

Dawggles

David Gelody, Aidan Lincke, Pat Lucas

Use Case



- Affordable & open source smart goggles
- Heads up display with partial transparency to the real world
- Specific use case: real time text translation and display
- Including a camera and controller that can run text and language recognition models, and translate from/to several languages
- Directly applicable to those traveling
 - e.g. languages with characters have to be imaged for fast translation
 - Skiing, running, hiking, etc. in foreign countries where pulling out your phone is inconvenient
- Affordable, open source pipeline for those looking to expand on more complicated models (e.g. lip reading, facial recognition, etc.)

Use-Case Requirements

Goggles

- Strap must be comfortable and adjustable
- Lens must have at least 70% VLT for daytime and nighttime use



Display

- Display must appear farther away so as not to strain the user's eyes
- Display must be able to show 20 readable characters at once
- Display must scroll through letters in order to complete translations
- Display must be virtually see through when turned off and partially see through when turned on
- Display must be reflective (so that it's hardware does not obstruct the users vision)

Use-Case Requirements

Camera

- Camera must have high enough resolution ($\geq 5\text{Mp}$) to run high end character recognition NNs
- Camera must be mounted firmly on top of the goggles and centered between the users eyes
- Camera must be compatible with selected controller and not a WiFi camera

Controller

- Controller must be compatible with selected camera
- Controller must be able to run character/language recognition in near real time ($< 3\text{s}$)
- Controller must be able output translation to display in real time ($< 0.5\text{s}$)

Use-Case Requirements

Software

- Must scrape words from image that are at the center of eyesight (height and angle adjusted to camera mount position) at 1 sec / word
- Must run language detection software and query translations to a variety of languages within 3 sec / word
- Must pack and resize words so that they are easily readable to the user
- Must update display with adequate refresh rate (60Hz) and scroll translations

Technical Challenges

Reflective Display

- Design
- Transparency & Contrast
- Coverage
- Comfortability

Character / Language Recognition

- Open source models?
- Query openAI?

Technical Challenges

Electrical Housing

- Physically mounting display and reflective components
- Physically mounting camera and microcontroller, hiding wiring, battery

Camera / CV

- Calibrating camera and adjusting eye center in image
- Will likely need our own CV to box the centermost words in image
- Ordering, queuing, and normalizing these images for a NN

Technical Challenges

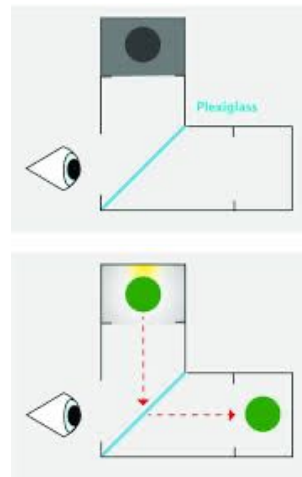
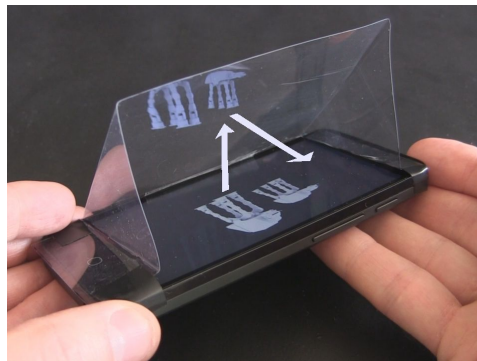
Risk Mitigation

- Display strain
 - Buy multiple angles of transparency, buy multiple concave lenses, buy multiple displays?
- Display bulkiness
 - Could downsize character and resolution requirements to make housing more minimal
- Language detection
 - Could stick to one language translation – Chinese to English

Solution Approach

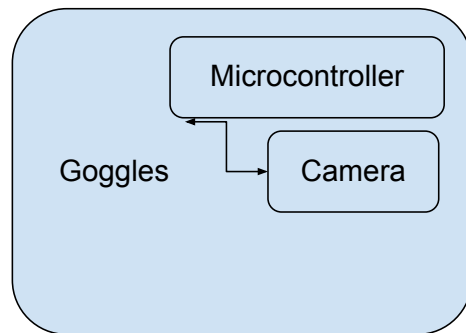
Reflective Display

- Pepper's Ghost illusion with beam splitter
- concave lens to shrink image
- micro OLED Display (great contrast, transparent when off, maybe a low blue light model)



Other Hardware

- Raspi4, raspi camera, battery, wires,



Testing, Verification, and Metrics

Reflective Display - Test & rate combinations (for comfortability) of 3 major variables

- Different beamsplitter angles → different display / transparency ratio (45 deg angle for 50% display and 50% transparency)
- Different height and angle positions of OLED/beamsplitter combination relative to eye
- Different concave lenses and positions for different perceived distances

Software - speed, performance, and efficiency

- Speed and performance of both image transfer and text recognition models at different resolutions (may need to downsample images)
- Necessary character size and scroll rates for fast but readable translations

Tasks and Division of Labor

- Reflective Display part selection (OLED, concave lens, beam splitter) - Aidan & Pat
- Microcontroller and camera selection - Aidan & David
- Camera → Controller → Display Hardware/Software interfacing - Aidan & Pat
- Computer Vision for text detection, center selection, and image resizing - Pat
- Image queue and dequeue management for NN processing - David
- NN (CNN) model selection, testing, and tuning - David & Pat
- Translation language selection and querying - Aidan
- Translation queue and deque management for display - David
- Text → display scrolling speed and size configuration - Aidan & Pat
- Testing - David, Aidan, & Pat

Schedule

Concluding Remarks

- These goggles will be a general-purpose heads-up display containing several functions useful for day-to-day life.
- By developing the software and hardware of the goggles in open-source, others may freely develop additional features.

