Pinholes and pinspecks:

Recovering unseen pictures through reflexive light

Yiqian Wu 3170104915
Shangguan Yue 3170104168
Ye Yan 3170104149
Zhejiang University

IAVI 2019

1. Motivation

Light propogates everywhere in our daily life and contains various information. However, due to the limited sensitivity of humans' eyes and unawareness, we may neglect many useful details of objects behind the superficies of some photos.

In our investigation, we aims to recover the unseen scene behind the photos through the reflexive light. We consider this project both interesting and meaningful, considering its ability in discovering potential knowledge and its application in search-and-rescue, anti-terrorism and transportation.

The idea about pinholes and pinspeks is based on two facts. Firstly, pinhole imaging reflects clear but reverse pictures of objects, for light transports in straight lines. There are countless pinholes in our lives, although not as ideal, presenting light information indirectly, and windows are vivid examples. On the other hand, shadows can carry light and color information, which is often ignored, following the principle shown as below. Therefore, we hope that through 'pinhole' images caused by windows, and shadows caused by intentional occluders, we can extract and recover more useful unseen information.

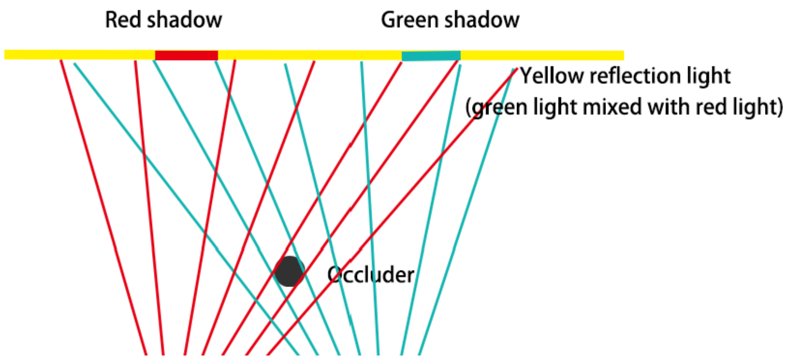


Figure 1.1 principles of colored shadows

2. Related work

There exists several ways in computer graphics to achieve NLOS target. The NLOS, none-line-of-sight imaging, has several ways to reconstruct the image with different facility setting and results. And their method varies also, which can be classified as *active methods* and *passive methods*. These approaches range from using naturally-occurring pinholes or pinspecks to using edges resolve the scene while light field can also be useful in reconstructing too.

For different goals in application, methods and the complexity of setup may vary greatly. In our case, we aim to recover simple and unrestricted data that can be gained from everywhere in our daily life, which is quite random and cause difficulty to specific advancement. If the experiment condition can be more concrete and more explorative techniques can be applied, matrix solving and accurate method like wave-transport can be implemented in to recover a higher-level of accuracy too.

3. The Math and principle

3.1 Pinhole

Accidental pinhole cameras happen everywhere by accidental arrangement of surfaces in the world. The images formed are generally too faint and blurry to be noticed or misinterpreted as shadows or inter-reflections.

It differs from ideal pinhole in several ways: large non-circular aperture, a complex projected surface compared with white flat lambertian surface, multiple apertures, inter reflections etc. We will make some strong assumption to avoid several of these, for example, we use a dark box to reduce inter reflections and make the box empty except necessary equipment.

A simple illustration of math in pinhole is

$$L(x, y) = T(x, y) - S(x, y)$$
$$I(x, y) = \rho(x, y) * L(x, y)$$

Function L is the light information projected on the wall the function ρ is about the reflexive characteristic of the white wall. Function I represents the light information gained by the camera.

3.2 pinspeck

General idea of pinspeck imaging can be presented simply through a subtraction.

$$I\ occluder\ (x, y) = I\ background\ (x, y) - I\ pinhole\ (x, y)$$
$$I\ pinhole\ (x, y) = I\ background\ (x, y) - I\ occluder\ (x, y)$$

The condition of performing pinspecks is that, we already gain a background image, which is caused by lights through the hole without any coverings, and an occluder image, which is gained by placing an occuler in the route of lights, and importantly, under same exposure and shutter speed. Then by simply subtract the two images, we obtain a pinspeck image which theoretically presents the same result image as with a pinhole of the occluder's size.

As to relate the pinspeck with the pinhole, we take the diffuse reflectance of the material of the display screen, which is the wall, and the size of the aperture, which is the window, under consideration. Finally, the formula is presented as below, where *ρ(x, y)* is the albedo variation, *L* is the illumination image, *T hole* is the aperture and *S(x, y)* stands for the ideal pinhole image of the scene to be captured.

$$I\,pinhole\,(x, y) = \rho(x, y)(L(x, y) - L\,occluder\,(x, y))$$
$$= \rho(x, y)(T\,hole\,(x, y) * S(x, y))$$

4. Our Setup

Based on the principles explained in Chapter 2, we design a set of portable equipments to simulate the condition of a ideal darkroom. The main material is a paperbox of a size of 530*290*370mm. White papers are used to act like a white wall.

Our experiment consists of two part. In the first part, we experimented with indoor light from an iPad, light with strong energy and few loss of information, so as to provide proof for the validity of pinspecks. For this part, occluders are hung inside the darkroom before the wall, but does not cover their shadows. Our setup for this part are designed as below.

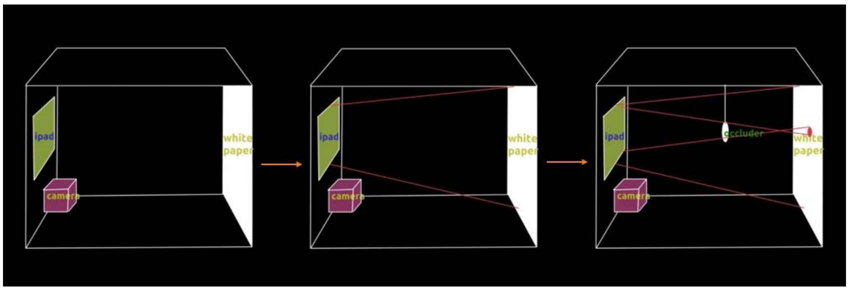


Figure 4.1 setup for light from iPad

In the second part, we experimented with outdoor light, which is much noiser, weaker, but resembles more of daily scenes. For this part, occluders are placed outside the darkroom and are presented as the passers-by, moving objects, hands, etc, as to simulate a more actual scene.

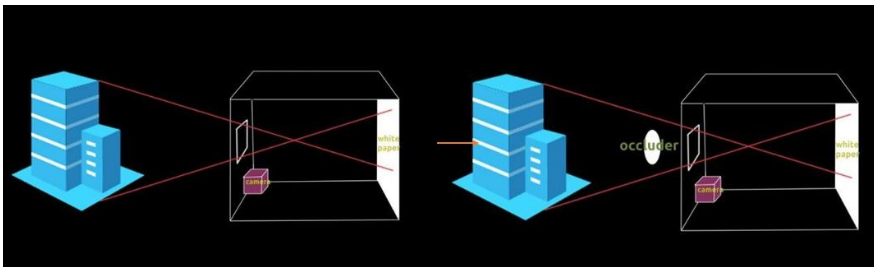


Figure 4.2 setup for light from outside

Fig 4.3 our actual equipment

5. Data acquisition and analysis

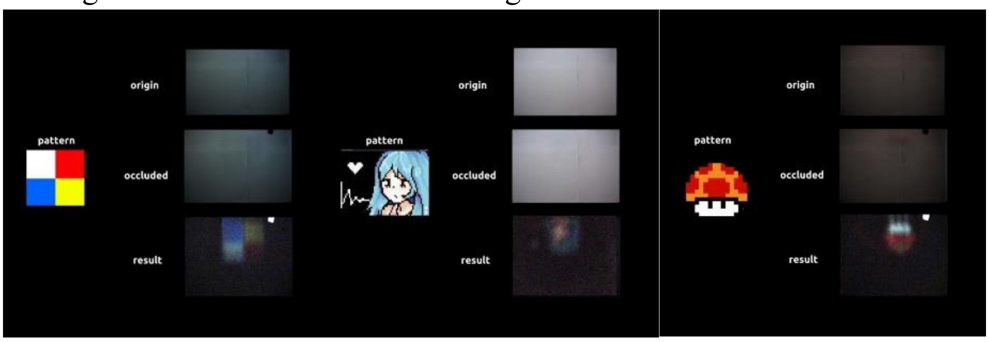
We perform pinhole and pinspeck experiments in vaious situations.

a) different environment: noon time, late afternoon time,
   different light conditions : cloudy and shiny daytime.
b) Different object reflexive characteristcs: objects with different diffuse reflectance, including normal buildings, trees, metal metal materials like cars, glass, etc.
c) acquisition tools: industrial cameras, the front camera of iPhone with fixed lens, and a 24.2 megapixel SLR camera, for capturing single images and serial videos.

Results are presented as follows. Analysis of effects are also given in this section too.

5.1 Data with light from iPad

With light from iPad we obtained some good results.



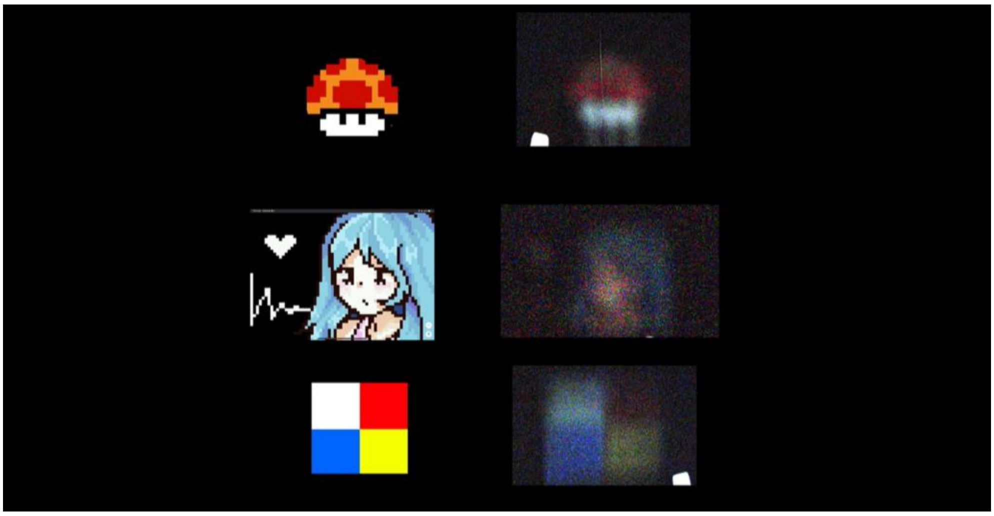


Figure 5.1 pinspeck results of light from iPad

In this experiment, ipad light are projected onto the wall. *Origin* is taken while there are no occluder while the *occluded* is taken while the occluder blocks some of the light. We can see that few information seems to be contained in the data acquired directly and we can not see what is in the ipad. However, when we subtract this two and enlarge its effect, we can see that some information does exist in those seemingly useless pictures.

We actually captured the occluder while experimenting, but it does not matter since the shadows and the pinspeck images are not covered.

From the images we can see that light from iPad gives ideal results. The reason is that while the distance between the wall and the light source is short (530mm), there is few loss on the way of the light transportation, and the light from iPad is so strong that the patterns can be reflected clearly.

Based on this analysis, it is quite sure that results with light from outside scenes would contain much more noise and suffer more loss.

5.2 Data with light from outside

In this part, we discover some factors that greatly influence the quality of results.

5.2.1 Direct sunlight and diffuse sunlight

We discover that diffuse sunlight (that is, in cloudy environment) performs better results, for direct strong ligh t(in sunny environment) will cause accidential reflection halation on the wall, leaving unexpected colored regions on the pinspeck images.Also, strong reflection on the walls are possible to wipe out details even the main character of the scene.

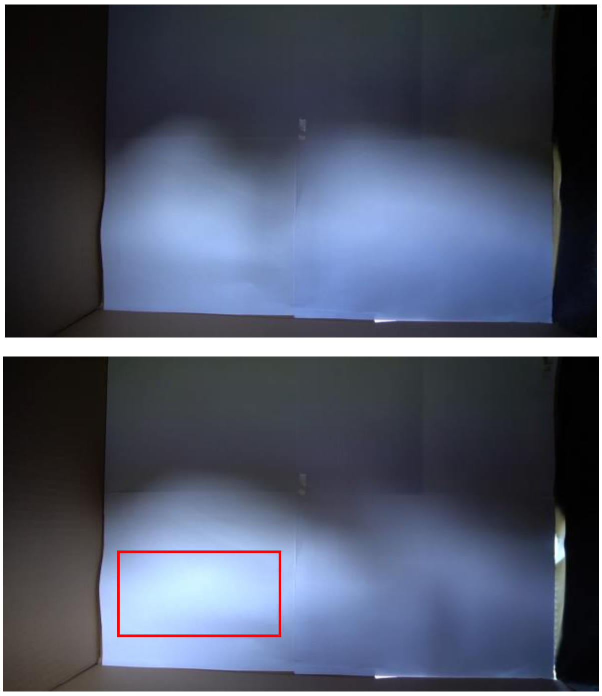


Figure 5.2.1 captured image frames of diffuse sunlight and direct sunlight

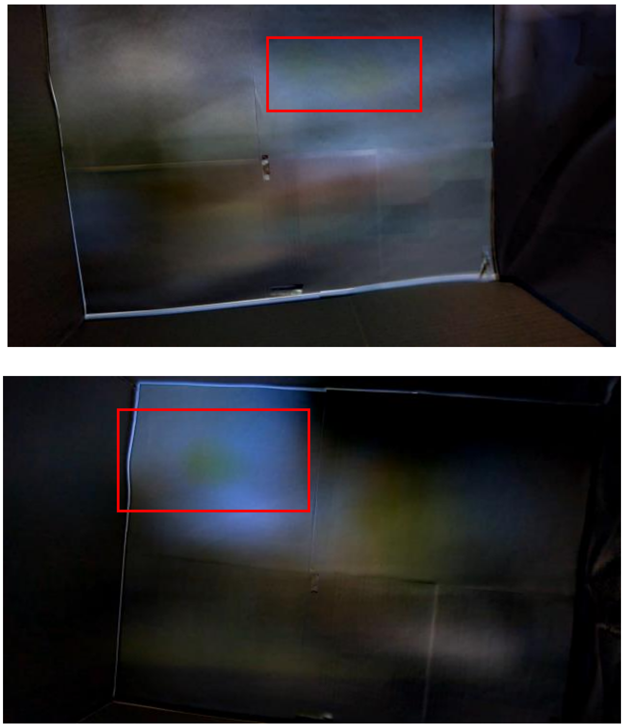


Figure 5.2.1 single pinspeck frame with light halation on captured images

5.2.2 Objects with large and little reflectance

It is clear that objects with large reflectance is likely to introduce great distortation since the light projected is no longer representing the expected region and contains noise information which comes from other objects.

Figure 5.2.2 Objects with large reflectance and its result

Figure 5.2.2 Objects with little reflectance and its result (the right one need to flip horizonally to fit the truth)

5.3 Data using phone cameras

We gained bad results with data from phone cameras, for that compression on phone applications and low resolution front camera produce relatively noisy pictures.

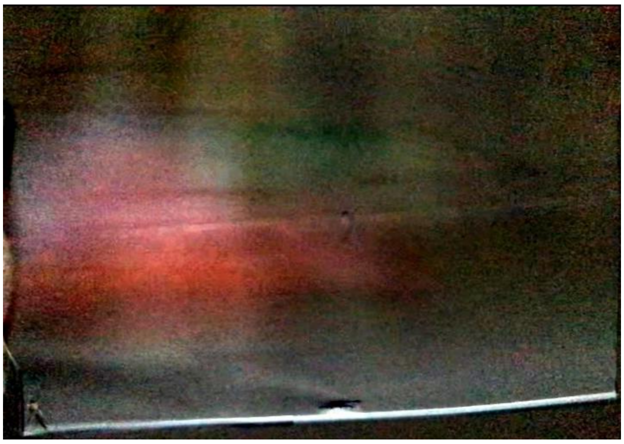


Figure 5.3 single pinspeck frame of a red car captured by phone camera

6. Data processing

We experiment several ways to obtain better results of pinhole and pinspeck images. Meanwhile, we discover that the two methods have distinct and different strengths and drawbacks.

6.1 Image enhancement on pinhole and pinspeck images

Pinhole images in life are often vague and unclear on the outline of the objects. Using MSRCR image enhancement method, we gain a clearer picture of the structure of the outside scene. Here we present 5 of our enhanced pinhole results, where the left one is the original image and the right one is the processed image.

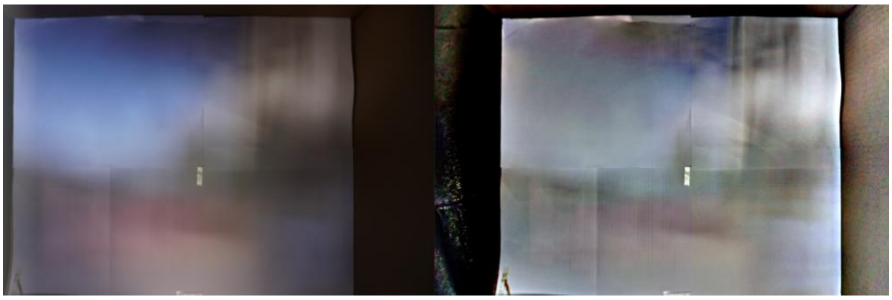


Figure 6.1.1 enhanced pinhole result of a building and a red car on the left

Figure 6.1.2 enhanced pinhole result of a building and a tree in the front

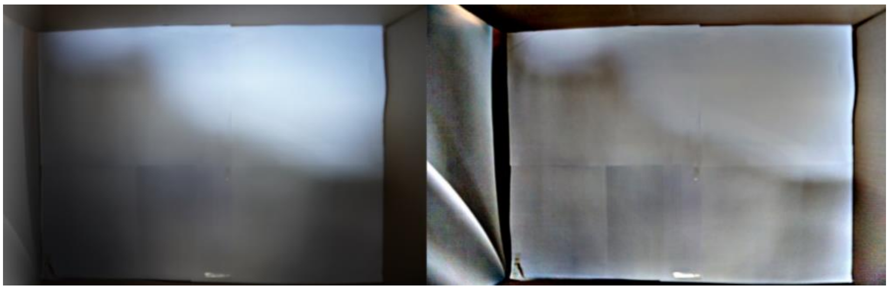


Figure 6.1.3 enhanced pinhole result of a range of buildings-1

Figure 6.1.4 enhanced pinhole result of a range of buildings-2

However, the same method cannot be applied to pinspeck images, for that classical image enhancement method focuses on the original reclection on the object, but aims to remove any effects of extern light source. And as for enhancement on pinhole images, the outlines are strengthened but the general color fades a bit. For pinspeck images, who reflects stronger outside colors, the results are quite disastrous.

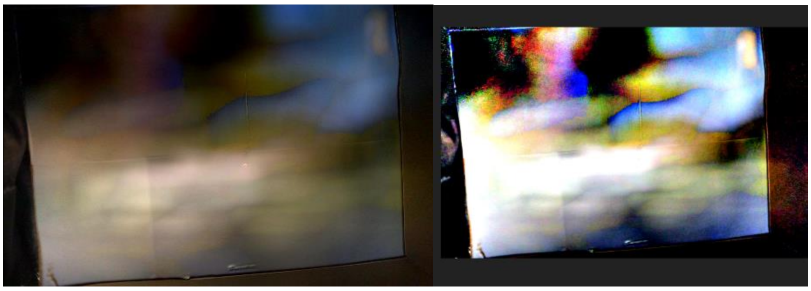


Figure 6.1.5 enhanced pinspeck result of a normal street scene

6.2 Denoise on pinspeck images

Single pinspeck image are very noisy due to the continuous random noise caused by environment and camera status, and it is almost impossible to get a pinhole image and a occluded image at the exact same time. We tried several methods to dig out information hidden in pinspecks. Since classical image enhancement methods fail on pinspecks, we try to perform denoise by taking the mean value of continuous frames from a video as the reference frame

As there is continous different envirenmental noise at every neighbouring frame, we take several frames which are captured without occluders, and take the mean of them to be a reference pinhole image for the subtraction. For convenience of calculation, we choose 50 neighbouring frames for the mean.

We can see from the result, that the mean reference produces smoother pinspeck image.

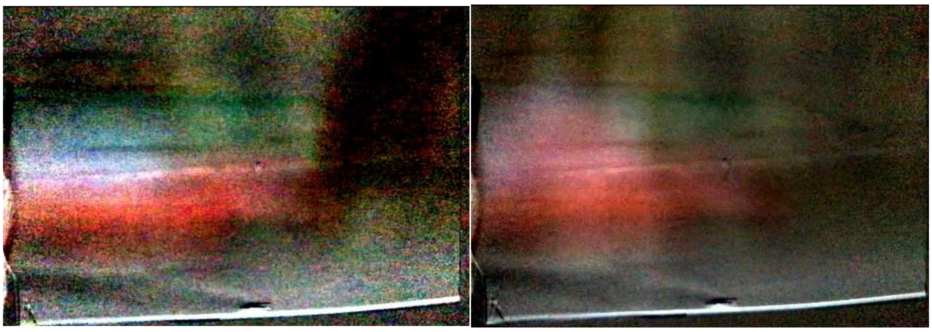


Figure 6.2.1 pinspek with reference of: left-single frame; right-with mean of 50 frames

6.3 Merge for a more complete pinspeck image

When processing the video data, we discovered that because the ‘occluders’ in frame were rapidly and violently changed, causing loss of certain parts in the subtracted pinhole images. Thus, we hope to compose a complete image with the good parts of every subtracted frame.

6.3.1 Select frames and result pixel values indepently on RGB channels

We select several key frames of the video, of which each presents certain part of relatively high-frequency details, and merge them together for a more complete pinspeck result. For each pixel to choose, we compare all the chosen frames and select the maximum value, indepently on R, G, and B channels.

This means that parts of the pixels on the result image are not completely from one of the frames. But by selecting each channel indepently, we ensure the visual smoothness of the colors of the image. We find that setting points' values in this way ((0,255,75) and (5,100,60) -> (5,255,75)) has great priviledge over setting points'values by choosing the largest pixel of sum of RGB ((0,255,75) and (5,100,60) -> (0,255,75)) since it introduces more continuity (see figure 6.3.0) , and that selecting pixel values using the sum of RGB values makes the whole image 'dirty'.

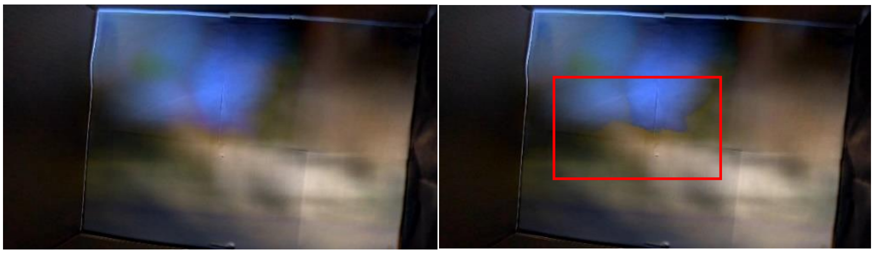


Figure 6.3.0 compare between two differement mix methods

Compared with single pinspeck image, the merged image are contains more information and details.

Figure 6.3.1 the image above is merged from the six frames listed below

Figure 6.3.2 comparison of: original images; merged pinspecks; single pinhole image

Also, we search for ways to select 'good' frames, which supposed to contains more high-frequency imformation and are relatively complemented, automatically rather than by hand.

Firstly, we cut the frame into 9 parts of the same size. Firstly, we compute the number of pixels that have a higher value than a set bound to ensure that it is not a black picture and contains certain information.

For each part, we select one frame that has the maximum sum of pixel values of this part, and finally we have nine frames for the merge.

The result is rather satisfactory.

Figure 6.3.3 left: with frames selected automatically; right: with frames selected by hand

Figure 6.3.4 result of another scene with automatically selected frames

6.3.2 Select frames and result pixel values based on illumination value

Apart from selecting based on RGB values, we try to see what the results would be when choosing illumination values as the judge, and discovered that it also performed much worse than selecting in RGB channels independently.

We consider the reason is that by selecting RGB independently, neighbouring pixel values may keep their connection in the least of one color channel, which smoothen the final image to some extent. However, by comparing based on a single dimension I, the neighbouring pixesl do not have such strong relationship among each other, thus causing the color of the final image to break in blocks.

Figure 6.3.5 details of result of: based on RGB independently; based on I; based on sum

We also show you the table of results of choose frames or pixels use brightness or rgb, we can see there is just a little difference. We use the 7.1.MOV as an example:

	Choose frames use brightness	Choose frames use rgb
Choose pixels use brightness		
Choose pixels use rgb		

6.4 An attempt to optimize pinhole results

We use a image enhancement method to optimize the result of pinhole. However, we can see that even though more details are introduced but the color is reduced.

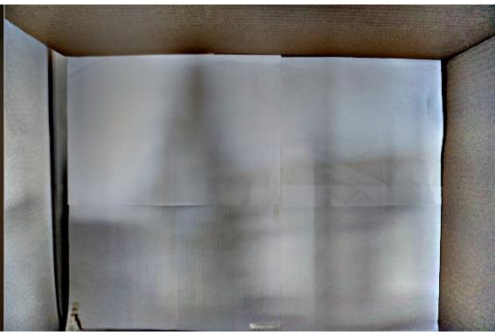



Figure 6.4 pinhole after enhancement and before

Motivated by this result, we consider can we combine two pictures together (with both color and the details). We think of graph-cut immediately. If the algorithm can correctly extract out the a component in the enhancement image (with more details) and integrate the rgb in the unprocessed pinhole image (with more color), it can produce better result.

However, after trying several algorithms like Grab etc, we find that partly due to our data and partly due to the algorithm, the current popular way of graph-cut can not correctly identify the componenet as we expected.

Therefore, either can we outline the component manually to add the componenet correponded in the other pictures together, or we should explore some new 'graph_cut' algorithms specific for our application.

7. List of changes

1. General data analysis, including data acquisition and data process
2. Realize the goal of select video frames automatically to merge pinspeck results
3. Several attempts to improve the result, like denoise, optimize, etc.

8. Further work

In our project we explore the effect of pinhole and pinspeck and try to find ways to analyze and further optimze it. However, the result can still be improved from several aspects:

1. During acquisition

   Further explore how the location of occluders and number of photos taken affect the result.

   How the camera's parameters matters in our acquisition like shutter speed etc.
2. During data processing:

   2.1 the acquisited picture is generally of lower resolution ratio and saturation level, which differ its from general denoise method. Specific modification must be made to satisfy this situaion.

   2.2 Extended method to integrate different results to get more complete information.

   2.3 Give some prerequisite to better estimate the parameters used in solving.

9. Timeline and goals

The deadline is about 11.6 and the project start date is at 10.17. There are about 3 weeks of time.

First week: read the related paper and use its idea to develop our first stage of code. Design our experiment and get the raw data.

Second week: on the base of the first week, adjust the code from the data and try to repeat the result as the paper said. Improve our way of getting the data and discuss the factors that may interfere with our result in the experiment.

Third week: perfect the result from week2 and try some improvement. We can inject some creativity into our project in several aspects: simplify the input, simplify the experiment environment, improve the quality, discuss some solution to fit the original shortages, its application in other fields, etc.

10. References

(1) Inferring Light Fields from Shadows
(2) Accidental pinhole and pinspeck cameras: revealing the scene outside the picture
(3) Turning corners into cameras: Principles and methods.
(4) Anti-pinhole imaging. Optica Acta: International Journal of Optics
(5) Detection and tracking of moving objects hidden from view
(6) Occluded imaging with time-of-flight sensors
(7) Linear view synthesis using a dimensionality gap light field prior
(8) Human detection using depth information by kinect.
(9) Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging.
(10) Learning to synthesize a 4d rgbd light field from a single image.
(11) Estimating motion and size of moving non-line-of-sight objects in cluttered environments