https://doi.org/10.31891/csit-2024-4-14
UDC 004.5

Serhii ZELINSKYI, Yuriy BOYKO
Taras Shevchenko National University of Kyiv

USING FACIAL EXPRESSIONS FOR CUSTOM ACTIONS: DEVELOPMENT AND EVALUATION OF A HANDS-FREE INTERACTION METHOD

This study explores a novel facial expression-based interaction method designed to provide an accessible and hands-free alternative for performing precision tasks. Traditional input devices, such as keyboards and mice, are often unsuitable for individuals with limited mobility or for hands-free environments. The proposed system leverages standard computer hardware and machine learning-based facial landmark detection to map customizable facial expressions to specific actions, making it a low-cost and adaptable solution. This study evaluates the usability, learnability, and potential applications of this interaction method through a task-based user study. Sixteen participants aged 19–34 completed a series of five trials, performing the same color and number-matching task on an interactive grid. This approach allowed the evaluation of the learning curve by analyzing how participants improved their skills and reduced task completion times with each subsequent trial. Participants also provided feedback on challenging facial expressions to identify usability challenges. The evaluation focused on task completion time, participant-reported challenging actions, and qualitative feedback to assess system usability, user adaptability, and potential applications. The results indicate a clear learning curve, with participants improving task completion times over repeated trials. Feedback highlighted the potential of this interaction method for assistive technologies, gamified facial exercises, and as a supplementary input tool, while also identifying challenges such as facial fatigue and action complexity. The findings demonstrate the system's promise as an accessible and adaptable alternative interaction method, with opportunities for future refinement and broader application. The proposed interaction method demonstrates significant promise as an accessible, flexible alternative for hands-free interaction. This research contributes to the growing field of alternative interaction methods and supports the development of inclusive and adaptable human-computer interaction technologies.

Keywords: facial expression-based interaction, facial blendshapes, hands-free interaction, human-computer interaction (HCI), accessibility, adaptive input methods, usability evaluation, assistive technology, gamified interaction, customizable actions.

Сергій ЗЕЛІНСЬКИЙ, Юрій БОЙКО
Київський національний університет імені Тараса Шевченка

ВИКОРИСТАННЯ ВИРАЗІВ ОБЛИЧЧЯ ДЛЯ ВИКОНАННЯ КОРИСТУВАЦЬКИХ ДІЙ: РОЗРОБКА ТА ОЦІНЮВАННЯ МЕТОДУ ВЗАЄМОДІЇ БЕЗ ВИКОРИСТАННЯ РУК

Це дослідження представляє новий метод взаємодії, заснований на виразах обличчя, що пропонує доступну та безконтактну альтернативу для виконання точних завдань. Традиційні пристрої введення, як-от клавіатури та миші, часто не підходять для людей з обмеженою рухливістю або у середовищах, де потрібна безконтактна взаємодія. Запропонована система використовує стандартне комп'ютерне обладнання та алгоритми машинного навчання для розпізнавання орієнтирів обличчя, дозволяючи зіставляти настроювані вирази обличчя зі специфічними діями, що робить її недорогою та гнучкою. У ході користувацького дослідження 16 учасників віком від 19 до 34 років п'ять разів поспіль виконали одне й те саме завдання із зіставлення кольорів і чисел на інтерактивній сітці. Такий підхід дозволив оцінити криву навчання, аналізуючи, як учасники покращували свої навички та скорочували час виконання завдання з кожною наступною спробою. Відгуки учасників також допомогли ідентифікувати складність відтворення виразів обличчя та інші виклики зручності використання. Оцінювання охоплювало час виконання завдань, складність відтворення виразів обличчя, зазначену учасниками, та якісні відгуки для аналізу зручності, адаптивності системи та її потенційного застосування. Результати демонструють позитивну криву навчання: учасники поступово скорочували час виконання завдань із кожною наступною спробою. Учасники також підкреслили потенціал цього методу взаємодії в контексті допоміжних технологій, гейміфікованих вправ для м'язів обличчя та як додаткового інструменту до традиційних пристроїв введення. Водночас було виявлено такі виклики, як втома м'язів обличчя та складність відтворення окремих виразів. Отримані результати підтверджують перспективність системи як доступної та адаптивної альтернативи, з можливостями подальшого вдосконалення та широкого спектру застосування. Запропонований метод взаємодії демонструє значні перспективи як доступна, гнучка альтернатива для взаємодії без допомоги рук. Це дослідження є внеском у зростаючу галузь альтернативних методів взаємодії та підтримує розвиток інклюзивних та адаптивних технологій взаємодії людини з комп'ютером.

Ключові слова: взаємодія на основі виразів обличчя, блендшейпи обличчя, безконтактна взаємодія, взаємодія людина-комп'ютер (HCI), доступність, адаптивні методи введення, оцінка зручності використання, допоміжні технології, гейміфікована взаємодія, настроювані дії.

Introduction

As web interfaces become increasingly complex and are used in diverse settings — ranging from home applications to industrial environments — and by a broad audience, including people with disabilities, the demand for innovative, hands-free, and adaptive interaction methods has grown significantly.

Motivation. Traditional input devices, such as a mouse or keyboard, often fail to accommodate users with limited mobility or those operating in environments where hands-free interaction is necessary. Moreover, current technologies lack the ability to provide customizable and intuitive methods for interacting with digital systems. Facial expression-based interaction offers a promising alternative, leveraging natural gestures to enable precise, adaptable, and accessible input.

Objective. This study aims to develop and evaluate a facial expression-based interaction method that relies on standard computer hardware, making it an affordable and widely accessible solution. The system is designed to provide a practical, inclusive, and adaptable interaction method, addressing the needs of diverse users, including those with physical limitations. The study evaluates the system's effectiveness in performing precision tasks, such as matching specific colors and numbers, and its potential as a supplementary tool for traditional input methods.

Approach. The proposed system uses standard computer hardware and a machine learning model to detect facial landmarks and analyze facial blendshapes [1] in real time. Customizable facial expressions are mapped to specific actions, enabling users to perform tasks like navigation, selection, and manipulation without the need for physical devices. The system was evaluated through a user study in which participants performed the same task across five trials, with their performance and feedback recorded.

Metrics. To quantify and qualify the system's usability, the study measured:

1. **Task completion time.** To observe learning effects and assess efficiency across trials.
2. **Challenging actions.** Based on user feedback, specific facial expressions and combinations were evaluated to identify those participants found difficult to perform.
3. **User feedback.** To capture qualitative insights into the usability, intuitiveness, and challenges of the interaction method, as well as potential applications of the suggested interaction method.

By integrating these metrics, the study evaluates the system's learnability, usability, and potential real-world applications.

Related Work

Facial expressions are one of the promising modalities for human-computer interaction that can offer new ways of interacting with digital systems. During the past year, a number of studies have used facial expressions for human-computer interaction.

The use of facial expressions for "natural interaction" with devices like smartphones and tablets was explored in [2]. The study presented an eBook reader use case where users could control the device using facial expressions. Their system interprets user facial expressions and emotions to obtain user input. While this approach demonstrated the feasibility of natural interaction using facial expressions, it was limited to basic gestures, lacked evaluation in precision tasks, and did not adapt to diverse user needs, reducing its applicability for complex scenarios like object manipulation. Another study discussed facial expressions and hand gestures for healthcare and smart environment applications [3], suggesting that facial expression recognition could help physically impaired people control devices. While the study demonstrates the applicability of facial expressions and hand gestures for smart environment interactions, it did not specifically evaluate the system’s effectiveness for precision tasks, which limits its application to more complex interactions. Similarly, [4] discussed using facial expression recognition and computer vision for facial analysis in human-computer interaction, focusing on improving the interaction between humans and machines, particularly for individuals with limited mobility. This work improves facial recognition and analysis but focuses on evaluating expressions and predicting muscle activity, with limited attention to practical usability or integration into real-world interaction systems.

Another study proposed a human-computer interaction system based on facial feature localization, expression recognition, and gaze tracking [5]. The approach integrates gaze tracking but does not address the complexity of mapping facial expressions to precise actions. A transfer learning approach for facial emotion recognition in human-computer interaction was utilized in [6]. The authors noted that emotion recognition is useful for many tasks, such as customer satisfaction, criminal justice, e-learning, social robots, and smart cards. However, this study focused on recognizing affective states rather than enabling actionable user interactions, limiting its relevance for hands-free control systems. While facial expressions might be useful, another research that studied affective state recognition using facial expressions captured during active computer tasks suggests that facial expressions alone might not accurately reveal user feelings during computer interaction [7].

A digital human system that uses synchronized and realistic facial expressions for friendly human-machine interaction was developed in [8]. A parameterized 3D model for facial animations was built using blendshapes theory to effectively generate facial expressions. The system tracks the user's facial movements and synchronizes the expressions of the digital human. Although this system focuses on realistic facial animations, it does not address interaction tasks that require precise control of objects. Similarly, a proof-of-concept interface that uses electromyography (EMG) to detect facial muscle activity and control a 3D avatar was developed [9]. The user's expressions control the avatar's facial movements, conveying emotions in a virtual environment. Blendshape deformers on 3D objects were used to create animations, which allows for precise control of the avatar's facial muscles. This system, while innovative, requires specialized equipment (EMG sensors), reducing its accessibility compared to systems relying on standard hardware like webcams. Another study introduced a wearable device called OCOsense glasses, equipped with sensors that monitor facial gestures and expressions for augmented human-computer interaction [10]. The glasses use optomyographic (OMG) technology to recognize facial movements. Similarly to [9], this system relies on specialized hardware, making it less accessible for general users.

Another study developed FaceSwitch, an open-source accessibility software that enhances gaze interaction with facial gestures [11]. The system allows users to select targets with their gaze and trigger actions with facial gestures. However, its focus on gaze interaction and basic gesture-triggered actions does not extend to complex task control. Other studies explored the emulation of standard computer input devices, such as a mouse and keyboard. In [12], a vision-based system was designed for a child with a degenerative neuromuscular disease to interact with a computer through facial expression recognition. The system uses a standard webcam and PCA for facial expression classification and recognizes "smile" and "kiss" expressions to activate mouse functions. While effective for basic mouse actions, this system lacks support for advanced interaction tasks. [13] developed a system to emulate mouse and keyboard functions using facial gestures and head movements. The system used head movement for mouse control and eye-closing gestures for clicking. Similarly, [14] developed a face gesture-based virtual mouse using MediaPipe. The study aimed to make computer interaction more accessible to people with disabilities. These studies focus on basic input emulation, providing limited functionality for more complex applications.

In our previous study [15], a blendshape-based human-computer interaction method was developed and evaluated in a preliminary usability study. Some participants reported difficulties in remembering which facial expressions corresponded to specific actions, suggesting the use of more intuitive expressions. However, they appreciated the benefits of the system's "hands-free" interaction. In the previous study, users were required to hold the "Shift" key while applying actions using facial expressions to ensure reliability. This limitation reduced the system's intuitiveness and accessibility for users with limited mobility.

Based on the aforementioned analysis, existing systems demonstrate limitations such as reliance on specialized hardware, focus on basic input emulation, or lack of support for precise task control. These gaps underscore the need for a system that leverages standard hardware to provide intuitive, hands-free interaction for advanced tasks like object manipulation. In this study, we aim to address these challenges by introducing a new set of potentially more intuitive facial expressions and eliminating the need to hold the "Shift" key, making the interaction fully hands-free.

Methodology

In this study, we used the MediaPipe Face Landmarker [16], a machine-learning model designed to detect and analyze facial landmarks in real time. One of the key advantages of this model is that it can operate directly in a web browser and requires only a standard webcam for capturing and processing facial data, which makes it accessible and practical for a variety of applications. The Face Landmarker outputs a comprehensive set of data points, including key facial landmarks and blendshape scores, which represent specific facial expressions based on the movements of different facial muscles.

System Design and Workflow. The proposed system implements a structured pipeline for real-time facial expression-based interaction. This pipeline transforms raw input data into actionable system responses through a sequence of well-defined stages, as illustrated in Fig 1. The pipeline is divided into two main stages: the Processing Stage and the Action Execution Stage. The pipeline includes the following steps:

1. **Input.** Real-time video input is captured from a standard webcam, which is used for facial expression detection.
2. **Processing Stage:**
   - **Facial Expression Detection and Scoring.** MediaPipe Face Landmarker detects key facial landmarks and computes blendshape scores.
   - **Threshold Filtering.** Blendshape scores are filtered to identify active expressions for further action mapping.
3. **Action Execution Stage:**
   - **Action Mapping.** Recognized expressions are mapped to predefined system actions (e.g., navigation, selection).
   - **System Response and Interaction Execution.** The system executes the mapped action in real time.
4. **Output.** The process concludes with the execution of the mapped action, providing feedback to the user.

Selecting Facial Blendshapes for This Study. Facial blendshapes serve as the basis for defining custom actions in this study. MediaPipe Face Landmarker's output includes a set of blendshape scores in the range [0-1], each corresponding to a distinct facial expression or movement. In this implementation, a blendshape score is considered active if it reaches or exceeds a threshold of 0.5. This threshold value was determined empirically through a series of pilot tests, balancing sensitivity and accuracy to ensure that subtle facial expressions did not unintentionally trigger actions. The model outputs 52 blendshapes, but only a subset was selected for defining actions to maintain simplicity and reliability in the interaction system.

Prior to full implementation, the blendshape scores were evaluated in a series of tests with a separate group of users to ensure that the selected blendshapes were easy to activate and consistently recognized across different users. Based on these evaluations, only a subset of blendshapes was incorporated into the final design. Excluded blendshapes involved subtle or less consistent movements that participants found challenging to replicate reliably.

To create a reliable interaction method, eye blinks and gaze directions were intentionally excluded from the implemented actions. Since blinking occurs involuntarily, users might accidentally trigger actions if blinks were used as an activation signal. Moreover, closing the eyes restricts users from seeing the screen, potentially disrupting the experience. However, the system can be configured to use a single eye closure or more complex expressions, such as combining a single eye closure with another facial expression, like a smile. Additionally, while not implemented in this study, gaze directions (e.g., looking left or right) could be incorporated in other use cases, such as changing viewpoints in 3D virtual spaces.

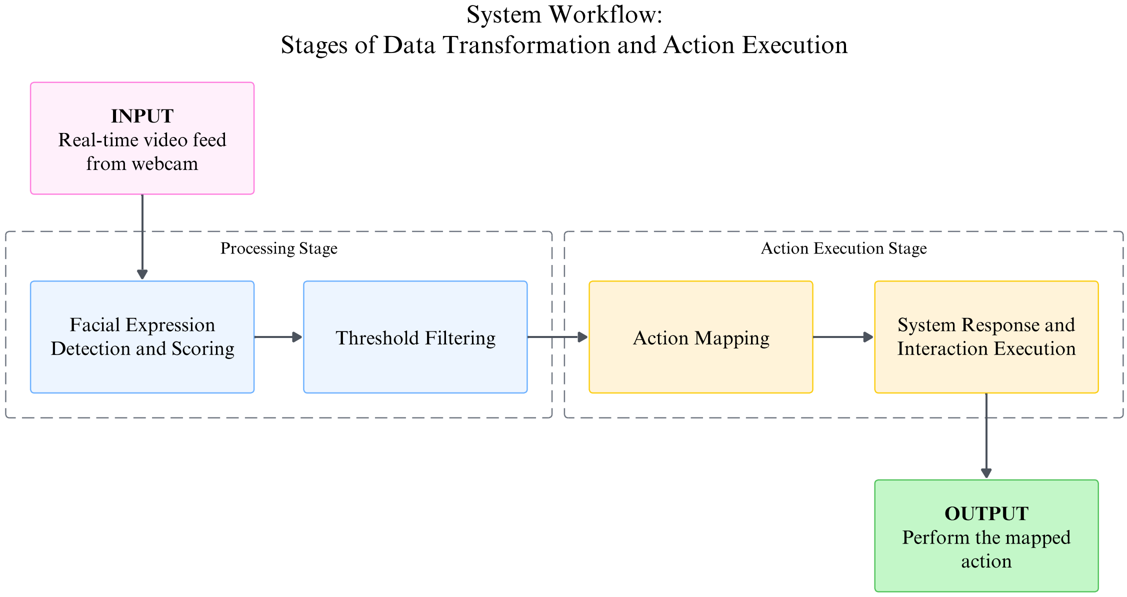


Fig. 1. Stages of the proposed system workflow, starting with real-time input from the webcam, processing facial expressions, mapping actions, and producing the final output

To formalize the definition of facial expressions used in this study, a function $F(blendshape)$ is introduced:

$$F(blendshape) = \begin{cases} 1, & blendshape\ score \geq 0.5 \\ 0, & blendshape\ score < 0.5 \end{cases}$$

The function returns true if the blendshape score is greater than or equal to 0.5, indicating that the blendshape is active.

The selected facial expressions are summarized in Table 1, where each row contains the expression name along with the corresponding formula.

Table 1

Facial Expressions Definition

Facial Expression	Facial Blendshapes Formula
Brows Up	$F(browInnerUp) \wedge F(browOuterUpLeft) \wedge F(browOuterUpRight)$
Brows Down	$F(browDownLeft) \wedge F(browDownRight)$
Smile	$F(mouthSmileLeft) \wedge F(mouthSmileRight)$
Smile Left	$F(mouthSmileLeft) \wedge \neg F(mouthSmileRight)$
Smile Right	$\neg F(mouthSmileLeft) \wedge F(mouthSmileRight)$
Mouth Pucker	$F(mouthPucker)$
Jaw Open	$F(jawOpen)$

System Overview and Evaluation Setup. The developed system for evaluation includes an interactive grid designed to assess users' ability to match specified colors and numbers using facial expression-based interactions. The setup is intended to test the effectiveness and usability of the interaction methods by engaging users in a color and number-matching task.

Interactive grid. The interface consists of a grid with multiple cells, where some cells are designated as interactable. Each interactable cell contains a color (one of four possible colors) and a number. At the top of each interactable cell, a circular icon displays a target color and a number. This circle serves as a reference for users to match each cell's properties according to the task requirements.

Task goal. Starting at the top-left corner of the interactive grid, users are required to navigate to each interactable cell on the grid and match the color and number of the cell with the specified target color and number shown in the circle. Users need to adjust each cell's color and number until both attributes align with the target.

Task completion. The task is considered complete once all interactable cells in the grid match their respective target properties.

Table 2 provides a facial expression formula for each custom action.

Fig. 2 contains a screenshot of the developed system's user interface. The following key elements are numerically highlighted:

1. **Video feed.** Displays the user's webcam feed and the detected facial landmarks in real time.
2. **Actions configuration.** Contains a list of available actions and their corresponding facial expressions.
3. **Interactive grid.**
4. **Color palette.** Contains four available colors users can pick from.
5. **Stopwatch.** Displays the time spent since the user started completing the task. It stops once all interactable cells are matched, displaying the completion time.
6. **Reset.** A button to reset the interactive grid state and start completing the task again.

Table 2

Custom Actions Definition

Custom Action	Facial Expression Formula
Move Up	Brows Up
Move Down	Brows Down
Move Left	Smile Left
Move Right	Smile Right
Select Previous Color	Jaw Open ∧ Brows Up
Select Next Color	Jaw Open ∧ Brows Down
Decrease Number	Mouth Pucker ∧ Brows Down
Increase Number	Mouth Pucker ∧ Brows Up

Interaction Workflow. The action recognition workflow is illustrated in Fig. 3. The mechanics of applying facial expression actions work similarly to "keydown" events, where users perform an action by making a specific facial expression, and the action is triggered immediately upon detecting the expression and only once.

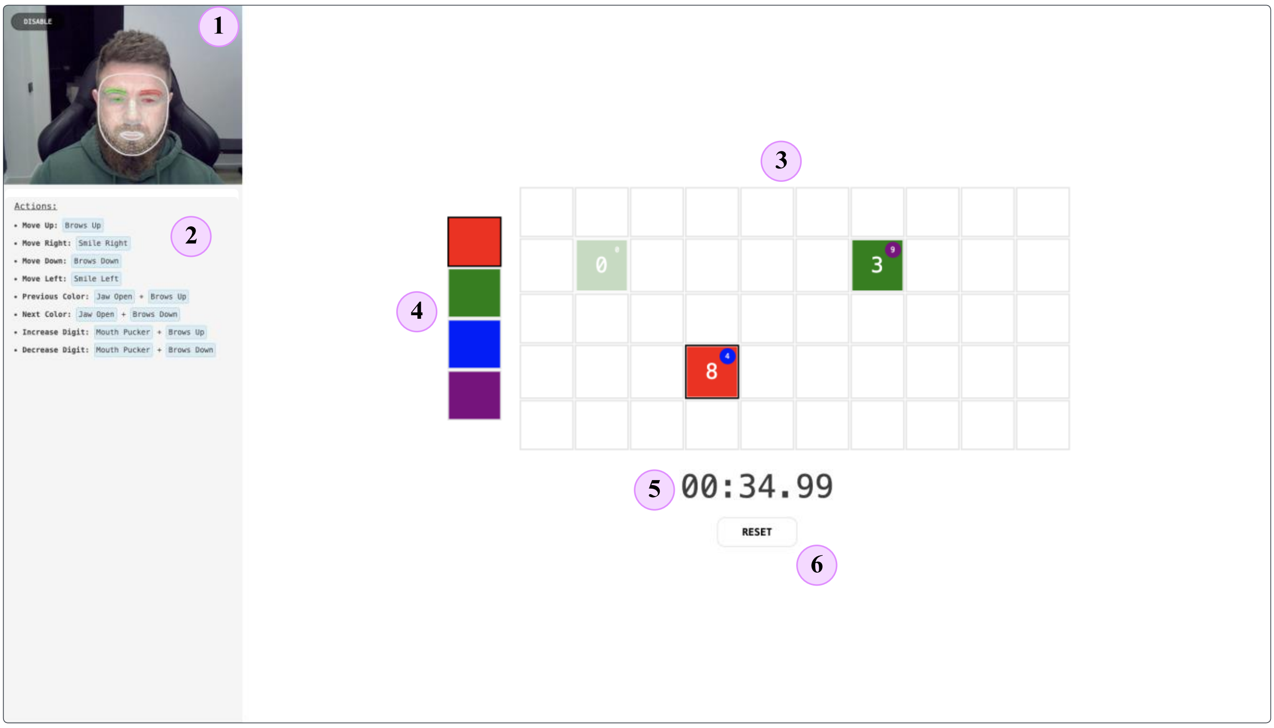


Fig. 2. Screenshot of the developed system user interface

To perform another action, users must return their brows or mouth to the neutral position. To keep the diagram clean, it doesn't reflect steps when the system waits while the user's brows or mouth returns to the neutral position.

For actions that involve two consecutive facial expressions, such as selecting colors or manipulating numbers, the system first recognizes the initial expression (e.g., a smile for color selection), which "activates" the selection mode. Once activated, the user can modify the selection by raising or lowering their brows to adjust the color or number. In these cases, only the brows need to return to neutral after each selection, while the initial expression (such as "Jaw Open" or "Mouth Pucker") can be maintained.

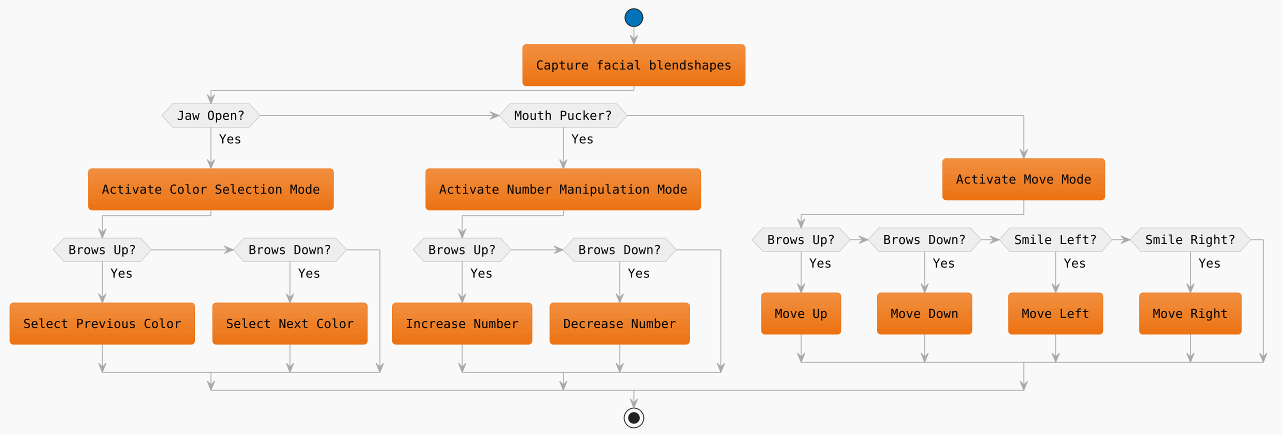


Fig. 3. Facial expression actions recognition workflow

Fig. 4 illustrates all custom actions and their respective facial expressions from Table 2.

Fig. 4. Facial expressions corresponding to each supported action in the system, illustrating the mapping between expressions and their respective functions

Evaluation Metrics and Justification. To evaluate the proposed system, a combination of quantitative and qualitative metrics was utilized. These metrics were selected based on established methodologies in human-computer interaction (HCI) to comprehensively assess the system's usability and effectiveness:

1. **Task completion time.** Task completion time was measured to evaluate the efficiency of the interaction method. This metric is widely used in usability studies (e.g., ISO 9241-11) to assess how quickly users can complete tasks, particularly when adapting to a new system. Observing changes in task completion time across trials also provides insights into the system's learnability.

2. **User feedback.** Qualitative feedback was collected from participants to capture subjective insights into the system's usability, intuitiveness, and experienced challenges. This feedback is particularly important for identifying usability issues that may not be evident in quantitative metrics. Additionally, participants provided suggestions for potential applications of the interaction method in various scenarios and contexts.

By using these metrics, the evaluation captures both objective performance data (e.g., efficiency and consistency) and subjective user experience, which provides a comprehensive assessment of the system's usability and potential for real-world applications.

Evaluation and Results

The study involved 16 participants aged between 19 and 34 years. Each participant completed a task involving matching cell properties on an interactive grid. The task was repeated five times (five trials), and the completion time for each trial was recorded to analyze the learning curve.

Learning Curve Analysis. The analysis of task completion times across trials revealed important insights into participant performance and system usability. As shown in Fig. 5, participants demonstrated a clear learning effect, with progressively faster task completion times in subsequent trials. This trend highlights the learnability of the system, as most participants improved with practice.

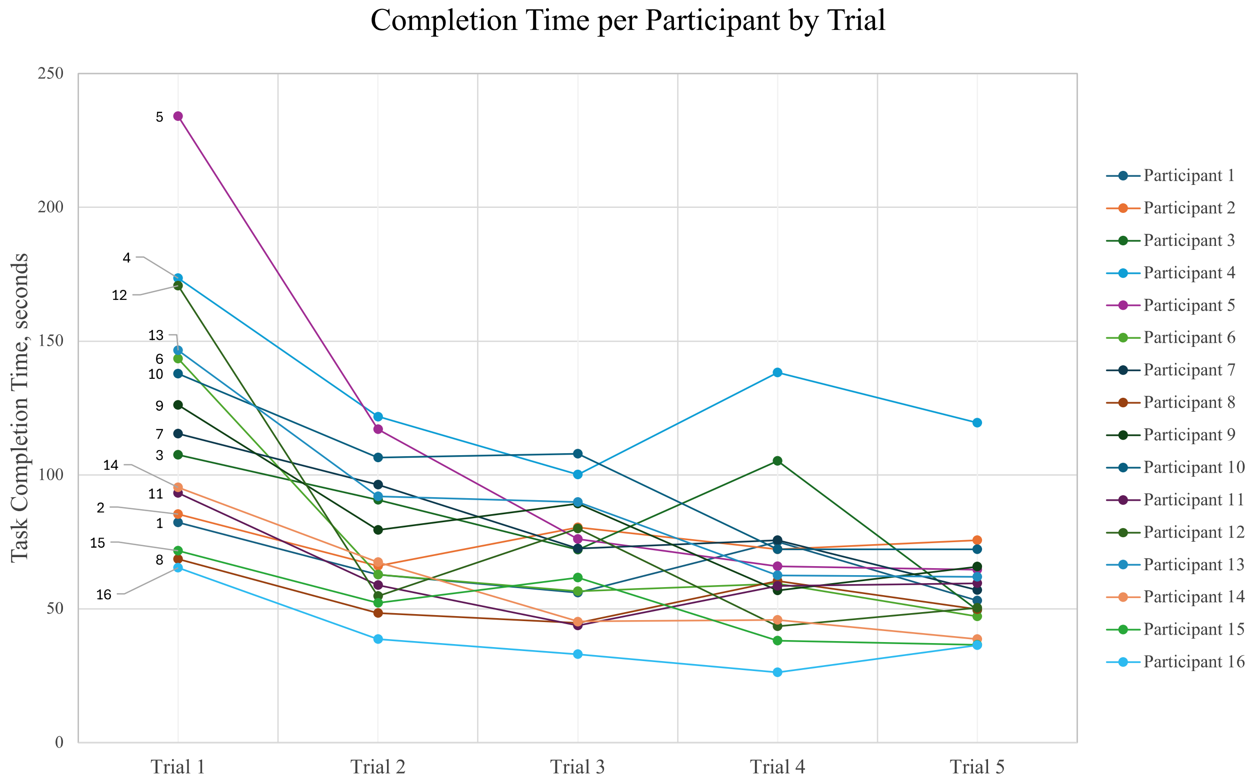


Fig. 5. Task completion time across five trials for all participants

However, performance fluctuations were observed in the third and fourth trials, where some participants demonstrated slightly degraded completion times. Participant feedback attributed this to two main factors: facial muscle fatigue, which made actions physically demanding, and overconfidence, where participants reported being less focused after feeling they had mastered the system.

Performance variability was also evident across participants. For instance:

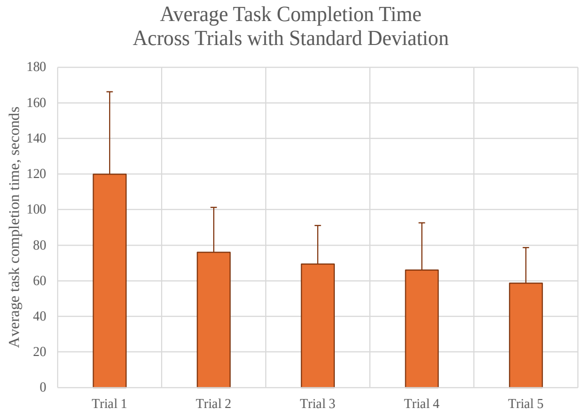
- **Participants 4 and 10.** Consistently higher completion times suggest these individuals faced challenges adapting to the system, potentially due to individual differences in learning rates or usability difficulties. These outliers may indicate areas for further investigation, such as refining the system to better accommodate diverse user needs.
- **Participant 16.** Consistently recorded the fastest times, reflecting a strong ability to adapt quickly to the interaction method.

By the fourth and fifth trials, many participants' completion times had stabilized or plateaued, indicating that they had reached a level of proficiency with the system. This plateau suggests that the system enables users to achieve consistent performance after sufficient practice. While the overall results demonstrate the system's usability and learnability, the observed variability across participants highlights potential areas for improvement. For example, a calibration process to personalize recognition thresholds could help address individual differences and ensure the system is equally effective for a broader range of users.

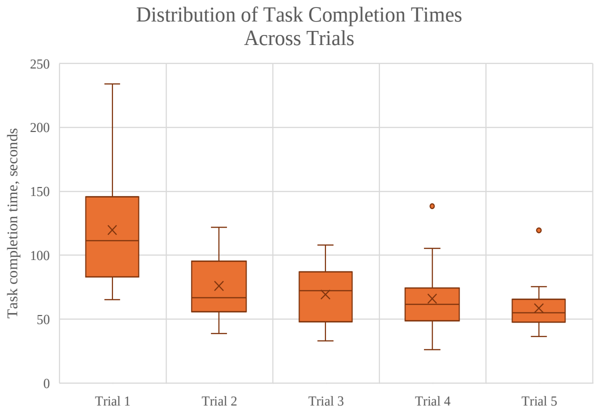
To assess the system's learnability, task completion times were examined across five consecutive trials. As shown in Fig. 6(a), the average completion time decreased consistently, with error bars indicating the standard deviation. This steady decline highlights a clear learning effect, as participants became increasingly proficient with the interaction method. The average completion time dropped from approximately 120 seconds in the first trial to 59 seconds in the fifth, representing a 51% improvement.

To validate this observed trend, a paired t-test was conducted to compare task completion times between the first and fifth trials. The results revealed a statistically significant decrease in completion times ($t(15) = 5.99, p < 0.0001$), confirming the presence of a strong learning effect.

Fig. 6(b) provides the distribution of task completion times across trials, showing the performance variability among participants. The data reveals a notable reduction in variability over time. This trend highlights improved consistency in participant performance, supporting the system's learnability.



(a)



(b)

Fig. 6. Learning curve analysis of task completion times across trials

Challenging Facial Expression Actions. Participants were asked to provide feedback on which facial expressions they found challenging to perform. The number of participants who reported difficulty with each specific action is shown in Fig. 7. To analyze these ratings, a Chi-square test was conducted, revealing a statistically significant difference in the perceived difficulty of actions ($\chi^2(7) = 18.0, p = 0.012$). This confirms that certain actions were disproportionately challenging for participants.

The most challenging action reported by participants was "Jaw Open + Brows Down", with 9 participants finding it difficult. This suggests that compound actions requiring the coordination of multiple expressions are more demanding. Conversely, simpler actions such as "Brows Up", "Brows Down", and "Smile Left" were reported as challenging by only one participant each, supporting their relative ease of execution.

"Smile Right" was reported as challenging by 3 participants, with one participant mentioning difficulty controlling the right side of their face due to a condition they've had since childhood, which prevents them from closing their right eye independently of their left eye. This highlights how individual physical differences can influence the performance of specific actions, underscoring the need for flexibility in system design to accommodate diverse users.

These results suggest that multi-step or compound actions, such as "Jaw Open + Brows Down", may require additional optimization to enhance recognition and ease of use. Addressing challenges related to asymmetrical facial control could further improve the system's usability and inclusivity.

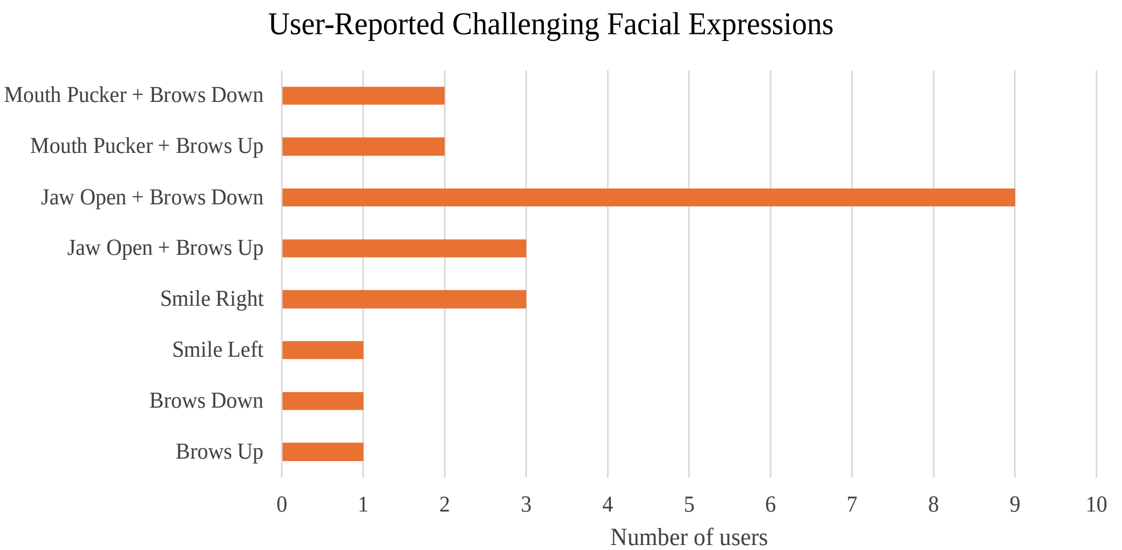


Fig. 7. Challenging facial expressions reported by participants (number of participants per action)

Participant Feedback and Observations. Overall, participants highly valued the suggested interaction method. They mentioned its potential usefulness in several contexts:

1. **Assistive technology.** As a hands-free interaction method, it may benefit people with upper limb disabilities or limited mobility.

2. **Facial exercise tool.** The system could be adapted into a gamified format for face exercises, particularly for individuals requiring rehabilitation or physical therapy.

3. **Supplementary shortcuts.** Participants suggested that facial expression actions could serve as shortcuts to complement traditional input devices, adding versatility to user interactions.

Discussion

This study demonstrated the usability and learnability of the proposed facial expression-based interaction method, highlighting its strengths and identifying areas for improvement. Participants showed a clear learning effect, with task completion times improving over trials, which is consistent with established usability principles that emphasize the importance of learnability in human-computer interaction (ISO 9241-11). However, some participants demonstrated slightly degraded performance during the third and fourth trials, which they attributed to facial muscle fatigue or overconfidence that led to reduced focus. These findings suggest that while the system is intuitive and easy to adopt, factors such as user fatigue and concentration must be addressed for prolonged use.

Advantages. One of the system's key advantages is its accessibility, as it relies on standard hardware, such as a webcam, making it a low-cost solution for hands-free interactions. This aligns with the criteria for accessible design, as it reduces the need for specialized equipment. The ability to define custom actions using facial expressions enhances its flexibility, allowing for diverse use cases. Participants highlighted its potential applications in assistive technologies, particularly for individuals with upper limb disabilities or limited mobility. Additionally, some participants saw value in gamifying the system for facial exercises or using it as a supplementary shortcut tool alongside traditional input devices. These findings underscore the versatility of the system, making it suitable for both accessibility-focused and recreational use cases.

Disadvantages. Despite these strengths, the study revealed several challenges. Recognition accuracy varied across participants due to differences in facial structures and movement, suggesting the need for a calibration process to replace the fixed threshold. This finding supports existing research on the importance of personalization in interaction systems to accommodate diverse user needs. Environmental factors, such as lighting conditions, also influenced system performance, highlighting a need for robustness in varied settings. Participants noted that complex, multi-step actions, such as "Jaw Open + Brows Down", were particularly difficult, while simpler actions, like "Brows Up", were easier to perform. Prolonged use was also associated with facial muscle fatigue, a known limitation in systems requiring continuous physical effort, which restricts the system's suitability for extended interactions.

Future work. Future work should address these limitations by introducing a calibration process to personalize recognition thresholds, as recommended in usability guidelines for adaptive systems. Optimization of challenging actions and enhancements to system robustness under varying environmental conditions are necessary to improve overall usability. Additionally, larger-scale studies are needed to validate the system across more diverse user groups and explore its long-term effectiveness. Participants' feedback suggests opportunities for further innovation, including integration into gamified tools or assistive technologies for broader accessibility.

Conclusion

This study explored the potential of a facial expression-based interaction method as a hands-free alternative for performing precision tasks. By leveraging standard hardware, such as a webcam, the system offers a low-cost, accessible solution for diverse use cases, including assistive technology and supplementary input methods.

The evaluation demonstrated significant learning progress, with participants reducing task completion times by an average of 51% between the first and fifth trials ($t(15) = 5.99, p < 0.0001$). Participants also highlighted the system's ease of use, with simpler actions, such as "Brows Up" and "Smile Left", identified as easy to perform, while complex multi-step actions, such as "Jaw Open + Brows Down", were reported as more challenging ($\chi^2(7) = 18.0, p = 0.012$). These results confirm the system's usability and learnability, while also identifying areas for improvement.

Feedback from participants underscored the system's value, particularly its potential to benefit individuals with upper limb disabilities or limited mobility, as well as its application in gamified tools for facial exercises. However, the study revealed several limitations:

- Variability in recognition accuracy due to differences in users' facial structures and movements.
- Fatigue during prolonged use, which may limit its effectiveness for extended interactions.
- Challenges with complex multi-step actions, such as compound movements, highlighting the need for improved design and usability optimization.
- Environmental factors, such as lighting conditions, affecting system reliability.

These findings emphasize the need for further refinement. Future work will focus on introducing calibration processes to personalize thresholds, optimizing challenging actions, and enhancing system robustness under varied conditions. Larger-scale studies involving diverse user groups are needed to validate the system's applicability and long-term effectiveness in various contexts.

In conclusion, the proposed interaction method demonstrates significant promise as an accessible, flexible alternative for hands-free interaction. This research contributes to the growing field of alternative interaction methods and supports the development of inclusive and adaptable human-computer interaction technologies.

References

1. J. P. Lewis, K. Anjyo, T. Rhee, M. Zhang, F. Pighin, and Z. Deng, 'Practice and Theory of Blendshape Facial Models', 2014, *The Eurographics Association*. doi: 10.2312/EGST.20141042.

2. B. Anand, B. B. Navathe, S. Velusamy, H. Kannan, A. Sharma, and V. Gopalakrishnan, 'Beyond touch: Natural interactions using facial expressions', *2012 IEEE Consumer Communications and Networking Conference (CCNC)*, pp. 255–259, Jan. 2012, doi: 10.1109/CCNC.2012.6181097.

3. Q. Chen, M. D. Cordea, E. M. Petriu, A. R. V. Koczy, and T. E. Whalen, 'Human Computer interaction for smart environment applications using hand gestures and facial expressions', *IJAMC*, vol. 3, no. 1/2, p. 95, 2009, doi: 10.1504/IJAMC.2009.026854.

4. S. Ren, 'Computer Vision for Facial Analysis Using Human-Computer Interaction Models', *Journal of Interconnection Networks*, vol. 22, no. Supp03, p. 2144005, May 2022, doi: 10.1142/S0219265921440059.

5. Y. Shi, Z. Zhang, K. Huang, W. Ma, and S. Tu, 'Human-computer interaction based on face feature localization', *Journal of Visual Communication and Image Representation*, vol. 70, p. 102740, Jul. 2020, doi: 10.1016/j.jvcir.2019.102740.

6. M. K. Chowdary, T. N. Nguyen, and D. J. Hemanth, 'Deep learning-based facial emotion recognition for human–computer interaction applications', *Neural Computing and Applications*, vol. 35, no. 32, pp. 23311–23328, Nov. 2023, doi: 10.1007/s00521-021-06012-8.

7. A. Samara, L. Galway, R. Bond, and H. Wang, 'Affective state detection via facial expression analysis within a human–computer interaction context', *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 6, pp. 2175–2184, Jun. 2019, doi: 10.1007/s12652-017-0636-8.

8. A. Condegni, W. Wang, and R. Li, 'A Digital Human System with Realistic Facial Expressions for Friendly Human-Machine Interaction', D.-S. Huang, P. Premaratne, B. Jin, B. Qu, K.-H. Jo, and A. Hussain, Eds., in Lecture Notes in Computer Science, vol. 14086. Singapore: Springer Nature Singapore, 2023, pp. 787–798. doi: 10.1007/978-981-99-4755-3_68.

9. F. French *et al.*, 'Expressive Interaction Design Using Facial Muscles as Controllers', *Multimodal Technologies and Interaction*, vol. 6, no. 9, p. 78, Sep. 2022, doi: 10.3390/mti6090078.

10. H. Gjoreski *et al.*, 'OCOsense Glasses – Monitoring Facial Gestures and Expressions for Augmented Human-Computer Interaction: OCOsense Glasses for Monitoring Facial Gestures and Expressions', *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–4, Apr. 2023, doi: 10.1145/3544549.3583918.

11. D. Rozado, J. Niu, and M. Lochner, 'Fast Human-Computer Interaction by Combining Gaze Pointing and Face Gestures', *ACM Transactions on Accessible Computing*, vol. 10, no. 3, pp. 1–18, Aug. 2017, doi: 10.1145/3075301.

12. A. Matos, V. Filipe, and P. Couto, 'Human-Computer Interaction Based on Facial Expression Recognition: A Case Study in Degenerative Neuromuscular Disease', *Proceedings of the 7th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion*, pp. 8–12, Dec. 2016, doi: 10.1145/3019943.3019945.

13. H. Mosquera, H. Loaiza, S. Nope, and A. Restrepo, 'Identifying facial gestures to emulate a mouse: Control application in a web browser', *2016 XXI Symposium on Signal Processing, Images and Artificial Vision (STSIVA)*, pp. 1–6, Aug. 2016, doi: 10.1109/STSIVA.2016.7743345.

14. A. Kumar Raja, C. Sugandhi, G. Nymish, N. Sai Havish, and M. Rashmi, 'Face Gesture Based Virtual Mouse Using Mediapipe', *2023 IEEE 8th International Conference for Convergence in Technology (I2CT)*, pp. 1–6, Apr. 2023, doi: 10.1109/I2CT57861.2023.10126453.

15. S. Zelinskyi and Y. Boyko, 'Facial Blendshapes for Custom Action: An Experimental Evaluation Using Facial Expressions for Human-Computer Interaction', *Modeling, Control and Information Technologies: Proceedings of International Scientific and Practical Conference*, (7), pp. 174–176, Dec. 2024, doi: 10.31713/MCIT.2024.050.

16. 'Face landmark detection guide', Google AI for Developers. Accessed: Sep. 06, 2024. [Online]. Available: https://ai.google.dev/edge/mediapipe/solutions/vision/face_landmarker

Serhii Zelinskyi **Сергій Зелінський**	PhD Student of the Department of Computer Engineering, Faculty of Radiophysics, Electronics and Computer Systems, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine, https://orcid.org/0009-0003-6271-2601 e-mail: serhii.zelinskyi@knu.ua Scopus Author ID: 59400981000	аспірант кафедри комп’ютерної інженерії факультету радіофізики, електроніки та комп’ютерних систем, Київський національний університет імені Тараса Шевченка, Київ, Україна.
Yuriy Boyko **Юрій Бойко**	Candidate of Physical and Mathematical Sciences, Associate Professor, Head of the Department of Computer Engineering, Faculty of Radiophysics, Electronics and Computer Systems, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine, https://orcid.org/0000-0003-1417-7424 e-mail: yuriyboyko@knu.ua Scopus Author ID: 24722552300	кандидат фізико-математичних наук, доцент, завідувач кафедри комп’ютерної інженерії факультету радіофізики, електроніки та комп’ютерних систем, Київський національний університет імені Тараса Шевченка, Київ, Україна.