February 2020

Personalized Augmented Reality Content Delivery Using Computer Vision

Aiko Nakano

Diane Wang

Personalized Augmented Reality Content Delivery Using Computer Vision

ABSTRACT

This disclosure describes techniques that personalize, with user permission, a user's real surroundings with augmented reality content that is matched to user preferences and history. With user permission, an AR device, e.g., AR glasses, a head-mounted device, etc., discovers physical objects in the user's surroundings. The device downloads AR content relevant to the physical objects that match user preferences and prior user history. The downloaded AR content is delivered to the user via text-to-speech, audio, visual, or audiovisual mediums. A next point-of-interest in the user's journey is suggested.

KEYWORDS

- Augmented reality
- Virtual reality
- Computer vision
- Personalized content
- Point of interest
- Smart glass
- Head-mounted device

BACKGROUND

Immersive experiences incorporate large corpora of data that are difficult to navigate through and disambiguate. Augmented reality (AR) content that is indiscriminately sourced and dispersed within a user's field of vision can lead to information overload and confusion.

DESCRIPTION

This disclosure describes techniques that, with user permission, leverage contextual understanding of a user to prioritize both the immersive content to be surfaced and the mode of surfacing such content in the user's world space. The techniques enable a seamless, quick, and user-friendly experience that conveys AR information that is relevant to the user.

With user permission, the techniques use cameras in the AR device, e.g., head-mounted device (HMD), smart glasses such as AR or VR glasses, smartphones, or other wirelessly connected, programmable device with a camera, etc. to determine the user's interest, based on various factors such as the amount of time the user engages with a particular physical object. The physical object can be gleamed (marked with an AR marker). The user's interest is used to determine personalized recommendations of AR content and of the next point-of-interest to explore. The AR content can be delivered using various modes, e.g., the AR device can read textual content out loud using text-to-speech (TTS); via a companion app on the user's smartphone if the experience benefits from a larger screen surface or touch-based user interactions; etc.

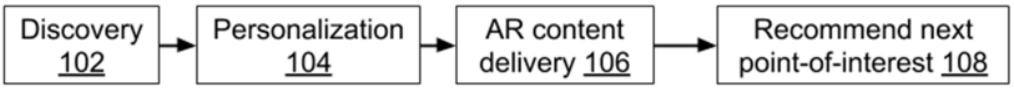


Fig. 1: Personalized AR content delivery using computer vision

Fig. 1 illustrates personalized AR content delivery using computer vision, per techniques of this disclosure.

Discovery of physical objects with AR content (102): With user permission, a camera on the AR device captures images and uses computer vision (CV) techniques to detect the presence of objects in the user's world-space vicinity that have associated AR content. In addition, GPS or

other positioning techniques in conjunction with a database of preexisting AR content in the user's vicinity can be used to detect physical objects of interest with associated AR content.

Personalization (104): Detected and identified objects are checked against a database of objects that are of interest to the user to determine if the presently-detected object is of interest to the user. The database of objects that are of interest to the user (also known as preference table) can be based on previous sessions and/or the current session. For example, with user permission, the length of time spent by a user in observing an object can be used as a factor in determining if that object is an object of interest to the user. Objects that are determined to be of interest are added to the database.

AR content delivery (106): As mentioned above, the identified object is determined to be of interest to the user if the user observes it continuously for at least a certain duration, e.g., five seconds. A user can move onto content delivery without experiencing the latency of user-interest determination by tapping on the wearable device. For objects detected to be of interest to the user, AR content is delivered using various modes that are selected based on the type of content that is attached to the object. The AR content may be delivered in one the following example modes: text-to-speech, e.g., audio, mode for text-rich content; video mode for content that is optimally presented on a larger screen interface; interactive mode for content that is optimally presented using touch-based user interactions; audiovisual mode for content that has substantial audio and video content; etc. Audio content can be delivered directly to a wearable AR device or to another device such as a smartphone. Video or interactive content can be delivered by a companion app on the user's smartphone. The techniques are agnostic to device form factors, and automatically pick the best available content delivery platform, e.g., wearable, smartphone, etc., based on the content. For example, if the user is wearing an audio-only wearable, the

techniques provide AR content, including guidance and navigation, using audio instructions. Audio-only navigation can be provided, e.g., by using statements such as “The painting is at your 12 o’clock.” If the user has a smartphone, AR content is delivered using AR overlay on the smartphone screen, by automatically redirecting the user to a website, etc. The user can specify preferences for the mode of AR content delivery as well as for the granularity of information that is delivered.

Recommend next point-of-interest (108): After the AR content is delivered to the user, recommendations of other points of interest are provided to the user. The recommendation is based on the updated user-preference table and on a match of the user-preference table against the database of points of interest in the vicinity of the user.

The techniques are illustrated by an example of their use at an art museum.

Example: Personalized AR content delivery in a museum: A user visits a museum. As the user looks at the exhibits, the AR techniques described herein discover exhibits of interest to the user, deliver personalized AR content to the user, and recommend a next exhibit. Each of these features is illustrated in greater detail below.

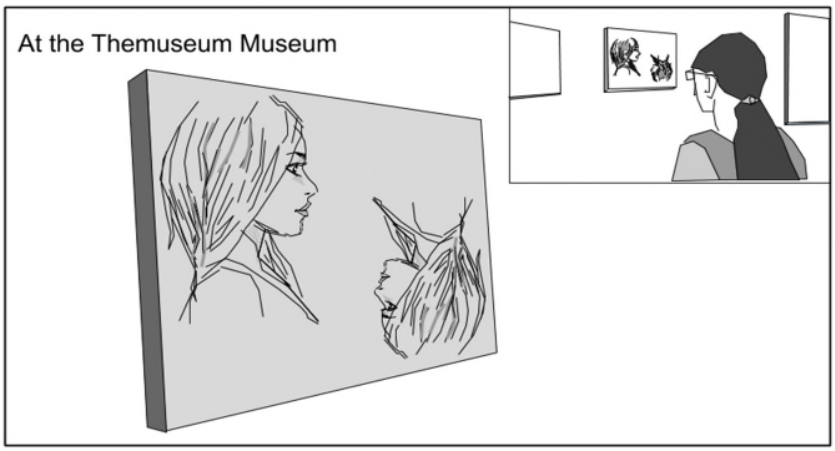


Fig. 2: Discovery - the augmented reality view, as seen by the user through an AR device; Inset: The real world space of the user

- **Discover exhibits of interest to the user:** As illustrated in Fig. 2, with user permission, a camera on the AR device captures images and uses computer vision techniques to detect exhibits or artworks in the user's world-space vicinity that have associated AR content.

Fig. 3: Personalization. Augmented reality content personalized to the user; Inset: The real world space of the user

- **Personalization:** Detected and identified objects are checked against a database of objects that interest the user to determine if the presently-detected object is of interest to the user. As illustrated in Fig. 3, one or more exhibits are recommended based on matches between artworks detected in the vicinity of the user and the user's preference table, which is itself based on the user's previous sessions. Matches can be based on the properties of the objects. For example, for artworks, matches can be based on artist; art category, e.g., American, Contemporary, Renaissance, etc.; art style, e.g., impressionism, pointillism, cubism; etc. The extent of match between an object and the user preference is measured by a score, and matches are ranked and prioritized based on the score. Score computation can include differential weighting of object properties.

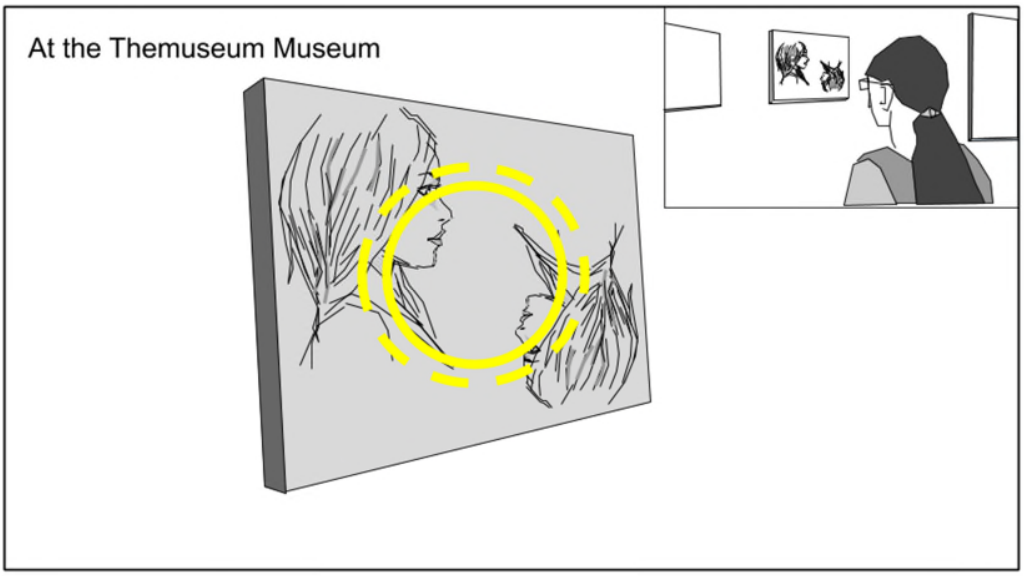
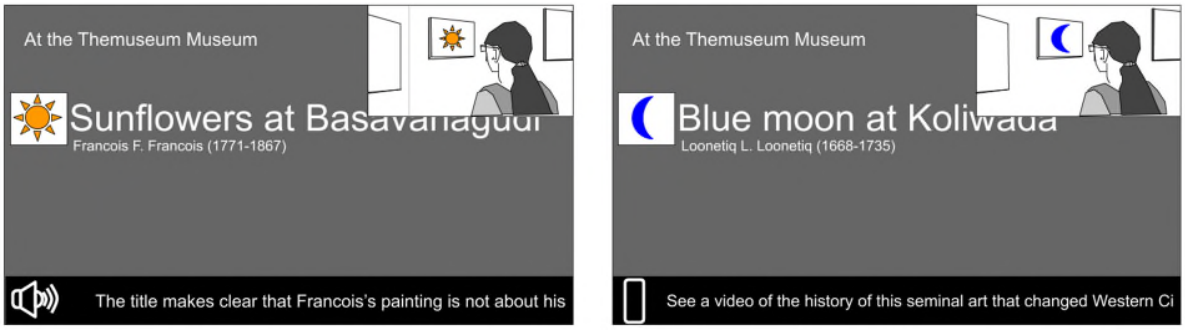


Fig. 4: Using AR overlay to identify or recommend objects of interest to the user; Inset: The real world space of the user

As illustrated in Fig. 4, if a match is detected in the user's field of view, a notification can be provided to the user, e.g., by overlaying an AR marker (shown in yellow) on top of the recommended object.



(a) **(b)**

Fig. 5: AR content delivery (a) TTS for a wearable AR device (b) Offloading AR content to a smartphone

- **AR content delivery:** AR content may be delivered in one the following example modes: text-to-speech, e.g., audio, mode for text-rich content; video mode for content that is optimally presented on a larger screen interface; interactive mode for content that is

optimally presented using touch-based user interactions; audiovisual mode for content that has substantial audio and video content; etc. Fig. 5 illustrates two example modes of AR content delivery, e.g., audio using text-to-speech (Fig. 5(a)), and video via a smartphone (Fig. 5(b)).

Fig. 6: Recommending the next point-of-interest

- As illustrated in Fig. 6, after the AR content is delivered to the user, recommendations of other artworks are provided to the user. The recommendations can be based on artworks that the user has liked, including artworks from the present visit, e.g., "your trip today" section in Fig. 6, and from previous visits.

Additional examples of personalized AR content delivery in other situations are below:

Example: Personalized AR content delivery in stores: The described techniques enable a user to be guided to products that suit their taste and fit their style. A user contemplating the purchase of an item at a store can be given AR information about the item. With user permission, it can be detected that a user has shown interest in a certain product, e.g., by observing it for an extended

duration. The user can then be guided using AR audiovisual channels to accessories that match that product. In this manner, the techniques help a customer reach a purchasing decision fast and help retailers move inventory.

Example: Personalized AR content during travel: The techniques enable a user to receive a personalized, AR-based, guided tour of a city based on their interests, e.g., an architectural tour, a history walk, etc. The AR-guided tour can also translate signs into the user's native language.

Example: Personalized AR content in out-of-home ads: The techniques enable a user to discover relevant media in world space that fit the profile of the user. Also, an advertiser can present advertisements that are personalized to the user. For example, an ad for a car company can display an SUV in the AR field-of-view for one user and a minivan for another.

Example: Personalized AR content for wayfinding in events: The techniques enable a user to receive personalized guidance and navigation during events. For example, at a conference, the user can be provided a schedule of sessions of interest via AR audiovisual channels. At a trade fair, the user can be guided via AR audiovisual channels to stalls that match their interests.

In this manner, the techniques of this disclosure can turn everyday journeys, e.g., through a downtown, a building, an office, a museum, a store, etc., into a highly personalized journey with high-density and relevant information.

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 interests, gaze direction, social actions or activities, a user's preferences, or a user's current location), and if the user is sent content or communications. 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 that personalize, with user permission, a user's real surroundings with augmented reality content that is matched to user preferences and history. An AR device, e.g., a smart-glass, a head-mounted device, etc., discovers physical objects in the user's surroundings. The device downloads AR content relevant to physical objects that match user preferences and prior user behavior. The downloaded AR content is delivered to the user via, e.g., text-to-speech, audio, visual, or audio-visual mediums. A next point-of-interest in the user's journey is suggested.