Mobile Gaze Mapping: A Python package for mapping mobile gaze data to a fixed target stimulus

Jeff J MacInnes[1, 2]**, Shariq Iqbal**[3]**, John Pearson**[2]**, and Elizabeth N. Johnson**[2, 4]

1 Institute for Learning and Brain Sciences, University of Washington, Seattle, WA **2** Center for Cognitive Neuroscience, Duke University, Durham, NC **3** University of Southern California, Los Angeles, CA **4** Wharton Neuroscience Initiative, University of Pennsylvania, Philadelphia, PA

Summary

Mobile eye-trackers allow for measures like gaze position to be recorded under naturalistic conditions where an individual is free to move around. Gaze position is typically recorded relative to an outward facing camera attached to the eye-tracker and approximating the point-of-view of the individual wearing the device. As such, gaze position is recorded relative to the individual's position and orientation, which changes as the participant moves. Since gaze position is recorded without any reference to fixed objects in the environment, this poses a challenge for studying how an individual views a particular stimulus over time.

This toolkit addresses this challenge by automatically identifying the target stimulus on every frame of the recording and mapping the gaze positions to a fixed representation of the stimulus.

It does this by identifying matching keypoints between the reference stimulus and each frame of the video. Keypoints are obtained using the Scale Invariant Feature Transform algorithm (SIFT)(Lowe, 2004), and matches between keypoints are found using the Fast Approximate Nearest Neighbor search algorithm (FLANN)(Muja & Lowe, 2009), both implemented in OpenCV(Bradski, 2000). (We note that while use of the SIFT algorithm is free for non-commercial, research, and educational purposes, all commercial applications require a purchased license.)

Once matching keypoints have been identified, we determine the 2D linear transformation that maps keypoints from the video frame to key points on the reference image. Once determined, this same transformation is applied to the gaze position sample corresponding to the given video frame, resulting in gaze position being expressed in terms of the pixel coordinate system of the reference image.

For a more detailed overview of this process, and the results of using this toolkit to compare gaze accuracy and precision across 3 popular models of mobile eye-trackers, please see (MacInnes, Iqbal, Pearson, & Johnson, 2018)

Acknowledgments

References

Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools.*

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, *60*(2), 91–110. doi:10.1023/B:VISI.0000029664.99615.94

MacInnes, J. J., Iqbal, S., Pearson, J., & Johnson, E. N. (2018). Wearable eye-tracking for research: Automated dynamic gaze mapping and accuracy/precision comparisons across devices. *bioRxiv.* doi:10.1101/299925

Muja, M., & Lowe, D. G. (2009). FAST approximate nearest neighbors with automatic algorithm configuration. INSTICC; SciTePress. doi:10.5220/0001787803310340