


# World Labs Marble AI

How-To Guide, Tips & Best Practices  
*Building 3D Worlds with Spatial Intelligence AI*

## What Is a World Model?

A world model is a system that learns and understands the geometry, physics, lighting, and real-world spatial properties of a scene — similar to how the human brain builds an internal map of its environment.

Unlike traditional image generators (like Sora) that predict the next pixel, world models understand 3D structure. They can infer what the rest of a room looks like from partial data.

 **Key Insight:** World Labs' Marble AI, created under the direction of Dr. Fei-Fei Li, is one of the first commercially available world models that produces true 3D output — not just visual imagery.

## The 6 Input Methods in Marble

Marble accepts six different types of inputs to generate a 3D world. Each method trades off ease of use against accuracy and detail.

Method	Input Required	Accuracy	Best For
1. Text Prompt	Descriptive text	Low	Creative, imaginative scenes
2. Single Image	1 photo	Medium	Quick reconstructions
3. Multi-Image (up to 8)	2–8 photos	Medium-High	More complete coverage
4. Video Walkthrough	30-second video	Medium	Wide scene capture
5. 360 Panorama	360° camera image	High	Faithful spatial capture
6. Chisel Mode	Manual 3D layout	Variable	Custom/designed scenes

### Method 1: Text Prompt Only


Type a description of the scene you want to create. Marble will generate a 3D world based purely on the text.

### Tips

- Use detailed, specific prompts. Vague prompts produce generic results.
- Try using another AI (like ChatGPT) to generate a rich scene description from your photos, then paste that into Marble.
- Include specific details: cabinet colors, flooring type, room dimensions, architectural features.

### Limitations

- The AI has no awareness of your actual space — it will invent details freely.
- Complex features like kitchen islands may be omitted or misinterpreted.

 **Best Use:** Text prompts are ideal when you want an inspired, AI-designed scene rather than a faithful recreation of a real space.

## Method 2: Single Image

---


Upload one photo and let Marble infer the full 3D scene from that single viewpoint.

### Tips

- Choose a photo that captures as much of the room as possible from one angle.
- Include important architectural elements (doorways, windows, distinctive features) in frame.
- The AI will predict what's outside the frame — expect creative filling of unseen areas.

### Limitations

- Areas not visible in the photo will be guessed and may not match reality.
- Scene quality degrades quickly when navigating far from the original camera angle.
- The 'Expand' feature to enlarge the navigable area can be unreliable — try it multiple times if it fails.

 **Best Use:** Single image works well for capturing the primary view of a room when exact accuracy beyond that view is not needed.

## Method 3: Multi-Image (Up to 8 Photos)

---

Upload multiple photos (up to 8) from different angles to give Marble more spatial context.

## Tips

- Vary your angles significantly — include shots from opposite sides of the room.
- Capture specific features you care about from close range (countertops, shelving, windows).
- Don't worry about capturing every inch — the AI will fill in gaps.
- Including photos of floors and ceilings helps ground the geometry.

## Limitations

- Freestanding objects in the middle of a room (like kitchen islands) may still be problematic.
- The AI may take creative liberties with materials and textures despite clear photo references.
- Editing inaccuracies afterward is limited — you can select an area and provide a reference image, but results may be minor.

**Recommended:** Multi-image is the sweet spot for most users — significantly better accuracy than a single image, without the complexity of video or 360 camera setup.

## Method 4: Video Walkthrough

---


Record a short video (up to 30 seconds) walking through your space and upload it to Marble.

## Tips

- Start your video facing the area you most want to capture accurately — Marble prioritizes the initial direction.
- Move slowly and steadily to avoid motion blur.
- Keep your video focused on the primary space; avoid sweeping across adjacent rooms unless you want them included.
- Consider using video as a 180–360 pan from a fixed point rather than walking — stand in the center of the room and rotate.

## Limitations

- The model seems to weight the initial camera direction heavily, so rooms you look at briefly may not render well.
- 30-second limit constrains how much you can capture.
- More frames does not always mean better accuracy — it can introduce conflicting data.

 **Caution:** Video can produce inconsistent results. If your video starts looking at the wrong area, that area may dominate the output. Consider your starting direction carefully.

## Method 5: 360 Panorama

---

Use a 360° camera (or a 360 panoramic image) to capture the full environment in a single shot.

### Tips

- Place the camera at the center of the space you want to capture — ideally where you'd want a virtual viewer to stand.
- Eliminate moving objects (people, pets) from the scene when possible.
- Ensure adequate, even lighting to avoid shadows that confuse depth estimation.
- The panorama excels at capturing the perimeter of a room — walls, windows, and upper features.

### Limitations

- Objects beneath the camera (floors directly under it, freestanding islands directly below) may not capture well.
- The mesh collider output from panoramas can be coarse and uneven — not ideal for robotics simulation or game engines without further processing.
- Adjacent rooms or spaces outside the camera's field will be blank.

**Recommended:** 360 panorama produces the most spatially faithful single-shot world. Best for virtual walkthroughs, virtual staging, and background environments.

## Method 6: Chisel Mode (3D Layout Design)

---


Chisel is Marble's advanced mode that lets you manually define a 3D spatial layout by extruding shapes and building a scene from scratch before the AI 'fills it in.'

### Tips

- Think of Chisel as a rough architectural sketch tool — define volumes, not details.
- Extrude shapes from floor to ceiling to represent walls and large objects.
- Use it when you want a custom-designed space rather than a recreation of a real environment.
- Pair Chisel with a text prompt (e.g., 'modern kitchen') to guide the aesthetic fill.

### Limitations

- Controls can feel unintuitive for users without 3D modeling experience.
- Occlusion objects (freestanding items in the middle of the room) are difficult to represent faithfully.
- The AI may interpret your layout unexpectedly — e.g., converting short walls into full-height structures.

 **Best Use:** Chisel is powerful for simulation environments, game design prototyping, or when you want to design a fictional space and have AI bring it to life visually.

## Editing and Expanding Your World

After generating a world, Marble provides tools to refine and extend it.

### Expanding the Scene


---

- Use the 'Expand' feature to extend the navigable area beyond the original generation boundary.
- Expansion can fail — retry multiple times if it doesn't work on the first attempt.
- Expansion works best when the scene boundary is clearly defined and not overly complex.

### Editing via Panorama Mode

---

- Open your world in Panorama mode to make targeted edits.
- Select a region of the panorama and upload a reference image to guide changes.
- Include a text description of what you want updated (e.g., 'update the cabinet style to dark wood').
- Note: You can only edit from the fixed panorama viewpoint — you cannot navigate and edit simultaneously.

 **Known Limitation:** Editing results can be subtle or unpredictable. Marble's editing tools are still maturing. For major changes, regenerating from better inputs is often more effective.

## Downloading and Exporting Your World

Marble supports multiple export formats for different use cases:

### Export Options

---

- Screenshot / Image — Export a still view from any angle within the world.
- Gaussian Splats — A 3D representation format ideal for immersive rendering and virtual environments.

- Collider Mesh — A simplified 3D mesh suitable for game engines and simulation platforms (e.g., NVIDIA Isaac Sim/Lab).

## Using Exports


- For virtual staging or background use: Gaussian Splat is the best choice.
- For robotics training or physics simulations: export the Collider Mesh, then import into your simulation environment (note: mesh quality may require cleanup).
- Collider meshes from panoramas tend to be coarser — the floor may be uneven and require flattening in a 3D tool like MeshLab.
- Meshes often import upside down — rotate 180° on the X-axis to correct orientation.

## Understanding Marble's True Purpose

If you're hoping to create a perfect replica of a real room, you may be disappointed — that's not Marble's primary strength. Dedicated scanning tools (like Portal Cam) with an iPhone are better for exact replication.

Marble's power is in generating plausible, immersive variations of spaces for:

- Simulation & Robotics Training — Generate 100 variations of a kitchen to train a robot on diverse data sets without rebuilding scenes manually.
- Game Development Prototyping — Rapidly create explorable environments from minimal inputs.
- Virtual Production / Film — Create inspired environments that look similar to a reference location but with creative modifications.
- Architectural Visualization — Explore 'what if' spatial designs quickly.

 **Reframe Your Expectations:** Think of Marble not as a 3D scanner, but as a spatial imagination engine. It generates worlds inspired by your inputs — not copies of them. That distinction unlocks its true potential.

## Quick Reference: Best Practices


Situation	Recommended Approach
I want to recreate a real room accurately	Use 8 photos from varied angles + 360 panorama if you have the camera
I want a custom designed space	Use Chisel Mode + text prompt for aesthetic guidance

Situation	Recommended Approach
I need scenes for simulation/robotics	Generate multi-image or 360 worlds, export Collider Mesh, clean up in MeshLab
I need many variations of a space	Generate one world from a 360 or multi-image, then use it as a base for further text-guided variations
I want a quick creative environment	Use text prompt — accept that it will be inspired by your description, not exact
My scene is missing an area	Try Expand (retry multiple times) or regenerate with a photo covering the missing area
I want to fix a specific part of the scene	Use Panorama edit mode — select the area, provide a reference image and text instruction
My Collider Mesh is upside down	Rotate 180° on the X-axis in MeshLab or your 3D tool

## Known Limitations (V1)

Marble is a first-generation product. Keep these current limitations in mind:

- Occlusions (freestanding objects in the center of a space) are consistently difficult — kitchen islands, display cases, and similar objects often render poorly or are omitted.
- The 'Expand' feature is unreliable — multiple retries may be needed, and it sometimes makes no change or degrades the scene.
- Panorama editing is position-locked — you can only edit from the original panorama viewpoint.
- Collider meshes are coarse and require cleanup for use in physics simulations or game engines.
- The model takes creative liberties with materials, textures, and features not clearly visible in inputs.

 **Future Outlook:** As a V1 product, Marble is expected to improve significantly. Think of ChatGPT's early days — current limitations reflect early-stage development, not the ceiling of what world models can achieve.