Computational Photography (CS-413)
Project Proposals

February 2025

IVRL lab - EPFL

Below is a list of project proposals for the CS 413 course, Spring semester 2025. Clarifications about each proposal can be obtained from the corresponding supervisor TA. The deliverables explain what is expected from you to submit by the end of the semester, aside from presentations/reports.

Table of Contents

1 Automatic and aesthetic image cropping

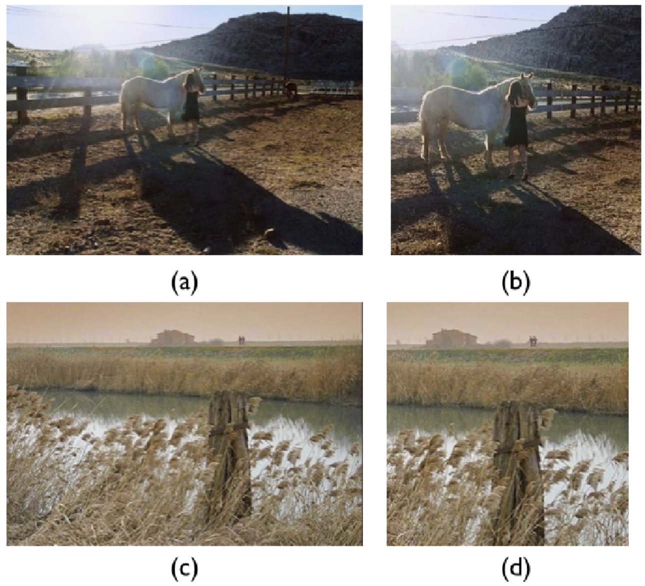


Fig. 1: Example of automatic image cropping (image source: [2]). Here, an algorithm evaluates the aesthetics of 1000 random square crops from an input image (a and c), and returns the crop with highest aesthetics score (b and d).

Description: This project focuses on creating an automatic image cropping algorithm that both preserves essential image content and optimizes the aesthetics. By leveraging existing models (colorfulness metric [1], CLIP image similarity [6], aesthetics assessment models [3][4][5]), the goal is to develop a model that can intelligently identify the most important regions of an image and crop it in a way that enhances visual aesthetics. A baseline example is shown in Figure 1.

Tasks (non exhaustive list):

- You will propose algorithms to automatically crop images based on aesthetics, content, and other factors you consider relevant (*e.g.,* size of the crop, similarity with original image, etc).
- Among other important aspects, your algorithm should ideally be interpretable, by having the possibility to detail various aesthetics scores for each crop, and highlighting which combination of these scores is optimized for.
- You will analyze your model's output (in terms of speed, quality, control, flexibility, interpretability, etc), qualitatively and quantitatively, and compare it with relevant prior works.

- Based on your analysis, you will improve your model.
- Implement an online demo (*e.g.*, a HuggingFace Space) of your approaches.

Prerequisites: Basics of Machine Learning. Basics of Python programming. Basics of image processing.

Learning objectives: Computational aesthetics, Machine Learning, Foundation models (CLIP), Image processing.

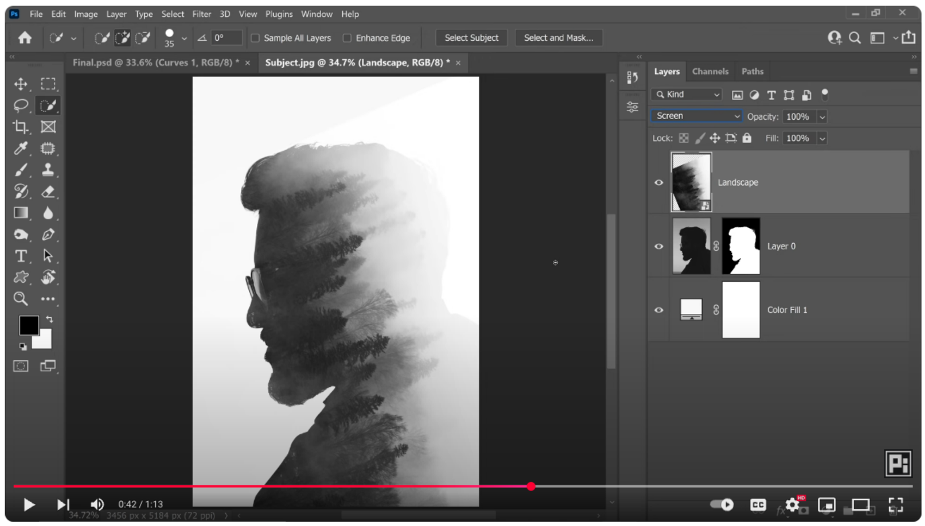
Deliverables: Code, well cleaned up and easily reproducible. Written report, explaining the literature and steps taken for the project.

Supervised by: Martin Nicolas Everaert (martin.everaert@epfl.ch)

References

1. David Hasler and Sabine Süsstrunk. Measuring colorfulness in natural images. In *Human vision and electronic imaging VIII*, volume 5007, pages 87–95. SPIE, 2003.
2. Bin Jin, Maria V Ortiz Segovia, and Sabine Süsstrunk. Image aesthetic predictors based on weighted cnns. In *2016 IEEE International Conference on Image Processing (ICIP)*, pages 2291–2295. Ieee, 2016.
3. Aditya Khosla, Atish Das Sarma, and Raffay Hamid. What makes an image popular? In *Proceedings of the 23rd international conference on World wide web*, pages 867–876, 2014.
4. Jana Machajdik and Allan Hanbury. Affective image classification using features inspired by psychology and art theory. In *Proceedings of the 18th ACM international conference on Multimedia*, pages 83–92, 2010.
5. Kuan-Chuan Peng, Tsuhan Chen, Amir Sadovnik, and Andrew C Gallagher. A mixed bag of emotions: Model, predict, and transfer emotion distributions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 860–868, 2015.
6. Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In *International conference on machine learning*, pages 8748–8763. PMLR, 2021.

2 Double exposure effect with diffusion models



Create a Double Exposure in 74 Seconds with Photoshop

Fig. 2: Manually creating a "Double Exposure" effect on PhotoShop. Image source: this YouTube video.

Description: It is possible to create an effect mimicking double exposure [3] by combining a silhouette image and a content image. However, this process is manual and requires time to select appropriate images and combine them. The goal of this project is to automate this process. The suggested direction for this project is to modify the inference pipeline of a diffusion model to jointly generate the silhouette and content images, using a similar approach as [1]. You are not limited to this particular effect, if you are interested in additional visual effects, feel free to suggest them.

Tasks (non exhaustive list):

- Perform a literature review on possible modifications of the inference pipeline of a text-to-image model. How did people manage to generate optical illusion, panoramas, or videos without retraining the model?
- Modify the inference pipeline of Stable Diffusion [2] (or another text-to-image diffusion model if appropriate) to generate effects like Double Exposure. Ideally, the user should be have the possibility to enter one textual prompt for the silhouette and one textual prompt for the image content, or provide one of these two image as input too.

- Evaluate qualitatively and quantitatively the images produced by your approach.
- Explore other visual effects that can be done using the same approach.
- Implement an online demo (*e.g.*, a HuggingFace Space) of your approaches.

Prerequisites: Basics of Machine Learning. Basics of Python programming. Basics of Diffusion models.

Learning objectives: Diffusion Models, Computational Aesthetics, Optical illusions, Multiple Exposure.

Deliverables: Code, well cleaned up and easily reproducible. Written Report, explaining the literature and steps taken for the project.

Supervised by: Martin Nicolas Everaert (martin.everaert@epfl.ch)

References

1. Daniel Geng, Inbum Park, and Andrew Owens. Visual anagrams: Generating multi-view optical illusions with diffusion models. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024. `https://arxiv.org/abs/2311.17919`.
2. Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10684–10695, 2022.
3. Wikipedia contributors. Multiple exposure, 2024. `https://en.wikipedia.org/wiki/Multiple_exposure`.