AUTOMATIC VIDEO CONTENT ITEM CREATION FROM STATIC CONTENT ITEMS

Eli Danziger

Amit Sharma

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

Recommended Citation
Danziger, Eli and Sharma, Amit, "AUTOMATIC VIDEO CONTENT ITEM CREATION FROM STATIC CONTENT ITEMS", Technical Disclosure Commons, (April 22, 2016)
http://www.tdcommons.org/dpubs_series/194

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.
AUTOMATIC VIDEO CONTENT ITEM CREATION FROM STATIC CONTENT ITEMS

In a computer networked environment such as the Internet, entities such as people or companies provide information for public display on web pages or other online documents. The documents can include information provided by the entities via a computing device for display on the Internet. Content items can also be provided by third parties for display on the documents together with the information provided by the entities. Thus, a person viewing a document can access the information that is the subject of the document, as well as third party content items that may or may not be related to the subject matter of the document. Many of these content items (e.g., advertisements) on the Internet are static in nature, and only contain a combination of text and images that describe the products or services. Compared to video content items, static content items are generally less engaging to users. As a result, the products or services depicted in such static content items do not receive as much awareness as they could with video content items. Users typically spend more time on the Internet on activities besides searching for products or services. To such users, video content items can offer a more native experience (e.g., video content items on video online documents). Furthermore, video content items may allow users to engage with the product or service depicted for some amount of time. In contrast, users may scroll past static content items without much engagement. Video content items can also be more enticing to users with a more robust canvass than static content items. Having video content items could increase the conversion rates for the third party associated with the content items and consequently the awareness for their products or services. However, for many third parties, creating video advertisements could be challenging, particularly because of their lack of resources to put together such content items. Thus, having an automatic video content item creation system that can render readily available and preexisting static content items could significantly lessen the burdens faced by these parties in creating video content items.

This paper discusses an automatic video content item creation system that entities such as product vendors and service providers can use to be able to create video content items from preexisting static content items. To this end, the automatic video content item creation system can: retrieve static content items by accessing a database; extract content item features from the
static content items using various algorithms; determine which video templates would be the most suited for the extracted set of content item features; and render a video content item based on the video template selected and the set of content item features extracted.

Referring to Figure 1, an example of a system for automatically creating video content items from static content items (100) is shown. The content storage (105) can store content item records. Content item records can include the static content items and also other information associated about the item (e.g., products and services). Static content items can include text and images describing the item. The text can include name or title of the item, reviews or testimonials about the item by users, and item descriptions by the third party. Images can include screenshots of the item, item or third party logo, or any graphic associated with the item.

The retrieval module (110) can access the content item records from the content storage (105). The retrieval module (110) can also access the static content items from the content storage (105). The retrieval module (110) can identify the static content items from the content item records retrieved or accessed from the content storage (105). The retrieval module (110) can also convert the static content items from one format to another compatible for further processing by the automatic video content item creation system (100).

The extraction module (115) can extract some content item features from the static content items based on whether the static content item is text or an image. If the static item content is text, the extraction module (115) can extract phrases from the text static content item as the content item feature using natural language processing algorithms. If the static item content is an image, the extraction module (115) can select some of the image static content item as the content item feature using image processing techniques.

For text static content items, the extraction module (115) can first remove some of the text static content items based on their ratings or ratings by other users. For example, the extraction module (115) can remove reviews and testimonials that have given low evaluation scores for the item (e.g., 3/5 stars). The extraction module (115) can also remove reviews and testimonials that have been voted as not helpful by other users.

Second, the extraction module (115) can extract phrases from the text static content items using key phrase extraction algorithms. For example, the extraction module (115) can use a supervised learning key phrase extraction algorithm that has been trained on corpora that can include other text static content items with the key phrases tagged. Using the trained key phrase
extraction algorithm, the extraction module (115) can find candidate key phrases from the text static content items.

Lastly, the extraction module (115) can then determine the text quality scores of the candidate key phrases based on a number of parameters. The parameters can include phrase length, use of lengthy words within the phrase, grammatical correctness, and idiomatic correctness. The extraction module (115) can then select the candidate key phrases as the content item features based on the rankings of the text quality scores determined.

For image static content items, first the extraction module (115) can remove substantially similar or duplicate images. The extraction module (115) can use various image comparison algorithms to determine which images are substantially similar to one another. For example, the extraction module (115) can use a nearest neighbor algorithm. In this example, the extraction module (115) can identify certain interest points or features within the images, map the parameters of these interest points to an \( n \)-dimensional space, and group the images based on clustering in the \( n \)-dimensional space. The extraction module (115) can repeat these steps until the change of the mean of the groups between iterations is within a certain threshold. The extraction module (115) can then determine that the images within these groups are substantially similar to another and remove some of the images within each group.

The extraction module (115) can then determine the visual quality scores for each image static content item based on attributes of the image. Attributes can include, for example, resolution of the image, features or interest points found in the image, and aesthetics of the image. The extraction module (115) can assign a higher visual quality score for image static content items with higher resolutions, based on type or number of features, and those with higher contrast or matching colors. The extraction module (115) can then select some of the image static content items as the static content item features based on the rankings of the visual quality scores determined.

The video template storage (120) can store video templates. Each video template can include a set of scenes. Each scene can include a number of scene parameters that specify the number, type, and placement of static content items in the scene, effects in the scene, and the transitions from the current scene to the next.

The template matching module (125) can match the scenes of the video template to the static content item features based on suitability. The suitability can be a measure based on the
static item features extracted and the scenes of video templates. The template matching module (125) can determine the suitability based on a supervised learning algorithm. The supervised learning algorithm can be trained with data that can include scenes of the video template along with the type of static content items suitable for the scene labeled. Using the trained algorithm, the template matching module (125) can determine the suitability of the static content item features extracted to the scenes of the video template. Based on this determination, the template matching module (125) can assign scenes of the video template to the static content item features extracted. The trained algorithm of the template matching module (125) can also take feedback from the user. For example, if a number of users of the automatic video creation system (100) reject the assignment of the scene of the video template to the static content item feature, the template matching module (125) can label the assignment as unsuitable and adjust the model accordingly.

The matched template storage (130) can store the assignment of the scenes of the video template to the static content item features. The matched template storage (130) can also create and store an identifier for the assignment, such as a URL. The template retrieval module (140) can retrieve the static content item features stored in the matched template storage (130) by using the identifier.

The video render module (140) can render the video content based on the static content item features assigned to the scenes of the video template retrieved by the matched template storage (130). The video render module (140) can apply the specifications of the scene parameters in each scene of the video template to create the video content. The video storage (145) can store the video content rendered by video render module (140). The video storage module (145) can also serve the render video for display by a computing device via a network in response to a user request from the computing network.
ABSTRACT

This paper discloses an automated video creation system by processing static content, such as text and images. The system can apply natural language processing algorithms such as key phrase extraction to extract phrase from a database of text. The system can execute image processing algorithms to select images from a database of images. The system can then determine which video templates are most suited for the extracted text and the selected images, and render the video according to the specifications of the selected video template.