Screenless

Sean Zhu
NYU ITP

Fall 2021

I am writing this paper in part at a table at school[1], in part at my desk at home, and in part on my couch. These locations are all standard fare; these, plus, during non-pandemic times, coffee shops, are probably virtually all of the places where people write papers. While these might appear to be a fairly diverse set of work locations, they actually have a lot in common: in all of these I'm sitting in one place, stationary, at a computer. Any perception of these locations being wildly different is a product of some kind of unspoken restriction on the nature of digital work that we have accepted as fact.

I don't like this monotony. If I could have it my way, I would have liked to have written the first draft of my paper while on the subway — I spend a good chunk of my day en route to and from school. I would have liked to brainstorm ideas while at a park, perhaps, without needing to look up "parks with good wifi" or to expose my hands to the chilly weather for hours. I would have liked to have sent the draft out to peers for review while running my errands, composing the request emails and messages without straining my eyes and thumbs to adapt to the small screen of my phone. And in the coming weeks, I would like to be able to work on the code for my prototypes while I'm on the go as well, without needing to increase the chances of colliding with a car or person on the street.

This isn't possible today, but it's the kind of future I want to explore through my thesis.

Background: The issue with screens

My thesis originated from a simple observation. Most activities throughout human history require the activity-doer to move around quite a bit. This tendency is more than just human: moving around a lot is a key difference between animals and plants. On the other hand, not only does modern "creative" and "productive" work not require a lot of movement, but it also appears enhanced by the lack of motion. The stereotype of a productive person is someone who can sit at their desk all day to get things done, while people who can't stop moving around are seen as "fidgety".

---

[1] This is actually false: I can't write at school because it's too distracting. But I included it for effect, as school is a common place for many people to get things done.

I am one of these fidgety people. Over the years, society has tried to convert me to a sedentary worker, and it's quite tempting to be normal. However, I've noticed some drawbacks with the normal approach.

The first one, as alluded to above, is that screens are unhealthy. We simply did not evolve for our eyes to be glued to a small port of our field of vision — like a dog watching the outdoors through a hole in the fence — for a quarter to half of our waking lives. To attempt to address this, we've created things like curved monitors and blue light glasses and f.lux. We also did not evolve to be sitting in one place, moving hands in the same restrictive motions ceaselessly, and when we do this, we develop back problems and repetitive hand injuries. To attempt to address these issues, we've created comfortable chairs, standing desks, and all kinds of fancy keyboards and mice. All of these solutions are just attempts to approximate the variety and ergonomics of the outside world, and they are imperfect and expensive approximations. Why can't we just have the real outside world?

A second problem with screens is that they are distracting. The modern digital display is very good at capturing attention, potentially to the detriment of things in real life that need attention. The color and contrast in computer and phone user interfaces is exceptionally vivid, and this subconsciously stresses to the user the importance of what's on screen. To add to this, any third party that has engagement as a goal — advertisers, app makers, and content creators — will seek to maximize the eye-catching-ness of their content. Remember a decade or two ago when we were afraid of a future where ads floated around every street and in our homes? Well, that future is already here, except these ads are floating in front of us because we're willingly holding them up to our faces.

Yet another issue with screens (although fixing this might be out of scope for my project) is that they are not conducive to collaboration. This occurs for no reasons other than the ones already mentioned, but the effect is distinct in this application. The unnaturalness of interacting with screens litters a "modern" collaboration session with phrases like "can you see my screen", "where are you in the doc", "just move your mouse to the left a bit, yes that one", and "oh sorry, sound was my cat," phrases that would be quite uncommon in an in-person collaborative session. And the distracting nature of screens has led to, at least in part I suspect, the convention of requiring remote participants to keep their video on, and the convention of requiring laptops to be closed in class — lest the participants get distracted by something else on that dazzling black mirror of theirs.

Let's first take a look at how everyone has tried to address these issues.

Prior art

There are quite a few existing projects that attempt to make displays healthier, to make it easier to work not at a desk, to make the digital world less distracting, and to make it more natural to collaborate online. Some of them are mentioned above, but I wanted to list out a few more here.

In the past five years or so, there have been a proliferation of voice assistants — Siri, Alexa, Google Assistant, and so on — that have provided a promising take on interacting with computing devices without looking at a screen. They work quite well for their intended use cases and definitely reduce reliance on the screen: there is a novel category of device powered by these assistants, smart home speakers, that don't have any sort of display! One issue with these devices is that the intended use cases are limited: it's abysmally hard to provide any sort of nuance, clarification, or correction to a previously issued command, and smaller micro-commands (like entering a keystroke) seem entirely out of scope. The other issue is with what the intended use cases are: all of the three major voice assistants have their almost entire use case revolving around internet-based subscription services and expensive Internet-of-Things products. It's entirely commercial, and when people primarily interact with voice assistants as a consumer, not a creator.

Voice assistants seem closely related to a technology that fewer people have used, but has been around for much longer, created for users for whom the screen is more than a mere inconvenience: voice assistive technology for the blind. In particular, Apple's VoiceOver is a particularly comprehensive voice user interface that parses what's on screen and allows the user to interact with it through keyboard or touchscreen.[2] Assistive technology is more fine-grained than voice assistants because they have to be, and it's great that they enable people with disabilities to be more than just consumers.[3] VoiceOver and similar technologies still suffer from some problems. One problem is that, in my opinion, VoiceOver feedback is too verbose; using this best-in-class tool is still akin to interacting with the most chatty and unintelligent secretary ever. Every sentence of voice feedback is littered with words describing the type of every kind of control — "button", "menu", etc. This is bad for one of the same reasons why screens are bad: this interface requires complete concentration to be able to use. I don't think this can be fundamentally avoided, because I think the underlying issue with assistive technology is that it's translating an interface that was completely

[2] I was introduced to VoiceOver and spent a lot of hands-on time with it through the course Looking Forward, which I took last semester (Spring 2021).
[3] Apple's short commercials, such as this one, do a great job of showcasing this: https://www.youtube.com/watch?v=tEEPpdct5t8

optimized for a different medium. The only way to solve this would be to create a user interface from scratch that didn't assume visual interaction, and for an audience that third-party developers would not consider second-class. Voice assistants are one example of this.

This semester (Fall 2021), I took the course Textile Interfaces, which introduced me to interfaces that can be woven into clothing and other flexible substrates. Making input more physically flexible might not seem to reduce our need for screens, but I think it actually does. One of the reasons why we need to look at a screen is because our input into computing devices is symbolic, so we need a screen to confirm the actual action performed. Allowing input and output on more surfaces makes it easier for a user's hands and eyes to be where the actual action is performed, without needing a screen to correlate the two. Flexible technology looks promising, but currently it appears to be just that — a promise. Many wires and sensors can be truly integrated into clothing now; however, batteries and actual microprocessors can't be made flexible yet and must be integrated into hard components like buttons, glasses, or jewelry. In class we also looked at Google's Project Jacquard, a project to weave flexible controls into mainstream fabric products like Levi's jeans. Project Jacquard has not seen a public update since 2020.

Finally, I found a very different kind of existing screen-less interaction in an everyday device: an alarm clock. The big snooze button on alarm clocks is specifically designed to be activated by a groggy user before they open their eyes for the first time each day.[4] Modern smartphones have alarm clocks with a snooze button too, but anyone who has used one knows that it is a lot harder to activate with eyes closed.

The existing solutions in this space offer a lot of inspiration, but none of them completely address the focus of my thesis.

Research questions and focus

The following questions will drive my thesis project.

Much of the motivation for my thesis project comes from my own needs and the needs I've observed of people around me. With my project, I want to reach a broader audience, the first two questions are designed to gain that broader context:

---

[4] I explored this interaction in a project where I created a big button on a wall that a user can tap with their eyes closed. https://itp.nyu.edu/shows/winter2020/walltap/ or https://github.com/szhu/WallTap

> Q1. To what extent do screens pose a hindrance to people today?
>
> Q2. Are there people disproportionately and unfairly negatively impacted by screens?

With this knowledge in hand, to start solving the problem. When it comes to futuristic technologies, it's quite common for people to dream of science fiction futures that aren't practical because they were created without regard to technical feasibility. I believe that practical concerns are disregarded not because people like being unrealistic, but simply because it's hard to understand what practical constraints exist without actually building something and getting feedback on it. Because of this, my project will heavily involve making, and I will use the next question as my guide:

> Q3. As a low-hanging fruit, what is a computer interface that can be created in a month or two, using technologies that exist today, that can fundamentally alleviate some of the issues with doing creative work on screens?

My use of "fundamentally" is meant to differentiate my project from projects like f.lux, which try to simulate the benefits of the outside world without being the outside world. I want to avoid these projects because I think they will be playing catch-up to all the benefits of the outside world for a long time to come. My use of "creative work" is meant to differentiate my project from voice assistants, which don't allow the user to stray from intended flows and create something fundamentally new.

> Q4. Is there a family of interfaces that this can inspire?

Ideally, I want to create not a single device or interface, but to invent a family of them, and to have the device I make simply a (usable!) proof of concept. I would only consider this project particularly successful if the interface I make is something someone would want to copy and modify.

After creating something that is provably useful, I would be open to using the insights gained to think about the further future:

> Q5. What is a 20-50 year vision of what the future would look like in terms of display and touch technology? Are screens prominent in them? Should they be?

I intend the bulk of my work to be addressing Question 3, in which I would build a single, practical device that can allow someone to do many typical computer or phone tasks without looking at a screen for a majority of the time. I want to focus my time on building

a simple, particular product because I believe this is what I believe is needed the most in this space: there are already a lot of people doing research into problems that screens have caused, and a lot of people building things that only address the concerns of a very small set of people. It is crucial to connect the two.

Current progress and implementation plan

I've already started prototyping some interfaces. They might just be tools to help me understand the nuances of being able to accurately and naturally interact with computers, or they might morph into my final created interface; I won't know for sure until later.

First, I created an example of a screen-less input method: a text field that speaks the last typed word.[5] It was simple to implement, and it made for a great demo when I used it — with laptop lid closed — in my thesis presentation.[6] However, after extended usage, I noticed a persistent issue, which is that it was easy for me to give the computer input that would lead to changes that are hard to describe via voice. For example, when editing text, I'm used to using my mouse or arrow keys to change things around, but navigating this way without looking at a screen for confirmation was a recipe for disaster. I was surprised how much the interaction design of even a basic, plain-text text field revolved around visual feedback. To alleviate some of the navigation problems, I added a feature to allow for sections of text that the user can jump back and forth between, but it did not solve the issue of stray inputs.

To resolve some of these issues, I experimented with creating a locked-down version of the text editing widget.[7] In this new widget the cursor can only be used to jump entire words, no selection is allowed, and there are dedicated shortcuts for navigating by words, sentences, or paragraphs. Even though all the input aspects of this interface have been created from scratch, it looks and feels similar to a regular text field when typing new text at the end of the field. I intend on using this interface as a tool for experimenting with nonvisual user interactions.

My current plan for my final device is a pair of handheld keyboards that are small enough to each fit into a coat pocket, with the intention that the user can use this input device while walking around or sitting outside, potentially even keeping the devices in

---

[5] Demo: https://szhu-thesis-2022.glitch.me/1-speak-typing/ Note: There is a server-side save/load component to this app, but it is disabled for this demo and will speak "Error" on load as a result. It does not affect the typing and speaking functionality.
[6] See https://photos.app.goo.gl/ee6Zc1fiJ4dE5imw8 at around the 2:20 mark.
[7] Demo: https://szhu-thesis-2022.glitch.me/2-structured-text-editing/

their pockets the entire time. The lack of a screen means that users will be able to interact with the world around them. They won't be able to fully use their hands for other purposes while they are typing, but that is a compromise I find acceptable, since multi-tasking often leaves one's hands occupied with just one of the tasks. What's less acceptable is for a task to occupy the entirety of one's sensory input, and this device won't do that. People can use it to jot down thoughts while talking with their colleagues, to navigate their email (hearing it through earbuds) while on the subway, or to message a friend to tell them that they're running late, without this task causing them to be even later.

For this keyboard to be usable without the user looking at it, it needs to have not many keys, but it ideally should have the full range of inputs expressible through a normal keyboard. This problem has been solved by chorded keyboards[8], which have extra modifier keys[9] to allow for more combinations with a smaller number of keys. Two kinds of chorded keyboards with active communities are them include the stenotype, which is used by court reporters and closed captioners to quickly transcribe spoken words[10], and braille keyboards, which are used to input braille or text.[11]

There is a category of one-handed chorded keyboards called keyers that are designed to be used in one hand,[12] and there are some keyers created specifically to solve the problem of being able to type not at a desk. The devices on the market today may or may not be sufficient for my project. If they are, then I will focus my project on reducing screen reliance, which is something they don't solve; and if they are not, I will build my own keyer as well.

My plan for implementing the interface is to have a device that is possibly usable for daily use within a few weeks of the new semester, by the end of February 2022 at the latest. I want to be able to spend much of my semester (attempting to) use the device for my everyday tasks, and using this feedback and feedback I get from other users to iterate on the device and to build out the interface on the computer end.

I assume that interacting without a screen is a hard problem to solve, and that the iteration process might take the entire semester. In case somehow I arrive at a usable device and interface with time to spare, some further avenues of exploration include

---

[8] https://en.wikipedia.org/wiki/Chorded_keyboard
[9] This is a simplification. In a chorded keyboard, all the keys are modifier keys, and key combinations activate when the keys are released, which means that a non-modifier key is not required.
[10] https://en.wikipedia.org/wiki/Stenotype
[11] To explore this kind of interaction, I created a simplified braille keyboard. Demo: https://szhu-thesis-2022.glitch.me/3-braille-typing/
[12] https://en.wikipedia.org/wiki/Keyer

creating other kinds of devices to solve problems in this space and making sure the device can work with both phones and computers. Or if I feel like the device is perfect and just needs a little word of mouth to spread, then I will spend some time crafting a better presentation of the use cases of this device, and working on pitching the vision of a screen-less future to consumers.

There is a lot of uncertainty left in my plans for this project, and this reflects the amount of insight waiting to be discovered as I build screenless interfaces; I believe that any plan more certain would be unrealistic. I can't wait to get closer to certainty and having actual devices to demo, but for now, I'm just excited that I finished writing this paper so I can stop looking at my screen.