Enhancing User Experience Using a Framework Integrating Emotion Recognition and Eye-Tracking

Jianyi Wang
School of Design, Hunan University, Changsha, China
Jianyi830@hnu.edu.cn

Abstract. In a rapidly evolving digital landscape, ensuring a seamless and engaging user experience (UX) has become paramount. This study delves into the intricate realm of web interaction design, aiming to enhance user satisfaction and engagement. Through the integration of emotion recognition and eye-tracking technologies, a profound relationship between user emotions and web interface design is unveiled. The central theme of this research revolves around the integration of emotion recognition and eye-tracking technologies to evaluate web interaction designs. An experimental study involving 24 participants from diverse backgrounds and age groups was conducted. These participants navigated web interfaces that featured both emotion recognition and eye-tracking technologies. Three distinct tasks were carefully crafted to represent a spectrum of web interactions, ranging from basic navigation to complex decision-making. Data collection involved real-time valence and arousal scores, video recordings, visual attention patterns, task completion times, and questionnaires. The analysis revealed a series of significant discoveries. Users who engaged with simplified web versions consistently exhibited elevated valence values, indicative of heightened positive emotional feedback, diverging starkly from their counterparts navigating complex web iterations. The dynamic ebb and flow of emotional states, underscored by arousal levels, underscore the pivotal role of real-time emotional assessment in the critical evaluation of web interfaces. Moreover, the study unveiled a persistent preference for a state of emotional calmness during interactions, demonstrating a universal need for user-centered design principles that prioritize minimal cognitive load and emotional tranquility. In summation, this research contributes a robust framework to the design community and academia for the comprehensive evaluation of web interaction designs. The findings underscore the paramount significance of simplicity, real-time emotional evaluation, and unwavering adherence to user-centered design principles in the realm of web interaction. This study constitutes an invaluable repository for designers, developers, and researchers, steering their endeavors towards the relentless pursuit of optimized user experiences within the ever-evolving digital landscape.

Keywords: User Experience (UX); Emotion Recognition; Eye-tracking; Human-Computer Interaction (HCI); Valence and Arousal.

1. Introduction

As technology continues to reshape our world, the way people interact with digital platforms and applications is evolving rapidly. From communication to education, web-based interfaces play a crucial role in our daily lives. Ensuring a smooth and enjoyable user experience (UX) has become vital. This transition isn't just about individual experiences; it's revolutionizing industries and redefining success criteria.

1.1. Research Background and Significance

The realm of human-computer interaction has undergone substantial evolution in recent years, with a marked focus on elevating user experiences (UX). The foundations of UX can be traced back to the 'Experience Economy' theory, as propounded by American economists Joseph Pine and Gilmore in their influential work 'The Experience Economy' [1]. As depicted in Figure 1, The original concept of user experience originated in the field of product design and later gradually extended to interface and interaction design. It refers to the subjective comprehensive perception established by users while manipulating user interfaces. User Interface (UI) functions as the pivotal link connecting
users with a product's functionalities, encompassing visual components like colors, typography, and layout. It facilitates interactions, strategically positioning buttons to catalyze actions [2]. In this context, grasping and enhancing user experiences within web interfaces assumes paramount importance. The integration of emotion recognition and eye-tracking technologies presents a promising avenue to this end. These technologies empower us to decode user emotions and pinpoint their focal points during digital content navigation. Nonetheless, achieving a cohesive and integrated incorporation of these technologies into web design remains an ongoing challenge.

1.2. Current Research Landscape

Recent research has explored the realms of emotion recognition and eye-tracking in the domain of web interaction design. Moreover, advancements in user-interface elements such as gesture, speech, context, and affect have paved the way for the creation of environments that emanate a natural essence, transcending strict computerized interfaces. This evolution, including the reintegration of all four senses, psychological insights into the pivotal role of emotions in human cognition [3]. The study by Hussain, J., Ul Hassan, A., Muhammad Bilal, H.S. et al. introduces a domain and device-independent model-based adaptive user interface methodology, implemented as an A-UI/UX-A tool, which outperforms existing approaches by leveraging user context and experience [4]. And Pu, Chen, and Hu propose ResQue, an evaluation framework for recommender systems. It combines user-centric criteria, validates them through psychometrics, and yields a 32-question, 15-construct model for cost-effective assessment and development guidance [5].

1.3. Research Gap and Motivation

This research aims to bridge this gap by introducing an integrated evaluation framework. The analysis of emotional responses and visual attention during web interactions will provide nuanced insights into how design elements impact user experiences. This research aspires to offer a fresh perspective that guides the creation of web interfaces prioritizing user satisfaction and engagement.

1.4. Research Objectives

The primary objective of this research is to develop and validate a framework for assessing web interaction designs through the integration of emotion recognition and eye-tracking. This framework aims to provide designers and developers with a method for understanding how design choices influence user emotions and visual engagement. Through rigorous experiments, a deep understanding of the relationship between design decisions and user perceptions will be gained. In the subsequent sections, we will detail the research methodologies, present concrete findings, and engage in discussions. The study aims to uncover the intricate connections between emotion recognition, eye-tracking, and web interaction design, contributing to the improvement of human-computer interaction.
practices. It offers a roadmap for future design enhancements firmly rooted in user-centered principles and supported by solid evidence.

2. Literature Review

The evolution of User Experience (UX) has a rich historical trajectory that traces its roots back to product design, where its origins can be found. Over time, UX has transcended its initial domain and gradually expanded its influence to encompass interface and interaction design. This transformation has led to a holistic approach that incorporates diverse visual elements, including but not limited to color schemes, typography, and layout. As UX has evolved, it has become an interdisciplinary field, fostering collaboration among designers, psychologists, and technologists to create user-centered and visually engaging experiences across various digital platforms and applications. This historical development underscores the significance of UX as an integral part of contemporary design practices, emphasizing the paramount importance of user satisfaction and engagement [6].

One of the fundamental challenges in HCI today lies in the seamless integration of emotion recognition and eye-tracking technologies to enhance UX within web interfaces. Emotion recognition technology enables the identification and analysis of users' emotional states during digital interactions. On the other hand, eye-tracking technology offers insights into users' visual attention patterns and focal points while navigating web content. The integration of these two technologies holds great promise for creating web interfaces that can adapt to users' emotional states and optimize their visual engagement.

Recent research has explored the applications of emotion recognition and eye-tracking in web interaction design. These studies have demonstrated how the real-time detection of user emotions can lead to adaptive interfaces that tailor content based on emotional feedback [7]. Additionally, eye-tracking data has been used to identify areas of interest on web pages, allowing designers to optimize the placement of critical elements [8]. An illustrative example can be observed in Figure 2.

![Fig. 2. An example of eye-tracking heatmap](image-url)

However, the current research landscape lacks a comprehensive study that combines these dimensions to provide a holistic evaluation of user experiences within web environments. Existing studies often focus on either emotion recognition or eye-tracking, missing the synergistic potential of combining both technologies.

The motivation behind this research stems from the need to bridge this gap and introduce an integrated evaluation framework for web interaction design. This framework aims to analyze emotional responses and visual attention patterns simultaneously, offering nuanced insights into how design elements impact user experiences. By unifying emotion recognition and eye-tracking data, we
intend to provide designers and developers with a method to understand how design choices influence user emotions and visual engagement.

The primary research objective is to develop and validate a framework that assesses web interaction designs using both emotion recognition and eye-tracking technologies. Through research-based experiments, a deep understanding of the relationship between design decisions, user emotional responses, and visual attention patterns will be gained. This study contributes to the improvement of human-computer interaction practices, offering a roadmap for future design enhancements grounded in research-based evidence and user-centered principles.

In the following sections, the research methodologies will be detailed, more findings will be presented, and discussions will shed light on the connections between emotion recognition, eye-tracking, and web interaction design.

3. Method

3.1. Experimental Design

An experimental study was conducted to explore the impact of the combined application of emotion recognition and eye-tracking technologies on user experience within web interaction design. In this study, all participants engaged with web interfaces that featured both emotion recognition and eye-tracking technologies.

3.2. Participants

A total of 24 participants were recruited, representing diverse backgrounds and age groups. Participants ranged in age from 18 to 33 years (mean age = 22.8 years, standard deviation = 3.6 years). All participants had prior experience using web interfaces, ensuring a basic level of familiarity with the tasks.

3.3. Procedure

Participants engaged in a multifaceted exploration of web interfaces while concurrently interacting with emotion recognition and eye-tracking technologies. This work designed two versions of websites to do the experiment. In Figure 3, which can see my simplified and complex versions of the web surface.

![Figure 3. Complex Version of the Website (Left) Simplified Version of the Website (Right)](image)

To comprehensively evaluate user experiences, three distinct tasks were designed within the web interfaces. These tasks encompassed: 1) User Registration and Login: Participants initiated their interaction by registering and logging into the web interface. 2) Membership Understanding and Application: In this task, participants explored the website's membership mechanism and applied for membership. 3) Shopping Cart Management and Checkout: The third task involved adding specified items to the shopping cart and completing the checkout process.

These tasks were carefully selected to represent a spectrum of web interactions, from basic navigation to complex decision-making. To minimize any order effects, the sequence in which participants encountered the interface versions and technology conditions was meticulously counterbalanced.
As participants engaged with the web interfaces, a rich dataset was collected in real-time. This dataset included valence and arousal scores, capturing emotional intensity and valence, sampled at regular intervals, typically every 3 seconds. These scores provided indirect insights into participants' emotional fluctuations during web interactions. To enhance the granularity and accuracy of these assessments, valence and arousal scores were synchronized with concurrent video recordings of participants' interactions with the web interfaces. This video analysis enabled a nuanced understanding of how users' feedback and emotional states correlated with specific interface functionalities.

Furthermore, participants' visual attention patterns were comprehensively tracked using eye-tracking technology, generating heatmaps that vividly illustrated where their attention was predominantly focused on the web pages. Analyzing these heatmaps deepened the comprehension of how interface elements influenced users' visual engagement. Additionally, the time taken by participants to complete each of the designated tasks was recorded, providing insights into their efficiency and task-completion performance.

3.4. Data Collection and Analysis

Data analysis embraced a multifaceted approach, underscoring the diverse data streams acquired during the experiment. Valence and arousal data were used to create two line graphs, with time intervals of approximately 5 seconds on the x-axis. These graphs served as dynamic indicators of participants' emotional states, enabling the discernment of how specific functionalities of the web interfaces influenced their emotional responses. Video analysis further enriched these insights, linking emotional states to users' interactions with interface elements.

The eye-tracking heatmaps provided essential visualizations of participants' attention distribution across web pages, offering a visual narrative of their focal points. These heatmaps were instrumental in correlating attention patterns with emotional fluctuations and user feedback. Efficiency was assessed by analyzing task completion times, shedding light on participants' proficiency in navigating the web interfaces. Quantitative and qualitative data from post-interaction questionnaires furnished subjective feedback, granting participants an avenue to express their perceptions and sentiments regarding their experience.

These diverse data sources underwent a thorough analysis, which allowed for the uncovering of the intricate interplay between emotional states, visual attention, task performance, and user feedback. The holistic approach to data collection and analysis enhanced the comprehensiveness and robustness of the study, offering valuable insights into the user interaction design landscape.

3.5. Enhancing Methodological Validity

3.5.1 Facial Expression Recognition

The experimental methodology drew inspiration from well-established techniques in the field of facial expression recognition, particularly the insights and VGG19-based architecture proposed in [9]. The methodologies outlined in [4] significantly influenced the deep learning approach, providing essential guidance on model architecture and training procedures. The integration of core principles from this body of work enhanced the methodological soundness of the study, ultimately contributing to a more comprehensive examination of facial expression recognition.

3.5.2 Eye-Tracking Technology Support

In addition to the previously described methodologies, this study benefited from the incorporation of eye-tracking technology. This technology allowed for the precise capture of visual attention patterns during participants' interactions with the web interfaces, further enhancing our understanding of user experience in the context of facial expression recognition. To maintain methodological validity, the guidelines established by GazeRecorder Inc. (2019) for conducting eye-tracking experiments were strictly adhered to, ensuring calibration accuracy and reliable data collection [10].
The integration of facial expression recognition and eye-tracking technology played a pivotal role in strengthening the overall methodological framework of the study. Through the integration of these technologies, the goal was to provide a more holistic analysis of user experience, recognizing the intricate interplay between facial expression recognition and visual attention patterns. This comprehensive approach enhances the methodological rigor of the research, yielding invaluable insights into the landscape of user interaction design.

4. Results

Having completed the experiments, essential data has been collected to delve into user experiences in web interaction design. This section presents the findings, including user emotional responses, visual attention patterns, task completion times, and feedback.

4.1. Analysis of Heatmap

Based on the analysis of the heatmaps, several noteworthy patterns emerge. In the heatmaps obtained from the complex version of the web interface, longer average interaction times were observed among users, resulting in a more extensive and dispersed distribution of colored areas. This distribution is attributed to the increased presence of images and text in the complex version, leading to overall darker-shaded regions compared to the simplified version. Notably, in areas requiring significant user interaction, such as navigation or decision-making tasks, more time was spent by users, resulting in deeper coloration. For two typical examples illustrating these patterns, please refer to Figure 4.

![Fig. 4. Two heatmap examples of the complex website](image)

Conversely, the heatmaps derived from the simplified version of the web interface displayed a shorter average interaction time, corresponding to a smaller and more concentrated distribution of colored areas. The concentration primarily centered around critical information areas. Interestingly, within these key information regions, the coloration depth was relatively similar to that observed in the complex version's heatmap. More heatmaps of simplified web interface was shown in Figure 5.

![Fig. 5. Two Heatmap Examples of the Simplified Website](image)

4.2. Analysis of emotion intensity recognition

In the study, users' emotional responses were assessed using two primary dimensions: arousal and valence. Arousal measured the level of emotional activation experienced during web interface interactions, differentiating between heightened states (e.g., excitement or agitation) and subdued states (e.g., calm or relaxation). Valence, on the other hand, measured the emotional tone, indicating
whether emotions tended toward positivity (e.g., joy or satisfaction) or negativity (e.g., sadness or anger). More detailed explanations of arousal and valence can be found in Figure 6 [11]. These metrics were instrumental in unraveling users' emotional states during web interface interactions. Further details on arousal and valence can be found in [12].

![Two-dimensional Valence-arousal Space](image)

**Fig. 6. Two-dimensional Valence-arousal Space**

In this study, arousal and valence data were recorded at three-second intervals for two selected participants during their interactions with both the simplified and complex web versions. Each data point on the scatter plots represented a participant's emotional state at a specific time, with valence values on the x-axis and arousal values on the y-axis. Data points from the complex web interactions were indicated in red, while those from the simplified web were represented in blue.

Additionally, line graphs were generated for each participant to visualize their emotional trajectories while interacting with both web versions. The x-axis represented time, while the y-axis depicted emotional intensity and valence, ranging from -1 to 1. Each graph featured two lines: a red line for arousal values and a blue line for valence values, offering dynamic insights into participants' emotional states over time.

For a closer look at the data representations for two sample users, please refer to Figures 7, 8 and 9.

![Emotion Scatter Plots For Two Users](image)

**Fig. 7. Emotion Scatter Plots For Two Users**

In the scatterplots, it is evident that blue points (representing the simplified web version) tend to cluster a little more towards the right side of the screen compared to the red points (representing the complex web version). This clustering suggests that users interacting with the simplified web version had higher valence values, indicating slightly more positive emotional feedback. On the other hand, the red points exhibit a wider distribution, suggesting greater emotional fluctuations among users.
when using the complex web version. This variability in emotional responses may be attributed to potential distractions caused by extraneous elements on the website.

Overall, both sets of points are centered around the middle and slightly rightward positions, indicating that users, in general, experienced relatively calm emotions during their interactions and were not notably frustrated.

Next, a more detailed analysis of the line graphs will be conducted to gain further insights into the temporal aspects of user emotions.

Analyzing the valence lines from the line graphs reveals a general trend of initially decreasing and then increasing valence values throughout the user interaction period. This pattern suggests that users’ emotional states tended to dip at the start of the interaction, potentially due to adjusting to the interface, before gradually improving over time. However, it's important to note that these trends are interspersed with multiple fluctuations.

On the other hand, when examining the arousal lines, frequent oscillations are observed, indicating that users’ arousal levels were quite dynamic during the interaction. However, when aggregating and averaging arousal values across all participants, it becomes evident that arousal values tend to cluster around a value close to zero. This suggests that, on the whole, users maintained a relatively calm emotional state throughout the testing process, despite occasional fluctuations.

When considering both lines together, it becomes apparent that during testing with the simplified web version, the trends of both valence and arousal lines tend to be more similar, with less pronounced fluctuations. In contrast, during testing with the complex web version, both lines exhibit more unpredictable patterns, indicating that user emotions were more volatile and challenging to predict in this context.
5. Discussion

In this section, a comprehensive discussion of the findings and their implications will be presented, drawing connections to existing literature and shedding light on the multifaceted nature of user experiences in web interface design, particularly concerning the integration of emotion recognition and eye-tracking technologies.

5.1. Interpretation of Results

This study unveiled several intriguing insights into the intricate interplay between users' emotional responses and visual attention patterns during web interface interactions. Through the analysis of arousal and valence dimensions, valuable understanding of users' emotional states was gained.

The scatterplots (Figure 7) illustrate that users interacting with the simplified web version generally exhibited higher valence values, suggesting more positive emotional feedback, compared to those using the complex web version. This observation aligns with prior research emphasizing the importance of user-friendly, uncluttered interfaces in eliciting positive emotional responses [13].

Conversely, the complex web version led to greater emotional fluctuations, indicated by the wider distribution of valence points. This could be attributed to potential distractions stemming from the inclusion of extraneous elements on the website, reinforcing the significance of interface simplicity in maintaining user engagement and satisfaction.

The trends observed in the line graphs (Figures 8 and 9) further illuminate users' emotional trajectories over time. The initial dip in valence values, followed by an upward trend, suggests users' gradual adaptation to the interface. The frequent oscillations in arousal levels indicate dynamic emotional responses during interactions, yet the aggregation of these values highlights the overall calm emotional states of users.

5.2. Comparison with Existing Literature

The findings presented here align with existing literature that emphasizes the relevance of emotional design principles in web interface development [13]. The positive valence values associated with the simplified web version corroborate prior studies highlighting the role of user-friendly, aesthetically pleasing interfaces in fostering positive emotional experiences [14]. Furthermore, the dynamic nature of arousal levels underscores the significance of real-time emotional assessment in web interface evaluation, aligning with research that advocates for the incorporation of emotion recognition technologies into user experience research [15].

5.3. Implications

The study carries significant implications for web interface design. The preference for simplified interfaces with higher valence values emphasizes the importance of minimalistic design and streamlined content presentation. Designers should strive to minimize distractions and optimize user engagement by providing clear, intuitive navigation paths.

The consistent trend of maintaining relatively calm emotional states throughout interactions has broader implications. It suggests that users value web interfaces that do not induce frustration or agitation, underscoring the importance of user-centered design principles in minimizing cognitive load and enhancing user satisfaction [15].

5.4. Limitations and Future Research

Despite the valuable insights gained, this study has obvious limitations. The relatively small sample size warrants caution in generalizing the results. Future research should consider larger and more diverse participant groups to ensure broader applicability. Furthermore, this study focused on specific tasks within web interfaces. Future investigations could explore a wider range of tasks and incorporate additional variables, such as user demographics and prior experience, to provide a more nuanced understanding of user experiences.
In conclusion, this study underscores the intricate relationship between user emotions and web interface design, emphasizing the significance of simplicity and real-time emotional assessment. Through the integration of emotion recognition and eye-tracking technologies, a valuable framework is provided for enhancing user experiences, ultimately contributing to the advancement of web interface design practices.

5.5. Implications for Future Research

The implications of the findings extend to several dimensions of web interaction design and user experience research. This study contributes to the existing body of knowledge by emphasizing the importance of integrating emotion recognition and eye-tracking technologies. Future researchers may find this work instrumental in enhancing their studies. The primary contribution is a comprehensive framework for evaluating web interaction designs through emotion recognition and eye-tracking. This framework can provide designers and developers with valuable insights into how design choices influence user emotions and visual engagement. Researchers can build upon this framework to delve deeper into specific aspects of web interaction design and assess its applicability in various contexts.

5.6. Outlook for Future Research

While this research has yielded valuable insights, it is essential to acknowledge its limitations. The relatively small sample size limits the generalizability of these results. Future research should consider larger and more diverse participant groups to ensure broader applicability. Additionally, this study focused on specific tasks within web interfaces. Subsequent investigations could explore a wider range of tasks and incorporate additional variables, such as user demographics and prior experience, to provide a more nuanced understanding of user experiences. In conclusion, this study has shed light on the intricate relationship between user emotions and web interface design. By integrating emotion recognition and eye-tracking technologies, a valuable framework for enhancing user experiences has been offered, ultimately contributing to the advancement of web interface design practices. These findings emphasize the importance of simplicity, real-time emotional assessment, and user-centered design principles in web interaction, opening doors for further investigations and practical applications in this dynamic field.

6. Conclusion

While this research has yielded valuable insights, it is essential to acknowledge its limitations. The relatively small sample size limits the generalizability of these results. Future research should consider larger and more diverse participant groups to ensure broader applicability. Additionally, this study focused on specific tasks within web interfaces. Subsequent investigations could explore a wider range of tasks and incorporate additional variables, such as user demographics and prior experience, to provide a more nuanced understanding of user experiences.

In conclusion, this study has shed light on the intricate relationship between user emotions and web interface design. By integrating emotion recognition and eye-tracking technologies, a valuable framework for enhancing user experiences has been offered, ultimately contributing to the advancement of web interface design practices. These findings emphasize the importance of simplicity, real-time emotional assessment, and user-centered design principles in web interaction, opening doors for further investigations and practical applications in this dynamic field.

Secondly, the dynamic nature of arousal levels observed during our study implies that real-time emotional assessment is a valuable asset in web interface evaluation. This aligns with previous research advocating for the incorporation of emotion recognition technologies into user experience research. By recognizing the potential of emotional fluctuation during user interactions, designers can tailor content to adapt to users’ emotional states in real-time, offering personalized experiences that resonate more deeply with users’ emotional needs. This recommendation serves as a response to questions regarding the dynamic emotional responses of users during web interactions.
Third, the preference for calm emotional states throughout interactions, regardless of the interface's complexity, signifies the importance of user-centered design principles. Interfaces should aim to minimize cognitive load and emotional agitation. This aligns with prior literature highlighting the significance of interface simplicity in maintaining user engagement and satisfaction. Designers should prioritize creating interfaces that do not induce frustration or agitation to ensure a harmonious user experience. This addresses questions concerning users' emotional states and their significance in web interaction design.

References


