

Technical Disclosure Commons

Defensive Publications Series

September 16, 2019

IMAGE-BASED SLIDE PRESENTATION RETRIEVAL

Matthew Sharifi

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

Recommended Citation

Sharifi, Matthew, "IMAGE-BASED SLIDE PRESENTATION RETRIEVAL", Technical Disclosure Commons, (September 16, 2019)
https://www.tdcommons.org/dpubs_series/2488



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.

IMAGE-BASED SLIDE PRESENTATION RETRIEVAL

A cloud-based content platform provides a variety of online applications to users for a comprehensive online service. For example, the content platform may provide a presentation application to enable users to generate and/or present a presentation document, such as a slide presentation document, to one or more audience members. The content platform may also provide a cloud storage application to enable users to store the presentation document and make the presentation document available for access by users of the content platform. Often, a presenter and/or an owner of a presentation document may wish to distribute the document to each audience member of the live presentation. However, the presentation may be given to hundreds or thousands of audience members and therefore, the presentation document may not easily distributed. In some examples, the presentation document may be accessible via the cloud storage application and one or more content platform users may be permitted to access the presentation document. In such examples, an electronic address (i.e., a link) to access the document via the cloud storage application may be provided in one or more portions of the document so the audience members may access the presentation document via the cloud storage application.

During the live presentation, an audience member may generate an image of a portion of the presentation document using a camera application of a user device. For example, an audience member may capture an image of a slide from the presentation document as the slide is being presented. The slide may include the address to access the presentation document via the cloud storage application, or may include content that is of interest to the audience member. The audience member, referencing the slide image, may search for the address and/or extracted portions of text included in the slide of interest, to identify the slide presentation document of the

cloud storage application. However, this can be a time intensive process for the audience member, as the audience member may not easily access the presentation document based only on referencing only the address and/or the extracted text from the slide image.

Therefore, a technique is proposed for assisting an audience member in retrieving a slide presentation document based on at least one generated image of a slide from a live presentation of the slide presentation document. The following description uses a slide presentation as an example document type to which the proposed technique can be applied. It should be noted, however, that the proposed technique can be applied to various other types of documents, including, for example, text documents, spreadsheets, etc.

Figure 1 illustrates a flow diagram of a method for identifying and retrieving a slide presentation based on a provided slide image. The method may be performed by at least a presentation indexing component and/or a presentation identification component of the content platform. In some embodiments, the presentation indexing component and/or the presentation identification component may be part of the cloud storage application. In other embodiments, the presentation indexing component and/or the presentation identification component may be part of another online application of the content platform.

At block 110, the presentation indexing component can receive, from a user device, a slide presentation document including one or more slides of a slide presentation. The user device can be any computing device that can access the content platform. In some embodiments, the slide presentation document may be generated using the presentation application of the content platform. The owner of the slide presentation document may make the document available for access by one or more content platform users. For example, the slide presentation document may be made public through the content platform (i.e., any user of the content platform may access

the slide presentation document). In another example, the slide presentation document may be made available to specific users of the content platform.

At block 120, each slide of the slide based presentation is rendered as an image. The presentation indexing component may parse the slide presentation document to separate each slide and render each slide as an image. In some embodiments, each slide may be rendered as an image responsive to receiving an indication that the slide presentation document is available to one or more users. In other embodiments, each slide may be rendered as an image responsive to receiving a command from the owner of the slide presentation document.

At block 130, each rendered slide image may be associated, by the presentation indexing component, with a node of an index structure. In one embodiment, each rendered slide image may be converted into an embedding using a trained machine learning model. Each embedding may be a representation (e.g., a feature vector) of a rendered image that has a smaller dimensionality than the image and defines one or more image attributes (i.e., gradient magnitude, color, grayscale intensity, edges, areas, etc.). The machine learning model may be trained to extract features from an image of a slide to be used in generating a feature vector. The machine learning model may be trained using feature vectors generated for a set of training images to identify one or more attributes corresponding to the training feature vectors. Based on a correlation of the identified image attributes and the attributes of the training feature vectors, the machine learning model may identify attributes from an image of a slide to be used in generating the feature vector for the image. Each embedding, along with metadata extracted from the slide presentation document (i.e., document title, date and/or time of live presentation, geographic location live slide presentation, etc.), may be associated with a node of an index structure, such as a k-dimensional tree structure.

In some embodiments, the index structure may be an inverted index structure. In such embodiments, text may be extracted from each rendered slide image by an optical character recognition (OCR) model and associated with an index entry of the inverted index structure. Each index entry of the inverted index structure may correspond to one or more words, or an anagram of one or more words, of the text extracted from each rendered slide image. The inverted index structure may include a mapping from each index entry to an identifier for each slide of the slide presentation document.

At block 140, a request for the slide presentation document is received by the presentation identification component. The request may include an image of a slide from a live presentation. In some embodiments, the image may include only the content of the slide as was presented. In other embodiments, the image may include the content of the slide as was presented and additional objects surrounding the slide (e.g., one or more presenters, a podium, a stage, etc.). A segmentation model may be applied to the image to detect the portion of the image that includes the slide. In some embodiments, the segmentation model may be a machine learning model that is trained to identify slides of a slide presentation from an image. The machine learning model may be trained using a training set of images, where one or more of the set of images depict slides of a slide presentation, and a bounding box to identify where, in the image, the slide is depicted. In other embodiments, the segmentation may be a simple heuristic model that identifies an object from an image that occupies a large portion of the image with two roughly vertical parallel lines and roughly horizontal parallel lines surrounding the object. Responsive to detecting the portion of the image that includes the content of the presentation, the image may be segmented to only include the portion of the image that includes the slide.

At block 150, an identification object may be generated for the image included in the request. In some embodiments, the identification object may be generated by converting the image into an embedding, as described previously herein. In other embodiments, the identification object may be generated using text extracted from the image using an OCR model.

At block 160, the slide presentation document associated with the image is identified, based on the generated identification object. As discussed previously, the identification object may be generated by converting the image into an embedding.. A nearest neighbor search (i.e., a search to identify a point in a given set that is closest to a given point) may be performed to determine a node of the k-dimensional tree structure that is associated with an embedding that is the same or similar to the embedding indicated by the identification object. . Metadata associated with the identified node may indicate a slide presentation document of the cloud storage application that includes the slide corresponding to the identified node.

In one example, a user may request a slide presentation document by providing an image of a slide from a live presentation. An embedding of the image may be generated using the machine learning model. A nearest neighbor search may identify at least one node of the k-dimensional tree structure that corresponds to the identification object, where the identified node includes metadata indicating the slide presentation document title, "Slide Presentation A." The slide presentation document title, "Slide Presentation A," may correspond with a slide presentation document of the cloud storage application.

As discussed above, the identification object may be generated using text extracted from the image provided in the request. In such embodiments, the inverted index structure may be searched to identify one or more index entries that correspond to the identification object. Based

on the identified index entries, a slide identifier may be determined which may correspond to a presentation document of the cloud storage application.

In some embodiments, the presentation identification component may identify more than one matching nodes and/or index entries that correspond to the identification object for the provided image. In such embodiments, additional metadata (i.e., date and/or time the image was generated, geographic location where the image was generated, etc.) may be used to identify the node of the index structure, or the index entry of the inverted index structure, that most closely matches the identification object. For example, a date and/or time the image was generated may be extracted from metadata associated with the image included in the request. The processing device may compare the extracted date and/or time with a date and/or time corresponding with each matching node identified from the index structure. The slide presentation document associated with the node that most closely matches the identification object may be determined based on the comparison of the dates and/or times.

At block 170, the user is provided access to the slide presentation document of the cloud storage application. In some embodiments, an address to the slide presentation document may be provided to the user to access the slide presentation via the cloud storage application.

In some embodiments, the slide presentation document may not be made public through the content platform. In such embodiments, the presentation identification component may verify that a user is permitted to access a requested slide presentation document. Responsive to providing the slide presentation document to the content storage application, the owner may define a group of users permitted to access the slide presentation document by creating a set of permissions associated with the slide presentation document. For example, the owner may only permit audience members of the live presentation to access the slide presentation document.

Responsive to receiving a request for the slide presentation document and identifying the slide presentation document associated with the image in the request, the presentation identification component may verify that the user requesting access to the slide presentation document is permitted to access the document. In some embodiments, a verification identifier may be included in the request for the slide presentation document indicating whether the user is permitted to access the slide presentation document. The verification identifier may be a unique identifier that is specific to the user transmitting the request. In some embodiments, the verification identifier may be an identifier associated with the provided image that indicates the date and/or time, and location where the image was generated. Prior to providing access to the slide presentation document, the presentation identification component may verify, based on the verification identifier and the permissions defined by the owner of the slide presentation document, whether the user is permitted to access the slide presentation document.

In additional embodiments, responsive to receiving a request to access the presentation document, a notification may be transmitted to the owner of the presentation document indicating that a user of the content platform is requesting access to the presentation document. The notification may include an identifier of the user requesting access, such as a user name or an electronic mail address. Responsive to receiving the notification, the owner may allow or deny access of the slide presentation document to the user that transmitted the request.

In some embodiments, the presentation document may not be available for access by one or more users of the content platform at the time the request for the slide presentation document is transmitted and instead may be available for access after the request for the slide presentation document is transmitted. In such embodiments, the presentation identification component may perform the above described method and determine that no slide presentation document is

associated with the image included in the request. As additional slide presentation documents are added to the content platform, the presentation identification component may periodically search the index structure to identify a subsequently provided presentation document that may be associated with the image included in the previous request. Responsive to identifying a presentation document that is associated with the image, a notification may be transmitted to the user requesting the slide presentation document indicating that a slide presentation document corresponding to the image was identified.

ABSTRACT

A technique is proposed for assisting a user with accessing a slide presentation document on a cloud based content platform. A presenter of the slide presentation document provides the slide presentation document to the content platform to be accessed by one or more users. A presentation component of the content platform renders each slide of the slide presentation document as an image and associates a node of an index structure with each rendered slide image. The index structure is referenced by the content platform to identify slide presentation documents. A user transmits a request for a slide presentation document where the request includes an image of a slide from the slide presentation document. An identification object for the image included in the request is generated and the index structure is searched to identify a node corresponding to the identification object. Responsive to identifying an node corresponding to the identification object, a slide presentation document is identified and access to the slide presentation document is provided to the user.

Keywords: presentation assistant, machine learning, optical character recognition, intelligent document identification

