January 2021

Face Mask with Microphone and Speaker for Speech Enhancement

N/A

Face Mask with Microphone and Speaker for Speech Enhancement

ABSTRACT

Wearing a mask muffles the wearer's voice thus making it difficult for the wearer to communicate and be understood by others. This disclosure describes techniques to improve the output speech from a person that is wearing a mask. Per the techniques, a microphone and a speaker is provided (e.g., as part of the mask) and are configured to add missing frequencies to the speech. Missing frequencies/sounds are identified based on a comparison with normal speech of the user when the user is not wearing a mask and are added by outputting sounds from the speaker.

KEYWORDS

- Face mask
- Muffled speech
- Garbled speech
- Speech enhancement
- Audio compensation
- Machine learning

BACKGROUND

Wearing a face mask is common in certain conditions, e.g., individuals who suffer from an infection such as a cold can reduce or prevent the spread of disease by wearing a mask. Certain professionals, e.g., surgeons, certain industrial workers, etc., are required to wear face masks as part of their job. However, wearing a mask muffles the wearer's voice thus making it difficult for the wearer to communicate and be understood by others.

DESCRIPTION

This disclosure describes techniques to improve the output speech from a person that is wearing a mask. Per the techniques, a microphone and a speaker is provided (e.g., as part of the mask) and are configured to add missing frequencies to the speech. Missing frequencies/sounds are identified based on a comparison with normal speech of the user when the user is not wearing a mask and are added by outputting sounds from the speaker.

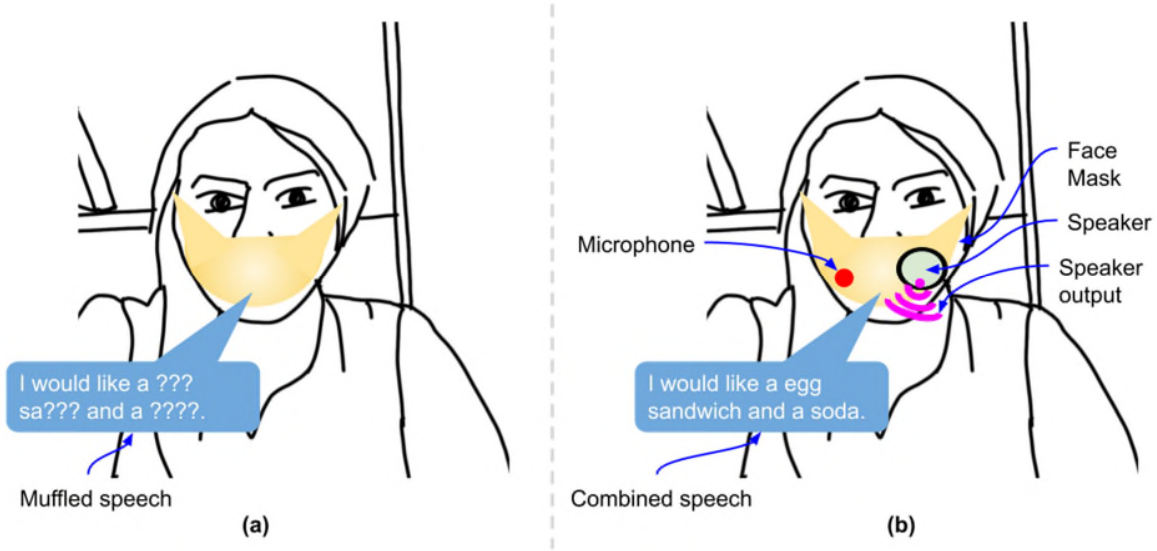

Fig. 1: (a) muffled speech due to face mask; (b) speech enhancement using mic and speaker

Fig. 1(a) illustrates a person wearing a face mask. Due to the face mask, their speech is muffled and may be not be recognized by a listener. Fig. 1(b) illustrates the use of a microphone and a speaker built into the face mask. The microphone recognizes the user's speech and provides it to a processor (not shown). The processor determines appropriate compensation (e.g., based on missing frequencies) and outputs appropriate audio from the speaker. The combined audio is easier for the listener to parse, since it more closely resembles speech without the face mask.

Per the described techniques, in a training phase, the individual user provides speech input to a device with a microphone, e.g., a smartphone, smart display, or other device. The speech input is provided under two conditions - first where the user does not wear a mask and second where the user wears a mask. For example, the user can be instructed to utter a training phrase under both conditions. An algorithm is trained to detect the difference between the received speech input under the two conditions, and can then determine the missing sounds in the speech received from the user while wearing the mask. The processor of the face mask can implement this algorithm.

For example, deep learning (or other machine learning techniques) can be used to provide playback of the missing frequencies via the speaker when the user speaks while wearing a mask. Machine learning techniques such as federated learning can be implemented with user permission to eliminate the requirement for each user to individually train the algorithm. Provision of the missing sounds is similar to active noise cancellation, except that missing sounds are added to improve audio, rather than canceling noise frequencies.

The described techniques can be utilized for any user that is wearing a face mask. For example, speech from surgeons or others wearing a surgical mask can be improved by adding missing frequencies. In case of a situation where the general public is required to wear masks, e.g., as is the case with the Covid-19 pandemic, the described techniques can improve communication between users and help maintain mask compliance.

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 speech frequencies, a user's preferences), 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

Wearing a mask muffles the wearer's voice thus making it difficult for the wearer to communicate and be understood by others. This disclosure describes techniques to improve the output speech from a person that is wearing a mask. Per the techniques, a microphone and a speaker is provided (e.g., as part of the mask) and are configured to add missing frequencies to the speech. Missing frequencies/sounds are identified based on a comparison with normal speech of the user when the user is not wearing a mask and are added by outputting sounds from the speaker.

REFERENCES

1. Kroo, Laurel, Anesta Kothari, Melanie Hannebelle, George Herring, Thibaut Pollina, Ray Chang, Samhita P. Banavar et al. "Pneumask: Modified Full-Face Snorkel Masks as Reusable Personal Protective Equipment for Hospital Personnel." *medRxiv* (2020).
2. "Bluetooth Masks." https://www.bioppe.com/bluetooth-masks/
3. "ReddyPort Microphone." https://www.reddyport.com/reddyport-microphone