INDIAN INSTITUTE OF TECHNOLOGY KANPUR

UNDER GRADUATE PROJECT

PROJECT REPORT

REFLECTION REMOVAL FROM PHOTOGRAPHS

Submitted By
Santwana Verma
Dept. of Electrical Engineering,
Indian Institute of Technology, Kanpur
Email: santwana@iitk.ac.in

Supervised By
Dr. K. S. Venkatesh
Dept. of Electrical Engineering,
Indian Institute of Technology, Kanpur
Email: venkats@iitk.ac.in

TABLE OF CONTENTS

ACKNOWLEDGEMENT

I would like to present my heartiest gratitude to my supervisor, Dr. K. S. Venkatesh for giving me the opportunity to undertake the project and to guide me throughout the project. This would have not been possible without the constant encouragement that he provided me throughout the project.

I would like to further thank him to have faith in me for the future plans the project holds. It has been immense pleasure to work under his able guidance. The opportunity helped me realize a lot of things better, both on technical as well as personal levels.

Special Thanks to Luthra Milind Vikas and Subhdeep Saha for helping me with small details of both the project and the report.

I would further like to extend my gratitude towards the reviewing committee to bear with any mistakes I would have committed in the report and to give me the opportunity to present my work.

ABSTRACT

Photography has been around since very long and has evolved to include advances like cameras on the mobile phone, flash etcetera. Use of flash in photography comes with its own disadvantages. Pictures which are taken using flash usually contain reflections from various surfaces thus leading to a loss of information from the image. Thus the need for the removal of reflections arising due to flash is very high. The aim of the project is to remove reflections from the images as far as possible using multiple images with a minimum of two images. The project involves the removal of the reflection by taking two or more images of the area of interest from different angles, with the underlying assumption that the reflection is contained at different regions in the area of interest, aligning the images and then taking the part with no reflection from both the images. The report has been structured to give a tour of the methods used for various components of the process, the problems that arrived during the implementation of various algorithms and then finally the best-suited method.

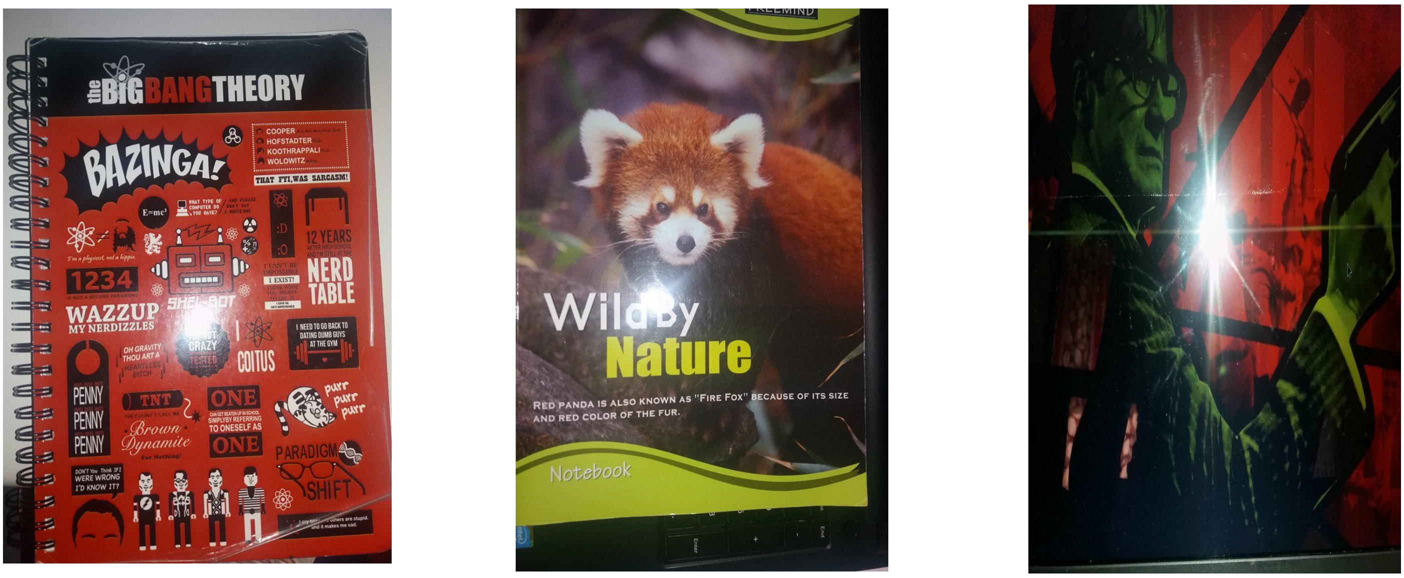


Fig 1. Examples of reflection caused by Flash in Mobile Phones

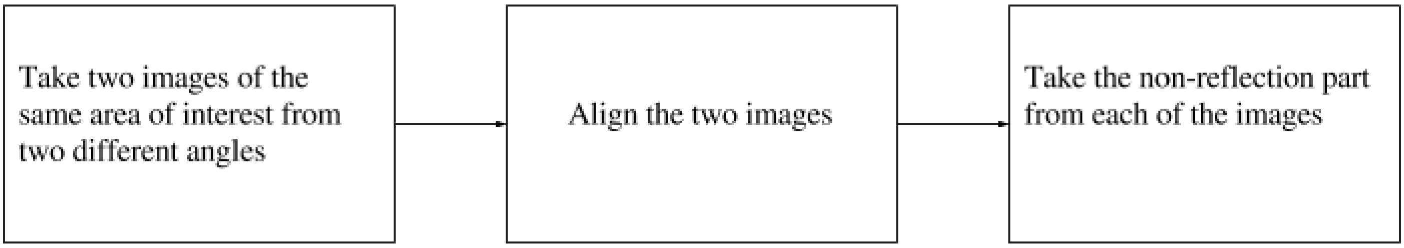
INTRODUCTION

Removal of reflection is important since it causes loss of images. One could suggest turning the flash off in mobile cameras but the reflections are still present due to the nature of material such as glass. Turning off of flash will only avoid large concentrated areas of reflection. This is needed in various day-to-day life incidents such as when taking a picture of a form behind a glass enclosing or taking pictures of course notes on the blackboard. Another day-to-day example could be in case of window shopping where people might want to click photographs of items kept inside the display glasses.

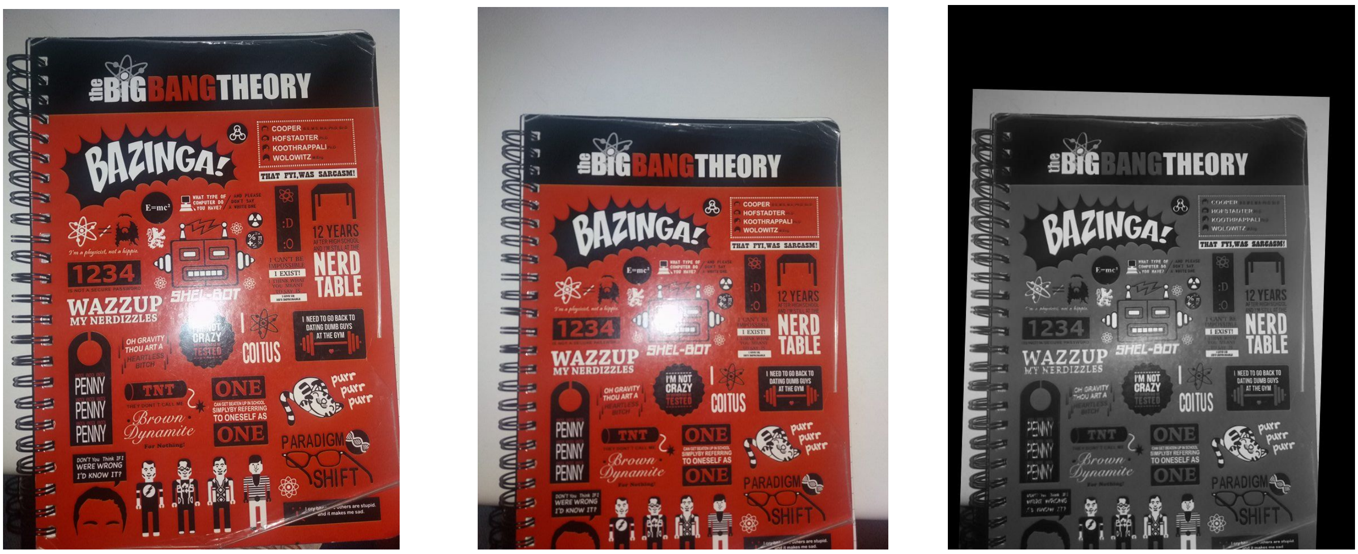
There have been various attempts in this field by researchers around the globe which have been mentioned in the references section.

The report presents a way to remove the reflection by taking two images of the same object from different angles (assuming that the area of interest of the user is contained in both the images). The algorithm used will combine two images in an attempt to remove the reflection.

OUTLINE



PROCEDURE IN DETAIL



(a) First Input Image (b) Second Input Image (c) Result

Fig 2. Input images along with the results

Figure 2 shows two input images of a notebook. As we can see the two images contain areas of bright flash which degrade both the images and thus none of them is individually sufficient to provide us with complete information. The first step is to align images as far as possible to recognize the areas which need to be taken from the other image.

ALIGNMENT

The process of alignment can be broken into two parts, the first part comprises of deciding the features, also known as *key features,* which should be chosen as a reference in both the images for alignment and then to align them using the transformation matrix calculated based on the *key features.*

The two methods which were considered for the detection of *key features* were **SURF Feature Detection** and **Harris Feature Detection.**

Both Harris Feature Detection and **SURF** feature detection were tested on a few images and compared on the basis of a number of key features both the methods detected. The Harris detection detects corners and edges in the photograph while the detector detects both the corners as well as the blobs in the image.

Based on results of the comparison, **SURF** was selected as the optimal method for feature detection.

Fig 3. Key Features Extraction from the two images using **SURF** Feature Detection

FEATURE MATCHING

The next step was to match the features obtained by the detector which was done using the MSAC algorithm.

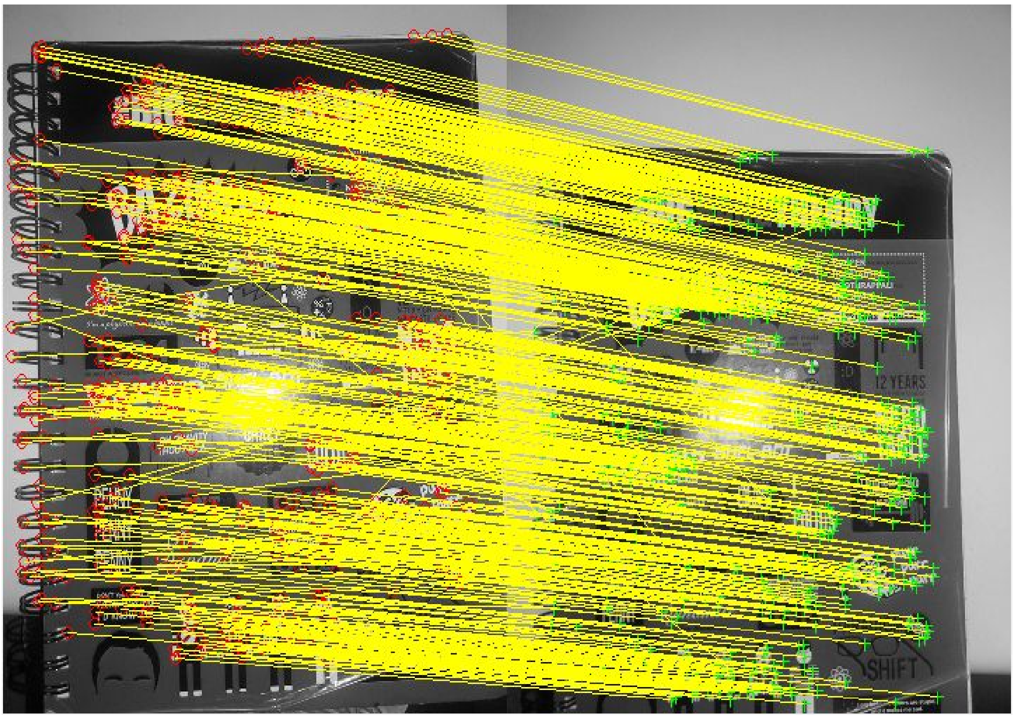


Fig 4. Matched Features (Including Outliers)

Initial mapping contains outliers points as well. The MSAC algorithm estimates the geometric transform excluding the outliers and including only the inliers.

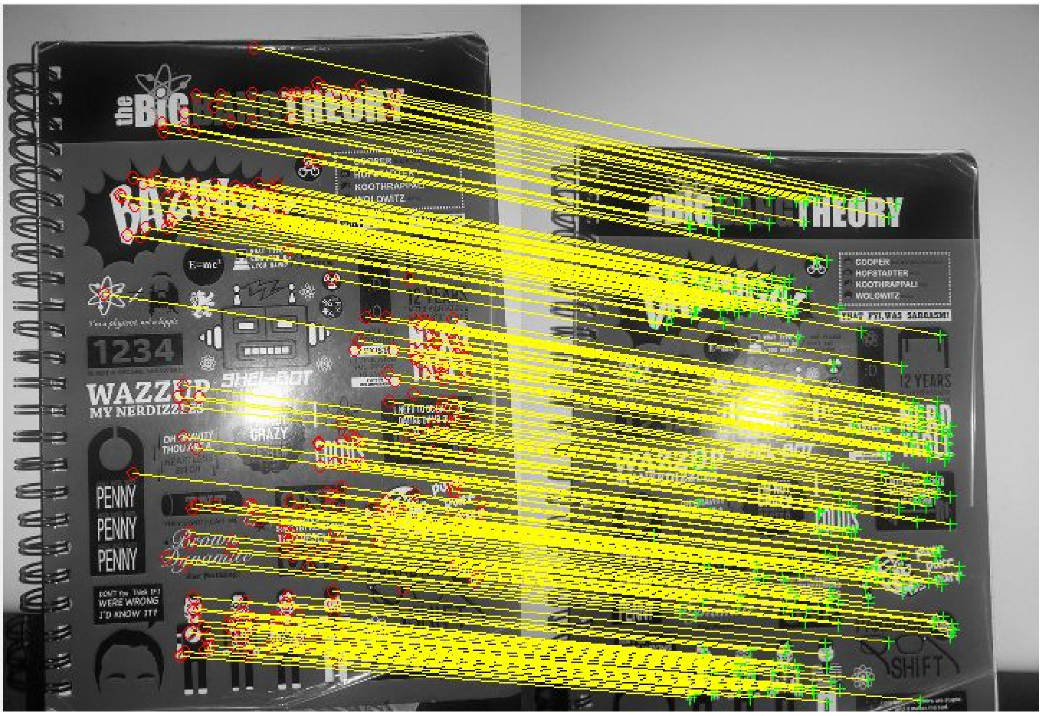


Fig 5. Matched Points (Excluding Outliers)

OVERLAPPING

A geometric transformation matrix is then calculated based on the matched features which will align one image with the other. The transformation matrix is applied to the image matrix to get a transformed image.

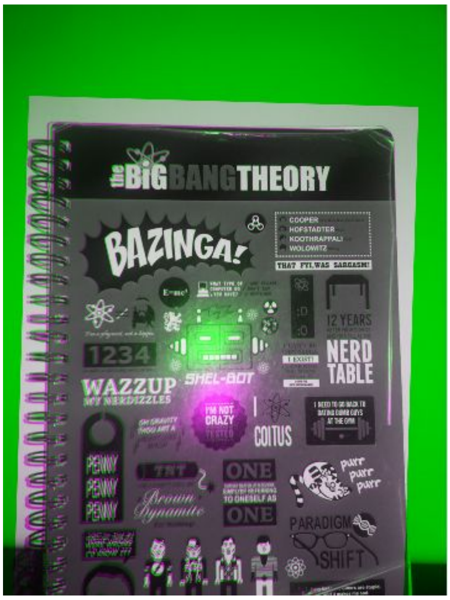


Fig 6. Image Illustrating the alignment of two images

COMPUTATION OF FINAL PIXEL INTENSITIES

Once we have the two aligned images the next task is to generate an image without the reflection. Now since both are images are aligned, we have the pixel intensity at each point. The task is to remove the reflection. Thus taking the minimum of the two intensities available does seem to work as the region having the reflection/flash will definitely have greater intensity than the corresponding intensity at the same pixel location.

Assume the images to be **A** and **B.**

Assume the results obtained from homography (SURF detection and alignment) to be **B`** i.e the alignment of B according to picture **A.**

Then to calculate the final intensity,

$$A(i,j) = \min \{ A(i,j) , B`(i,j) \}$$

Fig 7. Results of the above given examples

SOME MORE EXAMPLES

This section contains a few more examples on which the algorithm was run and the results were found in good accordance.

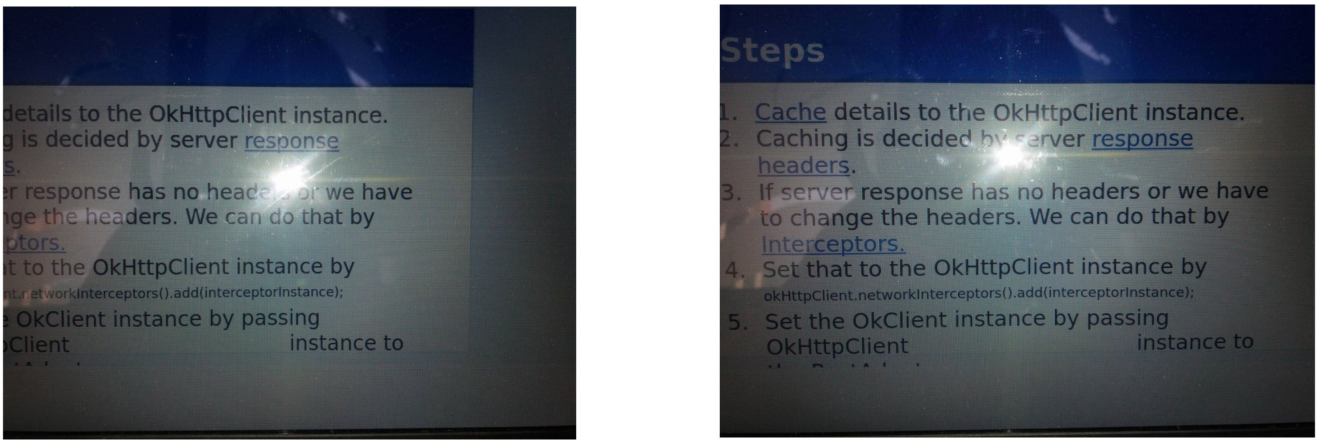


Fig 8 a. Input images

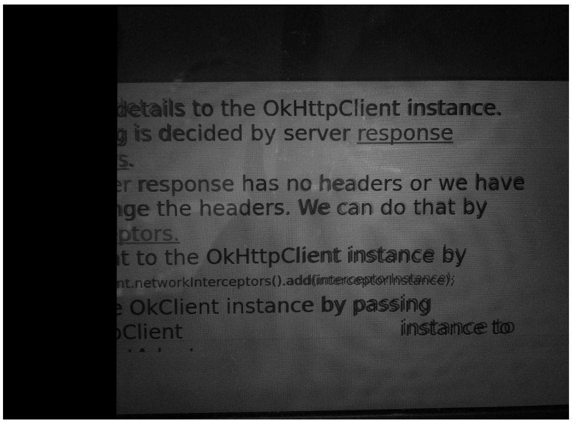


Fig 8 b. Result for images in 8a.

Fig 9a. Input Images

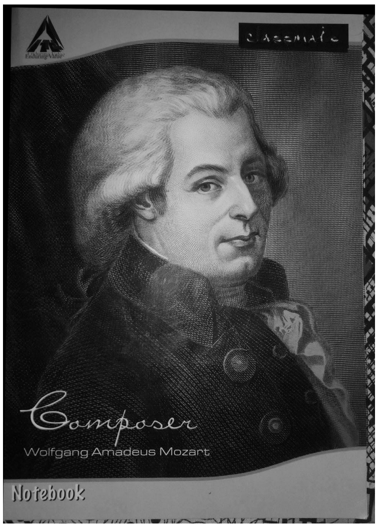


Fig 9b. Final Result

PROBLEMS AND FUTURE PLANS

This approach demands for the ‘Area of Interest’ to be present in both the images because if any one of the images has nothing corresponding to any of the pixel location, it will be zero value and since the algorithm requires a grayscale image which has intensity range between 0-255, both inclusive, the minimum is always zero and the part appears as black in the resulting image. Another problem that can be thought of with the approach is in case the region of reflection is very closely placed in both the images. In any such case minimum of the region will also give rise to a flash region

Thus we came across two problems with the above-mentioned approach :

1. The resulting image will contain only the area common to both the images.
2. The resulting image will contain minor reflection in areas of the common diffused flash area.

The problems need more work and are planned ahead. With the current focus, the problems are small but require some efforts on the part of the user. The user needs to try to include the area of interest in both the images.

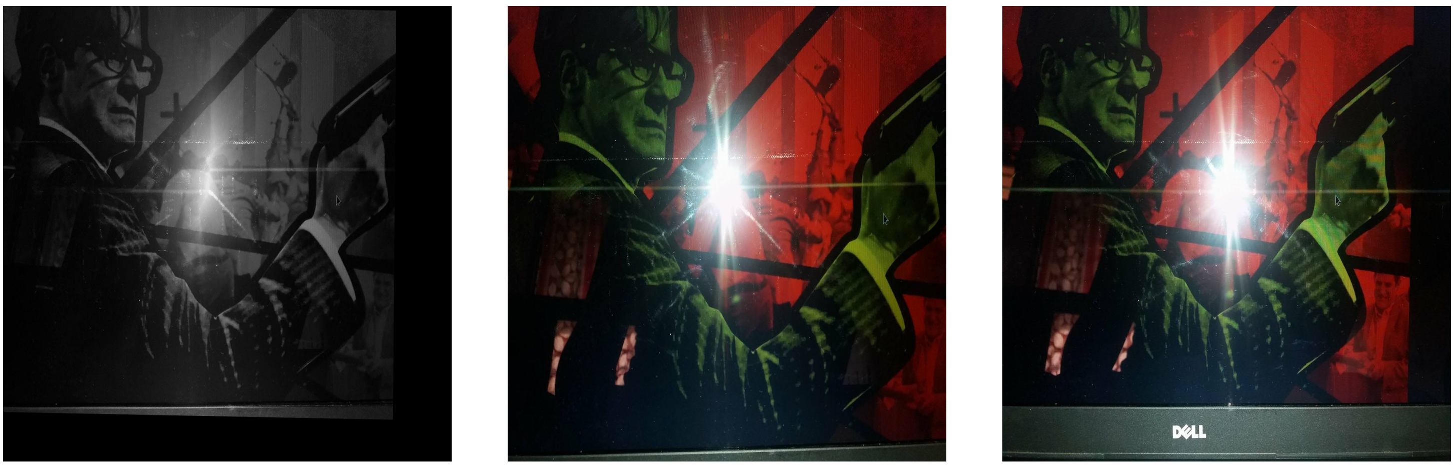


Fig 10. Case when the computation fails

Other problem that can arise in the approach is when **Alignment** fails.

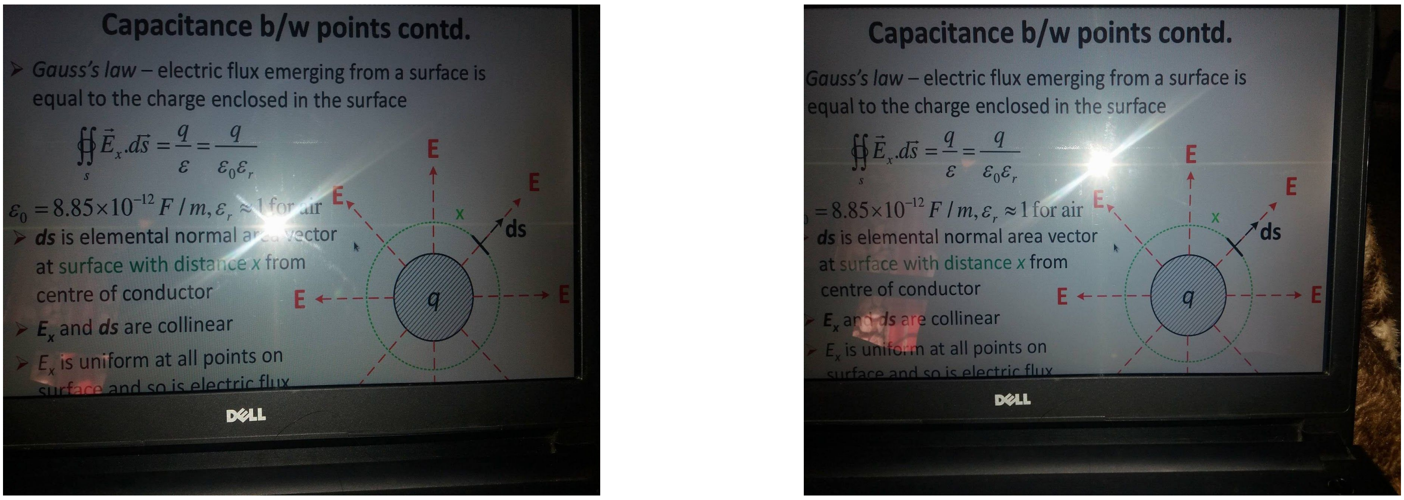


Fig 11. Examples to show the failure of alignment

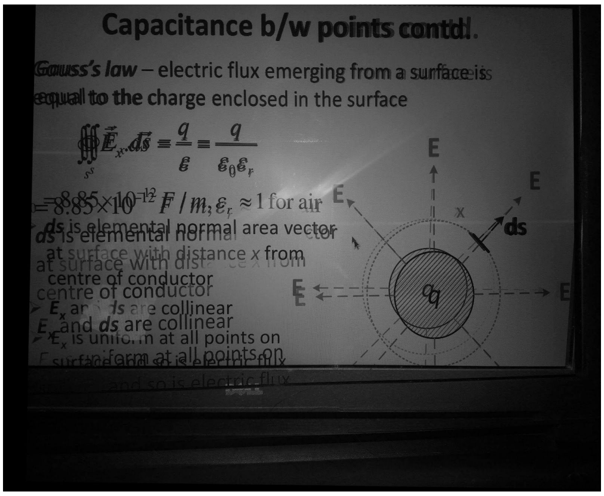


Fig 12. Result with misalignment

In the example above we can see the case in which alignment fails. We can also observe that not the entire image is misaligned. Instead some part of the text is perfectly aligned while the others aren't. One possible solution to this could be applying homography again in parts of image and then combining them into a single photograph. If there exists a linear model which will cause the images to align properly, it will be displayed as better results after one iteration. If the image does not get better with the iterative approach, then the image contains radial distortions which can be stated as one of the shortcomings of the cameras of mobile phones.

Another solution involves scaling up of the current solution i.e using **n images** instead of two images and getting $^{\mathbf{n}}\mathbf{C}_{\mathbf{2}}$ solutions. The next step with this approach is to decide which of the $^{\mathbf{n}}\mathbf{C}_{\mathbf{2}}$ images is the best solution. The solution to this issue lies in the fact that if the images are not aligned, they will have double edges in some parts of the image. This increases the number of edges the image contains. Keeping this in mind, image with least number of edges can be safely declared as the best possible solution among the images of interest.

To get the number of edges in a image **SOBEL** edge detector is applied on the images which returns a matrix having 1 if the pixel is an edge point and 0 otherwise. The total number of edge points is given by sum of the elements of the matrix. The method was tested on a sample which has highly misaligned images. The initial data contains three images.

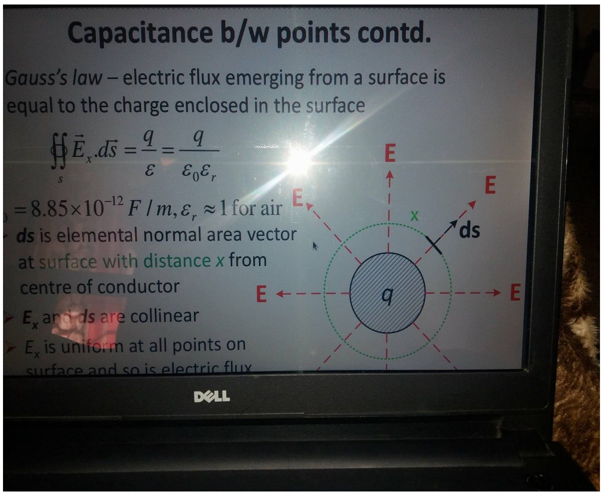


Fig 13.a

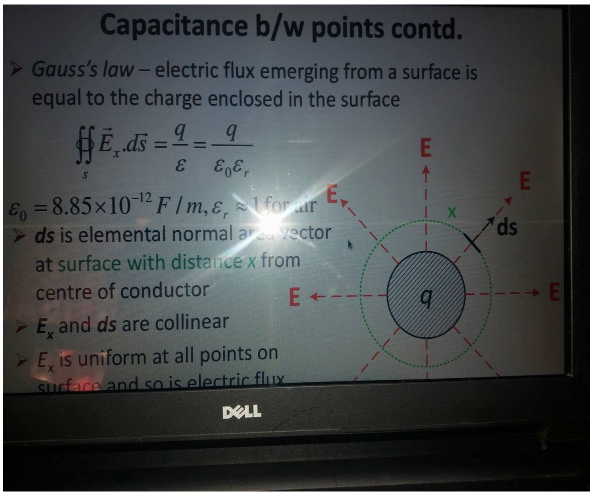


Fig 13.b

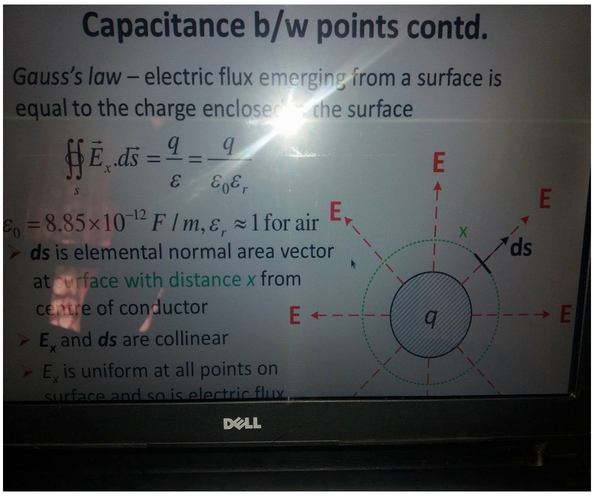


Fig 13. c.

Fig 13. Three Input Images

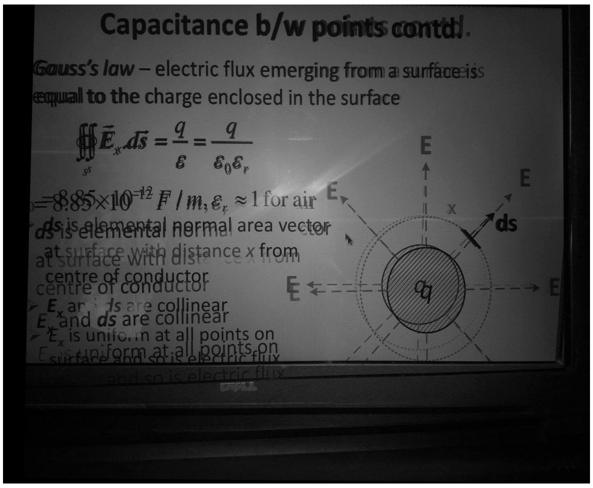


Fig 14.a

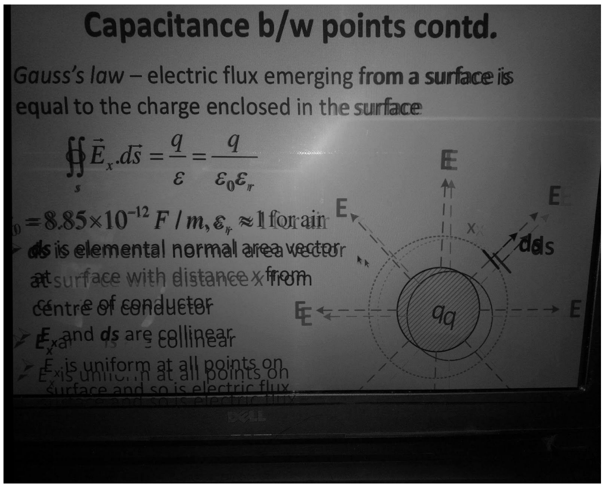


Fig 14.b

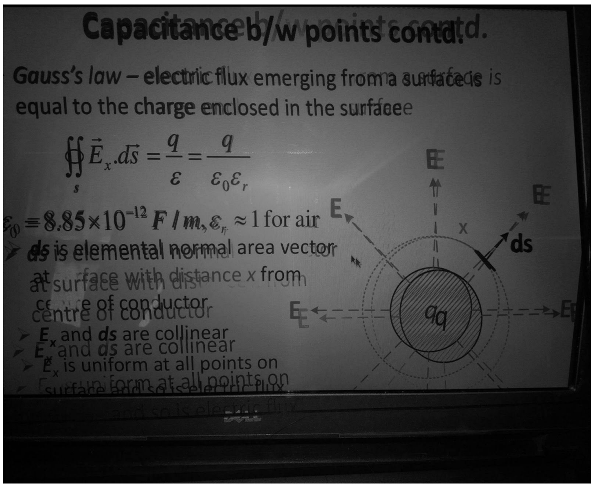


Fig 14. c.

Fig 14. Three misaligned results taking two out of three inputs at a time

One way to decide the best out of the results is to manually decide which to keep and then discard the rest of the images. But the manual method is cumbersome and not feasible.

The result obtained from the edge analysis is in accordance with the one obtained with manual comparison.

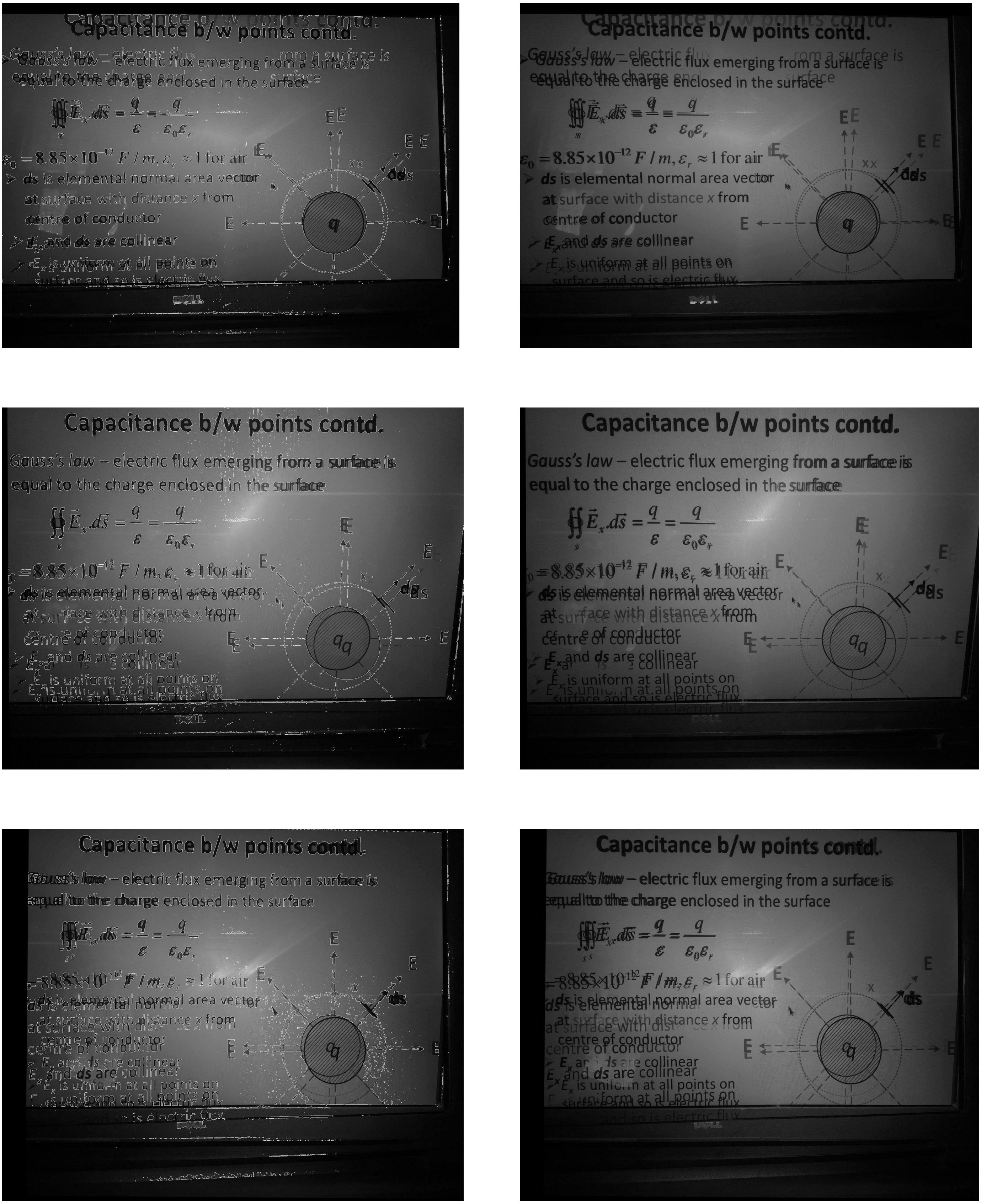


Fig 15. Image results and their corresponding results with edge point

The edge points for the above images (Fig 15.) are as follows : **114399**, **121405** and **128275**. The edge points suggests that the **first** result is the best result out of the results .

<u>CONCLUSION</u>

The results obtained from the present approach are not perfect, but present promising work for future. The removal of reflection can help in a lot of things as stated in the introduction. The present approach comprises of detection of features by **SURF** and **MSAC**, and then a simple algorithm to get the pixel intensities at each pixel location. The future work consists of optimizing the approach even further by using edge detection techniques to identify the best result. Further optimization can be done by the applying the homography twice in parts.

REFERENCES

1. [1]Pooja Ghosh, Achala Pandey* and Umesh C. Pati, “Comparison of Different Feature Detection Techniques for Image Mosaicing”, ACCENTS Transactions on Image Processing and Computer Vision, Volume-1 Issue-1 November-2015.
2. [2]Yilong Geng and Zizhen Jiang, “Reflection Removal on Mobile Devices”.
3. [3]Matthew Hu, “Reflection Removal Algorithms”.

[1] http://accentsjournals.org/PaperDirectory/Journal/TIPCV/2015/11/1.pdf
[2] https://stacks.stanford.edu/file/druid:np318ty6250/Geng_Jiang_Reflections_Removal_on_Mobile_Devices.pdf
[3] http://stanford.edu/class/ee367/Winter2016/Hu_Report.pdf