Mark Felea-Motet (V00904671)
Bob Liu (V00937824)
Levi Bare (V00965748)

Video Composition Techniques
CSC 461 Final Report

Introduction

With the current swell of evolving digital media, video composition exists as an important component of visual storytelling now more than ever. In industries such as film and advertising, many challenges are presented with the creation of a piece of video such as having a high bar of difficulty to work with in things such as VFX, and also challenges with hardware and rendering enormous amounts of video footage. Leading companies in the world of video composition such as Adobe and Blackmagic Design have introduced products to aid digital designers, editors, and others within this space to give them more creativity and control when it comes to digital video manipulation. Within this report, we will introduce video composition topics involving datamoshing, colour correction and grading, frame interpolation, and AI inpainting/outpainting. With each of these topics we will explain how they operate from a low-level multimedia perspective, and explain their uses in expanding the creativity in the ever-changing space of digital video media.

Datamoshing

Overview

Datamoshing is an artistic video editing effect that produces glitch-like results. The technique involves manipulating the I-frames and P-frames of a compressed video to achieve the desired effect.There are two major subtypes of datamoshing: one were you delete certain I-frames to apply the predicted motion data of the video to the "wrong" frame (I-frame deletion), and one where you duplicate P-frames to create a swirling or sliding glitch effect (P-frame duplication).

Most datamoshing video edits are achieved using an old, buggy version of Avidemux. Avidemux is an outdated video editing software that had an issue in older versions that let the user manipulate these frames. A caveat is that in the versions that allow for datamoshing, crashing is prevalent which makes the process tedious and prone to progress loss. Glitch aesthetic was popularized in the 2000s, with a noteworthy mainstream example being in Kanye West's *Welcome To Heartbreak* music video [1]

Explanation

Intra-Coded Frames

Intra-coded frames (I-frames) are keyframes that contain complete image data. They serve as reference points for the frames that follow.

Predictive Frames

Predictive frames (P-frames) are keyframes that store only the changes from the previous frame. P-frames rely on the data from I-frames to make sense.

I-Frame Deletion

I-frame deletion involves removing certain I-frames from the video. When an I-frame is deleted, the video decoder tries to construct the image using only P-frames. Since P-frames are designed to work with specific I-frames, their misinterpretation leads to the application of incorrect motion or transformation data. The result is a visual distortion where elements from previous scenes bleed into the current scene, creating a dreamlike, surreal quality. This technique can make stationary objects appear to move or transform in weird ways. **Figure 1** is an example of I-frame deletion.

P-Frame Duplication

P-frame duplication involves duplicating P-frames within the video. By repeating these frames, the small motion changes are exaggerated. This can create a swirling or sliding visual effect, as if the image is melting or shifting continuously. It emphasizes motion and transformation, often creating a hypnotic, fluid motion. If the frames of the audio are also duplicated, the result sounds and looks severely corrupted. **Figure 2** is an example of P-frame duplication.

Figure 1. I-frame deletion example [1]

Figure 2. P-frame duplication example [2]

Demo

For the project, we implemented an I-frame deletion program with a UI. The program uses Python for the code, FFMPEG for video manipulation, and Gradio for the UI. It is based on a command line implementation by */u/justhadto* on Reddit [3].

The first step to use the program is to select the video you wish to apply the effect to as shown in **Figure 3**. After this, clicking the *Extract I-Frames* button will call the *datamosh_prep()* function. This function converts the video to AVI, which is a video format that uses the Group of Pictures (GOP) compression technique necessary for datamoshing. After the conversion, the *extract_frames()* function is called, which uses FFMPEG to extract the I-frames from the video.
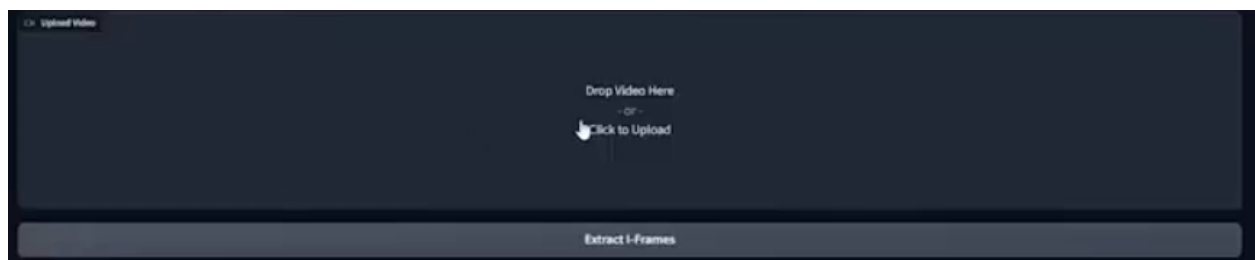

Figure 3. Drag and drop video upload with extract I-frames button

Once the I-frame extraction is complete, the I-frames show up in the Gradio gallery viewer as shown in **Figure 4**. The user can then select which I-frame they wish to remove and click the *Delete I-Frame* button.
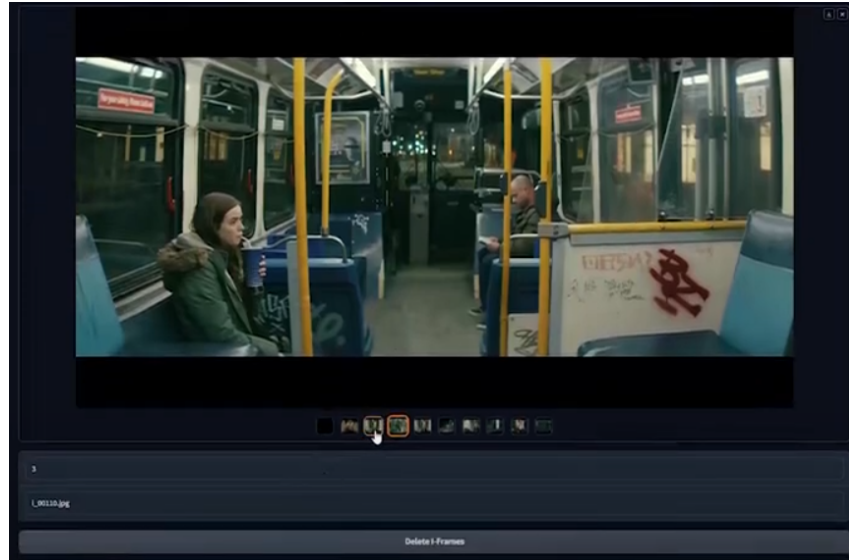
Figure 4. I-frame viewer with delete I-frame button

The final step is to click the *Datamosh* button as shown in **Figure 5** which merges all of the raw frame data back into an AVI video except without the deleted I-frames. This video is unplayable by most video players, so as a final extra step the intermediary AVI video is then converted to an MP4. The final video is then displayed and can be saved to the user's computer.
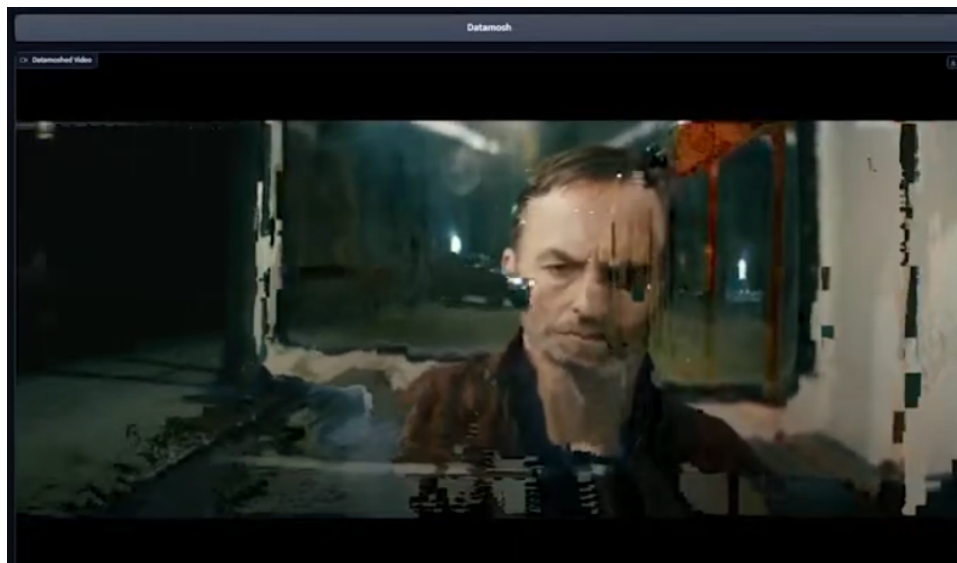

Figure 5. Datamosh button with result video player

Colour Correction

<u>Overview</u>

The science of colour comes from how computers deal with representing the spectrum of visible light. Digital colours can be stored and represented as 3 integer values that represent red, green, and blue. A 16-bit image can represent 281 trillion colours using 16-bits for each R, G and B channel, which will seemingly cover every colour visible to the naked eye, but not truly every single possible colour as the eye can only distinguish an estimated 10 million colours. When manipulating digital images, vector and matrix operations are a primary technique that compositing programs will make use of for computationally changing the colour of an image.

<u>Colour Correction</u>

Colour correction as a concept and technique has been around since before the widespread use of digital cameras. Though, within this section we will focus on the context of digital video editing when it comes to these techniques. colour correction at its basic is the process of making mathematical changes and adjustments to an image to allow accurate representation of how the scene looked at the time of filming. When doing so, editors and colourists often adjust parameters such as exposure, white balance, shadows, and highlights (among others) to achieve a generally neutral image before being passed on to the grading process. This correction is applied in order to allow images from various cameras, lenses and shooting conditions to match for better continuity. Colour correction is often one of the first steps in video editing as it allows for more flexibility when it comes to things like visual effects and CGI.

Figure 6. Original image on the left, colour corrected image on the right

<u>Colour Grading</u>

While "under the hood" colour grading is a somewhat identical process to colour correction, it allows the editor more flexibility and creativity when it comes to video manipulation. Colour grading is an important step when a video will be given a certain look with the purpose of creating a more immersive visual story. To help speed up the editing process, look-up tables (LUTs) can be applied to videos in most editing applications as a one click "preset". These look-up tables are a pre-calculated string of numbers ranging from 1 to 3 dimensional that allows for a computer to quickly process thousands of image values to produce a resulting graded image. While LUTs are not necessary to use, they are just a tool to help with editing

efficiency. When colour grading manually, typical composition software applications offer a wide range of tools, often including the basics as mentioned with colour correction, as well as curves for precise adjustment of luminance, hue and saturation, colour wheels for highlights, shadows and midtones, to name a few.

Figure 7. Colour corrected image on left, final graded image on the right

Colour Profiles

High Dynamic Range (HDR): High Dynamic Range images offer a more comprehensive range of colours and contrasts compared to Standard Dynamic Range (SDR). This allows for darker blacks, brighter whites, and a wider spectrum of colours. HDR achieves this by increasing the bit

depth of images, often to 10 or 12 bits, allowing for more detailed colour gradations and an expanded colour gamut.In HDR, the increased dynamic range allows for more detail in both the brightest and darkest parts of an image that can typically get washed out. This feature is particularly beneficial in colour correction, as it enables colourists to retrieve and enhance details that would otherwise be lost in shadows or highlights. HDR's wider colour gamut enables more accurate colour representation, especially in areas of extreme brightness and darkness. This allows for more natural and nuanced colour adjustments in post-production. Colour grading HDR images can give more flexibility in colour grading, which can be pushed to extremes.

Figure 8. Example of HDR image, showing the preserved details in the shadows and highlights[9]

Log Colour Profiles: LOG video recording is a powerful tool in the realm of colour grading and correction, designed to maximize the dynamic range captured by a camera similar to HDR. Unlike standard video, which compresses the dynamic range to fit within the limitations of conventional display technology, log video captures footage with a flat, gray look that preserves details in both the highlights and shadows, providing a much wider dynamic range. The flat,

desaturated profile of log video serves as a neutral starting point, allowing colourists to change

a broader range of tones and colours without compromising image quality. This is particularly

crucial for scenes requiring detailed adjustments in colour or contrast, as log footage retains

more detail and offers more control over these aspects. When grading log footage, it's essential

to apply an appropriate colour space transformation or a look-up table to convert the flat log

image into a more visually appealing and realistic image. This step is the foundation of the

grading process, enabling colourists to creatively enhance the image while maintaining natural

skin tones, balanced exposure, and a harmonious colour palette.

Figure 9. Example of an ungraded log colour profile frame[10]

AI Assisted Tools

Many video compositing applications are starting to implement AI-assisted features with further

promises of time-saving and creative improvements on a typical editor's workflow. One

mainstream video editor by Adobe, called Premiere Pro, has implemented the following AI

colour tools in their recent update:

- Colour match - simple colour correction that can be applied from different images to
  ensure continuity of colours

- Face detection - allowing faces to be masked and have separate adjustments applied

- Auto colour - making full image adjustments with a click

Similarly, a wide range of other standalone tools are offered and will surely continue to be developed in the years to come.

Overview + Introduction

In the changing world of video production the quest for stunning content has always been a driving force behind technological advancements and creative breakthroughs. One crucial element that significantly impacts the quality of video content is frame rate. Frame rate refers to the frequency at which consecutive images or frames appear on a screen. It has played a role in filmmaking and broadcasting influencing how motion and realism are perceived by viewers. Traditionally cinema has adhered to the standard of 24 FPS, striking a balance between limitations and creating an aesthetic that audiences associate with films. However with the evolution of technologies, new possibilities for manipulating frame rates have emerged, opening up prospects for enhancing video quality and enriching viewer experiences.

Frame interpolation has completely transformed the representation and enhancement of motion, in video content. Frame interpolation involves generating frames between existing ones in a video sequence. It goes beyond simply filling gaps; it aims to understand and recreate motion in a way that adds smoothness reduces blurring and enhances quality. This technology finds applications in a range of areas including creating quality slow motion effects without the need for high speed cameras and converting content between different frame rates to meet international broadcasting standards. This report aims to delve deep into the intricacies of frame interpolation, offering a comprehensive exploration of its techniques, applications, and the challenges it presents.

Figure 10. Frame Generation Example [4]

Motion interpolation of a horse and rider galloping.
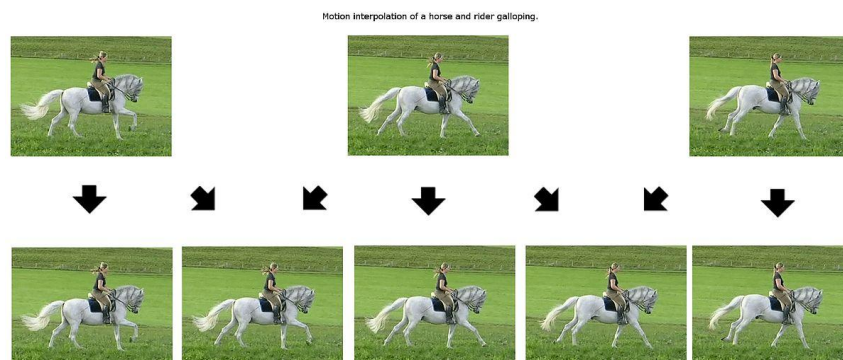
Techniques

Temporal Frame Blending: This is one of the most fundamental techniques in frame interpolation. This approach involves generating frames by blending pixels, from back to back frames. The concept is simple; for every pixel in the interpolated frame its position is determined by taking an average of its positions in the frames. This method is not used often as other methods are preferred for their better results.

Pros: The main advantage lies in its simplicity and speed of implementation.It requires minimal computational resources, making it an ideal choice for applications where processing power is limited or real-time rendering is necessary.

Cons: It can result in unrealistic motion blur, especially in scenes with fast or complex movements. Since the linear model fails to accurately capture the nuances of motion, it can lead to a less realistic and sometimes jarring viewing experience

Motion-Compensated Interpolation (MCI): Motion-Compensated Interpolation (MCI) is a significant leap in sophistication. MCI operates by estimating and compensating for the motion between frames. This involves tracking the movement of pixels/objects across frames and uses these motion vectors to create frames that better depict the movement. Video editing applications use this method

Pros: MCI is highly effective in producing more realistic motion, particularly in scenes with complex movements. It is an excellent tool for upscaling frame rates in videos, ensuring that the resulting motion appears natural and fluid.

Cons: It's more demanding compared to interpolation techniques as it necessitates sophisticated algorithms and processing capabilities. If motion estimation is imprecise, distortions within the interpolated frames can occur, diminishing video quality

Optical Flow-Based Interpolation: This method is well known for its ability to handle nonlinear motions. The process involves calculating the movement of objects within a scene referred to as flow and using this information to create frames. To achieve this the direction and speed of movement between frames are tracked, providing an understanding of how the scene evolves over time. Initially, the motion vector of each pixel or object is estimated by analyzing the changes in positions across frames. These motion vectors are then employed to predict the movement and speed of each pixel in frames, ensuring that the generated frames accurately represent the flow of motion while preserving the original trajectory. Video applications, like Adobe Premiere Pro, also use this method.

Pros:Handling Complex Motions: One of the most significant advantages of optical flow-based interpolation is its capability to manage complex and non-linear motions, making it highly effective for videos with a lot of movement and varying speeds.

High-Quality Results: This technique often produces superior results in terms of motion realism and fluidity, significantly improving the visual quality of the interpolated video.

Cons:Computational Demands: The complexity of calculating motion vectors for every pixel in a frame makes it resource-intensive, which can be a limitation for real-time applications or lower power devices. **Accuracy of Estimation:** If the optical flow is inaccurately estimated, it can lead to artifacts in the video, such as smearing or incorrect motion paths. This can detract from the overall viewing experience
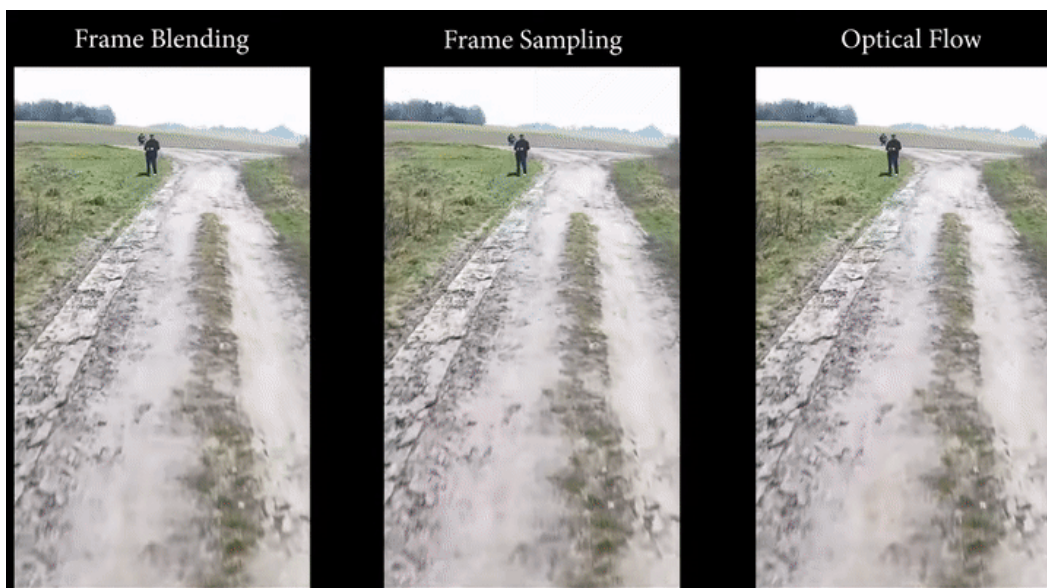

Figure 11. [5] Example of Various Techniques

1. **Deep Learning-Based Interpolation:**

Deep Learning-Based Interpolation represents a cutting-edge approach in frame interpolation, leveraging the advancements of artificial intelligence (AI) and machine learning, particularly through the use of convolutional neural networks (CNNs). This method goes beyond traditional interpolation techniques by using AI algorithms to predict and generate intermediate frames. At its core, this technique involves training neural networks to understand and replicate the patterns of motion in videos. By feeding large datasets of video sequences into these networks, the AI learns complex motion patterns and applies this knowledge to interpolate new frames. The neural networks effectively analyze the existing frames and predict what the in-between frames should look like, taking into account various factors such as speed, direction, and the nature of object movements. Applications include DLSS Frame Generation in games.

Pros: High Accuracy and Quality: Deep learning-based methods can achieve high levels of accuracy in frame interpolation, handling complex and varied motions with greater efficacy than many other traditional methods. **Adaptive to Different Scenarios:** This technique can be tailored and fine-tuned for specific types of video content, making it versatile for various applications, like video games.

Cons:Computational Power and Resources: Deep learning approaches require significant computational power and resources, making them less accessible for real-time applications or lower-end devices. **Training Data and Overfitting:** The effectiveness of this method heavily depends on the quality and quantity of training data. Additionally, there is a risk of overfitting to specific types of motion or video styles, which can limit its generalizability

<u>**Applications of Frame Interpolation**</u>

The use of frame interpolation spans across a range of applications all benefiting from its ability to improve video quality and create smoother motion. It has become a tool in various sectors of the video and entertainment industry.

Video Editing Software

Leading video editing software such as Adobe After Effects DaVinci Resolve and FFmpeg have incorporated frame interpolation tools significantly enhancing their capabilities. For example Adobe After Effects employs frame interpolation for effects like motion and time remapping allowing editors to create seamless transitions and realistic slow motion effects even from footage shot at standard frame rates. DaVinci Resolve offers motion estimation based frame interpolation, which is crucial for high end colour grading and post production workflows. FFmpeg, a software provides an accessible platform for frame interpolation techniques enabling a wider range of users to apply these methods in their projects.

Real-time Applications

In the realm of video gaming, technologies like NVIDIAs Deep Learning Super Sampling (DLSS) and AMDs FidelityFX Super Resolution (FSR) have transformed the gaming experience. These technologies can perform real time frame interpolation resulting in smoother gameplay on less powerful hardware. DLSS, for example, utilizes learning powered by AI to enhance the resolution of lower quality images in real time. This effectively increases frame rates without compromising the quality, which is especially advantageous in demanding gaming scenarios where a smooth experience relies on a high frame rate. However, one of the challenges for frame interpolation in video games is input latency as frame interpolation requires 2 frames to interpolate. This is tackled by technology like Nvidia's ultra low latency mode .

Specialized Uses

Frame interpolation also finds specialized applications in areas such as slow motion video creation, virtual and augmented reality (VR/AR), video restoration, and sports broadcasting. In slow motion video, interpolation allows for the creation of high-quality slow-motion effects without requiring high-speed cameras. In VR and AR applications, maintaining a high frame rate is crucial for preventing motion sickness; frame interpolation helps achieve this, ensuring a seamless and immersive experience. In video restoration, it can be used to repair damaged or missing frames, or to upgrade the frame rate of older films to modern or international standards. In sports broadcasting, interpolation can be used for a smoother experience, especially in high action sequences. Frame Interpolation can also be used to create smooth slow-motion replays for detailed analysis of players' movements or actions, enhancing both viewer experience and analytical capabilities.

Impact Analysis

The inclusion of frame interpolation in applications has had a significant impact on enhancing video quality and improving the overall viewer experience. In the realm of editing software it opens up creative avenues for filmmakers and content creators enabling them to tell more captivating stories and apply engaging post production effects. When it comes to real time applications like gaming frame interpolation plays a role in finding the right balance between performance and visual fidelity thereby making immersive gaming experiences more accessible to a wider audience. Moreover in areas such as VR/AR and sports broadcasting frame interpolation contributes to creating a heightened sense of realism and immersion in the content. Nonetheless achieving frame interpolation in these applications relies on striking a delicate equilibrium between interpolation quality and addressing challenges like computational requirements and potential visual distortions. As technology progresses further we can expect frame interpolation scope to expand more significantly continuously pushing the boundaries of video quality and captivating viewers, with enhanced engagement.

AI Inpainting and Outpainting

<u>Overview</u>

The recent advancements in generative AI has opened many creative opportunities in all facets of multimedia production. One particular area in generative AI that has seen immense improvement in recent months is AI image generation. Consumer-grade tools such as Dalle-2, Dalle-3, Stable Diffusion, Midjourney, and Photoshop's generative fill/expand make compositing techniques that were previously only possible for high budget productions accessible to small production teams and solo creators. Two especially noteworthy compositing techniques that stemmed from AI image generation are outpainting and inpainting. Out of these image generators, only Dalle-2, Stable Diffusion, and Photoshop's generative fill/expand support inpainting and outpainting.

<u>Explanation</u>

Outpainting

Outpainting is the use of generative AI to extend the boundaries of an image or video beyond its original frame, creating a seamless expansion of the scene that appears naturally consistent with the existing content. **Figure 6** shows an example of outpainting. Outpainting has many uses: modifying an image's aspect ratio without cropping, help center subjects without cropping, generate higher resolution images, and many more outlined on our project website.
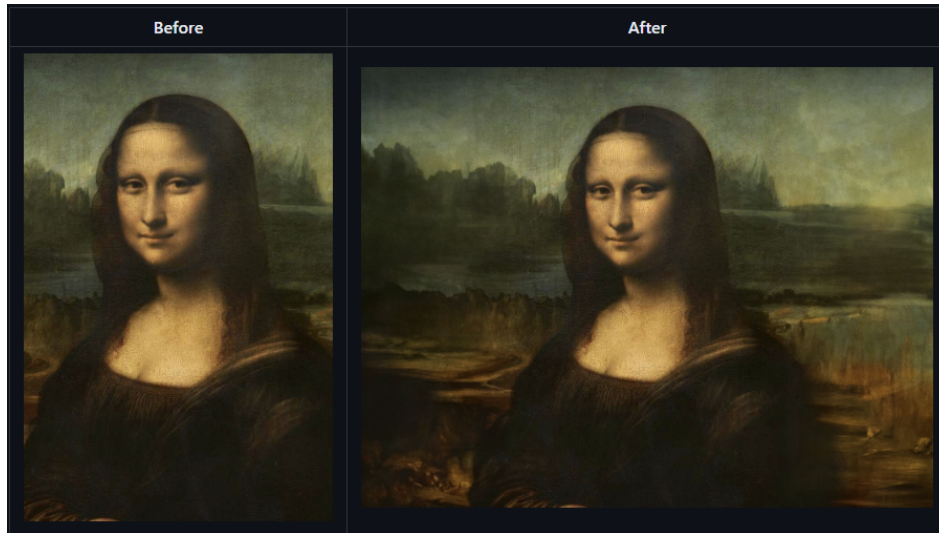
Figure 12. Outpainting example

Inpainting

Inpainting involves AI filling in parts of an image, such as removing objects, repairing scratches, or restoring corrupted areas to make them look as if they were never altered. Some examples of inpainting's uses include seamlessly compositing various images into one scene, restoring old photos by removing scratches, removing unwanted objects and many more outlined on our project website. **Figure 12** shows how inpainting can be used to make an image of the Mona Lisa wearing a hat.
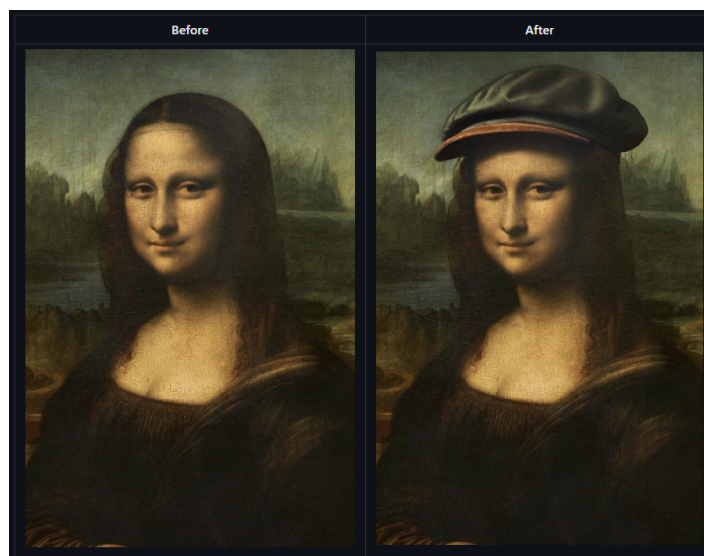

Figure 8. Inpainting example

<u>**Demo**</u>

For the project, we covered a real use case of using AI outpainting and inpainting together in video compositing to create a video effect that would otherwise not be possible. Our task was to make Black Widow in a scene from The Winter Soldier look like she disappeared after getting shot.

To achieve this video effect, the first step was to take a screenshot of a frame from the movie scene that we wished to remove Black Widow from. Next, we used inpainting to remove her from the still frame and then outpainting to expand the border of the frame. Finally, using the video editor Premiere Pro we used motion tracking to track our edited frame over the original camera motion from the source movie scene. Having the borders expanded by outpainting was crucial because the camera motion would otherwise cause the still frame we edited to move off screen. **Figure 13** shows the before frame and after frame with the outpainting and inpainting as well as the final result. The final result is a video which you can find on our project website.
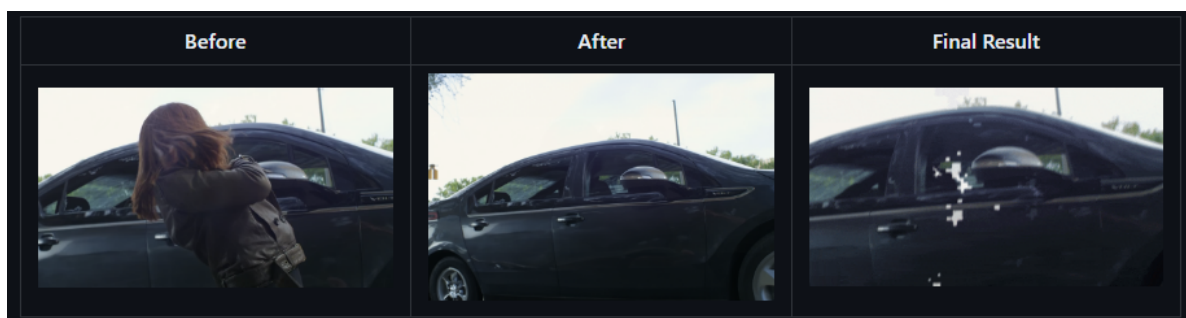

Figure 13. Outpainting and inpainting demo

Conclusion

In conclusion, our project highlights advanced techniques in modern video compositing.

Techniques like datamoshing, colour correction, frame interpolation, and AI

inpainting/outpainting all are highly advanced and fascinating methods video editors can use to

create high quality and engaging videos. Datamoshing lets the video editor apply a unique

glitch-like look. Colour correction is crucial for maintaining visual consistency and mood. Frame

interpolation enhances fluidity in motion especially in converting frame rates to match between

different sources. Lastly, AI inpainting/outpainting is a cutting edge technique with many uses in

image manipulation and VFX. Without these techniques, much of the video media we consume

today would not be possible.

Feedback Received

Unfortunately we did not receive any feedback from our classmates on our final video

presentation.

Individual Contributions

Mark Felea-Motet

Project contributions:
- Website setup
- Wrote project proposal
- Wrote week of october 23rd project update
- Datamoshing page
- Generative AI Inpainting/Outpainting page
- Final presentation slides format
- Splicing final video presentation
- datamosh_iframe.py implementation

Report contributions:
- Report formatting and outline
- Datamoshing section
- AI Inpainting and Outpainting section

- Conclusion section
- Individual references

Bob Liu

Project contributions:
- Contributed to project proposal
- Frame Interpolation Github Page
- Week of Nov. 6th + Nov 20th update
- Frame interpolation slides + presentation

Report contributions:
- Frame Interpolation Section
- References
- General Formatting

Levi Bare

Project contributions:

- Github Wiki Setup
- Colour Correction/Grading page and presentation slides
- Nov 6th and Nov 20th project update

Report contributions:
- Introduction section
- Colour Correction/Grading section
- Formatting
- References

References

---

[1]   Kanye West, "Welcome To Heartbreak," *YouTube*, [Online Video]. Available:
        https://youtu.be/wMH0e8kIZtE?si=N6IMtGPMEpiw3Kfg. [Accessed: Nov. 20,
        2023].
[2]   Shaun Ray Edits, "Datamoshing Tutorial Video - Shaun Ray," YouTube, [Online
        Video]. Available:https://www.youtube.com/watch?v=fw5QnRevqX4. [Accessed:
        Dec. 8, 2023].
[3]   /u/ihadto, "Datamoshing with FFmpeg (How-To in comments)," *Reddit* [Online].
        Available:
        https://www.reddit.com/r/datamoshing/comments/t46x3i/datamoshing_with_ffmp
        eg_howto_in_comments/. [Accessed: Dec. 8, 2023]
[4]   Wikimedia Foundation. (2021, January 12). *Motion interpolation (computer graphics)*.
        Wikipedia.
        https://en.wikipedia.org/wiki/Motion_Interpolation_(computer_graphics)

[5]   Venet, J. (2021). Frame blending vs frame sampling vs optical flow in Adobe Premiere Pro. YouTube. Jeremy Venet. Retrieved December 9, 2023, from https://www.youtube.com/watch?v=tenXF82Bv3U.

[6]   Color Grading 101 - Everything You Need to Know. YouTube, 2022. Available: https://youtu.be/pAh83khT1no?si=MmNOSUpeUdU0q4EB

[7]   Video color correction, https://www.adobe.com/products/premiere/color-correction.html (accessed Nov. 7, 2023).

[8]   "High Dynamic Range (HDR) Video," Amazon Web Services, https://aws.amazon.com/media/tech/what-high-dynamic-range-hdr-video/ (accessed Nov. 7, 2023).

[9]   K. Torres, "HDR vs SDR" https://www.winxdvd.com/convert-4k-video/hdr-vs-sdr.htm 2023.

[10]   K. Graham, "What is Log Footage?" https://filmpac.com/what-is-log-footage/ FilmPac.