

An Appliance Display Reader for People with Visual Impairments

Giovanni Fusco¹

Ender Tekin²

James Coughlan¹

1 THE SMITH-KETTLEWELL
EYE RESEARCH INSTITUTE

2 UNIVERSITY OF WISCONSIN – MADISON



Waisman Center

Motivation

More and more everyday appliances have displays that must be read in order to operate them:

- microwave oven
- thermostat
- DVD player
- exercise equipment
- glucose meter
- etc.



Motivation

These displays are often hard to read (low contrast, small font, glare issues) even with normal vision...

For those with a visual impairment they are even harder to read, and may be completely inaccessible!



Past work and existing tools

Some past research on this topic [1-2], but nothing commercially available

Commercially available solutions:

- Optical Character Recognition (OCR), such as KNFB Reader for iOS/Android, is very useful for reading printed documents, but not well suited to reading displays:
 - Non-standard fonts aren't always read correctly
 - How can a blind user aim the camera properly?

Existing tools (con't)

Apps based on crowdsourcing (and in some cases also computer vision), including:

- VizWiz
- TapTapSee
- Be My Eyes



This approach is very promising, and we will discuss Be My Eyes later

The main limitations are delays (latency), aiming the camera properly and privacy concerns

Display Reader

Research project conceived at Smith-Kettlewell to make appliance displays accessible [3-4]

Goals:

- make a smartphone app that helps visually impaired users acquire usable images of a display
- have the contents read aloud (or displayed in magnified/contrast-enhanced form)

System components

(1) Paper markers must be affixed around display region.

We are exploring performance trade-offs of using only one marker rather than four.

Note that this marker may also serve as a useful tactile symbol.



System components

(2) Template of the appliance: a model of the display region specifying precise locations of text fields in display (and optionally also buttons).

Templates may be generated by a sighted friend or a crowdsourcing process.



Current prototype

Android smartphone, held by the user, sends video information to a nearby laptop, where it is analyzed in real time.

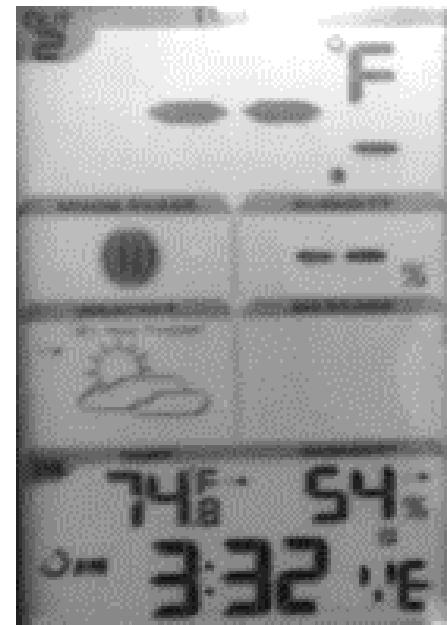
A new version of the prototype has been implemented entirely on the smartphone, without using any laptop. (Work in progress.)

Aiming the camera

System helps users aim the camera properly even if they can't see the display, using audio feedback

Challenge: user may be aiming camera properly towards the display, but the display may be obscured by glare/reflections

→ System helps user move camera to find a better viewpoint with less glare



User interface (UI)

Wanted to strike a balance between making the UI informative yet also easy to use

After several iterations and testing the system with multiple visually impaired users, we arrived at the following UI...

User interface (UI) overview

The user must complete these main steps:

- 1) Launch app and begin by laying smartphone flat against the surface of the display
- 2) “Back-away strategy”: slowly back the smartphone camera away from the display
- 3) Listening to audio feedback, move camera so as to capture a usable image
- 4) Once a glare-free, usable image has been captured, the contents are read aloud

User interface (UI) overview

The two most important tasks for the user:

- 1) Move the camera into “the zone,” i.e., move it to a location (and hold it in a direction) such that the display is well framed in the image
- 2) Even if the camera is in the zone, there may be glare that makes display unreadable. In this case the user must move somewhere else in the zone to avoid glare

UI feedback

Audio feedback issued by app to help user enter the zone:

- “Up” / “Down” / “Left” / “Right” / “Back” / “Close”
- Ambient (“heartbeat”) tone rate is *slow* or *fast*:
 - slow (every sec.) means app is running but you’re not in the zone
 - fast (~twice per sec.) means you’re in the zone but the display is unreadable because of glare

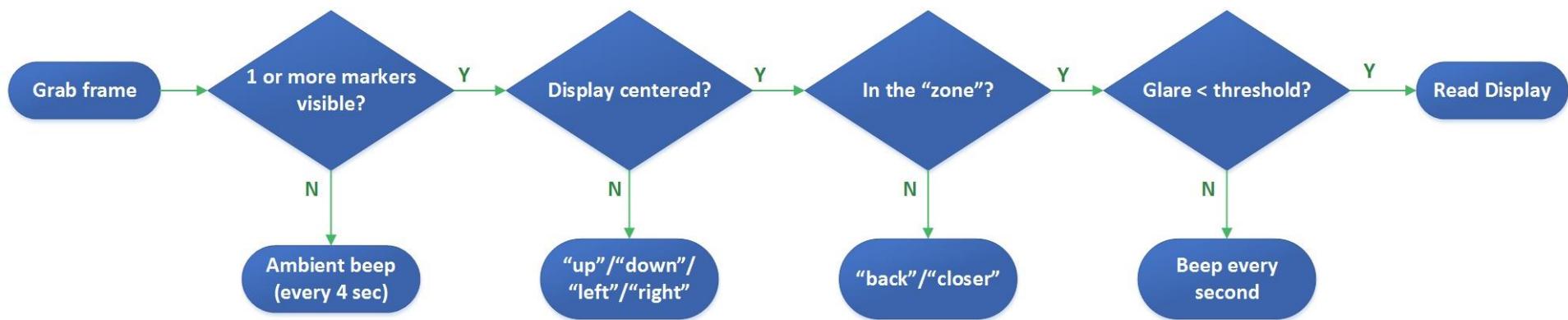
Display is read aloud once a usable image acquired

UI feedback

“Up” / “Down” / “Left” / “Right” helps user center the display in the camera’s field of view

“Back” / “Close” helps user acquire images from the proper distance (not too close or too far)

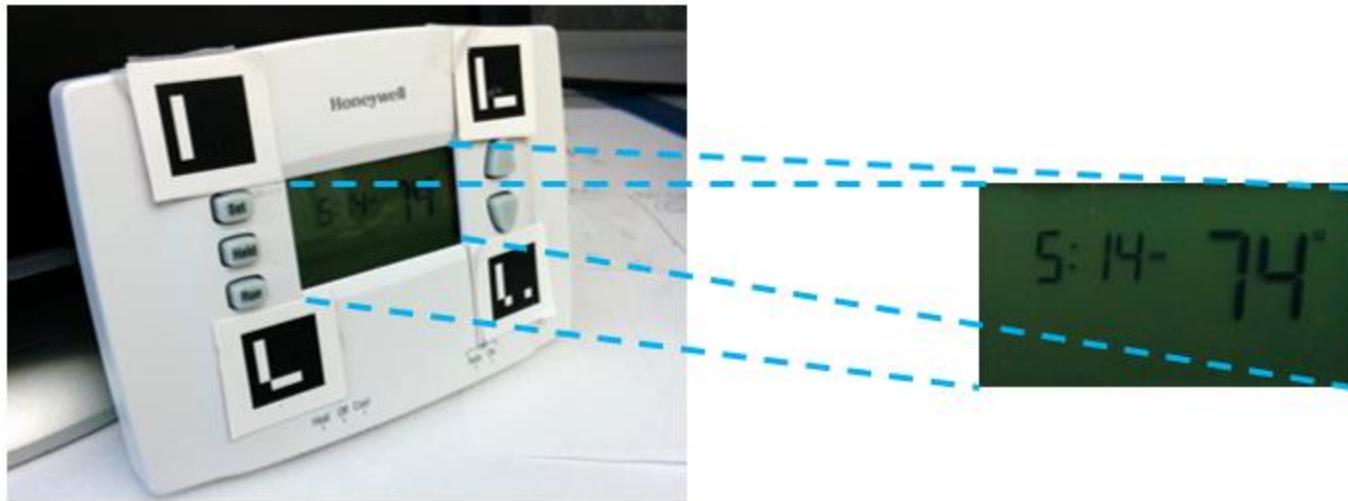
UI flowchart



Good image acquired

Once the user has acquired a good-quality image of the display, his/her work is done!

Image of the display area alone is extracted from full image (note that it is *rectified*, meaning any perspective warping has been undone):



Final step: reading the display

We have devised a simple computer vision algorithm for reading seven-segment digits

These characters are standard on many appliances, yet they are not handled well by OCR because they are a non-standard font

Input image (red LED digits) and output of our algorithm in yellow:

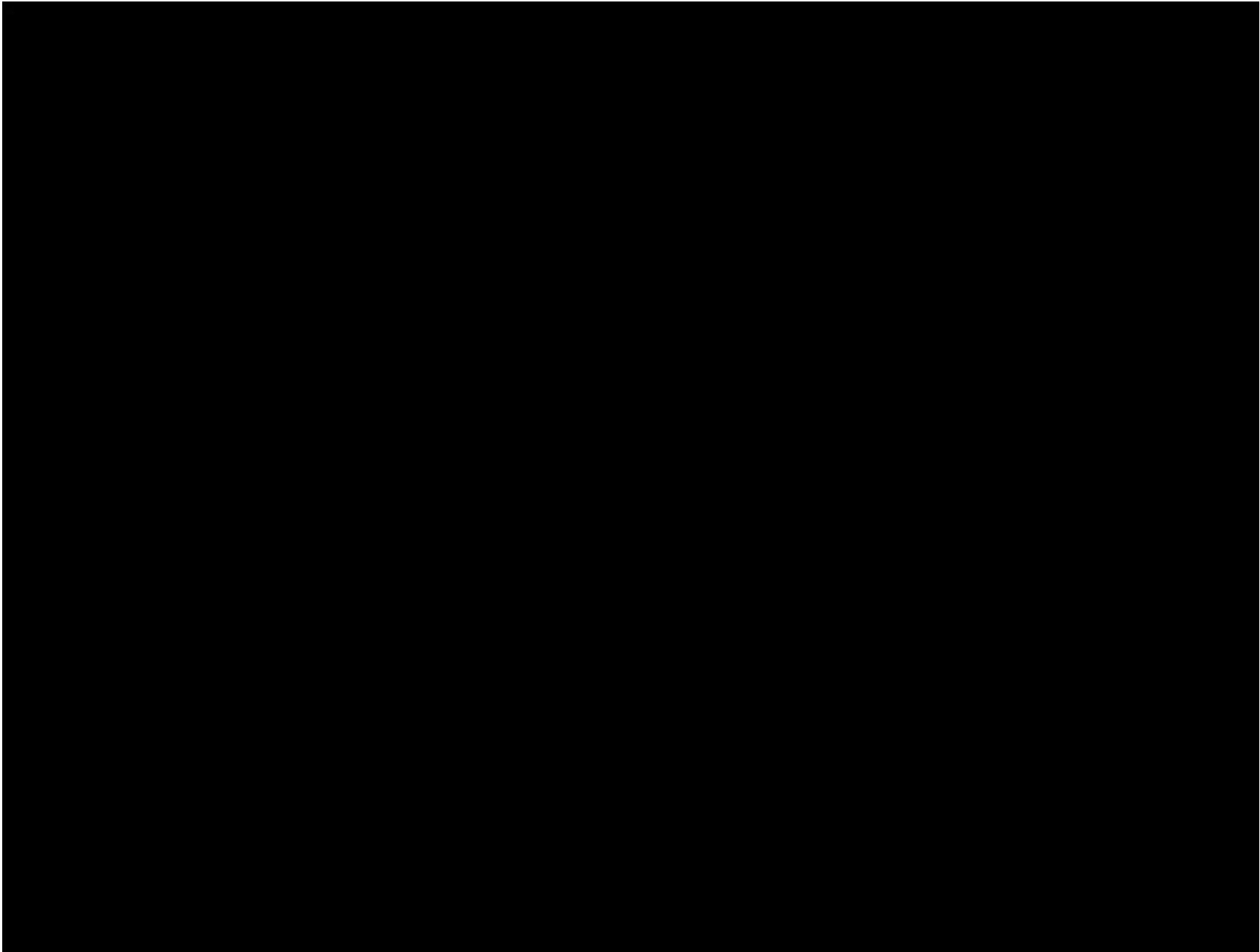


Final step: reading the display

The output is then either:

- read aloud using text-to-speech, or else
- displayed on screen in magnified/high-contrast form (work in progress)

Video demo



Experiments with visually impaired volunteers

We conducted formative studies with volunteers who are either blind or visually impaired

User training is required to explain:

- the concept of the camera's field of view
- the need to move the camera slowly to avoid motion blur
- how glare can impair display readability
- how to move the camera to minimize glare!

Experiments with visually impaired volunteers

Our studies:

- showed us how to improve our UI
- demonstrate the feasibility of our approach:
 - volunteers were almost always able to acquire good-quality, usable display images
 - these usable display images can be acquired quickly (roughly 10-30 sec)
- suggest the kinds of improvements and additions we need to make

Work in progress: anecdote

We asked one blind volunteer to use the Be My Eyes (BME) app to read the displays instead of using our Display Reader system



(BME sends live video from user's smartphone to a Remote Sighted Assistant, who can talk with user to answer his/her question)

Preliminary observations

- BME is a great tool that allows a visually impaired user to get feedback on almost any kind of visual information in the environment
- However, the Remote Sighted Assistant (RSA) has to guide the user to capture usable video, which can be challenging
- It can take some time to connect with an RSA and get the needed information (e.g., a minute or more)

Preliminary observations (con't)

- One time the RSA made an interesting mistake:
He/she read the display aloud to the user but *failed to include one digit* at the edge of the display area that lay just outside the camera's field of view!
- This highlights the importance of good training of the RSAs

Preliminary conclusion

We will explore combining computer vision with crowdsourcing/RSAs to harness the advantages of both:

- Computer vision is fast and can be used for rapid, real-time feedback to find a good-quality image of the display
- Crowdsourcing/RSAs can read displays with all different fonts, and may cope better with glare and other disturbances if they can't be completely eliminated

Ongoing and future work

- Use current system to capture a good-quality image of the display region, and send this cropped-out image to a crowdsourcing/RSA service
- Advantages of this hybrid approach:
 - cropping out the image reduces privacy concerns
 - doesn't require much bandwidth (compared with sending a live video feed)

Ongoing and future work (con't)

- Improve the automatic detection and measurement of glare
- Incorporate other measures of factors that can impair readability, such as motion blur and low contrast



Ongoing and future work (con't)

- Add a function that guides the user's finger towards a desired button location (related to [5])
- Explore the possibility of using a wearable camera (e.g., Google Glass, Vuzix), which may make it:
 - easier for users to aim the camera properly, and
 - more practical to point to an appliance while listening to audio feedback



Conclusion

- Display Reader is a smartphone app under development to make appliance displays accessible to people with visual impairments
- Its user interface explicitly guides the user to point the camera at the desired target
- We will open source the code so that anyone can use, incorporate or build on it for free:

<http://www.ski.org/project/display-reader>

References

[1] Morris, T., Blenkhorn, P., Crossey, L., Ngo, Q., Ross, M., Werner, D., & Wong, C. (2006). Clearspeech: A display reader for the visually handicapped. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 14(4), 492-500.

[2] Rasines, I., Iriondo, P., & Díez, I. (2012). *Real-Time display recognition system for visually impaired* (pp. 623-629). ICCHP 2012.

[3] Fusco, G., Tekin, E., Giudice, N. A., & Coughlan, J. M. (2015, October). Appliance Displays: Accessibility Challenges and Proposed Solutions. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility* (pp. 405-406). ACM.

[4] Fusco, G., Tekin, E., Ladner, R. E., & Coughlan, J. M. (2014, October). Using computer vision to access appliance displays. In *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility* (pp. 281-282). ACM.

[5] Guo, A., Chen, X. A., & Bigham, J. P. (2015, April). ApplianceReader: A Wearable, Crowdsourced, Vision-based System to Make Appliances Accessible. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 2043-2048). ACM.

Thanks to...

- Programming assistance: Dr. Huiying Shen (Smith-Kettlewell)
- Helpful feedback on UI design from Dr. Joshua Miele (Smith-Kettlewell) and Dr. Nicholas Giudice (Univ. Maine)
- Very patient volunteer testers!!!
- Funding from NIH and NIDRR / National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR)