

Recovery and Reconstruction of Blackboard or Whiteboard images with occlusions

Vijayaraj Gopinath

vgopinat@stanford.edu

Computer Vision: From 3D reconstruction to recognition

CS231A

Stanford University

Abstract

We have all taken the pictures of blackboard for reference purpose, but the problem is most of the time those pictures comes with occlusion such as professor standing in between or other students. Even if you can take multiple image, there will one or other occlusion obstructing the blackboard information. By the time one student or professor has moved out another will often have moved in, making it difficult and time-consuming to get the perfect blackboard picture. Also we can't wait for the right moment, since the blackboard can be erased multiple time which makes this problem even more difficult. In this project we will make use of available multiple images and automatically recover the blackboard by removing all the occlusions.

1 Review of previous work

We can find many works related to generic scene reconstruction with occlusion using multiple images. Especially the work [1] from Microsoft research which tries to remove any occlusion from landmarks or monuments from multiple pictures. Generic scene construction has lot of assumptions and work around with that, but here in this work, we are only concerned about a specific scenario, reconstructing the blackboard images from occlusions such as person (professors or students) or any other objects.

2 Proposed solutions

In this work, the multiple pictures of the blackboard can be in any orientation and we assume the first picture dictate our required orientation for result purpose and all other pictures would be orientated towards our first picture. Also the proposed algorithm will find out whether the given pictures would able to recover the blackboard (or whiteboard) or not otherwise report an error. Since we have a specific scenario here, we would tap on the properties of this scenario (blackboard only) to segment the occlusion with a better techniques such as background subtraction and image morphology techniques which provides better results than without worrying about other more complex segmentation techniques.

3 Summary of the technical solution

To solve this problem, we have proposed the following solution to this.

- 1, Since we are dealing with multiple image, find homography for first image and to all other image and using the found homography warp all other images to the first image.
- 2, Detect, segment and label all the occlusions in all the images. Use image subtraction and other morphology

techniques to detect and segment. 3, Once labelled, we need to find out, which occlusions comes from which original image and decide whether the scene can be recovered or not. 4, Using the label and identification of occlusion from the original image, recover the scene by copying the pixels from non occluded region to the occluded region. 5, Finally blend the image to complete the reconstruction.

4.1 Finding Homography

It is important to have an accurate homography since we later use image subtraction to identify occlusions. It is common nowadays we have more whiteboards than blackboards which makes finding homography more challenging. With our experiments, we found that detecting rich features proven very difficult around whiteboards because of homographic surface nature of the scene. Also most of the found features are also around the occlusion since we often see a steep change in intensity levels near the occlusions.

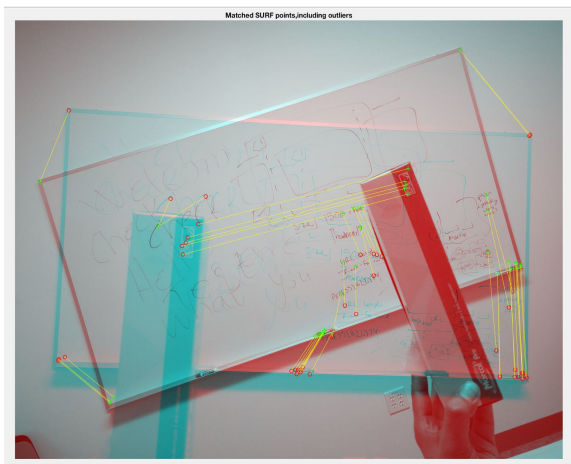


Figure 1 - with outlier features.

In figure 1, you can see that lot of features (more outliers than inliers) are found around occlusion since those regions have more intensity changes. Also, we will not get enough features on the corners always since the color of the whiteboards sometimes closely match the background of the scene itself. For all these reason, we need user feedback to select four corner of the blackboard for us. We use these four points to calculate our initial homography and we remove all the outliers using RANSAC or other such methods.

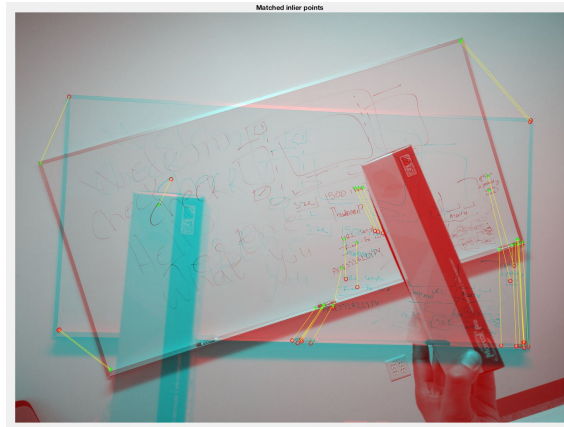


Figure 2 - Scene after the removing all outlier features.

Only using the inliers set which we found using the last step, we recalculate the homography which becomes our final homography. We need these two step approach to get our homography since we need accurate homography and also we assume that the user selection of four points will be not accurate enough for homography calculation but accurate enough to remove outliers. We have used SURF features here. Using the final homography we warp all the images to match our base first image.

4.2 Detection Segmenting and Labeling occlusion.

The first attempt we made in detecting the occlusion was face detection. Since most of the time, our occlusion will be a person occluding the blackboard it would be very interesting to take this approach. We can find the bounding box for face after detecting it and We can use face to body ratio to define a segment box to cover entire occlusion.

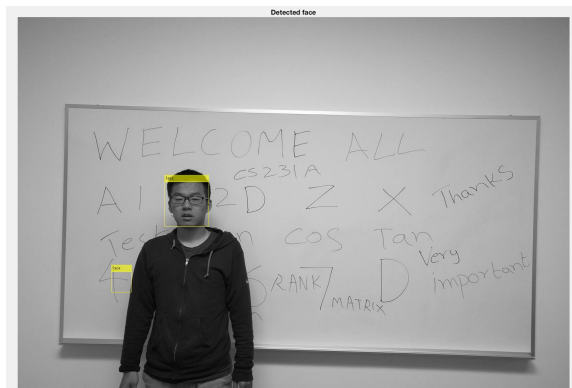


Figure 3 - Face detection

This idea has clearly many issues, We will get little noise in terms of detecting other small or non faces in the images also the occlusion need not to be a person here or even if the occlusions are assumed to be only person, finding bounding box would be wrong to do since person can take other forms. After experimenting with lot other methods, we found out that the morphological techniques would suit this scenario more perfectly.

Once we have accurate homography and good warping, we can use image subtraction between base image and all other image to remove everything except the occlusion. Though we found out

accurate homography, results after image subtraction will have noises around it which we will deal later. After image subtraction, we can use global image threshold using Otsu's method [2], to compute the global threshold level which is normalized to [0 1]. Using the found threshold level, we can convert the intensity image to a binary image.



Figure 4 - After image subtraction

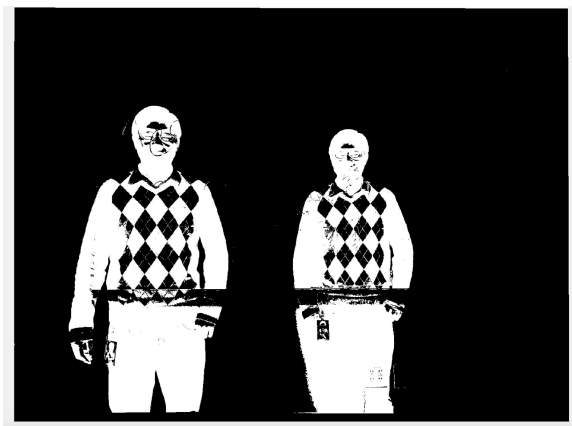


Figure 5 - Binary image with noise

To remove noises in the binary image we can use morphological techniques. Use structural element of 'Disk' with radius of 3 and erode the picture. The above step helps to remove the noise which we got after image subtraction and thresholding.

The binary image we found from the last step will have lot of discontinuity. In order to get the complete segment, we need to dilate the binary image to get supersize of the occlusion. We can use the structural element of 'Disk' with radius of 25 to get the supersize of the occlusion. Figure 6 shows the binary image after dilation.

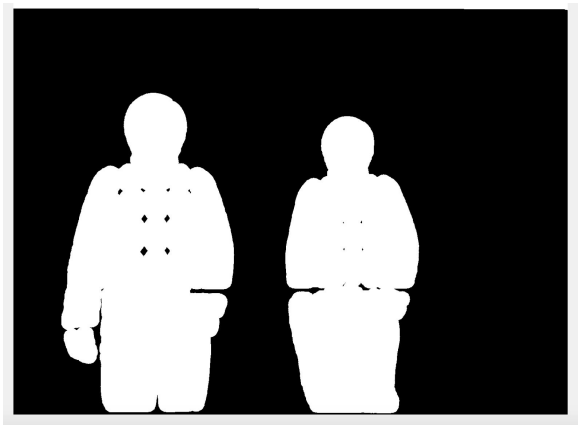


Figure 6 - Image dilation, Disk radius 25

Depends upon the intensity level of occlusion in comparison to the blackboard or whiteboard, the found occlusion potentially can have holes inside it. Which can be noticed from the figure 6. Here occlusion originally had intensity level similar to the background so after image subtraction and thresholding will have holes in it. We can find holes by finding the connected components in the binary image and finding the missing pixel inside it. Find such missing pixel and fill it. Figure 7 shows the binary image after filling up the hole. Once finding the supersize occlusion, erode the binary image with SE of radius of 22 to get back the original size of the occlusion. Figure 8 - shows the final binary image with the occlusion having same size.



Figure 7 - After filling the holes

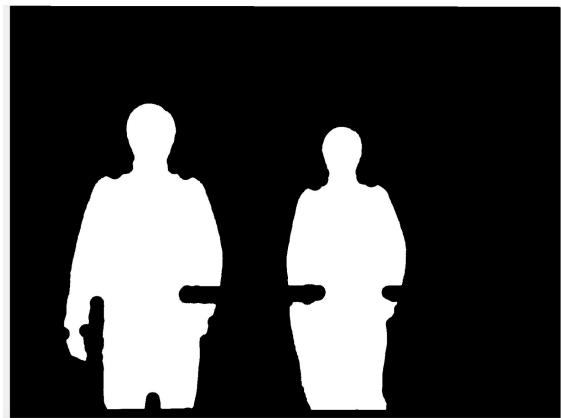


Figure 8 - Erode to get original size

After finding the binary image of the occlusions, label all the independent connected components to get the accurate boundary for every occlusions in the scene.



Figure 9 - Color map of the found label.

4.3 Mapping occlusions to the original image region.

After the segmenting and labeling, it is important to map each found occlusions to the original image. We assume that the object which is occluding has totally different average intensity value in comparison to the blackboard itself. We can build a model based on the intensity level around the occluded object and occluded region in the original images, which can be used to map the objects.



Figure 10 - Recovered image

4.4 Recovering and Blending

Once we identified, labeled and mapped the occlusions we can recover the scene by copying the pixels from non occluded portion of the images to the occluded portion of the images. Since, we copy pixels come from different image to recover the scene, the boundary of the recovered region in the final image will be visible, so we need to blend the image to complete our reconstruction. Figure 10 shows the final recovered image by removing all the occlusions.

5 Conclusion

Dealing with similar background in computer vision is very challenging, human eyes has evolved to detect this seamlessly.

We address this in the future works section. Blackboard reconstruction by removing occlusion is very interesting project and has an important application in the education domain. This can be pretty much a cool app which students can install in the smartphone and take multiple pictures and app can automatically recover the entire blackboard scene in a self contained manner.

6 Future works

1, As previously discussed sometimes we won't get enough features to find homography, and so need user selecting corners for us. We can work on sophisticated rectangle detection to avoid user input. 2, Fine tuning segmentation using other techniques such as Fast marching method, 3, For mapping occlusion to the original image, currently we build model based on intensity value, this can be improved to other sophisticated model using features. 4, After the scene has been reconstructed, we can try recognizing the text in it and create a doc.

REFERENCES

1. <http://research.microsoft.com/pubs/69386/peoplemover.pdf>
2. <http://ijarcet.org/wp-content/uploads/IJARCET-VOL-2-ISSUE-2-387-389.pdf>
3. <http://www.eiti.uottawa.ca/school/research/viva/papers/homographie.pdf>
4. http://www.cescg.org/CESCG-2016/papers/Jarabka-Generation_of_lecture_notes_as_images_from_recorded_whiteboard_and_blackboard_based_presentations.pdf
5. http://visual.cs.ucl.ac.uk/pubs/learningOcclusion/CVPR_2011_learning-occl.pdf