Development of an Augmented Reality System

Image Processing and Computer Vision – Project Work

Master Degree in Artificial Intelligence

DISI - University of Bologna, Bologna

Prof. Luigi Di Stefano

Introduction:

The object of this project is to develop an application that can create an Augmented Reality (AR) video. Given a video file, containing the shot of a book, we want to superimpose an augmented reality layer on top of the cover.

Inside the Data folder we can find:

Multiple View: target video on which the layer will be projected

Reference frame: first frame of the video used as reference

Object Mask: mask of the reference frame

Augmented layer: layer that we want to superimpose on the target video

Augmented layer mask: mask of the layer

We can use two different approaches to apply the AR layer, “Frame to reference” and “Frame to frame”. F2R employs the first frame of the video as reference to find the right transformation to be applied to the layer to each frame.

It is necessary to pre-process the data before proceeding with one of the two methods.

Pre-processing

The data pre-processing concerns the addition of the alfa channel, both in the reference frame and in the layer. The images are split in 3 channels RGB. Then, these 3 channels are reunited with an additional fourth channel, called alpha, created through the image mask.

A resize of the layer follows, which will take the dimensions of the reference frame.

At this point we create the feature detector, in particular SIFT, which finds the coordinates of the keypoints and the descriptors of every keypoints through these functions:

```
kp_reference = sift.detect(reference_rgba)
kp_reference, des_reference = sift.compute(reference_rgba, kp_reference)
```

After these operations we enter inside a while that allows us to read all the target video frames ad, based upon the chosen approach, F2R o F2F, we can superimpose the AR layer.

Frame to Reference approach

The first action inside the while is to read the next frame of the video. Next, we proceed to find the frame keypoints and descriptors with sift detector. Then we use a flann based matcher to find the correspondence between the keypoints of the reference and the current frame under examination. We test the correspondence as defined by Lowe in his SIFT paper. If these correspondences are more than 4 we proceed to apply the AR layer.

A this point we see how the reference is oriented in the frame through the homography H. We then calculate

```
dst = cv2.perspectiveTransform(pts, H)
```

used to find the points useful for changing the perspective. With these 4 points we can change perspective calculating another homography:

```
H = cv2.getPerspectiveTransform(pts_layer, dst)
```

Now we can use H to warp the augmented reality layer with:

```
warped = cv2.warpPerspective(layer_rgba_resized, H, (w_frame, h_frame))
```

At the end of this process, we append the new frame to the new video and we iterate again on the next frame.

Frame to Frame approach

This approach may seem more immediate than the previous one but it's actually trickier.

We find the keypoints and the descriptors of the current frame. Using the flann based matcher we find the best matches and we test the good correspondence as defined by Lowe.

At this point we find the homography between the previous frame and the current with:

```
H_current, _ = cv2.findHomography(src_pts, dst_pts, cv2.RANSAC, 5.0)
```

Then we accumulate the current transformation with the previous ones calculating the product between the current homography with the previous one. Then we proceed again with all the passage that we have made also in the F2R approach

As last step, before the ending of the while we compute the new perspective of the reference mask and we pre-process the current frame for the next iteration.

Conclusions

At the end of the while we proceed to save the new video with the super imposed layer respectively “Augmented_Multiple_View_F2R.avi” fot the Frame to Reference approach and “Augmented_Multiple_View_F2F.avi” for the Frame to Frame approach.

In the F2F approach we couldn’t superimpose the layer as precisely as with the F2R approach. This could be due to the fact that in the later iteration, the resulting transformation matrix can be affected from the cumulative error caused by the matrices product.

Project Members

Simona Scala 0001033460 – simona.scala6@studio.unibo.it

Sara Vorabbi 0001026226 – sara.vorabbi@studio.unibo.it