Identifying Interesting Video Segments

Victor Carbune
Alexandru Damian

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

Recommended Citation
Carbune, Victor and Damian, Alexandru, "Identifying Interesting Video Segments", Technical Disclosure Commons, (May 07, 2020)
https://www.tdcommons.org/dpubs_series/3221

This work is licensed under a Creative Commons Attribution 4.0 License.
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.
IDENTIFYING INTERESTING VIDEO SEGMENTS

ABSTRACT

Disclosed herein is an improved mechanism for identifying interesting video segments. The mechanism can identify a segment of a video being viewed by a user. The mechanism can determine whether the video segment is likely to be of interest to the user using a trained classification model. The mechanism can then, for example, save an indication of whether the segment is likely to be of interest to the user. The system can then loop back and identify a second segment of the video being viewed by the user.

BACKGROUND

It can be useful to identify interesting portions of a video, such as scenes that contain important information, scenes that include crucial action points, etc. For example, by identifying interesting portions of a video, a video content sharing service that hosts the video may be able to create a compelling summary of the video. As another example, by identifying interesting portions of a video to a viewer of the video, the viewer may be able to more easily jump to an interesting portion. However, it can be difficult to accurately identify interesting portions of videos. Thus, there is a need for an improved mechanism for identifying interesting video segments.

DESCRIPTION

The systems and techniques described in this disclosure relate to identifying interesting video segments. The system can be implemented on a server, such as a server associated with a video content sharing service.
Note that the systems and techniques described in this disclosure can be used in any suitable application. For example, in identifying interesting video segments, the system can enable a video creator to identify a portion of a video that can be removed, thereby reducing noisy data and/or irrelevant scenes. In another example, in identifying interesting video segments, the system can enable a user consuming the video to skip to interesting video segments that are of interest to the user (e.g., most interesting scenes), thereby reducing noisy data and/or irrelevant scenes. In yet another example, in identifying interesting video segments, the system can generate improved summaries of videos, thereby reducing noisy data and/or irrelevant scenes.

FIG. 1 illustrates an example process for identifying interesting video segments.

At 102, the system can identify a segment of a video being viewed by a user. In some implementations, the system can identify the segment of the video in any suitable manner. For example, in some implementations, the system can identify a first portion of the video that corresponds to a beginning of the video. As another example, in some implementations, the
system can identify a portion in the middle of the video. In some implementations, the segment can be of any suitable duration (e.g., three seconds, five seconds, and/or any other suitable duration).

At 104, the system can determine whether the segment is likely to be of interest to the user using a trained classification model. In some implementations, a segment determined as likely to be of interest to the user can be determined as likely to be of interest for any suitable reason. For example, in some implementations, the system can determine that the segment is likely to be of interest to the user based on a current interest of the user. As a more particular example, in an instance where the system determines that the user is interested in learning to play a particular video game or a particular portion of the video game, and where the system determines that the segment includes content relevant to the particular video game or the particular portion of the video game, the system can determine that the segment is likely to be of interest to the user.

In some implementations, the system can determine whether the segment is likely to be of interest to the user using any suitable technique or combination of techniques. For example, as described above, in some implementations, the system can use the segment as an input to a classification model that has been previously trained on other videos and that returns as an output a classification that indicates whether the segment is likely to be of interest to the user. In some implementations, the model can additionally take as input any suitable information, such as information about the user (e.g., a watch history of the user indicating videos the user has previously viewed, topics or genres of media content the user has indicated interest in, a geographic location of the user, and/or any other suitable user information), information about the video (e.g., keywords of the video, a name of a creator of the video, a duration of the video, a
number of viewers of the video, and/or any other suitable information). Additionally, in some implementations, the model can take as an input user interaction information related to the video, such as indications of portions of the video other viewers of the video have expressed interest in (e.g., portions of the video that other users have paused, portions of the video that other users have rewound and/or viewed multiple times, and/or any other suitable user interaction information). More detailed techniques related to training of the classification model are described below in connection with FIG. 2.

At 106, the system can save an indication of whether the segment is likely to be of interest to the user. In some implementations, the system can save the indication in any suitable manner. For example, in an instance where the system determined that the segment is likely to be of interest, the system can store an indication of timing information corresponding to the segment (e.g., frames of the video corresponding to the segment, time interval information corresponding to the segment, and/or any other suitable timing information) in connection with an identifier of the video.

In some implementations, the system can then loop back to 102 and can identify a second segment of the video. For example, in some implementations, the system can identify a second segment of the video that is subsequent to the segment of the video.

In some implementations, the system can perform any suitable action(s) using segments that are determined to be of interest to the user. For example, in some implementations, the system can generate a summary video of the video that includes segments of the video that have been identified as likely to be of interest to the video appended together. As another example, in some implementations, the system can select a segment that has been determined to be likely of
interest to the user and can recommend the segment to a second user (e.g., a connection of the user, a second user with similar interests as the user, and/or any other suitable second user).

Note that, in some implementations, the identified segments determined to likely be of interest to the user currently viewing the video can be used by any other suitable user or entity. For example, in some implementations, identified segments can be recommended to another user, as described above. As another example, a summary video created from the identified segments can be presented to another user as a recommendation for the full video.

FIG. 2 illustrates an example process for training a classification model to identify interesting video segments. The process illustrated in FIG. 2 can be executed by any suitable device, such as a server associated with a video sharing service, as described above.

At 202, the system can identify a group of videos. In some implementations, the identified group of videos can include any suitable videos. For example, in some
implementations, the identified group of videos can include videos each relating to a similar topic (e.g., a particular game, a particular video game, and/or any other suitable topic), videos that have recently been uploaded to a video sharing service, videos associated with a particular group of channels of a video sharing service, and/or any other suitable videos. Note that, in some implementations, the system can identify any suitable number of videos (e.g., one hundred, one thousand, ten thousand, and/or any other suitable number).

At 204, the system can extract user information corresponding to users who have viewed videos in the group of videos. In some implementations, the user information can include any suitable type of user information.

For example, in some implementations, the user information can include information about the users who have viewed videos in the group of videos, such as demographic information, geographic information indicating a location of a user, interests of a user, and/or any other suitable user information.

As another example, in some implementations, the user information can include user interaction information corresponding to a user viewing a video in the group of videos. As a more particular example, the user interaction information can include portions of a video that a user or multiple users rewound (e.g., to view multiple times). As another more particular example, the user interaction information can include portions of a video that a user or multiple users paused. As yet another more particular example, the user interaction information can include portions of a video that a user or multiple users fast-forwarded or skipped over.

At 206, the system can train a classification model using features of the videos in the group of videos and the user information identified at 204. In some implementations, the classification model can produce an output that indicates a classification of whether an input
video segment is likely to be of interest to a particular user, as described above in connection with FIG. 1.

Note that the system can train the classification model in any suitable manner. For example, in some implementations, the system can construct a training set using features of the videos in the group of videos (e.g., metadata of a video indicating a title of a video, keywords of a video, a creator of a video, and/or any other suitable information) and the user information. As a more particular example, in some implementations, each sample in a training set can correspond to a particular video or a segment of a particular video in the group of videos and can include one or more features of the video and one or more items of user information corresponding to one or more users who have viewed the video. In some implementations, each training sample can be associated with an indication of whether the video or the segment of the video corresponding to the training sample was of interest to a user (e.g., based on manual annotation of the video or the segment of the video, and/or determined in any other suitable manner).

In some implementations, the classification model can include any suitable classification algorithm (e.g., a Support Vector Machine, a logistic regression, a neural network, and/or any other suitable algorithm). In some implementations, the classification algorithm can be trained using the training set described above using any suitable technique or combination of techniques.

Accordingly, a mechanism for identifying interesting video segments is provided.