Computational Imaging and Photography

Lecture #12: Wednesday, 14 Oct 2009
Lecturer: Ravi Ramamoorthi
Scribe: Yiding Jia

1 Introduction

Imaging is an integral tool in computer graphics, people have been using real images to create new images. Now that digital cameras are commonplace, can we do more with digital cameras than with film cameras?

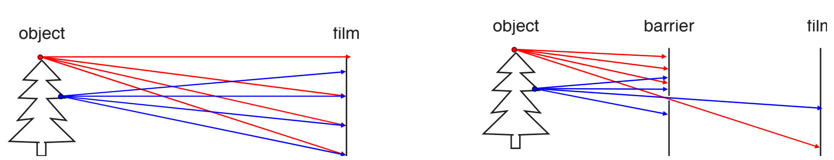
Many novel possibilities exist in the fields of computer graphics, optics, and image processing, building off of ideas in rendering and in particular, light fields. There has been a constant shift in applications for computer graphics, from virtual environments, to CAD, to movies and video games, and now to photography, which is a bigger market and has many more consumers than the previous fields

2 Image Formation, Lens Based Camera

The first question is: how do we see the world and how do we design a camera to capture images of the world? The first idea is to put film in front of the object, however since each point receives light from all directions in the hemisphere, the light is not focused and it does not produce an image.

2.1 Pinhole Camera

Using barriers and masks in front of the film produces interesting effects. The pinhole camera model involves putting a barrier with a pinhole in front of the film. Most rays are stopped by the barrier, ray from only one direction hit each point of the film. This is the model typically used in ray tracer and path tracers.



(a) No barrier **(b)** A pinhole barrier

Figure 1: Pinhole Camera Model

An inverted image of the 3D world is projected onto the film. In this process of doing the persepective projection from 3D to 2D, angle and length is not preserved. Parallel lines are no longer parallel, and further objects appear smaller. Humans have adapted to this by not perceiving lengths in the image plane, but rather, by estimating the geometry of the scene.

2.2 Lens Based Camera

While in computer graphics we typically assume a perfectly small pinhole, making the aperature narrower limits the amount of light and may introduce diffraction patterns. On the other hand, making the aperature larger causes blurriness in the image. Real cameras therefore combines a large aperature with lenses in order to let through more light and allows for the light to be focused on the image plane.

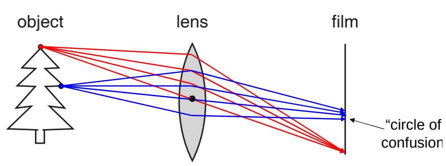


Figure 2: Lenses and Focus

Lenses introduce the depth of field effect. Only light from a specific depth will be focused onto a point in the image plane. At other depths, the light is out of focus, and projects into a "circle of confusion" around the point. The lenses can be changed to change the focal distance.

The thin lens equation allows one to compute the focal distance given the distance between the lens, the object, and the image plane.

$$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}$$

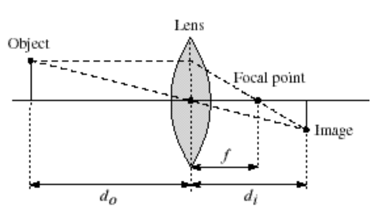


Figure 3: Thin Lens Model

3 Light Field Camera

In a previous lecture the applications of capturing a light field was discussed. But how do we modify a camera in order to acquire a full lightfield?

3.1 Microlens Camera

Typically, when light arrives at a point on the image sensor, the light from every direction is integrated together. If instead we put an array of microlenses in front of the image sensor. The light that is focused on this microlens array is then propagated to different parts of the sensor, allowing the camera to capture the full 4D light field that varies across space and direction.

(a) Regular Camera **(b)** Microlens Camera

Figure 4: Microlens Camera Model

The stanford plenoptic camera uses a 16 megapixel sensor (4000x4000 pixels), and an 292x292 microlens array. Each microlens captures a 14x14 image. The final image resolution is reduced since the image captured is the 4D lightfield instead of a 2D image.

3.2 Digital Refocusing (Application)

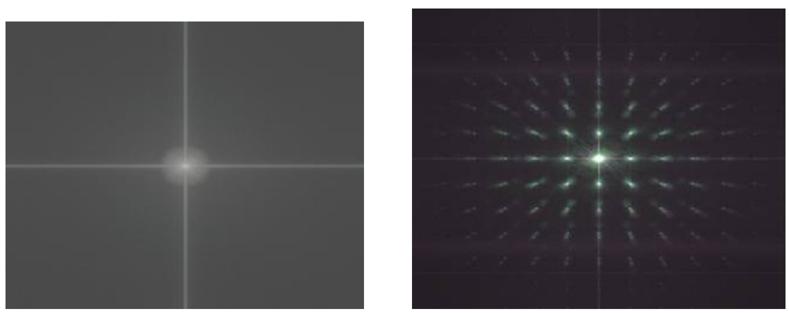
One application for a captured light field is digital refocusing. If we capture the whole light field, we can choose which set of rays we use, corresponding to focusing at different depths, and thus digitally refocus the image after capturing it. This is a key concept in computational photography: capture richer information about the image, and then using it to create different images.

3.3 Mask Based Lightfield Camera

Another method of making a light field camera is to put a mask before the sensor. By introducing a barrier, parts of the light field is blocked and the image sensor effectively does not integrate the entire light field, similar to how a pinhole camera only allowing through a single ray of light onto the film.

From a signal processing point of view, a regular camera convolves light from the entire aperature of the lens. If we introduce an invertible filter to convolve the signal by, it makes it possible to recover the information.

Figure 5: Digital Refocusing with microlens camera: The focus can be shifted from the faces from the front to the faces in the back



(a) FFT, without mask

(b) FFT, with mask

Figure 6: Mask Based Lightfield Camera: More information about the image is retained in the frequency domain

The cosine mask used tiles on the frequency domain. Each tile encodes information corresponding to different incoming angular directions. To recover the information we isolate each tile and apply an inverse fourier transform.

Both techniques, microlens and masks, represents the idea of multiplexing, that is, encoding high dimensional information in lower dimensions. The microlens technique multiplexes the light field into the spacial domain of the image, while the mask based technique multiplexes the light field into the frequency domain of the image.

3.4 Coded Aperature Depth of Field

A similar idea to using a mask at the image sensor, we can put a mask at the aperature in order to get more frequency information. The incoming light is blocked in certain regions instead of simply being a low pass filter, and thus allows some depth information to be recovered.

There are other simpler techniques such as acquiring a set of photos focused at dif-

(a) Focused **(b)** Unfocused **(c)** Coded Aperature

Figure 7: Coded aperature images

ferent distances to allow for refocusing but without capturing the full light field.

A drawback of these techniques is that the viewpoint can't be changed very far since all the recorded light is still coming through the aperature of the camera. Using an array of cameras allow for greater viewpoint changes as well as focusing at different depths to see through occluders blocking some of the cameras.

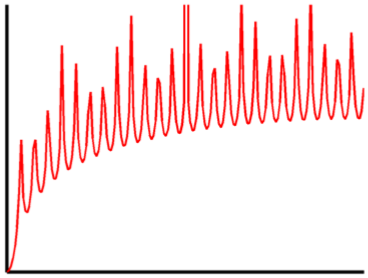
4 Flutter Shutter

Another technique called "flutter shutter", allows for recovering scenes that vary in time, i.e. undo motion blur. In a traditional camera, the shutter is open within a duration of time, the equivalent of convolving the incoming signal with a box filter. There is a loss of high frequency information as the shutter is open for a long time, but it allows for a stronger signal (more light) than using a faster shutter, which tend to be more noisy. With flutter shutter, the shutter is opened and closed to during the interval to create a different signal, with the ultimate goal of producing an image that is unblurred while maintaining good signal to noise ratio.

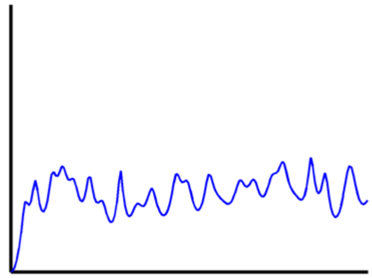
More precisely, the result is a convolution over the duration of the shutter $h = f \star g$, where $f(t)$ is the incoming light signal, and $g(t)$ is whether the shutter is open. In the frequency domain, convolution is multiplication, and thus $H = FG$. We wish to find $F = \frac{H}{G}$. If g is the delta function, i.e. the shutter is open for an instant in time, $G(t) = 1$, and finding F is easy. But with a traditional shutter, g is a box filter, and G is the *sinc* function. In this case $g(t) = 0$ for many values of t, and thus it is difficult to invert.

The flutter shutter can be seen as a filter that is open half the time, using a sequence such that in the frequency domain the filter approximates $g(t) = 1$. The particular filter used is probably derived experimentally in simulation, and is constrained by the physics of the shutter mechanism. The pattern should be designed such that there is no frequency bias.

A limitation of the system is that the velocity of the object needs to be known, and

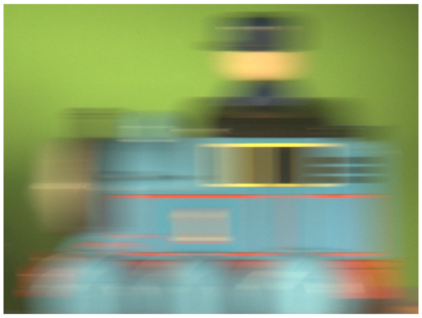


(a) Standard Shutter

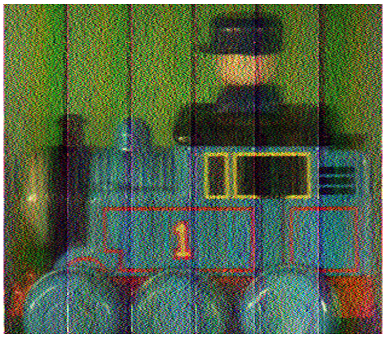


(b) Flutter Shutter

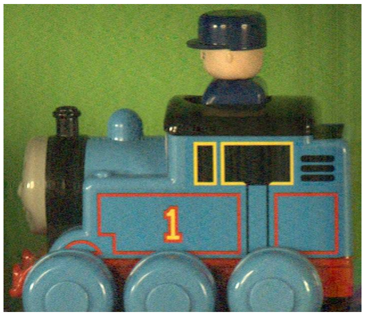
Figure 8: Inverse signal in the frequency domain of standard shutter is unstable, while that of the flutter shutter is stable



(a) Blurred, Standard Shutter



(b) Unblurred, Standard Shutter



(c) Unblurred, Flutter Shutter

Figure 9: Flutter Shutter Unblur Result Comparison: Notice the reduced noise and banding in the flutter shutter version

the image must be segmented to areas with different velocities. A similar method uses a camera that has a velocity that vary quadratically with time, but is limited to 1D motion within a min and max range of velocities.