Input Video and Recording Environment Recommendations

FacePulseRate: A Suite of Features to Facilitate the Use of iPhys Toolbox

Douglas Magill

October 2020

Input Video Requirements and Recommendations

Video Format

Color videos. Grayscale videos are not suitable for the pulse-rate algorithms used by FacePulseRate.

Supported file extensions include …

- .mp4
- .avi

Other file extensions may be supported. Use Matlab function VideoReader.getFileFormats() to return a list of supported file extensions.

Note that videos with these file extensions may not be supported depending on the specific format (encoding) of the video.

Whether a format is supported depends upon two factors. The first factor is the Matlab release as FacePulseRate uses Matlab video readers. Format support from the Matlab video readers may change between releases. The second factor is the availability of codecs installed on a particular computer. FacePulseRate will verify whether a video is supported when the function is run. FacePulseRate uses the Matlab video readers to check the support of the video readers and to check whether the appropriate codecs are available on a particular computer.

Maximum Video Frame Size

The input videos can be up to 5000 pixels in either height or width.

Minimum Recommended Video Length

The duration of the input video must be at least 2 seconds. This requirement is necessary for minimal functioning of the pulse-rate algorithms. However, not that a longer duration may be necessary for accurate results; for example, Poh, McDuff, and Picard (2011) captured samples of 60 seconds in duration.

Consider using input videos of at least 3 minutes. The skin-detection algorithm and tailored skin-segmentation algorithm use skin-color samples from the face in the video to enhance their classifications. These algorithms require a minimum number of samples to be collected in order to be activated. For videos where face detections are not readily made, a longer length will be needed. These algorithms do not need to be enabled for proper operation, but their use may improve accuracy. A warning message will indicate when these algorithms were not activated due to an insufficient number of skin-color samples.

Maximum Video Length

There is no established maximum length, but the RAM demands for videos over 1.5 hours may cause some systems to become unstable.

Beyond the RAM used by Matlab, this function may use around 1 gibabyte of RAM for a 20-minute input video. The RAM use will gradually increase for longer videos. RAM use facilitates efficiency by reducing frame re-reading and duplicate processing. For longer videos, RAM used should be monitored to determine whether it is within desired levels. If operation becomes noticeably slower, this could indicate RAM use has exceeded RAM capacity and is being allocated to the hard drive. If RAM use is undesirable, consider breaking the video down to smaller segments. The function has been tested on a system with 8 gigabytes of RAM.

Display of Legend

If convenient, consider using frame dimensions of 1280 x 720 or 1920 x 1080 and recording in a landscape orientation. Doing so will allow a legend to appear on the output video. The legend is currently only implemented for these specifications. Note that most of the indications on the legend are also printed to the Command Window regardless of whether the legend is displayed.

Function Speed Influenced by Start Time

Faster read performance will occur when argument StartTime is less than 10% the length of the full video.

Time used in Output Video

The seek time of the output video may not match the seek time of the input video. This because the output video is produced using a fixed frame rate whereas the input video may use a variable frame rate. However, the timestamp annotations on the output video will always match the seek time of the input video. That is, to seek to a point in the output video based on a time in the input video, use the timestamp annotations rather than the seek controls of the output video. Note that all other output (variables and csv files) correspond exactly to the seek time of the input video.

Recording Environment

Although the environment cannot always be changed, skin detection and skin segmentation will be more successful in some recording environments than others.

Best recording environments:

- The face is not more than a few feet from the video camera.

- The colors of the background are distinct from the colors of the face. Skin tends to have a red hue -- this is the case across individuals with different skin colors -- so care might be taken to avoid backgrounds with colors similar to skin that also have a red hue.

- The brightness of the face is not too dark. Skin segmentation will sometimes segment out pixels that are close to black, so ensuring proper brightness in the area of the face can avoid this. Note that, assuming brightness of the face is adequate, the skin-segmentation algorithm is robust across individuals with different skin colors as no skin color would be expected to approach pure black. Cases of inadequate brightness are often the result of video cameras automatically reducing their brightness or exposure when a bright light source is present in an area of the frame other than the face. For example, lamps, windows, and solid white objects, including clothing, can sometimes lead to a video automatically reducing the brightness of the face. FacePulseRate uses a built-in operation to check for and compensate for low brightness of the face, but this operation is applied conservatively. Reliance on this operation is not recommended.

- There is even illumination of the face. For example, a shadow should not partially cover the face. Uneven illumination can reduce the accuracy of skin segmentation.

References

Poh, M. Z., McDuff, D. J., & Picard, R. W. (2011). Advancements in noncontact, multiparameter physiological measurements using a webcam. *IEEE Transactions on Biomedical Engineering, 58*(1), 7-11.