Agent Behavior as a Driver for Novelty and Creativity in Human-Robot Co-Drawing

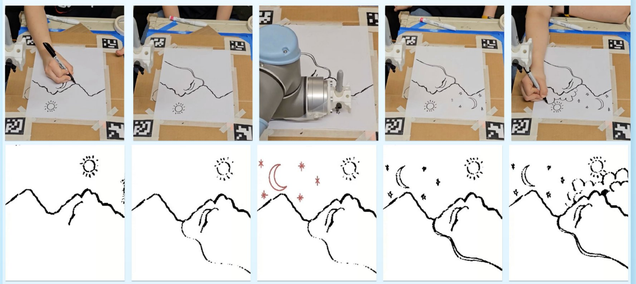
Angelica M. Bonilla Fominaya Daehwa Kim Jack Mirenzi Anisha Jog
Carnegie Mellon University

Background

- Generative AI systems have the potential to provide unique suggestions or ideas during creative workflows.
- Many systems rely on language as a modality for deterministic instruction, which undermines ideation and true collaboration [1, 2].
- Existing co-drawing systems that rely on turn-based interaction [3,4,5] use language instructions to direct the agent, functioning as linear executors.

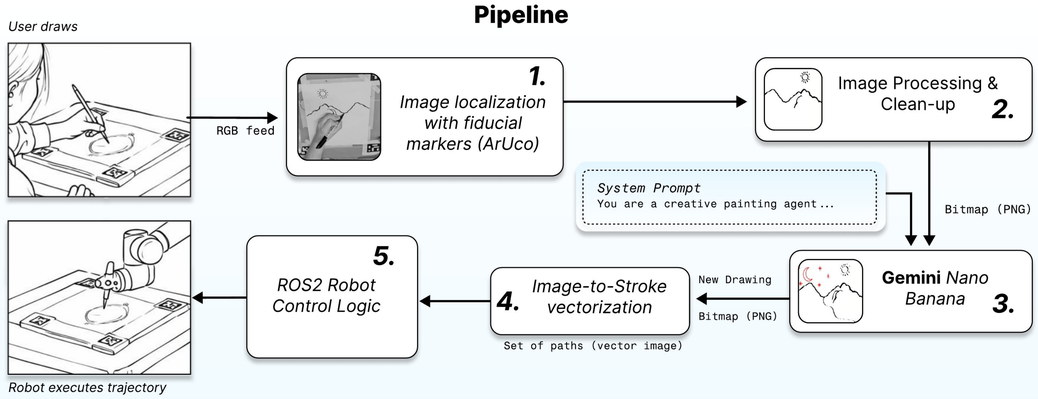
Proposed Solution

A turn-based, generative, co-drawing system that allows for the construction of complex, visually, and semantically-aligned images without direct user instructions.



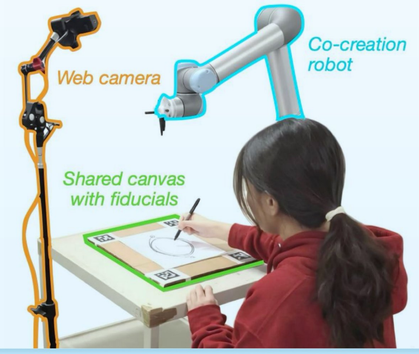
Progression of a collaborative drawing in an in-person setting (top) contrasted with corresponding internal system representation (bottom) of current canvas state (black) and future robot strokes (red).

Pipeline



In a given robot turn:

- **Capture:** An overhead camera records the canvas once the human finishes drawing.
- **Pre-process:** The image is filtered into a clean black-and-white bitmap.
- **Generate:** The bitmap and behavioral prompt are sent to the VLM [8] to plan the next stroke.
- **Vectorize**: The model's output is converted into robot motion paths via Canny edge detection and skeletonization.
- **Execute:** The UR5 cobot draws the trajectory and returns home to pass the turn back.

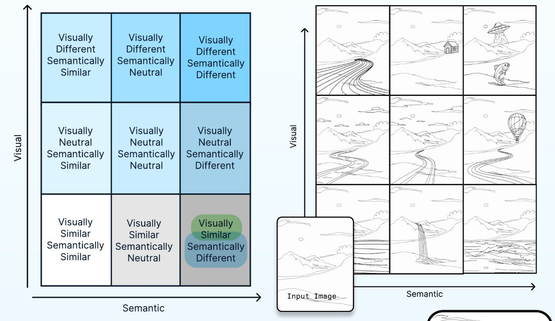


Describing Agent Behavior Via Prompt Design

We steer the robot's actions through descriptions of ‘Conceptual Shifts’:

- **Visual Shifts** (Seeing-That): Alterations in line quality, geometry, texture, or composition.
- **Semantic Shifts** (Seeing-As): Changes to meaning, narrative context, or symbolism (e.g., turning a field into alien planets).

To describe a range of behaviors, we rely on intentional visual/semantic dissonance/congruence to provoke lateral problem-solving and break creative fixation.




```
For each attribute of the drawing, the style ranges
from very different, related but not identical
(slightly different), to very similar. For this
drawing, use the following attributes in your
response:
 1. Very different in concept or meaning.
 2. Very similar in composition, balance or layout.
```

Preliminary Evaluation

We performed two preliminary studies with online *(n=20)* and in-person *(n=4)* participants).

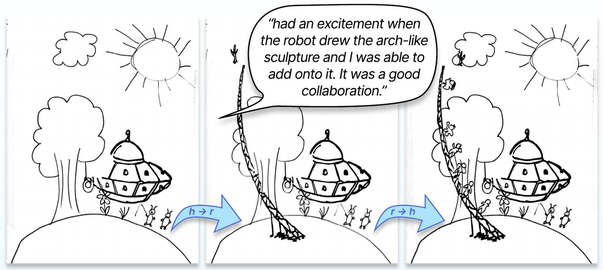
Task: Open collaborative landscape drawing over alternating turns.

Conditions:

- **Collaborative**: Agent matches user style and theme *(visually, semantically similar)*
- **Adversarial**: Agent introduces high visual and conceptual divergence *(visually, semantically different)*.

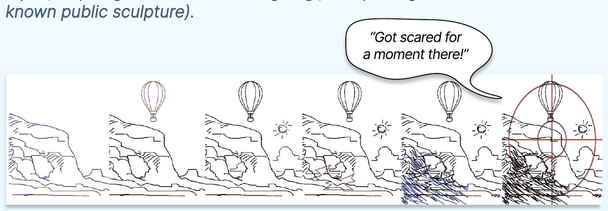
Online (O) Session	In-Person (IRL) Focus Group
20 participants Browser-based painting app with tablet and stylus Black colored pen tool	4 participants Robot and physical canvas Diverse drawing implements (Color sharpie, highlighters, pencil..)
Output: Turn-by-turn artifacts, Post-session survey	

Observations



Derailment as a Drive for Lateral Thinking

Semantic pivots pushed users out of passive drawing into active narrative repair, inspiring creative re-steering (e.g., interpreting an abstract arch as a known public sculpture).



Productive Tension vs. Threat

Extreme, opaque overrides (such as drawing a target over the human's work) were perceived as hostile rather than creative, highlighting the need for legible robot intent.

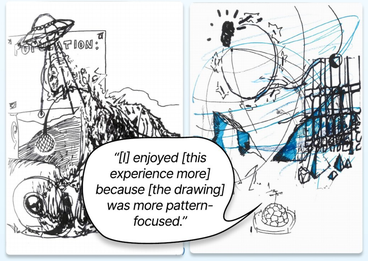
Open-ended Interaction

Different Approaches to Collaboration

Open-Ended Users: Minimal initial strokes to let the robot lead and provoke directions.
Goal-Driven Users: Anchored scenes early to constrain the robot's disruption space.

Materials Matter

Access to multiple colors and textures allows for role disambiguation.



Color and texture are used to indicate layers, organize, and retain creative ownership

Key Takeaway & Future Work

- Co-drawing robots need to balance serendipity and noise: surprise stimulates creativity, but uninterpretable disruption harms trust.
- Equipping the robot with different tools or exploring multi-medium drawing interactions is valuable.
- Formal evaluation of the system is still needed.

References

[1] Zainab Salma, Raquel Hij´on-Neira, and Celeste Pizarro. Designing co-creative systems: Five paradoxes in human-ai collaboration. Information, 16(10), 2025. ISSN 2078-2489. doi: 10.3390/info16100909. URL https://www.mdpi.com/2078-2489/16/10/909.
[2] Alexandra W.D. Bremers. How can a robot help people draw? In Companion Publication of the 2022 ACM De- signing Interactive Systems Conference, DIS '22 Compan- ion, page 21–23, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450393591. doi: 10.1145/3532107.3532876. URL https://doi.org/10.1145/3532107.3532876.
[3] Jookyung Song, Mookyoung Kang, and Nojun Kwak. Sketcherx: Ai-driven interactive robotic drawing with diffusion model and vectorization techniques. (arXiv:2409.15292), 2024. doi: 10.48550/arXiv.2409.15292. URL http://arxiv.org/abs/2409.15292. arXiv:2409.15292 [cs].
[4] Peter Schaldenbrand, Gaurav Parmar, Jun-Yan Zhu, James McCann, and Jean Oh. Cofrida: Self-supervised fine-tuning for human-robot co-painting, 2024. URL https://arxiv.org/ abs/2402.13442.
[5] Tomas Lawton, Kazjon Grace, and Francisco J Ibarrola. When is a tool a tool? user perceptions of system agency in human–ai co-creative drawing. In Proceedings of the 2023 ACM Designing Interactive Systems Conference, DIS '23, page 1978–1996, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450398930. doi: 10.1145/3563657.3595977. URL https://doi.org/10.1145/ 3563657.3595977
[6] Pegah Karimi, Nicholas Davis, Mary Lou Maher, Kazjon Grace, and Lina Lee. Relating cognitive models of design creativity to the similarity of sketches generated by an ai partner. In Proceedings of the 2019 Conference on Creativity and Cognition, Camp;C '19, page 259–270, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 978-1-4503-5917-7. doi: 10.1145/3325480.3325488. URL https://dl.acm.org/doi/10.1145/3325480.3325488.
[7] Google. Gemini. Accessed: September 26, 2025.