Technical Disclosure Commons

November 29, 2017

Smart assistance for meetings

Sandro Feuz

Thomas Deselaers

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

Smart assistance for meetings

ABSTRACT

The described techniques provide smart assistance to video conference meeting attendees. With user permission, such assistance includes replaying, transcribing, translating and/or summarizing all or part of a meeting, interpreting specified voice commands, and identifying topics or items of interest from the meeting recording or transcript. For example, meeting participants can provide keywords corresponding to topics of interest and seek notifications when such topics are discussed in the meeting. Also, users can query meeting recordings or transcripts for topics of interest or meeting action items. Present techniques can be utilized for video conferences, audio conferences, recorded talks, and face-to-face meetings.

KEYWORDS

- Video conference replay
- Meeting transcription
- Meeting summary
- Automatic summarization
- Meeting assistant

BACKGROUND

Video conference meetings typically include participants from different remote locations. Video conference participants are sometimes distracted with other activities and do not focus continuously on the video conference meeting. For example, during a video conference, participants perform tasks unrelated to the meeting using their laptops or other personal devices, e.g., check emails/messages, browse the internet, etc. In these situations, participants are distracted from the meeting and may miss important information.

DESCRIPTION

      Techniques described herein leverage advances in video conferencing and artificial intelligence to improve the effectiveness of video conferences, audio conferences, face-to-face meetings, and recorded talks. Replay of a prior portion of a meeting, e.g., a portion of a predetermined duration such as a few seconds/minutes, is integrated into software and hardware that provides the video conference. The replays can be executed at a higher speed (e.g., 1x/1.25x /1.5x/2x) and displayed in a picture-in-picture (PIP) mode during the video conference. The replays can help participants review portions of the meeting. The PIP mode can be implemented by displaying the live video conference meeting in a larger view and the video replay within a smaller window or vice versa.

      The present techniques, with user permission and express consent, transcribe speech from the meeting using automatic speech recognition. Replays of the meeting can also be initiated by clicking on a corresponding part of the transcript. Replays of relevant excerpts of a multi-speaker video conference meeting can be shown in a picture-in-picture (PIP) mode during the meeting.

      In large meetings, e.g. staff meetings of large organizations, some aspects of the meeting may not be of interest to the entire audience, which can lead to some of the meeting participants losing interest in the meeting or becoming distracted with other non-meeting related activities. To improve user experience during such meetings, the described techniques allow participants to configure keywords of their interest. Upon performing automatic speech recognition, if a selected keyword is identified, participants are notified. In addition to keywords chosen by the participant, other factors, e.g., the participant's name, the participant's project and/or team name can also be included as additional keywords, with permission from the participant.

Meeting summaries or transcripts are automatically generated using machine learning. For instance, a meeting participant may choose to have summaries prepared for a part of the meeting during which the participant had to leave the meeting. With user permission, a machine learning model that is trained to perform summarization is provided the transcript of the portion of the meeting when the participant was away (and optionally, transcript of the entire meeting and metadata about the meeting). The summary generated by the model based on these inputs, e.g., a few sentences, is provided to the participant.

The meeting transcript can include features to facilitate more efficient note taking. For instance, a dedicated note taking user can annotate a different view of the transcript with meeting notes. These notes, represented in the form of a document, can provide additional details and context. The meeting summarization model can also facilitate automatic meeting note taking, and along with the meeting recording and transcript, make note taking easier and seamless.

Also, generation of meeting transcripts includes support for pre-configured voice commands that can be used by meeting participants. For example, if a participant speaks a phrase that corresponds to a command, e.g., "Thomas, can you take this action item," followed by a description of the action item, the generated transcript includes the action item, along with the description. The description is editable, e.g., by the assignee of an action item, by a note taker, etc.

In meetings where participants speak different languages, automatic translations of the different languages are included the transcript. Further, even for meetings conducted in a single language, automatic transcript can enable participants that are not fully fluent in the language to look up meanings of unfamiliar words or phrases.
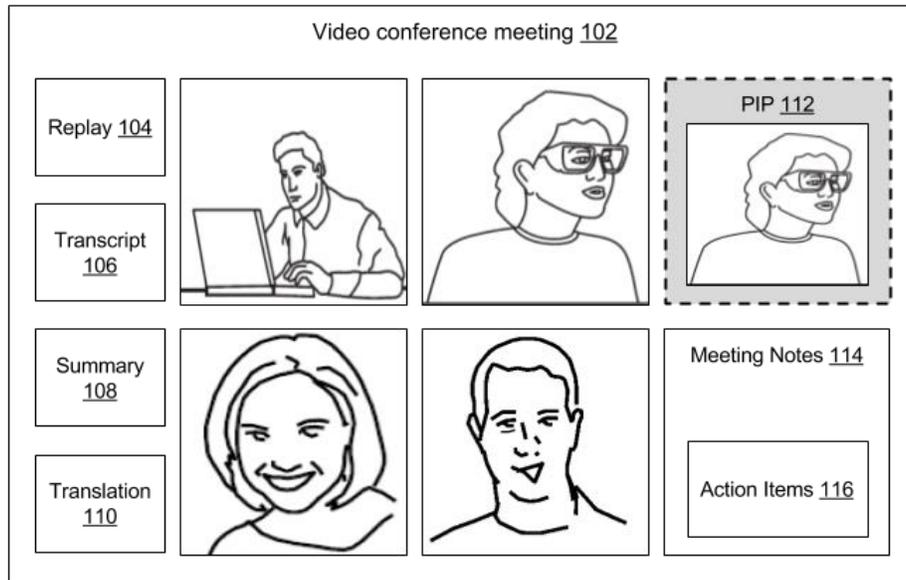
Fig. 1: Video conference meeting with smart assistance

Fig. 1 illustrates an example user interface for a video conference meeting (102) with smart assistance to participants via features such as replay (104), transcription (106), summary (108), and translation (110). The features are enabled with permission from the meeting participants. Participants are provided with options to turn off one or more of the features during a particular meeting, or a portion of the meeting. Recording, transcription, summarization, translation, etc. are performed using computing devices that the users have granted permission to access the meeting.

When a participant chooses to replay a portion of the meeting, replays are shown in PIP (picture-in-picture) mode (112) in the user interface. Participants can take meeting notes (114) and track action items (116) assigned during the conference. The present techniques enable users that are unable attend a meeting or that miss portions of the meeting to review a summary of the meeting, or the entire recording or transcript. Users can query the meeting recording or the

corresponding transcript for topics of interest discussed within a meeting or action items assigned to a specific meeting participant.

The described techniques can also be used in face-to-face meetings and can function seamlessly on user devices, e.g., phone, tablet, or other computing devices. The techniques can be especially beneficial to make meetings more accessible, e.g., the automatic transcription and associated described features can assist participants with hearing disability, etc. The techniques described in this disclosure can be incorporated in software and/or hardware products and services that provide audio and/or video conferencing. Such products can include meeting room hardware, client software, server software, etc.

In situations in which certain implementations discussed herein may collect or use personal information about users (e.g., user data, information about a user's social network, user's location and time at the location, user's biometric information, user's activities and demographic information), users are provided with one or more opportunities to control whether information is collected, whether the personal information is stored, whether the personal information is used, and how the information is collected about the user, stored and used. That is, the techniques discussed herein collect, store and/or use user personal information specifically upon receiving explicit authorization from the relevant users to do so.

For example, a user is provided with control over whether programs or features collect user information about that particular user or other users relevant to the program or feature. Each user for which personal information is to be collected is presented with one or more options to allow control over the information collection relevant to that user, to provide permission or authorization as to whether the information is collected and as to which portions of the information are to be collected. For example, users can be provided with one or more such

control options over a communication network. 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. As one example, a user's identity may be treated so that no personally identifiable information can be determined. As another example, a user's geographic location may be generalized to a larger region so that the user's particular location cannot be determined.

CONCLUSION

The described techniques provide smart assistance to video conference meeting attendees. With user permission, such assistance includes replaying, transcribing, translating and/or summarizing all or part of a meeting, interpreting specified voice commands, and identifying topics or items of interest from the meeting recording or transcript. For example, meeting participants can provide keywords corresponding to topics of interest and seek notifications when such topics are discussed in the meeting. Also, users can query meeting recordings or transcripts for topics of interest or meeting action items. Present techniques can be utilized for video conferences, audio conferences, recorded talks, and face-to-face meetings.