

Review

Open Access



Review of electrooculography-based human-computer interaction: recent technologies, challenges and future trends

Linkai Tao^{1,#} , Hao Huang^{2,#}, Chen Chen^{3,*} , Loe Feijs¹ , Jun Hu¹ , Wei Chen^{2,3,*} 

¹Department, Industrial Design, Eindhoven University of Technology, Eindhoven 5600 MB, Netherlands.

²The Center for Intelligent Medical Electronics, School of Information Science and Technology, Fudan University, Shanghai 200433, China.

³The Human Phenome Institute, Fudan University, Shanghai 201203, China.

#Authors contributed equally.

Correspondence to: Prof./Dr. Wei Chen, The Center for Intelligent Medical Electronics, School of Information Science and Technology, No.2 Jiaocha Building Section B, Jiangwan Campus, Fudan University, Shanghai 200433, China. E-mail: w_chen@fudan.edu.cn; Prof./Dr. Chen Chen, The Human Phenome Institute, Administrative Building, Zhangjiang Campus, Fudan University, Shanghai 201203, China. E-mail: chenchen_fd@fudan.edu.cn

How to cite this article: Tao L, Huang H, Chen C, Feijs L, Hu J, Chen W. Review of electrooculography-based human-computer interaction: recent technologies, challenges and future trends. *Conn Health Telemed* 2023;2:2000010. <https://dx.doi.org/10.20517/chatmed.2023.05>

Received: 9 Aug 2023 **First decision:** 24 Aug 2023 **Revised:** 14 Sep 2023 **Accepted:** 25 Sep 2023 **Published:** 28 Sep 2023

Academic Editor: Stefano Omboni **Copy Editor:** Yanbing Bai **Production Editor:** Yanbing Bai

Abstract

Electrooculography-based Human-Computer Interaction (EOG-HCI) is an emerging field. Research in this domain aims to capture eye movement patterns by measuring the corneal-retinal potential difference. This enables translating eye movements into commands, facilitating human-computer interaction through eye movements. This paper reviews articles published from 2002 to 2022 in the EOG-HCI domain, aiming to provide a comprehensive analysis of the current developments and challenges in this field. It includes a detailed and systematic analysis of EOG signal electrode arrangement, hardware design for EOG signal acquisition, commonly used features, and algorithms. Representative studies in each section are presented to help readers quickly grasp the common technologies in this field. Furthermore, the paper emphasizes the analysis of interaction design within the EOG-HCI domain, categorizing different interaction task types and modalities to provide insights into prevalent interaction research. The focus of current research in the field is revealed by examining commonly used evaluation metrics. Lastly, a user-centered EOG-HCI research model is proposed to visually present the current research status in the EOG-HCI field from the perspective of users. Additionally, we highlight the challenges and opportunities in this



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.



field.

Keywords: Electrooculography, user interface, human-computer interaction, interaction design

INTRODUCTION

The scope of human-computer interaction (HCI) research lies in studying the process of interaction between humans and machines, which constitutes a loop^[1]. In a typical HCI loop, the process is as follows: Step 1, the human brain autonomously or non-autonomously generates interaction intentions; Step 2, the interaction intentions lead to activities (such as limb movements), and this category of activity-generated information is defined as "trigger" in this review study; Step 3, the human-computer interface translates the "trigger" into control information and delivers it to the machine; Step 4, the machine reacts to the control information; Step 5, the human-computer interface acquires the machine's reaction and converts it into "feedback"; Step 6, the human actively or passively receives the "feedback" through sensors; Step 7, the human brain processes the feedback information. Through one interaction loop after another, the user ultimately achieves their goal of using the interactive system.

With the development of biomedical engineering technology in areas such as biosensing and biofeedback, the introduction of biomedical engineering technology into HCI research allows for more flexible and context-appropriate selection of triggers and feedback in the design of interaction forms.

The eye, as an important organ of the human body, serves both as a sensor and a trigger. Most of the information in the environment is perceived through visual information received by the eyes, and eye movements can also reflect some intentions, such as changes in attention. Therefore, using eye movements as triggers in the HCI loop has direct, simple, and efficient advantages over using other types of triggers, especially for tasks that involve primarily visual information as feedback. Additionally, for users in scenarios where limb movements are restricted, such as individuals whose hands and feet are occupied by other tasks or patients with physical disabilities, providing an eye-based interaction interface can be a better choice.

There exists a stable potential difference between the cornea and retina of the eye^[2]. Mowrer, Ruch, and Miller (1935)^[3] found that placing electrodes around the eyes allowed measurements of potential difference changes caused by eye movements. These studies are considered to have laid the foundation for Electrooculography (EOG) research.

EOG signals in non-invasive brain-computer interfaces (BCIs) offer distinct advantages over Electroencephalograph (EEG) signals. They have significantly higher voltage levels (exceeding 1 mV vs. < 100 μ V for EEG), making them easier to capture and analyze. EOG also enables voluntary control, enhancing user experience by generating specific signal patterns through eye movements. Moreover, EOG is less susceptible to interference, in contrast to EEG, where EOG artifact removal is one of the many challenges. As a result, EOG is an attractive modality when it comes to BCIs. In the field of eye tracking, EOG offers unique advantages. Firstly, in terms of data acquisition: (1) The data source is a physiological electrical signal that reflects the underlying characteristics of the body; (2) It is less affected by lighting conditions and not restricted by makeup, contact lenses, or eyelashes; (3) The data collection apparatus is simple, small, and flexible in use; (4) It can monitor eye movements even when the eyes are closed; (5) It can capture a wide range of eye movements. Moreover, in terms of data, the data format is simple, with low storage, transmission, and computational costs, making it suitable for long-term monitoring.

Based on the EOG-HCI loop, as shown in [Figure 1](#), the human is on the left side of the interactive interface, and the machine with computational capabilities is on the right side of the interactive interface. The entire interaction loop starts from the human's central nervous system generating interaction intentions and controlling eye movements to emit triggers, and ends with the human's central nervous system receiving and processing visual and other feedback information. The interactive interface serves as an intermediary connecting humans and machines, responsible for information acquisition, information parsing, information integration, and information exchange between humans and machines. In eye-tracking-based computer interaction, the interactive interface generally includes an EOG information extraction module, a signal processing module, an interactive task module, and an interactive form that matches the interactive task. The interactive interface plays a central role in the entire interaction loop and is a major research focus in the field of EOG-HCI.

This review study conducted a systematic search, review, and summary of EOG-based human-computer interaction research conducted between 2002 and 2022. The aim was to showcase the development trends in this field, summarize the research methods, identify commonalities in the approaches, and present some distinctive research outcomes.

DATA SCREENING PROCESS

Peer-reviewed journal articles and conference papers that met the criteria from four databases, namely ACM, PubMed, IEEE Xplore, and Science Direct, are involved. The overall process of the literature research strategy is illustrated in [Figure 2](#).

- Inclusion round:

The data set was selected based on the following criteria. Firstly, the titles must include EOG or its related synonyms, specifically: "EOG", "electrooculogram", "electro-oculogram", "electrooculography", or "electro-oculography". Secondly, the titles or abstracts must mention HCI or related terms, such as: "HCI", "human-computer interaction", "interaction", "interface", "VR", "Virtual Reality", "AR", "Augmented Reality", "MR", or "Mixed Reality". Thirdly, the research must be published between 2002 and 2022. Fourthly, the studies must be written in English.

- First exclusion round:

During the second round of screening, we excluded review articles that did not align with the research objectives and any literature that did not match the intended focus. We specifically focused on human-computer interaction research with EOG signals as the primary modality and excluded studies that utilized other types of signals.

- Second exclusion round:

In the third round of screening, we carefully examined the content of the literature and retained those that presented a complete and well-defined interaction loop. We further excluded studies that utilized other modalities and studies that did not use EOG signals as the primary modality.

After the third exclusion round, we finally have retained 113 eligible papers, of which 35 are journal articles. As one of the objectives of this review study is to quantify the research quality in this field, and conference

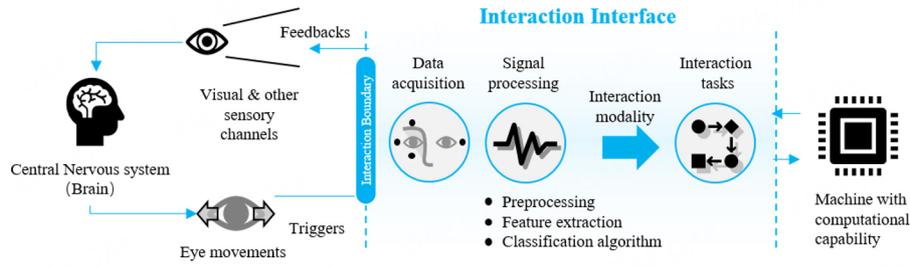


Figure 1. EOG-HCI loop.

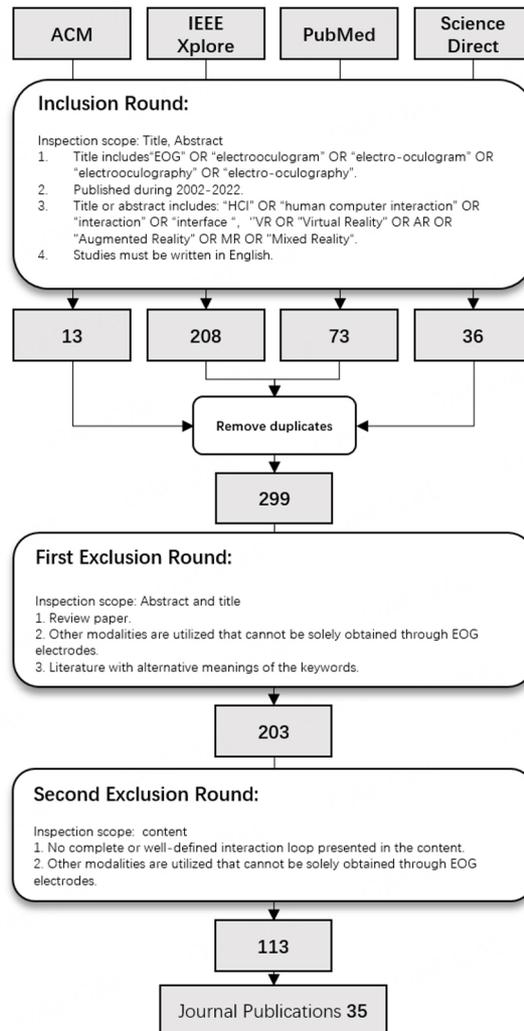


Figure 2. Overall Data Screening Process.

papers are challenging to assess in terms of research quality, the following summary and analysis will focus solely on the scope of journal research.

In order to clarify the trend of the amount of research input and the quality of research output in the research field, we demonstrate the number of publications per 5-year period and the average impact factor per publication per 5-year period (AIFP5) from 2002 to 2022. Shown in [Figure 3](#).

The AIFP5 is calculated as follows:

$$AIFP5 = \frac{\sum_{i=1}^N IF_i}{N} \quad (1)$$

Where N indicates the number of researches published in this five-year period, and IF indicates the impact factor score of the journal in the year of each research published.

As can be seen from [Figure 3](#), EOG-HCI is a vibrant and promising research field.

[Table 1](#) presents information from all the surveyed literature regarding five aspects: Device, electrode arrangement, preprocessing method, feature, and algorithm. Since many studies employ various approaches in some aspects, such as performing preliminary experiments before choosing the best algorithm, we have listed only the most representative information for each article in [Table 1](#).

The rest of the paper is organized as follows. In Section 3, we summarize the EOG signal acquisition techniques used in EOG-HCI research, including electrode arrangement and acquisition hardware system design. Section 4 provides an overview of EOG signal preprocessing, feature extraction, and classification algorithms. In Section 5, we review commonly used evaluation metrics in the research field. Section 6 focuses on summarizing and analyzing interaction design within the research field. Finally, in Section 7, we discuss and conclude based on the research literature covered in this review.

ACQUISITION OF EOG FOR HCI

Electrode arrangement

In the field of EOG-based human-computer interaction, the arrangement and configuration of electrodes are greatly influenced by the research objectives. Based on the motivations of various studies, the electrode configurations in EOG-HCI research can be broadly categorized into three types: standard arrangement, task-oriented arrangement, and wearable-oriented arrangement. We have conducted a statistical analysis of the advantages, disadvantages, and the number of literature using each type of electrode configuration. The results are presented in [Table 2](#).

[Figure 4](#) shows the EOG standard electrode arrangement, with a and b representing horizontal EOG electrodes which produce the horizontal EOG signal (HEOG) and c and d representing vertical EOG electrodes which produce the horizontal EOG signal (VEOG).

In task-oriented electrode arrangement, Li *et al.* used only the EOG signal related to blinking to perform commands on a graphical user interface (GUI)^[5,18-20,32]. For this task, they measured only the vertical EOG signal using a single electrode. Kim *et al.* designed an interaction task where users answered questions by using left-right horizontal eye movements^[37]. Therefore, the interaction system involved a pair of electrodes placed approximately 2.5 cm outside the corner of each eye.

Table 1. Table of Methodologies in EOG-HCI Journal Papers from 2002 to 2022

Reference	Device		Electrode arrangement			Preprocessing method			Feature			Algorithm				
	Custom-designed devices	Commercial devices	Standard	Task-Oriented	Wearable-Oriented	Bandpass filters	Wavelet transform	Others	Original signal	First-order differential signal	Others	Threshold detection	Waveform matching	Neural network	Traditional machine learning	Others
[4]	√		√			not mentioned					√	√		√		
[5]		√		√		√				√		√				
[6]	√		√			√				√		√				
[7]		√		√		√	√			√		√				
[8]		√	√			√					√					√
[9]		√	√			√			√			√				
[10]		√		√		√			√		√					
[11]	√		√				√			√					√	
[12]		√	√			√			√							√
[13]		√			√			√	√		√					
[14]	√		√			not mentioned			√			√				
[15]		√	√			√			√		√					
[16]		√	√					√		√		√				
[17]		√		√		√				√		√		√		
[18]	√			√		√				√		√				
[19]	√			√		not mentioned				√		√				
[20]		√		√				√	√	√		√				
[21]		√	√			√			√		√					
[22]		√	√			√			√	√		√				
[23]	√		√			√			√		√					
[24]		√	√					√		√		√				
[25]	√		√													√
[26]	√		√			√				√		√				
[27]	√		√					√		√		√				
[28]	√		√				√		√		√	√		√		
[29]		√		√		√			√	√		√				
[30]	√		√					√			√			√		
[31]		√	√			not mentioned			√					√		
[32]	√				√	√			√	√		√				

[33]	√		√	√		√	√		√	
[34]	√		√			√	√	√		√
[35]	√		√		√		√	√		√
[36]		√	√			√	√			√
[37]		√		√		√				√
[38]	√		√		not mentioned				√	√

In wearable-oriented electrode arrangement, Ryu *et al.* adopted a glasses-based electrode positioning scheme that adapts to the structure of the glasses frame^[13]. Heo *et al.* designed a wearable EOG system in the form of a headband, placing four electrodes on the forehead with vertical channel electrodes arranged vertically and horizontal channel electrodes arranged horizontally, reducing the number of electrodes by sharing a common positive electrode^[32].

Acquisition hardware devices

EOG signals are non-periodic, with a frequency range typically between 0 to 50 Hz^[30,35,39-41]. According to the Nyquist criterion, the sampling frequency must be at least twice the maximum frequency component of the signal to avoid aliasing. Moreover, EOG signals are susceptible to interference from muscle activity and movement noise, necessitating amplification and filtering. For interactive hardware devices, user safety and comfort are also important considerations.

In general, EOG acquisition hardware can be categorized into custom-designed devices and commercial devices. In this review, 19 studies opted to develop custom-designed devices tailored and optimized to meet specific research requirements and experimental scenarios. In contrast, other studies utilized commercial devices such as NuAmps(Compumedics Neuroscan, Inc., Abbotsford, Australia), NeuroSky MindWave Mobile Headset (NeuroSky, San Jose, CA, USA), Neuroscan (Neuroscan Ltd., TX, USA), Enobio-32 EEG system (Neuroelectrics, Barcelona, Spain), BIOPAC MP150 (BIOPAC Systems, Goleta, CA, USA), EOG Mindo devices(Mindo, National Chiao Tung University Brain Research Center, Taiwan, China), gUSBamp amplifier (g.tec Medical Engineering GmbH, Austria) and the ActiveTwo biopotential recording system(Biosemi, Amsterdam, Netherlands). Self-designed EOG acquisition hardware involves systems carefully designed by researchers to align with their research goals. Many researchers integrated circuits into wearable devices to achieve greater comfort and portability. For instance, Perez Reynoso *et al.* designed an analog processing stage supplemented with a digital filtering module^[4]. The designed acquisition system was embedded in portable glasses with a PCB board, ensuring user safety and comfort. Milanizadeh *et al.* employed the ADS1299 biopotential amplifier to record EOG signals, utilizing upgraded safety glasses as the electrode support for easy fixation on users' faces^[23]. Ianez *et al.* integrated all electronic and mechanical components into a pair of EOG glasses, aiming for an aesthetically appealing design^[27]. Barea *et al.* used a commercially available model (Vuzix Wrap 230, Vuzix, Rochester, NY, USA) of glasses with head-mounted displays capable of fixing dry electrodes for EOG signal acquisition and serving as a medium for projecting user interfaces^[28]. Heo *et al.* designed a wearable EOG system based on a headband, placing Ag/AgCl electrodes inside the headband^[32]. Lin *et al.* proposed a wearable HCI-based system that can detect ten types of eye movements^[34]. The system integrates the EOG signal acquisition device and the foam type acquisition dry electrodes into the glasses.

Table 2. Types of Electrode Arrangements in EOG-HCI Research

Electrode layout types	Rationale for adoption	Advantages	Disadvantages
Standard Electrode Arrangement [4,6,8,9,11,12,14,16,27,28,30,31,34-36,38]	The most commonly used electrode arrangement in the EOG field, validated through extensive research, ensures consistency and comparability across different studies, facilitating data sharing and result comparison among researchers	Signal Reliability and Stability	High number of required electrodes
Task-Oriented Electrode Arrangement [5,7,10,17-20,29,33,37]	Research adopting this type of electrode arrangement focuses on exploring interaction tasks and interaction modalities. It allows for the design of the minimum number and configuration of electrodes based on the task type	High specificity, low signal processing difficulty	Poor scalability, requires higher rationality in interaction task design
Wearable-Oriented Electrode Arrangement [13,32]	The application of this type of research aims to implement research outcomes in practical settings, with a strong emphasis on user comfort. Therefore, compromises need to be made in signal quality to achieve higher levels of comfort	User friendly High generality	High cost for processing signal

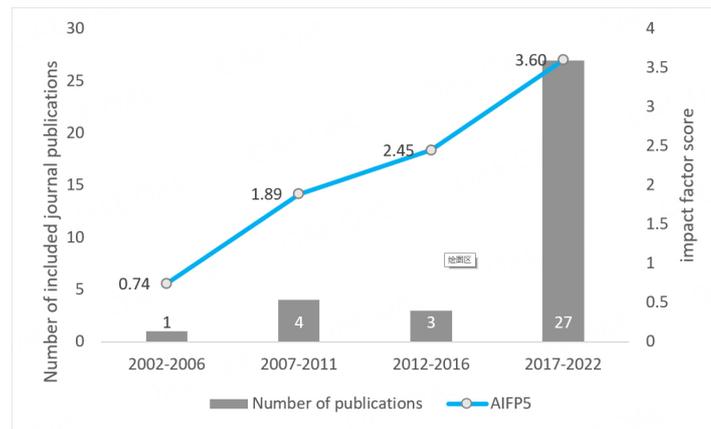


Figure 3. The number of publications and AIFP5 per 5-year period.

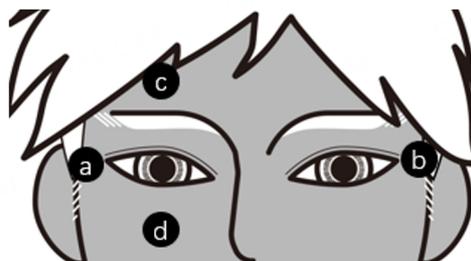


Figure 4. Standard Electrode Arrangement for EOG.

To assess hardware performance, Usakli *et al.* used Common-Mode Rejection Ratio (CMRR) during testing to measure the circuit's suppression capability of common-mode signals^[38]. CMRR is a parameter used to evaluate the ability of a circuit or signal processing system to suppress common-mode signals. The calculation of CMRR is typically represented in decibels (dB) and is defined as follows:

$$CMRR = 20 * \log_{10} \left(\frac{V_c}{V_{cm}} \right) \tag{2}$$

Where V_c represents the amplitude of the differential mode signal, and V_{cm} represents the amplitude of the common-mode signal. A higher value of CMRR indicates a stronger noise suppression capability of the system against common-mode signals.

In eye movement signal acquisition devices, a high CMRR helps reduce interference from power line frequencies and motion artifacts, preserving signal accuracy and stability, thus providing more reliable input for subsequent signal processing and interactive tasks.

Another commonly used hardware evaluation metric is the Signal-to-Noise Ratio (SNR), which is used to measure the relative strength or clarity between the measured signal and the noise in the system. It reflects the accuracy and reliability of the signal obtained by the EOG acquisition system. The calculation of SNR is typically represented in decibels (dB) and is defined as follows:

$$SNR = 10 * \log_{10}(P_{signal}/P_{noise}) \quad (3)$$

Where P_{signal} and P_{noise} are the power of the signal and noise, respectively. Signal and noise power are calculated as the square of the amplitudes in the frequency domain after performing Fourier transform. According to Pleshkov *et al.*, based on the research by Lopez *et al.*^[42], the effective frequency range of EOG signals is within [0, 30] Hz, while the bandwidth occupied by noise signals falls within [30, 122.5]^[43]. In addition to improving hardware design and implementing appropriate filters, Milanizadeh *et al.* proposed that the SNR can be improved by subtracting the signals of electrode pairs (up/down and left/right) from each other in a standard electrode arrangement^[23].

Self-designed devices typically consist of front-end amplifiers, hardware filters, second-stage amplifiers, analog-to-digital converters, and microprocessors, complemented by Bluetooth and battery modules. The system configuration and signal quality requirements determine the different settings for amplifier gain, filter cutoff frequencies, and sampling frequency.

In the context of interactive design, factors such as user comfort, safety, and device cost are crucial considerations for researchers. Some studies have used dry or semi-dry electrodes^[26] to improve comfort. Furthermore, wet electrodes, due to the use of conductive gel, initially provide signal quality superior to that obtained by dry electrodes within a short time. However, wet electrodes also face the issue of signal quality degradation as the conductive gel gradually dries out. In contrast, dry electrodes, which do not require the use of conductive gel, yield more suitable for long-term signal acquisition scenarios. Additionally, wearable designs consider human ergonomics^[44], offering users greater freedom, comfort, and safety. System safety is also a priority, adhering to International Electrotechnical Commission standards^[45] and incorporating current-limiting measures and protective components to ensure user and device safety.

EYE MOVEMENT INFORMATION PROCESSING AND CLASSIFICATION BASED ON EOG SIGNAL

In the field of EOG-HCI research, the recognition and classification of EOG signals are of great significance. It follows the steps of bio-signal processing, which include preprocessing of raw data, feature engineering, and algorithms.

Preprocessing

Before achieving an efficient and reliable EOG-HCI system, preprocessing steps are crucial. The main goals of preprocessing are to remove noise and baseline drift from the EOG signals, enabling more accurate and efficient feature extraction and interaction command recognition.

The EOG signal bandwidth is 0-50 Hz. Within this frequency range, EOG signals are often affected by various forms of interference, such as 50/60Hz power line noise and muscle activity^[46]. Additionally, baseline drift is another significant factor that affects the quality of EOG signals. Baseline drift, unrelated to eye movements, can cause large amplitude offsets in the signals, leading to the loss of characteristic features or waveforms associated with normal eye movements^[47].

During the hardware acquisition stage, hardware-based filtering is commonly employed for signal preprocessing. However, further preprocessing on the computer is still necessary to reduce noise and eliminate baseline drift, preparing the signals for subsequent signal processing tasks^[48].

To address noise and baseline issues in EOG signals, the most direct approach is to use bandpass filters for signal filtering. In the literature review, A total of 14 studies directly applied bandpass filters to the signals, with lower cutoff frequencies ranging from 0.1 to 0.5 Hz and upper cutoff frequencies ranging from 10 to 62.5 Hz. After passing through the bandpass filter, the signals were regarded as the signal with noise and baseline removed. To further ensure the removal of 50/60Hz power line noise, additional 50/60 Hz notch filters were used on top of the bandpass filters in studies^[12,15]. Additionally, moving average filters and median filters were commonly used for further signal smoothing^[16,20,24,27,34].

Given the non-stationary nature of EOG signals, wavelet transform has also become a preferred choice for many researchers. López *et al.* used thresholding to select appropriate thresholds for wavelet transform coefficients to remove or attenuate low-weight spectral components^[35]. They tested different mother wavelets and thresholding rules and found that using bior5.5 (biorthogonal) wavelet with the maximum-minimum thresholding method yielded the best signal-to-noise ratio. Choudhari *et al.* used the least squares method to fit and remove baseline effects, followed by denoising the signals using the db-10 mother wavelet and a fourth-order Butterworth filter with a frequency range of 1 to 10 Hz^[7]. In^[30], Perez-Reynoso *et al.* proposed a digital notch filter using Fast Fourier Transform(FFT) to eliminate unknown noise frequencies, achieving an undisturbed noise-free EOG signal by analyzing the EOG spectrum in real-time and selecting cutoff frequencies.

Ryu *et al.* introduced a novel method for removing baseline drift and noise from differential EOG signals, termed the Differential of the Sine-Based Fitting Curve (DOSbFC) method^[13]. This method involved calculating the DOSbFC signal by computing the voltage values in increments of 1° of eye gaze direction and the time required to stop eye movement after eye movement completion, using a formula to select and average the most similar potentials within a sliding window. Lee *et al.* used a linear function to fit and remove baseline drift and combined the amplitude-increasing waveform components in the horizontal and vertical channels with the EOG's maximum amplitude threshold to eliminate muscle artifacts^[24].

In addition to noise and baseline, blink signals are also considered artifacts in some interactive tasks and need to be removed during preprocessing. Chang *et al.* detected blinks based on a blink threshold and removed corresponding values^[36]. They also calculated a mean value for baseline every 100 ms, using linear interpolation with adjacent signals to fill in the missing data due to blink removal and performed linear regression to eliminate low-frequency drift.

Effective and rational preprocessing methods will significantly improve the performance and user experience of eye-movement-based interaction systems, promoting their widespread application in practical scenarios.

Features and algorithms

In HCI research, algorithms serve the interaction tasks, which often require real-time performance. Before applying algorithms, feature extraction is usually conducted, involving the extraction and selection of appropriate features, to enhance the performance and effectiveness of the algorithms. The quality of feature extraction often determines the final performance of the algorithms.

We have ranked the top four algorithms used in the 35 articles and conducted statistics on the features employed in the literature using these algorithms. The results are presented in [Table 3](#).

In the field of EOG signal analysis, feature extraction is a critical step, and common methods can be categorized into two types: features extracted from preprocessed clean signals and features in the form of first-order differentials. The original EOG signals typically exhibit distinct temporal patterns, directly reflecting waveform variations, which make them perform excellently in fast real-time processing. However, some researchers believe that certain features are not easily distinguishable in the original signal but become more prominent in the first-order differential forms of the original signal. Therefore, they opt to apply first-order differentiation to enhance feature discrimination.

Regarding feature extraction methods, researchers have proposed various innovative algorithms. For instance, Lv *et al.* introduced a spatial domain feature extraction method based on the Independent Component Analysis (ICA) algorithm^[8]. They used ICA separation matrix to construct a set of ICA spatial filters, mapping the experimental signals to different eye movement direction spaces through linear projection to extract spatial domain features. Fang *et al.* focused on continuous eye-writing recognition tasks and obtained extended EOG features suitable for different sampling rates through preprocessing, downsampling, and feature concatenation, used for subsequent neural network classification^[31].

Commonly used algorithms for eye movement signal classification include thresholding, waveform matching, neural networks, and traditional machine learning classifiers. The choice of algorithm largely depends on the research objectives and data characteristics. In this regard, neural networks have shown excellent performance in the field of biosignal processing, especially deep neural networks. However, traditional machine learning retains a distinct advantage in terms of interpretability that neural networks cannot replace. To illustrate the impact of research objectives and data characteristics on algorithm selection, we have separately summarized neural networks and traditional machine learning classifiers.

The thresholding method is the simplest classification approach, where a series of corresponding thresholds are set for eye movement signal classification using different feature combinations. It is worth noting that the thresholding method requires prior calibration of thresholds using collected eye movement signals to improve the accuracy when applied in formal usage. In research^[21,35], thresholds were set based on empirical rules, which may influence the accuracy for different subjects or within the same subject.

Waveform matching is primarily applied in interactive tasks involving eye-writing recognition of numbers, letters, *etc.*. Among them, the dynamic threshold detection method proposed by Ding *et al.* segments the eye movement signal into stroke segments and then uses the Dynamic Time Warping (DTW) algorithm to calculate similarity for handwriting eye movement signal classification^[9]. Kim *et al.* proposed a method that

Table 3. Feature and algorithm

Features		Threshold detection	Waveform matching	Neural network	Traditional machine learning
Original signal(s)	Amplitude	9, 10, 13, 14, 15, 20, 21, 23, 28, 33	24, 34, 36		
	Gradient	21, 22, 29	24, 37	31	
	Average EOG amplitude		9		
	delta-delta			31	
	Wavelet coefficient	36			
First-order differential forms of the original signal(s)	duration	5, 6, 7, 17, 18, 19, 20, 27, 29, 32, 33, 34			17
	Energy	5, 6, 17			11
	Amplitude	6, 7, 16, 18, 19, 22, 26, 27, 29, 33, 34			17
	waveform pattern	6			
Others			4	4, 28, 30,	8

uses the gradient value of the subject's horizontal EOG signal waveform to determine the smooth eye movement direction^[37]. And by comparing the waveform gradient with the waveform gradient of the sound source surrounding the subject, it is a method to judge whether the subject pays attention to one of the multiple sound sources.

The Support Vector Machine (SVM) is the most commonly used classifier in machine learning. In addition to SVM, López *et al.* used AdaBoost for classification, which combines weak learners to form a strong classifier through linear combination^[11].

Moreover, classification methods based on neural networks have been gaining popularity. Perez Reynoso *et al.* used an optimized genetic algorithm to obtain a multilayer neural network model for modeling and reduced the cost of pretraining users by dynamically calibrating thresholds^[4]. Barea *et al.* employed the Extreme Learning Machine (ELM) algorithm to train a Radial Basis Function (RBF) network to determine if a valid saccade occurs, an algorithm that is simple, has small training errors, and rapidly responds to users' eye movements^[28].

In addition to the above-mentioned four types of algorithms, Barbara *et al.* proposed a novel dual-channel input linear regression model that considers features extracted from horizontal and vertical EOG signal components, mapping eye potential displacements to gaze angle displacements^[12]. Barea *et al.* based their EOG system model (BiDimensional Dipole Model EOG, BiDiM-EOG) on physiological and morphological data of EOG to separate saccades and slow eye movements and determine fixations^[25].

Feature extraction and algorithm design play a pivotal role in EOG-based human-computer interaction tasks. The selection of appropriate features and algorithms is critical for building efficient and reliable eye movement human-computer interaction systems. Researchers need to customize feature selection and algorithm design based on the specific task requirements and interaction effectiveness to provide high-quality user experience and interaction efficiency. For example, for simple four-directional physical control tasks, the thresholding method may be more suitable as it can quickly distinguish eye movement signals from non-eye movement signals, generating simple control commands. For complex virtual keyboard designs or alphanumeric writing tasks, waveform matching may be more applicable, as it can combine

waveforms of EOG signals generated by a series of eye movements to expand the number of executable commands.

From [Table 3](#), it can be observed that the majority of research in the EOG-HCI field still relies on algorithms related to thresholds. Most of these studies were conducted in controlled laboratory environments, and a significant portion of them employed calibration techniques during data collection to ensure data quality. This helps to reduce individual differences and minimize signal variations over time, thereby lowering the robustness requirements for recognition technology. Furthermore, the time-domain information of EOG signals is easily interpretable by human researchers. This allows for the development of algorithms that combine threshold features with human prior knowledge to tailor recognition algorithms for specific eye movement classification tasks with different requirements. The high flexibility of threshold detection methods makes them widely applicable in various EOG-HCI research scenarios.

However, threshold detection methods are sensitive to parameter changes, leading to poor robustness. Furthermore, the features derived from thresholds are often designed based on human observations, resulting in low-dimensional, concrete features that cannot capture some high-dimensional, abstract features of different eye movements that are difficult for humans to understand. Hence, threshold detection methods often encounter issues related to suboptimal precision and robustness. In fact, nearly every study utilizing threshold detection methods necessitates the implementation of calibration techniques during data acquisition. In summary, while threshold detection methods offer simplicity and versatility, they tend to lack the desired level of robustness.

Additionally, four studies utilized neural network technology to construct eye movement classification models. Study^[4] used a multilayer neural network (MNN) to model the EOG signal as a mathematical function and optimize genetic algorithms for automatic signal calibration to obtain the threshold of each subject's EOG signal. Barea *et al.*^[28] employed an Extreme Learning Machine (ELM) algorithm to train a radial basis function (RBF) network with an internal structure comprising 20 neural neurons to determine whether the user had performed a valid eye jump. Pérez-Reynoso *et al.*^[30] used a multilayer perceptron (MLP) to establish a neural network with four hidden nodes for eye movement classification. Fang *et al.*^[31] employed deep learning techniques (DL) to construct a Deep Neural Network-Hidden Markov Model (DNN-HMM). This network model comprises four hidden layers, with 132 units in the input layer, 200 units in the first hidden layer, 100 units in the second hidden layer, and the number of units in the output layer equal to the number of HMM states. The ultimate output of the network represents the probability distribution of eye movement classifications.

It can be observed that DL techniques have not received sufficient attention in the EOG-HCI field. On one hand, based on the studies surveyed in this review, researchers in this field have placed more emphasis on application and have not explored EOG recognition algorithms extensively. On the other hand, there is currently no openly available EOG dataset tailored for HCI, making it challenging to train high-performance DL networks based solely on the samples collected from various studies. Chang *et al.*^[49] used a database of ten Arabic digit EOG eye-writing tasks collected in their previous work^[35], consisting of 540 eye-written digit samples for training their model.

In summary, current research is more inclined to explore the feasibility of using EOG for HCI, and threshold-based methods, due to their high generality, have become a popular recognition technology for exploring EOG-HCI feasibility. However, once the feasibility of EOG-HCI is established, higher demands will be placed on recognition technology in terms of robustness, classification precision, the number of

classes, *etc.* At that point, recognition technology developed based on deep learning, for example, will be developed to meet these requirements.

EVALUATION METRICS IN EOG-HCI RESEARCH

In the 35 literature papers, a total of 21 experimental evaluation metrics were identified and summarized, as shown in [Table 4](#). The summary of these evaluation metrics helps to understand the current research focus of researchers in the EOG-HCI field.

Based on the metrics in the literature, we categorized them into three types: system performance-related (SPR) metrics, interaction efficiency-related (IEF) metrics, and user-related (UR) metrics. SPR metrics focus on evaluating the usability of the EOG-HCI system, with accuracy (ACC) being the most commonly used evaluation metric. IEF metrics are used to assess the efficiency of the system's use, with information transfer rate (ITR) and task completion time being the most commonly used metrics. This indicates that many studies not only ensure the availability of the proposed interaction system but also emphasize its usability. Finally, some studies emphasize the usability of the proposed system from the perspective of users and introduce UR metrics in the research validation process, such as user proficiency^[19], game scores^[11], user experience^[25,33], *etc.*

Overall, system performance-related metrics were used 36 times, interaction efficiency-related metrics were used 36 times, and user-related metrics were used 18 times. Researchers are more concerned about system engineering performance and usage efficiency, while a smaller number of studies also emphasize the user perspective.

INTERACTION DESIGN

Classification of research results in EOG-HCI

For clarity, we define four concepts within the scope of this review:

1. Interaction Application: A program or system with interactive functionality and a complete interactive loop, designed to meet user needs.
2. Interaction Task: Tasks that need to be accomplished by users via interaction to fulfill their requirements.
3. Control Command: The command translated by the interaction interface based on the user's operational intent, used to complete the interaction task.
4. Interaction Modality: The mapping between trigger (specifically referring to eye movement in this review study) emitted by users' interaction intention and control commands. In the field of EOG-HCI (Electrooculography-based Human-Computer Interaction), the interaction application represents the final presentation form of research. Users achieve their goals of using the interaction application by completing interaction tasks, where the type of interaction task determines the interaction modality. The interaction modality significantly influences the engineering solutions for EOG signal acquisition, processing, and algorithm design. [Figure 5](#) classifies all interaction applications over the 21 years based on their interaction tasks, dividing them into eye movement control and eye movement information input categories, and further subdividing them according to the interaction modality. As some literature describes multiple types of interaction applications, this review only selects more typical or distinctive applications for statistical analysis. This summarization provides a more comprehensive understanding of the entire EOG-HCI domain.

Table 4. List of Evaluation Metrics Used in EOG-HCI Research

Types	Evaluation metrics	Number of applications
System performance-related	Accuracy	22 ^[5-9,13,17-21,22,24,27,31-38]
	Recall	4 ^[21,24,27,36]
	Precision	3 ^[14,29,36]
	F1 score	3 ^[21,24,36]
	False positive rate	2 ^[5,20]
	Error rate	2 ^[26,27]
	Interaction efficiency-related	Input speed
Response time		5 ^[5,17,19,20,29]
Information transfer rate		9 ^[6,7,10,13,17,20,33,35,37]
Total time		11 ^[4,8,12,15-18,20,23,26,30]
Average time per command		2 ^[8,27]
Total number of operations		5 ^[15,18,19,28,33]
User-related		Number of interaction failures
	Number of successful interactions	2 ^[11,28]
	False operation rate	3 ^[18-20]
	Number of interaction errors	4 ^[11,27,32,33]
	User control level	4 ^[4,11,19,25]
	User experience	3 ^[15,25,33]

*Error Rate: the proportion of all test samples that are mismarked; its formula is as follows: $ErrorRate = (FP+FN)/(TP+TN+FP+FN)$; Input Speed: the number of characters entered per minute in writing system; Response time (RT): the time used to generate a command; Information transfer rate (ITR): bits of information transferred per minute; Total time: The total time cost to complete a task, for example, typing a text in a keyboard interaction, steering a wheelchair to complete a route; False operation rate: the number of false output commands per minute when the subject has no intention of selecting any buttons; User Control Level: reflecting in the learning cost of users and the improvement of interaction efficiency with the increase of the number of uses; User Experience: including the use of scales to reflect user satisfaction and compare the comfort of different GUI guidance strategies

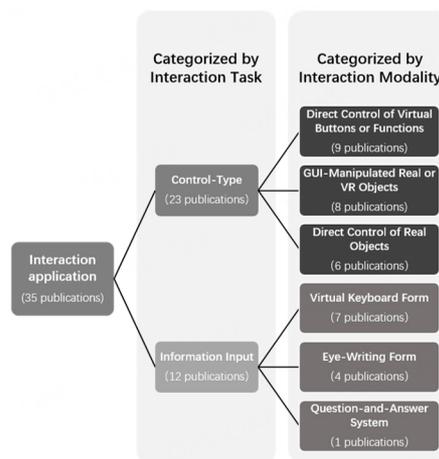


Figure 5. Classification of EOG-HCI Publications (2002-2022).

Control-type interaction applications

Control-type interaction applications in eye movement control allow users to control virtual or real-world objects to achieve specific goals. In such interaction tasks, the interaction itself can be the goal, and the results brought about by the user's actions are usually deterministic. Users can quickly obtain various forms of visual and sensory feedback.

Based on different interaction modalities, control-type interaction applications can be divided into three categories:

Direct control of virtual buttons or functions:

In this form of interaction, users control virtual software, desktop, or mobile applications through a GUI interface without involving real objects. For example, Lv *et al.* developed an electronic book reading controller using a MATLAB/JAVA joint platform, where users can achieve six basic control commands, such as page flipping, pause/resume reading, and opening/closing the book, through six predefined eye movement patterns^[8]. Similarly, Ryu *et al.* designed desktop and mobile applications that utilize different eye movement patterns to act as an auxiliary controller for Windows applications and perform seven functions for multimedia control on mobile devices^[13]. These applications provide users with a convenient way to control virtual software and applications rapidly through eye movements.

GUI-manipulated real or VR objects:

In this form of interaction, users typically select control commands for a target object in a GUI which serves as a control panel, before sending the control commands to the object for execution. For instance, Li *et al.* designed a GUI interface that included wheelchair switch converter buttons^[5]. Users activate the EOG switch by blinking synchronously with the flashing, thereby sending control commands to navigate the powered wheelchair.

Direct control of real objects:

In this form of interaction, researchers usually directly send user-generated eye movement commands to a controller, which implements high-level control and generates specific operation commands for interactive objects, primarily wheelchairs and robotic arms. In Barea *et al.*'s research^[25], recorded eye gaze information is processed by an onboard computer to formulate the wheelchair's control strategy and generate linear and angular velocity commands. These commands are then transformed into angular velocities for each wheel based on the wheelchair's kinematic model and sent to the low-level control module, achieving two closed-loop velocity controls. Consequently, users can navigate the wheelchair's movement and steering through eye movement commands, enabling autonomous navigation.

On the other hand, Ubeda *et al.*'s research^[26] focused on designing a robotic arm application. This application converts users' eye gaze information into control commands for the robotic arm's end effector, replacing human arms to operate various tools. This interaction modality allows users to perform complex robotic arm operations through simple eye movement commands, improving their self-care abilities and work efficiency.

Information input-type interaction applications

Information input-type interaction applications differ from control-type applications as they focus on using eye movement information as an input method, with more open-ended interaction purposes. According to the interaction modality, they can be divided into three categories: virtual keyboard form, eye-writing form, and question-and-answer system. These applications convert eye movement information into other signal forms and cannot complete interaction tasks independently. Compared to control-type interaction applications, they are more oriented towards using the interaction behavior as a means, with open-ended goals.

Virtual keyboard form:

In this form of eye movement interaction, virtual keyboards with different character layouts are used, allowing users to select characters they need through eye movements to complete input. To simplify the

number of eye movements required for character selection, researchers explore different stimulation modes. One mode involves selecting a button by blinking users' eyes in synchrony with the flash of the button^[17]. Another mode is the row/column paradigm activation mode^[29], where characters are arranged in rows and columns. Users efficiently determine the character's position by selecting the row and column through eye blinking. This method streamlines the user's selection process, making eye movement commands more efficient.

Eye-writing form:

In information input-type eye movement interaction applications, the eye-writing form aims to use different eye movement patterns to write symbols, such as characters. Researchers combine different eye movement types into different eye movement patterns, with each pattern corresponding to a stroke. By calculating the similarity between the user's input eye movement pattern and predefined templates, the system outputs the character the user has written. For example, Ding *et al.* designed a stroke-based Chinese eye-writing system that maps five basic strokes to predefined eye movement patterns, enabling continuous writing of Chinese characters^[9]. Similarly, Lee *et al.* predefined 12 eye movement templates for outputting 26 lowercase letters and 3 control commands^[24]. These systems determine the most similar character as the output result by comparing the differences between the input eye movement trajectory and the character templates. This eye-writing form of interaction system provides a novel and efficient way of inputting information.

Question-and-answer system:

Kim *et al.* developed a novel EOG-based Augmentative and Alternative Communication (AAC) system aimed at helping users engage in binary yes/no communication^[37]. The interaction system plays two rotating sound stimuli to the user and asks them to focus on the specified sound source. During this process, horizontal EOG is recorded in response to the involuntary auditory ocular reflex, which answers the designed questions. This system provides a means for users with communication barriers to engage in simple communication through eye movements, bringing convenience and improvement to their lives.

Summary

We classified the interaction applications into six subcategories within two main categories based on the categories presented in the previous section and proposed two metrics to compare the related research of the two interaction application types:

- Degree of Interaction Freedom (DOF): Reflecting the availability of the application, i.e., the number of available predefined commands. This metric is based on the number of available commands.
- Interaction Efficiency (EEF): Reflecting the usability of the application in terms of how quickly users can complete commands. This metric is based on the maximum number of eye movements required to complete a command.

There is a correlation between DOF (Degree of Freedom) and EEF (Efficiency). In essence, as DOF increases for the same application, the efficiency of interaction tends to decrease. This correlation is referred to as the Interaction Benefit score, represented by C . The formula for calculating C is as follows:

$$C = \frac{\log_2 DOF}{EEF} \quad (4)$$

Using $\log_2 DOF$ instead of DOF in the calculation helps mitigate the sensitivity of C to changes in DOF , considering that increasing DOF is relatively easier. All the information is summarized in [Table 5](#).

Research on control-type interaction applications primarily focuses on user-directed interaction tasks and designs corresponding interaction modalities based on the purpose of using the interaction application. This type of interaction modality covers a wide range of application scenarios, including appliance control in home IoT settings^[14,29], machine control^[23,25], navigation^[4,5,26], web browsing^[16], e-book reading^[21], eye-tracking games^[22], and various other aspects. The number of research papers in this category is nearly twice as much as that of information input-type interaction applications.

Among the existing control-type interaction applications, 17 out of 23 papers used Graphical User Interface (GUI) as an intermediate medium. GUIs offer high extensibility, allowing the functionality of control-type applications to be easily expanded by adding more interfaces, placing new functions in separate sub-interfaces, or adopting layouts similar to P300 speller to integrate multiple functions on the same interface, thus enhancing the freedom and extensibility of the interaction interface. Therefore, in theory, there is no upper limit to the number of functions in control-type interaction applications.

However, it is essential to consider user experience. As the number of functions increases, users' cognitive load may also increase, leading to higher error rates and reduced operational efficiency. This aspect is particularly critical for the target users of EOG-HCI, as a significant portion of them may have mobility impairments, such as Amyotrophic lateral sclerosis (ALS) patients. Therefore, research on control-type interaction applications needs to pay special attention to a reasonable analysis of interaction tasks to ensure that users can easily understand and operate the interaction interface.

Information input-type applications have more explicit interaction task purposes, namely, user communication with the outside world. However, since human communication purposes are challenging to predict, information input-type applications require higher freedom to meet users' varied needs. From [Table 5](#), it can be observed that among the 35 EOG-HCI literature samples, the top five interaction applications designed in literature^[6,12,16,35,38] have the highest DOF . Among them, applications in literature^[16] is a comprehensive human-computer interaction application that includes both pages browsing through eye movements and information input through a virtual keyboard. Although categorized as a control-type application, it overall encompasses the functionalities of an information input-type application. The interactive applications mentioned in all these five studies have a virtual keyboard, which indicates a high requirement of freedom for information input type applications.

In research of information input-type applications, ensuring efficiency while maintaining degrees of freedom is essential. Therefore, investigations into information input-type interaction applications may encounter some challenges, but they have the potential to make significant contributions to the field.

It is worth noting that the evaluation of interaction design should be conducted within a unified context. This review covers a wide variety of application types, and each interaction design is tailored to different users and scenarios. Therefore, the evaluation metric C proposed in [Table 5](#) can only serve as a reference point. Performance assessment needs to be carried out specifically for the selected user groups and within the context of each specific application type.

USER-CENTERED EOG-HCI

Looking back at history, it is not difficult to observe that the development of most application-oriented

Table 5. Interactions Tasks and Interaction Performance

Reference	Interaction applications						Interaction performance		
	Control-type			Information input-type			DOF	EEF	C score
	Direct control of virtual buttons or functions	GUI-manipulated real or VR objects	Direct control of real objects	Virtual keyboard	Eye-writing	Question-and-answer system	Number of available commands	Maximum number of eye movements required to complete a command	
[12]				√			80	2	3.16
[35]				√			73	5	1.24
[38]				√			54	17	0.34
[16]	√						50	15	0.38
[6]				√			42	5	1.08
[17]				√			40	6	0.89
[18]	√						32	6	0.83
[33]		√					23	2	2.26
[13]	√						18	2	2.08
[19]		√					14	2	3.58
[4]			√				12	1	1.90
[24]					√		12	5	0.72
[31]					√		12	3	1.19
[34]				√			10	4	0.83
[22]	√						9	2	1.58
[29]		√					9	2	1.58
[36]					√		9	4	0.79
[20]		√					8	3	1.00
[28]	√						8	4	0.75
[30]			√				8	2	1.50
[9]					√		7	4	0.70
[10]			√				7	2	1.40
[32]				√			7	1	2.81
[8]	√						6	4	0.65
[15]		√					6	1	2.58
[11]	√						5	3	0.77
[27]	√						5	4	0.58
[7]			√				4	3	0.67
[14]		√					4	1	2.00
[23]		√					4	1	2.00
[25]			√				4	1	2.00
[26]		√					4	1	2.00
[21]	√						3	2	0.79
[5]		√					2	2	0.50
[37]						√	2	1	1.00

technologies requires a collaborative effort between function-oriented research and user-oriented research directions. For instance, during the 1920s and 1930s, pioneers dedicated to radio technology research focused on aspects such as receiver sensitivity, transmitter power, and channel separation. Meanwhile, another line of development ensured that users would enjoy radio broadcasts through features like easy-to-use dials and switches, appealing equipment design, engaging programs, and reliable broadcast schedules.

EOG-HCI research is an interdisciplinary field that blends the expertise of biomedical engineering and interaction design. Interaction design focuses on understanding the relationship between humans and machines, while biomedical engineering aims to implement this relationship as defined by interaction design. In this integrated cross-disciplinary research domain, researchers need to recognize the significance of biomedical engineering as the foundation and also possess the top-down research perspective required by HCI studies. Clarifying the interplay between these two aspects helps us comprehend how collaboration should be approached in user-centered EOG-HCI research. To this end, we propose a 5-level model for EOG-HCI research, as illustrated in [Figure 6](#). Additionally, leveraging the insights from the model and the information provided in [Table 4](#) and [Table 5](#), we shed light on the current research landscape from a user-centered standpoint.

In this model, different levels of research focus on different aspects of users, with higher-level research emphasizing more on users, resulting in a better user experience.

1. The first level focuses on the usefulness of research results, ensuring that research is carried out based on the real needs of target users.
2. The second level focuses on the availability of research results, achieving effective implementation of functionalities by utilizing research outcomes from the field of biomedical engineering, including hardware development and algorithm research, to bridge the gap between user needs and implementation.
3. The third level focuses on the usability of research results, providing low cognitive load and efficient interaction methods to help users accomplish their tasks effectively.
4. The fourth level focuses on the intelligence of research results, understanding user interaction intentions, predicting user behavior outcomes, and providing innovative capabilities and additional information. As defined by Fengliu and Yong Shi in the Liufeng-Shiyong architecture^[50], one of the characteristics of an intelligent system is the ability to generate demands based on external data, information, and knowledge and utilize the acquired knowledge to achieve innovation, including association, creation, guessing, and discovering patterns. The application of this ability leads to the formation of new self-acquired knowledge.
5. The fifth level focuses on the enjoyment and amusement of research results, creating a pleasant overall experience through interactive design and adding emotional value. This level of research primarily relies on interactive design methods, taking changes in user emotions towards the interactive interface with the attributes of the previous levels as the research entry point, to add emotional value to primary functional interactions.

In EOG-HCI research, the five levels should be followed progressively without reversal. A research framework must examine each level, ensuring the fulfillment of the preceding ones. The foundation is set at the usefulness level, which all research must satisfy. The availability level requires functional implementation, relying on relevant technologies from biomedical engineering. The usability and intelligence levels demand research from both user and interaction interface perspectives, utilizing techniques from interaction design and biomedical engineering. The enjoyment and amusement interactive user experience level emphasizes design's role in adding emotional value to the interaction behavior.

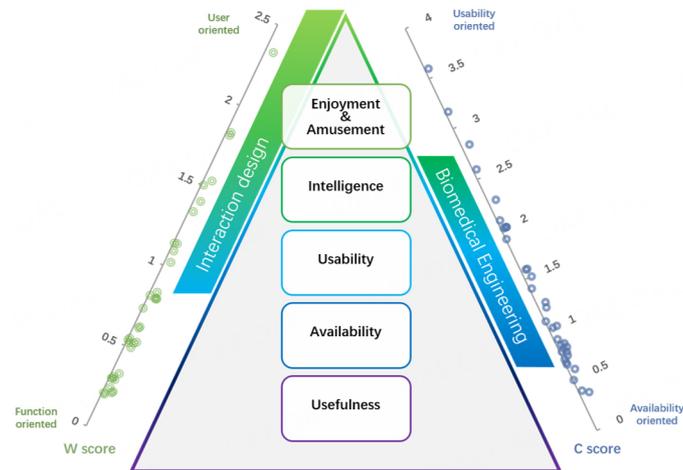


Figure 6. 5-level model for EOG-HCI research.

In Figure 6, we use scatter plots to represent C score and W score. C score reflects the benefits users gain from interaction applications, measured by the ratio between functionalities and efforts. W score represents the research's emphasis on user experience, quantifying the importance given to user experience in the literature. The W score is calculated as follows:

$$W_n = M_{SPR_n} * \alpha_{SPR} + M_{IER_n} * \alpha_{IER} + M_{UR_n} * \alpha_{SPR} \tag{5}$$

Where W_n is the W score of the n th research, and M_{SPR} , M_{IER} , M_{UR} are the quantities of SPR, IER, and UR category evaluation metrics used in the research. Based on the correlation between the three categories of evaluation metrics and user experience, we propose three weight coefficients: $\alpha_{SPR} = 0.2$, $\alpha_{IEF} = 0.3$, $\alpha_{UR} = 0.5$.

While the evaluation criteria for W score and C score may not directly correspond across different levels of the model, the meanings of both indicators align with the model's trend. Specifically, at the base level, they reflect a more function-oriented research perspective, whereas at the top level, they reflect a more user-oriented research perspective. Combining these two indicators with the model allows us to visualize insights in the domain.

FUTURE RESEARCH AND TRENDS

As an interdisciplinary field, approaches of EOG-HCI research can be broadly categorized into two main approaches: (1) Exploration of the feasibility of EOG-HCI: Researchers taking this approach primarily focus on exploring the feasibility of applying EOG technology in HCI. Their emphasis is on identifying suitable HCI tasks/applications for EOG-based eye movement classification technology; (2) Task-driven EOG classification technology development: In this approach, researchers begin with specific HCI tasks and then develop EOG-based eye movement classification technology tailored to these designed HCI tasks. The emphasis here is on meeting the requirements of specific HCI tasks.

In general, the exploration of feasibility offers the baseline for the entire EOG-HCI research, while EOG eye movement classification technology development sets the upper limit. A complete EOG-HCI research is a spiral upward iterative process. It should begin by using established EOG recognition technologies, such as

threshold detection methods, to design reasonable HCI tasks. Subsequently, as insights from HCI task designs emerge, new demands are placed on EOG classification technology, leading to technical breakthroughs. Eventually, suitable EOG recognition algorithms are developed for these well-suited HCI tasks, completing a cycle of EOG-HCI research iteration.

Threshold detection methods, as an easily usable and highly versatile EOG eye movement classification technology, are friendly for research aimed at exploring the feasibility of EOG-HCI. In fact, studies using threshold detection methods for eye movement classification are far more abundant than other EOG-based technologies. This fact indirectly indicates that the majority of research in this field is currently focused on identifying suitable HCI tasks/applications for eye movement classification technology based on EOG signals.

DL is an excellent technical approach that has not been extensively utilized in EOG-HCI. One possible reason is the limited availability of open EOG datasets tailored for HCI. Relying solely on data collected in controlled laboratory environments makes it challenging to train high-performance EOG eye movement classification models using DL technology. Additionally, current research in the field appears to be primarily focused on exploring the feasibility of EOG-HCI. In other fields, such as EEG, the interesting idea is to combine computational vision and BCI^[51]. The extra BCI (EEG) signals are shown to improve the performance of a trained recognizer. The same combination could be considered the other way around: it would be a good idea that an EOG-based BCI system not only looks at the eye movements, but also at the image as seen by the user.

In contrast to the high versatility, ease of use, but lower accuracy and robustness of threshold-based methods, DL methods often yield highly accurate and robust classification models but with relatively lower versatility and higher development costs (e.g., data and time cost). Therefore, in research aimed at exploring the feasibility of EOG-HCI, threshold-based methods are preferred.

As EOG-HCI research advances and concepts like the metaverse gain acceptance, especially with the development of head-mounted displays such as AR and VR, there is a growing demand to incorporate EOG-HCI into these emerging research and application domains. This demand calls for more precise eye movement classification algorithms for extensive eye movements, eliminating the need for frequent recalibration, and requiring robust EOG eye movement multiclass technology. However, threshold detection methods, due to their sensitivity to signal variations and heavy reliance on prior knowledge, are ill-suited to meet the requirements of these trends for multiclass eye movement recognition tasks. As one advantage of DL, the capability of capturing high-dimensional features that are challenging for humans to comprehend, is more likely to detect the sample-dependent and sample-independent features under various eye movement activities. This makes DL a promising approach for developing more mature EOG multiclass classification algorithms.

In fact, in other eye movement detection domains, DL has already shown remarkable success. Tang *et al.* proposed the concept of representing biosignals as time-dependent graphs and presented GraphS4mer, a general graph neural network (GNN) architecture that improves performance on biosignal classification tasks by modeling spatiotemporal dependencies in biosignals^[52]. Gupta *et al.* applied DL to capture the mechanism of visual asymmetry through psychological experiments^[53]. In the future, researchers may be able to use EOG-HCI in psychological and neurological research to contribute to the monitoring and treatment of some visual diseases.

To support the application of DL technology in the EOG-HCI field, one challenge is the need to establish open EOG datasets tailored for HCI. The eye movements of individuals and the EOG signals are constantly influenced by a variety of physiological and psychological factors. Currently, the correlation between the variations in individual resting-state EOG signals and their physiological and psychological factors remains largely unexplored. These unexplored factors, to some extent, impact the development of DL in the field of EOG-HCI. Therefore, in order to harness deep learning more effectively as a powerful tool for developing EOG-HCI, we should draw insights from mature fields such as sleep monitoring, as exemplified by the template provided in reference^[54], for creating multimodal databases tailored for EOG-HCI.

High-quality EOG eye-tracking databases contain a wealth of information on eye movements across different tasks, as well as various physiological, psychological, and potential phenotypic information about the subjects. Leveraging DL's high adaptability and automated feature learning capabilities, as well as its ability to handle complex data forms, we can process eye movement data, EOG signals, and other multimodal signals together. This can result in highly accurate, generalized, and robust EOG-based eye-tracking analysis models. Ultimately, this will result in a technological breakthrough in EOG-HCI, liberating researchers in the field from the constraints of laboratory environments and allowing them to design more complex HCI tasks/applications without being limited to just a few simple but robust eye movements, such as blinking.

Furthermore, from the perspective of HCI task/application design, it can be seen that both the interaction benefit brought by the research's interaction design itself and the evaluation system used in the research indicate that there is still untapped potential in EOG-HCI research from the user's perspective. From the many experimental set-ups reviewed, it can be seen that there is great variety in the types of stimuli offered. However, there are some mismatches regarding the user experience. Some stimulus types are audio-based (such as the rotating sounds^[37]), while others are image-based (such as the target groups and target letters^[6]). These stimuli may obstruct or hinder the user's perception of natural sounds and images in their environment. Whether such a hindrance makes sense or not depends on the application type and the user's specific needs. When designing HCI tasks, it is essential not only to consider the technical feasibility of EOG-HCI but also to consider the user, equipment, environment, and context as a whole. Moreover, most of the research requires users to use additional eye movements to operate the EOG-HCI system, which introduces significant changes in the user's visual field. Vision is a crucial source of information for humans, and forcing users to shift their visual focus for interaction can be inconvenient and even risky. When the user's actions and attention cannot be separated in specific scenarios, eye movement-based operations can become frustrating. For instance, when using eye movements to control a wheelchair in an unfamiliar environment, users are likely to have difficulty controlling their movements due to unintended eye movements.

To strike a balance between user needs and technical requirements, user-centered design methods such as context mapping and user values mapping are essential. We anticipate a trend where HCI work for EOG-HCI will evolve into subcategories, each tailored to a specific application domain. In the future, with the increasing adoption of a top-down approach in EOG-HCI research, it is foreseeable that more user-centric evaluation metrics such as the Usability Metric for User Experience (UMUX) score^[55] and the System Usability Scale (SUS)^[56] will be used to assess research outcomes. As a result, the interaction benefits for end-users will improve, the integration of technology and interaction design will become more rational, and better user experiences will result from the outcomes of EOG-HCI research.

CONCLUSION

This review provides an overview of the research status in the EOG-HCI field from 2002 to 2022. The study reveals that the research vitality in this field has been continuously increasing. The research summarizes and categorizes various aspects, including hardware design, electrode arrangement, preprocessing methods, feature extraction, classification algorithms, and interaction design in the domain. Based on commonly used evaluation metrics, the research focus in the current stage is analyzed.

One notable finding is that the existing research lacks a user-centric and user-driven research design concept, which may lead some studies to deviate from the primary goal of HCI. Consequently, the outcomes of such research might be hampered by rejecting users. This presents both challenges and opportunities for researchers in this field. In response to this limitation, the paper proposes a 5-level model for EOG-HCI research. It is hoped that this model will assist future researchers in formulating research plans that align with their knowledge background and prioritize user perspectives.

DECLARATIONS

Acknowledgments

Special thanks to Zheng Z for helping with the revisions of the article.

Authors' contributions

Collaborated closely in the literature review, content writing, and revision of this review article: Tao L, Huang H

Contributed to the overall structure of the article, information extraction and summary of the research findings: Tao L

Played a significant role in data organization, information retrieval, and formatting adjustments for the article: Huang H

Provided supervision for this research and also contributed to the manuscript writing and revision: Feijs L, Hu J, Chen W, Chen C

Availability of data and materials

Not applicable.

Financial support and sponsorship

None.

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2023.

REFERENCES

1. Bachmann D, Weichert F, Rinkenauer G. Review of three-dimensional human-computer interaction with focus on the leap motion controller. *Sensors* 2018;18:2194. DOI PubMed PMC
2. DAVIS JR, SHACKEL B. Changes in the electro-oculogram potential level. *Br J Ophthalmol* 1960;44:606-18. DOI PubMed PMC
3. Mowrer OH, Ruch TC, Miller NE. The corneo-retinal potential difference as the basis of the galvanometric method of recording eye movements. *American Journal of Physiology-Legacy Content* 1935;114:423-8. DOI
4. Perez Reynoso FD, Niño Suarez PA, Aviles Sanchez OF, et al. A custom EOG-based hmi using neural network modeling to real-time for the trajectory tracking of a manipulator robot. *Front Neurobot* 2020;14:578834. DOI PubMed PMC
5. Li Y, He S, Huang Q, Gu Z, Yu ZL. A EOG-based switch and its application for “start/stop” control of a wheelchair. *Neurocomputing* 2018;275:1350-7. DOI
6. Saravanakumar D, Ramasubba Reddy M. A high performance asynchronous EOG speller system. *Biomed Signal Proces* 2020;59:101898. DOI
7. Choudhari AM, Porwal P, Jonnalagedda V, Mériaudeau F. An Electrooculography based human machine interface for wheelchair control. *Biocybern Biomed Eng* 2019;39:673-85. DOI
8. Lv Z, Wang Y, Zhang C, Gao X, Wu X. An ICA-based spatial filtering approach to saccadic EOG signal recognition. *Biomed Signal Proces* 2018;43:9-17. DOI
9. Ding X, Lv Z. Design and development of an EOG-based simplified Chinese eye-writing system. *Biomed Signal Proces* 2020;57:101767. DOI
10. Sharma K, Jain N, Pal PK. Detection of eye closing/opening from EOG and its application in robotic arm control. *Biocybern Biomed Eng* 2020;40:173-86. DOI
11. López A, Fernández M, Rodríguez H, Ferrero F, Postolache O. Development of an EOG-based system to control a serious game. *Measurement* 2018;127:481-8. DOI
12. Barbara N, Camilleri TA, Camilleri KP. EOG-based eye movement detection and gaze estimation for an asynchronous virtual keyboard. *Biomed Signal Proces Control* 2019;47:159-67. DOI
13. Ryu J, Lee M, Kim DH. EOG-based eye tracking protocol using baseline drift removal algorithm for long-term eye movement detection. *Expert Syst Appl* 2019;131:275-87. DOI
14. Deng LY, Hsu CL, Lin TC, Tuan JS, Chang SM. EOG-based human-computer interface system development. *Expert Syst Appl* 2010;37:3337-43. DOI
15. Postelnicu CC, Gîrbacia F, Talaba D. EOG-based visual navigation interface development. *Expert Syst Appl* 2012;39:10857-66. DOI
16. Lledó LD, Úbeda A, Iáñez E, Azorín J. Internet browsing application based on electrooculography for disabled people. *Expert Syst Appl* 2013;40:2640-8. DOI
17. He S, Li Y. A single-channel EOG-based speller. *IEEE Trans Neural Syst Rehabil Eng* 2017;25:1978-87. DOI
18. Xiao J, Qu J, Li Y. An electrooculogram-based interaction method and its music-on-demand application in a virtual reality environment. *IEEE Access* 2019;7:22059-70. DOI
19. Huang Q, He S, Wang Q, et al. An EOG-based human-machine interface for wheelchair control. *IEEE Trans Biomed Eng* 2018;65:2023-32. DOI
20. Zhang R, He S, Yang X, et al. An EOG-based human-machine interface to control a smart home environment for patients with severe spinal cord injuries. *IEEE Trans Biomed Eng* 2019;66:89-100. DOI
21. Ouyang R, Lv Z, Wu X, Zhang C, Gao X. Design and implementation of a reading auxiliary apparatus based on electrooculography. *IEEE Access* 2017;5:3841-7. DOI
22. Lin C-T, King J-T, Bharadwaj P, et al. EOG-Based eye movement classification and application on HCI baseball game. *IEEE Access* 2019;7:96166-76. DOI
23. Milanizadeh S, Safaie J. EOG-based HCI system for quadcopter navigation. *IEEE Trans Instrum Meas* 2020;69:8992-9. DOI
24. Lee KR, Chang WD, Kim S, Im CH. Real-time "eye-writing" recognition using electrooculogram. *IEEE Trans Neural Syst Rehabil Eng* 2017;25:37-48. DOI PubMed
25. Barea R, Boquete L, Mazo M, López E. System for assisted mobility using eye movements based on electrooculography. *IEEE Trans Neural Syst Rehabil Eng* 2002;10:209-18. DOI PubMed
26. Ubeda A, Iáñez E, Azorín JM. Wireless and portable EOG-based interface for assisting disabled people. *IEEE/ASME Trans Mechatron* 2011;16:870-3. DOI
27. Iáñez E, Azorín JM, Perez-Vidal C. Using eye movement to control a computer: a design for a lightweight electro-oculogram electrode array and computer interface. *PLoS One* 2013;8:e67099. DOI PubMed PMC
28. Barea R, Boquete L, Rodríguez-Ascariz JM, Ortega S, López E. Sensory system for implementing a human-computer interface based on electrooculography. *Sensors* 2011;11:310-28. DOI PubMed PMC
29. Laport F, Iglesia D, Dapena A, Castro PM, Vazquez-Araujo FJ. Proposals and comparisons from one-sensor EEG and EOG human-machine interfaces. *Sensors* 2021;21:2220. DOI PubMed PMC
30. Pérez-Reynoso FD, Rodríguez-Guerrero L, Salgado-Ramírez JC, Ortega-Palacios R. Human-machine interface: multiclass classification by machine learning on 1D EOG signals for the control of an omnidirectional robot. *Sensors* 2021;21:5882. DOI PubMed PMC

31. Fang F, Shinozaki T. Electrooculography-based continuous eye-writing recognition system for efficient assistive communication systems. *PLoS One* 2018;13:e0192684. DOI PubMed PMC
32. Heo J, Yoon H, Park KS. A novel wearable forehead EOG measurement system for human computer interfaces. *Sensors* 2017;17:1485. DOI PubMed PMC
33. Huang Q, Chen Y, Zhang Z, et al. An EOG-based wheelchair robotic arm system for assisting patients with severe spinal cord injuries. *J Neural Eng* 2019;16:026021. DOI
34. Lin CT, Jiang WL, Chen SF, Huang KC, Liao LD. Design of a wearable eye-movement detection system based on electrooculography signals and its experimental validation. *Biosensors* 2021;11:343. DOI PubMed PMC
35. López A, Ferrero F, Yangüela D, Álvarez C, Postolache O. Development of a computer writing system based on EOG. *Sensors* 2017;17:1505. DOI PubMed PMC
36. Chang WD, Cha HS, Kim DY, Kim SH, Im CH. Development of an electrooculogram-based eye-computer interface for communication of individuals with amyotrophic lateral sclerosis. *J Neuroeng Rehabil* 2017;14:89. DOI PubMed PMC
37. Kim DY, Han CH, Im CH. Development of an electrooculogram-based human-computer interface using involuntary eye movement by spatially rotating sound for communication of locked-in patients. *Sci Rep* 2018;8:9505. DOI PubMed PMC
38. Usakli AB, Gurkan S. Design of a novel efficient human-computer interface: an electrooculogram based virtual keyboard. *IEEE Trans Instrum Meas* 2010;59:2099-108. DOI
39. Enderle JD. Observations on pilot neurosensory control performance during saccadic eye movements. *Aviat Space Environ Med* 1988;59:309-13. PubMed
40. Enderle J. Eye Movements. In: Akay M, editor. Wiley Encyclopedia of Biomedical Engineering. Wiley; 2006.
41. Ülkütaş HÖ, Yıldız M. Computer based eye-writing system by using EOG. Medical Technologies National Conference (TIPTEKNO); 2015 October 1-4; Bodrum, Turkey. DOI
42. Lopez A, Ferrero FJ, Valledor M, Campo JC, Postolache O. A study on electrode placement in EOG systems for medical applications. 2016 IEEE International symposium on medical measurements and applications (MeMeA); 2016 May 1-5; Benevento, Italy. DOI
43. Pleshkov M, Zaitsev V, Starkov D, Demkin V, Kingma H, van de Berg R. Comparison of EOG and VOG obtained eye movements during horizontal head impulse testing. *Front Neurol* 2022;13:917413. DOI PubMed PMC
44. Champaty B, Jose J, Pal K, Thirugnanam A. Development of EOG based human machine interface control system for motorized wheelchair. AICERA/iCMMD: 2014 Annual International Conference on Emerging Research Areas: Magnetics, Machines and Drives; 2014 July 24-26; Kerala, India. DOI
45. International Organization for Standardization (ISO). Part 1-11: General requirements for basic safety and essential performance. Available from: http://www.iso.org/iso/catalogue_detail.htm?csnumber=65529[Last accessed on 26 Sep 2023].
46. Bulling A, Herter P, Wirz M, Tröster G. Automatic artefact compensation in EOG signals. Available from: https://www.perceptualui.org/publications/bulling07_eurossc.pdf[Last accessed on 26 Sep 2023]
47. Barbara N, Camilleri TA, Camilleri KP. A comparison of EOG baseline drift mitigation techniques. *Biomed Signal Proces* 2020;57:101738. DOI
48. Bhatnagar S, Gupta B. Acquisition, processing and applications of EOG signals. Available from: <https://ieeexplore.ieee.org/document/10009179>[Last accessed on 26 Sep 2023].
49. Chang WD, Choi JH, Shin J. Recognition of eye-written characters using deep neural network. *Applied Sciences* 2021;11:11036. DOI
50. Liu F, Shi Y. A study on artificial intelligence IQ and standard intelligent model. *arXiv* 2015;Online ahead of print:1512.00977. DOI
51. Cudlenco N, Popescu N, Leordeanu M. Reading into the mind's eye: boosting automatic visual recognition with EEG signals. *Neurocomputing* 2020;386:281-92. DOI
52. Tang S, Dunmon JA, Liangqiong Q, Saab KK, Baykaner T, Lee-Messer C, Rubin DL. Modeling multivariate biosignals with graph neural networks and structured state space models. Available from: <https://proceedings.mlr.press/v209/tang23a.html>[Last accessed on 26 Sep 2023]
53. Gupta SK, Zhang M, Wu CC, Wolfe J, Kreiman G. arXiv 2021. pp. arXiv-2106. Visual search asymmetry: deep nets and humans share similar inherent biases. *arXiv* 2016;Online ahead of print:02953 DOI
54. Liu M, Zhu H, Tang J, et al. Overview of a sleep monitoring protocol for a large natural population. *Phenomics* 2023;3:1-18. DOI PubMed PMC
55. Finstad K. The usability metric for user experience. *Interacting with Computers* 2010;22:323-7. DOI
56. Brooke J. Sus: a quick and dirty usability. Usability evaluation in industry. 1996;189: 189-94. Available from: http://www.tbstafftraining.info/smartphones/documents/b5_during_the_trial_usability_scale_v1_09aug11.pdf[Last accessed on 26 Sep 2023]