

Assignment 1: Basic image processing

Computer Vision (H02A5a)

February 20, 2019

The goal of this first assignment is to apply some basic image processing techniques to a one minute video. It's your job to first add special effects to modify video content and finally add some magic using an object of your choice. You will have to submit this video output and your code by **Monday 4 March midnight**.

1 Overview

You will be the director, scenario writer, producer and actor of this video: start by thinking about what will happen in your movie and make sure you have the object(s) ready for action. Next, you should shoot a one minute video with your webcam and store it on your hard drive. You are going to process this video frame by frame by building Python code (and using the OpenCV library) to implement the techniques described in Sect. 2. You will thus process your video offline and ignore any explicit temporal information to process a certain frame (for your final part, the part where the magic happens, you are free to experiment using temporal information).

1.1 The video scheme

You should split your video into 5s time intervals and apply a different technique in each interval to end up with a video according to the scheme in Fig. 1. More information will be given under Sect. 2.

1.2 The video scenario

Once you have applied a few basic processing techniques in the first intervals, you can be creative in the following intervals by *grabbing* an object of interest (in code: detect all pixels of this object) and then applying some magic to it. An example could be to detect and follow a coloured ball in the video. Another example is using a coloured towel to make parts that are covered by the towel invisible. You can of course invent your own *trick* but focus on trying to demonstrate how you are able to grab an object as accurately as possible and influence the video scene using this object.

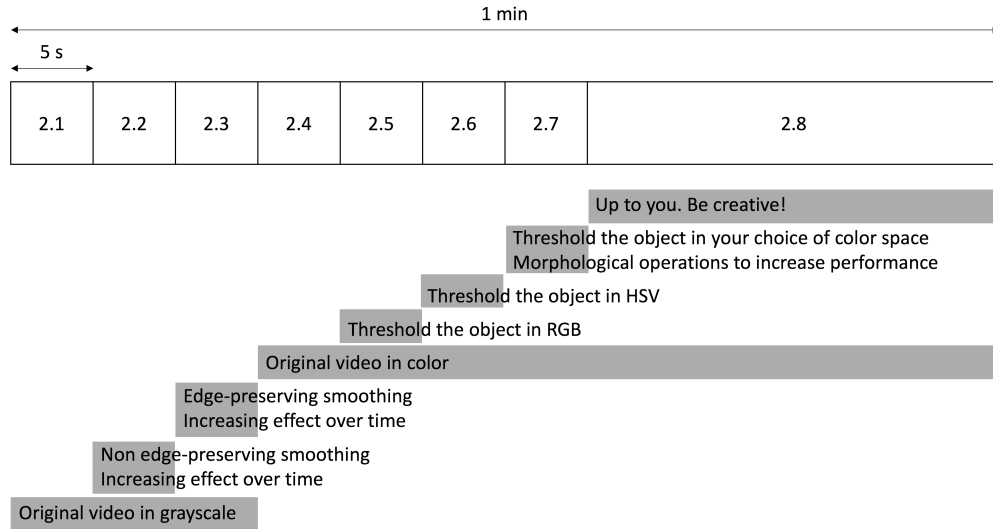


Figure 1: Your one minute video.

2 Processing

While most of the computation will be rather efficient, it can be a good idea to downsample your video right from the start. You will anyway have to make sure your video is smaller than 30 MB (after MPEG compression) before you upload it on Toledo. It's up to you.

2.1 0 - 5 s

Start by showing your video in grayscale as a baseline. This grayscale video will be processed in the next two 5s time intervals.

2.2 5 - 10 s

Smoothing or blurring is a simple image processing technique often used to reduce noise. In this 5s window, choose one simple filter (e.g. Gaussian) and increase the effect of the blurring by widening your filter kernel (and increasing the high-pass threshold accordingly).

2.3 10 - 15 s

You should have observed that simple smoothing filters remove noise but also the edges. In this time interval, change your smoothing filter to preserve the edges (e.g. bilateral filter). What happens if you apply this filter multiple times?

2.4 15 - 20 s

Next, just show your video as it is and establish the baseline for next time windows.

2.5 20 - 25 s

Grab your object in RGB color space. If you carefully choose your object (i.e. with a distinct color compared to the rest of the scene), this can be a simple thresholding operation. Can you find the optimal threshold without any "post"-processing? Consider visualizing the intensities present in your frames to understand how far this simple technique can bring you. In this window, your video should only show the grabbed object on a black/white background.

2.6 25 - 30 s

Do the same, but now try to grab your object in HSV color space. Again, if you carefully chose your object, your video should only show the *grabbed* object. Can you find the optimal threshold without any "post"-processing? Pay attention to the differences when working in RGB color space.

2.7 30 - 35 s

Choose a color space and try to improve your *grabbing* (e.g. fill holes, undetected edges) by using binary morphological operations. Your video is still a black/white background with the grabbed object. Put the improvements in a different color.

2.8 35 - 60 s

Now you are free to show off by adding extra magic to your scene (e.g. build an invisibility towel). Are you able to detect other object(s) of interests (e.g. your eyes)? Can you manipulate objects of interest (e.g. change color, move them to other places in your video, draw boxes around them, ...). Are you able to sharpen an object that was captured out-of-focus? What happens to your tricks if the scene changes (e.g. an object of similar color as your object of interest appears, scene lighting diminishes)?

3 Submission

Submit your code and the video as one single zip-file on Toledo. Make sure to compress the video under MPEG format. The video you upload must be smaller than 30 MB in size. Remember you can choose to downsample your video right from the start, or work on full resolution and downsample it just before compressing and uploading it on Toledo. **Videos larger than 30 MB and in a different format than MPEG will not be reviewed.**