International Journal of Human–Computer Interaction

ISSN: 1044-7318 (Print) 1532-7590 (Online) Journal homepage: www.tandfonline.com/journals/hihc20

Interaction Techniques for Taking Selfies: A Review

Shaowei Chu & Huawei Tu

To cite this article: Shaowei Chu & Huawei Tu (13 Jun 2025): Interaction Techniques for Taking Selfies: A Review, International Journal of Human–Computer Interaction, DOI: 10.1080/10447318.2025.2513582

To link to this article: https://doi.org/10.1080/10447318.2025.2513582

Published online: 13 Jun 2025.

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
https://doi.org/10.1080/10447318.2025.2513582

Interaction Techniques for Taking Selfies: A Review

Shaowei Chu[a] and Huawei Tu[b]

[a]College of Media Engineering, Communication University of Zhejiang, Hangzhou, Zhejiang, China; [b]Department of Computer Science & Information Technology, La Trobe University, Melbourne, Victoria, Australia

ABSTRACT

In this work, we present an overview of current selfie techniques and state-of-the-art selfie systems that aim to simplify and enhance the experience of capturing excellent selfies. We contribute a comprehensive framework that summarizes the process of selfie interaction, categorizing the diverse techniques employed at each phase of the selfie-capturing procedure. Following our review of selfie techniques, we offer recommendations for the design of selfie cameras, interaction methods, and evaluation criteria for selfie interaction techniques. Our reviews on selfie-related technologies hold the potential to enrich human-computer interaction research in this domain and can serve as a valuable resource for informing marketing strategies for both hardware and software selfie applications.

KEYWORDS

Selfie; self-portrait; natural user interface; photography; social networking service

1. Introduction

Selfies emerged with the widespread adoption of smartphones and the development of social networking services (SNS). In 2013, "selfie" was registered as a new word by the Oxford dictionary. It is defined as a self-portrait photograph, typically taken with a digital camera or camera phone held in the hand or supported by a selfie stick, usually shared via social media.

The world's first selfie was captured in 1839 by Robert Cornelius, utilizing a single-lens reflex (DSLR) camera, which required over a quarter of an hour to complete the exposure (Sachse, 1893). After that, with the updating and iteration of photographic equipment and the improvement of interaction methods, people began to use digital cameras, tripods, timers, remote controls, etc. for self-portrait photography. Over the years, various camera hardware and software have been developed to facilitate the capture of high-quality selfies with ease, enabling users to record their best moments effortlessly and share them instantly with friends and family (Casio, 2011; Chen et al., 2024; Chu & Tanaka, 2011, 2012, 2015; Kim & Lee, 2019; Mei-Chen & Hsiao-Wei, 2014; Okabe et al., 2006; Party-shot, 2010; Wang et al., 2015).

With the advancement of SNS, people have the opportunity to showcase their selfies on various platforms, including online social networks and circles, which allows them to share personal updates, foster communication, convey emotions, enhance their overall sense of well-being, and alter perceived life satisfaction (Chen et al., 2017; Chua & Chang, 2016; Hogan & Wellman, 2014). The sheer popularity of selfies is staggering. According to Photutorial's data, 92 million selfies were taken each day in 2022 (Broz, 2024; Chen et al., 2024). Furthermore, the prevalence of selfie-taking is particularly high among young adults. Over 82% of them have taken a selfie using smartphones, and over 90% post selfies on platforms like Facebook and Instagram (Pounders et al., 2016). This trend is further bolstered by the billions of likes and comments generated by selfie postings on SNS. These factors, combined, provide significant motivation for individuals to continue capturing and sharing selfies on SNS daily. This has sparked a meteoric rise in selfie culture, which has become an integral part of modern social media engagement. A notable milestone is the 2015 listing of Meitu, a company specializing in selfie photography, whose market value peaked at 127.9 billion USD.

With the advent of smartphones and the evolution of cameras towards miniaturization and high-quality imaging, an increasing number of selfie interactive technologies have emerged. For example, as early as 2011, Chu and Tanaka (2011) proposed to use directional motion gestures and air touch to control a Pan-Tilt-Zoom (PTZ) selfie camera. Compared to traditional button- or timer-based interaction, their technique had higher positive feedback in terms of selfie-taking efficiency and user preference, mainly because of the freedom of hand when preparing body poses. In their follow-up studies, they proposed the use of hand motion gestures, head nodding and shaking gestures, and facial gestures, such as mouth opening and raising eyebrow, to control a self-portrait camera of various parameters such as the aperture size, ISO setting, or shutter speed (Chu et al.,

CONTACT Huawei Tu H.Tu@latrobe.edu.au Department of Computer Science & Information Technology, La Trobe University, Melbourne Victoria 3086, Australia

2017a; Chu et al., 2017b; Chu & Tanaka, 2012, 2013a, 2013b, 2015). Results showed air gestures achieved higher interaction efficiency and user satisfaction than remote control and contact-based interaction methods. Recent studies have investigated the user-defined methodology (Wobbrock et al., 2009) to generate more appropriate gesture sets for taking selfies. For example, Peng and Xu (2020) found that "make a victory sign" and "swipe up" are the highest rated user preferred gestures to trigger the selfie camera shutter. (Mao et al., 2023) specifically compared panel and pie menus for air gestures in a photo booth condition. Their results indicated that the pie menu offered notable advantages by reducing muscular loads, achieving shorter task completion times, and delivering an overall more enjoyable user experience.

Selfie-supporting devices, include selfie sticks (Arif et al., 2017; DJI, 2015), gimbals (Chu et al., 2017b; Party-shot, 2010), drones (Ahmed et al., 2021), and AI robots (Wang et al., 2024), have significantly enhanced the flexibility and diversity of self-portrait photography, allowing individuals to capture a diverse array of selfies across various times and locations. For example, Arif et al. (2017) investigated the usability of different types of selfie sticks and gave recommendations such as a shorter arm, adding a counterweight to the handle and applying automatic angle adjustment apparatus to support selfie shooting. Joo et al. (2022) proposed an estimation of relative pose for cameras on selfie sticks for refocusing and stylization of selfie images. Using drones to take selfies is a new emerging trend. For instance, Jane et al. (2017) proposed such as T-shape and palm-out gestures for stop; beckoning gestures for summoning; and frame gestures for taking selfies. Ahmed et al. (2021) proposed that utilizing a user-provided selfie template, the drone can fly to the designated location to take a selfie automatically.

Over the past few years, advances in deep learning, large language models, and artificial intelligence, along with the maturation of natural language recognition and comprehension technologies, image editing tools (Wang et al., 2023; Zhang et al., 2019), aesthetic assistance (Fang et al., 2018; Hu et al., 2019), pose estimation (Liu et al., 2023), facial expression recognition (Liu et al., 2024), and image generation (Chen et al., 2024; Jin et al., 2023) techniques have greatly impacted the selfie capturing and quality of images.

These endeavors have culminated in a diverse array of technological aids that enhance self-photography. Both academia and industry have been vigorously engaged in exploring and refining the user experience and photo quality of selfies, ensuring that individuals can capture memorable and professional-looking self-portraits.

However, selecting suitable techniques to capture an aesthetically pleasing selfie remains challenging for the average user, often leading to a cumbersome and lengthy process with selfie cameras. For example, the most popular selfie-taking method of extended arms has the disadvantages of limited field of view, distorted facial shapes, and difficulty in maintaining a steady hand (Chen et al., 2024; Li & Vogel, 2017). It is hard to shoot full body selfies with extended arms or a selfie stick, and the unergonomic selfie stick would result in chronic fatigue (Arif et al., 2017). Gimbals and drones provide good framing and composition; however, they are not always available and have the limitations of high weight to carry and time-consuming to set up (Ahmed et al., 2021; Roudaki et al., 2016).

Existing studies have discussed selfie technologies with advantages and shortcomings. For example, Mao et al. (2023) and Peng and Xu (2020) reviewed gesture interfaces for taking selfies and suggested a user-defined gesture set for shutter control. Hu et al. (2019) investigated a number of approaches that measure the aesthetic score of facial attractiveness, such as using geometric features, angles, and lighting, for taking well-posed selfies. Fang et al. (2018) examined many selfie systems and suggested facial features, such as face size, face position, facial expressions, and head postures, as shutter control functions. Kim and Lee (2019) reviewed guide-assisted techniques for taking photos, encompassing triangular compositions and the rule of thirds, and analyzed both 2D and 2.5D visual guidance to facilitate the user in positioning the camera at the intended location and orientation. Faimau (2020) provided an overview of the most common theoretical approaches that have been used by researchers to understand the phenomenon of the selfie.

Although these studies provide an understanding of selfie-related techniques, they usually focus on limited aspects of selfie techniques. We are unaware of a systematic review of the full body of selfie interaction techniques such as camera positioning, framing, composition, parameter control and shutter control techniques. It is also unclear how these techniques are adopted in different contexts. These factors are all important in the full process of taking selfies (Chen et al., 2024; Faimau, 2020; Wang et al., 2023). Therefore, it is necessary to conduct a comprehensive review of existing selfie-related technologies, assessing their strengths and weaknesses, and exploring the crucial elements of selfie technology which could be beneficial for both end users and selfie camera designers.

Therefore, the main objective of this work is to identify and analyze relevant studies investigating how selfie interaction techniques are offered to optimize the process of capturing selfies, such as framing, composition, parameter control, shutter control in various contexts. Specifically, we identify three research questions below:

1. What are the currently available and appropriate techniques that individuals employ when capturing selfies?
2. What selfie interaction framework can researchers adhere to, and how can we customize a selfie procedure to fit within that framework?
3. What are the prevailing research trends and directions related to selfies?

To answer these questions, we did a literature review on interaction techniques for taking selfies. The primary contribution of this article is to clarify selfie camera interaction technologies by establishing a framework for selfie systems. Based on this framework, we delve into aspects including the inherent characteristics of selfie cameras, such as camera positioning, framing techniques, parameter control, and shutter control methods within the interaction techniques for selfie systems. Finally, we discuss and provide suggestions for the design of selfie cameras, interaction design, and the evaluation method relating to selfie techniques. This review serves as a definitive reference for both the research community and the industry, offering a directional guideline for future research endeavors and technological applications.

2. Method

We followed the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) method (Moher et al., 2009) to conduct a systematic review. Since PRISMA guidelines were originally developed for health care and may not be directly suitable for HCI studies (Rogers & Seaborn, 2023), particularly for interaction techniques for taking selfies, we modified the PRISMA method to make it suitable for HCI studies. We excluded participant analysis since patients and diseases are special for medical research but not for HCI. Likewise, the risk of bias in individual studies, and meta-analysis is not included in our study.

As the purpose of the study is to catalogue various selfie interaction techniques and uncover the foundational framework of selfie interaction systems, we have structured the results section of the PRISMA into distinct, dedicated subsections: Framework of Selfie Interactions, Interaction Techniques, User Experience in Taking Selfies, and Real-World Cases or Techniques. This segmentation aims to provide a comprehensive presentation of our findings.

2.1. Literature search

We primarily used keywords such as selfie(s) and self-portrait(s) to search for related studies in the digital libraries of ACM, IEEE, Google Scholar, and DBLP. Once a relevant paper was identified, we continued our search using its references and citations, iterating this process until no new papers were discovered. The criteria for excluding papers were: (1) social science papers without technical development and implementation; (2) papers only focusing on image processing of selfie photos.

In addition, we employed search engines like Google and Bing to locate technological articles related to selfies, particularly focusing on selfie-oriented products such as the Casio TR series, Oppo N series selfie cameras, and various selfie accessories like selfie sticks, drones, and selfie booths.

Through this process, we ultimately identified a total of 27 papers related to selfie interaction technology. Table 1 summarizes the current representative selfie studies (N = 27, years from 2011 to 2024), along with an overview of their respective application characteristics, arranged in ascending order based on the year of introduction.

2.2. Framework of selfie interactions

According to Faimau (2020) and Shah and Tewari (2016), selfie can be divided into three parts: front-facing mobile cameras, photo-editing software, and multiple social media platforms. And based on the above literature collection, and the photography and interaction technology provided (Rand & Meyer, 2014), we further break down the camera-related part into camera positioning, previewing, composition, parameter control, and shutter control for the detail review.

Therefore, the comprehensive framework for the process of selfie interaction encompasses eight distinct steps, as depicted in Figure 1. These steps include camera positioning, previewing the frame, composition, parameters control, shutter control, selfie image editing, sharing selfies on social media platforms, and receiving feedback or motivation that inspires the next selfie session. This schematized process is often repeated until a sufficient result is obtained, which requires to focusing multiple aspects at once.

In this work, we primarily focus on interaction techniques, specifically reviewing the preceding five sections: camera positioning, preview, composition, parameter control, and shutter control. As selfie image editing is primarily associated with image processing, it falls outside the scope of our review. Similarly, the posting of images on social networking sites is also excluded from our review. Next, we present a detailed review on the five interaction aspects of the framework separately.

3. Interaction techniques

Selfie photography encompasses distinct steps, requiring selfie takers to navigate a series of operations independently. While adjusting various camera parameters, they must simultaneously consider their positioning, composition, pose, facial expression management, and eye contact within the frame. This introduces challenges for interaction design. Consequently, selfie takers must not only possess a grasp of photography skills but also leverage selfie assistance technology to enhance the quality of their selfies. This has led to the development of more convenient, efficient, and high-quality technologies for selfie photography. In the subsequent sections, we undertake a comprehensive review of the interaction techniques.

3.1. Camera positioning

The positioning of the camera holds immense significance in capturing a compelling selfie, encompassing critical elements such as the imaging angle, the proximity to the face, and the overall composition. Given the distinctiveness of selfies, which requires the selfie taker to exercise independent control over the camera and precisely maintain the facial position within the frame, the flexibility and ease of adjusting the camera's placement become paramount in achieving a successful selfie. Listed below are typical techniques for positioning the camera during a selfie session.

3.1.1. Extended arm

This method entails the selfie taker positioning the camera securely in their hand, extending their arm—typically with one arm raised at a 45° angle—and aiming the camera lens directly at themselves, as illustrated in Figure 2. Once

Table 1. Representative studies of selfie taking systems in recent years.

Techniques	Interactions	Key technologies	Devices	Characteristics
Gesture self-portrait (Chu & Tanaka, 2011)	Air motion gestures, air touch	Image recognition	PTZ camera, large screen	Air shutter, camera pan & tilt control
Head gesture (Chu & Tanaka, 2012)	Head motions	Image recognition	DSLR camera, tripod	Zoom control
Head gesture (Chu & Tanaka, 2013a)	Head motions, mouth motions	Image recognition	DSLR camera, tripod	Zoom control, mouth opening gesture
Yes, right there (Lo et al., 2013a, 2013b)	Voice prompts	Image recognition, IMU sensors	Smartphone rear camera	Guidance while no camera preview
Motion gesture (Chu & Tanaka, 2013b, 2015)	Air motion gestures	Image recognition	DSLR camera, tripod, projector	Gesture menu interface for various controls
Jump selfie 2014 (Garcia et al., 2014)	Automatic	Image recognition, face detection	Smartphone	Take jumping selfies
Selfie drone (Chen et al., 2015)	Smartphone remote control	Drone control	Drone, AR tag	Compose and framing control of a drone
Stitched groupies (Lu et al., 2016)	Touch screen	Image stitching	Photo Kiosk	Group selfies over temporal and spatial distances
Multi-Photo Fusing (Xia et al., 2016)	Smartphone takes 3 photos	Image stitching	Smartphone	Improve selfie FOV
Portrait aesthetics (Li & Vogel, 2017)	Visual guidance	Image recognition	Smartphone	Aesthetic rating based on face size, position, and lighting direction
LORA (Ngo & Kim, 2017)	Automatic	Image recognition	Sony Super-zoom camera	Take long-range selfies from a distance
Pan & tilt selfie system (Chu et al., 2017b)	Air motion gestures	Image recognition	Pan & tilt web camera	Gesture control of camera pan & tilt
Motion gesture (Chu et al., 2017a)	Air motion gestures	Image recognition	Oppo N1 smartphone with rotating camera	Waving and circling gesture for camera control
Head pose (Fang et al., 2018)	Visual guidance, voice prompts	Image recognition, facial landmarks	Smartphone	Accurate headshot using extended arm
PicMe (Kim & Lee, 2019)	Others' help, Visual guidance	Image recognition	Smartphone	Composition guide from requester
Cake Me (Lusk & Jones, 2019)	Others' help, Visual guidance	Image recognition, IMU sensors	Smartphone	Composition guide from requester
Virtual Portraitist (Hu et al., 2019)	Voice prompts	Image recognition, facial features	Smartphone	Suggestions on the angle of the face
Garuda (Virk & Dhall, 2019)	Visual guidance	Image recognition	Smartphone	Capturing selfies safely
Multi-person (Lin & Yang, 2020)	Touchscreen, AR	Image recognition	Smartphone	Two persons group selfie
SelfieDroneStick (Alabachi et al., 2020)	Smartphone taking a selfie	Image recognition, IMU sensors	Smartphone, Drone	Control a drone takes selfies from a preset posture template
Unselfie (Ma et al., 2020)	Automatic	Image recognition	Smartphone	Selfie photo transform to move raised arms
Forward Selfies (Trenz et al., 2021)	Touch screen	Image recognition	Smartphone, RGB-D camera	Combine photos of the front and the rear camera
Selfie drone (Ahmed et al., 2021)	Smartphone remote control	Image recognition	Drone	Control a drone takes selfies from a preset posture template
Virfie (Im et al., 2022)	Mouse	Image recognition	Desktop Computer, Webcam	Group selfies for video conference
Camera pose estimation (Joo et al., 2022)	Selfie stick	Image recognition, Camera pose estimation	Smartphone, selfie stick	Image refocusing and stylization
Total selfie (Chen et al., 2024)	Touch screen	Image recognition, image generation	Smartphone	Generate full-body selfies from a set of selfie images

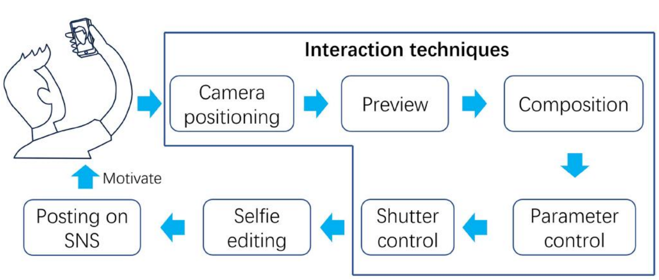

Figure 1. The interactive framework for selfie cameras. Within this framework, receiving likes, comments, and other feedback after posting selfies on social media further motivates her to capture more selfies, thus achieving a logical closed loop.

aligned, they press the shutter button to capture the moment. This approach is not only the most straightforward and expedient way to capture a selfie, but it also marks the seminal form of selfie interaction (Bevan, 2017).

The resulting images exhibit distinctive characteristics, notably a prominent head occupying a significant portion of the frame, with elements of the background scenery seamlessly incorporated. This style of selfie photography is ubiquitous, particularly during travels and social gatherings.

To address crucial factors such as the size, positioning, and angle of the head within the frame, Li and Vogel (2017) introduced a method that guides the movement of the arm to precisely adjust the camera's position. This, in turn, facilitates the refinement of the face size and head orientation. This guidance is achieved through the utilization of radial arrows strategically positioned on the face, which provide intuitive cues on how to manipulate the distance and positioning of the camera in the hand. Similarly, Fang et al. (2018) devised a graphical and voice-guided interface that instructs users on how to adjust their head orientation effectively while extending their arm for a selfie.

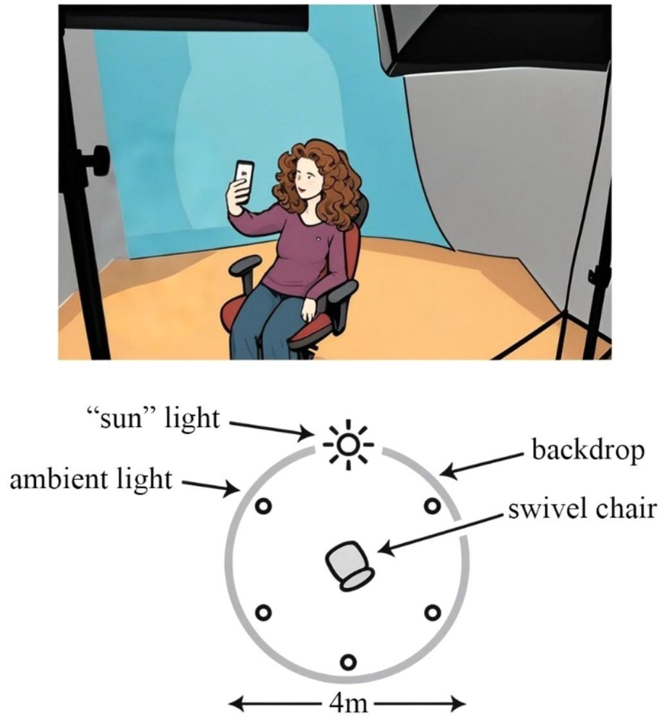

Figure 2. Using extended arm to take selfies. The lower part describes the controlled environment for user study. This image has been redrawn based on the original as presented in Li and Vogel (2017).

3.1.2. Selfie stick

This approach entails mounting a camera onto a selfie stick and elevating it to a certain height to align the camera lens with the selfie taker's perspective. This technique allows for capturing richer background details in the frame. Typically, selfie sticks incorporate buttons to trigger the camera's shutter, facilitating photo capture (Arif et al., 2017; Bevan, 2017). This method surpasses the traditional arm-extended selfie, offering greater flexibility regarding angles and the distance between the camera and the face. Furthermore, it enables capturing expansive background scenes and even group selfies, introducing more diversity and creativity to selfie photography. Additionally, selfie sticks with anti-shake technology, like the DJI Osmo (DJI, 2015), feature motorized gimbals, providing advanced functionalities like subject tracking.

Currently, the market boasts a diverse range of selfie sticks from brands such as Xiaomi, Huawei, GoPro, and Insta360. The Selfie Stick Market size was estimated to be USD 627.30 million in 2023, rising to USD 648.22 million in 2024, and is forecasted to grow at a Compound Annual Growth Rate (CAGR) of 3.42%, reaching USD 794.25 million by 2030 (Patel, 2023).

Ergonomics, a pivotal aspect in the usability of selfie sticks, has garnered significant research attention. Arif et al. (2017) analyzed the usability of various selfie sticks and discovered that the combined weight of the stick and camera, leading to arm fatigue, is a critical factor for users, as shown in Figure 3. Lee and Kosasih (2015) outlined the design considerations for selfie sticks, highlighting the importance of adjustable lengths to accommodate open-space scenarios, as most selfies are captured in such environments. In essence, the key factors that influence selfie stick photography encompass adjustable length, portability, Bluetooth connectivity, and anti-shake functionality.

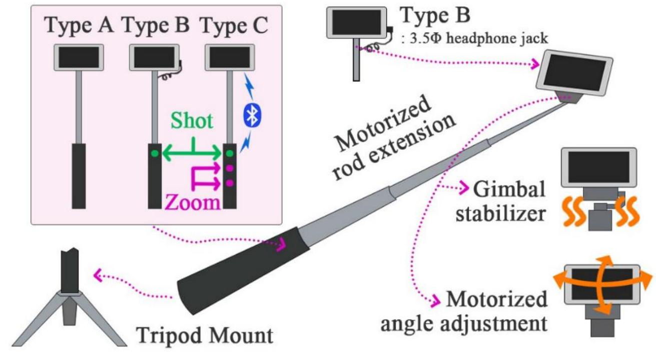

Figure 3. The three common types of selfie stick. This image has been redrawn based on the original as presented in Arif et al. (2017).

3.1.3. Tripod

This method involves mounting the camera on a tripod, which can either be set up on the ground or a platform, or be a hanging-style tripod such as an octopus tripod. Then, the camera's countdown function or a remote control is utilized to capture the selfie. Currently, many selfie sticks, like the DJI Osmo (DJI, 2015), have handle sections that can be unfolded into tripods. Typically, the selfie-taker needs to preset the camera parameters, activate the countdown shutter, and then run to the predetermined position in front of the camera to wait for the countdown to complete the selfie.

The advantage of this method is that it frees up both hands, allowing for almost unlimited posing for selfies. It even enables creative selfies, such as capturing jumping selfies (Garcia et al., 2014). However, the disadvantage is that since the photographer cannot see their actual position in the frame while setting up the camera, they often need to repeatedly run back and forth multiple times to capture an ideal selfie. Especially when using a DSLR camera, due to the lack of dynamic autofocus for the lens, as shown in Figure 4, a stand-in is often needed to assist in focusing, which can be time-consuming and laborious (Chu & Tanaka, 2013b). Additionally, tripods tend to be bulkier and less portable compared to selfie sticks, thus they are less frequently used in selfie photography interactions.

3.1.4. Gimbals

This method involves mounting the camera on an electric gimbal, which can automatically rotate to capture a person's face, ensuring the accuracy of the selfie. Typically, gimbal cameras can be combined with automatic shutter activation features such as face detection, smile recognition, and expression recognition to trigger the shutter for photography. For instance, the DJI Osmo Mobile (DJI, 2015) realizes person tracking to ensure that the camera captures the face in real-time, guaranteeing the accuracy and stability of the selfie. Similarly, the Sony Party-shot (Party-shot, 2010) is equipped with a pan and tilt gimbal that tracks and detects smiles to automatically capture faces during parties. As

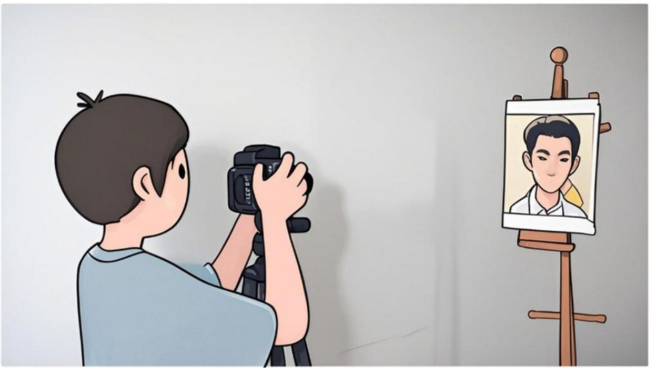

Figure 4. Place the camera on a tripod for capturing selfies. To enable the camera to autofocus, the user deploys a stand-in. Once the autofocus is achieved and the camera's self-timer is initiated, the user swiftly moves to the front of the camera, replacing the stand-in and assuming their desired pose. This image has been redrawn based on the original as presented in Chu and Tanaka (2013b).

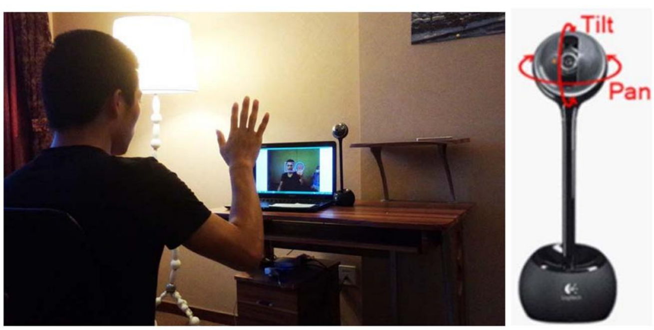

Figure 5. PTZ camera selfie system. Hand gestures are used to control camera functions. The figure is used with permission from IEEE license (Chu et al., 2017b).

shown in Figure 5, Chu et al. (2017b) proposed using gestures to control a PTZ (Pan-Tilt-Zoom) camera in a desktop environment. Another example is the movable PTZ camera designed by Roudaki et al. (2016), which utilizes sound wave recognition and positioning to locate and capture speakers in a meeting in real-time.

The advantage of using a gimbal is that the camera can automatically track users and capture selfies even when the selfie-taker is not looking directly at the camera. This is particularly useful during family gatherings, where the selfie gimbal can conveniently capture the natural expressions of all family members, enabling the capture of group selfies as well.

3.1.5. Drone

This method uses a drone to track and shoot around oneself. One significant advantage is that drones are not limited by distance or angle, allowing for facial close-ups, full-body shots, or capturing people from a long distance within the scene. Drones are considered a promising platform for capturing selfies. Typically, selfie-takers utilize smartphone applications and joystick-based controllers to manipulate the drone. For instance, Chen et al. (2015) designed a system that utilizes a smartphone to preview and control the drone for selfie photography. Similarly, Ahmed et al. (2021) and Alabachi et al. (2020) proposed methods to enable drones to automatically recognize human bodies and shooting angles for selfies.

Due to the unique nature of selfies, it is undesirable to have the remote controller restricting both hands or interfering with eye contact with the camera lens. Therefore, drone vision-based gesture recognition technology presents itself as a promising approach. For instance, as illustrated in Figure 6, Jane et al. (2017) leveraged user-defined interactions to design gestures such as Palm out, Beckon, T-shape, and others to control the drone's movement (closer, further, hover), as well as a framing gesture for capturing a selfie.

3.1.6. Others' help

This method involves handing over the camera to a companion or stranger nearby to take a picture for oneself, often for souvenir selfie photos at tourist attractions. However, it can be challenging for others to fully understand one's specific shooting needs, such as composition, resulting in significant uncertainty when asking someone else to take the photo. To address this issue, as illustrated in Figure 7, Kim and Lee (2019) proposed an assisted composition technique that utilizes Canny's edge detector to calculate image contours and framing guides, matching pre-captured selfie composition photos to guide others in taking pictures according to the selfie-taker's preferences. Similarly, Lusk and Jones (2019) introduced Cake Cam, a method that calculates the camera pose using Apple's iOS AR Kit to determine the image composition, enabling the person assisting in taking the photo to accurately capture the selfie-taker's desired composition.

3.2. Previewing the frame

During the process of taking a selfie, the selfie-taker needs to observe their position, camera angle, pose, facial expression, etc., in the camera's viewfinder in real-time. At the same time, the preview screen should be as close to the lens as possible to ensure eye contact. Therefore, designing the real-time preview function for selfie cameras is a crucial aspect of selfie photography. Below is a summary of the related technologies.

3.2.1. Front camera

Utilizing the front camera of a smartphone and capturing photos in real-time through the screen preview has become the most prevalent method for selfies. Holding the phone, as shown in Figure 2, selfie taker can instantly observe their appearance in the lens, like gazing into a mirror. This approach offers both convenience and precision, significantly elevating the quality of selfie photography (Fang et al., 2018). Moreover, leveraging the phone screen's luminous properties, selfie takers can utilize it to illuminate their faces during selfies, achieving a professional effect resembling that of a flashlight. Apple's iPhone X, released in 2017, introduced a 3D structured light front camera, revolutionizing mobile selfie photography by enabling 3D face modeling and recognition. For instance, Bao et al. (2017) harnessed

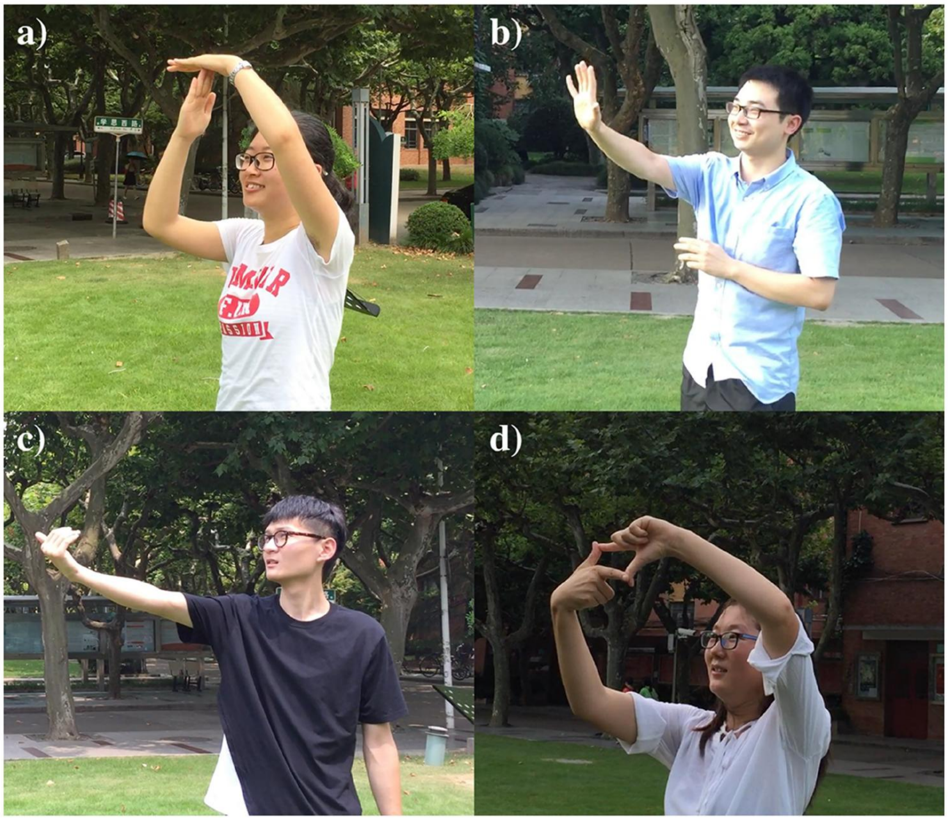

Figure 6. Interacting with a drone. (a) T-shape gesture for stop; (b) Palm out stop gesture; (c) beckon; (d) frame gesture for taking photos. The figure is used with permission from authors (Jane et al., 2017).

the 3D camera of the iPhone X to capture high-fidelity 3D digital human head selfies. Currently, nearly all smartphones feature increasingly high-resolution front cameras, catering to the diverse selfie demands of users.

3.2.2. Flip camera

The camera can be manually or automatically rotated to align with the screen's orientation, allowing the screen to serve as a real-time preview. For instance, Casio's TR series of rotating camera selfie cameras, released in 2011 shown in Figure 8, featured a lens and body frame that could rotate to offer various holding options or be hung on a wall for real-time preview and photography (Casio, 2011). Similarly, Oppo introduced the N series rotating camera smartphone in 2013, where the camera can rotate freely between the front and rear positions, providing a new approach for high-definition selfies using the front camera (Chu et al., 2017a; Oppo, 2013).

Likewise, the preview screen can be folded or rotated to align with the camera's orientation, making it convenient to observe the preview image from the front of the lens. This tilting screen design is commonly seen in professional photography equipment, such as Canon's EOS series of DSLR cameras.

3.2.3. Remote screen

This method entails wirelessly connecting a remote screen, such as a smartphone, to a camera to facilitate real-time transmission of the camera's preview. However, a significant limitation arises from the requirement for the selfie taker to simultaneously hold a remote-control device and observe the device's screen, often constraining the freedom of posing. For instance, in a comparative study conducted by Chu and Tanaka (Chu & Tanaka, 2013a), it was found that the utilization of remote preview significantly hindered posing freedom due to the need to occupy both hands, as illustrated in Figure 9. Furthermore, the inability to maintain eye contact with the camera while observing the remote screen impedes the user's ability to accurately assess the selfie's effect. These challenges contribute significantly to the limited adoption of remote preview methods in selfie photography.

3.2.4. Mirror

A selfie taker stands in front of a mirror, utilizing the mirror to preview their pose and expression, and then captures a photo of themselves reflected in the mirror. The advantage of this method is that the mirror provides a 1:1 preview, allowing even an estimation of the selfie taker's body size based on the pixel ratio of the phone held in the mirror's reflection (Sheth & Srivastava, 2019). However, as the selfie taker typically needs to align the phone's camera directly towards the mirror for the shot, it is inevitable that the phone will appear in the selfie photograph.

3.2.5. Large screen

This method typically involves establishing a dedicated selfie space, such as a selfie booth or a photography studio, where a large screen displays the camera's preview, enabling selfie

Figure 7. PicMe: Ask others to take selfie pictures. A requester captures a photo composition and delivers it to the photographer (user), the user grasps the target image roughly, aligns the current camera frame and a virtual frame, and adjusts the detail elements. The figure is depicted in accordance with the description provided in Kim and Lee (2019).

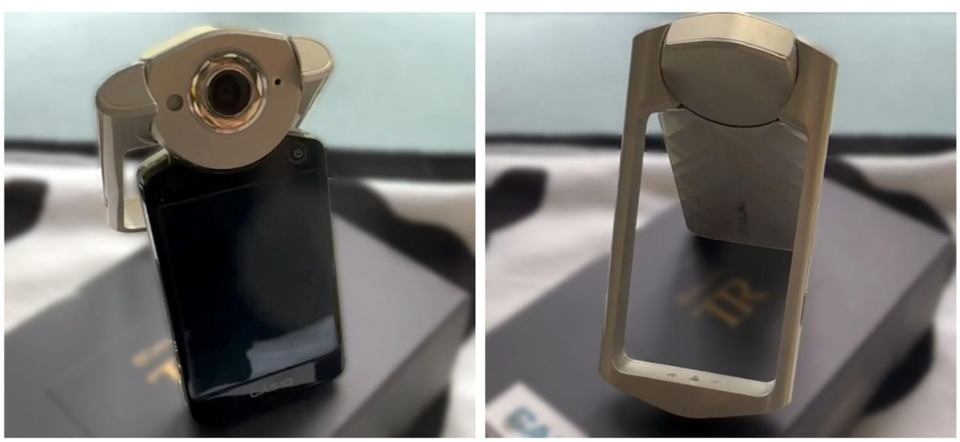

Figure 8. Casio TRYX camera, allowing the frame to be used as a stand and for the screen to point in all kinds of random directions (Casio, 2011).

takers to capture even the most nuanced facial changes. Originating in Japan in the 1990s, selfie booths quickly gained popularity as an entertainment trend, offering users a private enclosure equipped with a camera and a large screen. On this screen, users can preview the camera's view in real-time, adjust settings like beauty enhancements and facial AR effects through an intuitive touchscreen interface, and trigger a self-timer by pressing a button to capture their selfie (Okabe et al., 2006).

Nowadays, selfie booths are not solely confined to entertainment venues; they have also become a standard fixture in government offices, airports, and other public spaces, catering to the need for convenient ID photo capture (Dugan et al., 2015). However, the primary limitation of selfie booths lies in the fixed distance between the user and the camera, restricting them primarily to capturing headshots.

On the other hand, in professional settings like photography studios and laboratories, utilizing large screens or

Figure 9. Using a smartphone as the remote controller to preview the frame of the DSLR camera. (a) Selfie system; (b) smartphone as remote controller. The figure is used with permission from authors (Chu & Tanaka, 2013a).

projected previews offers selfie takers an expanded viewfinder, enabling them to observe finer facial expressions and poses for more professional close-up and full-body selfies, as shown in Figure 10 (Chu & Tanaka, 2011, 2013b, 2015). As the large screen preview tends to captivate the user's attention, visual cues are often displayed on the screen once the shutter countdown is activated, guiding the selfie taker to redirect their gaze towards the camera lens.

3.2.6. No preview

When using the rear high-definition camera of a smartphone for selfies or in special scenarios such as drone selfies, selfie takers rely on the camera's visual algorithms for automatic positioning, framing, and capturing, without the aid of any preview screen. For instance, as shown in Figure 11, Lo et al. (2013a) designed "Yes, Right There!" application that utilizes the rear camera of a smartphone for selfies, allowing users to adjust the head's position in the frame through voice prompts to complete the selfie shot.

3.3. Composition

Once the camera position and preview mode are established, the subsequent step is to determine the optimal self-portrait composition for an appealing selfie. Composition holds a pivotal role in photography, as a well-executed composition highlights the selfie taker's distinctive personality and aligns with visual aesthetic principles. The techniques related to composition include methods recommended by classic self-portrait composition experience, real-time assisted composition technology, and AI-driven automatic capturing capabilities.

3.3.1. Photography techniques

In portrait photography, there are numerous techniques that can be directly referenced for selfies, such as shooting from a top-down angle to make the face appear smaller, positioning the face at the center of the frame for a visually impactful effect, placing the face at the one-third position to enhance the transparency of the image, and utilizing the rule of thirds composition (Rand & Meyer, 2014). Posing App (Mendo, 2024) is a mobile application that compiles classic portrait photography compositions and corresponding selfie pose diagrams, such as the classic poses of scissors hands, chin resting, and eye-covering, to assist users in obtaining shooting guidance and creative inspiration for selfies. However, these methods often fail to provide real-time shooting assistance, requiring users to select and learn a specific composition independently and then imitate it to complete the selfie.

3.3.2. Intelligent assistance

This method applies image processing and recognition algorithms to capture the position of the subject in the frame, match classic selfie patterns, and assist the selfie taker in adjusting the camera position, angle, as well as their own position and pose in the frame, thus guiding users in capturing ideal selfies in real-time. There are two mainstream approaches. The first one considers both the subject and the background in composition, which often requires the help of others. For example, methods that involve finding someone to operate the camera can match pre-shot compositions to guide the operator in capturing the selfie according to the selfie taker's preferences (Kim & Lee, 2019; Lusk & Jones, 2019; Xu et al., 2014). Alternatively, basic photography principles such as central and golden-ratio composition can be utilized to aid in composition (Balata et al., 2015). Similar approaches can also achieve creative selfies such as double exposures or multiple selves in a single photo (Lin & Yang, 2020; Lu et al., 2016). Additionally, in a human position recommendation method, composition rules are learned from human-scenery aesthetic datasets (Wang et al., 2015; Xu et al., 2014).

The second approach focuses specifically on headshots for selfies. It involves building a selfie dataset and utilizing expert or crowdsourcing scoring to identify outstanding selfie patterns. Then, during the shooting process, an

Figure 10. Air gesture menu interface for taking selfies. The figure is used with permission from authors (Chu & Tanaka, 2015).

Figure 11. Yes, Right There! Prevent faces from being cut out of the camera frame by giving suggestions to users until they are in a suitable position in the frame. The figure is used with permission from authors (Lo et al., 2013b).

algorithm estimates the head posture score to guide the capture. For instance, Fang et al. (2018) used crowdsourcing to score the attractiveness of head postures and camera distances in a constructed selfie dataset. Based on facial feature states, it provided visual guidance and voice prompts to achieve good composition selfies, as illustrated in Figure 12. Li and Vogel (2017) considered factors such as face size, face position, and lighting direction in composition. They established aesthetic rating criteria for selfie compositions based on these three factors using crowdsourcing scores from a selfie dataset. During selfies, it can evaluate the score in real time and guide users to adjust the composition to achieve a high aesthetic score. Similarly, Hsieh and Yeh (2017), Hu et al. (2019), Mei-Chen and Hsiao-Wei (2014), and Yeh and Lin (2014) focused on the angle of the face. They trained a model using a selfie dataset to pattern the size and spatial arrangement of facial features, automatically estimating the head pose aesthetic score based on the face angle, and suggesting pose adjustments when taking a selfie.

In addition, there are assistant technologies that target aspects like grid guidelines, environmental lighting, and avoiding clutter in the background. For instance, Jane et al. (2020) proposed a saliency-based algorithm for selecting composition grids that aid in capturing more well-composed images. This method differs from traditional composition gridlines by calculating adaptive armatures in real time to provide shooting

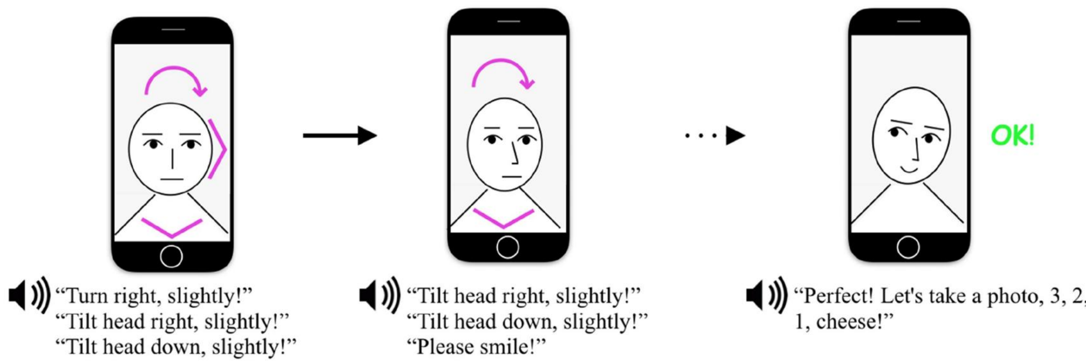

Figure 12. Visual and voice-guided instructions for the common user to frame and pose, that to achieve satisfying selfies. The figure is used with permission from authors (Fang et al., 2018).

references and inspire new composition ideas. Jane et al. (2019) focuses on environmental lighting during shooting, calculating the brightness and angle changes of facial lighting in real time to guide the composition of facial angles. Jane et al. (2021) introduced a decluttering technique to remove unwanted objects in the frame. It utilizes image edge detection algorithms to highlight object edges and guides the user to switch camera angles or compose the shot to block or remove clutter from the frame.

3.3.3. Automatic capturing

This method is typically implemented using devices like gimbals or drones, leveraging technologies like image-based face detection and facial expression analysis, along with advanced intelligent composition techniques, to automate portrait photography. For instance, in Party-shot (2010), intelligent portrait capturing at parties was achieved through face and smile recognition. Additionally, Roudaki et al. (2016) introduced a movable PTZ camera that utilizes sound wave recognition and positioning to precisely locate and capture speakers during conferences in real time. Furthermore, Garcia et al. (2014) developed a method for capturing jumping selfies, as illustrated in Figure 13, which predicts the peak of the jump and the optimal photo-taking moment based on the trajectory detected through image recognition. The key benefit of intelligent capturing lies in its ability to capture natural expressions and postures of individuals, as well as diverse facial angles, including side profiles.

3.4. Parameter control

Parameter control is concerned with adjusting settings such as camera lens zoom, ISO sensitivity, aperture, shutter speed, and selfie filters during self-portrait photography. In summary, the main parameters are: (1) Zoom: Switching between facial and full-body shots, as well as switching to wide-angle lenses. Wide-angle lenses can capture more background and are widely used in travel selfies. (2) Brightness: Adjusting the lens sensitivity or using the phone's screen as a fill light during photography. (3) Filters: including facial smoothing, face slimming, facial feature enhancement, makeup effects, and adjusting the intensity levels of these filter effects. (4) Augmented Reality (AR) filters: Adding features like glasses, head accessories (stickers), cartoon face swaps, distortion effects, ambience, and more. Controlling these parameters requires the selfie-taker to face the camera and make adjustments in real-time while viewing the preview screen. The mainstream methods of interaction include touchscreen controls, air gestures, and facial expression interactions.

3.4.1. Touchscreen interface

When capturing selfies by extending their arm and utilizing the front-facing camera of a smartphone, users can conveniently navigate the touchscreen interface with their thumb, seamlessly adjusting settings like zoom ratio, ISO sensitivity, and selecting various photo filters via intuitive virtual buttons. Furthermore, in the traditional selfie booth setup, large-screen touchscreen interaction emerges as the preferred and highly convenient method of user interaction. This touchscreen interface offers an array of virtual buttons, enabling users to easily zoom in and out, manage the flash, customize filters, and apply attractive AR effects.

3.4.2. Air gesture user interface

Air gesture interaction leverages the image recognition capabilities of cameras to identify users' hand gestures in mid-air, offering a non-contact, distance-based interaction method that is ideal for selfie photography (Chu & Tanaka, 2011; Peng & Xu, 2020). Chu and Tanaka (2011) first proposed to utilize the optical flow method to recognize dynamic hand gestures in four cardinal directions: up, down, left, and right. This enables precise control of the PTZ selfie camera's panning and tilting for optimal framing. Zooming capabilities have also been enhanced, allowing users to manipulate the camera's zoom function through recognition of continuous nodding and shaking motions of the head, as illustrated in Figure 14, ensuring an ideal face size for the composition (Chu & Tanaka, 2012, 2013a).

Research has demonstrated that the gesture-based pie menu in a photo booth setting significantly reduces muscular strain, speeds up task completion, and provides a more enjoyable user experience, making it the preferred choice for mid-air gesture interaction in such environments (Mao et al., 2023). By integrating eight-directional motion gestures with a pie menu interface, as shown in Figure 15, real-time adjustments to the DSLR selfie camera's aperture size, ISO,

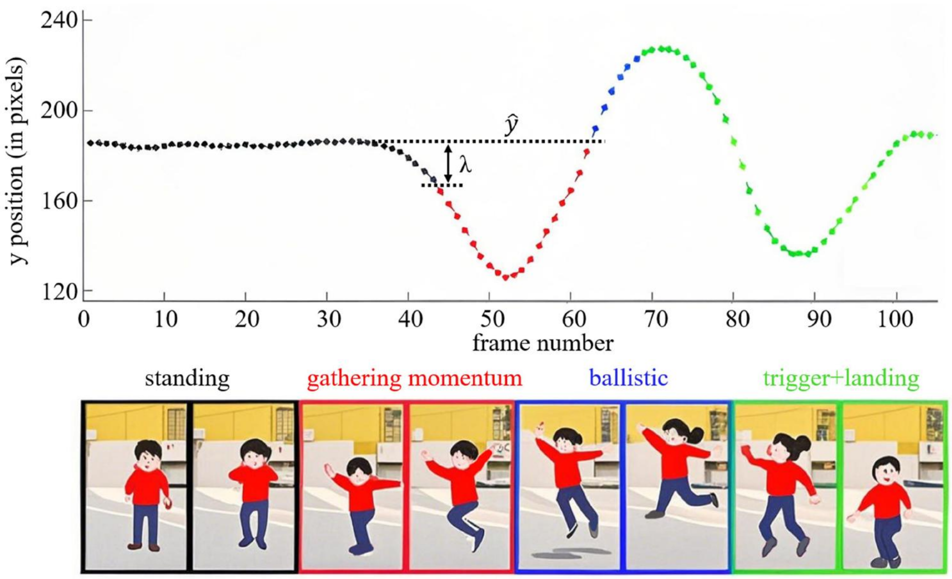

Figure 13. Using jump trajectory and jump phases automatically shoot jumping selfies. The figure is depicted in accordance with the description provided in Garcia et al. (2014).

color balance, shutter speed, and other settings have been proven to be highly efficient (Chu & Tanaka, 2015).

While gesture recognition can be affected by factors such as lighting variations, advancements in artificial intelligence algorithms and the increasing popularity of RGB-D depth cameras, like the LiDAR (Light Detection And Ranging) depth camera on the iPhone, have greatly improved image capture precision and gesture interaction accuracy. In summary, mid-air gesture interaction, which facilitates distance-based control and allows users to pose freely with both hands and navigate complex menus, is widely considered as the most promising interaction method for selfie camera applications.

3.4.3. Voice user interface

Voice interaction leverages speech recognition technology to control cameras via natural language. Over the past few years, advances in deep learning, large language models, and artificial intelligence, along with the maturation of natural language recognition and comprehension technologies, have contributed to the increasingly widespread use of voice interaction in diverse application domains (Seaborn et al., 2021). When equipped with the capability to comprehend human semantics and instructions, the voice interface offers remote control that does not hinder posing during selfies, thereby enhancing the freedom and flexibility of interaction with selfie cameras. For example, Park (2020) designed voice commands for editing selfie AR stickers and conducted experiments among people with visual impairments, proving its efficacy. It is convincingly reasonable to believe that voice interaction serves as an effective approach for controlling selfie cameras.

3.4.4. Remote control

By leveraging a smartphone or a dedicated remote controller, selfie takers can remotely preview and fine-tune camera parameters. For example, Canon has tailored a mobile application specifically for remote control, facilitating seamless real-time interconnection and comprehensive functional operation with Canon cameras. Furthermore, the Canon SDK empowers users to preview camera footage in real time and exercise control over camera functions directly from a computer. Additionally, in a desktop environment, users can effortlessly manipulate camera parameters using a mouse or keyboard.

3.5. Shutter control

Despite being just one step in self-portrait photography, shutter control holds significant importance due to the distinctive character of selfies. It must minimize any interference with the selfie taker's ability to pose freely while also ensuring steady eye contact with the camera lens during the capture process. Typically, following shutter operation, there are two primary photo-taking approaches: immediate shutter and countdown timer. When holding the camera or a selfie stick, direct capture tends to be the preferred method. However, with the countdown method, the selfie taker gains the flexibility to step away from the camera, pose with more variety, and overcome challenges such as device shaking and maintaining eye contact while operating the camera shutter. Notably, the interaction methods for shutter control exhibit numerous similarities to those for parameter control discussed in the previous section, enabling the application of most general parameter control techniques to shutter control as well. In the following, we will primarily highlight the specific interaction methods tailored for activating the shutter.

3.5.1. Physical contact buttons

Contact buttons, whether they are physical buttons on the camera device itself or virtual buttons displayed on a touchscreen, serve as the means to activate the shutter or

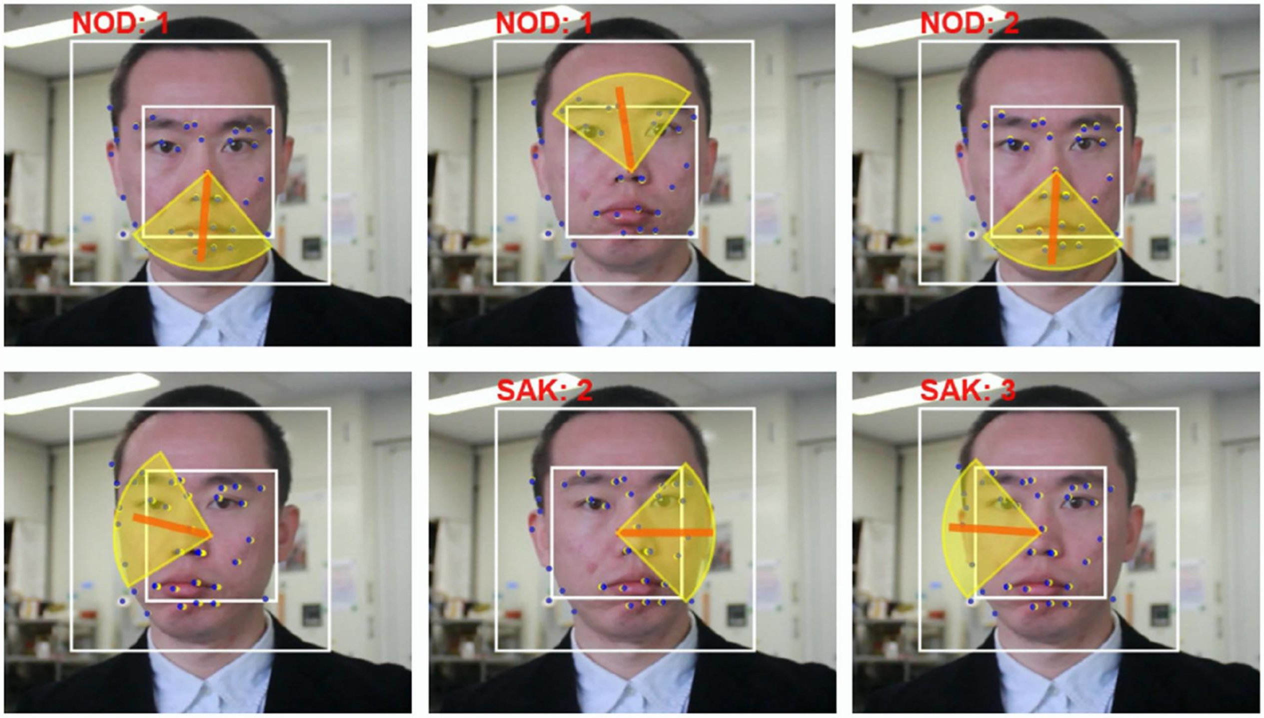

Figure 14. Head motion gestures for controlling camera face zoom. The figure is used with permission from authors (Chu & Tanaka, 2013a).

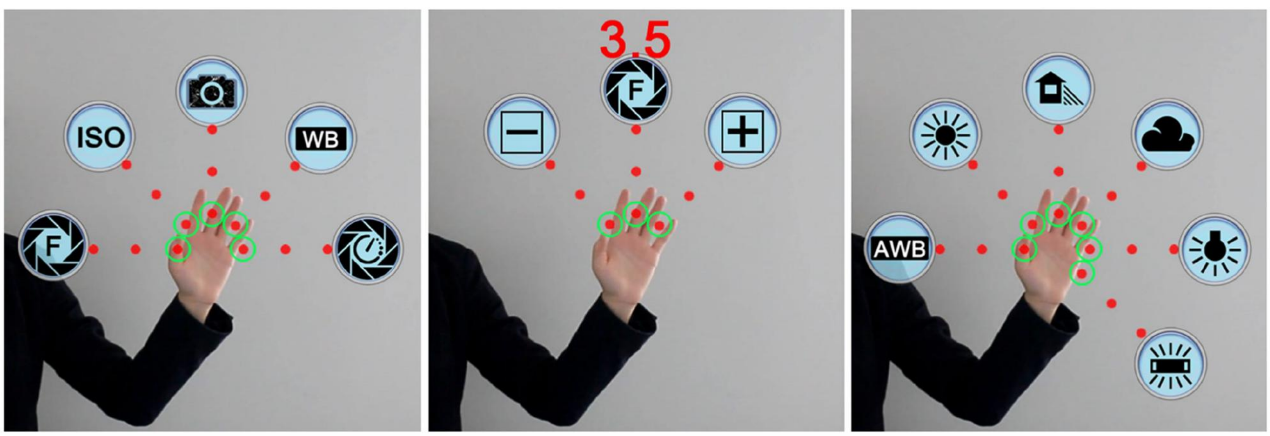

Figure 15. The air gesture menu interface for operating selfie camera parameters. The figure is used with permission from authors (Chu & Tanaka, 2015).

initiate a timer to capture photos. This approach offers the most intuitive and direct form of interaction, making it a popular choice for selfies captured with an extended arm (Fang et al., 2018; Li & Vogel, 2017) or in selfie booth applications that leverage touchscreen buttons for shutter activation (Dugan et al., 2015; Lu et al., 2016; Okabe et al., 2006). However, a limitation of this method is that it requires positioning the camera close to the subject, resulting primarily in close-up photographs. Alternatively, when utilizing equipment such as a tripod, users typically set the timer and subsequently move in front of the camera to pose for their selfie once the timer is activated (Chu & Tanaka, 2013b).

3.5.2. Remote control

A remote controller is commonly paired with a tripod to control the camera, either via the remote controller itself, a smartphone, or a mouse and keyboard in a desktop setting. In portrait photography, the classic 85 mm focal length typically necessitates a distance of 2–5 meters between the subject and the camera (Rand & Meyer, 2014). One advantage of using a remote controller is its capability to capture half-body or full-body portraits from a distance, while also incorporating more background details and perspective effects. However, a disadvantage lies in the fact that the hands are occupied by the remote controller, thereby limiting the subject's freedom to pose naturally (Chu & Tanaka, 2015).

3.5.3. Air gesture control

Air Gestures utilize image processing technology to recognize users' hand movements, enabling remote shutter control. For instance, as shown in Figure 16, Chu and Tanaka (2011) designed a fingertip recognition algorithm to activate

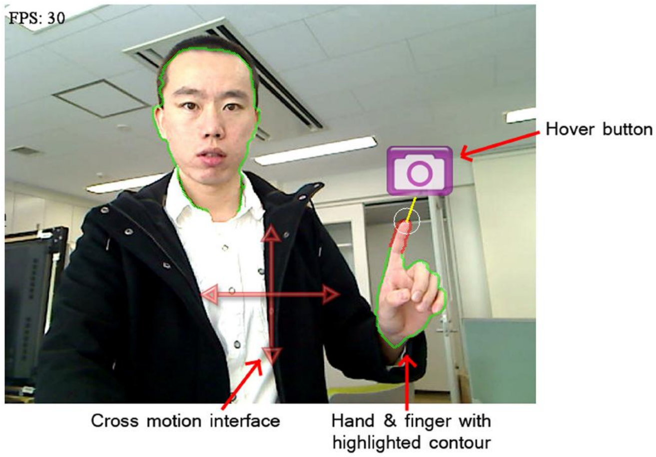

Figure 16. Using a fingertip to activate an air shutter control button. The figure is used with permission from authors (Chu & Tanaka, 2011).

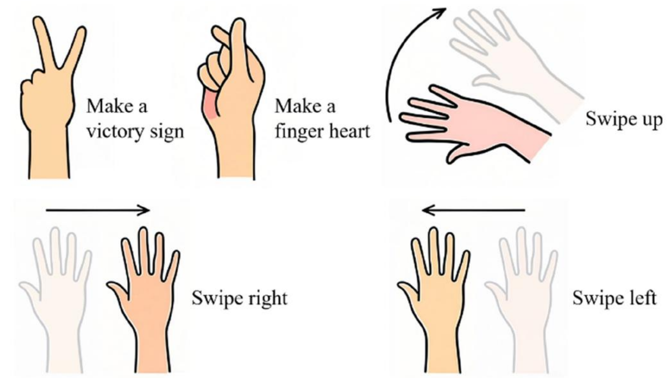

Figure 17. User-defined gestures for shutter control. The figure is depicted in accordance with the description provided in Peng and Xu (2020).

an air shutter control button. Chu and Tanaka (2013b) designed the waving and directional motion gestures to activate the shutter. Peng and Xu (2020) further explored user-defined gestures, identifying a set of 24 suitable for selfies, and discovered that over 94% of users prefer single-hand gestures for shutter operation. Typical selfie shutter gestures include making a victory sign, forming a finger heart, swiping up, and swiping left and right, as illustrated in Figure 17. Further experiments highlighted the victory sign and upward swipe as particularly effective selfie shutter triggers.

Moreover, the utilization of head gestures and micro-expressions for shutter activation in selfie photography has proven effective. Chu and Tanaka (2012, 2013a) introduced the double nod gesture as a means to activate the camera shutter. Unlike hand gestures, methods such as opening the mouth, raising and lowering eyebrows, and blinking do not interfere with posing, thereby enhancing the user's overall experience (Chu & Tanaka, 2013a).

3.5.4. Automatic shutter

As detailed in the composition Section 3.3.3, an automatic shutter entails the automated capture of selfies based on predefined compositions and other specific conditions (Garcia et al., 2014; Party-shot, 2010). Furthermore, techniques such as pose recognition, as introduced by Anderson et al. (2013), and continuous self-timer photography, which automatically captures multiple poses, are also available.

4. User experience of taking selfies

Existing studies have examined user experience for taking selfies accompanied by their proposed selfie systems. The primary and universal method employed to evaluate selfie systems is ISO-9241-11, which examines three aspects of effectiveness, efficiency, and user satisfaction for an interactive system. As for selfie systems, the effectiveness is generally evaluated based on task completion rates and error rates. For instance, Chu and Tanaka (2011, 2015) assessed the effectiveness of their proposed selfie systems by calculating the task completion rate and errors that occurred when users were unable to use gestures to control the selfie systems for completing instructed experimental tasks. They instructed users to operate a PTZ selfie camera to frame and focus on the selfie-taker before capturing a selfie shot. The efficiency is measured by task completion time, i.e., how long a user needs to take a selfie for a specific combination of lens zoom, ISO sensitivity, aperture, and shutter speed (Chu & Tanaka, 2013b, 2015; Mao et al., 2023). The satisfaction can be assessed subjectively through the use of Likert scale questionnaires, which evaluate factors such as accuracy, ease of use, comfort levels, fatigue levels, user preferences, and emotional responses. Depending on the focus of the experiment, these factors can be combined to enhance assessment (Chu et al., 2017a; Chu & Tanaka, 2015; Fang et al., 2018; Kim & Lee, 2019; Lusk & Jones, 2019; Mao et al., 2023).

The distinct characteristic of selfie systems is that they not only require the user to control the selfie camera to take photos, but also concern the user with the quality of the photos they have taken. Consequently, there are distinct user experiences related to selfies compared to traditional human-computer systems. The following metrics are typically adopted frequently when evaluating selfie systems.

4.1. Accuracy

It measures the preciseness of the captured selfie shot in comparison to the originally intended selfie or the degree of accuracy in operating camera parameters. For instance, Kim and Lee (2019) and Lusk and Jones (2019), by utilizing visual or voice guidance interfaces, assist users in positioning the camera towards a target frame and assess the accuracy of camera composition. To ensure a precise evaluation of accuracy, the shooting location and initial position of participants were kept as consistent as possible by presenting target images and aligning the camera pose accordingly. Specifically, in error-prone interactions, such as vision-based gesture recognition, the degree of accuracy in user operations affects muscle activity, leading to fatigue and significantly impacting the user experience during selfie interactions (Chu & Tanaka, 2015; Mao et al., 2023).

4.2. Pose score

It measures the pose of a captured selfie based on professional aesthetic assessments in photographic studies, considering factors such as face angle, geometry compositions, and the rule of thirds. There are two main approaches to assessing the pose score: one involves obtaining evaluations from experienced photographers on the quality of selfie images (Hu et al., 2019), while the other relies on a trained model to evaluate selfie quality. Although the former approach can deliver professional results, it is often time-consuming and may not always be practical due to the limited availability of photographers (Hu et al., 2019). In contrast, the latter method, which has gained traction recently, uses selfie-related features such as body pose, head pose, facial expressions, facial geometry, eye movement, gaze direction, mouse motion, wrinkles, and face texture features to train a model. The trained model can then assess the score of a new portrait (Fiss et al., 2011; Hu et al., 2019). This approach offers relatively objective evaluations and allows the model to provide real time guidance during the selfie-taking process (Fang et al., 2018).

4.3. Non-interruptive

It means that users have the freedom during the interaction procedure to prepare various poses. Thus, the interaction should minimally disrupt the selfie-taker's posing process (Chu & Tanaka, 2012). Various evaluation methods have accounted for this factor by contrasting proposed interaction techniques with traditional selfie interaction methods to assess their benefits. For example, Chu and Tanaka (2013a) compared a gesture interface with traditional methods such as timers and remote controls to evaluate the level of disruption during the selfie-taking process.

4.4. Process fluency

It means smooth and easy-to-follow guidance that users can readily understand and follow to capture intended selfies with desired compositions significantly impacts the user experience. This is particularly relevant in scenarios where users seek assistance from others. The process involves requesting help, clearly describing the requirements, the helper accurately understanding those requirements, and then utilizing the visual guidance or graphic interfaces provided by the selfie systems to achieve specific selfies with the original framing and compositions (Kim & Lee, 2019; Lusk & Jones, 2019).

4.5. Likert scale

It gathers questionnaire responses using a 5- or 7-point Likert scale to assess subjective user experiences with selfie techniques. Commonly designed questions address aspects such as interaction accuracy, efficiency, ease of use, confidence in use, naturalness or intuitiveness, and preference for frequent system use. Specifically, Fang et al. (2018) explored user satisfaction with the visual interface, the clarity and appropriateness of head pose guidance, and the application's effectiveness in helping users take better selfies. Other specific factors include communication accuracy when seeking assistance and feelings of resignation due to social anxiety (Lusk & Jones, 2019).

4.6. Subjective response

It refers to users' comments after completing the experiments, which provide valuable insights into their experiences with the system. These include concerns about comparisons with traditional methods, communication when seeking assistance, and social pressure associated with specific selfie techniques. For instance, using selfie sticks or new, potentially disruptive methods can make users feel awkward or intrusive in public spaces (Bevan, 2017; Lusk & Jones, 2019).

5. Real-world cases or techniques

Many successful real-world applications leveraging selfie techniques have emerged. A notable application is the selfie booth, which has gained global popularity, especially in Japan where it is known as Purikura (print club) (Okabe et al., 2006). Selfie booths usually feature large displays, touch-based interfaces, and timers to facilitate selfie-taking. They are commonly found in private rooms within entertainment venues (Okabe et al., 2006; Sandbye, 2018), office spaces (Dugan et al., 2015), and public landscapes and streets (Memarovic et al., 2015; Parker et al., 2020). Their immense success can be attributed to the high demand for entertainment in public spaces and the enthusiasm and narcissistic tendencies of many users, particularly young girls (Cheng et al., 2024; Sandbye, 2018).

Selfie sticks, particularly the light-weight versions as studied by Arif et al. (2017) and Xu (2023), have evolved, with the more recent and promising addition being the Osmo Selfie Stick (DJI, 2015). This device utilizes PTZ gimbals to ensure precise camera tracking of humans. It offers ant-shake capabilities, stability, and person tracking, making it ideal for capturing selfies and videos. Despite its relatively higher price than traditional selfie sticks, the Osmo has gained widespread adoption among professional photographers and online influencers.

DJI selfie drones, such as DJI Neo Mini Drone (DJI, 2025), are compact and lightweight drones designed to capture hands-free selfie photos and videos. These drones enable flight control via a smartphone application, voice commands, and gestures, and are equipped with advanced features such as person tracking and face recognition.

Companion robots, such as PepperPose (Wang et al., 2024), optimize captured selfie videos to estimate the pose of a user as they move and act diversely in an open space, enabling interactive applications in fields like sports, fitness, and healthcare. Similar applications include detecting hypomimia symptoms (Grammatikopoulou et al., 2019), assisting users with the insertion of hearing aids (Yang et al., 2022), and enabling self-diagnostic to improve the quality of care (Siddiqui et al., 2022) by analyzing selfie photos.

Sony's Party-shot (Party-shot, 2010) utilizes face and smile detection along with a PTZ camera to automatically capture selfies. It has achieved great success at family parties and public gatherings. Its key advantage lies in its ability to take selfies automatically without human intervention, capturing intuitive and natural facial expressions.

Casio's TR camera (Casio, 2011) features a rotating frame that allows for selfies to be taken from various angles. Its notable benefits include a lightweight design, the ability to position the camera on different platforms, walls, and extended arms, enabling creative selfies from numerous perspectives. Similarly, the subsequently designed Oppo N1 smartphone (Oppo, 2013) offers a high-quality rotatable camera component for shooting innovative selfies from various angles.

Various smartphone apps, a multitude of smartphone applications offer selfie-taking capabilities along with guidance. For instance, the Posing App (Mendo, 2024) presents poses to guide users to imitate professional selfies using aesthetic poses and frame compositions. Other selfie-focused apps, such as Meitu, YouCam Perfect, and Sweet Selfie, among others, provide beauty filters and leverage AI-powered editing to enhance selfies.

6. Discussion

Based on the review, we discuss our findings regarding interaction techniques for taking selfies in this section. We comment on the literature and suggest aspects to consider when designing selfie systems. We further suggest and speculate about future research directions and reflect upon the limitations of our work. Our findings provide an overview of the topic and some guidance for future research.

6.1. Interaction techniques

As stated in the introduction, our research focuses on identifying and analyzing relevant studies that investigate how selfie interaction techniques can streamline the process of capturing selfies. We summarize the review results as follows. Firstly, it is reasonable to ascertain that the most prevalent and suitable methods for capturing selfies involve using smartphones and front cameras for previewing, with extended arms or selfie sticks aiding in camera positioning and composition. This is evidenced by the daily influx of millions of selfies uploaded on SNS, facilitated by smartphones (Broz, 2024; Chen et al., 2024), as well as the projected substantial growth of the selfie stick market, valued in millions of USD, both currently and extending to 2030 (Patel, 2023).

Furthermore, applications that incorporate image recognition technologies have emerged as the most suitable interactive means for selfie-taking, particularly in terms of parameter control and shutter control. These technologies detect and recognize the user's body, guiding the selfie-taker into aesthetically pleasing poses. Additionally, they leverage air gestures, such as swiping upwards or nodding, to seamlessly operate camera functions, significantly enhancing the efficiency of capturing satisfying selfies. For instance, face detection algorithms can instruct selfie-takers to position their heads for optimal compositional aesthetics (Fang et al., 2018; Hu et al., 2019; Li & Vogel, 2017), and air gestures enable the camera shutter to be triggered with minimal disruption during pose preparation (Chu & Tanaka, 2013a, 2015).

Second, the selfie interaction framework can be comprehensively categorized into seven distinct steps: camera positioning, previewing the frame, composition, parameters control, shutter control, selfie image editing, sharing selfies on SNS, and receiving feedback or motivation that inspires the next selfie session (Faimau, 2020; Rand & Meyer, 2014; Shah & Tewari, 2016), as depicted in Figure 1. Notably, within the realm of interaction techniques, the first five aspects are particularly pertinent. Designers and researchers have devoted sophisticated attention to these aspects, employing diverse image recognition technologies, including face detection (Chu & Tanaka, 2013a; Fang et al., 2018; Hu et al., 2019) and gesture recognition (Chu & Tanaka, 2015; Mao et al., 2023), to delve into each facet, enhance selfie quality and overall interaction experience, thereby paving the way for improvements throughout the selfie-taking process.

Third, when designing selfie systems, four related research trends can be summarized: selfie camera design, interaction design, evaluation methodologies, and interdisciplinary design. We delve into these trends in detail, accompanied by design recommendations and visions for future endeavors, in the subsequent subsections.

6.2. Selfie camera design

Smartphones continue to reign supreme as the selfie device, primarily due to their ubiquitous availability and the fact that their screens offer the ideal preview window. Beyond just elevating camera quality, we recommend designing smartphones equipped with flip-fold screens or innovative cases that enhance grip and make it easy to place and capture photos from various angles, such as the design of Casio's TR cameras (Casio, 2011).

To enrich the smartphone selfie experience, we advocate the development of complementary accessories designed specifically for selfies. These include selfie sticks, gimbals, robotic arms, selfie booths, and even drones, each offering unique perspectives and fostering creativity. The design of these accessories must prioritize portability and ease of use, ensuring that they are not just functional but also convenient to carry and operate. As emphasized in Arif et al. (2017), selfie sticks, for example, should adhere to ergonomic principles, be lightweight, and incorporate anti-shake technology to facilitate swift and effortless photography whenever the moment strikes.

Also, we suggest embedding these selfie accessories within the environment, offering them as public rentable items or integrating them into selfie booths, thereby ensuring their quick availability whenever to capture a selfie.

6.3. Interaction design

Existing interaction methods for selfie cameras encompass physical buttons, touchscreen, voice commands, gestures, facial expressions, and automatic pose matching based on presets. Among these, employing visual recognition for gesture recognition technology stands out as an advantageous strategy. It utilizes the camera's inherent image-capturing and processing capabilities to recognize human bodies and gestures, thereby eliminating the need for additional devices and minimizing disruptions to the selfie taker's pose.

When designing gesture-based interactions for taking selfies, we recommend considering the following factors in the interaction design:

6.3.1. User-defined gestures

Adopt methods like those in Peng and Xu (2020) and Wobbrock et al. (2009) to define a set of natural and intuitive gestures tailored specifically for selfie interactions. Complement this with field studies to evaluate user experiences in actual usage scenarios.

6.3.2. Subtle interaction gestures

Develop subtle gestures as triggers for actions, such as facial movements like opening the mouth, raising eyebrows, or blinking. These subtle interactions minimize distractions during posing (Chu & Tanaka, 2013a), allowing selfie takers to maintain their desired pose.

6.3.3. Menu interface for complex operations

When designing interfaces for complex operations involving camera controls like aperture, shutter speed, ISO, and special effects filters, consider implementing an air gesture-supported menu interface (Chu & Tanaka, 2015; Mao et al., 2023). This approach is suitable and efficient, especially when paired with large screens, such as in selfie booths or projection screens for previewing camera footage and displaying operational interfaces.

6.3.4. Automated operation and shooting

When utilizing gimbals, robotic arms, drones, and other supportive devices with image recognition technology, we can design features like automatic face tracking, focusing, framing composition, and shooting. Allow users to automate their shooting process based on their predefined selfie concepts and preferences as outlined by Alabachi et al. (2020). Consequently, it facilitates the capturing of more creative poses, such as jumping selfies (Garcia et al., 2014), thereby enhancing the overall selfie experience.

6.4. Evaluation method

The evaluation of selfie interaction technology is similar to that of traditional human-computer interaction, which can be conducted from the aspects of effectiveness, efficiency, and user satisfaction as outlined in ISO-9241-11. For instance, the effectiveness of interaction technology can be assessed by evaluating the task completion time, accuracy, and user subjective rating of gesture operations (Chu & Tanaka, 2015; Mao et al., 2023).

Furthermore, given the unique nature of selfie interactions, it is crucial to incorporate the evaluation of selfie image quality. This can be achieved through the establishment of an aesthetic evaluation model for selfies based on a large dataset of high-quality selfies (Fang et al., 2018; Hu et al., 2019; Li & Vogel, 2017; Zhu et al., 2014). Such a model not only assesses the image quality of selfies but also provides real-time guidance to users on composition, lighting, angles, and other selfie settings.

Therefore, we suggest that the evaluation of selfie camera interaction technology should build upon the traditional ISO-9241-11 framework by adding an aesthetic assessment of image quality. Carefully selecting a dataset of excellent selfie photos to train a beautiful selfie model can enable the system to provide interactive scores and guidance during the process of taking selfies.

6.5. Interdisciplinary design

After reviewing the literature, we found that numerous deep learning and AI (artificial intelligence) techniques, VR (Virtual Reality), human behaviors, and social science methods have been developed for portrait image editing, recognition, aesthetics score estimation, image generation, and understanding the selfie culture. Although we did not investigate these aspects extensively as we were focusing on interaction techniques, they offer new perspectives for enhancing selfie quality.

First, with the rapid advancement of deep learning in computer vision, large language models, and AI, these technologies can be integrated into selfie systems to improve face detection, facial expression recognition, pose estimation, as well as image aesthetics and generation. These new technologies have significantly impacted the way selfies are captured and the quality of the images.

Second, as VR and the metaverse continue to evolve, the integration of high-quality selfies, 3D portraits (Bao et al., 2017), and AI technology enables the calculation and generation of diverse visual effects, encompassing varying shooting angles, lighting conditions, facial expressions, poses, attire, and makeup. This reshapes the way we interact with digital worlds. The advancements in these technologies empower us to swiftly and effortlessly create numerous high-quality selfies, or digital humans, thereby promising an immersive experience and enhancing our ability to cultivate a comprehensive, diverse, and high-quality self-image within social networks in the VR and metaverse.

Third, selfies possess characteristics typical of social activities, allowing them to be uploaded and shared on SNS. Consequently, we must consider security factors, such as privacy concerns, misuse, and the psychological and behavioral motivations behind selfie cultural trends (Attia & Zaghloul, 2025; Hu et al., 2022; Ma & Zheng, 2024). Current research has also explored various aspects of selfie-taker behavior, including their self-concept, self-confidence, self-disclosure, narcissism, covert narcissism, neuroticism, self-discrepancy, and self-esteem (Barry et al., 2015; Yang et al., 2021). It's crucial to be mindful of the potential implications of excessive selfie-taking and posting, as it may lead to issues like body image dissatisfaction and social comparison (Chua & Chang, 2016). Balancing the use of selfies with a healthy digital lifestyle is essential.

Fourth, as selfies become a social activity, specialized applications have been developed for specific purposes. For instance, selfie systems can assist users in learning dance moves, fitness routines, monitoring health conditions, and various other activities (Anderson et al., 2013; Wang et al., 2024). Blind individuals may find taking selfies an effective way to share their status, appearance, expressions, and other relevant information with their sighted friends or to receive remote assistance (Kawai & K.K.a.S.I.a.N, 2024; Penuela et al., 2022).

These techniques will, in turn, provide feedback for the design of interaction techniques. Expanding the current research framework to include selfie editing, AI image generation, VR, and the metaverse could lead to a more comprehensive improvement of selfie systems.

7. Future directions

To advance the field toward technology development of selfie systems and gain more general insights, we present implications for future research and practice.

7.1. Implications for future research

We summarize the following implications for future research.

- Future research should conduct comparative experiments across various techniques to assess their effectiveness in real-world applications.
- With the emergence of VR and metaverse technologies, coupled with advancements in deep learning and AI development, the capture and creation of high-quality selfies, as well as the generation of digital humans in VR, could emerge as thrilling application directions for selfie systems.
- User-defined gestures ought to be employed to devise user-satisfying and intuitive interaction paradigms for selfie cameras across diverse scenarios with different supporting devices. Simultaneously, the gesture recognition method should offer high accuracy and incorporate an error recovery mechanism to ensure smooth and efficient interaction.
- A well-trained dataset of beautiful selfies should be established to provide aesthetic guidance for evaluating captured selfies and to offer interactive assistance during the selfie-taking process.
- When evaluating selfie systems or specific selfie interaction techniques, it is imperative to not only consider the traditional ISO-9241-11 framework for human-computer interaction but also integrate aesthetic assessments or users' original desires.
- Interdisciplinary designs that incorporate selfie editing, image generation, VR, the metaverse, and social science should be taken into account to broaden the scope of selfie studies.

7.2. Implications for future practice

We summarize implications for future practice as follows.

- Utilizing extended arms and the front camera of smartphones is the most accessible method for capturing typical selfies. By utilizing supporting apps (Fu et al., 2013; Mendo, 2024), users can learn and adhere to specific guidance to capture well-posed selfies.
- Users have the option to utilize selfie sticks to broaden their view and capture selfies from various angles. In advanced scenarios requiring stability and automatic camera framing and composition, more professional gimbal-equipped selfie sticks can be employed.
- Flip-fold screens or innovative framing cases for smartphones are available, enhancing grip and facilitating selfie capture from diverse angles.
- Supporting equipment, such as selfie sticks, gimbals, robotic arms, selfie booths, and drones, can be integrated into various environments to facilitate selfie-taking in diverse locations and creative contexts.
- Subtle motion gestures can be utilized when interacting with selfie cameras to reduce distractions while posing.
- A pie menu interface can be employed for designing complex gesture operations, making it easier for users to navigate and execute intricate commands.
- Automated operation and shooting should be developed for non-eye-contact selfies, such as profile and jumping selfies, ensuring that high-quality images can be captured even in challenging poses.

We acknowledge the necessity for cross-disciplinary and interdisciplinary research among human-computer interaction, image art, and social science to address these factors and enhance the positive impact of selfie systems, ensuring both interaction effectiveness and user satisfaction.

8. Conclusion

In this work, we present a literature review on interaction techniques for selfie cameras. We contribute a comprehensive framework that summarizes the intricate process of selfie interaction, categorizing the diverse techniques involved at each stage of the selfie-capturing procedure. We provide a summary of existing selfie camera systems, delving deep into their distinctive characteristics. In addition, we discuss selfie camera design, interaction design, evaluation aspects, and the future directions relating to selfie techniques. Our reviews on selfie-related technologies hold the potential to enrich HCI research on selfies and inform marketing strategies for both hardware and software selfie applications.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Open Fund of Provincial Key Laboratory of Film and TV Media Technology and the Intelligent Media Engineering Research Center of Zhejiang Province.

ORCID

Huawei Tu http://orcid.org/0000-0001-9689-9767

References

Ahmed, D., Qureshi, W. S., Aijaz, S. A., Imran, B. M., Naqvi, S. M. A., & Lin, C. Y. (2021). Towards selfie drone: Spatial localization and navigation of drone using human pose estimation. *2021 International Conference on Robotics and Automation in Industry (ICRAI)* (pp. 1–7).

Alabachi, S., Sukthankar, G., & Sukthankar, R. (2020). *SelfieDroneStick: A natural interface for quadcopter photography* [Paper presentation]. 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA (pp. 10684–10690). IEEE Press.

Anderson, F., Grossman, T., Matejka, J., & Fitzmaurice, G. (2013). YouMove: enhancing movement training with an augmented reality mirror.

Arif, A. S., Kim, S., & Lee, G. (2017). Usability of different types of commercial selfie sticks. *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, Vienna, Austria (p. Article 10). Association for Computing Machinery.

Attia, M. H., & Zaghloul, N. M. (2025). Selfie accidents and crimes in Egypt: A comprehensive discussion of the medico-legal implications. *Medicine, Science, and the Law*, *65*(2), 113–119. https://doi.org/10.1177/00258024241257100

Balata, J., Mikovec, Z., & Neoproud, L. (2015). BlindCamera: Central and golden-ratio composition for blind photographers. *Proceedings of the Mulitimedia, Interaction, Design and Innnovation*, Warsaw, Poland (p. Article 8). Association for Computing Machinery.

Bao, L., Lin, X., Chen, Y., Zhang, H., Wang, S., Zhe, X., Kang, D., Huang, H., Jiang, X., Wang, J., Yu, D., & Zhang, Z. (2017). High-fidelity 3D digital human head creation from RGB-D selfies. *ACM Transactions on Graphics*, *41*(1), 1–21. https://doi.org/10.1145/3472954

Barry, C. T., Doucette, H., Loflin, D. C., Rivera-Hudson, N., & Herrington, L. L. (2015). "Let me take a selfie": Associations between self-photography, narcissism, and self-esteem. *Psychology of Popular Media Culture*, *1*(1), 1–13. https://doi.org/10.1037/ppm0000089

Bevan, J. L. (2017). Perceptions of selfie takers versus selfie stick users: Exploring personality and social attraction differences. *Computers in Human Behavior*, *75*, 494–500. https://doi.org/10.1016/j.chb.2017.05.039

Broz, M. (2024). *Selfie statistics, demographics, & fun facts (2024)*, https://photutorial.com/selfie-statistics/

Casio. (2011). *TR camera*. https://www.casio-intl.com/asia/en/dc/products/ex_tr80/

Chen, B. H., Curless, B., Kemelmacher-Shlizerman, I., & Seitz, S. M. (2024). *Total selfie: Generating full-body selfies* [Paper presentation]. 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA (pp. 6701–6711).

Chen, T., Chen, Y., & Luo, J. (2017). A selfie is worth a thousand words: Mining personal patterns behind user selfie-posting behaviours. *Proceedings of the 26th International Conference on World Wide Web Companion. International World Wide Web Conferences Steering Committee*, Perth, Australia (pp. 23–31).

Cheng, L., Wang, X., & Zhang, J. (2024). She is seeking attention: Women's self-objectification and hostility toward peer selfies. *Computers in Human Behavior*, *156*, 108214. https://doi.org/10.1016/j.chb.2024.108214

Chen, C.-F., Liu, K.-P., & Yu, N.-H. (2015). *Exploring Interaction Modalities for a Selfie Drone, SIGGRAPH Asia 2015 Posters*, Kobe, Japan (p. Article 25). Association for Computing Machinery.

Chu, S., & Tanaka, J. (2011). *Hand gesture for taking self portrait* [Paper presentation]. 14th International Conference on Human-Computer Interaction, Human-Computer Interaction, Part II, U.S.A (pp. 238–247). Springer.

Chu, S., & Tanaka, J. (2012). *Head nod and shake gesture interface for a self-portrait camera* [Paper presentation]. ACHI 2012, the Fifth International Conference on Advances in Computer-Human Interactions, Spain (pp. 112–117).

Chu, S., & Tanaka, J. (2013a). Development of a head gesture interface for a self-portrait camera. *The Transactions of Human Interface Society*, *15*(3), 247–259. https://doi.org/10.11184/his.15.3_247

Chu, S., & Tanaka, J. (2013b). Interacting with a self-portrait camera using motion-based hand gestures. *Proceedings of the 11th Asia Pacific Conference on Computer Human Interaction*, India (pp. 93–101). ACM.

Chu, S., & Tanaka, J. (2015). Design of a motion-based gestural menu-selection interface for a self-portrait camera. *Personal and Ubiquitous Computing*, *19*(2), 415–424. https://doi.org/10.1007/s00779-014-0776-1

Chu, S., Lu, D., & Pan, R. (2017a). Motion gesture interface for self-portrait camera. *Computer Engineering and Applications*, *53*(10), 181–186. https://doi.org/10.3778/j.issn.1002-8331.1512-0158

Chu, S., Zhang, F., Ji, N., Jin, Z., & Pan, R. (2017b). *Pan-and-tilt self-portrait system using gesture interface* [Paper presentation]. 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS) (pp. 599–605).

Chua, T. H. H., & Chang, L. (2016). Follow me and like my beautiful selfies: Singapore teenage girls' engagement in self-presentation and peer comparison on social media. *Computers in Human Behavior*, *55*, 190–197. https://doi.org/10.1016/j.chb.2015.09.011

DJI. (2015). Osmo.

DJI. (2025). *DJI Neo*. https://www.dji.com/cn/neo

Dugan, C., Laumer, S., Erickson, T., Kellogg, W., & Geyer, W. (2015). The #selfiestation: Design and use of a kiosk for taking selfies in the enterprise. In J. Abascal, S. Barbosa, M. Fetter, T. Gross, P. Palanque, & M. Winckler (Eds.), *Human-computer interaction – INTERACT 2015* (pp. 38–46). Springer International Publishing.

Faimau, G. (2020). Towards a theoretical understanding of the selfie: A descriptive review. *Sociology Compass*, *14*(12), 1–12. https://doi.org/10.1111/soc4.12840

Fang, N., Xie, H., & Igarashi, T. (2018). Selfie guidance system in good head postures. ACM IUI 2018 Workshops (pp. 1–8).

Fiss, J., Agarwala, A., & Curless, B. (2011). Candid portrait selection from video. ACM Transactions on Graphics (TOG), 128. https://doi.org/10.1145/2024156.2024162

Fu, H., Han, X., & Phan, Q. H. (2013). *Data-driven suggestions for portrait posing, SIGGRAPH Asia 2013 Technical Briefs* [Paper presentation]. Association for Computing Machinery, Hong Kong, Hong Kong (p. Article 29).

Garcia, C., Bazin, J.-C., Lancelle, M., & Gross, M. (2014). *Automatic jumping photos on smartphones, Image Processing (ICIP)* [Paper presentation]. 2014 IEEE International Conference on (pp. 2522–2526). IEEE.

Grammatikopoulou, A., Grammalidis, N., Bostantjopoulou, S., & Katsarou, Z. (2019). Detecting hypomimia symptoms by selfie photo analysis: For early Parkinson disease detection. *Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, Rhodes, Greece (pp. 517–522). Association for Computing Machinery.

Hogan, B., & Wellman, B. (2014). The relational self-portrait: Selfies meet social networks. *Society and the Internet: How Networks of Information and Communication Are Changing Our Lives*, 53–66. https://doi.org/10.1093/acprof:oso/9780199661992.003.0004

Hsieh, Y.-T., & Yeh, M.-C. (2017). head pose recommendation for taking good selfies. *Proceedings of the Workshop on Multimodal Understanding of Social, Affective and Subjective Attributes*, California, USA pp. 55–60). Association for Computing Machinery, Mountain View.

Hu, C.-S., Hsieh, Y.-T., Lin, H.-W., & Yeh, M.-C. (2019). Virtual portraitist: An Intelligent Tool for Taking Well-Posed Selfies. *ACM Transactions on Multimedia Computing, Communications, and Applications*, *15*(1s), 1–17. https://doi.org/10.1145/3288760

Hu, X., He, L., & Liu, J. (2022). The power of beauty: Be your ideal self in online reviews—an empirical study based on face detection. *Journal of Retailing Consumer Services*, *67*, 102975. https://doi.org/10.1016/j.jretconser.2022.102975

Im, H., Kim, T., Jung, E., Ku, B., Baek, S., & Lee, T. Y. (2022). *Virfie: Virtual group selfie station for remote togetherness* [Paper presentation]. Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems, New Orleans, LA, USA (p. Article 382). Association for Computing Machinery.

Jane, J. E., Fried, O., & Agrawala, M. (2019). Optimizing portrait lighting at capture-time using a 360 camera as a light probe. *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, New Orleans, LA, USA (pp. 221–232). Association for Computing Machinery.

Jane, J. E., Fried, O., Lu, J., Zhang, J., Mech, R., Echevarria, J., Hanrahan, P., & Landay, J. A. (2020). Adaptive photographic composition guidance. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, Honolulu, HI, USA (pp. 1–13). Association for Computing Machinery.

Jane, J. E., Ilene, I. E., Landay, J. A., & Cauchard, J. R. (2017). Drone & Wo: Cultural influences on human-drone interaction techniques. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, Denver, Colorado, USA (pp. 6794–6799). Association for Computing Machinery.

Jane, J. E., Zhai, K. Y., Echevarria, J., Fried, O., Hanrahan, P., & Landay, J. A. (2021). Dynamic guidance for decluttering photographic compositions. *The 34th Annual ACM Symposium on User Interface Software and Technology*, Virtual Event, USA (pp. 359–371). Association for Computing Machinery.

Jin, X., Fan, C., Wang, B., Xiao, C., & Xia, C. (2023). Human pose recommendation and professionalization. *Proceedings of the 1st International Workshop on Multimedia Content Generation and Evaluation: New Methods and Practice*, Ottawa ON, Canada (pp. 91–98). Association for Computing Machinery.

Joo, K., Li, H., Oh, T. H., & Kweon, I. S. (2022). Robust and efficient estimation of relative pose for cameras on selfie sticks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(9), 5460–5471. https://doi.org/10.1109/TPAMI.2021.3085134

Kawai, K.K.a.S.I.a.N. (2024). Selfie taking with facial expression recognition using omni-directional camera. arXiv:2405.15996.

Kim, M., & Lee, J. (2019). PicMe: Interactive visual guidance for taking requested photo composition. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems – CHI '19* (pp. 1–12).

Lee, T. P., & Kosasih, W. (2015). Exploring the distance, location, and style of "selfies" with the self-pano mobile application. *New Contributions in Information Systems and Technologies* (pp. 875–884). Springer.

Li, Q., & Vogel, D. (2017). Guided selfies using models of portrait aesthetics. *Proceedings of the 2017 Conference on Designing Interactive Systems*, Edinburgh, United Kingdom (pp. 179–190). Association for Computing Machinery.

Lin, J., & Yang, C.-K. (2020). A multi-person selfie system via augmented reality. *Computer Graphics Forum*, *39*(7), 553–564. https://doi.org/10.1111/cgf.14167

Liu, H., Zhang, C., Deng, Y., Liu, T., Zhang, Z., & Li, Y.-F. (2023). Orientation cues-aware facial relationship representation for head pose estimation via transformer. *IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society*, *32*, 6289–6302. https://doi.org/10.1109/TIP.2023.3331309

Liu, H., Zhou, Q., Zhang, C., Zhu, J., Liu, T., Zhang, Z., & Li, Y. (2024). MMATrans: Muscle movement aware representation learning for facial expression recognition via transformers. *IEEE Transactions on Industrial Informatics*, *20*(12), 13753–13764. https://doi.org/10.1109/TII.2024.3431640

Lo, C.-C., Huang, S.-P., Ren, Y., & Tseng, Y.-C. (2013a). *Design of a self-portrait application with sensor-assisted guiding for smart devices* [Paper presentation]. Communications and networking (BlackSeaCom), 2013 first international black sea conference on (pp. 99–101). IEEE.

Lo, C.-C., Huang, S.-P., Ren, Y., & Tseng, Y.-C. (2013b). Yes, right there!: A self-portrait application with sensor-assisted guiding for smartphones. *Proceeding of the 11th Annual International Conference on Mobile Systems, Applications, and Services* (pp. 505–506). ACM.

Lu, D., Dugan, C., Farzan, R., & Geyer, W. (2016). Let's stitch me and you together!: Designing a photo co-creation activity to stimulate playfulness in the workplace. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, Santa Clara, California, USA (pp. 3061–3065). ACM.

Lusk, C., & Jones, M. D. (2019). Cake cam: take your photo and be in it too. *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*, Taipei, Taiwan (p. Article 12). Association for Computing Machinery.

Ma, L., & Zheng, L. (2024). Ego-networks, SNSs affordances, and personalities: Understanding individuals' selfie posting on SNSs based on Actor-Network Theory. *Behaviour & Information Technology*, *43*(3), 507–522. https://doi.org/10.1080/0144929X.2023.2177824

Ma, L., Lin, Z., Barnes, C., Efros, A. A., & Lu, J. (2020). Unselfie: Translating selfies to neutral-pose portraits in the wild. *ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XVII*, Glasgow, United Kingdom (pp. 156–173). Springer-Verlag.

Mao, L., Ono, K., Qi, M., Watanabe, M., & Huang, J. (2023). Comparison of panel and pie menus for mid-air interaction in a photo booth: Muscle activity, performance, and user experience. *International Journal of Human–Computer Interaction*, *40*(23), 1–15. https://doi.org/10.1080/10447318.2023.2279412

Mei-Chen, Y., & Hsiao-Wei, L. (2014). Virtual portraitist: Aesthetic evaluation of selfies based on angle. *Proceedings of the 22nd ACM International Conference on Multimedia*, Orlando, Florida, USA (pp. 221–224). ACM.

Memarovic, N., Schieck, A., Schnädelbach, H. M., Kostopoulou, E., North, S., & Ye, L. (2015). Capture the moment: "In the wild" longitudinal case study of situated snapshots captured through an urban screen in a community setting. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, Vancouver, BC, Canada (pp. 242–253). Association for Computing Machinery.

Mendo, S. (2024). Posing App.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G., PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Ngo, T.-T., & Kim, D.-S. (2017). *Design and implementation of an automated, long-range selfie camera system using LORA* [Paper presentation]. 2017 17th International Conference on Control, Automation and Systems (ICCAS), Jeju, Korea (South) (pp. 1768–1772). IEEE Press.

Okabe, D., Ito, M., Chipchase, J., & Shimizu, A. (2006). The social uses of purikura: Photographing, modding, archiving, and sharing. Proc. PICS Workshop, UbiComp (pp. 2–5).

Oppo. (2013). *Oppo N1*. https://www.cnet.com/reviews/oppo-n1-review/

Park, S. (2020). Supporting selfie editing experiences for people with visual impairments. *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility* (p. Article 106). Association for Computing Machinery.

Parker, C., Fredericks, J., Yoo, S., & Hoggenmueller, M. (2020). Let's take a selfie: Design considerations for public selfie booths. *Proceedings of the 31st Australian Conference on Human-Computer-Interaction*, Fremantle, WA, Australia (pp. 560–562). Association for Computing Machinery.

Party-shot. (2010). *Sony party-shot*. https://www.sony.jp/cyber-shot/party-shot/

Patel, D. (2023). *Global selfie stick market*. https://dataintelo.com/report/global-selfie-sticks-market/

Peng, Z., & Xu, J. (2020). User-defined gestures for taking self-portraits with smartphone based on consistency. In *Advances in Usability, User Experience, Wearable and Assistive Technology*, Cham (pp. 316–326). Springer International Publishing.

Penuela, R. E. G., Vermette, P., Yan, Z., Zhang, C., Vertanen, K., & Azenkot, S. (2022). Understanding how people with visual impairments take selfies: Experiences and challenges. *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility*, Athens, Greece (p. Article 63). Association for Computing Machinery.

Pounders, K., Kowalczyk, C. M., & Stowers, K. (2016). Insight into the motivation of selfie postings: Impression management and self-esteem. *European Journal of Marketing*, *50*(9/10), 1879–1892. https://doi.org/10.1108/EJM-07-2015-0502

Rand, G., & Meyer, T. (2014). *The portrait: Understanding portrait photography*. Rocky Nook.

Rogers, K., & Seaborn, K. (2023). The systematic review-lution: A manifesto to promote rigour and inclusivity in research synthesis. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (pp. 1–11).

Roudaki, A., Kong, J., & Reetz, S. (2016). SmartCamera: A low-cost and intelligent camera management system. *Multimedia Tools and Applications*, *75*(13), 7831–7854. https://doi.org/10.1007/s11042-015-2700-8

Sachse, J. F. (1893). Philadelphia's share in the development of photography. *Journal of the Franklin Institute*, *135*(4), 271–287. https://doi.org/10.1016/0016-0032(93)90284-2

Sandbye, M. (2018). Selfies and Purikura as affective, aesthetic labor. In J. Eckel, J. Ruchatz, & S. Wirth (Eds.), *Exploring the selfie: Historical, theoretical, and analytical approaches to digital self-photography*. Springer International Publishing (pp. 305–326).

Seaborn, K., Miyake, N. P., Pennefather, P., & Otake-Matsuura, M. (2021). Voice in human–agent interaction: A survey. *ACM Computing Surveys*, *54*(4), 81. https://doi.org/10.1145/3386867

Shah, R., & Tewari, R. (2016). Demystifying 'selfie': A rampant social media activity. *Behaviour & Information Technology*, *35*(10), 864–871. https://doi.org/10.1080/0144929X.2016.1201693

Sheth, M., & Srivastava, N. (2019). Predicting body size using mirror selfies. Intelligent Human Computer Interaction: 11th International Conference, IHCI 2019, Proceedings, Allahabad, India, December 12–14, 2019 (pp. 151–162). Springer-Verlag.

Siddiqui, H., Rattani, A., Cure, L., Keene Woods, N., Lewis, R., Twomey, J., Smith-Campbell, B., & Hill, T. (2022). Toward on-device weight monitoring from selfie face images using smartphones. In C. Comito, A. Forestiero, E. Zumpano (Eds.), *Integrating artificial intelligence and IoT for advanced health informatics: AI in the healthcare sector* (pp. 53–67). Springer International Publishing.

Trenz, P., Pasewaldt, S., Klingbeil, M., Döllner, J., & Trapp, M. (2021). Forward Selfies, ACM SIGGRAPH 2021 Appy Hour, Virtual Event, USA (p. Article 3). Association for Computing Machinery.

Virk, J. S., & Dhall, A. (2019). Garuda: A deep learning based solution for capturing selfies safely. *Companion Proceedings of the 24th International Conference on Intelligent User Interfaces*, Marina del Ray, California, USA (pp. 43–44). Association for Computing Machinery.

Wang, B., QvZhen, B., Zhang, Z., Li, Y., Yao, W., Wei, S., & Xing, S. (2023). MAGICROP: A cropping system for novice users to make artful portrait images. *International Journal of Human–Computer Interaction*, *40*(19), 1–16. https://doi.org/10.1080/10447318.2023.2247570

Wang, C., Zheng, S., Zhong, L., Yu, C., Liang, C., Wang, Y., Gao, Y., Lam, T. L., & Shi, Y. (2024). PepperPose: Full-body pose estimation with a companion robot. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, Honolulu, HI, USA (p. Article 586). Association for Computing Machinery.

Wang, Y., Song, M., Tao, D., Rui, Y., Bu, J., Tsoi, A. C., Zhuo, S., & Tan, P. (2015). Where2Stand: A human position recommendation system for souvenir photography. *ACM Transactions on Intelligent Systems and Technology*, *7*(1), 1–22. https://doi.org/10.1145/2770879

Wobbrock, J. O., Morris, M. R., & Wilson, A. D. (2009). User-defined gestures for surface computing. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Boston, MA, USA (pp. 1083–1092). Association for Computing Machinery.

Xia, S., Yang, S., Liu, J., & Guo, Z. (2016). Novel self-portrait enhancement via multi-photo fusing. *2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)* (pp. 1–4).

Xu, P., Yao, H., Ji, R., Liu, X.-M., & Sun, X. (2014). Where should I stand? Learning based human position recommendation for mobile photographing. *Multimedia Tools and Applications*, *69*(1), 3–29. https://doi.org/10.1007/s11042-012-1343-2

Xu, X. (2023). Handle sliding structure and selfie stick using same. U.S. Patent, United States.

Yang, D., Wu, T.-Y., Atkin, D. J., Ríos, D. I., & Liu, Y. (2021). Social media portrait-editing intentions: Comparisons between Chinese and American female college students. *Telematics and Informatics*, *65*, 101714. https://doi.org/10.1016/j.tele.2021.101714

Yang, Q., Papakostas, M., Scott, J. M., O'Neill, E. R., Kondrashov, K. S., Mateevitsi, V. A., Olsen, G., & Dittberner, A. B. (2022). CheckMyFit: Ear selfie to assist user insertion of hearing aids. *Proceedings of the ACM on Human-Computer Interaction*, *6*(MHCI), 1–21. https://doi.org/10.1145/3546721

Yeh, M.-C., & Lin, H.-W. (2014). Taking good selfies on your phone. *Proceedings of the ACM International Conference on Multimedia* (pp. 751–752). ACM.

Zhang, X., Li, Z., Constable, M., Chan, K. L., Tang, Z., & Tang, G. (2019). Pose-based composition improvement for portrait photographs. *IEEE Transactions on Circuits and Systems for Video Technology*, *29*(3), 653–668. https://doi.org/10.1109/TCSVT.2018.2817231

Zhu, J.-Y., Agarwala, A., Efros, A. A., Shechtman, E., & Wang, J. (2014). Mirror mirror: Crowdsourcing better portraits. *ACM Transactions on Graphics*, *33*(6), 1–12. Article 234. https://doi.org/10.1145/2661229.2661287

About the authors

Shaowei Chu is an associate professor in College of Media Engineering, Communication University of Zhejiang in China. He received his PhD from the University of Tsukuba in Japan. His research interests lie in human-computer interaction, user interface usability engineering, and tactile user interface design.

Huawei Tu is a senior lecturer at La Trobe University, Australia. He has 15 years of research experience in Human-Computer Interaction (HCI) and Virtual Reality (VR). He conducts both fundamental and applied research, combining theory and experimentation to explore how to enhance user experience in VR with HCI technologies.