

Napkinmatic

Yosun Chang
yc@AIMagical.com
AIMagical Permute.xyz
San Francisco, CA, USA



Figure 1: Screenshots (left to right): Napkin Sketch, proverbial; Generated image; Inverted plane indicating augmentation surface

ABSTRACT

Napkinmatic is a seamless spatial computing platform connecting AI to the real world – and back again; currently, it is an iOS and Android app that lets you (automatically) turn your napkin sketch into whatever it is you actually want it to be – from wireframe to vivid web design or sketch to masterpiece painting – or, using AppMode we can create our own mini-apps, for example, we can pipe our napkin idea into a 3D model creation endpoint [Chang 2023a].

CCS CONCEPTS

• **Human-centered computing** → Mixed / augmented reality; Collaborative interaction; Graphical user interfaces; Natural

language interfaces; HCI theory, concepts and models; Pointing; Text input; Gestural input; Auditory feedback; Touch screens; Ubiquitous and mobile computing; Ubiquitous computing; Mobile computing; Ambient intelligence; Smartphones; Interactive whiteboards; Tablet computers; Mobile devices; Mobile phones; Ubiquitous and mobile computing systems and tools; Collaborative and social computing; • Software and its engineering; • Information systems; • Computing methodologies → Artificial intelligence; • Computer systems organization → Embedded systems; Redundancy; Robotics; • Networks → Network reliability;

KEYWORDS

augmented reality, AI, spatial computing, ubiquitous computing

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
SIGGRAPH '23 Appy Hour, August 06-10, 2023, Los Angeles, CA, USA
© 2023 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0156-6/23/08.
<https://doi.org/10.1145/3588427.3595357>

ACM Reference Format:

Yosun Chang. 2023. Napkinmatic. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Appy Hour (SIGGRAPH '23 Appy Hour)*, August 06-10, 2023. ACM, New York, NY, USA, 3 pages. <https://doi.org/10.1145/3588427.3595357>

1 INTRODUCTION

Napkinmatic is a portmanteau of the word Napkin and automatic. Similar in name and in the spirit of DrawmaticAR [Chang 2020a] [Chang 2020c], it's about enabling the potential of your napkin sketch – “automagically.”

1.1 Technical Implementation

The app extends upon my studies on AR creation from paper – it “digitizes from a photo snap” [Chang and Grandhi 2019] [Chang 2019] [Chang 2020c] [Chang 2020d], extracting content from the paper while turning the paper into the augmentation image target. The relevant content is extracted, from text to sketch, and made into generalized objects, which the app's high level scripting environment can interact with, or connect to external endpoints, etc. In the screenshots shown, we use a controlnet-scribble [Zhang and Agrawala 2023] endpoint.

2 THE PROVERBIAL NAPKIN SKETCH AS THE CREATION INTERFACE

A napkin is a form of proto creation medium: it's where you roughly spew out your idea – and in that sense Napkinmatic aims to be the connecting platform that takes your napkin sketch and turns it into the real thing.

2.1 Expressive Ideas and Notes on Paper

Ideas expressed as words can be turned into buttons, for example, or as denoted by other contextual menu items. Different functionality on different planes can be connected as nodes. Additional interactivity can be scripted via either connecting with other APIs (in a standardized format defined below) or typescript using AppMode. For example, entire pages of text can be summarized into content that can be queried.

3 THE HYBRID SPACE BETWEEN SOFTWARE AND PAPER

The app connects the “hybrid space” between software and paper, utilizing the paper as a stabilizing medium to both extract user input and to augment content on. The app turns the paper into a surface.

3.1 API Response Data Format

The app posts an image and a prompt and receives and processes a response in the format as defined in. [Chang 2023b]

3.2 When anything can become an interface, make it so.

Napkinmatic places computing on any surface in reality, and thus, we can turn any surface into a computing platform.

4 INTERACTIVITY MADE INTUITIVE AND EXPRESSIVE USING NATURAL LANGUAGE

What if you could simply write in natural language to add interactivity - or build your app.

4.1 Parsing Grammar

Prior to LLMs, parsing any English sentence into its constituent grammatical parts was a trivial task. Once that is done, then map each part to the corresponding analogue in our 3D stage: identify nouns as 3D models, verbs as animations, adjectives as render property modifications, and sentence structures for character interaction.

4.2 DrawmaticAR - a prehistoric “baby AGI” version

In 2020, I presented the pre-generative-AI era product-working-market-fit version of that as DrawmaticAR [Chang 2020a], an iOS/Android app that lets children turn what they write on real paper into 3D animated AR scenes. It used a classical programming dictionary to look up an internal database of several hundred 3d models, and augmented with Google Poly [Zvinakis 2017], a 3D models database (with solid Unity integration) that Google discontinued shortly after I won my first SIGGRAPH Real-Time Live Audience Choice Award [Chang 2020b].

4.3 Natural Language kind of works but is ambiguous

Natural language is itself a programming language – grammar contains control structures: nouns that are solid well-defined objects, verbs that traverse to places, adjectives that describe nouns, etc. The only problem in using it as an explicit programming language is the latent ambiguity of natural language. However, with LLMs, and the desires of most humans on the normal curve, it seems that common requests on this spectrum are fully learned and thus understood.

5 APPMODE

The insights from the previous section enable the natural language programming creation feature of Napkinmatic AppMode.

We describe 3D mode in detail in [Chang 2023a].

6 TOWARDS TRANSPARENT USER INTERFACES

The best UI is a transparent user interface – a user experience that is naturally intuitive. Barring that, it needs to not hamper the user with additional learning curves or unnecessary tedium time in needing to use interfaces that do not scale. We believe that with the transformative power of AI and the many more parameters of control yet to come, we can achieve the fully transparent user interface.

REFERENCES

- Yosun Chang. 2019. *SIGGRAPH '19: ACM SIGGRAPH 2019 Appy Hour* (July 2019). <https://dl.acm.org/doi/10.1145/3305365.3329730>
- Yosun Chang. 2020a. *Special Interest Group on Computer Graphics and Interactive Techniques Conference Real-Time Live! (SIGGRAPH '20 Real-Time Live!)* (Aug. 2020). <https://doi.org/10.1145/3407662.3407755>
- Yosun Chang. 2020b. *SIGGRAPH Blog* (Aug. 2020). <https://blog.siggraph.org/2020/09/real-time-live-audience-choice-award-winner-talks-one-woman-projects-interactive-storytelling-and-corgis.html/>
- Yosun Chang. 2020c. DrawmaticAR. <https://DrawmaticAR.com>.
- Yosun Chang. 2020d. PaperMATA. <https://history.siggraph.org/experience/papermata-habitats-for-ai-by-chang/>.
- Yosun Chang. 2023a. 50, 1 (Aug. 2023). <https://doi.org/10.1145/3588430.3597253>

Yosun Chang. 2023b. *Napkinmatic - Open Framework*. Retrieved May 23, 2023 from <https://github.com/yosun/napkinmatic>

Yosun Chang and Uttam Grandhi. 2019. *SIGGRAPH '19: ACM SIGGRAPH 2019 Appy Hour* (July 2019). <https://dl.acm.org/doi/10.1145/3305365.3329729>

Lvmin Zhang and Maneesh Agrawala. 2023. Adding Conditional Control to Text-to-Image Diffusion Models. arXiv:2302.05543 [cs.CV]

Andrea Zvinakis. 2017. *Poly: Browse, discover and download 3D objects and scenes: 2017*. Accessed: 2023-05-20. Retrieved May 20, 2023 from <https://blog.google/products/google-ar-vr/poly-browse-discover-and-download-3d-objects-and-scenes/>