

Designing Generative AI for Iteration

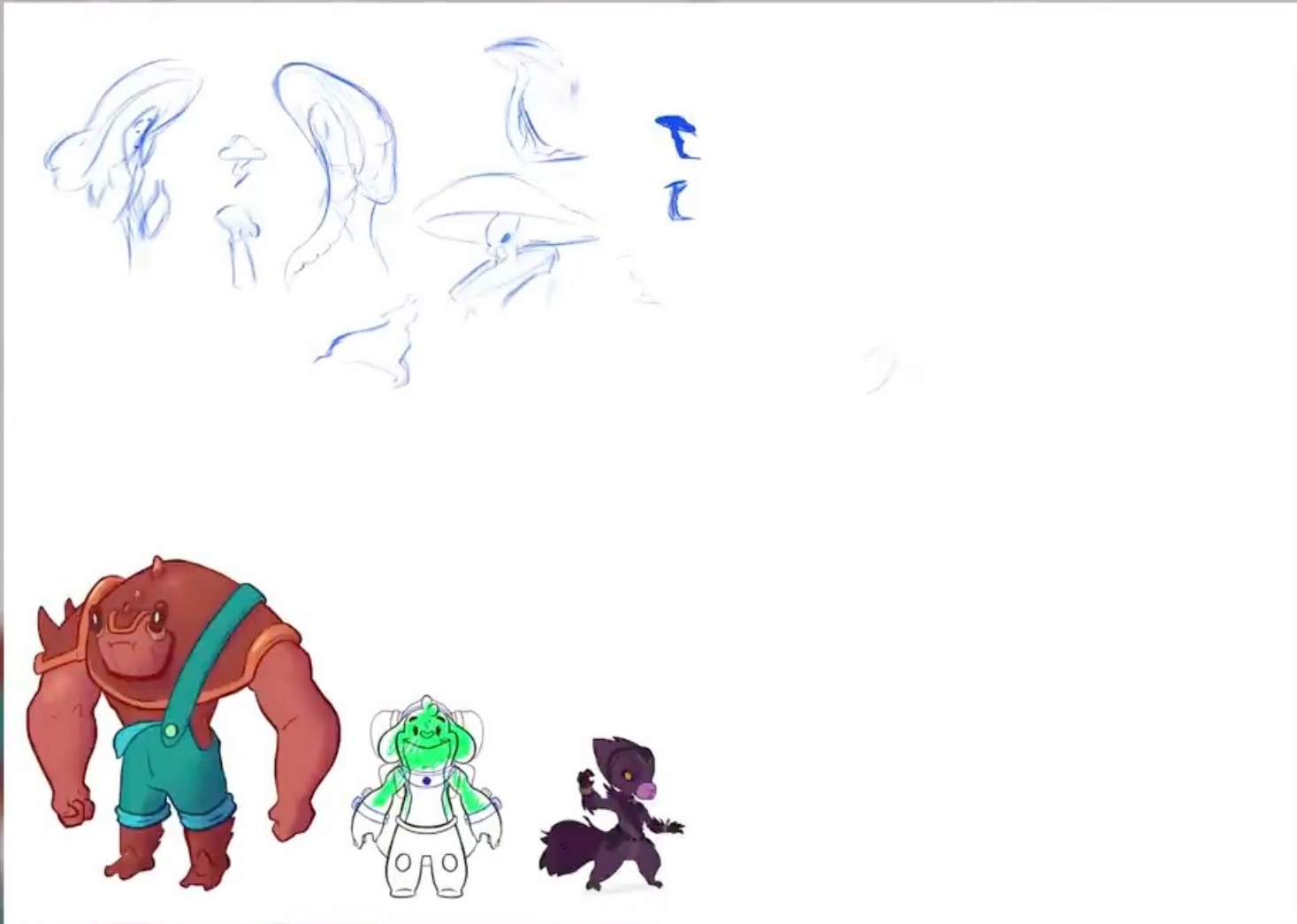
Maneesh Agrawala
Stanford University and Roblox

November 5, 2023

@magrawala
@magrawala@fediscience.org

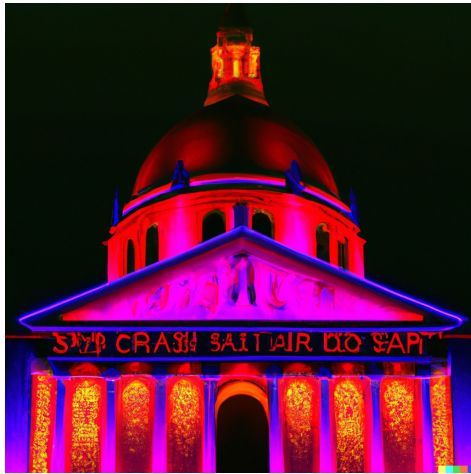
Creating quality content requires iteration

e given a low
 ration
 who is this?
 defines racism to be "the belief that a particular
 person's social and moral traits are predetermined by his or her inbo
 s." Racism is not just labeled as one hate between races like white and black. Many
 nities can have racism within itself. Racism can also be just blind hatred between
 duals. It can also be due to skin color, background, sex, language, birth places, or even
 toms. Racism can influence many things like slavery or the formation of countries and
 s. The belief of inferiority was not this automatic creation. Not all skin types or colorings
 riginally considered inferior. Many years ago the Portuguese discovered how much mo
 anced Africans than themselves. An important feature of race is that how on





stanford memorial church with neon signage in the style of bladerunner



Iteration 1

stanford memorial church **and main quad with palm trees** in the style of bladerunner



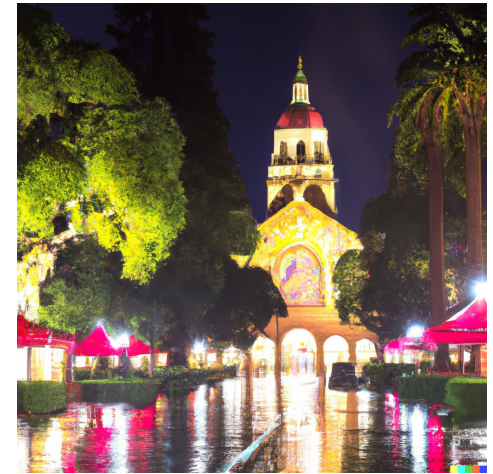
Iteration 3

nighttime rain stanford memorial church and main quad with palm trees, **night market food stalls and neon signs** in the style of bladerunner



Iteration 8

nighttime rain stanford memorial church and main quad with palm trees, night market food stalls and neon signs **like downtown tokyo**

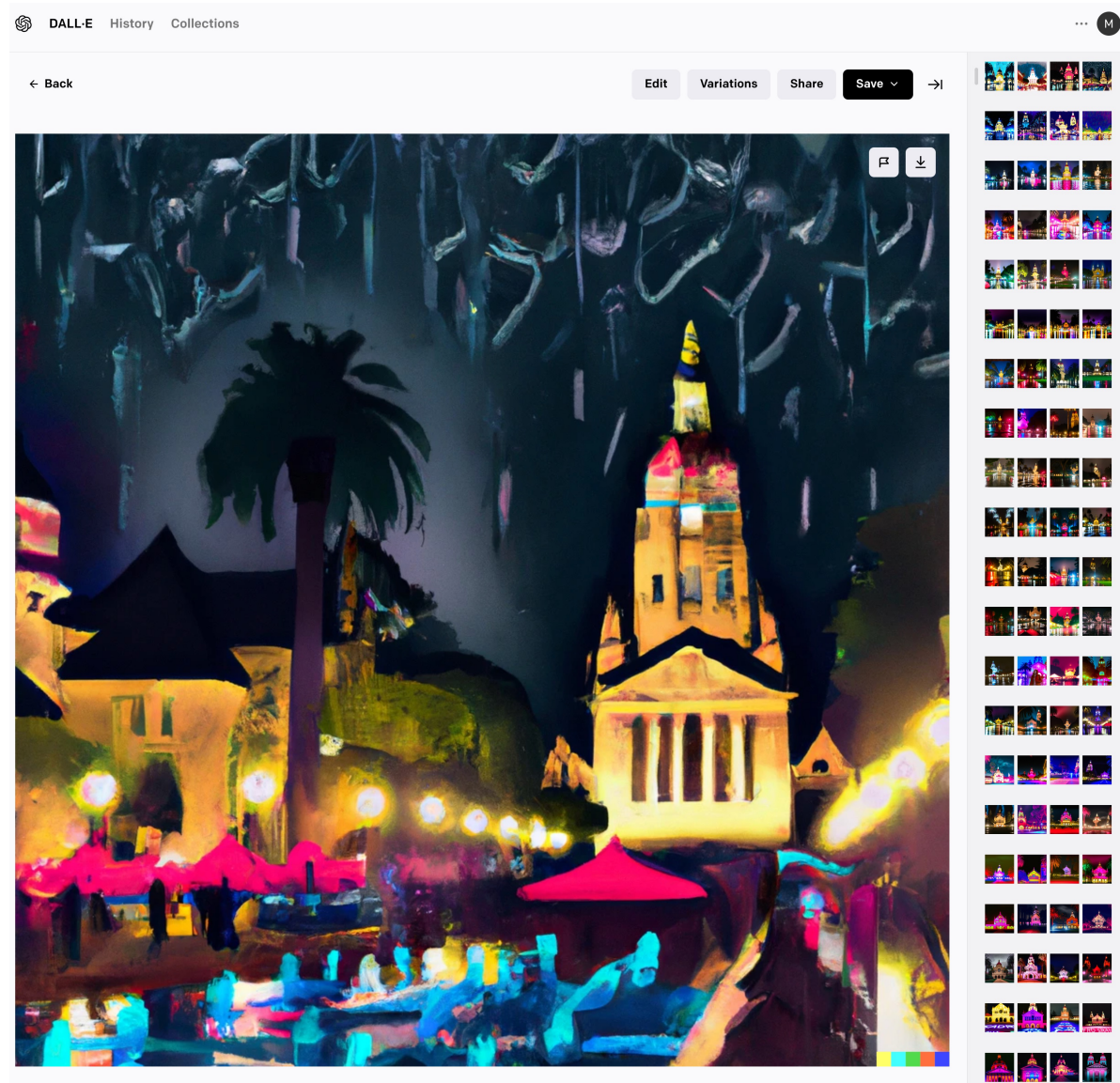


Iteration 17

nighttime rain stanford
memorial church and
main quad with palm
trees, night market
japadog food stalls and
neon signs, **neo** tokyo
bladerunner style **film**
still illustration

Iteration 21

<https://magrawala.substack.com/>



Real **artistic tools should act as extensions of the artist**, the way a paintbrush adds capabilities to a painter's hand, **rather than a slot machine that may or may not give you something useful.**

Aaron Hertzmann

A **good conceptual model** let's users **predict** how **input controls** affect the **output**

When the **conceptual model** is **not predictive**, users resort to **iterative trial-and-error**

It is ***our* job** as AI tool builders **to provide interfaces** that **let users build predictive conceptual models**

Supporting iterative ~~trial-and-error~~
refinement



**Rapid, reversible,
incremental actions**



Night from Day



Before

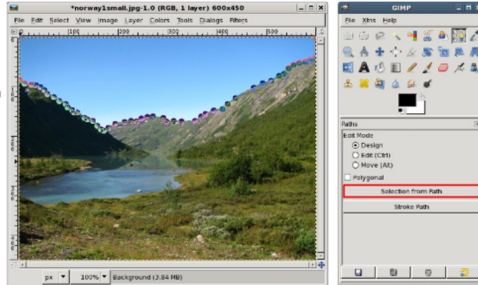


After

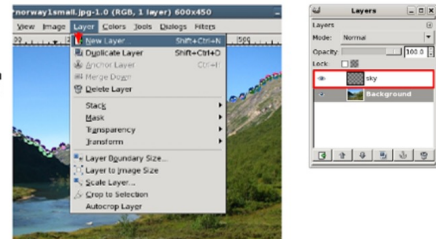
1 Select the path tool from the toolbar to create and edit paths.



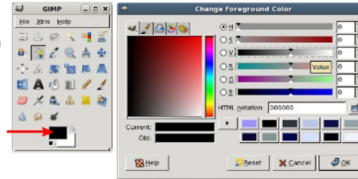
2 Use the path tool to select the sky. Then click on "Create Selection from Path" in the toolbar.



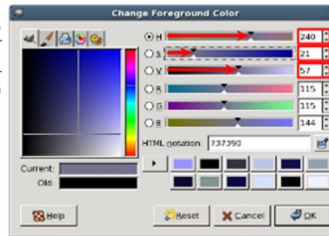
3 Select Layer > New Layer from the image menu bar to create a new layer and add it to the image. You will see a new layer appear in the layer dialog.



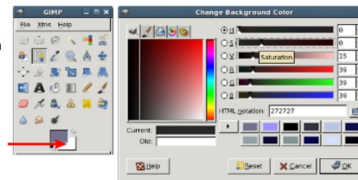
4 Select the Foreground color tool from the toolbar to change the foreground color. You will see a new dialog appear.



5 Set the V slider from 0 to 57, the H slider from 0 to 240 and the S slider from 0 to 21 in the Change Foreground/Background Color dialog.



6 Select the Background color tool from the toolbar to change the background color. You will see a new dialog appear.



Incremental actions

Maintain shared structure before and after action

Break into sequence of simpler actions/steps

Incremental: Maintaining Shared Structure



Stable Diffusion Inpainting Model

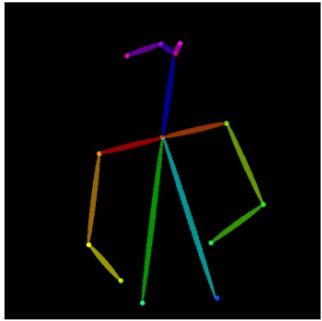
Incremental: Maintaining Shared Structure



“a building in a city street”



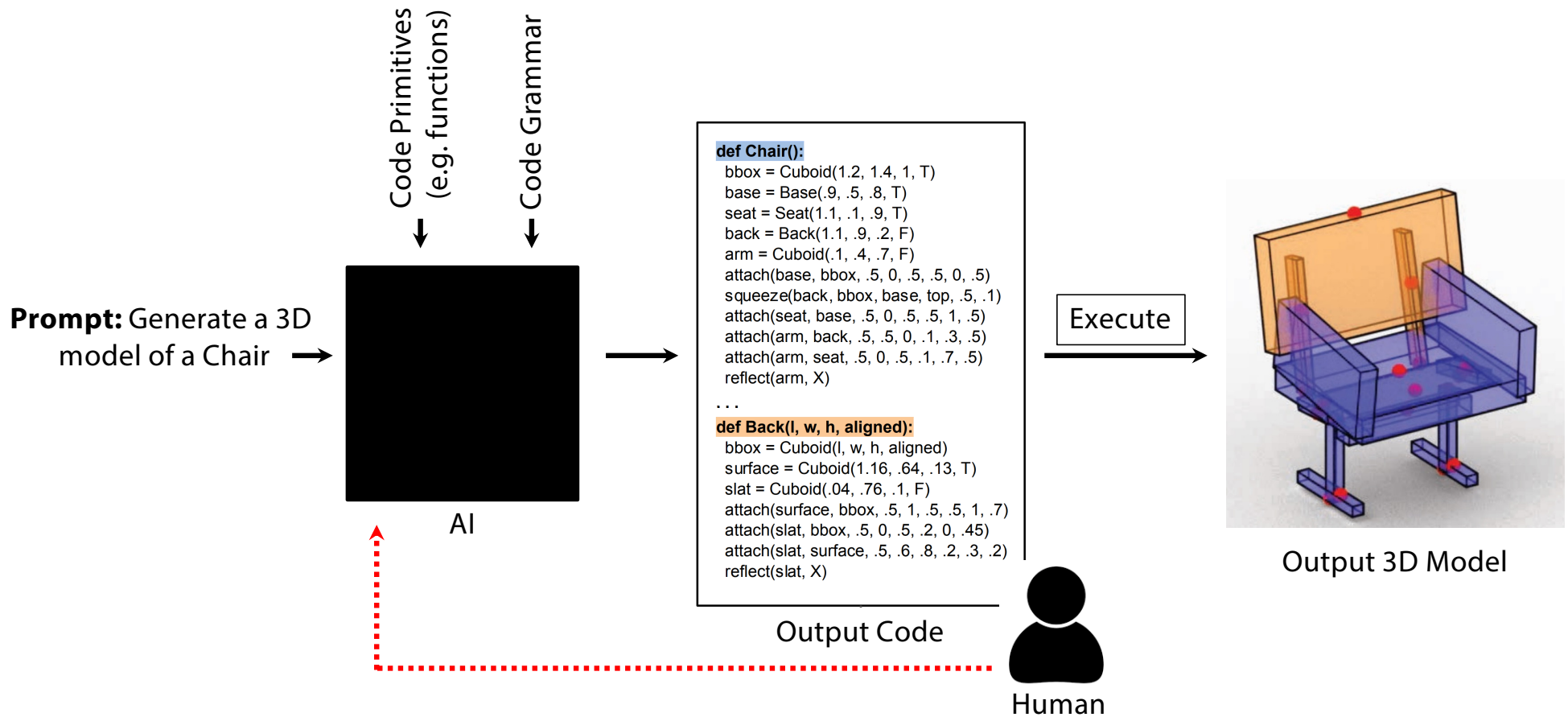
“inside a gorgeous 19th century church”



“chef in the kitchen”

Adding Conditional Control to Text-to-Image Diffusion Models [Zhang 2023]

Incremental: Break into Simpler Actions/Steps

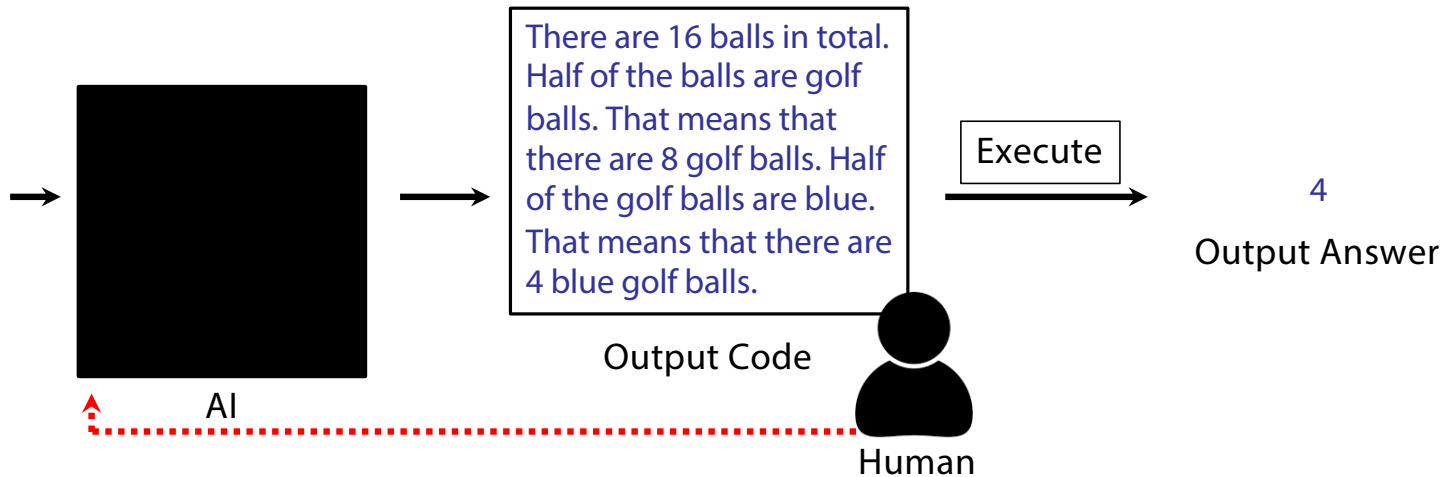


Neurosymbolic Methods for Computer Graphics [Ritchie 2023]

Incremental: Break into Simpler Actions/Steps

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A: Let's think step by step.



Large Language Models are Zero-Shot Reasoners [Kojima 2023]

Takeaways

Content creation requires *iterative refinement*

Generative AI tools primarily support *iterative trial-and-error*

For iterative refinement generative AI tools should allow *rapid, reversible, **incremental actions***

Open challenges

How to identify structure to maintain across actions

How to break actions into simpler actions/substeps (should align with human cognition/workflows)

MANEESH AGRAWALA



Stanford University
Computer Science Department
Brown Institute for Media Innovation
Graphics | HCI | Visualization

CONTACT INFORMATION

Gates Computer Science, Room 364
Stanford, CA 94305-1248
Tel: (650) 723-2642

Admin: Wen Audrey Chen
Tel: (650) 724-3770

Office Hours: By appointment



TEACHING

TRAVEL

Apr 19-22 2023 Northwestern, Chicago, IL
May 30-Jun 2 2023 GI, Victoria, BC

STUDENTS, POSTDOCS & VISITORS

Jean-Peic Chou, Terrell Ibanez, Jeongyeon Kim,
Jingyi Li, Sean Liu, Jiaju Ma, Jacob Ritchie,
Sofia Wyetzner, Lvin Zhang, Sharon Zhang

ALUMNI

Sean Arietta, Adrien Bousseau, Robert Carroll,
Lydia Chilton, Will Crichton, Ross Daly,
Abe Davis, Ohad Fried, Shiry Ginosar,
Ahna Girshick, Madeleine Grunde-McLaughlin,
Alex Hall, Jonathan Harper, Jeffrey Heer,
Robert Held, Enamul Hoque, Jessica Hullman,
David Jacobs, Jennifer Jordan, Alan Karim, ...

Maneesh Agrawala is the **Forest Baskett Professor of Computer Science** and Director of the **Brown Institute for Media Innovation** at Stanford University. He works on **computer graphics**, **human computer interaction** and **visualization**. His focus is on investigating how cognitive design principles can be used to improve the effectiveness of audio/visual media. The goals of this work are to discover the design principles and then instantiate them in both interactive and automated design tools. Honors include an **Okawa Foundation Research Grant** (2006), an **Alfred P. Sloan Foundation Fellowship** (2007), an **NSF CAREER Award** (2007), a **SIGGRAPH Significant New Researcher Award** (2008), a **MacArthur Foundation Fellowship** (2009), an **Allen Distinguished Investigator Award** (2014), induction into the **SIGCHI Academy** (2021), and being named an **ACM Fellow** (2022).

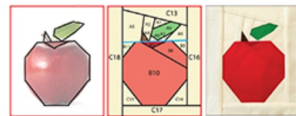
ANNOUNCEMENTS

- MAR 29, 2023 Blog post on generative AI. **Unpredictable Black Boxes are Terrible Interfaces.**
- AUG 17, 2020 **Stanford Cable TV News Analyzer** released. Examine who and what appears in the last decade of CNN, FOX News and MSNBC.
- JAN 28, 2020 **Data and code for visual question answering with explanations for charts and graphs.**
- DEC 13, 2018 Source code for **dancifying** video released, based on our **Visual Rhythm and Beat** project.
- JUL 1, 2015 Moved to Stanford!
- JAN 28, 2015 Released **speecheditor** tool and **source code**. Edit audio stories using cut/copy/paste on text transcripts.
- NOV 11, 2014 Released **D3 Deconstructor** and **source code**. Extract data from D3 visualizations.
- OCT 5, 2012 Released **3D Puppetry** system and **source code**. Make 3D animations using toys and a Kinect.
- OCT 2, 2012 Released source code for our **Proton Multitouch Framework**. Declaratively specify multitouch gesture as regular expressions.

PUBLICATIONS



SlideSpecs: Automatic and Interactive Presentation Feedback Collation
Jeremy Warner, Amy Pavel, Tonya Nguyen, Maneesh Agrawala and Björn Hartmann
Intelligent User Interfaces (IUI), Mar 2023. pp. 695-709.
[PDF](#) | [YouTube](#) | [Demo](#) | [Code](#) | [BibTeX](#)



Sketch-Based Design of Foundation Paper Pieceable Quilts
Mackenzie Leake, Gilbert Bernstein and Maneesh Agrawala
User Interface Software and Technology (UIST), Oct 2022. pp. 21:1-21:11.
[PDF](#) | [MP4](#) | [YouTube](#) | [Supplemental Results](#) | [BibTeX](#)



Measuring Compositional Consistency for Video Question Answering
Mona Gandhi, Mustafa Omer Gul, Eva Prakash, Madeleine Grunde-McLaughlin, Ranjay Krishna and Maneesh Agrawala
Computer Vision and Pattern Recognition (CVPR), 2022. pp. 5046-5055.
[PDF](#) | [Dataset](#) | [BibTeX](#)



Disentangled3D: Learning a 3D Generative Model with Disentangled Geometry and Appearance from Monocular Images
Ayush Tewari, Mallikarjun B R, Xingang Pan, Ohad Fried, Maneesh Agrawala and Christian Theobalt
Computer Vision and Pattern Recognition (CVPR), 2022. pp. 1516-1525.
[PDF](#) | [Supplemental PDF](#) | [MP4](#) | [BibTeX](#)



More Information

graphics.stanford.edu/~maneesh

[@magrawala](https://twitter.com/magrawala)

[@magrawala@fediscience.org](mailto:magrawala@fediscience.org)