Dual Photography

Sen *et al* (2005)

Presented by Daniel L. Lu and Jonathan Shen

Overview

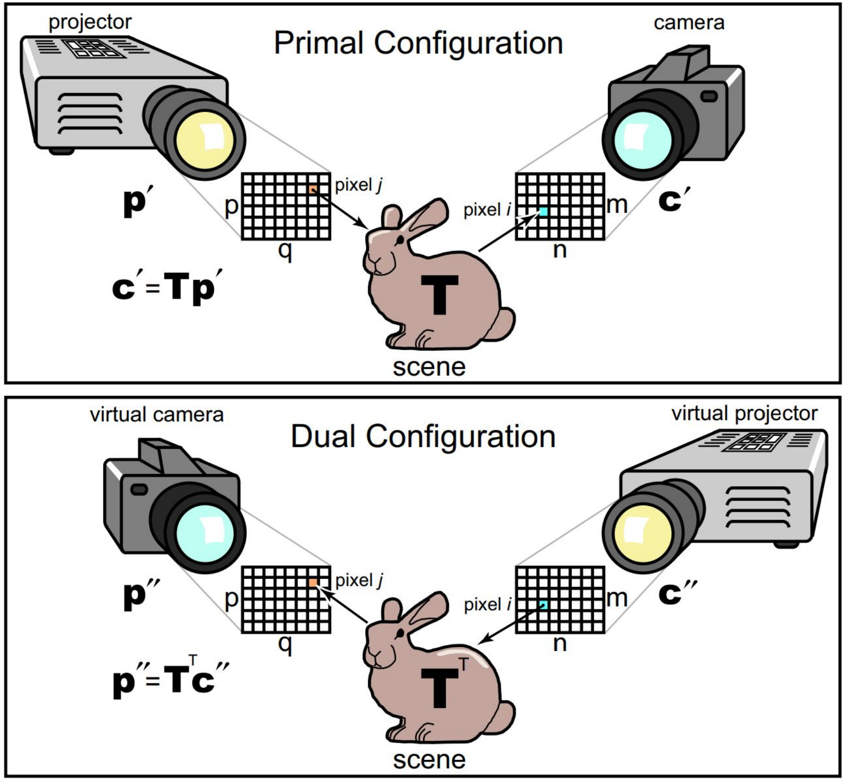
The concept of Dual Photography is based on efficiently capturing the light transport between a camera and a projector looking at a scene. By treating the projector as a “camera in reverse”, we can:

- Synthesize images from the point of view of the projector
- Relight the scene with 2D incident illumination
- Relight the scene with arbitrary light fields (if we have many cameras, or a light field camera, instead of just one)

The method is image-based and does not require knowledge about scene geometry or surface properties, while capturing global effects such as mirrored reflections, caustics, diffuse inter-reflections and subsurface scattering.

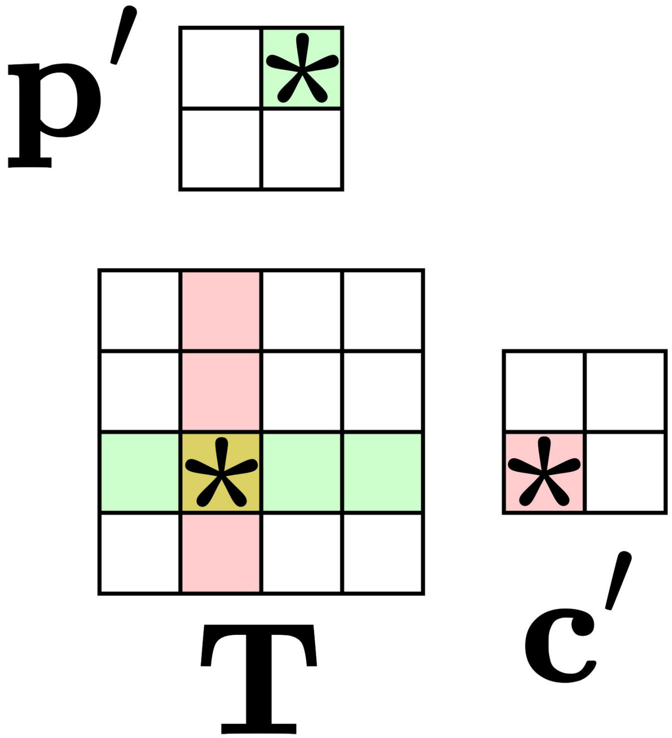
Helmholtz reciprocity

Each ray of light can be reversed without altering its transport properties. This means that we can effectively exchange the positions of the camera and the projector.



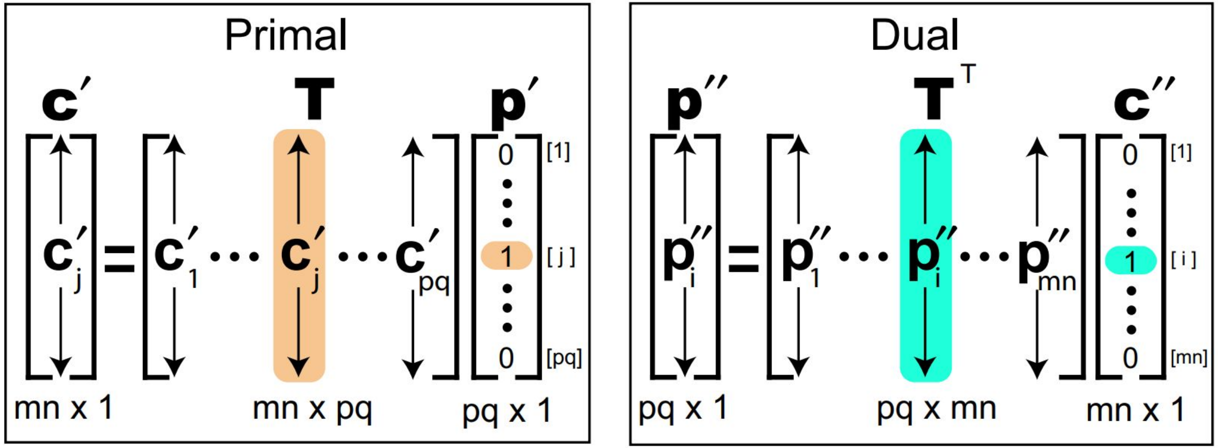
The light transport matrix **T**

Each element of **T** describes the transmission in the optical path between a pixel in the projector and a pixel in the camera.



The light transport matrix **T**

Each column of **T** is the image seen by the camera when only one pixel of the projector is turned on.



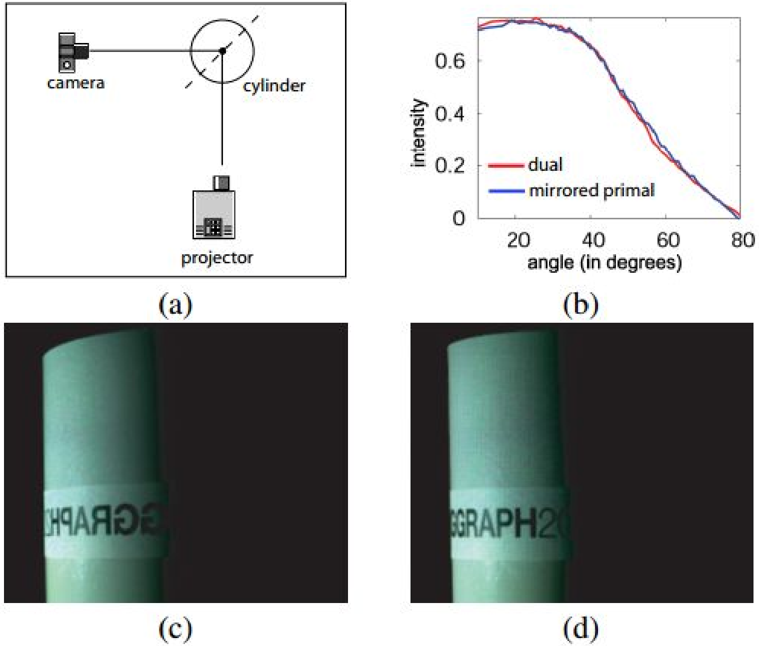


Figure 13: Experimental Validation of Dual Photography. (a) Experimental setup. (b) Plot of intensity vs. surface normal angle for scanlines in the mirrored primal (c) and dual (d) images.

Measuring **T** by brute force

The trivial way to construct the **T** matrix is to turn on one pixel of the projector at a time. Each picture you take is a column of **T**. This is called the “brute-force” pixel scan. Unfortunately,

- need to take as many pictures as there are projector pixels
- the image when only one pixel is lit can be quite dim

Measuring **T** efficiently

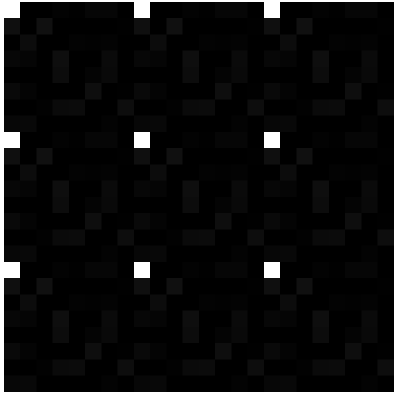
The idea is to use **structured light**, where you project known patterns with many pixels turned on at once (multiplexing) in such a way that you can easily separate the contributions from each pixel (demultiplexing).

The authors propose two ways:

- Fixed pattern scanning. Project a set of patterns known in advance.
- Adaptive multiplexed illumination. Project increasingly finer patterns to resolve detail as needed.

Measuring **T** efficiently: fixed pattern

One method is to divide the projected image into blocks of, say, 8×8 and then illuminating one pixel from each block.



To recover which block each pixel in the image corresponds to, we can project an obvious pattern for each block over a number of frames (e.g. the block number in binary).

Measuring **T** efficiently: fixed pattern

The very strong assumption is that pixels far away in *projector space* should also be far away in *camera space*.

This only holds if there are no global illumination effects.

But in the general case, two pixels illuminated far away in the projector space can lead to the same pixels in the camera image being lit, leading to ambiguities.

Measuring **T** efficiently: adaptive patterns

The adaptive pattern recursively refines blocks into 4 quadrants.

When a block is subdivided, the quadrants are illuminated in sequence to detect *conflicts*, which happen when the same region in the camera image is illuminated by two sub-blocks. If two sub-blocks have a conflict, they will have to be investigated one after another; otherwise, they can be subdivided in parallel.

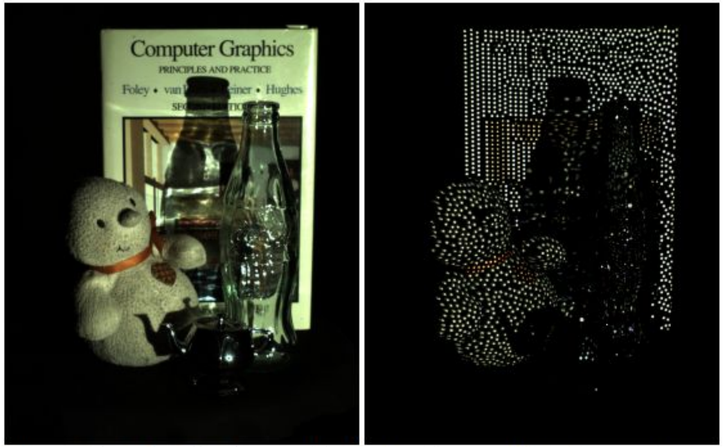


Figure 6: Adaptively parallelized patterns. This figure demonstrates that how our algorithm adapts to the scene content. Because of the complex light transport within the bottle visible in the left image, only a few pixels can be investigated in parallel in this region. Thus, the bottle remains relatively dark when we project an adaptive pattern as shown on the right.

Measuring **T** efficiently: Problems

In scenes where diffuse inter-reflections or subsurface scattering dominates the appearance, each projector pixel could be spread across large areas in the camera image, causing adaptive method to degrade to brute-force.

Energy can be lost if the contribution from one projector pixel to one camera pixel falls below the noise threshold. Many entries have low energies, and the sum of their conclusions is significant.

Measuring **T** efficiently: hierarchical adaptive patterns

Instead of capturing pixel level matrix **T**, capture a sequence of matrices $\mathbf{T}_k$ at different scales, then upscale and combine them.

$$\mathbf{p}'' = \sum_k f(\mathbf{T}_k^T \mathbf{c}'')$$

Experimentally demonstrated that this works in $O(\log \# \text{ projector pixels})$

Measuring **T** efficiently: hierarchical adaptive patterns

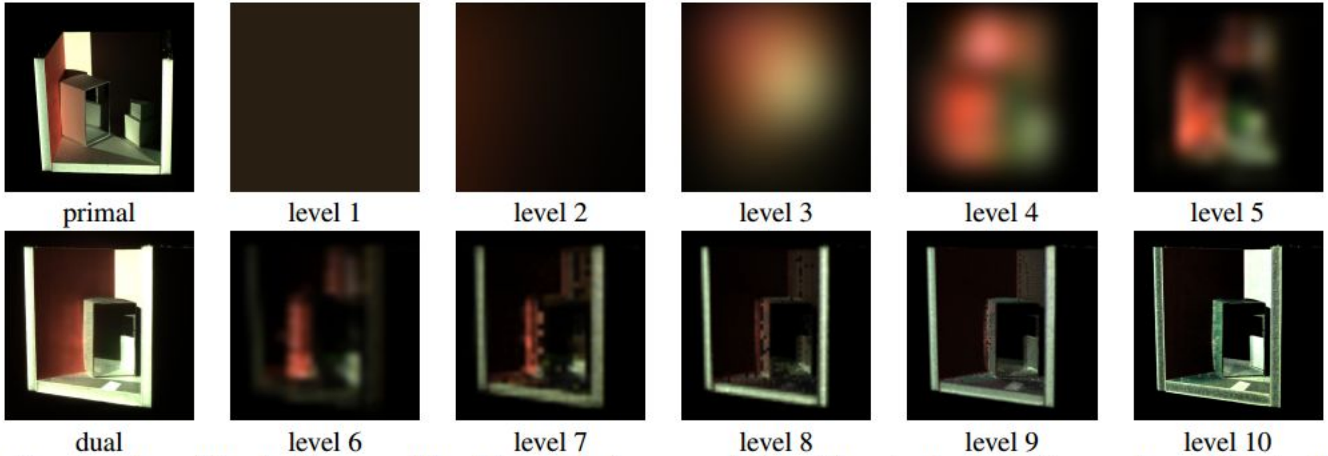


Figure 14: Construction of the dual image with a hierarchical representation. The primal and dual image show diffuse to diffuse inter-reflections which could only be captured by use of the hierarchical acquisition. Energy that might have been lost when further subdividing a block is deposited at a coarse level of the **T** matrix. To synthesize the dual image, the levels are individually reconstructed by applying the appropriate basis functions, then added together to obtain the image on the left. In this figure the intensity of the images for level 1 to 9 has been increased to visualize their contribution.

Measuring **T** more efficiently: compressive sensing

Compressive sensing exploits sparsity to recover images using few random samples. In 2009, the authors wrote “sequel” paper that greatly speeds up the dual photography process.

Compared to the adaptive methods, compressive sensing does not require time-consuming computation in real time between captures since it uses a fixed set of patterns.

- Peers, Pieter, et al. “Compressive light transport sensing.” *ACM Transactions on Graphics* (TOG) 28.1 (2009): 3.
- Sen, Pradeep, and Soheil Darabi. “Compressive dual photography.” *Computer Graphics Forum*. Vol. 28. No. 2. Blackwell Publishing Ltd, 2009.

Scene relighting (2D)

Knowing **T**, we can exchange the camera and projector.

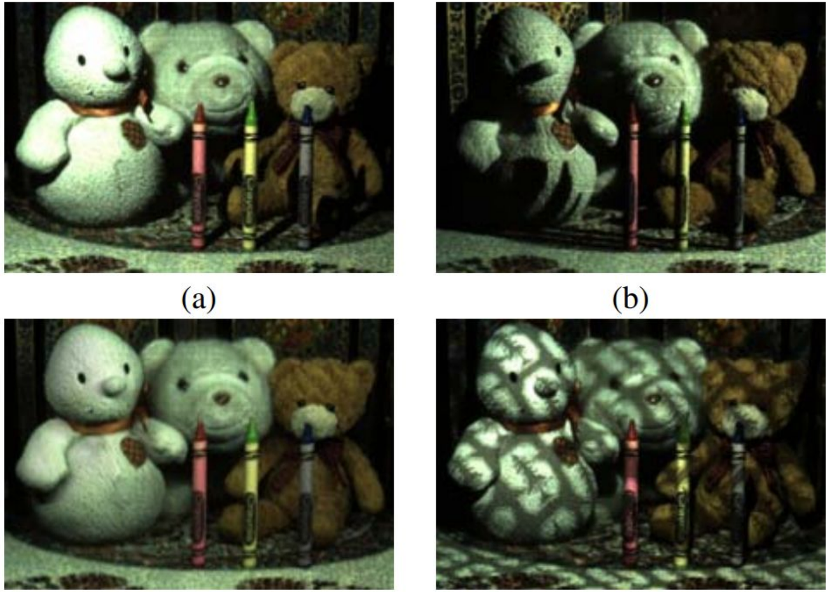
The image $\mathbf{p}''$ "seen" by the projector, given 2D illumination pattern $\mathbf{c}''$ from the camera position would be:

$$\mathbf{p}'' = \mathbf{T}^T \mathbf{c}''$$

Scene relighting (4D)

With many cameras, we can simulate illumination from a light field. If the ith camera sees the image $\mathbf{c}_i = \mathbf{T}_i \mathbf{p}$ then we can relight the scene from the projector point of view by summing over all cameras:

$$\mathbf{p}'' = \sum_i \mathbf{T}_i^T \mathbf{c}_i''$$



(a) (b) (c) (d)

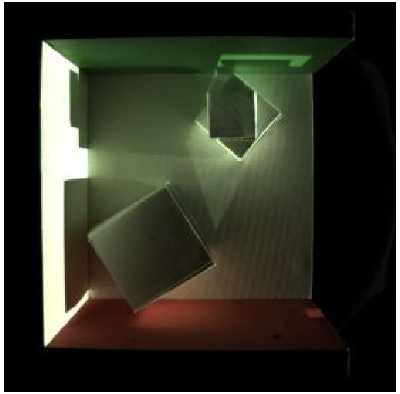
Scene relighting (4D)

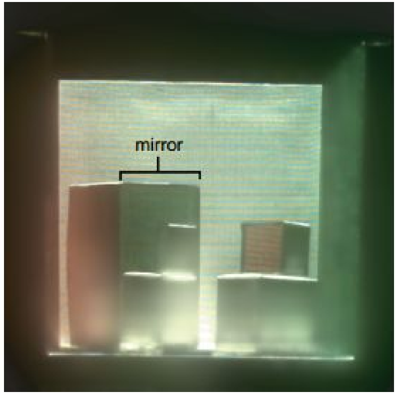
Instead of using many cameras and a single projector, previous work used a single camera and many projectors. The advantages of the dual photography approach are:

- Many cameras can take pictures at once, but many projectors will interfere with the patterns each other produces.
- Cameras are much cheaper and more compact than projectors.

Some disadvantages include the fact that the final image's resolution is limited by that of the projector, which is usually low.

Limitations





- If there is not much direct light transport between the projector and camera, most pixel contributions would be close to noise level, so transport matrices at higher scales need to be used, leading to blurring.

- Dual image is produced from projector, which typically has much shallower depth of field and lower resolution than camera.

- Scenes with significant global illumination still cannot be processed efficiently.

- It takes a long time to capture **T**, and scene must be static.

- **T** is very large and uses a lot of memory.

Conclusion and score

Overall rating: 2 (accept).

- The contributions are novel, useful, and simple to understand.
- The experimental setup is clearly described.
- Results are backed with objective evidence.

- The paper is overly verbose.
- Focused more on efficient capture of T instead of strengths and limitations or applications of dual photography.