Contents

1 Telescope

Kevin Poli, Phil Vitale, Connor O'Hara and Brendan Von Hofe

2 Intro

Telescope is a machine learning assisted toolkit for digital video compositors with applications in visual effects, matte painting and diverse use cases accross the video post production pipeline. Tools from existing compositing packages will interact with a novel ML core to assist or completely automate the rotoscoping process. Rotoscoping is the process of masking and segmenting poritons of an image accross multiple moving frames, any feature length movie will often consist of hundreds of rotoscoped shots with multiple tracked mattes per image, and this is the primary job of thousands of roto artists accross the world.

We hope to make this process, fast, intuitive, and accesible to alleviate the manual and time consuming process that makes up a huge chunk of the man hours required to produce even low budget features. We believe that machine learning is in the process of revolutionizing image processing, and that user driven toolkits rather than black box command line workflows will bring our intelligent core into the hands of the artists where they can thrive.

2.1 Demonstration

Rotoscoping is the process of frame by frame selecting and isolating a given feature (usually an object or person) in a video, such that you can produce a video clip of exclusively that selection on a transparent background

Lets walk through this step by step:

- First, our source image at frame 1, of Marceu the Mime

- Lets start by creating a selection just of the Mime's face and hand - these are the features that are actually being "rotoscoped" out

- This purple selection represents a 'mask' which are the points and curves that make up the boundary of what we are looking to isolate. Traditionally, artists will digitally paint this selection in a software of their choice, by hand.
- This selection or mask is different from a matte, which is another important piece of terminology. A matte is a single channel image; meaning rather than pixels having red,green,blue values, they only contain 1 value from 0-255 called 'alpha'. 'Alpha' will often be displayed in software as white. The Matte of this selection is an image where only the pixels corresponding to the selection are white, and all other pixels are black.

- this is so that, under the hood, all we need to do is pixel-wise 'multiply' the source image to the matte, meaning any pixels with a black 'zero value' in the matte will become transparent, and any pixels in the white '255 value' in the matte will remain.

- Here is the result of that multiply, an image containing only the pixels we selected before

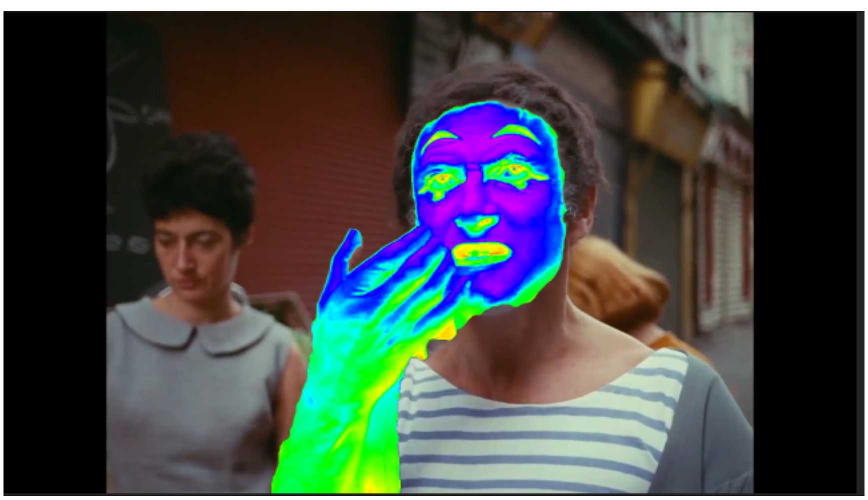
2.1.1 Frame By Frame

Much of the challenge and tedium of rotoscoping comes from repeating the above process for every frame, traditionally, artists will go frame by frame through the video and manually adjust their selections to match the feature they are isolating, here is the next frame of that video, with an adjusted selection for clarity

2.1.2 Use Cases

With our selection isolated, we can start to play with the image accordingly

By layering the source footage and our rotoscoped hand and face, we can apply an effect, like the 'colorama' effect to only the pixels we roto'd previously



1. Compositing The most popular use case for rotoscoping is Composit-

ing, which is the process of combining multiple images into one. Consider three layers to see how this is done.

Say we want this red square video clip to appear 'behind' the Mime's face and hand (note what appears black is acutally transparent)

We can grab our source clip and place the square image on top

Then grab our rotoscoped face and hand and place that on top

And place it on top for the desired effect

3 Technical Plan

3.1 Components

Telescope as a product will consist of two primary modules, the Telescope Core, which is a machine learning core assisted by traditional algorithmics that implements the novel functionality of Telescope, and an exchange plugin that allows existing professional compositing tools to interact with our proccesses. Telescope For Nuke is our chosen example exhange plugin, designed to demonstrate how the Telescope core can interact with existing artist workflows - but the separation of core and plugin is designed such that Telescope can be implemented into other software packages like Adobe After Effects or Blackmagic Design Fusion at a later date.

Category	What are we using?
Communication	
Email	Gmail
Web Conferencing	Facebook Video
Instant Messaging	GroupMe
Collaboration	
Document Collaboration	Google Drive
File Sharing/Data Tracking	GitHub
Plugin Development	
OS Supported	Windows, Mac OS, Linux
Host Application	Nuke
Development Language	C++
Machine Learning Development	
Development Language	Python
Packages	PyTorch

3.2 Algorithmics

The algorithmic core of our plugin will take images (frames of videos) as input and output segmentation masks (mattes) as output. The goal of the masks is to identify all the discrete objects in the image. It is class-agnostic and therefore does not need to determine what the objects are (e.g. cat or dog) but rather the fact that they are discrete. Our criteria for determining how well our model is accomplishing the task is the Intersection-over-Union metric (IoU). We have yet to determine what an acceptable IoU score is for industry applications. The model will be a convolutional neural network. Specifically, we will begin with the UNet model (`https://arxiv.org/abs/1505.04597`). Initially, our primary dataset to train the model with will be the Panoptic Detection COCO dataset, modified for a class-agnostic task. Further iterations of the model will take advantage of the additional information in EXR images to refine object mattes and the DAVIS video object segmentation dataset.

3.3 Dependency Model

[width=.9]./DGraph

4 Team

4.1 Roles

- Connor O'Hara: Image Processing (cohara1@stevens.edu)
- Kevin Poli: Application/ Artist Tools Developer (kpoli@stevens.edu)
- Philip Vitale: Application & Systems Developer (pvitale@stevens.edu)
- Brendan von Hofe: Machine Learning (bvonhofe@stevens.edu)

Advisors: Hong Man (hman@stevens.edu), Jeff Thompson (JThomps4@stevens.edu)

4.2 Delegation of Tasks

Connor O'Hara

4.2.1 Last Week

- Continue contact with Venture Center
  - we have a primary contact, but still waiting to be met with

4.2.2 This week

- Research Generative Ladder Networks

Kevin Poli

4.2.3 Last Week

- Acquire developer license for Nuke
  - Working on trial version until license is acquired

4.2.4 This week

- Follow along with Nuke developer tutorials, implement Nuke boilerplate

Philip Vitale

4.2.5 Last week

- Nuke API research
  - Noted and shared video tutorials and downloaded the manual

4.2.6 This week

- Follow along with Nuke developer tutorials, implement Nuke boilerplate

Brendan Von Hofe

4.2.7 Last Week

- Define image processing model

4.2.8 This Week

- Researching how to improve older partial solutions DeepMask/SharpMask
- `https://github.com/facebookresearch/deepmask`

Team

4.2.9 Updates

- Meet with Visual Arts department
  - Move forward taking them on as our client, Hong Man will remain Advisor
- Attend tech meetup for capital/business opportunity
  - Machine learning meetup October 4th