Computer Vision Seminars Project Instructions and Links

Objective: The goal of the project is to apply several methods for each topic seen in class to 3 videos, and analyse/compare results. It is realised in groups of 2 or 3 students, and there is an optional (but recommended) intermediate submission to evaluate the progress of your project. We are not interested in the quality of the resulting videos, but rather their analysis.

Delivery: The optional (but recommended) intermediate submission should be delivered by 23rd May, and the final submission by 9th June. Projects should preferably be submitted to the Aula Global, but if some other format is preferred, such as a link to Google Drive, or there is some other problem with the submission of the project, please make sure you contact the Seminars teachers and inform them accordingly. Please make sure the submitted folder is well organised, with everything clearly named.

Topics: The contents of seminars 2-5: colourisation, video matting, video editing, and anomaly/fake detection.

Methods: For each topic, use algorithms corresponding to (at least) two methods presented or recommended in seminars.

Videos: Apply each method to the same 3 videos. These should be of different types (e.g. human with plain background, human with complex background and movement, non human video with movement). These videos can be by/of yourselves or taken from the Internet, and should last around 5 seconds each.

Expected contents: A report presenting the methods used and, if relevant, how they were adapted, as well as a detailed discussion of the results. We recommend roughly 5 pages excluding images and references. We also expect all input and output videos and any code that was useful for the realisation of the project, with everything adequately named and referenced in the report.

Methodology: We do not need detailed explanations of how the used methods work if they have been covered in class or recommended by professors, only the reference to the model used. However, if some editing of the video (e.g. reshaping, sampling, separating into frames, scribbles) or of the model was done, it should be briefly presented with the reasoning behind it (algorithm takes images as input, reduce computation time, etc). In general, this part should be quite brief.

Discussion: This should be the essence of the report. We expect a presentation of the interesting aspects of the results obtained, and for them to be compared/contrasted according to the type of video and the method applied. Errors, encountered challenges, possible improvements, comparisons between methods (weaknesses, advantages, quality of results...) can be included in this part. As a reminder, we are not interested in the quality of the resulting videos, but rather their analysis.

Topic 1: Colourisation

Scribble-based method:

Colorization using Optimization
Fast image video colorization using chrominance blending

Example-based method:

Image Colorization Using Similar Images
Transferring color to greyscale images

Learning-based method:

Time-Travel Rephotography
Image-to-Image Translation with Conditional Adversarial Networks
ChromaGAN: Adversarial picture colorization with semantic class distribution

Topic 2: Matting

Trimap Models: require input image/video + trimap:

User-guidance/Scribble based: A closed-form solution to natural image matting
Manual Trimap: ViTMatte: Boosting Image Matting with Pretrained Plain Vision Transformers

Segmentation Models: only require the input image/video, they use a segmentation model to obtain an initial mask to initialize the trimap to obtain later a refined alpha matte.

Matte Anything: Interactive Natural Image Matting with Segment Anything Models
MatAnyone: Stable Video Matting with Consistent Memory Propagation

Encoder-decoder models: only require the input image/video, it solves semantic segmentation and matting simultaneously.

Robust Video Matting (RVM)
Real-Time High-Resolution Background Matting