February 05, 2019

User authentication via lip reading

Horacio Hernan Moraldo

User authentication via lip reading

ABSTRACT

This disclosure describes techniques to prevent unauthorized access to information via a facial recognition interface. With user permission, a live video of the user attempting to access the information is obtained via a camera to perform facial recognition and determine the presence of the user. The user is instructed to dictate a randomly generated passphrase while facing the camera. Lip reading techniques are applied to validate the dictation of the randomly generated passphrase. Access to the information is granted only upon successful validation.

KEYWORDS

- user authentication
- facial recognition
- live video
- lip reading
- fake video detection
- presence verification

BACKGROUND

Various authentication techniques, e.g., passwords, patterns/pins, or biometric-based techniques such fingerprint or iris scanning, facial recognition, etc., can be utilized, e.g., on a mobile device, to authenticate a user prior to granting access to the device or information accessed via the device. Although convenient and intuitive to use, authentication techniques such as facial recognition have certain weaknesses. For example, there is the possibility of malicious

actors fraudulently using a photograph (or other model) of the user or pointing the device camera towards the user (e.g., while the user is asleep or otherwise unaware) to gain access.

DESCRIPTION

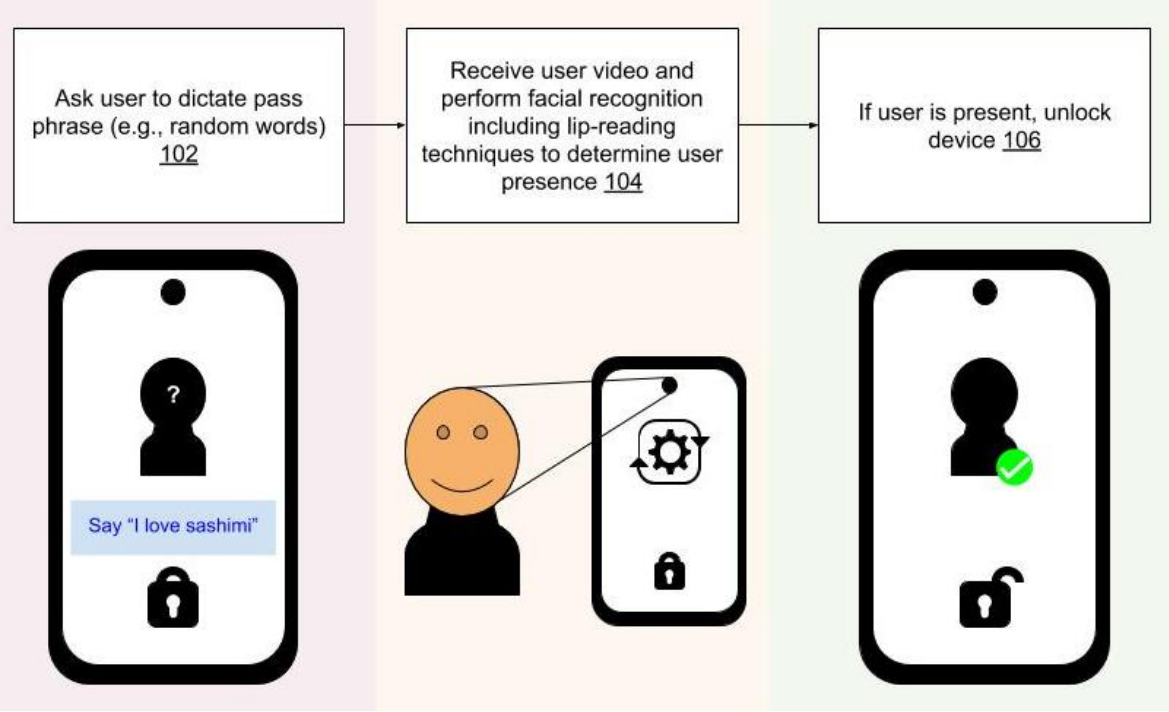

Fig. 1: Mobile authentication via lip reading

This disclosure describes techniques to prevent unauthorized access to information via a facial recognition interface. Fig. 1 illustrates an example process of user authentication using lip reading. The authentication process works as follows:

1. A user requests access to the device. A prompt to dictate a passphrase is displayed (102) on the device screen. For example, the phrase can be randomly generated and can include one or more words. For example, a combination of words from a dictionary can be displayed. In the example illustrated in Fig. 1, the phrase is "I love sashimi."

2. With user permission, a live video of the user is captured (104) as the user reads the phrase or moves their lips as if silently dictating the phrase.

3. Facial recognition is performed on the obtained video to match the captured face against pre-stored information to determine whether the user is authorized to access the device. Per techniques of this disclosure, if the user's face is recognized as authentic, lip deep learning based techniques are applied to the video to detect whether the user's lip movements match lip movements for the displayed phrase. Further, to thwart attackers that attempt to use a spoofed or machine-generated video (e.g., generated by use of generative neural networks) of the user dictating the passphrase, fake-video detection techniques are used to detect the spoofed video.

4. If the lip movements in the video match and the video is not detected as a spoofed video, the user is authenticated, e.g., the device is unlocked(106).

In this manner, the described techniques perform three different checks - facial recognition to detect that the user is a valid user of the device, lip reading to verify live presence of the user, and fake video detection to confirm that the video is not a spoofed video. Attackers that attempt to access devices via user photographs or spoofed videos are thus thwarted. Further, the user interface, e.g., "Say I love sashimi," is easy for the user to understand.

The described techniques can be implemented in any device that supports use of facial recognition techniques for user authentication. The described techniques are implemented with specific user permission. If the user denies or restricts permission for use of facial data or capture of live video, the techniques are not implemented. In these situations, the user can utilize other techniques such as passwords or patterns to access the device.

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 social network, social actions or activities, profession, a user's preferences, or a user's current location), and if the user is sent content or communications from a server. 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, or a user's geographic location may be generalized where location information is obtained (such as to a city, ZIP code, or state level), so that a particular location of a user cannot be determined. Thus, the user may have 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 techniques to prevent unauthorized access to information via a facial recognition interface. With user permission, a live video of the user attempting to access the information is obtained via a camera to perform facial recognition and determine the presence of the user. The user is instructed to dictate a randomly generated passphrase while facing the camera. Lip reading techniques are applied to validate the dictation of the randomly generated passphrase. Access to the information is granted only upon successful validation.

REFERENCES

1. K. Sai Sowjanya, Y. Aruna Suhasini Devi, and Sandeep K. "User Authentication Using Lip Movement as a Password." In International Journal of Advanced Research in Computer Science and Software Engineering Volume 5, Issue 10, October 2015. Available online at

http://ijarcsse.com/Before_August_2017/docs/papers/Volume_5/10_October2015/V5I10-0343.pdf.

2. https://www.newscientist.com/article/2113299-googles-deepmind-ai-can-lip-read-tv-shows-better-than-a-pro/ accessed November 21, 2016.
3. Marra, Francesco, Diego Gragnaniello, Davide Cozzolino, and Luisa Verdoliva. "Detection of GAN-generated fake images over social networks." In 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), pp. 384-389. IEEE, 2018.