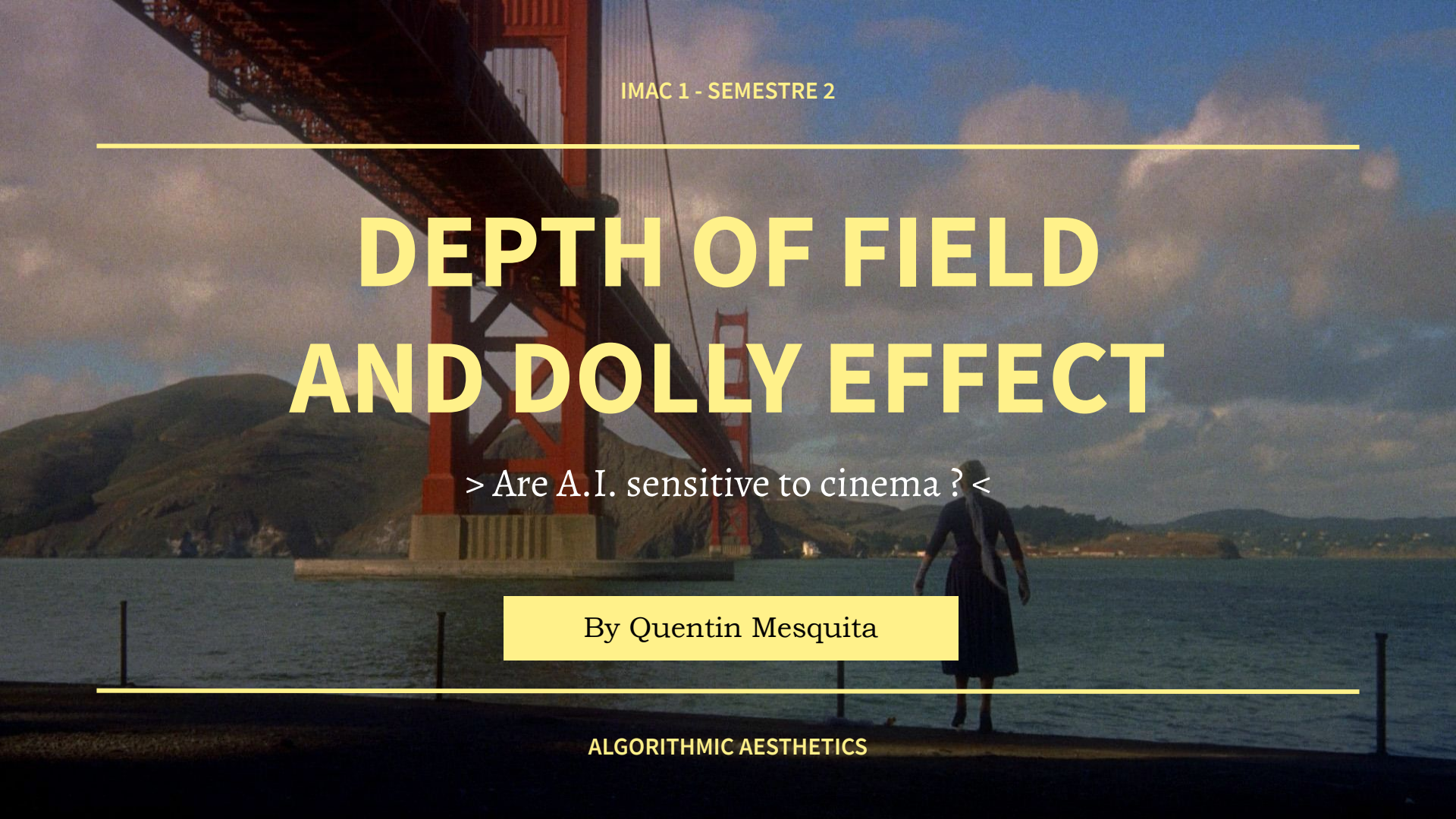

IMAC 1 - SEMESTRE 2

DEPTH OF FIELD AND DOLLY EFFECT

> Are A.I. sensitive to cinema ? <

By Quentin Mesquita

ALGORITHMIC AESTHETICS

PROJECT TOPIC

A GAN project

Among the three topics proposed, my attention was drawn to the GAN (Generative Adversarial Network). There are several reasons for this: I don't like developing websites, and animation and movies are things I enjoy.

Today, AIs can be found in any field, although their usefulness and effectiveness vary greatly. But there is one area in which artificial intelligence has greatly improved over the last few years: **image analysis and processing.**

For this project, I used an AI that estimates the depth of an image: MiDaS. From an existing photo, MiDaS creates a new monochromatic image of the first one. The different shades of grey then indicate the proximity of the objects to the camera.

I GOT AN IDEA

The aim of this project is to create an animation using images produced by an artificial intelligence. **However, the files returned by MiDaS are difficult to use in this sense**: at best, it is possible to give a new aspect to an existing animation, but that' all.

I then had a question: **is MiDaS capable of capturing cinematic settings?**

In cinema, it is common to vary the scale of a shot using a zoom, a dolly, or any other camera movement. However, by varying the scale of the shot, we also vary the depth of field! So would MiDaS be able to perceive this change in focus?

The aim of this project is therefore to test the ability on MiDaS to perceive the variation of depth of field in motion.

The art of directing is what determines whether or not a film will touch its viewer. Thanks to this we are able to feel the cinema; to feel it.

So in addition to simply testing the capabilities of an AI, **this project will also test whether, in some way, an artificial intelligence is sensitive to what makes cinema !**

DOLLY EFFECT OR VERTIGO ZOOM

The Dolly Effect or the Vertigo shot or the Dolly Zoom is an iconic camera movement. Popularised by director Hitchcock in the 1958 film Vertigo, it has now become a cliché due to its overabundance in productions of the 2000s.

It is characterised as follows: the camera moves quickly towards a character, while the background scenery moves away even faster. This paradox often breaks with the classic staging of the film from which it originates, so that it is very easily recognisable.

This technique therefore has an enormous effect on the depth of field. To achieve this, you have to get closer to the central character while gradually zooming out. The character is then completely torn away from the background of the image, which allows us to convey an intense feeling of unease.

This effect can be found in iconic scenes from Jaws and The Lion King for example.

To test MiDaS, I chose the Vertigo zoom of Poltergeist, released in 1982, because it is long enough to be close to 30 seconds as requested in the subject.

LET’S CODE IT

To realise the Poltergeist scene, we first had to cut it frame by frame. Knowing that the extract is in 30 frames per second and lasts more than 20 seconds, I found myself with more than 600 different images to process: which was far too many! I therefore proceeded as in a classic animation, selecting only one image out of two. My short animation is finally composed of 259 different frames, which is still a lot but it can be processed.

Then it was simply a matter of calling MiDaS 259 times through a simple loop in the code. I did have several problems with the different image formats, but fortunately it worked out in the end.

So, this part was not particularly difficult, but it took a very long time. **All the new images then had to be reassembled together to form the Depth animation.**

I was afraid that the final result would not look good: the brightness of the scene being relatively low, the AI could have had trouble processing the information. But once again, everything went well, and I finally got the answer to my very first question ...

CONCLUSION AND RESULTS

The comparison between the original clip and the new animation is clear: MiDaS does not perceive the camera effect. Where the background clearly moves away from the woman in the original clip, the shades of grey hardly change in the depth animation. There is only a slight movement, due to the lack of camera stability, but that is all.

This result was predictable, when you think about it: MiDaS only processes the frames one by one, it averages the different grey values of the final image within the frame: what will be perceived as “farthest away” will always be black, even if "farthest away" has changed between two different frames.

That’s the reason why the door is always in the shadow and the character don’t seems to be moving.

However, making this experiment was fun and interesting. I’m pretty sure that AI dedicated to motion (and not static pictures) can quite easily detect Dolly Zoom and other camera movements by calculating the depth of field. MiDaS may not have prove it, but it would be really surprising if the final result was “no”.

Link to the animation (without sound thanks to Youtube) :

https://youtu.be/TWrifRq8k00

Link to Github (Source) :
https://github.com/MesQuentin/Esthetique_Algorithmique/tree/AI_Creative_Project

Link to MiDaS :
https://app.runwayml.com/models/anastasis/MiDaS