Removing Artifacts Due To Frequency-Domain Processing of Light-Fields

Chintan Intwala[1], Todor Georgeiv[1],

[1] Adobe Systems

Abstract

In previous works, light-field capture has been analyzed in spatio-angular representation. A light-field camera samples the optical signal within a single photograph by multiplexing the 4D radiance onto the physical 2D surface of the sensor.
Besides sampling the light field spatially, methods have been developed for multiplexing the radiance in the frequency domain by optically mixing different spatial and angular frequency components. The mathematical method for recovering the multiplexed spatial and angular information from the frequency representation is very straightforward. However, the results are prone to lots of artifacts due to limitations inherent to frequency-domain processing of images. In this paper, we try understand the characteristics of these artifacts. Furthermore, we study the effect and sources of artifacts that affect the quality of the results and present various methods for the removal of artifacts.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Digital Photography

1. Introduction

A central area of research in computational photography is capturing "light itself" as opposed to capturing a flat 2D picture. Advantages of this light-field or integral photography are gaining information about the 3D structure of the scene, and the new ability of optical manipulation or editing of the images, like refocusing and novel view synthesis.

As demonstrated by Levoy and Hanrahan [LH96] and Gortler et. al. [GGSC96], capturing the additional two dimensions of radiance data allows us to re-sort the rays of light to synthesize new photographs.

Ng et al. [NLB*05] showed that a full 4*D* light field can be captured even with a hand-held "plenoptic" camera. This approach makes light field photography practical, giving the photographer the freedom and the power to make adjustments of focus and aperture *after* the picture has been taken.

Along with techniques for analyzing radiance in the frequency domain [CCST00, DHS*05, Ng05, VMA*07], a number of good results have been derived, like application of Poisson summation formula to depth representation, of general light fields and displays, optical transforms, Fourier slice theorem applied to refocusing, and others. However, one area that hasnt achieved full attention is the quality of the resultant images. This is because achieving high-quality results by doing frequency domain processing as a primary method of decoding the 4D light field, is non-trivial. Although the overall method of decoding the light field is straightforward, there are various limitations in getting things *just right*. In this paper, we demonstrate the presence of factors responsible for degrading artifacts and furthermore, we present techniques for removing artifacts.

2. Artifacts

The artifacts are best visible as an intensity wave across the various angular views. These waves appear as random low-frequency waves that are present spatially within a single angular view, but they travel across views. Despite the fact that waves show up in individual angular views, they completely disappear when the views are mixed for the an application like refocusing. Look at Figure 1 to see an example of intensity wave. Also the angular views seem to be

Figure 1: *The right column shows the original images from 'Beach Seagull' and 'Apple Tree' data set. The left column shows the corrected results using techniques described in 4.3 and 4.4 respectively.*

Figure 2: *Chromatically bad result.*

warped in a peculiar way such that when you view them in a movie, there is a wind-like movement. Other artifacts appear as chromatic aberration which arise to incorrect alignment of the color channels after frequency-based processing. Refer to Figure 2 for an example of this artifact. We strongly encourage to view the attached electronic material as they best illustrate these effects.

3. Source of the waves

As detailed by Veeraraghavan et al. [VMA*07] and [NLB*05], the process of light field capture through mask-based or lens-based multiplexes the angular information within a 4D light field visible in the 2D Fourier transform of the acquired image. This modulated data, or referred to as slices [Ng05], are demodulated (or simply cropped) and rearranged into a 2D stack [VMA*07]. As a result, a crucial step involved in processing, is to detect the center of these slices. Incorrect detection of the centers of these slices directly corresponds to modulation of these cropped slices by a very low-frequency. Now since DFTs sample the underlying DTFT of the image, detection of the correct location of the peak can be correct only within 1-pixel neighborhood. Also, there are can be cases where the detected center can be completely incorrect, due to the fact that the original signal contains large low-frequency side-lobes. This is responsible for one of the most prominent artifacts that arise in the results obtained by light field cameras – the traveling intensity waves; see Figure 1.

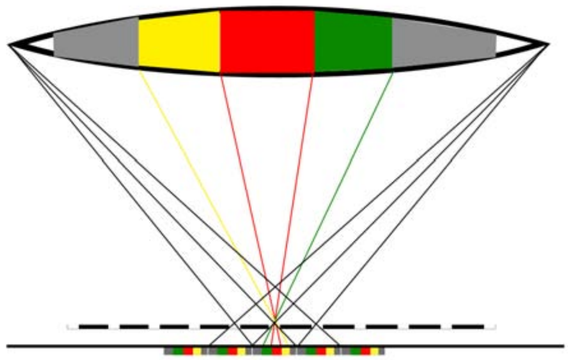

Figure 3: *When the F-numbers are matched, the sub images are tightly spaced together.*

Now the other form of artifacts, are seen as the wind-like movement. One theory that explains such artifacts is related to the circular aperture of cameras. But before that, lets first look at how each pinhole captures different parts of the main aperture. Based on the design of Veeraraghavan et al. [VMA*07], it can been seen that the sensor sees different slices of the main aperture as illustrated in Figure 3. Now this is the case when the F-number of the optical system consisting of the mask and the sensor is equal to that of the system consisting of the sensor and the main lens. Now each slice of

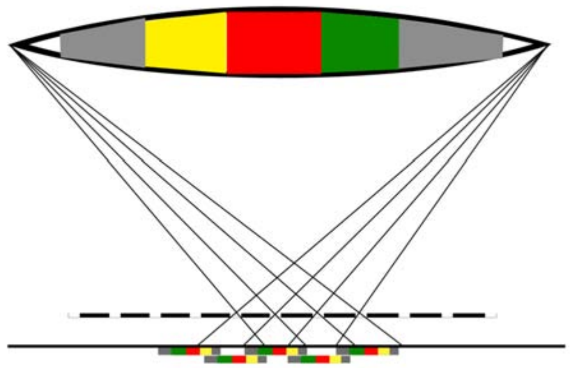

Figure 4: *When the F-numbers are not matched, the subimages overlap with neighboring subimages.*

the aperture, because it is circular, contains different amount light; the central one with the highest amount of light. As a result each subimage contains different amount of light. Finally, when these subimages are put together to form one coherent angular view, all views contain different intensity. Moreover these artifacts are worsened when considering the case of Figure 4 . This is the case when the F-numbers of the above mentioned optical systems are different. As a result, the subimages overlap each other onto the image sensor. This would mean the edges of the subimages will contain information about the neighboring subimages. However, as evident from the Figure 3 , this spatially neighboring information isn't well neighboring from an angular information point of view. This means that there will be redundant information in farthest angular views. Also, this is the best case scenario when the physical mask is optically high quality and is placed equidistant from the sensor.

Distance of the mask from the sensor directly controls this amount of overlapping angular information. (This can also be looked as spatial aliasing whose frequency counterpart can be understood in the angular dimension.) So if the mask was spaced unevenly from the sensor, it would correspond to diffusion of angular information across subimages. This diffusion can be confined to different regions of the image as it is simply controlled by the distance of the mask from the sensor. Finally, when all the angular views are extracted, each pixel doesn't necessarily correspond to the same view. This like wind going through the stereo set. See the attached movie to observe this effect. The movie plays the angular views 1,2,3...7,1,2... sequentially; however, there is no visual jump going from view 7 to 1. This is because of the above described phenomena. This observation leads to the conclusion that the true angular view is not a simple frame in this space-time volume of angular views, but rather a surface within the volume.

Another source of artifacts is due to the optical transfer function of the lens (or the pre-mask optical system). Figure 6 , shows an exaggerated effect of this transfer function to the signal in frequency domain. It is evident that the angular slices are not symmetric anymore causing the overall result of the entire process to be complex instead of real. The effect of this attenuation function is harder to correct as we approach high frequencies, as it gets harder to invert. This problem is similar to a traditional signal processing problem of amplitude demodulation in presence of a memoryless channel.

4. Methods of removal

4.1. Oversampling

As described above, determining the center of the Fourier slices is extremely important. In an extremely well-calibrated camera, the location of centers can be predetermined. However, most of the times it is hard to achieve. Another way to see this problem is to find the location of the carrier frequencies (of the mask or microlens array) in the observed image. One of the ways to correctly detect the center of the slice is to oversample the DFT of the image, by atleast 3 times. This is based on our experiments. However, due to practical limitation of memory on various OSes, this is hard to achieve with straightforward FFT implementations.

4.2. Phase Multipliers

Figure 5: *(a) Shows the original result and (b) shows the result after multiplying by the correct phase.*

If oversampling is not possible due to hardware/software problems, there are other ways of detecting the exact centers of the Fourier slices. For most practical examples, our neighborhood of certainty lies within 1x1 pixel and we are looking for subpixel accurate location of the center of the slices. Assume we located the center to be (x, y) however the real location is $(x+\delta x, y+\delta y)$ where $|\delta x, \delta y| < 1$.

Now instead of shifting the entire slice by $(\delta x, \delta y)$ we can use the following property of Fourier-transform, $\mathbf{F}\{x(x+\delta t)\} = \mathbf{F}(\omega)e^{-i\delta t}$. So for slice S_i we multiply it by the phase factor $e^{-i(\frac{2\pi\delta x}{N}x+\frac{2\pi\delta y}{M}y)}$; then compute its inverse Fourier transform to get $S_{i,mid}$. We re-arrange S_i, mid in 2D and then perform another 2D-inverse Fourier-transform in order to obtain the angular views. As a result the overall process $4DIFT(arrangein4D(R(\omega_{\mathbf{x}}, \omega_{\theta})))$ is replaced with $2DIFT(arrangein2D(correctphase(arragein2D(R(\omega_{\mathbf{x}}, \omega_{\theta})))))$. One can use various methods to detect the motion $(\delta x, \delta y)$, however, we used brute force method in order to detect the shift to illustrate the use of this method. Figure 5 shows the result obtained from this technique.

4.3. Skip the lowest slice

In our observation and experiments with various camera designs we have noticed that there are cases when the effect of low intensity waves is severe. That is, major regions of all angular views are completely black. This would mean that the intensity wave is fairly stationary across all the angular views. This would corresponds to a very low-frequency component that corrupts the results in angular dimension. Intuitively speaking, one would expect such corruption introduced through the carrier (or modulator). In order to test this intuition, we excluded the lowest modulated slices (or the ones closes to the central slice) from the entire demultiplexing process. As a result, the wave artifacts, almost disappeared. The results shown in Figure 1, demonstrate the usefulness of this technique.

4.4. Cosmetic Correction

Despite all the above mentioned corrections, one might find that there is a fair amount of vignetting-type artifacts or again slowly varying color-shifts or other wavy artifacts in image intensity. As a result, we have developed a cosmetic correction technique that tackles exactly this problem. The intuition here is that, all the artifacts are present in low-frequency content of the image. As a result, the corrections need to be targeted to the low-end of the image spectrum. The corrections are simply additive in nature, however, they great increase the visual quality of the extract angular views. Its essentially a color/intensity matching process, to start, pick a view that the rest of the views will have to match to. Call this S_base. Now we adjust the rest of the angular views using,

$$S_{i,new} = S_i + Filter(S_{base} - S_i)$$

Here, $Filter$ if a low-pass filter. For the sake of speed, we accomplish this by,

$$Filter(x) = gaussian(\sigma) * downsample(x)$$

Results, in Figure 1 and Figure 7, are obtained using this algorithm as a primary means of wave-correction. The visual quality is good. However, one draw back of this corrective procedure is that, it replaces darker regions with blurry content, despite the fact that its chromatically correct. See Figure 7.

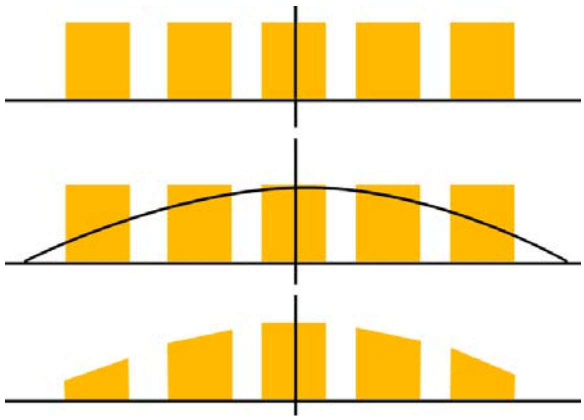

Figure 6: *From top to bottom, (a) shows illustration of the Fourier transform of original light field before it passes through the camera optical system, (b) shows hypothetical transfer function of the camera's optical system (c) shows the resulting light field as captured by the sensor.*

Figure 7: *Left image is original. Right image is cosmetically corrected.*

5. Conclusion and future work

The idea of multiplexing the $4D$ light field onto a $2D$ sensor is extremely note-worthy and powerful. More specifically, the idea that the use of a simple mask can turn a conventional camera into a ligthfield camera is very practical. This approach requires processing the captured lightfields in frequency domain. The general concept of processing data in frequency domain with traditional tools like Fourier transform is fairly straightforward, but there are a lot of implementation details and caveats that one needs to be aware of in order to obtain visually pleasing results.

In this paper, we have shown the effect of various factor which need to be correct in order to get high-quality results. By developing intuition about these factors and sources of artifacts, we enable upcoming researchers to continue investigating techniques that allow extremely high-quality processing of light fields data sets via frequency domain based signal processing.

References

[CCST00] CHAI J., CHAN S., SHUM H., TONG X.: Plenoptic sampling. *ACM Trans. Graph.* (2000), 307–318.

[DHS*05] DURAND F., HOLZSCHUCH N., SOLER C., CHAN E., SILLION F.: A frequency analysis of light transport. *ACM Trans. Graph.* (2005), 1115–1126.

[GGSC96] GORTLER S. J., GRZESZCZUK R., SZELISKI R., COHEN M. F.: The lumigraph. *ACM Trans. Graph.* (1996), 43–54.

[Ive03] IVES F.: Patent us 725,567.

[LH96] LEVOY M., HANRAHAN P.: Light field rendering. *ACM Trans. Graph.* (1996), 31–42.

[Lip08] LIPPMANN G.: Epreuves reversible donnant la sensation du relief. *J. Phys. 7* (1908), 821–825.

[Ng05] NG R.: Fourier slice photography. *ACM Trans. Graph.* (2005), 735–744.

[NLB*05] NG R., LEVOY M., BRĹĘDIF M., DUVAL G., HOROWITZ M., HANRAHAN P.: Light field photography with a hand-held plenoptic camera. *Tech. Rep.* (2005).

[VMA*07] VEERARAGHAVAN A., MOHAN A., AGRAWAL A., RASKAR R., TUMBLIN J.: Coded aperture and optical heterodyning : A mask-based approach for digital refocusing and light field acquisition by conventional cameras. *ACM Trans. Graph.* (2007), 69–12.