June 27, 2019

Machine learning techniques to automatically advance slides in a presentation

Archie Russell

Machine learning techniques to automatically advance slides in a presentation

ABSTRACT

Slideshow presentations require the presenter or other human to click on a button to go to the next slide, to go back a few slides, to stop presenting, etc. This doesn't work well in some situations, e.g., if the presenter is differently abled, if there is no remote control available, etc. This disclosure describes machine-learning techniques that automatically advance slides in a presentation (with user permission) based on user-permitted factors such as slide content, sentiment analysis, speech understanding, etc.

KEYWORDS

- Slide deck
- Slide progression
- Speech recognition
- Machine learning
- Sentiment analysis
- Presentation

BACKGROUND

Slideshow presentations require the presenter or other human to click on a button to go to the next slide, to go back a few slides, to stop presenting, etc. This doesn't work well in some situations, e.g., if the presenter is differently abled, if there is no remote control available, etc. Slideshows can be recorded to change slides at precise intervals; however, this is extra work for a presenter, and fails in case of delay or in case of a change in the slides.

DESCRIPTION

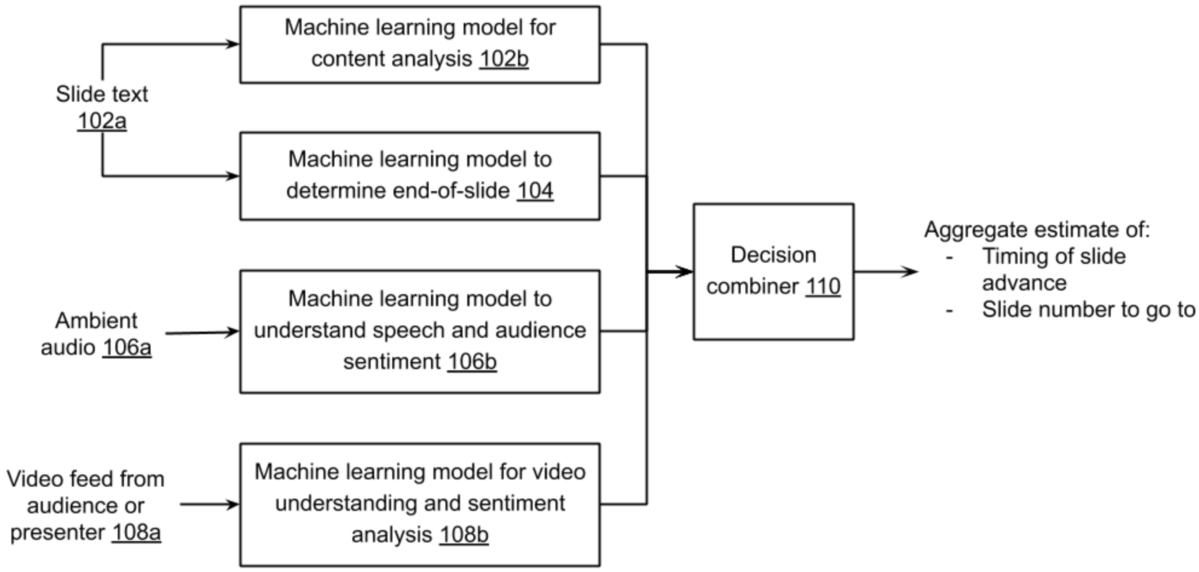

Fig. 1: Automatically advancing slides in a presentation

Fig. 1 illustrates automatically advancing or moving back slides in a presentation, per techniques of this disclosure. Several components are used to determine the time instant at which a presentation is to be advanced and the next slide number, as follows.

A first component is a machine learning model (content analyzer) that performs content analysis (102b) on slide text (102a). With presenter permission, the content analyzer uses factors such as the amount of text, number of questions, number of figures, the position of the slide in the presentation, etc. to estimate the length of time a particular slide is likely to take during a presentation. The machine learning model is trained on corpus of presentations (that are available for use as training data) to learn these factors. Human tuning of the model can also be performed.

A second component is a machine learning model (104) that determines the final slide in a presentation based on the slide text. The final slide is determined so that the presentation does not go beyond an end-slide to supplementary materials that may be included in a slide deck.

A third component is a machine learning model (106b) for understanding speech and audience/speaker sentiment. With user permission, the component analyzes ambient audio (106a) to estimate when the presenter is ready to go to the next slide. With user permission, the component identifies the primary people speaking, and detects if the speaker has completed the last line of text on a slide, whether the speaker has paused speaking, whether the audience is interacting with the speaker, whether the speaker has said anything indicative of slide advancement, e.g., “next,” “ok,” “previous slide,” etc. The component uses speech and natural language understanding techniques to detect the presenter’s intent to move forward or back a few slides, e.g., by detecting utterances such as “go back to the last slide.” The machine learning model is trained on audio and corresponding presentations (that are available for use as training data) to learn factors that are indicative of slide advancement. Human tuning of the model can also be performed.

A fourth component is a machine learning model (108b) for video understanding and sentiment analysis that accepts as input the video feed from the presenter or from the audience (108a). With the presenter’s permission, the machine learning model analyzes the video feed and estimates the time instant for slide advancement and the next slide to go to, using factors derived from the video such as the presenter’s gestures, appearance, etc. The machine learning model is trained on videos and corresponding presentations (that are available for use as training data) to learn factors that are indicative of slide advancement. Human tuning of the model can also be performed.

A decision combiner (110) combines the estimates of the time instant to advance a slide and the next slide to go to, as provided by each of the components and determines an aggregate estimate (112) for the time instant to advance the slide and the next slide to go to. When the presenter enables automatic control, the output of the decision combiner is used to advance the presentation. The techniques of this disclosure can be exposed to users via menus and automatic prompting. If the presenter uses a clicker in addition to the automatic slide-advancement techniques described herein, the machine learning models can utilize such user input with user permission, e.g., to determine the difference between the model predictions and the presenter's actual next clicks, to improve model performance. The presenter can be apprised of the model performance by recapping how closely the techniques match the presenter's actual next clicks.

User data such as content of the user's slide decks, slide advancement input, audio and/or video feeds from the actual presentation, etc. are utilized with specific user permission. While Fig. 1 illustrates multiple factors, any subset of the factors that are permitted by the user can be utilized to predict slide advancement. For example, when audio or video from the presentation is unavailable or if the user denies permission, only the first and second component are utilized. While Fig. 1 illustrates separate machine learning models and a combiner, a single model that utilizes multiple inputs can be utilized instead. Automatic slide advancement can be performed in a subset of situations, e.g., when the model prediction meets a threshold confidence.

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 machine-learning techniques that automatically advance slides in a presentation (with user permission) based on user-permitted factors such as slide content, sentiment analysis, speech understanding, etc. The techniques enable a presenter to present a slide deck without need for a clicker or human assistant.

REFERENCES

[1] Arriaga, Emmanuel M., “Automatic slide progression during a presentation”, Technical Disclosure Commons, (December 05, 2017) http://www.tdcommons.org/dpubs_series/862

[2] Chen, Xiaohui, and Nathan Cordeiro. “Prioritized display of visual content in computer presentations.” U.S. Patent 9,870,755, granted Jan. 16, 2018.