Interactive Dynamic Video

Nishanth Reddy, Rishith Reddy

I. Introduction

Images and videos play an important role in our daily lives. But they are cannot be interacted with. We get a sense of objects only by interacting with them, by doing things like poke, move, rotate, etc. The goal of this project is to take such static images and normal videos and try to make them slightly interactive.

II. Scope

Our scope is limited to a rigid body motion with a single pivot.

III. Methods

A. Image Segmentation

The main goal of this module is to extract the object of interest from the image. In the static image we are assuming that the background is plain (since we have no other information unlike video). We used an edge filter (Sobel) and morphological operators to separate the object.

In the case of videos, the technique used is different, which is much better and independent of how the object and background is. Using multiple frames of the video we detect the pixels which have moved and get the object mask from this. This is done by subtracting two different frames and thresholding. Then we use morphological operations to get a good object mask.

B. Background Separation

From the previous image segmentation we have our object. But the background has a hole where the object was supposed to be. We want to create a background image without the object so that it can simply be combined with the transformed object.

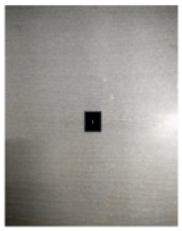
In the case of static images, since we do not have any other information other than the given image itself, we try to produce a background image from the original itself. This is done by removing the object using seam carving (object mask from previous module). And the image is again expanded to the original dimensions using seam carving again.

For videos, in multiple frames the object is at different places. Hence by using multiple frames we can recover the background information which is missing in a particular frame using another frame (i.e intersection).

C. Finding the Pivot

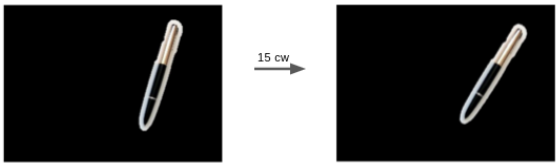
For static images, we take the pivot to the center of the object by default.

In the case of videos, while doing background separation using intersection of multiple frames, we notice that some part of the object is always covering the pivot. Hence, when we try to create a background image, we find that some of the portion is still empty. This portion is where the pivot lies. By defining a bounding box for this pivot hole and finding the center we can get a good approximation of the location of the pivot.



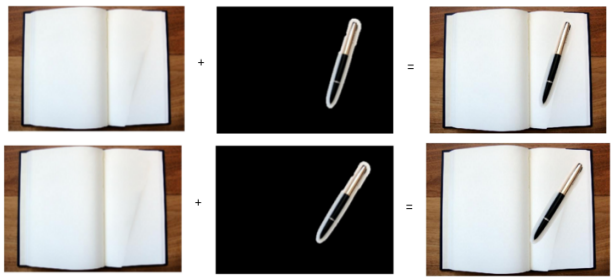
D. Affine Transform

Once we have a separate object and background image, we take the input from the user to find out how he wants to interact with the object, i.e, if he wants to rotate it, we find the angle of rotation. Then we use affine transform with reverse mapping to find the transformed object and also the transformed object mask.



E. Blending

While coming into this module what we have with us is a plain background without an object. The transformed object, and the transformed mask. Our final task is to blend these and give the output. We use gaussian and laplacian pyramids of the images up to a few levels and combine the background and object by using the mask at all these levels. Then we use the laplacian pyramid of the combined image and reconstruct our final output image. Doing this gives us a smooth blending without rough edges.



IV. Conclusion

With some more optimisations and better computing power, this can be done in realtime which give the user more perception of interaction. This can be used in animations to animate virtual objects interacting with real objects or static placement of objects, eg adjust position of paintings on the wall, or items on a desk, etc.