December 2019

Travel itinerary mapping and planning using image search

Diane Wang

Aiko Nakano

Travel itinerary mapping and planning using image search

ABSTRACT

Users refer to online content such as travel blogs, travel videos, etc. for inspiration regarding places to visit and things to do. However, such content often does not include place annotations. Due to this, the user needs to perform multiple web searches to identify points referred to in the content. This disclosure describes techniques to provide users with information based on analyzing an image. The user can provide an input image or video, which is analyzed to recognize depicted objects or places. Information regarding the identified objects, e.g., location, knowledge base information, additional images of the object, place suggestions based on the object, is identified via map applications, video applications, knowledge bases, etc. A user interface is provided that enables the user to view the information and store it in their personal collection, e.g., a collection of points of interest. When the user engages in trip planning, points of interest from the collection are surfaced to the user.

KEYWORDS

- Travel planning
- Trip planning
- Landmark
- Point of interest
- Visual search
- Travel inspiration
- Travel itinerary

BACKGROUND

Many users refer to videos, e.g., video travelogues, and/or other image content, e.g., photos on travel blogs, etc. from online influencers as a source of travel inspiration and to inform their decisions on visiting points of interest (POI). However, it is difficult for a user to gather destination and specific POI information from a video or a phone unless such information is explicitly mentioned by the content creator. Therefore, to obtain such information, users currently either engage in social interaction, e.g., by asking questions via a comments section for the video or other content to obtain guidance regarding specific content and associated POI, or by performing multiple web searches to uncover the destination.

Currently, maps applications allow users to perform text searches for points of interest and save those on a map. However, a user needs to know the keywords in order to perform such searches. Users can also use online collections to save images or video content to perform deep-dive research at a later time. Content curation platforms enable users to generate trip itineraries and/or curated videos based on user selected POI.

DESCRIPTION

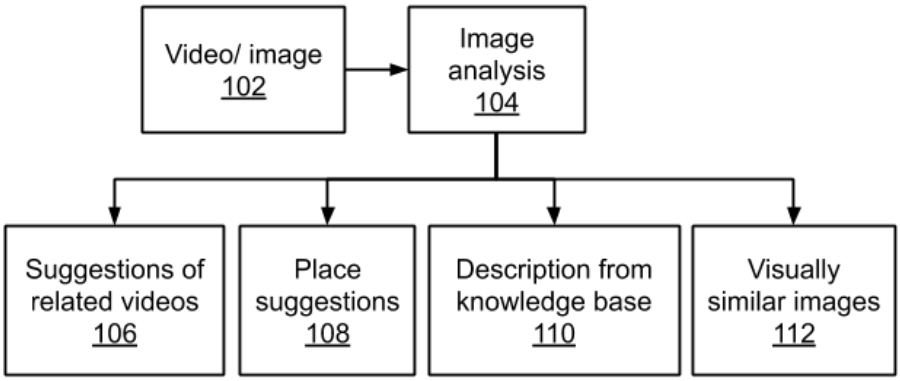

Fig. 1: System architecture

Fig. 1 illustrates an example system architecture to enable video or image based identification of points of interest and to perform travel itinerary mapping and planning. A video or image (102), e.g., from a travel video or blog, is provided to for image analysis (104). For example, computer vision techniques can be used to identify entities that are depicted in the image. Identification of the entities can include determining the name of the object and associated confidence score for the detection.

If the image is a user-selected video frame and no entity is detected with a threshold level of confidence from the frame, additional frames near the user-selected frame - having timestamps before or after the frame - are also analyzed. If the image is a static image with detection score below a threshold, a cropping tool is provided to the user to enable the user to select a region of the image that depicts the content of interest. This is useful, e.g., when there are multiple objects in the frame. Upon user selection, another round of image analysis is performed.

Based on the identification of objects in the image, related information, e.g., suggestions of related videos (106); place suggestions (108), e.g., of places recognized as being depicted in the image; description of the object from a knowledge base (110); and/or visually similar images (112) is identified and provided to the user. Such content is retrieved from various platforms such as video, maps, image databases, and knowledge bases by querying such platforms using the name of the object.

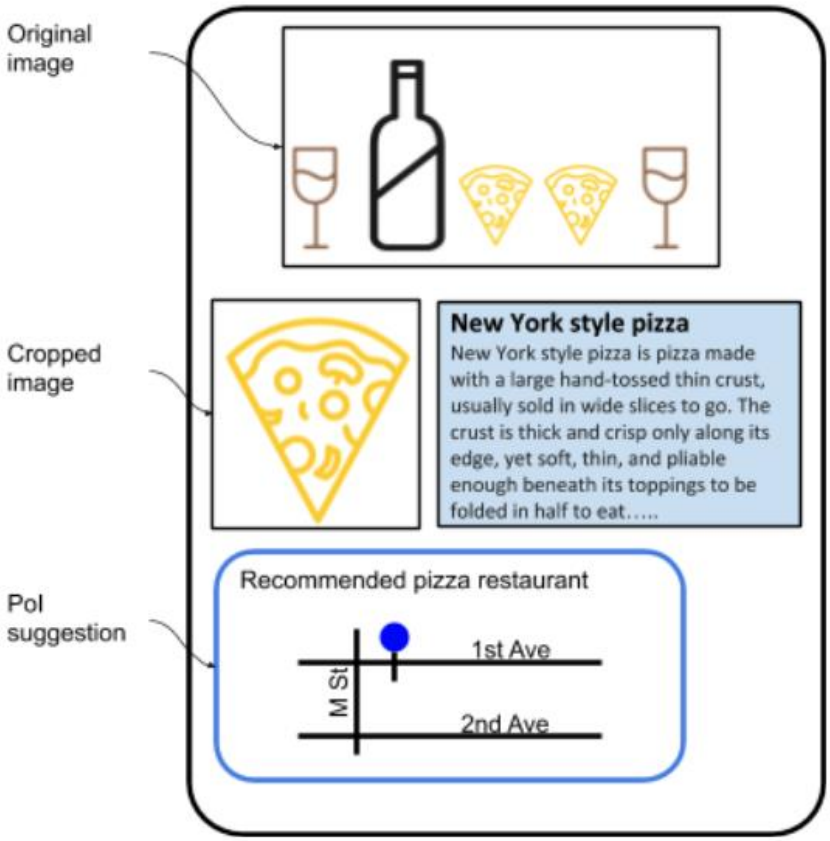

Fig. 2: Example results panel

One example of results displayed to the user is shown in Fig. 2. In this example, the user \ watches a video related to pizza and selects a frame of the video (original image). The cropped image of the pizza is analyzed to determine that it includes a New York style pizza (highest score) and possibly other objects (e.g., wine bottle, wine glass - with lower scores).

Results are displayed for the object with the highest score, e.g., in a results panel, as illustrated in Fig. 2. The results panel includes a cropped version of the input image, the name of the object, and a description of the object from a knowledge base. The results panel can also include additional images of the object, e.g., from the knowledge base or other resources, or other visually similar images (e.g., if the object was not identified). The results panel also includes point of interest suggestions - in this instance, a recommendation for a pizza restaurant

that sells New York style pizza that is near to the user (or to a previous search result for a POI) is provided. Videos related to New York style pizza can also be provided (not shown). Multiple points of interest can be saved on the map from a single frame. The results panel is automatically updated based on user selection, even when the object has a lower confidence score.

Object identification and corresponding recommendations are applicable to points of interest (e.g., the Eiffel Tower, the Statue of Liberty, etc.) as well as objects (e.g., New York style pizza) that are not associated with a specific location. If a specific location does not exist for the object, content recommendations from the map or video platforms are used to show suggestions.

Further, a collection of places that the user has previously indicated interest in, e.g., based on images that have been analyzed, can be provided to the user in a maps application to assist with travel planning. The map can be centered to particular locations that have been identified, e.g., based on a video. A point of interest identified using image analysis can be saved to a maps application. The collections of places can be shared with other users directly from the session (e.g., a temporary list) or from a maps application (in which case it can include saved data from multiple similar sessions).

If no object is recognized, visually similar images can be provided as results or additional images for the user perform searches from. The video platform, maps platform, knowledge base, etc. can be accessed by using APIs provided by the respective platform to query based on the object identified in the image.

The described techniques leverage computer vision techniques in combination with maps and video platforms and knowledge bases to automatically identify points of interest from an image or video. While watching a video, a user can select a video frame to obtain more

information about the depicted content, e.g., to identify and save a point of interest in a maps application. The user can then utilize the maps application to plan a trip; the saved points of interest are automatically shown to the user.

While the disclosure describes techniques for providing users with a point of interest and supplementary information from a cropped image from video content, the described techniques can be utilized for visual content from any source such as social media sites, webpages, images, etc. For example, the techniques can enable a user to find the location of a landmark shared on a social media platform by a travel blogger. Further, the techniques can be used for language acquisition (e.g., by performing optical character recognition on foreign language text in an image) and/or for educational purposes for subjects such as art history (e.g., by searching for images of art) or other subjects. Users can store the search results in custom collections.

Further to the descriptions above, a user may be provided with controls allowing the user to make a selection 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 techniques to provide users with information based on analyzing an image. The user can provide an input image or video, which is analyzed to recognize depicted objects or places. Information regarding the identified objects, e.g., location, knowledge base information, additional images of the object, place suggestions based on the object, is identified via map applications, video applications, knowledge bases, etc. A user interface is provided that enables the user to view the information and store it in their personal collection, e.g., a collection of points of interest. When the user engages in trip planning, points of interest from the collection are surfaced to the user.

REFERENCES

1. Jing, Yushi, and Jian Yuan. "Inferring locations from an image." U.S. Patent 8,472,664, issued June 25, 2013.