FINAL SCRIPT: APPROVED FOR FILMING

Submission ID #: 68386
Scriptwriter Name: Sulakshana Karkala
Project Page Link: https://review.jove.com/account/file-uploader?src=20857778

Title: Photorealistic Learned Landscapes for Augmented Reality

Authors and Affiliations:
David Martínez-Miranzo*, Sergio Suescun-Ferrandiz*, Miguel Cazorla*, Francisco Gomez-Donoso*

University Institute for Computer Research, University of Alicante

***All authors contributed equally to this work**

Corresponding Authors:
Miguel Cazorla (miguel.cazorla@ua.es)
Email Addresses for All Authors:
David Martínez-Miranzo (david.martinez@ua.es)
Sergio Suescun-Ferrandiz (sergio.suescun@ua.es)
Miguel Cazorla (miguel.cazorla@ua.es)
Francisco Gomez-Donoso (fgomez@ua.es)

Author Questionnaire

1. We have marked your project as author-provided footage, meaning you film the video yourself and provide JoVE with the footage to edit. JoVE will not send the videographer. Please confirm that this is correct.

X Correct

2. Microscopy: Does your protocol require the use of a dissecting or stereomicroscope for performing a complex dissection, microinjection technique, or something similar? **No**

3. Software: Does the part of your protocol being filmed include step-by-step descriptions of software usage? **Yes, all done**

4. Proposed filming date: To help JoVE process and publish your video in a timely manner, please indicate the proposed date that your group will film here: 06/18/2025

When you are ready to submit your video files, please contact our Content Manager, Utkarsh Khare.

Current Protocol Length

Number of Steps: 15
Number of Shots: 26

Introduction

1.1. **Sergio Suescun-Ferrandiz:** The research focuses on creating photorealistic 3D reconstructions from 360° images to build immersive virtual environments. It explores how these can support therapy, education, and industrial validation.

1.1.1. INTERVIEW: Named Talent says the statement above in an interview-style shot, looking slightly off-camera. *Suggested B.roll**:4.1***

What technologies are currently used to advance research in your field?

1.2. **Sergio Suescun-Ferrandiz:** Common approaches include Structure-from-Motion and Multi-View Stereo, often using tools like COLMAP. Recent methods like Gaussian Splatting offer faster, more visually realistic results for immersive applications.

1.2.1. INTERVIEW: Named Talent says the statement above in an interview-style shot, looking slightly off-camera.

What are the current experimental challenges?

1.3. **David Martínez Miranzo:** Challenges include the computational cost of Gaussian Splatting, handling low-texture scenes, and ensuring accurate camera pose estimation, which directly affects reconstruction quality.

1.3.1. INTERVIEW: Named Talent says the statement above in an interview-style shot, looking slightly off-camera. *Suggested B.roll:2.9*

What research gap are you addressing with your protocol?

1.4. **David Martínez-Miranzo:** We address the lack of accessible, photorealistic 3D reconstruction systems that integrate smoothly into VR and support dynamic, user-driven interaction without costly sensors.

1.4.1. INTERVIEW: Named Talent says the statement above in an interview-style shot, looking slightly off-camera.

Protocol

2. 360° Image Acquisition and Processing for 3D Reconstruction

Demonstrator: Sergio Suescun-Ferrandiz

2.1. To begin, place a 360-degree camera on a tripod that has adjustable height **[1]**. Select a series of positions within the environment, for scanning, using a square mesh pattern where each edge is spaced 1.5 meters apart **[2]**.

2.1.1. WIDE: Talent adjusting the tripod to a suitable height and mounting the 360-degree camera.

2.1.2. Talent marking or moving to the designated mesh points in the scanning environment.

2.2. At each mesh point, capture images at three different heights of approximately 0.4 meters, 1.2 meters, and 2 meters **[1]**.

2.2.1. Talent capturing images at the specified heights using the camera, adjusting its position accordingly.

2.3. Convert the 360-degree images to equirectangular format using a tool like the **Insta360** *(Instah-Three-Sixty)* app **[1]**. Select the image, press the **Export** button, choose **Export 360 photo** mode, and export it as a 2 to 1 ratio image **[2]**.

2.3.1. SCREEN: 2.3.1.mp4 00:00-00:08

2.3.2. SCREEN: 2.3.2.mp4 00:00-00:23

2.4. Use the Equi2Pers.py *(E-kwi-Two-Purrs-dot-P-Y)* script to extract 16 to 9 format perspective images from each equirectangular image with a 90-degree horizontal field of view **[1]**. Apply horizontal angles of 0, 45, 90, 135, 180, 225, 270, and 315 degrees, and vertical angles based on height **[2-TXT]**.

2.4.1. SCREEN: 2.4.1,-2.4.2.mp4. 00:16-00:35

2.4.2. SCREEN: 2.4.1,-2.4.2.mp4 00:00-00:16 **TXT: Vertical angle: 0, 50 for 0.4 m, -50, 0,50 for 1.2 m, and -50, 0 for 2 m**

2.5. Next, click on **File** and **New Project** to create a new COLMAP *(coal-map)* project **[1].** Specify the path to the images and create a new database **[2].**

2.5.1. SCREEN: 2.5.1.mp4 00:00-00:06

2.5.2. SCREEN: 2.5.2.mp4 00:00-00:13

2.6. Click on **Processing** followed by **Feature Extraction** to extract features for each image **[1].** Select **PINHOLE** as the camera model and share all the images. Leave remaining parameters as default **[2].**

2.6.1. SCREEN: 2.6.1.mp4 00:00-00:04

2.6.2. SCREEN: 2.6.2.mp4 00:00-00:08

2.7. Compute Structure from Motion by clicking on **Reconstruction** and **Start Reconstruction** to obtain the camera positions and orientations, using the default COLMAP parameters **[1].** Click on **Reconstruction** and choose **Bundle Adjustment** to minimize the reprojection errors **[2].**

2.7.1. SCREEN: 2.7.1.mp4 00:00-00:22

2.7.2. SCREEN: 2.7.2.mp4 00:00-00:11

2.8. Now generate a dense 3D scene representation by choosing **Dense Reconstruction** with outputs including camera poses and reconstructed points **[1].**

2.8.1. SCREEN: 2.8.1.mp4 00:00-00:14

2.9. For photorealistic 3D scene reconstruction using Gaussian splatting, execute the train.py *(Train-dot-P-Y)* script using the parameters -s *(minus-S)* , -m, and -r **[1-TXT]**. Locate the generated .ply *(dot-P-L-Y)* file within the specified output directory for subsequent import into Unity **[2]**.

2.9.1. SCREEN: 2.9.1.mp4 00:00-00:10
TXT: -s : COLMAP projection path; -m: path output for Gaussian splatting; -r : rescales image between 2 and 4

2.9.2. SCREEN: 2.9.2.mp4 00:00-00:07

3. Rendering in Virtual Reality with Unity and Gaussian Splatting

Demonstrator: David Martínez-Miranzo

3.1. Connect the virtual reality headset to the computer, following the specific instructions for the headset model used **[1]**.

3.1.1. Talent plugging in the VR headset and confirming connection on the desktop interface.

3.2. Use Unity Hub to create a 3D project with version 2022.3.44f1 *(Two-thousand-Twenty-Two-Dot-Forty-Four-F-One)* **[1-TXT]**. Navigate to **Projects**, click **New Project**, select the **3D (Built-In Render Pipeline)** *(3-D-Built-In-Render-Pipeline)* template, set the project name and location, and click **Create Project [2]**.

3.2.1. SCREEN: clip_3_2_1.mp4 00:00-00:06, 00:15-00:23 **TXT: Install Unity ver. 2022.44f1 if necessary**

3.2.2. SCREEN: clip_3_2_2.mp4 00:00-00:18

3.3. To manage the VR headset and simplify development tasks, install a plugin from the Unity Asset Store via the **Package Manager** by clicking **Window** and **Package Manager [1]**. Use the **Unity Gaussian Splatting** plugin to convert the Gaussian Splatting output into a usable asset **[2]**.

3.3.1. SCREEN: clip_3_3_1.mp4. 00:00-00:28

3.3.2. SCREEN: clip_3_3_2.mp4 00:37-00:54

3.4. Improve hand tracking by installing the **UltraLeap** *(Ultra-Leap)* plugin via the **Package Manager** from the Unity Asset Store **[1]**. Transcribe audio from the VR headset microphone using the **whisper.unity** *(whisper-dot-Unity)* plugin. Install it using **Package Manager [2]**.

3.4.1. SCREEN: clip_3_4_1.mp4. 00:09-00:33

3.4.2. SCREEN: clip_3_4_2.mp4 00:04-00:16,00:47-00:52

3.5. Enable response generation using a large language model by installing the **LLMUnity** *(L-L-M-Unity)* plugin. Install it through the **Package Manager** as demonstrated earlier **[1]**.

3.5.1. SCREEN: clip_3_5_1.mp4 00:43-00:55, 01:12-01:24

3.6. Generate speech from LLM-generated responses using the **Meta - Voice SDK** *(Meta-Voice-S-D-K)*. Install a Text-to-Speech plugin from the Unity Asset Store via the **Package Manager** by clicking **Window** and **Package Manager [1]**. Finally use the VR headset to experience and interact with the immersive environment **[2]**.

3.6.1. SCREEN: clip_3_6_1.mp4 00:11-00:28, 00:51-00:57,

3.6.2. Talent wearing the VR headset, looking around and interacting with the rendered 3D scene.

Results

4. Results

4.1. Groups of camera positions derived from shared equirectangular origins were used to generate dense point clouds for scene reconstruction, revealing a consistent spatial mapping of capture angles **[1].**
4.1.1. LAB MEDIA: Figure 3. *Video editor: Highlight the red camera icons*

4.2. The proposed method using Gaussian Splatting produced a photorealistic reconstruction closely resembling the real environment **[1].** Users could effectively interact with the reconstructed environment through virtual reality, maintaining immersion and spatial awareness, with visuals presented inside the headset matching the room setup **[2].**
4.2.1. LAB MEDIA: Figure 4. *Video editor: Please highlight 4B*
4.2.2. LAB MEDIA: Figure 5. *Video editor: Show Figure 4 A and then B*

4.3. Familiar and unfamiliar virtual environments were developed with therapeutic goals in mind, based on feedback from professional therapists **[1].** A virtual agent was rendered in the reconstructed space, allowing users to engage in realistic interactive scenarios through VR, with the agent appearing as a lifelike figure in the headset view **[2].**
4.3.1. LAB MEDIA: Figure 6. *Video editor: Highlight A and C then B and D*
4.3.2. LAB MEDIA: Figure 7. *Video editor: Emphasize B*

4.4. Virtual reconstructions replicated specific viewpoints accurately when based on input images **[1],** but deviations in perspective resulted in noticeable rendering limitations **[2].**
4.4.1. LAB MEDIA: Figure 8. *Video editor: Highlight the top images (A and B), showing a real and virtual view of the same room angle.*
4.4.2. LAB MEDIA: Figure 8. *Video editor: Highlight the bottom images (C and D), comparing a new real view with its less accurate virtual rendering.*

4.5. Compared to COLMAP point cloud output, Gaussian Splatting produced more visually continuous and lifelike reconstructions suitable for real-time interaction, albeit with reduced metric precision **[1].**
4.5.1. LAB MEDIA: Figure 9. *Video Editor: Please show Figure A first and then B*

Pronunciation Guide:

1. Gaussian

- **Pronunciation link**: https://www.howtopronounce.com/gaussian
- **IPA**: /ˈgaʊ.si.ən/
- **Phonetic Spelling**: GOW-see-uhnyoutube.com+9youglish.com+9howtopronounce.com+9dictionary.cambridge.org

---

2. COLMAP

- **Pronunciation link**: https://www.howtopronounce.com/colmap
- **IPA**: /ˈkɒl.mæp/
- **Phonetic Spelling**: KOL-map

---

3. Equirectangular

- **Pronunciation link**: https://www.howtopronounce.com/equirectangular
- **IPA**: /ˌɛk.wɪ.rɛkˈtæŋ.gjə.lɚ/
- **Phonetic Spelling**: EK-wih-rek-TANG-gyuh-lurnameslook.com+6howtopronounce.com+6forvo.com+6

---

4. Unity

- **Pronunciation link**: https://dictionary.cambridge.org/us/pronunciation/english/unity
- **IPA**: /ˈjuː.nə.t̬i/
- **Phonetic Spelling**: YOO-nuh-teecollinsdictionary.com+5dictionary.cambridge.org+5dictionary.cambridge.org+5englishgrammarzone.com+2dictionary.cambridge.org+2dictionary.cambridge.org+2

---

5. Ultraleap

- **Pronunciation link**: No confirmed link found
- **IPA**: /ˈʌl.trə.liːp/
- **Phonetic Spelling**: UL-truh-leapsupport.ultraleap.com+3support.ultraleap.com+3robotshop.com+3