Technical Disclosure Commons

April 2022

ADJUSTING VIRTUAL BACKGROUND TO CONVEY EMOTIONS

HP INC

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

Adjusting Virtual Background to Convey Emotions

Abstract

This disclosure describes a system that uses machine learning algorithms and computer vision techniques to detect user sentiment and dynamically adjust the virtual background content. As part of this disclosure scope, a curated list of visual assets (images or short videos) is associated with the potential psychological states. The output of the proposed system will replace, for a pre-defined amount of time, the virtual background content with the visual asset that better represents the user's current psychological state or emotion.

I. Introduction

Emotions are efficacious and self-explanatory in day-to-day human interactions. The most noticeable human emotion is through their facial expressions. During online events, such as virtual meetings and webinars, relevant facial expressions might not be perceived by the audience.

This disclosure describes a method that uses machine learning and computer vision to dynamically adjust virtual background images based on the detected emotion or facial expression. As one example, when installed and enabled, the attendee happiness will be detected and *augmented* through the usage of a happy and colorful background image, for a pre-defined amount of time.

II. Method

The idea described herein consists of a software component that will periodically run machine learning models to infer the overall sentiment of the detected face(s). Based on this inference result, the proposed algorithm will apply a combination of machine learning and computer vision techniques to apply a virtual background that matches the detected sentiment.

The process starts with the interception of the meeting video contents, so it can be decomposed into frames for processing. As shown with the help of Figure 1, each frame is processed aiming to detect faces.

When a face is detected, the proposed system will use the state-of-the-art machine learning models to infer the psychological state and to apply selfie-segmentation to remove the subject current background pixels.

The outcome of the inferred sentiment (or psychological state) is then used to find the associated background image. As described below, the proposed method considers the existence of a curated list of images (or short videos) associated with a list of possible psychological states (sentiments).

By having the selected visual asset (image or short video that matches the detected sentiment state) and the detected prominent human in the scene, the proposed system will stack both contents. At this point, the produced outcome will be the original frame with the background content replaced with the selected visual asset that matches the subject current psychological state (sentiment).

This modified frame is then released to be used by the different meeting provider client applications, such as Zoom, MS Teams, Google Meet, or any other software application that uses content captured from the user's web-camera.

i. Detecting Sentiment

Deep learning techniques has made facial expression recognition one of the most trending research fields in computer vision area. As such, there are plenty of options that can be used to achieve this goal. To the scope of this disclosure, we are interested in the produced outcome from the state-of-the-art in
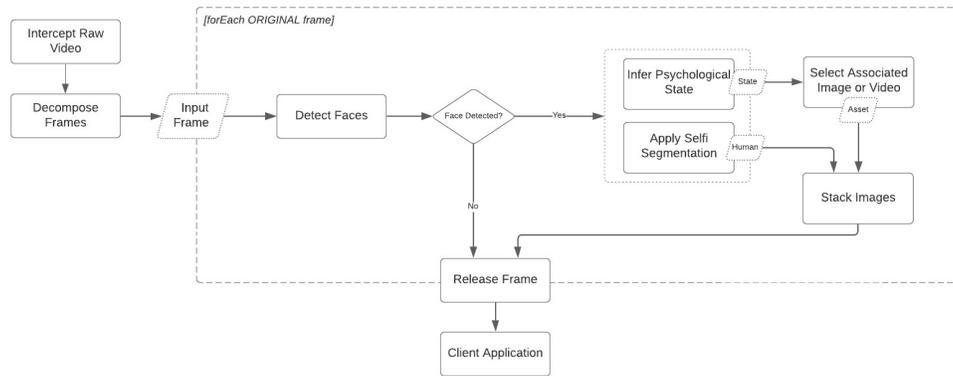
1

Figure 1: *Proposed solution execution flow*

facial expression and emotion recognition.

ii. Changing Background

Based on the detected emotion, the system will select, from the available options, a representative image or a short video. This media asset will then be placed as the user background for a given amount of time.

Similar to sentiment detection, the scope of this disclosure aims to use the produced outcome of existent methods and techniques. As one example, state-of-the-art body segmentation models can be used to detach the user from its current background images.

Once the region of interest (segmented human) is extracted, computer vision techniques can be applied to place the prominent human (segmented area) on top of the selected media asset. replacing the background.

iii. Background Configurations

As part of the solution, a set of curated backgrounds or short videos can be made available. These backgrounds will be offered as the default options and will be associated the the different emotions (or sentiments). To enable a more personalized experience, such backgrounds and short videos can be replaced with options that are specific to the user's needs (custom media assets).

Along with the actual images or short videos to be used, the user can also configure the amount of time that such temporary images will be displayed as background images.

Ideally, the temporary image will be added to the user's background to emphasize the current emotion and as such will be removed (fade away) after a while.

III. Discussion

Allowing the usage of customized images or short videos as virtual backgrounds is now a commodity for the major virtual meeting providers and services. Zoom and MS Teams, for example, allow users to manually select an image (or short video) to be used in replacement of the actual environment background.

What is unique about the approach described herein, is the automatic adjustment of the media asset to be used, based on the user's current emotional state. Different from existent solutions, the proposed method will have a curated list of backgrounds mapping to different sentiments. During the virtual event, backgrounds are dynamically replaced based on the captured facial expression or sentiment.

The main advantage of this method is allowing all attendees to better understand sentiments and emotions while attending virtual events. As one example, a *colorful and fun* background image that is temporarily displayed can help signalize that a particular meeting attendee is happy.

Offering this as a way to express and highlight sentiments and emotions, the overall meeting experience can be improved leading to a more productive environment.

2

Disclosed by Rafael Zotto,
HP Inc.

3