MemoVis: A GenAI-Powered Tool for Creating Companion Reference Images for 3D Design Feedback

Chen Chen[1], Cuong Nguyen[2], Thibault Groueix[2], Vladimir Kim[2], Nadir Weibel[1]
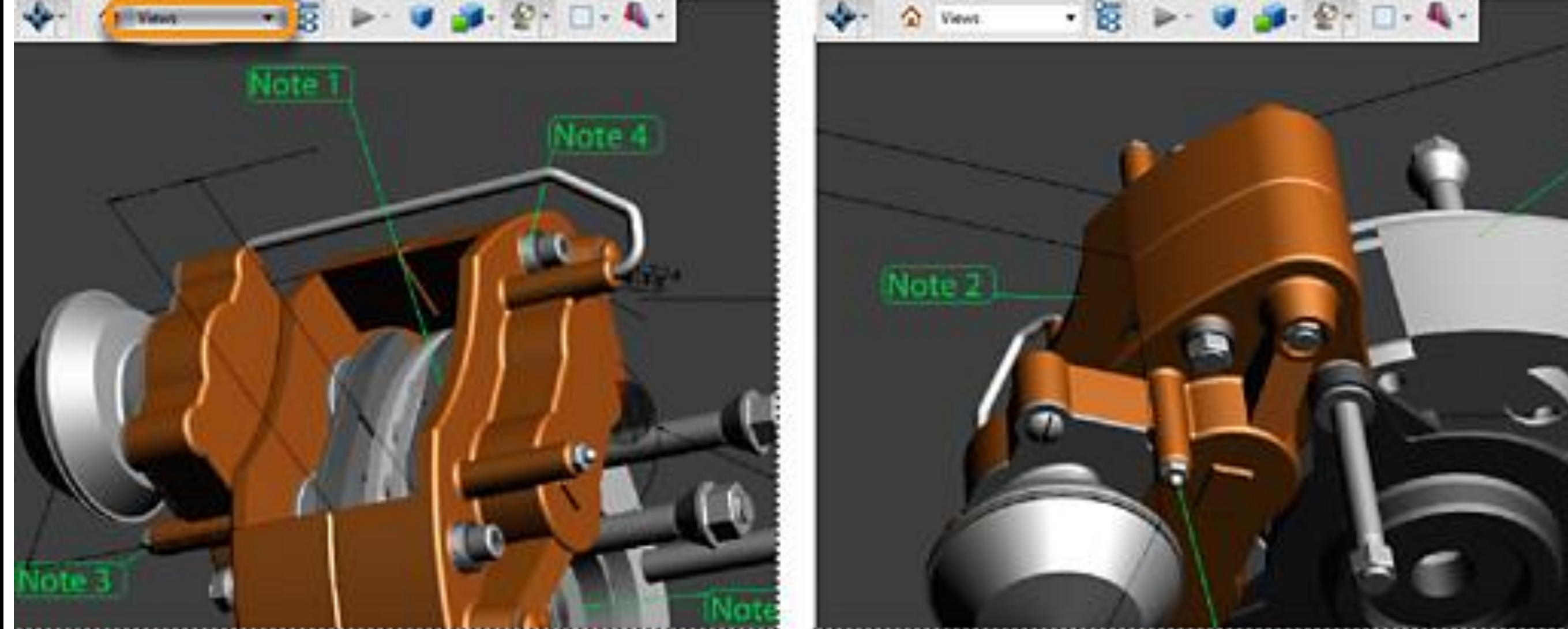
[1]University of California San Diego , [2]Adobe Research

Email: chenchen@ucsd.edu | Website: chen-chen.me

📖 Paper   ▶ Video

Creating Feedback for 3D Design Review

Feedback in 3D design review workflow (Acrobat Example 📕)

- Navigating viewing camera by mouse can be challenging.
- Texts alone can be difficult to convey the gist of feedback.
- ❓ How to create reference image(s) for 3D design feedback?

Formative Studies

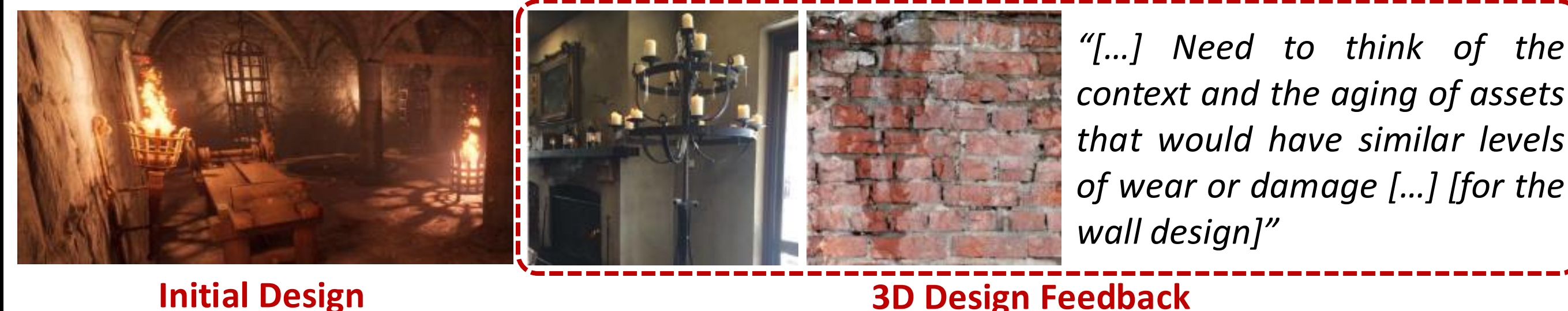### Study 1: Preliminary Needs Finding Study (N = 2 Designers)
- Creating reference image starts with finding supporting views.
- Conveying changes requires extra work on the reference images.
- Control for reference image generations.

Study 2: Analysis of Real-World Online 3D Design Feedback

99 posts from 15 creators and 36 feedback providers. 😀 **polycount** *The Videogame Art Community*

Key findings:
- Reference images are important to complement textual comments, but creators might need additional "imaginations" to transfer the gist of visual imagery.

Initial Design   **3D Design Feedback**

"[…] Need to think of the context and the aging of assets that would have similar levels of wear or damage […] [for the wall design]"

- Suggestions conveyed by feedback providers can be either the revision of specific parts or redesign of the entire assets.

Initial Design   **3D Design Feedback**   **Initial Design**   **3D Design Feedback**

"[…] Maybe use a belt instead, with the blades going in, or some knife block as can be found in some kitchens"

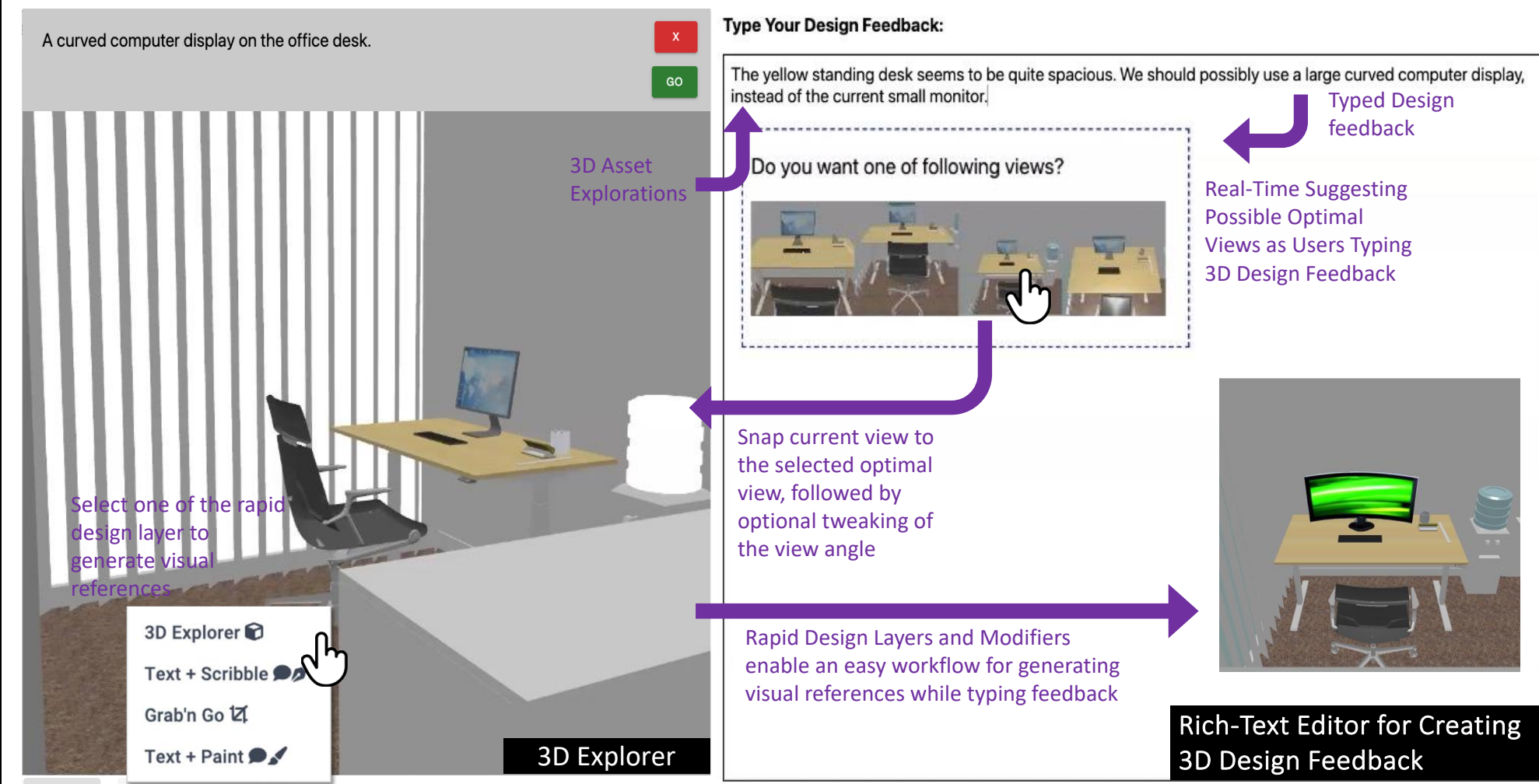*"just think about how this gun is used, and where wear and damage would be […]"*

- Although some design feedback provides actionable solutions, others offer potential exploratory directions.

Key Design Considerations
- Controlled visualization for both local & global changes
- Creativity support for exploratory design feedback
- Offer ways for feedback providers without 3D/image editing skills to create in-context visualizations

⚙ MemoVis System

- GenAI driven
- Browser-Based Rich Text Editor Interface
- Can be integrated into 3D software, web app, & document tools.

Real-Time Viewpoint Suggestions

f_{image}: CLIP Image Encoder
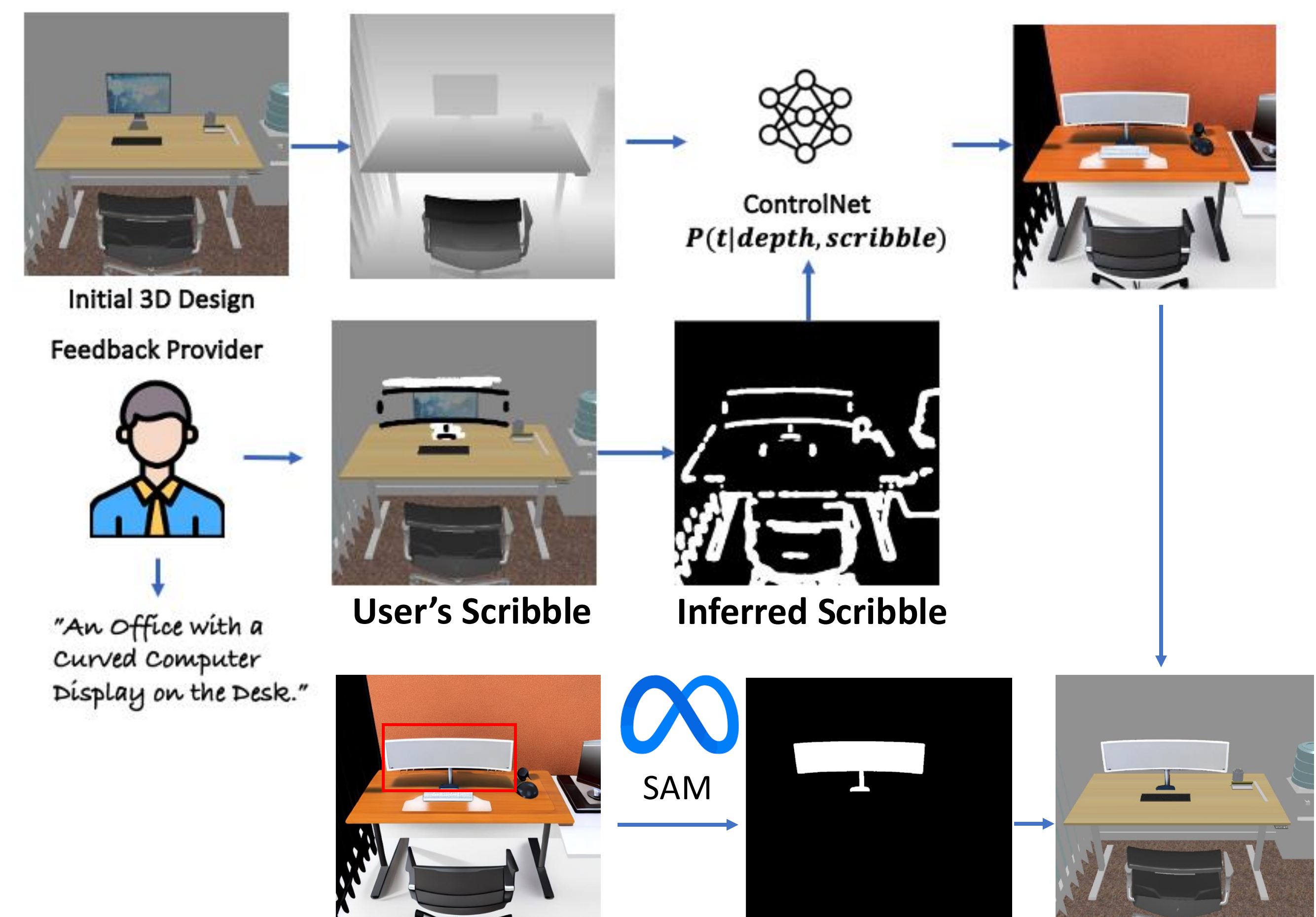f_{text}: CLIP Text Encoder

- **Preprocessing**

$$EncodeI_v = f_{image}(I_v), v = \{\alpha, \beta, r, t_x, t_y, t_z\} \quad EncodeI \in R^{27K \times 512}$$

- **Real-Time Inference**

$$v = argtopk_{v, k=4} \cos\{f_{text}(t), EncodeI_v\} \quad t: \text{textual feedback comment}$$

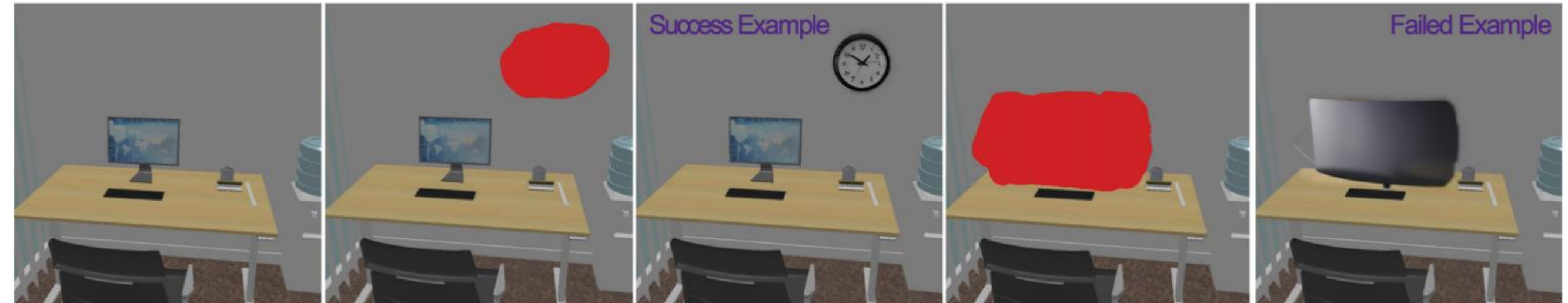### Creating Reference Images with Rapid Image Modifiers

- **Text + Scribble Modifier with Scribble Design Layer**

Initial 3D Design   ControlNet $P(t|depth, scribble)$

Feedback Provider

"An Office with a Curved Computer Display on the Desk."

User's Scribble   **Inferred Scribble**

SAM

- **Grab'n Go Modifier with GenAI Design Layer**

- **Text + Paint Modifier with Painting Design Layer**

Success Example   Failed Example

Results

Creating Reference Images (N = 14)

Navigating the View (Baseline)
Navigating the View (MemoVis)
Image Creations (Baseline)
Image Creations (MemoVis)
Image Explicitness (Baseline)
Image Explicitness (MemoVis)
Creativity Support (Baseline)
Creativity Support (MemoVis)

Participant Count

Assessing Reference Images (N = 8)
How well the reference image could convey the gist of textual comments?

	Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree
Baseline	18.18%	22.73%	20.45%	20.45%	18.19%
MemoVis	20.51%		28.21%		38.46%

% of Design Feedback

Strongly Agree | Agree | Neither Agree Nor Disagree | Disagree | Strongly Disagree

Examples with MemoVis   Reference Images / Anchored View of the Initial Design

Add an Indoor Plant | Modify the wall design | Add a wall poster | Change to suspended bed | Change the color to be more metallic | Add a pet seat

Examples without MemoVis   Reference Images / Anchored View of the Initial Design