# Automation of Rough Rotoscopy

Sam Hodge: Kognat Software

July 1, 2018

## Rotobot AI Based Masks: OpenFX Plugin for Nuke and more

Rotobot is the first to market AI product for compositing packages which use the plugin standard OpenFX, including industry standard Foundry Nuke. Rotobot has been tested by a variety of professionals around the world on both Linux and macOS platforms.

Rotobot is able to isolate instances of pixels that belong to semantic classes. This means that by analysing a photo with recurrent convolutional neural networks the pixels that belong to 81 categories (including 'person', 'car', 'sports ball' etc.) Up to 100 instances of each of these categories can be isolated at a time. This is useful to the visual effects industry, because as a typical use case, the pixels that belong to 'person' may need to be treated differently to the pixels that are in the background, thus categorised as 'not person'. For instance, an explosion may need to be put behind people. Normally this task would be completed by a process called rotoscopy, in which a skilled operator will create an animated outline of the person to create a mask to separate it into a different layer. This is a specialised role, as even in developing economies, artists are charged out at around one hundred United States Dollars per day. Even if an artist is capable of rotoscoping at a rate of one frame per minute, one second of footage will take 24 minutes to process, meaning that in one day, an artist can process 20 seconds of footage featuring one instance of a person. Thus, two people in shot for twenty seconds would have a minimum turn around time of one day, as two people in the footage means twice the amount of work, costing at the very least two hundred dollars and requiring two trained artists to work concurrently and have the cost of two interactive licenses of Silhouette or similar.

That is the cost associated with the actual work of rotoscopy. Artists requiring the *bash roto* to complete their tasks are unable to start the work until the masks are available for the footage. As a result, often work is completed quickly in a substandard way, decreasing the quality in order to release a task for downstream departments. This work is called *bash roto*, and is to be used in *slap composites*. These constructs are typically internal and are not shown outside of visual effects companies. The cost of a *bash roto* can be as much as forty percent of the total cost of the final roto. Often the *bash roto* cannot be reused to form the foundation for the finished masks, so in this way it is throw away work. The cost of bash roto for a typical visual effects shot can be two

to sixty hours of labour depending on the length and complexity of the shot and number of elements that need masks.

Until Rotobot can reach the goal of finished masks for final composites it can focus on automating internal *bash roto* to be used in internal *slap comps*. Consequently, Rotobot can automate as much as forty percent of the total rotoscopy cost of a shot. Automation of this process can have great saving for a visual effects company. Turn around time can be condensed from a 24 hour turn around, to a 5 minute calculation on a sufficiently scaled render farm with appropriate licenses, meaning that downstream artists can start their tasks earlier and there can be more parallel work flow where the rotoscopy artists are spending time that would have previously spent on *bash roto* finishing the final quality work, while the effects or lighting artist are working on the elements of the shot that will integrate with the rotoscopy mask layer.

The OpenFX plugin can be processing footage 24 hours a day without breaks, and the number of workers processing in parallel is only limited to the number of licenses and hardware resources to run the software. This is where Rotobot excels, as it can process a frame in as little as 12 seconds on standard laptop hardware and can isolate up 100 instances of 81 categories within a single calculation. No specialised GPU hardware is required, proved during test in which we ran the software successfully on a low end computer manufactured in 2007 with 2GB of RAM, the result being successful albeit slow. Rotobot will occupy as much CPU resource as it can to minimise the license usage, with no extra cost for multiple elements. Standard hardware for 64 bit OS: Linux or Mac is needed.

The license manager used by Rotobot is industry standard Reprise License Manager RLM which allows for floating licenses.

The quality of the rotoscopy acheived by Rotobot is not temporally stable, and it will flicker between frames. It is suggested to keep the footage that Rotobot produces internal to release time within the schedule. Kognat is working to increase the fidelity of the result to a much higher standard. However, given it can produce masks in a fraction of the time for a fraction of the cost, the reduction in quality allows it to be useful to downstream artist who are more concerned with a presence or absence of a mask rather than how precisely it matches an edge of the feature it is producing a mask to separate from surrounding pixels.

There may be other use cases for Rotobot in bidding rotoscopy work, where Rotobot can estimate on how many frames there are how many instances of a class. For example it could answer the question: "How many frames of how many people will need to be rotoscoped?" A text field puts out data of how many of each detected class, for example: "Horse: 5 Person: 7 Car: 2". Having a person calculate this by hand is error prone, and expensive.

This could also be used for *garbage mattes* for green/blue screen footage. This could be automatically calculated in advance for a batch of chroma key based shots.

Please consider the benefit AI based masks offer to your visual effects process.

Sam Hodge, Director, Kognat Pty Ltd: email: sam@kognat.com cell: +61417801006