March 22, 2019

Selective lifelogging to record missed events of interest

Victor Cărbune

Thomas Deselaers

Selective lifelogging to record missed events of interest

ABSTRACT

A user can sometimes miss content or events of interest due to distraction or inattention. Rewinding and replaying to view the missed content or event may not be possible when the missed situation is live and not pre-recorded. Further, the user may view the content as a real-world experience rather than via a computing device. This disclosure describes techniques that make it possible to replay moments from the immediate past, thus enabling an experience for the real world similar to that of the rewind functionality for recorded content. With the user's permission, a recording is triggered when it is determined that the user is likely to miss an upcoming event of interest, enabling the user to retrieve and play the missed event. A trained event interpretation model is used to determine the event of interest from the recorded data.

KEYWORDS

- Event capture
- Event replay
- Event interpretation
- Lifelogging
- User distraction
- Memorable moment

BACKGROUND

When using a device to watch content, such as movies or talks, a user's attention can get diverted by external events. As a result, the user may miss a part of the content being watched. When the content being viewed is a recording, the user can compensate for the disruption by rewinding to a last watched content location prior to the occurrence of the attention diverting

event. However, needing to rewind often due to frequent occurrence of attention diverting events can be inefficient and frustrating. Moreover, rewinding is sometimes not possible, e.g., when the content being watched is live and not pre-recorded. Further, the user may view the content as a real-world experience rather than via a computing device.

DESCRIPTION

This disclosure describes techniques that make it possible to replay moments of life from the immediate past, thus enabling an experience for the real world similar to that provided by rewind functionality for recorded content.

With the user's permission, a trained lightweight scoring model is executed on a user device to detect potential real world situations that are likely to be missed by the user. Such situations include unexpected situations, such as a sudden sound, or rapid events, e.g., fast disappearing pop-ups within computing applications. If the user permits, contextual information obtained from the user device, e.g., currently running apps, values of sensors, user activities on the device, etc., is provided as input to the scoring model. The input is compared with threshold values to determine if the user is likely to miss a situation in the near future.

When the scoring model output indicates that the user is likely to miss an upcoming event, e.g., a real world event, a recording is automatically triggered. Such user permitted recording is achieved by turning on appropriate device sensors, such as camera, microphone, screen capture, etc. The specific sensors used for recording are dependent on the capabilities of the respective user devices.

For example, lifelogging cameras can record the physical world; smartphones can buffer screen activity, etc. Depending on the circumstances, not all device sensors may be useful. For instance, the camera of a smartphone is unlikely to record anything of interest if the device is in

the user's pocket. Therefore, with the user's permission, sensor information is used to determine the sensors that can serve to record information of likely interest to the user. Only those sensors that are deemed useful for recording are used, thus ensuring effective use of device resources, such as battery, memory, and processing power.

If the situation where the user misses an event of likely interest materializes, the user is then able to know about the situation by viewing recorded events from the recent past. Such requests to view situations of interest from the recent past can be issued, e.g., via appropriate voice commands, such as "What just happened?" The voice commands can be processed by voice-enabled assistive technologies, such as smart speakers or voice assistants on user devices.

When the user expresses a desire to know about a situation from the recent past, a trained event interpretation model is used to analyze available recordings to identify the event of interest to the user that was missed. The interpretation model can employ heuristics and/or classification models based on knowledge about user's attention, e.g., obtained by analyzing contextual information, such as sound, activities, etc., gathered from the user device.

In addition to requesting playing of the recording of the missed situation, the user can further request that the event be retained by storing to an archive. The ability to request storage enables the user to archive memorable events, e.g., a baby taking its first steps.

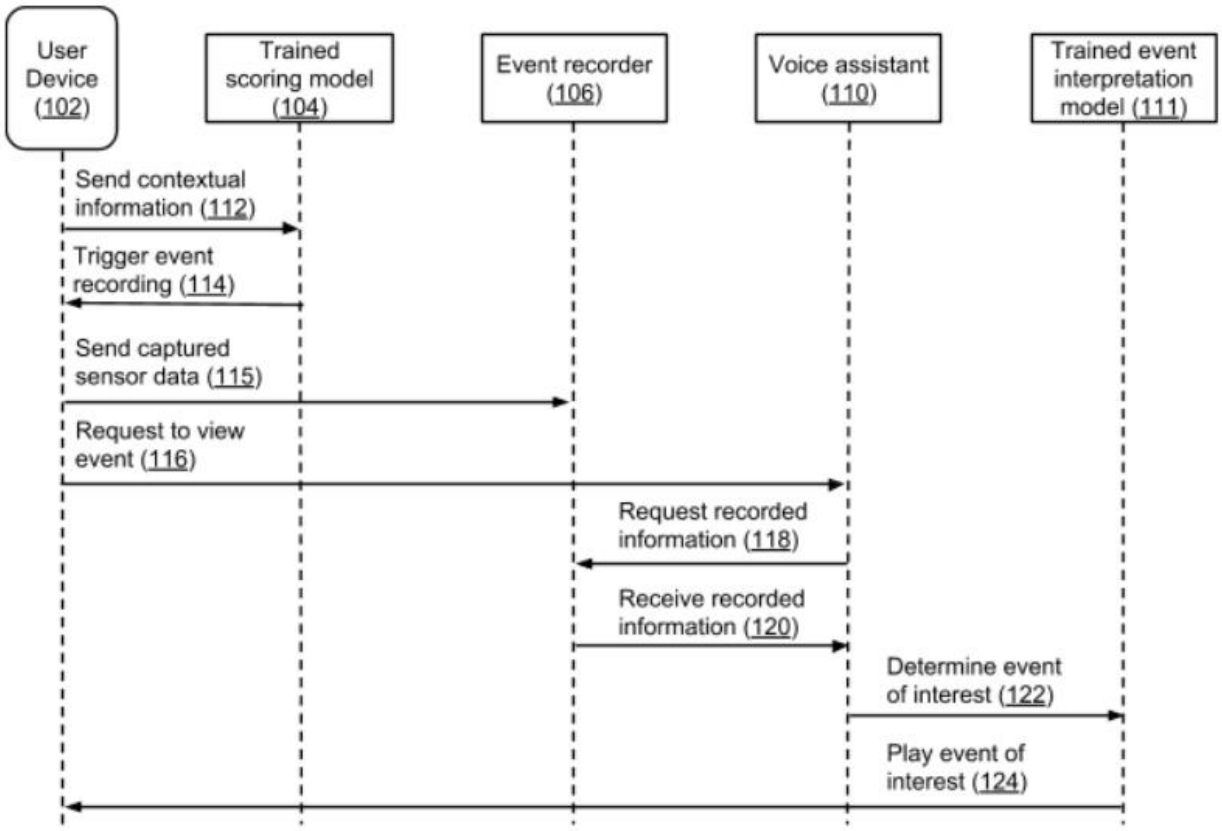

Fig. 1: Automatic capture and playback of events that a user is likely to miss

Fig. 1 shows an example operational implementation of the described techniques, implemented with user permission. With the user's permission, contextual information gathered from a user device (102) is provided (112) to a trained scoring model (104). If the output of the scoring model indicates that the user is likely to miss a situation of interest in the near future, recording is triggered (114) such that information from relevant device sensors is sent for recording (115) by an event recorder (106).

The user can then request, e.g., via a voice-activated assistant application (110) to view a missed event from the recent past. Upon receiving the request (116), recordings of the recent past are requested (118) and are obtained (120) from the event recorder. The received recorded information (120) is provided (122) to a trained event interpretation model (111) that determines

(122) the event of interest that was missed by the user. The available recording of the event is then played back (124) for the user. Although shown separately, the trained models, the event recorder, and the voice assistant can be implemented with the same user device. Alternatively, one or more of these components can be implemented on one or more separate devices, if permitted by the user.

The scoring and interpretation models can be implemented using neural networks that can handle a diversity of inputs. The output of the scoring model that indicates whether a situation occurring in the near future is likely to be missed by the user can be a threshold value, a confidence score, a Boolean classification, or a combination. The relevant threshold values can be specified by stakeholders, such as the model designer, the device manufacturer, the user, etc. Alternatively, the threshold values can be determined dynamically during operation. In addition to the training data, the learned scoring model can be enhanced and updated using appropriate user input.

The described techniques can be enhanced by incorporating heuristics that can indicate upcoming situations that are likely to be missed by the user. Such heuristics can be used in addition to, or in lieu of, the scoring and interpretation models. The techniques of this disclosure can be implemented as features of specific devices such as lifelogging cameras, as a part of the operating system for devices such as smartphones, wearable devices, smart home appliances, smart speakers, cameras, etc.

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 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 that make it possible to replay moments from the immediate past, thus enabling an experience for the real world similar to that of the rewind functionality for recorded content. With the user's permission, a recording is triggered when it is determined that the user is likely to miss an upcoming event of interest, enabling the user to retrieve and play the missed event. A trained event interpretation model is used to determine the event of interest from the recorded data. Optionally, the user can request archiving the event for memory. The techniques of this disclosure can be implemented as features of specific devices such as lifelogging cameras, smartphones, wearable devices, etc.