

What are Rotoscoping & Image Matting ?

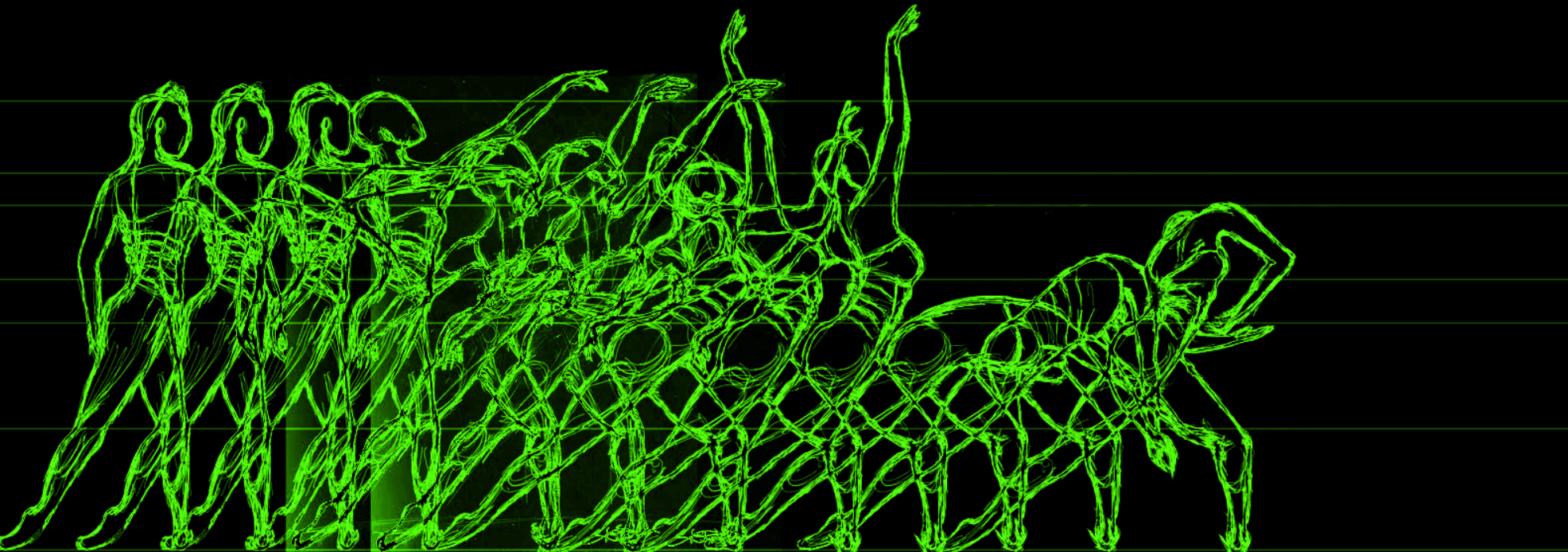
Rotoscoping is the process of frame by frame selecting and isolating a given feature (usually an object or person) in a video, such that you can produce a video clip of exclusively that selection on a transparent background. Any feature length movie will often consist of hundreds of rotoscoped shots with multiple tracked mattes per image, and this is the primary job of thousands of roto artists accross the world.



many objects individually rotoscoped in a single frame of "Conan The Barbarian"

Typical roto tasks include, selecting and tracking objects like on-set wires and rigs to remove them from the movie, and isolating objects for *compositing* - which is how effects artists create 3d interactions by layering many flat, 2d clips. The Dinosaurs in *Jurassic Park* are all 2d clips rendered out by other graphics software. Talented artists painstakingly used rotoscoping on the actors and many objects in the scene, so the Dinosaurs could appear to pass behind and around them in the massively multi-layered clips

Rotoscoping can be an incredibly challenging and time-consuming process, as the artist must adjust their selection, or mask for every frame to match the new position and movement of the subject.



The dancer's silhouette moves drastically from frame to frame, a tricky job to paint

These selections are represented by computers in the form of a matte - a single channel image (black & white), that is white at the location of the desired selection, and black everywhere else



Just as these objects and people move across the screen, so can our matte - as an animated video clip. Herein lies the challenge of Telescope, a task uniquely suited to machine learning - take rough, quick selections that can be approximations of the desired object across the video, and feed them frame by frame into a model that zeroes in on the hard to reach edges, creating stunningly 'ripped from the video' images.

How We Tackled the Problem

Telescope Core

- Written in Python 3.7
- Uses the newly available, advanced machine learning library PyTorch
- Exports to a C language 'productionized' model that is fast and fully available in the C++ written plugin
- Trained once on a powerful GPU so its ready to use out of the box

Telescope for Nuke*

- Written in C and C++, native to the host application
- Takes advantage of the Nuke NDK, the framework for all plugins
- Uses the native timeline and row-by-row image processing to handle multi frame operations, and implement fast multi-threading by processing multiple frames at once
- Easily compiled for use on Microsoft Windows™, Apple's MacOS™ and Linux

Telescope

Machine Learning Toolkit for Motion Picture Visual Effects

A Project in Computer Science, Computer Engineering and Visual Art Technology by

- Kevin Poli
- Brendan Von Hofe
- Philip Vitale
- Connor O'Hara

With Help From Our Advisors

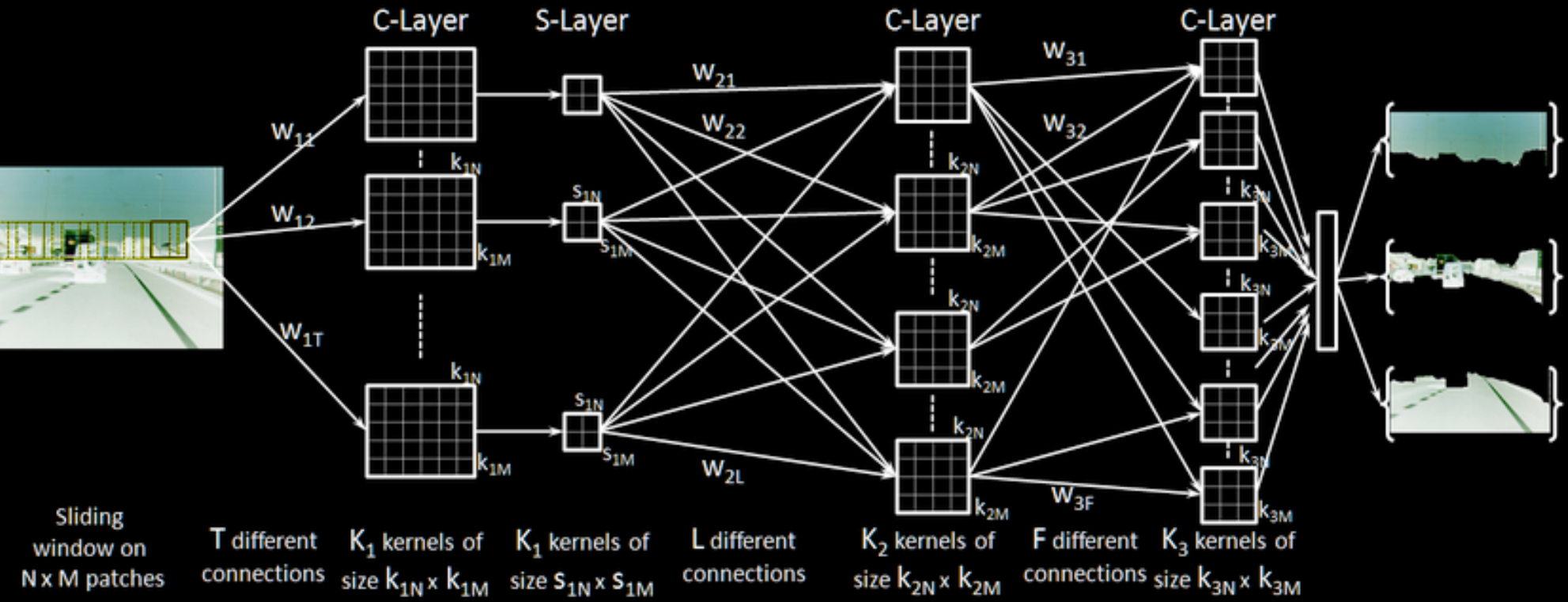
- Hong Man - Computer Engineering
- Jeff Thompson - Visual Art & Technology

What is Telescope?

- 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.
- The core of our rotoscoping program is the deep learning model that takes the image to be cropped and an associated TriMap, and outputs the cropped portion. The image to be cropped can be of anything the user wishes. This is a new and essential feature, as previous ML solutions use algorithms that must be told what they are isolating (what kind of object, is it a person?). All Telescope needs is a rough visual of what it should cut out, and our model takes it from there.

Convolutional Neural Networks

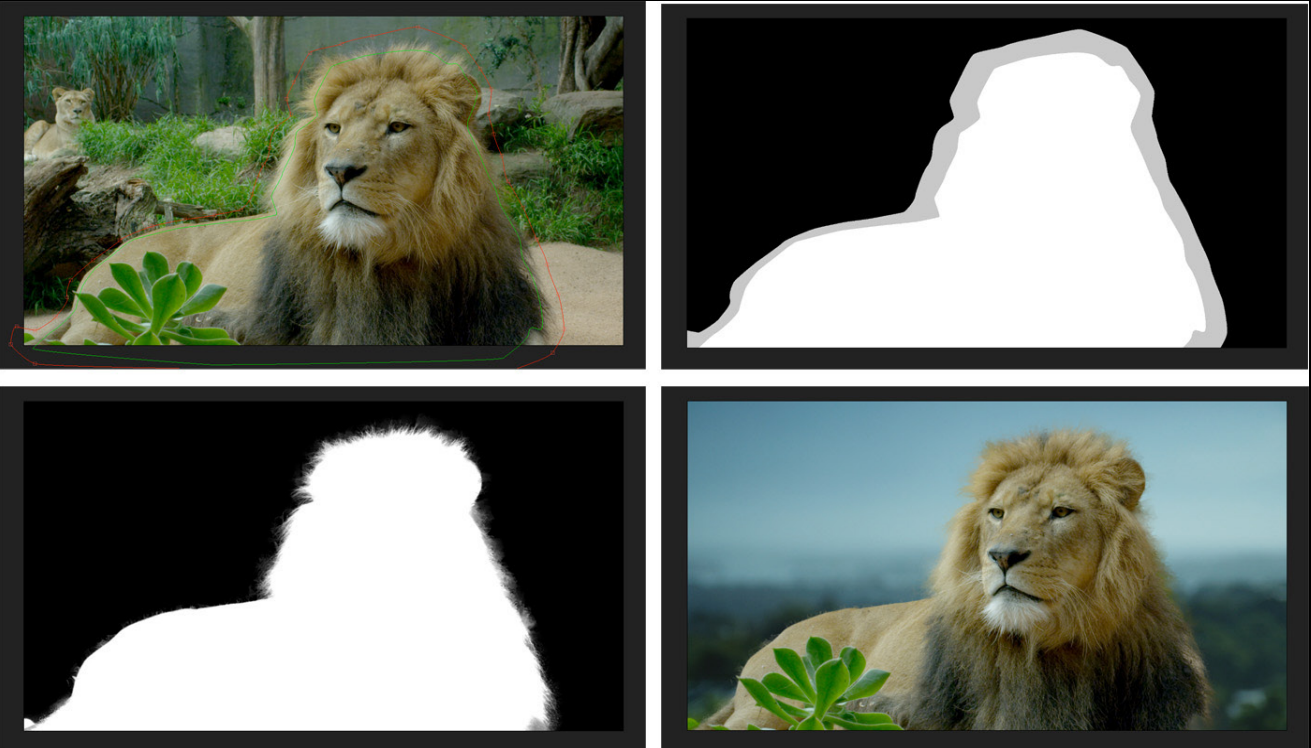
- Telescope, at its core is powered by a Convolutional Neural Network, a powerful machine learning algorithm, modeled after biological processes that is capable of analyzing patterns in visual imagery and learning complex relationships amongst its inputs.



The TriMap workflow

Telescope works by taking two inputs, an RGB source image and a single channel trimap which indicates definite foreground of the feature to be extracted (white), definite background (black), and the unknown region (gray)

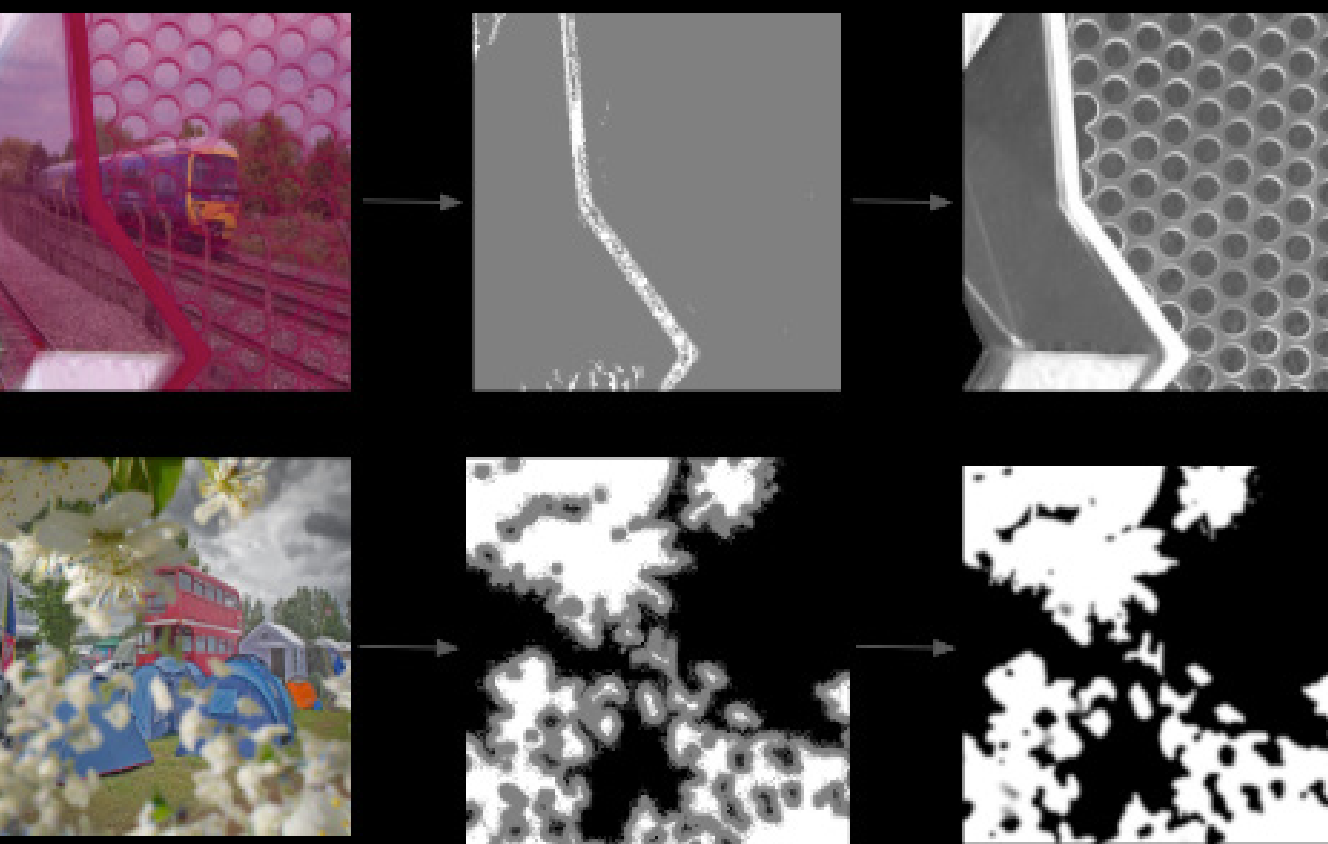
1. Source image with artist created mask
2. TriMap generated from that mask
3. Matte outputted by Machine Learning Model
4. Example composite of that lion on a different background



The Telescope program processes these two inputs and outputs a single channel alpha matte, which is used to extract the feature from the RGB source image

Training Computers to be Artists

Training Dataset



- TriMaps are automatically generated from the ground-truth alpha-mattes during training so that thousands can be processed each minute without human intervention.
- TriMaps are dilated and eroded randomly to reduce information leakage. The model will never see the same TriMap twice.

How Does the Network Learn?

- Gradient Descent: We specify Loss functions that are differentiable with respect to the neural network. We update the weights of the neural network by following the gradient of the Loss functions to find a minimum.
- Alpha Loss: This is the squared error between the pixels of the predicted alpha matte and the ground-truth alpha matte.
- Compositional Loss: We use the predicted and ground-truth alpha mattes to composite the subject on a different background and calculate the pixel-wise squared error between the two composites.



Telescope in the film industry

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.

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 processes.

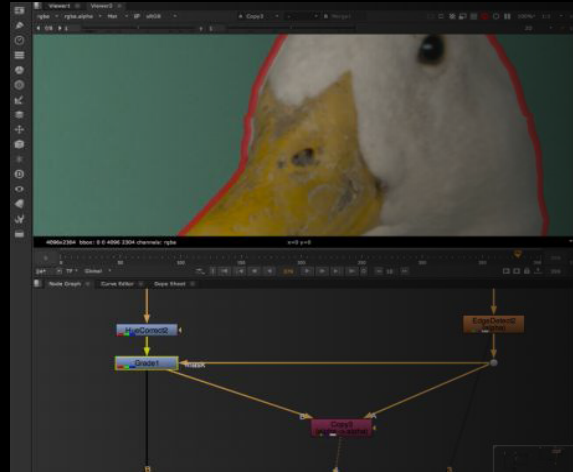
Telescope Core

- Free and Open Source
- Works as a standalone application in the command line only
- Contains the fundamental machine learning technology that allows Telescope to segment still images
- Does not provide an interface for creating TriMaps, or making selections in videos

Exchange Plug-ins

- Proprietary, licensed to production studios
- Allows industry standard video compositing software to interact with Telescope Core seamlessly
- Implements the innovative artist tools and interfaces for quickly creating TriMaps and adjusting roto selections across many frames in applications and with existing tools artists already know and use

Telescope for Nuke™



For our first exchange plugin to demonstrate the use of Telescope, we have picked Nuke™ the multi award winning industry standard video compositing tool from The Foundry. Nuke is in the pipeline of the majority of major motion pictures, and brought powerful tools to the artists of *Into The Spider-Verse*, *Sorry to Bother You*, *Get Out* and nearly the entire box office for close to a decade.

Nuke's NDK framework for developing plugins allows us to work in fast, robust C++, and take advantage of the systems and tools integrated into Nuke that Artists love

Seamless Integration, Zero Learning Curve

enhance any standard roto

Any mask or roto selection used in traditional rotoscoping methods can be easily piped into Telescope and instantly refined with deep learning

create your own workflows by combining built-in tools with Telescope

Planar Tracking is a popular feature in recent Nuke versions that allows the application to examine details in a selection and track how that selection moves in a video. using planar tracking to automatically move selections, and Telescope to automatically detect and refine selection edges makes even the most tricky rotoscopes available at speeds never achievable before



The Future of Telescope

The Separation of the Telescope core and its Exchange Plug-in system means that Telescope can be adapted into a variety of video compositing software with minimal re-implementation. Here are some applications that with a relatively small developer commitment, could be made to work with telescope.



Adobe After Effects



Blackmagic Design Fusion



Blender