pp. 911–928 (2023).

[3] Kurasaki, S. and Kanaoka, A.: Image Movement Attacks on Optical See-Through HMDs: Covert Gaze Manipulation and Privacy Risks in AR/MR Systems, *2025 IEEE International Conference on Metaverse Computing, Networking, and Applications (MetaCom)* (2025).

[4] Otsuka, K. and Kanaoka, A.: Auditory Stimulus Attack in XR: Stimulus Characteristics and Technical Background Considerations, *2025 IEEE International Conference on Metaverse Computing, Networking, and Applications (MetaCom)* (2025).

[5] Mori, R., Kakizaki, Y., Nakatani, H., Kanaoka, A. and Ohigashi, T.: Attack based on Operational Error caused by Stop-Signal Reaction Time in VR space , *2025 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 881–883 (online), DOI: 10.1109/VRW66409.2025.00179 (2025).

[6] Wang, X., Lee, L.-H., Fernandez, C. B. and Hui, P.: The Dark Side of Augmented Reality: Exploring Manipulative Designs in AR, *International Journal of Human–Computer Interaction*, Vol. 40, No. 13, pp. 3449–3464 (online), DOI: 10.1080/10447318.2023.2188799 (2024).

[7] Valluripally, S., Gulhane, A., Hoque, K. A. and Calyam, P.: Modeling and Defense of Social Virtual Reality Attacks Inducing Cybersickness, *IEEE Transactions on Dependable and Secure Computing*, Vol. 19, No. 6, pp. 4127–4144 (online), DOI: 10.1109/TDSC.2021.3121216 (2022).

[8] Nair, V., Guo, W., Mattern, J., Wang, R., O'Brien, J. F., Rosenberg, L. and Song, D.: Unique Identification of 50,000+ Virtual Reality Users from Head & Hand Motion Data, *32nd USENIX Security Symposium (USENIX Security 23)*, pp. 895–910 (2023).

[9] de Guzman, J. A., Thilakarathna, K. and Seneviratne, A.: A First Look into Privacy Leakage in 3D Mixed Reality Data, *Computer Security – ESORICS 2019: 24th European Symposium on Research in Computer Security, Luxembourg, September 23–27, 2019, Proceedings, Part I*, p. 149–169 (online), DOI: 10.1007/978-3-030-29959-0_8 (2019).

[10] Farrukh, H., Mohamed, R., Nare, A., Bianchi, A. and Celik, Z. B.: LocIn: Inferring Semantic Location from Spatial Maps in Mixed Reality, *32nd USENIX Security Symposium (USENIX Security 23)*, pp. 877–894 (2023).

[11] Gopal, S. R. K., Shukla, D., Wheelock, J. D. and Saxena, N.: Hidden Reality: Caution, Your Hand Gesture Inputs in the Immersive Virtual World are Visible to All!, *32nd USENIX Security Symposium (USENIX Security 23)*, pp. 859–876 (2023).

[12] Zhang, Y., Slocum, C., Chen, J. and Abu-Ghazaleh, N.: It's all in your head(set): Side-channel attacks on AR/VR systems, *32nd USENIX Security Symposium (USENIX Security 23)*, pp. 3979–3996 (2023).

[13] Slocum, C., Zhang, Y., Abu-Ghazaleh, N. and Chen, J.: Going through the motions: AR/VR keylogging from user head motions, *32nd USENIX Security Symposium (USENIX Security 23)*, pp. 159–174 (2023).

[14] Luo, S., Hu, X. and Yan, Z.: HoloLogger: Keystroke Inference on Mixed Reality Head Mounted Displays, *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 445–454 (online), DOI: 10.1109/VR51125.2022.00064 (2022).

[15] Meteriz-Yildiran, U., Yildiran, N. F., Awad, A. and Mohaisen, D.: A Keylogging Inference Attack on Air-Tapping Keyboards in Virtual Environments, *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 765–774 (online), DOI: 10.1109/VR51125.2022.00098 (2022).

[16] Arafat, A. A., Guo, Z. and Awad, A.: VR-Spy: A Side-Channel Attack on Virtual Key-Logging in VR Headsets, *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 564–572 (online), DOI: 10.1109/VR50410.2021.00081 (2021).

[17] Ye, Z., Mahdad, A. T., Wang, Y., Shi, C., Chen, Y. and Saxena, N.: BPSniff: Continuously Surveilling Private Blood Pressure Information in the Metaverse via Unrestricted Inbuilt Motion Sensors , *2025 IEEE Symposium on Security and Privacy (SP)*, pp. 4356–4374 (online), DOI: 10.1109/SP61157.2025.00049 (2025).

[18] Denning, T., Dehlawi, Z. and Kohno, T.: In Situ with Bystanders of Augmented Reality Glasses: Perspectives on Recording and Privacy-Mediating Technologies, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pp. 2377–2386 (online), DOI: 10.1145/2556288.2557352 (2014).

[19] Roesner, F., Kohno, T. and Molnar, D.: Security and Privacy for Augmented Reality Systems, *Commun. ACM*, Vol. 57, No. 4, pp. 88–96 (online), DOI: 10.1145/2580723.2580730 (2014).

[20] Fujita, M., Kurasaki, S. and Kanaoka, A.: Investigating 3D Object Spoofing on Fundamental and Custom Objects in Virtual Reality, *HCI for Cybersecurity, Privacy and Trust* (Moallem, A., ed.), pp. 171–187 (2025).

[21] Fujita, M., Kurasaki, S. and Kanaoka, A.: Securing Cross Reality: Unraveling the Risks of 3D Object Disguise on Head Mount Display, *Proceedings of the 13th International Conference on the Internet of Things*, IoT '23, pp. 281–286 (online), DOI: 10.1145/3627050.3631570 (2024).

[22] Mukherjee, C., Mohamed, R., Arunasalam, A., Farrukh, H. and Celik, Z. B.: Shadowed Realities: An Investigation of UI Attacks in WebXR, *34th USENIX Security Symposium (USENIX Security 25)* (2025).

[23] Sajid, M., Shah Bukhari, S. I. M., Ji, B. and David-John, B.: "Just stop doing everything for now!": Understanding security attacks in remote collaborative mixed reality, *2025 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pp. 623–633 (online), DOI: 10.1109/VR59515.2025.00085 (2025).

[24] Lee, H., Lee, J., Kim, D., Jana, S., Shin, I. and Son, S.: AdCube: WebVR Ad Fraud and Practical Confinement of Third-Party Ads, *30th USENIX Security Symposium (USENIX Security 21)*, pp. 2543–2560 (2021).

[25] Chandio, Y., Bashir, N. and Anwar, F. M.: Stealthy and Practical Multi-modal Attacks on Mixed Reality Tracking, *2024 IEEE International Conference on Artificial Intelligence and eXtended and Virtual Reality (AIxVR)*, pp. 11–20 (online), DOI: 10.1109/AIxVR59861.2024.00010 (2024).

[26] Vondráček, M., Baggili, I., Casey, P. and Mekni, M.: Rise of the Metaverse' s Immersive Virtual Reality Malware and the Man-in-the-Room Attack & Defenses, *Computers & Security*, Vol. 127, p. 102923 (オンライン), DOI: https://doi.org/10.1016/j.cose.2022.102923 (2023).

[27] Slocum, C., Zhang, Y., Shayegani, E., Zaree, P., Abu-Ghazaleh, N. and Chen, J.: That Doesn't Go There: Attacks on Shared State in Multi-User Augmented Reality Applications, *33rd USENIX Security Symposium (USENIX Security 24)*, pp. 2761–2778 (2024).

[28] Odeleye, B., Loukas, G., Heartfield, R. and Spyridonis, F.: Detecting framerate-oriented cyber attacks on user experience in virtual reality (2021).

[29] : Time to check if you ran any of these 33 malicious Chrome extensions - Ars Technica, (online), available from ⟨https://arstechnica.com/security/2025/01/dozens-of-backdoored-chrome-extensions-discovered-on-2-6-million-devices/⟩.

[30] : Google and Microsoft Trusted Them. 2.3 Million Users Installed Them. They Were Malware. — Koi Blog, (online), available from ⟨https://www.koi.security/blog/google-and-microsoft-trusted-them-2-3-million-users-installed-them-they-were-malware⟩.

[31] : A new playground: Malicious campaigns proliferate from VSCode to npm — ReversingLabs — ReversingLabs, (online), available from ⟨https://www.reversinglabs.com/blog/a-new-playground-malicious-campaigns-proliferate-from-vscode-to-npm⟩.

[32] : 1/6 — How We Hacked Multi-Billion Dollar Companies in 30 Minutes Using a Fake VSCode Extension — Koi Blog, (online), available from ⟨https://www.koi.security/blog/1-6-how-we-hacked-multi-billion-dollar-companies-in-30-minutes-using-a-fake-vscode-extension⟩.

[33] Kanaoka, A. and Isohara, T.: Enhancing Smishing Detection in AR Environments: Cross-Device Solutions for Seamless Reality, *2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 565–572 (online), DOI: 10.1109/VRW62533.2024.00108 (2024).

[34] Bonnail, E., Tseng, W.-J., Mcgill, M., Lecolinet, E., Huron, S. and Gugenheimer, J.: Memory Manipulations in Extended Reality, *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, (online), DOI: 10.1145/3544548.3580988 (2023).

[35] Anne de Haas, E. H. and Lee, L.-H.: Deceiving Audio Design in Augmented Environments : A Systematic Review of Audio Effects in Augmented Reality, *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 36–43 (online), DOI: 10.1109/ISMAR-Adjunct57072.2022.00018 (2022).

[36] Vallasciani, G., Schinoppi, A., Cascarano, P., Marfia, G. and Donatiello, L.: Handling privacy and security aspects in a collaborative AR session, *2024 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 20–22 (online), DOI: 10.1109/ISMAR-Adjunct64951.2024.00013 (2024).

[37] Abraham, M., Saeghe, P., Mcgill, M. and Khamis, M.: Implications of XR on Privacy, Security and Behaviour: Insights from Experts, *Nordic Human-Computer Interaction Conference*, NordiCHI '22, (online), DOI: 10.1145/3546155.3546691 (2022).

[38] He, Y., Gu, Y., Su, P., Sun, K., Zhou, Y., Wang, Z. and Li, Q.: A Systematic Study of Android Non-SDK (Hidden) Service API Security , *IEEE Transactions on Dependable and Secure Computing*, Vol. 20, No. 02, pp. 1609–1623 (online), DOI: 10.1109/TDSC.2022.3160872 (2023).

[39] Wang, C., Zhang, Y. and Lin, Z.: Uncovering and Exploiting Hidden APIs in Mobile Super Apps, *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security*, CCS '23, p. 2471–2485 (online), DOI: 10.1145/3576915.3616676 (2023).