

VR ゴーグル対応 Web ベース多地点 360° VR 動画フレームワーク

岡田 義広, 永山 寛二, 木佐貫 浩司

九州大学 データ駆動イノベーション推進本部 次世代型オープンエデュケーション推進部門

okada.yoshihiro.520@m.kyushu-u.ac.jp

Framework for Web-based Multi-Location 360-degree VR Videos Supporting VR Goggles

Yoshihiro Okada, Kanji Nagayama, Hiroshi Kisanuki

Division of Next generation Open Education promotion, Data-Driven Innovation Initiative, Kyushu Univ.

概要

本稿では、360° VR 画像・動画の Web ベース VR コンテンツのフレームワークを取り上げる。最近では、多くの企業が VR ゴーグルを製造し、多くのゲーム会社が VR ゴーグルに対応した VR ゲームをリリースしている。しかし、教育用 VR コンテンツはまだ少ないのが現状である。また、多くの企業が 360° VR カメラを製造している。実験や演習科目において、それら実験室や演習室に関する VR ゴーグルに対応した 360° VR 画像・動画コンテンツは、事前に実験や演習の進め方を深く学ぶことができるため教育用途として有用である。このような 360° VR 画像・動画コンテンツを容易に開発するために、著者らはすでに 360° VR 動画のナビゲーションコンテンツや 360° VR 画像のウォークスルーコンテンツなど、Web ベース VR コンテンツのフレームワークをいくつか提案している。本稿で、著者らは VR ゴーグル対応 Web ベース多地点 360° VR 動画のための新しいフレームワークを提案する。自由視点動画や自由視点ライブストリーミング動画のサービスは数多く存在する。ただし、これらは複数地点の 360° VR 動画ではない。VR ゴーグル対応の多地点 360° VR 動画は、通常の 360° VR 動画や自由視点動画よりも没入感が高く教育効果が高いと考えられる。著者らは、実際にこのフレームワークを使用して開発した VR ゴーグル対応の複数地点 360° VR 動画を示し、提案したフレームワークの有用性を述べる。

1 はじめに

九州大学は 2022 年 4 月、学内の DX を推進するためデータ駆動イノベーション推進本部を設立した。当本部はいくつかの部門で構成されており、そのうちの 1 つが著者らの所属する次世代型オープンエデュケーション推進(NOE)部門である。NOE 部門の使命は、大学の教育 DX を推進することである。NOE 部門の活動として、XR (VR、AR、MR) などの最新 ICT を活用した電子教材の開発や、教員が教材作成を行う際の支援などがある。本稿では、VR ゴーグル対応 Web ベース多地点 360° VR 動画のフレームワークを提案する。

近年多くの企業が VR ゴーグルを製造しており、多くのゲーム会社が VR ゴーグルに対応した VR ゲームをリリースしている。しかし、教育用 VR コンテンツはほとんどないのが現状である。一方、これまで多くの企業が、LiDAR(Light Detection And Ranging) カメラ、360° VR カメラなど、

さまざまなスキャンデバイスを製造している。NOE 部門では教育用 XR コンテンツ開発用にこれらのスキャンデバイスを保有している。これらのスキャンデバイスを使用することで、点群データ (PCD: Point Cloud Data)、3D モデルデータ、360° VR 画像/動画などの 3D データを取得することができる。このような 3D データを教育用 XR コンテンツに活用するために、NOE 部門ではすでに Web ベースの教育 XR コンテンツや 360° VR 画像/動画の Web ベース VR ツアーのフレームワーク [1] を提案している。さらに、OpenVSLAM (Open Visual SLAM: Simultaneous Localization and Mapping)[2] を採用してフレームワークを拡張し、キーワード検索と多言語字幕を可能にする OpenVSLAM ベースのフレームワーク [3,4] を提案した。さらに、本稿では、VR ゴーグル対応 Web ベース多地点 360° VR 動画フレームワークを提案する。自由視点動画や自由視点ライブストリーミング動画のサ

ービスは数多く存在する。しかし、複数地点 360° VR 動画やそのためのサービスはほとんどない。多地点 360° VR 動画の制作には複数の 360° VR カメラが必要であるが、VR ゴーグル対応コンテンツは通常の 360° VR 動画や自由視点動画よりも高い没入感が得られるため、VR ゴーグル対応複数地点 360° VR 動画は、教育効果が高いと考えられる。例えば、実験・演習科目では、実験・演習シーンを VR ゴーグル対応多地点 360° VR 動画で視聴することで、事前に実験・演習の様子を深く学ぶことができる。ただし、このような複数地点 360° VR 動画の制作は、一般の教師のようなエンドユーザにとって簡単ではない。そこで、著者らは、そのための開発フレームワークを提案する。提案されたフレームワークを使用して実際に開発された VR ゴーグル対応複数地点 360° VR 動画を示すことにより、提案するフレームワークの有用性を述べる。

2 関連研究

LiDAR や 360° VR カメラなどのセンシングデバイスをリリースしている企業がいくつかある。一部は、ユーザが PCD および 360° VR 画像/動画の対話的な Web ベース VR アプリケーションを作成できる商用サービスを提供している。RICOH のサービス (<https://www.theta360.biz/>) は 360° VR 画像に対応しているが、360° VR 動画や PCD には対応していない。このサービスの注目すべき点は、エンドユーザが独自の Web ベース 360° VR コンテンツを作成できる、著者らの提案フレームワークにはない Web 編集機能を提供していることである。Matterport のサービス (<https://matterport.com/>) は 360° VR 画像と PCD をサポートしているが、360° VR 動画はサポートしていない。このサービスの注目すべき点は、Matterport キャプチャ アプリケーションと呼ばれる AI 技術をベースとした強力なツールを備えており、360° VR 画像と PCD から位置マップを自動生成することである。著者らの OpenVSLAM ベース・フレームワークでも、移動パスなどの位置情報を自動生成できる。RICOH も Matterport も、複数地点 360° VR 動画サービスは提供していない。

一方、自由視点動画やライブストリーミング動画のサービスも数多くある。SwipeVideo は、

こうした自由視点動画サービスのパッケージの 1 つ(<https://swipevideo.jp/>)である。ただし、360° VR 画像・動画には対応していない。現在、YouTube などでも視聴できる 360° VR 動画が世の中に数多く存在する。しかし、複数地点での 360° VR 動画やそのようなサービスはほとんどないのが現状である。著者らのコンテンツと他のコンテンツとの違いは、著者らのコンテンツが複数の 360° VR カメラを同時に使用して複数の異なる場所で同期的に記録された複数の 360° VR 動画であることである。著者らと同様のアプローチとして、日本の企業 connekiss (<https://connekiss.com/>) は、Meta Quest 2 などの VR ゴーグルに対応した POV 観光アプリケーションをベースとした多地点 360° VR 動画コンテンツを提供している。これらのコンテンツは Web ベースではなく、著者らの提案するフレームワークは、複数地点の 360° VR 動画の Web コンテンツである点で異なる。

3 既存フレームワーク

ここでは、著者らの提案する 360° VR 画像・動画の Web ベース VR コンテンツの既存フレームワークを紹介する。

3.1 360°VR カメラと専用ソフトウェア

360° VR 画像・動画の Web ベース VR コンテンツを開発するには、360° VR カメラを使用してデータを得る必要がある。図 1 の左側に示すように、著者らの NOE 部門 には、最大 8K 解像度の高品質 360° VR 画像/動画をキャプチャできる 6 個の魚眼レンズを備えた Insta360 Pro がある。また、図の右側に示される Insta360 ONE X2 もある。2 個の魚眼レンズを備え、最大 5.7K 解像度の 360° VR 画像/動画をキャプチャできる。Insta360 ONE X2 は Insta360 Pro よりもポータブルで安価である。NOE 部門 には Insta360 ONE X2 が 5 台あるため、複数の場所で 360° VR 動画の同期録画が可能である。図 2 の上部に示される 2 個の魚眼レンズでキャプチャした画像/動画を、Insta360 Studio 等の専用ソフトウェアを使用して、図の中段に示される 正距円筒画像/動画に変換する。変換された正距円筒画像/動画をテクスチャマッピング技術で 3D 空間の球体モデルの内面にマッピングすることにより、360° VR 画像/動画の視聴が可能となる。ユーザ視点は、球

体モデルの中心にあり、マウスデバイスなどを使用して、インタラクティブに視点の向きを垂直または水平方向に移動できる。



図 1. Insta360 Pro (左)と Insta360 ONE X2 (右).



Two fish-eye lens images/videos of Insta360 ONE X2



converted by dedicated software



Equirectangular image/video



mapped onto the inner surface of a sphere model



Changing the user's view direction vertically or horizontally

図 2. 専用ソフトウェアを用いた魚眼レンズ画像・動画から正距円筒図画像・動画への変換と閲覧メカニズム。

3.2 360° VR 画像・動画の Web ベース VR コンテンツの既存フレームワーク

360° VR 画像/動画の Web ベース VR コンテンツのフレームワークは 2 つのシステムで構成される。360° VR 画像のウォークスルー VR コンテンツ用と 360° VR 動画のナビゲーション VR コンテンツ用である。

1) 360° VR 画像の Web ベースウォークスルーVR コンテンツ

図 3 に 360° VR 画像のウォークスルーVR コンテンツ用システムのファイル構成を示す。1 つの HTML ファイル (walk_map.html)、主要な JavaScript ファイル (つまり walk_map.js)、および 補助的な JavaScript ファイル (つまり walk_map_controls.js) がある。

```
walk_map.html
JS { walk_map.js, walk_map_controls.js }
JSM { Several JS files from Three.js library }
```

Assets

```
Images { 360VR images, optional images }
Models { optional 3D model files }
Movies { optional movie files }
Sounds { optional sound files }
```

図 3. 360°VR 画像のウォークスルーシステムのファイル構成

walk_map.html は、最初に HTML タグと Web コンテンツの外観を定義する。walk_map.js は、WebGL、WebXR、および Three.js ライブラリを使用して、Web コンテンツの 3D グラフィックス部分とその対話インターフェイスを定義している。Three.js は、WebGL の強力なライブラリである。walk_map_controls.js は、Web コンテンツで使用されるアセットを定義している。ウォークスルーVR コンテンツを開発する場合、walk_map.js を変更する必要はない。360° VR 画像のファイル名を、walk_map_controls.js で設定し、walk_map.js プログラムからそれらのメディアファイルが適切に読み込まれるようにすれば良い。

図 4 は、CT 演習室の 360° VR 画像 (上段) と 2 つの異なる視点位置画像 (中段および下段) のウォークスルーVR コンテンツのスクリーンショットである。メイン画面上でマウスをドラッグ

することにより、視点の方向を上下左右に変えられる。その仕組みは図 2 に示した閲覧メカニズムの通りである。各図の左側は部屋のマップである。マップ上の赤いマーカーを見ることで、現在地と方向を把握できる。360° VR 画像が撮影された場所を示す小さなオレンジ色の点がマップ上に表示されている。ユーザは、マップ上の対応するオレンジ色の点をマウスクリックすることで、現在の 360° VR 画像を、行きたい場所で撮影した画像に変更できる。メイン画面では、360° VR 画像が撮影された場所を示す灰色の円盤が表示されている。それらの 1 つをマウスクリックすると、ユーザは現在の 360° VR 画像をその場所で撮影された画像に変更できる。このようにして、ユーザは自分の位置を対話的に変更し、フロア内をウォークスルーできる。

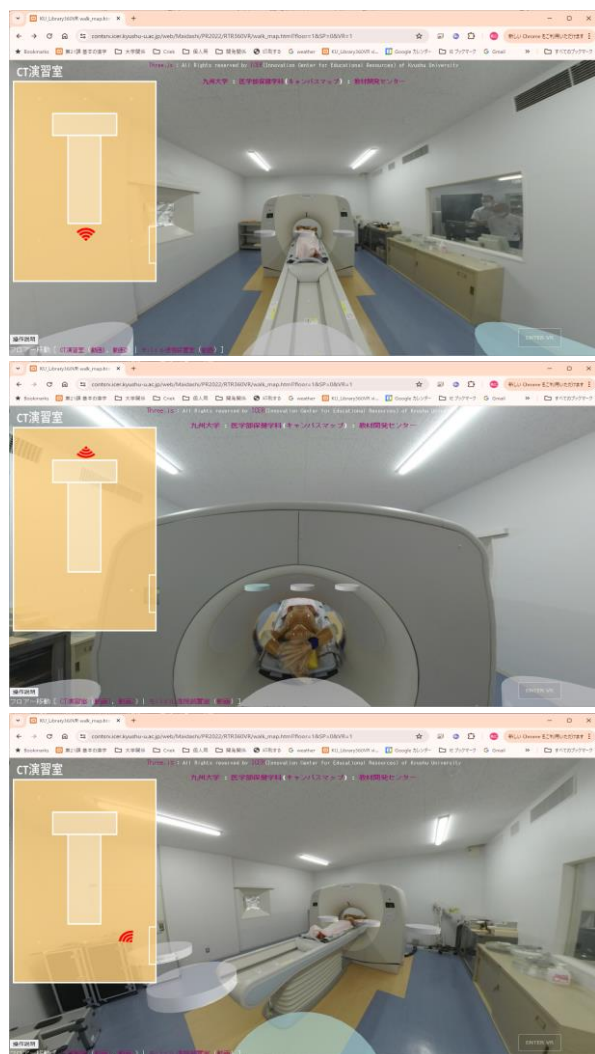


図 4. 360°VR 画像の Web ベースウォークスルー VR コンテンツ例

図 5 は、図 4 と同じコンテンツであるが、VR ゴーグルモードのスクリーンショットである。図 4 の各スクリーンショットの右下に [Enter VR] ボタンが表示されており、このボタンを押すと、通常モードから VR ゴーグルモードに切り替わる。ユーザが VR ゴーグルの Web ブラウザを使用している場合、このボタンを押すと図 5 に示すようなステレオ画像として当該のシーンが VR ゴーグル内に表示される。頭の動きに追従して視線方向が変化し、高い没入感が得られる。

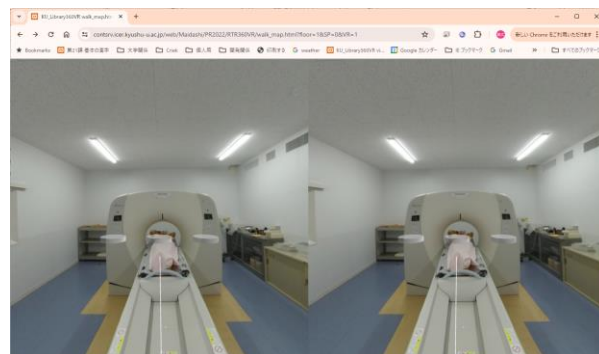


図 5. 360°VR 画像の Web ベースウォークスルー VR コンテンツ例-VR ゴーグルモード

2) 360° VR 動画の Web ベースナビゲーション VR コンテンツ

図 6 は、360° VR 動画のナビゲーション VR コンテンツ用システムのファイル構成を示している。1 つの HTML ファイル (video_map.html)、主要な JavaScript ファイル (video_map.js)、および補助的な JavaScript ファイル (video_map_controls.js) がある。

video_map.html
JS { video_map.js, video_map_controls.js }
JSM { Several JS files from Three.js library }
Assets
Images { optional images }
Videos { 360VR videos }
Models { optional 3D model files }
Movies { optional movie files }
Sounds { optional sound files }

図 6. 360°VR 動画のナビゲーションシステムのファイル構成

video_map.html は、最初に HTML タグと Web コンテンツの外観を定義する。

video_map.js は、WebGL、WebXR、および Three.js ライブラリを使用して、Web コンテンツの 3D グラフィックスとその対話インターフェイスを定義する。video_map_controls.js は、Web コンテンツで使用されるアセットを定義する。ナビゲーション VR コンテンツを開発する場合、video_map.js を変更する必要はない。360° VR 動画のファイル名を、video_map_controls.js で設定し、video_map.js プログラムからそれらのメディアファイルが適切に読み込まれるようにすれば良い。



図 7. 360°VR 動画の Web ベースナビゲーション VR コンテンツ例

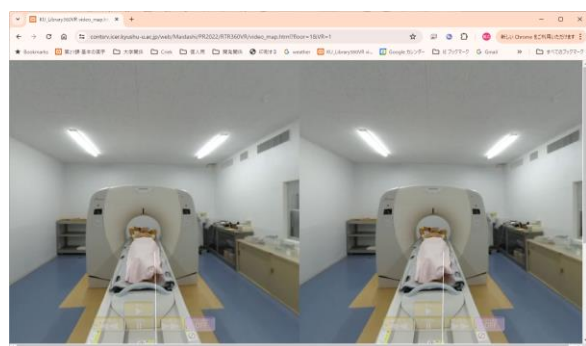


図 8. 360°VR 動画の Web ベースナビゲーション VR コンテンツ例-VR ゴーグルモード

図 7 は、360° VR 動画のナビゲーション VR コンテンツ（上段）と CT 演習室の異なる視線方向（下段）のスクリーンショットである。メイン画面上のマウสดラッグにより、視線の方向を上下左右に変更できる。その仕組みは図 2 に示した関

覧メカニズムの通りである。各図の左側は部屋のマップである。このコンテンツは、マップ上の赤いマーカーで示された場所で撮影された 1 つの 360° VR 動画のみで構成されているため、マップ上ではオレンジ色の点は表示されない。各スクリーンショットの中央下部には、360° VR 動画を制御する再生、一時停止、逆方向、早送りなどホームビデオ機器と同様のコントロールパネルが表示されている。

図 8 は、図 7 と同じコンテンツのスクリーンショットであるが、VR ゴーグルモードである。図 7 の各スクリーンショットの右下に [Enter VR] ボタンが表示されている。このボタンを押すと、通常のモードから VR ゴーグルモードへ切り替わる。ユーザが VR ゴーグルの Web ブラウザを使用している場合、このボタンを押すと図 8 のような部屋のステレオ画像が VR ゴーグル内に表示される。頭の動きに追従して視線方向が変化し、高い没入感が得られる。重要な点は、VR ゴーグルモードでは HTML DOM のオブジェクトが表示されないため、画面中央下に表示されているコントロールパネルは HTML DOM ではないことである。代わりに、VR ゴーグルモードのコントロールパネルは、WebXR および Three.js の機能によって定義された 3D オブジェクトとして実現されている。VR ゴーグルモードでは、ユーザは VR ゴーグルのコントローラーでこのコントロールパネルの各ボタンを操作する。

4 提案フレームワーク

このセクションでは、VR ゴーグルに対応した Web ベース多地点 360° VR 動画のためのフレームワークについて説明する。フレームワークのすべての JavaScript ファイルと HTML ファイルを図 9 に示す。

フレームワークには、1 つの HTML ファイル (walk_video_map.html)、主要 JavaScript ファイル (walk_video_map.js)、および補助的 JavaScript ファイル (walk_video_map_controls.js) が含まれている。これらの JavaScript ファイルと HTML ファイルは、前のセクションで説明したフレームワークのファイルを統合して作成したものである。厳密に言えば、walk_video_map.html は walk_map.html と video_map.html を統合して

作成した。walk_video_map.js は walk_map.js と video_map.js を統合して作成した。さらに、walk_video_map_controls.js は walk_map_controls.js と video_map_controls.js を統合して作成した。

walk_video_map.html JS { walk_video_map.js, walk_video_map_controls.js } JSM { Several JS files from Three.js library }
Assets Images { optional images } Videos { 360VR videos } Transcripts { vtt files } Models { optional 3D model files } Movies { optional movie files } Sounds { optional sound files }

図 9. Web ベース多地点 360°VR 動画システムのファイル構成

既存のフレームワークと同様に、新しい VR コンテンツを作成する際、walk_video_map.js を変更する必要はない。まず、必要なメディア ファイル、つまり Assets ディレクトリに保存される 360° VR 動画ファイルを準備する。それら動画のファイル名を指定するなど walk_video_map_controls.js に適切な変更を加える。これで、必要なメディアファイルが主要 JavaScript プログラム (walk_video_map.js) から適切に読み込まれるようになる。さらに追加した機能は、字幕を表示することである。実際、OpenVSLAM ベースの開発フレームワーク [4] ではすでに字幕の表示をサポートしている。AssemblyAI サービスを使用して、動画ファイルから VTT 形式の日本語トランスクリプトファイルを生成した。生成された日本語トランスクリプトには、360° VR 動画に含まれる音声ナレーションとほぼ同じ内容が含まれている。ただし、完璧ではないため、手作業による修正が必要である。

図 10. は、CT 演習室の Web ベース多地点 360° VR 動画(上段)と異なる 2 地点動画(中段と下段)のスクリーンショットである。これらのスクリーンショットは、図 4 と図 7 を結合した画像のように見える。部屋には 5 台の 360° VR カメラが見える。メイン画面には、灰色の円盤が

5 つ見える。それぞれの位置は 5 台の 360° VR カメラのそれぞれとほぼ同じである。また、各スクリーンショットの左側に示されるマップには、オレンジ色の点が 5 つある。各位置は、5 つの 360° VR カメラ位置のそれぞれに対応している。

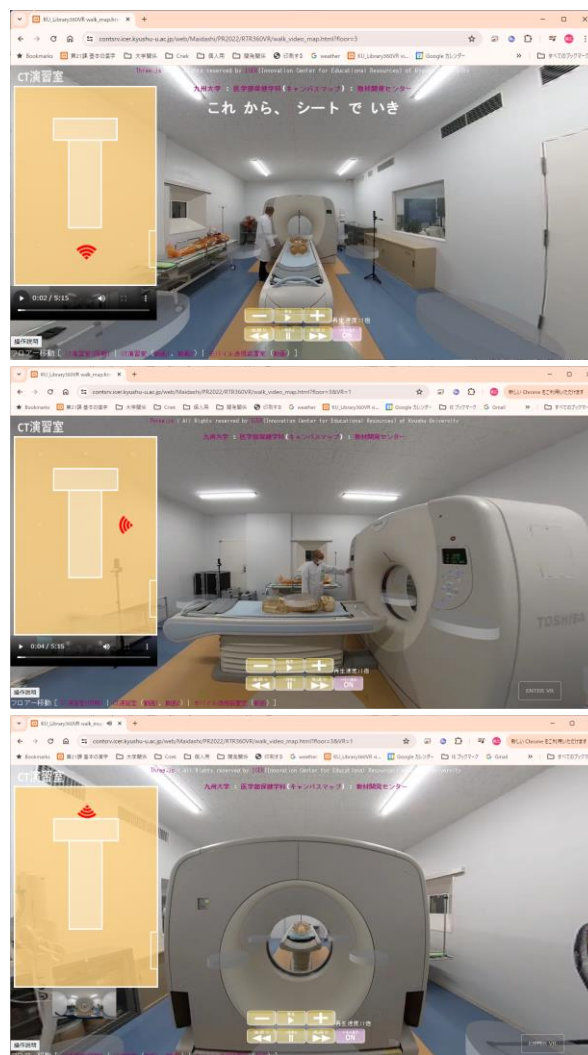


図 10. Web ベース多地点 360°VR 動画 VR コンテンツ例

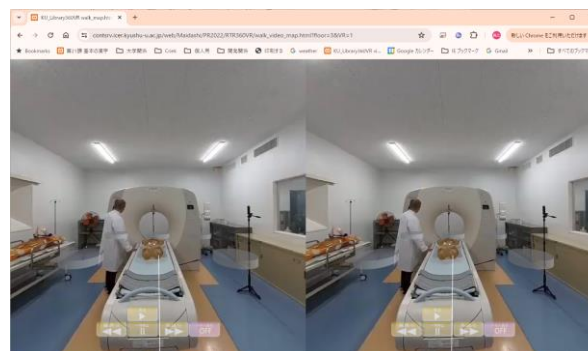


図 11: 360°VR 動画の Web ベースナビゲーション VR コンテンツ例-VR ゴーグルモード

図 4 と同様に、マップ上の赤いマーカーを見る

ことで、自分の現在位置と方向を把握できる。ユーザは、マップ上の対応するオレンジ色の点をクリックすることで、現在の 360° VR 動画を、指定した場所で撮影された動画に変更できる。メイン画面で、灰色の円盤の 1 つをマウスクリックすると、現在の 360° VR 動画がその場所で撮影された動画に変更される。このようにして、ユーザは自分の位置を対話的に変更しながら動画を視聴できる。

図 7 と同様に、各スクリーンショットの中央下部には、再生、一時停止、逆再生、早送りなどの 360° VR 動画を制御するコントロールパネルが表示される。図 11 は、図 10 と同じコンテンツであるが、これは VR ゴーグルモードのスクリーンショットである。図 4、図 7 と同様に、図 10 の各スクリーンショットの右下に[Enter VR]ボタンが表示されている。このボタンを押すと、通常モードから VR ゴーグルモードへ切り替わる。ユーザが VR ゴーグルで Web ブラウザを使用している場合、ボタンを押した後、VR ゴーグル内に図 11 のようなステレオ視の部屋が表示される。図 8 で説明したように、重要な点は、VR ゴーグルモードでは HTML DOM のオブジェクトを表示できないため、画面中央下に表示されるコントロールパネルは HTML DOM でないことである。代わりに、VR ゴーグルモードのコントロールパネルは、WebXR および Three.js の機能によって定義された 3D オブジェクトとして実現されている。VR ゴーグルモードでは、ユーザは VR ゴーグルのコントローラーでこのコントロールパネルの各ボタンを押すことができる。

5 Web ベース多地点 360° VR 動画コンテンツ例

このセクションでは、著者らの提案するフレームワークを使用して実際に作成された Web ベース多地点 360° VR 動画コンテンツを紹介する。図 12、図 13 は、宇宙線測定装置、透過型電子顕微鏡の説明コンテンツである。これらの多地点 360° VR 動画コンテンツの作成には、Insta360 ONE X2 を 5 台使用した。ユーザは、図のスクリーンショットでスタンドを備えたカメラのいくつかを見ることができ、各マップ上の小さな紫色の点を見ることができ、それらの位置を把握できる。また、字幕生成機能によって字幕が表示されている。

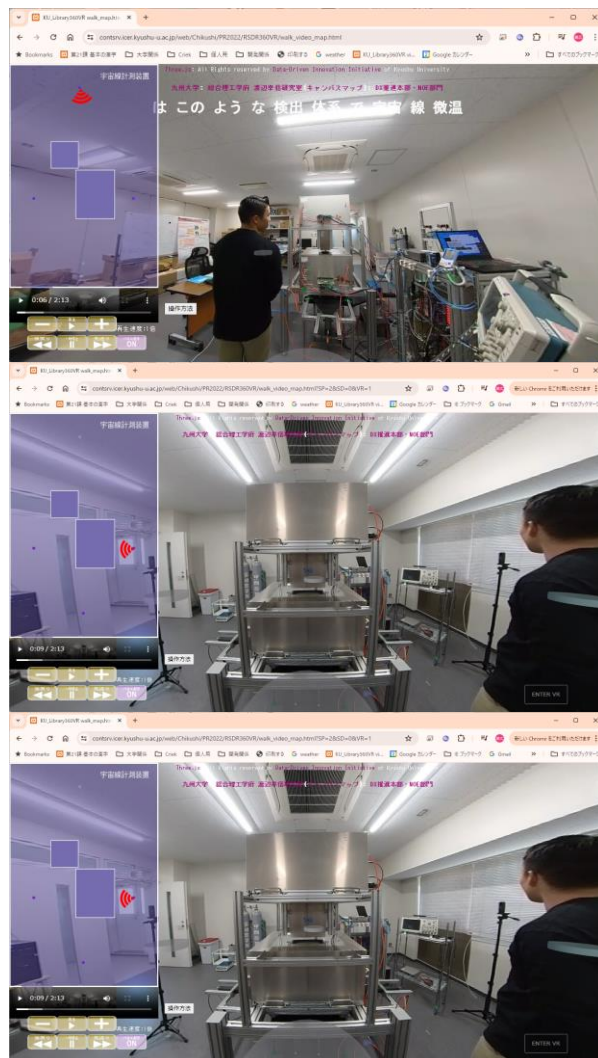


図 12: 宇宙線測定装置の説明コンテンツ

6. おわりに

LiDAR や 360° VR カメラなど、さまざまな種類のスキャンデバイスが研究開発されている。これらのデバイスを使用すると、3D モデル、点群データ (PCD: Point Cloud Data)、360° VR 画像/動画などの 3D データを取得できる。このようなデータを活用した教育用 XR コンテンツの制作に向けて、360° VR 画像/動画の Web ベース VR コンテンツや Web ベースの教育 XR コンテンツの開発フレームワークをこれまでにいくつか提案してきた。本稿では、VR ゴーグル対応 Web ベース多地点 360° VR 動画コンテンツのための新しいフレームワークを提案した。さらに、360° VR 動画のナレーションに字幕機能を実装した。また、提案したフレームワークを利用して実際に開発したコンテンツ例を紹介した。VR ゴーグル対応 Web ベース多地点 360° VR 動画により、実験や

演習の対象を事前に学習できるため、実験や演習を行う多くの学生にとって有用であると考えられる。今後は、手動作業を可能な限り自動化する等のフレームワークの機能強化を図る予定である。

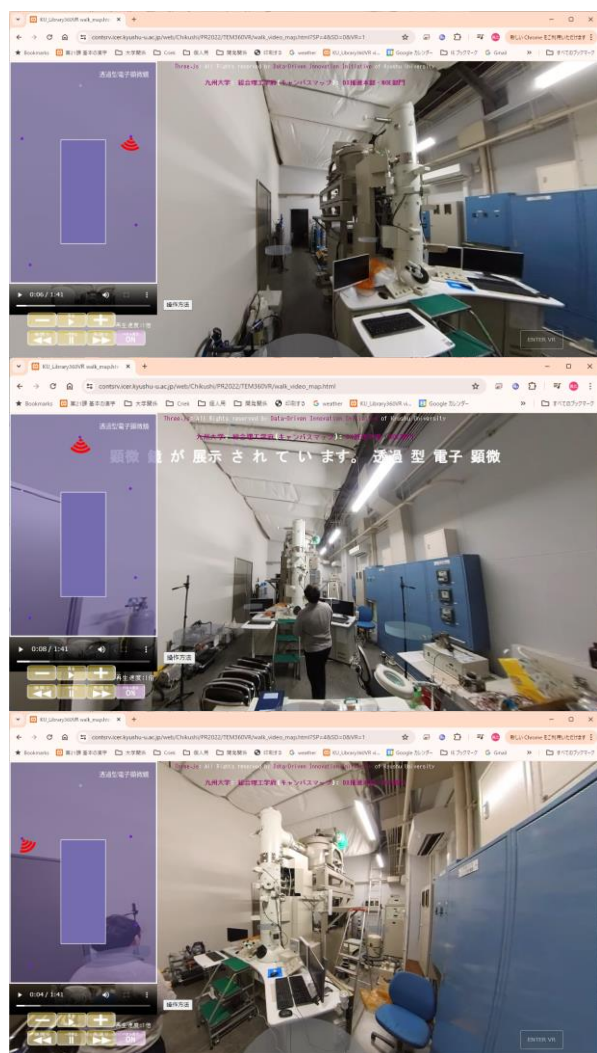


図 13: 透過型電子顕微鏡の説明コンテンツ

謝辞

本研究の一部は、JSPS 科研費 JP22H03705、JP23K24960 および九州大学 VISION EXPO プロジェクトの支援を受けた。

参考文献

- [1] Okada, Y., Kaneko, K., Shi, W. (2023). Development Framework for Web-based VR Tours and Its Examples. 27th International Conference Information Visualisation (IV), pp. 420-425.
- [2] Sumikura, S., Shibuya, M. and Sakurada, K., (2019), OpenVSLAM: A Versatile Visual SLAM Framework. MM '19: Proceedings of the 27th ACM International Conference on Multimedia,

pp. 2292-2295.

- [3] Okada, Y., Kaneko, K., Shi, W. (2023) DEVELOPMENT FRAMEWORK BASED ON OPENVSLAM FOR WEB-BASED VR TOURS USING 360VR VIDEOS, International Conference on WWW/Internet 2023, pp. 109-116.
- [4] Okada, Y., Shi, W., Kaneko, K. (2024), OpenVSLAM-Based Development Framework for Web-Based VR Tours Using 360VR Videos and Its Extensions. Proc. of 12th International Conference on Emerging Internet, Data & Web Technologies (EIDWT 2024), pp. 31-42.