May 2022

Real-Time Feedback to Improve Selfies for Low-Vision Users

N/A

Real-Time Feedback to Improve Selfies for Low-Vision Users

Abstract:

This publication describes systems and techniques, implemented on computing devices, directed to improving self-portrait (i.e., selfie) images of a user captured on a front-facing camera using real-time feedback. The systems and techniques may track the user's face and provide haptic, audio, and/or visual feedback to guide the user to position the computing device so that the user becomes positioned in a center of frame of the camera. In an aspect, a user interface may display visual indicators that flash over a viewfinder image of a camera displayed on a computing device display. The visual indicators may increase in brightness near a user's face to guide a user to the center of the frame of the camera. In another aspect, the user interface may display a high-contrast outline of the user's face and/or torso on the display to provide feedback to the user of their position in the frame. In another aspect, the user may receive an audio detail description of what is in the viewfinder to confirm desired faces and objects are included. Through such systems and techniques, a user can take a high-quality self-portrait even when they have limited or no ability to see a display screen.

Keywords:

real-time feedback, facial tracking, front-facing camera, image sensor, self-portrait photographs, selfies, center of frame, auditory feedback, haptic feedback, visual/motion feedback, visual indicator flash, selfie optimization, audio detail description, high-contrast outline, low vision user, screen reader, assistive technology, accessible technology, audio assistance, latency, accessibility, voice-based accessibility service

Background:

Front-facing cameras on computing devices allow a user of the computing device to capture self-portrait photographs (i.e., selfies). When the user takes a selfie, they may look at a display of the computing device to include themselves and any other desired objects in an image. However, the user must look away from the display to the camera lens and then press a shutter button to take the selfie. This may induce movement of the computing device and degrade the quality of what the user intended to capture in the selfie. Low-vision users who use a voice-based accessibility service (e.g., screen reader) currently receive limited audio assistance when taking a selfie, making it a guessing game for a good photo. The latency associated with the limited audio assistance makes it even less useful. Further, if it is difficult for the low-vision user to see the display screen, the quality of the selfie may be poor with undesirable framing and angles.

Description:

This publication describes systems and techniques directed to improving selfie images of a user captured on a front-facing camera of a computing device using real-time feedback. The systems and techniques may track the user's face and provide haptic, audio, and visual feedback to guide the user to position the computing device so that the user becomes positioned in a center of frame of the camera. By so doing, the techniques described herein may enable the user to capture an attractive selfie without relying on the user to preview the selfie on a display of the computing device. Figure 1 illustrates an example computing device in which the described systems and techniques can be implemented.

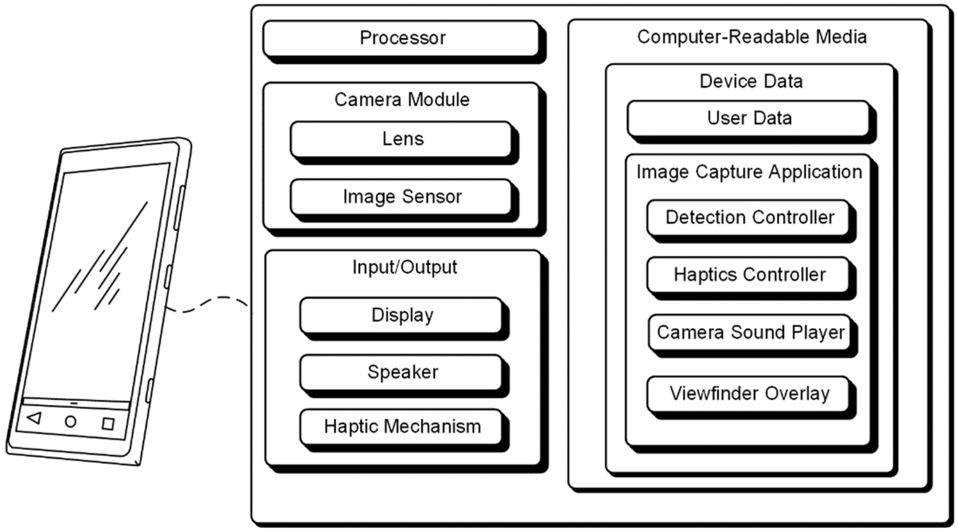

Figure 1

As illustrated in Figure 1, the computing device is a smartphone. In some implementations, the computing device can be a variety of other electronic devices (e.g., smartphones, tablets, cameras, tablet computers, computers, smart watches, and so forth). The computing device may include a processor, an input/output device (e.g., a display, a speaker, a haptic mechanism (e.g., DC motor that creates vibrations)), and a camera module. The camera module may include a lens and an image sensor. The computing device may also include a computer-readable media (CRM) that stores device data (e.g., user data, multimedia data, applications, an operating system). The device data may include instructions of an image capture application that, responsive to execution by the processor, cause the processor to perform operations described in this publication to utilize the image sensor, image processing capabilities, and feedback capabilities of the computing device for real-time selfie optimization of the front-facing camera. The image capture application may include a detection controller, a haptics controller, a camera sound player, and a viewfinder overlay.

For example, assume that a user initiates a front-facing camera of a smartphone (computing device) to take a selfie. The image capture application is initiated, and the detection controller may detect a face or a portion of a face in the viewing area of the front-facing camera. The detection controller may calculate a distance from the face to the center of the camera frame and provide it to the haptics controller.

The haptics controller may segment the camera's viewfinder into zones. As a user's face transitions between zones, the haptic controller may cause haptic feedback (i.e., vibrations) and audio feedback to change states. For example, the vibration of the phone and/or an audio pattern produced can become more rapid, intense, and high-pitched as the face approaches a zone at the center of the camera frame that is desirable for selfies. Haptics may indicate a direction (e.g., up, down, etc.) to move the smartphone and/or a distance (e.g., near, far). The haptics controller may compose haptic feedback for the haptic mechanism in combination with audio feedback. Audio feedback may be communicated to the user via the camera sound player and the speaker. A viewfinder overlay initiates communication of visual information via a user interface (UI) to a user. As discussed below, the UI may overlay a display and include a visual indicator that flashes, a checkmark, and/or a high-contrast outline. The user may receive the haptic, audio, and/or visual feedback and continue to adjust the position of the smartphone. The image capture application and its modules may repeat the process (e.g., every 100-200 milliseconds) to track the face, calculate a distance, and communicate feedback.

Figures 2a, 2b, and 2c illustrate an example UI overlaid on the viewfinder of a camera that is displayed on a smartphone display.

Figure 2a **Figure 2b** **Figure 2c**

As illustrated in Figure 2a, the UI controlled by the viewfinder overlay may display a visual indicator flash to help guide a user to the center of the camera frame. The visual indicator flash may be displayed in the UI as several translucent dashes that approximate a circle and increase in brightness or color when a user's face is nearby. Figure 2a illustrates a visual indicator flash closest to the face (in the lower right quadrant) highlighted to provide visual cues to the user. The user can adjust the position and angle of the smartphone to center themselves in the dashed circle. Figure 2b illustrates the UI displaying a confirmation to the user that they are in the center of the camera frame with a checkmark. Figure 2c illustrates that the UI (e.g., checkmark and dashed circle) may fade once the center of the camera frame is reached and the camera may take the photo without additional user input.

The UI may be displayed in addition to auditory feedback and/or haptic feedback to provide a rich user experience. Table 1 illustrates example combinations of auditory, haptic, and visual feedback provided to a user at three different moments: when the user is finding the center of the camera frame, when the center of the frame is achieved, and when a photo is taken by the camera module.

Key moments to provide feedback	Auditory feedback	Haptic feedback	Visual/motion feedback
01 State change	Upward Downward	Rising Falling	
02 Confirmation	Confirm	Confirm	
03 Take a photo	Shutter	Align with shutter sound	

Table 1

As illustrated in Table 1, the smartphone may be performing audio feedback of an increasingly high-pitched sound and/or the haptic feedback may increase as the user's face becomes more prominent in the center of the frame. For example, the haptic feedback may be a haptic pattern that increases in strength with perceived intensity increasing as the user's face approaches the center of the frame. Alternatively, the haptic pattern may be temporal with an increase in the pulse speed as the user's face approaches the center of the camera frame. When the smartphone UI displays a checkmark, the smartphone may provide confirmation sounds and confirmation haptic feedback. Finally, audio and haptic feedback can confirm to a low-vision user that a photo has been taken.

Figure 3 illustrates an example of using a high-contrast outline to guide a user to the center of the frame.

Figure 3

As illustrated in Figure 3, the image capture application may provide a high-contrast outline in a UI displayed over the viewfinder of the camera. For example, assume that a user initiates a front-facing camera of a smartphone to take a selfie. The image capture application is initiated, and the detection controller may detect a face. A UI controlled by the image capture application on the smartphone may display a bright, high-contrast outline around a face and/or torso (displayed as a red line). The outline may allow a low-vision user to locate themselves in the center of the camera frame more easily. The outline may update as the image capture application repeats the tracking process (e.g., every 100-200 milliseconds) and may be provided in addition to the audio, haptic, and/or visual feedback discussed above and illustrated in Table 1.

The image capture application may automatically take more than one photo (e.g., from different angles) when a user tilts a smartphone and automatically select the best photo for the user.

Figure 4 illustrates an example of providing an audio detail description of the content present in a viewing frame of a smartphone camera.

Figure 4

As illustrated in Figure 4, the image capture application may perform auditory assistance via a speaker to announce to a user what is in the camera frame. For example, assume that a user initiates a front-facing camera of a smartphone to take a selfie. The image capture application is initiated, and the detection controller may detect a face and one or more objects in the viewing area. The image capture application may rely on a timer to determine the user's intent to take a selfie when the user holds a viewing frame for a period of time (e.g., 1 second). After the timeout, the image capture application may initiate a smartphone speaker to announce faces and objects in the foreground and/or background. The audio detail description provides confirmation of the content present in the viewfinder for the user. The audio detail description would be helpful to a user who is low-vision or who is otherwise unable to clearly see the viewfinder of the camera on the display. The audio detail description may be provided in addition to the audio, haptic, and/or visual feedback discussed above and illustrated in Table 1.

The systems and techniques discussed above utilize assistive technology to make accessible technology and better selfies available to everyone.

References:

[1] Patent Publication: US20150334292A1. System and Method for Providing Haptic Feedback to Assist in Capturing Images. Priority Date: May 13, 2014.

[2] Patent Publication: US20140192247A1. Method for Controlling Camera Operation Based on Haptic Function and Terminal Supporting the Same. Priority Date: January 7, 2013.

[3] Patent Publication: US20190208115A1. Identifying User Intent for Auto Selfies. Priority Date: December 28, 2017.

[4] "Take motion photos & use Top Shot on your Pixel phone." Google Camera Help. Accessed (March 4, 2022). https://support.google.com/googlecamera/answer/9937175?hl=en. Published at least as early as October 22, 2020. Internet Archive. https://web.archive.org/web/20201022041322/https://support.google.com/googlecamera/answer/9937175?hl=en.

[5] White, Samuel, Ji, Hanjie and Bigham, Jeffrey P. "EasySnap: real-time audio feedback for blind photography." UIST '10: Adjunct proceedings of the 23rd annual ACM symposium on User interface software and technology. October 3, 2010. https://dl.acm.org/doi/10.1145/1866218.1866244.