

Technical Disclosure Commons

Defensive Publications Series

August 2020

Multimedia Slides with Attached Video Clips and Transcript from Presentation Video

Safwan Samla

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

Recommended Citation

Samla, Safwan, "Multimedia Slides with Attached Video Clips and Transcript from Presentation Video", Technical Disclosure Commons, (August 20, 2020)
https://www.tdcommons.org/dpubs_series/3533



This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).

This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.

Multimedia Slides with Attached Video Clips and Transcript from Presentation Video

ABSTRACT

Presentation slides by themselves include neither the spoken delivery of the presenter nor the interactive discussion with the audience during and after the presentation. As a result, the content of the slides can sometimes miss important information conveyed during the presentation. Those who simply read slide content without access to the accompanying live presentation may not develop a complete understanding of the material. This disclosure describes techniques to create multimedia slides augmented by corresponding video clips and/or audio transcript from the presentation.

KEYWORDS

- Presentation slides
- Multimedia slides
- Lecture
- Seminar
- Webinar
- Online learning
- Remote learning
- Education
- Learning and Development

BACKGROUND

When presenting content, people often use slides that accompany their delivery of the material. During live presentations, the slides provide a rich and powerful visual aid that supports the presenter's speech. Oftentimes, such slides are provided to people after the presentation,

including to those who did not attend the presentation. However, by themselves, the slides include neither the spoken delivery of the presenter nor the interactive discussion with the audience during and after the presentation.

As a result, the content of the slides can be missing important information that was conveyed during the presentation, such as background and context, illustrative examples, answers to audience questions, etc. Those who simply read the slide content without having attended the accompanying presentation may not develop a complete understanding of the material. Even those who attended the live presentation may not remember the details of the presentation when going through the slides at a later time.

For instructional presentations, such as lectures, seminars, training sessions, etc., the above shortcomings have a negative impact on learning based on the slides. Learners can try to mitigate the negative impact with approaches such as reading from textbooks and other materials, reviewing notes taken during the presentation by self, and/or others, examining presenter notes in the slides, viewing a video of the presentation, etc.

However, each of these approaches has limitations that make it unsuitable for fully overcoming the shortcomings of the slides. For instance, reading materials are stale and unlikely to include specific clarifications provided during the presentation; detailed lecture notes are tedious to capture and distract focus from learning during the presentation; presenter's notes on slides are often missing, incomplete, or terse; and presentation video is typically disconnected from the slides, thus making it inefficient for accessing presentation content related to specific slides.

DESCRIPTION

This disclosure describes techniques to automatically augment slides to generate multimedia slides based on the video and slides of a presentation. With permission of the presenter, the video of the portion of the presentation corresponding to the content of a slide is extracted and attached to the slide as a video clip. In addition, the audio content within the video is converted to text and appended to the notes field of the slide. A user that views such a multimedia slide is provided with options to view the video clip of the corresponding portion of the video presentation or read the text transcript of the corresponding presentation audio.

Each frame of the video of the full presentation is mapped to a slide using computer vision approaches. The frames corresponding to each of the slides are merged to create a video clip of the portions of the presentation during which the presenter covered the contents of the slide. The slide content is augmented by attaching the video clip that the user can play whenever needed.

Further, in addition to attaching the video clip to a slide, audio corresponding to the slide is transcribed using speech-to-text techniques. The text transcription is appended to the slide, e.g., added to the notes field of the slide, which the user can consult as necessary.

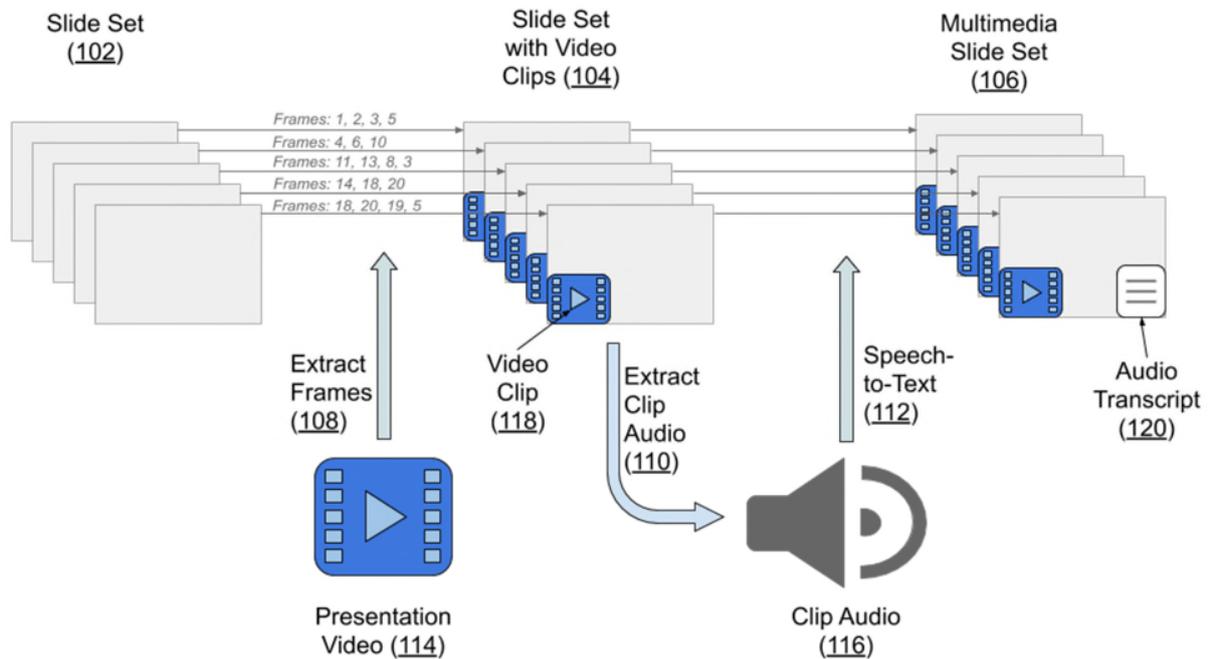


Fig. 1: Attaching presentation video clips and audio text to corresponding slides

Fig. 1 shows an operational implementation of the techniques described in this disclosure. A set of slides (102) for a presentation is obtained and analyzed in conjunction with a video of that presentation (114). The video is processed to extract frames (108) and each frame is mapped to a corresponding slide. The frames associated with each slide are assembled into video clips (118) attached to the associated slide to create a slide set augmented with the video clips (104). Audio (116) from each clip is also extracted (110) and transcribed using speech-to-text techniques (112). The generated transcript (120) is appended to the notes field of the corresponding slide to create a multimedia slide set (106) with each slide including the video clip and transcript of the corresponding portion within the original presentation.

The above techniques can also handle situations when only the video of a presentation is available without the corresponding set of slides. In such cases, computer vision approaches can be utilized to mark and extract video frames that contain the slides shown during the

presentation. The frames can then be merged to recreate the set of slides which are then augmented with video clips and text transcripts as described above.

Analysis of video frames, including automatic detection and extraction of slides, as well as mapping each frame to a corresponding slide can be performed using computer vision techniques, e.g., by using any commercially available computer vision application programming interface (API) as provided by cloud service providers. Similarly, speech-to-text APIs provided by cloud service providers can be used to transcribe audio for attaching to each slide. Further, if the slide presentation is available via online office productivity software suites, the API provided by such tools can be utilized to generate or augment the set of slides.

The techniques described in this disclosure enhance the user experience (UX) for viewing the contents of a set of presentation slides on their own. Moreover, the techniques generate the augmented presentation in an automated manner that imposes no additional burden on the presenter or other users. The enhanced UX can improve the effectiveness of learning for slides that contain instructional material regardless of whether the learner attended the live presentation during which the slides were shown. The augmented slides can improve the efficiency of learning by enabling learners to quickly locate and consult presentation portions pertinent to specific slides of interest.

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 slide presentation documents, videos of presentations, 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. 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 create multimedia slides augmented by corresponding video clips and/or audio transcript from a live presentation. With permission of the presenter, a video of the portion of the presentation corresponding to the content of a slide is extracted and attached to the slide as a video clip. The audio content within the video is converted to text and appended to the notes field of the slide. When reading such a multimedia slide, the user has the option to view the video clip and/or read the text transcript. The techniques enhance the user experience (UX) for viewing slides without imposing additional burden on the presenter. The enhanced UX can improve the effectiveness and efficiency of learning from instructional slides regardless of whether the learner attended the live presentation.