Understanding Face and Eye Visibility in Front-Facing Cameras of Smartphones used in the Wild

Mohamed Khamis[1], Anita Baier[1], Niels Henze[2], Florian Alt[1,3], Andreas Bulling[4]

[1]LMU Munich, Germany (firstname.lastname@ifi.lmu.de)
[2]University of Stuttgart, Germany (firstname.lastname@vis.uni-stuttgart.de)
[3]Munich University of Applied Sciences, Germany (firstname.lastname@hm.edu)
[4]Max Planck Institute for Informatics, Saarland Informatics Campus, Germany (bulling@mpi-inf.mpg.de)

Figure 1. In our dataset of photos taken from front-facing cameras of smartphones used in the wild, the face is visible in only 29% of the cases. However, the eyes, and not the whole face, are visible 48% of the time. We derive multiple implications for face and eye detection on mobile devices. For example, our analysis suggests that gaze estimation on mobile devices should rely less on full-face images but estimate gaze based on the eyes only instead.

ABSTRACT

Commodity mobile devices are now equipped with high-resolution front-facing cameras, allowing applications in biometrics (e.g., FaceID in the iPhone X), facial expression analysis, or gaze interaction. However, it is unknown how often users hold devices in a way that allows capturing their face or eyes, and how this impacts detection accuracy. We collected 25,726 in-the-wild photos, taken from the front-facing camera of smartphones as well as associated application usage logs. We found that the full face is visible about 29% of the time, and that in most cases the face is only partially visible. Furthermore, we identified an influence of users' current activity; for example, when watching videos, the eyes but not the entire face are visible 75% of the time in our dataset. We found that a state-of-the-art face detection algorithm performs poorly against photos taken from front-facing cameras. We discuss how these findings impact mobile applications that leverage face and eye detection, and derive practical implications to address state-of-the art's limitations.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

Author Keywords

Mobile Device; In the wild study; Face Detection; Gaze Estimation; Eye Tracking; Front-facing camera

INTRODUCTION

The availability of high-resolution front-facing cameras and an ever-increasing processing power of commodity smartphones pave the way for a wide range of applications impossible before. These include, for example, face detection (e.g., when taking pictures), facial expression analysis (e.g., for emotion detection), person identification (e.g., FaceID for authentication on iPhone X), attention analysis (e.g., for usability testing), or gaze estimation (e.g., for eye-based input). However, state-of-the-art computer vision methods for face detection and gaze estimation require full-face images [5, 20, 36, 41, 48, 52]. Despite the significant potential of these methods, it currently remains unclear how often users hold their device in a way that allows capturing their face or eyes and how different mobile tasks influence face and eye visibility.

To fill this gap, and thus inform the development of future mobile apps, we conducted a 2-weeks field study in which we continuously collected photos from the front-facing camera of 11 Android users along with usage logs. Our application collected 30,194 photos. 25,726 photos were approved by the participants and analyzed to find that users' entire face is visible only about 29% of the time, and that in the majority of cases only part of the face is visible (38.2%). In about 48% of the photos in our entire dataset, the whole face is not visible, but both eyes are. We also found that the type of app currently in use, has an impact on whether or not the face is visible in the front facing camera's view. For example, compared to other apps, the user's face is visible more often (49%) when using social networking apps. All of the collected meta data, usage logs, and 20,656 photos are available[1].

---

[1]Please contact the first author for a link to the dataset. Some photos included bystanders, and since we do not have approvals from them, we had to exclude these photos from the dataset, in addition to P11's photos who did not agree to publicly share his photos.

We evaluated a state-of-the-art face detection algorithm on our dataset and found that it succeeds in finding a face only one third of the time. Even in cases where an entire face was visible in our manual inspection (29% of the dataset), it only succeeded in detecting the user's face in 75% of those. We discuss the implications for mobile applications that rely on detecting the user's face in the front-facing camera view, and how researchers and practitioners can overcome the limitations of state-of-the-art approaches for face and eye detection.

This work makes the following contributions: (1) we provide a dataset of 20,656 photos taken from front-facing cameras on users' phones during daily use with time-synchronized smartphone usage logs. (2) We provide an analysis of face and eye visibility in our dataset; we gather insights about the impact of running applications and user behavior, as well as the performance of state-of-the-art face detection algorithms against our dataset. (3) We derive implications for designing future mobile applications that require face and eye detection.

RELATED WORK
Our research is related to previous work on (1) how users hold their phones, and (2) using the front-facing cameras of smartphones for face detection and gaze estimation.

How Users Hold Their Phones
Prior work suggested the way users hold their phones is influenced by the currently running app [10, 47]. Kim et al. identified eight holding postures for five mobile apps based on brainstorming sessions and interviews [32]. Le et al. reported on a study of smartphone holding postures when writing and reading text, and watching a video [34]. Knowledge about holding posture can be leveraged to improve the user experience, e.g., to switch between portrait and landscape mode. Taylor and Bove proposed using the holding pattern as an additional input modality [42] while Wimmer suggested a mobile phone could switch from camera to typing mode depending on whether it was held by one or two hands [47]. Finally, Kim et al. directed users to the intended mobile app by recognizing their grip based on an array of capacitive touch sensors [32].

Although we do not investigate how users hold their smartphones, this body of previous work tells us that the currently running application, the smartphone holding posture and the context in which the user is interacting are all interrelated. This inspired us to examine how these factors influence face and eye detection on smartphones. Since the holding posture influences the angle between the camera and the face, we expect that similar factors that influence the holding posture also influence whether or not the user's face is seen from the front-facing camera's perspective.

Huang et al. collected a dataset in the lab consisting of 51 subjects, each holding a *tablet* in four defined body postures: standing, sitting, slouching, and lying, and 35 on-screen gaze locations [27]. In contrast, our dataset was collected in a two-weeks in-the-wild study and participants consequently behaved naturally. Furthermore, we look into the relationship between the activity and the visibility of the user's face, and we investigate smartphones, which are known to require different hand postures than those for tablets [10].

Wang et al. used photos from front-facing cameras for face logging to infer mental health [46]. We build on that work by collecting 5x more photos, classify them based on face visibility, and make them publicly available.

Face and Eyes Detection on Commodity Smartphones
An increasing number of applications use the integrated front-facing camera. For example, Cheng et al. improved changing the orientation of the screen content substantially by leveraging face detection from the front-facing camera [12]. In a followup work they proposed to rely on the user's grasp of the phone instead, citing drawbacks of face detection on smartphones caused by fingers blocking the front cameras, and device tilt [13]. Cui et al. proposed tagging videos with the viewer's emotions, captured using the front-facing camera [15], while SeeSaw captured reactions to video messages [45], and Pulse rated movies and videos by capturing the user's reaction through the front facing camera [6]. Mc-Duff et al. used digital cameras to detect cognitive stress [39] and photoplethysmographic systolic [38]. Yoo et al. proposed capturing the user's facial expressions through the front-facing camera for lifelogging applications [50]. EyeProtector and SmartPose leveraged users' face size in the front-facing camera view to warn users if they were too close to the screen [22, 35]. There is also a large body of work about authentication via face recognition on mobile devices [40], such as FaceID.

Early works on gaze estimation on mobile devices used an external camera [16]. More recently, gaze estimation on mobile devices has become feasible using front-facing cameras. Holland et al. introduced eye tracking for commodity mobile devices [24, 25]. Wood and Bulling improved over that in EyeTab where they used a model-based approach for gaze estimation that did not require calibration [49]. Hohlfeld et al. then evaluated EyeTab for multiple use cases [23]. Krafka et al. and Huang et al. introduced gaze estimation datasets for users of mobile devices in controlled settings [27, 33]. Huang et al. introduced ScreenGlint, which exploits the reflection of the screen on the user's cornea for gaze estimation on mobile phones [26]. In addition to gaze estimation for monitoring purposes, several works proposed interacting with smartphones using gaze. For example, previous work investigated eye gesture detection on unmodified tablets [18, 28, 36, 43, 51], and multiple authentication schemes used gaze gestures detected from the front-facing camera [29, 30, 31].

In most of these works, the device was not held naturally by users' but was fixed by using a stand [18, 23, 24, 26, 28, 33, 43, 49], by using a headset [37], or by having another person hold it [51]. Even in works where the user was holding the phone, the authors reported that detection failures were often due to users not holding the phone in a way that would show their face and eyes in the camera's view [27, 29, 30, 31, 36]. However, our work is first to study this problem and identify how and which user behaviors impact face and eye detection.

APPROACH AND IMPLEMENTATION
To investigate which aspects influence the visibility of the user's face and eyes in the front-facing camera's view, we implemented an Android application with three main components: (1) Photo and Data collection component which also
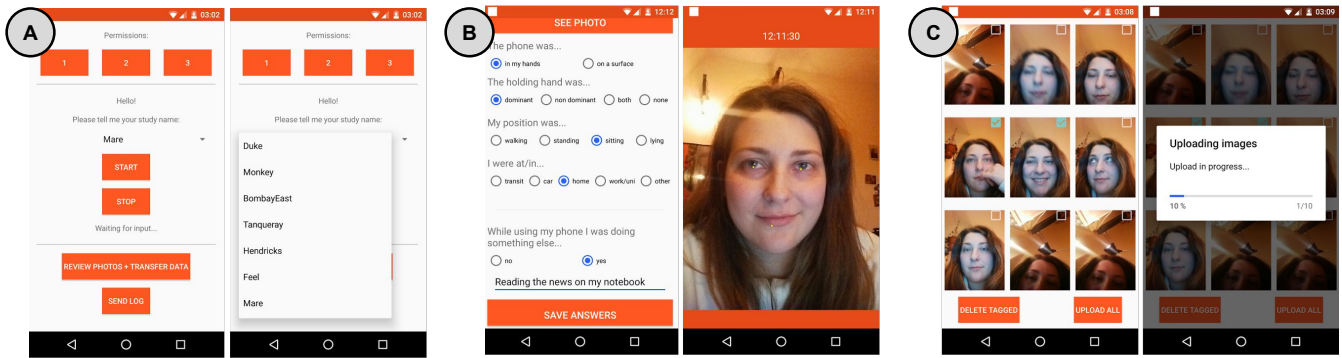
Figure 2. We used an Android app to study the aspects that influence the appearance of the user's face in the front-facing camera view. (A) shows the start page from which the participant gives our app the permission to launch background services that collect photos and data. (B) shows the experience sampling survey which asked participants to provide more context for the currently taken photo. (C) shows the photo review view from which participants selected which photos to share with us.

decided when to collect data, (2) Experience Sampling component for gathering additional information from the user, and (3) Photo Review component that allowed participants to review and filter out photos they did not wish to share.

Photo and Data Collection Component

This component consists of three background services: (1) A controller service for deciding when a photo should be taken, (2) a photo capture service that took photos from the front-facing camera, and (3) a data collection service that monitored and logged the sensor data. Once installed, the app asked for the permissions necessary to collect camera and usage data, and initialized the services. We allowed participants to manually stop and start the photo and data collection if desired. The photos were taken and the data was logged according to the *data collection triggers* that we explain below. The services kept running in the background as long as the phone screen was on, and were reinitialized on startup whenever the phone rebooted. Participants were neither interrupted by the photo and data collection, nor shown the camera feed while taking the photo. This also means that participants were not aware when a photo was going to be taken to prevent bias. In some cases, participants became aware after the photo was taken due to the experience sampling approach that we describe in the following section.

Triggers

Our Android service took photos at the following occasions:

- **ScreenOn trigger:** A photo was taken whenever the screen was turned on, e.g., right before entering a PIN to unlock the phone, and again 3 seconds afterwards. Since several authentication schemes for mobile phones rely on detecting the face or the eyes [29, 40], this was to study whether the face and eyes were visible right before users unlocked their phones. We were also interested to investigate whether face visibility stayed the same throughout the interaction session.

- **Notification trigger:** A photo was taken whenever a notification was received while the screen is on. Applications can make use of face or eye analysis when users review a new notification. For example, similar to how Cui et al. tagged videos [15], applications could tag the notification's content based on users' emotion inferred from facial expressions.

- **Application trigger:** A photo was taken whenever an application was launched or brought to the foreground (e.g., by clicking on a notification). As long as the application remained in the foreground additional photos were taken 3, 15 and 30 seconds afterwards. This was motivated by the fact that knowledge about which tasks (e.g., typing) influence the visibility of the user's face would present opportunities as to which tasks designers can leverage face and/or eye detection in. Also it would inform designers of computer vision applications if there is a need to guide users to holding the phone in an appropriate way before using their applications.

- **Orientation trigger:** A photo was taken whenever the phone orientation changed, and again three seconds later. Not only could the change of orientation be triggered based on the user's face [12], a change of orientation could also indicate that the user changed the context, task, or both.

- **Random trigger:** In addition, photos were taken at random occasions while the phone was active between 10 am and 10 pm, since this is the time period in which phones are used most frequently [9] and for the longest duration [44]. Up to six random triggers were activated daily during that period. The reason behind limiting the period and number of times the trigger was activated is to (1) avoid overwhelming the participants with many photos to review every day, and (2) reduce potential technical implications such as running out of space, battery consumption, etc.

Our selection of the intervals at which photos are taken is inspired by previous work by Böhmer et al. about smartphone apps usage [9]. We treat an application session as a "continuous period of time in which an application is both active and visible" [11, 19, 44]. This means that if a participant brought an application to the foreground after exiting it, it is considered a new session and hence photos are taken as described earlier.

Data Logging

We logged (1) the timestamp, (2) the full file path of the taken photo, (3) the trigger that caused the capture event, (4) the application in the foreground, (5) location data (latitude, longitude, road, PLZ), (6) readings from the ambient light sensor, and the accelerometer and gyroscope (x,y,z), (7) the phone's

orientation, (8) the screen's brightness, (9) battery level and whether or not the phone is charging; tethered charging of the phone could influence the user's posture.

Experience Sampling Component

We followed the experience sampling method [8] to query participants about some other aspects that we expected to influence the visibility of their face and eyes. The questionnaire asks for some additional information about the photo that was taken immediately before the questionnaire was shown (see Figure 2B). Participants were allowed to see the said photo, and were asked (1) if the phone was held or if it was lying on a surface. In the former case, the participant was asked (2) whether it was held with the dominant, non-dominant or both hands. Participants were also asked about (3) their posture, (4) their location (e.g., transit, work, etc.), and (5) whether they were exclusively using their smartphone or doing something else in parallel. In the latter case, participants were asked to describe the parallel activity. To this end, we showed participants a questionnaire via push notifications that they were allowed to dismiss or submit only partially filled [14]. In the former case, the questionnaire stayed in the notifications bar until it was filled. We showed a questionnaire after each *random trigger*, as explained before.

Photo Review Component

At 10 pm every evening, the app prompted the user to review the photos taken during that day and pick the ones they approved (see Figure 2C). This was necessary to ensure that participants sent us only the photos they were willing to share. After reviewing the photos, the photos and data logs were uploaded on a secured university FTP server, to which only the experimenter had access. The app required uploading the photos once a day and recommended using a WiFi connection.

FIELD STUDY

Although it was challenging to find participants willing to share daily photos taken inconspicuously from the front-facing cameras of their phones, we were able to recruit 11 participants (5 females). Their ages ranged from 19 to 34 years (Mean=26.4, StDev=4.2). Participants included a hotel manager, a waiter, a hospital employee, a project manager, a primary school teacher, three software developers, and three students. All participants were right-handed Android users and were compensated with an online shopping voucher. To avoid influencing their behavior, we installed the application on their own personal smartphones. Seven used smartphones with a front-facing camera in the upper right corner, the others were in the upper left corner. The exact models and front-facing camera resolutions of the smartphones are in Table 1.

Limitations

Having participants use their own smartphones was necessary to avoid unnatural behavior. While this means that they used different phones with different configurations and camera placements. It also means that the collected data is more ecologically valid, covers a larger diversity of phones, and represents users' natural behavior in the wild. Another limitation is that all participants were right-handed. Although this

ID	Phone Model	Camera Position	Resolution
1	Sony Xperia Z3 Compact	Right	2.2 MP
2	Sony Xperia Z	Left	2.2 MP
3	Sony Xperia Z3	Right	2.2 MP
4	Samsung Galaxy S7	Right	5.0 MP
5	Sony Xperia Z5 Compact	Left	5.0 MP
6	Sony Xperia Z2	Right	2.2 MP
7	Sony Xperia Z3	Right	2.2 MP
8	Samsung Galaxy S7	Right	5.0 MP
9	Nexus 5X	Left	5.0 MP
10	LG G3	Left	2.1 MP
11	Samsung Galaxy S5 Mini	Right	2.1 MP

Table 1. Model and camera specifications of participants' smartphones.

means we cannot draw generalizable quantitative conclusions from the data, we are nevertheless able to draw general qualitative insights. For example, we cannot claim that the eyes, but not the entire face, are generally visible 75% of the time when watching videos because that was the case in our dataset, however we can claim based on the trends among participants, there are many cases in which the user's face is not entirely visible when watching videos, and that this issue needs to be addressed if the face detection is required in this context.

Procedure

We invited each participant individually to our lab, where we first explained the experiment. Participants signed a consent form, then we installed the application on their phones. While all participants allowed us to analyze their data, 10 out of 11 allowed us to share the data and photos publicly. We then explained the application: we showed them the experience sampling questionnaires and told them to expect them multiple times a day, and we explained how to use the photo review component. Participants were instructed to answer the questionnaires as soon as possible and to choose the most suitable answers or leave them empty if they did not remember. Participants were asked to delete photos only if they did not wish us to see them or have them in a publicly available dataset. We strictly asked them not to delete any photos that they perceived to be redundant or irrelevant. Participants were asked to contact us in case of any problems or questions.

The data collection ran for two weeks. Afterwards, we had another meeting with the participants in which we uninstalled the application, collected demographic data, and conducted a semi-structured interview.

Photo Classification

In total, our application took 30,194 photos from the front-facing camera of the participants. Out of those, participants shared with us 25,726 photos. This means we analyzed 25,726 photos in addition to the associated logged data. The collected photos were classified into one of the following categories:

1. Whole Face: the case where the whole face is visible. This can be further categorized into:

    (a) Whole Face (all landmarks): the user's eyes and mouth are "detectable", i.e., they are not hidden (Figure 3A)

    (b) Whole Face (some landmarks): although the face is visible, not all eyes and mouth are visible e.g., obscured by hand or hair (Figure 3B and 3C)
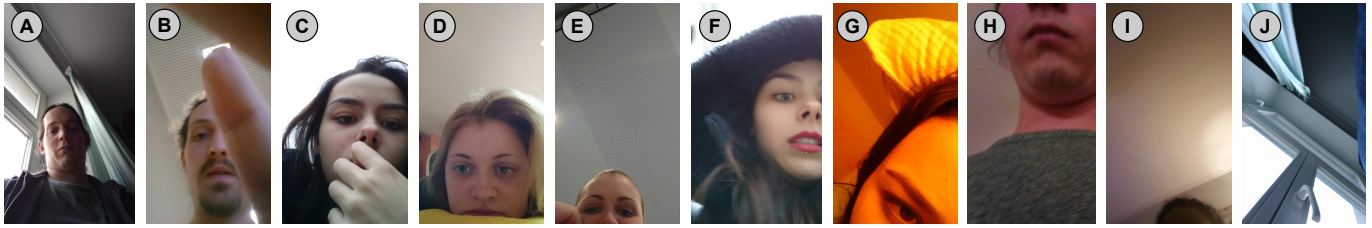
Figure 3. We collected photos where the whole face is visible (A), the whole face is in the view but some landmarks are obscured (B and C), both eyes and mouth are visible yet the entire face (e.g., the chin) is not visible (D), eyes are visible but mouth is not (E), one eye and the mouth are visible (F), only one eye is visible (G), only the mouth is visible (H), part of the face is visible but eyes and mouth invisible (I) and cases were no face is visible at all (J).

2. Partial Face: the case where *part* of the face can be seen. This case can be classified into further subcategories:

   (a) Partial Face (2 eyes, mouth): Although the face is partially out of the camera's range, both eyes and the user's mouth can be seen (Figure 3D).

   (b) Partial Face (2 eyes, no mouth): Both eyes are visible but the user's mouth is not (Figure 3E).

   (c) Partial Face (1 eyes, mouth): Only one eye and the mouth are visible (Figure 3F).

   (d) Partial Face (1 eyes, no mouth): Only one eye is visible, the mouth is not visible (Figure 3G).

   (e) Partial Face (0 eyes, mouth): None of the eyes are visible, but the mouth is visible (Figure 3H).

   (f) Partial Face (0 eyes, no mouth): Although the face is partially visible, neither the eyes nor the mouth are visible, e.g., only the forehead is visible (Figure 3I).

3. NoFace: the case where no face is detected at all (e.g., only the ceiling is visible in Figure 3J).

4. Both Eyes: this category includes cases where both eyes are visible, i.e., it combines 1a, 2a and 2b.

RESULTS

Overall, the whole face is visible 28.8% of the time (see Figure 4). This is in line with the Rice TabletGaze dataset that was collected in a lab and found that the whole face is visible in 30.8% of the cases [27]. Although there are many cases where the face is not visible at all (29.3%), there are even more cases where the face is partially visible (38.2%). More importantly, there are many cases where the entire face is not visible, but one or more of the user's eyes are visible. Figure 4 shows that the face and eyes visibility is almost consistent across Screen On, Notification, and Random triggers. But visibility is generally low when there is a switch in orientation. We also found that both eyes are overall visible 47.9% of the time in our dataset, and 54% of the time when using an application.

Investigating application usage further, we found that the face and eyes are more likely to be visible in social apps (e.g., Facebook, Instagram), messaging apps (e.g., Facebook Messenger, Whatsapp), reading and writing apps (e.g., pdf reader and text processing apps), video apps (e.g., YouTube), and games (see Figure 5). Many of these apps, such as video apps and games, typically run in landscape orientations. But when looking into the distribution of face and eye visibility across the landscape and portrait orientations, we find that using the smartphone in portrait mode is more likely to result in the face and eyes appearing in the front-facing camera's view (see Figure 6).

Users could be holding their phones with their dominant, non-dominant, or both hands. We found that the face and eyes are more likely to be visible when holding the hand with both hands (see Figure 7). Holding the smartphone with the dominant hand seems to result in higher visibility of the face and the eyes. Finally the face and eyes are rarely visible when the smartphone is not held by the user, but rather put on a surface (e.g., watching a video or playing music while doing something else). Figure 8 shows that the whole face is rarely completely visible when lying, yet visibility of eyes and whole face is fairly consistent when standing or walking. The face was slightly more likely to be visible when exclusively using the phone (32%) compared to when doing something in parallel (27%) such as reading news, watching T.V., etc.

In the interviews, we asked participants to indicate on a 5-point scale whether their participation in the study influenced their regular smartphone usage behavior (5=significantly changed my behavior;1=did not change my behavior at all). Participants indicated that they have not changed their behavior significantly (Median=2, StDev=1.3). P2 mentioned not using the smartphone as usual to avoid having to delete many pictures. P8 said that he avoided taking the smartphone to the bathroom during the study, while P11 sometimes avoided it intentionally when others were around. Other than that, there was no evidence that participants' sense of security influenced their behavior and hence the results. Overall participants deleted 4,468 out of 30,194. The vast majority of those (3,348 photos) were deleted by P11, who was very careful not to share photos of others. Otherwise, participants mainly deleted photos that were taken in private or awkward contexts.

Faces of Other Persons

In several cases more than one face was (partially) visible. This can be attributed to screen sharing, shoulder surfing [17], or due to a passerby unintentionally coming in the field of view of the camera. In the interviews, multiple participants reported feeling obliged to delete photos in which someone else appeared, since that other person was not a participant in our study and hence did not approve sharing the photo. In our dataset, we have 99 photos in which the face of another person was (partially) visible to the camera. We included these photos in the analysis, but excluded them from the public dataset. The interviews revealed that out of those, only 11 were strangers while the rest were acquaintances of the participants.

Reasons Behind Obscured Faces

The major reason for missing facial landmarks was that these were out of the camera's field of view. While future smart-
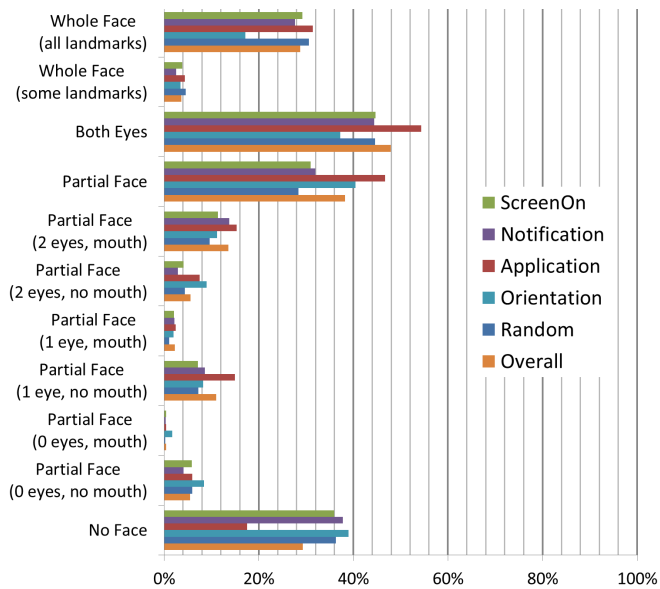
Figure 4. The figure shows the distribution of photos for each trigger. The face is almost equally visible when turning the screen on, running an application, and in random events. Visibility is generally low when there is a switch in orientation. However both eyes are visible in many cases when an application is being used.

phones are likely to adopt cameras with wider angle lenses, a multitude of other objects were found to obscure the user's facial landmarks when interacting with their smartphones. We logged many cases where part of the user's face was obscured by the user's hand or fingers (see Figures 3B and 3C). The second most occurring reason was due to wearing high collars or scarfs. On the other hand, one participant was on vacation during the study and was wearing sun glasses for a long time, which in turn obscured his eyes. This means that external factors such as weather conditions could heavily influence the visibility of the user's faces and eyes in the front facing camera view. We also logged cases where the face was partially obscured due to smoking, drinking, holding a telephone's headset (e.g., making calls at work), or eating.

Face Detection Performance on the Dataset

We evaluated OpenFace [5], a state-of-the-art framework for face detection, against our dataset. OpenFace uses a state-of-the-art constrained local neural field model, and was developed with the aim of detecting faces in poor lighting conditions and in the presence of extreme poses or occlusions [4]. As shown in Figure 9, the OpenFace framework failed to detect faces in the vast majority of the dataset; faces were detected in approximately one third of the dataset (33.44%). In the many cases where the face was present but some of its landmarks were hidden (e.g., by fingers, like in Figure 3B), the OpenFace framework successfully identified a face in only 18.87% of the cases. Performance was also very low when the mouth, or one or both eyes were hidden (between 8.21% and 34.53%). Even in cases where the whole face was visible, only 75.91% of the faces were successfully detected, due to motion blur and poor light conditions. While this is a relatively high value compared to the other cases, it still means that in one out of four times, OpenFace will fail to detect the face appearing in the view of the front-facing camera.
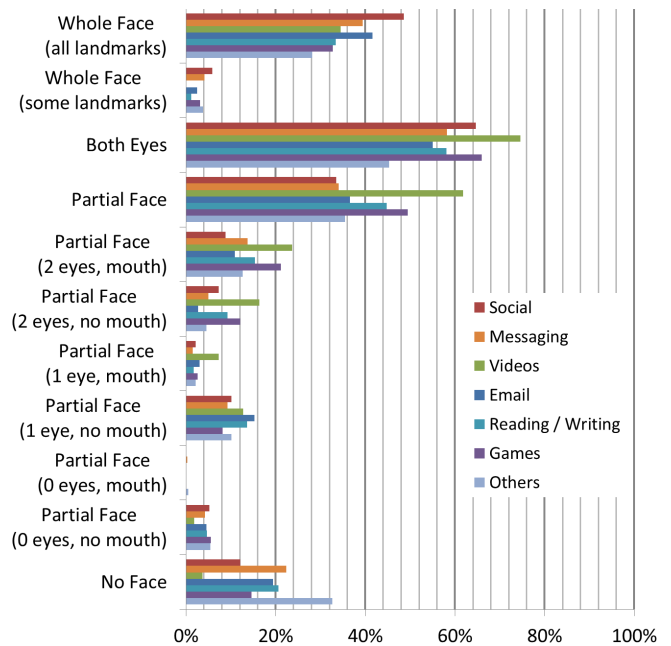

Figure 5. The figure shows the apps that are associated with the highest visibility of the face and eyes. Compared to other apps, both eyes are more likely to be visible when gaming, using social networking apps, messaging an watching videos.

These results suggest that to exploit the pictures from a front-facing camera for the suggested scenarios, there is a need for, either, novel face detection and gaze estimation approaches that do not rely on a full-face, but rather make use of the partially visible features of the user's face; or means to make the user hold their phone appropriately need to be investigated.

DISCUSSION

While in some cases the user might be aware that they should make their face visible to the camera (e.g., explicit authentication), in many other applications such as emotion detection, gaze estimation, and continuous authentication the detection needs to happen implicitly. Our results indicate that the whole face is rarely visible in the front-facing camera's view. In our dataset, that's the case only 28.8% of the time. We found that there are many cases in which the face is not visible, but both eyes are clearly visible. We also found that the user's face and eyes were visible the most when our participants were using certain applications, in particular, watching videos, using social apps, messaging apps, email apps, reading/writing apps, and when gaming. We found that the face and eyes are more likely to be visible when holding the phone with both hands. Finally, we found that OpenFace, a state-of-the-art face detection framework, succeeds only one third of the time in our dataset. Even when the whole face is shown in the front-facing camera view, OpenFace is able to detect a face in about 75% of the time. The reason behind the poor performance of OpenFace against our dataset is that in-the-wild photos of faces taken from from front-facing cameras are taken from diverse angles, have very diverse lighting conditions, could contain faces that are partially occluded because of the limited camera view or due to the user's hair, hands, fingers, or clothes, or distorted due to motion blur.
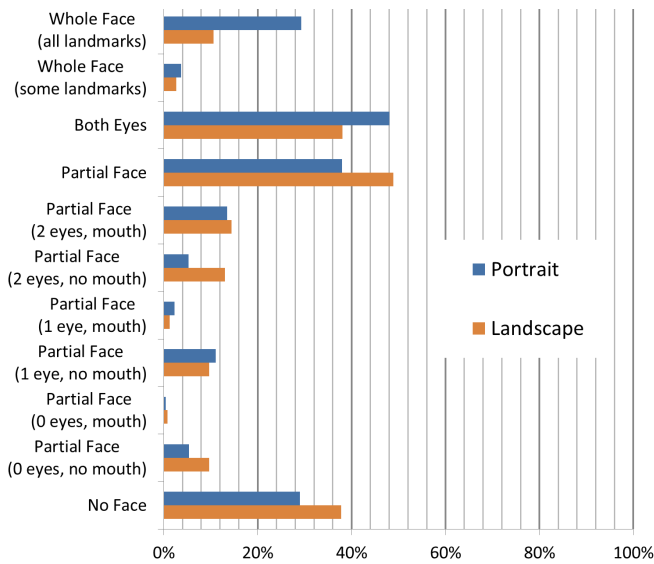
Figure 6. Compared to the portrait orientation, holding the phone in landscape orientation seems to result in more cases in which the face is not visible in the camera's view. However the difference in the visibility of both eyes is not as big.
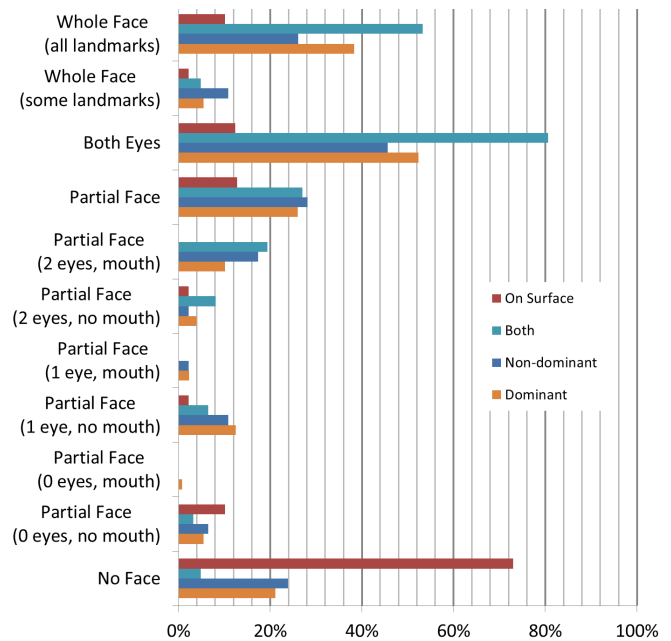

Figure 7. Face and eyes are rarely visible when the device is on a surface. There is a higher probability the face is visible to the camera when holding the phone with the dominant hand, and even higher probability that the face and eyes are visible when it is held with both hands.

Gaze Estimation Should not Depend on Face Detection

One of the major conclusions from this study is that the user's whole face is not always visible in the front-facing camera's view. This finding is not only important for applications that require face detection, but also for those that employ eye detection or tracking. Many of state of the art algorithms for gaze estimation require detecting the entire face as a prerequisite for eye tracking [20, 36, 41, 48, 52]. On the other hand, in many cases where the face was not visible, one or more eyes were visible. More specifically, both eyes were visible in 47.9% of our dataset. Given that the face is not always visible to front-facing cameras of smartphones, gaze estimation for mobile devices should not rely entirely on detecting the full face, but fall back to using the eyes that are more often visible.

Therefore, while the aforementioned methods can be useful for many other contexts, such as public displays or desktop settings, they are less suited for mobile devices. We recommend using approaches that do not require full face detection for eye tracking on mobile devices. For example, an alternative could be gaze estimation based on the whites of the eyes [7].

Guide the User to a Suitable Holding Posture

While an obvious solution is to use cameras with wider angle lenses to increase the chances that the user's face is shown, there are still cases where a wide-lens angle might not capture the face, for example if the user is lying, or if the face is obscured by an object or the user's hair or hand (see Figure 3B).

Another solution that would work on off-the-shelf mobile devices, is to guide the user to holding the smartphone in an optimal posture that would make the face visible. This can be achieved by borrowing concepts from public display research, in which a common problem is that users need to position themselves in an optimal spot at the outset of interaction. For example, GazeHorizon guides its users to the optimal position for interaction via gaze by using mirror video feed overlaid

with a face outline [53]. Similarly, a mobile app that requires a visible face could show a video feed and overlay a face outline to guide the user. While this approach is likely accurate in guiding users to hold the phone in a way that shows their faces, this form of explicit guidance might not always be feasible due to, for example, the visual clutter caused by overlaying the video feed on the the current view. In situations where showing a video feed is not optimal because, for example, the user is typing, reading text, etc. an implicit guidance method could be employed. For example, similar to GravitySpot [2] which guides users of public displays to certain positions, the interface of a mobile device could implicitly guide the user's holding posture by blurring the view when the user's face is no longer detected by the front-facing camera, and deblurr when the face is visible.

Know When to Track Eyes and Faces

Our results show that the user's face and eyes are more likely to be visible when using particular types of apps. This presents an opportunity for researchers and practitioners to exploit this for building affective applications that respond to users' facial expressions. For example, there has been a recent growing interest in augmenting messaging apps with context and emotional awareness [21]. Future work could leverage facial expressions to detect the user's emotions to share them on social media websites, respond to posts and messages, or log them for reflection. Our results also indicate that both eyes are visible the majority of the time while watching videos (see Figure 5). Gaze estimation could be leveraged in advertisement videos to collect insights about the user's interests.

Furthermore, our results could be useful for designers of computer vision mobile applications by knowing when *not* to employ face and eye detection. Face and eye detection tasks
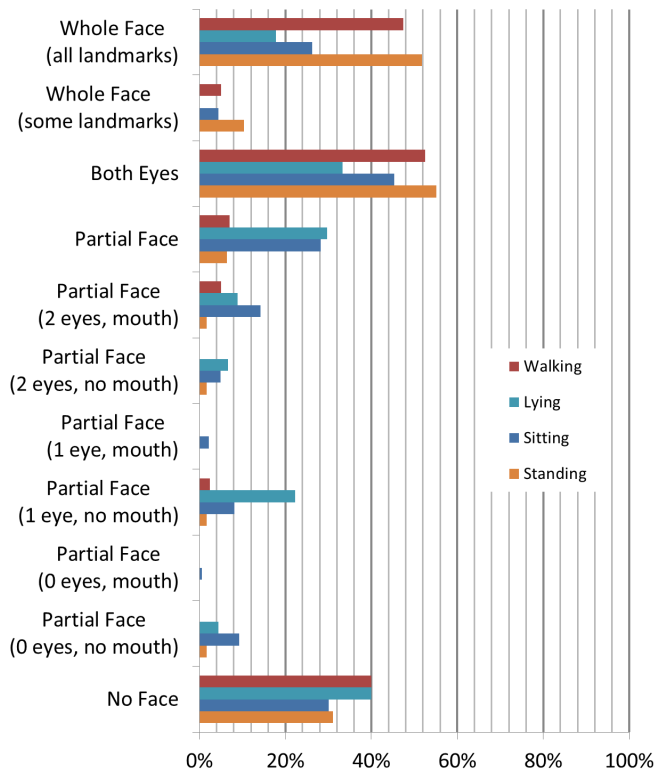
Figure 8. The face and eyes are most likely to be visible while standing and while walking.
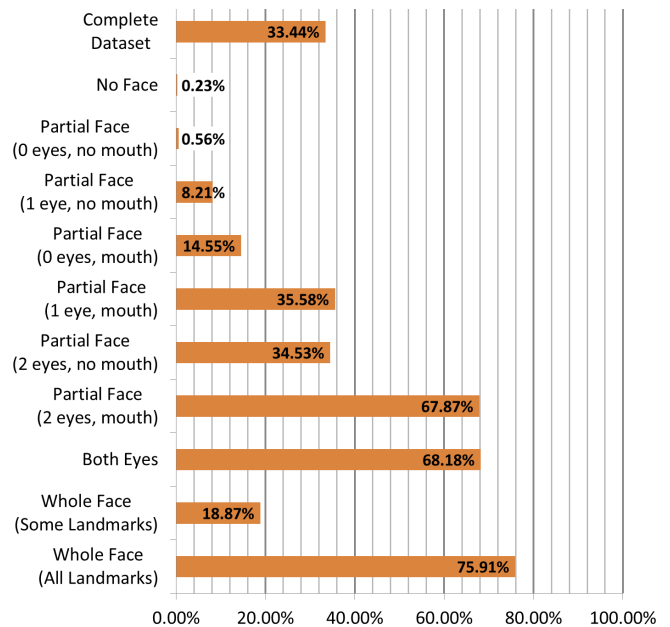

Figure 9. OpenFace is a framework that employs state-of-the-art methods for detecting faces in the wild. Yet, it detects faces in only one third of our dataset. Even in case where the entire face is visible to the front-facing camera, it succeeds only 75.9% of the time, which means that it would fail once every four times even though the face is perfectly visible.

are computationally expensive tasks that can drain the battery or even overheat the device. Our results indicate that there are many cases where it is very unlikely that the user's face or eyes will be visible. For example, when the device is on a surface (Figure 7), or when the user is lying (Figure 8). Mobile apps could use the inertial sensors or knowledge about usage patterns to determine when *not* to track the eyes.

Design with Privacy Implications in Mind
"My Facebook app has a permission to use my camera, does this mean that it could be taking pictures all along?". Some of our participants were surprised to see that our application could inconspicuously take photos without showing the camera feed. As commodity devices become more capable of performing visual computation tasks, they also become platforms for collecting private data. This highlights the importance of making computer vision mobile apps transparent by, for example, making them open source. While the average user might not be well equipped to check if the currently used mobile app violates privacy, developers and enthusiasts have found vulnerabilities in open source software in the past, which in turn resulted in average users learning about these issues. We argue that we need a mechanism, similar to that of certificate authorities used in HTTPS, to *certify* mobile apps that do not violate the user's privacy.

Several approaches have been proposed to nudge users when certain permissions or private data is shared with an app [1, 3]. These approaches can be adopted to warn the user when the camera feed is being processed by an app, and in turn encourage using mobile apps that employ face and eyes detection.

Furthermore, our participants tried to delete most of the photos that included people other than them, due to their fear of violating that person's privacy. This also has implications on face and eye detection apps on mobile devices; if a user allows a device to take photos of her face, it does not mean that photos of other people's faces could also be unintentionally taken. Hence, future apps that employ face and eye detection should automatically filter out photos in which a person, other than the user, is visible in the camera view.

Rather than Storing them, Process Photos on the Fly
It was challenging to find participants willing to participate. This could be attributed to the effort needed for reviewing the photos before sharing them, but could also be due to privacy concerns. We recommend designers of systems that employ face and eye detection to process the frames in real time rather than storing them locally or on the cloud. For future in-the-wild studies that involve photos taken from the front facing cameras, we recommend researchers to allow participants to review the data before sharing it, offer them to opt out from and pause/resume participation at any time during the study.

One possible way to overcome privacy issues in future work is to log the positions of facial landmarks rather than storing the actual photo. However this means that the detection has to be highly accurate for the data to be reliable. Future work should investigate how to collect accurate yet privacy-respectful data.

Another interesting direction for future work is to study the factors that influence adoption of apps that leverage face and eye detection. For example, while Android offered FaceUnlock long ago, studying whether or not users will adopt the FaceID on the long run could reveal factors that influence the user's trust in face and eye detection apps.

CONCLUSION

In this work we collected a dataset of 25,726 photos taken from front-facing cameras of smartphone in the wild and analysed it with respect to face and eye visibility. We are happy to share the dataset for the community to build over our work; please contact the first author for a link to the dataset. Being the first dataset of its kind, we gained several insights. Most importantly, we showed that the user's face and eyes are not always visible in the front-facing camera view, and not always detectable by state-of-the-art face detection methods. We found that the currently running application, the hand used to carry the phone, and the state of the user influenced the visibility of the user's eyes and face. There are many cases where the user's eyes, but not the whole face, were visible. We discussed how the results motivate the need for new computer vision algorithms that rely on the eyes only and not necessarily on the visibility of the full face. We also discussed how applications could guide the user to revealing her face to the front-facing camera, directions for exploiting the dataset and the results of this study, privacy implications, and challenges of this type of research.

REFERENCES

1. Hazim Almuhimedi, Florian Schaub, Norman Sadeh, Idris Adjerid, Alessandro Acquisti, Joshua Gluck, Lorrie Faith Cranor, and Yuvraj Agarwal. 2015. Your Location Has Been Shared 5,398 Times!: A Field Study on Mobile App Privacy Nudging. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 787–796. DOI:http://dx.doi.org/10.1145/2702123.2702210

2. Florian Alt, Andreas Bulling, Gino Gravanis, and Daniel Buschek. 2015. GravitySpot: Guiding Users in Front of Public Displays Using On-Screen Visual Cues. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (UIST '15)*. ACM, New York, NY, USA, 47–56. DOI:http://dx.doi.org/10.1145/2807442.2807490

3. Rebecca Balebako, Jaeyeon Jung, Wei Lu, Lorrie Faith Cranor, and Carolyn Nguyen. 2013. "Little Brothers Watching You": Raising Awareness of Data Leaks on Smartphones. In *Proceedings of the Ninth Symposium on Usable Privacy and Security (SOUPS '13)*. ACM, New York, NY, USA, Article 12, 11 pages. DOI:http://dx.doi.org/10.1145/2501604.2501616

4. Tadas Baltrusaitis, Peter Robinson, and Louis-Philippe Morency. 2013. Constrained Local Neural Fields for Robust Facial Landmark Detection in the Wild. In *Proceedings of the 2013 IEEE International Conference on Computer Vision Workshops (ICCVW '13)*. IEEE Computer Society, Washington, DC, USA, 354–361. DOI:http://dx.doi.org/10.1109/ICCVW.2013.54

5. Tadas Baltruåąaitis, Peter Robinson, and Louis-Philippe Morency. 2016. OpenFace: An open source facial behavior analysis toolkit. In *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*. 1–10. DOI:http://dx.doi.org/10.1109/WACV.2016.7477553

6. Xuan Bao, Songchun Fan, Alexander Varshavsky, Kevin Li, and Romit Roy Choudhury. 2013. Your Reactions Suggest You Liked the Movie: Automatic Content Rating via Reaction Sensing. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '13)*. ACM, New York, NY, USA, 197–206. DOI:http://dx.doi.org/10.1145/2493432.2493440

7. Pascal Bérard, Derek Bradley, Maurizio Nitti, Thabo Beeler, and Markus Gross. 2014. High-quality Capture of Eyes. *ACM Trans. Graph.* 33, 6, Article 223 (Nov. 2014), 12 pages. DOI:http://dx.doi.org/10.1145/2661229.2661285

8. Niels Van Berkel, Denzil Ferreira, and Vassilis Kostakos. 2017. The Experience Sampling Method on Mobile Devices. *ACM Comput. Surv.* 50, 6, Article 93 (Dec. 2017), 40 pages. DOI:http://dx.doi.org/10.1145/3123988

9. Matthias Böhmer, Brent Hecht, Johannes Schöning, Antonio Krüger, and Gernot Bauer. 2011. Falling Asleep with Angry Birds, Facebook and Kindle: A Large Scale Study on Mobile Application Usage. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '11)*. ACM, New York, NY, USA, 47–56. DOI:http://dx.doi.org/10.1145/2037373.2037383

10. Christina Bröhl, Alexander Mertens, and Martina Ziefle. 2017. *How Do Users Interact with Mobile Devices? An Analysis of Handheld Positions for Different Technology Generations*. Springer International Publishing, Cham, 3–16. DOI:http://dx.doi.org/10.1007/978-3-319-58536-9_1

11. Juan Pablo Carrascal and Karen Church. 2015. An In-Situ Study of Mobile App & Mobile Search Interactions. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2739–2748. DOI:http://dx.doi.org/10.1145/2702123.2702486

12. Lung-Pan Cheng, Fang-I Hsiao, Yen-Ting Liu, and Mike Y. Chen. 2012. iRotate: Automatic Screen Rotation Based on Face Orientation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 2203–2210. DOI:http://dx.doi.org/10.1145/2207676.2208374

13. Lung-Pan Cheng, Meng Han Lee, Che-Yang Wu, Fang-I Hsiao, Yen-Ting Liu, Hsiang-Sheng Liang, Yi-Ching Chiu, Ming-Sui Lee, and Mike Y. Chen. 2013. IrotateGrasp: Automatic Screen Rotation Based on Grasp of Mobile Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 3051–3054. DOI: http://dx.doi.org/10.1145/2470654.2481424

14. Sunny Consolvo and Miriam Walker. 2003. Using the Experience Sampling Method to Evaluate Ubicomp Applications. *IEEE Pervasive Computing* 2, 2 (April 2003), 24–31. DOI: http://dx.doi.org/10.1109/MPRV.2003.1203750

15. Yanqing Cui, Jari Kangas, Jukka Holm, and Guido Grassel. 2013. Front-camera Video Recordings As Emotion Responses to Mobile Photos Shared Within Close-knit Groups. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 981–990. DOI: http://dx.doi.org/10.1145/2470654.2466125

16. Heiko Drewes, Alexander De Luca, and Albrecht Schmidt. 2007. Eye-gaze Interaction for Mobile Phones. In *Proceedings of the 4th International Conference on Mobile Technology, Applications, and Systems and the 1st International Symposium on Computer Human Interaction in Mobile Technology (Mobility '07)*. ACM, New York, NY, USA, 364–371. DOI: http://dx.doi.org/10.1145/1378063.1378122

17. Malin Eiband, Mohamed Khamis, Emanuel von Zezschwitz, Heinrich Hussmann, and Florian Alt. 2017. Understanding Shoulder Surfing in the Wild: Stories from Users and Observers. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 4254–4265. DOI: http://dx.doi.org/10.1145/3025453.3025636

18. Hanene Elleuch, Ali Wali, Anis Samet, and Adel M. Alimi. 2016. *A Real-Time Eye Gesture Recognition System Based on Fuzzy Inference System for Mobile Devices Monitoring*. Springer International Publishing, Cham, 172–180. DOI: http://dx.doi.org/10.1007/978-3-319-48680-2_16

19. Denzil Ferreira, Jorge Goncalves, Vassilis Kostakos, Louise Barkhuus, and Anind K. Dey. 2014. Contextual Experience Sampling of Mobile Application Micro-usage. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services (MobileHCI '14)*. ACM, New York, NY, USA, 91–100. DOI: http://dx.doi.org/10.1145/2628363.2628367

20. Jari Hannuksela, Pekka Sangi, Markus Turtinen, and Janne Heikkilä. 2008. *Face Tracking for Spatially Aware Mobile User Interfaces*. Springer Berlin Heidelberg, Berlin, Heidelberg, 405–412. DOI: http://dx.doi.org/10.1007/978-3-540-69905-7_46

21. Mariam Hassib, Daniel Buschek, Paweł W. Wozniak, and Florian Alt. 2017. HeartChat: Heart Rate Augmented Mobile Chat to Support Empathy and Awareness. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 2239–2251. DOI: http://dx.doi.org/10.1145/3025453.3025758

22. Jimmy Ho, Reinhard Pointner, Huai-Chun Shih, Yu-Chih Lin, Hsuan-Yu Chen, Wei-Luan Tseng, and Mike Y. Chen. 2015. EyeProtector: Encouraging a Healthy Viewing Distance when Using Smartphones. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. ACM, New York, NY, USA, 77–85. DOI:http://dx.doi.org/10.1145/2785830.2785836

23. Oliver Hohlfeld, André Pomp, Jó Ágila Bitsch Link, and Dennis Guse. 2015. On the Applicability of Computer Vision Based Gaze Tracking in Mobile Scenarios. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. ACM, New York, NY, USA, 427–434. DOI: http://dx.doi.org/10.1145/2785830.2785869

24. Corey Holland, Atenas Garza, Elena Kurtova, Jose Cruz, and Oleg Komogortsev. 2013. Usability Evaluation of Eye Tracking on an Unmodified Common Tablet. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13)*. ACM, New York, NY, USA, 295–300. DOI: http://dx.doi.org/10.1145/2468356.2468409

25. Corey Holland and Oleg Komogortsev. 2012. Eye Tracking on Unmodified Common Tablets: Challenges and Solutions. In *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '12)*. ACM, New York, NY, USA, 277–280. DOI: http://dx.doi.org/10.1145/2168556.2168615

26. Michael Xuelin Huang, Jiajia Li, Grace Ngai, and Hong Va Leong. 2017a. ScreenGlint: Practical, In-situ Gaze Estimation on Smartphones. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 2546–2557. DOI: http://dx.doi.org/10.1145/3025453.3025794

27. Qiong Huang, Ashok Veeraraghavan, and Ashutosh Sabharwal. 2017b. TabletGaze: dataset and analysis for unconstrained appearance-based gaze estimation in mobile tablets. *Machine Vision and Applications* 28, 5 (01 Aug 2017), 445–461. DOI: http://dx.doi.org/10.1007/s00138-017-0852-4

28. Jari Kangas, Deepak Akkil, Jussi Rantala, Poika Isokoski, Päivi Majaranta, and Roope Raisamo. 2014. Gaze Gestures and Haptic Feedback in Mobile Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 435–438. DOI: http://dx.doi.org/10.1145/2556288.2557040

29. Mohamed Khamis, Florian Alt, Mariam Hassib, Emanuel von Zezschwitz, Regina Hasholzner, and Andreas Bulling. 2016. GazeTouchPass: Multimodal Authentication Using Gaze and Touch on Mobile Devices. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. ACM, New York, NY, USA, 2156–2164. DOI: http://dx.doi.org/10.1145/2851581.2892314

30. Mohamed Khamis, Regina Hasholzner, Andreas Bulling, and Florian Alt. 2017a. GTmoPass: Two-factor Authentication on Public Displays Using Gaze-touch Passwords and Personal Mobile Devices. In *Proceedings of the 6th ACM International Symposium on Pervasive Displays (PerDis '17)*. ACM, New York, NY, USA, Article 8, 9 pages. DOI: http://dx.doi.org/10.1145/3078810.3078815

31. Mohamed Khamis, Mariam Hassib, Emanuel von Zezschwitz, Andreas Bulling, and Florian Alt. 2017b. GazeTouchPIN: Protecting Sensitive Data on Mobile Devices using Secure Multimodal Authentication. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction (ICMI 2017)*. ACM, New York, NY, USA, 5. DOI: http://dx.doi.org/10.1145/3136755.3136809

32. Kee-Eung Kim, Wook Chang, Sung-Jung Cho, Junghyun Shim, Hyunjeong Lee, Joonah Park, Youngbeom Lee, and Sangryong Kim. 2006. Hand Grip Pattern Recognition for Mobile User Interfaces. In *Proceedings of the 18th Conference on Innovative Applications of Artificial Intelligence - Volume 2 (IAAI'06)*. AAAI Press, 1789–1794. http://dl.acm.org/citation.cfm?id=1597122.1597138

33. Kyle Krafka, Aditya Khosla, Petr Kellnhofer, Harini Kannan, Suchendra Bhandarkar, Wojciech Matusik, and Antonio Torralba. 2016. Eye Tracking for Everyone. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2176–2184. DOI: http://dx.doi.org/10.1109/CVPR.2016.239

34. Huy Viet Le, Sven Mayer, Katrin Wolf, and Niels Henze. 2016. Finger Placement and Hand Grasp During Smartphone Interaction. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. ACM, New York, NY, USA, 2576–2584. DOI: http://dx.doi.org/10.1145/2851581.2892462

35. Hosub Lee, Young Sang Choi, Sunjae Lee, and Eunsoo Shim. 2013. Smart Pose: Mobile Posture-aware System for Lowering Physical Health Risk of Smartphone Users. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13)*. ACM, New York, NY, USA, 2257–2266. DOI: http://dx.doi.org/10.1145/2468356.2468747

36. Yinghui Li, Zhichao Cao, and Jiliang Wang. 2017. Gazture: Design and Implementation of a Gaze Based Gesture Control System on Tablets. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 74 (Sept.

2017), 17 pages. DOI: http://dx.doi.org/10.1145/3130939

37. Alex Mariakakis, Jacob Baudin, Eric Whitmire, Vardhman Mehta, Megan A. Banks, Anthony Law, Lynn Mcgrath, and Shwetak N. Patel. 2017. PupilScreen: Using Smartphones to Assess Traumatic Brain Injury. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 81 (Sept. 2017), 27 pages. DOI: http://dx.doi.org/10.1145/3131896

38. Daniel McDuff, Sarah Gontarek, and Rosalind W. Picard. 2014. Remote Detection of Photoplethysmographic Systolic and Diastolic Peaks Using a Digital Camera. *IEEE Transactions on Biomedical Engineering* 61, 12 (Dec 2014), 2948–2954. DOI: http://dx.doi.org/10.1109/TBME.2014.2340991

39. Daniel J. McDuff, Javier Hernandez, Sarah Gontarek, and Rosalind W. Picard. 2016. COGCAM: Contact-free Measurement of Cognitive Stress During Computer Tasks with a Digital Camera. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 4000–4004. DOI: http://dx.doi.org/10.1145/2858036.2858247

40. Tolga Soyata, Rajani Muraleedharan, Colin Funai, Minseok Kwon, and Wendi Heinzelman. 2012. Cloud-Vision: Real-time face recognition using a mobile-cloudlet-cloud acceleration architecture. In *2012 IEEE Symposium on Computers and Communications (ISCC)*. 000059–000066. DOI: http://dx.doi.org/10.1109/ISCC.2012.6249269

41. Yusuke Sugano, Yasuyuki Matsushita, and Yoichi Sato. 2014. Learning-by-Synthesis for Appearance-Based 3D Gaze Estimation. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*. 1821–1828. DOI: http://dx.doi.org/10.1109/CVPR.2014.235

42. Brandon T. Taylor and V. Michael Bove, Jr. 2009. Graspables: Grasp-recognition As a User Interface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA, 917–926. DOI: http://dx.doi.org/10.1145/1518701.1518842

43. Vytautas Vaitukaitis and Andreas Bulling. 2012. Eye Gesture Recognition on Portable Devices. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*. ACM, New York, NY, USA, 711–714. DOI: http://dx.doi.org/10.1145/2370216.2370370

44. Niels van Berkel, Chu Luo, Theodoros Anagnostopoulos, Denzil Ferreira, Jorge Goncalves, Simo Hosio, and Vassilis Kostakos. 2016. A Systematic Assessment of Smartphone Usage Gaps. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 4711–4721. DOI: http://dx.doi.org/10.1145/2858036.2858348

45. Gina Venolia, John C. Tang, and Kori Inkpen. 2015. SeeSaw: I See You Saw My Video Message. In *Proceedings of the 17th International Conference on*

Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15). ACM, New York, NY, USA, 244–253. DOI:
http://dx.doi.org/10.1145/2785830.2785847

46. Rui Wang, Andrew T. Campbell, and Xia Zhou. 2015. Using Opportunistic Face Logging from Smartphone to Infer Mental Health: Challenges and Future Directions. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers (UbiComp/ISWC'15 Adjunct)*. ACM, New York, NY, USA, 683–692. DOI:
http://dx.doi.org/10.1145/2800835.2804391

47. Raphael Wimmer. 2011. Grasp Sensing for Human-computer Interaction. In *Proceedings of the Fifth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '11)*. ACM, New York, NY, USA, 221–228. DOI:
http://dx.doi.org/10.1145/1935701.1935745

48. Erroll Wood, Tadas Baltruaitis, Xucong Zhang, Yusuke Sugano, Peter Robinson, and Andreas Bulling. 2015. Rendering of Eyes for Eye-Shape Registration and Gaze Estimation. In *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV) (ICCV '15)*. IEEE Computer Society, Washington, DC, USA, 3756–3764. DOI:
http://dx.doi.org/10.1109/ICCV.2015.428

49. Erroll Wood and Andreas Bulling. 2014. EyeTab: Model-based Gaze Estimation on Unmodified Tablet Computers. In *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '14)*. ACM, New York, NY, USA, 207–210. DOI:
http://dx.doi.org/10.1145/2578153.2578185

50. Chungkuk Yoo, Jaemyung Shin, Inseok Hwang, and Junehwa Song. 2013. FaceLog: Capturing User's Everyday Face Using Mobile Devices. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication (UbiComp '13 Adjunct)*. ACM, New York, NY, USA, 163–166. DOI:
http://dx.doi.org/10.1145/2494091.2494145

51. Xiaoyi Zhang, Harish Kulkarni, and Meredith Ringel Morris. 2017. Smartphone-Based Gaze Gesture Communication for People with Motor Disabilities. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 2878–2889. DOI:
http://dx.doi.org/10.1145/3025453.3025790

52. Xucong Zhang, Yusuke Sugano, Mario Fritz, and Andreas Bulling. 2017. It's Written All Over Your Face: Full-Face Appearance-Based Gaze Estimation. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 2299–2308. DOI:
http://dx.doi.org/10.1109/CVPRW.2017.284

53. Yanxia Zhang, Jörg Müller, Ming Ki Chong, Andreas Bulling, and Hans Gellersen. 2014. GazeHorizon: Enabling Passers-by to Interact with Public Displays by Gaze. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '14)*. ACM, New York, NY, USA, 559–563. DOI:http://dx.doi.org/10.1145/2632048.2636071