**Reading Report**

**Unlocking Eye Gestures with Earable Inertial Sensing for Accessible HCI**

**1.Introduction**

In the realm of Human-Computer Interaction, the quest for more accessible and inclusive interfaces has led to the exploration of hands-free interaction methods. Traditional input mechanisms, such as keyboards and pointing devices, often pose significant barriers for individuals with motor impairments or those whose hands are otherwise occupied. Eye gestures have emerged as a promising alternative, offering a natural and intuitive means of interaction. These gestures are already utilized in various applications, including device control, biometrics, visual analytics, and healthcare, particularly for conditions like Alzheimer's disease.

However, conventional methods for detecting eye gestures typically rely on visual information processing devices or smart glasses, which can be intrusive, costly, and privacy-invasive. Recent advancements in wearable technology, particularly earable devices equipped with Inertial Measurement Unit sensors, present a novel and compelling alternative. Earables, which are increasingly integrated into daily life for hands-free communication and entertainment, offer a familiar and non-intrusive platform for eye gesture detection. By leveraging IMU data, these devices can capture subtle eye movements and translate them into actionable commands, enabling hands-free interaction for users with limited mobility.

Despite the potential of earable-based inertial sensing for eye gesture detection, several challenges remain. Motion-induced interference from other head movements, facial expressions, or lower body activities can complicate the accurate isolation and interpretation of specific eye gestures. Additionally, distinguishing between spontaneous and intended eye blinks is crucial, as intentional blinks can be interpreted as commands for further actions. Temporal variability within an individual further exacerbates these challenges, as the same gesture may vary in timing and intensity over time.

This paper introduces a novel framework that addresses these challenges by utilizing IMU sensors embedded in earable devices for real-time eye gesture detection. The proposed framework demonstrates high accuracy in recognizing eye gestures across diverse activities, such as sitting, walking, running, and driving, achieving a macro F1 score of 0.85 on a self-collected dataset. By translating specific eye movements into actionable commands, this framework enhances the accessibility and inclusivity of HCI, offering a practical solution for continuous eye gesture recognition without compromising user comfort or privacy.

**2.Related Work**

1. **SkinTrack: Using the Body as an Electrical Waveguide for ContinuousFinger Tracking on the Skin[1].**

It brought up a method for continuous finger tracking using the skin as an electrical waveguide. The system utilizes the propagation characteristics of electric current on the skin to achieve high-precision finger position detection.On one hand,it provides a new input method that allows for finger tracking without the need for additional devices and offers high accuracy and responsiveness, suitable for various interaction scenarios.But on the other hand,its performance might be affected by the user's physical condition, such as variations in skin moisture.Also,it requires users to maintain certain postures during use, which might impact comfort and naturalness.

**(2)EyeMU Interactions: Gaze + IMU Gestures on Mobile Devices[2]**

Gaze + IMU Gestures on Mobile Devices combines eye tracking with inertial measurement unit (IMU) gestures to create a new type of interaction method, allowing users to operate devices through gaze and gestures simultaneously.The advantage is that it Enhances user interaction capabilities with mobile devices, especially in complex tasks and Improves user experience, making operations more intuitive and efficient.However,the accuracy of eye tracking may be affected by environmental lighting and user fatigue.And users may require some learning and adaptation, which could pose initial difficulties.

**(3)Eye, Head and Torso Coordination During Gaze Shifts in Virtual Reality[3]**

The study examines the coordination between the eyes, head, and torso during gaze shifts in virtual reality environments, analyzing how these factors influence users' gaze behavior and interaction experience.The study provides important physiological insights for user experience design in virtual reality.And it helps optimize interaction design in virtual environments, enhancing immersion and naturalness.But the sample size of the experiment is relatively small, which may limit the generalizability of the results.What’s more,it primarily focuses on physiological coordination, without sufficiently exploring the psychological factors affecting gaze behavior.

**(4)Headar: Sensing Head Gestures for Confirmation Dialogs on Smartwatches with Wearable Millimeter-Wave Radar[4]**

This paper proposes a wearable device based on millimeter-wave radar that can recognize head gestures for confirmation dialogs on smartwatches.It offers a contactless interaction method, enhancing user convenience and privacy protection and effectively operates in various environments, unaffected by lighting and other interference factors.But the disadvantages are that the complexity of technological implementation may increase costs, limiting widespread application and that recognition of head gestures may be influenced by the speed and habits of user movements, leading to a certain risk of misrecognition.

**3.Research Methods**

The research methodology for this study involves a combination of data collection, preprocessing, feature engineering, and machine learning techniques to accurately detect eye gestures using earable IMU sensors. The following sections outline the key components of the proposed framework.

**(1)Data Collection and Preprocessing**

The study involved 14 participants who performed a variety of activities, including sitting, walking, running, driving, and eating, while wearing earable devices equipped with IMU sensors. The IMU data, which includes accelerometer and gyroscope readings, was collected in real-time and stored for further analysis. To ensure data consistency, each activity was repeated four times by each participant.

Preprocessing steps included noise reduction using Butterworth filtering and data segmentation into fixed-length windows. The Fast Fourier Transform (FFT) was applied to convert the data into the frequency domain, capturing the magnitude of frequency components over time. Kalman filtering was also employed to correct any orientation issues that could lead to misclassification.

**(2)Feature Engineering**

Feature engineering played a crucial role in enhancing the accuracy of eye gesture detection. Statistical features derived from both the accelerometer and gyroscope data, such as mean, standard deviation, median, and skewness, were included. Additionally, features related to facial muscles, such as jerk, magnitude, and cross-correlation, were incorporated to capture the nuances of eye movements.

**(3)Machine Learning Models**

The proposed framework utilized several machine learning models, including Decision Tree (DT), Random Forest (RF), Gated Recurrent Unit (GRU), and Recurrent Neural Network (RNN), to classify and detect eye gestures. The GRU model, in particular, demonstrated superior performance in capturing temporal dependencies within the IMU data, achieving an F1 score of 85.0%.

**(4)Thresholding and Frequency Counting**

Thresholding was employed to distinguish between intended eye gestures and spontaneous blinks. A threshold value was set based on the IMU signature window, and gestures were recognized if the length of the interval exceeded the threshold. The frequency counter module tracked the number of recognized eye gestures over a given period, providing a measure of gesture rate within predefined time intervals.

**(5)Experimental Setup and Evaluation**

The core framework was developed using PyTorch, and the user interface was implemented as an iOS application. Training was conducted over 200 epochs with a batch size of 64, and the dataset was split into 80% training and 20% testing sets. The model's performance was evaluated using the macro F1 score, which accounts for the imbalance in the dataset.

The results demonstrated that the proposed framework achieved high accuracy in eye gesture recognition across diverse activities, with an F1 score of 0.85 on the self-collected dataset. The GRU model outperformed other models, highlighting its effectiveness in capturing temporal dependencies within the IMU data.

**4.Applicability**

**(1)Broad Application Scenarios**

The proposed framework for eye gesture detection using earable IMU sensors is highly applicable in various domains, including device control, biometrics, visual analytics, and healthcare (e.g., assisting individuals with Alzheimer's or mobility impairments).Its hands-free interaction capability makes it particularly suitable for users with limited mobility or in situations where hands are occupied.

**(2)High Accuracy and Real-Time Performance**

The framework achieves a macro F1 score of 0.85 on a self-collected dataset, demonstrating its effectiveness in recognizing eye gestures across diverse activities such as sitting, walking, running, and driving.The use of supervised machine learning models (e.g., GRU) ensures robust performance even in dynamic environments.

**(3)Privacy and Cost-Effectiveness**

Unlike vision-based systems, earable devices offer enhanced privacy protection and are more cost-effective, making them a practical solution for continuous eye gesture recognition.The discreet and non-intrusive design of earables supports extended use without compromising user comfort.

**(4)Complementary to Existing Technologies**

The framework can work alongside visual information processing devices, providing a complementary approach to eye gesture detection.It addresses the limitations of traditional methods, such as occlusion, high cost, and privacy concerns associated with camera-based systems.

**(5)Real-World Usability**

The framework is validated in real-world scenarios, including noisy environments and activities like running and driving, showcasing its adaptability and robustness.The proposed pipeline is lightweight and suitable for edge devices, ensuring efficient performance without significant computational overhead.

**5.Limitations**

**(1)Motion-Induced Interference**

The biggest challenge is isolating specific eye gestures from mixed signals caused by head movements, facial expressions, or lower body activities.While the framework incorporates noise filtration techniques, achieving consistent accuracy across all real-world conditions remains a challenge.

**(2)Distinguishing Intended vs. Spontaneous Gestures**

Differentiating between intentional eye gestures (e.g., winks) and spontaneous blinks is crucial but difficult, as unintentional blinks may be misinterpreted as commands.Temporal variability in individual gestures further complicates this task.

**(3)Dependence on Sensor Placement and Fit**

The accuracy of gesture recognition is influenced by the placement and physical fit of earable devices. Variations in sensor positioning or fit can introduce noise and reduce reliability.Ensuring consistent sensor placement across users is essential but may not always be feasible.

**(4)Limited Gesture Vocabulary**

The current framework focuses on a limited set of eye gestures (e.g., blinking, winking). Expanding the gesture vocabulary to include more complex interactions (e.g., head gestures, facial expressions) requires further research.The system may struggle to recognize subtle or nuanced gestures.

**(5)Dataset Limitations**

The framework is trained and tested on a self-collected dataset with 14 participants, which may not fully represent the diversity of users or real-world conditions.The dataset imbalance (fewer non-target class labels) could affect the model's generalizability and performance in detecting unintended gestures.

**(6)Response Time**

While the framework achieves competitive response times (e.g., 1150 ms in silent environments), it is slower than video-based systems (e.g., 250 ms).Further optimization is needed to reduce latency and improve real-time interaction.

**(7)Energy Consumption**

Continuous use of IMU sensors in earable devices may lead to increased energy consumption, potentially limiting battery life.Balancing performance with energy efficiency is a critical consideration for practical deployment.

**6.Experiment result**

**(1)existing research results**

Innovative Input Methods: Research indicates that utilizing the skin as an electrical waveguide (as demonstrated by SkinTrack) can achieve high-precision finger tracking, providing users with a new interaction method. This innovation not only enhances user experience but also expands the application scenarios of wearable devices.

Multimodal Interaction: Combining different interaction modalities (such as eye tracking and gesture control) can significantly enhance user experience. The EyeMU research shows that the integration of gaze and inertial gestures makes mobile device operations more intuitive and efficient.

Coordination of Physiological and Psychological Factors: Studies in virtual reality emphasize the importance of coordination between the eyes, head, and torso for users’ gaze behavior and interaction experience. This finding provides crucial physiological insights for future interaction design.

Potential of Contactless Interaction: The application of millimeter-wave radar technology (as seen in Headar's research) demonstrates the possibilities of contactless interaction, which not only enhances user convenience but also strengthens privacy protection.

**(2）Evaluation of Specific Research Directions Related to Eye Gestures**

Interaction Design Combining Eye Movements and Gestures: Future research can further explore how to integrate eye gestures with other input methods (such as touch and voice) to create a more flexible and natural interaction experience.

Environmental Adaptability: Research should focus on the performance of eye gestures in different environments (such as varying lighting conditions and user states) to ensure the technology's generalizability and stability.

User Adaptability and Learning Curve: Studies need to evaluate the adaptation process of users when utilizing eye gestures and develop corresponding training and guidance programs to reduce the learning curve.

**（3）Potential Challenges in the Future**

Technical Complexity: The technical implementation of wearable devices may face issues of complexity, particularly when integrating multiple sensors and algorithms, which could lead to increased costs and development difficulties.

Accuracy and Reliability: The accuracy of eye tracking may be influenced by environmental factors (such as lighting and reflections) and individual differences among users, making it a challenge to improve the system's reliability.

User Privacy and Security: As more biometric technologies (like eye tracking) are employed, concerns regarding user privacy and data security will become increasingly important, necessitating the establishment of appropriate protective measures.

**（4）Outlook**

Future research should focus on optimizing the user interaction experience of wearable devices and exploring more natural and intuitive interaction methods. With technological advancements, the application of eye gestures is expected to become more widespread, potentially playing a significant role in fields such as healthcare, gaming, and education. Additionally, researchers need to pay attention to user feedback and needs to ensure the usability and satisfaction of the technology. Through continuous innovation and iteration, the wearable HCI field is likely to achieve higher user experience and broader application prospects.

**7.Conclusion and inspiration**

The study on unlocking eye gestures with earable inertial sensing is a remarkable innovation that bridges the gap between technology and accessibility. As I reflect on its implications, I am struck by how it redefines the boundaries of human-computer interaction while addressing critical issues like privacy and inclusivity. The idea of using earable devices to detect and interpret eye gestures is both ingenious and practical, offering a hands-free solution that could transform the lives of individuals with limited mobility. This technology not only empowers users but also demonstrates how innovation can be harnessed to create tools that are both functional and respectful of personal privacy.

One aspect that particularly resonates with me is the framework’s adaptability across various real-world scenarios, such as walking, running, or driving. This versatility highlights its potential to seamlessly integrate into daily life, making it more than just a niche solution but a universal tool for enhancing human capabilities. However, I also recognize the challenges that come with this innovation, such as motion-induced interference and the difficulty in distinguishing between intentional and spontaneous gestures. These hurdles remind me that even the most groundbreaking technologies require continuous refinement to reach their full potential.

What impresses on me most is the broader impact this research could have beyond HCI. In healthcare, for instance, it could provide a lifeline for patients with conditions like ALS or spinal cord injuries, enabling them to communicate and interact with their environment in ways that were previously unimaginable. This potential to improve quality of life underscores the profound societal value of such technological advancements.

As I think about the future, I am inspired by the possibilities this research opens up. Expanding the gesture vocabulary, integrating multimodal inputs, and refining the framework to handle more complex interactions could unlock even greater potential. This study is a testament to the power of interdisciplinary collaboration and the importance of designing technology with empathy and inclusivity at its core.

In conclusion, this research is not just a technical achievement but a step toward a more accessible and equitable world. It reminds us that technology, when thoughtfully designed, can be a force for good, empowering individuals and fostering a more inclusive society. As this framework evolves, I am eager to see how it will continue to shape the future of human-computer interaction and beyond.

**References:**

The topic of the reading report is based on

G. Chugh, S. Chakraborty and S. Chakraborty, "Unlocking Eye Gestures with Earable Inertial Sensing for Accessible HCI," 2025 17th International Conference on COMmunication Systems and NETworks (COMSNETS), Bengaluru, India, 2025, pp. 828-832,doi:10.1109/COMSNETS63942.2025.10885707.<https://ieeexplore.ieee.org/document/10885707>

1. Ludwig Sidenmark and Hans Gellersen. 2019. Eye, Head and Torso Coordination During Gaze Shifts in Virtual Reality. ACM Trans. Comput.-Hum. Interact. 27, 1, Article 4 (February 2020), 40 pages. <https://doi.org/10.1145/3361218>
2. Andy Kong, Karan Ahuja, Mayank Goel, and Chris Harrison. 2021. EyeMU Interactions: Gaze + IMU Gestures on Mobile Devices. In Proceedings of the 2021 International Conference on Multimodal Interaction (ICMI '21). Association for Computing Machinery, New York, NY, USA, 577–585. <https://doi.org/10.1145/3462244.3479938>
3. Xiaoying Yang, Xue Wang, Gaofeng Dong, Zihan Yan, Mani Srivastava, Eiji Hayashi, and Yang Zhang. 2023. Headar: Sensing Head Gestures for Confirmation Dialogs on Smartwatches with Wearable Millimeter-Wave Radar. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 7, 3, Article 138 (September 2023), 28 pages. <https://doi.org/10.1145/3610900>
4. Yang Zhang, Junhan Zhou, Gierad Laput, and Chris Harrison. 2016. SkinTrack: Using the Body as an Electrical Waveguide for Continuous Finger Tracking on the Skin. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16). Association for Computing Machinery, New York, NY, USA, 1491–1503. <https://doi.org/10.1145/2858036.2858082>