Technical Disclosure Commons

Defensive Publications Series

February 2020

Classifying Augmented Reality (AR) Effects

Anonymous Anonymous

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

Recommended Citation

Anonymous, Anonymous, "Classifying Augmented Reality (AR) Effects", Technical Disclosure Commons, (February 11, 2020)
https://www.tdcommons.org/dpubs_series/2947

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.

Classifying Augmented Reality (AR) Effects

Abstract

The present disclosure describes a method for classifying augmented reality (AR) effects using tags derived from content of the AR effects. A tag generator (for example, an image classifier, a natural language classifier) derives the tags from images or thumbnails with the AR effects that are given as an input to the tag generator. The tag generator includes an AR comparison unit, a confidence score generation unit and a dictionary database, which are utilized to generate the tags for the AR effects. Each of the tags includes a text string, a confidence score and metadata about how the tag was created.

Problem statement

The AR effect is a collection of modalities (for example, audio, visual, object actions/reactions) that enables users to augment and enhance images and videos in a manner they desire. But it is difficult for the user to effectively search the desired AR effect, if the AR effects are not properly tagged in an augmented reality (AR) application.

The present disclosure provides a novel solution to overcome the aforementioned problem.

System and working

The present disclosure proposes a method for classifying augmented reality (AR) effects using tags derived from content of the AR effects. A tag generator (for example, an image classifier, a natural language classifier) derives the tags from images or thumbnails with the AR effects that are given as an input to the tag generator. For example, the natural language classifier derives the tags from the images with the AR effects having text descriptions, that are given as an input to the natural language classifier.

The tag generator includes an AR comparison unit, a confidence score generation unit and a dictionary database. The AR comparison unit derives a set of attributes from the AR effect and compares the set of attributes to a set of text strings stored in the dictionary database. The text strings are provided by specialized knowledge sources (referred to as sources hereafter) for the AR effect. The AR comparison unit then generates a similarity index for each pair of the set of attributes. The similarity index is a measure of correlation between the set of attributes and the text string. The similarity index ranges from 0 to 1. If all of the set of attributes matches to the text string, the similarity index is assigned a value of 1. For example, if the AR effect depicts a "red heart" and the text string is "red heart", the similarity index is

Classified as Confidential-External

assigned the value of 1. However, if none of the set of attributes matches to the text string, the similarity index is assigned the value of 0. For example, if the AR effect depicts a “red heart” and the text string is “smiling face”, the similarity index is assigned the value of 0.

The AR comparison unit further sends the similarity index and the set of text strings to the confidence score generation unit. The confidence score generation unit generates a sourcemap, and metadata associated with the sourcemap based on the set of text strings and the similarity index. The sourcemap is a dictionary that lists all the sources for the tag generator, the set of text strings, and confidence scores that are assigned to each of the text strings. The confidence score indicates a likelihood of the text string matching the AR effect. The confidence score indicates a maximum percentage of number of sources that suggest the text string for the AR effect. For example, if 90% of the sources suggest that the AR effect depicts the “red heart”, the tag generator assigns the confidence score of 90% to the text string “red heart”. Further, the metadata provides information about how the confidence scores were generated. For example, the metadata includes a list of the sources that provides the text string matching the AR effect. Finally, the confidence score generation unit generates an output that includes the tag for the AR effect. The tag includes following components:

- the text string related to the content of the AR effect,
- the confidence score, and
- the metadata about how the tag was created.

The AR effect is assigned the tag, corresponding to which the confidence score is maximum. For example, the tag generator shown in Figure 1, takes the AR effect and creates the tag that includes the text string “heart” for the AR effect and the confidence score of 95%. This means that the tag generator is 95% confident about the presence of “heart” in the AR effect.

Figure 1: The tag generator

In another example, the tag generator may create the tag that further includes the metadata, for example, the list of the sources that predicted the text string “heart” for the AR effect. This way, the tag generator classifies the AR effect, and thus enables a user to effectively search the AR effect using the text string.

Additional embodiments

In an additional embodiment, the tag generator is implemented as a concrete tag generator. The concrete tag generator depends on a single source of information and does not depend on any other source for generating the tag for the AR effect. The concrete tag generator accesses fields and edges of the AR effect, and gathers information from the fields and the edges for deriving the tag for the AR effect rather than relying upon the multiple sources. In one example, the concrete tag generator is an XRAY classifier. The XRAY classifier does not directly receive the AR effect for generating the tag. Instead, an image generating function first creates the images from the AR effect that is to be classified. The images are then passed to the XRAY classifier. In one example, instead of directly passing the images to the XRAY classifier, a uniform resource identifier (URI) of the image, a string representing raw JPEG data of the image, or an ever store handle of the image is passed to the XRAY classifier. In another example, the thumbnail of the AR effect is passed to the XRAY classifier.

In yet another embodiment, the user may want to add a plurality of strings as the text strings for the AR effects with a 100% confidence score. To achieve this, the tag generator implements a strings collection framework that gathers the plurality of strings that are to be tagged with 100% confidence score from the user. The tag generator further stores the plurality of strings in the dictionary database, wherein each string is annotated to the corresponding AR effect with the confidence score of 100%. If the user searches for the AR effect tagged with 100% confidence score, the user needs to search by the string added earlier and the user is presented with the only one AR effect corresponding to the string.

In yet another embodiment, the tag generator is implemented as a compound tag generator that combines multiple tag generators. The main purpose of the compound tag generator is to hide the multiple tag generators inside a blackbox and provide a specialized logic to combine a series of confidence scores from all the tag generators for the tag. The compound tag generator receives the sourcemap from each of the individual tag generators. The compound tag generator implements a fusion framework that combines the series of confidence scores from all the tag generators into a final confidence score. The compound tag generator derives a final tag for the AR effect corresponding to which the final confidence score is maximum.

Classified as Confidential-External

In yet another embodiment, the fusion framework combines the series of confidence scores by assuming that each of the series of the confidence scores is a probability of a strictly independent event and computes a probability of all the sources being wrong. Thereafter, the fusion framework returns an inverse of the probability of all the sources being wrong. The inverse is a probability of at least one of the sources being right. For example, individual probabilities of sources A, B and C being right are 90%, 50% and 50% respectively. Therefore, individual probabilities of the sources A, B and C being wrong are 10%, 50% and 50% respectively. Further, the probability of all the sources being wrong at the same time is: 10% * 50% * 50% = 2.5%. So, the probability of at least one of the sources A, B and C being right is 100% -2.5% = 97.5%.

Conclusion

AR effect is an amazing feature available in augmented reality (AR) applications that lets a user to search a fantastical AR environment and then capture pictures, videos and selfies. But for providing swift experience to the user, the AR app should enable the user to effectively search a desired AR effect. The present disclosure provides a mechanism for classifying and tagging AR effects so that the user can effectively search the desired AR effect.

Classified as Confidential-External