

Dynamatte: A dynamic matting method to generate in scene video fusion

Richa Mishra
richmis@andrew.cmu.edu

Achleshwar Luthra
achlesh1@andrew.cmu.edu

Christopher Klammer
cklam@andrew.cmu.edu

March 1, 2023

1 Motivation

Imagine standing on a sandy beach, the wind whipping through your hair as you gaze out at the horizon. The sky is alive with an explosion of hues - pinks, oranges, yellows, and reds - that dance and meld together in a symphony of color. The sun, a giant ball of fire, seems to kiss the ocean as it sinks lower and lower, casting a warm glow across everything it touches.

As you take out your camera, you realize that this moment is too perfect not to capture. You begin to record, the lens capturing every detail - the frothy waves lapping at the shore, the gulls circling overhead, and the distant hum of laughter and conversation from the people around you.

But then, as you pan the camera to the horizon, you wish there was something even more to accentuate the colors in the sky. Maybe a rainbow kite, soaring high above the waves, its vibrant colors a stark contrast to the soft pinks and oranges of the sunset. And then, just beyond it, you imagine spotting a surfer riding a wave, his silhouette outlined against the fiery sky.

You think of the lost potential, the combination of the stunning sunset, the playful kite, and the daring surfer would have been nothing short of magical. As you watch the footage back, you can't help but feel a sense of awe at the beauty of the world around you while wildly dissatisfied with the limitations of current technology and you couldn't make your picturesque moment that much more vibrant.

In video editing, much care and attention must be dedicated towards creating seamless and temporally consistent video with synthetically inserted objects or elements. These efforts consume an obscene amount of time to execute and perfect in post production. However, what if that burden could be ameliorated with a method that could add dynamic objects with photorealistic qualities into a live scene?

Overall, we would like to stress upon the impact that this project can have in video editing. This work has the potential of making VFX much more streamlined as getting clean foreground and background mattes can be a painful experience especially when there are different scene complications such as dust, reflections, and shadows but our neural network will potentially be able to perform the same task in seconds.

2 Prior Work

Many works have have looked at the task of matte creation. Sengupta et al. [2] capture image and background pairs to estimate an alpha matte and foreground in order to place humans in scenes with novel backgrounds. This work primarily looks at using context switching to combine different cues and a self supervised adversarial loss to to take a blended image with the novel background, generating realistic images.

Lu et al. [1] propose Omnimatte, a method to associate pixels in the screen with masked objects in a given scene. This method allows complex effects such as shadows, reflections, smoke, and water ripples to appear in the Omnimatte. This method takes videos as input to train a 2D U-Net that processes the video frame by frame. Each object is processed through a different layer in the model with dense optical flow fields calculated to maintain temporal consistency in the RGBA output.

(Third potential paper: Layered Neural Atlases)

3 The Idea

We plan to take the result of Omnimatte and seamlessly take one to many omnimattes and seamlessly insert them into an existing video. By doing this, we plan to use existing videos and allow dynamic objects to be added to the videos. The main challenge of this work will be to seamlessly perform domain transfer and conform things like noise and background to the objects.

In Background Matting [2], the authors have worked on a similar problem of compositing the foreground mattes to a novel background. This method requires an additional photo of the background without the subject at the time of capture which limits us from applying this method to a random image where there is unavailability of subject-free background and it also demands for a certain amount of foresight during the process of data collection. Our project idea aims to overcome this limitation by taking inspiration from [1] that just needs a video and segmentation masks of objects of interest, and is capable of estimating omnimattes — an alpha matte and color image that includes the subject along with all its time-varying scene elements.

4 Experiements and Timeline

5 Baselines

References

- [1] Erika Lu et al. “Omnimatte: Associating Objects and Their Effects in Video”. In: *CVPR*. 2021.
- [2] Soumyadip Sengupta et al. “Background Matting: The World is Your Green Screen”. In: *Computer Vision and Pattern Recognition (CVPR)*. 2020.