Exploration of Dual Photography

Rachel Tang
Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA
rachelt1@andrew.cmu.edu

December 14, 2021

Abstract

Dual photography is a technique of generating a virtual image using a projector-camera setup. This virtual image (interchanging the position of the projector and the camera) is created by obtaining a light transport matrix using the images captured by the camera under different lighting by the projector. In this paper, I implemented the processing pipeline described by Sen and Darabi in their discussion on Compressed Dual Photography. I experimented with two different sets of projector patterns, fixed scanning patterns and Bernoulli binary patterns, to attempt to create dual images.

1 Introduction

The theory of Helmholtz Reciprocity is the idea that given a light source and a sensor, the measurement of a light ray coming from the light source to the sensor should be able to be identical if the locations light source and sensor were swapped. This is represented by the equation $f(\omega_i \longrightarrow \omega_o) = f(\omega_o \longrightarrow \omega_i)$. This concept can be scaled up to generate photorealistic images from the perspective of a light source without having a sensor there; this is Dual Photography.

In this paper, I explored concepts described by Sen and Darabi in their paper on Compressed Dual Photography [1]. This was an extension on the work of Sen et al. on Dual Photography using adaptive algorithms [2]. Rather than requiring computation in between each image capture, compressed dual photography is able to use pre-captured images in order to produce similar results.

2 Background

2.1 Dual Photography

The goal of dual photography is to be able to use one photo setup to obtain images from two different perspectives: one captured by a camera in the set up and one generated using information about the lighting and the captured camera images. The setup we use to do this involves using a projector and a camera both pointed at a scene; an image captured by the camera with lighting from the projector is referred to as the primal image. We can use these primal images in order to obtain a light transport matrix and generate a dual image, an image from the perspective of the projector as if the camera were lighting the scene. Refer to Figure 1 for a visualization of what the setup looks like.

The light transport matrix that describes the way light travels from light source to sensor. In the context of dual photography, the light transport matrix describes the contribution of a single projector pixel to the camera pixels. Thus, we can write the relationship between a projector-camera pair as $\boldsymbol{c} = \boldsymbol{T}\boldsymbol{p}$ where $\boldsymbol{c}$ is the flattened camera image (dimensions $m \times n$) of size $m * n \times 1$, $\boldsymbol{p}$ is the flattened projector pattern (dimensions $p \times q$) of size $p * q \times 1$, and $\boldsymbol{T}$ is the light transport matrix of size $m * n \times p * q$. This transpose of the light transport matrix, $\boldsymbol{T}^T$, can be used to generate a dual image from the perspective of the projector as if the camera was lighting the image.

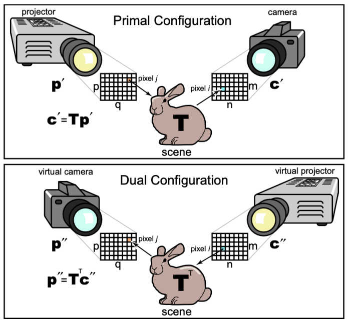

Figure 1: Principle of Dual Photography
This visual from Sen et al. [2] shows the setup the primal image capturing. The bottom image labeled "Dual Configuration" shows the configuration of virtual image we would like to generate.

2.2 Reconstruction of Light Transport Matrix

If our projector matrix transposed ($\boldsymbol{P}^T$) meets the Restricted Isometry Condition (RIC), the theory of compressed sensing shows that we can reconstruct the signal with a relatively small number of linear samples. RIC is the condition that for the z-sparse matrix $\boldsymbol{P}^T$, the subsets of z columns of $\boldsymbol{P}^T$ are all approximately orthogonal. Under these conditions, we are able to pre-capture a set of images and still be able to get a good reconstruction of the light transport matrix with a reasonable amount of samples. This proposed method by Sen and Darabi [1] improves upon the dual photography method proposed by Sen et al, which uses an adaptive algorithm, requiring computation between image captures.

Under RIC, we are able to transform our images and projected patterns into linear equations of the form $b = Ax$. For one camera-projector image pair, we write their relationship as:

$$\boldsymbol{c} = \boldsymbol{T}\boldsymbol{p}$$

Then, the collection of all camera-projector image pairs is:

$$\boldsymbol{C} = \boldsymbol{T}\boldsymbol{P} \longrightarrow \boldsymbol{C}^T = \boldsymbol{P}^T\boldsymbol{T}^T$$

The transpose property of matrices effectively switches the order of the multiplied matrices, so we are able to turn the unknown value of $b = Ax$ from A to x, a linear equation we can solve. $\boldsymbol{C}^T = \boldsymbol{P}^T\boldsymbol{T}^T$ yields the equations:

$$\boldsymbol{c}_i = \boldsymbol{P}^T\boldsymbol{t}_i$$

where $\boldsymbol{c}_i$ is a row of matrix $\boldsymbol{C}$ and $\boldsymbol{t}_i$ is a row of matrix $\boldsymbol{T}$.

3 Method

3.1 Setup and Implementation

My setup for image capture can be seen in Figure 2. I used black paper as my backdrop, an iPhone 12 camera tethered with wired earbuds to ensure no movement between captures and with auto-adjustments turned off, and an Optoma ML500 projector. To ensure a sparse matrix, I also chose convex objects with little inter-reflection.

To calculate the matrix $\boldsymbol{T}$, I used the Orthogonal Matching Pursuit algorithm (OMP). This algorithm is able to recover a sparse signal from noisy measurements, but does not perform to recover the exact signal because it is a greedy algorithm. Sen and Darabi [1] suggest using Regularized Orthogonal Matching Pursuit (ROMP), which I also implemented but did not utilize due to inefficiency and time constraints.

Please see `https://github.com/racheltang/15463final` for the code I used for this project

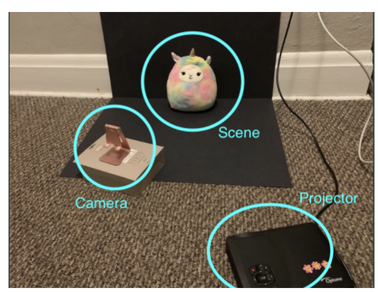

Figure 2: Setup
This is the setup I used to capture images. As labeled, there is a stand to place the phone camera on the left, a scene on top of a black backdrop, and a projector placed in the bottom-right corner of the screen

3.2 Projected Patterns

I implemented the dual photography pipeline using two sets of images: fixed pattern scanning and random Bernoulli patterns, both with a projector resolution of 256×256 pixels. The fixed pattern scanning image set contains 64 images each with every eighth pixel lit in both the x and y directions. The Bernoulli pattern image set contains 512 images each with pixels that are either lit or not lit based on a Bernoulli distribution. I decided to capture more images for the Bernoulli patterns based on what was suggested in Sen and Darabi's paper on compressed dual photography [1]. As previously mentioned in section 2.2, it is important that our projector images meet the Restricted Isometry Condition. Both of these projected patterns meet this condition: the fixed pattern scanning will create a projector matrix where the columns of the projector matrix transposed creates an orthogonal set and Bernoulli binary matrices have been proven to meet RIC.

After capturing the images, I did need to process the images so that I could get as much of a sparse image as possible. To do so, I subtracted out a "black image" which is an image of the scene with the projector lighting it a black screen. I also applied a threshold as well as increased the contrast to get rid of more ambient lighting that the subtraction of the black image did not capture. See Figure 3 to see projected patterns with their respective captured images.

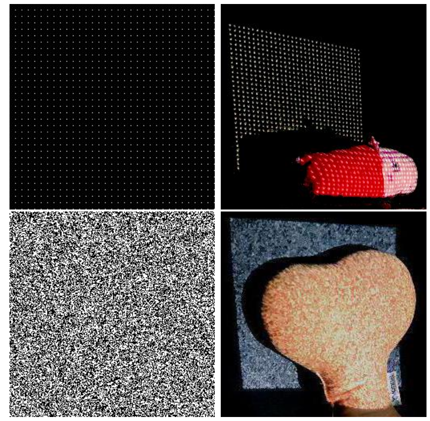

Figure 3: Projector-Camera Image Pairs
These are examples of projected images along with their captured images (After post processing) from my fixed pattern scanning and Bernoulli binary pattern image sets.

4 Results

I was not able to produce dual images from my captured image sets. I believe this is in part due to my image capturing conditions and likely more attributed to my choice of reconstruction algorithm. If I had utilized the ROMP algorithm as suggested in the paper rather than the OMP algorithm, I believe I would have obtained better results. However, as mentioned in 3.1, I was not able to efficiently implement the ROMP algorithm and did not have the time to run my slower implementation.

To see whether or not my algorithm recovered a light transport matrix that reasonably describes light going from the projector to the image, I applied the transport matrix to a checkerboard pattern as the projector matrix. This produced an image that looked as if it were completely lit with a dim light for both the fixed scanning and Bernoulli binary patterns. This demonstrates that my reconstructed transport matrix was incorrectly assessing contributions from the different projector pixels. See Figures 4 and 5 for my results as well as the expected results from the paper on compressed dual photography [1].

More discussion of what could have gone wrong and what I tried throughout this process is discussed in Section 5.

5 Discussion

5.1 Imaging Conditions

The first of several issues with my attempts at obtaining the Light Transport Matrix is non-ideal imaging conditions. The images presented in the paper showed very little light captured on anything other than the target object itself. My images, on the other hand, captured light in areas surrounding the target object. I attempted to account for this in post-processing by subtracting out a black image (image with no pixels lit), applying an empirically determined threshold, and increasing the contrast of the image. With these steps, I was able to obtain an image that was closer to the images from the paper by Sen and

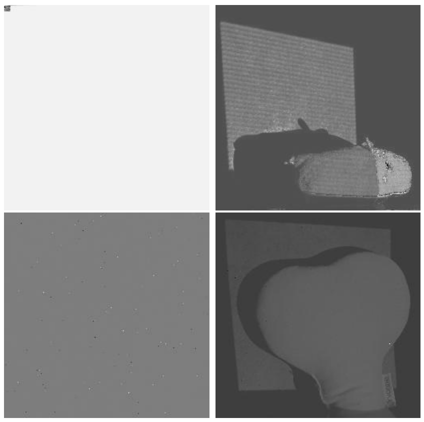

Figure 4: My Results
These are my results from running my algorithms on the grayscale of the camera images. The top row uses the fixed pattern scanning and the bottom row uses the Bernoulli binary patterns with the left being the produced dual image and the right being the image relit with a checkerboard pattern. Both image sets produced similar results in that the relit scenes look like they have been illuminated with dim lighting and the dual images are visually indistinguishable

Figure 5: Results From Reference Paper
This is an example result from the paper I referenced [1]. They used a similar setup and also used 512 Bernoulli binary patterns, but ran the ROMP algorithm and applied a compression basis to their projector matrix on top of the patterns already meeting RIC.

Darabi. However, in the processed image, the projected pattern still reflects off the back wall despite having little ambient lighting elsewhere. In the future, I would use a projector that has better contrast between light and dark pixels, create a setup that is better shielded from ambient light as well as has a minimally reflective background.

I also had to resort to using a phone camera to capture images. Initially, I had planned to use the Nikon D3300 DSLR camera. However, I found that the camera captured colors in the images that were unexpected. To attempt to fix this, I experimented with changing the white-balancing technique utilized by the camera, applying gamma decoding in post-processing, changing the brightness, white-balancing, and gamma settings on the projector, but all attempts to alleviate the issue just lead to the same image with slightly different colors in the captured image. I found that the phone captured more accurate images of the scene. In using both the DSLR and phone cameras, I would need to crop and resize the images to a lower resolution, which could also be a source of error. If I had discovered this issue earlier, I would experiment with different cameras to see what works the best: I suspect that a webcam would be a good option since I would be able to easily control it using a computer script and it would be able to capture lower resolution images.

5.2 Limitations in Implementation

In the paper on compressed dual photography [1], the authors talked about using Regularized Orthogonal Matching Pursuit (ROMP) rather than simple Orthogonal Matching Pursuit (OMP). OMP is fast since it is a greedy algorithm, but does not perform as well for exact recovery as some other ℓ_1 methods. To alleviate this, the Sen and Darabi propose using ROMP since it recovers multiple coefficients per iteration for a better recovery. I was not able to find an existing Python package that implements this, so I implemented it myself following the algorithm in Zhang et al.'s discussion on Matching Pursuit algorithms [3]. Unfortunately, since I don't have a means of parallelizing and am not sure what I could do to optimize the code to a large degree, running my algorithm to recover the light transform matrix would take an amount of time on the order of several days. Due to the lack of time, I resorted to using OMP in hopes that it would recover decent results but sadly was proven insufficient.

I also did try to use the least square solver to recover the light transform matrix, but that method also did not

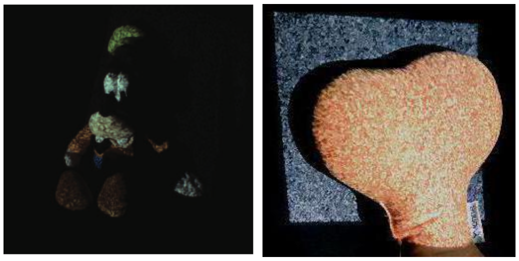

Figure 6: My Captured Images vs. Paper's Captured Images
The images the Sen and Darabi [1] captured in their paper is significantly darker than the ones I captured. Their images also do not show the projected pattern in the area behind their scene, which I was not able to reproduce using the settings on my projector/camera nor in post-processing.

Figure 7: Images Captured With Nikon DSLR vs. iPhone 12 Camera
This is a comparison of the images captured by the DSLR camera vs. the camera on an iPhone 12 both with a white screen projected onto the scene. The image on the left is the DSLR image– it has several colors that likely come from the way that the projector combines lights to create different colors. The phone camera is able to capture images much closer to what we expect.

succeed in finding a sufficiently exact and sparse matrix. One thing that I tried to improve upon the least squares method is prune the results so that each row contains only its 100 largest values so that I could get a sparser matrix and eliminate the matrix values that might have been affected by ambient lighting– this method did not produce promising results.

6 Conclusion

In this paper, I made attempts at generating dual images using compressed sensing reconstruction algorithm for sparse matrices. While unsuccessful, this was a good experience in that I had an opportunity to learn about a topic that interests me. Additionally, I was able to do an analysis of what could have gone wrong and find a direction for future work I might want to do in the future regarding the topic of dual photography.

References

[1] Pradeep Sen and Soheil Darabi. Compressed Dual Photography. *EUROGRAPHICS*. 2009.

[2] Pradeep Sen, Billy Chen, Gaurav Garg, Stephen R. Marschner, Mark Horowitz, Marc Levoy, and Hendrik P. A. Lensch. Dual Photography. *ACM SIGGRAPH*. 2005.

[3] Hanfei Zhang, Shungen Xiao, and Ping Zhou. A Matching Pursuit Algorithm for Backtracking Regularization Based on Energy Sorting. *Symmetry*. 2020.

[4] Elaine Crespo Marques, Nilson Maciel, Lirida Naviner, Hao Cai, and Jun Yang. A Review of Sparse Recovery Algorithms. *IEEE Access*. 2018.

[5] Ivan Ralasic, Matea Donlic, and Damir Sersic. Dual Imaging–Can Virtual Be Better Than Real? *IEEE Access*. 2020.