August 2020

Automatically Pausing Caller Video When the Caller's Face is Out-of-frame

Garrett Frederick Berg

Zac Livingston

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

Automatically Pausing Caller Video When the Caller's Face is Out-of-frame

<u>ABSTRACT</u>

Many video communication applications offer a picture-in-picture mode that shows the caller's video feed that usually includes the caller's face in addition to that of others that participate in the call. The impulse to check one's own image as provided in the user interface can distract callers from focusing on the other participants. Further, if the caller moves, the camera may capture and provide a video feed that includes unexpected content, e.g., the room background. This disclosure describes the use of face detection at the caller end of a video call that ensures that the caller's camera feed includes a face and automatically switches to a static picture being displayed to the call recipient if the face is not detected. At the caller end, the caller's video feed is not displayed if the face is present. In this manner, the described techniques simultaneously safeguard against inadvertent display of other content while removing the distraction of one's own face from the user.

<u>KEYWORDS</u>

- Video call
- Videoconference
- Front camera
- Selfie camera
- Face detection
- Picture-in-picture
- Automatic pause

BACKGROUND

When conducting an in-person conversation, one cannot see one's own face with the rare exception of situations when one is standing in front of a mirror. Many video communication applications offer a picture-in-picture mode that shows the caller's video feed that usually includes the caller's face in addition to that of others that participate in the call. Reasons to show callers their own video include enabling the caller to ensure that the call participants are receiving the video and providing feedback regarding how the caller and their environment appears to the call participants.

However, a user interface that provides a self-view makes video calls made via video communication applications differ from an in-person conversation. The impulse to check one's own image as provided in the user interface can distract callers from focusing on the other participants. The impulse exists due to the human tendency for positive self-presentation that leads callers to look at their own camera feed to ensure that their looks and expressions are appropriate. Moreover, people often consult their own video to confirm that the video does not show undesired personal poses or their environment (e.g., clutter in the background). Further, if the caller moves, the camera may capture and provide a video feed that includes unexpected content.

DESCRIPTION

This disclosure describes the use of face detection at the caller end of a video call that ensures that the caller's feed is hidden from them while their camera feed includes a face and automatically switches to a static picture for other participants if their face is not detected, thus safeguarding against inadvertent display of other content. Such a user interface facilitates video

calls that show callers only the camera feed of other call participants mimicking the experience of an in-person conversation.
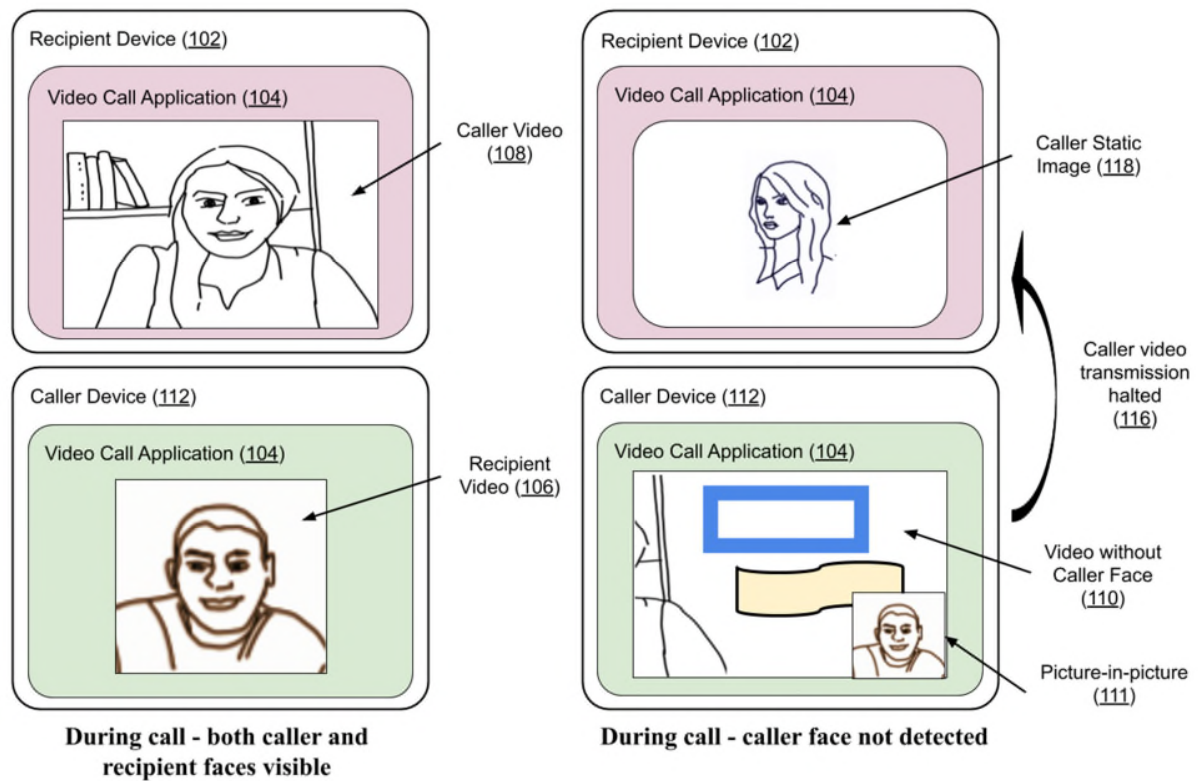

Fig. 1: Automatically halting caller video when the caller's face is out of the frame

Fig. 1 shows an operational implementation of the techniques described in this disclosure. A caller uses a video call application (104) on a device (112) to make a video call to a recipient that uses a recipient device (102). The recipient device shows the caller's video (108) as captured by a camera on the caller's device while the caller device shows recipient video (106) if enabled by the recipient. Neither device shows its own video feed, thus mimicking real life conversation.

With the caller's permission, the video call application continually performs face detection during the call to determine whether the caller's face is present in the video feed. When the caller's face is no longer in the video, transmission of the video feed to the recipient device is

automatically halted (116) and the caller can again see their own face (110), along with a picture-in-picture view (111) of the recipient. As can be seen in Fig. 1, the caller has moved such that their body is partially visible as are some background objects, without the face being in the frame. The recipient device then displays the caller's chosen static image (118) instead of the caller's live video feed. While the foregoing discussion refers to the caller's video feed being halted upon no face being detected, the use of face detection to halt transmission of video feed is equally applicable to all parties in a video call that enable such feature.

Before starting a video call, each participant is provided a user interface in which they can see their own face as captured by their device camera, allowing them to ensure that they look presentable, to adjust the camera angle to avoid capturing unwanted personal poses and environmental objects, to ensure proper lighting, etc. With permission of the user, each participant device performs face detection (e.g., which may be a part of the call application) to ensure that the user's face is within the frame in the captured video feed. The user interface shows an indication to the user to convey that face detection is turned on, and that the presence of their face in the video is being continually and automatically detected. The user has the option to disable the continual automatic face detection. When the user is satisfied with their own video feed, she can select an option to initiate or join the video call.

Once the video call starts, users no longer see their own camera feeds, instead seeing only the video of other participants. When a user has turned on face detection, other participants in the call see the user's video feed as long as the user's face is in the frame. If the user's face is not detected in the frame at any point during the call, transmission of the user's video to other participants is automatically halted. At this time, the other participants see the user's static image, such as a profile picture. Automatically stopping video transmission in this manner

prevents the other participants from inadvertently seeing background objects or undesired poses in the video feed.

When a user's video is automatically halted during an ongoing video call, the user interface on their device is automatically switched to the camera that shows the user's self-view. When the face is detected again, the caller is provided an option to resume sending video to the called party or to temporarily disable the face detection feature. For example, this may be useful when the user wants to show other objects such as a whiteboard, a piece of paper, other objects in the vicinity, etc.

With user permission, the techniques described in this disclosure can be integrated within any application or platform used for video communication. The techniques can support one-to-one as well as many-to-many video calls. The techniques ensure that the video feed of a user is presented only when the user's face is detected, without needing to echo the video feed on the user interface. As a result, the user experience (UX) of video communication is improved by helping callers avoid distraction and focus on the other participants in the video call, thus creating a conversational experience that more closely mimics in-person conversation.

Further to the descriptions above, a user may be provided with controls allowing the user to make an election as to both if and when systems, programs or features described herein may enable collection of user information (e.g., information about a user's camera feed, a user's preferences), and if the user is sent content or communications from a server. If permitted by the user, face detection may be performed to detect the presence of a human face in the video frame and without storing data of the video feed. In addition, certain data may be treated in one or more ways before it is stored or used, so that personally identifiable information is removed. For example, a user's identity may be treated so that no personally identifiable information can be

determined for the user. Thus, the user has control over what information is collected about the user, how that information is used, and what information is provided to the user.

CONCLUSION

This disclosure describes the use of face detection at the caller end of a video call that does not show a caller their own face while ensuring that the caller's camera feed includes a face and automatically switches to a static picture if the face is not detected, thus safeguarding against inadvertent display of other content. Prior to a call, a user can examine and adjust their own video feed via a user interface. During the call, the user's own video feed is not shown in the user interface, unless no face is detected in the video feed. The described techniques can be integrated within any application or platform used for video communication and can improve the user experience of video communication by helping participants avoid distraction from their own video feed and focus on other participants in the call to create a conversational experience that more closely mimics in-person conversation.