Unmasked

Hye Yeon Nam
Louisiana State University
Baton Rouge, Louisiana
hyenam@lsu.edu

Iyleah Hernandez
Dominican University
River Forest, Illinois
ihernandez@my.dom.edu

Brendan Harmon
Louisiana State University
Baton Rouge, Louisiana
baharmon@lsu.edu

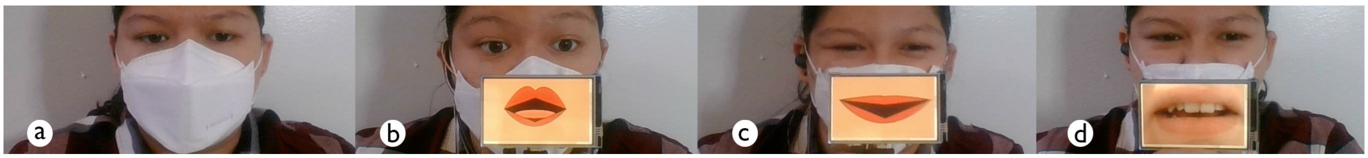

Figure 1. Unmasked is an expressive lip tracking interface for clear communication while wearing a mask. In the user study, a user expresses a) sadness while wearing a regular mask, b) surprise while wearing Unmasked with accelerometers, c) happiness while wearing Unmasked with LED tracking, and d) disgust while wearing Unmasked with streaming video.

ABSTRACT
Due to the COVID-19 pandemic, wearing a mask to cover one's mouth is recommended in public spaces to prevent the spread of the virus. Wearing masks hinders our ability to express ourselves, as it is hard to read facial expressions much less lips behind a mask. We present Unmasked, an expressive interface using lip tracking to enhance communication while wearing a mask. Unmasked uses three different methods – either accelerometers, LEDs with a camera tracking, or streaming video – to make speaking while wearing a mask more expressive. Unmasked aims to improve communication during conversations while wearing a mask. This device will help people express themselves while wearing a mask by tracking their mouth movements and displaying their facial expressions on an LCD mounted on the front of the mask. By enhancing communication while wearing a mask, this prototype makes social distancing less disruptive and more bearable, metaphorically closing some of the distance between us.

Author Keywords
COVID-19; Emotions; Expressive interface; Lip tracking; Mask; Physical computing

CCS Concepts
•**Human-centered computing** → **Human computer interaction (HCI);** *Haptic devices;* User studies;

INTRODUCTION
Face masks have become ubiquitous around the world with the COVID-19 pandemic. While wearing a mask is a critical measure for preventing the spread of viruses like COVID-19, covering the bottom two-thirds of the face profoundly impacts our ability to communicate. Wearing a mask significantly reduces the transmission of verbal and non-verbal information, lowering the quality of communication. For verbal communication, masks attenuate the high frequencies and lower the decibel (db) level [7]. For non-verbal communication, masks hide the mouth, the primary channel for non-verbal communication. Since the expression of many emotions relies heavily on subtle mouth movements [14], masks hinder our ability to understand each other's expressions and emotions. The aim of this project, Unmasked, is to enable clear communication while wearing a mask. Unmasked uses lip tracking sensors, either accelerometers, LEDs, or video, as input. It is implemented in JavaScript so that it can easily be deployed on web platforms in the future. The design team conducted a counterbalanced user study comparing the three versions of Unmasked with a regular face mask.

RELATED WORK
Research on the human body as a user interface has been focused on voluntarily movable body parts such as eyes, fingers, and arms and on the body as a whole [1, 2, 3, 8, 9, 10, 11, 12, 13, 15]. With the prevalence of masks during the COVID-19 pandemic there is an urgent need to rethink the mouth as an interface as our ability to communicate is seriously impacted. Recent research on the lips as an interface includes methods based on video and audio. Machine learning has been used to read lips in video footage for automatic captioning [4]. Masks, however, hide the lips in video footage. A DIY voice activated mask with a builtin microphone has been developed that renders moving lips on a flexible LED array based on decibel level [6]. This interactive mask can abstractly represent lips opening, closing, or smiling. Audio, however, is not enough information to reconstruct lips' full range of motion. Research is needed to track lips that are occluded by obstacles such as masks.

UIST '20 Adjunct, October 20–23, 2020, Virtual Event, USA

ACM ISBN 978-1-4503-7515-3/20/10.
https://doi.org/10.1145/3379350.3416137

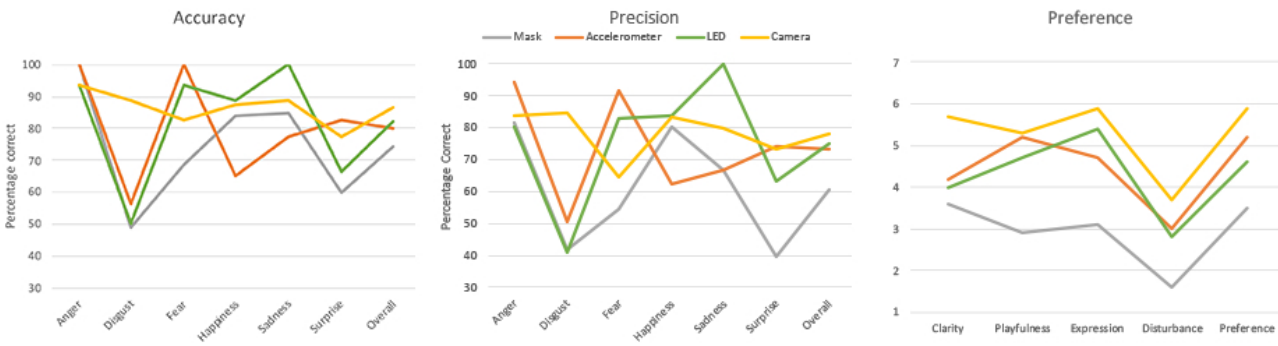

Figure 2. Study results for accuracy (number of the correct answers), precision (rate of the correct answers), and preference.

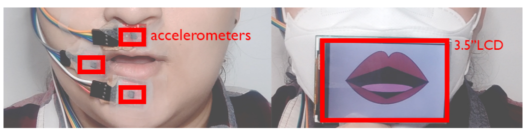

Figure 3. Accelerometers inside and the LCD outside of the mask.

IMPLEMENTATION

There are three versions of Unmasked using different methods to sense and display the lips. The first version has accelerometers attached to the top, side, and bottom of the lips that track the position of the lips (Fig. 3). The accelerometers are attached using surgical transparent tape for adhesion and insulation. The sensors need to be calibrated based on their initial position. Since accelerometers are compact analog sensors that can measure the acceleration of key points on the lips, they do not require much space. The second version uses a small camera mounted inside the mask to track LEDs that are attached to the top, side, and bottom of the lips. The camera needs to be mounted far enough from the mouth to capture the full range of motion of the lips. In the first and second versions, the coordinates of the sensors are interpolated in P5.js on an Arduino microcontroller as curves to represent the lips. The animated lips are rendered in real time on a 3.5" LCD. The third version uses a small camera mounted inside the mask to stream video directly to the LCD mounted on the front of the mask.

EVALUATION

We conducted a user study comparing the three versions of Unmasked against a regular mask as a control. Eighteen volunteers (6 f., mean: 21.8y) were recruited for the user study and separated into two groups (n=10, n=8). For each case, a performer read Lorem ipsum placeholder text, while expressing a sequence of six basic emotions – anger, disgust, fear, happiness, sadness, and surprise [5] – for 15 seconds each in random order. Participants were given a multiple choice questionnaire and asked to select which emotion was being expressed. They also ranked each case in terms of clarity, playfulness, expression, disturbance, and overall preference on a scale from 1 to 7 (from lowest to highest).

Unmasked was developed to enable users to share their emotions while wearing a mask. In the user study a plain mask was used as a baseline to test for significance. Since participants could choose more than one emotion per question, the results were measured both for accuracy and precision. Thus, the closer a measure of precision is to its accuracy, the less confusion there was for the participant on average. While participants were able to accurately identify the emotion portrayed an average of 74.38% (SD=2.71) of the time when only a mask was worn, their precision was lower with only a 60.75% (SD=2.11) rate of correctness. Unmasked with accelerometers averaged 80.21% for accuracy and 72.29% for precision. Unmasked with LEDs averaged 82.08% for accuracy and 75.28% for precision. Unmasked with a camera averaged 86.46% for accuracy and 78.25% for precision. All versions of Unmasked showed a significant increase in accuracy as well as precision in comparison to a regular mask. All values can be considered significant with a p-value of < 0.016. In terms of overall preference, masks were rated 3.5 out of 7 (SD=0.49), Unmasked with accelerometers was rated 5.2, Unmasked with LEDs was rated 4.6, and Unmasked with a camera was rated highest at 5.9. Unmasked with accelerometers, LEDs, and a camera were all rated higher than the plain mask in terms of clarity, playfulness, expression, and disturbance.

CONCLUSION

While we hypothesized that Unmasked with streaming video would be considered uncanny this design was rated the highest. There is an artistic quality to this disembodiment. Unmasked reassembles and contextualizes the body like some works of contemporary video art by artists like Nam June Paik or Tony Oursler. Unmasked with sensors may be a better choice for speakers who feel uncomfortable with their face on display and would rather be portrayed with a graphical abstraction of their expression. We plan to develop a new version of Unmasked for use in classrooms and conduct a larger user study. Our goal is to design a more portable, precise system with more sensors and more sophisticated graphics that can be streamed to external displays, shared via the web, or act as an Internet of Things (IoT) interface.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under award OAC-1852454 with additional support from the Center for Computation & Technology at Louisiana State University.

REFERENCES

[1] Jason S Babcock and Jeff B Pelz. 2004. Building a lightweight eyetracking headgear. In *Symposium on Eye Tracking Research & Applications*. ACM Press, 109–114. DOI: http://dx.doi.org/doi.org/10.1145/968363.968386

[2] Andreas Bulling and Hans Gellersen. 2010. Toward mobile eye-based human-computer interaction. *IEEE Pervasive Computing* 9, 4 (oct 2010), 8–12. DOI: http://dx.doi.org/10.1109/MPRV.2010.86

[3] Liwei Chan, Rong-Hao Liang, Tsai Ming-Chang, Cheng Kai-Yin, Su Chao-Huai, Chen Mike Y., Cheng Wen-Huang, and Chen Bing-Yu. 2013. FingerPad: Private and Subtle Interaction Using Fingertips. In *Proceedings of the 26th annual ACM symposium on User interface software and technology (UIST '13)*. ACM Press, 255–260. DOI: http://dx.doi.org/10.1145/2501988.2502016

[4] Joon Son Chung, Andrew Senior, Oriol Vinyals, and Andrew Zisserman. 2017. Lip reading sentences in the wild. In *30th IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 3444–3453. DOI: http://dx.doi.org/10.1109/CVPR.2017.367

[5] Paul Ekman. 1992. Are there basic emotions? *Psychological Review* 99, 3 (1992), 550–553. DOI: http://dx.doi.org/10.1037/0033-295x.99.3.550

[6] Tyler Glaiel. 2020. How to make a voice activated LED facemask (DIY Guide). (2020). https://bit.ly/2ZS9aVs

[7] Alexander Goldin, Barbara E. Weinstein, and Nimrod Shiman. 2020. How do medical masks degrade speech perception? *Hearing Review* 27, 5 (2020), 8–9.

[8] Hsin-Liu (Cindy) Kao, Christian Holz, Asta Roseway, Andres Calvo, and Chris Schmandt. 2016. DuoSkin: rapidly prototyping on-skin user interfaces using skin-friendly materials. In *Proceedings of the 2016 ACM International Symposium on Wearable Computers (ISWC '16)*. ACM Press, 16–23. DOI: http://dx.doi.org/10.1145/2971763.2971777

[9] Arshad Khan, Joan Sol Roo, Tobias Kraus, and Jürgen Steimle. 2019. Soft Inkjet Circuits: Rapid Multi-Material Fabrication of Soft Circuits using a Commodity Inkjet Printer. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (UIST '19)*. ACM Press, 341–354. DOI: http://dx.doi.org/10.1145/3332165.3347892

[10] Dongheng Li, Jason Babcock, and Derrick J. Parkhurst. 2006. openEyes: a low-cost head-mounted eye-tracking solution. In *Symposium on Eye Tracking Research & Applications*. ACM Press, 95–100. DOI: http://dx.doi.org/10.1145/1117309.1117350

[11] Masa Ogata, Yuta Sugiura, Yasutoshi Makino, Masahiko Inami, and Michita Imai. 2013. SenSkin: Adapting skin as a soft interface. In *Proceedings of the 26th annual ACM symposium on User interface software and technology (UIST '13)*. ACM Press, 539–543. DOI: http://dx.doi.org/10.1145/2501988.2502039

[12] Qifan Pu, Sidhant Gupta, Shyamnath Gollakota, and Shwetak Patel. 2013. Whole-home gesture recognition using wireless signals. In *Proceedings of the 19th annual international conference on Mobile computing & networking (MobiCom '13)*. ACM Press, 27–38. DOI: http://dx.doi.org/10.1145/2500423.2500436

[13] Michael Rohs and Albrecht Schmidt. 2013. OMG!: a new robust, wearable and affordable open source mobile gaze tracker. In *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services (MobileHCI '13)*. ACM Press, 408–411. DOI: http://dx.doi.org/10.1145/2493190.2493214

[14] Martin Wegrzyn, Maria Vogt, Berna Kireclioglu, Julia Schneider, and Johanna Kissler. 2017. Mapping the emotional face. How individual face parts contribute to successful emotion recognition. *PLoS ONE* 12, 5 (may 2017). DOI: http://dx.doi.org/10.1371/journal.pone.0177239

[15] Martin Weigel, Tong Lu, Gilles Bailly, Antti Oulasvirta, Carmel Majidi, and Jürgen Steimle. 2015. iSkin: Flexible, stretchable and visually customizable on-body touch sensors for mobile computing. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM Press, 2991–3000. DOI:http://dx.doi.org/10.1145/2702123.2702391